Clustering Optimization Method for Liquid Metal Battery Screening Requirements

A. E Zhang[1], B. Kangli Wang[1*], and C. Kai Jiang[1]

[1] State Key Laboratory of Advanced Electromagnetic Engineering and Technology, School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China

Abstract-- **Liquid metal battery (LMB) exhibits the potential to appear as a cost-effective solution for grid-scale energy storage to improve the stability and flexibility of new power systems. In order to improve the consistency and reliability of liquid metal battery packs, a novel classification method is proposed in this paper. Firstly, a feature vector for battery screening is established from the key influencing factors of battery pack performance. Secondly, the density-based spatial clustering of applications with noise (DBSCAN) algorithm whose input parameters have been optimized by the elbow method is used to detect outliers among sample data. Then, the remaining data are screened by the Mean shift algorithm and the Davies-Bouldin Index (DBI) as the optimized objective. Finally, an experimental of 216 laboratory-made LMBs with the proposed method is presented and the clustering optimization results verify that the proposed method can automatically determine the number of clustering and improve the accuracy of clustering with outlier detection.**

Index Terms-- **Liquid metal battery, battery screening, DBSCAN algorithm, outliers, Mean shift algorithm.**

I. INTRODUCTION

LMB is constructed with a novel three-liquid-layer structure, the negative electrode, the molten salt electrolyte, and the positive electrode from top to bottom as shown in Fig. 1. Liquid electrodes do not have the effects of deformation or dendrite growth of conventional solid electrodes, which contributes to LMB's long cycle life [1, 2]. Simpson et al. [3] compared the performance specifications for LMB, Li-ion, and Lead-acid battery systems configured for a large-scale application, showing that LMB options have much lower cost and higher cycle life than the traditional battery options. As a new type of electrochemical energy storage technology, those advantages make LMBs promising for applications in stationary energy storage [4, 5]. Furthermore, Ambri, a company producing LMB, has received customer orders to design and develop energy storage systems based on LMB technology. As can be seen, LMB is approaching commercial production.

However, due to differences in manufacturing processes, there are inevitable inconsistencies between individual batteries [6]. As a result, battery packs that are series/parallel-connected by lots of batteries have a clear "barrel effect" where the overall performance of the pack is limited by the lowest-performing individual battery. Furthermore, the further expansion of the inconsistency will accelerate the performance degradation of the battery

pack during its operation [7]. The establishment of battery screening methods is one of the key technologies to solve these problems.
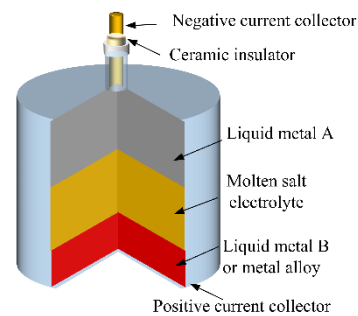

Fig. 1. The structure of LMB.

In the current research on battery screening, there are two main aspects: one is to choose indexes that can accurately characterize the battery performances and another is the corresponding screening algorithm [8]. Zheng et al. [9] tests the self-discharge rate of batteries by connecting them in parallel and measures the consistency among batteries based on the self-discharge test results. However, a single screening index cannot fully reflect the characteristics of the battery. Lai et al. [10] extracted internal resistance and capacity from partial discharge curves, and then these parameters are taken to screen batteries. But this method could not ensure the batteries' consistent performance during further operation. Indeed, screening indices should consider battery life trajectories that represent the battery's future operation state. Therefore, Jiang et al. [11] proposed a screening strategy based on the aging mechanism of lithium-ion batteries (LIBs) employing peak values on the incremental capacity (IC) curve to identify the capacity loss variations. Aihua et al. [12] explored the mechanisms of the pulse test to reflect the aging degree of LIBs and then took the pulse voltage as one of the features to cluster. However, different types of batteries have different aging behavior owing to their inherent electrochemical characteristics. Thus each type of battery is in great need of targeted multidimensional screening indices that can comprehensively assess the battery or battery pack states over the full life cycle.

After screening indexes extracted from battery characteristics, the key question is how to use these indexes to sort them. This is a classic clustering problem or multidimensional classification. Lai et al. [13] suggested an improved K-means algorithm which was achieved by modifying the Euclidean distance formula in

the conventional K-means algorithm for the two-dimensional classification of indices to consider echelon utilization scenarios. Zhou et al. [14] trained the multi-class model based on a support vector machine to accurately screen the retired LIBs with good consistency in capacity. Lai et al. [15] constructed a neural network model to screen the retired LIBs in large sample cases. However, these algorithms are of the complete clustering algorithm, ignoring the effect of ill-behaved batteries on the clustering results.

Most existing battery screening methods take LIBs as the research objects. However, the inherent characteristics of LMBs are somewhat different from those of LIBs. Therefore, it is necessary to research screening strategies for the practical application of LMBs. The screening indexes need to consider the typical characteristics of LMBs, and the clustering process requires an optimization strategy considering the characteristics of the data set constructed by screening indexes. This paper proposes a combinatorial clustering optimization algorithm based on battery performance characteristics to achieve effective screening of LMBs, laying the foundation for subsequent large-scale LMBs applications.

The remainder of this paper is organized as follows: in Section II, the cycling performance for the LMB pack is analyzed, and multidimensional indices are established for screening. In Section III, the screening algorithm is proposed for effectively screening internal resistance, capacity, and CE. The experiment results are introduced in Section IV. Section V presents the main conclusions.

II. EXPERIMENT DATA AND PROCESSING

A. Sorting index

The purpose of battery screening is to select amounts of consistent cells to optimize the initial performance of the battery pack consisting of cells connected in series/parallel. The capacity of a battery pack is one of the indicators that can reflect the state of health (SOH) of the battery pack [16]. In order to ensure the safe operation of the battery pack, the battery pack charging cut-off condition is usually set to one of the batteries in the battery pack (recorded as Cell A) reaching the charging cut-off voltage (marked as V_C). Likewise, the discharge cut-off condition is set to one of the in-pack batteries (recorded as Cell B) reaching the discharge cut-off voltage (marked as V_D). Therefore, it can be seen that the charging capacity of the battery pack is equal to the rechargeable electric quantity of Cell A (marked as E_C) because Cell A has the minimum rechargeable electric quantity, the discharge capacity of the battery pack is equal to the dischargeable electric quantity of Cell B (marked as E_D) because Cell B has the minimum dischargeable electric quantity as shown in Equation (1).

$$\begin{cases} Q_{\text{Pack\_C}} = E_{\text{Cell A\_C}} = \int_{t|V=V_D}^{t|V=V_C} |I| \, dt \\ Q_{\text{Pack\_D}} = E_{\text{Cell B\_D}} = \int_{t|V=V_C}^{t|V=V_D} |I| \, dt \end{cases} \quad (1)$$

where $Q_{\text{Pack\_C}}$ is the charge capacity of the battery pack and $Q_{\text{Pack\_D}}$ is the discharge capacity of the battery pack. Assuming the coulombic efficiency of the battery is 100% (conventional method of calculating the capacity of a battery pack), Equation (1) can be simplified as follows:

$$\begin{aligned} Q_{\text{Pack}} &= \int_{t|V=V_D}^{t_0} |I| \, dt + \int_{t_0}^{t|V=V_C} |I| \, dt \\ &= \min\{E_D\} + \min\{E_C\} \end{aligned} \quad (2)$$

Unlike LIB, many other batteries, such as Li Metal-based batteries, LMBs, etc. do not have very high coulombic efficiency. Thus, the accurate definition of the battery pack capacity should consider the factor of coulombic efficiency. Taking the working current as I and the changed electric quantity as ΔE, the changed electric quantities for all series-connected batteries should be identical, because they share the same working current. Thus. Considering the coulombic efficiency is independent of SOC which is reasonable for LMB, the battery pack charge/discharge capacity can be represented as

$$\begin{cases} Q_{\text{Pack\_C}} = \int_{t|V=V_D}^{t_0} |I| \, dt + \int_{t_0}^{t|V=V_C} |I| \, dt \\ \quad = \min\{E_D\}/\eta_{\text{minE}_D} + \min\{E_C\} \\ Q_{\text{Pack\_D}} = \int_{t|V=V_C}^{t_0} |I| \, dt + \int_{t_0}^{t'|V=V_D} |I| \, dt \\ \quad = \eta_{\text{minE}_C} \times \min\{E_C\} + \min\{E_D\} \end{cases} \quad (3)$$

where η_{minE_C} represents the coulombic efficiency of the in-pack cell with minimum E_C (Cell A), η_{minE_D} represents the coulombic efficiency of the in-pack cell with minimum E_D (Cell B). Furthermore, the coulombic efficiency of the battery pack can be shown as follows:

$$\begin{cases} \eta_{\text{Pack}} = \dfrac{E_D}{E_C} = \eta_{\text{minE}_C} \left(Q_{\text{Pack\_C}} \text{measured after } Q_{\text{Pack\_D}} \right) \\ \eta_{\text{Pack}} = \dfrac{E_D'}{E_C} = \eta_{\text{minE}_D} \left(Q_{\text{Pack\_C}} \text{measured before } Q_{\text{Pack\_D}} \right) \end{cases} \quad (4)$$

Fig. 2 shows the cycling performance of an LMB, which has a relatively low coulombic efficiency but negligible capacity loss during the cycle. However, it is pretty normal that there exist coulombic efficiency differences among in-pack batteries. Fig. 3 shows the cycling performance of an LMB pack with an evident pack dischargeable capacity loss. It can be observed that, unlike single LMB, the characteristics of the actual capacity of the LMB pack are significantly affected by the coulombic efficiency inconsistency of the in-pack LMBs. The most direct factors affecting the performance of a battery pack are capacity, internal resistance, and coulombic efficiency. Battery capacity determines the output initial capacity of the battery pack, internal resistance is the main parameter to assess the power performance of the battery, and the cycling performance of the liquid metal battery pack is mainly affected by the coulomb efficiency of the battery. Therefore, in this paper, multidimensional screening indices (capacity, internal resistance, and coulombic

efficiency) are extracted from the key factors affecting the performance of the battery pack. The capacity and internal resistance can represent the battery's current state, and coulomb efficiency can reflect the future state of the battery pack. These indices can be used to comprehensively assess the LMB states over the full life cycle in different dimensions, which will improve the safety and economy of LMB pack utilization.
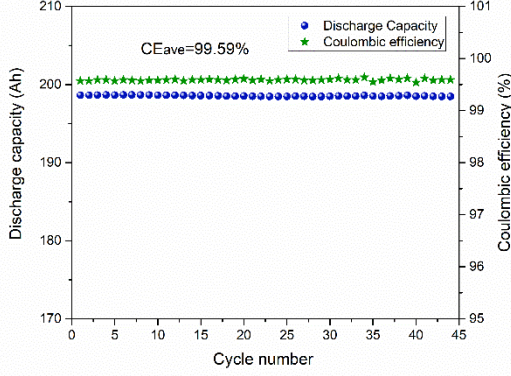

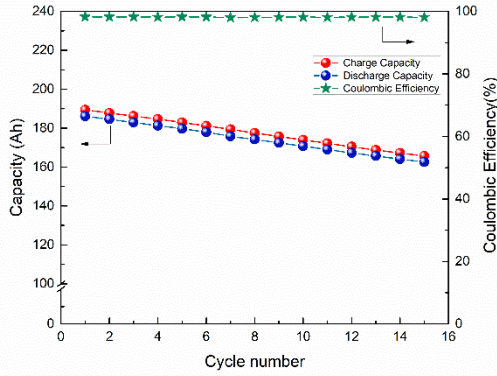Fig. 2. The cycling performance of an LMB.

Fig. 3. The cycling performance of an LMB pack.

B. Data processing

There exist differences in magnitude and order of magnitude between cell capacity, internal resistance, and coulombic efficiency. In order to balance the influence of the above screening indexes on the clustering results, it is necessary to standardize them before clustering. This paper adopts zero-mean standardization as follows:

$$x^* = \frac{x - \mu}{\sigma} \tag{5}$$

where x represents the cell capacity x_C, internal resistance x_R, and coulombic efficiency x_{CE}, respectively, μ and σ are the mean and variance of each index. After normalization, each cell corresponds to a 3-dimensional feature vector, denoted as $x_i^* = (x_C^{i*}, x_R^{i*}, x_{CE}^{i*})$ where i represents the battery serial number. The entire set of sample data S is denoted as $\{x_1^*, x_2^*, x_3^*, \cdots, x_N^*\}$, where N represents the total number of batteries.

III. CLUSTERING METHODS

In general, the data aggregation pattern consisting of multiple parameters of the battery has no clear distribution boundaries and the density is unevenly distributed. For the classification of the above sample data, it is more suitable to choose the density-based clustering algorithm. The DBSCAN algorithm is a typical density-based clustering algorithm, which is good at finding outliers [17]. However, its clustering results are highly dependent on the setting of the density threshold (*Eps*, *MinPts*). The Mean Shift is another density-based non-parametric clustering algorithm, which requires only one input parameter (*Bandwidth*) and does not require setting the number of clusters [18]. This paper combines the advantages of the above two algorithms and proposes a combined clustering algorithm. The flow chart of the combined clustering algorithm is shown in Fig. 4.
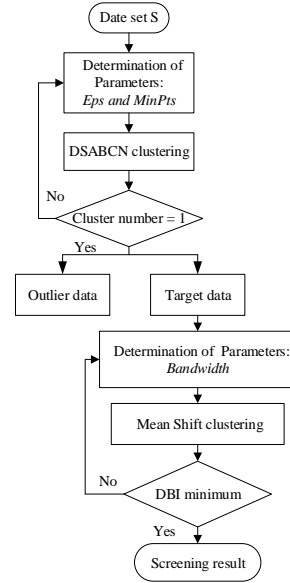

Fig. 4. The flow chart of the combined clustering algorithm.

A. DBSCAN for detecting outliers

The DBSCAN algorithm is very sensitive to the density threshold parameters. In this paper, the elbow method combined with a specified cluster number is adopted to achieve the optimal search for the density threshold parameters. Specific algorithm steps are as follows:

a) Calculate the k-distance. The distance between each point and the remaining points is first calculated by the formula：

$$d(x_i^*, x_j^*) = \sqrt{(x_C^{j*} - x_C^{i*})^2 + (x_R^{j*} - x_R^{i*})^2 + (x_{CE}^{j*} - x_{CE}^{i*})^2}$$
$$j \in \{1, 2, \cdots, i-1, i+1, \cdots, N\}$$
$$\tag{6}$$

Then, sort the distance sets ascending：$D(i) = \{d(i,1), d(i,2), \ldots, d(i,k), \ldots d(i,N)\}$, $d(i,k)$ is referred to as the k-distance, where $k = MinPts$. The initial value of *MinPts* is set to the data dimension plus 1 in this paper, i.e. *MinPts* = 4.

b) Determination of optimal *Eps* value. Sort the k-distances of all points ascending and plot it on a k-distance

graph, *Eps* corresponds to the maximum curvature in the curve.

c)   DBSCAN clustering. If the number of clusters is 1, then skip to step d), otherwise, update *MinPts* and return to step a), the *MinPts* update formula is as follows:

$$MinPts = MinPts + 1 \qquad (7)$$

d) Output the clustering result, i.e. the target sample data.

B. Battery screening by Mean Shift clustering

In order to evaluate the clustering effect more scientifically, while optimally dividing the target sample data, for the density-based clustering algorithm, this paper chooses the Davies-Bouldin Index (DBI) as the objective function, as shown in Equation (8):

$$DBI = \frac{1}{n}\sum_{i=1}^{n} \max_{j \neq i} \left(\frac{S_i + S_j}{d_{ij}} \right) \qquad (8)$$

where *n* is the number of clusters of points after clustering, S_i and S_j are the average distances of data points within cluster *i* and cluster *j* to the respective cluster centroids, and d_{ij} is the distance between the centroids of cluster *i* and cluster *j*. The algorithm steps are shown as follows:

a)   Set the initial value of *Bandwidth* parameter.

b)   Perform Mean shift clustering, output the clustering result and calculate the *DBI* according to Equation (8), denoted as DBI_m, *m* indicates the number of iterations.

c)   If $DBI_m < DBI_{m+1}$, then skip to step d), otherwise update *Bandwidth* and return to step b), the update formula of *Bandwidth* is:

$$Bandwidth = 0.98 * Bandwidth \qquad (9)$$

d) Output the results of the last iteration of clustering.

IV. EXPERIMENTS AND RESULTS

The experimental platform consists of a Land battery test system (CT3001A) and a computer. The accuracy of the current in our test bench is 0.05% which is considered to be accurate enough for capacity measurement. The laboratory-made 200Ah-level LMBs are used in the experiments. The detailed parameters of the battery are shown in Table 1.

Table 1
Liquid metal battery parameters.

Item	parameter
Nominal current	0.2 C
Charging cut-off voltage	1.2 V
Discharging cut−off voltage	0.55 V
Operation temperature	550 ℃

The constant current constant voltage (CCCV) profile is used to extract battery capacity and coulombic efficiency. The Hybrid Pulse Power Characteristic (HPPC) test is used to extract battery internal resistance. 216 LMBs were selected as samples, and Fig. 5 (a), (b), (c) show the battery capacity, internal resistance, and coulombic

efficiency of 216 LMBs, respectively. The parameters of this batch of batteries are very scattered, for example, capacity ranging from 198.25 Ah to 206.75 Ah, so clustering is important. The data sample after standardization is shown in Fig. 5(d).
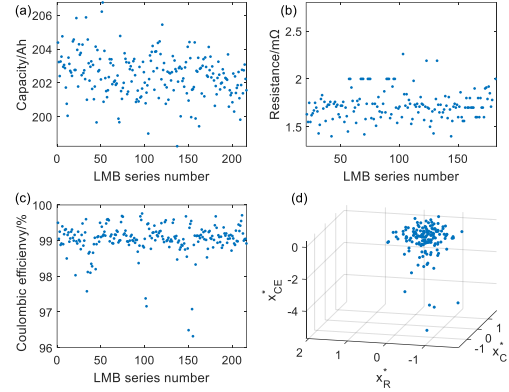

Fig. 5. (a) Capacity, (b) internal resistance, (c) coulombic efficiency of 216 LMBs. (d) The sample data after standardization.

Fig. 6 shows optimized density threshold parameters. The k-distance curve is calculated when *MinPts* = 5, and the best value of *Eps* that can be obtained by the elbow detection method is 0.6236.
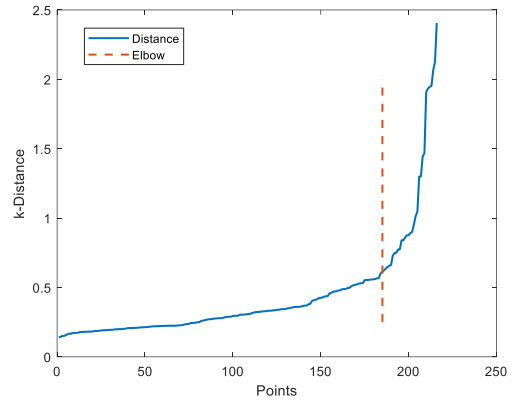

Fig. 6. k-distance curve to find suitable values of Eps.

The final DBSCAN algorithm clustering results are shown in Fig. 7, in which the algorithm classifies 28 noise points regarded as outliers. Then, the remaining data are clustered by Mean shift and the clustering results are shown in Fig. 8, with a total of 4 clusters.
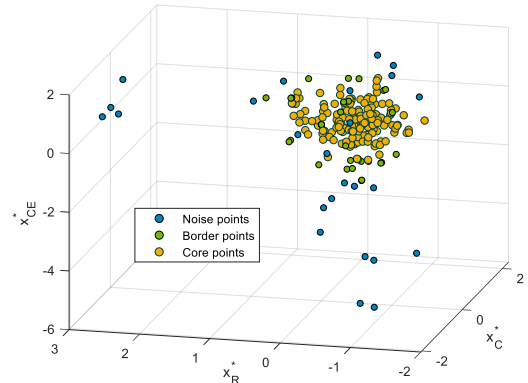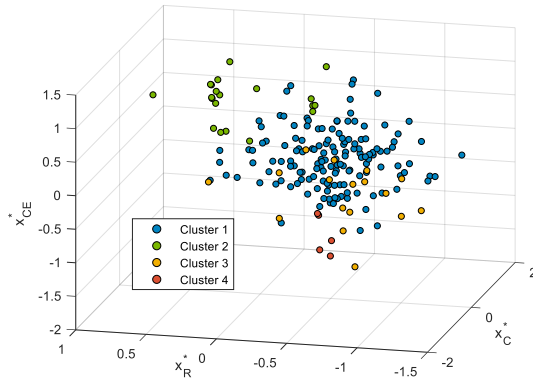

Fig. 7. DBSCAN clustering result.

Fig. 8. Battery screening results.

To further quantify the intra-cluster consistency, the range and mean value of the corresponding indicators are calculated for each cluster separately, as shown in Table 2. The results of the four clusters are analyzed, the internal resistance and coulombic efficiency of Clusters 1 and 3 are near, just the capacity of Cluster 1 is higher than that of Cluster 3. And Cluster 2 has the highest internal resistance and coulombic efficiency, Cluster 4 has the lowest coulombic efficiency, just the capacity of Cluster 4 is higher than that of Cluster 2. It can be seen that a higher degree of consistency is found in the battery characteristic indexes of each cluster. It can be concluded that the proposed screening method can improve the consistency of the batteries in the cluster and can help ensure the uniformity of batteries during long-term service.

Table 2
Comparative analysis of battery data before and after battery screening.

	Number of cells	Capacity (Ah)		Internal Resistance (mΩ)		Coulombic efficiency (%)	
		Range	Mean value	Range	Mean value	Range	Mean value
Target data	216	8.5	202.45	1.73	1.75	3.44	99.02
Cluster 1	146	4.42	202.84	0.7	1.71	0.82	99.07
Cluster 2	20	2.46	201.78	0.49	1.93	0.46	99.53
Cluster 3	17	1.67	200.60	0.6	1.61	0.77	99.01
Cluster 4	5	1.89	202.66	0.06	1.73	0.31	98.55

V. Conclusions

In this paper, battery screening indexes are extracted from the key influencing factors of battery pack performance. Then, a combined clustering algorithm is proposed: DBSCAN algorithm combinded with elbow method is used for detecting outliers and the Mean shift algorithm is used to divide the remaining data into different clusters. Finally, a clustering experiment of 216 LMBs is performed and statistical analysis is provided. The results show that the combined algorithm is an effective approach to obtain the best division of the original data and thus improve the clustering performance.

Acknowledgment

References

[1] H. Zhou, H. Li, Q. Gong, S. Yan, X. Zhou, S. Liang, W. Ding, Y. He, K. Jiang, and K. Wang, "A sodium liquid metal battery based on the multi-cationic electrolyte for grid energy storage," *Energy Storage Materials,* vol. 50, pp. 572-579, 2022.

[2] X. Zhou, H. Zhou, S. Yan, Y. He, W. Zhang, H. Li, K. Wang, and K. Jiang, "Increasing the actual energy density of Sb-based liquid metal battery," *Journal of Power Sources,* vol. 534, p. 231428, 2022.

[3] J. G. Simpson, G. Hanrahan, E. Loth, G. M. Koenig, and D. R. Sadoway, "Liquid metal battery storage in an offshore wind turbine: Concept and economic analysis," *Renewable and Sustainable Energy Reviews,* vol. 149, p. 111387, 2021.

[4] D. Agarwal, R. Potnuru, C. Kaushik, V. R. Darla, K. Kulkarni, A. Garg, R. K. Gupta, N. Tiwari, and K. S. Nalwa, "Recent advances in the modeling of fundamental processes in liquid metal batteries," *Renewable and Sustainable Energy Reviews,* vol. 158, p. 112167, 2022.

[5] H. Li, K. Wang, H. Zhou, X. Guo, S. Cheng, and K. Jiang, "Tellurium-tin based electrodes enabling liquid metal batteries for high specific energy storage applications," *Energy Storage Materials,* vol. 14, pp. 267-271, 2018.

[6] L. Wildfeuer and M. Lienkamp, "Quantifiability of inherent cell-to-cell variations of commercial lithium-ion batteries," *eTransportation,* vol. 9, p. 100129, 2021.

[7] Y. Lu, K. Li, X. Han, X. Feng, Z. Chu, L. Lu, P. Huang, Z. Zhang, Y. Zhang, F. Yin, X. Wang, F. Dai, M. Ouyang, and Y. Zheng, "A method of cell-to-cell variation evaluation for battery packs in electric vehicles with charging cloud data," *eTransportation,* vol. 6, p. 100077, 2020.

[8] S.Ma, J. Li, Y. Li, M.Chen, B. Zhang, " Customized clustering optimization method for battery reutilization screening requirements," Proceedings of the CSEE, vol. 42, pp. 6208-6220, 2022.

[9] Y. Zheng, H. Wu, W. Yi, X. Lai, H. Dai, X. Han, and M. Ouyang, "A novel classification method of commercial lithium-ion battery cells based on fast and economic

detection of self-discharge rate," *Journal of power sources,* vol. 478, p. 229039, 2020.

[10] X. Lai, C. Deng, J. Li, Z. Zhu, X. Han, and Y. Zheng, "Rapid sorting and regrouping of retired lithium-ion battery modules for echelon utilization based on partial charging curves," *IEEE Transactions on Vehicular Technology,* vol. 70, pp. 1246-1254, 2021.

[11] T. Jiang, J. Sun, T. Wang, Y. Tang, S. Chen, S. Qiu, X. Liu, S. Lu, and X. Wu, "Sorting and grouping optimization method for second-use batteries considering aging mechanism," *Journal of energy storage,* vol. 44, p. 103264, 2021-01-01 2021.

[12] A. Ran, Z. Liang, S. Chen, M. Cheng, C. Sun, F. Ma, K. Wang, B. Li, G. Zhou, X. Zhang, F. Kang, and G. Wei, "Fast clustering of retired lithium-ion batteries for secondary life with a two-step learning method," *ACS Energy Letters,* vol. 7, pp. 3817-3825, 2022.

[13] X. Lai, Y. Huang, C. Deng, H. Gu, X. Han, Y. Zheng, and M. Ouyang, "Sorting, regrouping, and echelon utilization of the large-scale retired lithium batteries: A critical review," *Renewable and Sustainable Energy Reviews,* vol. 146, p. 111162, 2021.

[14] Z. Zhou, B. Duan, Y. Kang, Y. Shang, N. Cui, L. Chang, and C. Zhang, "An efficient screening method for retired lithium-ion batteries based on support vector machine," *Journal of Cleaner Production,* vol. 267, p. 121882, 2020.

[15] X. Lai, D. Qiao, Y. Zheng, M. Ouyang, X. Han, and L. Zhou, "A rapid screening and regrouping approach based on neural networks for large-scale retired lithium-ion cells in second-use applications," *Journal of cleaner production,* vol. 213, pp. 776-791, 2019.

[16] A. Cordoba-Arenas, S. Onori, Y. Guezennec, and G. Rizzoni, "Capacity and power fade cycle-life model for plug-in hybrid electric vehicle lithium-ion battery cells containing blended spinel and layered-oxide positive electrodes," *Journal of Power Sources,* vol. 278, pp. 473-483, 2015.

[17] N. Rahmah and I. S. Sitanggang, "Determination of optimal epsilon (Eps) value on dbscan algorithm to clustering data on peatland hotspots in sumatra," *IOP conference series. Earth and environmental science,* vol. 31, pp. 12012-12015, 2016.

[18] A. E. Ezugwu, A. M. Ikotun, O. O. Oyelade, L. Abualigah, J. O. Agushaka, C. I. Eke, and A. A. Akinyelu, "A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects," *Engineering Applications of Artificial Intelligence,* vol. 110, p. 104743, 2022.