

Article

Using Self Organizing Maps to Achieve Lithium-Ion Battery Cells Multi-Parameter Sorting Based on Principle Components Analysis

Bizhong Xia ¹, Yadi Yang ^{1,*}, Jie Zhou ¹, Guanghao Chen ¹ , Yifan Liu ¹, Huawen Wang ², Mingwang Wang ² and Yongzhi Lai ²

¹ Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China

² Sunwoda Electronic Co. Ltd., Shenzhen 518108, China

* Correspondence: yangyd16@mails.tsinghua.edu.cn; Tel.: +86-1324-386-7516

Received: 28 June 2019; Accepted: 29 July 2019; Published: 1 August 2019



Abstract: Battery sorting is an important process in the production of lithium battery module and battery pack for electric vehicles (EVs). Accurate battery sorting can ensure good consistency of batteries for grouping. This study investigates the mechanism of inconsistency of battery packs and process of battery sorting on the lithium-ion battery module production line. Combined with the static and dynamic characteristics of lithium-ion batteries, the battery parameters on the production line that can be used as a sorting basis are analyzed, and the parameters of battery mass, volume, resistance, voltage, charge/discharge capacity and impedance characteristics are measured. The data of batteries are processed by the principal component analysis (PCA) method in statistics, and after analysis, the parameters of batteries are obtained. Principal components are used as sorting variables, and the self-organizing map (SOM) neural network is carried out to cluster the batteries. Group experiments are carried out on the separated batteries, and state of charge (SOC) consistency of the batteries is achieved to verify that the sorting algorithm and sorting result is accurate.

Keywords: lithium-ion battery; cell sorting; multi-parameters sorting; principal component analysis; self-organizing maps clustering

1. Introduction

The initial differences among batteries which lead to inconsistency after charging and discharging are the main reasons for the shortened life and low safety in the use of battery packs. It is not feasible to completely eliminate the inconsistency of battery packs, but there are some ways to reduce inconsistency's negative impact [1,2]. Cell sorting in lithium-ion battery industry is an indispensable process to assure the reliability and safety of cells that are assembled into strings, blocks, modules and packs [3]. In the current lithium-ion power battery pack production line, cell sorting refers to the selection of qualified cells from raw ones according to quantitative criteria in terms of accessible descriptors such as battery resistance, open circuit voltage (OCV), charging/discharging capacity, etc. Correspondingly, resistance sorting, voltage sorting and capacity sorting are the main single parameter sorting methods used in battery pack production lines at present [4–7]. The single-parameter sorting method can quickly screen batteries whose parameters are in the qualified range from raw ones [8], but this method ignores the relationship among battery parameters. In contrast, the multi-parameter sorting method which combines static and dynamic characteristics of batteries is a more accurate and comprehensive solution. In order to realize multi-parameter sorting of lithium batteries in the battery pack production line, it is necessary to test battery parameters, such as the resistance, voltage and capacity, using the existing equipment.

We note that a certain number of open publications focusing on sorting methods can be found, and clustering algorithms [9,10], including the fuzzy C-means algorithm (FCM) [11,12], k-means algorithm [13], and self-organizing maps (SOM) neural networks [14–16], are the main research direction. Nevertheless, the sorting parameters and their relationships and constraints adopted in the sorting process, which are of fundamental significance, have not been investigated. Therefore, in this study, a series of parameters as sorting criteria which can characterize the performance and difference of batteries are investigated, and they are divided into three categories: general parameters, static performance parameters and dynamic characteristics.

General parameters refer to the appearance parameters such as mass (m) and volume (v) of batteries and the derived parameters such as mass specific energy and volume specific energy caused by them. Static performance parameters include battery resistance, voltage, charge/discharge capacity, self-discharge rate, charge/discharge efficiency and charge/discharge time. They describe the performance of batteries under non-working conditions, but their acquisition needs certain conditions. For instance, charging and discharging [17] experiments are needed to obtain the capacity data of batteries. As for static characteristics of a battery, charge-discharge voltage curve and electrochemical impedance spectroscopy (EIS) [18] are typical methods to show the dynamic performance. Some previous studies utilize an overall voltage-current or voltage-time curve comparing the differences among batteries to complete sorting. However, drawing an overall voltage-time curve or EIS takes a long time to complete, so it is unrealistic to use it in production line.

Under such circumstances, this study compares several parameters as descriptors for battery sorting and finally selects the following sorting indicators: battery mass (m), volume (v), voltage (V), capacity (C) and surface temperature (T) [19] as static indicators. As aforementioned, an overall voltage-time curve or EIS spectrum takes a long time to obtain, therefore, in this study the impedance characteristics of batteries at several key frequencies in EIS spectrum [20] are extracted, and the AC (Alternative Current) impedance (R_A) and reactance (X) of batteries are taken as the evaluation criteria of dynamic characteristics. These parameters can be obtained by testing equipment now available on the production line. The principle components analysis (PCA) [21] method is used to pre-process the data of battery parameters (clarifying the relationship between parameters), and obtain the principal components that can reflect the characteristics of the battery. The clustering and sorting process of batteries is accomplished by a self-organizing map (SOM) neural network.

The rest of this paper is organized as follows: The acquisition of battery parameters and data analysis are completed in Section 2, followed by the results of data processing using PCA. Section 3 introduces the principle of SOM algorithm and its application in battery clustering. Section 4 carries out the battery sorting experiments and operation tests to obtain the state of charge (SOC) curves of the batteries to be sorted under special conditions. Finally, in Section 5 the performance of batteries after grouping is studied, and the results of battery sorting are verified. Section 6 concludes the major findings.

2. Parameter Acquisition and Data Processing

This section describes the process to obtain the parameters of the batteries to be sorted, so as to get a sufficient number of parameter inputs to start a follow-up clustering algorithm.

A set of 58 lithium iron phosphate (LiFePO_4) batteries are considered in this work. The batteries to be sorted are INR 18650-33G cylindrical batteries produced by Samsung SDI Company (Suwon, South Korea). The cathode is made of lithium iron phosphate and the anode is made of graphite. Tests were performed in laboratories in Tsinghua University Shenzhen Graduate School and Sunwoda electronics Co., Ltd. in Shenzhen to obtain experimental data such as mass, volume, voltage and currents under charge/discharge and capacity, etc. and to identify internal resistance. These 58 Samsung 18650 cells with a nominal capacity of 2700 mAh and nominal voltage of 3.6V were tested under certain circumstances. Table 1 shows the cell specification.

Table 1. Cell Specification.

Item	Value
Manufacturer	Samsung INR18650-33G
Nominal capacity	2700 mAh
Charging cut-off voltage	4.2 V
Discharging cut-off voltage	2.6 V
Nominal voltage	3.6 V
Standard weight	47 ± 2.0 g
Dimensions	Height: 65.35 ± 0.15 mm Diameter: φ18.33 + 0.1/−0.13 mm
Working temperature	Charge: −10~45 °C Discharge: −20~60 °C

2.1. General Parameters

The general parameters of batteries can be directly reflected in their physical attributes, such as the mass, volume and surface temperature of battery cells, as well as the derived mass specific energy and volume specific energy. Because all the batteries selected in this paper are of the same type, the battery energy is the same by default. Therefore, the mass (m), volume (v) and surface temperature (T) of the battery cells are selected as the general parameters.

Parameters m and v of batteries were tested with Keyence pressure sensor and vision measurement sensor. The quality measurement with 10 μg accuracy can be achieved by pressure conversion measured by pressure gauge. The volume of the battery is calculated by measuring the geometric size. The surface temperature (T) of battery is measured by thermocouple.

2.2. Static Parameters

The voltage (V) of the batteries to be sorted can be measured with a high precision voltmeter. The total capacity (C) of batteries is a unique characteristic, which is different between different batteries. It can be obtained from the charge/discharge test, by means of the capacity data provided from the Coulomb counting method, shown as Equation (1), directly implemented in the testing equipment. The charge/discharge test is carried out using Arbin BT-5HC battery testing equipment, with battery cells (to be sorted) in a Sanwood constant temperature and humidity box. Figure 1 indicates the curve of voltage and current during the charge/discharge test.

$$C = It. \quad (1)$$

C denotes charging/discharging capacity, I and t represent charging/discharging current and time. In the process of capacity testing, the lithium battery that has been stationary for 2 h is discharged to the discharge cut-off voltage (2.6 V) at a rate of 1 C. Then the battery is charged by means of CC-CV (Constant Current-Constant Voltage) at a rate of 0.5 C to the charging cut-off voltage (4.2 V), and then is discharged at a rate of 1 C until the discharge cut-off voltage. The capacity test curve as shown in Figure 1 is obtained after several cycles, and the capacity is calculated using Equation (1).

There are several ways to measure DC (Direct Current) internal resistance of batteries. In this study, according to Equation (2), the DC internal resistance is derived by dividing the voltage drop during the transition to rest by the constant discharge current:

$$R_D = \frac{\Delta V}{I} = \frac{V_2 - V_1}{I}. \quad (2)$$

Due to the polarization phenomenon, the voltage of lithium batteries presents a dynamic process of rapid rise, slow rise, rapid fall and slow fall during pulse charging. The change of voltage ΔV during discharging is contrary to that during charging. In this process, the sharp voltage drop is caused by the

ohmic polarization inside the battery, so the ohmic internal resistance of the battery can be identified by using the voltage and current data.

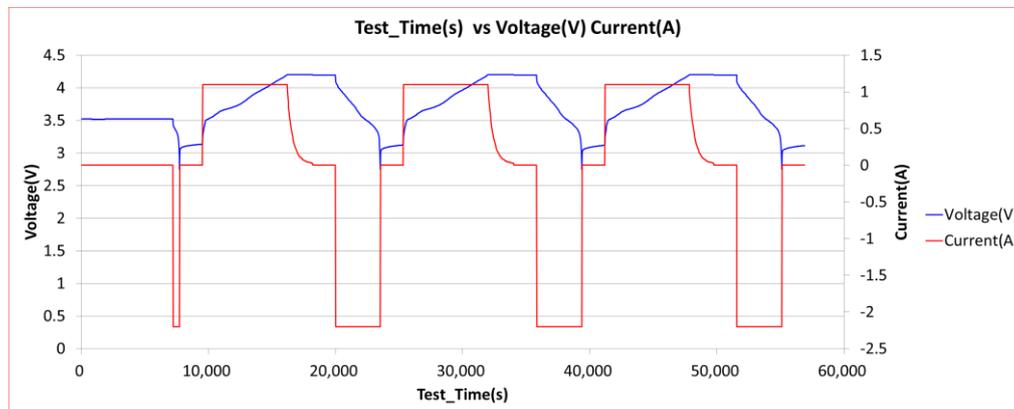


Figure 1. Capacity testing process.

2.3. Impedance Characteristics

The electrochemical impedance spectroscopy (EIS) method is used to measure the change of the impedance or phase angle of the battery system with frequency, by applying sinusoidal alternating currents of different frequencies to the battery. The EIS of a battery is an inherent characteristic. It is a feasible method to use the similarity of EIS drawings to sort batteries. The EIS method is a quasi-dynamic process, and small sinusoidal alternating current will not cause great disturbance to the battery, so the EIS sorting method is more accurate than the static method.

In the EIS of lithium-ion batteries, the different parts corresponding to different frequencies can reflect the process of lithium ion removal and embedding in the embedded electrode during the charging and discharging process of lithium-ion batteries.

- (1) Ultra high frequency (UHF) part: The section where the impedance curve intersects the transverse axis in EIS, reflecting the ohmic impedance (R_b) of lithium ion batteries.
- (2) High frequency part: A semicircle related to the diffusion and migration of lithium ions through the insulating layer on the surface of active material particles. This process can be represented by a R_{sei}/C_{sei} parallel circuit.
- (3) Intermediate frequency (IF) part: A semicircle reflecting the charge transfer process. The process is represented by a R_{ct}/C_{dl} parallel circuit. R_{ct} is the charge transfer impedance, or the polarization impedance of the battery, and C_{dl} is double-layer capacitor.
- (4) Low frequency part: A line with a slope of 45 degrees, which is related to the solid diffusion process of lithium ions in active material particles. In the equivalent circuit model, a Warburg impedance Z_W describing diffusion can be used, also known as concentration polarization impedance.
- (5) Ultra low frequency: A semicircle and a vertical line, reflecting the change in crystal structure of active material particles and the accumulation and consumption of lithium ions in active material respectively.

Figure 2a indicates the equivalent circle model for EIS used for curve fitting. In this study, using the HIOKI BT4560 Impedance Tester (HIOKI, Shanghai, China), the impedance experiments are carried out on one of the lithium battery cells using 0.1–1050 Hz AC current, and the impedance spectra are plotted, as shown in Figure 2b, which reflects the intermediate frequency part mentioned above. BT4560 can realize simultaneous measurement of battery voltage and impedance with high accuracy. The measurement accuracy of voltage and impedance can reach 10 μ V and 1 $\mu\Omega$ respectively. It can also measure the surface temperature of batteries by connecting thermocouples, and the accuracy can reach 0.1 $^{\circ}$ C.

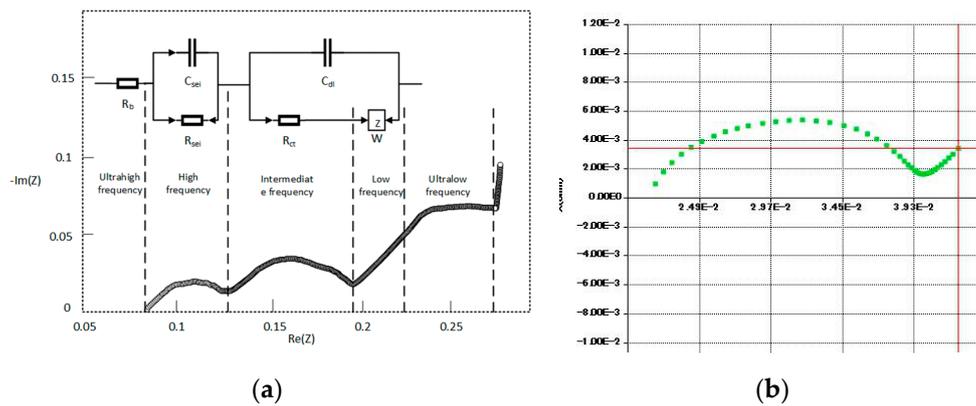


Figure 2. Electrochemical impedance spectroscopy (EIS) of lithium-ion battery: (a) EIS and equivalent circle model of Li-ion battery; (b) EIS in intermediate frequency (IF) part of one 18,650 cell using BT4560.

As described in Section 1, we collect the impedance characteristics of batteries at several key frequencies in EIS spectrum, so that the measurement process can be rapid. In this study, two characteristic frequencies, 0.1 Hz and 1000 Hz, are selected to obtain the impedance (R_A) and reactance (X) of the battery as the sorting parameters. These two frequencies are between high, medium and low frequencies. The results of impedance can reflect the polarization impedance and Warburg impedance of batteries, respectively.

2.4. Principle Components Analysis

The parameters of each battery cell to be sorted are tested using the test equipment suitable for the production line. The results are shown in Table 2; two cells are selected to show the input parameters.

Table 2. Examples of parameters for two batteries.

No.	1	2
Manufacturer	Samsung INR18650-33G	Samsung INR18650-33G
Mass (g)	46.92499999	47.60953785
Volume (mm ³)	16,540	16,540
Capacity (mAh)	2600	2650
Voltage (V)	3.524017	3.568177
Impedance R1 (1000 Hz) (Ω)	0.0219427	0.0223717
Reactance X1 (1000 Hz) (Ω)	-0.0002701	-0.0006503
Impedance R2 (0.1 Hz) (Ω)	0.0355982	0.0401869
Reactance X2 (0.1 Hz) (Ω)	-0.0022978	-0.0024677
Surface temperature ($^{\circ}$ C)	26.2	26.5

The input data matrix of 58×9 is obtained after the test of 58 battery cells. Principle component analysis (PCA) is used to analyze data. PCA is a statistical method, which uses linear transformation to achieve dimension reduction. Its basic principle is to use the idea of dimension reduction and linear transformation to transform multiple indicators into several unrelated comprehensive indicators without losing much information, which are called the principle components. In this study, the data of 58 batteries are analyzed by principal component analysis in the following steps.

1. Arrange the raw data in rows to form a matrix X .
2. Standardize the data of matrix X , change its mean to zero.
3. Compute Covariance Matrix C .
4. Calculate eigenvalues according to the covariance matrix C . The eigenvectors are arranged from large to small eigenvalues, and the first k matrices are composed of rows P .

5. Calculate by Equation (3) to get the reduced dimension data matrix Y .

$$Y = PX. \quad (3)$$

6. Compute the contribution rate V_i of each eigenvalue as shown in Equation (4)

$$V_i = \frac{x_i}{x_1 + x_2 + \dots + x_n}. \quad (4)$$

7. Sort the principal components according to their contribution rate and explain the physical significance of principal components based on eigen roots and their eigenvectors.

The result of PCA shows that three principle components can cover more than 90% of data variability, as shown in Table 3. By analyzing the proportion of principal components of each parameter, the proportion results are obtained.

$$t_1 = -0.09011m - 0.09044v - 0.08985C - 0.34997V + 0.461975R_1 + 0.458163X_1 + 0.465659R_2 + 0.456563X_2 + 0.06728T. \quad (5)$$

$$t_2 = 0.564575m + 0.564293v + 0.564558C - 0.14541V + 0.046857R_1 + 0.079282X_1 - 0.00029R_2 + 0.083406X_2 + 0.086721T. \quad (6)$$

$$t_3 = -0.05338m - 0.0557v - 0.05278C - 0.17591V - 0.08143R_1 - 0.08759X_1 - 0.04988R_2 - 0.08818X_2 + 0.967342T. \quad (7)$$

Table 3. (a) Principal component analysis of 58 batteries with nine parameters. (b) Proportional relationship between the original parameters and principal components.

(a)				
Principle Component	Eigenvalue	D-Value	Contribution Rate	Cumulative Contribution Rate
1	4.513363	1.502216	50.14848	50.14848
2	3.011147	2.013297	33.45719	83.60566
3	0.99785	0.546226	11.08722	94.69288
4	0.451624	0.427816	5.018046	99.71093
5	0.023808	0.022719	0.264529	99.97546
6	0.001089	0.000355	0.012096	99.98755
7	0.000733	0.000383	0.008149	99.9957
8	0.000351	0.000315	0.003898	99.9996
9	3.6×10^{-5}	0	0.000399	100

(b)			
Standardized Variables	Feature Vector (t_1)	Feature Vector (t_2)	Feature Vector (t_3)
Mass (g)	-0.09011	0.564575	-0.05338
Volume (mm ³)	-0.09044	0.564293	-0.0557
Capacity (mAh)	-0.08985	0.564558	-0.05278
Voltage (V)	-0.34997	-0.14541	-0.17591
Impedance R1 (1000 Hz) (Ω)	0.461975	0.046857	-0.08143
Reactance X1 (1000 Hz) (Ω)	0.458163	0.079282	-0.08759
Impedance R2 (0.1 Hz) (Ω)	0.465659	-0.00029	-0.04988
Reactance X2 (0.1 Hz) (Ω)	0.456563	0.083406	-0.08818
Surface temperature ($^{\circ}$ C)	0.06728	0.086721	0.967342

Equations (5)–(7) use eigenvectors to show the relationship with nine initial characteristic parameters. It can be seen that these nine characteristics are confirmed and can be divided into three

categories: appearance parameters, impedance characteristics and safety parameters. The characteristic vector t_1 is mainly positively correlated with the impedance characteristics of batteries, and negatively correlated with mass, volume and capacity, so t_1 can be regarded as the principal component reflecting the impedance characteristics of batteries; similarly, the vector t_2 can be regarded as the principal component reflecting the appearance parameters of batteries. The characteristic vector t_3 mainly reflects the temperature parameters of the battery, which is related to the safety and thermal management of the battery.

3. SOM Neural Networks and SOM Clustering

3.1. Introduction of SOM Neural Networks

The self-organizing map (SOM) neural network is an unsupervised learning clustering algorithm which realizes high-dimension visualization. It is an artificial neural network developed by simulating the characteristics of human brain to signal. The model was first proposed by Teuvo Kohonen [22], a professor at Helsinki University in Finland in 1981, so it is also called the Kohonen network.

The application of SOM neural networks in clustering is mainly based on the following advantages:

1. Dimension reduction can be achieved, and clustering results have good visibility. High dimensional input space can be mapped to low dimensional output space maintain the original topological relationship.
2. With self-organizing and unsupervised learning, it can be applied to situations where the characteristics of input data are not fully understood.
3. The algorithm is clear and the calculation is simple.

These features enable SOMs to be widely used in clustering, helping to recognize homogeneous groups of generous and complex inputs. Reference [14–16] indicates research and applications where SOMs are used in clustering processes in the last decades. Compared with traditional clustering algorithms, such as the k-means algorithm and FCM algorithm, SOM is more accurate and not affected by the selection of initial clustering centers, although its operation process is relatively complex.

3.2. Principle and Learning Algorithm of SOM

The structure of a SOM neural network is shown in Figure 3. It consists of an input layer and output layer (competition layer). The number of neurons in the input layer is determined by the number of vectors in the input network, and the output layer is arranged into a two-dimensional node matrix by neurons. The neurons in the input layer and the neurons in the output layer are fully connected by weights W . When the network receives the external input signal, the neurons in the output layer excite and distribute the neurons in the region with the highest input spatial density through competition.

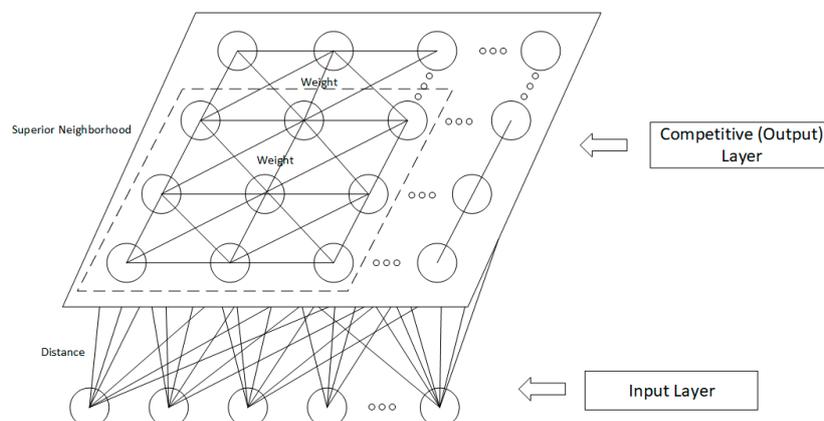


Figure 3. The structure of a self-organizing map (SOM) neural network.

A SOM is a competitive neural network, which follows the rules of competitive learning (referring to lateral inhibition of neurons in human brain). For all neurons in the output layer (competitive layer), the winner neurons are found according to the principle of minimum distance, then only the winner neurons and the neurons in the winning neighborhood are adjusted by the weight vector. Based on the WTA (winner-takes-all) rule, the specific algorithm steps of the SOM are as follows:

Step 1: Data initialization. The weight vectors of the output layer are given smaller random numbers and normalized, which are recorded as $\hat{W}_j (j = 1, 2, \dots, m)$, m is number of neurons in the output layer. The initial optimal neighborhood is established as $N_j^*(0)$, and the initial value of learning efficiency α and the number of iterations T are set.

Step 2: Input data and normalize. Normalize the input vector $X^p (p = 1, 2, \dots, n)$ and the weight vector W_j in the competition layer as shown in Equation (8), where n is the dimension of the input data.

$$\hat{X} = \frac{X}{\|X\|}, \hat{W}_j = \frac{W_j}{\|W_j\|} \quad (8)$$

The normalized input vector is denoted as $\hat{X}^p (p = 1, 2, \dots, n)$, where n is the dimension of the input data.

Step 3: Find the winning node. Calculate the distance between input \hat{X}^p and all output neurons, and select the winning neuron with the smallest distance. Euclidean distance is chosen for distance formula, as shown in Equation (9). Set the winning node as the center and determine the weight adjustment region at t -time. Setting the neighborhood distance d , the region within the D range of the distance from the winning node is regarded as the winning neighborhood.

$$d_j = \|\hat{X} - \hat{W}_j\| = \sqrt{\sum_{j=1}^m [X - W_j]^2}. \quad (9)$$

Step 4: Definition of superior neighborhood. Taking the winning node j^* as the center, the weight adjustment region of t time is determined. Setting the neighborhood distance d , the region within the d range of the distance from the winning node is regarded as the winning neighborhood.

$$N_j(t) = \{j, d_{j,j^*} \leq d\} \quad (10)$$

Step 5: Weight adjustment. Based on the gradient descent method, the connection weights between the winning node and all other nodes in the winning neighborhood are adjusted. Equation (11) indicates the adjustment process as follows:

$$W_{j^*}(t+1) = \hat{W}_{j^*}(t) + \Delta W_{j^*} = \begin{cases} \hat{W}_{j^*}(t) + \alpha(\hat{X}^p - \hat{W}_{j^*}(t)) & j = j^* \\ \hat{W}_{j^*}(t) + \alpha N(t)(\hat{X}^p - \hat{W}_{j^*}(t)), & j \in N_j(t) \quad i = 1, 2, \dots, n, j \in N_j^*(t) \\ \hat{W}_{j^*}(t) & j \notin N_j(t) \end{cases} \quad (11)$$

where, $w_{ij}(t)$ denotes the weight of neuron i to neuron j at t time; $0 < \alpha \leq 1$ denotes the learning rate, which affects the convergence and stability of the algorithm and decreases with time; $N(j, t)$ denotes the topological distance between the j th neuron and the winning neuron j^* in the neighborhood at training time t .

Step 6: End Judgment: When the learning rate $\alpha(t) \leq \alpha_{\min}$, or iterations $t \geq T$, the training process is terminated, otherwise return back to Step 2 and continue the iteration.

The advantages of SOM unsupervised learning and visualization of results enable it to be well applied in clustering analysis, as well as in lithium cell sorting.

3.3. SOM Clustering in Battery Sorting

The SOM is able to map any high-dimensional inputs to low-dimensional outputs, such as one-dimensional linear array or two-dimensional grid. Therefore, this feature of the algorithm provides

the possibility of sorting battery cells from raw cell groups according to single or multiple parameters or characteristics. As aforementioned in Section 2, multiple parameters tested by testing equipment in a production line are obtained, and they can be regarded as input data. The output layer of the SOM map shows the sorting types each cell belongs to.

Nine parameters are tested including mass, voltage, impedance and parameters data. These are pre-processed using PCA and transformed to three components in Section 2. These parameters are used as inputs of the SOM neural network for battery cell sorting and the results of classification are the output.

This study uses a SOM neural network to sort battery cells. The data of m battery cells with n parameters are input in the form of matrix of $m \times n$, and finally the cells are classified into k classes. The learning rate α and neighborhood radius r of the network are updated in the way shown in Equations (12) and (13), respectively.

$$\alpha = \alpha_0 \times \left(1 - \frac{t}{T_{\max}}\right), \quad (12)$$

$$r = r_0 \times \left(1 - \frac{t}{T_{\max}}\right) \quad (13)$$

where, T_{\max} means the maximum numbers of iteration. $\alpha_0 = 0.99$ and $r_0 = 1$ are the initial learning rate and winning neighborhood radius respectively, and these two parameters decrease with time.

Figure 4 demonstrates the process of battery sorting using SOM in this study and the specific operation will be introduced in next section.

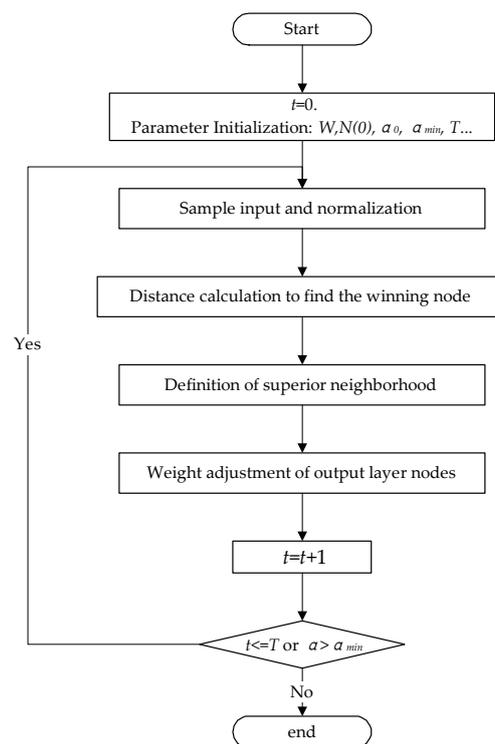


Figure 4. Procedure of battery sorting using a SOM.

4. Cell Sorting Using a SOM

As described in Section 2, the vector of nine parameters is considered as input for SOM neural network clustering. We tested 58 battery cells and recorded nine parameters of each cell. The input of the SOM is the parameter matrix with 58×9 . After PCA, in order to reduce the data dimension, the matrix can be transformed to 58×3 as the nine parameters are replaced with the three principle

components. Therefore, the first group of input data including nine parameters is marked as Group 1, and the second group of input data including three principle components is marked as Group 2. Table 4 denotes part of the input data (eight cells).

Table 4. Examples of principal component scores for eight batteries (Group 2).

Battery Number	Principal Component Score First (t_1)	Principal Component Score Second (t_2)	Principal Component Score Third (t_3)
01	-1.42814	7.525299	-0.02753
02	-1.38087	7.602607	0.678699
03	-0.97269	2.441534	-0.89297
04	-0.96143	2.449494	-0.78799
05	-0.96094	2.466339	-0.71301
06	-0.95142	2.465672	-0.68406
07	-0.90951	-1.02728	-1.6356
08	-0.88487	2.546314	0.351155

This section compiles the algorithm of a SOM neural network and sets its initial algorithm parameters by using the Neural Network Toolbox in MATLAB. The initial number of neurons in the output layer, is set to four, as shown in Figure 5.

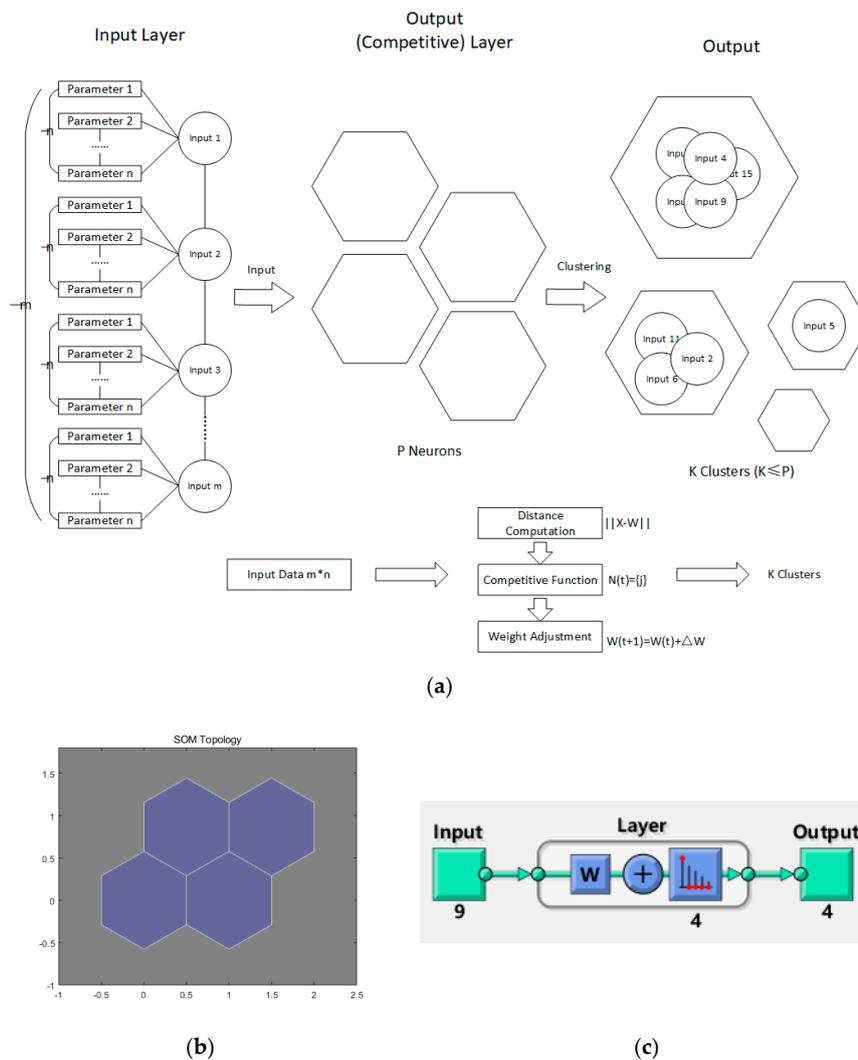


Figure 5. Example of SOM mapping and clustering process: (a) SOM mapping and learning process in clustering; (b) SOM topology in MATLAB; (c) SOM structure in view in MATLAB (e.g., nine inputs in the input layer and four neurons in the output layer).

Figure 6 shows the SOM clustering results of Group 1 and Group 2. In both sets of cells, eight battery cells are clearly separated from the whole ones, which means that these eight cells are different from the others.

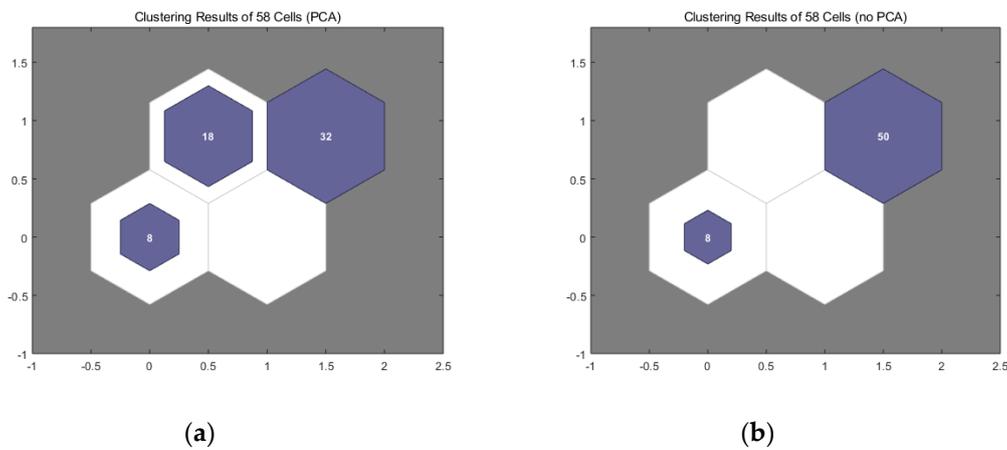


Figure 6. SOM clustering results (four neurons), (a) Group 1, (b) Group 2.

The same procedure using a SOM is carried out to obtain the sorting results from the remaining 50 battery cells and the number of sorting is adjusted to six, as shown in Figure 7. Six groups are obtained after being classified. Table 5 shows the categories of 50 numbered batteries based on the SOM clustering results, both in Group 1 and Group 2.

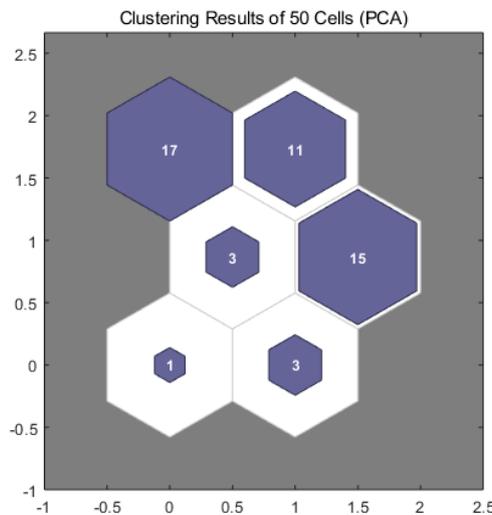


Figure 7. SOM clustering results of Group 1 (six neurons in the output layer).

Table 5. Categories of 50 battery cells.

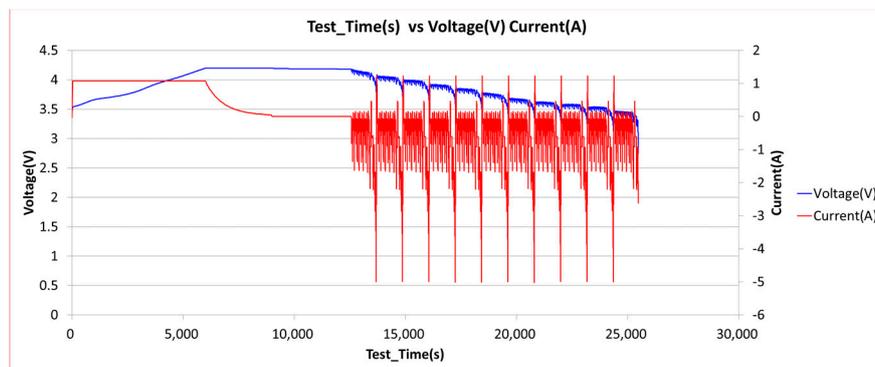
Group No.	Battery Cell Number in Group 1	Battery Cell Number in Group 2
Category 1	1 2 7 15 21 22 24 27 32 34 35 36 38 40 45	2 7 15 21 22 34 35 36 38 40 45
Category 2	3 4 5 6 9 13 16 23 28 33 41 42 43 46 47 48 49	3 4 5 6 9 13 28 33 41 42 43 46 47 49
Category 3	8 10 14 19 20 29 30 31 37 39 44	8 10 14 19 20 29 30 31 37 39 44
Category 4	11 17 25	11 17 25 24 27 32
Category 5	12 18 26	12 16 18 23 26 48
Category 6	50	1 50

The contents mentioned above in this section illustrates a qualitative result of SOM battery cell sorting, therefore a validation experiment of the sorting results must be operated.

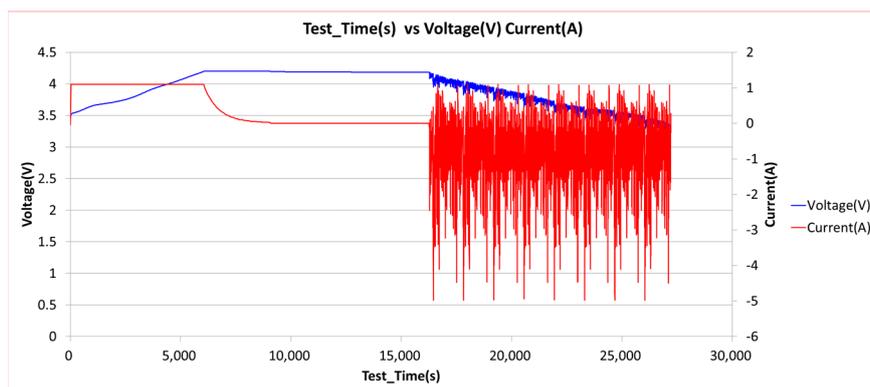
It can be noted that the sorting result of Group 2 is basically coincident with Group 1, which means SOM sorting, using the three principle components obtained from nine parameters after PCA, is feasible. Consequently, pretreatment of PCA is also a workable method when the sorting batteries have multiple dimension parameters. It also effectively reduces the workload of the subsequent scoring algorithm such as SOM. According to the classification results, eight batteries with good consistency (i.e., classified into the same category) can be selected from each group to form a module. This result will be validated by experiments in the next section.

5. Verification for Sorting Results

In response to different operating conditions, the SOC of lithium-ion batteries will vary with the voltage and current. Despite of their different dynamic responses, battery cells with good consistency should have less SOC variation differences. Therefore, two typical power load profiles of electric vehicles, the New European Driving Cycle (NEDC) and the Urban Dynamo-meter Driving Schedule (UDDS), shown in Figure 8, are chosen to test the battery modules sorted and formed from Section 4. The changes of the SOC curve and differences in the module under working conditions are recorded and calculated in the module under working conditions to test the clustering algorithm.



(a)



(b)

Figure 8. Power load profile: (a) New European Driving Cycle (NEDC) working conditions; (b) Urban Dynamo-meter Driving Schedule (UDDS) working conditions.

Eight battery cells of Module 1 and Module 2 are chosen from the same categories of Group 1 and Group 2 in Section 4, respectively. Eight cells (No. 3, 4, 5, 6, 9, 13, 16, 23), which are all classified to category 2 in Group 1, are connected in series to form Module 1. The same procedure is carried out to category 1 in Group 2, so that eight cells (No. 2, 7, 15, 21, 22, 34, 35, 36) are connected in series to form Module 2. For validation of consistency, another eight cells are chosen randomly from the whole cell

stack to form Module 3. To ensure the randomness of the selection of the eight batteries, one cell is selected from each seven cells according to number order, which means the first cell is selected from No.1 to No.7 randomly, the second cell is selected from No. 8 to No. 14 and so on. The final number of these eight cells in Module 3 are: No. 1, 8, 17, 26, 32, 37, 45, 54.

The battery module is charged and discharged using battery charging and discharging equipment under the working conditions shown in Figure 8, and the current and voltage changes of each battery are recorded.

The main result of battery module inconsistency is that the state of charge (SOC) of each cell is not uniform, but differs from one another, because every cell is different in available capacity, dynamics, and imbalance of the individual cells in a series or parallel chain. So, the SOC curve and variability can be used to determine whether the battery module has good consistency or not. Currently, a lot of research has been done on the estimation methods of SOC. Among them, the most typical methods are the time integration method, open-circuit voltage method and Kalman filter method. In this section, extended Kalman filtering (EKF) [23,24] is used to estimate SOC change under two working conditions.

The Kalman filter is a common algorithm in SOC estimation. It can realize the optimal estimation of the state of a discrete-time linear system based on the minimum mean square error. The Kalman filter method mainly includes two parts: prediction and correction. The prediction process refers to updating the state estimates at the last time based on the state equation obtained from the battery model. The correction process is to update the state predicted value, according to the observed value obtained from the test. The extended Kalman filter (EKF) is based on the Kalman filter, which expands the non-linear function of the system into a Taylor expansion of the first order, and obtains the linearized system equation to complete the filtering and estimation of the state parameters.

In this study, EKF is one of the SOC estimation methods used for example, so the specific process of EKF is not discussed here.

In order to save time, a section of NEDC working conditions is selected as shown in Figure 9. Figure 10 shows the SOC of these three modules subject to the power profile of the NEDC working conditions shown above. The SOC in the battery module appears as a consistent trend over time, but the SOC between each cell at the same time is clearly different, especially in different modules. The curves of cells in Module 1 and Module 2 have relatively better consistency, while in Module 3 the difference between the SOC of batteries increases obviously, showing that cells after clustering and sorting have better homogeneity than those chosen randomly.

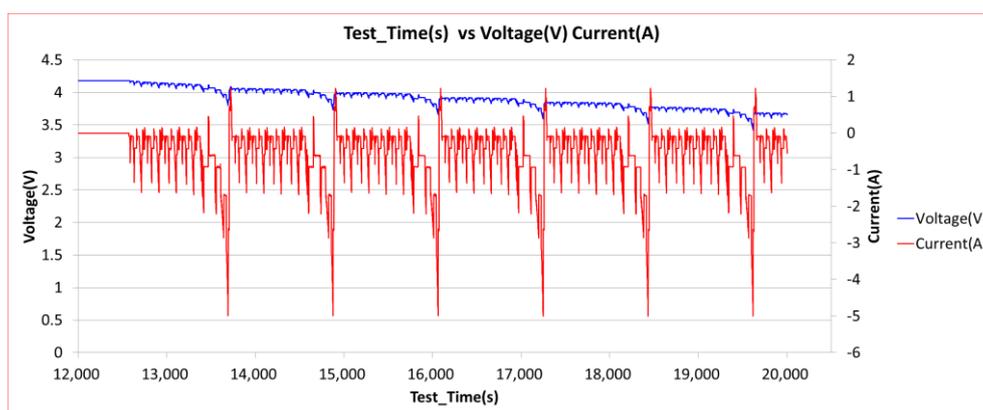


Figure 9. Local display of NEDC working conditions.

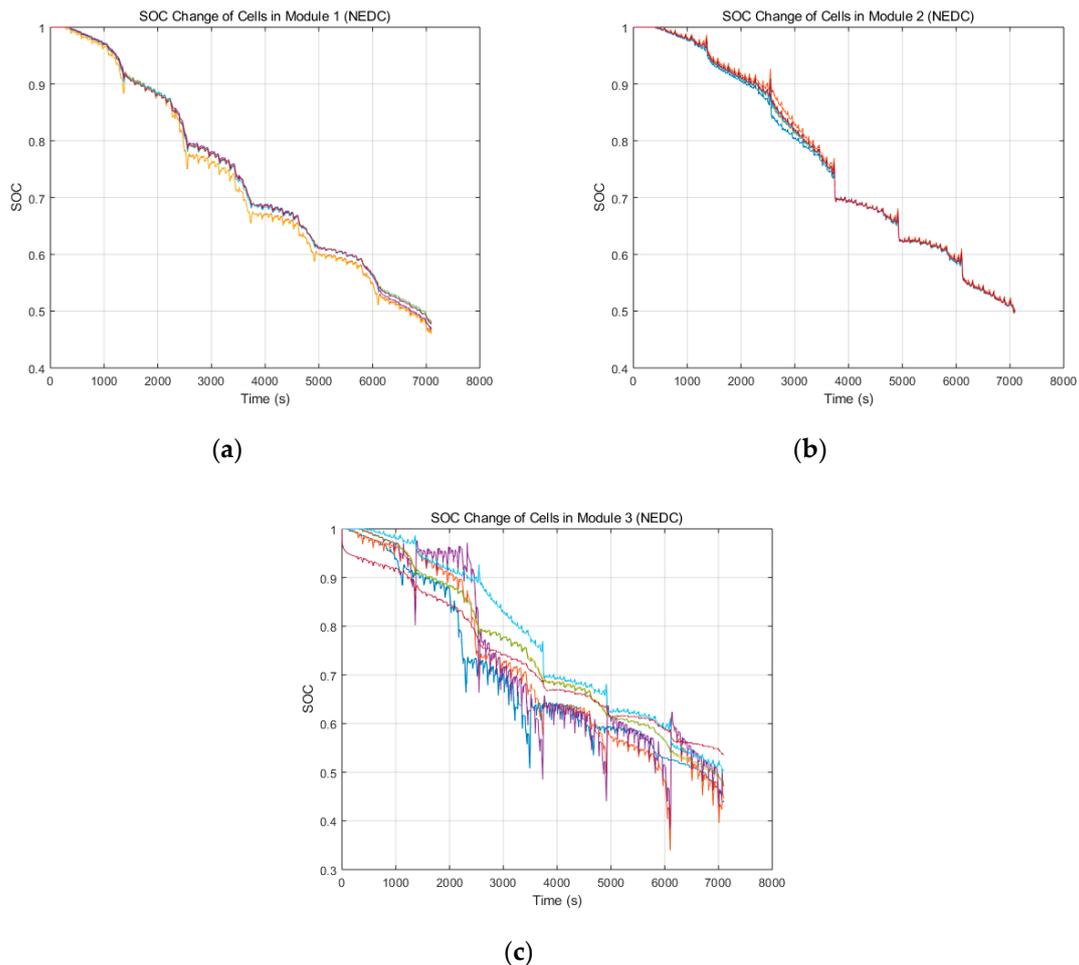


Figure 10. State of charge (SOC) curve of cells in modules: (a) Module 1, (b) Module 2, (c) Module 3.

The same validation procedure is performed under UDDS working conditions to verify the SOC changes in battery modules. The standard deviation of the SOC after load profile of Figure 11 is chosen as another testing criterion. Figure 11, and Tables 6 and 7 illustrate the standard deviation of state of charge of these cells in respective modules.

Table 6. SOC standard deviation of cells in modules under NEDC working conditions.

Module No.	Mean Standard Deviation	Maximum Standard Deviation
Module 1	0.0058	0.0208
Module 2	0.0026	0.0142
Module 3	0.0356	0.1274

Table 7. SOC standard deviation of cells in modules under UDDS working conditions.

Module No.	Mean Standard Deviation	Maximum Standard Deviation
Module 1	0.0069	0.0413
Module 2	0.0034	0.0184
Module 3	0.0472	0.1591

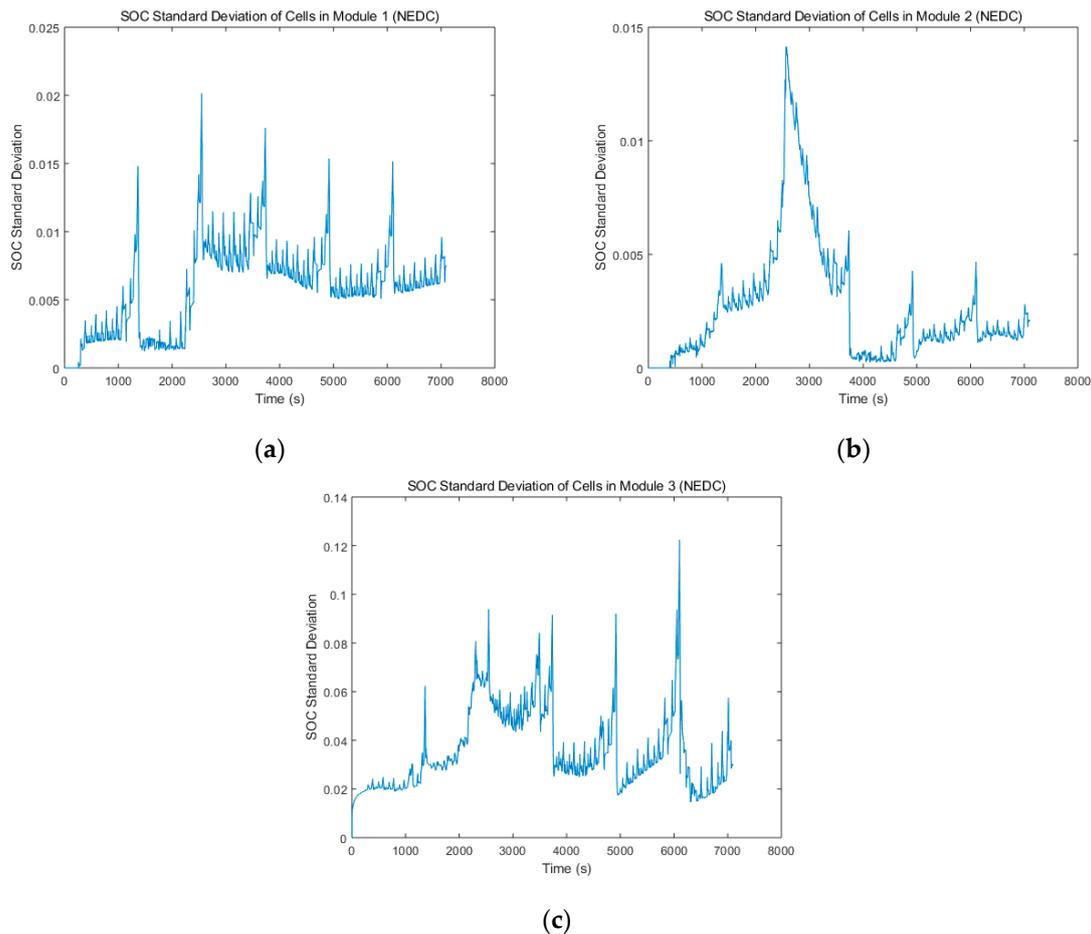


Figure 11. SOC standard deviation curve of cells in modules: (a) Module 1, (b) Module 2, (c) Module 3.

The results clearly show a reduction in the SOC variability after power profile charge and discharge. The SOC standard deviation of Module 1 or Module 2 is notably lighter than that of Module 3, which means there is a significant improvement in the consistency of the battery module.

At the same time, it can be seen that the standard deviation of Module 2 is slightly less than Module 1, indicating the clustering based on PCA gives the best results overall.

6. Conclusions

Effective sorting of lithium batteries is a means to eliminate the inconsistency of battery modules and battery modules. Selecting appropriate sorting parameters and using appropriate sorting algorithms can effectively improve the accuracy and efficiency of battery sorting. This work analyzes the static and dynamic performance of 18650-cylinder lithium battery cell and selects appropriate parameters to form feature characteristics. Based on the unsupervised learning and self-organization characteristics of SOM neural networks and the dimensionality reduction characteristics of PCA, the multi-parameter sorting of lithium battery combining dynamic and static characteristics was carried out. The results show that sorting significantly improves the SOC consistency of batteries in the module after grouping. The consistency after grouping can be further validated by other methods, and the elimination of inconsistency during the using process is also the author's future research direction. Whether different numbers of batteries will affect the accuracy of sorting results, and whether the sorting method can ultimately achieve the purpose that batteries are more accurately sorted, also remains to be further explored by the author and his team.

Author Contributions: B.X. and Y.Y. proposed the SOM sorting algorithm based on PCA analyzed the data; Y.Y. designed the testing experiment and performed the data processing; J.Z. performed testing and simulation; G.C.

and Y.L. (Yifan Liu) provided experiment guidance; H.W., M.W. and Y.L. (Yongzhi Lai) prepared necessary data and equipment; Y.Y. wrote the paper.

Funding: This research was funded by the Shenzhen Economic, Trade and Information Commission of Shenzhen Municipality Strategic Emerging Industries and Future Industrial Development “Innovation Chain + Industrial Chain” Project (2017) and National Natural Science Foundation of China (Grant No. 51877120).

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

References

- Paul, S.; Diegelmann, C.; Kabza, H.; Tillmetz, W. Analysis of ageing inhomogeneities in lithium-ion battery systems. *J. Power Sources* **2013**, *239*, 642–650. [[CrossRef](#)]
- Moore, S.W.; Schneider, P.J. A review of cell equalization methods for lithium ion and lithium Polymer battery systems. In Proceedings of the SAE 2001 World Congress, Detroit, MI, USA, 5–8 March 2001; SAE International: Detroit, MI, USA, 2001.
- Wei, X.; Zhu, B. The research of vehicle power Li-ion battery pack balancing method. In Proceedings of the International Conference on Electronic Measurement and Instruments (ICEMI), Beijing, China, 16–19 August 2009; pp. 2–498.
- Li, J.; Wang, Y.; Tan, X. Research on the Classification Method for the Secondary Uses of Retired Lithium-ion Traction Batteries. *Energy Procedia* **2017**, *105*, 2843–2849. [[CrossRef](#)]
- Kim, J.; Cho, B.H. Screening process-based modeling of the multi-cell battery string in series and parallel connections for high accuracy state-of-charge estimation. *Energy* **2013**, *57*, 581–599. [[CrossRef](#)]
- An, F.; Huang, J.; Wang, C.; Li, Z.; Zhang, J.; Wang, S.; Li, P. Cell sorting for parallel lithium-ion battery systems: Evaluation based on an electric circuit model. *J. Energy Storage* **2016**, *6*, 195–203. [[CrossRef](#)]
- Li, X.; Wang, T.; Pei, L.; Zhu, C.; Xu, B. A comparative study of sorting methods for Lithium-ion batteries. In Proceedings of the 2014 IEEE Transportation Electrification Conference and Expo, ITEC Asia-Pacific 2014, Beijing, China, 31 August–3 September 2014; Institute of Electrical and Electronics Engineers Inc.: Beijing, China, 2014.
- Jong-Hoon, K.; Jong-Won, S.; Chang-Yoon, J.; Bo-Hyung, C. Screening process of Li-Ion series battery pack for improved voltage/SOC balancing. In Proceedings of the 2010 International Power Electronics Conference (IPEC), Sapporo, Japan, 21–24 June 2010; pp. 1174–1179.
- Hartigan, A.J. *Clustering Algorithms*; John Wiley & Sons: Hoboken, NJ, USA, 1975; Volume 31, p. 793.
- Miyamoto, S. Data clustering algorithms for information systems. In Proceedings of the 11th International Conference on Rough Sets, Fuzzy Sets, Data Mining, and Granular Computer, RSFDGrC 2007, Toronto, ON, Canada, 14–17 May 2007; Springer: Toronto, ON, Canada, 2007; pp. 13–24.
- Guo, L.; Liu, G.W. Research of Lithium-ion Battery Sorting Method Based on Fuzzy C-Means Algorithm. In *Progress in Power and Electrical Engineering, Pts 1 and 2*; Zhang, H., Fu, Y., Tang, Z., Eds.; Trans Tech Publications Ltd.: Stafa-Zurich, Switzerland, 2012; Volume 354–355, pp. 983–988.
- Zhang, R.; Zhou, Y.; Li, R. A Battery Sorting Scheme Based on Fuzzy C-mean Clustering, Taking Advantage of the Flatness of Discharge Voltage Curve. *Qiche Gongcheng Automot. Eng.* **2017**, *39*, 864.
- Kumar, U.A.; Dhamija, Y. Comparative Analysis of SOM Neural Network with K-means Clustering Algorithm. In Proceedings of the 2010 IEEE International Conference on Management of Innovation & Technology (ICMIT 2010), Singapore, 2–5 June 2010; pp. 55–59.
- He, F.X.; Shen, W.X.; Song, Q.; Kapoor, A.; Honnery, D.; Dayawansa, D. Self-organising map based classification of LiFePO₄ cells for battery pack in EVs. *Int. J. Veh. Des.* **2015**, *69*, 151–167. [[CrossRef](#)]
- Fang, K.Z.; Chen, S.; Mu, D.B.; Wu, B.R.; Wu, F. Investigation of nickel-metal hydride battery sorting based on charging thermal behavior. *J. Power Sources* **2013**, *224*, 120–124. [[CrossRef](#)]
- Raspa, P.; Frinconi, L.; Mancini, A.; Cavalletti, M.; Longhi, S.; Fulimeni, L.; Bellesi, P.; Isidori, R. Selection of Lithium cells for EV battery pack using Self-Organizing maps. In Proceedings of the 25th World Battery, Hybrid and Fuel Cell Electric Vehicle Symposium and Exhibition: Sustainable Mobility Revolution, EVS 2010, Shenzhen, China, 5–9 November 2010; Electric Drive Transportation Association: Shenzhen, China, 2010.
- Miyatake, S.; Susuki, Y.; Hikiyama, T.; Itoh, S.; Tanaka, K. Discharge characteristics of multicell lithium-ion battery with nonuniform cells. *J. Power Sources* **2013**, *241*, 736–743. [[CrossRef](#)]

18. Huang, J.; Li, Z.; Zhang, J.B. Dynamic electrochemical impedance spectroscopy reconstructed from continuous impedance measurement of single frequency during charging/discharging. *J. Power Sources* **2015**, *273*, 1098–1102. [[CrossRef](#)]
19. Fang, K.Z.; Mu, D.B.; Chen, S.; Wu, B.R.; Wu, F. A prediction model based on artificial neural network for surface temperature simulation of nickel-metal hydride battery during charging. *J. Power Sources* **2012**, *208*, 378–382. [[CrossRef](#)]
20. Wang, H.; Dong, K.; Li, G.; Wu, F. Research on the consistency of the power battery based on multi-points impedance spectrum. In Proceedings of the 2010 International Forum on Strategic Technology, IFOST 2010, Ulsan, Korea, 13–15 October 2010; IEEE Computer Society: Washington, DC, USA, 2010; pp. 1–4.
21. Jolliffe, I. *Principal Component Analysis and Factor Analysis*; Springer: Berlin, Germany, 1986; pp. 115–128.
22. Kohonen, T. *Self-Organizing Maps*, 2nd ed.; Springer: Berlin, Germany, 1995.
23. He, H.W.; Xiong, R.; Zhang, X.W.; Sun, F.C.; Fan, J.X. State-of-Charge Estimation of the Lithium-Ion Battery Using an Adaptive Extended Kalman Filter Based on an Improved Thevenin Model. *IEEE Trans. Veh. Technol.* **2011**, *60*, 1461–1469.
24. Xia, B.Z.; Guo, S.K.; Wang, W.; Lai, Y.Z.; Wang, H.W.; Wang, M.W.; Zheng, W.W. A State of Charge Estimation Method Based on Adaptive Extended Kalman-Particle Filtering for Lithium-ion Batteries. *Energies* **2018**, *11*, 2755. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).