Gravitational search algorithm with Gaussian process for lithium-ion batteries state of health (SOH) estimation

Jing Ye[1, a *], Santong Zhang[1,b] and Wei Yang[1,c]

[1]School of Electronic Information Engineering, Beijing Jiaotong University, Beijing, China

[a]jingye92@126.com, [b]stzhang@bjtu.edu.cn, [c]wyang@bjtu.edu.cn

Keywords: Gaussian process regression, gravitational search algorithm optimization, lithium-ion battery, state of health.

Abstract. State of health (SOH) estimation plays a significant role in battery prognostics. In this paper, Gaussian Process Regression (GPR) is used as a data-driven approach to perform SOH estimation, which supports uncertainty representation and management. At present, the hyper- parameters of GPR are optimized by conjugate gradient algorithm. However, the conjugate gradient algorithm has the shortcomings of too strong dependence on initial value and easily falling into local optimum. In order to improve the prediction precision and generalization ability of GPR, we utilized Gravitational Search Algorithm (GSA) replace of conjugate gradient to search the optimal hyper-parameters during the training process automatically then formed the GSA-GPR algorithm. Experimental results confirm that the proposed method can be effectively applied to lithium-ion batteries SOH estimation by quantitative comparison with the standard GPR algorithms, Genetic Algorithm (GA)-GPR and Particle Swarm Optimization (PSO)-GPR algorithms.

Introduction

Lithium-ion batteries are core components in a wide variety of systems due to their high energy density, high galvanic potential, lightness of weight and long lifetime [1]. However, failure of a battery could lead to reduced performance, operational impairment and even catastrophic failure, especially in aerospace systems [2]. Therefore, Batteries Prognostics and Health Management (PHM) is a subject of great interest to the electronics industry. Conventionally, PHM is encapsulated in the terms state-of-health (SOH), state-of-charge (SOC), and state-of-life (SOL), respectively [3]. Here, we discussed on SOH estimation. SOH is a measure of its health condition relative to a pristine state as an indicator of the age and condition of the battery in its life cycle [4]. As a critical part of battery PHM, a good SOH prognostic would indicate the performance degradation, prevent possible accidents and provide a significant value addition to the management of any operation involving electrical systems.

Traditionally, there exist several methods to derive a battery SOH, such as comparing the internal resistance [5], impedance analysis [6], voltage drop [7], self-discharge etc. However, traditional methods cannot provide online and precision estimated for SOH. In recent years, many modern theories, which can be divided into model-based categories and data-driven categories, have been proposed to improve the estimation of the SOH. In the model-based approach domain, Bhangu [8] used the extended Kalman filter (EKF) in real-time prediction of SOC and SOH. Saha [3] applied Particle Filters (PF) to predict the RUL of lithium-ion batteries based on Bayesian framework. However, model-based approach is formidable for the lithium-ion battery SOH prognostics due to the model hardly take all the complicated on-line conditions of battery system into consideration. Therefore, data-driven prognostics algorithm is relatively flexible for the SOH prediction of lithium battery. Kozlowski [9] proposed a data-driven approach that combines three predictors, auto-regressive moving average (ARMA), neural networks, and fuzzy-logic to predict the battery SOH. Guo [10] applied an improved nonlinear degradation Autoregressive to perform batteries SOH prognostics. However, both of the two methods lacked uncertainty expression and management for prognostics.

In this study, Gaussian Process Regression (GPR) [11] is applied as a data-driven approach to perform battery SOH prognostics. GPR algorithm holds many advantages such as programming easily

and prediction with probability interpretation. Due to its advantages, GPR has been applied in many areas [12-16], including the domain of battery prognostics [2,14]. In the above studies, the hyper-parameters of GPR are optimized by maximizing likelihood function of training samples based on conjugate gradient algorithm. However, the conjugate gradient algorithm has the shortcomings of too strong dependence on initial value in optimization effect and easily falling into local optimum, which caused the poor extrapolation performance of GPR when test data are ''distant'' from the training data. To address this issue, Genetic Algorithm (GA) [15] and Particle Swarm Optimization (PSO) [16] have been applied to hyper-parameters optimization of GPR. However, falling into local optimum in different degrees was still occurred in the process of optimization. In this study we present Gravitational Search Algorithm (GSA) [17] replace of conjugate gradient to search the optimal hyper-parameters during the training process automatically, combined with cross validation then formed the GSA-GPR algorithm. Experimental results from the processing of a NASA battery data set proved the effectiveness of the GSA-GPR algorithm compared with the basic GPR algorithms, GA-GPR and PSO-GPR algorithms.

This paper is organized as follows. First, in Section 2, two main methods for calculating the SOH of batteries are introduced. The prediction methods of GPR, GSA, and hybrid GSA-GPR algorithm are presented in Section 3. The prognostics for SOH of lithium-ion batteries are described in detail in Section 4. Finally, conclusions are discussed in Section 5.

The SOH of lithium-ion batteries.

Battery operation is a dynamic complex process, and its performance is strongly influenced by ambient environmental and load conditions [18]. There are two main methods for calculating the SOH of batteries [19]. First, battery impedance may be introduced to determine the battery SOH. Its definition can be given by:

$$SOH = \frac{R_i}{R_0} \times 100\% .$$

(1)

where R_0 is the initial impedance and R_i is the i th impedance measurement that is varied with the charge and discharge cycles.

Second, battery capacity can also be used to indicate battery SOH. This method can be expressed as:

$$SOH = \frac{C_i}{C_0} \times 100\% .$$

(2)

where C_0 is the initial capacity and C_i is the i th capacitance value degenerated with cycles. In this study, we choose the battery capacity to measure the SOH.

Methods

Gaussian Process Regression. The Gaussian Process (GP) is defined as a probability distribution over a number of variables $\{f(x_i)|x_i \in x\}$ indexed by a set x [11]. It can be fully specified by its mean function $m(x)$ and its covariance function $k(x,x')$ .and described as $f(x) \square GP(m(x),k(x,x'))$ .When a Gaussian process is applied to a general regression problem, we describe the target y_{target} given the effect of noise, which is $y_{target} = f(x) + e$. Where e is the white Gaussian noise, and $e \square N(0, \sigma_n^2)$. Now let y be the known target values of the training set X , and let y_* be a set of function values

corresponding to the test set inputs x_*. The goal of GPR is to compute posterior estimates by constraining the prior distribution to fit the available training data.

Prior:

$$\begin{pmatrix} y \\ y_* \end{pmatrix} \sim N\left(0, \begin{bmatrix} K(X,X) + \sigma_n^2 I & K(X,x_*) \\ K(x_*,X) & K(x_*,x_*) \end{bmatrix}\right).$$ (3)

Posterior:

$$y_* / X, y, X_* \sim N(\overline{y_*}, cov(y_*)).$$ (4)

where: $\overline{y_*} = E[y_* | X, y, X_*] = K(X_*,X)\left[K(X,X) + \sigma_n^2 I\right]^{-1} y$.

$cov(y_*) = K(X_*,X_*) - K(X_*,X)\left[K(X,X) + \sigma_n^2 I\right]^{-1} K(X,X_*)$.

where I is the unit matrix, $K(X,X)$, $K(X,x_*)$ and $K(x_*,x_*)$ are covariance matrix for the training cases, the training-testing cases, and the testing cases, respectively. There are many choices of covariance functions. In this study, combined kernel function K_{MP} was adopted.

$$\begin{cases} K_{Matern}(x,x') = \sigma_{f1}^2 \dfrac{2^{1-v}}{\Gamma(n)}(\dfrac{\sqrt{2n}r}{l_1})^n K_n(\dfrac{\sqrt{2n}r}{l_1}) , \ r = |x - x'|. \\[3mm] K_{PER}(x,x') = \sigma_{f2}^2 \exp[-\dfrac{2}{l_2^2}\sin^2(\pi r / p)], r = |x - x'|. \\[3mm] K_{MP}(x,x') = K_{Matern} \times K_{PER}. \end{cases}$$ (5)

where let $\theta = [\sigma_{f1}, \sigma_{f2}, l_1, l_2, v, p]$, of which the σ_{f1} and σ_{f2} are the signal variance, l_1 and l_2 reflect the length-scale, and v is a positive parameter , and p is the periodic parameter.

A GPR model can approximate any system by appropriate hyper-parameter θ . But the hyper-parameters are unknown and need to be optimized in training process. In previous work, hyper-parameters of GP is obtained by maximizing likelihood function of training samples based on conjugate gradient algorithm, as showed in Eq. 6.

$$\frac{\partial}{\partial q_j}\log(p(y|X,q)) = \frac{1}{2}tr\left\{\left[\alpha\alpha^T - (K + s_n^2 I)^{-1}\right]\frac{\partial(K + s_n^2 I)}{\partial q_j}\right\}.$$ (6)

However, conjugate gradient algorithm has the shortcomings of too strong dependence on initial value in optimization effect and easily falling into local optimum. In this paper, a hybrid GSA-GPR algorithm was proposed for searching the optimal hyper-parameters during the training process automatically.

A standard GSA. GSA is one of the newest heuristic algorithms introduced by Rashedi et al in 2009[17]. It is inspired by the Newtonian laws of gravity and motion. The fruitful ability of GSA in finding and converging to an optimum infers from the results of experiments undertaken previously , the results obtained by GSA in various standard benchmark functions provide more superior results than PSO and GA. Compared with conjugate gradient method, GSA is a powerful global searcher which is also flexible for the non-differential objective functions.

To describe the GSA consider a system with N agents in which position of the ith agent is defined as follows:

$$X_i = (x_i^1, x_i^2, ..., x_i^d, ..., x_i^n) , i = 1, 2, \mathbf{L}, N. \tag{7}$$

where x_i^d is position of ith agent in the dth dimension and n is dimension of the search space. Based on [17], the mass of each agent is calculated by the fitness evaluation which is as follows:

$$\begin{cases} m_i(t) = \dfrac{m_{fit_i}(t) - m_{worst}(t)}{m_{best}(t) - m_{worst}(t)}, \\ M_i(t) = \dfrac{m_i(t)}{\sum_{j=1}^{N} m_j(t)}. \end{cases} \tag{8}$$

where $M_i(t)$ and $m_{fit_i}(t)$ represent the mass and the fitness value of the agent i at time t, $m_{best}(t)$ and $m_{worst}(t)$ are defined as follows (for a minimization problem):

$$\begin{cases} m_{best}(t) = \min_{j \in \{1,2,\mathbf{L},N\}} m_{fit_j}(t), \\ m_{worst}(t) = \max_{j \in \{1,2,\mathbf{L},N\}} m_{fit_j}(t). \end{cases} \tag{9}$$

To compute the acceleration a_i^d of an agent, total forces F_i^d from other agents that applied on it should be considered based on the law of gravity (Eq. 10), which is followed by calculation of agent acceleration using law of motion (Eq. 11). Afterwards, next velocity v_i^d and position x_i^d of an agent could be calculated using Eq. 12 and Eq. 13, respectively.

$$F_i^d(t) = \sum_{j=k_{best}, j \neq i} R_1 \times F_{ij}^d(t) = \sum_{j=k_{best}, j \neq i} R_1 \times G(t) \frac{M_i(t) \times M_j(t)}{R_{ij}(t) + e} (x_j^d(t) - x_i^d(t)). \tag{10}$$

$$a_i^d(t) = \frac{F_i^d(t)}{M_{ii}(t)}. \tag{11}$$

$$v_i^d(t+1) = R_2 \times v_i^d(t) + a_i^d(t). \tag{12}$$

$$x_i^d(t+1) = x_i^d(t) + v_i^d(t+1). \tag{13}$$

where R_1 and R_2 is two uniform random variable in the interval [0, 1], e is a small constant, and $R_{ij}(t)$ is the Euclidian distance between two agents i and j, $G(t) = G_0 e^{-bt/T}$ is the gravitational constant at t of which G_0 is an initial value, T is the biggest iteration and b is a constant, k_{best} is the set of first K agents with the best fitness value and biggest mass.

Hybrid GSA-GPR algorithm. The hybrid GSA-GPR method mentioned above is illustrated by the diagram in Fig. 1. In order to improve the generalization ability and the extrapolation performance of GPR when test data are ''distant'' from the training data, the *cross validation* method was integrated into the algorithm. Training cases X was divided into two parts, one part was off-line data, which named X1 and used as the learning samples of GPR model, and the other was online data, which named X2 and used as the testing sample of GPR model. Then, X1 was further divided into two parts of S1 and S2. In the process of searching the optimal hyper-parameters based on GSA, S1 was used as

training set , S2 was used as *validation set*, optimization targets of GSA was minimizing the mean square error of S2.
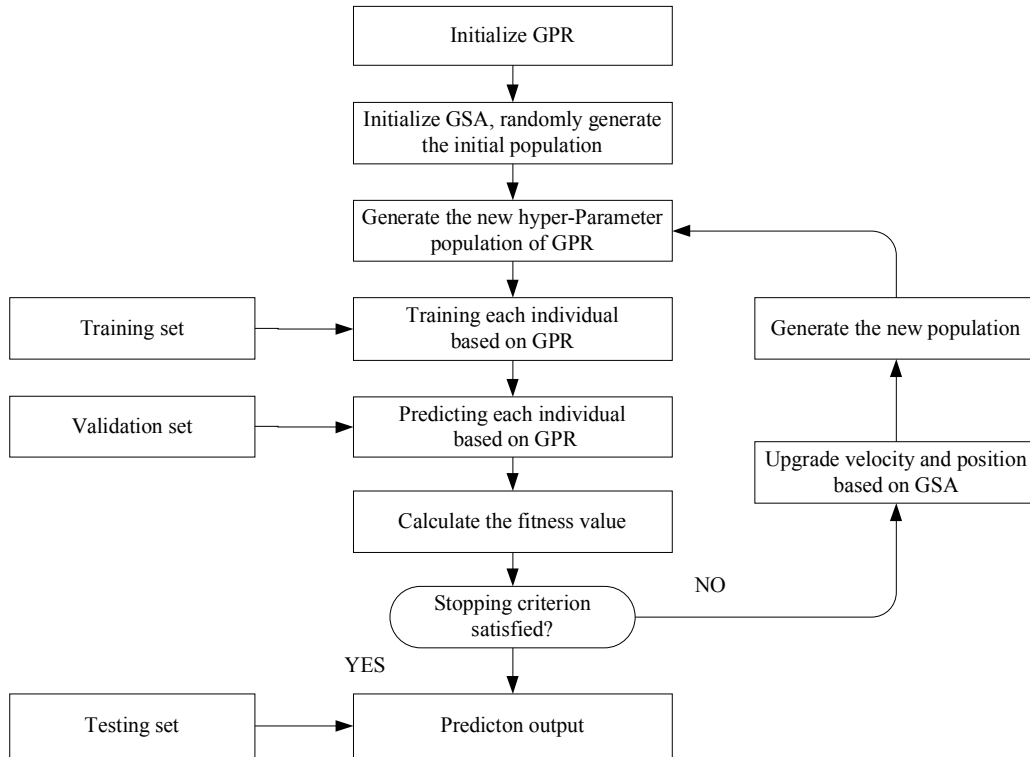

Fig. 1. The prediction framework based on the GSA-GPR model.

Experiments and results

Raw data. The lithium-ion batteries data that we used to make prognostics was obtained from the data repository of the NASA Ames Prognostics Center of Excellence (PCoE) [2]. The batteries (No.5, No.6, No.7 and No.18) were run through three different operational profiles (charge, discharge and impedance) at room temperature. Repeated charge and discharge cycles result in accelerated aging of the batteries while impedance measurements provide insight into the internal battery parameters that change as aging progresses. The experiments were stopped when the batteries reached end-of-life criteria, which was a 30% fade in rated capacity (from 2 [Ahr] to 1.4 [Ahr]).

The data from battery No.6 was selected to the experiments, including training and testing. The test results of SOH conducted for battery No.6 is shown in Fig. 2, which containing aging information about battery SOH values from cycle 1 to cycle 168. From Fig. 2, it is showed that the SOH is higher at cycle 90 than in previous cycles due to significant regeneration during the resting period, and similar regeneration phenomenon also emerges in other cycles. Therefore, the prognostics model must be selected to respond to the degradation trend and regeneration phenomenon in lithium-ion batteries.

Lithium-ion batteries SOH estimation base on Hybrid GSA-GPR. The prediction results based on a hybrid GSA-GPR model are shown in Fig. 3, and the prognostics results also include 95% confidence bounds for uncertainty representation. The SOH data from cycle 1 to cycle 100 were selected as the training data sets, and the rest served as testing data sets. Population size of GSA was set to 60, maximum iteration was set to 100, G_0 was set to 100 and b was set to 20.

As seen in Fig. 3, the black line represents the actual SOH, the mean prediction output marked with the red line is quite close to the actual SOH for the whole testing data sample, even if the test data is ''distant'' from the training data set. The confidence interval with the grey region is narrow, which indicates the high reliability of the prediction result. We can find that both the point prediction result and the uncertainty representation are quite satisfied. Moreover, as shown in Fig. 3, the prediction

curve can well capture the self-recharge phenomenon, in other words, the hybrid GSA-GPR model can respond to the regeneration phenomenon of battery degradation.
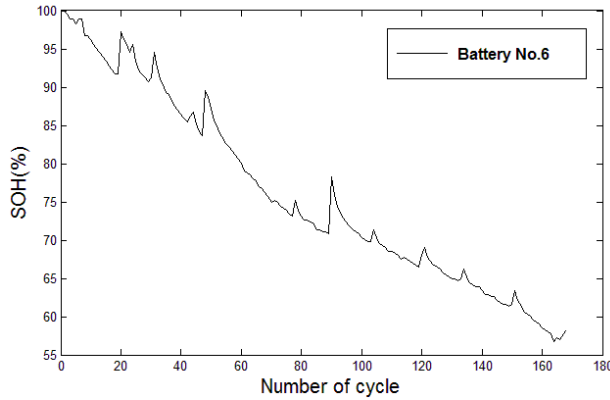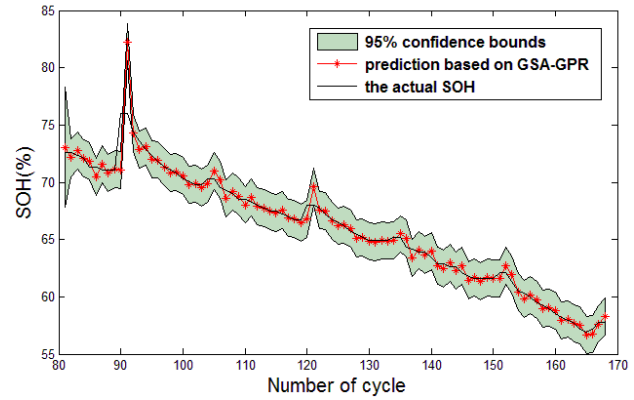

Fig. 2. The SOH of Battery No. 6.

Fig. 3. Battery health prognostics based on GSA-GPR with 95% confidence bounds.

To evaluate the performance of GSA-GPR algorithm, GA(crossover and mutation probability values of 0.8 and 0.2 respectively) and PSO(inertia factor $\omega(i)=0.9\text{-}0.5(i/i_{max})$, acceleration coefficients c1=c2=1.4945) were also used to model the same examples and the experimental results are compared in Fig. 4.

From Fig. 4, it is shown that PSO is poor in finding good solutions since it cannot explore well, but GSA and GA are more successful in exploring and exploiting the search space, especially GSA. Compare to GA, GSA is better in exploiting, which significantly caused GSA is more robust than GA in finding the global optimum and obtain good results.

In order to evaluate the three models as well as the basic GPR model, we compared the predictions, as shown in Fig. 5, in which only mean prediction values are involved.
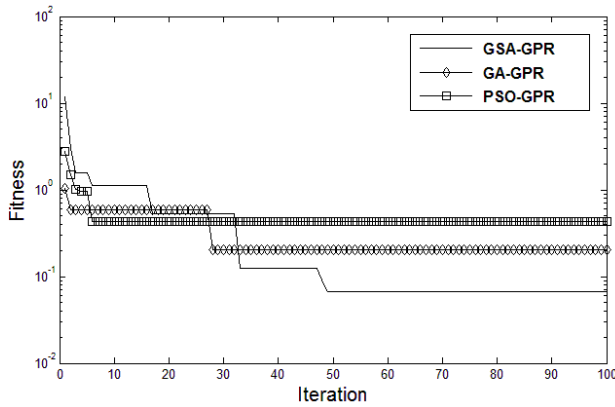

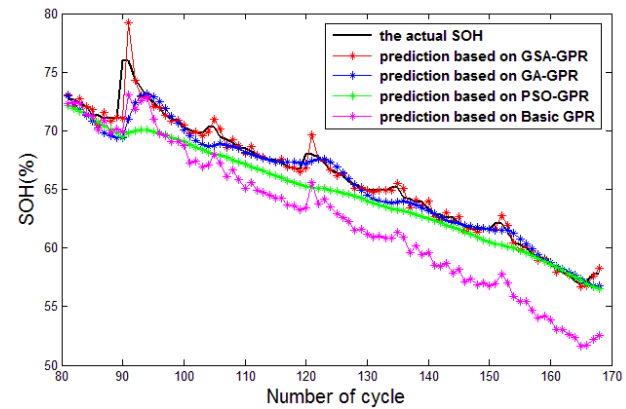Fig.4 Comparison of performance of GSA-GPR, GA-GPR and PSO-GPR.GSA-GPR,

Fig. 5. Battery health prognostics compared with GA-GPR, PSO-GPR and basic GPR.

As seen in Fig. 5, prediction based on the GSA-GPR model produced better results than the GA-GPR model, PSO-GPR model and basic GPR. Compared to the basic GPR, the prediction precision with the GSA-GPR improved greatly. In addition, we also made prognostics for batteries No. 5 and No. 7, the quantitative analysis results of the experiment on batteries No. 5, No.6 and No.7 based on the four different models are shown in Tables 1.Here, two criteria, including the root mean square error (RMSE) and mean absolute percentage error (MAPE), were introduced to evaluate the prediction performance.

$$RMSE = \sqrt{\sum_{i=1}^{n}\left(\overline{f}_{*i} - y_i\right)^2 / n}. \tag{14}$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left[\left| \overline{f}_{*i} - y_i \right| / y_i \right].\tag{15}$$

where y_i is the actual SOH, $\overline{f}_{*i}$ is the prediction value of SOH, n is the number of prediction points.

Table.1 Comparison of four different modesl for batteries.

Batteries	Method	RMSE	MAPE
Battery No.5	GSA-GPR	0.3451	0.0049
	GA-GPR	0.5632	0.0079
	PSO-GPR	1.2192	0.0137
	Basic GPR	9.5658	0.0429
Battery No.6	GSA-GPR	0.5876	0.0059
	GA-GPR	1.0036	0.0095
	PSO-GPR	2.9339	0.0197
	Basic GPR	7.7477	0.0499
Battery No.7	GSA-GPR	0.2037	0.0043
	GA-GPR	0.4741	0.0058
	PSO-GPR	2.1045	0.0184
	Basic GPR	9.2117	0.0383

As shown in Table 1, the prognostic RMSE and MAPE for three batteries based on the GSA-GPR were the smallest compared to the basic GPR model, GA-GPR model and PSO-GPR, in other words, prediction based on the GSA-GPR model is much more precise than the other three models. For example, the prediction RMSE on battery No. 6 based on the basic GPR was 7.7477, while the prediction RMSE based on the GSA-GPR was only 0.5876. This indicates that GSA-GPR obtains much better prediction performance than the basic GPR model. The results were similar for Battery No.5 and Battery No.7. From the above experiments, it can be concluded that the prediction precision with the proposed method GSA-GPR improved greatly compared to the basic GPR, meanwhile, GSA-GPR obtains much better prediction performance than GA-GPR and PSO-GPR.

Conclusions

In this paper, we presented a battery SOH estimation approach based on GSA-GPR algorithm, obtained results confirm the efficiency of the proposed method. This method has the following characteristics:

1. GSA-GPR, being a purely data-driven method, incorporate any physics of the process into the computation, and does not require large amount of historical data to perform offline training, the application is flexible and versatile, which is suitable for the prediction of the SOH of the lithium battery.

2. The hybrid GSA-GPR model obtains satisfactory prediction ability and uncertainty representation to respond to the degradation trend and regeneration phenomenon in lithium-ion batteries.

3. The GSA-GPR model effectively solve the problem of GPR that too strong dependence on initial value and the poor generalization ability which is caused by the conjugate gradient method. Experiments confirmed that the proposed approach GSA-GPR achieved much better prediction performance than the basic GPR.

4. Compared with the other two hybrids of intelligent algorithm of GA-GPR and PSO-GPR, GSA-GPR has stronger adaptability in the application of lithium-ion batteries SOH estimation.

References

[1] W. He, N. Williard, M. Osterman, et al. Journal of Power Sources, 2011, 196(23):10314-10321.

[2] K. Goebel, B. Saha, A. Saxena, et al. IEEE Instrumentation and Measurement Magazine, 2008, 11(4):33-40.

[3] B. Saha, K. Goebel, S. Poll, et al. Instrumentation and Measurement, IEEE Transactions on, 2009, 58(2):291-296.

[4] B.X. Sun, J.C. Jiang, F.D. Zheng, et al. Journal of Power Sources, 2015, 282:146-157.

[5] K. Jonghoon, B.H. Cho, Vehicular Technology, IEEE Transactions 60 (2011) 4249-4260.

[6] A. Eddahech, O. Briat. Electrical Power and Energy Systems, 2012, 42(1):487-494.

[7] Y. Zhang, C.Y. Wang, X. Tang, Journal of Power Sources 196 (2011) 1513-1520.

[8] B.S. Bhangu, P. Bentley, D.A Stone, et al. IEEE Transactions on Vehicular Technology, 2005, 54:783-94.

[9] J.D. Kozlowski. Electrochemical cell prognostics using online impedance measurements and model-based data fusion techniques. Aerospace Conference, 2003. Proceedings. 2003 IEEE. IEEE, 2003(7):3257-3270.

[10] L.M. Guo, J.Y. Peng, D.T. Liu, et al. Data-driven framework for lithium-ion battery remaining useful life estimation based on improved nonlinear degradation factor. Electronic Measurement & Instruments (ICEMI), 2013 IEEE 11th International Conference on. IEEE, 2013, 2:1014-1020.

[11] C.E. Rasmussen, C.K.I. Williams. Gaussian Processes for Machine Learning, The MIT Press, London. 2006.

[12] C. Park, J.Z. Huang, Y. Ding. The Journal of Machine Learning Research, 2011, 12:1697-1728.

[13] L. Wang, B.J. Zou, X.N. Peng, et al. Acta Electronica Sinica, 2007, 35(11):2087-2091. In Chinese

[14] D.T. Liu, J.Y. Pang, J.B. Zhou, et al. Microelectronics Reliability, 2013, 53(6):832-839.

[15] K.Y. Liu, B.G. Liu, C. Xu. Chinese Journal of Rock Mechanics and Engineering, 2009, 28(10):2128-2134. In Chinese

[16] F.W. Zhu, C. Xu, G.S. Dui. Particle swarm hybridize with Gaussian process regression for displacement prediction. Proc of the 5th IEEE Int Conf on Bio-Inspired Computing: Theories and Applications. Changsha, 2010:522-525.

[17] E. Rashedi, H. Nezamabadi-Pour, S. Saryazdi. Information sciences, 2009, 179(13):2232-2248.

[18] B. Saha, K. Goebel, S. Poll, et al. An integrated approach to battery health monitoring using Bayesian regression and state estimation. In autotestcon, 2007 IEEE. IEEE, 2007:646-653.

[19] L. Daniel, X.D. Tang. Lithium-ion battery state of health estimation using Ah-V characterization. Proceedings of the Annual Conference of Prognostics and Health Management (PHM) Society, Montreal, QC, Canada. 2011, 2529:367373.