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

Online Battery State-of-Charge Estimation Based on Sparse Gaussian Process Regression

Ozcan, G.; Pajovic, M.; Sahinoglu, Z.; Wang, Y.; Orlik, P.V.

TR2016-097 July 2016

Abstract

This paper presents a new online method for state-of charge (SoC) estimation of Lithium-ion (Li-ion) batteries based on sparse Gaussian process regression (GPR). Building upon sparse approximation of the regular GPR, the proposed method is computationally more efficient. The battery SoC is estimated based on measured voltage, current and temperature. The accuracy of the proposed method is verified using LiMn2O4/hardcarbon battery data collected from a constant-current discharge test. In addition, the estimation performance of the proposed method is compared with a SoC estimation method using regular GPR with different covariance functions.

IEEE Power & Energy Society General Meeting

Online Battery State-of-Charge Estimation Based on Sparse Gaussian Process Regression

Gozde Ozcan
Department of Electrical Engineering and Computer Science
Syracuse University
Syracuse, NY 13244, USA
gozcan@syr.edu

Milutin Pajovic, Zafer Sahinoglu, Yebin Wang, and Philip V. Orlik
Mitsubishi Electric Research Laboratories
201 Broadway, Cambridge, MA 02139, USA
{pajovic, zafer, yebinwang, porlik}@merl.com

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

***Abstract*—This paper presents a new online method for state-of charge (SoC) estimation of Lithium-ion (Li-ion) batteries based on sparse Gaussian process regression (GPR). Building upon sparse approximation of the regular GPR, the proposed method is computationally more efficient. The battery SoC is estimated based on measured voltage, current and temperature. The accuracy of the proposed method is verified using LiMn2O4/hard-carbon battery data collected from a constant-current discharge test. In addition, the estimation performance of the proposed method is compared with a SoC estimation method using regular GPR with different covariance functions.**

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

I. Introduction

State-of-Charge (SoC) estimation is a fundamental component for battery management system, which is used to describe the actual charging status of the battery. Accurate SoC estimates are particularly important not only to avoid permanent damage from over discharge and over charge events, but also to improve the life-time of the battery by optimizing the energy usage.

Different approaches have been developed to determine the SoC of the battery, which are commonly divided into two categories: electrochemical model-based methods and data-driven methods. The model-based methods applied Kalman filter (KF) [1], extended KF [2], [3] and unscented KF [4], [5] to estimate the SoC of the battery with an improved accuracy. However, Kalman filter based SoC estimation methods require precise battery models in order to achieve higher level of the estimation accuracy at the expense of high computational complexity. Therefore, data-driven methods have been proposed to eliminate the dependency of SoC estimation on physical battery models. The authors in [6] estimated the SoC of Lithium-ion batteries using neural network (NN) based on current, voltage and temperature of the battery. The estimation accuracy was further improved by applying unscented KF to the artificial NN output, which reduces the noise in SoC estimates. SoC estimation methods based on support vector machine were developed in [7], [8], [9]. Recently, in [10], relevance vector machine, which uses sparse Bayesian learning was employed to estimate the the capacity of Li-ion battery as a function of voltage and current measurements during charge.

In this paper, we apply sparse Gaussian process regression (GPR) to model the relationship between voltage, current, temperature and the SoC of the battery. After being trained offline, sparse GPR is used to estimate the SoC. Instead of using all training dataset as in regular GPR, a subset of training data points, called inducing points are used for training the regression model. Therefore, the proposed method significantly reduces the computational complexity when the size of the training dataset exceeds a few thousand. To the best of our knowledge, this is the first work which applies sparse GPR for estimating the SoC of the battery. The performance comparison between SoC estimation method based on regular GPR in [11] and the proposed method is carried out. The effects of covariance functions on the estimation accuracy are further analyzed. In order to assess the reliability of the SoC estimates, we provide uncertainty representation owing to non-parametric probabilistic modeling of the proposed method.

II. Preliminaries

In this section, we first briefly review the fundamentals of regular GPR and sparse GPR before introducing the SoC estimation method.

A. Regular Gaussian Process Regression

Let $\mathcal{D} = \{(\mathbf{x_1}, y_1), \ldots, (\mathbf{x_N}, y_N)\}$ denote a training dataset in which $\mathbf{x}_i$ where $i \in \{1, \ldots, N\}$ corresponds to a D-dimensional training input vector and y_n represents a scalar training output. We assume that there is an underlying latent function $f(.)$, which maps the inputs, $\mathbf{x}_n$, to their corresponding output values, y_n, as follows

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

Above, ε_n denotes additive Gaussian noise with mean zero and variance σ_n^2, i.e., $\varepsilon_n \sim \mathcal{N}(0, \sigma_n^2)$ and $\{\varepsilon_n\}_{n=1}^N$ is identically distributed and independent of each other. It is assumed that

$\mathbf{f} = [f(\mathbf{x}_1), f(\mathbf{x}_2), \ldots, f(\mathbf{x}_n)]^T$ is distributed according to a multivariate Gaussian distribution

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

where $\mathbf{0}$ is a $N \times 1$ vector with all 0 entries and $\mathbf{K}$ is a covariance matrix whose elements $\mathbf{K}_{i,j} = k(\mathbf{x}_i, \mathbf{x}_j)$ are determined by evaluating the covariance function at each pair of training inputs. The prior assumptions about the properties of the latent function (i.e., smoothness, periodicity and non-stationary) are encoded in the covariance function. In this work, we use Matèrn and rational quadratic (RQ) covariance functions listed in Table I. In order to represent different structures of the dataset, we also consider the sum of Matèrn and RQ covariance functions, which is again a valid covariance function. As it is seen from the table, the covariance functions depend on hyperparameters, Θ, where σ_1 and σ_2 quantify variation of the underlying latent function from its mean; and the characteristic length scales, ρ_d and η_d determine the relative importance of the input variables in estimating the target output. In particular, a larger values of the characteristic length scale indicates that the corresponding input variable has a smaller impact on estimation of the target output, hence it is less relevant.

We include the additive noise component of the output in (1) into the covariance functions listed in Table I as follows:

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

where $\delta_{i,j}$ denotes the Kronecker delta, i.e., $\delta_{i,j}$ is 1 if $i = j$ and 0 otherwise.

Then, the distribution of $\mathbf{y}$ conditioned on the values of the latent function $\mathbf{f}$ and the input $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$, is given by

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

where $\mathbf{y} = \{y_1, \ldots, y_N\}$ and $\mathbf{I}$ is an $N \times N$ identity matrix. By using (2) and (4) and integrating over the latent function values $\mathbf{f}$, the marginal distribution of $\mathbf{y}$ is obtained as

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

From (5), the marginal log-likelihood of $\mathbf{y}$ is expressed in the following:

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

where $|.|$ is the determinant of a matrix. The optimal values of the hyperparameters are determined by maximizing the marginal log-likelihood function in (6). Therefore, we calculate the gradient of (6) with respect to the hyperparameters as follows:

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

With the above characterization, any gradient-based optimization method can be employed in order to find the maximum value of the marginal log-likelihood function in (6), which is in general a nonconvex function of the hyperparamaters. Therefore, a gradient based method may converge to a stationary point. In order to alleviate this issue, gradient-based optimization can be performed with different initial conditions; and the optimal hyperparameters having the largest marginal log-likelihood can be chosen. After finding the optimal hyperparameters, the joint distribution of the training outputs $\mathbf{y}$ and the test output y_* can be written as

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

where the asterisk, $*$ is used as a shorthand for f_*, which is the corresponding latent function value at the test input. Based on the training dataset $\mathcal{D}$, and given a new input vector, $\mathbf{x}_*$, the main objective of GPR is to find the predictive distribution of the corresponding output y_*. Hence, we can obtain the predictive distribution of the output, y_* by marginalizing the joint distribution (7) over the training dataset output $\mathbf{y}$, which results in a Gaussian distribution as follows

$$p(y_*|\mathbf{X}, \mathbf{y}, \mathbf{x}_*, \Theta) = \mathcal{N}(\mu_*, \Sigma_*) \tag{8}$$

where

$$\mu_* = \mathbf{K}_{*,\mathbf{f}} (\mathbf{K}_{\mathbf{f},\mathbf{f}} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{y} \tag{9}$$

$$\Sigma_* = \sigma_n^2 + K_{*,*} - \mathbf{K}_{*,\mathbf{f}} (\mathbf{K}_{\mathbf{f},\mathbf{f}} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{K}_{\mathbf{f},*}, \tag{10}$$

where μ_* gives the estimate of the test output, which is a linear combination of the noisy output $\mathbf{y}$, and the variance Σ_* provides a measure of the uncertainty in the estimate of the test output. Note that regular GPR takes $\mathcal{O}(N^3)$ time due to the inversion of the $N \times N$ matrix $\mathbf{K} + \sigma_n^2 \mathbf{I}$. Once the inversion is performed, estimation costs $\mathcal{O}(N)$ and $\mathcal{O}(N^2)$ for calculating the mean μ_* and the variance, Σ_*, respectively. Therefore, when the size of the training dataset, N exceeds a few thousand, regular GPR based estimator becomes computationally prohibitive. In order to overcome this problem, sparse GPR is proposed, which we will review in the following subsection.

B. Sparse Gaussian Process Regression

The computational cost of a regular GPR is reduced by introducing inducing variables and modifying the joint prior distribution, $p(f_*, \mathbf{f})$. Let $\mathbf{u} = [u_1, \ldots, u_m]^T$ denote the inducing variables which correspond to a set of input locations $X_{\mathbf{u}}$ called inducing points. The inducing points are chosen as a subset of the data points. Given the inducing points, the joint prior distribution, $p(f_*, \mathbf{f})$ can be rewritten as

$$p(f_*, \mathbf{f}) = \int p(f_*, \mathbf{f}|\mathbf{u}) p(\mathbf{u}) d\mathbf{u}, \tag{11}$$

where $p(\mathbf{u}) = \mathcal{N}(\mathbf{0}, K_{\mathbf{u},\mathbf{u}})$. It is assumed that f_* and $\mathbf{f}$ are conditionally independent given $\mathbf{u}$ for the appromixation of

Covariance Functions	Hyperparameters, Θ
Matèrn covariance function : $k_s(\mathbf{x}_i,\mathbf{x}_j)=\sigma_1^2\frac{1}{\Gamma(\nu)2^{\nu-1}}\left[\sqrt{2\nu}\sum_{d=1}^{D}\left(\frac{x_{id}-x_{jd}}{\rho_d}\right)\right]^{\nu}K_\nu\left(\sqrt{2\nu}\sum_{d=1}^{D}\left(\frac{x_{id}-x_{jd}}{\rho_d}\right)\right)$	$\Theta=[\sigma_1,\nu,\rho_1,\ldots,\rho_D]^T$ σ_1^2:signal variance, ν:smoothness parameter ρ_d: characteristic length scale
Rational Quadratic covariance function: $k_s(\mathbf{x}_i,\mathbf{x}_j)=\sigma_2^2\left(1+\frac{1}{2\alpha}\sum_{d=1}^{D}\left(\frac{x_{id}-x_{jd}}{\eta_d}\right)^2\right)^{-\alpha}$	$\Theta=[\sigma_2,\alpha,\eta_1,\ldots,\eta_D]^T$ σ_2^2:signal variance, α: smoothness parameter η_d: characteristic length scale

TABLE I: Covariance functions and their corresponding hyperparameters.

$p(f_*,\mathbf{f})$ in the following [13]

$$p(f_*,\mathbf{f})\approx q(f_*,\mathbf{f})=\int q(f_*|\mathbf{u})q(\mathbf{f}|\mathbf{u})p(\mathbf{u})d\mathbf{u}. \tag{12}$$

Subsequently, it is assumed that the training conditional $q(\mathbf{f}|\mathbf{u})$ is fully independent and the test conditional remains exact as

$$\begin{aligned} q(\mathbf{f}|\mathbf{u}) &= \prod_{n=1}^{N} p(f_n|\mathbf{u}) \\ &= \mathcal{N}(K_{\mathbf{f},u}K_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{u},\mathrm{diag}[K_{\mathbf{f},\mathbf{f}}-Q_{\mathbf{f},\mathbf{f}}]), \qquad (13)\\ q(f_*|\mathbf{u}) &= p(f_*|\mathbf{u}), \qquad (14)\end{aligned}$$

where diag$[A]$ denotes the diagonal matrix in which all of the diagonal elements equal the corresponding elements of A and other elements are zero. By inserting above distributions into (12) and integrating over $\mathbf{u}$, the joint prior is given by

$$q(\mathbf{f},f_*)=\mathcal{N}\left(\mathbf{0},\begin{bmatrix}Q_{\mathbf{f},\mathbf{f}}-\mathrm{diag}[Q_{\mathbf{f},\mathbf{f}}-K_{\mathbf{f},\mathbf{f}}] & Q_{\mathbf{f},*}\\ Q_{*,\mathbf{f}} & K_{*,*}\end{bmatrix}\right) \tag{15}$$

where $Q_{\mathbf{a},\mathbf{b}}=K_{\mathbf{a},\mathbf{u}}K_{\mathbf{u},\mathbf{u}}^{-1}K_{\mathbf{u},\mathbf{b}}$ is a low-rank matrix (i.e., rank M). Using the above joint prior distribution, the predictive distribution is obtained as

$$q(y_*|\mathbf{X},\mathbf{y},\mathbf{x}_*,\Theta)=\mathcal{N}(\tilde{\mu}_*,\tilde{\Sigma}_*) \tag{16}$$

where

$$\tilde{\mu}_*=K_{*,\mathbf{u}}\Omega K_{\mathbf{u},\mathbf{f}}\Lambda^{-1}\mathbf{y} \tag{17}$$

$$\tilde{\Sigma}_*=\sigma_n^2+K_{*,*}-Q_{*,*}+K_{*,\mathbf{u}}\Omega K_{\mathbf{u},*}. \tag{18}$$

Above, $\Omega=(K_{\mathbf{u},\mathbf{u}}+K_{\mathbf{u},\mathbf{f}}\Lambda^{-1}K_{\mathbf{f},\mathbf{u}})^{-1}$ and $\Lambda=\mathrm{diag}[K_{\mathbf{f},\mathbf{f}}-Q_{\mathbf{f},\mathbf{f}}+\sigma_n^2 I]$. It is seen that the only matrix requiring inversion is the $N\times N$ diagonal matrix Λ, which yields a significant reduction in computational complexity. The computational cost of training becomes $\mathcal{O}(NM^2)$ that is linear in N and a larger M leads to better accuracy at the expense of increased computational requirements. Also, testing time complexity is $\mathcal{O}(M)$ and $\mathcal{O}(M^2)$ for calculating the mean and the variance, respectively.

III. SoC Estimation Method Based on sparse GPR

In this section, we first provide a definition of the SoC of a battery before introducing a new SoC estimation method using sparse GPR.

The SoC at time t is defined as

$$\mathrm{SoC}(t)=\frac{Q(t)}{Q_r}\times 100\%, \tag{19}$$

where $Q(t)$ represents the residual capacity of the battery at time t, and $Q(t)\in[0,Q_r]$. Also, Q_r denotes the maximum amount of charge that can be drawn for the battery in terms of the ampere-hours (Ah). It should be noted that SoC $\in[0\%,100\%]$, where 0% and 100% indicate the fully discharged and the fully charged states, respectively.

We now give the details of the proposed SoC estimation method based on sparse GPR. As shown in Fig. 1, voltage, current and temperature measurements of the battery are the inputs to the sparse GPR and the output is the estimated SoC. The proposed method consists of training and estimation parts.

Fig. 1: SoC estimation using sparse GPR.

For training, SoC values in the training dataset are normalized to have a zero mean by subtracting their sample mean. The optimal hyperparameters of the chosen covariance function are obtained by employing a conjugate gradient method. Consequently, online SoC estimation of a battery is performed based on voltage, current and temperature measurements of the battery and the optimal hyperparameters. In particular, the mean of the predictive distribution gives the SoC estimate. In order to represent the uncertainty in the estimates, the variance of the predictive distribution is used to construct a confidence interval as follows

$$[\tilde{\mu}_*-z_{(1-\alpha)/2}\tilde{\Sigma}_*,\tilde{\mu}_*+z_{(1-\alpha)/2}\tilde{\Sigma}_*], \tag{20}$$

where $\alpha\in[0,1]$ denotes the confidence level and $z_{(1-\alpha)/2}$ is the critical value of the standard normal distribution. The confidence interval returns a range of values which likely includes the true value of the test output. As the variance decreases, the confidence interval gets narrower, which implies

a more accurate estimate of the test output. The SoC estimation procedure is described in more detail in Algorithm 1.

Algorithm 1 can be modified for SoC estimation based on regular GPR [11] in such a way that the maximum log-likelihood function in (6) is used in training for finding the optimal hyperparameters and, $\tilde{\mu}_*$ and $\tilde{\Sigma}_*$ are replaced by the expressions in (9) and (10), respectively.

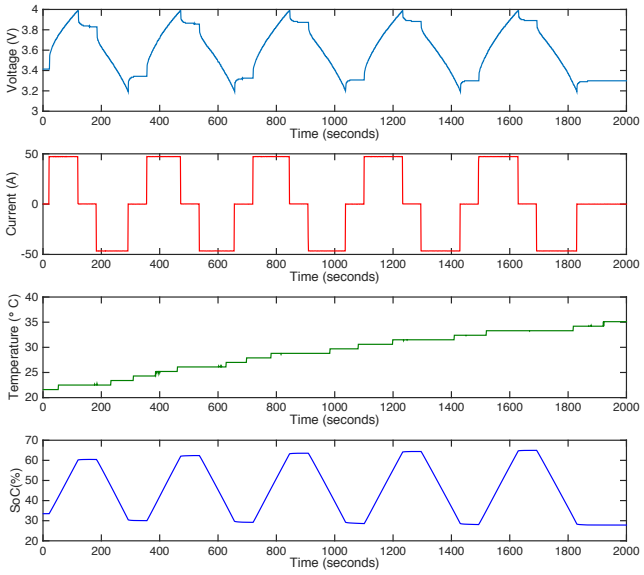


Fig. 2: Experimental dataset: voltage, current, temperature and SoC of the battery vs. time.

Algorithm 1 The flow chart of SoC estimation method using sparse GPR

1: **Traning part:**
2: **Step 1:** Determine the training dataset, $\mathcal{D} = (\mathbf{X}, \mathbf{y})$, where $\mathbf{X}$ contains voltage, current and temperature measurements of the battery, and $\mathbf{y}$ are the corresponding SoC values.
3: **Step 2:** Choose the covariance function and initialize the corresponding hyperparameters by setting Θ to zero.
4: **Step 3:** Apply a conjugate gradient method to find the optimal values of the hyperparameters that minimize the negative marginal log-likelihood function or equivalently maximize the marginal log-likelihood function.
5: **Estimation part:**
6: Obtain the mean and the variance of predictive distribution given optimal hyperparameters, training dataset, $\mathcal{D}$, test input $\mathbf{x}_*$ as follows:

$$\tilde{\mu}_* = K_{*,\mathbf{u}} \Omega K_{\mathbf{u},\mathbf{f}} \Lambda^{-1} \mathbf{y}$$
$$\tilde{\Sigma}_* = \sigma_n^2 + K_{*,*} - Q_{*,*} + K_{*,\mathbf{u}} \Omega K_{\mathbf{u},*},$$

where $\tilde{\mu}_*$ corresponds to the SoC estimate.

IV. Results and Discussion

In this section, we validate the proposed SoC estimation method based on sparse GPR using experimental data from a constant-current discharge test of a LiMn2O4/hard-carbon battery. We compare the performance of regular GPR and sparse GPR. Also, we analyze the effect of the covariance function selection on the estimation performance. The performance evaluation is carried out in terms of the root mean square error (RMSE) and maximum absolute error (MAE), which are defined, respectively as follows

$$\text{RMSE} = \sqrt{\frac{1}{N_t} \sum_{i=1}^{N_t} (y_{*,i}^{\text{true}} - \hat{y}_{*,i}^{\text{est}})^2}, \tag{21}$$

$$\text{MAE} = \max |\mathbf{y}_*^{\text{true}} - \hat{\mathbf{y}}_*^{\text{est}}|. \tag{22}$$

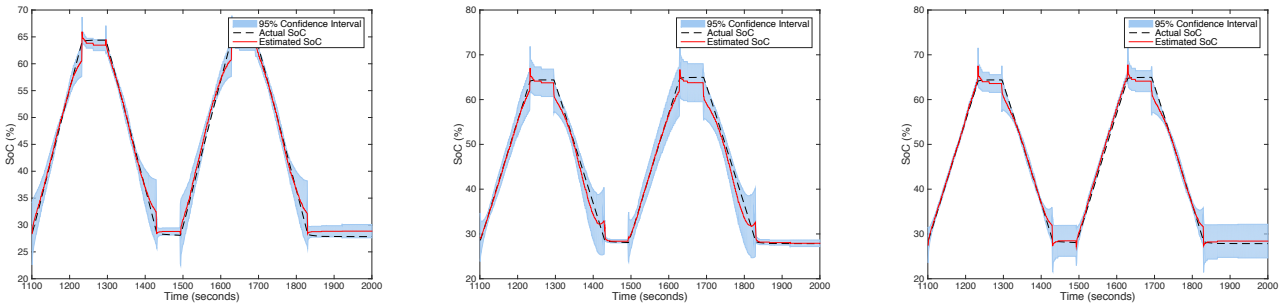
Particularly, N_t is the size of test data, $\mathbf{y}_*^{\text{true}}$ and $\hat{\mathbf{y}}_*^{\text{est}}$ are vectors of length N_t, which include SoC values of the test data and the estimated SoC values, respectively.

We first describe the experimental setup, and then present the SoC estimation results of the proposed method. Fig. 2 shows the experimental dataset including voltage, temperature, current and SoC of a LiMn2O4/hard-carbon battery with a nominal capacity of 4.93 Ah tested in the Advanced Technology R&D Center, Mitsubishi Electric Corporation. As it is seen from the figure, five consecutive cycles of charging and discharging at 10 C-rates were performed using a rechargeable battery test equipment produced by Fujitsu Telecom Networks. The sampling period was set to 1 second. The first 1100 samples are used for training in order to determine the optimal hyperparameters whereas the remaining 900 samples are used to verify the performance of the proposed method.

In Fig. 3, we plot the actual SoC, the estimated SoC values and 95% confidence interval for Matèrn, RQ, and sum of Matèrn and RQ covariance functions using sparse GPR. The shaded blue area represents the 95% confidence interval. 14 inducing points are randomly chosen from the training dataset. In Matèrn covariance function, ν is set to $3/2$. The RMSE and MAE values of SoC estimation using regular GPR and sparse GPR are listed in Table II.

Although slightly higher RMSE and MAE values are observed for SoC estimation based on sparse GPR compared to regular GPR, the computational cost is significantly reduced, i.e., time complexity is $\mathcal{O}(1100 \times 14^2)$ for sparse GPR and $\mathcal{O}(1100^3)$ for GPR in training. In the case of the sparse GPR, Matèrn and RQ covariance functions have RMSE below 1% and 1.35%, and MAE below 3.7% and 4.7%, respectively, which result in resonable SoC estimates. On the other hand, combining the two covariance functions provides a better fit to the data. In particular, the sum of Matèrn and RQ covariance functions achieves better estimation performance compared to Matèrn and RQ covariance functions with less than 0.5% RMSE in regular GPR and less than 1% RMSE in sparse GPR.

In the same figure, a larger confidence interval indicates higher gap between the actual and estimated SoC values. On the other hand, more precise SoC estimates have lower



(a) (b) (c)

Fig. 3: SoC estimation based on GPR with (a) Matèrn covariance function, (b) RQ covariance function, (c) sum of Matèrn and RQ covariance functions.

	Regular GPR		Sparse GPR	
Covariance Functions	RMSE (%)	MAE (%)	RMSE (%)	MAE (%)
Matèrn	0.7273	2.1796	0.9955	3.6133
Rational Quadratic (RQ)	1.1233	3.6897	1.3479	4.6650
Sum of Matèrn and RQ	0.4588	1.5502	0.9161	3.7661

TABLE II: RMSE and MAE values of SoC estimation using GPR and sparse GPR with Matèrn, RQ and the sum of Matèrn and RQ covariance functions.

uncertainty, thus smaller confidence interval. Moreover, the relative importance of the inputs can be identified through the values of the optimal hyperparameters. In particular, smaller values of the characteristic length scales indicate that the corresponding input is more important and relevant for SoC estimation. In the case of a sparse GPR with Matèrn covariance function, the optimal values of the characteristic length scales for voltage, current and temperature are 0.5326, 84.7906 and 58.1731, respectively, which implies that voltage has more impact than temperature, and temperature has more impact than current on the SoC estimate. We observe the same relative importance order for other covariance functions, hence we have not included the corresponding optimal hyperparameter values for the sake of clear presentation.

V. Conclusion

In this paper, we have introduced a new method based on sparse GPR for estimating the SoC of Li-ion batteries as a function of voltage, current and temperature measurements of the battery. In comparison to the method based on regular GPR, the proposed method requires smaller amount of training data. Through experimental results, it is shown that the proposed method achieves high accuracy with RMSE below 1% and MAE below 3.8%. Also, it is observed that the sparse GPR significantly reduces the computational time at the expense of minor performance degradation compared to regular GPR.

References

[1] 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)*, pp. 1-5, Sept. 2011.

[2] O. Barbarisi, F. Vasca, and L. Glielmo, "State of charge Kalman filter estimator for automotive batteries," *Control Eng. Pract.*, vol. 14, no. 3, pp. 267-275, March 2006.

[3] 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, March 2013.

[4] 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.

[5] 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.

[6] 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.

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

[8] J.C. Á. Antón, P. J. G. Nieto, C. B. Viejo, and J. A. V. Vilán, "Support vector machines used to estimate the battery state of charge," *IEEE Trans. on Power Electron.*, vol. 28, no. 12, pp. 5919-5926, Dec. 2013.

[9] 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.

[10] C. Hu, G. Jain, C. Schmidt, C. Strief, and M. Sullivan, "Online estimation of lithium-ion battery capacity using sparse Bayesian learning," *Journal of Power Sources*, vol. 289, pp. 105-113, April 2015.

[11] G. Ozcan, M. Pajovic, Z. Sahinoglu, Y. Wang, P. V. Orlik, and T. Wada, "Online state of charge estimation for Lithium-ion batteries using Gaussian process regression," to be submitted.

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

[13] J. Q. Candela and C. E. Rasmussen, "A unifying view of sparse approximate gaussian process regression," *Journal of Machine Learning Research*, vol. 6, pp. 11939-1959, 2005.