The Application Of Gaussian Process Regression In State Of Health Prediction Of Lithium Ion Batteries

Yanqin Zhang, Huafeng Zhang, Zhiwei Tian
College of Environmental and Energy Engineering, Beijing University of Technology
Beijing, China
zhangyanqin@bjut.edu.cn, 1147667027@emails.bjut.edu.cn, 1720058817@qq.com

***Abstract*—To predict the available capacity and state of health of lithium ion batteries by using Gaussian process regression, one of the crucial tasks is to choosing the covariance function. This paper proposes a method which can fulfill the Gaussian process regression by using most proper covariance functions or the optimal combination. First, a variety of typical functions are tried to fit the battery experimental data points with least square method, which can give us a valuable interpretation of the properties of the data; next, the three functions with minimum root mean square errors are selected to guide the choosing of the patterns of the covariance functions; then the Gaussian process regression is applied on the training data to determine the ultra-parameters included; finally , we use the Gaussian process model to predict the latter cycle capacities within the test data. Experiments show that the combination of selected covariance functions is effective and can be applied on predictions with different batteries. Also, the method can reduce the time in applying Gaussian process regression by determination of the covariance function quickly.**

Keywords—lithium ion battery; state of health; prediction; Gaussian process regression; covariance functions

I. INTRODUCTION

In recent years, the lithium ion batteries are widely used in mobile devices, electric vehicles and aerospace industry as the most promising power batteries. Although the lithium ion battery has numerous benefits, the uncertainty in performance life hinders the advantages of energy density and reliability. The process of capacity fade is complex and strongly depends on cycling scheme, operating temperature and storage conditions which cannot distinguish respectively in real circumstances [1~3]. Therefore, prediction of the available capacity and the state of health (SOH) remains a challenging problem for the long life battery applications.

The SOH of lithium battery can be estimated by the capacity, the internal resistance and other performance parameters, and then combined with a certain threshold to determine the invalidity of the battery, in practice some people estimate the SOH from a single measurement of either the cell impedance or the cell capacity [4, 5]. Existing SOH estimation methods can be divided into model-based and data-driven approaches. The equivalent circuit model has been proposed to describe the cell's electrochemistry and the parameters are identified and used to indicate the cycling life [6]. Except the performance model, the cycle life model [7, 8] and the use of neural network [9, 10] for characterizing the battery cycling pattern has been applied to predict the available capacity. However, as the dispersion of battery's internal parameters due to manufacturing and the operating condition cannot be neglected, the pattern which applied for the test battery may not applied to the battery in use even they are of the same kinds. Since the data-driven approaches are more flexible, the support vector machine [4] and Gaussian process regression (GPR) [5] have been used to develop the regression models and to forecast the battery capacity and the battery SOH. These methods not only are capable of predicting different batteries at different conditions, but also can give the uncertainty of the predictions which can guide proper decision of battery maintenance. Despite that these methods can obtain sound predictions, the determination of kernel functions and covariance functions is empirical and hard to follow. Especially in situation when the experimental data is very large, the training time is very long.

In this study, we try to find the route of determining the proper covariance functions in the application of Gaussian process regression. The rest sections of the paper are arranged as follows: in section 2, the Gaussian process regression is introduced briefly. The data fitting functions are tested and the covariance functions are chosen, and the prediction algorithm by using GPR is developed in section 3. The final conclusions are given in section 4.

II. INTRODUCTION OF GAUSSIAN PROCESS REGRESSION

A. Gaussian process for regression

The Gaussian process is defined as a collection of finite number of random variables , and it can be fully described by the mean function and co-variance function , the mean function and covariance function are defined in (1),

$$\begin{cases} m(x) = E[f(x)] \\ k(x,x')=E[(f(x)-m(x))(f(x')-m(x'))] \end{cases} \quad (1)$$

Thus a Gaussian process can be written in (2),

$$f(x) \sim GP(m(x), k(x,x')) \quad (2)$$

In this case, the function $f(x)$ are the random variables represent the value of training points at x.

Suppose the target value y can be modeled as observed experimental data plus Gaussian white noise as follows:

$$y = f(x) + \varepsilon \tag{3}$$

Where $\varepsilon \sim N(0, \sigma^2)$, since $f(x)$ is a Gaussian process, y is also a Gaussian process, and described as in (4)

$$y \sim GP(m(x), k(x, x') + \sigma_n^2 \delta_{ij}) \tag{4}$$

Where, δ_{ij} is Kronecker delta function, $\delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$. The application of Gaussian process regression is mainly to find the proper covariance function $k(x, x')$.

From the training points $\{(x_i, y_i) \mid i = 1, 2, 3...n\}$, and the test points $\{x_i \mid i = n+1, n+2, n+3....\}$, the joint distribution of the training outputs f and the test outputs f_* according to the prior is given by (5):

$$\begin{bmatrix} f \\ f_* \end{bmatrix} \sim N(0, \begin{bmatrix} K(X,X)+\sigma_n^2 I & K(X,X_*) \\ K(X_*,X) & K(X_*,X_*) \end{bmatrix}) \tag{5}$$

Then $K(X, X_*)$ denotes the $n \times n_*$ matrix of the co variances evaluated at all pairs of training and input points, and similarly for the other matrixes $K(X, X)$, $K(X_*, X_*)$, and $K(X_*, X)$.

The posterior distribution of the predicted value f_* is given in (6) according to conditioning Gaussian distribution on the observations [11].

$$f_* \mid X, y, X_* \sim N(m_*, \mathrm{cov}(f_*)) \tag{6}$$

Where m_* and $\mathrm{cov}(f_*)$ are the mean and variance matrix of f_* respectively, and they can be described as the follows:

$$\begin{aligned} m_* &= E[f_* \mid X, y, X_*] \\ &= K(X_*, X)(K(X, X) + \sigma_n^2 I)^{-1} y \end{aligned} \tag{7}$$

$$\begin{aligned} \mathrm{cov}(f_*) = {} & K(X_*, X_*) - \\ & K(X_*, X)(K(X, X) + \sigma_n^2 I)^{-1} K(X, X_*) \end{aligned} \tag{8}$$

From the above equations, we can see that the covariance function plays crucial role in a Gaussian process predictor as it compromises our knowledge about the relationship between the input and the output.

B. Determination of ultra parameters

Covariance function acts as a model in Gaussian process regression, once it is set; the training process is mainly focusing on determination of ultra-parameters. There exists a variety of covariance functions, each of these functions have a number of free parameters called ultra–parameters, which specify the property of the covariance function and should be determined.

The two covariance functions we used in this paper are given below:

1) Matérn covariance function [12]

$$\begin{aligned} & K(x, x') = \sigma_f^2 f_d(r_d) \exp(-r_d) \\ & r_d = \sqrt{\frac{d}{l^2}(x - x')^T (x - x')} \\ & f_1(t) = 1,\ f_3(t) = 1 + t,\ f_5(t) = f_3(t) + \frac{t^2}{3} \end{aligned} \tag{9}$$

2) Periodic covariance function [12]

$$K(x, x') = \sigma_f^2 \exp\left(-\frac{2}{l^2} \sin^2 \left[\frac{\omega}{2\pi} (x - x') \right] \right) \tag{10}$$

While the noise variance covariance function is constant:

$$K(x, x') = \sigma_n^2 \tag{11}$$

Where l is length-scale, σ_f^2 is signal variance, σ_n^2 is noise variance, ω is angular frequency. Parameters set $\theta = \{l, \sigma_f^2, \sigma_n^2, \omega\}$ are called ultra-parameters, generally determined by the maximum likelihood method. The likelihood functions of GP model with logarithmic is:

$$\begin{aligned} & \lg p(y \mid X, \theta) = \\ & \quad -\frac{1}{2} y^T K_y^{-1} y - \frac{1}{2} \lg |K_y| - \frac{n}{2} \lg 2\pi \end{aligned} \tag{12}$$

Then the partial derivative of θ is：

$$\begin{aligned} & \frac{\partial}{\partial \theta_j} \lg p(y \mid X, \theta) = \\ & \quad -\frac{1}{2} y^T K_y^{-1} \frac{\partial K}{\partial \theta_j} K^{-1} y - \frac{1}{2} tr(K^{-1} \frac{\partial K}{\partial \theta_j}) \end{aligned} \tag{13}$$

Where $K_y = K + \sigma_n^2 I$ is the covariance matrix for the noisy targets y .

Optimal ultra-parameters solution can be obtained by minimizing the partial derivative using the method of conjugate gradient or Newton method and other optimization methods. The mean and variance of prediction value f_* on test point X_* can be obtained respectively according to the (7) and (8).

III. AVAILABLE CAPACITY AND SOH PREDICTION WITH GPR

This section gives detailed analysis of experimental data through function fitting; In the light of the analysis of the

experimental data, now we can do the regression and prediction with the GPR method.

A. Experimental data and analysis

The experimental data come from the Center for Advanced Life Cycle Engineering (CALCE) in USA [13]. Two lithium ion batteries that numbered CS2-34 and CS2-37 are selected，each has nominal capacity of 1100mAh. They both underwent the same charging profile: the charging was carried out in a constant current (CC) mode at 0.5C until the battery voltage reached 4.2V and then continued in a constant voltage (CV) mode until the charge current dropped to 0.05A. Discharge was carried out at a constant current (CC) level of 0.5C and 1C respectively until the battery voltage fell to 2.7V. These data are divided into two parts: training part and test part， the training points which take up 80% of the total samples are for data fitting and regression, while the remaining test part is for prediction validation and error analysis.

In general, the battery capacity decreases with the cycle number, there exist always fluctuations in the tendency curves. Although they are of the same kind lithium ion batteries, they show some kind of different properties as for different discharge currents.

Given the experimental data we want to find the relationship between the cycle number and the corresponding available capacity, which can give us basic information for prediction and choosing the covariance functions for Gaussian process regression. In this circumstance, data fitting is always the choosing method. Data fitting uses functions that can describe the characteristics of the data curves; the task is to find the parameter values for the function which best fits the data. The least square method is employed to adjust the parameters included in the function. A variety of function families have been tried to do this, and three of them are among the best fitting candidates considering our experimental data: the polynomial, exponentials and the trigonometric functions.

To measure the degree of best fit with these three functions, the sum of squares (SSE) of the vertical deviations of all experimental data points, root mean squared error (RMSE) and coefficient of determination (R-square) are calculated and compared. Table I and II show the calculated results and comparisons applied on the two batteries.

TABLE I FITTING FUNCTIONS ERRORS WITH BATTERY CS2-34

CS2-34	polynomial	exponentials	trigonometric functions
SSE	0.0837	0.06596	0.07002
RMSE	0.01138	0.01014	0.01043
R-square	0.9872	0.9899	0.9893

Table II FITTING FUNCTIONS ERRORS WITH BATTERY CS2-37

CS2-37	polynomial	exponentials	trigonometric functions
SSE	0.07733	0.0405	0.04008
RMSE	0.01094	0.007948	0.007901
R-square	0.9654	0.9819	0.9821

According to the characteristics of the experimental data and fitting results, now we can choose the covariance functions for Gaussian process regression. As for the experimental data of both batteries are not smooth as the squared exponential covariance function implies, the Matérn covariance function is selected, which is a combination a polynomial and an exponential function; the periodic covariance function is also the candidate to account for the small fluctuation of data points. These two covariance functions are introduced in section II.

B. Available capacity and SOH prediction

Parameters which change significantly with cycle numbers are always been used as basis for providing an indication of the SOH, such as available capacity and cell resistance. In practice some people estimate the SOH from a single measurement of either the cell impedance or the cell capacity. In this paper, SOH of the battery is defined as the ratio of capacity available to the nominal capacity in standard conditions; the ratio is a reflection of the health status of the battery.

$$SOH = \frac{C_i}{C_0} \times 100\% \tag{14}$$

Where, C_0 is rated capacity, C_i is the discharge capacity of cycle i. In general, when capacity falls from 100% to 60~70% of the rated capacity, battery power cannot fulfill the needed requirements.

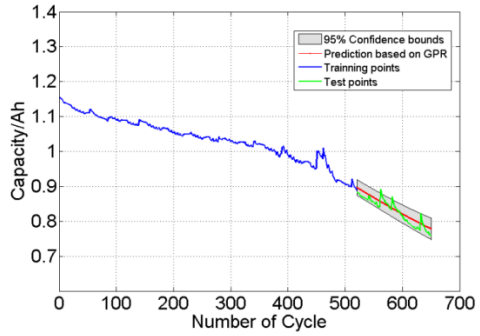
From A in section III, we have found that the two kinds of functions can interpret the relationship between the input and output variables. In parallel, there are two covariance functions with the similar forms: the Matérn and the periodic covariance function, the two related covariance functions are introduced in section II. Based on our experiences, the combination of two functions work well in prediction, so the final covariance function we use is given as follows, as a comparison the Matérn covariance function is also used.

$$\begin{aligned} K(x,x') &= K_{\text{Matérn}}(x,x') + K_{\text{Periodic}}(x,x') \\ &= \sigma_f^2 f_d(r_d)\exp(-r_d) + \sigma_f^2 \exp\left(-\frac{2}{l^2}\sin^2\left[\frac{\omega}{2\pi}(x-x')\right]\right) \end{aligned} \tag{15}$$

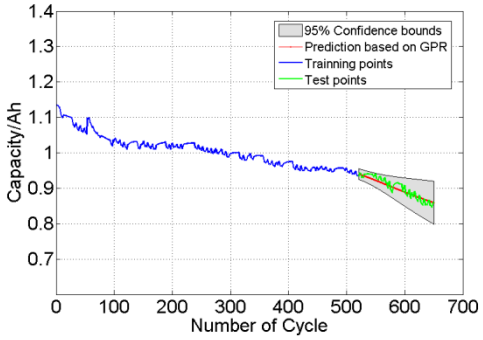
The mean function $\mathrm{m}_\phi : x \to R$ (with ultra-parameters ϕ) of a GP f is a scalar function defined over the whole domain x that computes the expected value $m(x) = E[f(x)]$ of f for the input x. Since the polynomial functions can interpret the tendency of the data curve, we use the polynomial mean function。

$$m(x) = \sum_d a_d{}^T \mathrm{x}^d \quad (a \in R^{D\times d}) \tag{16}$$

After determination of the main elements, the GPR method can run the available capacity regression and prediction. Firstly, the Matérn covariance function is applied and next the combination covariance function is used as the covariance function. The ultra-parameters are estimated by maximized the marginal likelihood which is introduced in section II. Fig. 1 and fig. 2 show the regression and prediction result.

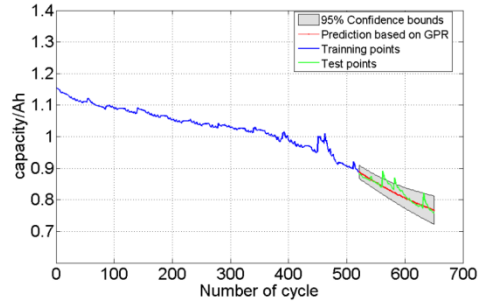


(a) Battery CS2-34

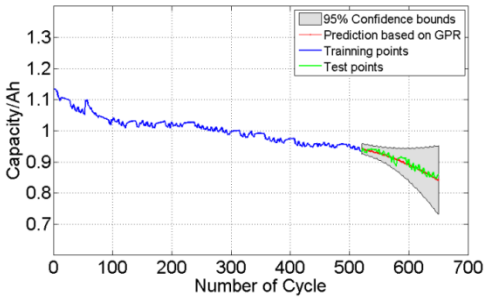


(b) Battery CS2-37

Fig. 1 GPR with Matérn covariance function



(a) Battery CS2-34



(b) Battery CS2-37

Fig.2 GPR with the combination of Matérn and periodic covariance functions

In order to evaluate the accuracy of these two covariance function applications, the following two indicators are chosen.

$$\sigma_{MAE} = \frac{1}{N}\sum_{i=1}^{N} | \bar{y}(i) - y(i) | \tag{17}$$

$$\sigma_{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (\bar{y}(i) - y(i))^2} \tag{18}$$

Where σ_{MAE} is mean absolute error; σ_{RMSE} is the root mean square error; N is the number of samples for prediction; $\bar{y}(i)$ is prediction results; $y(i)$ is the actual value. Table III and IV show the prediction errors of the two cases applied on the two batteries experimental data.

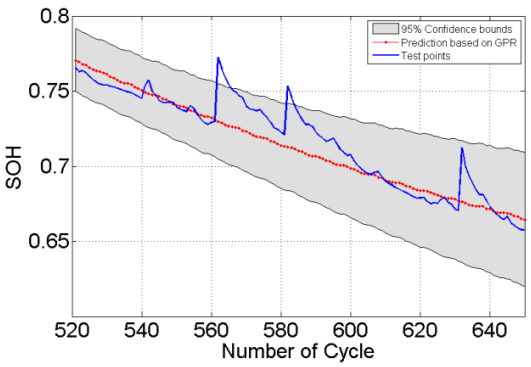
Table III PREDICTION ERRORS CONCERNING CS2-34

CS2-34	Matérn	Matérn+Periodic
MAE	0.01331	0.01067
RMSE	0.01494	0.01488

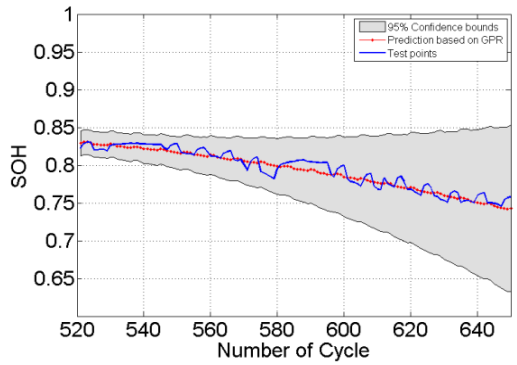
Table IV Prediction errors concerning CS2-37

CS2-37	Matérn	Matérn+Periodic
MAE	0.00932	0.00738
RMSE	0.01071	0.00896

From table III and IV, we find that both Matérn and the combined covariance function can fulfill the task of predictions, while the latter has the smaller errors in both batteries applications. With the predicted capacity data with cycle numbers, the SOH of lithium ion battery thus can be calculated. The expected cycle life can be predicted, SOH and remaining useful life are predicted as shown in fig. 3.



(a) Battery CS2-34



(b) Battery CS2-37

Fig.3.The SOH estimation results of battery cs-2-34 and cs2-37 based GPR

IV. CONCLUSIONS

Gaussian process regression and prediction is a flexible data-driven method, it can be applied in different batteries and situations theoretically. To make it more practical, the route of finding the covariance functions needs to be limited. In fact, the training data can give us basic information about the relationship between the input and the target values. By data fitting, the potential functions with similarities with the experimental data curves can be found and used to limit the pattern of the covariance functions. Our work with the real batteries cycling data verifies the effectiveness the method.

ACKNOWLEDGMENT

We gratefully acknowledge support from "Industry-University cooperation funding project under Grant No.40005015201206.

REFERENCES

[1] XM. Li. Research and design of lithium batteries management system. South China University of Technology 2012. (In Chinese)

[2] X. Hui. The research on SOH estimation methods of the lithium-ion battery. Jilin University 2013. (In Chinese)

[3] WX. Shen. State of available capacity estimation for lead-acid batteries in electric vehicles using neural network. Energy Conversion and Management 2007;48(2): 433-442.

[4] A. Nuhic , T. Terzimehic , T. Soczka-Guth , M. Buchholz, K. Dietmayer. Health diagnosis and remaining useful life prognostics of lithium-ion batteries using data-driven methods. Journal of Power Sources, 2013;239(0): 680-688.

[5] D .Wang, Q. Miao, M. Pecht. Prognostics of lithium-ion batteries based on relevance vectors and a conditional three-parameter capacity degradation model. Journal of Power Sources 2013;239(0): 253-264.

[6] DT Liu. Prognostics for state of health estimation of lithium-ion batteries based on combination Gaussian process functional regression. Microelectronics Reliability, 2013;53(6): 832-839.

[7] J.H. Heinbockel. Introduction to tensor calculus and continum mechanics, Victoria, B.C. Canada: Trafford Publishing. 2001; 14-31, ISBN 1-55369-1333-4

[8] J. Kang. The overview of particle filter technical based on Bayesian principles. Modern Radar 2004;01:34-36.

[9] XY. Deng, DH.Yang. The research of Mercer conditions in support vector regression with Ricker wavelet kernel. Chinese Journal of Geophysics 2009;09:2335-2344.

[10] ZK He, GB Zhao, XJ Zhao. A review of methods of Gaussian process regression, Control and Decision 2013;8: 1121-1129 .(In Chinese)

[11] B Sun, HT Yao, T Liu. Short-term wind speed prediction based on Gaussian process regression,Chinese Society for Electrical Engineering 2012;29: 104-109. (In Chinese)

[12] B.E. Olivares, M.A.C. Muñoz, M.E. Orchard, J.F. Silva. Particle-filtering-based prognosis framework for energy storage devices with a statistical characterization of state-of-health regeneration phenomena. Instrumentation and Measurement, IEEE Transactions on 2013; 62(2): 364-376.

[13] DT, Liu., JY Pang.,JB Zhou, Y Pang. Data-driven prognostics for lithium-ion battery based on Gaussian Process Regression. in Prognostics and System Health Management (PHM), 2012 IEEE Conference on. 2012. Beijing