A Battery Capacity and RUL Prognostic Approach Based on ARIMA and PF

Yun Yang, Haodong Liu, Qiang Zhang, Meng Huang, Jing Wu, Xuan Zhang and Chaolong Zhang

School of Electronic Engineering and Intelligent Manufacturing, Anqing Normal University, Anhui, Anqing, 246011, China

Abstract

It is important to predict the future available capacity and the remaining useful life (RUL) of battery accurately for the stable driving of electric vehicles in long terms. Therefore, a predicting technique integrating ARIMA and PF are presented in the paper. The ARIMA learns the past ten available battery capacities, which is used to predict the next available capacity. Particle Filter (PF) builds a model based on the past long-term available capacities, which is used to predict the decline law of battery capacity and discover the RUL of battery. The method proposed in this paper is tested by several batteries of NASA. The mean absolute errors of battery5, battery6 and battery7 are 0.006871 Ah, 0.011197631 Ah, and 0.005769204 Ah, respectively. The RUL error of battery B5, B6 and B7 by PF algorithm are 9 cycles, 2 cycles and 6 cycles, respectively.

Keywords

Battery; Capacity; RUL prognostic; ARIMA; PF.

1. INTRODUCTION

In recent years, environmental pollution and energy crisis have become more and more serious, so the automotive industry has focused on developing the energy-saving and emission-reducing electric vehicles. Lithium battery has advantages of small size, lightness of weight, high energy density, a wide temperature range, low self-discharge rate and long lifetime. Currently, it is the ideal alternative of power source for electric vehicles. Therefore, the future capacity change and the remaining useful life (RUL) of batteries [1, 2] are the current focus. Since the aging velocity of the battery cannot be accurately measured, it is necessary to make a reliable prediction of its RUL and remaining battery capacity. Therefore, the battery reasonable replacement time and battery state are provided for the users, which can ensure the electric vehicles functioning normally by predicting the RUL and the future available capacity of battery accurately.

RUL Prognostic approaches can be normally classified into two categories: data-driven approach and model-based approach. Data-driven approaches mainly discover the variation laws based on the monitoring data of equipment or system during the process from normal to failure, which needs huge amount of data. However, little data is commonly collected from equipment or system in use. Model-based approaches establish a mathematical or physical model of the aging process of the equipment or system. Particle Filter (PF) is a novel nonlinear filtering method that combine Bayesian learning techniques with importance sampling, and the method provides good state tracking performance. Therefore, PF is used to battery RUL prognostics. However, the method’s accuracy is not high in short-term prediction.

The Auto Regressive Integrated Moving Average (ARIMA) [3,4] model is the most widely used model in mathematical statistics methods, which uses the lagged variable itself and the random
error to interpret the variable. In the paper, a novel RUL prognostic method integrating ARIMA and PF is proposed. The ARIMA is used in short-term prediction and the PF is used in long-term prediction. The ARIMA learns the past ten available battery capacities, which is used to predict the next available capacity. The PF builds a model based on the past long-term available capacities, which is used to predict the decline law of battery capacity and discover the RUL of battery. The proposed prognostic method combines the advantages of ARIMA and PF. It can predict the remaining life and the future available capacity of battery accurately.

The remainder of this paper is organized as follows. The lithium battery capacity degradation process is illustrated in Section 1. ARIMA algorithm and PF algorithm are given in Section 2. Experimental process and result analysis are presented in Section 3. Finally, conclusions are drawn in Section 4.

2. THE PROCESS ANALYSIS LITHIUM BATTERY CAPACITY DEGRADATION

Lithium battery is a device for energy storage and conversion. Its capacity naturally decreases with constant using, which makes the capacity insufficient and voltage drop. Therefore, it is particularly critical to analyze the reasons for the degradation of battery capacity.

The lithium battery is composed of a positive electrode, a diaphragm, a negative electrode, an organic electrolyte and a battery shell. The graphite and the lithium cobaltate (\( \text{LiCoO}_2 \)) are applied to negative electrode and positive electrode respectively.

The total chemical equation of lithium battery is as following

\[
\text{LiCoO}_2 + 6C = \text{Li}_{(1-x)}\text{CoO}_2 + \text{Li}_x\text{C}_6
\]  

(1)

The reaction on the positive electrode is as following

\[
\text{LiCoO}_2 \rightleftharpoons \text{charge} \rightleftharpoons \text{Li}_{(1-x)}\text{CoO}_2 + x\text{Li}^+ + xe^-
\]  

(2)

The reaction on the negative electrode is as following

\[
6C + x\text{Li}^+ + xe^- \rightleftharpoons \text{discharge} \rightleftharpoons \text{Li}_x\text{C}_6
\]  

(3)

The reasons for the capacity degradation of lithium batteries [5, 6] include the abuse of lithium batteries, high or low temperature, structural destruction and deactivation of positive electrode and cathode materials, overcharging and irreversible self-discharge of lithium-ion batteries, consumption of lithium-ions, growth of solid electrolyte membrane, decomposition of electrolyte and dissolution of active substances, etc. The lithium-ions are consumed by these chemical reactions which reduces the amount of recyclable lithium and create other precipitates. Because lithium and other precipitates block the diaphragm, the battery's internal resistance increases and capacity gradually decreases.

3. METHOD

3.1. ARIMA Algorithm

3.1.1 Principle of algorithm

The analysis of time series is arranged in Chronological order. [7] Mathematical statistics method is used to deal with these data and predict their development. The Auto Regressive
Integrated Moving Average (ARIMA) model is the most widely used model in mathematical statistics methods, which uses the lagged variable itself and the random error to interpret the variable. ARIMA model can be expressed as ARIMA (p, d, q) and it is considered as a generalization of Autoregressive-Moving Average (ARMA) method. In this work, ARIMA model is employed to predict the future capacity of battery.

### 3.1.2 Modeling process

Setting these series $y_t$ as a stationary series with performing d-order differences, the formula is as following

$$u_t = y^d_t = (1-B)^d y_t$$  \hspace{1cm} (4)

The series $u_t$ are stationary, which can be expressed by ARMA (p, q) model. The model is as following.

$$u_t = c + h u_{t-1} + B + h_p u_{t-p} + \varepsilon_t + \varphi \varepsilon_{t-1} + B + \varphi_q \varepsilon_{t-q}$$  \hspace{1cm} (5)

The ARMA (p, q) model is called ARIMA (p, d, q) model after d-order differences, where p is the order of autoregressive model; q is the order of the moving average; $\varepsilon_t$ is a white noise process; h and $\varphi$ are the undetermined parameter, and c is a constant.

### 3.2. Model Establishment Steps

Stationarity test: Import the experimental data of battery discharge capacity and judge whether the series is stationary. There are some methods to test the series, which include sequence diagrams test, autocorrelation diagrams test and unit root test [8].

Data processing: If the series are not stationary, the difference methods which are first-order difference or d-order differences are used to generate a stationary time series. Then, its stationarity is checked again. When the sequence becomes stabilized, the data processing is over.

The selection of model order: According to Akaike Information Criterion (AIC) criterion [9], the parameters of the model are determined.

Model test: The residual series of the model are tested by a white noise test. If it is white noise series, the established model is effective. The methods of establishing include QQ diagrams test, LB test and D-W test.

Model prediction: The calculated model is used to predict the battery capacity.

### 3.3. Particle Filtering

#### 3.3.1 particle filtering algorithm

The particle filtering [10, 11] is a sequential Monte Carlo method based on recursive Bayesian estimation.

The principle is utilizing the set of weighted samples to represent the probability densities, which are available for any state-space model. As the sample size ($N_s$) becomes very large, the characteristics of Monte Carlo (MC) become the equivalent of the post-pdf functional description. [12, 13] Thus, $\{x^i_{0:k}, w^i_k\}$ represents the random metric that characterizes the posterior, where $\{x^i_{0:k}, i = 0, ..., N_s\}$ is a set of associated weights point. The weights are normalized as $\sum w^i_k = 1$. Thus, the posterior filtered can be approximated as
\[ p(x_k|y_{1:k}) \approx \sum_{i=1}^{N} w_k^i \delta(x_k - x_k^i) \]  

(6)

The normalized weights \( w_k^i \) are based on the principle of importance sampling. [14, 15]

Through recursive relation, the weights are given by

\[ w_k^i \propto w_{k-1}^i \frac{p(z_k|x_k^i)p(x_k^i|x_{k-1})}{q(x_k^i|x_{k-1}, z_k)} \]

(7)

The effective sample [16] size \( N_{eff} \) is introduced for resampling in order to solve degeneracy of the particles

\[ N_{eff} \approx \frac{1}{\Sigma_{i=1}^{N} \left( \frac{1}{w_k^i} \right)^2} \]

(8)

Resampling is performed when \( N_{eff} \) is lower than the threshold \( N_T \).

3.3.2 Battery degradation model

Battery capacity \( Q \) can be expressed by discharge current \( I \) and full discharge time \( t \):

\[ Q = \int I \, dt \]

(9)

When lithium-ion batteries are used, the capacity of the batteries becomes smaller and smaller. If it falls to a threshold, the battery is failure and it needs to be replaced.

The first step is to build a degraded model of the battery [17, 18]. Models can generally be set up based on known system physics knowledge and experience. Suppose the battery’s rated capacity \( Q_{rated} = 1 \) and define \( Q = Q_{rated} \cdot \hat{Q}/Q_{rated} \). Then the experience degradation equation for the capacity of lithium-ion batteries [19,20] is as following

\[ Q = a \cdot \exp(b \cdot k) + c \cdot \exp(d \cdot k) \]

(10)

Where \( a, b, c, d \) are the model parameters; \( k \) is a variable of the model that represents the number of charge and discharge cycles.

4. EXPERIMENT PROCESS AND RESULT ANALYSIS

4.1. Experiment Data and Process

The data set of lithium battery aging used in the work are obtained from NASA Research Center [21]. Then these data are used to verify the fusion algorithm. 18650 battery with 2 Ah rated capacity are used to obtain the data.

Test process:

the experimental temperature is set to 24 °C.

Charge step: firstly, charging is conducted at constant current (CC) 1.5 A until the charge voltage reaches 4.2 V. Then, charging is turned into constant voltage (CV) mode until the charge current drops to 20 mA, and the charging is completed.

Discharge step: discharging is conducted at constant current (CC) 2 A until the discharge voltage reaches predefined cutoff voltages are 2.7, 2.5, 2.2.
NASA battery 5, battery 6 and battery 7 are shown in Figure 1.

![Figure 1. NASA battery data](image)

4.2. Battery Capacity Prediction

The predictions are classified into two stages in this work: short-term predictions and long-term predictions. The experiments terminate when the battery capacity decreases to 70%. In this work, 1.38 Ah is set as the threshold value of the life of lithium battery. The data of lithium battery 5, battery 6 and battery 7 are selected as an example.

Short-term prediction is used to predict available capacity of battery, the ten battery capacities of the latest charge and discharge cycles are applied to predict the next battery capacity of charge and discharge cycles by ARIMA model. After that, the ten battery capacities of the latest charge and discharge cycles are used to predict the capacity again. When the battery lifetime is finished, the prediction is over.

Long-term prediction is used to predict the RUL of battery, the ten battery capacities of the latest charge and discharge cycles are applied to establish battery degrading model by PF algorithm, so the track of battery capacity degradation is predicted. At last the RUL of battery is obtained.

4.3. Short-term Predictions

4.3.1 experimental process

In the prediction of future battery capacity, there are three sets of experimental data which are collected from battery 5, battery 6 and battery 7. There are 168 cycle data in each group, and ten successive data in each group is chosen to predict the 11th battery capacity $u_{p(t+1)}(0<t<159)$.

The experiment programming software is Jupyter Notebook in Anaconda3 [22].

At first, the data stationarity is judged by its stationarity of AR(p). There are various methods to test the stationarity of model, such as autocorrelation function, run test, unit root test and parameter test, etc, and the unit root test is used in this work. In python, the Augmented Dickey-Fuller (ADF) [23] function is used to analyze a set of observed values which are selected, then the returned value can be obtained. If returned $p$ is greater than the significant level parameter
(0.05), the original series is unstable, then the difference is used to deal with the time series. If the p is less than the significant level parameter $\alpha$ (0.05), the series is stable.

In the first group of raw data, the result of ADF test is $p=0.8815268242790534$. The result of first-order difference is $p=0.01624398516948173$. With the 158th group of raw data, the result of the 158th of raw data by ADF test is $p=0.33779948953080063$. The result of first-order difference is $p=0.1638246870802363$, the result of second-order difference is $p=5.2645295116266845e-09$.

According to the result of ADF test, the $p$ in raw data is greater than 0.05, the series is non-stationary. In the first group, if the result of first-difference is less than 0.05, the series is stationary after first-order difference. Then the $d$ is set to 1. In the 158th group, it is stationary after second-order difference. Then the $d$ is set to 2.

It is different that the parameters of ARIMA model are established by different groups. Because a large amount of data needs to be fitted, the AIC algorithm is used directly to determine the optimal $p$ and $q$.

AIC algorithm is used to obtain the $p$ and $d$ in the first group of data, which are as shown in Table 1.

<table>
<thead>
<tr>
<th>Name</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-46.204831</td>
<td>-46.494072</td>
<td>NaN</td>
<td>NaN</td>
</tr>
<tr>
<td>1</td>
<td>-44.874499</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
</tr>
<tr>
<td>2</td>
<td>-43.858611</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
</tr>
<tr>
<td>3</td>
<td>-42.858613</td>
<td>NaN</td>
<td>NaN</td>
<td>NaN</td>
</tr>
</tbody>
</table>

AIC algorithm is used to obtain the $p$ and $q$ in the 158th group of data, which are as follows:

<table>
<thead>
<tr>
<th>Name</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-39.789235</td>
<td>-44.231525</td>
<td>NaN</td>
</tr>
<tr>
<td>1</td>
<td>-45.063187</td>
<td>NaN</td>
<td>NaN</td>
</tr>
<tr>
<td>2</td>
<td>-43.207172</td>
<td>NaN</td>
<td>NaN</td>
</tr>
</tbody>
</table>

Then the optimal values of $p$ and $q$ are determined by AIC algorithm. The results of the first group are $p=0$ and $q=1$. The results of the 158th group are $p=1$ and $q=0$.

4.3.2 Establishment of ARIMA model and residual error test

The ARIMA model function is used for three sets of data of batteries in python. Then the $p$, $d$ and $q$ are determined by the above steps which are used to make predictions in ARIMA. The battery5, battery6 and battery7 serial observed values of three sets of data are used to establish the ARIMA model by python. The model is tested by the residual test [24] to ensure the ARIMA model residual is the Gaussian white noise by LB test. Otherwise, then the ARIMA model may not be a suitable model for samples. Meanwhile, the model is established by increasing the number of observed values or changing the difference, then the $\mu_{p,q (+1)}$ is predicted by new model.

The result of first group data by residual test is $p=0.61714754$. The result of 158th group data by residual test is $p=0.52644526$. The results of two groups data by LB test are greater than the
significant level parameter $\alpha (0.05)$, so the residual is determined as the Gaussian white noise [25]. The results prove that the data adapt ARIMA model.

4.3.3 Experimental evaluation and error analyzing

After the residual test is finished, the model is used to predict a future discharge capacity $u_p(t+1)$. According to the loop algorithms, 10 sets of data are chosen to establish ARIMA model by each time $t$ adding one and then the future $u_p(t+1)$ is predicted.

A total of 158 data are predicted, and the results of using python are shown in Figure 2-4, and given as follows.

![Figure 2. Battery 5 capacity prediction procedure](image1)

![Figure 3. Battery 6 capacity prediction procedure](image2)
The measured values and the predicted values are almost match. In order to confirm the effectiveness, the Mean Absolute Error (MAE) [26] and Maximum Mean Absolute Error (MMAE) can be calculated to evaluate the performance of ARIMA as follows:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |u_p - u_i|$$ \hspace{1cm} (11)

$$\text{MMAE} = \text{MAX} \left( \frac{1}{n} \sum_{i=1}^{n} |u_p - u_i| \right)$$ \hspace{1cm} (12)

**Table 3. Battery measurement error**

<table>
<thead>
<tr>
<th>The battery</th>
<th>Mean Absolute Error/Ah</th>
<th>Maximum Mean Absolute Error /Ah</th>
</tr>
</thead>
<tbody>
<tr>
<td>Battery5</td>
<td>0.006871</td>
<td>0.093497</td>
</tr>
<tr>
<td>Battery6</td>
<td>0.011197631</td>
<td>0.103943488</td>
</tr>
<tr>
<td>Battery7</td>
<td>0.005769204</td>
<td>0.0404219</td>
</tr>
</tbody>
</table>

According to Table 3, the following results can be drawn.

Battery5’s maximum mean absolute error: is 0.093497 Ah and mean absolute error is 0.006871 Ah.

Battery6’s maximum mean absolute error: is 0.103943488 Ah and mean absolute error is 0.011197631 Ah.

Battery7’s maximum mean absolute error: is 0.0404219 Ah and mean absolute error is 0.005769204 Ah.
So the battery 6’s data is the biggest among them.

4.4. Batteries’ RUL Prognostic

4.4.1 Experimental process

When the particle filter algorithm predicts the RUL of lithium-ion batteries [27], B 5, B6, B7 batteries are predicted to start at the 68th cycle and the threshold at the end of battery life is 1.47 Ah. The number of particles is 500. The process noise co-variance $Q$ is 0.0001, and the noise co-variance $R$ is observed to be 0.0001. The value of $a,b,c$ and $d$ are shown in Table 4.

<table>
<thead>
<tr>
<th>Battery</th>
<th>$a$</th>
<th>$b$</th>
<th>$c$</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Battery 5</td>
<td>1.830331</td>
<td>-0.002687</td>
<td>0.098931</td>
<td>0.001694</td>
</tr>
<tr>
<td>Battery 6</td>
<td>1.830302</td>
<td>-0.003867</td>
<td>0.1258354</td>
<td>0.002094</td>
</tr>
<tr>
<td>Battery 7</td>
<td>1.830308</td>
<td>-0.002498</td>
<td>0.119602</td>
<td>0.001824</td>
</tr>
</tbody>
</table>

The data of battery 5, battery 6 and battery 7 are tracked by the particle filter, and the result are showed in Figure 5-7. The red horizontal line represents the failure threshold of the battery. The blue dotted line represents the experimental measurement data for the battery. The green dotted line represents the estimated value by using the particle filtering algorithm. After 68 cycles, the prediction algorithm starts to execute. As can be seen from Figure 5-7, the value obtained by the particle filtering algorithm is very close to the experimental measurement value, which shows its effectiveness in state tracking.

![Figure 5. Battery 5 RUL prediction procedure](image-url)
Figure 6. Battery 6 RUL prediction procedure

Figure 7. Battery 7 RUL prediction procedure

Table 5. RUL prediction results of particle filter algorithm

<table>
<thead>
<tr>
<th>NASA experiment</th>
<th>Start</th>
<th>Eop</th>
<th>Eol</th>
<th>RUL_prediction</th>
<th>Rul_error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Battery 5</td>
<td>68</td>
<td>115</td>
<td>106</td>
<td>47</td>
<td>9</td>
</tr>
<tr>
<td>Battery 6</td>
<td>68</td>
<td>86</td>
<td>84</td>
<td>18</td>
<td>2</td>
</tr>
<tr>
<td>Battery 7</td>
<td>68</td>
<td>133</td>
<td>139</td>
<td>65</td>
<td>6</td>
</tr>
</tbody>
</table>
Prediction results are recorded in Table 5, "Start" represents the predict period during the experiment, "EoP" (end of the prediction period) represents the end of the predicted life, and "EoL" (end of life) represents the true life of the battery used in the experiment. $RUL_{erro}$ represents the absolute error value of the RUL prediction.

$$RUL_{erro} = |RUL_{prediction} - RUL_{true}|$$

Where $RUL_{prediction}$ represents the remaining cycle life value of the predicted lithium-ion battery, and $RUL_{true}$ represents the remaining true cycle life of the battery after the forecast starting point is set.

5. CONCLUSION

Experimental results have shown that the ARIMA model predicts the discharge capacity of a group future values well. On account of establishing the time series model must depend on the historical series of battery observed values, the results of the discharge capacity being simulated have large error by comparing with the measured values. But the total average error is small. What's more, the simulated values have larger error compared with several observed values of suddenly changing. The mean absolute errors of battery5, battery6 and battery7 are 0.006871 Ah, 0.011197631 Ah, and 0.005769204 Ah, respectively. The maximum mean absolute errors of battery5, battery6 and battery7 are 0.093497 Ah, 0.103943488 Ah, and 0.0404219 Ah, respectively. The remaining life of the lithium-ion battery was effectively predicted using Particle filtering. The $RUL_{erro}$ of battery B5, B6 and B7 by PF algorithm are 9 cycles, 2 cycles and 6 cycles respectively.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China under Grant No. 51607004, Collaborative Innovation Project of Anhui Universities, No.GXXT-2019-002, Natural Science Research Key Project of Education Department of Anhui Province No.KJ2018A0369 and College Students Innovation and Entrepreneurship Training Program of Anhui Province No.S202010372125X.

REFERENCES


