

# A Copula-based battery pack consistency modeling method and its application on the energy utilization efficiency estimation

Yan Jiang<sup>a, b</sup>, Jiuchun Jiang<sup>a, c</sup>, Caiping Zhang<sup>a, \*</sup>, Weige Zhang<sup>a</sup>, Yang Gao<sup>a</sup>, Chris Mi<sup>b, \*\*</sup>

<sup>a</sup> National Active Distribution Network Technology Research Center (NANTEC), Beijing Jiaotong University, Beijing, 100044, China

<sup>b</sup> Department of Electrical and Computer Engineering, San Diego State University, San Diego, CA, 92182, USA

<sup>c</sup> Shenzhen Precise Testing Technology CO.,LTD., Shenzhen, 518100, China

## ARTICLE INFO

### Article history:

Received 30 April 2019

Received in revised form

5 September 2019

Accepted 25 September 2019

Available online 25 September 2019

### Keywords:

Battery pack

Consistency modeling

Energy utilization efficiency

Copula

## ABSTRACT

The consistency of battery cells directly influences the maximum available energy and the efficiency of the battery pack, and the energy utilization efficiency (EUE) is a key parameter for the balancing of batteries. Therefore, this paper focuses on the consistency modeling and state estimation of battery packs. In this study, a Copula-based battery pack consistency modeling method is developed. The proposed method shows superiority compared with two existing methods, because it can describe the statistical characteristics of the battery consistency parameters, and the dependence structure between parameters. The squared Euclidean distances between the marginal empirical cumulative distribution functions of the test data and that of the proposed model for capacity, resistance, and SOC are 0.029, 0.169, and 0.025, respectively. The errors of the correlation coefficients between the proposed model and the test data are within 0.12. Then the framework of battery pack EUE estimation using the consistency model is proposed. The accuracy of the proposed method is verified based on the test results of a battery pack with 95 cells connected in-series. The EUE estimation error is within 0.6% at various discharge current rates. The EUE estimation results could provide support for the performance evaluation and balancing of battery packs.

© 2019 Published by Elsevier Ltd.

## 1. Introduction

In recent years, lithium-ion batteries have been widely used as energy storage elements in energy storage systems (ESSs) and electric vehicles (EVs), because of their high energy density, high power density, high efficiency, and long service life [1–3]. Owing to voltage and power limits, several hundred or more lithium-ion battery cells are sometimes connected in-parallel and in-series in the applications [4]. In a battery pack, because of the non-uniform thermal field or the different locations where the battery cells have been placed [5,6], the battery cells in the battery pack may have different aging patterns, causing the battery cells to have different capacity, internal resistance, and state of charge (SOC) during the aging process. As a result, the lifetime of the battery pack is much shorter than that of a single battery cell.

To evaluate the effect of battery inconsistency on the

performance of a battery pack, a battery pack consistency model is needed. One of the most important ways to model the battery pack consistency is to depict the statistical characteristics of battery consistency parameters. Normally, probability distribution fitting methods are used to describe the probability distribution functions (PDFs) of such parameters. In Ref. [4], the authors compared the statistical characteristics of new and aged battery cells in terms of the battery capacity and impedance. The test results showed that the capacity and impedance of new battery cells are normally distributed. Along with the aging of the battery pack, the distributions are altered from normal to Weibull distributions. In Ref. [7], the researchers tested a retired lithium-ion battery pack. The test results showed that Weibull and normal distributions are more suitable to describe the battery capacity distribution and internal resistance distribution, respectively. As for the SOC of the battery cells, a skewed distribution is exhibited. In Ref. [8], the researchers tested retired LiFePO<sub>4</sub> battery cells and reported that both the battery capacity and internal resistance are normally distributed. In general, battery consistency parameters may have different types of distributions under different situations. However, in the existing literature, only normal distributions are used to describe the

\* Corresponding author.

\*\* Corresponding author.

E-mail addresses: [zhangcaiping@bjtu.edu.cn](mailto:zhangcaiping@bjtu.edu.cn) (C. Zhang), [cmi@sdsu.edu](mailto:cmi@sdsu.edu) (C. Mi).

parameter distributions when evaluating the effects of battery inconsistency on the performance of the battery pack [9–11]. An incorrect modeling of the battery consistency parameters will result in an inaccurate estimation of the state of the battery pack. Another aspect of battery pack consistency modeling is the description of the correlation among the battery consistency parameters, because such parameters are always coupled [11]. In Ref. [12], the correlation between battery capacity variations and SOC variations was studied. It was found that a loss of variation in the lithium inventory at the anodes of the battery cells is the dominating factor causing SOC variations of the battery cells. The correlation between battery coulombic and capacity fade was studied in Ref. [13]. The low coulombic efficiency not only causes the batteries to incur a faster rate of capacity fade, it also increases the SOC inconsistency among the battery cells. The coupling mechanisms between battery capacity fade, and the increase in the internal resistance of the battery, were shown in Refs. [14,15]. Normally, the correlation coefficient between the battery capacity and internal resistance is within the range of  $-0.2$  and  $-0.8$  [7,16,17]. Owing to the extremely complex coupling mechanisms among the parameters, these parameters are assumed mutually independent in existing battery pack consistency models. Therefore, an effective battery pack consistency modeling method that can describe both the statistical characteristics of the battery consistency parameters and the correlation between the parameters is needed.

With the battery pack consistency model, the state of health (SOH) of the battery pack can be estimated. The battery pack SOH indicators can either be defined as the battery pack capacity or the battery pack internal resistance [11,18–20]. In Ref. [18], the battery pack capacity is defined as the minimum capacity of the battery cells. Considering the SOC variations of the battery cells, researchers in Refs. [11,21] further defined the battery pack capacity as the sum of the minimum remaining available capacity and the minimum chargeable capacity of the cells. The authors of [19,20] developed simple and effective methods to estimate the resistances of the battery cells. However, these two battery pack SOH indicators are not practical for a battery pack energy management (to a certain extent), because they do not reflect the effect of battery cell inconsistency on the performance of the battery pack. The energy utilization efficiency (EUE) is used as a battery pack SOH indicator in Refs. [9,22]. The advantage of this indicator is that it can be used for an equilibrium diagnosis of the battery pack, because the inconsistency of the battery cell directly affects the energy state of the battery pack itself. However, the proposed EUE calculation method in Ref. [22] is only suitable for small current applications, and the accuracy of the methods in Refs. [9,10] has not been verified. Therefore, an accurate battery pack EUE estimation method needs to be studied.

Facing the two challenges mentioned above, this study provides the following original contributions. First, a Copula-based battery pack consistency model is developed, which realizes the description of the statistical characteristics of the battery consistency parameters and the correlation between parameters simultaneously. Second, an accurate battery pack EUE estimation method is developed. Given the battery pack consistency model, the statistical characteristics of the battery pack EUE are obtained using the Monte Carlo (MC) method. The expected value of the EUE simulation results is then used as the estimated EUE. The remainder of this paper is organized as follows. Section 2 illustrates the processes used in the battery pack consistency modeling. Section 3 introduces the battery pack EUE estimation method. The battery tests conducted are introduced in Section 4. The superiority of the proposed battery pack consistency modeling method and the accuracy of the battery pack EUE estimation method are discussed in Section 5.

Finally, some concluding remarks are provided in Section 6.

## 2. Battery pack consistency modeling

### 2.1. Definition of Copula

The Copula is defined as the joint cumulative distribution function of standard uniform random variables, and is used to describe the dependence between random variables. According to Sklar's theorem, every multivariable distribution function

$$H(x_1, \dots, x_m) = P(X_1 \leq x_1, \dots, X_m \leq x_m) \quad (1)$$

of a random vector  $(X_1, X_2, \dots, X_m)$  can be expressed in terms of its marginal distribution functions and a Copula  $C$ . Normally, it is expressed in the following form:

$$H(x_1, \dots, x_m) = C(F_1(x_1), \dots, F_m(x_m)) \quad (2)$$

Therefore, if given the Copula and the marginal distributions of the input data, a set of random data can be generated that have the same marginal distributions and structural dependence as those of the input data.

### 2.2. The optimal Copula selection

The selection of an optimal Copula is crucial for accurately modeling the joint cumulative distribution function (CDF) of the battery consistency parameters. In this section, an analytical method used to select the optimal Copula is introduced. Denote  $(x_1^i, x_2^i, \dots, x_m^i)$ ,  $i = 1, \dots, n$  as the observations from a random vector  $(X_1, X_2, \dots, X_m)$ . Then, the processes to determine the optimal Copula are as follow:

Step 1: Determine the empirical CDFs  $F_k^n$ ,

$$F_k^n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(x_k^i \leq x) \quad (3)$$

where  $\mathbf{1}$  is the indicator function of the events. Taking event  $(x_k^i \leq x)$  as an example,  $\mathbf{1}(x_k^i \leq x)$  is defined as follows:

$$\mathbf{1}(x_k^i \leq x) = \begin{cases} 1 & \text{if } x_k^i \leq x \\ 0 & \text{if } x_k^i > x \end{cases} \quad (4)$$

Step 2: Determine the empirical Copula corresponding to the observations,

$$\widehat{C}(u_1, \dots, u_m) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(F_1^n(x_1^i) \leq u_1, \dots, F_m^n(x_m^i) \leq u_m) \quad (5)$$

where  $(u_1, \dots, u_m)$  are random variables within the range  $[0, 1]$ .

Step 3: Fit different types of Copulas to the observations and obtain

$$C_j(u_1, \dots, u_m) \quad (6)$$

where  $j$  is the number of Copulas that can be fitted.

Although there are numerous Copulas that can model the structural dependence of bivariate data, only a few can address multivariable data. Therefore, only a Gaussian Copula and t Copula

are used to fit the observations in this study. The unknown parameters of Copulas are estimated using the maximum likelihood estimation method [23,24].

Step 4: Calculate the squared Euclidean distance  $d_j^2$  between the empirical and fitted Copulas,

$$d_j^2 = \sum_{i=1}^n \left| \widehat{C}(u_1^i, \dots, u_m^i) - C_j(u_1^i, \dots, u_m^i) \right|^2 \quad (7)$$

where  $(u_1^i, u_2^i, \dots, u_m^i), i = 1, \dots, n$  are the standard uniform random variables transformed from  $(x_1^i, x_2^i, \dots, x_m^i)$  according to a probability integral transform.

Step 5: The optimal Copula has the minimum squared Euclidean distance  $d_j^2$ .

### 2.3. Copula-based sampling of the battery pack consistency model

Given the dataset of battery consistency parameters and the optimal Copula, a series of new data can be generated. The processes of generating new random data can be described as follows:

Step 1: Obtain the marginal empirical CDFs of the input data, and then obtain the marginal CDFs ( $F_k(x_k)$ ) using a kernel smoothing method.

Step 2: Transform the data into the Copula scale (standard uniform space) using the CDFs:  $u_k^i = F_k(x_k^i)$ .

Step 3: Determine the optimal Copula according to  $u$ .

Step 4: Generate a set of standard uniform distributed random samples  $u'$  from the optimal Copula.

Step 5: Transform a random sample back into the original scale of the data using the inverse CDFs of the input data:  $x_k' = F_k^{-1}(u_k')$ .

Therefore,  $(X_1', X_2', \dots, X_m')$  has the same marginal CDFs, as well as the dependence structure, as  $(X_1, X_2, \dots, X_m)$ .

## 3. Battery pack EUE estimation

### 3.1. Theoretical analysis

Denote  $Q_i$ ,  $R_i$ , and  $SOC_{start,i}$  as the battery capacity, internal resistance, and SOC of the  $i$ th battery cell in the battery pack consistency model mentioned in Section 2. It should be stated that  $SOC_{start,i}$  in this study denotes the SOC value of the  $i$ th battery cell when the battery pack is fully charged. To estimate the EUE of a battery pack, the Thevenin model is used, which is shown in Fig. 1.

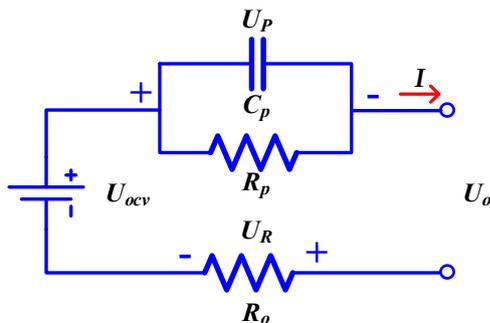


Fig. 1. Thevenin model of a battery cell.

The terminal voltage  $U_o$  is given by Eq. (8):

$$U_o = U_{ocv} - U_R - U_P \quad (8)$$

where  $U_{ocv}$  is the open circuit voltage of a battery cell, and  $U_R$  and  $U_P$  are the voltages across the ohmic resistance  $R_o$  and polarization resistance  $R_p$ , respectively. Let the internal resistance  $R$  denote the sum of the ohmic and polarization resistances. If the dynamic characteristics of the polarization voltage are neglected, the terminal voltage of a battery cell under a steady state can be estimated through the following:

$$U_o = U_{ocv} - I(R_o + R_p) = U_{ocv} - IR \quad (9)$$

where  $I$  denotes the discharge current.

The EUE of a battery pack is the ratio of the discharge energy of the battery pack at a certain discharge rate from battery pack SOC = 100%, to battery pack SOC = 0, to the sum of the maximum available energy of the battery cells in the pack. The mathematical expression of a battery pack EUE is shown in Eq. (10). The discharge energy of the battery pack is affected by the discharge current, because a portion of the energy is consumed by the internal resistances of the battery cells. Therefore, the EUE of a battery pack with  $N$  series-connected battery cells can be calculated using Eq. (11).

$$EUE = \frac{E_{pack\_dis}}{E_{pack\_max}} \quad (10)$$

$$EUE = \frac{\sum_{i=1}^N \int_{t_0}^{t_0+\Delta t} (U_{ocv,i}(t) \cdot I - R_i(t) \cdot I^2) dt}{\sum_{i=1}^N Q_i \cdot U_{av}} \quad (11)$$

where  $U_{av}$  is the average open circuit voltage during the time in which SOC changes from 100% to 0,  $\Delta t$  is the discharge time of the battery pack, and  $U_{ocv,i}(t)$  and  $R_i(t)$  are the change in the OCV and internal resistance of the  $i$ th battery cell over the discharge time  $t$ , respectively. Normally, it is assumed that the SOC-OCV curve of all battery cells are the same despite the battery cells having different capacities. Therefore, the relationship between OCV and SOC can be expressed as follows:

$$U_{OCV} = f(SOC) \quad (12)$$

Because all batteries are series-connected, the discharge time of the battery pack can be determined by finding the battery cell with the minimum residual electric quantity in the battery pack. Denote  $Q_{pack}$  as the battery pack capacity,  $\Delta SOC_i$  and  $SOC_{end,i}$  as the SOC range and SOC of  $i$ th battery cell when discharged to the cutoff voltage with current  $I$ , respectively. The discharge time of the battery pack is then given as follows:

$$\Delta t = \frac{Q_{pack}}{I} = \min_{1 \leq i \leq N} \frac{\Delta SOC_i \cdot Q_i}{I} = \min_{1 \leq i \leq N} \frac{(SOC_{start,i} - SOC_{end,i}) \cdot Q_i}{I} \quad (13)$$

### 3.2. Implementation of EUE estimation

As shown in Eq. (11), the internal resistance is expressed as a function of the discharge owing to the fact that the internal battery resistance varies at different SOC values [25,26]. In real applications, it is assumed that the internal resistance of a battery can only be tested when the battery cell SOC is at 50%. To establish the

relationship between the internal resistance of the battery and the SOC, the following is adopted:

$$R_i(SOC) = R_i|_{SOC=50\%} \cdot g(SOC) \quad (14)$$

where  $R_i(SOC)$  and  $R_i|_{SOC=50\%}$  are the internal resistance at different values of SOC, and the resistance tested at 50% SOC of the  $i$ th battery cell, respectively. For all battery cells, their  $g(SOC)$  are considered to be the same. Because  $g(SOC)$  can be obtained through offline measurements, the internal resistance of all battery cells over the entire SOC range can be determined if their resistances tested at 50% SOC are given. Then, the evolution of the internal resistance with discharge time can be expressed as Eq. (15):

$$R_i(t) = R_i|_{SOC=50\%} \cdot g\left(SOC_{start,i} - \frac{I \cdot t}{Q_i}\right) \quad (15)$$

In addition, the discharge time of the battery pack relies on the estimation of  $SOC_{end,i}$  in Eq. (13). Denote  $OCV_{end,i}$  as the OCV of the  $i$ th battery cell after discharging to the cutoff voltage. The  $SOC_{end,i}$  can then be determined using the inverse function of Eq. (16):

$$SOC_{end,i} = f^{-1}(OCV_{end,i}). \quad (16)$$

Here,  $OCV_{end,i}$  is related to the internal resistance, discharge current rate, and the discharge cutoff voltage of the battery. Because the discharge cutoff voltage is a fixed value,  $OCV_{end,i}$  can be expressed as follows:

$$OCV_{end,i} = h(I, R_i|_{SOC=50\%}). \quad (17)$$

Based on the derivations above, the simplified functions of the EUE estimation and the discharge time of the battery pack are obtained, which are shown in Eqs. (18) and (19), respectively.

$$EUE = \frac{\sum_{i=1}^N \int_{t_0}^{t_0+\Delta t} \left( U_{ocv,i}(t) \cdot I - R_i|_{SOC=50\%} \cdot g\left(SOC_{start,i} - \frac{I \cdot t}{Q_i}\right) \cdot I^2 \right) dt}{\sum_{i=1}^N Q_i \cdot U_{av}} \quad (18)$$

$$\Delta t = \min_{1 \leq i \leq N} \frac{\left( SOC_{start,i} - f^{-1}\left(h\left(I, R_i|_{SOC=50\%}\right)\right) \right) \cdot Q_i}{I} \quad (19)$$

It should be mentioned that  $EUE$  in Eq. (18) is a random variable with a complicated PDF. Therefore, the numerical results of the EUE cannot be solved through a mathematical derivation. To address this problem, a Monte Carlo (MC) simulation is used. The idea of an MC simulation is to repeat the simulation numerous times to obtain the statistical characteristics of the model outputs. As for the model inputs for each simulation, they can be obtained through random sampling according to the probability distributions of the parameters. Therefore, the MC can propagate the input uncertainties into the output uncertainties [27,28]. To express the EUE simulation results, the expected value of  $EUE$  is used. With  $m$  simulations, the estimated value of the EUE can be calculated using Eq. (20):

$$EUE = \frac{1}{m} (EUE_1 + \dots + EUE_m). \quad (20)$$

## 4. Battery tests

### 4.1. Battery pack and battery cell test

To verify the proposed battery pack consistency modeling method and its application on a battery pack EUE estimation, a retired LiFePO<sub>4</sub> lithium-ion battery is tested. The battery pack had been used in a pure electric passenger vehicle for more than three years. The battery pack consists of 95 series-connected battery cells, and its rated capacity is 60 Ah. After the battery pack was disassembled from the vehicle, the battery pack underwent a series of tests to determine the cell-to-cell variations of the battery cells. First, a battery pack rate capability test was conducted. The battery pack was first charged at 0.1C, and then discharged at 0.1C, 0.3C, 0.7C, and 1.2C. After testing the rate capability of the battery pack, the battery pack was disassembled into 95 individual battery cells. Battery cell capacity and battery cell internal resistance tests were then conducted. In the battery cell capacity test, all battery cells were charged and discharged at 0.1C within the cutoff voltages (3.65 V and 2.5 V for the charge and discharge processes, respectively). An internal resistance test of the battery cells was conducted when the SOC of each battery cell was 50%. During the test, the battery cells were discharged at 0.5C for 10 s. The ratio of the voltage change to the discharge current during this period is defined as the internal resistance of the battery cell. Based on the results of the battery pack rate capability and the battery cell capacity tests, the SOC variations of the battery cells were also determined. A detailed description of the experiment processes was introduced in our previously published paper [7].

### 4.2. Complementary tests

To make the EUE estimation more accurate, 5 of the 95 battery

cells were selected to conduct the following three complementary experiments.

#### (1) SOC-OCV measurement

From SOC = 100% to SOC = 0, the five battery cells were discharged by 5% SOC each time and set to rest for 2 h. The OCV was then measured after the rest time. Fig. 2(a) shows the SOC-OCV curves of the five battery cells. The five curves exhibit little

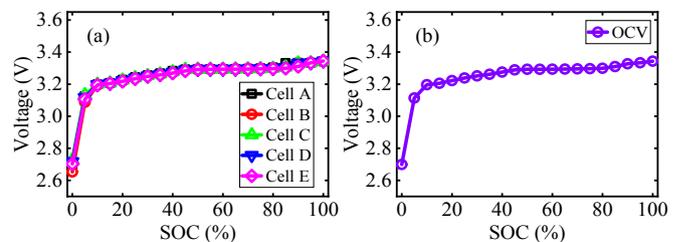


Fig. 2. (a) SOC-OCV curves of the five selected battery cells and (b) unified SOC-OCV curve used in the EUE estimation.

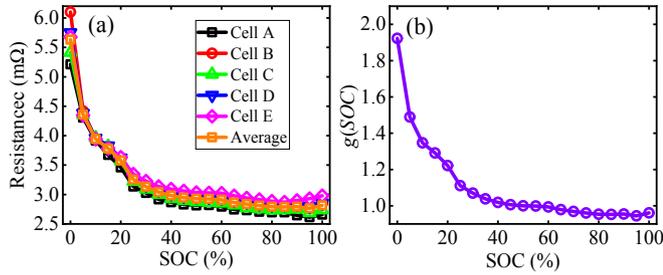


Fig. 3. (a) The internal resistances of the five selected battery cells over the whole SOC range and their average curve and (b) the evolution of  $g(SOC)$  over the entire SOC range.

difference over the entire SOC range. The unified SOC-OCV curve used in the EUE estimation is then obtained by averaging the five curves, as shown in Fig. 2(b).

(2) Internal resistance tests of battery cells over the entire of SOC range

Internal resistance tests were conducted on the five battery cells over the entire SOC range. From SOC = 100% to SOC = 5%, the battery resistances were tested using the same method described in Section 4.1. To obtain the resistances at 0% SOC, a 10 s charge constant current pulse was applied. The resistances of this charging process could be calculated, and were considered the resistances at 0% SOC. Fig. 3(a) shows the resistances of the five battery cells and their average resistance curve along with the change in the battery SOC. The internal resistance of the battery increases quite slowly as the SOC decreases from 100% to 40%. However, when the SOC is lower than 40%, the internal resistance increases significantly, and the resistance almost doubles when the SOC approaches zero. Then,  $g(SOC)$  is obtained according to the average resistance curve and Eq. (14). The result is shown in Fig. 3(b), and this piecewise linear curve is used to estimate the battery resistances at other SOC values if their resistances measured at 50% SOC are given.

(3) OCV measurement of battery cells after the discharge process

As shown in Eq. (17), the OCV is a function of the discharge current and the internal resistance. To obtain this function, the five battery cells were discharged to a lower cutoff voltage of 2.5 V, from 0.1C to 1.3C. Then, the OCV of the battery cells after a 2 h rest were measured. Based on the test results, Eq. (17) is obtained by building a response surface between the OCV of the battery cells after discharge to the cutoff voltage, internal resistances of the battery at 50% SOC, and the discharge current rates. The response surface is shown in Fig. 4. The OCV at a specific discharge current rate can then be estimated through an interpolation if the internal resistance at 50% SOC is given.

5. Validation and discussion

5.1. Copula-based battery pack consistency model

Fig. 5 shows the histograms of the battery capacity, internal resistance, and SOC of the battery cells mentioned in Section 4.1. The distribution of the battery capacity appears normal because the histogram of the battery capacity is almost symmetrical. However, the histograms of the internal resistance of the battery and the SOC are asymmetrical, namely, the former is right-skewed and the latter is left-skewed.

To show the statistical dependence between consistency

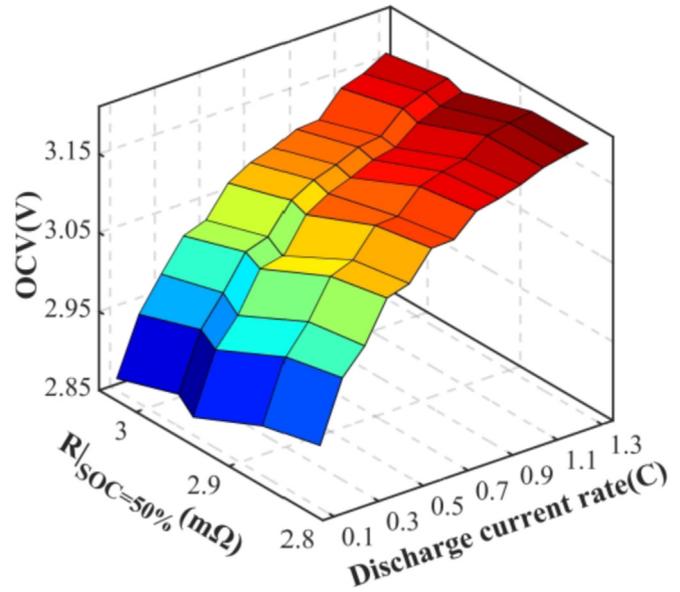


Fig. 4. The response surface of battery OCV after discharge to the cutoff voltage with respect to the internal resistance of the battery at 50% SOC and the discharge current rate.

parameters of the battery cells, a scatter plot among parameters is presented in Fig. 6(a)-(c). In Fig. 6(a) and (c), the correlation between parameters is weak because the points are discrete. However, there is a significant correlation between the battery capacity and the SOC in Fig. 6(b), meaning the initial SOC of the battery increases as the battery capacity increases.

Through the steps described in Section 2, a three-dimensional battery pack consistency model is established, as shown in Fig. 7.

5.2. Comparison of the proposed battery pack consistency model and the traditional models

To show the advantages of the Copula-based battery pack consistency modeling method, the proposed battery pack consistency model is compared with two other battery pack consistency models built without Copula.

Battery pack consistency model #1 (Model #1):

In battery pack consistency model #1, the model is built using an inverse transform sampling method. To describe the distributions of the battery consistency parameters, the cumulative distribution functions are used. Let  $(x_1^i, x_2^i, \dots, x_m^i), i = 1, \dots, n$  be the  $n$  observations. In general, the processes of battery pack consistency modeling can be concluded through the following steps:

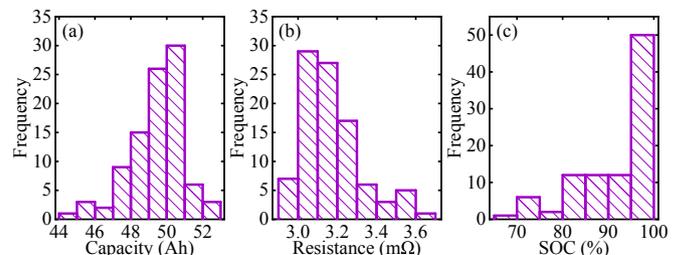


Fig. 5. Histograms of battery consistency parameters: (a) battery capacity, (b) internal resistance of the battery, and (c) the SOC.

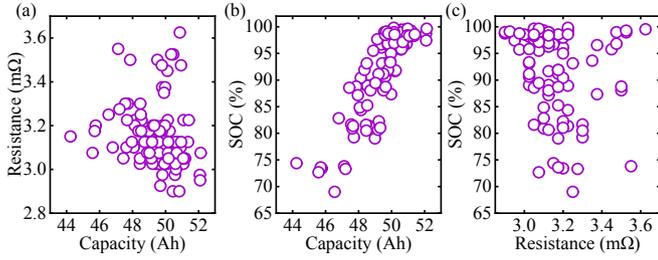


Fig. 6. Scatter plots of battery consistency parameters: (a) capacity vs. resistance, (b) capacity vs. SOC, and (c) resistance vs. SOC.

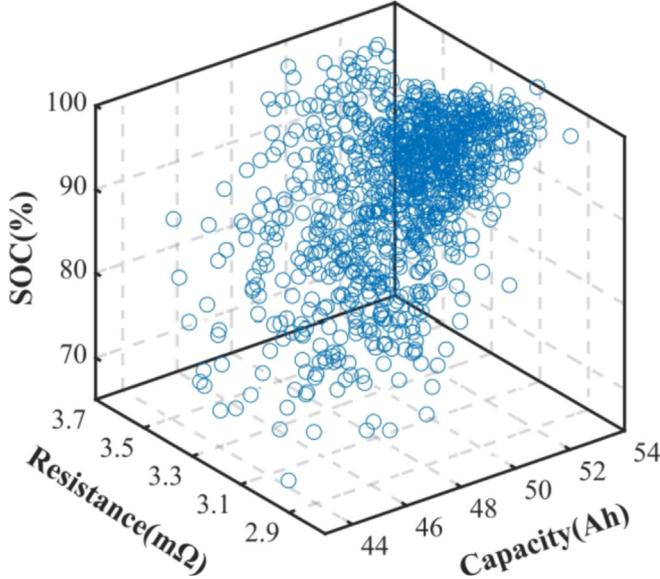


Fig. 7. The 3-dimensional battery pack consistency model.

- Step 1: Generate a set of random numbers  $u' = [u_1^i, \dots, u_m^i]$  from the standard uniform distribution in the interval  $[0, 1]$ .
- Step 2: Find the empirical CDFs  $F_k^n(x_k)$  of  $(x_1^i, x_2^i, \dots, x_m^i)$  and the related inverse empirical CDFs  $F_k^n^{-1}(x_k)$ .
- Step 3: Calculate  $x_k' = F_k^{-1}(u_k')$  and obtain  $(x_1', \dots, x_k')$ .

Although the battery pack consistency modeling processes of Model #1 is similar to that in the Copula-based modeling processes (because they both use an inverse transform sampling method), the following two differences between them should be noted:

- (1) In the Copula-based model, the CDFs of the battery pack consistency parameters are used. The CDFs are obtained according to the empirical CDFs and the kernel smoothing methods. However, in Model #1, only the empirical CDFs of the battery consistency parameters are used. The reason for using CDFs instead of empirical CDFs in the Copula-based modeling method is that only the CDFs can transform the data into standard uniformly distributed data, which is required by the definition of Copula.
- (2) The random variables  $u'$  in the Copula-based model are generated according to the optimal Copula but  $u'$  in Model #1 is generated randomly.

Battery pack consistency model #2 (Model #2):

Previously [9,10], the capacity distribution, internal resistance distribution, and SOC distribution were described using three

independent normal distributions. This battery pack consistency model can be further improved as a multivariable normal distribution of a three-dimensional vector  $\mathbf{X} = [Q_i, R_i, SOC_i]^T$ , which is written in the following form:

$$\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (21)$$

using the three-dimensional mean vector

$$\begin{aligned} \boldsymbol{\mu} &= E(\mathbf{X}) \\ &= [E(Q_i), E(R_i), E(SOC_i)]^T \\ &= [\mu_Q, \mu_R, \mu_{SOC}]^T \end{aligned} \quad (22)$$

and  $3 \times 3$  covariance matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_Q^2 & \rho_{1,2}\sigma_Q\sigma_R & \rho_{1,3}\sigma_Q\sigma_{SOC} \\ \rho_{1,2}\sigma_Q\sigma_R & \sigma_R^2 & \rho_{2,3}\sigma_R\sigma_{SOC} \\ \rho_{1,3}\sigma_Q\sigma_{SOC} & \rho_{2,3}\sigma_R\sigma_{SOC} & \sigma_{SOC}^2 \end{bmatrix} \quad (23)$$

where  $\mu_Q, \mu_R, \mu_{SOC}$  are the mean values of the battery capacity, the internal resistance, and the SOC;  $\sigma_Q, \sigma_R, \sigma_{SOC}$  are the standard deviations of the battery capacity, the internal resistance, and the SOC; and  $\rho_{1,2}, \rho_{1,3},$  and  $\rho_{2,3}$  are the correlation coefficients between the capacity and resistance, the capacity and the SOC, and the resistance and the SOC, respectively.

The probability density function of this three-dimensional multivariable normal distribution is as follows:

$$\begin{aligned} y &= f(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) \\ &= \frac{1}{\sqrt{|\boldsymbol{\Sigma}|(2\pi)^3}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right) \end{aligned} \quad (24)$$

where  $|\boldsymbol{\Sigma}|$  is the determinant of the  $\boldsymbol{\Sigma}$ . Then the battery pack consistency model can then be built using this probability density function.

Three battery pack consistency models (with 1000 sampling points) are built using the same battery test data from Section 4. The three models are then compared in terms of the statistical characteristics of the battery consistency parameters and the dependence structure between these parameters.

Fig. 8 shows the empirical CDFs of the battery capacity, internal resistance, and SOC of the test data and those of the three battery pack consistency models. The CDFs of the test data and Model #1 are almost the same because the sampling points in Model #1 are generated according to the inverse empirical CDFs of the test data. When comparing the CDFs of Copula-based model and Model #2 with those of the test data, the three curves are highly overlapped in Fig. 8(a). In Fig. 8(b) and (c), the difference between the CDF of the Copula-based model and that of the test data is quite small, whereas significant differences exist between the CDF of Model #2 and that of the test data.

Fig. 8 shows the modeling results for a single sampling process. To achieve a quantitative evaluation of the three models, the squared Euclidean distances between the marginal empirical CDFs are calculated according to the following:

$$d^2 = \sum_{i=1}^n \left| F_{k0}^n(x_k^i) - F_k^n(x_k^i) \right|^2, \quad (25)$$

where  $F_{k0}^n(x_k^i)$  is the marginal empirical CDF of the test data and  $F_k^n(x_k^i)$  is the marginal empirical CDF of a specific battery pack consistency model. To eliminate the randomness in the sampling

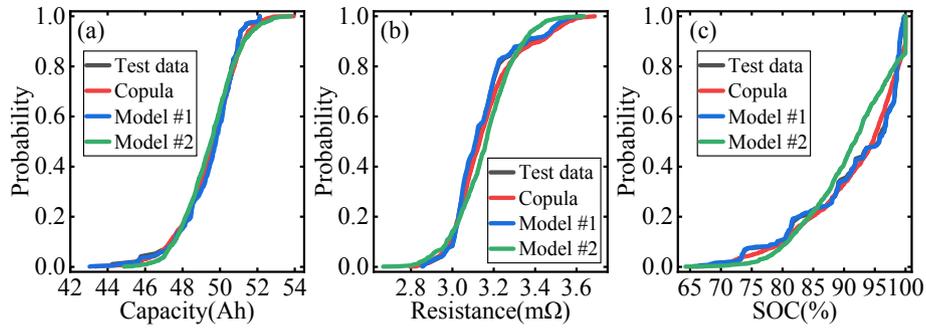


Fig. 8. Empirical CDFs of (a) battery capacity, (b) internal resistance, and (c) SOC of the test data and three battery pack consistency models.

process, the sampling processes for each model are repeated 1000 times and their average value is regarded as the final calculation result. The calculation results of the squared Euclidean distance between the marginal empirical CDFs are shown in Table 1. Model #1 has the smallest  $d^2$  values in all cases. The value of  $d^2$  of Copula-based model is slightly larger than that of Model #1 but the difference is negligible. The value of  $d^2$  of Model #2 is much larger than that of the other two models. Therefore, it can be concluded that the Copula-based model and Model #1 are better than Model #2 when describing the statistical characteristics of the battery pack consistency parameters.

To determine the dependence structure between parameters of the three battery pack consistency models, the correlation coefficient between parameters is calculated according to Eq. (26).

$$r_{X,Y} = \frac{COV(X,Y)}{\sigma_X \sigma_Y} \quad (26)$$

Similar to the above, the sampling processes are also repeated 1000 times, and the average values are considered to be the final calculation results, which are shown in Table 2. For the test data, although there is a strong positive correlation between the battery capacity, the correlation between the battery capacity and resistance, and between the resistance and SOC, is weak. The correlation coefficients of the Copula-based model and Model #2 are extremely close to that of the test data. As for Model #1, the correlation coefficients are almost zero, which means that there is no correlation between the parameters. To explain the results more intuitively, the scatter plots between the battery capacity and the SOC from the test data and the three models are presented in Fig. 9. In Fig. 9(b) and (d), battery cells that have larger capacity values also have higher SOC values, which is close to the behaviors shown in Fig. 9(a).

Table 1

The squared Euclidean distance  $d^2$  between the marginal empirical CDF of the test data and the empirical CDFs of the three battery pack consistency models.

Item	Capacity	Resistance	SOC
Copula-based model	0.029	0.169	0.025
Model #1	0.015	0.015	0.016
Model #2	0.301	0.825	0.852

Table 2

Correlation coefficient between parameters in the different battery pack consistency models.

Item	Capacity-Resistance	Resistance-SOC	Capacity-SOC
Test data	-0.188	-0.221	0.856
Copula-based model	-0.279	-0.217	0.746
Model #1	-0.001	-0.001	0.003
Model #2	-0.189	-0.222	0.855

However, in Fig. 9(c), the dots are dispersive and there is no correlation between the two parameters. Therefore, it can be concluded that the Copula-based model and Model #2 can describe the dependence structure between variables more accurately.

To summarize, Model #1 is good at describing the statistical characteristics of the battery consistency parameters because the model is built using an inverse transform sampling method and the inverse empirical CDFs in the model are the same as those of the test data. Therefore, Model #1 has the highest accuracy when describing the battery consistency parameter distributions. In addition, because only the empirical CDFs of the battery consistency parameters are used, fitting the distributions of these parameters is not needed. However, a method based on inverse transform sampling is not suitable for modeling multiple variables, particularly when a correlation exists among the parameters. This is because the traditional sampling method cannot generate two sets of standard uniformly distributed random variables with a specific correlation. Hence, the correlation coefficient between the parameters in Model #1 cannot be ensured. As for Model #2, it is good at describing the dependence structure between variables because a covariance matrix (Eq. (23)) used in the model. However, one disadvantage of this model is that it assumes that all battery

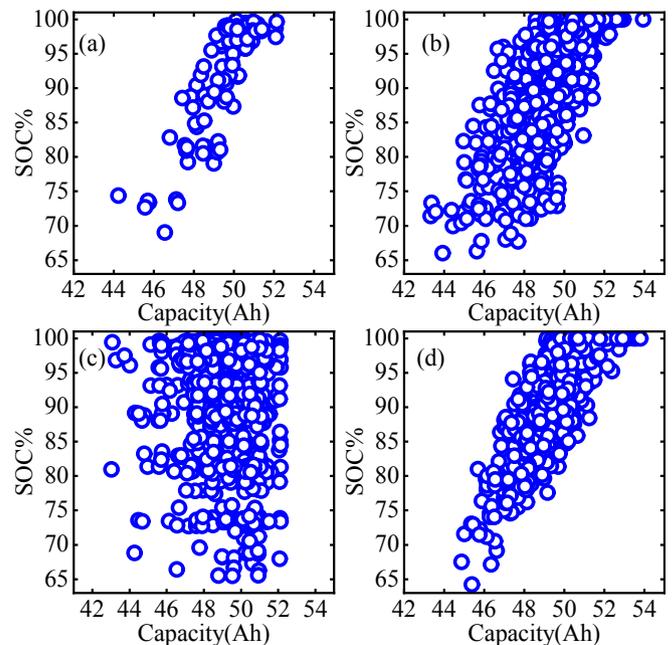


Fig. 9. Scatter plot of battery capacity and SOC from (a) test data, (b) the Copula-based model, (c) Model #1, and (d) Model #2.

consistency parameters are normally distributed. If a parameter is not normally distributed, the accuracy of describing the statistical characteristics of the parameters will be significantly reduced. For example, the SOC distribution in Fig. 5(c) is clearly not normally distributed. Therefore, the squared Euclidean distance between the marginal empirical CDF of Model #2 and that of the test data is much larger than that of the other two models. For the Copula-based battery pack consistency model, it integrates the advantages of Model #1 and Model #2, and can accurately describe the statistical characteristics of the variables as well as the correlation between variables. In addition, similar to Model #1, this model is also not limited by the distribution type of the battery consistency parameters. Therefore, the Copula-based model is the best among the three battery pack consistency models.

### 5.3. Verification of battery pack EUE estimation

Using the proposed Copula-based battery pack consistency model, the battery pack EUE at different current rates can be estimated according to Eq. (18) and Eq. (19). The true values of the battery pack EUE can be obtained from the battery pack rate capability and battery cell capacity tests. According to the results of the battery pack rate capability test, the battery pack discharge energy at different current rates can be obtained. Similarly, the maximum available energy that the battery pack can store can be calculated according to the test results of the battery cell capacity test. According to Eq. (18) the EUE of the battery pack at different current rates is calculated, the results of which are shown in Table 3. As the discharge current rate increases from 0.1C to 1.2C, the estimated battery pack EUE decreases from 62.01% to 54.47%. For the given four discharge current rates, the estimation errors are all within 0.6%, demonstrating that the proposed battery pack consistency modeling and the EUE estimation method have high levels of accuracy.

## 6. Conclusion

This paper proposes a novel battery pack consistency modeling method, and then the model is applied to estimate the EUE of a battery pack.

First, a Copula-based battery pack consistency modeling method is proposed. This model is then compared with two other models. The squared Euclidean distance  $d^2$  between the marginal empirical CDF of the test data and that of three battery pack consistency models are used to compare the ability of describing the statistical characteristics of the battery parameters. The results of the Copula-based model and Model #1 are all close to 0 (0.029, 0.169, and 0.025, respectively, for the Copula-based model, and 0.015, 0.015, and 0.016, respectively, for Model #1), which means that these two models are good at describing the statistical characteristics. The results of Model #2 (0.301, 0.825, and 0.852, respectively) are much larger than the results of the other two models, showing the weakest ability among the three models. The correlation coefficients between the parameters are used to compare the ability of describing the dependence structure between parameters. Model #1 shows no ability to describe the

correlation between parameters because the results are almost zero. The Copula-based model and Model #2 have a similar capability to describe the correlation between the parameters because the results of the two models ( $-0.279$ ,  $-0.217$ , and  $0.746$ , respectively, for Copula-based model, and  $-0.189$ ,  $-0.222$ , and  $0.855$ , respectively, for Model #2) are extremely close to that of the test data ( $-0.188$ ,  $-0.221$ , and  $0.856$ , respectively). Compared with the two existing battery pack consistency modeling methods, the proposed method exhibits good comprehensive performance both in describing the statistical characteristics of battery consistency parameters and the dependence between them.

Then, based on the battery pack consistency model, a framework to estimate the battery pack EUE is proposed. The EUE estimation value is obtained using the MC method. The proposed battery pack consistency modeling and EUE estimation method is verified based on the test results of a retired lithium-ion battery pack with 95 in-series connected LiFePO<sub>4</sub> battery cells. The EUE estimation errors at various discharge current rates are within 0.6%.

Because this study focus on the modeling of the battery pack consistency and its application to a battery pack EUE estimation, only the feasibility and accuracy are discussed herein. The battery consistency parameters in the study are obtained using an offline measurement. Our future work will focus on battery pack consistency modeling and EUE estimation using an on-board estimation of the battery consistency parameters.

## Acknowledgment

This work is supported by the National Key R&D Program of China [2018YFB0905304]; National Natural Science Foundation of China [U1664255]; and the Key Project of National Natural Science Foundation of China [61633015]. The authors would like to thank the members at the National Active Distribution Network Technology Research Center (NANTEC), Beijing Jiaotong University to help perform all the experiments of lithium-ion batteries mentioned in the paper.

## References

- [1] Liu K, Li K, Peng Q, Zhang C. A brief review on key technologies in the battery management system of electric vehicles. *Front Mech Eng* 2019;14:47–64.
- [2] Li Y, Liu K, Foley AM, Zülke A, Berecibar M, Nanini-Maury E, Van Mierlo J, Hoster HE. Data-driven health estimation and lifetime prediction of lithium-ion batteries: a review. *Renew Sustain Energy Rev* 2019;113:109254.
- [3] Lu L, Han X, Li J, Hua J, Ouyang M. A review on the key issues for lithium-ion battery management in electric vehicles. *J Power Sources* 2013;226:272–88.
- [4] Schuster SF, Brand MJ, Berg P, Gleissenberger M, Jossen A. Lithium-ion cell-to-cell variation during battery electric vehicle operation. *J Power Sources* 2015;297:242–51.
- [5] Klein M, Tong S, Park JW. In-plane nonuniform temperature effects on the performance of a large-format lithium-ion pouch cell. *Appl Energy* 2016;165:639–47.
- [6] Liu K, Zou C, Li K, Wik T. Charging pattern optimization for lithium-ion batteries with an electrothermal-aging model. *IEEE T Ind Inform* 2018;14:5463–74.
- [7] Jiang Y, Jiang J, Zhang C, Zhang W, Gao Y, Guo Q. Recognition of battery aging variations for LiFePO<sub>4</sub> batteries in 2nd use applications combining incremental capacity analysis and statistical approaches. *J Power Sources* 2017;360:180–8.
- [8] Xu X, Mi J, Fan M, Yang K, Wang H, Liu J, Yan H. Study on the performance evaluation and echelon utilization of retired LiFePO<sub>4</sub> power battery for smart grid. *J Clean Prod* 2019;213:1080–6.
- [9] Diao W, Xue N, Bhattacharjee V, Jiang J, Karabasoglu O, Pecht M. Active battery cell equalization based on residual available energy maximization. *Appl Energy* 2018;210:690–8.
- [10] Zhang C, Jiang Y, Jiang J, Cheng G, Diao W, Zhang W. Study on battery pack consistency evolutions and equilibrium diagnosis for serial-connected lithium-ion batteries. *Appl Energy* 2017;207:510–9.
- [11] Zhou L, Zheng Y, Ouyang M, Lu L. A study on parameter variation effects on battery packs for electric vehicles. *J Power Sources* 2017;364:242–52.
- [12] Zheng Y, Ouyang M, Lu L, Li J. Understanding aging mechanisms in lithium-ion battery packs: from cell capacity loss to pack capacity evolution. *J Power Sources* 2015;278:287–95.

**Table 3**  
EUE estimation results.

Rate	EUE (%)	Estimated EUE (%)	Error (%)
0.1C	62.56	62.01	−0.55
0.3C	60.38	60.10	−0.27
0.7C	57.70	57.46	−0.23
1.2C	54.54	54.47	−0.07

- [13] Yang F, Wang D, Zhao Y, Tsui K, Bae SJ. A study of the relationship between coulombic efficiency and capacity degradation of commercial lithium-ion batteries. *Energy* 2018;145:486–95.
- [14] Yang X, Leng Y, Zhang G, Ge S, Wang C. Modeling of lithium plating induced aging of lithium-ion batteries: transition from linear to nonlinear aging. *J Power Sources* 2017;360:28–40.
- [15] Yang X, Wang C. Understanding the trilemma of fast charging, energy density and cycle life of lithium-ion batteries. *J Power Sources* 2018;402:489–98.
- [16] Ma Z, Jiang J, Shi W, Zhang W, Mi CC. Investigation of path dependence in commercial lithium-ion cells for pure electric bus applications: aging mechanism identification. *J Power Sources* 2015;274:29–40.
- [17] Schuster SF, Brand MJ, Campestrini C, Gleissenberger M, Jossen A. Correlation between capacity and impedance of lithium-ion cells during calendar and cycle life. *J Power Sources* 2016;305:191–9.
- [18] Mathew M, Kong QH, McGrory J, Fowler M. Simulation of lithium ion battery replacement in a battery pack for application in electric vehicles. *J Power Sources* 2017;349:94–104.
- [19] Mathew M, Janhunien S, Rashid M, Long F, Fowler M. Comparative analysis of lithium-ion battery resistance estimation techniques for battery management systems. *Energies* 2018;11:1490.
- [20] Lievre A, Sari A, Venet P, Hijazi A, Ouattara-Brigaudet M, Pelissier S. Practical online estimation of lithium-ion battery apparent series resistance for mild hybrid vehicles. *IEEE T Veh Technol* 2016;65:4505–11.
- [21] Feng X, Xu C, He X, Wang L, Gao S, Ouyang M. A graphical model for evaluating the status of series-connected lithium-ion battery pack. *Int J Energ Res* 2019;43:749–66.
- [22] Diao W, Jiang J, Liang H, Zhang C, Jiang Y, Wang L, Mu B. Flexible grouping for enhanced energy utilization efficiency in battery energy storage systems. *Energies* 2016;9:498.
- [23] Xi Z, Jing R, Wang P, Hu C. A copula-based sampling method for data-driven prognostics. *Reliab Eng Syst Saf* 2014;132:72–82.
- [24] Valizadeh Haghi H, Lotfifard S. Spatiotemporal modeling of wind generation for optimal energy storage sizing. *IEEE T Sustain Energ* 2015;6:113–21.
- [25] Bao Y, Dong W, Wang D. Online internal resistance measurement application in lithium ion battery capacity and state of charge estimation. *Energies* 2018;11:1073.
- [26] Zhang Y, Shang Y, Cui N, Zhang C. Parameters identification and sensitive characteristics analysis for lithium-ion batteries of electric vehicles. *Energies* 2018;11:19.
- [27] Zhang Y, Xiong R, He H, Pecht MG. Lithium-ion battery remaining useful life prediction with box-cox transformation and Monte Carlo simulation. *IEEE T Ind Electron* 2019;66:1585–97.
- [28] Kroese DP, Brereton T, Taimre T, Botev ZI. Why the Monte Carlo method is so important today. *Wiley Interdiscip Rev: Comput Stat* 2014;6:386–92.