

## State-of-health estimation of lithium-ion batteries based on semi-supervised transfer component analysis

Yuanyuan Li<sup>a</sup>, Hanmin Sheng<sup>a</sup>, Yuhua Cheng<sup>a,\*</sup>, Daniel-Ioan Stroe<sup>b</sup>, Remus Teodorescu<sup>b</sup>

<sup>a</sup> School of Automation Engineering, University of Electronic Science and Technology of China, No. 2006, Xiyuan Ave, West Hi-Tech Zone, Chengdu 611731, China

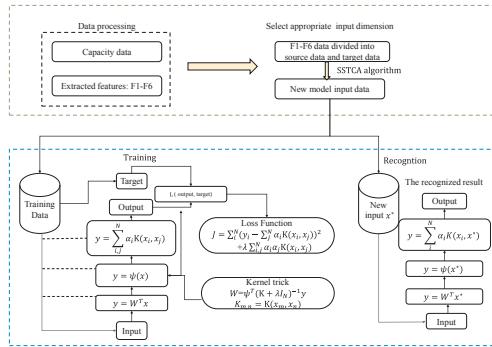
<sup>b</sup> Department of Energy Technology Aalborg University, Aalborg 9220, Denmark

### HIGHLIGHTS

- The idea of knowledge transfer is proposed for estimating battery state of health.
- A small amount of samples are used to achieve accurate state of health estimation.
- Multi-Features indication of battery state of health are extracted.
- The correlation between features and targets is measured by mutual information.
- The state of health is estimated effectively with less than 2.5% by various errors.

### GRAPHICAL ABSTRACT

Describes how the semi-supervised transfer component analysis method estimates the battery SOH based on the KRR model. Shows the process of the t semi-supervised transfer component analysis method.



### ARTICLE INFO

#### Keywords:

Lithium-ion battery  
Transfer learning  
Transfer component analysis  
Mutual information  
State of health estimation

### ABSTRACT

Accurate state-of-health estimation can ensure the safe and reliable operation of Lithium-ion batteries in any given application. Nevertheless, most of the state-of-health estimation methods require a large amount of laboratory aging data to offer precise results. As obtaining battery aging data under laboratory conditions requires a considerable amount of time and incurs high economic costs, in this paper, a method based on transfer learning is proposed to monitor state-of-health of batteries. A novel data processing method based on maximum mean discrepancy is considered to eliminate redundant information and minimize the difference between different data distributions. Then, mutual information is used to prove that the correlation between processed data is not decreased. To validate the developed transfer learning method, the data sets of four batteries in different working conditions are considered. Different error-detection methods, maximum average error, mean squared error and root mean squared error, which are utilized to evaluate the proposed model. The state of health is estimated effectively with less than 2.5% error considering the aforementioned errors after processed by using semi-supervised transfer component analysis algorithm, although the training set only accounts for about 35% of the entire set. The results indicate that transfer learning plays an important role in improving the estimation accuracy of a battery state-of-health.

\* Corresponding author.

E-mail addresses: [liyuanyuanfy@163.com](mailto:liyuanyuanfy@163.com) (Y. Li), [hmsheng8911@hotmail.com](mailto:hmsheng8911@hotmail.com) (H. Sheng), [yhcheng@uestc.edu.cn](mailto:yhcheng@uestc.edu.cn) (Y. Cheng), [dis@et.aau.dk](mailto:dis@et.aau.dk) (D.-I. Stroe), [ret@et.aau.dk](mailto:ret@et.aau.dk) (R. Teodorescu).

<b>Nomenclature</b>	
CC	Constant current.
CV	Constant voltage.
ECM	equivalent circuit model.
EIS	electrochemical impedance spectroscopy.
EVD	Equal voltage drop.
GPR	Gaussian process regression.
IC	Incremental capacity.
ICA	Incremental capacity analysis.
KRR	Kernel ridge regression.
MAE	Maximum average error.
MMD	Maximum mean discrepancy.
MSE	Mean squared error.
PCA	Principal Component Analysis.
RMSE	Root mean squared error.
SampEn	Sample entropy.
SOC	State of charge.
SOH	State of health.
SSTCA	Semi-supervised transfer component analysis.
TCA	Transfer component analysis.

## 1. Introduction

Lithium-ion batteries have better performances in terms of high energy density [1], high galvanic potential [2], long lifetime [3] and low self-discharge rate [4] in comparison to more traditional battery technologies such as lead-acid or nickel metal hydride. Subsequently, they have become ubiquitous for high-power energy storage applications like communications and aerospace. Furthermore, they have become to key energy storage technology for powering electric vehicles [5]. Independent of the aforementioned applications, Lithium-ion battery cells are connected in packs in order to reach the desired power levels. However, they tend to perform inconsistently due to manufacturing tolerances and inhomogeneous temperature distributions. These result in different performance and degradation behaviors which directly affects the power-supply performance [6]. Battery management systems can effectively supervise the power supplies for equipment, maintain the normal use of power equipment, and avoid accidents [7]. One of the main roles of battery management systems is to provide accurate knowledge regarding the internal states of batteries, such as its state-of-health (SOH) [8], state-of-charge (SOC) [9], state-of-power (SOP) [10]. The SOH characterizes the health status of a battery, which is usually represented in terms of the capacity or power loss. Thus, because it cannot be directly measured, the prediction of the battery SOH has become of great importance. In addition, owing to the series and parallel connections in each battery pack, predicting the SOH is a challenging task.

Battery SOH estimation has gained considerable research attention in recent years, and many estimation methods have been reported. These methods can be divided into two main categories: (1) model-based methods and (2) data-driven methods.

The model-based methods consider the battery load condition, material properties, and degradation mechanism to predict the SOH for batteries. Many theoretical results have been reported for model-based methods, such as electrochemical models, electrochemical impedance spectroscopy models, and equivalent circuit models, where the battery degradation models are established based on electrochemical or equivalent circuit models. For instance, to predict the overall behavior and the cell-to cell variation of a battery layout, the method based on Pseudo-Two-Dimensional(P2D) porous electrode model has been reported [11]. To deal with the dispersion effect, Tian et al. [12] has proposed a fractional-order model based on the EIS for degradation state recognition of lithium-ion batteries. Ma et al. [13] has presented a novel data-model fusion battery SOH estimation approach based on open-circuit-voltage parametric modeling, which considers the correlation between capacity degradation and the open-circuit-voltage changes. Consider the robustness of the proposed method, the estimation method based on parameter adaptive sliding mode observers has been reported [14], which also has been verified through the urban dynamometer driving schedule driving cycle. Xiong et al. [15] has proposed an online estimation of remaining capacity based on SEI resistance. Although the electrochemical model has a clear physical meaning, the chemical system of the battery is definitely complex and it

also depends on the specific experiment equipment. Owing to the coupling effect of multiple factors in the degradation process, it is difficult to accurately grasp the internal aging mechanism and identify the related parameters, which are rather used for the design and improvement of battery material and processes. In addition, the EIS measurement is sensitive to noise in practical applications and relies on specific experimental equipment. Moreover, although the method based on ECM are using conventional circuit parameters, these parameters are affected greatly by the working conditions, complicating the parameter identification, and thus the ability of this method to describe the comprehensive dynamic and static characteristics of a battery is weak.

Data-driven methods have become important in battery field owing to their superior nonlinear mapping capacity. Compared with the model-based method, the data-driven method, which considers the battery as a "black box" uses machine learning to analyze data. This method has good data integrity and does not need the tedious manual model process, which can autonomously learn the nonlinear relationship between battery health status and external characteristics through the data. The commonly used data-driven methods include neural networks (NNs) [16], support vector machines (SVMs) [17], Gaussian process regression (GPR)[18], and dynamic Bayesian networks (DBNs) [19]. You et al. [20] has developed a data-driven framework to estimate SOH on the fly, and it validated by actual EV driving conditions. Meng et al. has [21] found the optimal voltage ranges for SOH estimation with a partial voltage profile. Sui et al. [22] has estimated the battery SOH based on SVM algorithm. Compared with NNs and SVMs, GPR and DBNs are probability estimation methods that can assess the model confidence while providing the prediction result; they control the prediction model. For example, He et al. [19] has developed an online method based on DBNs for estimating the SOH of lithium-ion batteries. In addition, owing to their strong impact on the performance of the trained model, features are considered to be fundamental parts of most practical data-driven methods. Most of the previously propose data-driven methods analyzed features related to the battery-aging process to estimate the SOH. These methods are aimed at establishing a mapping between the healthy features of the battery and its SOH, e.g., the differential voltage method [23], internal resistance analysis [24], and incremental capacity (IC) method [25]. For instance, the IC peaks, IC valleys, and their corresponding voltage values are used as the feature to estimate battery SOH [26]. Guo et al. [27] has selected the area under the current curve/ temperature curve of CC phase/ CV phase/ whole charge process, CC time, CV time, the ratio of CC time to CV time, the ratio of area under temperature curve to the corresponding area under the current curve, and the maximum slope of current curve in CC phase/CV phase as the feature. Hu et al. [28] has used the sample entropy of voltage as the feature to estimate the battery SOH. The discharging voltage difference of equal time interval as the feature were proposed in [29]. Guha et al. [30] has chose the internal resistance analysis to establish the correlation with capacity fade. Duo et al. [31] has selected some health features to obtain a charging curve in order to map the relationship with the battery SOH by using GPR, the features are time of CC/CV mode duration, the slope of the curve at the end of

CC charging mode, and the vertical slope at the corner of the CC charging curve. However, this method does not consider the redundancy of the health feature information and the correlations with the target SOH, thereby resulting in a low accuracy rate and over-fitting owing to a lack of versatility and comprehensiveness.

The training and updating of machine learning models depend on a large volume of labeled data. In practical applications, obtaining battery-aging data under experimental conditions requires large amounts of time and economic costs. Although massive sensor data provide a large amount of data for the data-driven model, it is not possible to obtain reliable labeled data, as accurate measurement is difficult to be achieved using battery-health indicators. When the label data of the training set is insufficient, this will directly affect the training effect of the model. Another situation is in the application process, for a new operating environment and a new type battery, machine learning assumption will be affected, data must be recollected and the model must be retrained. Traditional machine learning is generally based on the same distribution assumption, that is, the training set and the testing set are consistent. Thus, when the battery label data of training set is insufficient or the battery faces a new working environment, the increase of the difference between the data distribution will affect this assumption. Therefore, the previously obtained label data is already not suitable for the changed scene. In other words, the previous label data is used as the training set, and the result is not satisfactory.

In this case, knowledge extracted from a related domain can be employed to help a machine learning algorithm perform better in the target domain. The knowledge transfer concept is a new approach for addressing the lack of labeled data in a small volume of data, and this approach has been applied to computer vision tasks [32], object recognition [33], WiFi Localization [34], image classification [35], bioinformatics [36], and handwriting recognition [37]. We treat related domain as training sets, the target domain as the testing sets. This method can achieve information sharing between the training set and the testing set. If the data between the training sets and the testing sets is highly correlated, then we can effectively apply the information of the training sets. If the correlation between these two sets is small, let the data with the smallest correlation affect the test set as much as possible. Its main purpose is to enhance the correlation between the training sets and the testing sets. We can measure the correlation by the difference between the data sets distribution.

This paper aims at establishing a mapping between the healthy features of the battery and its SOH. As mentioned before, multiple

features are proposed in [26], but these methods do not consider the redundancy of the health feature information and the correlations with the target SOH. In other words, it is not clear if or how much each feature plays a role in the training of the SOH estimation model. Guo et al. [27] has considered redundant information between features, and used principal component analysis (PCA) algorithm to reduce the dimension of features, but the proportion of its training set accounts for 70% of all the data. Moreover the training set test set is randomly divided, which means the training set will cover most of the information in the entire data set. Duo et al. [31] has chosen the first 49% of a data set as the training set. Even if the training set is not randomly divided, it also accounts for half of the entire data set. However, in this paper, we not only employ the first 35% of a data set to train the model, but also consider the redundant information between features.

Hence, in this paper, we propose a battery-health-assessment model based on a semi-supervised transfer component analysis (SSTCA) algorithm. The training of the model depends on feature selection; however, the redundant information between features is often ignored and the distributions of the features, thereby directly affecting the effectiveness of the model training process. Therefore, it is necessary to develop a method that eliminates redundant information between features and minimizes the distance between the data distributions of the features. By using semi-supervised learning, labeled data from a small volume of experimental data are employed to complete the mapping process in the latent feature space. In addition, we introduce the maximum mean discrepancy (MMD) to minimize the distance between the feature distributions in different periods and preserve the local geometric structures in the latent space. During the transferring phase, we not only employ the first 35% of a data set to train the model, but also consider the redundant information between features. This means that using a small sample data is used to perform the battery SOH estimation, which significantly improves the practicability of the proposed model. The specific framework for the proposed method is shown in Fig. 1.

The remainder of this paper is organized as follows. In Section 2, we introduce the concept of transfer learning. The processing method for experimental data is proposed in Section 3; features that can reflect the SOH of a battery are extracted from these data. The SOH estimation results are presented and discussed in Section 4. Finally, we draw our conclusions and the future work in Section 5.

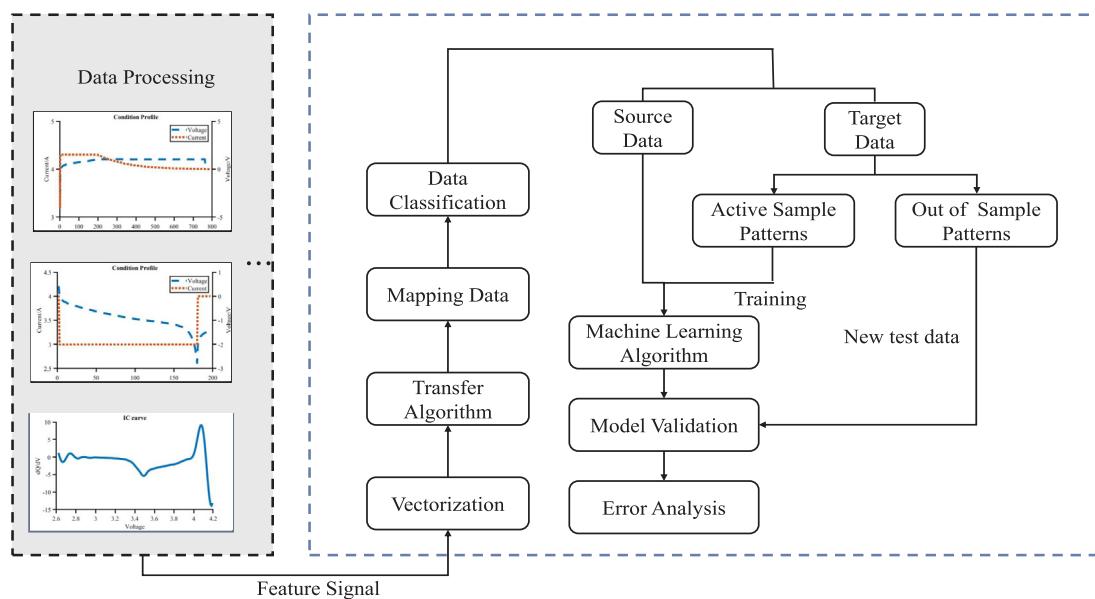


Fig. 1. SOH monitoring framework.

## 2. Transfer learning analysis

As a cross-domain technique, transfer learning usually begins with inductive, unsupervised, and transductive transfer learning. Table 1 shows the relationships between the conventional machine learning and transfer learning settings.

For batteries, the feature-distribution data can vary in different periods; thus, the number of feature-distribution data that overlap is low and the amount of knowledge learned by the model is limited; this directly affects the accuracy of the predicted results. Moreover, obtaining experimental data under different conditions is expensive, and therefore achieving the goal of predicting the SOH of a battery based on a small sample size is difficult. Thus, it is important to identify the latent factors in different time intervals that minimize the distances between the domain data distributions, as well as improve the overlapping between domain data and preserve the properties of the original data.

### 2.1. SSTCA algorithm

According to the aforementioned discussions, it is necessary to reduce the distances between feature-distributions data in different time intervals. Many previous studies have focused on the design and optimization of objective functions by using the Euclidean distance to measure the dissimilarity between instances. However, the Euclidean distance may be inappropriate for capturing the intrinsic similarity or dissimilarity between instances in some normal cases. Thus, we employed the MMD [38] to measure the differences in the distributions between the source and target domains in the proposed method. Let the kernel-induced feature map be  $\phi$ . The empirical estimate of MMD between  $X = \{x_i, \dots, x_n\}$  and  $Y = \{y_1, \dots, y_m\}$  is  $MMD(X, Y) = \left\| \frac{1}{n} \sum_{i=1}^n \phi(x_i) - \frac{1}{m} \sum_{j=1}^m \phi(y_j) \right\|_{\mathcal{H}}^2$ , where  $\|\cdot\|_{\mathcal{H}}$  is the reproducing kernel Hilbert space (RKHS) norm. The distance between two distributions is now the distance between the two mean elements in an RKHS. According to a previous study [39], RKHS is universal while the MMD asymptotically approaches zero if and only if the two distributions are the same.

We consider a large volume of labeled source domain data  $D_S = \{(x_{S_i}, y_{S_i}), \dots, (x_{S_n}, y_{S_n})\}$ , where  $x_{S_i} \in \mathcal{X}$  is the input and  $y_{S_i} \in \mathcal{Y}$  is the corresponding output, and certain unlabeled target domain data  $D_T = \{x_{T_l}, \dots, x_{T_m}\}$ , where input  $x_{T_l} \in \mathcal{X}$ . We assume  $P(X_S)$  and  $Q(X_T)$  are the marginal distributions of input sets  $X_S$  and  $X_T$  from the source and target domains, respectively. In general,  $P$  is different from  $Q$ . Our task is to predict label  $y_{T_j}$  that corresponds to input  $x_{T_j}$  in the target domain. We assume that  $P \neq Q$ ; however, transformation  $\phi$  exists such that  $P(\phi(X_S)) \approx P(\phi(X_T))$  and  $P(Y_S | \phi(X_S)) \approx P(Y_T | \phi(X_T))$ . Therefore, based on [34], our aim is to find transformation  $\phi$ , which must satisfy the following: (1) the distance between marginal distributions  $P(\phi(X_S))$  and  $P(\phi(X_T))$  is small; (2) the source domain label dependence on the latent space is maximum; and (3)  $\phi(X_S)$  and  $\phi(X_T)$  preserve some important properties of  $X_S$  and  $X_T$ .

The distance between two distributions  $P$  and  $Q$  can be written

$$Dist(X_S, X_T) = \left\| \frac{1}{n} \sum_{i=1}^n \phi(x_i) - \frac{1}{m} \sum_{j=1}^m \phi(y_j) \right\|_{\mathcal{H}}^2 \quad (1)$$

Instead of finding transformation  $\phi$  explicitly, we employed a kernel trick to replace  $\phi$ . Assume that  $\langle \phi(x_i), \phi(x_j) \rangle = k(x_i, x_j)$ , where  $k$  is the corresponding kernel; then, (2) can be rewritten as

$$\begin{aligned} Dist(X_S, X_T) &= \frac{1}{n_S^2} \sum_{i,j=1}^{n_S} k(x_{S_i}, x_{S_j}) + \frac{1}{n_T^2} \sum_{i,j=1}^{n_T} k(x_{T_i}, x_{T_j}) \\ &\quad - \frac{2}{n_S n_T} \sum_{i,j=1}^{n_S n_T} k(x_{S_i}, x_{T_j}) \\ &= tr(KL) \end{aligned} \quad (2)$$

where  $K$  is a composite kernel matrix, and  $K_S$  and  $K_T$  are the kernel matrices defined by  $k$  based on the data in the source and target domains, respectively.

$$K = \begin{bmatrix} K_{S,S} & K_{S,T} \\ K_{T,S} & K_{T,T} \end{bmatrix} \in \mathcal{R}^{(n_S+n_T) \times (n_S+n_T)} \quad (3)$$

$$L = \begin{cases} \frac{1}{n_S^2} & x_i, x_j \in X_S, \\ \frac{1}{n_T^2} & x_i, x_j \in X_T, \\ -\frac{2}{n_S n_T} & otherwise. \end{cases} \quad (4)$$

In the proposed method, a dimensionality reduction-based domain adaptation method called MMD embedding (MMDE) is employed. Thus, the objective function of MMDE can be written as

$$\min_{K \geq 0} tr(KL) - \lambda tr(K) \quad subject to constraints on K, \quad (5)$$

where the first term in (6) is the objective that minimizes the distance between the distributions, the second term preserves the properties of the feature space, and  $\lambda > 0$  is the tradeoff parameter. Now, we let kernel matrix  $K$  be decomposed as  $K = (KK^{-1/2})(K^{-1/2}K)$ , which is often referred to as the empirical kernel map [40]. And the distance between two distributions  $P$  and  $Q$  can be rewritten as follows.

$$Dist(X_S, X_T) = tr((KWW^T K)L) = tr(W^T KLW) \quad (6)$$

Furthermore, to maximize the dependence between embedding and labels, we maximally align the embedding with

$$\widetilde{K}_{yy} = \gamma K_l + (1 - \gamma) K_v, \quad (7)$$

where  $\gamma \geq 0$ ,  $[K_l]_{ij} = k_{yy}(y_i, y_j)$  if  $i, j \geq n_S$ ; otherwise,  $[K_l]_{ij} = 0$ , which serves to maximize the dependence on the labeled data, and  $K_v = I$ , which serves to maximize the variance on both the source and target domain data. The preservation of the local geometry of the original data also need to be considered.

Thus, according to the aforementioned discussion, the final optimization problem can be written as follows.

$$\begin{cases} \min_W tr(W^T KLW) + \mu tr(W^T W) + \frac{\lambda \times tr(W^T L \mathcal{L} KW)}{(n_S + n_T)^2} \\ s. t. W^T K H \widetilde{K}_{yy} H K W = I \end{cases} \quad (8)$$

where  $\mathcal{L} = D - M$ ,  $D$  is the diagonal matrix with entries  $d_{ii} = \sum_{j=1}^{n_S+n_T}$ , and  $M = [m_{ij}]$ ,  $m_{ij} = \exp(-d_{ij}^2/2\sigma^2)$ .

To use the Lagrange multiplier methods and Karush-Kuhn-Tucker condition to solve problem (10), problem (10) can be formulated as the

**Table 1**  
Relationship between conventional machine learning and transfer learning settings.

Learning Settings	Labeled Samples for Training	Source&Target Domain	Source&Target Task
Conventional Machine Learning	Large	the same	the same
Inductive Transfer Learning	Small	the same	different but related
Unsupervised Transfer Learning	No	different but related	different but related
Transductive Transfer Learning	Small	different but related	the same

following trace problem:

$$\max_W \text{tr} \left\{ (W^T K (L + \lambda \mathcal{L}) K W + \mu I)^{-1} \left( W^T K H \tilde{K}_{yy} H K W \right) \right\}. \quad (9)$$

Similar to the principal component analysis, the problem can be solved by eigendecomposing  $(K(L + \lambda \mathcal{L})K + \mu I)^{-1} K H \tilde{K}_{yy} H K$ . The SSTCA procedure was introduced in [34]. The implementation of the transfer algorithm is shown in Fig. 2.

## 2.2. SOH estimation based on kernel ridge regression

This paper proposed the kernel ridge regression model based on the SSTCA algorithm to predict the battery SOH. Kernel ridge regression (KRR) is a combination of ridge regression (linear least squares L2 norm regularization) and kernel tricks. The implementation of the KRR algorithm is shown in Fig. 3.

We consider the regression problem,

$$y = W^T x + \varepsilon, \quad \varepsilon \sim N(0, \sigma), \quad W_i \sim N(0, 1) \quad (10)$$

where  $x$  is the sample matrix, and  $y$  is the label vector. We construct the loss function to get the values for parameter  $W$ ,

$$J(W) = (y - W^T x)^T (y - W^T x) + \lambda \|W\|^2 \quad (11)$$

In order to get the optimal value for parameter  $W$ , using the  $W$  as a variable to derive (12), and then get the

$$W = (x^T x + \lambda I)^{-1} x^T y \quad (12)$$

We now replace all data with their feature vector:  $x_i \rightarrow \psi_i = \psi(x_i)$ , this equation can be rewritten as:  $W = \sum_i \alpha_i \psi(x_i)$  with  $\alpha = (\psi^T \psi + \lambda I)^{-1} y$ . When we have a new input  $x^*$ , this is computed by projecting it onto the solution  $W$ ,

$$\begin{aligned} y &= W^T \psi(x^*) = y(\psi^T \psi + \lambda I)^{-1} \psi^T \psi(x^*) = y(K + \lambda I)^{-1} \mathcal{K}(x^*) \\ &= \sum_i \alpha_i K \left( x_i, x^* \right) \end{aligned} \quad (13)$$

where  $K = \psi^T \psi$  and  $\mathcal{K}(x^*) = K(x_i, x^*)$ .

The SSTCA algorithm is a method of data processing. It is the most important step in turning features into model inputs. After the features are processed by SSTCA, we take the processed features as model input, and this model is the KRR model. We divide the battery data into a training set and a testing set. Firstly, we input the training set based on KRR model to obtain the parameters in the KRR model, such as  $K$  and  $\alpha_i$ . Then we input the testing set, that is, the data we need to estimate

into the trained model, so that we get the estimated result.

## 3. Data processing

In the following, we introduce the data set employed in this study and the experimental conditions for the battery. These data cannot be mapped directly to the battery SOH, and hence it is necessary to identify the health features that can indicate the battery SOH in the battery data set. Thus, health features are extracted from the data set in the second step. Finally, mutual information is employed to reflect the correlations between the health features and the battery SOH.

### 3.1. Experimental data analysis

In this study, we used data sets from the NASA Ames Prognostics Center of Excellence; these data sets were harvested from a battery prognostic test bed at NASA. To determine the SOH degradation trends for batteries under different working conditions, we selected the B0005, B0006, B0007, and B0034 data sets. These data were obtained with four operational profiles at room temperature. Charge is carried in the constant current (CC) mode at 1.5 A until the battery voltage reaches 4.2 V, and then continues in the constant voltage (CV) mode until the charging current drops to 20 mA. Discharge is conducted at a CC level of 2 A until the battery voltage drops to 2.7, 2.5, and 2.2 V for the B0005, B0006, and B0007 data sets, respectively. In contrast, for the B0034 data set, discharge is conducted at 4 A until the voltage drops to 2.2 V. The experiments were stopped when the batteries satisfied the end-of-life criteria comprising a 30% fade in the rated capacity for B0005, B0006, and B0007, and a 20% fade in the rated capacity for B0034. The specific experimental conditions for the batteries are shown in Table 2, and Fig. 4 shows the aging trends for the four data sets.

### 3.2. Feature extraction

Owing to the limitations of the battery data variables, it is necessary to select the best health features from the charging and discharging curves to establish links with the battery SOH. Thus, based on previous studies [18,28,29,31], we analyze the experimental data for the batteries to obtain the health features by using general mathematical calculations. Simple analyses are conducted based on the existing data sets before modeling the SOH of batteries. In the analyses, we consider data set B0005 as an example. Figs. 5 (a) and (b) show the charging and discharging curves for battery, respectively. Fig. 5 (c) shows the IC curves and Fig. 5 (d) illustrates the sample entropy for the discharging

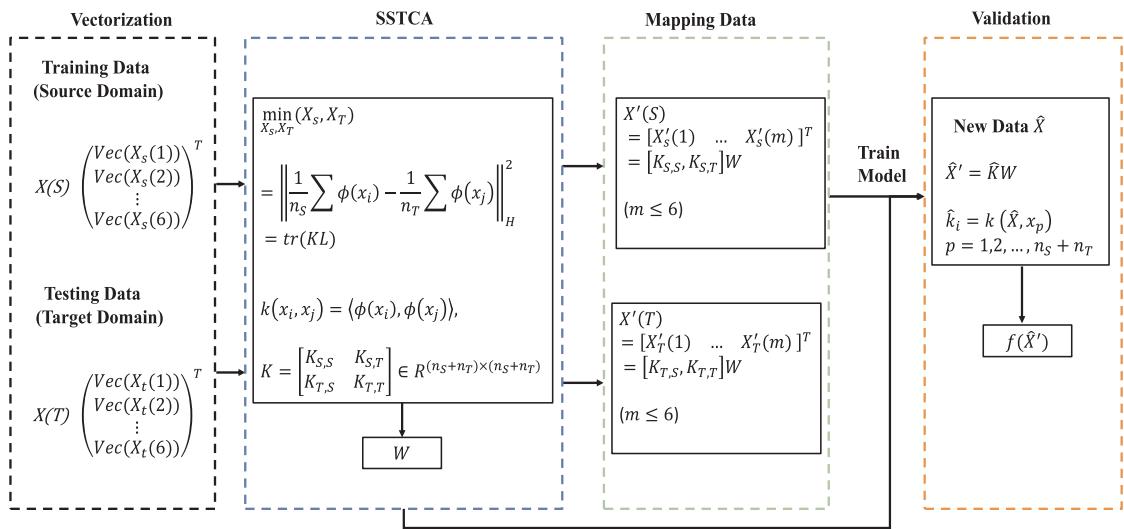


Fig. 2. SSTCA algorithm implementation.

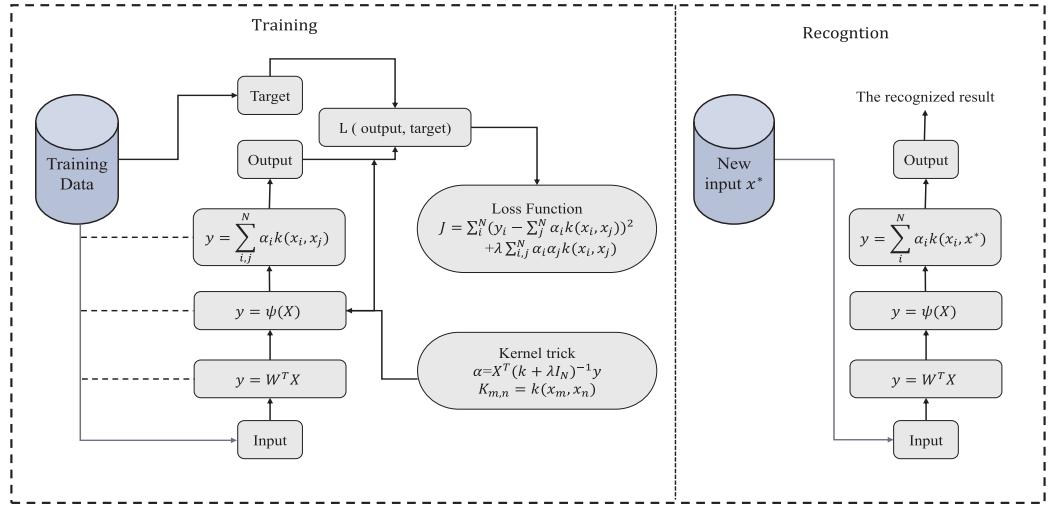


Fig. 3. KRR algorithm implementation.

**Table 2**  
Experimental conditions for batteries.

Battery Number	Discharge Current	Voltage Upper	Voltage lower	Normal Capacity
B0005	2 A	4.2 V	2.7 V	2 Ah
B0006	2 A	4.2 V	2.5 V	2 Ah
B0007	2 A	4.2 V	2.2 V	2 Ah
B0034	4 A	4.2 V	2.2 V	2 Ah

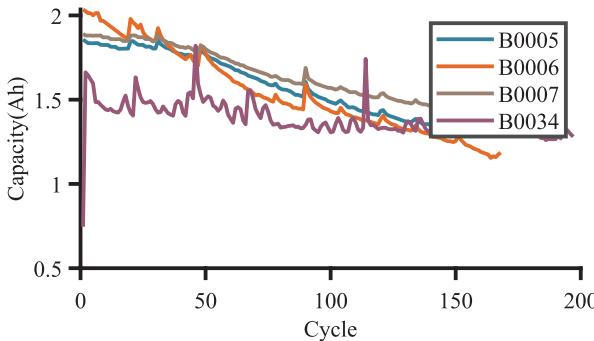


Fig. 4. Aging trends for four batteries.

voltage of the battery (The sample entropy analysis procedure was introduced by [41]); these figures demonstrate regular trends rather than irregular changes with increase in the number of cycles. According to these figures, the color of each curve varies in the pattern "red–green–blue" as the number of cycles increases.

### 3.2.1. Ratio of the CC mode

According to Fig. 5 (a), each battery has two charging modes under the charging conditions comprising the CC and CV modes. As the number of cycles increase, the time requires for the battery to reach the cut-off voltage when charged in the CC mode become shorter. If we let  $T_{cc}$  represent the CC charging time and  $T_{cv}$  denotes the CV charging time, then we can use (14) to calculate the ratio of the CC mode.

$$\text{Ratio} = \frac{T_{cc}}{T_{cc} + T_{cv}} \quad (14)$$

### 3.2.2. Equal voltage drop

As shown in Fig. 5 (b), as the number of cycles increase, the time spends in the selected equidistant voltage interval gradually decrease.

We can quantify battery degradations according to the time periods with equal voltage drops using (15).

$$EVD = |t_{2i} - t_{1i}| \quad (15)$$

### 3.2.3. Characteristics of IC curves

As described in a previous study [18], the feature information related to the SOH for the battery is extracted starting from the IC curve. Fig. 5 (c) shows that the IC curve has a significant peak point. We analyze the changes in the peak point coordinate values and the areas near the peak point. Thus, these three variables comprise the health features used to assess the SOH for the battery. The specific calculations are shown in (16).

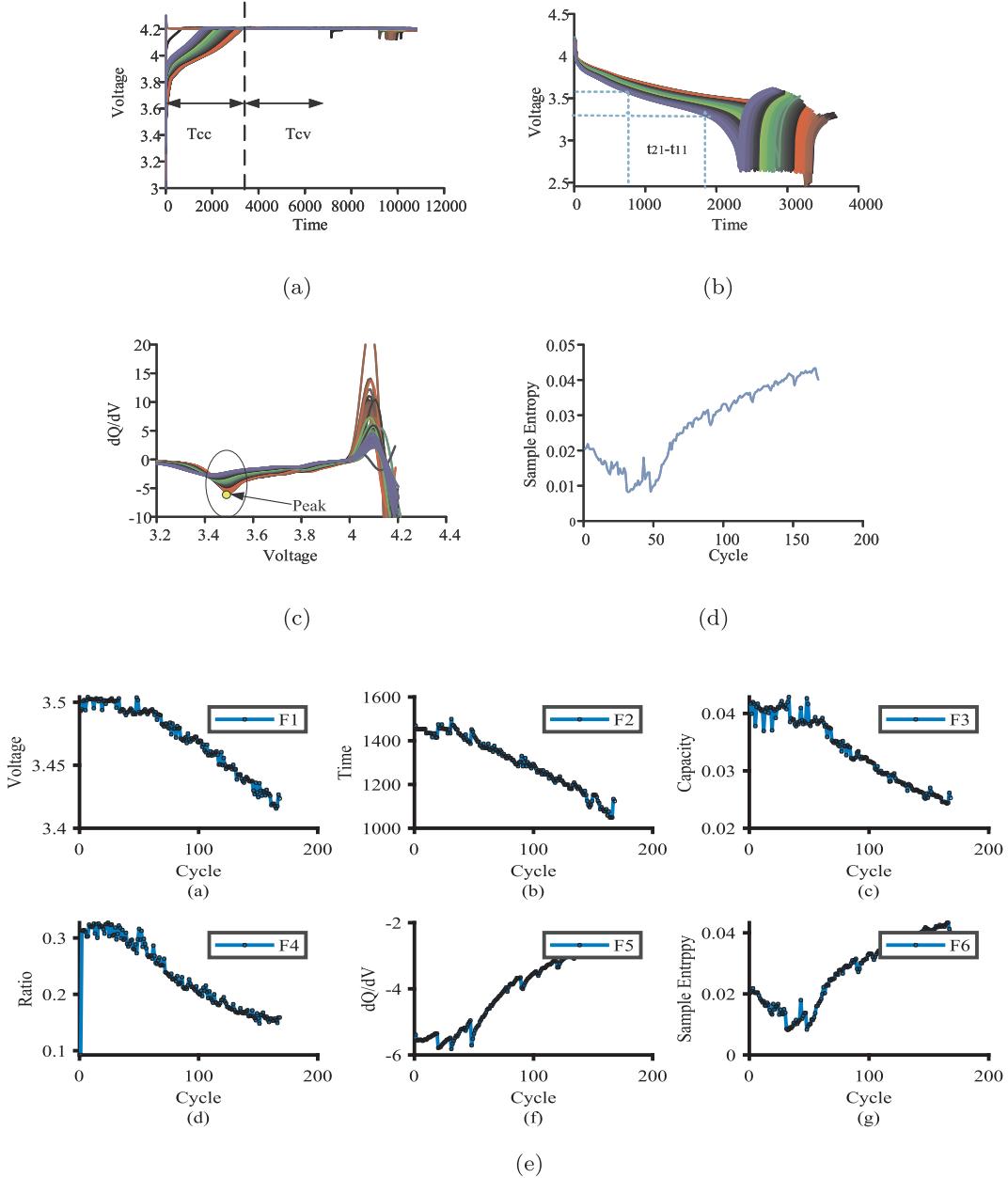
$$\begin{cases} \frac{dQ}{dV} = \frac{\Delta Q}{\Delta V} = \frac{Q_t - Q_{t-1}}{V_t - V_{t-1}}, \\ Q = \int I_c dt, \end{cases} \quad (16)$$

We use the information of the peak point in the IC curve as the features. In this paper, we use the abscissa, ordinate of the peak point of the IC curves, and the area near the peak point. Regarding how determine the area under the peak point, first we find the peak point  $x_p$ , and then use the peak point to take the small neighbourhood  $\delta$ , that is  $x \in (x_p - \delta, x_p + \delta)$ , finally we integrate the IC curve in this interval  $(x_p - \delta, x_p + \delta)$ , and the integrated value is the area near the peak point.

### 3.2.4. Sample entropy of discharge voltage

The sample entropy is a nonlinear dynamic parameter use to quantify the regularity and unpredictability of time series. The sample entropy  $SampEn(m, r, N_c)$  is defined as a nonnegative number representing the complexity of a time series, which reflects the occurrence of new information in the time series. The sample entropy is larger when the time series is more complex. Thus, SampEn is an estimate of the conditional probability that windows of length  $m$  (subseries of a time series of length  $N_c$ ) remain similar within a tolerance  $r$  and they also match the next point. Fig. 5 (d) shows the clear trend in the sample entropy of the discharge voltage, which we also use as a health feature to assess the SOH for the battery. The sample entropy analysis procedure is introduced by [41].

In this paper, six features are selected as the model inputs, i.e., F1 is the abscissa of the peak point of the IC curves, F2 is the equal voltage drop, F3 is the area near the peak point, F4 is the ratio of the CC mode, F5 is the ordinate of the peak point of the IC curves, and F6 is the sample entropy of the discharge voltage. Fig. 5 (e) (1–6) show that each health feature exhibits its own specific trend, and some of these trends



**Fig. 5.** Visualization of dataset B0005, (a) Charge curve for battery, (b) Discharge curve for battery, (c) IC curves for battery, (d) Sample entropy for battery, (e) Health features of battery.

are similar.

### 3.3. Mutual information analysis

The degree of association is the measurement of the magnitude of the association between two systems that vary over time or from different objects. Mutual information represents the volume of information shared between two or more variables. The mutual information value is greater when the correlation between the variables is stronger. Therefore, mutual information analysis can provide a quantitative measurement of the system's evolution, which is suitable for dynamic process analysis. The formal definition of mutual information for two random variables  $X$  and  $Y$  is given by

$$I(X, Y) = \int \int_S f_{x,y}(x, y) \log \frac{f_{x,y}(x, y)}{f_x(x)f_y(y)} dx dy \quad (17)$$

where  $S$  is  $X, Y \subseteq S$ ,  $f_{x,y}(x, y)$  is the joint distribution of  $X$  and  $Y$ , and  $f_x(x)$  and  $f_y(y)$  are the marginal distributions of  $X$  and  $Y$ , respectively.

### 4. Estimation results

In this section, we present the estimated SOH results of the battery. First, we select the data sets to obtain training and testing sets. Second, starting from the data characteristic, the data distribution difference is analyzed. Third, we use the SSTCA algorithm to process the original health features to achieve the aims of minimizing the distance between the marginal distributions, maximizing the source domain-label dependence on the latent space, and preserving the local geometric structures in the latent space. We also compare the health feature degrees obtained using the SSTCA method and the original health feature degrees based on the mutual information. Finally, the estimated results are obtained through kernel ridge regression (KRR), in general, we also use the Gaussian process regression (GPR) model training results under

the same conditions to compare with the KRR results, where they are measured based on the maximum average error (MAE), mean squared error (MSE), and root mean squared error (RMSE). In addition, we use the coefficient of determination  $R^2$  to measure the performance of model fitting.

Battery data sets B0005, B0006, B0007, and B0034 are utilized to determine whether the method is effective and accurate for predicting the SOH. We select label number cycles 20–60 as the source domain (training data), which is represented by  $D_s$ , and a large volume of unlabeled data from cycle 61 to the end is considered as the target domain (test data) represented as  $D_t$ . Furthermore, we randomly split  $D_t$  into  $D_t^u$  and  $D_t^o$ . All of the source domain data ( $D_s$ ) are used for training. The proposed methods are aimed at learning a model from  $D_s$  and  $D_t^u$ , before evaluating the model based on  $D_t^o$ . The data training framework is shown in Fig. 6.

Based on the methods described in Section 3, the health features are calculated to establish links with the battery SOH. The original health features are then processed using the SSTCA algorithm, which selects latent health features with lower dimensions to represent the original health features while preserving the properties as much as possible. By considering the three-dimensional data from the original health features as an example, Fig. 7 (a), (c), (e), and (g) show that the training and testing domain data have relatively different distributions. We process the data by using the SSTCA algorithm, and the three-dimensional data are used to represent the six dimensions of the data of the original health features. Fig. 7 (b), (d), (f), and (h) show that the distributions are significantly reduced. According to mutual information analysis, the results show that the latent health features produced by the SSTCA algorithm have higher degrees of relevance, as shown in Tables 8–11. Tables 12–15 show that the method based on knowledge transfer obtained more accurate predictions, as demonstrated in Fig. 12. According to these figures, the SSTCA results are better at tracing the battery SOH and obtain more accurate predictions. Thus, the estimated results based on the SSTCA method are more accurate and stable. The MAS, MSE, RMSE and  $R^2$  were calculated as follows.

$$MAE = \frac{1}{N} \sum_{i=1}^N \left| SOH_{i,estimate} - SOH_{i,real} \right| \quad (18)$$

$$MSE = \frac{1}{N} \sum_{i=1}^N (SOH_{i,estimate} - SOH_{i,real})^2 \quad (19)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (SOH_{i,estimate} - SOH_{i,real})^2} \quad (20)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (SOH_{i,estimate} - SOH_{i,real})^2}{\sum_{i=1}^N (SOH_{i,estimate} - SOH_{mean})^2} \quad (21)$$

#### 4.1. Distribution difference analysis

From the analysis of the original data, we could analyze the distribution differences between the selected training data and test data. By using the expectation maximization algorithm and Gaussian mixture model, the overlapping area of the two data sets is calculated to quantify the difference in the distribution. Machine learning assumes that the training and test data are stable and consistent; however, in practical applications, the characteristics of battery degradation will change with the increase in the number of cycles, while the difference in the data distribution is increasing. In this study, we use the battery historical data to predict the future battery health status. By selecting the label number cycles 20–60 as the training data, and from cycle 61 to the end as the test data. Compared to the general method of selecting data sets, the training and test data are cross-sampled. Fig. 7 show the distribution diagrams of battery data, and Table 3 shows the values of

overlapping areas of data distribution. From these figures and table, it can be concluded that the overlapping area of the two data sets used in this study is significantly reduced, indicating that there is a difference in the distribution of these data. After extracting the features of the data set, the difference in the feature-distribution data is particularly obvious in the feature space; however, after SSTCA processing, the difference is definitely reduced. In this paper, D1 and D2 are used to represent the cross-sampling and data-sampling methods.

#### 4.2. Feature redundancy analysis

Based on the analysis in Section 4.1, there is a difference in the distribution of these raw data. Similar to PCA, SSTCA algorithm has the effect of reducing the dimension. We take the feature dimension from 1 to 6 successively as the model input. And we choose the appropriate input dimension to train the mode based on the error value in different dimensions. The estimated error results are shown in Tables 4–7. Then, we add Fig. 8–10 to visualize the content of Table 4–7. Further, we construct a 3D graph of the feature input and showed that if the data are not processed by SSTCA, the distribution difference is definitely obvious. After SSTCA processing, according to Fig. 6 and Table 3, it can be seen that the difference in feature distribution is obviously reduced. In this paper, in order to clearly explain the prediction results, we use symbol type 1 and type 2 to represent the method of without using SSTCA and the method using SSTCA. Before the original feature data is processed through SSTCA, we use the mutual information to quantify the correlation between features and target. The correlation is arranged in the order of F1–F6. We use different dimension feature as the mode input to estimate the battery SOH, because of the SSTCA algorithm has the effect of reducing the dimension, so according to the error value in different dimensions, we choose the appropriate input dimension to train the model. Similar to PCA, SSTCA sorts the correlation of processed feature data while reducing the distribution difference. We use MAE, MSE and RMSE to measure the accuracy of the estimated result. Based on Table 4–7 and Figs. 8–10, we find out that the result, when the dimension is equal to three, is definitely stable and accurate under the four battery data sets. Furthermore, the estimated result based on KRR under type 2 is more accurate than the estimated result under type 1, although the training set only accounts for about 35% of the entire training set. This is the reason why "three-dimensional" features are selected for analysis in this study.

To train an accurate model through machine learning, the alignment of the distributions between domain data in the feature space should be maximized as much as possible or the training effect of the model will not be satisfactory. In the proposed method, the aim is to obtain knowledge from a small volume of historical battery data to predict the battery SOH. We determine that the training and test data

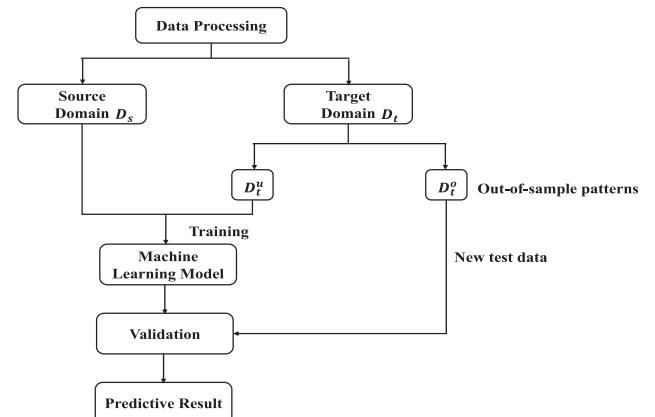
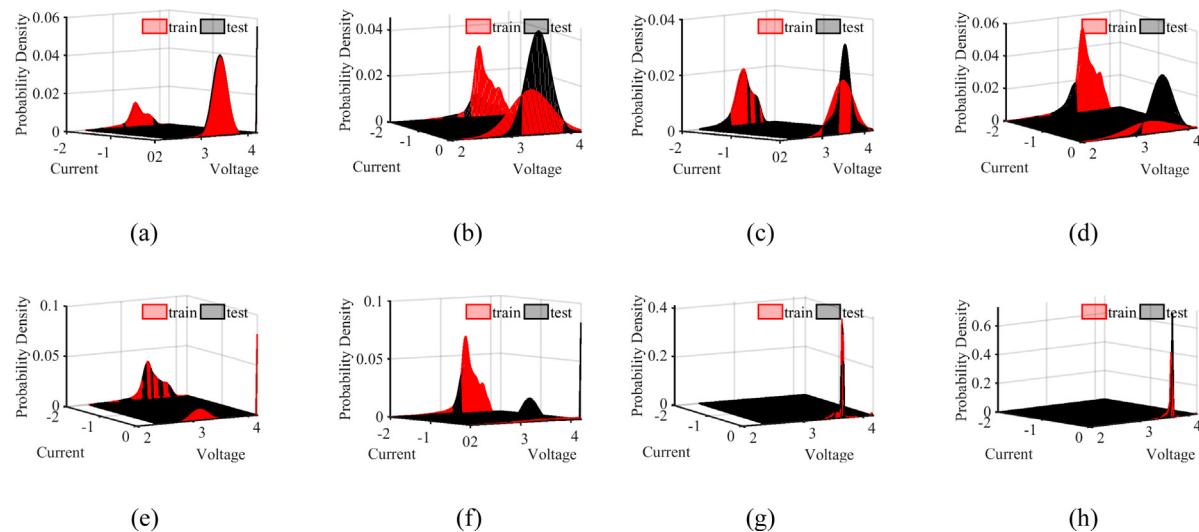


Fig. 6. Model training based on semi-supervised transfer component analysis algorithm.



**Fig. 7.** Two different data distribution for battery data set, (a) Battery B0005 distribution for D1, (b) Battery B0005 distribution for D2, (c) Battery B0006 distribution for D1, (d) Battery B0006 distribution for D2, (e) Battery B0007 distribution for D1, (f) Battery B0007 distribution for D2, (g) Battery B0034 distribution for D1, (h) Battery B0034 distribution for D2.

**Table 3**  
Overlapping area of datasets.

Overlapping area	B0005	B0006	B0007	B0034
D1	0.9818	0.8856	0.9787	0.9228
D2	0.6241	0.4795	0.6224	0.4728

**Table 4**  
Error results for different feature dimension and types of B0005.

Feature dimension	Type 1			Type 2		
	MAE	MSE	RMSE	MAE	MSE	RMSE
$n = 1$	>0.6	>0.6	>0.6	0.0643	0.0067	0.082
$n = 2$	0.3867	0.2584	0.5083	0.0131	0.0002	0.0159
$n = 3$	0.0456	0.0046	0.0678	0.0064	8.5e-05	0.0092
$n = 4$	0.068	0.0094	0.0969	0.006	7.0e-05	0.0083
$n = 5$	0.0763	0.0089	0.0945	0.0072	9.9e-05	0.01
$n = 6$	0.1165	0.0198	0.1406	0.0091	0.0001	0.0126

**Table 5**  
Error results for different feature dimension and types of B0006.

Feature dimension	Type 1			Type 2		
	MAE	MSE	RMSE	MAE	MSE	RMSE
$n = 1$	0.1591	0.0334	0.1827	0.1195	0.0254	0.1594
$n = 2$	0.167	0.0423	0.2058	0.0950	0.012	0.1097
$n = 3$	0.0376	0.0023	0.0484	0.0129	0.0005	0.0221
$n = 4$	0.0319	0.0018	0.0426	0.0248	0.001	0.0319
$n = 5$	0.0401	0.0023	0.0535	0.0114	0.0002	0.0169
$n = 6$	0.0286	0.0015	0.0393	0.0167	0.0004	0.022

distributions are different in the original feature space, as shown in Fig. 11 (a), (c), (e), and (g). Thus, we use the SSTCA algorithm to process the original data to address this issue. Furthermore, in order to avoid redundant feature information, and based on the above analysis content of choosing the appropriate input dimension, we employ a lower dimensional vector to represent the original health information. The obtained results are presented in Fig. 11 (b), (d), (f), and (h). In addition, we employ mutual information to determine the SSTCA features with much higher degrees of relevance to the battery SOH compared with that of the original features data. Tables 8–11 show that

**Table 6**  
Error results for different feature dimension and types of B0007.

Feature dimension	Type 1			Type 2		
	MAE	MSE	RMSE	MAE	MSE	RMSE
$n = 1$	>0.6	>0.6	>0.6	0.0167	0.0005	0.0224
$n = 2$	>0.6	>0.6	>0.6	0.0155	0.0004	0.0215
$n = 3$	0.09	0.01	0.1042	0.0039	3.1e-5	0.0056
$n = 4$	0.061	0.0052	0.0726	0.0107	0.0002	0.0158
$n = 5$	0.0156	0.0003	0.02	0.0126	0.0002	0.0156
$n = 6$	0.0212	0.0006	0.0248	0.0055	7.2e-5	0.008

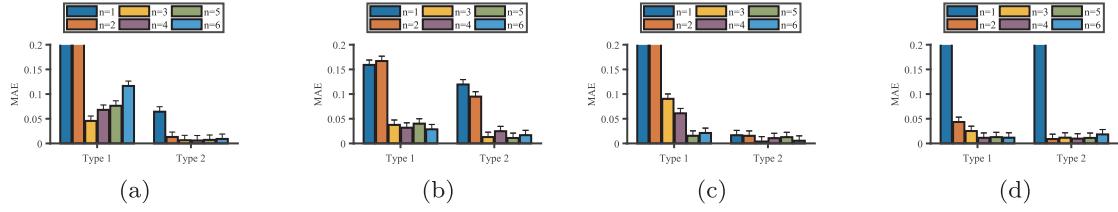
**Table 7**  
Error results for different feature dimension and types of B0034.

Feature dimension	Type 1			Type 2		
	MAE	MSE	RMSE	MAE	MSE	RMSE
$n = 1$	0.4211	0.2156	0.4643	0.3866	0.1872	0.4326
$n = 2$	0.0433	0.0023	0.0483	0.0089	0.0001	0.0112
$n = 3$	0.0251	0.001	0.0308	0.0117	0.0002	0.0144
$n = 4$	0.0114	0.0002	0.0142	0.0009	0.0001	0.0125
$n = 5$	0.0127	0.0002	0.0157	0.0112	0.0002	0.01413
$n = 6$	0.0117	0.0002	0.0143	0.0183	0.0004	0.0217

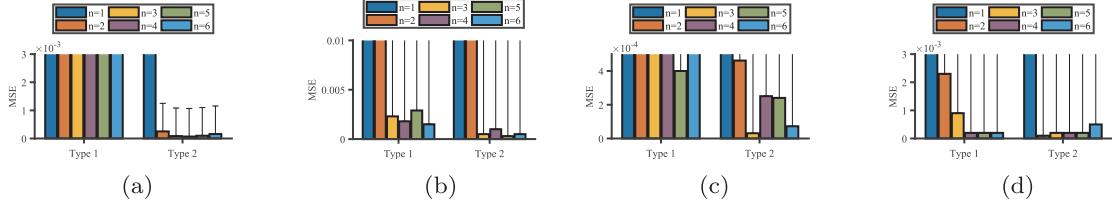
the features data of the SSTCA are strongly correlated with the battery SOH; however, their correlations with the original features data are not decreased.

#### 4.3. The SOH estimation

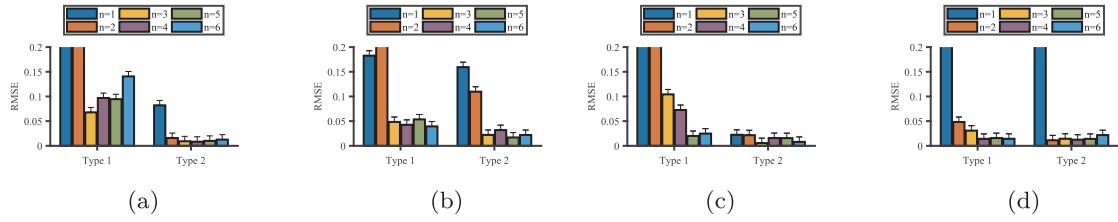
KRR is used to train the model based on the processed data in this paper. In order to analyze the estimated result clearly, we use the GPR model training results under the same conditions to compare with the KRR results. In this part, we give the estimated results based on GPR, and compare the estimated results of KRR and GPR under type 1 and type 2. As can be seen from the Fig. 12, the estimated result based on KRR is more accurate under type 2, which means that the feature input processed by the SSTCA algorithm, and the estimated result using KRR as the model is the closest to the true value. Tables 12–15 have given the specific error value under each condition of the four battery data sets. As shown in Figs. 12 (a), (c), (e), and (g), the SSTCA method based on



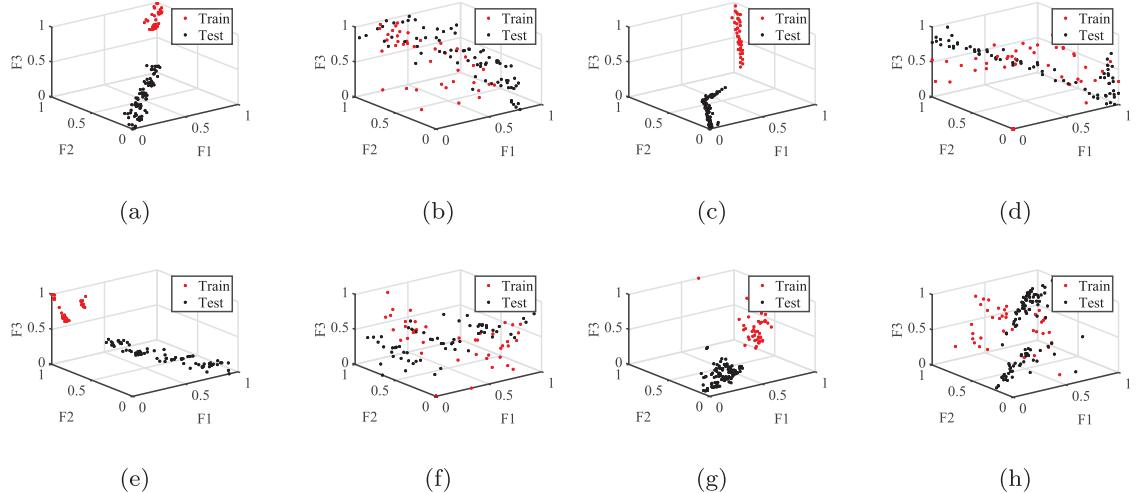
**Fig. 8.** MAE results for different feature dimension and types, (a) Battery B0005 data set, (b) Battery B0006 data set, (c) Battery B0007 data set, (d) Battery B0034 data set.



**Fig. 9.** MSE results for different feature dimension and types, (a) Battery B0005 data set, (b) Battery B0006 data set, (c) Battery B0007 data set, (d) Battery B0034 data set.



**Fig. 10.** RMSE results for different feature dimension and types, (a) Battery B0005 data set, (b) Battery B0006 data set, (c) Battery B0007 data set, (d) Battery B0034 data set.



**Fig. 11.** Two different feature space for battery dataset, (a) Battery B0005 in original feature space, (b) Battery B0005 in SSTCA feature space, (c) Battery B0006 in original feature space, (d) Battery B0006 in SSTCA feature space, (e) Battery B0007 in original feature space, (f) Battery B0007 in SSTCA feature space, (g) Battery B0034 in original feature space, (h) Battery B0034 in SSTCA feature space.

**Table 8**  
Degree values for dataset B0005.

Coefficient	F1	F2	F3
SSTCA Features	0.8711	0.9101	0.8735
Original Features	0.8619	0.7993	0.8904

KRR track the true values more effectively. However, the estimated results under type 1 show that the trained model could not learn all the knowledge from the test data because some of the test data could not be tracked, thereby resulting in poor prediction results. The MAE values for type 2 based on KRR are less than 2%, the MSE values for type 2 based on KRR are less than 0.05%, while the RMSE values for type 2 based on KRR are less than 2.5%. These three type error values are all smaller than the based on GPR under all four battery data sets. The absolute errors are shown in Figs. 12 (b), (d), (f), and (h), which

**Table 9**  
Degree values for dataset B0006.

Coefficient	F1	F2	F3
SSTCA Features	0.8381	0.9018	0.9049
Original Features	0.8942	0.8279	0.8956

**Table 10**  
Degree values for dataset B0007.

Coefficient	F1	F2	F3
SSTCA Features	0.8782	0.8711	0.8680
Original Features	0.8680	0.8612	0.8645

**Table 11**  
Degree values for dataset B0034.

Coefficient	F1	F2	F3
SSTCA Features	0.8738	0.8582	0.8753
Original Features	0.8567	0.8406	0.8004

demonstrate that the estimated errors based on KRR under type 2 are more stable than the errors under type 2. Figs. 12 and Tables 12–15 show the calculated MAS, MSE, and RMSE values to summarize the error results under different model and types of the four battery data sets.

Further, we use the coefficient of determination  $R^2$  to measure the performance of model fitting. Through Fig. 12 and Tables 12–15, we find that the estimated results obtained based on the KRR model are more accurate, so we use the coefficient of determination  $R^2$  to analyze this reason. The coefficient of determination  $R^2$  is a method to measure the fitting performance of the model. The better the model fitting performance, the more its value is close to 1, through Fig. 13, we find that the  $R^2$  of the KRR is higher than  $R^2$  of the GPR under type 2, which can also explain why the KRR model estimated results are better than GPR estimated results.

Thus, based on the above analysis, we use four aspects to prove our method can get high accuracy result.

(1) We take the feature dimension from 1 to 6 successively as the model input. Based on algorithm, we choose the appropriate input

**Table 12**  
Error results for different model and types of B0005.

Model	Type 1			Type 2		
	MAE	MSE	RMSE	MAE	MSE	RMSE
KRR	0.0456	0.0046	0.0678	0.0064	8.5e-05	0.0092
GPR	0.0486	0.0029	0.0538	0.0175	0.0005	0.0234

**Table 13**  
Error results for different model and types of B0006.

Model	Type 1			Type 2		
	MAE	MSE	RMSE	MAE	MSE	RMSE
KRR	0.0376	0.0023	0.0484	0.0129	0.0005	0.022
GPR	0.0356	0.0017	0.0411	0.0537	0.0044	0.0662

**Table 14**  
Error results for different model and types of B0007.

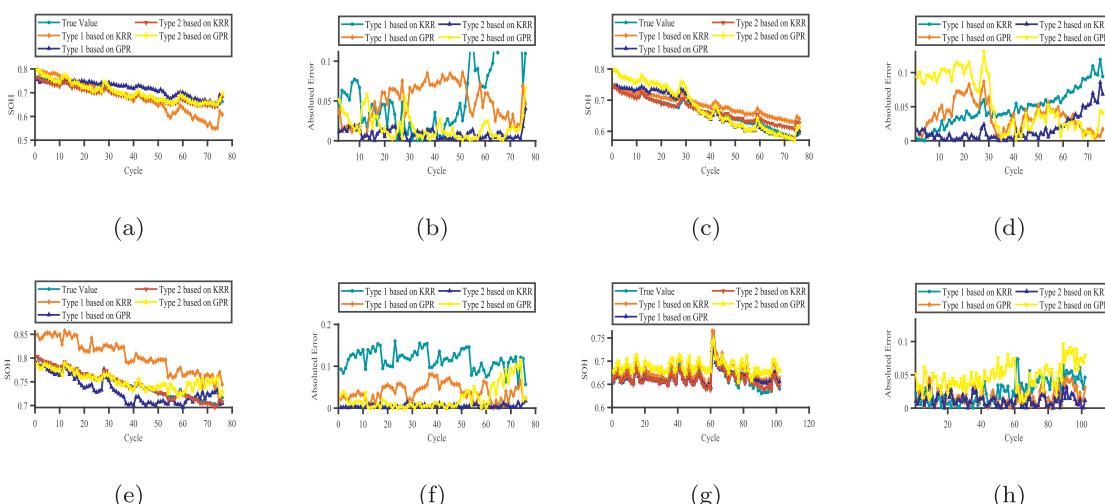
Model	Type 1			Type 2		
	MAE	MSE	RMSE	MAE	MSE	RMSE
KRR	0.09	0.01	0.1042	0.0039	3.1e-05	0.0056
GPR	0.0414	0.0025	0.0496	0.0237	0.0013	0.0367

**Table 15**  
Error results for different model and types of B0034.

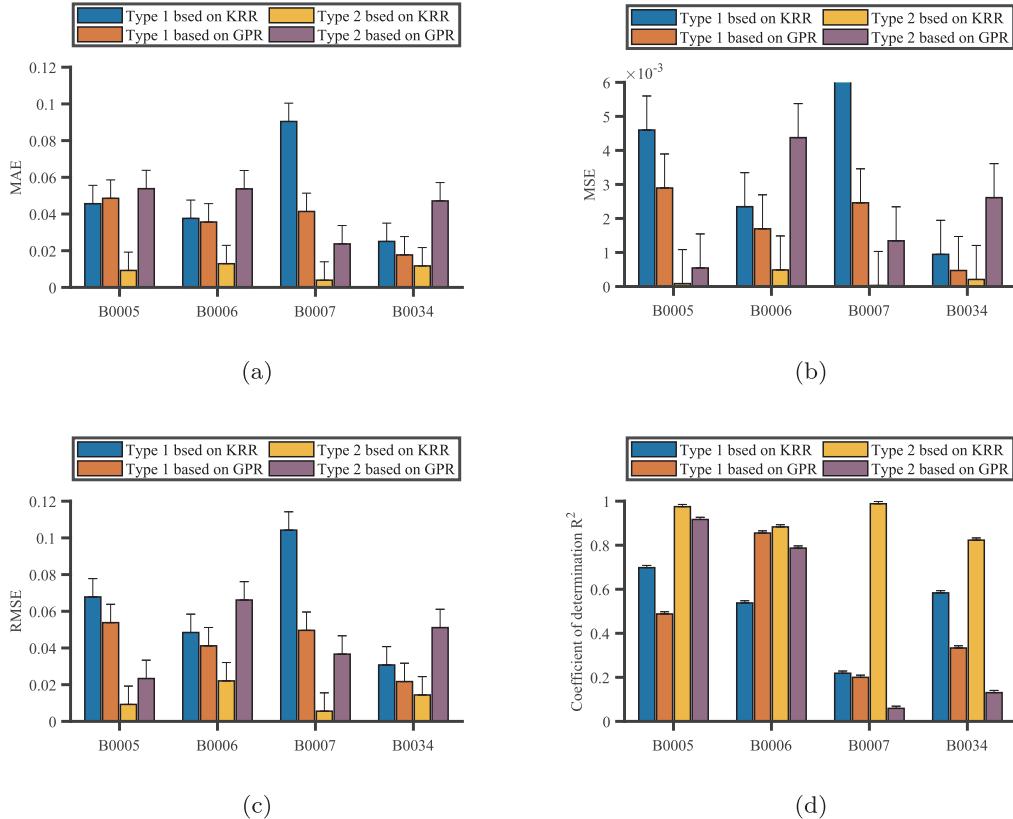
Model	Type 1			Type 2		
	MAE	MSE	RMSE	MAE	MSE	RMSE
KRR	0.0251	0.001	0.0308	0.0117	0.0002	0.0144
GPR	0.0177	0.0004	0.0217	0.0471	0.0026	0.0511

dimension to train the mode based on the three types error value in different dimensions. It can be seen that the higher dimension features do not get better performance.

(2) The comparison model that combines various errors to depict the performance of the method, we use the estimated results based on gaussian process regression (GPR) model under the same conditions



**Fig. 12.** Estimation and error results for battery data sets, (a) SOH estimation for battery B0005, (b) Estimated error for battery B0005, (c) SOH estimation for battery B0006, (d) Estimated error for battery B0006, (e) SOH estimation for battery B0007, (f) Estimated error for battery B0007, (g) SOH estimation for battery B0034, (h) Estimated error for battery B0034.



**Fig. 13.** Error results and coefficient of determination for different model and types, (a) MAE results, (b) MSE results (c) RMSE results (d) The coefficient of determination.

to compare with the kernel ridge regression (KRR) estimated results. The various errors results show that the estimated results based on SSTCA algorithm are better than these results without using SSTCA algorithm.

- (3) The performance of model fitting is measured by the coefficient of determination  $R^2$ . The better the model fitting performance, the more its value is close to 1. The values of the coefficient of determination  $R^2$  proves that the estimated results based on KRR are better than the estimated result based on GPR.
- (4) During transferring phase, we not only employ the first 35% of a data set to train the model, but also consider the redundant information between features. This means that using a small sample data is used to perform the battery SOH estimation, which significantly improves the practicability of the proposed model.

## 5. Conclusion and future work

In this study, starting from the distribution of data, we propose a battery health assessment model based on the semi-supervised transfer component analysis algorithm. To improve the accuracy when predicting the state-of-health for lithium-ion batteries, the proposed method calculates specific parameters from the charging and discharging curves in each cycle as health features. Then, the maximal alignments of the distributions of the training and testing domain data are achieved in the latent feature space using the semi-supervised transfer component analysis algorithm. The health features are selected through a mutual information analysis to train the battery health-assessment model. Finally, a kernel ridge regression learner is employed to estimate the battery state-of-health. Based on three aspect, we find that under the same conditions, the accuracy of the prediction results in this paper is improved compared with other models; on the other hand, the fitting ability of the kernel ridge regression model is also higher than that of

other models. The NASA battery data sets are used to verify the performance of the proposed method. At present, only few studies have been conducted on regression under transfer learning.

Taking into account the random dynamics of the actual battery operating conditions, it is very meaningful to study the battery state-of-health under actual dynamic operating conditions. Our team has started to analyze the battery health status under dynamic operating conditions. We want to learn the knowledge of data in condition A to complete the prediction (estimated battery state-of-health) under condition B. As for the feature extraction part, in fact, if feature extraction can be achieved automatically, the limitation of feature acquisition under certain conditions is avoided. For example, the deep learning method has the function of automatic feature extraction. Based on this analysis, our future work is to explore and verify the battery state-of-health estimation under the combination of two methods of deep learning and transfer learning.

## CRediT authorship contribution statement

**Yuanyuan Li:** Validation, Data curation, Formal analysis, Writing - original draft, Writing - review & editing. **Hanmin Sheng:** Methodology, Conceptualization. **Yuhua Cheng:** Supervision. **Daniel-Ioan Stroe:** Conceptualization, Writing - review & editing. **Remus Teodorescu:** Supervision.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgement

This work was supported in part by the National Natural Science Foundation of China under Grant 61903066.

## References

- [1] Deng Y, Ying H, Jiaqiang E, Zhu H, Wei K, Chen J, et al. Feature parameter extraction and intelligent estimation of the state-of-health of lithium-ion batteries. *Energy* 2019;176:91–102.
- [2] Xiong R, Li L, Tian J. Towards a smarter battery management system: A critical review on battery state of health monitoring methods. *J Power Sources* 2018;405:18–29.
- [3] Yang D, Wang Y, Pan R, Chen RY, Chen Z. State-of-health estimation for the lithium-ion battery based on support vector regression. *Appl Energy* 2018;227:273–83.
- [4] Stroe DI, Swierczynski M, Kær SK, Teodorescu R. Degradation behavior of lithium batteries during calendar ageing—the case of the internal resistance increase. *IEEE Trans Ind Appl* 2018;54:517–25.
- [5] Liu G, Ouyang M, Lu L, Li J, Han X. Online estimation of lithium-ion battery remaining discharge capacity through differential voltage analysis. *J Power Sources* 2015;274:971–89.
- [6] Baumhöfer T, Brühl M, Rothgang S, Sauer DU. Production caused variation in capacity aging trend and correlation to initial cell performance. *J Power Sources* 2014;247:332–8.
- [7] Xu Z, Wang Y, Chang L, Chen Z. A novel approach of battery pack state of health estimation using artificial intelligence optimization algorithm. *J Power Sources* 2018;376:191–9.
- [8] Cai L, Meng J, Stroe D-I, Luo G, Teodorescu R. An evolutionary framework for lithium-ion battery state of health estimation. *J Power Sources* 2019;412:615–22.
- [9] Sheng H, Xiao J, Wang P. Lithium iron phosphate battery electric vehicle state-of-charge estimation based on evolutionary gaussian mixture regression. *IEEE Trans Industr Electron* 2016;64(1):544–51.
- [10] Zhang Xu, Wang Y, Wu J, Chen Z. A novel method for lithium-ion battery state of energy and state of power estimation based on multi-time-scale filter. *Appl Energy* 2018;216:442–51.
- [11] Ashwin T, McGordon A, Jennings P. Electrochemical modelling of li-ion battery pack with constant voltage cycling. *J Power Sources* 2017;341:327–39.
- [12] Tian J, Xiong R, Yu Q. Fractional-order model-based incremental capacity analysis for degradation state recognition of lithium-ion batteries. *IEEE Trans Industr Electron* 2019;66(2):1576–84.
- [13] Ma Z, Yang R, Wang Z. A novel data-model fusion state-of-health estimation approach for lithium-ion batteries. *Appl Energy* 2019;237:836–47.
- [14] Ning B, Cao B, Wang B, Zou Z. Adaptive sliding mode observers for lithium-ion battery state estimation based on parameters identified online. *Energy* 2018;153:732–42.
- [15] Xiong R, Tian J, Mu H, Wang C. A systematic model-based degradation behavior recognition and health monitoring method for lithium-ion batteries. *Appl Energy* 2017;207:372–83.
- [16] Bai G, Wang P, Hu C, Pecht M. A generic model-free approach for lithium-ion battery health management. *Appl Energy* 2014;135:247–60.
- [17] Nuhic Adnan, Terzimehic Tarik, Soczka-Guth Thomas, Soczka-Guth Thomas, Buchholz Michael, Dietmayer Klaus. Health diagnosis and remaining useful life prognostics of lithium-ion batteries using data-driven methods. *J Power Sources* 2013;239:680–8.
- [18] Yi L, Mohamed A-M, Rahul G, Maitane B, Elise N-M, Noshin O, et al. A quick on-line state of health estimation method for li-ion battery with incremental capacity curves processed by gaussian filter. *J Power Sources* 2018;373:40–53.
- [19] He Z, Gao M, Ma G, Liu Y, Chen S. Online state-of-health estimation of lithium-ion batteries using dynamic bayesian networks. *J Power Sources* 2014;267(3):576–83.
- [20] You G won, Park S, Oh D. Real-time state-of-health estimation for electric vehicle batteries: A data-driven approach. *Appl Energy* 2016;176:92–103.
- [21] Meng J, Cai L, Stroe D-I, Luo G, Sui X, Teodorescu R. Lithium-ion battery state-of-health estimation in a vehicle using optimized partial charging voltage profiles. *Energy* 2019;185:1054–62.
- [22] Sui X, Stroe D-I, He S, Huang X, Meng J, Teodorescu R. The effect of voltage dataset selection on the accuracy of entropy-based capacity estimation methods for lithium-ion batteries. *Appl Sci* 2019;9(19):4170.
- [23] Honkura K, Takahashi K, Horiba T. Capacity-fading prediction of lithium-ion batteries based on discharge curves analysis. *J Power Sources* 2011;196(23):10141–7.
- [24] Chen L, Lü Z, Lin W, Li J, Pan H. A new state-of-health estimation method for lithium-ion batteries through the intrinsic relationship between ohmic internal resistance and capacity. *Measurement* 2018;116:586–95.
- [25] Zheng L, Zhu J, Lu DC, Wang G, He T. Incremental capacity analysis and differential voltage analysis based state of charge and capacity estimation for lithium-ion batteries. *Energy* 2018;150:759–69.
- [26] Stroe DI, Schatzl E. Lithium-ion battery state-of-health estimation using the incremental capacity analysis technique, *IEEE Trans Ind Appl*. doi: 10.1109/TIA.2019.2955396.
- [27] Guo P, Cheng Z, Yang L. A data-driven remaining capacity estimation approach for lithium-ion batteries based on charging health feature extraction. *J Power Sources* 2019;412:442–50.
- [28] Hu X, Li SE, Jia Z, Bo E. Enhanced sample entropy-based health management of li-ion battery for electrified vehicles. *Energy* 2014;64(1):953–60.
- [29] Liu D, Zhou J, Liao H, Peng Y, Peng X. A health indicator extraction and optimization framework for lithium-ion battery degradation modeling and prognostics. *IEEE Trans Syst Man Cybernet Syst* 2015;45(6):915–28.
- [30] Guha A, Patra A. State of health estimation of lithium-ion batteries using capacity fade and internal resistance growth models. *IEEE Trans Transp Electrific* 2018;4(1):135–46.
- [31] Yang D, Zhang X, Pan R, Wang Y, Chen Z. A novel gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve. *J Power Sources* 2018;384:387–95.
- [32] Wang M, Wang X. Automatic adaptation of a generic pedestrian detector to a specific traffic scene. 2011 Computer vision and pattern recognition. 2011. p. 3401–8.
- [33] Aytar Y, Zisserman A. Tabula rasa: Model transfer for object category detection. 2011 International conference on computer vision. 2011. p. 2252–9.
- [34] Pan SJ, Tsang IW, Kwok JT, Yang Q. Domain adaptation via transfer component analysis. *IEEE Trans Neural Netw* 2011;22(2):199–210.
- [35] Long M, Wang J, Ding G, Sun J, Yu PS. Transfer feature learning with joint distribution adaptation. 2013 IEEE international conference on computer vision. 2013. p. 2200–7.
- [36] Scholkopf B, Platt J, Hofmann T. A kernel method for the two-sample-problem. 2007 Conference on advances in neural information processing systems. 2007. p. 513–20.
- [37] Xu Y, Pan SJ, Xiong H, Wu Q, Luo R, Min H, et al. A unified framework for metric transfer learning. *IEEE Trans Knowl Data Eng* 2017;29(6):1158–71.
- [38] Long M, Wang J. Learning transferable features with deep adaptation networks. In: Proceedings of the 32nd international conference on machine learning, vol. 37; 2015. p. 97–105.
- [39] Smola A, Gretton A, Song L, Scholkopf B. A hilbert space embedding for distributions. 2007 international conference on algorithmic learning theory. 2007. p. 13–31.
- [40] Smola A. Nonlinear component analysis as a kernel eigenvalue problem. MIT Press; 1998.
- [41] Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circul Physiol* 2000;278(6):H2039–49.