State-of-Charge Dependent Battery Model Identification Using Gaussian Process Regression

Kesen Fan* Yiming Wan*

* Key Laboratory of Image Processing and Intelligent Control, Ministry of Education China, and School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan 430074, China (E-mail: ywan@hust.edu.cn)

Abstract: This paper investigates the parameter identification of a state-of-charge dependent equivalent circuit model (ECM) for Lithium-ion batteries. Different from most existing ECM identification methods, we focus on identifying the functional relations between ECM parameters and state-of-charge (SOC). By transforming the ECM into an ARX model, a Gaussian process regression (GPR) approach is proposed, without using parametric functions to describe the SOC dependence of ARX coefficients. The proposed approach derives the posterior distributions of ECM parameters, thus is capable to quantify the estimation uncertainties. Another advantage lies in the flexibility of incorporating the knowledge of batteries into the prior distributions used in GPR, which enhances the estimation performance in the presence of noises. The effectiveness of the proposed GPR approach is illustrated by simulation examples under both low and high noise levels.

Keywords: Gaussian process regression; non-parametric approach; Lithium-ion battery; equivalent circuit model; state-of-charge dependent circuit parameter

1. INTRODUCTION

Due to their high energy density and long cycle life, Lithium-ion batteries (LIBs) are widely used in various applications (Zhang et al., 2014). To ensure safe and reliable operations, an advanced battery management system is needed to monitor state-of-charge (SOC), state-of-health, and state-of-power of batteries. All these monitoring functions rely on a sufficiently accurate Lithium-ion battery model. Therefore, how to obtain such a battery model has attracted tremendous attention in recent years (Zhang et al., 2014).

Existing battery models in literature can be classified into three categories: electrochemical model, data-driven model, and equivalent circuit model (ECM) (Zhang et al., 2014). Among them, the ECM is a gray-box model that combines partial knowledge and data. The ECM is widely used due to its effectiveness in trade-off between model complexity and accuracy. Since battery dynamics significantly varies with operating conditions, the ECM parameters are often described as functions of SOC, temperature, and current load (Huria et al., 2012).

To account for the above time-varying characteristic of ECM parameters, a vast number of recursive least-squares or Kalman filtering based identification methods have been reported, see Zhang et al. (2018), Wang et al. (2018) and Partovibakhsh and Liu (2014) for a non-exhaustive list of references. Such methods update the ECM parameter estimates as the operating condition varies. In addition, it is also of practical interest to identify functional relations between the ECM parameters and the time-varying operating conditions. By collecting data around finite number of separated SOC conditions, a straightforward approach in Huria et al. (2012) first identifies a set of ECM parameters under each SOC condition, and then interpolates these ECM parameters to generate lookup tables. Alternatively, with a comprehensive dataset consisting of various SOC conditions, different optimization based methods are proposed in Malik et al. (2014) and Hu et al. (2009) to determine the nonlinear SOC-dependence of ECM parameters. By assuming affine dependence on SOC and temperature, a linear parameter varying (LPV) ECM model is identified in Hu and Yurkovich (2011) by a subspace method. A common feature of the methods in Malik et al. (2014), Hu et al. (2009) and Hu and Yurkovich (2011) is that the functional relations between the ECM parameters and the operation conditions are specified a priori in a parametrized form. Improper parametrization of such dependencies often leads to increased bias or variance of the identified ECM parameters.

Inspired by Golabi et al. (2017), a non-parametric approach is proposed in this paper to identify the SOC-dependent ECM parameters. From a first-order ECM, the proposed approach first derives an ARX model (autoregression model with exogenous inputs) with SOC-dependent coefficients. Such SOC dependency is described by non-parametric Gaussian processes (GPs), which is different from the parametrized functional forms adopted
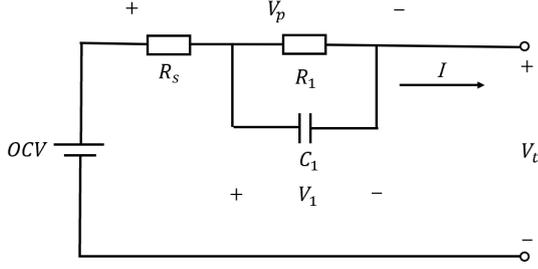
Fig. 1. First-order ECM of the LIB

in Malik et al. (2014), Hu et al. (2009) and Hu and Yurkovich (2011). By exploiting Gaussian process regression (GPR), the posterior Gaussian distributions of the SOC-dependence of ECM parameters are obtained. This naturally provides maximum posterior estimates with associated uncertainty quantification. Moreover, the proposed GPR approach is flexible in using an estimate of parametrized SOC dependence in Malik et al. (2014), Hu et al. (2009) and Hu and Yurkovich (2011) to define the priori distribution of GPs, which achieves performance improvement in the simulation study.

The rest of this paper is organized as follows. Section 2 describes the SOC-dependent battery model and states the parameter identification problem. Section 3 presents the GPR approach to identify SOC-dependent ECM parameters. Section 4 demonstrates the effectiveness of the proposed method via simulation examples. The conclusions are given in Section 5.

2. MODEL DESCRIPTION AND PROBLEM STATEMENT

In this section, we introduce the continuous-time first-order ECM, and derive the corresponding discrete-time LPV-ARX model. Then, we describe the SOC-dependent ECM parameter identification problem to be solved in this paper.

2.1 Continuous-time equivalent circuit model

Due to its simplicity and accuracy, the first-order ECM in Fig.1 is widely adopted in LIB applications. The open circuit voltage (OCV) is denoted by V_{oc}. The internal resistance R_s is used to represent the ohmic polarization phenomenon. The polarization resistance R_1 and polarization capacitance C_1 are used to denote the phenomena of electrochemical polarization and concentration polarization, which can reflect the transient dynamics under current excitation. Based on the circuit theory, the first-order ECM is expressed as

$$I = C_1 \frac{dV_1}{dt} + \frac{V_1}{R_1}, \qquad (1)$$

$$V_{oc} = IR_s + V_1 + V_t, \qquad (2)$$

where I is the load current. V_t is the terminal voltage and V_1 represent the voltage of C_1. The ECM parameters R_s, R_1 and C_1 vary with SOC, while their dependence on temperature is left to future research. For convenience, z is used to denote the SOC throughout this paper. Both the above ECM parameters and V_{oc} have nonlinear

dependence on SOC, which are denoted by $R_s(z)$, $R_1(z)$, $C_1(z)$ and $V_{oc}(z)$.

The OCV-SOC relationship $V_{oc}(z)$ can be obtained from the OCV-SOC test (Zheng et al., 2016). In this paper, SOC is assumed to be available, which can be obtained by Coulomb counting with an accurate initial SOC. For more methods of SOC estimation, please refer to Krewer et al. (2018).

2.2 Discrete-time LPV-ARX model

In order to identify ECM parameters from sampled measurements, the ECM in (1) needs to be transformed into a discrete-time model. Let T_s denote the sampling interval. By assuming constant current over the time interval $[(k-1)T_s, kT_s]$, it can be then derived from (1) that

$$V_{1,k} = R_1(z_k)(1 - e^{-\frac{T_s}{R_1(z_k)C_1(z_k)}})I_{k-1}$$
$$+ e^{-\frac{T_s}{R_1(z_k)C_1(z_k)}} V_{1,k-1}, \qquad (3)$$

where $V_{1,k}$, I_k, z_k denote $V_1(kT_s)$, $I(kT_s)$, z(kT_s).

Substituting

$$V_{p,k} = V_{t,k} - V_{oc,k} = V_{1,k} + R_s(z_k)I_k \qquad (4)$$

into (3) where $V_{oc,k}$ denotes $V_{oc}(kT_s)$, we obain the following LPV-ARX for ECM:

$$V_p(k) = \theta_1(z_k)V_{p,k-1} + \theta_2(z_k)I_{k-1} + \theta_3(z_k)I_k + e_k, \quad (5)$$

where $\theta_1(z_{k-1})$, $\theta_2(z_{k-1})$, and $\theta_3(z_k)$ are SOC-dependent ARX coefficients defined as

$$\theta_1(z_k) = e^{-\frac{T_s}{R_1(z_k)C_1(z_k)}}, \qquad (6)$$

$$\theta_2(z_k) = R_1(z_k)\left(1 - e^{-\frac{T_s}{R_1(z_k)C_1(z_k)}}\right)$$
$$- R_s(z_k)e^{-\frac{T_s}{R_1(z_k)C_1(z_k)}}, \qquad (7)$$

$$\theta_3(z_k) = R_s(z_k). \qquad (8)$$

Note that e_k is included as a zero-mean white Gaussian noise with variance σ^2 to account for measurement noises and approximation errors. Using the LPV-ARX model (5) for ECM parameter identification implies that $V_{p,k}$ needs to be computed from $V_{oc,k}$ and $V_{t,k}$ according to (4). $V_{oc,k}$ is determined by the estimated SOC and the OCV-SOC relationship.

2.3 Problem statement

Due to their dependence on SOC, the ECM parameters vary with time. Various recursive identification methods have been reported in literature to estimate these time-varying ECM parameters (Zhang et al., 2018; Wang et al., 2018; Partovibakhsh and Liu, 2014). However, such methods do not quantify the functional relations between these ECM parameters and SOC. In this paper, we aim at identifying the nonlinear SOC-dependence of the internal resistance R_s and the time constant τ by applying a GPR approach to the LPV-ARX model in (5).

3. IDENTIFICATION OF SOC-DEPENDENT ECM PARAMETERS

In this section, the identification of SOC-dependent ECM parameters is divided into two steps. Firstly, a GPR

approach is exploited to identify the LPV-ARX model coefficients; then, the SOC-dependent ECM parameters are derived by using (6)-(8).

3.1 The re-scaled LPV-ARX model

For convenience, y_k, $x_{1,k}$, $x_{2,k}$, and $x_{3,k}$ are used to denote $V_{p,k}$, $V_{p,k-1}$, I_{k-1}, and I_k, respectively. Then, the LPV-ARX model can be rewritten as

$$y(k) = \theta_1(z_k)x_{1,k} + \theta_2(z_k)x_{2,k} + \theta_3(z_k)x_{3,k} + e_k. \quad (9)$$

However, the LPV-ARX model coefficients in (9) take values at vastly different orders of magnitude, which results in a poorly scaled problem for numerical optimization. For this reason, the model identification are performed on a re-scaled LPV-ARX model

$$y(k) = \theta_1^*(z_k)x_{1,k}^* + \theta_2^*(z_k)x_{2,k}^* + \theta_3^*(z_k)x_{3,k}^* + e_k, \quad (10)$$

$$\theta_1^*(z_k) = \theta_1(z_k), \ \theta_2^*(z_k) = \beta_2\theta_2(z_k), \ \theta_3^*(z_k) = \beta_3\theta_3(z_k), \quad (11)$$

$$x_{1,k}^* = x_{1,k}, \ x_{2,k}^* = \frac{1}{\beta_2}x_{2,k}, \ x_{3,k}^* = \frac{1}{\beta_3}x_{3,k}, \quad (12)$$

which introduces scale factors β_2 and β_3 such that $\{\theta_i^*(z)\}_{i=1}^3$ are of the same order of magnitude.

3.2 GPR approach to LPV-ARX model identification

Inspired by Golabi et al. (2017), the SOC-dependent ARX model coefficients $\{\theta_1^*(z_k)\}$, $\{\theta_2^*(z_k)\}$ and $\{\theta_3^*(z_k)\}$ in (10) are described as three independent GPs whose mean and covariance functions are defined as

$$\mathbb{E}[\theta_i^*(z_j)] = \mu_i(z_j), \quad (13)$$

$$\mathrm{cov}[\theta_i^*(z_j), \theta_i^*(z_l)] = K^i(z_j, z_l) = \lambda_i e^{-\frac{(z_j-z_l)^2}{2\delta_i^2}}, \ i = 1, 2, 3 \quad (14)$$

with λ_i and δ_i being hyperparameters of the Gaussian kernel function $K^i(z_j, z_l)$. The values of these hyperparameters are determined from training data, as will be explained in Section 3.3. According to (10), it can be seen that $\{y_k\}$ is also a Gaussian process with its mean and covariance function as follows:

$$\mathbb{E}(y_j) = v(z_j) = \sum_{i=1}^3 \mu_i(z_j)x_{i,j}^*, \quad (15)$$

$$\mathrm{cov}(y_j, y_l) = \sum_{i=1}^3 x_{i,j}^* K^i(z_j, z_l)x_{i,l}^* + \sigma_{jl}^2, \quad (16)$$

$$j = 1, 2, \ldots, N, \quad l = 1, 2, \ldots, N,$$

where σ_{jl}^2 is a function that is equal to σ^2 only if $j = l$ and zero otherwise. With the above derivation, the joint distribution of a sequence of measured outputs $Y = [y_1, \ldots y_N]^\top$ and a SOC-dependent ARX coefficient $\theta_i^*(z_k)$ is a Gaussian distribution

$$\begin{bmatrix} Y \\ \theta_i^*(z_k) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} v \\ \mu_i(z_k) \end{bmatrix}, \begin{bmatrix} K & k_i \\ k_i^\top & K^i(z_k, z_k) \end{bmatrix}\right). \quad (17)$$

where the Gram matrix K, the mean vector v and the vector k_i are defined as

$$[K]_{jl} = \mathrm{cov}(y_j, y_l),$$

$$v = [v(z_1) \quad v(z_2) \ldots v(z_N)]^\top,$$

$$k_i = [x_{i,1}^* K^i(z_k, z_1) \quad x_{i,2}^* K^i(z_k, z_2) \ldots x_{i,N}^* K^i(z_k, z_N)]^\top.$$

From (17), given the measured sequence Y, the posterior distributions of the re-scaled LPV-ARX model coefficients

$\{\theta_i^*(z_k)\}_{i=1}^3$ still form GPs, and their posterior mean and variance are

$$\mathbb{E}[\theta_i^*(z_k)|D] = \mu_i(z_k) + k_i^\top(K + \sigma^2 I)^{-1}(Y - v), \quad (18)$$

$$\mathrm{cov}[\theta_i^*(z_k)|D] = K^i(z_k, z_k) - k_i^\top(K + \sigma^2 I)^{-1}k_i, \quad (19)$$

with D representing a dataset $\{x_{1,j}^*, x_{2,j}^*, x_{3,j}^*, z_j, y_j\}_{j=1}^N$. To compute the above posterior mean in (18), the priori mean function $\mu_i(z)$ needs to be specified. If no priori knowledge is available, the prior mean $\mu_i(z)$ is usually set to zero. If the SOC-dependent ECM parameters are parametrized and estimated using the methods proposed in Hu and Yurkovich (2011), Hu et al. (2009) and Malik et al. (2014), these estimated SOC dependencies can be used as the mean functions $\mu_i(z)$. Such incorporation of additional a prior knowledge into non-zero mean functions helps enhance the estimation performance, as illustrated by the simulation example in Section 4.2.

3.3 Determination of hyperparameters

With a training dataset, the hyperparameters $\{\lambda_i, \delta_i\}_{i=1}^3$ in (14) and the noise standard deviation σ can be determined by maximizing the marginal likelihood. The logarithm of the marginal likelihood to be optimized is shown below after removing the constant term:

$$g(\lambda_i, \delta_i, \sigma) = -\frac{1}{2}(Y - v)^\top(K + \sigma^2 I)^{-1}(Y - v) \\ -\frac{1}{2}\log|K + \sigma^2 I|. \quad (20)$$

However, this paper does not solve (20) directly due to a numerical problem. The second term in (20) involves the calculation of the determinant, with carries the risk of numerical underflow, which means the calculation result of determinant is less than the smallest number that the computer can represent. To address this issue, we introduces an amplification factor α to rewrite the objective function (20) as

$$g(\lambda_i, \delta_i, \sigma) = -\frac{1}{2}(Y - v)^\top(K + \sigma^2 I)^{-1}(Y - v) \\ -\frac{1}{2}\log|\alpha(K + \sigma^2 I)|. \quad (21)$$

This does not affect the optimal solution, since the difference between (20) and (21) are only a constant.

3.4 From LPV-ARX model coefficients to ECM parameters

According to (6)-(8) and (11), from the LPV-ARX model coefficients, the internal resistance R_s and the time constant τ can be expressed as

$$R_s(z_k) = \frac{1}{\beta_3}\theta_3^*(z_k), \quad (22)$$

$$\tau(z_k) = -\frac{T_s}{\ln\theta_1^*(z_k)}. \quad (23)$$

Since $\theta_1^*(z_k)$ and $\theta_3^*(z_k)$ are Gaussian distributed random variables obtained in Section 3.2, both $R_s(z_k)$ and $\tau(z_k)$ are also random. Note that $\tau(z_k)$ is non-Gaussian distributed due to the nonlinear transformation in (23). To avoid the expensive computation in deriving the accurate distribution of τ, an Gaussian approximation is adopted

by exploiting the first-order Taylor expansion of (23). Let $\hat{\mu}_1(z_k)$ denote $\mathbb{E}[\theta_1^*(z_k)|D]$, then (23) is approximated by

$$\tau(z_k) \approx -\frac{T_s}{\ln\hat{\mu}_1(z_k)} + \frac{T_s}{(\ln\hat{\mu}_1(z_k))^2}[\theta_1^*(z_k) - \hat{\mu}_1(z_k)]. \quad (24)$$

Therefore, according to (22) and (24), the posterior distributions of $R_s(z_k)$ and $\tau(z_k)$ are

$$\mathbb{E}[R_s(z_k)|D] = \frac{1}{\beta_3}\mathbb{E}[\theta_3^*(z_k)|D], \quad (25)$$

$$\mathrm{cov}[R_s(z_k)|D] = \frac{1}{\beta_3^2}\mathrm{cov}[\theta_3^*(z_k)|D], \quad (26)$$

$$\mathbb{E}[\tau(z_k)|D] = -\frac{T_s}{\ln\hat{\mu}_1(z_k)}, \quad (27)$$

$$\mathrm{cov}[\tau(z_k)|D] = \frac{T_s^2}{(\ln\hat{\mu}_1(z_k))^4}\mathrm{cov}[\theta_1^*(z_k)|D]. \quad (28)$$

3.5 Summary of the proposed identification algorithm

Based on the available data V_t and I, the proposed identification algorithm is summarized below.

(1) Compute SOC by Coulomb counting, then determine the functional relation between OCV and SOC from the OCV-SOC test.
(2) Select the appropriate scale factors β_2 and β_3 in (11) and compute the re-scaled regressors in (12).
(3) Determine the prior mean functions $\mu_i(z)$ for the GPs. If no priori knowledge is available, $\mu_i(z)$ is usually set to zero. If the SOC-dependent ECM parameters are parametrized and estimated using the methods proposed in Malik et al. (2014), Hu et al. (2009) and Hu and Yurkovich (2011), these estimated SOC dependencies can be used as the mean functions $\mu_i(z)$.
(4) Determine the amplification factors α in (21), then estimate hyperparameters $\{\lambda_i, \delta_i\}_{i=1}^3$ and the noise standard deviation σ by maximizing (21).
(5) Obtain the posterior distributions of the re-scaled LPV-ARX model coefficients $\{\theta_i^*(z)\}_{i=1}^3$ according to (18)-(19).
(6) Obtain the approximated posterior means and variances of the ECM parameters R_s and τ using (25)-(28). The parameter estimates are set to be the posterior means, and their uncertainties are quantified by the posterior variances.

4. SIMULATION EXAMPLE

This section presents the identification results of the proposed GPR approach using a simulated first-order ECM. In the simulation setup, the capacity, the OCV-SOC relationship and the SOC-dependent parameters are configured as follows according to Shen et al. (2016): the capacity is $8Ah$,

$$R_s(z) = \sum_{k=0}^{5} b_1(k)z^k, \quad R_1(z) = \sum_{k=0}^{5} b_2(k)z^k,$$

$$C_1(z) = \sum_{k=0}^{5} b_3(k)z^k, \quad V_{oc}(z) = \sum_{k=0}^{11} b_4(k)z^k,$$
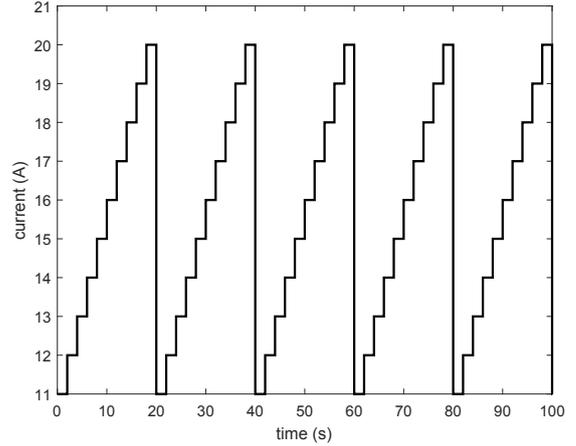
Fig. 2. Current excitation signal for the first 100 seconds

$b_1 = [0.005, 0.022, -0.109, 0.215, -0.192, 0.065]$,
$b_2 = [0.015, -0.128, 0.576, -1.180, 1.114, -0.396]$,
$b_3 = [12200, 19423, -40000, 40000, 1317, -4000]$,
$b_4 = [3.04, 8.58, -53.31, 125.72, 148.51, -1426.11,$
$\quad 3000, -2309, -934, 3000, -2045, -486]$.

The discharge current excitation is periodic, whose curve over the first 100 seconds is depicted in Fig.2. The total simulation time is 1518 seconds, such that the SOC starts at 80% and ends at zero. The sample time T_s is set to 2 seconds. To evaluate the identification performance, the following Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) of the identified ECM parameters are used:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(h_k - \hat{h}_k)^2}, \quad (29)$$

$$\mathrm{MAPE} = \frac{1}{N}\sum_{k=1}^{N}\left|\frac{h_k - \hat{h}_k}{h_k}\right| \times 100\%, \quad (30)$$

where h_k is the true value, $\hat{h}_k$ is the estimate, and N is the number of samples. In the following, two scenarios with a low and high noise level are considered in the simulation study. it is demonstrated that the proposed GPR approach with zero-mean priori distributions is effective at a low noise level, while encoding additional knowledge into nonzero-mean prior distributions improves estimation performance under both two noise levels.

4.1 Identification results at a low noise level

In the first simulation scenario, the measurement noise variance σ^2 is set to 10^{-8}. In the summarized identification algorithm of Section 3.5, the factors in Step 2 and Step 4 are selected as 1000, 1000, and 5×10^7; zero-mean priori distributions are adopted in Step 3 for the GPs. In Step 4, the optimization solver minFunc (Schmidt, 2005) is used to maximize (21) by using the Newton method, and the obtained hyperparameters are

$\lambda_1 = 0.91, \lambda_2 = 0.13, \lambda_3 = 0.97, \sigma = 1.49 \times 10^{-4}$,
$\delta_1 = 0.18, \delta_2 = 0.64, \delta_3 = 0.21$.

After determining the hyperparameters, we can obtain the posteriori distributions of internal resistance R_s and time

constants τ by following Steps 4 and 5 of the algorithm summarized in Section 3.5. The posteriori means are used as the parameter estimates, while the posterior variances are exploited to determine the 2-σ confidence intervals. These estimates and their confidence intervals are shown in Fig.3. The obtained RMSE and MAPE for R_s and τ are shown in Tables 1-4 .The small estimation errors and narrow confidence intervals indicate the accuracy and reliability of the obtained estimates.

Now consider the introduction of nonzero-mean prior distributions in Step 3 of the algorithm summarized in Section 3.5. For this purpose, the mean functions $\{\mu_i(z)\}_{i=1}^3$ are first polynomially parametrized as

$$\mu_i(z) = \sum_{k=1}^{4} a_i(k)z^k, \qquad (31)$$

and the least-squares identification of these polynomial coefficients are

$a_1 = [1.00467, -0.27380, 0.73487, -0.61904, 0]$,
$a_2 = [-0.00494, -0.01701, 0.75487, -0.11742, 0.06204]$,
$a_3 = [0.00497, 0.01936, -0.08159, 0.12174, -0.06123]$.

Using the mean functions in (31), the obtained estimates and their confidence intervals are shown in Fig.4. The obtained RMSE and MAPE for R_s and τ are shown in Tables 1-4. The proposed algorithms with zero-mean and nonzero-mean priori distributions both achieve small estimation errors and narrow confidence intervals under the low noise level. The proposed algorithms with nonzero-mean priori distributions achieve slightly better performance.

Table 1. RMSE for R_s in different scenarios

Scenarios	Low noise level	High noise level
$\mu_i(z) = 0$	8.46×10^{-6}	1.04×10^{-4}
$\mu_i(z)$ in (31)	7.19×10^{-6}	4.17×10^{-5}

Table 2. RMSE for τ in different scenarios

Scenarios	Low noise level	High noise level
$\mu_i(z) = 0$	6.99	16.20
$\mu_i(z)$ in (31)	4.39	13.84

Table 3. MAPE for R_s in different scenarios

Scenarios	Low noise level	High noise level
$\mu_i(z) = 0$	0.108%	1.27%
$\mu_i(z)$ in (31)	0.095%	0.63%

Table 4. MAPE for τ in different scenarios

Scenarios	Low noise level	High noise level
$\mu_i(z) = 0$	3.37%	21.49%
$\mu_i(z)$ in (31)	2.53%	18.05%

4.2 Identification results at a high noise level

The second simulation scenario sets a higher noise level, i.e., the measured noise variance is increased to 10^{-6}. In
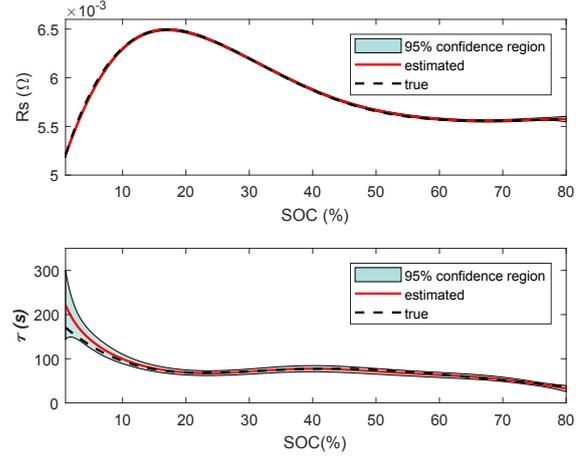
Fig. 3. Identification results of internal resistance R_s and time constant τ with zero-mean priori distributions at a low noise level
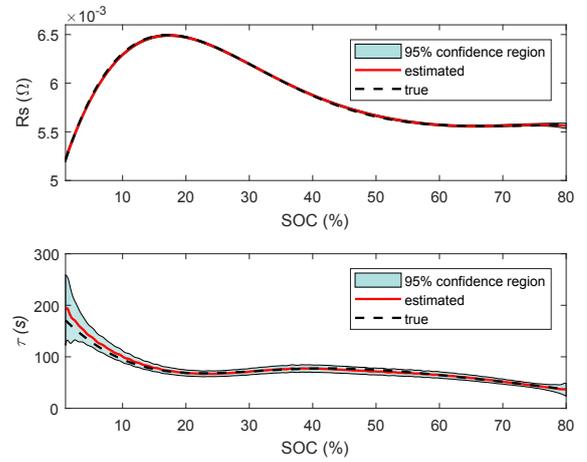
Fig. 4. Identification results of internal resistance R_s and time constant τ with nonzero-mean priori distributions at a low noise level

the implemented identification algorithm, the scale factors β_2, β_3 are the same as in Section 4.1, and the amplification factor α is set to 7×10^5. Firstly, the same GPR approach with the zero-mean priori distribution in Section 4.1 is examined. The obtained hyperparameters are

$\lambda_1 = 0.61, \ \lambda_2 = 0.83, \ \lambda_3 = 0.60, \ \sigma = 1.48 \times 10^{-3}$,
$\delta_1 = 0.41, \ \delta_2 = 0.97, \delta_3 = 0.35$.

Then the GPR approach with nonzero-mean priori distributions in Section 4.1 is examined. In this high noise case, the identified polynomial coefficients in (31) are

$a_1 = [0.99697, -0.17439, 0.36065, -0.28335, 0]$,
$a_2 = [-0.00473, -0.01675, 0.66215, -0.09277, 0.04436]$,
$a_3 = [0.00477, 0.01987, -0.07851, 0.11201, 0.05561]$.

The obtained hyperparameters are

$\lambda_1 = 0.95, \ \lambda_2 = 0.66, \ \lambda_3 = 0.88, \ \sigma = 1.50 \times 10^{-4}$,
$\delta_1 = 0.41, \ \delta_2 = 0.77, \delta_3 = 0.54$.

The identification results are shown in Fig.5 and Fig.6. The obtained RMSE and MAPE for R_s and τ are shown in Tables 1-4. Compared to the GPR approach with zero-
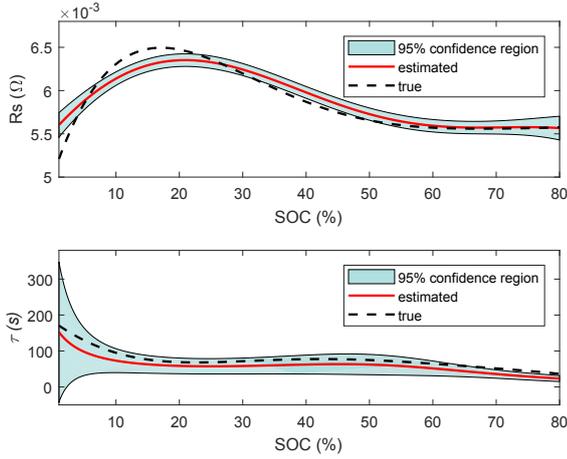
Fig. 5. Identification results of internal resistance R_s and time constant τ with zero-mean priori distributions at a high noise level
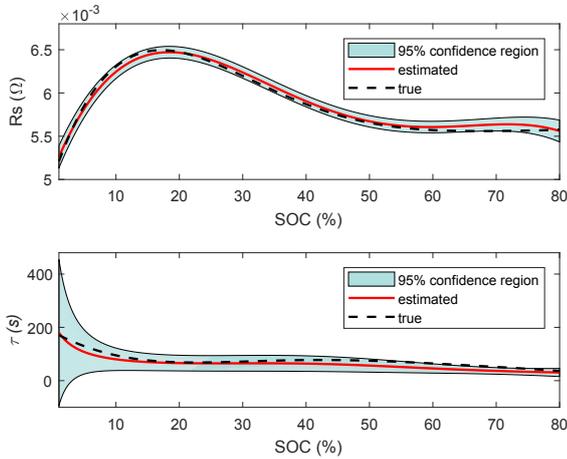
Fig. 6. Identification results of internal resistance R_s and time constant τ with nonzero-mean priori distributions at a high noise level

mean priori distributions, the use of nonzero-mean priori distributions achieves more than 50% decrease in RMSE and MAPE for R_s, and more than 14% decrease in RMSE and MAPE for τ. This shows that the estimation performance under a high noise level is significantly improved by encoding additional knowledge in the nonzero-mean priori distributions.

5. CONCLUSION

In this paper, GPR is employed in the identification of state-of-charge dependent ECM parameters of LIB. Unlike other methods, the method is able to obtain the posteriori distributions of the ECM parameters. The posterior means are used as the parameter estimates, and the posteriori variances quantify the associated estimation uncertainties. The proposed method can flexibly incorporate the priori knowledge of batteries by using nonzero-mean functions, which enhances its estimation performance under a high noise level. Future research effort will focus on the computation efficiency of the proposed GPR-based identification algorithm, in order to better cope with the large amount of data collected over a wide range of SOC.

REFERENCES

Golabi, A., Meskin, N., Tóth, R., and Mohammadpour, J. (2017). A Bayesian approach for LPV model identification and its application to complex processes. *IEEE Transactions on Control Systems Technology*, 25(6), 2160–2167.

Hu, Y., Yurkovich, S., Guezennec, Y., and Yurkovich, B. (2009). A technique for dynamic battery model identification in automotive applications using linear parameter varying structures. *Control Engineering Practice*, 17(10), 1190–1201.

Hu, Y. and Yurkovich, S. (2011). Linear parameter varying battery model identification using subspace methods. *Journal of Power Sources*, 196(5), 2913–2923.

Huria, T., Ceraolo, M., Gazzarri, J., and Jackey, R. (2012). High fidelity electrical model with thermal dependence for characterization and simulation of high power lithium battery cells. In *Proceedings of 2012 IEEE International Electric Vehicle Conference*, 1–8. Greenville, SC, USA.

Krewer, U., Röder, F., Harinath, E., Braatz, R.D., Bedürftig, B., and Findeisen, R. (2018). Dynamic models of li-ion batteries for diagnosis and operation: a review and perspective. *Journal of the electrochemical society*, 165(16), A3656.

Malik, A., Zhang, Z., and Agarwal, R.K. (2014). Extraction of battery parameters using a multi-objective genetic algorithm with a non-linear circuit model. *Journal of Power Sources*, 259, 76–86.

Partovibakhsh, M. and Liu, G. (2014). An adaptive unscented Kalman filtering approach for online estimation of model parameters and state-of-charge of lithium-ion batteries for autonomous mobile robots. *IEEE Transactions on Control Systems Technology*, 23(1), 357–363.

Schmidt, M. (2005). minfunc: unconstrained differentiable multivariate optimization in Matlab. `http://www.cs.ubc.ca/~schmidtm/Software/minFunc.html`.

Shen, J.N., He, Y.J., and Ma, Z.F. (2016). Simultaneous model selection and parameter estimation for lithium-ion batteries: A sequential MINLP solution approach. *AIChE Journal*, 62(1), 78–89.

Wang, Q., Kang, J., Tan, Z., and Luo, M. (2018). An online method to simultaneously identify the parameters and estimate states for lithium-ion batteries. *Electrochimica Acta*, 289, 376–388.

Zhang, C., Allafi, W., Dinh, Q., Ascencio, P., and Marco, J. (2018). Online estimation of battery equivalent circuit model parameters and state of charge using decoupled least squares technique. *Energy*, 142, 678–688.

Zhang, C., Li, K., Mcloone, S., and Yang, Z. (2014). Battery modelling methods for electric vehicles-A review. In *Proceedings of 2014 European Control Conference (ECC)*, 2673–2678. Strasbourg, France.

Zheng, F., Xing, Y., Jiang, J., Sun, B., Kim, J., and Pecht, M. (2016). Influence of different open circuit voltage tests on state of charge online estimation for lithium-ion batteries. *Applied Energy*, 183, 513–525.