A Uniform Estimation Framework for State of Health of Lithium-ion Batteries Considering Feature Extraction and Parameters Optimization

Xing Shu[1], Guang Li[2], Jiangwei Shen[1], Zhenzhen Lei[3], Zheng Chen[1, 2*] and Yonggang Liu[4*]
[1]Faculty of Transportation Engineering, Kunming University of Science and Technology, Kunming, 650500, China
[2]School of Engineering and Materials Science, Queen Mary University of London, London, E1 4NS, United Kingdom
[3]School of Mechanical and Power Engineering, Chongqing University of Science & Technology, Chongqing, 401331, China
[4]State Key Laboratory of Mechanical Transmissions & School of Automotive Engineering, Chongqing University, Chongqing, 400044, China
Email: shuxing92@kust.edu.cn, g.li@qmul.ac.uk, shenjiangwei6@163.com, 2010048@cqust.edu.cn, chen@kust.edu.cn, andylyg@umich.edu
Corresponding Author: Zheng Chen (chen@kust.edu.cn) and Yonggang Liu (andylyg@umich.edu)

Abstract: State of health is one of the most critical parameters to characterize inner status of lithium-ion batteries in electric vehicles. In this study, a uniform estimation framework is proposed to simultaneously achieve the estimation of state of health and optimize the healthy features therein, which are excavated based on the charging voltage curves within a fixed range. The fixed size least squares-support vector machine is employed to estimate the state of health with less computation intensity, and the genetic algorithm is applied to search the optimal charging voltage range and parameters of fixed size least squares-support vector machine. By this manner, the measured raw data during the charging process can be directly fed into the estimation model without any pretreatment. The estimation performance of proposed algorithm is validated in terms of different voltage ranges and sampling time, and also compared with other three traditional machine learning algorithms. The experimental results highlight that the presented estimation framework cannot only restrict the prediction error of state of health within 2%, but also feature high robustness and universality.

Key Words: state of health; lithium-ion batteries; fixed size least squares-support vector machine; genetic algorithm; feature extraction.

NOMENCLATURE

Abbreviations

EVs	electric vehicles	SVR	support vector regression
BMSs	battery management systems	LS-SVM	least square-SVM
SOH	state of health	BPNN	back propagation NN
SEI	solid electrolyte interface	CC	constant current
EIS	electrochemical impedance spectroscopy	IC	incremental capacity
EKF	extended Kalman filter	CV	constant voltage

AEKF	Adaptive EKF	GA	genetic algorithm
PF	particle filter	RBF	radial basis kernel function
HIF	H-infinity filter	DOD	depth of discharge
RLS	recursive least square	RMSE	root mean square error
SVM	support vector machine	PSO	particle swarm optimization
NN	neural networks	ME	maximum absolute error
ELM	extreme learning machine	MAE	mean absolute error
Symbols			
$\rho_{X,Y}$	Pearson correlation	α_i	Lagrange multiplier
X , Y	random variable matrices	$Z=h(x_i)^T h(x_j)$	kernel transformation matrix
$\mathrm{cov}(X,Y)$	covariance of X and Y	$K(x_i,x)$	kernel function
σ	standard deviation	γ	width of RBF
N	length of data	H_R	maximal quadratic Renyi entropy
$h(x)$	maps from low-dimensional training set data to a high-dimensional space	$p(x)$	density distribution of x
$y \in \Re$	estimation values	M	size of working set
ω	coefficients of LS-SVM	x^*	support vector
b	kernel parameter	x^{t*}	a random point
D	training set	SOH_{Real}	reference SOH
e_i	error variance	SOH_{Model}	output SOH of mode
c	regularization parameter	n	cycle number
l	sample number		

I. INTRODUCTION

Currently, lithium-ion batteries are main power sources of electric vehicles (EVs) [1]. To ensure safe reliable and efficient operation of lithium-ion batteries, battery management systems (BMSs) are usually deployed to monitor battery status and conduct necessary management including charging control [2], fault diagnosis and prognostic as well as thermal management [3]. An important function of BMS is to accurately estimate state of health (SOH) of batteries [4], which is vital to indicate their degradation status [5].

Degradation of lithium-ion batteries indispensably occur with operation due to formation and development of anode solid electrolyte interface (SEI), deposition of anode metal lithium, mechanical crushing of electrode active materials and other side reactions [6]. On the other hand, it can be aggravated by abused electrical and mechanical operations as well as improper surrounding thermal management [7]. Generally, battery aging often refers to decreased capacity and increased resistance, which are directly employed to characterize deterioration of electrical performance of batteries [8]. Due to intricate electrochemical reactions and various application conditions, it remains a challenging task to estimate SOH accurately [9]. To now, a

variety of estimation approaches have been put forward, and they can be simply classified into three categories: direct calibration methods, model-based estimation methods and machine learning methods.

Since battery aging can be reflected in decrease of capacity or increase of resistance, the most simple and direct manner of evaluating SOH is the accumulating current integration until full charge or discharge [10]. However, in practice, this method is susceptible to sampling accuracy of current, and rare occurrence of full charge/discharge leads to infeasibility of application. Measurement of electrochemical impedance spectroscopy (EIS) is another direct manner, as the impedance spectrum can indicate the aging variation trend. Nevertheless, EIS measurement requires specific pattern and excitation of current, and usually high frequency is indispensable [11], thus hindering its online application. Model-based estimation methods can predict SOH by means of building transfer functions with the inputs of measured voltage and current. A variety of model based estimation algorithms have emerged, including extended Kalman filter (EKF) [12] and its adaptive format (AEKF), particle filter (PF) [13], H-infinity filter (HIF), sliding model [14] and recursive least square (RLS) [15]. In [16], the Thevenin model is firstly established to characterize the dynamic and static behaviors of batteries, and then the battery capacity is considered as a model parameter and identified by the HIF. Compared with physical models, the electrochemical model that is composed of a series of partial differential equations can more accurately describe the electrical characteristics of the battery. In [17], a conventional pseudo-two-dimensional model is simplified by the finite analysis method, and the parameters in the whole lifespan are identified by the genetic algorithm (GA). Then, the quantitative relationship between the battery degradation and five inner parameters is constructed to help prediction of the battery capacity. These algorithms have already been intensively exploited to estimate SOH online. However, robustness of these algorithms cannot be guaranteed all the time, and credibility and reliability of battery models require in-depth understanding of degradation mechanism. To improve estimation accuracy, the model needs to contain more inner electrochemical information of batteries; whereas, the calculation intensity will no doubt increase, and the real-time fast application potential will be discounted [5].

With the development of computer science, machine learning methods, as alternative solutions to SOH estimation, attract substantial attention. Support vector machine (SVM) [18], Gaussian process regression [19], neural networks (NN) [20] and extreme learning machine (ELM) [21] have declared to estimate SOH by mapping external features/measurements into capacity loss [22]. Among these methods, NN algorithms feature

strong nonlinear fitting capability [23]; however, they require a large amount of sample data to excavate enough hidden information between SOH and electrical measures. As operation of lithium-ion batteries are stochastic, and it is a challenging task to consider all the operation aspects when selecting the training data. To lessen the dependence on data amount and mitigate the calculation intensity, SVM is introduced to attain SOH estimation [24]. Different from other machine learning methods with huge data requirement, SVM is more appropriate for small sample nonlinear problems and is less insensitive to dimensionality and variability of test data [25]. In [26], a support vector regression (SVR) based model is constructed to characterize the battery aging mechanism, and the PF is introduced to estimate the impedance degradation parameters. In [27], a multistage SVM algorithm is proposed to conduct SOH prediction. First, the SVM is leveraged to categorize object batteries into four types according to the cycle times. Then, the SVR is proposed to estimate the remaining useful life. The conventional SVM translates the controlling issue into a convex quadratic programming problem, and finally attains the optimal solution. Nonetheless, when the amount of data is large, intensive computation entails for finding solutions. To cope with it, the least square-SVM (LS-SVM) is introduced to transform the quadratic programming problem into a linear programming problem. Consequently, its solution speed is faster, and the computing labor becomes less, compared to the traditional SVM. On this account, the LS-SVM begins to be employed for SOH estimation. In [28], ten features after screening are selected as the inputs of LS-SVM, and the estimation results of SOH outperforms that of the conventional SVM and the back propagation NN (BPNN). In view of preferable performance of SOH estimation and less computation burden, the LS-SVM is continually employed and further investigated for the SOH estimation in this study.

To apply machine learning algorithms to estimate SOH, a critical task is to find proper feature variables that can capture the variation of SOH with acceptable precision. In addition, a high-quality feature variable should be easy to be extracted and contain abundant degradation information. To now, a number of feature variables have been employed for SOH prediction. In [29], the SOH is mapped by the time interval variation between the same discharge voltage intervals at each cycle. Since the discharging current profiles in EVs are usually stochastic, constant current (CC) discharging operations in a relative long duration rarely happen when driving. Similarly, it is difficult, and even impossible, to gauge the discharging time interval or the discharging voltage variation in a specially pre-determined range. Instead, the charging current is usually stably regulated

by chargers and the CC condition can be easily met. Consequently, CC charging voltage increase in a certain interval and/or capacity variation in between are often referred to highlight capacity variation. In [30], the second-order differential voltage signal is extracted to predict the SOH during the CC charging phase. In [25], partial incremental capacity (IC) curves are extracted as a health index, and the Gaussian process regression is employed to estimate the SOH. However, the testing sampling frequency may give rise to passive impact on accuracy of IC curves. In other words, the confidence interval of IC curves will reduce when the sampling frequency is low, thus undoubtedly deteriorating estimation precision of SOH. Furthermore, it is time consuming thanks to data processing and abstraction of health index [31]. In [3], four features including duration of the CC charging mode, duration of the constant voltage (CV) charging mode, the slope at the end of CC mode and the vertical slope at the corner of CC mode are regarded as the inputs of Gaussian process regression model. A common knowledge is that too many features will increase computational complexity in operation. In [32], two characteristic variables, i.e., the increase of ohmic internal resistance and polarized internal resistance, are regarded as the healthy indicators; and then ELM is applied to achieve the SOH estimation online. Although the discussed algorithms can estimate SOH with certain reliability and feasibility, it is still intractable to find effective features. For instance, to obtain IC, small current excitation and filtering operations are indispensable, leading to increase of estimation complexity; and improper selection of voltage ranges may arise deterioration of estimation accuracy [33]. In addition, it will be more practical to estimate SOH by partial charge voltage information, as full charge operations seldom happen in practice.

Motivated by this, the charging time for a fixed voltage range is quantitatively analyzed and selected as the feature variable in this study. The LS-SVM is further exploited to achieve the SOH estimation. Nevertheless, the support vectors in traditional LS-SVM are selected randomly, leading to lack of sparseness [34]. To compensate the defect, the quadratic Renyi criterion is developed to select the support vector, which forms a fixed size LS-SVM in this study. Meanwhile, the voltage range and crucial parameters of fixed size LS-SVM are simultaneously optimized in a unified framework solved by the GA, considering both estimation accuracy and practicability. The proposed approach contributes to existing manners of estimating SOH in the following three aspects: 1) The fixed size LS-SVM is applied to train offline parameters and furnish online SOH estimation with satisfactory precision, compared with the conventional SVM, LS-SVM and NN. 2) The charging time variation within a fixed voltage range in each cycle is extracted as the health indicator to quantify

the capacity degradation. Furthermore, GA is exerted to simultaneously find the optimal voltage range and the parameters of fixed size LS-SVM. 3) The proposed method is verified effective at different sampling steps and under the unacquainted test data set of the same type of battery cells, thereby proving its robustness and universality.

In the reminder of this study, Section II introduces the aging experiments and selects the feature variables. In Section III, the uniform framework for SOH prediction is elaborated and solved by the GA. The verification results are illustrated in Section IV. Some principal conclusions and future works are drawn in Section V.

II. Aging Experiment and Feature Extraction

A. Aging Experiment

The experiments are carried out on commercial 21700 NCM/graphite cells using a battery test system. The rated voltage and capacity are 3.6 V and 4 Ah, respectively; and the operating voltage ranges from 2.75 V to 4.2 V. A CC discharge test after the standard CC-CV charge action is scheduled to accomplish the cycle experiment. In the charging mode, the 0.5C current, wherein C denotes the rated capacity value of battery with the unit Ampere-hour, is applied to charge the battery until the terminal voltage reaches the pre-set voltage value. After that, the CV mode is applied to eliminate the polarization until the current declines to the cut-off current of 0.02C. After shelving for 5 minutes, the battery is discharged with the 1C current to its pre-set lower threshold. Note that all the experiments are conducted at room temperature (25 °C), which remains unchanged during the experiment.

B. Experiment Data Analysis and Feature Extraction

In actual applications, the discharging behavior of battery is a random and uncontrollable process; and instead, the CC-CV scheme is usually employed in battery's charging process. Given this, the electrical characteristics during the charging process is investigated to determine health features. The battery charging voltage evolutions at different SOH are depicted in Fig. 1. It can be seen the CC charging time is significantly shortened as the battery ages, this is because the polarization becomes more obvious with the degradation of battery capacity. The fresh battery takes 6270 s to reach its pre-set cut-off voltage, while the battery with 80% SOH merely requires about 4320 s, only 68.9% of that of fresh battery.

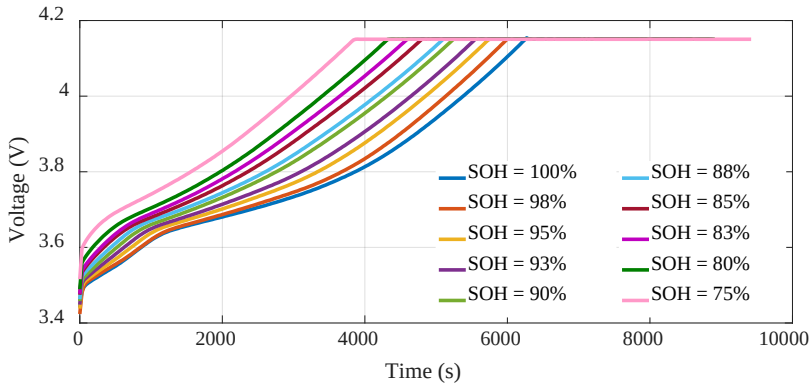


Fig. 1. The charging voltage evolution curves.

As described in [6], there is a strong correlation between battery life and depth of discharge (DOD). Ends users usually limit battery operations within a proper voltage range, and furthermore, to avoid range anxiety, batteries in EVs are seldom fully discharged. Thus, it is more reasonable to take advantage of partial charge and discharge data to estimate SOH. Apparently, a partial charging voltage profile is easier to acquire than the whole curve. Fig. 2 shows the charging time variation for a fixed voltage range variation when the battery ages. The blue and green lines respectively represent the charging time when the voltage increases from 3.4 V to 4.15 V and from 3.6 V to 4.1 V. It can be clearly found that the duration of selected voltage ranges shows a consistent agreement with the reduction of capacity.

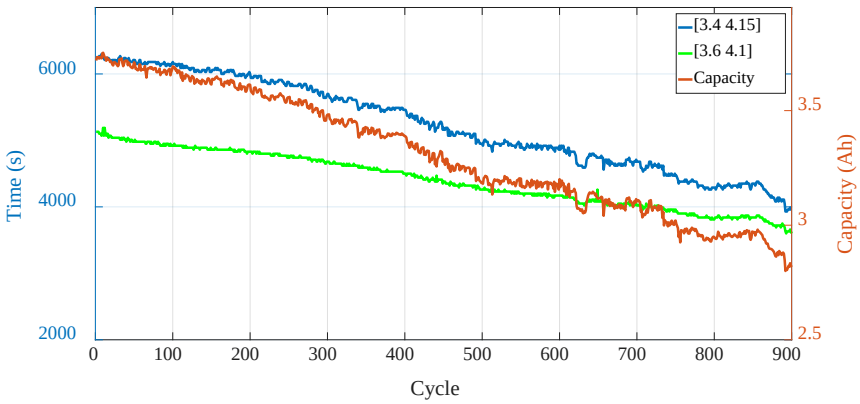


Fig. 2. Charging time for a fixed voltage range and capacity with cycle numbers.

To better assess the relational grade of selected partial charging duration and capacity variation, the statistic information is introduced. Although the covariance can characterize linear correlation of random variables, it cannot describe the degree of correlation. Here, the Pearson correlation [35] is employed, as:

$$\rho_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{\sum XY - \frac{1}{N}\sum X \sum Y}{\sqrt{\left(\sum X^2 - \frac{1}{N}(\sum X)^2\right)\left(\sum Y^2 - \frac{1}{N}(\sum Y)^2\right)}} \quad (1)$$

.

where X and Y represent the two random variable matrices, $\mathrm{cov}(X,Y)$ denotes the covariance of X and Y, σ is the standard deviation, and N denotes the length of data. The Pearson correlation grades between different voltage ranges and capacity of tested battery are listed in Table I. As can be found, the selected charging voltage ranges all show high relational grade, which means the feature extracted from the duration of selected voltage range is well-suited for establishing the degradation model and contributing to SOH estimation. In addition, the results clearly express that wider voltage range will bring higher correlation with the price of more intensive calculation. Therefore, proper selection of voltage range should be tailored to trade-off the correlation degree and calculation complexity, thereby ensuring the preferable aging characterization and acceptable calculation requirement. In the next step, GA is applied to find the optimal voltage range to ensure the estimation accuracy of SOH.

Table I. Pearson Correlation grades between different voltage ranges and capacity.

Voltage range (V)	Pearson Correlation	Voltage range (V)	Pearson Correlation
[3.4 3.5]	0.8368	[3.5 3.6]	0.9927
[3.4 3.7]	0.9983	[3.5 4.1]	0.9989
[3.4 3.9]	0.9985	[3.6 4.1]	0.9970
[3.4 4.1]	0.9994	[3.6 4.15]	0.9971

III. State of Health Estimation Algorithm Design

In this section, a brief derivation procedure of fixed size LS-SVM is provided, and then the GA is employed to search the optimal parameters and optimal voltage range.

A. Least Squares-Support Vector Machine

LS-SVM can be regarded as a special form of SVM with quadratic loss functions. In addition, LS-SVM uses equality constraints, rather than inequality constraints, and turns the solution process into a set of equations, thereby reducing the calculation amount [36]. For a nonlinear system,

$$y(x) = \omega^T h(x) + b \tag{2}$$

where $h(x)$ $\Re^n \to \Re^{n_h}$ maps low-dimensional training set data to a high-dimensional space, $y \in \Re$ denotes the estimation values, ω and b represent the coefficients. For a given training set $D = \{(x_i, y_i), i = 1,2,3,\ldots\}$, $x \in \Re^n$, $y \in \Re$, LS-SVM is employed to solve the loss function of optimization target, which is described by the quadratic term of error. The objection function can be formulated as:

$$\min J(\omega,\xi)=\frac{1}{2}\omega^T\omega+\frac{c}{2}\sum_{i=1}^{l}e_i \tag{3}$$

subject to:

$$y_i(x_i)=\omega^T h(x)+b+e_i,\qquad \forall i \tag{4}$$

where e_i is the error variance, c denotes a regularization parameter, and l represents the sample number. By introducing the Lagrange multiplier, the problem can be translated into an equivalent unconstraint problem, as:

$$L=\frac{1}{2}\omega^T\omega+\frac{c}{2}\sum_{i=1}^{l}e_i+\sum_{i=1}^{l}\alpha_i\omega^T h(x_i)+b-e_i \tag{5}$$

where α_i represents the Lagrange multiplier. The optimization conditions can be formulated, as:

$$\begin{cases}\dfrac{\partial L}{\partial \omega}=0\Rightarrow \omega=\sum_{i=1}^{l}\alpha_i h(x_i)\\ \dfrac{\partial L}{\partial e_i}=0\Rightarrow \alpha_i=ce_i\\ \dfrac{\partial L}{\partial \alpha_i}=0\Rightarrow y_i(x_i)=e_i-\omega^T h(x_i)-b=0\end{cases} \tag{6}$$

Thus, the dual problem can be formulated as:

$$\begin{bmatrix}0 & 1\\ 1 & Z+I/c\end{bmatrix}\begin{bmatrix}b\\ \alpha\end{bmatrix}=\begin{bmatrix}b\\ y\end{bmatrix} \tag{7}$$

where $Z=h(x_i)^T h(x_j)$ denotes the kernel transformation matrix. Now, the model can be reformulated as:

$$y(x_i)=\omega^T\alpha_i K(x_i,x)+b \tag{8}$$

where $K(x_i,x)$ represents the kernel function. Among all popular kernel functions, the radial basis kernel function (RBF) usually features the ability of mapping non-linear samples to high-dimensional space with less numerical calculation. Hence, the RBF is selected as the kernel function, as:

$$K(\boldsymbol{x},\boldsymbol{x}_i)=\exp\left(-\frac{\|\boldsymbol{x}-\boldsymbol{x}_i\|^2}{2\gamma^2}\right) \tag{9}$$

where γ denotes the width of RBF.

B. Fixed Size Least Squares-Support Vector Machine

To avoid lack of sparseness and robustness, the fixed size LS-SVM is employed to predict the SOH, where the support vectors are selected from the training set according to the quadratic Renyi criterion [34]. In

it, only a subset of the training data is utilized as support vectors. Moreover, the maximal quadratic Renyi entropy is exerted to optimize the selection of data in the working set [34], as:

$$H_R = -\lg \int p(x)^2 dx \tag{10}$$

where $p(x)$ represents the density distribution of x, and

$$\int p^2(x)dx \approx \int \hat{p}(x)dx = \frac{1}{M^2}\sum_{i=1}^{M}\sum_{j=1}^{M} K(x_i, x) \tag{11}$$

Furthermore, equation (10) can be rewritten as:

$$H_R \approx -\lg\left(\frac{1}{M^2}\sum_{i=1}^{M}\sum_{j=1}^{M} K(x_i, x)\right) \tag{12}$$

To apply the fixed size LS-SVM, a step-by-step process is listed as follows, and the implementation flowchart is depicted in Fig. 3.

1) Determine the training data set (x_i, y_i), $x \in \Re^n$, $y \in \Re^n$;

2) Select a working set with a fixed size of M, where $M << l$;

3) Randomly choose a support vector x^* from the working set;

4) Randomly select a point x^{t*} from the l training set, and replace x^{t*} by x^* in the working set;

5) Calculate the Renyi entropy. If the entropy increases, keep x^{t*} in the working set; otherwise, the point is abandoned and keep the support vector x^* in the original working set;

6) Terminate the process if the number of iterations is outstripped or the Renyi entropy is too small. Otherwise, go back to step 2).

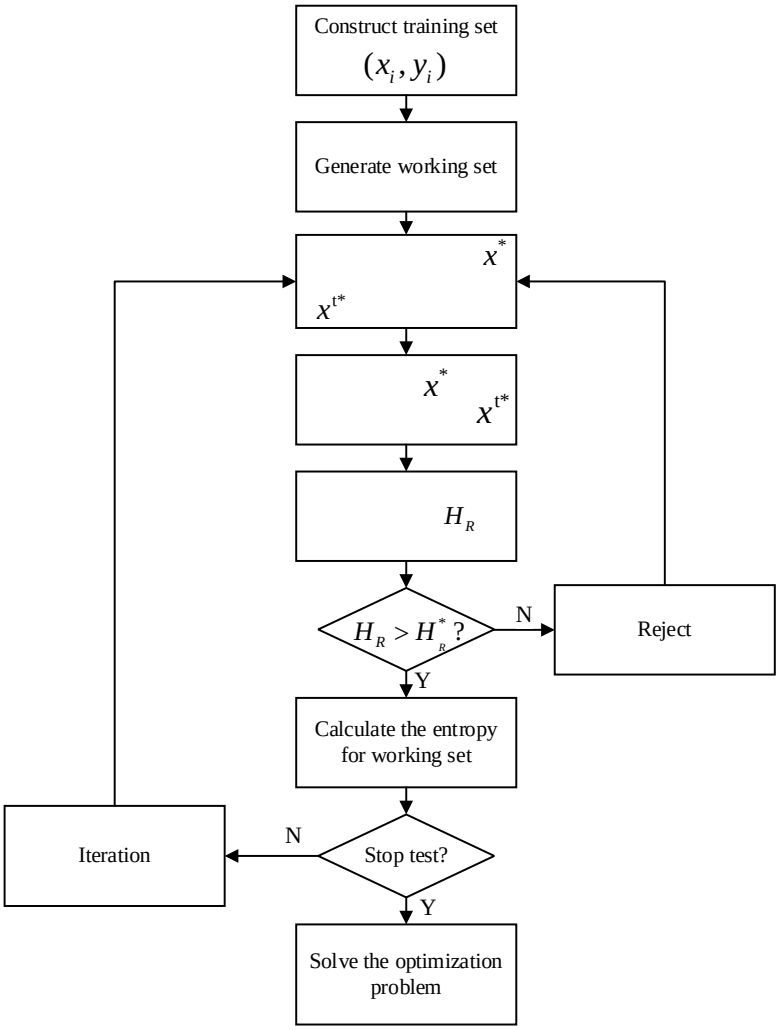


Fig. 3. The flowchart of fixed size LS-SVM algorithm.

C. The SOH Estimation Framework

When selecting the regression parameters, the parameters of fixed size LS-SVM should be appropriately determined to establish the SOH estimation model. The generalization ability is decided by the kernel parameter c, and the stability and complexity of this algorithm are influenced by the regularization parameter b [37]. By means of the arbitrary or empirical parameters selection, the performance of SOH estimation based on the fixed size LS-SVM cannot be ensured all the time. Here, GA is applied to simultaneously search the optimal voltage range, c and b. The overall framework of proposed SOH estimation method is illustrated in Fig. 4. As can be observed, an additional optimization procedure is added to find the optimal voltage range and parameters of fixed size LS-SVM. All the targets, including the SOH estimation, search of the optimal voltage range as well as determination of c and b, can be integrated into a unified optimization problem which can be solved simultaneously. After initializing the voltage range and parameters of the fixed size LS-

SVM, an initial SOH value will be set. Here, the root mean square error (RMSE) is regarded as the fitness criterion, as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(SOH_{Real}(i) - SOH_{Model}(i)\right)} \tag{13}$$

where SOH_{Real} is the reference SOH, SOH_{Model} is the output SOH of built model, and n denotes cycle number. According to the basic principle of GA, an initial population is generated randomly. After a series of selection, crossover and mutation evolution operations, the evolutionary mechanism will find the optimal (most of times sub-optimal) combination of feature variable and parameters [38]. The individuals that receive the best fitness assessment is utilized to estimate the SOH.

In the next step, the experimental validations and discussions are performed to manifest the feasibility of proposed algorithm.

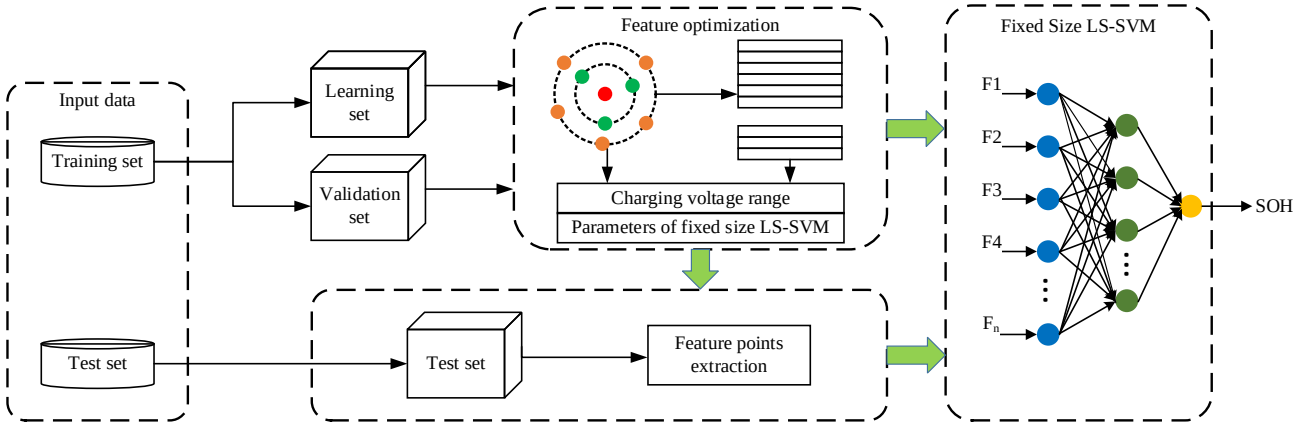


Fig. 4. The framework of proposed SOH estimation algorithm.

IV. Results and Discussion

In this section, a series of experiments were carried out to validate the effectiveness of proposed uniform estimation framework. Firstly, four groups of features and parameters extraction methods are compared to assess the impact of voltage range and parameters on the SOH estimation. After that, to evaluate the effectiveness of proposed fixed size LS-SVM algorithm, traditional algorithms including NN, SVM and conventional LS-SVM are employed to estimate the SOH based on the same healthy indicators. Afterwards, the results at different sampling time estimation results are addressed to demonstrate the bottom requirements of proposed algorithm on the sampling time. Finally, partial data is utilized for offline training and the

remaining unacquainted data is applied online to validate the model, thereby manifesting the universality capability of proposed algorithm.

A. Comparison with Random Voltage Range and the Optimized Voltage Range

To verify the effectiveness of developed feature and parameters optimization method, four groups of features and parameters are employed to predict the SOH.

1) The voltage range and the parameters of fixed size LS-SVM in the first group are designated randomly.

2) The voltage range in the second group is extracted randomly, whereas the parameters of fixed size LS-SVM are set the same as those optimized by the GA.

3) The voltage range and parameters in the third group are optimized by the particle swarm optimization (PSO) [39].

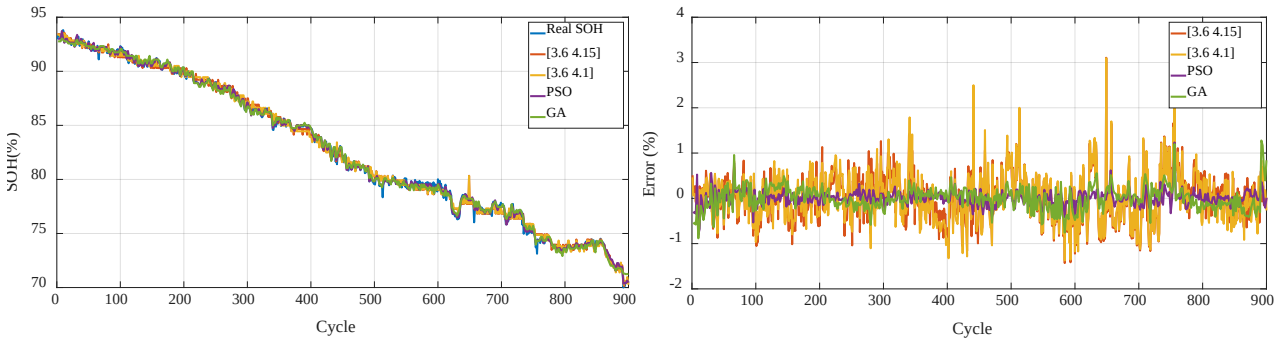
4) The last group of the voltage range and parameters are optimized by means of the GA.

In the first case, 50% experimental data set of the test battery are chosen as the training set to refine the fixed size LS-SVM algorithm, and the rest data are used to evaluate the prediction performance. The training and evaluation results are shown in Fig. 5. It is worth noting that the SOH in this paper is defined as the percentage of current maximum capacity over the rated value [3]. The initial SOH value is lower than 100% thanks to the operating environment and discharge conditions. As can be found from Fig. 5, although the estimation results with the voltage interval of [3.6 4.15] and [3.6 4.1] can roughly track the reference SOH, the estimation error is still more than that based on the PSO and GA. In contrast, the GA algorithm enables more accurate SOH estimation, which can track the reference value more precisely. The RMSE, maximum absolute error (ME) and mean absolute error (MAE) are calculated to quantify the prediction results. The statistical results of different voltage ranges listed in Table II exhibit that the GA can achieve the minimum ME, which is only one third of those by the randomly selected parameters. In addition, the MAE and RMSE by GA are both only half of those of the random selection approach of parameters. It can also be found that the MAE and RMSE of GA method are slightly higher than those of the PSO method. However, the voltage range obtained by the PSO is much wider than that by the GA. In practice, the voltage range is anticipated to be as narrow as possible for reducing the storage space and computation intensity in the premise of ensuring the precision. Moreover, the narrower voltage range under the constant current mode is easier to encounter in charge

operations. Hence, we can conclude that the proposed method shows satisfactory performance in terms of SOH estimation and parameters determination.

Table II. Statistical results of different voltage ranges.

Method	Voltage range (V)	c	b	ME (%)	MAE (%)	RMSE (%)
Randomly	[3.6 4.15]	0.75	1	3.1	0.39	0.495
	[3.6 4.1]	0.0257	14.2459	3.09	0.403	0.508
PSO	[3.4 4.15]	0.1173	16.1908	1.23	0.123	0.174
GA	[3.467 4.106]	0.0257	14.2459	1.11	0.18	0.23



(a) (b)

Fig. 5. SOH prediction results and errors at different voltage ranges: (a) SOH prediction curves; (b) SOH prediction errors.

B. Comparison with Different Algorithms

To further evaluate the performance of proposed algorithm, the NN, SVM and LS-SVM algorithms are respectively applied. For fair comparison, the feature variables remain the same; that said, the charging time ranging from 3.467 V to 4.106 V is selected as the inputs of all participant algorithms. The comparison results of different algorithms and experimentally measured values are illustrated in Fig. 6. As can be seen, the prediction results of proposed algorithm can track the actual profile most precisely, and the NN method leads to the largest fluctuation of estimation error. Table III lists the statistical results of different algorithms. It is clear that the proposed method provides the minimum MAE, ME, and RMSE. The ME, MAE and RMSE of the NN are respectively 1.93%, 0.22% and 0.29%; and nonetheless, the values by the fixed size LS-SVM are 1.11%, 0.18% and 0.23%. Particularly, the ME of proposed fixed size LS-SVM is only about 58% of that by the traditional NN algorithm, highlighting its preferable estimation capability. Moreover, the estimation results also demonstrate the superiority of the proposed fixed size LS-SVM in disposing of small sample nonlinear problems.

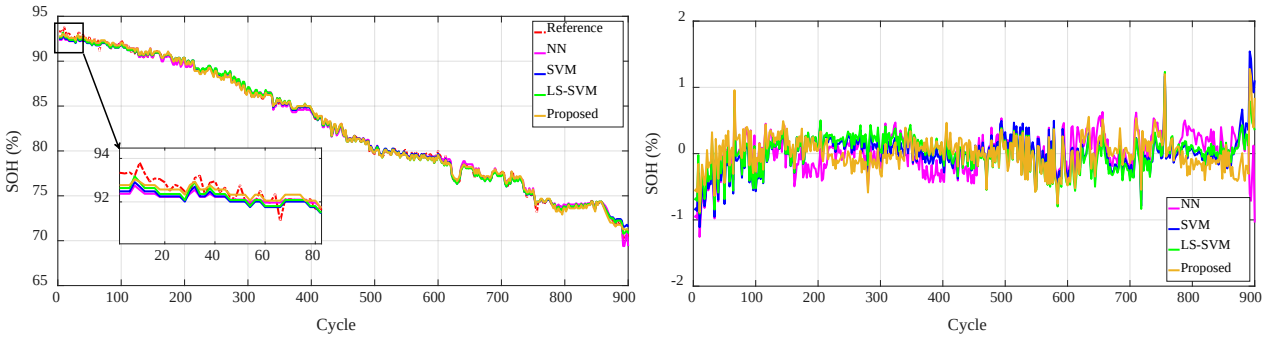


Fig. 6. SOH estimation results and errors at different estimation algorithms: (a) SOH prediction curves; (b) SOH prediction errors.

Table III. Statistical results of different algorithms.

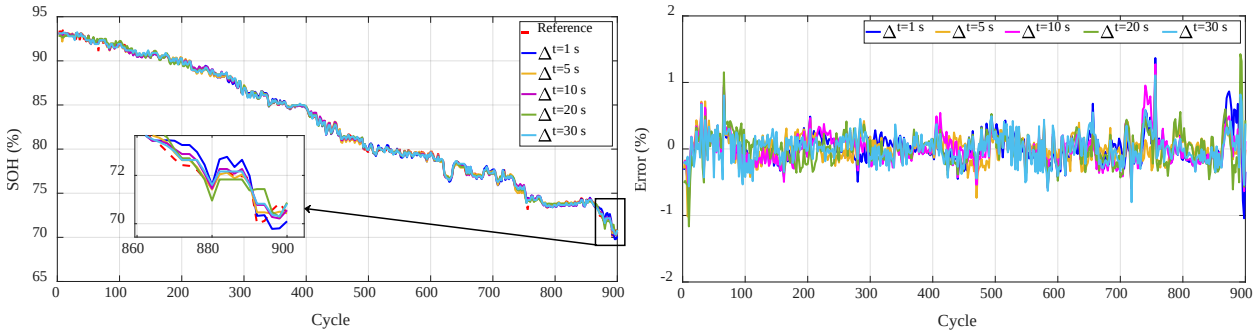
Algorithm	ME (%)	MAE (%)	RMSE (%)
NN	1.93	0.22	0.29
SVM	1.54	0.20	0.29
LS-SVM	1.23	0.20	0.26
Proposed	1.11	0.18	0.23

C. Verification on Different Sampling Time

To verify the practicability of proposed feature extraction and SOH estimation algorithms, different sampling time is taken into account to demonstrate their influence on estimation precision. To do so, the verification data at different sampling time are firstly extracted from the original measurement data. Then, the different sampling time data corresponding to the test battery is exerted to estimate the SOH evolution. Here, the feature of voltage range is still set to [3.467 4.106], and the parameters of fixed size LS-SVM are also the same as those in Section 4.1. Five SOH curves are attained based on the proposed method, wherein the sampling time are set to 1 s, 5 s, 10 s, 20 s and 30 s, respectively.

Fig. 7 demonstrates the estimation results and errors of SOH with different sampling time, and the statistical results are presented in Table IV. It is clear that the predicted SOH can track the reference values on the whole throughout the operating cycles, and it can also indicate from these similar error curves that they have the approximate estimation results. To be specific, the RMSEs when Δt equals 1 s, 5 s, 10 s, 20 s and 30 s are 0.24%, 0.21%, 0.24%, 0.26% and 0.23%, respectively. It can be intuitively found that higher sampling accuracy does not lead to better estimation performance all the time. The ME is 1.37% with the sampling time of 1 s; nevertheless, the ME is only 1.11% when the sampling time is 30 s. Thus, we can conclude that the proposed feature extraction method does not strictly rely on the sampling time. From these results, we can summarize that the sampling time of 5 s is more suitable for SOH estimation, as it can not only guarantee the

estimation performance, but also reduce the calculation burden and storage space to a great extent. From this point of view, the proposed estimation framework shows certain potential for practical application.



(a) (b)

Fig. 7. SOH estimation results and errors at different sampling time: (a) SOH prediction curves; (b) SOH prediction errors.

Table IV. Statistical results at different sampling time.

Sample time	ME (%)	MAE (%)	RMSE (%)
1 s	1.37	0.17	0.24
5 s	1.13	0.16	0.21
10 s	1.27	0.18	0.24
20 s	1.43	0.19	0.26
30 s	1.11	0.18	0.23

D. Universality Validation

After training the fixed size LS-SVM model, the test data set of other three batteries (labeled as Bat. 2, Bat. 3 and Bat. 4) are chosen as the testing set to validate the universality of proposed algorithm. The capacity attenuation profiles with cycle numbers of these three batteries are shown in Fig. 8. Note that the initial capacity is only 3.76 Ah as the battery is not fully charged. When the cycle number reaches approximately 650, the battery approaches the end of life (80% rated capacity).

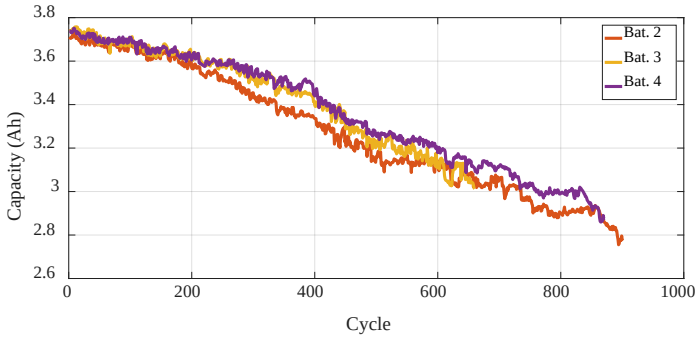
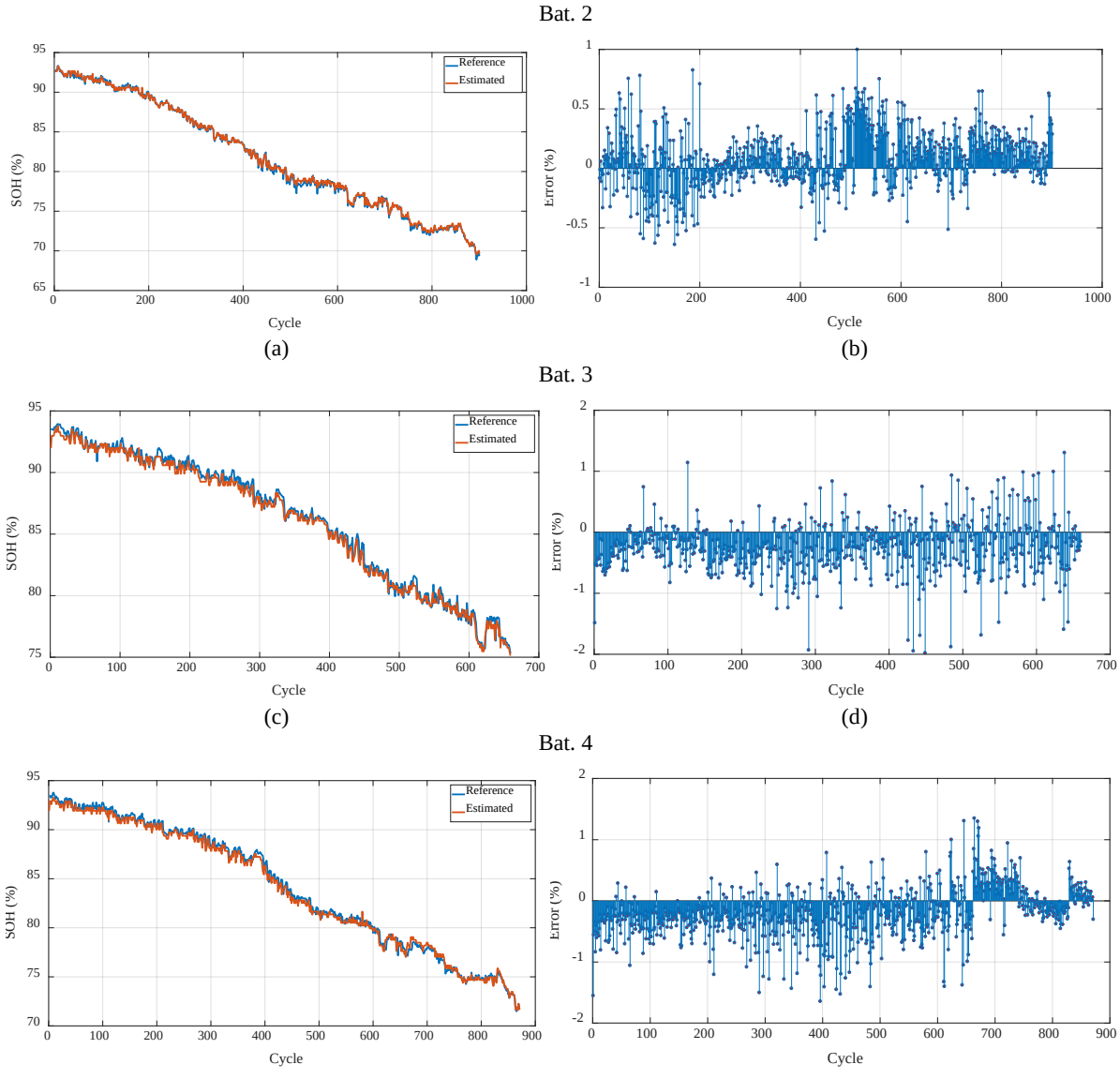


Fig. 8. Battery capacity degradation curves.

The SOH prediction curves and errors of Bats. 2 to 4 are plotted in Fig. 9, where the subfigures (b), (d) and (f) show the estimation errors of SOH, and it can be found that the error of the test battery is larger than that of the training battery. In particular, the output SOH of Bat. 2 based on the proposed estimation framework

can track the reference curve precisely in the whole process, and the estimated error is evenly distributed and less than 1% all the time. With respect to Bats. 3 and 4, the estimated SOH is slightly lower than the reference value and therefore leads to the negative errors between the estimated result and the reference value in most cases. As can also be found, the cycle times of Bat. 4 is much larger than that of Bat. 3, and after 650 cycles, the estimation error of Bat. 4 shows larger fluctuation, which can be reasoned by inconsistency of battery capacity and influence of temperature during the cycling experiment. Even so, the results in Bats. 2, 3 and 4 still indicate acceptable estimation precision, and furthermore the estimation error, as depicted in Fig. 9 (b), (d) and (f), is all less within 2%. Table V summarizes the RMSE, ME and MAE of the SOH estimation results. The maximum RMSE and MAE in the three testing batteries are 0.44% and 0.34%, respectively. All of them are fairly low, highlighting the proposed method can be extended to other batteries. From this point, we can conclude that the presented uniform SOH estimation framework shows favorable universality capability.



(e) (f)

Fig. 9. SOH prediction curves and errors: (a) Bat. 2 prediction SOH; (b) Bat. 2 prediction errors; (c) Bat. 3 prediction SOH; (d) Bat. 3 prediction errors; (e) Bat. 4 prediction SOH; (f) Bat. 2 prediction errors.

Table V. Statistical results of Bat. 2, Bat. 3 and Bat. 4.

Battery	ME (%)	MAE (%)	RMSE (%)
Bat. 2	1.00	0.19	0.25
Bat. 3	1.98	0.28	0.42
Bat. 4	1.64	0.34	0.44

V. CONCLUSIONS

In this paper, a framework is presented to achieve the state of health estimation for lithium-ion batteries using fixed size least square-support vector machine and genetic algorithm. In the proposed framework, the charging time for a fixed voltage range in the constant current charging process is taken as the input feature for the fixed size least square-support vector machine training. To prevent inaccurate prediction incurred by inappropriate parameters of fixed size least square-support vector machine, the genetic algorithm is harnessed to achieve the combined optimization of voltage range and algorithm parameters. Different voltage ranges and parameters are compared to verify the effectiveness of the developed framework and the estimation results prove that the voltage range and parameters optimized by the genetic algorithm is the best combination considering the estimation precision and practicability. We compare fixed size least square-support vector machine with three traditional methods and the fixed size least square-support vector machine is confirmed to achieve the satisfactory estimation performance. Further validations are carried out on different sampling time, comparative results indicate that the proposed feature excavation method does not rely on the sampling time. Additionally, the experimental validations are also conducted on other three unfamiliar batteries to check the universality of the proposed framework. The experimental results manifest that the present framework obtains desired estimation precision of state of health with a boundary of prediction error of 2%. In summary, the presented uniform state of health estimation framework features the advances of low sampling time requirement, high accuracy, robustness and universality capability.

In the future, we will endeavor to investigate state of health estimation principles at different environmental temperatures, and moreover finding an authentic manner to characterize the change of state of health in the battery packs is also the focus of our next step research.

ACKNOWLEDGEMENTS

This work was supported in part by the National Natural Science Foundation of China (No. 61763021 and No. 51775063), in part by the National Key R&D Program of China (No. 2018YFB0104000), and in part by the EU-funded Marie Skłodowska-Curie Individual Fellowships Project under Grant 845102-HOEMEV-H2020-MSCA-IF-2018.

REFERENCES

[1] Guo N, Shen J, Xiao R, Yan W, Chen Z. Energy management for plug-in hybrid electric vehicles considering optimal engine ON/OFF control and fast state-of-charge trajectory planning. Energy. 2018;163:457-74.

[2] Chen Z, Shu X, Xiao R, Yan W, Liu Y, Shen J. Optimal charging strategy design for lithium-ion batteries considering minimization of temperature rise and energy loss. International Journal of Energy Research. 2019;43(9):4344-58.

[3] Hu X, Liu W, Lin X, Xie Y. A Comparative Study of Control-Oriented Thermal Models for Cylindrical Li-Ion Batteries. IEEE Transactions on Transportation Electrification. 2019.

[4] Ma M, Wang Y, Duan Q, Wu T, Sun J, Wang Q. Fault detection of the connection of lithium-ion power batteries in series for electric vehicles based on statistical analysis. Energy. 2018;164:745-56.

[5] Xiong R, Li L, Tian J. Towards a smarter battery management system: A critical review on battery state of health monitoring methods. Journal of Power Sources. 2018;405:18-29.

[6] Gao Y, Jiang J, Zhang C, Zhang W, Jiang Y. Aging mechanisms under different state-of-charge ranges and the multi-indicators system of state-of-health for lithium-ion battery with Li (NiMnCo) O2 cathode. Journal of Power Sources. 2018;400:641-51.

[7] Li D, Danilov DL, Zwikirsch B, Fichtner M, Yang Y, Eichel R-A, et al. Modeling the degradation mechanisms of C6/LiFePO4 batteries. Journal of Power Sources. 2018;375:106-17.

[8] Lu L, Han X, Li J, Hua J, Ouyang M. A review on the key issues for lithium-ion battery management in electric vehicles. Journal of power sources. 2013;226:272-88.

[9] Yang Z, Patil D, Fahimi B. Online estimation of capacity fade and power fade of lithium-ion batteries based on input–output response technique. IEEE Transactions on Transportation Electrification. 2017;4(1):147-56.

[10] Chen Z, Xue Q, Xiao R, Liu Y, Shen J. State of Health Estimation for Lithium-ion Batteries Based on Fusion of Autoregressive Moving Average Model and Elman Neural Network. IEEE Access. 2019;7:102662-78.

[11] Eddahech A, Briat O, Bertrand N, Deletage J-Y, Vinassa J-M. Behavior and state-of-health monitoring of Li-ion batteries using impedance spectroscopy and recurrent neural networks. International Journal of Electrical Power & Energy Systems. 2012;42(1):487-94.

[12] Wassiliadis N, Adermann J, Frericks A, Pak M, Reiter C, Lohmann B, et al. Revisiting the dual extended kalman filter for battery state-of-charge and state-of-health estimation: A use-case life cycle analysis. Journal of Energy Storage. 2018;19:73-87.

[13] Su X, Wang S, Pecht M, Zhao L, Ye Z. Interacting multiple model particle filter for prognostics of lithium-ion batteries. Microelectronics Reliability. 2017;70:59-69.

[14] Kim I-S. A technique for estimating the state of health of lithium batteries through a dual-sliding-mode observer. IEEE Transactions on Power Electronics. 2009;25(4):1013-22.

[15] Zhang C, Allafi W, Dinh Q, Ascencio P, Marco J. Online estimation of battery equivalent circuit model parameters and state of charge using decoupled least squares technique. Energy. 2018;142:678-88.

[16] Chen C, Xiong R, Shen W. A Lithium-Ion Battery-in-the-Loop Approach to Test and Validate Multiscale Dual H Infinity Filters for State-of-Charge and Capacity Estimation. IEEE Transactions on Power Electronics. 2017;33(1):332-42.

[17] Xiong R, Li L, Li Z, Yu Q, Mu H. An Electrochemical Model Based Degradation State Identification Method of Lithium-Ion Battery for All-Climate Electric Vehicles Application. Applied Energy. 2018;219:264-75.

[18] Feng X, Weng C, He X, Han X, Lu L, Ren D, et al. Online State-of-Health Estimation for Li-Ion Battery Using Partial Charging Segment Based on Support Vector Machine. IEEE Transactions on Vehicular Technology. 2019;68(9):8583-92.

[19] Liu K, Hu X, Wei Z, Li Y, Jiang Y. Modified Gaussian Process Regression Models for Cyclic Capacity Prediction of Lithium-ion Batteries. IEEE Transactions on Transportation Electrification. 2019.

[20] Ma G, Zhang Y, Cheng C, Zhou B, Hu P, Yuan Y. Remaining useful life prediction of lithium-ion batteries based on false nearest neighbors and a hybrid neural network. Applied Energy. 2019;253:113626.

[21] Yang J, Peng Z, Wang H, Yuan H, Wu L. The remaining useful life estimation of lithium-ion battery based on improved extreme learning machine algorithm. Int J Electrochem Sci. 2018;13:4991-5004.

[22] Li X, Yuan C, Li X, Wang Z. State of health estimation for Li-Ion battery using incremental capacity analysis and Gaussian process regression. Energy. 2019:116467.

[23] Liu Y, Li J, Chen Z, Qin D, Zhang Y. Research on a multi-objective hierarchical prediction energy management strategy for range extended fuel cell vehicles. Journal of Power Sources. 2019;429:55-66.

[24] Yang D, Wang Y, Pan R, Chen R, Chen Z. State-of-health estimation for the lithium-ion battery based on support vector regression. Applied Energy. 2018;227:273-83.

[25] Chen Z, Sun M, Shu X, Xiao R, Shen J. Online state of health estimation for lithium-ion batteries based on support vector machine. Applied Sciences. 2018;8(6):925.

[26] Wei J, Dong G, Chen Z. Remaining useful life prediction and state of health diagnosis for lithium-ion batteries using particle filter and support vector regression. IEEE Transactions on Industrial Electronics. 2017;65(7):5634-43.

[27] Patil MA, Tagade P, Hariharan KS, Kolake SM, Song T, Yeo T, et al. A novel multistage Support Vector Machine based approach for Li ion battery remaining useful life estimation. Applied Energy. 2015;159:285-97.

[28] 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.

[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 Transactions on Systems, Man, and Cybernetics: Systems. 2015;45(6):915-28.

[30] Goh T, Park M, Seo M, Kim JG, Kim SW. Capacity estimation algorithm with a second-order differential voltage curve for Li-ion batteries with NMC cathodes. Energy. 2017;135:257-68.

[31] Li Y, Abdel-Monem M, Gopalakrishnan R, Berecibar M, Nanini-Maury E, Omar N, et al. A quick on-line state of health estimation method for Li-ion battery with incremental capacity curves processed by Gaussian filter. Journal of Power Sources. 2018;373:40-53.

[32] Pan H, Lü Z, Wang H, Wei H, Chen L. Novel battery state-of-health online estimation method using multiple health indicators and an extreme learning machine. Energy. 2018;160:466-77.

[33] Meng J, Cai L, Stroe D-I, Luo G, Sui X, Teodorescu R. Lithium-ion battery state-of-health estimation in electric vehicle using optimized partial charging voltage profiles. Energy. 2019;185:1054-62.

[34] Mahdjoubi A, Zegnini B, Belkheiri M, Seghier T. Fixed least squares support vector machines for flashover modelling of outdoor insulators. Electric Power Systems Research. 2019;173:29-37.

[35] Liu Y, Wang W, Ghadimi N. Electricity Load Forecasting by an Improved Forecast Engine for Building Level Consumers. Energy. 2017;139:18-30.

[36] Suykens JA, Vandewalle J. Least squares support vector machine classifiers. Neural processing letters. 1999;9(3):293-300.

[37] Cai L, Meng J, Stroe D-I, Luo G, Teodorescu R. An evolutionary framework for lithium-ion battery state of health estimation. Journal of Power Sources. 2019;412:615-22.

[38] Chen Z, Shu X, Li X, Xiao R, Shen J. LiFePO4 battery charging strategy design considering temperature rise minimization. Journal of Renewable and Sustainable Energy. 2017;9(6):064103.

[39] Aghajani G, Ghadimi N. Multi-Objective Energy Management in a Micro-Grid. Energy Reports. 2018;4:218-25.