Learning of Battery Model Bias for Effective State of Charge Estimation of Lithium-Ion Batteries

Zhimin Xi[®], *Member, IEEE*, Modjtaba Dahmardeh[®], Bing Xia, *Student Member, IEEE*, Yuhong Fu[®], and Chris Mi[®], *Fellow, IEEE*

Abstract-State of charge (SOC) estimation of lithium-ion batteries has been extensively studied and the estimation accuracy was mainly investigated through the development of various battery models and dynamic estimation algorithms. All battery models, however, contain inherent model bias due to the simplifications and assumptions, which cannot be effectively addressed through the development of various algorithms such as Kalman filtering (KF) or particle filtering (PF). Consequently, as observed in some study, battery SOC estimation using a typical extended KF in fact is not very accurate where the error could range from 5% to 10% or even more depending on the battery characteristics. This paper proposes bias characterization of the battery model, so that accuracy of the baseline model could be significantly improved and eventually SOC estimation could be much more accurate than the one only using the baseline model. This paper reports great potential for improving battery SOC estimation with the bias characterization and proposes two methods for actual bias modeling. In particular, the polynomial regression model and the Gaussian process regression model are proposed to examine the effects of the two methods on bias modeling and SOC estimation using a typical battery circuit model. Results are demonstrated in both simulation and lab testing using three battery charging/discharging profiles with the cross-validation technique.

Index Terms—Battery model bias, lithium-ion battery, SOC estimation, Gaussian process regression, battery circuit model.

I. INTRODUCTION

S OC estimation of lithium-ion batteries has been extensively studied through the development of various battery models and dynamic estimation algorithms. In general, battery models can be classified into three categories: i) electrochemical models [1]–[4], ii) equivalent circuit models (ECMs) [5]–[8], and iii) machine learning models [9]–[11]. Electrochemical models are typically computationally expensive where a set of non-linear differential equations should be solved to estimate the battery performances of interest. ECMs are simplified semi-physics-based models with high computational efficiency but the accuracy is compromised in addition to the limited capability of considering other important parameters such as

Manuscript received May 15, 2018; accepted July 7, 2019. Date of publication July 17, 2019; date of current version September 17, 2019. This work was supported in part by the National Science Foundation under award #1507198 and in part by DARPA – 2016 Young Faculty Award. The review of this paper was coordinated by Prof. L. Gauchia. (*Corresponding author: Zhimin Xi.*)

Z. Xi and M. Dahmardeh are with the Rutgers University, New Brunswick, NJ 08854 USA (e-mail: zhimin.xi@rutgers.edu; md1352@soe.rutgers.edu).

B. Xia, Y. Fu, and C. Mi are with the San Diego State University, San Diego, CA 92130 USA (e-mail: bixia@eng.uscd.edu; fuyuhong@gmail.com; cmi@sdsu.edu).

Digital Object Identifier 10.1109/TVT.2019.2929197

the operation temperature of the battery. Machine learning models are empirical-based models and the performance is largely dependent on the similarity between the battery training and actual operating conditions. Since battery SOC dynamically changes depending on the charging/discharging profile, various algorithms have been developed for estimating the hidden yet evolving model parameter (i.e., battery SOC) including the Kalman filter (KF) [12], the extend KF (EKF) [13], the unscented KF (UKF) [14], and the particle filter (PF) [15]. The major difference of these algorithms lies in the degree of simplifications of the battery model for the SOC estimation. KF applies only for linear battery model. EKF linearizes a nonlinear battery model at different time steps. UKF intends to accurately predict the nonlinearity of the battery model with a minimum set of chosen samples. PF directly runs sequential Monte Carlo simulation based on the actual nonlinearity of the battery model. Apparently, there is a tradeoff among the algorithm efficiency, accuracy, and complexity running on a chosen battery model.

As pointed out in authors' previous work [16], five types of uncertainty play a key role for reliable estimation of the battery performances of interest and they can be classified as: i) measurement uncertainty from sensors, ii) algorithm uncertainty based on algorithm characteristics, iii) environmental uncertainty such as temperature and various loading conditions, iv) model parameter uncertainty due to the cell-to-cell variability, and v) model uncertainty due to the inherent model inadequacy from model simplifications and assumptions. Related work in each category is briefly reviewed as follows.

- i) Sensor measurement uncertainty: Although typical filtering algorithms consider sensor measurement uncertainty, their mean values are usually assumed to be zero which may not be realistic considering actual sensor measurement errors. Battery SOC estimation error due to nonzero mean of the sensor measurement error (or bias) was recently studied using both the least square estimation and the EKF, and the estimation errors were theoretically derived and validated through experiments [17], [18].
- ii) Algorithm uncertainty: The essential purpose of various filtering algorithms [12]–[15] is to reduce the algorithm error for battery states estimation while maintaining high computational efficiency.
- iii) Environmental uncertainty: Temperature and loading conditions are the most important factors influencing the battery characteristics such as capacity, resistance, open circuit voltage (OCV), etc. [19]–[21].

0018-9545 © 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

- iv) Model parameter uncertainty: This type of uncertainty is mainly due to the cell-to-cell variability caused by manufacturing tolerance such that each cell should have its unique parameter set corresponding to a particular battery model. An early study conducted by Lin *et al.* [22] investigated battery SOC and voltage estimation errors with inaccurate model parameters and they found that SOC estimation error was unavoidable even though voltage error could be insignificant. Such parameter uncertainty was examined by Jing *et al.* [16] for EIG C020 cells and battery SOC was further estimated considering the parameter uncertainty. A recent study also investigated the battery pack SOC variability due to the cell-level parameter uncertainty [23].
- v) Model uncertainty: One of the basic objectives of developing various battery models [1]-[11] is to improve the model prediction accuracy, or in other words to reduce model prediction bias, under various conditions. For example, an extended ECM (EECM) based on the single-particle electrochemical model was proposed to improve the model estimation accuracy especially in the low SOC area [24]. Nevertheless, none of the model would be perfect without any errors. Hence, in addition to the model development, identification and characterization of model uncertainty of a baseline model could be an alternative option to dramatically improve the model estimation accuracy. In the early work conducted by Jing et al. [16], bias of individual battery cell was calculated and average of the bias from training cells was then applied to the testing cell under the same battery operating condition, in which noticeable SOC accuracy improvement had been reported. However, the major limitation is that the training and test cells were under the same charge/discharge profile. Mishra et al. [25] derived an exact model bias form based on the high-order and low-order RC circuit models and further verified such model bias using both simulations and experiments. It is worth noting that such bias derivation can only be conducted with an assumed "true" or high-fidelity reference model which may never exist in reality. Hence, learning of model bias through actual experiments under some training conditions would be more practical. Nevertheless, above work [16], [25] clearly demonstrated the benefits if the model bias could be identified or properly learned.

The contribution of this paper is three-fold: i) propose general bias learning and modeling framework for lithium-ion battery models so that fidelity of the baseline model could be significantly improved under different operating conditions without compromising the computational efficiency; ii) compare a typical machine learning method and a classical parametric regression model for the effectiveness of bias modeling; and iii) recommend the machine learning method for bias modeling due to its effectiveness and great potential for further improvement. In particular, two methods including the polynomial regression and the Gaussian process (GP) regression model are proposed for bias modeling and their effects on battery voltage and SOC estimation are studied under different battery operating conditions. Furthermore, based on an extensive comparison study conducted for battery circuit models [5], the first- and second- order RC models are chosen as two baseline battery models due to the overall high accuracy and efficiency in lithiumion battery modeling.

The rest of the paper is organized as follows. Section II briefly reviews the first- and second- order RC models [5] which are employed in this study as two baseline battery models. Section III first proposes general framework for considering the model bias in battery performance estimation, then elaborates two feasible methods for the bias characterization and approximation. Section IV demonstrates the effectiveness of the proposed method through both simulations and laboratory experiments. Section V concludes the paper and discusses future work on generalization of the bias modeling considering the cell-to-cell variability and aging effects.

II. REVIEW OF THE RC CIRCUIT MODEL

Battery hidden states such as SOC and state of health (SOH) are usually estimated using a discrete time state-space model as

$$\begin{cases} x_k = f(x_{k-1}, u_k) + w_{k-1} \\ y_k = g(x_k, u_k) + v_k \end{cases}$$
(1)

where x_k is the hidden battery state (e.g., SOC) at the k^{th} time step; u_k means the input vector (e.g., current); w_{k-1} is the process noise; y_k is the output vector (e.g., terminal voltage); v_k is the measurement noise of the output vector; $f(\cdot)$ is the state transition function; and $g(\cdot)$ is the battery model that relates the output vector with the input and hidden state vectors. With the Coulomb counting as the SOC transition function, the q^{th} order RC model can be written as

$$\begin{cases} x_k = x_{k-1} - \eta \Delta T i_k / C_r + w_{k-1} \\ y_k = OCV(x_k) + i_k R + \sum_q U_{q,k} + v_k \end{cases}$$
(2)

where $\eta \Delta T i_k$ is the coulomb accumulation for given charging/discharging efficiency η , current i_k and time accumulation ΔT ; C_r is the rated capacity of the battery; $OCV(x_k)$ is the open circuit voltage (OCV) of the battery cell as a function of the battery state (i.e., SOC); R is the battery charge/discharge internal resistance; and $U_{q,k}$ is the voltage of the q^{th} RC network and expressed as

$$U_{q,k} = \exp\left(-\Delta t/\tau_q\right)U_{q,k-1} + R_q\left[1 - \exp\left(-\Delta t/\tau_q\right)\right]i_k \tag{3}$$

where R_q and τ_q are the resistance and time constant of the q^{th} RC network, respectively. Parameter characterization of the battery model can be conducted based on training data sets such as hybrid pulse power characterization (HPPC) using either the EKF or optimization methods based on the least square estimation [26] or the particle swarm optimization [27]. The nonlinear relationship between OCV and SOC should also be identified through the battery characterization test.

III. FRAMEWORK AND METHODS FOR BIAS MODELING

Model bias is the inherent inadequacy of the model for representing the real physical systems due to the model assumptions



Fig. 1. Flowchart of the proposed framework for battery states estimation considering model uncertainty.

and simplifications. All battery models, whether electrochemical models, ECMs, or machining learning models, contain model errors or biases. The difference is the degree of errors at different battery operating conditions. Using the RC model in Eq. (2) as an example, model bias can be viewed as the error of the predicted terminal voltage compared to the true terminal voltage at any battery operating conditions defined by the charge/discharge current, SOC, temperature, maximum available capacity, etc. Battery model uncertainty stands for the uncertainty of model bias even at the same operating condition, which is mainly attributed by some unknown parameters that could affect the terminal voltage or some factors that cannot be fully described by the defined model parameters. The objective of the bias modeling is to accurately predict the model uncertainty (e.g., mean and variance) at any battery operating conditions. It is worth noting that bias modeling is different with model calibration, in which model parameters are typically tuned to maximize the agreement between the predicted and measured terminal voltage for specified operating conditions while ignoring the model form error. Consequently, a calibrated model may have inaccurate model prediction outside the calibration domain. Bias characterization and modeling have been studied and the benefits have been well demonstrated for general engineering design problems which are usually static systems [28]–[32]. This paper essentially extends the bias characterization and modeling to dynamic systems and addresses associated challenges.

A. Overall Approach and Assumptions

The proposed framework is shown in Fig. 1, where the first step is to choose a baseline battery model. The next step is to calibrate unknown model parameters through the battery characterization test such as the HPPC test. Depending on different battery models, the number of unknown model parameters could be different. Once the characterization test is completed, these model parameters are fixed for the battery states (e.g., SOC) estimation. In other words, prediction accuracy of the model output, i.e., terminal voltage, of the baseline model may vary under different battery operating conditions. The goal of the bias modeling is to further characterize the bias and its uncertainty of the baseline model under various battery operating conditions. Hence, training data sets are required at the 3rd step to model the bias and its uncertainty at different battery operating conditions, where the bias function form could be arbitrary, e.g., either parametric or non-parametric, defined by current, SOC, temperature, etc. The final step is to estimate the battery states (e.g., SOC) at any operating conditions considering the model uncertainty in addition to the baseline model.

Compared to majority of the publications for battery states estimation, key difference of the proposed work lies in model uncertainty characterization. Ideally, if model bias at any battery operating conditions could be truly determined, terminal voltage predicted from the corrected battery model would be exactly equal to the actual voltage measurement assuming no measurement errors. Apparently, such accuracy improvement would significantly improve the battery SOC estimation compared to only using the baseline model with the same filtering algorithm such as the EKF. In reality, accuracy of the bias modeling is dependent on the amount of training data and the method employed for the bias modeling. In addition, bias could exhibit uncertainty even at the same operating condition if other unknown factors are not considered in the bias model. For example, given a battery with a 2A instant discharge current at 50% SOC level, bias of the terminal voltage at the corresponding time step would not be a deterministic value because the charge/discharge history at previous time steps could also affect the bias. Therefore, uncertainty should be introduced to accommodate the unknown factors in the bias modeling, if necessary. If the RC circuit model is employed as the baseline model, a revised cell dynamic model considering model uncertainty can be formulated in Eq. (4),

$$\begin{cases} x_k = x_{k-1} - \eta \Delta T i_k / C_r + w_{k-1} \\ y_k^* = OCV(x_k) + i_k R + \sum_q U_{q,k} + \delta_k(\delta_{k-1}, i_k, x_k, ?) + v_k \end{cases}$$
(4)

where $\delta_k(\delta_{k-1}, i_k, x_k, ?)$ is the bias function at the k^{th} time step; and ? indicates all other factors that could affect the bias, e.g., temperature and aging effects of the battery.

While acknowledging that five types of uncertainty are all important in real battery states estimation, this paper focuses on bias modeling and thus simplifies other uncertainties in a reasonable way as follows. First of all, sensor measurement uncertainties are considered but with zero mean assumptions for both voltage and current sensors. Secondly, algorithm uncertainty is not considered and the EKF is employed for all case studies. Thirdly, temperature uncertainty is not considered and all experiments were conducted at constant room temperature in this study. Thus, effectiveness of the bias learning will be only demonstrated considering different loading (i.e., charging/discharging) profiles. Fourthly, model parameter uncertainty due to the cell-to-cell variability is not considered. Thus, only one battery cell with its unique parameter set will be used for the demonstration. However, generalization of the bias modeling considering model parameter uncertainty will be discussed in future study in Section V. Finally, while acknowledging the uncertainty of model bias δ , this paper focuses on the demonstration of the great potential for considering model bias in the battery SOC estimation. Hence, only the mean prediction of the bias is employed for correcting the baseline model. Given a certain amount of training data for bias characterization, it is expected that the new model output y_k^* (i.e., terminal voltage) in Eq. (4) would be much more accurate than the baseline model output y_k defined in Eq. (2) for various battery operating conditions. Consequently, SOC estimation would be more accurate as well. In particular, two methods are proposed and compared for the bias modeling and they are elaborated in subsections III.B and III.C.

B. Polynomial Regression Model for Bias Modeling

The polynomial regression model characterizes model bias as a polynomial function of battery operation parameters (e.g., SOC and current). This method is simple and straightforward for implementation. However, due to the high nonlinearity and complexity nature of the bias with respect to various battery operation conditions, this model would not be accurate for specific battery operations at given time steps, but it should be useful for modeling the overall trend of the bias in the battery operation domain. A suitable polynomial regression model should avoid potential overfitting issue while providing reasonable accuracy at different battery operation conditions. In this paper, a quadratic form of the bias model is adopted as defined in Eq. (5).

$$\delta_{k} = a_{0} + a_{1}\delta_{k-1} + a_{2}\delta_{k-1}^{2} + a_{3}i_{k} + a_{4}i_{k}^{2} + a_{5}x_{k}$$
$$+ a_{6}x_{k}^{2} + a_{7}\delta_{k-1}i_{k} + a_{8}\delta_{k-1}x_{k} + a_{9}i_{k}x_{k}$$
(5)

Model coefficients, i.e., a_0-a_9 , will be determined from available bias training data by minimizing the root mean square error (RMSE) between the predicted and true bias data.

C. Gaussian Process (GP) Regression Model for Bias Modeling

In the GP regression method [33], voltage biases are assumed to follow a multivariate Gaussian distribution such that the bias at a specific battery operation condition (e.g., given discharge current and SOC) follows a univariate Gaussian distribution. The mean and covariance functions should be determined in the GP regression model based on available bias training data. Technically, a bias GP model can be written as

$$\delta(z) \sim GP(m(z), \quad k(z, z')) \tag{6}$$

where z is a vector of battery operation parameters, e.g., $z = [\delta_{k-1}, i_k, x_k]$; m(z) is the mean function which could be a polynomial regression model defined in Eq. (5); k(z, z') is the covariance function which can take many forms. Among them, the squared-exponential covariance function is commonly used with the following form as

$$k(z, z') = \sigma_{\delta}^{2} \exp\left[-\frac{1}{2} \frac{(z - z')(z - z')^{T}}{l^{2}}\right]$$
(7)

where σ_{δ}^2 and *l* represent the bias variance and the characteristic length scale, respectively, and they must be larger than zero; *z*' is the dummy variable of *z* indicating the separate treatment of battery operation parameters. If *z* and *z*' happen to be at the same location defined by battery operation parameters, the covariance function reaches the maximum value σ_{δ}^2 ; otherwise if *z* and *z*' are far away from each other, the covariance function approaches zero indicating almost zero correlation among the two points.



Fig. 2. Current and voltage profile for model parameter calibration and OCV-SOC curve for Valence 26250 lithium-ion magnesium phosphate batteries.

Once the mean and the covariance function including the hyperparameters are determined based on the bias training data, the GP regression model can be used to predict the bias at any new battery operation conditions. Specifically, the mean of the bias at a set of new operation conditions Z_* is formulated as

$$\delta(Z_*) = m(Z_*) + K(Z_*, Z) K(Z, Z)^{-1} [\delta(Z) - m(Z)]$$
(8)

where $K(Z_*, Z)$ stands for $n_* \times n$ matrix of the covariance (i.e., Eq. (7)) calculated at all pairs of training and test points; nand n_* are the number of training and test points, respectively; and K(Z, Z) is the $n \times n$ covariance matrix at all training points. Compared to the polynomial regression model, the GP model is more accurate for predicting the local nonlinearity of the bias, which would further benefit the battery SOC estimation. The parameters in the GP model are set by maximizing the marginal likelihood [34] based on available bias training data. It is worth noting that the mean function, i.e., m(z), is not always necessary in the GP regression model. If a zero mean function was used, then the prediction in Eq. (8) is revised as

$$\delta(Z_*) = K(Z_*, Z) K(Z, Z)^{-1} \delta(Z)$$
(9)

Inclusion of an assumed mean function could, to some extent, improve the prediction accuracy of the GP regression model [34].



Battery parameter calibration based on the SPPC test and the 1st order RC model. Fig. 3.

D. EKF for SOC Estimation Considering Model Bias

Given a prior SOC estimation at the $(k-1)^{\text{th}}$ time step, the expectation of the SOC at the k^{th} time step is calculated from the state transition function $f(\bullet)$ as

$$x_{k|k-1} = f(x_{k-1}, i_k) \tag{10}$$

Based on the actual observation or evidence of the terminal voltage at the k^{th} time step, the mean of the posterior SOC is updated as

$$x_{k|k} = x_{k|k-1} + K_k \left(y_k^t - y_k^* \right) \tag{11}$$

where y_k^t is the actual voltage measurement; y_k^* is the model prediction considering the bias; and K_k is the Kalman gain at the k^{th} time step and is expressed in Eq. (12).

$$K_k = W_{k|k-1} \hat{H}_k V_k^{-1} \tag{12}$$

where $W_{k|k-1}$ is the SOC variance at the k^{th} time step given the process variance and the prior SOC variance at the $(k-1)^{\text{th}}$ time step; V_k is the overall voltage variance at the k^{th} time step considering the propagation of the SOC variance together with the measurement variance at the k^{th} time step; and \hat{H}_k is the first order derivative of the battery model with respect to the SOC given $x_{k|k-1}$. Considering the corrected battery model in Eq. (4),



Fig. 4. Battery parameter calibration based on the SPPC test and the 2nd order RC model.

 \hat{H}_k is computed as

$$\hat{H}_{k} = \frac{dy_{k}^{*}}{dx_{k}} = \left(\frac{dOCV(\bullet)}{dx_{k}} + i_{k}\frac{dR}{dx_{k}} + \sum_{q}\frac{dU_{q,k}}{dx_{k}} + \frac{d\delta_{k}(\bullet)}{dx_{k}}\right)$$
(13)

The first derivative term can be easily obtained from the battery OCV-SOC characteristic curve. The second and third terms can be calculated based on the parameter characteristic curves with respect to the SOC. The last term can be simply approximated using the finite difference method. It can also be analytically derived based on either Eq. (5) from the polynomial bias regression model or Eq. (8) from the GP regression model, which is expressed as

$$\frac{d\delta_k\left(\bullet\right)}{dx_k} = \frac{dm\left(\bullet\right)}{dx_k} + \frac{dK\left(\bullet, Z\right)}{dx_k}K(Z, Z)^{-1}\left[\delta\left(Z\right) - m\left(Z\right)\right]$$
(14)

Depending on the specific polynomial form of the mean function $m(\bullet)$, the first derivative term in Eq. (14) can be easily obtained. The second derivative term is a $1 \times n$ vector of the covariance, in which each element is computed as

$$\frac{dk\left(z_{k},z\right)}{dx_{k}} = \sigma_{\delta}^{2} exp\left[-\frac{1}{2}\frac{\left(z_{k}-z\right)\left(z_{k}-z\right)^{T}}{l^{2}}\right]\left(-\frac{x_{k}-x}{l^{2}}\right)$$
(15)

Finally, the overall voltage variance V_k can be formulated in Eq. (16) with the linear approximation of the battery model.

$$V_k = \hat{H}_k^2 W_{k|k-1} + var(v_k)$$
(16)

where $\operatorname{var}(v_k)$ is the measurement variance; and $W_{k|k-1} = \operatorname{var}(w_{k-1}) + \operatorname{var}(SOC_{k-1})$. In addition to the mean update of the SOC at the k^{th} time step in Eq. (11), its variance should be updated as well as shown in Eq. (17).

$$var(x_k) = W_{k|k-1} - K_k \hat{H}_k W_{k|k-1}$$
(17)

IV. EXPERIMENTS AND RESULTS

The objective of this section is to demonstrate the effectiveness of the proposed bias modeling method for battery SOC estimation. Valence 26250 lithium-ion magnesium phosphate batteries were employed in this study and all experiments were conducted at the room temperature. The static capacity test was conducted with 0.5C charge/discharge rates with 1-hour rest period in between. Three cycles of such tests were conducted and the average charge capacity was calculated as 2.5427 Ah. The self-defined pulse power characterization (SPPC) test as shown in Fig. 2a was conducted at every 5% change of the SOC and was combined with the OCV-SOC characterization test because they both need long rest periods.

The OCV was measured at every 5% change of the SOC with 0.5C discharge rate in between and a 2-hours resting period and the results are shown in Fig. 2b. It is worth noting that the OCV-SOC curve is fairly flat from 40%–60% and 75%–95% regions for this type of battery, which will make SOC estimation challenging using the traditional approach. For battery model bias training and validation, three cycle tests were conducted including the urban dynamometer driving schedule (UDDS) test,



(c) Mixture of UDDS and NYCC current profile and SOC trajectory

Fig. 5. Battery current profiles and true SOC trajectories.

the New York city cycle (NYCC) test, and the mixture of the two tests (i.e., UDDS followed by NYCC as a complete cycle). For these tests, the cycle currents were scaled such that the maximum current was 5C rate. The initial SOC was 90% and the cycles were repeated till the SOC reached 20%.

A. Parameter Identification of the Baseline Model

Given total twenty SPPC current and voltage profiles at every 5% SOC level reducing from 100% to 5%, model parameters of the 1st order and 2nd order RC model defined in Eqs. (2) and (3) were calibrated by minimizing the RMSE between the measured and predicted voltage as shown in Fig. 3a and Fig. 4a, respectively. Results of the voltage calibration error for the 1st



Fig. 6. SOC estimation using the EKF by assuming that the 2nd order RC model is the true model.

order RC model are shown in Fig. 3b, where most errors are bounded within ± 5 mV except at 5% SOC level where the error could reach ± 15 mV. This indicates that the RC model may have inherent inadequacy for representing the battery true behaviors at the low SOC range. As such, an extended ECM [24] may be needed if the objective is to estimate the battery SOC in the low SOC range. In this case study, however, SOC estimation is in the range from 20% to 90%. Hence, relatively large voltage calibration error at 5% SOC level is not a concern for demonstration of the proposed research. By employing the 2nd order RC model, the voltage calibration error can be further reduced to about ± 2 mV as shown in Fig. 4b except at 5% SOC level. It is worth noting that such calibration error could be further reduced if employing more RC pairs. However, it would be more likely to create an overfitting issue under real battery operating conditions. As such, only 1st and 2nd order RC models are employed in the case study. Results of the parameter calibration for the 1st and 2nd order RC model are shown in Fig. 3c–3f and Fig. 4c–4h, respectively.

B. SOC Estimation With an Assumed True Model

The goal of this section is to provide reference SOC estimation results using the EKF for three battery charging/discharging



(d) Current and bias profiles with bias RMSE of 0.018 $\rm V$

Fig. 7. Terminal voltage prediction and the voltage bias of the $1^{\rm st}$ order RC model under UDDS and NYCC loading profiles.



(a) Bias as a function of previous bias and current at 50% SOC level



(b) Bias as a function of current and SOC if pervious bias is fixed at 0

Fig. 8. Bias approximation using the polynomial regression model showing in a 3D space by fixing the 3^{rd} input variable.



(a) Bias as a function of previous bias and current at 50% SOC level



(b) Bias as a function of current and SOC if pervious bias is fixed at 0 (middle is the mean prediction with 95% confidence levels)





Fig. 10. Voltage comparison between model prediction and actual measurement under the mixed loading profile: (a) and (b) consider only the baseline model, (c) and (d) consider the corrected model using the polynomial regression, and (e) and f) consider the corrected model using the GP regression

profiles. In particular, the terminal voltage predicted from the 2nd order RC model was treated as actual voltage measurement such that there would be no model bias at all during the SOC estimation. The purpose is to firstly eliminate the influence of model bias in the SOC estimation and then examine the EKF performance under given parameter setting. The experiments were set up to discharge the battery from 90% SOC to 20% SOC and Fig. 5 shows three current profiles and corresponding SOC trajectories.

Without model bias, the only error source would be the EKF linearization error during the estimation. In addition,

the parameter setting in EKF including the process noise and measurement noise would influence the convergence rate. In this and following case studies, process noise was ignored by assigning an extremely small value (1e-8), and voltage measurement error was assumed to follow a Normal distribution with zero mean and 1 mV standard deviation. An initial guess of SOC was assigned as 80%, and the estimation performance is shown in Fig. 6 for all three current profiles. The results are as expected if model contains no bias at all, even though the OCV-SOC curve could be fairly flat as shown in Fig. 2b.



Fig. 11. Comparison of SOC estimation accuracy for the 1st order RC model and the corrected model with two bias modeling approaches for three battery loading profiles.

C. Bias Modeling and Approximation With the 1st Order *RC* Model

Even though accurate voltage prediction is observed under SPPC test with characterized model parameters, battery model prediction under various operating conditions could be inaccurate or exhibit different degree of voltage biases. The reason is because there is no one battery model (e.g., 1st order and 2nd order ECM) that can be applied to any battery operating conditions without any bias. As such, learning of model bias would increase the prediction accuracy of the baseline model under different operating conditions. It is worth noting that the bias model could have different dynamic model structure as the baseline model, hence, it would greatly increase the model prediction capability which is beyond simply conducting model parameter calibration.

This section conducts bias modeling and approximation using the polynomial regression model and the GP regression model. For demonstration purpose, bias training data were first obtained from two current loading profiles (e.g., UDDS and NYCC) based



Fig. 12. Terminal voltage prediction and the voltage bias of the 2nd order RC model under UDDS and NYCC loading profiles.

on the 1st order RC model, then bias models were constructed using both methods. Fig. 7 shows the voltage comparison between actual measurement and the model prediction from the 1st order RC model under UDDS and NYCC loading profiles. Overall, the model presents reasonable accuracy as shown in Fig. 7a and 7c. By plotting the bias vs. current profiles as shown in Fig. 7b and 7d, several facts can be observed. First of all, bias could be as low as zero or as high as about 50 mV (very rare) under UDDS profile, which is apparently current dependent. Under NYCC profile, bias is within 20 mV range for most of the time, but could increase to about -50 mV at the end of the profile. Secondly, since these current profiles are repeatedly applied to the battery and the bias does not show repeated pattern, the bias should also be dependent on the battery SOC level. Thirdly, biases are time series data and the value at current time step could be dependent on the value at previous time steps. Quantitatively, the bias RMSE is 14 mV and 18 mV for UDDS and NYCC, respectively.

Based on the bias training data and the model proposed in Eq. (5) for polynomial regression model and Eq. (8) or Eq. (9) for GP regression model, visualization of the bias approximation in a 3D space is shown in Fig. 8 and Fig. 9 for two different methods. In particular, the optimal fit for the polynomial regression model is shown in Eq. (18).

$$\delta_{k} = -0.0066 + 0.9711\delta_{k-1} + 2.1841\delta_{k-1}^{2} - 0.0002i_{k} + (1.95e - 5)i_{k}^{2} + 0.0001x_{k} - (1.42e - 6)x_{k}^{2} + 0.0055\delta_{k-1}i_{k} - 0.0013\delta_{k-1}x_{k} + (1.08e - 6)i_{k}x_{k}$$
(18)

For GP regression model, not only the mean prediction but also 95% confidence intervals are shown in Fig. 9. Compared to the polynomial regression model, the GP regression model captures not only the overall bias trend but also some local nonlinearity. In addition, confidence-based bias prediction could be conducted, if necessary, such that bias can be intentionally overestimated to a given confidence level. In this case study, however, only the mean prediction would be employed.

With the bias modeling using the polynomial and GP regression models, the corrected battery model shown in Eq. (4) can be obtained for any new battery operating conditions. Although such a statement is generally true, accuracy of the bias model depends upon the amount of training data. As a demonstration, terminal voltage prediction under the third loading profile, i.e., the mixed current profile, was calculated for the baseline model, the corrected model based on the polynomial regression, and the corrected model based on the GP regression. Fig. 10 shows detailed results and three important findings are observed. First of all, there are significant error reductions in terms of the RMSE for voltage estimation if model bias is included in the baseline model. Specifically, RMSE percentage reductions are 30.7% and 38.5% for the polynomial and GP regression model, respectively. Secondly, compared to the baseline model prediction, majority of voltage biases from the corrected model are distributed around zero as shown in Fig. 10d and 10f such that the corrected model



Fig. 13. Comparison of SOC estimation accuracy for the 2nd order RC model and the corrected model with two bias modeling approaches for three battery loading profiles.

could be overall considered as an unbiased model. Thirdly, there are some relatively large biases, e.g., up to 40 mV as shown in Fig. 10b, in the baseline model when the battery is subject to continuously large discharge current. The reduction of these biases in the corrected model is not significant using the polynomial model. The main reason is because these biases are not representative in the training data sets and the polynomial model simply minimizes the least square error from all data points. In the GP regression model, however, the model can better capture

such local nonlinearity if the data points are believed to be true value, i.e., not subject to any noise.

D. SOC Estimation Considering Model Bias With the 1st Order RC Model

To demonstrate the effectiveness of the bias modeling for SOC estimation, EKF was applied to three loading profiles (i.e., UDDS, NYCC, and the mixture of UDDS and NYCC) with the initial guess of the SOC as 70%. The true or reference SOC trajectory was calculated using the direct Coulomb counting knowing the true initial SOC as 90%. SOC estimations using the baseline model, i.e., the 1st order RC model, and the corrected model with two methods for bias modeling were compared to the true SOC till it degrades to 20%. In particular, cross validation was used in bias training and modeling such that the test profile is excluded from the training profile.

Fig. 11 shows the SOC comparison results for three current loading profiles, in which large estimation errors are observed for the baseline model. This is mainly due to the combination effect of the inherent bias of the baseline model and the flatness of the OCV-SOC curve. For the mixed loading profile as shown in Fig. 10b, the baseline model tends to slightly overestimate the terminal voltage in most of the time and underestimate the terminal voltage when there is a set of continuous large discharge currents. As such, even though the estimated SOC is continuously lower than the true SOC as shown in Fig. 11a, the predicted voltage due to the overestimation could perfectly match the measured voltage at higher SOC level. Consequently, the baseline model has relatively large SOC estimation error as shown in Fig. 11b. With the bias modeling of the baseline model using the polynomial and GP regression methods, terminal voltage estimation errors can be greatly reduced as shown in Fig. 10d and 10f. Consequently, both methods significantly improve the SOC estimation accuracy as shown in Fig. 11b. Furthermore, there is some degree of bias overcorrection for the polynomial regression model. As such, the estimated SOC is higher than the true SOC. The bias correction using the GP regression model performs consistently better than the polynomial regression model in this case study.

Similar observations can be made for the other two loading profiles as shown in Fig. 11c–11f. It is worth noting that SOC errors increase in Fig. 11f after about 18,000 seconds using the GP regression model. This is mainly because the baseline model increasingly overestimates the voltage for the NYCC profile after about 18,000 seconds as shown in Fig. 7d, and the bias model cannot capture such behaviors based on the training data sets from UDDS and the mixed profile. The bias correction using the polynomial regression, on the other hand, shows better performance in this region mainly because of the cancellation effect of previously overestimated SOC.

E. SOC Estimation Considering Model Bias With the 2nd Order RC model

This section repeats above work but changes the baseline model to the 2nd order RC model for two purposes. The first goal is to further demonstrate the effectiveness of the proposed work and the second goal is to obtain some insightful findings through the comparison of the performance difference.

With parameter characterization of the 2nd order RC model as shown in Fig. 4, bias of the baseline model can be again identified from the training data sets. Fig. 12 shows the voltage bias of the 2nd order RC model under UDDS and NYCC loading profiles. By comparing the results in Fig. 7 for the 1st order RC model, the 2nd order RC model shows much worse accuracy for the UDDS profile, but slightly better performance for the NYCC profile based on the bias RMSE. Even though a higher order RC model can fit the SPCC test profile better, it also has the risk of overfitting the data such that the accuracy becomes worse when applying the model to other loading profiles. In this case study, the bias RMSE is 27 mV for UDDS profile in the 2nd order RC model as compared to 14 mV in the 1st order RC model; and the bias RMSE is 13 mV for NYCC profile as compared to 18 mV using the 1st order RC model. As a briefly summary, a higher order RC model does not necessarily mean better accuracy. It is hence important to identify the most suitable battery model based on representative battery operating profiles. This study, however, does not focus on the identification of the most suitable baseline model but on the demonstration of the bias learning.

For simplicity, detailed results of bias modeling are not shown for the 2nd order RC model and the final results for the SOC estimations are shown in Fig. 13. Some observations are summarized as follows. First of all, SOC estimation using the 2nd order RC model performs worse for the mixed loading profile and the UDDS profile due to the overfitting issue, but performs better for the NYCC profile, as compared to the 1st order RC model. Secondly, the corrected model considering the model bias generally performs better than the baseline model, but the effect would be influenced by the baseline model. Basically, if there is an overfitting issue in the baseline model, the bias would be highly nonlinear, which makes bias modeling difficult for achieving the same accuracy. Thirdly, bias modeling using the GP regression generally performs better than the polynomial regression. GP regression is generally considered as a non-parametric model and its prediction is more data-driven compared to the polynomial regression.

V. CONCLUSION AND FUTURE WORK

This paper proposes model bias characterization for typical circuit-based lithium-ion battery models so that the baseline model accuracy can be significantly improved while maintaining similar computational efficiency. Under the same setting for battery SOC estimation, the corrected model incorporating the bias modeling consistently produces more accurate estimations than the baseline model. The GP regression model generally performs much better than the polynomial regression model. In addition to the previous bias history, current and SOC are two main factors used for the bias modeling in this paper. Even though the mean of the bias was actually used for the baseline model correction, noticeable uncertainties were observed from the training data, which indicates that other unknown factors should contribute to the bias as well. To further improve the confidence of the SOC estimation in particular applications, it may be necessary to include the uncertainty of the bias in the SOC estimation. To further enhance the expected SOC estimation accuracy, it is worthwhile to include more factors in the bias modeling.

For bias modeling considering temperature and degradation of the cell, the bias function in Eq. (4) can be extended by adding these two related parameters. As compared to the case study, there is no essential difference in terms of the bias learning process, but proper training data sets should be collected at different temperature and degradation levels. In addition, adding two more input variables with large amount of training data sets could pose accuracy and efficiency challenges for various machine learning methods including the GP regression model.

It should also be noted that the demonstration in Section IV is for one single battery cell which has its unique parameters after the SPPC test. In reality, different cells will exhibit parameter uncertainty due to the manufacturing tolerance. Hence, the bias learning should consider such cell-to-cell variability as well. It is almost impossible to track each cell's unique parameter set in real application. However, it is possible to conduct uncertainty modeling for battery model parameters such that each parameter may follow a certain distribution instead of a deterministic value over the SOC range. As such, any cell parameter realization can be treated as one random realization from the population distribution. Therefore, the deterministic parameter characterization such as the one shown in Fig. 3 and Fig. 4 would be replaced by probabilistic parameter representations. Then, the next step is to conduct bias modeling considering such parameter uncertainty. By picking one battery cell from the population and follow the proposed procedure in Section IV, bias of that cell can be identified such as the one shown in Fig. 8 and Fig. 9. If such process can be repeated for several cells, different bias realizations could be observed under different parameter settings observed from different cells. Then, the technical challenge is to construct such a model to build the relationship between model parameters and the model bias. Once this model is constructed, simulation can be conducted to accurately quantify the model uncertainty, i.e., uncertainty of the model bias, subject to the parameter uncertainty. In addition to the modeling technique for model uncertainty and parameter uncertainty characterization, data sufficiency issue, i.e., the number of cells used for uncertainty modeling, should be considered as well.

REFERENCES

- T. Fuller, M. Doyle, and J. Newman, "Simulation and optimization of the dual lithium ion insertion cell," *J. Electrochem. Soc.*, vol. 141, pp. 1–10, 1994.
- [2] K. A. Smith, C. D. Rahn, and C. Y. Wang, "Control oriented 1D electrochemical model of lithium ion battery," *J. Energy Convers. Manage.*, vol. 48, pp. 2565–2578, 2007.
- [3] L. Cai and R. E. White, "Reduction of model order based on proper orthogonal decomposition for lithium-ion battery simulations," *J. Electrochem. Soc.*, vol. 156, no. 3, pp. A154–A161, 2009.
- [4] C. W. Lee, Y. Hong, M. Hayrapetyan, X. G. Yang, and Z. Xi, "Derivation and tuning of a solvable and compact differential—Algebraic equations model for LiFePO₄—Graphite Li-ion batteries," *J. Appl. Electrochemistry*, vol. 48, no. 3, pp. 365–377, 2018.
- [5] X. Hu, S. Li, and H. Peng, "A comparative study of equivalent circuit models for Li-ion batteries," *J. Power Sources*, vol. 198, pp. 359–367, 2012.
- [6] S. Lee, J. Kim, J. Lee, and B. H. Cho, "State-of-charge and capacity estimation of lithium-ion battery using a new open-circuit voltage versus state-of-charge," J. Power Sources, vol. 185, no. 2, pp. 1367–1373, 2008.
- [7] B. Xia, X. Zhao, R. de Callafon, H. Garnier, T. Nguyen, and C. Mi, "Accurate lithium-ion battery parameter estimation with continuous-time system identification methods," *Appl. Energy*, vol. 179, pp. 426–436, 2016.
- [8] 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. Veh. Technol.*, vol. 62, no. 3, pp. 1020–1030, Mar. 2013.

- [9] J. C. Alvarez Anton, P. J. Garcia Nieto, C. Blanco Viejo, and J. A. Vilan Vilan, "Support vector machines used to estimate the battery state of charge," *IEEE Trans. Power Electron.*, vol. 28, no. 12, pp. 5919–5926, Dec. 2013.
- [10] J. Du, Z. Liu, and Y. Wang, "State of charge estimation for Li-ion battery based on model from extreme learning machine," *Control Eng. Practice*, vol. 26, no. 1, pp. 11–19, 2014.
- [11] Y. Jiang, B. Xia, X. Zhao, T. Nguyen, C. Mi, and R. A. de Callafon, "Databased fractional differential models for non-linear dynamic modeling of a lithium-ion battery," *Energy*, vol. 135, pp. 171–181, 2017.
- [12] F. Auger, M. Hilairet, J. M. Guerrero, E. Monmasson, T. Orlowska-Kowalska, and S. Katsura, "Industrial applications of the Kalman filter: A review," *IEEE Trans. Ind. Electron.*, vol. 60, no. 12, pp. 5458–5471, Dec. 2013.
- [13] C. Hu, B. D. Youn, and J. Chung, "A multiscale framework with extended Kalman filter for lithium-ion battery SOC and capacity estimation," *Appl. Energy*, vol. 92, pp. 694–704, 2012.
- [14] 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. J. Elect. Power Energy Syst.*, vol. 62, pp. 783–791, 2014.
- [15] X. Liu, Z. Chen, C. Zhang, and J. Wu, "A novel temperature-compensated model for power Li-ion batteries with dual-particle-filter state of charge estimation," *Appl. Energy*, vol. 123, pp. 263–272, 2014.
- [16] R. Jing, Z. Xi, X.G. Yang, and E. Decker, "A systematic framework for battery performance estimation considering model and parameter uncertainties," *Int. J. Prognostics Health Manag.*, vol. 5, no. 2, pp. 1–10, 2014.
- [17] S. Mendoza, J. Liu, P. Mishra, and H. Fathy, "On the relative contributions of bias and noise to lithium-ion battery state of charge estimation errors," *J. Energy Storage*, vol. 11, pp. 86–92, 2017.
- [18] X. Lin, "Theoretical analysis of battery SOC estimation errors under sensor bias and variance," *IEEE Trans. Ind. Electron.*, vol. 65, no. 9, pp. 7138–7148, Sep. 2018.
- [19] C. Forgez, D. V. Do, G. Friedrich, M. Morcrette, and C. Delacourt, "Thermal modeling of a cylindrical LiFePO₄/graphite lithium-ion battery," *J. Power Sources*, vol. 195, no. 9, pp. 2961–2968, 2010.
- [20] W. Waag, S. Käbitz, and D. U. Sauer, "Experimental investigation of the lithium-ion battery impedance characteristic at various conditions and aging states and its influence on the application," *Appl. Energy*, vol. 102, pp. 885–897, 2013.
- [21] K. Chatterjee, P. Majumdar, D. Schroeder, and S. R. Kilaparti, "Performance analysis of Li-ion battery under various thermal and load conditions," *J. Electrochem. Energy Convers. Storage*, vol. 16, no. 2, 2019, Art. no. 021006.
- [22] X. Lin, A. Stefanopoulou, P. Laskowsky, J. Freudenberg, Y. Li, and R. D. Anderson, "State of charge estimation error due to parameter mismatch in a generalized explicit lithium ion battery model," in *Proc. ASME Dyn. Syst. Control Conf. Bath/ASME Symp. Fluid Power Motion Control*, 2011, vol. 1, pp. 393–400.
- [23] M. Dahmardeh and Z. Xi, "State of charge estimation for lithium-ion battery packs considering cell-to-cell variability," ASCE-ASME J. Risk Uncertainty Eng. Syst., Part B. Mech. Eng., vol. 5, pp. 020909-1–020909-9, 2019.
- [24] M. Ouyang, G. Liu, L. Lu, J. Li, and X. Han, "Enhancing the estimation accuracy in low state-of-charge area: A novel onboard battery model through surface state of charge determination," *J. Power Sources*, vol. 270, pp. 221–237, 2014.
- [25] P. P. Mishra, M. Garg, S. Mendoza, J. Liu, C. D. Rahn, and H. K. Fathy, "How does model reduction affect lithium-ion battery state of charge estimation errors? Theory and experiments," *J. Electrochem. Soc.*, vol. 164, no. 2, pp. A237–A251, 2017.
- [26] G. L. Plett, "Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs—Part 3. State and parameter estimation," *J. Power Sources*, vol. 134, no. 2, pp. 277–292, 2004.
- [27] L. Wang and C. Singh, "Multicriteria design of hybrid power generation systems based on a modified particle swarm optimization algorithm," *IEEE Trans. Energy Convers.*, vol. 24, no. 1, pp. 163–172, Mar. 2009.
- [28] M. C. Kennedy and A. O'Hagan "Bayesian calibration of computer models," J. Roy. Statist. Soc., Ser. B., vol. 63, pp. 425–464, 2001.
- [29] H. Pan, Z. Xi, and R.-J. Yang, "Model uncertainty approximation using a copula-based approach for reliability based design optimization," *Structural Multidisciplinary Optim.*, vol. 54, no. 6, pp. 1543–1556, 2016.
- [30] L. Shi and S.-P. Lin, "A new RBDO method using adaptive response surface and first-order score function for crashworthiness design," *Rel. Eng. Syst. Saf.*, vol. 156, pp. 125–133, 2016.

- [31] M.-Y. Moon, K. K. Choi, H. Cho, N. Gaul, D. Lamb, and D. Gorsich, "Reliability-based design optimization using confidence-based model validation for insufficient experimental data," *J. Mech. Des.*, vol. 139, no. 3, 2017, Art. no. 031404.
- [32] Z. Xi, "Model-based reliability analysis with both model uncertainty and parameter uncertainty," J. Mech. Des., vol. 141, no. 5, 2019, Art. no. 051404.
- [33] J. Quiñonero-Candela and C. E. Rasmussen, "A unifying view of sparse approximate Gaussian process regression," J. Mach. Learn. Res., vol. 6, pp. 1939–1959, 2005.
- [34] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.



Zhimin Xi received the B.S. and M.S. degrees in mechanical engineering from the University of Science and Technology Beijing, Beijing, China, in 2001 and 2004, respectively, and the Ph.D. degree in reliability engineering from the University of Maryland, College Park, MD, USA, in 2010. He is currently an Assistant Professor with the Department of Industrial and Systems Engineering, Rutgers University, New Brunswick, NJ, USA. His research interests include reliability and safety for energy storage systems, prognostics and health management, model valida-

tion under uncertainty, and system reliability analysis.



Modjtaba Dahmardeh received the B.S. degree from the Iran University of Science and Technology, Tehran, Iran, in 2011, and the M.S. degree from the Sharif University of Technology, Tehran, Iran, in 2014. He is currently working toward the Ph.D. degree with Rutgers University, New Brunswick, NJ, USA. His research interests include machine learning, Bayesian statistics, and stochastic processes in uncertainty quantification of engineering systems.



Yuhong Fu received the M.S. degree in automotive systems engineering from the University of Michigan-Dearborn, Dearborn, MI, USA, in December 2011. She is currently with San Diego State University Research Foundation, San Diego, CA, USA. Her research interest is in battery management for electric vehicles.

Bing Xia (S'13) received the B.S. degree in mechanical engineering from the University of Michigan,

Ann Arbor, MI, USA, and the B.S. degree in electri-

cal engineering from Shanghai Jiaotong University,

Shanghai, China, in 2012, and the Ph.D. degree from

San Diego State University, San Diego, CA, USA, and

University of California, San Diego, CA, USA. His

research interests are in battery management systems.

Chunting Chris Mi (S'00–A'01–M'01–SM'03– F'12) received the B.S. and M.S. degrees in electrical engineering from Northwestern Polytechnical University, Xi'an, China, in 1985 and 1988, respectively, and the Ph.D. degree in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 2001. He is currently the Chair and a Professor in electrical and computer engineering, San Diego State University, San Diego, CA, USA. From 2001 to 2015, he was with the University of Michigan, Dearborn. His research interests include electric drives, power

electronics, battery management systems, and wireless power transfer.