Secondary Structural Ensemble Learning Cluster for Estimating the State of Health of Lithium-Ion Batteries
Si-zhe Chen, Hongtao Zhang, Long Zeng,*, Yuanliang Fan, Le Chang, and Yun Zhang

ABSTRACT: Accurate online state-of-health (SOH) estimation can improve the operational efficiency of lithium-ion batteries (LIBs) and ensure the safety of energy storage systems. However, the complex electrochemical properties of LIBs make accurate SOH estimation challenging. To overcome this challenge, we propose a secondary structural ensemble learning (SSEL) cluster. The proposed SSEL cluster includes multiple SSEL frameworks established separately within different SOH data intervals, allowing the identification of stable feature–SOH relationships. The adaptability and basic accuracy of each SSEL framework are guaranteed by various base learners and the corresponding stacking model and bagging model fusion. Each framework remains unique and specialized owing to the adoption of back propagation neural networks, which adjust learner weights based on the feature–SOH relationship at each interval. The effectiveness of the SSEL cluster was verified using the Oxford Battery Degradation Dataset 1. Comparisons showed that the proposed estimation method performs better than traditional machine learning methods.

1. INTRODUCTION
Lithium-ion batteries (LIBs) have several advantages, including high energy density, a long service life, and a lack of memory effect.1−6 Therefore, they are widely used in electric vehicles and portable devices and are also used for energy storage in power systems. However, non-emergency usage causes the continuous aging of LIBs, decreasing their actual capacity, affecting battery performance, and even causing serious accidents. Therefore, accurate state-of-health (SOH) estimation is critical for the safe use of LIBs.

Current SOH prediction methods for LIBs can be classified into three types: direct measurement methods, model-based methods, and data-driven methods.

1 In the direct measurement method,7,8 SOH is evaluated based on aging-related characteristics, such as the capacity and internal resistance of the LIB. Although this method is centered on basic mechanisms and can be used for several types of batteries, high measurement accuracy can be achieved only under regulated laboratory conditions.

2 To overcome the drawbacks of the direct measurement method, a model-based method was developed. This method uses empirical knowledge9−14 to establish the relationship between measured battery signals and the SOH. However, the model parameters need to be adjusted as the battery ages, making computations increasingly complex.

3 In the data-driven method, accurate SOH estimations can be made based on operational data, without analyzing internal chemical reactions. The data-driven method is widely used for SOH estimation. Hu et al.15 showed that the K-nearest neighbor (KNN) algorithm can illustrate the relationship between battery features and the SOH, even though its multi-dimensional data processing performance is poor. Li et al.16 developed a new random forest model that predicts LIB SOH using raw data without any preprocessing. Furthermore, because neural networks (NNs) have a strong nonlinear fitting ability, Shen et al.17 employed a NN to evaluate SOH. Although both methods could accurately evaluate the SOH, they were prone to overfitting. Hence, support vector regression (SVR) was adopted to estimate the SOH, and good generalization to an unknown dataset was observed.18 Nevertheless, the predicted performance of SVR remained sensitive to missing data.19

To predict SOH with satisfactory accuracy, data-driven methods have been combined. A hybrid model consisting of extreme learning machine and random vector functional link networks was applied to predict SOH, ensuring efficient and accurate estimation.20 In another study,21 the weight-sharing feature of convolutional NNs (CNNs) and the sequential correlation of long short-term memory were jointly applied to improve SOH estimation accuracy. In the study by Dai et al.,22 special features were used as inputs for a feed-forward NN, and

Received: March 16, 2022
Accepted: April 26, 2022
Published: May 10, 2022
2. DATA PREPROCESSING

2.1. SOH Definition. SOH is represented by the current capacity and rated capacity, as follows

\[
SOH = \frac{C_{\text{now}}}{C_{\text{new}}} \tag{1}
\]

where \(C_{\text{now}}\) and \(C_{\text{new}}\) are the current and rated capacity of the LIB, respectively.

2.2. Experimental Data. Oxford Battery Degradation Dataset 1, which contains long-term battery aging data from eight Kokam (SLPB533459H4) 740 mAh lithium-ion pouch batteries, was used to evaluate the algorithms. These batteries were tested using a constant-current–constant-voltage (CC–CV) charging profile and a drive cycle discharging profile obtained from the Artemis urban cycle. The testing temperature was maintained at 40 °C. Characteristic measurements were taken every 100 cycles, with a 1 C charge and a 1 C discharge current. The aging curves are shown in Figure 1.

2.3. Feature Extraction. The powerful estimation ability of the data-driven method comes from the close relationship between model features and SOH; hence, the selection of appropriate features is imperative. The voltage, current, and temperature of a LIB during CC–CV charging are closely related to battery degradation. Moreover, these physical quantities are also easy to measure. Therefore, these parameters were chosen as the features for the SOH estimation of LIBs.

The typical CC–CV charging curves of LIBs are shown in Figure 2. The voltage, current, and temperature curves were divided into 150 time-based subsections, and the average voltage, current, and temperature within each subsection were calculated as feature data.

2.4. GAN Data Expansion. The data-driven approach requires a large amount of experimental data for model training. A GAN can be used to overcome the problem of insufficient experimental data. Through mutual gaming of the generator and the discriminator, new data that followed the patterns of the original dataset were generated.

The generator \(G\) and discriminator \(D\) were trained through alternating iterations. The weights of the discriminator \(D\) were kept untrainable while training the generator \(G\) and vice versa. The GAN structure is shown in Figure 3.

The complete GAN training function was as follows...

Figure 1. SOH aging curves of eight lithium-ion batteries.
where $E(*)$ is the mathematical expectation, $X$ is the real data, $Z$ represents the random variables, $X \sim P_{data(X)}$ denotes the distribution of the real data $X$, $Z \sim P(Z)$ denotes the distribution of the random variable $Z$, $G(*)$ is the output of the generator $G$, and $D(*)$ is the output of the discriminator $D$.

As the values of $E_{X \sim P_{data(X)}}[\log(D(X))]$ and $E_{Z \sim P(Z)}[\log(1 - D(G(Z)))]$ increase, the discriminatory ability of the discriminator $D$ also increases. As the value of $E_{Z \sim P(Z)}[\log(1 - D(G(Z)))]$ decreases, the data generated by the generator $G$ become more realistic. Through alternate iterations, the distribution pattern of $G(Z)$ approaches the distribution pattern of the real data $X$.

The generator $G$ has six fully connected layers that transform the three-dimensional random variable $Z$ into 151-dimensional generated data. The first 150 dimensions are the features, and the 151st dimension is the SOH value. The generated data and real data were inputted into the discriminator separately. After running through four fully connected layers, the discriminator could verify the truth or falsity of each data point.

3. SSEL CLUSTER

3.1. Proposed SSEL Cluster for SOH Estimation. In this article, an SSEL cluster was proposed to improve the accuracy of SOH estimation. A host of SSEL frameworks with different SOH data intervals were established to identify a stable relationship between features and SOH. In each SSEL framework, the different data-driven models were considered base learners for the first-level stacking model and bagging model fusion. The outputs of the base learners, stacking model fusion, and bagging...
model fusion were used as inputs for the second-level back propagation neural network (BPNN) fusion model. By training the BPNN, a weight assignment mechanism was generated, improving the model’s function fitting ability.

3.1.1. First-level Stacking Model Fusion. First-level stacking model fusion was performed using different models to understand the feature–SOH relationships from different perspectives, which enabled the models to complement each other. The stacking model fusion ensured basic estimation accuracy in each SSEL framework. Its principle is shown in Figure 4.

![Figure 4. Principle of stacking model fusion.](image)

The implementation of the stacking model fusion is shown in Figure 5. The procedure was as follows.

Step 1. Based on the sub-training sets, each base learner was assessed by fourfold cross-validation. Three of the sub-training sets, S1, S2, S3, and S4, were used to train the base learner. Then, the remaining sub-training set and the Stest set were used to validate and test the trained base learners. The results of the four validations were recorded as the column vectors a1, a2, a3, and a4, and the corresponding testing results were recorded as the column vectors b1, b2, b3, and b4.

Step 2. A new column vector was generated by vertically merging the column vectors a1, a2, a3, and a4. This vector was defined as the new training set feature A1. The new testing set feature B1 was generated by averaging the elements in the corresponding positions of the column vectors b1, b2, b3, and b4.

Step 3. Steps 1 and 2 were looped until all base learners were trained, validated, and tested to obtain new training set features A1, A2, ..., AM and new testing set features B1, B2, ..., BM.

Step 4. The training set features of the meta-learner were formed by merging the column vectors A1, A2, ..., AM horizontally into a matrix. The real column vector Areal corresponding to A1, A2, ..., AM was jointly used as the training set label for the meta-learner. A1, A2, ..., AM and Areal were fed into the meta-learner for training.

Step 5. The testing set features of the meta-learner were formed by merging the column vectors B1, B2, ..., BM horizontally into a matrix. The real column vector Breal corresponding to B1, B2, ..., BM was jointly used as the testing set label for the meta-learner. B1, B2, ..., BM and Breal were fed into the meta-learner for testing.

3.1.2. First-level Bagging Model Fusion. Stacking model fusion has created large biases in some SOH intervals due to the use of the fusion. Given the obvious differences between the bagging and stacking models in terms of fusion principles, the bagging model fusion was added within each SSEL framework to improve estimation accuracy further. The principle of the bagging model fusion is shown in Figure 6.

Simple random sampling with replacement was performed in the training set to construct P sub-training sets. The P sub-training sets and the testing set were used to train each
base learner separately. Subsequently, the average of all predictions from these learners was calculated to obtain the final estimation.

3.1.3. Second-level BPNN Fusion Model. As the LIB ages, the relationships between features and SOH change. Owing to fixed internal weights, ordinary models cannot adapt to such continuous variation. Therefore, multiple SSEL frameworks were separately established within different SOH data intervals. The SSEL frameworks in the intervals comprise the proposed SSEL cluster. The relationships learned by the SSEL framework for each SOH interval were as follows

$$\text{SOH}_x = f_x(q_x)$$

$$q_x = [U_{x1}, U_{x2}, \ldots, U_{xH}, I_{x1}, I_{x2}, \ldots, I_{xH}, T_{x1}, T_{x2}, \ldots]$$

$$x \in [1, 2, \ldots, X]$$

(3)

where SOH$_x$, $q_x$, and $f_x$ represent our model's labels, features, and the learned relationship between SOH$_x$ and $q_x$ in the $x$-th SOH interval, respectively. $H$ represents the feature dimension number of each physical quantity, and $X$ represents the number of SOH intervals $U_{x1}, U_{x2}, \ldots, U_{xH}, I_{x1}, I_{x2}, \ldots, I_{xH},$ and $T_{x1}, T_{x2}, \ldots, T_{xH}$ are the voltage, current, and temperature features collected in the $x$-th SOH interval, respectively.

In each SSEL framework, the BPNN fusion model was trained to obtain the weights of the base learners.

As shown in Figure 7, the stacking model fusion, bagging model fusion, and base learners were incorporated while training the BPNN fusion model. The role of the base learners is to ensure that the different types of feature—SOH relationships are well-fitted by the SSEL framework. The stacking and bagging model fusion algorithms guarantee the basic estimation accuracy of the SSEL framework. Furthermore, the BPNN fusion model assigns weights for each learner. SOH estimation in the SSEL framework is performed as follows.

$$\text{SOH}_x = W_{x1} \times \text{SOH}_{x1} + \ldots + W_{xQ} \times \text{SOH}_{xQ}$$

$$+ \ldots + W_{xQ} \times \text{SOH}_{xQ}$$

(4)

where $Q$ is the number of learners; SOH$_{xi}$ and SOH$_{xq}$ are the outputs of the $i$-th learner and the SSEL framework in the $x$-th SOH interval, respectively; and $W_{xi}$ is the weight of the $i$-th learner in the $x$-th SOH interval.

3.2. Procedure for the SOH Estimation Method. The implementation procedure for the SSEL cluster is shown in Figure 8. The steps are as follows.

Step 1. The voltage, current, temperature, and SOH values from the CC—CV charging profile of the LIBs were used as the original dataset. Then, the data were normalized and divided into the training, validation, and testing sets.

Step 2. A GAN was used to expand the training set. The validation and testing sets were both divided into multiple intervals based on the SOH.

Step 3. The expanded training set was used for training various base learners, stacking model fusion, and bagging model fusion. The base learners included XGBoost, the light gradient boosting machine (LGBM), SVR, extra tree regressor, decision tree regressor, and CNN. The linear regressor was chosen as the base learner for bagging model fusion. The base learners for stacking model fusion were the extra tree regressor, decision tree regressor, and linear regressor. The meta-learner for stacking model fusion was the KNN.

Step 4. Each SSEL framework was trained using the validation set of the corresponding interval. In each SSEL framework, the outputs of the base learners, stacking model fusion, and bagging model fusion were used as inputs for the single-layer BPNN. Moreover, the weights of all learners were distributed based on the BPNN fusion model. SOH estimation for the SSEL cluster was performed based on eq 4.

Step 5. The SSEL cluster was tested using the testing set of the corresponding interval. The tested model could then be used for LIB SOH estimation.
4. RESULTS AND DISCUSSION

To prove the superiority of the SSEL cluster, the proposed model was compared with other models using data from the Oxford Battery Degradation Dataset 1. In addition, the effects of the dividing intervals method, GAN data expansion, and fusion models on LIB SOH estimation were evaluated. The root mean square error (RMSE) was used as the evaluation index for SOH estimation.

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\text{SOH}_{\text{est},i} - \text{SOH}_{\text{true},i})^2}$$  \hspace{1cm} (5)

where $\text{SOH}_{\text{true},i}$ and $\text{SOH}_{\text{est},i}$ are the true and predicted values, respectively, and $N$ is the number of training samples.

4.1. Performance of the SSEL Cluster. The Oxford Battery Degradation Dataset 1 included data from eight batteries. The data were collected in the SOH range of 70–100%, which was defined as the complete life cycle of the battery.

| SOH interval (%) | bagging (%) | stacking (%) | extra tree regressor (%) | CNN (%) | decision tree regressor (%) | XGBoost (%) | LGBM (%) | KNN (%) | SVR (%) | linear regressor (%) |
|------------------|-------------|--------------|--------------------------|--------|-----------------------------|------------|---------|--------|--------|---------------------|
| I 90–100         | 1.87        | 25.89        | -7.29                    | -16.91 | 15.82                       | 15.76      | 2.56    | 30.06  | -6.09  | 38.48               |
| 80–90            | -6.59       | 21.69        | -5.66                    | 22.21  | -23.46                      | 10.25      | 29.62   | 4.88   | 18.90  | 27.50               |
| 70–80            | 1.46        | 34.90        | -0.26                    | 13.04  | -10.07                      | 32.38      | 4.74    | -12.77 | -4.24  | 30.57               |
| II 90–100        | 37.62       | -4.50        | 38.96                    | 4.77   | 8.95                        | -5.29      | 17.76   | 13.59  | -6.08  | -5.96               |
| 80–90            | 35.31       | 7.16         | 26.70                    | -9.44  | 22.31                       | 12.89      | 4.19    | -13.62 | -7.29  | 30.57               |
| 70–80            | 6.45        | 25.82        | 5.94                     | -20.16 | -7.26                       | 27.25      | 21.10   | -13.27 | 32.43  | 17.50               |

Figure 8. Implementation procedure for the SSEL cluster.

Table 1. Weights of the SSEL Cluster for Case I and Case II
In case I, data from batteries no. 1−7 were divided into the training (80%) and validation sets (20%), and data from battery no. 8 were used as the testing set. The training set was expanded by using a GAN for training various base learners, stacking model fusion, and bagging model fusion. The validation and testing sets were divided into three intervals corresponding to the SOH ranges of 70−80, 80−90, and 90−100%. Each SSEL framework was trained using the validation set of the corresponding interval. Based on training results, the weights of learners in each SSEL framework were determined (Table 1). The SSEL cluster was used to estimate the SOH of battery no. 8; the estimates and errors are shown in Figure 9.

In case II, data from battery no. 4 were used as the testing set, and data from the remaining batteries were used as the training and validation sets. The above preprocessing, training, and testing processes were re-executed to obtain the weights of learners in each SSEL framework (Table 1). The SSEL cluster was used to estimate the SOH of battery no. 4; the estimates and errors are shown in Figure 10.

Table 1 presents the internal weights of the SSEL cluster for batteries no. 8 and no. 4. The weight of each learner was adjusted across different SOH intervals, enabling each SSEL framework to fit the corresponding feature−SOH relationships accurately. Notably, either the stacking or bagging model was always maintained as a heavyweight in all SOH intervals for batteries no. 8 and no. 4, ensuring the basic accuracy of SOH estimation.

The proposed model was compared with other models using the same training and testing sets as those of cases 1 and 2. XGBoost, LGBM, SVR, extra tree regressor, decision tree regressor, linear regressor, KNN, and CNN were used as comparative models. The results for cases 1 and 2 are shown in Table 2. The RMSE of all SSEL frameworks was within 0.6%, indicating that they outperformed the other models. SOH estimation results and error curves were plotted for the four comparative models with the highest estimation accuracy in cases 1 and 2 (Figures 11 and 12, respectively). As shown in Figures 9−12, other models could not guarantee estimation accuracy across the entire SOH interval, especially in specific SOH ranges.

4.2. Effect of the Dividing Interval Method. To explore the influence of the dividing intervals method, a single SSEL framework was trained and tested based on the validation and testing sets without any division. The results are presented in Table 3. By assigning weights to each learner, the fitting ability of single SSEL frameworks could be improved. However, the SOH estimation accuracy was substantially lower than that of comparative models. During the entire life cycle of the LIB, each learner within the single SSEL framework corresponded to only one fixed weight. Therefore, the changing relationship between the features and SOH was not accurately represented. Hence, the SOH estimation was less accurate than the SSEL cluster-based SOH estimation.

4.3. Effect of GAN Data Expansion. The effect of GAN data expansion on the accuracy of LIB SOH estimation was evaluated. Training for various base learners, stacking model fusion, and bagging model fusion were performed using the unexpanded original training set. Then, the SSEL cluster was trained and tested with the dividing interval method. The results are displayed in Table 4. It can be seen that GAN data expansion had a positive effect on the accuracy of SOH estimation. This was because the procedure prevented overfitting.

4.4. Effect of Model Fusion. The effect of model fusion on LIB SOH estimation was evaluated. Table 5 shows the RMSE
obtained after using stacking model fusion, bagging model fusion, and the SSEL cluster. In general, the SSEL cluster provided better accuracy for SOH prediction than the stacking model or bagging model fusion.

### Table 2. Estimation Results

| SSEL cluster comparative models | case item | 90−100% interval | 80−90% interval | 70−80% interval | XGBoost | LGBM | SVR | extra tree regressor | decision tree regressor | linear regressor | KNN | CNN |
|-------------------------------|-----------|------------------|-----------------|-----------------|---------|-----|----|-------------------|-------------------|---------------|------|-----|
| I  | RMSE (%)   | 0.3556           | 0.4221          | 0.5718          | 1.2466   | 0.6787 | 4.6774 | 0.8389           | 0.9290            | 0.7079        | 0.7112 | 1.3210 |
|    | accuracy (%) | 99.6444         | 99.5779         | 99.4282         | 98.7534  | 99.3213 | 95.3226 | 99.1611          | 99.0710           | 99.2921       | 99.2888 | 98.6790 |
| II | RMSE (%)   | 0.3568           | 0.3994          | 0.1049          | 0.9762   | 1.0142 | 3.0279 | 1.7710           | 1.8185            | 0.6895        | 0.9031 | 1.4010 |
|    | accuracy (%) | 99.6432         | 99.6006         | 99.8951         | 99.0238  | 98.9858 | 96.9721 | 98.2290          | 98.1815           | 99.3105       | 99.0969 | 98.5990 |

### Table 3. Comparison between the Performance of the Dividing Interval SSEL Cluster and the Single SSEL Framework

| dataset | SSEL cluster RMSE (%) | single SSEL framework RMSE (%) |
|---------|-----------------------|--------------------------------|
|         | 90−100% | 80−90% | 70−80% | 70−100% |
| battery no. 8 | 0.3556 | 0.4221 | 0.5718 | 0.6097 |
| battery no. 4 | 0.3568 | 0.3994 | 0.1049 | 0.5888 |

### Table 4. Effect of GAN Data Expansion on the SSEL Cluster

| dividing intervals RMSE (%) | 90−100% | 80−90% | 70−80% |
|-----------------------------|---------|--------|--------|
| dataset | GAN without GAN | GAN without GAN | GAN without GAN | GAN without GAN |
| battery no. 8 | 0.3556 | 0.6139 | 0.4221 | 0.5926 | 0.5718 | 0.8842 |
| battery no. 4 | 0.3568 | 0.4929 | 0.3994 | 0.5757 | 0.1049 | 0.5433 |

### Figure 11. SOH estimation and errors for battery no. 8. (a−d) LGBM, KNN, linear regressor, and extra tree regressor fitting curves for battery no. 8, respectively. (e−h) LGBM, KNN, linear regressor, and extra tree regressor error curves for battery no. 8, respectively.

### Table 5. Comparison between the Performance of Different Fusion Models for Battery no. 8 and no. 4

| battery no. 8 | SOH intervals (%) | stacking | bagging | SSEL cluster RMSE (%) |
|---------------|-------------------|----------|---------|-----------------------|
| 90−100%       | 0.4586            | 1.1210   | 0.3556  |
| 80−90%        | 0.6708            | 0.5556   | 0.4221  |
| 70−80%        | 0.7681            | 0.3956   | 0.5718  |
| battery no. 4 | 90−100%           | 0.4257   | 0.4521  | 0.3568                |
| 80−90%        | 0.7906            | 0.2927   | 0.3994  |
| 70−80%        | 0.2637            | 0.2640   | 0.1049  |

Tables 2 and 5 show that in case I, stacking model fusion outperformed all comparative models during the entire life cycle, except during the 70−80% SOH interval. Similarly, bagging model fusion performed all comparative models, except during the 90−100% SOH interval. In case II, the RMSE obtained after applying stacking model fusion was lower than that obtained with other comparative models, and it was inferior to the linear regressor only in the 80−90% SOH interval. Bagging model fusion provided the best estimation accuracy.
compared with all the other models. This showed that the stacking model and bagging model fusion play an important role in ensuring the basic estimation accuracy of the SSEL cluster.

5. CONCLUSIONS

This article presents an SSEL cluster for LIB SOH estimation based on voltage, current, and temperature measurements. The dataset, which was expanded using GAN, was divided into intervals based on true SOH values. Then, an independent SSEL framework was built into each interval to fit the feature−SOH relationships accurately. Various base learners enabled the SSEL framework to accommodate the complex relationships between features and SOH; stacking model fusion and bagging model fusion ensured basic estimation accuracy; the BPNN fusion model assigned weights to each learner to allow optimal cooperation among learners. Owing to these mechanisms, the SSEL cluster provided improved SOH estimation accuracy. Through comparisons with other models, the superiority of the SSEL cluster was verified. Moreover, the role of the dividing interval method, GAN data expansion, and fusion in improving the accuracy of SOH estimation was explored experimentally.

■ AUTHOR INFORMATION

Corresponding Author

Long Zeng — School of Automation, Guangdong University of Technology, Guangzhou 510006, China; orcid.org/0000-0002-3999-3850; Email: zenglong2404@gdut.edu.cn

Authors

Si-zhe Chen — School of Automation, Guangdong University of Technology, Guangzhou 510006, China
Hongtao Zhang — School of Automation, Guangdong University of Technology, Guangzhou 510006, China
Yuanliang Fan — Department of Equipment Technology Research Center, Electric Power Research Institute of Fujian Electric Power Company Limited, Fuzhou 350007, China; Fujian Provincial Enterprise Key Laboratory of High Reliable Electric Power Distribution Technology, Fuzhou 350007, China
Le Chang — School of Automation, Guangdong University of Technology, Guangzhou 510006, China
Yun Zhang — School of Automation, Guangdong University of Technology, Guangzhou 510006, China

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.2c01589

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China under grant 61802070 and the Industrial Guidance Key Project of Fujian Province under grant 2020H0043.

■ REFERENCES

(1) Liang, F.; Jiang, L.; Zhang, Z.; Wu, D.; Li, X.; Han, N.; Rui, Y.; Zhang, W.; Tang, B. Cobalt-doped TaOCl3 nanoparticles/carbon compounds with advanced specific capacity for lithium-ion batteries. J. Alloys Compd. 2022, 897, 163193.
(2) Jiang, L.; Liang, F.; Zhang, Z.; Wu, D.; Chai, J.; Luo, T.; Han, N.; Zhang, W.; Rui, Y.; Tang, B. A novel cobalt chloride hydrate modified Co-MOF derived carbon microspheres as anode materials for lithium ion batteries. Chem. Eng. J. 2022, 433, 133568.
(3) Ge, Z.; Jiang, L.; Liang, F.; Zhang, Z.; Wu, D.; Ye, W.; Han, N.; Rui, Y.; Zhang, W.; Tang, B. Nitrogen-Doped Porous Ag−C@Co3O4 Nanocomposite for Boosting Lithium Ion Batteries. Energy & Fuels 2022, 36, 2861−2871.
(4) Li, Q.; Du, J.; Chai, J.; Han, N.; Zhang, W.; Tang, B. Vanadium Metaphosphate V(PO3)5 Derived from V-MOF as a Novel Anode for Lithium-Ion Batteries. ChemistrySelect 2021, 6, 8150−8157.

Figure 12. SOH estimation and errors for battery no. 4. (a−d) LGBM, KNN, linear regressor, and XGBoost fitting curves for battery no. 4, respectively. (e−h) LGBM, KNN, linear regressor, and XGBoost error curves for battery no. 4, respectively.
(5) Chai, J.; Du, J.; Li, Q.; Han, N.; Zhang, W.; Tang, B. Recent Breakthroughs in the Bottleneck of Cathode Materials for Li–S Batteries. *Energy & Fuels 2021*, *35*, 15455–15471.

(6) Han, N.; Guo, X.; Cheng, J.; Liu, P.; Zhang, S.; Huang, S.; Rowles, M. R.; Fransaer, J.; Liu, S. Inhibiting in situ phase transition in Ruddlesden-Popper perovskite via tailoring bond hybridization and its application in oxygen permeation. *Matter 2021*, *4*, 1720–1734.

(7) Zhang, S.; Guo, X.; Dou, X.; Zhang, X. A rapid online calculation method for state of health of lithium-ion battery based on coulomb counting method and differential voltage analysis. *J. Power Sources 2020*, 249, 228740.

(8) Lavigne, L.; Sabatier, J.; Francisco, J. M.; Guillemand, F.; Noury, A. Lithium-Ion Open Circuit Voltage (OCV) curve modelling and its ageing adjustment. *J. Power Sources 2016*, *324*, 694–703.

(9) Guha, A.; Patra, A. State of health estimation of lithium-ion batteries using capacity fade and internal resistance growth models. *IEEE Trans. Transport. Electrif. 2018*, *4*, 135–146.

(10) Vichard, L.; Ravely, A.; Venet, P.; Harel, F.; Pelissier, S.; Hissel, D. A method to estimate battery SOH indicators based on vehicle operating data only. *Energy 2021*, *225*, 120235.

(11) Tan, X.; Zhan, D.; Lyu, P.; Rao, J.; Fan, Y. Online state-of-health estimation of lithium-ion battery based on dynamic parameter identification at multi timescale and support vector regression. *J. Power Sources 2021*, *484*, 229233.

(12) Xiong, R.; Li, L.; Tian, J. Towards a smarter battery management system: A critical review on battery monitoring systems. *J. Power Sources 2018*, *405*, 18–29.

(13) Li, J.; Adewuyi, K.; Lotfi, N.; Landers, R. G.; Park, J. A single particle model with chemical/mechanical degradation physics for lithium ion battery State of Health (SOH) estimation. *Appl. Energy 2018*, *212*, 1178–1190.

(14) Tang, X.; Zou, C.; Yao, K.; Chen, G.; Liu, B.; He, Z.; Gao, F. A fast estimation algorithm for lithium-ion battery state of health. *J. Power Sources 2018*, *396*, 453–458.

(15) Hu, C.; Jain, G.; Zhang, P.; Schmidt, C.; Gomadam, P.; Gorka, T. Data-driven method based on particle swarm optimization and k-nearest neighbor regression for estimating capacity of lithium-ion battery. *Appl. Energy 2014*, *129*, 49–55.

(16) Li, Y.; Zou, C.; Berecibar, M.; Nanini-Maury, E.; Chan, J. C.; W.; Van den Bossche, P.; Van Mierlo, J.; Omar, N. Random forest regression for online capacity estimation of lithium-ion batteries. *Appl. Energy 2018*, *232*, 197–210.

(17) Shen, S.; Sadoughi, M.; Chen, X.; Hong, M.; Hu, C. A deep learning method for online capacity estimation of lithium-ion batteries. *J. Power Sources 2019*, *25*, 100817.

(18) Cai, L.; Meng, J.; Stroe, D.-I.; Peng, J.; Luo, G.; Teodorescu, R. Multiobjective optimization of data-driven model for lithium-ion battery SOH estimation with short-term feature. *IEEE Trans. Power Electron. 2020*, *35*, 11855–11864.

(19) Ng, M.-F.; Zhao, J.; Yan, Q.; Conduit, G. J.; Seh, Z. W. Predicting the state of charge and health of batteries using data-driven machine learning. *Nat. Mach. Intell. 2020*, *2*, 161–170.

(20) Gou, B.; Xu, Y.; Feng, X. State-of-health estimation and remaining-useful-life prediction for lithium-ion battery using a hybrid data-driven method. *IEEE Trans. Veh. Technol. 2020*, *69*, 10854–10867.

(21) Ren, L.; Dong, J.; Wang, X.; Meng, Z.; Zhao, L.; Deen, M. J. A data-driven auto-cnn-lstm prediction model for lithium-ion battery remaining useful life. *IEEE Trans. Ind. Inf. 2020*, *17*, 3478–3487.

(22) Dai, H.; Zhao, G.; Lin, M.; Wu, J.; Zheng, G. A novel estimation method for the state of health of lithium-ion battery using prior knowledge-based neural network and Markov chain. *IEEE Trans. Ind. Electron. 2018*, *66*, 7706–7716.

(23) Meng, J.; Cai, L.; Stroe, D.-I.; Ma, J.; Luo, G.; Teodorescu, R. An optimized ensemble learning framework for lithium-ion battery state of health estimation in energy storage system. *Energy 2020*, *206*, 118140.

(24) Liu, K.; Shang, Y.; Ouyang, Q.; Widanage, W. D. A data-driven approach with uncertainty quantification for predicting future capacities and remaining useful life of lithium-ion battery. *IEEE Trans. Ind. Electron. 2020*, *68*, 3170–3180.

(25) Yu, J. State of health prediction of lithium-ion batteries: Multiscale logic regression and Gaussian process regression ensemble. *Reliab. Eng. Syst. Saf. 2018*, *174*, 82–95.

(26) Ungurean, L.; Cărtăuş, G.; Micea, M. V.; Groza, V. Battery state of health estimation: a structured review of models, methods and commercial devices. *Int. J. Energy Res. 2017*, *41*, 151–181.

(27) Birkl, C. Oxford Battery Degradation D 1, 2017.

(28) Creswell, A.; White, T.; Dumoulin, V.; Arulkumaran, K.; Sengupta, B.; Bharath, A. A. Generative adversarial networks: An overview. *IEEE. Signal. Proc. Mag 2018*, *35*, 53–65.

(29) Wolpert, D. H. Stacked generalization. *Neural Networks 1992*, *5*, 241–259.

(30) Breiman, L. Bagging predictors. *Mach. Learn 1996*, *24*, 123–140.