MITSUBISHI ELECTRIC RESEARCH LABORATORIES
https://www.merl.com

Online State of Charge Estimation for Lithium-Ion Batteries Using Gaussian Process Regression

Ozcan, Gozde; Pajovic, Milutin; Sahinoglu, Zafer; Wang, Yebin; Orlik, Philip V.; Wada, Toshihiro

TR2016-136 October 24, 2016

Abstract

This paper presents an application of Gaussian process regression (GPR) to estimate a state of charge (SoC) of Lithium-ion (Li-ion) batteries with different kernel functions. One of the practical advantages of using GPR is that uncertainties in the estimates can be quantified, which enables reliability assessment of the SoC estimate. The inputs of GPR are voltage, current and temperature measurements of the battery and the output is an estimate of SoC. First, training is performed in which optimal hyperparameters of a kernel function are determined to model data properties. Then, the battery SoC is estimated online based on the trained model. The kernel function is the key element in the GPR model since it encodes the prior assumptions about the properties of the function being modeled. Therefore, the impact of kernel function selection on the estimation performance is analyzed using both simulated data and experimental data collected from a LiMn2O4/hardcarbon battery with a nominal capacity of 4.93Ah operating under constant charge and discharge currents.

IEEE Industrial Electronics Society (IECON) 2016

Mitsubishi Electric Research Laboratories, Inc.
201 Broadway, Cambridge, Massachusetts 02139

Online State of Charge Estimation for Lithium-Ion Batteries Using Gaussian Process Regression

Gozde Ozcan, Milutin Pajovic, Zafer Sahinoglu, Yebin Wang, Philip V. Orlik, and Toshihiro Wada

***Abstract*—This paper presents an application of Gaussian process regression (GPR) to estimate a state of charge (SoC) of Lithium-ion (Li-ion) batteries with different kernel functions. One of the practical advantages of using GPR is that uncertainties in the estimates can be quantified, which enables reliability assessment of the SoC estimate. The inputs of GPR are voltage, current and temperature measurements of the battery and the output is an estimate of SoC. First, training is performed in which optimal hyperparameters of a kernel function are determined to model data properties. Then, the battery SoC is estimated online based on the trained model. The kernel function is the key element in the GPR model since it encodes the prior assumptions about the properties of the function being modeled. Therefore, the impact of kernel function selection on the estimation performance is analyzed using both simulated data and experimental data collected from a LiMn2O4/hard-carbon battery with a nominal capacity of $4.93Ah$ operating under constant charge and discharge currents.**

***Index Terms*—Battery management system, Gaussian process regression, Lithium-ion battery, state of charge estimation.**

I. Introduction

Lithium-ion (Li-ion) batteries have been widely deployed as a major energy storage component in numerous applications including consumer electronics, residential rooftop solar photovoltaic systems, electric vehicles, smart grid systems and etc. The main advantages of Li-ion batteries over other types of batteries with different chemistries are low self-discharge rate, high cell voltage, high energy density, lightweight, long lifetime, and low maintenance [1]. One of the key parameters for assessing battery's state is the state of charge (SoC), which is defined as the percentage of available charge remaining in the battery. The SoC indicates when a battery should be recharged. Hence, it enables battery management systems to improve the battery life by protecting the battery from over-discharge and over-charge events. Accurate SoC estimation of a Li-ion battery is still a challenging task due to nonlinear battery dynamics as well as variations of operating conditions such as temperature.

There have been many studies addressing SoC estimation of Li-ion batteries using equivalent circuit models of a battery. For instance, the authors in [2] applied a Kalman filter (KF) to estimate the SoC of Li-ion batteries. The studies in [3], [4] mainly focused on online SoC estimation using an extended Kalman filter (EKF). In order to improve the accuracy, the joint estimation of model parameters and the SoC based on an iterated EKF was performed in [5]. The authors in [6] estimated both the SoC of Li-ion batteries using the EKF and the internal resistance which is directly correlated with the batterys state-of-health (SoH). However, the EKF linearizes the nonlinear dynamic system, which leads to inevitable linearization error. To address this problem, the authors in [7], [8] estimated SoC of Li-ion batteries using an unscented Kalman filter (UKF), which avoids model linearization at the expense of higher computational complexity.

Although aforementioned KF based SoC estimation methods provide acceptable performance, the estimation accuracy strongly depends on the chosen battery model and parameters. The authors in [9] showed the impact of parameter uncertainties on the SoC estimation. In recent years, there has been a growing interest in machine-learning based SoC estimation methods, which do not require a detailed physical knowledge of the battery and learn the nonlinear relationship between the SoC and measurable battery quantities such as voltage, current, temperature. Mainly, the authors in [10] proposed an SoC estimation method for Li-ion batteries using a multi-layer feedforward neural network (NN). The authors in [11] introduced a hybrid SoC estimation method by combining a radial basis function NN, an orthogonal least-squares algorithm and an adaptive genetic algorithm. In [12], a fuzzy NN based SoC estimation method was proposed. The authors in [13] applied the NN to find the state-space model of the SoC, which is then employed by the EKF for estimating the SoC of Li-ion batteries. Moreover, SoC estimation methods based on a support vector machine (SVM) were introduced in [14], [15].

Notably, none of the above mentioned machine-learning based SoC estimation methods quantify estimation uncertainty which is crucial for evaluating the reliability of SoC estimates. On the other hand, an SoC estimation method based on the GPR provides a particularly flexible and powerful framework to obtain predictive probability distribution of the SoC rather than just a point estimate of the SoC [16]. In particular, the GPR is trained offline by using voltage, current and temperature measurements of the battery, and then used to infer the SoC values. One of the main advantages of the GPR is analytically tractable inference with elegant closed-form expressions. In addition, it is essentially a data-driven approach and eliminates the need for knowledge of the physical model of a battery. Therefore, it can be easily applied to estimate the

G. Ozcan is with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244, USA. gozcan@syr.edu

M. Pajovic, Z. Sahinoglu, Y. Wang, and P. V. Orlik are with Mitsubishi Electric Research Laboratories, 201 Broadway, Cambridge, MA 02139, USA. {pajovic, zafer, yebinwang, porlik}@merl.com

T. Wada is with the Advanced Technology R&D Center, Mitsubishi Electric Corporation, 8-1-1, Tsukaguchi-honmachi, Amagasaki City, 661-8661, Japan. Wada.Toshihiro@bx.MitsubishiElectric.co.jp

SoC of other types of batteries with different chemistries. In this paper, we analyze the performance of the SoC estimation method for Lithium-ion (Li-ion) batteries based on GPR and identify the impact of kernel function selection on the estimation performance.

The rest of the paper is organized as follows: Section II gives a brief overview of the theory of GPR. Section III presents an SoC estimation method based on GPR. The simulation and experimental results are provided and discussed in Section IV. Finally, main conclusions are drawn in Section V.

II. Gaussian Process Regression

In this section, we briefly review the theory of GPR before introducing an SoC estimation method based on a GPR framework.

A. Fundamentals of Gaussian Process Regression

We have a training data set $\mathcal{D} = (\mathbf{X}, \mathbf{y})$ comprising D-dimensional N input vectors $\mathbf{X} = \{\mathbf{x}_n\}_{n=1}^{N}$, where $\mathbf{x}_n \in \mathbb{R}^D$, and the corresponding outputs $\mathbf{y} = \{y_n\}_{n=1}^{N}$, where $y_n \in \mathbb{R}$. In this setting, the input-output relationship is written as

$$y_n = f(\mathbf{x}_n) + \varepsilon_n, \tag{1}$$

where $f(.)$ is the underlying latent function and ε_n denotes zero-mean additive Gaussian noise with variance σ_n^2, i.e., $\varepsilon_n \sim \mathcal{N}(0, \sigma_n^2)$. It is assumed that $\{\varepsilon_n\}_{n=1}^{N}$ form an independent and identically distributed (i.i.d) sequence. The main objective is to model the underlying function $f(.)$ which maps the inputs, $\mathbf{X}$ to their corresponding output values, $\mathbf{y}$. The key assumption in GPR is that any set of function values follows a multivariate Gaussian distribution [17]

$$p(\mathbf{f}|\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) = \mathcal{N}(\mathbf{0}, \mathbf{K}). \tag{2}$$

Above, $\mathbf{f} = [f(\mathbf{x}_1), f(\mathbf{x}_2), \ldots, f(\mathbf{x}_n)]^T$ and $\mathbf{0}$ is a $N \times 1$ vector whose elements are all 0. In addition, $\mathbf{K}$ is a covariance matrix, whose entries $\mathbf{K}_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$ are the values of the kernel function evaluated for all pairs of training inputs.

In Table I, we list the kernel functions that are adopted in this work. More specifically, we consider the squared exponential (SE), Matèrn, rational quadratic (RQ) and quasi-periodic kernels. As it is seen from the table, the kernel functions are characterized by hyperparameters, Θ. In particular, ϑ_i for $i \in \{0, 1, 2\}$ denotes the signal variance and quantifies the variation of the underlying latent function from its mean; and l_d, ρ_d and η_d represent the characteristic length scale for each input dimension, which determines the relative importance of each input variable in estimating the target output. Smaller value of a characteristic length scale implies that the corresponding input dimension has more impact on the output, hence it is more relevant. For the quasi-periodic kernel, λ_d is the period length which determines the distance between repetitions of the function. Additionally, $\nu > 0$ and $\alpha > 0$ correspond to the smoothness parameter. As the smoothness parameter increases, the function becomes more smooth. For the Matèrn kernel, we consider a special case where the value of ν is not too high, i.e., $\nu = 3/2$, such that the covariance is given by

$$k_s(\mathbf{x}_i, \mathbf{x}_j) = \vartheta_1^2 \left(1 + \sqrt{3} \sum_{d=1}^{D} \left(\frac{x_{id} - x_{jd}}{\rho_d}\right)\right) \times \exp\left(-\sqrt{3} \sum_{d=1}^{D} \left(\frac{x_{id} - x_{jd}}{\rho_d}\right)\right),$$

where x_{id} and x_{jd} correspond to d-th element of vectors $\mathbf{x}_i$ and $\mathbf{x}_j$, respectively. The kernel function plays an important role in GPR since they encode the prior assumptions about the properties of the underlying latent function (i.e., smoothness, periodicity and non-stationary) that we are trying to model. More specifically, the SE and RQ kernels are appropriate for modelling the function which exhibits a smooth behavior whereas the Matèrn kernel allows less stringent prior assumption about the smoothness or differentiability of the function to be modeled [17]. In addition, a quasi-periodic kernel, which is obtained by multiplying a periodic kernel with the SE kernel, is capable of modeling a repeating structure that is not strictly periodic.

Recall that the output in (1) is assumed to be corrupted by an additive Gaussian noise with variance σ_n^2. Therefore, we incorporate this additive white Gaussian noise term into the aforementioned kernel functions as follows:

$$k(\mathbf{x}_i, \mathbf{x}_j) = k_s(\mathbf{x}_i, \mathbf{x}_j) + \sigma_n^2 \delta_{ij}, \tag{3}$$

where δ_{ij} denotes the Kronecker delta, which takes value 1 if and only if $i = j$ and 0 otherwise. In this setting, the distribution of $\mathbf{y}$, given the latent function values $\mathbf{f}$ and the input $\mathbf{X}$, is written as

$$p(\mathbf{y}|\mathbf{f}, \mathbf{X}) = \mathcal{N}(\mathbf{f}, \sigma_n^2 \mathbf{I}), \tag{4}$$

where $\mathbf{I}$ is an $N \times N$ identity matrix. By using (2) and (4), the marginal distribution of $\mathbf{y}$ can be found to be

$$p(\mathbf{y}|\mathbf{X}) = \int p(\mathbf{y}|\mathbf{f}, \mathbf{X}) p(\mathbf{f}|\mathbf{X}) d\mathbf{f} = \mathcal{N}(\mathbf{0}, \mathbf{K} + \sigma_n^2 \mathbf{I}). \tag{5}$$

Based on (5), the marginal log-likelihood of $\mathbf{y}$ can be written

$$\log p(\mathbf{y}|\mathbf{X}, \Theta) = -\frac{1}{2} \mathbf{y}^T (\mathbf{K} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{y} - \frac{1}{2} \log |\mathbf{K} + \sigma_n^2 \mathbf{I}| - \frac{N}{2} \log 2\pi, \tag{6}$$

where $|.|$ is the determinant of a matrix. The hyperparameters are optimized by maximizing the marginal log-likelihood function in (6). In this regard, the gradient of (6) with respect to the ith element of Θ is calculated as

$$\frac{\partial \log p(\mathbf{y}|\mathbf{X}, \Theta)}{\partial \theta_i} = -\frac{1}{2} \mathrm{tr}\left((\mathbf{K} + \sigma_n^2 \mathbf{I})^{-1} \frac{\partial (\mathbf{K} + \sigma_n^2 \mathbf{I})}{\partial \theta_i}\right) + \frac{1}{2} \mathbf{y}^T (\mathbf{K} + \sigma_n^2 \mathbf{I})^{-1} \frac{\partial (\mathbf{K} + \sigma_n^2 \mathbf{I})}{\partial \theta_i} (\mathbf{K} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{y},$$

which allows the use of any gradient-based optimization method to find the optimal values of the hyperparameters that maximize the marginal log-likelihood function in (6). The objective function is not necessarily convex so that the gradient

Kernel Functions	Hyperparameters, Θ
Squared exponential kernel: $k_s(\mathbf{x}_i,\mathbf{x}_j)=\vartheta_0^2\exp\Big[-\frac{1}{2}\sum_{d=1}^{D}\Big(\frac{x_{id}-x_{jd}}{l_d}\Big)^2\Big]$	$\Theta=[\vartheta_0,l_1,\ldots,l_D]^T$ ϑ_0^2:signal variance l_d: characteristic length scale
Matèrn kernel: $k_s(\mathbf{x}_i,\mathbf{x}_j)=\vartheta_1^2\frac{1}{\Gamma(\nu)2^{\nu-1}}\Big[\sqrt{2\nu}\sum_{d=1}^{D}\Big(\frac{x_{id}-x_{jd}}{\rho_d}\Big)\Big]^{\nu}K_\nu\Big(\sqrt{2\nu}\sum_{d=1}^{D}\Big(\frac{x_{id}-x_{jd}}{\rho_d}\Big)\Big)$ $\Gamma(.)$: Gamma function $K_\nu(.)$: modified Bessel function of the second kind	$\Theta=[\vartheta_1,\nu,\rho_1,\ldots,\rho_D]^T$ ϑ_1^2:signal variance, ν: smoothness parameter ρ_d: characteristic length scale
Rational quadratic kernel: $k_s(\mathbf{x}_i,\mathbf{x}_j)=\vartheta_2^2\Big(1+\frac{1}{2\alpha}\sum_{d=1}^{D}\Big(\frac{x_{id}-x_{jd}}{\eta_d}\Big)^2\Big)^{-\alpha}$	$\Theta=[\vartheta_2,\alpha,\eta_1,\ldots,\eta_D]^T$ ϑ_2^2:signal variance, α: smoothness parameter η_d: characteristic length scale
Quasi-periodic kernel: $k_s(\mathbf{x}_i,\mathbf{x}_j)=\vartheta_3^2\exp\Big[-2\sum_{d=1}^{D}\Big(\frac{\sin\left(\frac{\pi(x_{id}-x_{jd})}{\lambda_d}\right)}{\varphi_d}\Big)^2-\frac{1}{2}\sum_{d=1}^{D}\Big(\frac{x_{id}-x_{jd}}{\mu_d}\Big)^2\Big]$	$\Theta=[\vartheta_3,\lambda_1,\ldots,\lambda_D,\mu_1,\ldots,\mu_D]^T$ ϑ_3^2:signal variance, λ_d: period length φ_d,μ_d: characteristic length scale

TABLE I: List of kernel functions used in the GPR model.

based method may converge to a local optimum. A possible approach to alleviate this problem would be to initialize multiple gradient based searches and then to choose the optimal point which yields the largest marginal log-likelihood. It should be also noted that computation of the marginal log-likelihood function and its gradient involves an inversion of a matrix, $\mathbf{K}+\sigma_n^2\mathbf{I}$ with size $N\times N$, which requires a computational time of $O(N^3)$. Thus, a simple implementation of GPR is suitable for data sets with up to a few thousands training examples. For larger data sets, sparse approximations to regular GPR based on choosing a small representative subset of training samples can be efficiently applied [18].

After determining the optimal hyperparameters, we express the joint distribution of $\mathbf{y}$ and y_* as follows

$$p(\mathbf{y},y_*|\mathbf{X},\mathbf{x}_*,\Theta)=\mathcal{N}\left(\begin{bmatrix}\mathbf{0}\\0\end{bmatrix},\begin{bmatrix}\mathbf{K}+\sigma_n^2\mathbf{I} & \mathbf{k}_*\\ \mathbf{k}_*^T & k_{**}+\sigma_n^2\end{bmatrix}\right), \quad (7)$$

where $\mathbf{k}_*=[k(\mathbf{x}_1,\mathbf{x}_*),\ldots,k(\mathbf{x}_N,\mathbf{x}_*)]^T$ and $k_{**}=k(\mathbf{x}_*,\mathbf{x}_*)$. The main goal of GPR is to find the predictive distribution for a new input vector, $\mathbf{x}_*$. In this regard, by marginalizing the joint distribution (7) over the training data set output $\mathbf{y}$, we obtain the predictive distribution of test output, y_*, which is a Gaussian distribution, i.e., $p(y_*|\mathbf{X},\mathbf{y},\mathbf{x}_*,\Theta)=\mathcal{N}(\mu_*,\Sigma_*)$ with the mean and covariance

$$\mu_*=\mathbf{k}_*^T(\mathbf{K}+\sigma_n^2\mathbf{I})^{-1}\mathbf{y} \quad (8)$$

$$\Sigma_*=\sigma_n^2+k_{**}-\mathbf{k}_*^T(\mathbf{K}+\sigma_n^2\mathbf{I})^{-1}\mathbf{k}_*. \quad (9)$$

It can be observed from (8) that the mean μ_* of the predictive distribution is obtained as a linear combination of the noisy output $\mathbf{y}$, and it is effectively the estimate of the test output. Once the inversion of a matrix, $\mathbf{K}+\sigma_n^2\mathbf{I}$ is precomputed, the computational complexity of the testing stage is relatively low, that is $O(N)$, which makes the GPR highly suitable for online SoC estimation.

In addition, the variance of the predictive distribution in (9) is a measure of the uncertainty. By using (8) and (9), the $100(1-\alpha)\%$ confidence interval is computed as

$$[\mu_*-z_{(1-\alpha)/2}\Sigma_*,\mu_*+z_{(1-\alpha)/2}\Sigma_*], \quad (10)$$

where $\alpha\in[0,1]$ represents the confidence level and $z_{(1-\alpha)/2}$ is the critical value of the standard normal distribution. The confidence interval provides a range of values which is likely to contain the true value of the test output. In particular, smaller variance results in a narrower confidence interval, and hence indicates a more precise estimate of the test output. Also, GPR gives a predictive probability distribution which is one of the practical advantages of GPR over the SVM, NN and other non-probabilistic machine learning methods.

III. SoC Estimation Method Based on GPR

In this section, we present the details of the SoC estimation method using GPR. The SoC of the battery is defined as the ratio of the amount of energy presently stored in the battery to its maximum capacity [19]. In particular, the fully discharged battery has an SoC of 0% and SoC increases while the battery is being charged. Consequently, the fully charged battery reaches 100% SoC.

As shown in Fig. 1, the method mainly consists of two parts, i.e., training and estimation. In training stage, the optimal hyperparameters of the chosen kernel function are determined by using conjugate gradient method based on a training data set, $\mathcal{D}=(\{\mathbf{x}_n\}_{n=1}^N,\mathbf{y})$ where $\mathbf{x}_n$ includes the voltage, current and temperature of the battery and $\mathbf{y}$ contains the corresponding SoC values. It should be noted that the SoC values in the training data set are normalized to have zero mean by subtracting their sample mean. Then, online SoC estimation of the battery is carried out based on present voltage, current and temperature measurements of the battery. More specifically, the mean of the predictive distribution corresponds to the SoC estimate. The entire process is described in Algorithm 1.

IV. Results and Discussion

In this section, we analyze the performance of the SoC estimation method for Li-ion batteries based on GPR with the simulation and experimental data obtained from testing the battery under constant charge and discharge current. The implementation of GPR was based on the GPML toolbox [20]. Subsequently, we identify the impact of kernel function

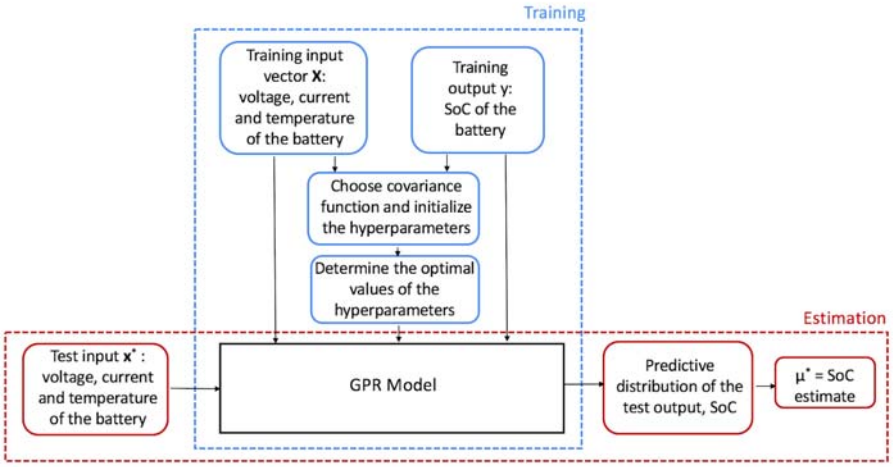


Fig. 1: Block diagram of the SoC estimation method using GPR.

selection on the estimation performance. The root mean square error (RMSE) and maximum absolute error (MAE) are chosen as the main performance metrics, which are respectively defined as follows

$$\text{RMSE} = \sqrt{\frac{1}{N_t}\sum_{i=1}^{N_t}(y_{*,i}^{\text{true}} - \hat{y}_{*,i}^{\text{est}})^2}, \tag{11}$$

$$\text{MAE} = \max_{i=1,\ldots,N_t} |y_{*,i}^{\text{true}} - \hat{y}_{*,i}^{\text{est}}|, \tag{12}$$

where N_t denotes the size of test data, $\mathbf{y}_*^{\text{true}}$ is a $1 \times N_t$ vector including SoC values of the test data and $\hat{\mathbf{y}}_*^{\text{est}}$ is a $1 \times N_t$ vector containing the estimated SoC values.

Algorithm 1 The flow chart of SoC estimation method

1: **Traning part:**
2: **Step 1:** Obtain the training data set, $\mathcal{D} = (\mathbf{X}, \mathbf{y})$, where $\mathbf{X}$ includes the voltage, current and temperature measurements of the battery, and $\mathbf{y}$ contains the corresponding SoC values.
3: **Step 2:** Initialize the hyperparameters of the given kernel function.
4: **Step 3:** Find the optimal hyperparameters that minimize the negative marginal log-likelihood function (equivalently maximize the marginal log-likelihood function) by using the conjugate gradient method.
5: **Estimation part:**
6: Obtain the mean and the variance of predictive distribution given optimal hyperparameters, training dataset, $\mathcal{D}$, and test input $\mathbf{x}_*$ (i.e., present voltage, current and temperature measurement of the battery) as follows:

$$\mu_* = \mathbf{k}_*^T(\mathbf{K} + \sigma_n^2\mathbf{I})^{-1}\mathbf{y}$$
$$\Sigma_* = \sigma_n^2 + k_{**} - \mathbf{k}_*^T(\mathbf{K} + \sigma_n^2\mathbf{I})^{-1}\mathbf{k}_*,$$

where μ_* is the SoC estimate.

A. Simulation Results

1) Simulation Dataset: The simulated battery data was generated by using a simple battery model based on equivalent circuit model including thermal equation. The initial temperature was set to $10°C$ and the initial SoC was 0.5.

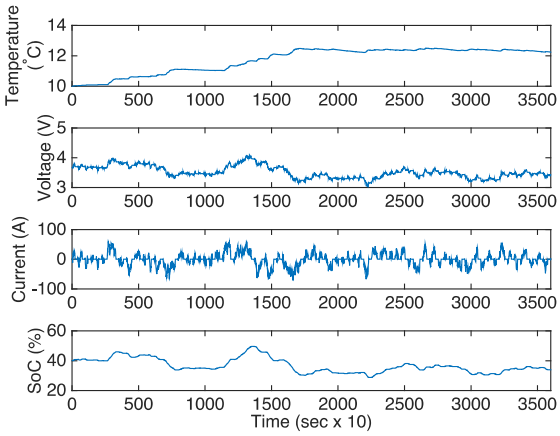
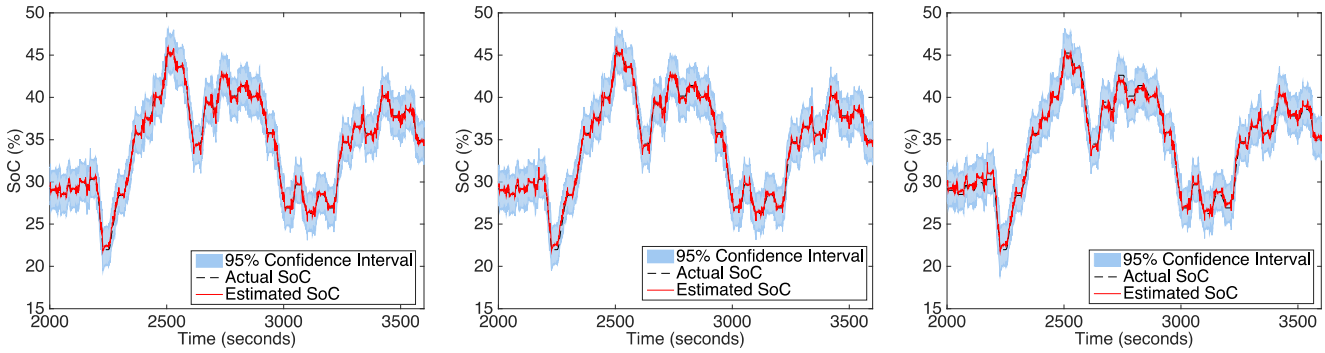


Fig. 2: Simulation dataset: temperature, voltage, current and battery SoC vs. time.

Fig. 2 shows the simulation dataset, which corresponds to a dynamic charging-discharging profile. In the figure, the negative values of the current indicate that the battery is being discharged. The GPR model is trained offline in which the optimal hyperparameters are determined for a given kernel function using the first 2000 samples of voltage, temperature and current measurements. The remaining 1600 samples are used to verify the performance of the SoC estimation method. We initialize hyperparameters in the training stage with ones.

2) Performance of SoC Estimation Method Based on GPR: In this subsection, we analyze the performance of SoC estimation using GPR in terms of RMSE and MAE. Fig. 3 displays the actual SoC, the estimated SoC values and 95% confidence interval for SE, RQ, and Matèrn kernels. The shaded blue area represents the 95% confidence interval. The corresponding RMSE and MAE values are listed in Table II.

It is observed from Fig. 3 that the selection of a kernel function has substantial impact on the estimation performance. In particular, the SE and RQ kernels provide better fit to the data compared to that attained with the Matèrn kernel



(a) (b) (c)

Fig. 3: SoC estimation based on GPR with (a) SE, (b) RQ, (c) Matèrn kernels.

Kernel Functions	RMSE (%)	MAE (%)
(a) SE	0.3808	1.8185
(b) RQ	0.3917	1.7467
(c) Matèrn	0.4965	2.0447

TABLE II: RMSE and MAE values of SoC estimation based on GPR with SE, RQ, Matèrn kernels.

since SoC varies smoothly over time. Therefore, SE and RQ kernels generally achieve higher accuracy with RMSE less than 0.4% and MAE less than 1.82%. The Matèrn kernel allows a less stringent prior assumption about smoothness of the function. Therefore, the Matèrn kernel yields slightly lower accuracy, where the RMSE and MAE are below 0.5% and 2.1%, respectively.

The optimal hyperparameters associated with each input variable enable us to infer the relative importance of the inputs. For example, in the case of GPR with the SE kernel, smaller values of the characteristic length scales imply that the corresponding input dimension is more important and relevant. The optimal values of the characteristic length scales for voltage, current and temperature are 1.1025, 278.27 and 109.76, respectively, which indicate that the voltage has more impact than the temperature, and the temperature has more impact than the current on the SoC estimate. The same relative importance order is observed for other two kernel functions.

B. Experimental Results

1) Experimental Dataset: The dataset was collected from a LiMn2O4/hard-carbon battery with a nominal capacity of 4.93 Ah in the Advanced Technology R&D Center, Mitsubishi Electric Corporation. In particular, five consecutive cycles of charging and discharging at 10 C-rates were performed using a rechargeable battery test equipment produced by Fujitsu Telecom Networks. Charging and discharging currents are expressed in terms C-rate. In this case, 10 C-rate corresponds to a current of 49.3 A and the battery will deliver its rated capacity for 6 minutes. The reference (true) SoC was obtained using the Coulomb counting method. The sampling period was chosen to be 1 second.

In Fig. 4, we display the temperature, voltage, current and SoC values as a function of time. The experimental dataset corresponds to a constant charge-discharge current profile.

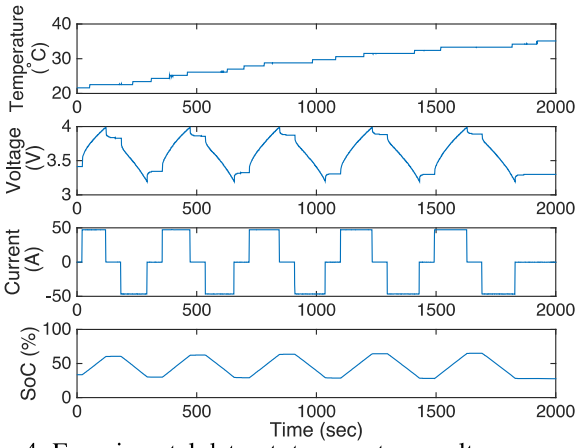


Fig. 4: Experimental dataset: temperature, voltage, current and battery SoC vs. time.

The first 1050 samples are used as training data to find the optimal values of the hyperparameters while the remaining 950 samples are used as testing data to verify the performance of the SoC estimation method.

2) Performance of SoC Estimation Method Based on GPR: In Fig. 5, we display the actual SoC and the estimated SoC values attained with the GPR for SE, RQ, quasi-periodic kernels. The resulting RMSE and MAE values are listed in Table III. In this case, the quasi-periodic kernel is well suited to the simulated dataset since it can capture the increasing and periodic trend but where the periods are not exactly identical. Hence, GPR with the quasi-periodic kernel achieves the highest accuracy (namely, RMSE=1.0648%, MAE=2.7701%). On the other hand, SE and RQ kernels still provide resonable SoC estimates with RMSE below 1.54% and MAE below 4.9%.

Kernel Functions	RMSE (%)	MAE (%)
(a) SE	1.5383	4.8503
(b) RQ	1.1803	3.7641
(c) Quasi-periodic	1.0648	2.7701

TABLE III: RMSE and MAE values of SoC estimation based on GPR with SE, RQ, quasi-periodic kernels.

It is also observed that we have a higher uncertainty, hence a larger confidence interval when the difference between the actual and the estimated SoC values is higher whereas accurate SoC estimates result in lower uncertainty, thus smaller

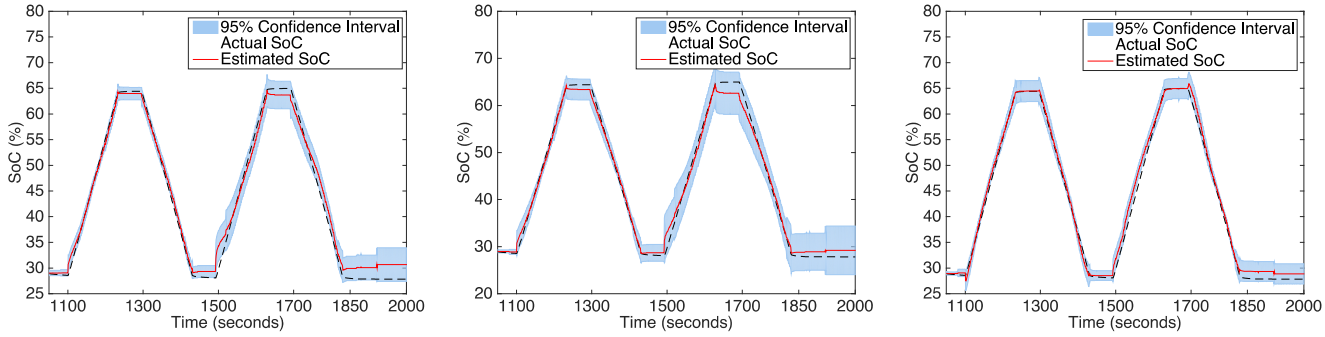


Fig. 5: SoC estimation based on GPR with (a) SE, (b) RQ, (c) quasi-periodic kernels.

confidence interval. This uncertainty characterization is one of the key advantages of the GPR-based methods over non-probabilistic machine learning methods such as the SVM, NN.

Our method is compared on the experimental dataset with finely tuned iterated EKF in [5] using third-order battery model in [21], which is, to our knowledge, one of the most accurate SoC estimation methods, and similar performance (i.e., RMSE=1.2% and MAE=2.16%) is observed. Nevertheless, we emphasize that in comparison with the benchmark, our method delivers one-time shot SoC estimates without accounting for SoC correlation in time, the development of which is our next research item.

V. Conclusion

In this paper, we have analyzed the performance of an SoC estimation method for Li-ion batteries based on GPR using different kernel functions. We have evaluated the impact of kernel function selection on the estimation performance. Simulation and experimental results reveal that the GPR provides accurate SoC estimates with RMSE less than 0.4% and MAE less than 1.82% when the SE kernel is used for a dynamic charging-discharging profile, and RMSE is below 1.1% and MAE is below 2.8% when the quasi-periodic kernel is used for a constant charge-discharge current profile. We have also provided uncertainty representation through 95% confidence interval, which enables us to evaluate the reliability of the SoC estimation. Moreover, we have identified the relative importance of the input variables on the estimation performance. In the future work, we will address the impact of battery aging by incorporating the battery capacity degradation into the SoC estimation method based on GPR framework.

References

[1] V. Ramadesigan, P. W. C. Northrop, S. S. S. De, R. D. Braatz, and V. R. Subramanian, "Modeling and simulation of lithium-ion batteries from a systems engineering perspective," *Journal of the Electrochem. Soc.*, vol. 159, no. 3, pp. R31–R45, 2012.

[2] M. W. Yatsui and B. Hua, "Kalman filter based state-of-charge estimation for lithium-ion batteries in hybrid electric vehicles using pulse charging," in *Proc. IEEE Vehicle Power and Propulsion Conf. (VPPC)*, Sep. 2011, pp. 1–5.

[3] O. Barbarisi, F. Vasca, and L. Glielmo, "State of charge Kalman filter estimator for automotive batteries," *Journal of the Electrochem. Soc.*, vol. 14, no. 3, pp. 267–275, Mar. 2006.

[4] Z. Chen, Y. Fu, and C. C. Mi, "State of charge estimation of Lithium-ion batteries in electric drive vehicles using extended Kalman filtering," *IEEE Trans. on Vech. Tech.*, vol. 62, no. 3, pp. 1020–1030, Mar. 2013.

[5] H. Fang, Y. Wang, Z. Sahinoglu, T. Wada, and S. Hara, "State of charge estimation for Lithium-ion batteries: an adaptive approach," *Control Eng. Pract.*, vol. 25, pp. 45–54, Apr. 2014.

[6] A. E. Mejdoubi, A. Oukaour, H. Chaoui, H. Gualous, J. Sabor, and Y. Slamani, "State-of-charge and state-of-health Lithium-ion batteries' diagnosis according to surface temperature variation," *IEEE Trans. on Ind. Electron.*, vol. 63, no. 4, pp. 2391–2402, Apr. 2016.

[7] W. He, N. Williard, C. Chen, and M. Pecht, "State of charge estimation for electric vehicle batteries using unscented kalman filtering," *Microelectronics Reliability*, vol. 53, no. 6, pp. 840–847, June 2013.

[8] S. Santhanagopalan and R. E. White, "State of charge estimation using an unscented filter for high power lithium ion cells," *Int. Journal of Energy Research*, vol. 34, no. 2, pp. 152–163, 2010.

[9] C. Zhang, L. Y. Wang, X. Li, W. Chen, G. G. Yin, and J. Jiang, "Robust and adaptive estimation of state of charge for Lithium-ion batteries," *IEEE Trans. on Ind. Electron.*, vol. 62, no. 8, pp. 4948–4957, Aug. 2015.

[10] W. He, N. Williard, C. Chen, and M. Pecht, "State of charge estimation for Li-ion batteries using neural network modeling and unscented Kalman filter-based error cancellation," *Int. Journal of Electrical Power & Energy Systems*, vol. 62, pp. 783–791, Nov. 2014.

[11] W. Y. Chang, "Estimation of the state of charge for a LFP battery using a hybrid method that combines a RBF neural network, an OLS algorithm and AGA," *Int. Journal of Eletrical Power & Energy Systems*, vol. 53, pp. 603–611, Dec. 2013.

[12] Y. S. Lee, W.-Y. Wang, and T. Y. Kuo, "Soft computing for battery state-of-charge (BSOC) estimation in battery string systems," *IEEE Trans. on Ind. Electron.*, vol. 55, no. 1, pp. 229–239, Jan. 2008.

[13] M. Charkhgard and M. Farrokhi, "State-of-charge estimation for Lithium-ion batteriesusing neural networks and EKF," *IEEE Trans. on Ind. Electron.*, vol. 57, no. 12, pp. 4178–4187, Dec. 2010.

[14] J. A. Antón, P. J. G. Nieto, F. C. Juez, F. S. Lasheras, M. G. Vega, and M. R. GutiÉrrez, "Battery state-of-charge estimator using the SVM technique," *Applied Mathematical Modelling*, vol. 37, no. 9, pp. 6244–6253, May 2013.

[15] H. Sheng and J. Xiao, "Electric vehicle state of charge estimation: Nonlinear correlation and fuzzy support vector machine," *Journal of Power Sources*, vol. 281, pp. 131–137, May 2015.

[16] G. Ozcan, M. Pajovic, Z. Sahinoglu, Y. Wang, P. V. Orlik, and T. Wada, "Battery state of charge estimation based on regular/recursive Gaussian process regression," *IEEE Trans. on Ind. Electron.*, submitted.

[17] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press, 2006.

[18] J. Q.-Candela and C. E. Rasmussen, "A unifying view of sparse approximate Gaussian process regression," *Journal of Machine Learning Research*, vol. 6, pp. 1939–1959, Dec 2005.

[19] S. Piller, M. Perrin, and A. Jossen, "Methods for state-of-charge determination and their applications," *Journal of Power Sources*, vol. 96, no. 1, pp. 113–120, June 2001.

[20] C. E. Rasmussen and H. Nickisch, "Gaussian processes for machine learning (gpml) toolbox," *The Journal of Machine Learning Research*, vol. 11, pp. 3011–3015, 2010.

[21] T. Kim, Y. Wang, Z. Sahinoglu, T. Wada, S. Hara, and W. Qiao, "A rayleigh quotient-based recursive total-least-squares online maximum capacity estimation for Lithium-ion batteries," *IEEE Trans. Energy Conversion*, vol. 30, no. 3, pp. 842–851, Sept. 2015.