



SOH estimation of lithium-ion batteries based on least squares support vector machine error compensation model

Ji'ang Zhang¹ · Ping Wang¹ · Qingrui Gong¹ · Ze Cheng¹

Received: 25 March 2021 / Revised: 27 August 2021 / Accepted: 29 August 2021 / Published online: 9 September 2021
© The Korean Institute of Power Electronics 2021

Abstract

Accurate estimation of the state of health (SOH) of lithium-ion batteries is an important determinant of their safe and stable operation. In this paper, a method for the SOH estimation of lithium-ion batteries based on the least squares support vector machine error compensation model (LSSVM-ECM) is proposed. This method achieves a combination of an empirical degradation model and a data-driven method. Battery degradation can be divided into overall trends and local differences, where the former can be described by an empirical degradation model (EDM) established by the historical data of the battery capacity, while the latter can be mapped by a least squares support vector machine (LSSVM). An LSSVM-ECM is established, where the input is the time interval of the equal charging voltage rising (DV_DT) and the output is the fitting error of the EDM, which represents the local difference of the capacity degradation to dynamically compensate the prediction results of the EDM that represents the global trend in terms of the capacity degradation. Validations are carried out with battery data provided by Oxford and NASA datasets. Results show that the proposed method has a high prediction accuracy and a strong robustness.

Keywords Lithium-ion battery · State of health · Least squares support machines · Error compensation

1 Introduction

Lithium-ion batteries have the advantages of low cost, high energy density, and long cycle life. Thus, recent years have witnessed its wide application in aerospace, electric vehicles (EV), photovoltaic power grids, and other fields [1]. A battery management system (BMS) can realize the scientific assessment, risk warning, and regular replacement recommendations of batteries to ensure their healthy and stable operation. Accurate estimation of the state of health (SOH), as a key part of the BMS in [2], is desired to provide information for safety management and charging/discharging

optimal control, which is defined as the ratio of available current capacity to factory rated capacity. Due to battery aging, the SOH of a battery takes on a decreasing trend. Generally, the end of life (EoL) of a battery is reached when the SOH has decreased to 80%, and it is necessary to replace the batteries before a battery failure makes the whole system crash [3].

Unfortunately, the SOH of a battery cannot be measured directly by a sensor. It can only be estimated by analyzing external measurements combined with a mathematical method. The SOH estimation methods for the lithium-ion battery can be divided into model-based methods and data-driven methods. Model-based methods mainly include equivalent circuit models [4], electrochemical models [5, 6] and empirical degradation models (EDM) [7–11]. Equivalent circuit models, with the merits of a simple structure and reduced calculations can simulate the external working state of the battery, which is combined with optimization and filtering algorithms for parameter identification and SOH estimation. However, the adaptability of equivalent circuit models is relatively poor, and they can be easily affected by working conditions and the convergence of the estimation algorithm. The electrochemical models, when combined

✉ Ze Cheng
chengze@tju.edu.cn

Ji'ang Zhang
sddxzja@163.com

Ping Wang
pingW@tju.edu.cn

Qingrui Gong
gongqingrui@tju.edu.cn

¹ Department of Electrical Engineering, Tianjin University, Tianjin, China

with a series of partial differential equations, can describe the physical and chemical mechanism of battery degradation, as represented by a pseudo two-dimensional (P2D) model based on the porous electrode theory [5] and its simplified scheme [6]. However, both the parameter identification and the equation calculations are complex, which makes them unsuitable for the on-line estimation of BMS systems. The empirical degradation models [7–11] can describe the overall trend of the capacity degradation of a battery in all of its cycles. They possess the advantages of good practicability and strong robustness, which makes them suitable for on-line applications. However, as a simplification of historical data, empirical degradation models are difficult to adapt to the local differences of battery degradation, which are mainly caused by the individual differences among batteries and the local fluctuations of capacity degradation that are not smooth.

The data-driven based SOH estimation methods do not need to analyze the internal mechanisms of a battery. Instead, they extract and analyze the external health features (HF) that are closely related to battery degradation. Then, they establish a nonlinear mapping relationship between HF and SOH through a machine learning algorithm. This avoids the need for parameter identification and physical modeling. As a result, they have strong flexibility. The rationality of the health features selected and the generalization capability of the training algorithm have a significant impact on accuracy. The health features include the primary features of the charge and discharge curves of a battery [12], the secondary features after processing [13], etc. In addition, the training algorithms mainly include neural networks [12–14], support vector machines [15, 16], relevance vector machines [17], etc. Generally, data-driven methods are easy to establish and implement, and the external features can be extracted to efficiently realize online estimation after the battery aging regression model is established off-line. Thus, they have strong nonlinear mapping capability and can capture the feature details of capacity degradation to describe the actual situation of battery degradation. However, adequate training samples are required to fully learn and remember the capacity degradation situations. In addition, the divergence of estimation results can occur when the training samples are insufficient, which indicates poor robustness.

Therefore, a SOH estimation method that combines an empirical degradation model with a data-driven method deserves consideration, since it is expected to maintain accuracy and enhance the robustness. A least squares support vector machine (LSSVM) [18, 19] is an efficient machine learning algorithm with strong nonlinear mapping capability. It has wide application in the field of nonlinear regression. Thus, this paper proposes a lithium-ion battery SOH estimation method based on a least squares support vector machine error compensation model (LSSVM-ECM), which realizes

the integration of an empirical degradation model and a data-driven method. Experimental results on battery data provided by NASA [20] and Oxford [21] datasets show that the proposed method has a high SOH estimation accuracy under different conditions. Meanwhile, it has the stronger robustness than purely data-driven methods and the better transplant ability than empirical degradation models.

2 Empirical degradation model

The relationship between battery capacity and SOH is shown in Eq. (1):

$$\text{SOH}(N) = \frac{Q(N)}{Q_N}, \quad (1)$$

where Q_N is the initial capacity, $Q(N)$ is the actual capacity under the cycle N . For the sake of simplicity, the degradation rate is considered as the function of Q and N :

$$\frac{dQ}{dN} = f(Q, N), \quad (2)$$

where $f(\cdot)$ is the nonlinear function with two variables. The first-order Taylor expansion of Eq. (2) is carried out to obtain Eq. (3):

$$\frac{dQ}{dN} = a_0 + a_1Q + a_2N + o(\sqrt{Q^2 + N^2}), \quad (3)$$

a_0 , a_1 , and a_2 are constant. $o(\cdot)$ refers to a high order remainder, which can be ignored since the capacity fluctuation is not considered here. See Eq. (4):

$$\frac{dQ}{dN} = a_0 + a_1Q + a_2N. \quad (4)$$

Discretize (4) into (5) as follows:

$$Q(N+1) - Q(N) = a_0 + a_1Q(N) + a_2N. \quad (5)$$

Then arrange it into (6):

$$Q(N+1) = (1 + a_1)Q(N) + a_2N + a_0. \quad (6)$$

Equation (6) is a first-order linear difference equation, whose general solution form is:

$$Q(N) = (1 + a_1)^N - \frac{a_2}{a_1}N - \frac{a_2}{a_1^2} - \frac{a_0}{a_1} \quad (7)$$

Let $1 + a_1 = a$, $-a_2/a_1 = b$, and $-a_2/a_1^2 - a_0/a_1 = c$. Then, (7) can be transformed into Eq. (8):

$$Q(N) = a^N + bN + c. \quad (8)$$

Combining Eq. (1) and considering $\text{SOH}(0) = 1$, (8) can be transformed into (9):

$$SOH(N) = k_1 k_2^N + k_3 N - k_1 + 1. \tag{9}$$

SOH (N) = $Q(N)/Q_N$, $k_1 = 1/Q_N$, $k_2 = a$, and $k_3 = b/Q_N$, which are the parameters to be identified with historical data of battery degradation by the least square method. Fitting the historical data of the capacity degradation from the Oxford and NASA datasets for partial cycles and the whole cycles with (9), Figs. 1 and 2 can be obtained.

Here the extrapolation method is applied to the cycle numbers to investigate the capacity degradation. Notably, extrapolation can also be applied to the voltage, temperature, or other physical quantities to achieve state observation, which is suitable for many power electronic components such as MOSFETs [22–24], where the extrapolation models of these variables or their reciprocals are established for lifetime estimation. Voltage is an important variable for the state estimation of a battery, as introduced in Sect. 3. As a critical factor, temperature significantly impacts the performance of batteries [25], which is also important in the field of batteries.

Table 1 shows parameter identification results for the whole cycle.

It can be seen from Figs. 1 and 2 that an EDM can describe the overall trend of capacity degradation. However, the main problems are as follows.

- (1) Using an EDM established by partial cycles to predict the remaining cycles tends to produce large errors, as shown in Fig. 1. This is due to the fact that formula (9) ignores the high-order terms and only contains historical information. It does not contain real information, which makes it difficult to track the actual degradation trend of capacity.
- (2) Using an EDM established by all of the cycles of a basis battery to predict the capacity degradation of other batteries tends to produce large errors, even when they are the same type of battery, as shown in Fig. 2. This is due to the fact that for different batteries of the same type, the actual capacity degradation can be quite different due to battery inconsistencies caused by manufacturing.

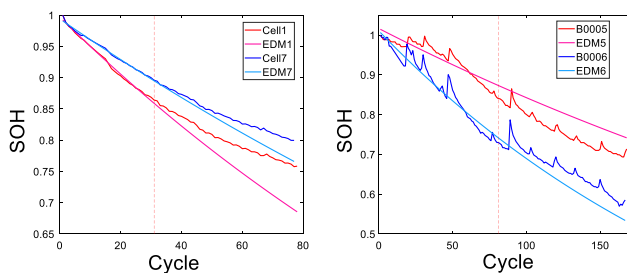


Fig. 1 Fitting of partial cycles of an EDM

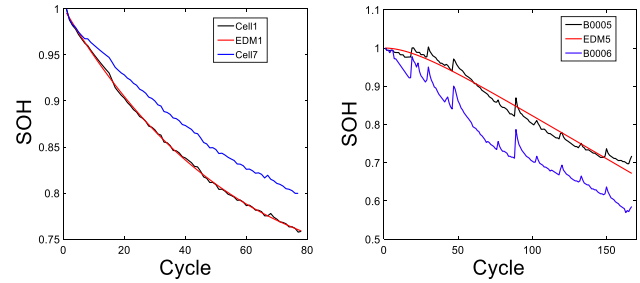


Fig. 2 Fitting of all the cycles of an EDM

Therefore, the transplant ability of an EDM between different batteries is poor.

An EDM is a simplification of actual capacity degradation that includes historical information of battery aging, and can describe the overall trend of capacity degradation. However, the above shortcomings make it necessary to combine an EDM with a data-driven method that contains more realistic information on capacity degradation to enhance accuracy.

3 Data-driven method

The basic idea of a data-driven method for SOH estimation is to describe and map the actual capacity degradation through HF. Thus, the selection of health features and a training algorithm are two key factors that affect the performance and practicability of this method. The data-driven method based on DV_DT and LSSVM are introduced in this section.

3.1 HF extraction

In this paper, B0005–B0007 and B0029–B0031 from the NASA randomized battery usage data set [20] and Cell1–Cell8 from the Oxford battery degradation data set [21] are used for the experiment. The actual capacity of each battery can be acquired by ampere–hour integration. In other words, it can be acquired by calculating the integration of the discharging current with the discharging time for every charge–discharge cycle, which is then divided by the rated capacity to get the actual SOH.

Table 1 Parameter identification results

| Battery | k_1 | k_2 | k_3 |
|---------|---------|--------|-----------|
| Cell1 | 0.6445 | 0.9888 | 0.001743 |
| B0005 | −0.0408 | 0.9088 | −0.002215 |
| B0029 | −0.0347 | 0.9896 | −0.003478 |

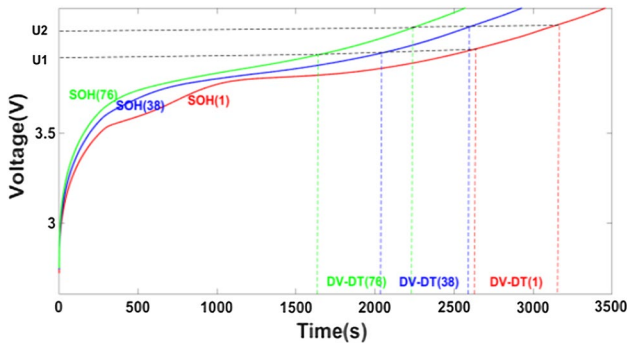


Fig. 3 Charging voltage curves

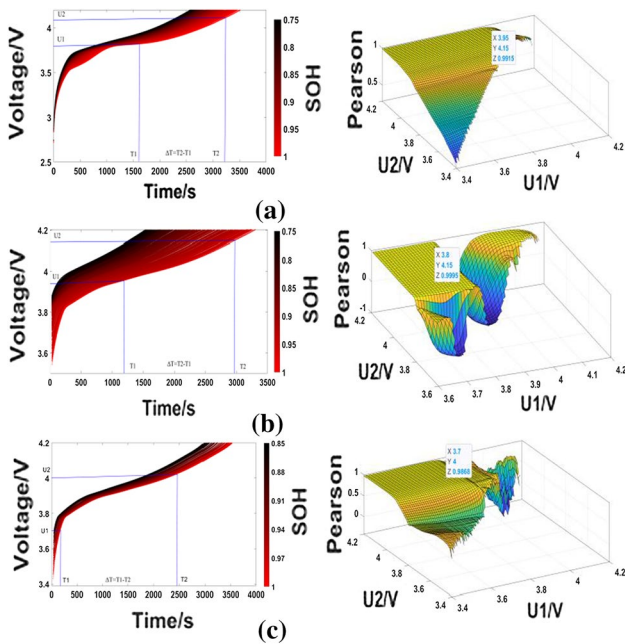


Fig. 4 Charging voltage and Pearson relevance between DV_DT and SOH for different voltage segments: a Cell1; b B0005; c B0029

The constant current (CC)–constant voltage (CV) mode [17] for charging is the most common. The voltage curves of the CC stage under different SOH values of a battery are shown in Fig. 4. In this figure, the curve color changes from bright to dark, which indicates that the battery aging is deepening and presents a strong relevance. In practice, the complete charging voltage curve is often difficult to obtain. Thus, the voltage segment can be selected instead. Considering that the charging time of the battery is difficult to determine, while the charging voltage is easy to measure, the time interval $\Delta T = T_2 - T_1$ of equal charging voltage rising $[U_1, U_2]$ is selected as a HF in this paper. This is denoted by DV_DT.

Figure 3 shows the charging voltages under 1st, 38th and 76th cycles. The charging voltages under the other cycles have the same trend. Thus, they are not shown here.

The first battery of every pack is taken as the reference battery to determine the optimal voltage segment $[U_1, U_2]$. For any given segment $[U_1, U_2]$, it is possible to obtain the corresponding time interval sequence $\{DV_DT(1), \dots, DV_DT(38), \dots, DV_DT(EOL)\}$. In addition, the SOH degradation sequence $\{SOH(1), \dots, SOH(38), \dots, SOH(EOL)\}$ of the reference battery can be acquired by the ampere-hour integral method. The traversing method is used to search for the optimal voltage segment $[U_1, U_2]$, whose corresponding DV_DT sequence has the highest Pearson coefficient value with respect to the SOH degradation sequence. This indicates that the DV_DT sequence has the highest relevance with respect to the SOH degradation sequence and is suitable as a health feature.

The formula of the Pearson coefficient is Eq. (10). Figure 4 shows the Pearson coefficient between ΔT and the SOH for each of the batteries under different $[U_1, U_2]$. It can be seen that the different voltage rising segments have a great impact on the correlation degree. The optimal voltage segment of the Oxford data set battery is $[3.8, 4.15]$, that of the B0005–B0007 data set is $[3.95, 4.15]$, and that of the B0029 ~ B0031 data set is $[3.7, 4]$, which are presented in the cursors in Fig. 4:

$$\text{Pearson} = \frac{\sum_i (DV_DT_i - DVT)(SOH_i - SOH_e)}{\sqrt{\sum_i (DV_DT_i - DVT)^2} \sqrt{\sum_i (SOH_i - SOH_e)^2}} \quad (10)$$

$$\text{GRC} = \frac{1}{n} \sum_{i=1}^n \frac{\min |SOH_i - DV_DT_i| + \rho \max_{V_i} |SOH_i - DV_DT_i|}{|SOH_i - DV_DT_i| + \rho \max_{V_i} |SOH_i - DV_DT_i|} \quad (11)$$

Formula (11) refers to the gray relevance coefficient (GRC). DV_DT_i is the time interval of the equal charging voltage rising of the i th cycle, SOH_i is the state of health of the i th cycle, n is the total number of cycles, and DVT and SOH_e are the mean values of $\{DV_DT_i\}$ and $\{SOH_i\}$ respectively. The Pearson coefficient and GRC between DV_DT and the SOH of the batteries are calculated for the voltage segment mentioned above, as shown in Table 2.

It can be seen from Table 2 that the Pearson coefficient of each battery is above 0.95, and that the GRC is above 0.8. This indicates that the health feature DV_DT has a strong correlation with battery aging, which is also easy to extract for online estimation.

3.2 Least squares support vector machine

Based on the VC dimension theory of the statistical learning theory and the principle of structural risk minimization, a support vector machine (SVM) is suitable for non-linear small sample problems. LSSVM is an improvement

Table 2 Relevance between SOH and DV_DT of batteries

| Data | Battery | Pearson | GRC |
|--------|---------|---------|--------|
| Oxford | Cell1 | 0.9995 | 0.9334 |
| | Cell2 | 0.9964 | 0.8791 |
| | Cell3 | 0.9993 | 0.9344 |
| | Cell4 | 0.9975 | 0.9390 |
| | Cell5 | 0.9981 | 0.8862 |
| | Cell6 | 0.9944 | 0.9280 |
| | Cell7 | 0.9989 | 0.9421 |
| | Cell8 | 0.9988 | 0.9302 |
| NASA | B0005 | 0.9915 | 0.8547 |
| | B0006 | 0.9898 | 0.8605 |
| | B0007 | 0.9917 | 0.8575 |
| | B0018 | 0.9876 | 0.8328 |
| | B0029 | 0.9868 | 0.8369 |
| | B0030 | 0.9792 | 0.8322 |
| | B0031 | 0.9736 | 0.8218 |
| | B0032 | 0.9876 | 0.8521 |

of SVM [26]. The LSSVM inherits the advantages of SVM, while replacing the insensitive function of SVM with the two norms of error, and replacing the inequality constraint with the equality constraint, to reduce the computational complexity.

Let the training sample set be: $\{(x_1, y_1) \dots (x_n, y_n)\}$, which can be fitted with the following high-dimensional linear mapping:

$$f(x) = w^T \phi(x) + b, \tag{12}$$

where w is the weight vector, b is the bias value, and ϕ is the nonlinear mapping function. According to the principle of structural risk minimization, the LSSVM regression problem can be transformed into the constrained optimization problem of (13):

$$\begin{cases} \min \frac{1}{2} w^T w + \gamma \sum_{i=1}^n e_i^2 \\ \text{s.t. } y_i = w^T \phi(x_i) + b + e_i, i = 1, 2 \dots n \end{cases}, \tag{13}$$

where e_i is the training error, γ is the regularization factor, and the Lagrange multiplier method is used to transform (13) into the dual problem of (14):

$$L = \frac{1}{2} w^T w + \gamma \sum_{i=1}^n e_i^2 + \sum_{i=1}^n \lambda_i (w^T \phi(x_i) + b + e_i - y_i), \tag{14}$$

where λ_i is the i th Lagrange multiplier, and the partial derivative of (14) is obtained by the KKT condition, listed as (15):

$$\begin{cases} \partial L / \partial w = 0 \\ \partial L / \partial b = 0 \\ \partial L / \partial e_i = 0 \\ \partial L / \partial \lambda_i = 0 \end{cases} \Rightarrow \begin{cases} w = \sum_{i=1}^n \lambda_i \phi(x_i) \\ \sum_{i=1}^n \lambda_i = 0 \\ \lambda_i = \gamma e_i \\ y_i = w^T \phi(x_i) + b + e_i \end{cases}. \tag{15}$$

The model parameters λ and b of the LSSVM can be determined by solving linear Eqs. (16), which are equivalent to (15):

$$\begin{cases} \begin{bmatrix} 0 & \mathbf{1}^T \\ \mathbf{1} & K + \mathbf{I}/\gamma \end{bmatrix} \begin{bmatrix} b \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \\ \mathbf{1} = (1, 1 \dots 1)^T \\ \lambda = (\lambda_1, \lambda_2 \dots \lambda_n)^T \\ y = (y_1, y_2 \dots y_n)^T \\ K(x_i, x_j) = \phi(x_i)^T \phi(x_j) \end{cases}, \tag{16}$$

where K is the kernel function matrix. In addition, the radial basis function (RBF) (17) is selected in this paper:

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right). \tag{17}$$

Then the LSSVM model is as follows:

$$f(x) = \sum_{i=1}^n (\lambda_i \cdot K(x, x_i)) + b. \tag{18}$$

The unknown parameters of (18) are the regularization factor γ and the hyperparameter δ of RBF, which are usually determined by cross validation or algorithm optimization. In this paper, particle swarm optimization (PSO) [26] is used to optimize the parameters of LSSVM, whose fitness function is the root mean square value between the actual value and the output value of the LSSVM (19):

$$\min \sum_{i=1}^m (y_i - f(x_i, \gamma, \delta))^2, \tag{19}$$

where m is the number of training samples, and x_i and y_i are the input and output of the training set, respectively.

The implementation steps of the SOH estimation method based on a pure data-driven method with LSSVM are as follows.

LSSVM is established with the offline HF sequence $\{DV_DT_i\}$ as the input and the SOH sequence $\{SOH_i\}$ as the output. When using this method online, the DV_DT of new cycle is extracted, and input into LSSVM. Then, the SOH estimation of this cycle can be output. Generally, this method needs adequate training samples. In other words, adequate actual capacity values, which can only be

obtained by regular calibration, which is time-consuming and laborious. If the capacity degradation details are not fully trained for lack of samples, the prediction results may diverge, which indicates poor robustness. It can be seen from Figs. 1 and 2 that the EDM does a good job of describing the overall trend of capacity degradation, which has good robustness. Therefore, the data-driven method can be combined with the EDM to enhance its robustness. Thus, the accuracy can be ensured while the number of training cycles can be reduced. In addition, the poor transplant ability of the EDM between different batteries due to battery inconsistencies can be improved, since the data-driven method can provide more information about the actual degradation.

4 LSSVM-ECM method

From the analysis in the previous sections, the capacity degradation of the battery can be divided into the overall trend and local differences. The EDM is used to describe

the overall trend of battery degradation, which represents the historical information of battery aging. Meanwhile, the local difference of capacity degradation quantified by the fitting error of the EDM can be reflected by the external health features, which contain actual information of battery aging. Based on this idea, it is possible to achieve a combination between the EDM and a data-driven method. Figure 5 show a flow chart of the proposed LSSVM-ECM.

The SOH estimation test can be divided into a single battery test and a different battery test. The former aims to estimate the SOH of the battery to be tested after the SPth cycle. When capacity cycle data before the SPth cycle is known, the SOH after the SPth cycle can be estimated through data before the SPth cycle. The latter, where the capacity degradation of the battery to be tested is unknown for the whole cycle, aims to estimate the SOH of the battery to be tested through the capacity degradation data of other batteries of the same type.

For the single battery test, as shown in Fig. 6a, the capacity data for the 1-SP cycle is used for the training test and the remaining cycles to be evaluated are used for the testing set, as

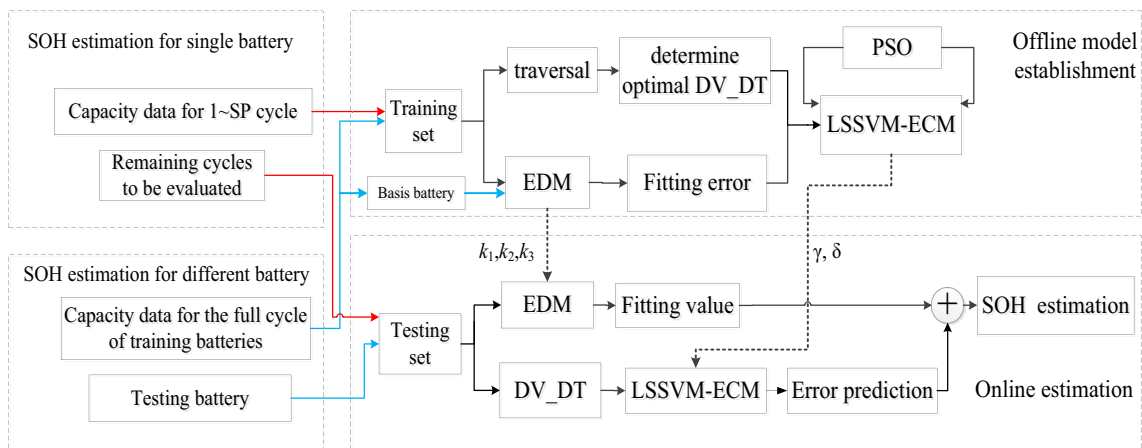


Fig. 5 Flow chart of the LSSVM-ECM

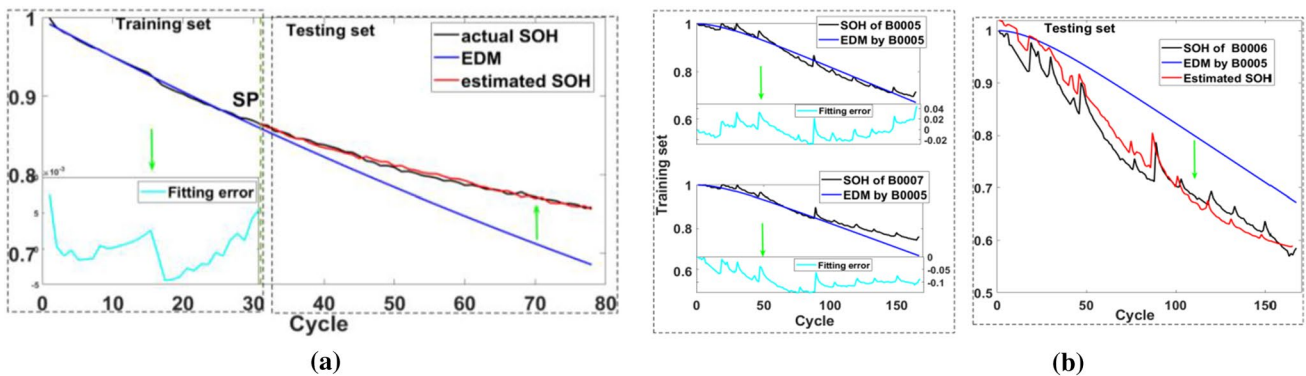


Fig. 6 Diagrams of: a single test; b different battery test

shown by the red arrow in Fig. 5. The traversal method is used to determine the optimal voltage segment and to extract the health feature DV_DT with a training test. Refer to Sect. 3 for specific steps. Meanwhile, fit the capacity data before the SPth cycle with (9) to establish the empirical degradation model (EDM), which contains the parameter k_1 , k_2 and k_3 . Calculate the fitting error between the actual SOH and the fitting value of the EDM before the SPth cycle, and (20) shows the training and testing set for a single battery test:

$$\begin{cases} \text{EDM}_i = k_1 k_2^i + k_3 i - k_1 \\ P_{\text{train}} = \{\text{DV_DT}_i\}_{i=1\dots\text{SP}} \\ T_{\text{train}} = \{\text{SOH}_i - \text{EDM}_i\}_{i=1\dots\text{SP}} \\ P_{\text{test}} = \{\text{DV_DT}_i\}_{i=\text{SP}+1\dots\text{EOL}} \\ T_{\text{test}} = \{\text{SOH}_i^e - \text{EDM}_i\}_{i=\text{SP}+1\dots\text{EOL}} \end{cases} \quad (20)$$

In (20), P_{train} and T_{train} serve as the input and output to train LSSVM model, where the parameters γ and δ can be determined by PSO, as shown in [26]. P_{test} and T_{test} are the input and output of the testing set. SOH_i^e refers to the estimated SOH, where the subscript i refers to the cycle number. When battery is in use under the i th cycle ($i > \text{SP}$), the health feature DV_DT_i ($i > \text{SP}$) can be easily extracted, which is put into the established LSSVM model to output the estimated value of the fitting error between the actual SOH and the EDM output. EDM_i ($i > \text{SP}$) is easy to calculate and summed with the estimated fitting error to acquire the estimated SOH_i , which achieves error compensation for the SOH estimation after the SPth cycle, as can be seen in (21):

$$T_{\text{test}} + \{\text{EDM}_i\}_{i=\text{SP}+1\dots\text{EOL}} = \{\text{SOH}_i^e\}_{i=\text{SP}+1\dots\text{EOL}} \quad (21)$$

For the different battery test, as shown in Fig. 6b, the capacity data for the battery to be tested is used for the testing set, and batteries of the same type as the battery to be test are used for the training set, as shown by the blue arrows in Fig. 5. First, select one battery of the same type as the battery to be tested as the basis battery. Then, fit its capacity data with (9) to identify the parameter k_1 , k_2 and k_3 . Next, the established EDM is applied to the other training batteries to calculate the fitting error sequence T_{train} (22) shows the training and testing set for the different battery test:

$$\begin{cases} \text{EDM}_i = k_1 k_2^i + k_3 i - k_1 \\ P_{\text{train}} = \{\text{DV_DT}_i^j\}_{i=1,2\dots\text{EOL}, j=1, 2\dots N} \\ T_{\text{train}} = \{\text{SOH}_i - \text{EDM}_i^j\}_{i=1,2\dots\text{EOL}, j=1, 2\dots N} \\ P_{\text{test}} = \{\text{DV_DT}_i\}_{i=1,2\dots\text{EOL}} \\ T_{\text{test}} = \{\text{SOH}_i^e - \text{EDM}_i\}_{i=1\dots\text{EOL}} \end{cases} \quad (22)$$

where N is the number of training batteries. P_{test} contains the health feature sequences of the battery to be tested. P_{train} and T_{train} are served as the input and output to train the LSSVM model, where the parameters γ and δ can be determined by the PSO in [26]. When in an application, P_{test} are put into the established LSSVM model to output the estimated value of the fitting error between the actual SOH and EDM output sequence of the battery to be tested. In addition, the estimated fitting error sequence is summed with the output value of the EDM of the battery to be tested to acquire the SOH estimation for the whole cycle of the battery to be tested. This achieves error compensation for the SOH estimation between different batteries. This process is expressed in (23):

$$T_{\text{test}} + \{\text{EDM}_i\}_{i=1,2\dots\text{EOL}} = \{\text{SOH}_i^e\}_{i=1,2\dots\text{EOL}} \quad (23)$$

5 Experimental results and analysis

5.1 Single battery test

In the single battery test, the LSSVM-ECM is established based on the capacity data of the first SP cycles to predict the SOH value of the battery after the SP cycle. To verify the accuracy and robustness of the LSSVM-ECM for SOH estimation, the larger SP and the smaller SP are set up, and contrastive experiments with the LSSVM method are carried out. All of the standardized cycles of Cell1, Cell2, Cell3, Cell7, and Cell8 in the Oxford dataset are about 80, and the smaller SP is 30 and the larger SP is 50. In addition, for Cell4, Cell5, and Cell6 whose total standardized cycles are about 50, the smaller SP is 15 and the larger SP is 30. For B0005–B0007, which have 168 cycles, the smaller SP is 80 and the larger SP is 120, and for B0029–B0032 which have 40 cycles, the smaller SP is 15 and the larger SP is 25. Results of SOH estimations and the relative error percentages for each of the batteries are shown in Figs. 7 and 8. The black curves in Figs. 7 and 8 represent the real values of the SOH, the green dotted lines represent the prediction starting points (SP) of the SOH, the green curves are the EDM established with the capacity data of the first SP cycles, and the red and blue curves represent the SOH estimation values using the LSSVM-ECM and LSSVM methods, respectively.

Figure 7 shows estimation results for the larger SP. It can be seen that the proposed method has a high SOH estimation accuracy. The relative error of the batteries (a)–(h) in the Oxford battery dataset is less than 1%, and for the NASA batteries (i)–(o), except for a few points, most of the points are less than 2%, which meets the requirements of BMS for estimation error. The values of the MAE and RMSE are given in Table 3. They are less than 1%, which indicates

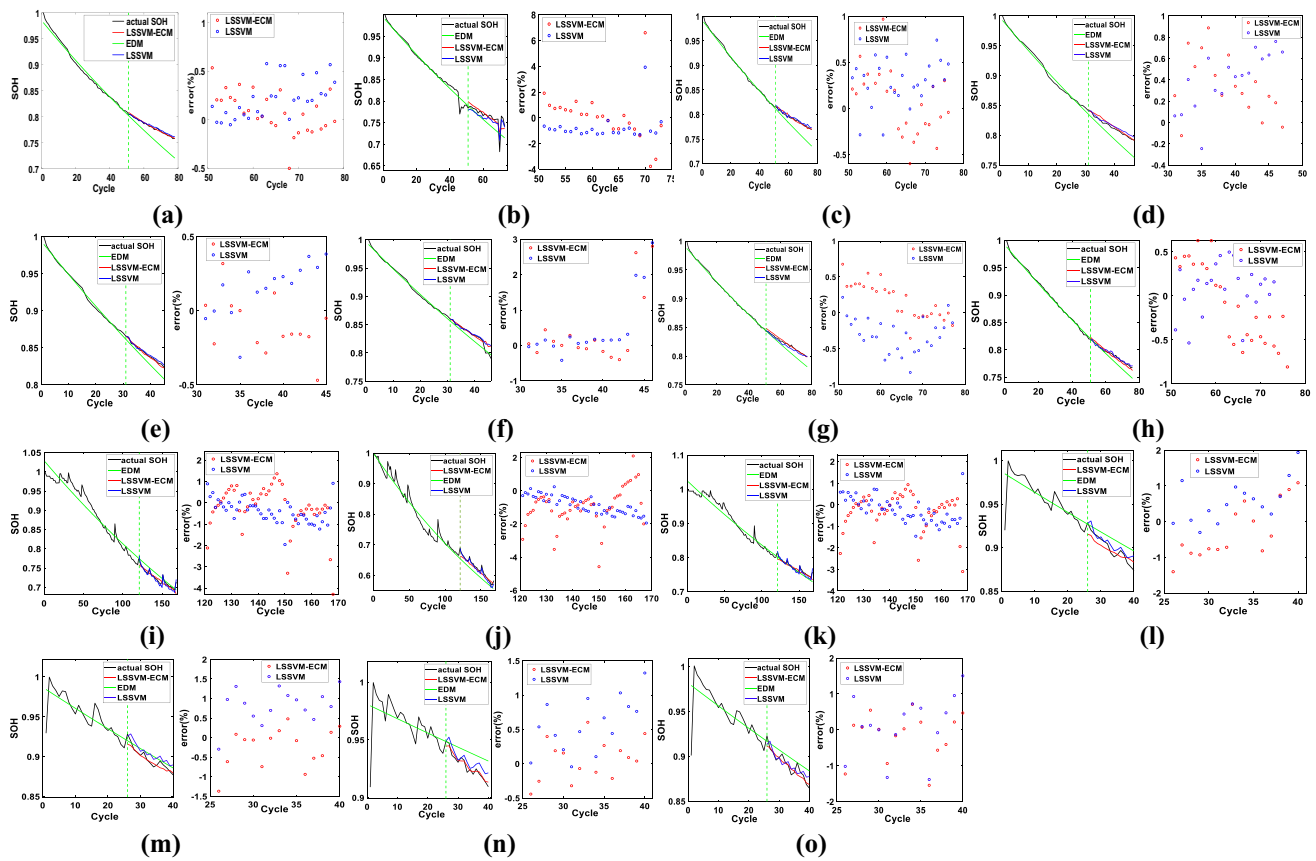


Fig. 7 Single battery test for a larger SP: **a** Cell1; **b** Cell2; **c** Cell3; **d** Cell4; **e** Cell5; **f** Cell6; **g** Cell7; **h** Cell8; **i** B0005; **j** B0006; **k** B0007; **l** B0029; **m** B0030; **n** B0031; **o** B0032

that the LSSVM-ECM proposed in this paper is suitable for various types of batteries. In addition, it can be seen from Fig. 7 and Table 3 that the LSSVM method can also achieve accurate SOH estimation when the larger SP is selected. Using this method, the relative errors of the Oxford batteries are less than 1%, and the NASA batteries, except for B0006 with 4%, are less than 2%.

Figure 8 shows estimation results for the smaller SP. It can be seen that the SOH estimation results with the LSSVM method show divergence at different degrees. For instance, the estimation errors of Cell11, Cell7, and B0005–B0007 reach more than 10%. Conversely, the LSSVM-ECM method still maintains a high accuracy. The estimation errors of the Oxford batteries are less than 1%. Except for B0006 and B0007 with 5%, the estimation errors of the batteries in the NASA dataset are less than 2%. It can be seen from Table 3 that the values of the MAE and RMSE of the LSSVM-ECM slightly increase or decrease on the basis of the same order of magnitude when the SP is larger or smaller. Meanwhile, that of the LSSVM increases dramatically. This is due to the fact that the latter, as a pure data-driven method, needs adequate training samples, which can fully learn and map the details of capacity degradation to establish the capacity

degradation model. In addition, the insufficient samples tend to result in bad performance. On the other hand, the proposed method can maintain a high accuracy for both the larger SP and the smaller SP without significant changes. This is due to the fact that the EDM ensures the overall trend of capacity degradation and enhances the robustness of this method. In addition, capacity fluctuation can be well-reflected by feedback and compensation of the error, which represents the actual difference of battery degradation. In practical applications, this method can ensure the estimation accuracy, and reduce the training samples. In other words, it reduces the number of regular capacity calibrations, which saves time and labor.

5.2 Different battery test

When the capacity degradation for all of the cycles of the battery to be tested is unknown, the SOH can be estimated by the HF and degradation information of other batteries of the same type. The degradation trends of the batteries are different due to battery inconsistencies, which results in the poor transplant ability of the EDM. Therefore, the LSSVM-ECM method is used to dynamically compensate

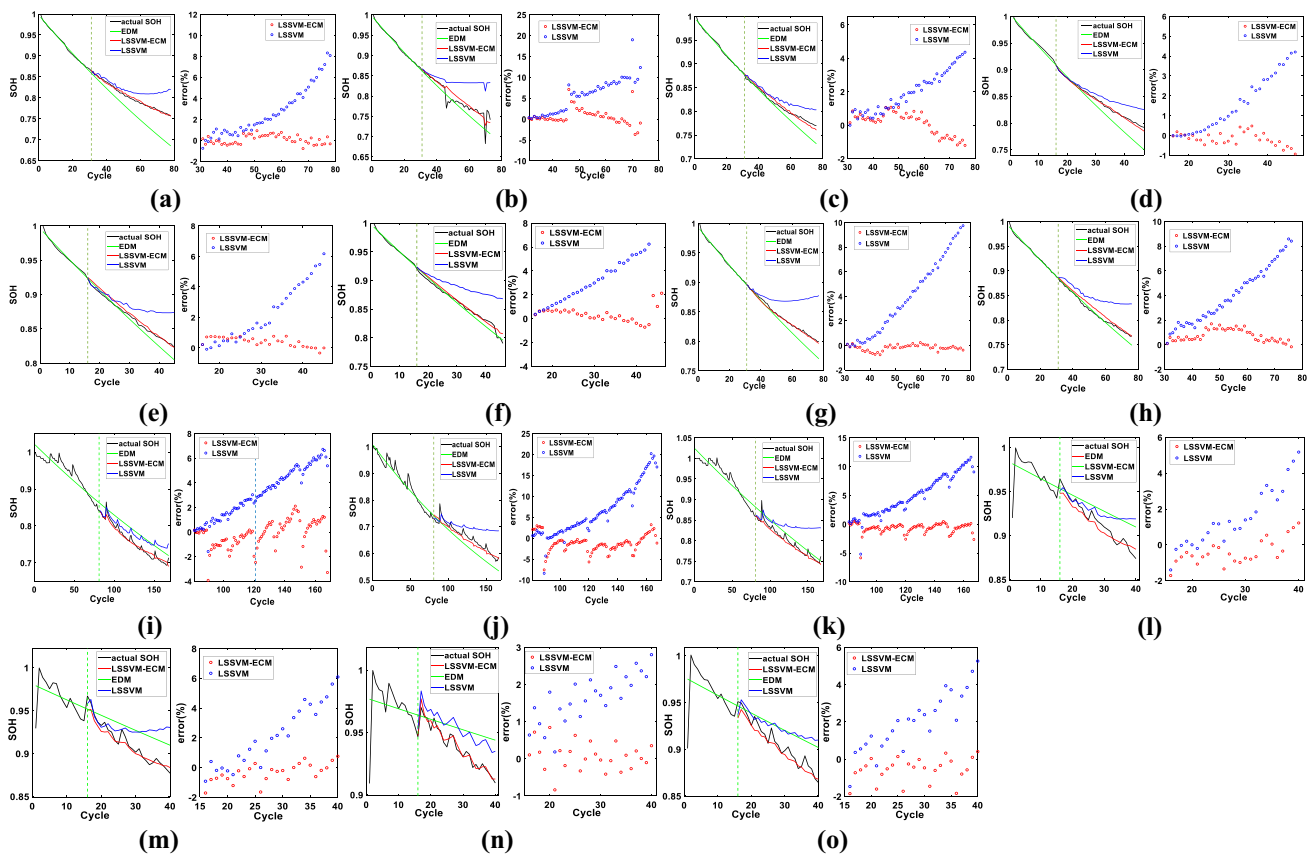


Fig. 8 Single battery test for a smaller SP: **a** Cell1; **b** Cell2; **c** Cell3; **d** Cell4; **e** Cell5; **f** Cell6; **g** Cell7; **h** Cell8; **i** B0005; **j** B0006; **k** B0007; **l** B0029; **m** B0030; **n** B0031; **o** B0032

Table 3 Estimation errors of a single battery test

| Battery | LSSVM-ECM | | | | LSSVM | | | |
|---------|-----------|--------|----------|--------|----------|--------|----------|--------|
| | Large SP | | Small SP | | Large SP | | Small SP | |
| | MAE | RMSE | MAE | RMSE | MAE | RMSE | MAE | RMSE |
| Cell1 | 0.0014 | 0.0018 | 0.0029 | 0.0034 | 0.0017 | 0.0023 | 0.0213 | 0.0274 |
| Cell2 | 0.0090 | 0.0137 | 0.0100 | 0.0159 | 0.0080 | 0.0092 | 0.0446 | 0.0535 |
| Cell3 | 0.0025 | 0.0030 | 0.0049 | 0.0055 | 0.0026 | 0.0030 | 0.0155 | 0.0180 |
| Cell4 | 0.0030 | 0.0036 | 0.0024 | 0.0030 | 0.0035 | 0.0039 | 0.0129 | 0.0170 |
| Cell5 | 0.0016 | 0.0019 | 0.0037 | 0.0045 | 0.0017 | 0.0019 | 0.0189 | 0.0244 |
| Cell6 | 0.0047 | 0.0083 | 0.0045 | 0.0059 | 0.0045 | 0.0081 | 0.0306 | 0.0361 |
| Cell7 | 0.0020 | 0.0025 | 0.0019 | 0.0025 | 0.0027 | 0.0031 | 0.0320 | 0.0403 |
| Cell8 | 0.0033 | 0.0036 | 0.0062 | 0.0073 | 0.0020 | 0.0024 | 0.0322 | 0.0365 |
| B0005 | 0.0055 | 0.0082 | 0.0066 | 0.0087 | 0.0045 | 0.0052 | 0.0230 | 0.0267 |
| B0006 | 0.0078 | 0.0098 | 0.0082 | 0.0166 | 0.0051 | 0.0060 | 0.0394 | 0.0530 |
| B0007 | 0.0048 | 0.0072 | 0.0060 | 0.0093 | 0.0039 | 0.0047 | 0.0379 | 0.0461 |
| B0029 | 0.0066 | 0.0078 | 0.0045 | 0.0081 | 0.0065 | 0.0087 | 0.0149 | 0.0217 |
| B0030 | 0.0047 | 0.0054 | 0.0025 | 0.0071 | 0.0073 | 0.0084 | 0.0204 | 0.0281 |
| B0031 | 0.0026 | 0.0030 | 0.0035 | 0.0040 | 0.0058 | 0.0070 | 0.0170 | 0.0181 |
| B0032 | 0.0048 | 0.0065 | 0.0067 | 0.0088 | 0.0064 | 0.0081 | 0.0199 | 0.0251 |

the fitting error of EDM to realize the SOH estimation of different batteries.

According to the total cycle times, the Oxford battery dataset is divided into two groups: Cell 1, Cell 2, Cell 3, Cell7, and Cell 8, whose total standardization cycles are around 80, and Cell 4–Cell 6, whose total cycles are about 50. For the NASA data sets, B0005–B0007 are in a group, and B0029–B0032 are in another group. The EDM model is established by taking one battery as the basis battery. The LSSVM-ECM is established by taking the offline HF data and the fitting error of the EDM of the M-1 training batteries to predict the capacity degradation of the testing battery. The setup of the training set and the test set are shown in

Table 4 Different battery test setup

| Battery | Label | LSSVM-ECM | | EDM | |
|---------|-------|-----------|--------|--------|--------|
| | | MAE | RMSE | MAE | RMSE |
| Cell2 | (a) | 0.0060 | 0.0091 | 0.0094 | 0.0165 |
| Cell3 | (b) | 0.0055 | 0.0065 | 0.0090 | 0.0096 |
| Cell7 | (c) | 0.0140 | 0.0147 | 0.0300 | 0.0323 |
| Cell8 | (d) | 0.0025 | 0.0029 | 0.0126 | 0.0139 |
| Cell5 | (e) | 0.0024 | 0.0031 | 0.0171 | 0.0194 |
| Cell6 | (f) | 0.0043 | 0.0050 | 0.0127 | 0.0146 |
| B0006 | (g) | 0.0233 | 0.0248 | 0.0989 | 0.1133 |
| B0007 | (h) | 0.0090 | 0.0113 | 0.0199 | 0.0274 |
| B0030 | (i) | 0.0015 | 0.0082 | 0.0136 | 0.0186 |
| B0031 | (j) | 0.0131 | 0.0168 | 0.0293 | 0.0353 |
| B0032 | (k) | 0.0062 | 0.0095 | 0.0126 | 0.0210 |

Table 4. The blue curves in Fig. 9 show the SOH estimation results for different batteries by the EDM established with the basis battery. It shows the significant error, which indicates poor transplant ability. The red curve shows SOH estimation results after error compensation by the proposed method. It can be seen that for the Oxford data set (a)–(f), except for individual points, the relative error percentage of each battery is less than 2%, and less than 5% for the NASA data set (g)–(l), which is greatly improved on the basis of the EDM estimation results. Meanwhile, local fluctuations of the capacity can be reflected well, which matches the actual capacity degradation. It can be seen from Table 5 that the values of the MAE and RMSE of the proposed method in the Oxford dataset are less than 1%, except for Cell7 with 1.5%. It can also be seen that they are less than 2% in the NASA dataset, except for B0006 with 2.5%, which are far less than that of the EDM method. Test results show that the LSSVM-ECM method has strong nonlinear mapping performance, which can establish the regression relationship between the HF and the fitting error of the EDM for different batteries. This is done to effectively predict and compensate the fitting error, and improve the transplant ability of the EDM to realize accurate SOH estimation for different batteries.

6 Conclusion

In this paper, a SOH estimation method based on the LSSVM-ECM is proposed, which realizes the fusion and complementation of an empirical model and a data-driven

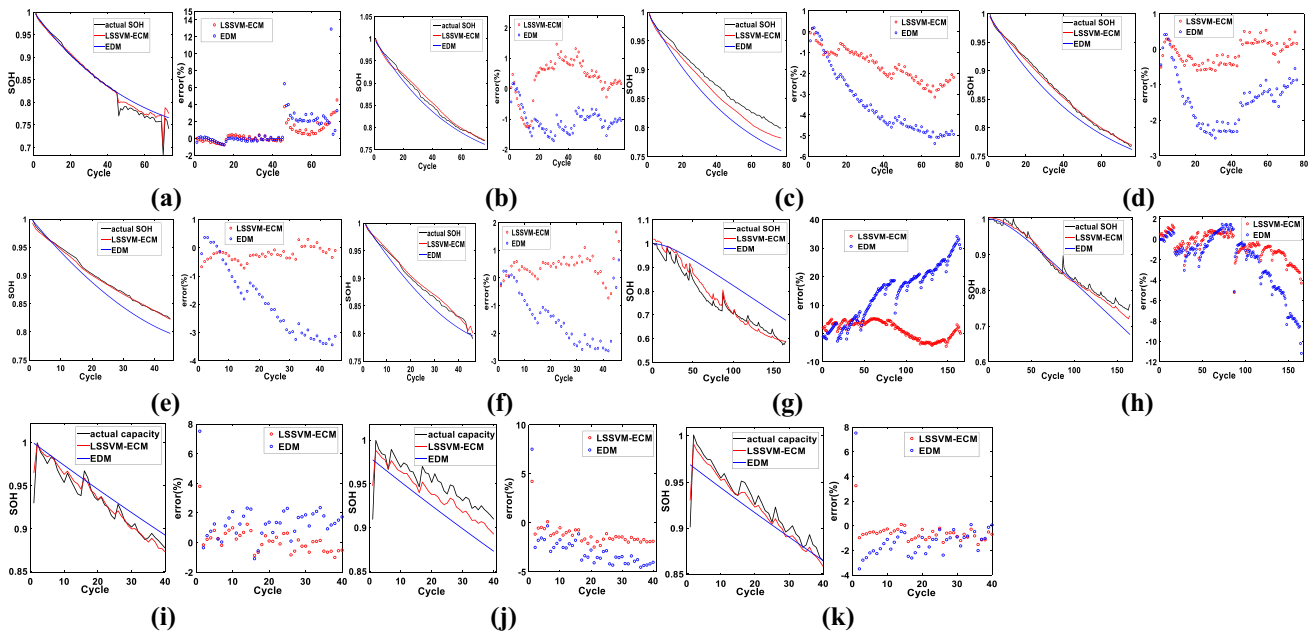


Fig. 9 Different battery test: **a** Cell2; **b** Cell3; **c** Cell7; **d** Cell8; **e** Cell5; **f** Cell6; **g** B006; **h** B007; **i** B0030; **j** B0031; **k** B0032

Table 5 Estimation errors of a different battery test

| Basis battery | Training battery | Testing | Label |
|---------------|----------------------------|---------|-------|
| Cell1 | Cell1, Cell3, Cell7, Cell8 | Cell2 | (a) |
| | Cell1, Cell2, Cell7, Cell8 | Cell3 | (b) |
| | Cell1, Cell2, Cell3, Cell8 | Cell7 | (c) |
| | Cell1, Cell2, Cell3, Cell7 | Cell8 | (d) |
| Cell4 | Cell4, Cell6 | Cell5 | (e) |
| | Cell4, Cell5 | Cell6 | (f) |
| B0005 | B0005, B0007 | B0006 | (g) |
| | B0005, B0006 | B0007 | (h) |
| B0029 | B0029, B0031, B0032 | B0030 | (i) |
| | B0029, B0030, B0032 | B0031 | (j) |
| | B0029, B0030, B0031 | B0032 | (k) |

method. The EDM is used to describe the overall trend of capacity degradation. The time interval of equal charging voltage rising is selected as a health feature to reflect the local difference in terms of capacity degradation. The LSSVM-ECM is established with the DV_DT as the input and the fitting error of the EDM as the output. This is done to dynamically compensate the prediction results of the EDM. Validations are conducted based on Oxford and NASA battery data obtained under different operating conditions. Test results demonstrate that the proposed SOH estimation method has high estimation performance, which can effectively reduce the number of training samples and improve robustness when compared to the traditional data-driven method. In addition, it can improve the transplant ability over the empirical model method.

References

- Sarmah, S.B., Kalita, P., Garg, A.: A review of state of health (SoH) estimation of energy storage systems: challenges and possible solutions for futuristic applications of Li-ion battery packs in electric vehicles. *J. Electrochem. Energy Convers. Storage*. **16**(4), 040801–040810 (2019)
- Park, S., Ahn, J., Kang, T.: Review of state-of-the-art battery state estimation technologies for battery management systems of stationary energy storage systems. *J. Power Electron*. **12**(1), 1–9 (2020)
- Kim, J., Lee, S., Cho, B.: Discharging/charging voltage-temperature pattern recognition for improved SOC/capacity estimation and SOH prediction at various temperatures. *J. Power Electron*. **20**(6), 1526–1540 (2012)
- Liu, C., Wang, Y., Chen, Z.: Degradation model and cycle life prediction for lithium-ion battery used in hybrid energy storage system. *Energy* **166**, 796–806 (2019)
- Doyle, M.: Modeling of galvanostatic charge and discharge of the lithium/polymer/insertion cell. *J. Electrochem. Soc.* **140**(6), 1526–1534 (1993)
- Pang, H., Mou, L., Guo, L.: Parameter identification and systematic validation of an enhanced single-particle model with aging degradation physics for Li-ion batteries. *Electrochim. Acta* **307**, 474–487 (2019)
- Xing, Y., Ma, E.W.M., Tsui, K.L.: An ensemble model for predicting the remaining useful performance of lithium-ion batteries. *Microelectro. Rel.* **53**(6), 811–820 (2013)
- Qin, W., Lv, H.C., Liu, C.L.: Remaining useful life prediction for lithium-ion batteries using particle filter and artificial neural network. *Ind. Manag. Data Syst.* **120**(2), 312–328 (2020)
- Tseng, K.H., Liang, J.W., Chang, W.: Regression models using fully discharged voltage and internal resistance for state of health estimation of lithium-ion batteries. *Energies* **8**(4), 2889–2897 (2015)
- Wang, D., Miao, Q., Pecht, M.: Prognostics of lithium-ion batteries based on relevance vectors and a conditional three-parameter capacity degradation model. *J. Power Sources* **239**, 253–264 (2013)
- Han, X., Ouyang, M., Languang, Lu.: A comparative study of commercial lithium ion battery cycle life in electric vehicle: capacity vehicle: capacity loss estimation. *J. Power Sources* **268**, 658–669 (2014)
- Yang, Q., Xu, J., Li, X.: State-of-Health estimation of lithium-ion battery based on fractional impedance model and interval capacity. *Elect. Power Energy Syst.* **119**, 105883–105891 (2020)
- Yi, L., Abdel-Monem, M., Gopalakris-Hnan, R.: A quick online state of health estimation method for Li-ion battery with incremental capacity curves processed by Gaussian filter. *J. Power Sources* **373**, 40–53 (2018)
- Che, Y., Liu, Y., Cheng, Z.: SOC and SOH identification method of Li-ion battery based on SWPSO-DRNN. *IEEE J. Emerg. Select. Top. Power Electron.* **9**(4), 4050–4061 (2021)
- Tian, J., Xiong, R., Shen, W.: State-of-health estimation based on differential temperature for lithium ion batteries. *IEEE Trans. Power Electron.* **35**(10), 10363–10373 (2020)
- Gao, D., Huang, M.: Prediction of remaining useful life of lithium-ion battery based on multi-kernel support vector machine with particle swarm optimization. *J. Power Electron.* **17**(5), 1288–1297 (2017)
- Peiyao, G., Ze, C., Lei, Y.: A data-driven remaining capacity estimation approach for lithium-ion batteries based on charging health feature extraction. *J. Power Sources* **412**, 442–450 (2019)
- Xiong, W., Mo, Y., Yan, C.: Online state-of-health estimation for second-use lithium-ion batteries based on weighted least squares support vector machine. *IEEE Access* **9**, 1870–1881 (2021)
- Meng, J., Luo, G., Gao, F.: Lithium polymer battery state-of-charge estimation based on adaptive unscented kalman filter and support vector machine. *IEEE Trans. Power Electron.* **31**(3), 2226–2238 (2016)
- Birkel, C.: Oxford Battery Degradation Dataset 1. Univ. Oxford, Oxford (2017)
- Saha, B. and Goebel, K.: “Battery data set,” NASA Ames Prognostics Data Repository. <http://ti.arc.nasa.gov/project/prognostic-data-repository> (2017)
- Danković, D., Manić, I., Djorić, S.: NBT stress-induced degradation and lifetime estimation in p-channel power VDMOSFETs[J]. *Microelectron. Reliabil.* **46**, 1828–1833 (2006)
- Choi, C., Lee, J.C.: Bulk and interface trap generation under negative bias temperature instability stress of p-channel metal-oxide-semiconductor field-effect transistors with nitrogen and silicon incorporated HfO₂ gate dielectrics. *Appl. Phys. Lett.* **98**, 063504–063508 (2011)
- Manić, D., Danković, A., Pijrić, NBTI related degradation and lifetime estimation in p-channel power VDMOSFETs under the static and pulsed NBT stress conditions. *Microelectron. Reliabil.* **51**, 1540–1543 (2011)

25. Ma, S., Jiang, M., Tao, P.: Temperature effect and thermal impact in lithium-ion batteries: a review. *Prog. Nat. Sci.: Mater. Int.* **28**, 653–666 (2018)
26. Yun, Z., Qin, W., Shi, W.: State-of-health prediction for lithium-ion batteries based on a novel hybrid approach. *Energies* **13**, 4858–4868 (2020)



Ji'ang Zhang was born in Shandong, China, in 1995. He received his B.S. degree in Electrical Engineering from Shandong University, Shandong, China, in 2018. He is presently working towards his M.S. degree in Electrical Engineering at Tianjin University, Tianjin, China. His current research interests include the modelling and state estimation of lithium-ion batteries.



Ping Wang was born in Tianjin, China, in 1959. She received her B.S., M.S., and Ph.D. degrees in Electrical Engineering from Tianjin University, Tianjin, China, in 1981, 1991, and 2005, respectively. Since 1981, she has been a Teacher and a Researcher at Tianjin University, where she is presently working as a Professor. Her current research interests include lithium-ion battery management systems, power electronic control of renewable energy sources, PWM converters, and intelligent detection and

control.



Qingrui Gong was born in Shanxi, China, in 1998. He received his B.S. degree in Electrical Engineering from the Taiyuan University of Technology, Shanxi, China, in 2020. He is presently working towards his M.S. degree in Electrical Engineering at Tianjin University, Tianjin, China. His current research interests include machine learning algorithms and their application in the state estimation of lithium-ion batteries.



Ze Cheng was born in China, in 1959. He received his M.S. and Ph.D. degrees from Tianjin University, Tianjin, China, in 1985 and 2002, respectively. He is presently working as a Professor in the School of Electrical and Information Engineering, Tianjin University. His current research interests include photovoltaic cell fault modeling, power generation prediction, lithium battery model parameter identification, wireless charging theory and technology development, and UAV track planning.