A Data-Driven Framework to Predict Lithium-Ion Battery Cell Imbalance for Real-Time Battery Management Systems

Chao Li and Assimina A. Pelegri *

Mechanical and Aerospace Engineering, Rutgers, The State University of New Jersey, Piscataway, NJ 08854-8058, USA
* Correspondence: pelegri@rutgers.edu; Tel.: +1-848-445-0691

Abstract: Models that can predict battery cells’ thermal and electrical behaviors are necessary for real-time battery management systems to regulate the imbalance within battery cells. This work introduces a Gaussian Process Regression (GPR)-based data-driven framework that succeeds the Multi-Scale Multi-Dimensional (MSMD) modeling structure. The framework can make highly accurate predictions at the same level as full-order full-distribution simulations based on MSMD. A pseudo-2D model is used to generate training data and is combined with a process that shifts computation burdens from real-time battery management systems to lab data preparation. The testing results highlight the reliability of the GPR-based data-driven framework in terms of accuracy and stability under various operational conditions.

Keywords: imbalance within battery cells; GPR-based data-driven framework; predict the thermal and electrical behaviors of battery cells

1. Introduction

In applications such as electric vehicles (EVs), the imbalance of electrical potential, temperature, and the State of Charge (SOC) within battery cells may cause problems such as thermal runaway or short usage life [1,2]. Some methods such as Barsukov’s [3] have been applied to regulate the imbalance, which monitors battery cells’ electrical and thermal behavior with sensors and balance the cells via circuits and cooling devices. Although this method is straightforward and effective, many necessary sensors and balancing circuits render battery systems complicated and bulky [4]. Without a breakthrough in developing new materials, recent relevant industrial improvements through better designs in sensors and circuits have been minimal [5]. Therefore, mathematical models that predict the behavior of battery cells becomes necessary for optimizing the battery design and eliminate the inherent imbalance within battery cells. Additionally, sensors and circuits may be reduced by providing real-time battery management that can act early through predictive computational algorithms to prevent battery imbalance.

This research aims to introduce a mathematical model for battery cells applied to real-time battery management systems. Due to the critical attributes of Li-ion battery cells (they provide one of the best tradeoffs in terms of power density, low weight, cell voltage, and low self-discharge according to [6]), they constitute the focus of this paper without any loss of generality.

Mathematical models for Li-ion battery dynamics generally fall within two main categories: the equivalent circuit models and the electrochemical models [7,8]. Popular equivalent circuit models include the Rin model and its later extension, the RC model, introduced by [9]. The equivalent circuit algorithms use electrical components to model battery cells’ dynamic behavior based on empirical knowledge. These algorithms are advantageous due to their simplicity and fast computational speed, but their straightforward construction also creates inherent shortcomings, resulting in an accuracy that hinges upon restricted working conditions or temperature ranges [1,10]. To the authors’ best knowledge,
the most recent update of the equivalent circuit model is the Thevenin model [11, 12], which enables State of Charge (SOC) estimations rather than improving the accuracy of the model while extrapolating beyond restrictions. Thus, equivalent circuit models are not adaptable enough to effectively be applied in real-time Li-ion battery management systems, especially considering all the coupled electrochemical/physical phenomena present in Electric Vehicles (EVs).

On the other hand, electrochemical models that describe detailed electrochemical/physical battery dynamics, for example, cf. [13], have the advantages of predicting the battery cells’ thermal, electrical, and mechanical responses under various conditions. To address the interplay among the physics in varied scales in large-format battery cells, which is necessitated when handling battery cell imbalances, Kim et al. [14] introduced the Multi-Scale Multi-Dimensional (MSMD) framework. MSMD incorporates the Pseudo-2D (P2D) model developed by [15] as a sub-mode and has been widely recognized as a well-suited framework to achieve the complexity of physical phenomena brought by the scale-up of the batteries. However, full-order full-distribution simulations based directly on the MSMD and its sub-models involve solving an extensive system of coupled nonlinear Differential-Algebraic Equations (DAEs), which is computationally expensive and challenging to implement in real-time simulations.

In [16], Guo and White summarized the advantages and disadvantages of various Reduced-Order Models (ROMs) [17–20] that can be applied within the MSMD framework and derived their own ROM to balance simulation time efficiency and accuracy. In that work, the group simplifies, in varying degrees, the sub-domains of the MSMD to reduce computation time while maintaining tolerable accuracy. Our work presents a Gaussian Process Regression (GPR)-based data-driven framework with many advantages over the ROMs in battery cell behavior simulations or predictions. It shifts computation burdens from real-time battery management systems to lab data preparation. Simultaneously, it inherits the accuracy from the full-order full-distribution simulations based on MSMD, provides confidence intervals for predictions, and can employ empirical knowledge just as equivalent circuit models do.

Data-driven approaches build models directly from observations instead of a theoretical understanding of underlying physical principles [21]. They circumvent the solution’s computational complexities common in physical-principle-based models, and their main tasks are on data measurements and data preparations, most of which can be conducted offline [22]. They are easier to update since no model simplification assumptions restrict them as in physical-principle-based models [23]. Essentially, data-driven approaches estimate functions at inputs given some observations, and the various data-driven approaches differ from each other in that they may be subject to different model classes and describe their estimations over functions differently. Since a data-driven approach for battery modeling needs to encode details within the complex battery dynamics, it should be subject to a model class that effectively considers a wide variety of candidate functions. Frequently used model classes where data fit polynomials would not be appropriate because of practical function selections’ limitations. In such approaches, a single function under a pre-determined model class is ultimately selected merely based on the “art” of setting regularization term parameters among infinite possibilities; furthermore, there might be many other model classes to be considered [24]. Another drawback is the lack of uncertainty representation in their function estimations [25]. Data-driven approaches for battery modeling should quantify the uncertainty to clarify the degree of trust we can place in its output while working in a battery management system. All these lead us to the application of Gaussian Process Regression (GPR).

This work establishes an iterative framework that predicts a battery cell’s thermal and electrical behaviors in a real-time battery management system. Efficient methods are produced to train GPRs with deterministic inputs, and the corresponding probabilistic Finite Element Analysis (FEA) scheme applies. The work considers a Li-ion electrode...
pair to exemplify the GPR-based data-driven framework application on a single Li-ion battery cell.

2. The GPR-Based Data-Driven Framework

Section 2.1 first introduces the battery cell structure (a Li-ion electrode pair) and the relevant model assumptions and produces a summary flowchart of the data-driven framework. The modeling scales and domains follow the Multi-Scale Multi-Dimensional (MSMD) framework definitions [14]. Section 2.2 focuses on the particle and electrode sub-domains and explains how the Gaussian Process Regression models are trained and how the predictions work parallel in the first timestep. Finally, in Section 2.3, we ascend to the battery cell domain, in which finite element analysis and dimension reduction methods are applied together to process the GPR predictions from the electrode level; the resultant electrical/thermal solution distributions over the cell are acquired before the next timestep.

2.1. The Modeling Scales

Following the MSMD framework by [14] for solving the electrical and thermal behaviors in one single battery cell, each battery cell is composed of the following three sub-domains: (i) cell domain, (ii) porous electrode domain, and (iii) particle domain.

Here, a single electrode plate pair constitutes the cell domain. The pair is composed of negative and positive electrodes, current collectors attached to their outside surfaces, and a separator sandwiched between them, as shown in Figure 1.

The summary of the governing equations and their solution variables of the sub-domains are given by Kim et al. [14] and are listed in Table 1. Details of the parameters discussed in Table 1 are summarized below.

**Table 1. Summary of solution variables and governing equations of the sub-models.**

| Domain | Governing Equation | Solution Variable |
|--------|-------------------|-------------------|
| Particle | \( \dot{c}_c = k_i(c_e)\alpha_0 (c_{i,max} - c_{i,eq})^\beta \) \{ \exp \left( \frac{\alpha_0}{\Theta} \right) - \exp \left( \frac{\alpha_0}{\Theta} \right) \} \left( \partial \frac{\Theta}{\partial x} \right)_{r=R_s} = -\frac{i_s}{\alpha_i} | \( \dot{c}_c(r, x, Y, Z) \) |
| Electrode | \( \frac{\partial (\alpha_0 c_e^\beta \phi)}{\partial y} = \frac{\partial (\alpha_0 c_e^\beta \phi)}{\partial y} + \frac{\partial (\alpha_0 c_e^\beta \phi)}{\partial y} \ln c_e \) + \( \frac{\partial \phi}{\partial x} \) \left( \partial \frac{\Theta}{\partial x} \right)_{x=0} = 0, \phi_e(x, Y, Z) | \( \phi_e(x, Y, Z) \) |
| Cell | \( \frac{\partial \phi}{\partial y} \left( \partial \frac{\Theta}{\partial y} \right)_{y=L} + \frac{\partial \phi}{\partial y} \left( \partial \frac{\Theta}{\partial y} \right)_{y=L} + \left. \frac{\partial \phi}{\partial y} \right|_{y=0} = 0, \phi_{e,c}(x, Y, Z) | \( \phi_{e,c}(x, Y, Z) \) |
Figure 1. Depictions of a battery cell constituted of coupled positive and negative plates. (a) Top view of plate partitioned into 16 subareas. The relatively large rectangular segment on the left side corresponds to the current collector tab of each plate. (b) 3D depiction of cell’s top view. Top layer (yellow) is the positive plate, and bottom layer (grey) is the negative plate. The scales show the dimension of the battery cell (Unit: m).

In the particle domain, $i_{ξ}''$ is the transfer current density at the particle surface, and $c_s$ is the lithium concentration within a solid electrode particle. In the electrode domain, $j_{x}'''$ is the volumetric reaction current at the electrode composite volume, which is evaluated with $\overline{i}_{ξ}''$, the averaged transfer current density over the active particle surfaces. In the cell domain, $q_{Θ}'''$ is the volumetric current source in the current collectors, which is determined by $j_{x}'''$, the current density in the electrode. $q_{Θ}'''$ is the heat source. Here the triple prime denotes volumetric terms (e.g., current per volume), while the double prime denotes density terms (e.g., current per area).

The cell is treated in 2D to be compared with the works by Kim [14] and Guo [16], and the electrodes and the current collectors attached to their outside surfaces are treated together to discuss the electrical and thermal behaviors, named positive electrode plate and the negative electrode plate, respectively.

Each of the electrode plates is partitioned into 16 subareas, cf. Figure 1a. Note that the current collector tabs, i.e., the relatively large rectangular segments on the left side of the electrode plates connecting with the external circuits, are not of interest in this work.
since no electrochemical process is active there. All the inputs are homogenized within each subarea before starting each timestep, while the temperature distribution is not. The temperature distribution is used as the initial condition to solve for the subsequent cell domain temperature distribution.

The Pseudo-2D (P2D) model, which describes the particle and electrode domain dynamics [15], is used to prepare the necessary data for training GPR models, cf. Figure 2. The process is separated into multiple timesteps to simulate an entire process of battery charge or discharge.

Figure 2. Each new timestep starts with training GPR models. GPR models and the cell models solve for the temperature and electrical potential distributions on the plates in 2D in each timestep, and they become inputs for the next timestep.
Each new timestep starts with training GPR models to predict electrical current source, heat source, and updated SOC (uSOC); these results are then fed to cell models to solve for the potential and temperature distributions through the cell structure in 2D (cf. Figure 1), which then become inputs for the next timestep.

In the first timestep, the inputs to the GPR models are deterministic. The predicted electrical current source, heat source, and uSOC are not correlated, and the predictions are parallel; see Section 2.2. However, in the second timestep and thereafter, the inputs to the GPR models are uncertain and subject to probability distributions resulting in non-parallel predictions and correlated predicted results.

2.2. Gaussian Process Regression Models

The lithium diffusion dynamics and charge transfer kinetics of the particle domain and the electrode domain can be solved with the P2D porous electrode model introduced by [15]. Solving the set of highly nonlinear and tightly coupled Differential-Algebraic Equations (DAEs) from this model is a computationally intensive process. Here, we will introduce GPR-based data-driven approaches to solve these DAEs only to generate the training data, and in real-time battery management processes, the recurring solution of the DAEs will be bypassed.

Data are obtained from the P2D model-based simulations with different combinations of inputs. The model simulates a potentiostatic process of the battery cell in the particle and electrode domains for a fixed time interval (timestep). For a specific battery cell under potentiostatic charge/discharge, we fix all the other changeable settings in the P2D model to constants and keep reference voltage, working temperature, and initial State of Charge (iSOC) as the three input variables. The combinations of these three variables compose the training input data pool. The output values of electrical current source, heat source, and uSOC after one timestep compose the training target data pool; these are obtained by solving the P2D model for a one-time step process. The training input data is then paired with the training target data from the observations we have available.

Gaussian Process Regression is categorized as a nonparametric regression method that gives probabilistic models through extracting the features within the data and provides confidence information [26]. The data space \( X = [x_1, x_2, \ldots, x_N]^T \) represents the \( N \) training input data selected from the training input data pool, and \( y \) represents the corresponding training target data. GPRs give the complete posterior distributions over possible functions under which the function values at any and every point are jointly Gaussian

\[
f(X) \sim \mathcal{N}(m(X), K(X, X)) \tag{1}
\]

where the matrix \( K \) has entries \( K_{ij} \) representing \( k(x_i, x_j) \), the covariance function takes two arguments and returns the covariance between their corresponding function values. The mean function \( m \) and the covariance function \( k \) fully determine the Gaussian Process. According to [27], the GPR is sufficiently expressive when specifying the mean function \( m(x) = 0 \).

This paper applies the squared exponential covariance function, the hyperparameters of which are denoted as \( \theta \). The covariance between function values is determined by the corresponding input locations and the covariance function with specified \( \theta \) hyperparameters and is not related to the actual values of the function. The process of optimizing the hyperparameters to make the covariance and thus the posterior distribution on functions agree most with the observed data is realized by minimizing the negative log marginal likelihood defined as \( \text{NLML} = -\log p(y|X, \theta) \) as in the work of [28].

We can find the hyperparameters through

\[
\hat{\theta} \in \arg\min_{\theta} -\log p(y|X, \theta) \tag{2}
\]
where, \( y \) is the observation, which is expressed by the distribution of functions in addition to the Gaussian distributed observation noise.

\[
y = f(x) + \varepsilon
\]  

(3)

where \( \varepsilon \sim \mathcal{N}(0, \sigma^2) \). The consequent log marginal likelihood is

\[
\log p(y|X, \theta) = -\frac{1}{2} \langle y - \mathbf{m}(X) \rangle^T (\mathbf{K} + \sigma^2 \mathbf{I})^{-1} \langle y - \mathbf{m}(X) \rangle - \frac{1}{2} \mathbf{m}(X)^T (\mathbf{K} + \sigma^2 \mathbf{I})^{-1} \mathbf{m}(X) - \frac{N}{2} \log 2\pi
\]  

(4)

At the given prediction (testing) inputs \( X^* \), we use the jointly Gaussian distribution to make predictions of the new states \( f^* \),

\[
p(f^*|X^*, X, y) = \mathcal{N}(\mathbf{m}, \mathbf{S})
\]  

(5)

where,

\[
\mathbf{m} = \mathbf{m}(X^*) + \mathbf{K}(X^*, X) \left( \mathbf{K} + \sigma^2 \mathbf{I} \right)^{-1} \langle y - \mathbf{m}(X) \rangle
\]  

(6)

and

\[
\mathbf{S} = \mathbf{K}(X^*, X^*) - \mathbf{K}(X^*, X) \left( \mathbf{K} + \sigma^2 \mathbf{I} \right)^{-1} \mathbf{K}(X, X^*)
\]  

(7)

The same set of training inputs \( X = [x_1, x_2, \ldots, x_N]^T, x_i \in \mathbb{R}^D \), but different training targets \( y_a = [y_{a1}, y_{a2}, \ldots, y_{aD}]^T, a = 1, \ldots, E \), can train \( E \) independent GPR models. In this work, the GPRs predict the one-timestep-after heat source, electrical current source, and usSOC, with the reference voltage, working temperature, and iSOC as the three input variables, so \( D = 3 \) and \( E = 3 \). As shown in the works of [26,29], the target values do not covary with each other since the training and prediction inputs are deterministic; that is, the predictions are parallel.

2.3. Probabilistic Finite Element Analysis of Battery Cells

The electrode plate potentials and the temperature in the cell domain are governed by three Partial Differential Equations (PDEs). As listed in Table 1, the governing PDEs are

\[
\frac{\partial}{\partial X} \left( \sigma \frac{\partial \Phi_\vdash}{\partial X} \right) + \frac{\partial}{\partial Y} \left( \sigma \frac{\partial \Phi_\vdash}{\partial Y} \right) - j^\varepsilon = 0
\]  

(8)

\[
\frac{\partial}{\partial X} \left( \sigma \frac{\partial \Phi_u}{\partial X} \right) + \frac{\partial}{\partial Y} \left( \sigma \frac{\partial \Phi_u}{\partial Y} \right) + j^m = 0
\]  

(9)

\[
\frac{\partial (\rho c_p T)}{\partial t} = \frac{\partial}{\partial X} \left( k_p \frac{\partial T}{\partial X} \right) + \frac{\partial}{\partial Y} \left( k_p \frac{\partial T}{\partial Y} \right) + \frac{\partial}{\partial Z} \left( k_i \frac{\partial T}{\partial Z} \right) + q^m
\]  

(10)

The cell is separated into 16 subareas in 2D, cf. Figure 1, and each grid has its heat source and electrical current source input. Thus, \( j^\varepsilon \) and \( q^m \), \( i = 1, 2, \ldots, 16 \), form the variable vectors \( j^\varepsilon \) and \( q^m \).

The heat source term \( q^\varepsilon \) and the electrical current source term \( j^m \) are predicted by GPRs. They are independent random variables following Gaussian distributions in the first timestep. The PDEs are to be solved with these random variables as input terms, and the solutions will thus be in the form of probability distributions. Here, the Multiplicative Dimensional Reduction Method (M-DRM) [30] is combined with finite element analysis (FEA) to find these solution distributions; hereafter this process is denoted as probabilistic FEA.

Noticing that the negative electrode plate potential \( \Phi_\vdash \), the positive electrode plate potential \( \Phi_\vdash \), and the cell temperature \( T \) are not coupled among the three governing PDEs Equations (8)–(10); thus, they can be solved independently of the probabilistic FEA. For example, the flowchart in Figure 3 explains how to use probabilistic FEA to find the solution distribution of the positive electrode plate potential \( \Phi_+ \) at a designated location.
The procedure of the M-DRM application in [31] is followed here, which incorporates the Maximum Entropy principle [32] with fractional moment constraints [33].

In this procedure, cut functions are defined at the beginning for each PDE. Normal FEAAs can solve their results, and their multiplication approximates the response of the PDE at the variable vector corresponding to the cuts.

Additionally, to find the statistical moments of the response, an adapted Gauss–Hermite quadrature method is derived; it allocates weighted cuts for the cut functions and splices the approximated moments. Finally, the Maximum Entropy principle is applied to find the solution distribution from the statistical moments. The details are shown step by step in the rest of this section.

**Figure 3.** Solution process flowchart of the probabilistic FEAs resulting in the solution distribution of the positive electrode plate potential.
Firstly, the solution of $\Phi_+(X = x, Y = y)$ in Equation (8) with $f^m$ as input is defined as function $h_p(f^m)$, the solution of $\Phi_-(X = x, Y = y)$ in Equation (9) with $f^m$ as input is defined as function $h_N(f^m)$, and the solution of $T(X = x, Y = y)$ in Equation (10) with $q^m$ as input is defined as function $h_T(q^m)$.

In this context, the $i$-th cut functions of the governing PDEs when $i \geq 1$ is defined as

$$h_p\left(f^m_{(i)}\right) = h_p\left(c_1, \ldots, c_{i-1}, f^m_{O}, c_{i+1}, \ldots, c_n\right)$$

$$h_N\left(f^m_{(i)}\right) = h_N\left(c_1, \ldots, c_{i-1}, f^m_{O}, c_{i+1}, \ldots, c_n\right)$$

$$h_T\left(q^m_{(i)}\right) = h_T\left(c_1, \ldots, c_{i-1}, q^m_{O}, c_{i+1}, \ldots, c_n\right)$$

With the zeroth cut functions described as

$$h_{P0} = h_p(c_1, \ldots, c_i, \ldots, c_n)$$

$$h_{N0} = h_N(c_1, \ldots, c_i, \ldots, c_n)$$

$$h_{T0} = h_T(c_1, \ldots, c_i, \ldots, c_n)$$

where, $c_i$s are mean values of the corresponding variables.

The response function of a PDE is the solution, in our case the numerical solution from the FEA to the PDE given the variable vector. The response functions of the governing PDEs in Equations (8)–(10) are approximated by

$$Y_P = h_p(f^m) \approx h_{P0}(1-n) \times \prod_{i=1}^{n} h_{P1}\left(f^m_{(i)}\right)$$

$$Y_N = h_N(f^m) \approx h_{N0}(1-n) \times \prod_{i=1}^{n} h_{Ni}\left(f^m_{(i)}\right)$$

$$Y_T = h_T(q^m) \approx h_{T0}(1-n) \times \prod_{i=1}^{n} h_{Ti}\left(q^m_{(i)}\right)$$

Next, the statistical moments of the response functions are calculated. From the definition of the moments, the $\alpha$-th statistical moments of the response functions are

$$M^P_{Y_P} = E[Y_P^\alpha] = \int_{Y_P} y^\alpha f_{Y_P}(y_P)dy_P$$

$$M^N_{Y_N} = E[Y_N^\alpha] = \int_{Y_N} y^\alpha f_{Y_N}(y_N)dy_N$$

$$M^T_{Y_T} = E[Y_T^\alpha] = \int_{Y_T} y^\alpha f_{Y_T}(y_T)dy_T$$

where $f_{Y_P}(y_P)$, $f_{Y_N}(y_N)$, and $f_{Y_T}(y_T)$ are the probability density functions (PDFs) of $Y_P$, $Y_N$, and $Y_T$, respectively.

To avoid possible computational difficulties from the integrations, we use the approximations in Equations (17)–(19) instead to further approximate the $\alpha$-th statistical moments of the response functions

$$M^P_{Y_P} \approx \left(\frac{h_{P0}(1-n)}{\prod_{i=1}^{n} h_{P1}\left(f^m_{(i)}\right)}\right)^\alpha \times E\left[\left(h_{P1}\left(f^m_{O}\right)\right)^\alpha\right] \times \ldots \times E\left[\left(h_{Pn}\left(f^m_{O}\right)\right)^\alpha\right]$$

$$M^N_{Y_N} \approx \left(\frac{h_{N0}(1-n)}{\prod_{i=1}^{n} h_{Ni}\left(f^m_{(i)}\right)}\right)^\alpha \times E\left[\left(h_{N1}\left(f^m_{O}, 1\right)\right)^\alpha\right] \times \ldots \times E\left[\left(h_{Nn}\left(f^m_{O}, n\right)\right)^\alpha\right]$$

$$M^T_{Y_T} \approx \left(\frac{h_{T0}(1-n)}{\prod_{i=1}^{n} h_{Ti}\left(q^m_{(i)}\right)}\right)^\alpha \times E\left[\left(h_{Ti}\left(q^m_{O}\right)\right)^\alpha\right] \times \ldots \times E\left[\left(h_{Tn}\left(q^m_{O}\right)\right)^\alpha\right]$$
To find the moments $E\left[\left(h_{P_1}\left(f''_{\Theta}\right)\right)^{\alpha}\right]$, $E\left[\left(h_{N_1}\left(f''_{\Theta}\right)\right)^{\alpha}\right]$, and $E\left[\left(h_{T_1}\left(q''_{\Theta}\right)\right)^{\alpha}\right]$, the Gauss–Hermite quadrature method can be applied. The steps from Equations (26) through (33) explain how the Gauss–Hermite quadrature method is derived for application in our case.

According to [34], an integral of the form $\int e^{-x^2}g(x)dx$ can be assessed as

$$
\int e^{-x^2}g(x)dx \approx \sum_{j=1}^{L} w_j g(x_j)
$$

where $w_j$ and $x_j$ are weights and nodes given by this approximation.

So, if the expectation of a function $H(Z)$ where $Z$ is standard normal is estimated, Equation (26) can be rewritten as

$$
E(H(Z)) = \int f_Z(z)H(z)dz
= \int \frac{1}{\sqrt{2\pi}} \exp\left\{-\left(\frac{z}{\sqrt{2}}\right)^2\right\} H(z)dz
= \int \frac{1}{\sqrt{2\pi}} \exp\left\{-\left(\frac{z}{\sqrt{2}}\right)^2\right\} H(z)\sqrt{2}d\left(\frac{z}{\sqrt{2}}\right)
= \int \exp\left\{-\left(\frac{z}{\sqrt{2}}\right)^2\right\} \tilde{g}\left(\frac{z}{\sqrt{2}}\right)d\left(\frac{z}{\sqrt{2}}\right)
$$

where $\tilde{g}\left(\frac{z}{\sqrt{2}}\right) = \frac{1}{\sqrt{2\pi}} H(z)$. Subsequently, the expectation is approximated by

$$
E(H(Z)) = \sum_{j=1}^{L} q_j \tilde{g}\left(\frac{z_j}{\sqrt{2}}\right) = \sum_{j=1}^{L} q_j H(z_j)
$$

With our proposed weights $q_j = \frac{1}{\sqrt{\pi}} w_j$, and nodes $z_j = \sqrt{2}x_j$. For a 5-node case, the nodes $z_j$ and weights $q_j$ are listed in Table 2.

Table 2. Weights and nodes of the five order Gaussian quadrature rule ($L = 5$).

| Gaussian Rules | $j$  | 1   | 2   | 3   | 4   | 5   |
|----------------|-----|-----|-----|-----|-----|-----|
| Gauss–Hermite  | $q_j$ | 0.01126 | 0.22208 | 0.53333 | 0.22208 | 0.01126 |
|                | $z_j$ | -2.85697 | -1.35563 | 0.0 | 1.35563 | 2.85697 |

On the other hand, for normally distributed random variables $f''_{\Theta} \sim \mathcal{N}(\mu_j, \sigma_j^2)$, which are the electrical current source terms predicted by GPRs, they can be related to the standard normal random variable $Z$ via

$$
j''_{\Theta} = \mu_j + \sigma_j Z
$$

Using integration by substitution, we find

$$
E\left[\left(h_{P_1}\left(f''_{\Theta}\right)\right)^{\alpha}\right] = \int_{\Theta} h_{P_1}(\mu_j + \sigma_j z) d\left(\mu_j + \sigma_j z\right)
$$

In other words, to find the expectation of $\left(h_{P_1}\left(f''_{\Theta}\right)\right)^{\alpha}$, a function of a normally distributed random variable $f''_{\Theta}$, we can instead calculate the expectation of $\left(h_{P_1}(\mu_j + \sigma_j z)\right)^{\alpha}$, which is a function of a standard normal random variable as $H(Z)$ in Equation (28). We thereby follow the Equation (28) to approximate the Equation (30),

$$
E\left[\left(h_{P_1}\left(f''_{\Theta}\right)\right)^{\alpha}\right] = E\left[\left(h_{P_1}(\mu_j + \sigma_j z)\right)^{\alpha}\right] \approx \sum_{j=1}^{L} q_j \left(h_{P_1}(\mu_j + \sigma_j z_j)\right)^{\alpha}
$$
Similarly, we also have
\[
E\left(\left(h_{N_i}(h_{\Theta_i})\right)^{\alpha}\right) = E\left(\left(h_{N_i}(\mu_{ji} + \sigma_{ji}z)\right)^{\alpha}\right) \approx \sum_{j=1}^L q_j h_{N_i}(\mu_{ji} + \sigma_{ji}z)^{\alpha} \tag{32}
\]
\[
E\left(\left(h_{T_i}(h_{\Theta_i})\right)^{\alpha}\right) = E\left(\left(h_{T_i}(\mu_{Qi} + \sigma_{Qi}z)\right)^{\alpha}\right) \approx \sum_{j=1}^L q_j h_{T_i}(\mu_{Qi} + \sigma_{Qi}z)^{\alpha} \tag{33}
\]
where the normally distributed random variables \(q_{\Theta_i} \sim N(\mu_{Qi}, \sigma_{Qi}^2)\) are the heat source terms predicted by GPRs.

Following the Maximum Entropy principle with fractional moment constraints \([32, 33]\), the estimated PDFs of \(Y_P, Y_N,\) and \(Y_T\) are in the form of
\[
f_{\hat{Y}}(y_k) = \exp\left(-\sum_{k=0}^m \lambda_k y_k^{\alpha_k}\right) \tag{34}
\]
where \(Y_k\) represents \(Y_P, Y_N,\) and \(Y_T,\) and \(y_k\) represents \(y_P, y_N,\) and \(y_T,\) correspondingly. \(\lambda = [\lambda_0, \lambda_1, \ldots, \lambda_m]^T\) are the Lagrange multipliers, and \(\alpha = [\alpha_0, \alpha_1, \ldots, \alpha_m]^T\) are the fractions associated with the fractional moments to be determined.

Based on the normalization condition that the integration of a single PDF must be equal to one, for \(k = 0, \alpha_0 = 0,\) and
\[
\lambda_0 = \ln\left(\int_{y_k} \exp\left(-\sum_{k=0}^m \lambda_k y_k^{\alpha_k}\right) dy_k\right) \tag{35}
\]
We find the aim solution distributions by
\[
\left\{\begin{array}{c}
\text{Find: } \{\alpha_k\}_{k=1}^m, \{\lambda_k\}_{k=1}^m \\
\text{Minimize: } I(\lambda, \alpha) = \ln\left(\int_{y_k} \exp\left(-\sum_{k=0}^m \lambda_k y_k^{\alpha_k}\right) dy_k\right) + \sum_{k=1}^m \lambda_k M_Y^{\alpha_k}
\end{array}\right. \tag{36}
\]
where \(M_Y^{\alpha_k}\) are the moments in Equations (23)–(25).

The solutions for the case \(Y_k = Y_P\) are noted as \(\alpha_P\) and \(\lambda_P,\) the solutions for the case \(Y_k = Y_N\) are noted as \(\alpha_N\) and \(\lambda_N,\) and the solutions for the case \(Y_k = Y_T\) are noted as \(\alpha_T\) and \(\lambda_T.\)

### 3. Simulation Experiments and Results

#### 3.1. Observation Data Creation

Data were obtained from the P2D model-based simulations with different input combinations, as explained in Section 2.1. Potentiostatic discharge processes of the battery cell in the particle and electrode domains with a timestep of 25 s. were performed. The parameters of the simulations are specified in Table 3. The reference voltage, working temperature, and iSOC were maintained as the three input variables. The observation points shown in Table 3 consist of values picked uniformly for each input variable within their upper and lower bounds.

| Input Variable [Unit]   | Lower Bound | Upper Bound | Number of Values Picked |
|-------------------------|-------------|-------------|-------------------------|
| Reference voltage [V]   | 3.74        | 3.83        | 51                      |
| Working temperature [K] | 298         | 324         | 51                      |
| iSOC [%]                | 53          | 88          | 25                      |
| Simulation timestep     |             | 25 s        |                         |

There are, thus, 65,025 input variable observation point combinations. The P2D model is solved under every combination by the tool LIONSIMBA [35] to generate observation data.
3.2. Training Data Selection and GPR Results

To prognosticate at any prediction input point $x^*$, we reference the observation data to find training input data $X = [x_1, x_2, \ldots, x_N]^T$ that satisfy

\[
\begin{align*}
    x^*(1) - 0.005 & < x_i(1) < x^*(1) + 0.005 \\
    x^*(2) - 3.2 & < x_i(2) < x^*(2) + 3.2 \\
    x^*(3) - 2 & < x_i(3) < x^*(3) + 2
\end{align*}
\]

where $x^*(1)$ and $x_i(1)$ represent the input variable reference voltage, $x^*(2)$ and $x_i(2)$ represent the input variable working temperature, and $x^*(3)$ and $x_i(3)$ represent the input variable iSOC.

Paired with the corresponding training target data $y_a = [y^a_1, y^a_2, \ldots, y^a_N]^T, a = 1, 2, 3$, we create independent GPRs as explained in Section 2.2, in which $\hat{y}^a_i$ represent the electrical current source, $\hat{y}^3_i$ represent the heat source, and $\hat{y}^3_i$ represent the uSOC.

The reliability of the GPR predictions was tested by selecting 20 random groups of testing inputs and comparing the GPR predictions with the P2D simulation results.

The box plots reflect the performance of the predictions on the above-selected testing inputs. In the plots, every two error bars and one box construct one column, and each of the columns describes a corresponding distribution of $[(f - m)|x^*, X, y]$, which is the transformed prediction result on a testing input. The distributions are still Gaussian but centered at zero. The upper and lower ending tips of the error bars in each column mark the 97.5th and 2.5th percentiles, respectively. The ceilings and floors of the boxes are respectively at the 75th and 25th percentiles. The bars at the center of the boxes are the means of the distribution, which here are all zeros. The “X” mark on each column shows the value of the difference between the testing target and the testing prediction mean $|y^* - m|x^*, X, y]$. The closer the “X” mark is to the center (0 value) of the column, the better the performance of that prediction.

In Figure 4, the box plot illustrates the distribution of the transformed prediction for the electrical current source, and the scatter plot denotes the testing prediction mean for the electrical current source. The horizontal axis denotes the 20 testing groups. The vertical axis on the left corresponds to the boxes denoting the distributions of the transformed predictions for the electrical current source, while the vertical axis on the right corresponds to the square marks denoting the testing prediction means. The bars at the center of the boxes, signifying the means of the distributions, are all zeros. The “X” mark on each column indicates the value of the difference between the testing target and the testing prediction mean for the electrical current source. The closer the “X” mark is to the center of the column, the better that prediction performance. For the electrical current source, the distributions of the transformed predictions are in the order of $10^3$. The testing prediction means are in the order of $10^5$.

The box plot in Figure 5 illustrates the distribution of the transformed prediction for the heat source, and the scatter plot denotes the testing prediction mean for the heat source. The horizontal axis denotes the 20 testing groups. The vertical axis on the left corresponds to the boxes denoting the distributions of the transformed predictions for the heat source, while the vertical axis on the right corresponds to the square marks denoting the testing prediction means. The bars at the center of the boxes, denoting the means of the distributions, are all zeros as well. The “X” mark on each column indicates the value of the difference between the testing target and the testing prediction mean for the heat source. The closer the “X” mark is to the center of the column, the better that prediction performance. For a heat source, the distributions of the transformed predictions are in the order of $10^5$. The testing prediction means are in the order of $10^5$. 
Figure 4. The box plot shows the distribution of the transformed prediction for electrical current source. The bars at the center of the boxes are the means of distribution. The “X” mark on each column indicates the value of the difference between the testing target and the testing prediction mean. The closer the “X” mark is to the center of the column, the better that prediction performance. The scatter plot denotes the testing prediction mean for electrical current source. The predictions for electrical current source are reliable given that prediction means are in the order of 10^7 while the transformed prediction distributions are in the order of 10^3.

Figure 5. The box plot illustrates the distribution of the transformed prediction for heat source. The scatter plot marks the testing prediction mean for heat source. The predictions for heat source are reliable given that prediction means are in the order of 10^5 while the transformed prediction distributions are in the order of 10^3.

The scatter plots illustrating the testing targets and the testing prediction means for the updated SOC (uSOC) are shown in Figure 6. The horizontal axis denotes the 20 testing groups. The vertical axis corresponds to the square marks denoting the uSOC of the testing targets, and the “+” marks denoting the uSOC of the testing prediction means. The prediction means are in the order of 10^2 while their Mean Squared Error (MSE) with the testing targets is in the order of 10^-4. Box plots are therefore not necessary here to see the obvious adequate accuracy of the uSOC predictions.
3.3. Probabilistic FEA Results

Two groups of FEA simulations were performed on two different initial conditions. In the first group, the cell started discharging initially from the universal temperature of 298 K, with a State of Charge (SOC) of 85.51%, the output voltage at 3.75 volts, and a simulation timestep of 25 s. The results were compared with those from another simulation under the same settings based on Guo’s method. In the second group, the previous results from Guo’s method in the first group served as the initial condition and started another 25-s timestep. The simulation results with the proposed method were compared with those from Guo’s method again. The other specifications about the cell are in Table 4.

Table 4. Summary of model parameters.

| Domain          | Parameter                              | Value                  |
|-----------------|----------------------------------------|------------------------|
| Electrode       | Electrolyte initial concentration, $c_e$ (mol m$^{-3}$) | 1000                   |
|                 | Thickness, $l_e$ (m)                   | $80 \times 10^{-6}$    | $25 \times 10^{-6}$ | $88 \times 10^{-6}$ |
|                 | Bruggeman factor, $p$                  | 4                      | 4                      | 4                      |
|                 | Volume fraction electrolyte, $\varepsilon_e$ | 0.385                   | 0.724                   | 0.485                   |
|                 | Volume fraction inert, $\varepsilon_f$  | 0.025                   | --                      | 0.0326                   |
|                 | Volume fraction active solid phase, $\varepsilon_s$ | 0.59                    | 0                      | 0.4824                   |
|                 | Specific surface area, $a_{s,x}$ (m$^2$ m$^{-3}$) | $0.885 \times 10^6$    | $7.236 \times 10^5$   |
|                 | Solid electronic conductivity, $\sigma_e$ (S m$^{-1}$) | 100                     | 100                     | 100                     |
|                 | Reaction rate constant (mol$^{-0.5}$ s$^{-1}$) | $2.334 \times 10^{-11}$ | $5.031 \times 10^{-11}$ |
|                 | Activation energy for reaction (J mol$^{-1}$) | 5000                   | 5000                   |
|                 | The universal gas constant, $R$ (J mol$^{-1}$ K$^{-1}$) | $8.314$                |
The 95% prediction confidence interval for the results and temperature results from Guo [16] and fall within our 95% prediction confidence interval.

### Table 4. Cont.

| Domain | Parameter | Value |
|--------|-----------|-------|
| The Faraday constant, \( F \) (C mol\(^{-1}\)) | | 96,487 |
| Electrolyte diffusion coefficient, \( D_e \) (m\(^2\)s\(^{-1}\)) | \( D_e = 5.84 \times 10^{-7} \exp[-2870/T](e/1000)^2 - 33.9 \times 10^{-7} \exp[-2920/T](e/1000) + 129 \times 10^{-7} \exp[-3200/T] \) | |
| Electrolyte ionic conductivity, \( \kappa \) (S m\(^{-1}\)) | \( \kappa = 3.45 \ exp[-798/T](e/1000)^3 - 48.5 \ exp[-1080/T](e/1000)^2 + 244 \ exp[-1440/T](e/1000) \) | |
| Effective solid phase conductivity, \( \sigma_{eff} \) (S m\(^{-1}\)) | \( \sigma_{ac} \exp(t_e^p) \) | |
| Effective liquid phase diffusion coefficient, \( D_{lep} \) (m\(^2\)s\(^{-1}\)) | \( D_{ac} \exp(t_e^p) \) | |
| Effective liquid phase ionic conductivity, \( \kappa_{lep} \) (S m\(^{-1}\)) | \( \kappa_{ac} \exp(t_e^p) \) | |
| Effective liquid phase diffusional ionic conductivity, \( \kappa_{D_{lep}} \) (S m\(^{-1}\)) | \( \kappa_{D_{ac}} \exp(t_e^p) \) | |
| Li\(^+\) transference number, \( t_{lep}^0 \) | \( t_{lep}^0 = -0.00026 \exp[883/T](e/1000)^2 + 0.00309 \exp[653/T](e/1000) + 0.517 \exp[-49.6/T](e/1000) \) | |

### 3.3.1. First Group Experiments

In the first group, three points in subarea block A and three points in subarea block B were randomly selected on a cell to observe the simulation results, as shown in Figure 7.

![Figure 7](image-url)

**Figure 7.** Two subarea blocks A and B are randomly selected to observe the simulation results. The cell demonstrated in 2D here is the same one as in Figure 1. The grids partitioned the cell into 16 subareas. The scales are to show the dimensions of the cell (Unit: m).

In block A, the probability distributions of the potential difference between the pairs at the three points are plotted in Figure 8A(a–c). The probability distributions of the temperature at the three points are plotted in Figure 8B(a–c).
Figure 8. The box plots display the predicted probability distributions of (A) potential differences, (B) temperature, between plates at three random locations selected in block A in the first simulation group, and (C) potential differences, (D) temperature at the three locations in block B in the first simulation group. The red “X” marks denote the corresponding results at these locations at the 25th s. using the method in work [16], and they all fall within the 95% prediction confidence interval presented in this paper. The scales are to show potential differences and temperatures as marked (Units: (A,C) Volt and (B,D) K).

The box plots display the predicted probability distributions of potential differences between plates at three random locations selected in block A in the first simulation group, cf. Figure 8A. The scales on the vertical axis indicate potential differences. The red “X” marks denote the corresponding potential difference results at these locations at the 25th s.
using Guo’s method [16]. All results fall within the 95% prediction confidence interval presented in this paper. Similarly, the predicted probability distributions of temperatures at the three locations in block A in the first simulation group are illustrated in Figure 8B. Note that the red marks indicate the corresponding temperature results using Guo’s work [16], and fall within our 95% prediction confidence interval.

In block B, the probability distributions of the potential difference and temperatures between the pairs at the three points are plotted in Figure 8C(a–c) and Figure 8D(a–c), respectively. The data in Figure 8C,D confirm that the corresponding potential difference results and temperature results from Guo’s method [16] are always within our 95% prediction confidence interval as expected.

The entire cell temperature results from the presented GPR-based model are depicted in Figure 9A(a) for the first simulation group. Figure 9A(b) illustrates the corresponding temperature results at the 25th s. using Guo’s work [16]. When the two are compared, they match. The temperature distribution of a cell in 2D does not show severe unevenness in the first simulation group. The end far from the current collector tab has a relatively lower temperature while the unevenness is growing. The horizontal and vertical axes scales indicate the cell dimensions as in Figure 1, while the color scales index the cell temperatures.

Figure 9B(a) illustrates the GPR model-based electrical potential distribution result on the cell’s negative plate in 2D in the first simulation group; the potential is relatively higher near the current collector tab area. Figure 9B(b) illustrates the GPR model-based electrical potential distribution result on the cell’s positive plate in 2D in the first simulation group; the potential is relatively lower near the current collector tab area. In addition to the scales on the horizontal and vertical axes showing the cell dimensions, the color scales index the electrical potentials.

The electrical potential distribution results at the 25th s. using the method in Guo’s work [16] is illustrated in Figure 9C(a,b). The outcomes correspond to the results from our GPR models in the first simulation group, as in Figure 9B. Figure 9C(a) is the result on the cell’s negative plate. Figure 9C(b) is the cell’s positive plate. They are identical to the results shown in Figure 9B.

3.3.2. Second Group Experiments

Three locations in block A and three locations in block B were randomly selected on a cell to observe the simulation results, as shown in Figure 7. In block A, the probability distributions of the potential difference between the pairs at the three points are plotted in Figure 10A(a–c). The probability distributions of the temperature at the three points are plotted in Figure 10B(a–c). In an identical manner as in the first group of experiments, the red “X” marks indicate results using Guo’s method [16] for the potential differences (Figure 10A) and temperatures (Figure 10B) all fall within the 95% prediction confidence interval by the method in this paper without exception. The experiments are repeated for block B, and the probability distributions of the potential difference between the pairs at the three points are plotted in Figure 10C(a–c) and the temperature ones in Figure 10D(a–c). Again, for all locations in block B, our 95% prediction confidence interval still covers the corresponding potential difference and temperature results compared to Guo’s simulations [16]. The results validate the presented method in the second simulation group.

The entire cell’s temperature results, of the second simulation group, from our GPR-based model, are depicted in Figure 11a. Figure 11b illustrates the results employing Guo’s work [16]. A comparison between the two reveals that the results of the two methods show an exact match. The temperature distribution is uneven in the second simulation group with the areas near collector tabs subjected to higher temperatures because of the more intensive electrical and chemical activities.

The electrical potential distribution results of the GPR model on the negative and positive plates are depicted in Figure 11B,C. Both illustrations are for the second simulation group. As seen, when compared to the first simulation group, cf. Figure 9B,C, the results are identical signifying the robustness of the proposed framework. Figure 11C(a,b) shows
the electrical potential distribution results using [16], which correspond to the results by our GPR models in the second simulation group as in Figure 11B(a–c).

Figure 9. 2D results of temperature and electrical potential for the first simulation group. The scales show the cell dimensions as in Figure 1. (Unit: m). Color distribution as noted below. (A) The temperature distribution of a cell. The end far from the current collector tab has relatively lower temperature while the unevenness is growing. (a) GPR-based model. (b) Result at the 25th s. using [16]. (a,b) Results are identical. Color scale index cell temperature in K. (B) GPR model-based electrical potential distribution on the (a) negative plate and (b) positive plate. Potential is relatively higher in (a) and lower in (b) near the current collector tab area). Color scales index the electrical potentials in Volts. (C) Electrical potential distribution at the 25th s. using [16], which correspond to the results by our GPR models in the first simulation group as in Figure 9B. Electrical potential distribution on the (a) negative plate, and (b) positive plate of the cell in 2D. They are identical to the results in Figure 9B. Color scales index the electrical potentials in Volts. Scales indicate cell dimensions (Unit: m).
Figure 10. The box plots display the predicted probability distributions of (A) potential differences, (B) temperature, between plates at three random locations selected in block A in the second simulation group, and (C) potential differences, (D) temperature at the three locations in block B in the second simulation group. The red marks denote the corresponding results at these locations by the method in Guo’s work [16], and they all fall within the 95% prediction confidence interval presented in this paper. The scales are to show potential differences and temperatures as marked (Units: (A,C) Volt and (B,D) K).
Figure 11. 2D results of temperature and electrical potential for the second simulation group. The scales show the cell dimensions as in Figure 1 (Unit: m). Color distribution as noted below. (A) The temperature distribution of a cell shows unevenness. Areas near collector tabs indicate higher temperatures. (a) GPR-based model. (b) Result using [16]. (a,b) results are identical. Color scale index cell temperature in K. (B) GPR model-based electrical potential distribution on the (a) negative plate and (b) positive plate. Potential is relatively higher in (a) and lower in (b) near the current collector tab area. Color scales index the electrical potentials in Volts. (C) Electrical potential distribution at the 25th s. using [16], which correspond to the results by our GPR models in the first simulation group as in Figure 11B. Electrical potential distribution on the (a) negative plate, and (b) positive plate of the cell in 2D. They are identical to the results in Figure 11B. Color scales index the electrical potentials in Volts. Scales indicate cell dimensions (Unit: m).
4. Discussion

The previous section tested the GPR predictions’ reliability at the beginning of the simulation experiments by randomly selecting 20 groups of testing inputs and comparing the GPR predictions with the P2D simulation results.

Predictions of the electrical current and heat sources are reliable as seen from comparing Figures 4 and 5. As seen, the transformed prediction distributions are in the order of $10^3$ and $10^0$, given that the prediction means are in the order of $10^7$ and $10^5$ for them. The prediction means on the uSOC are in the order of $10^{-2}$ while their Mean Squared Error (MSE) with the testing targets is in the order of $10^{-4}$. Based on the above, one may surmise that the proposed method is of adequate accuracy without using box plots that uSOC predictions, as shown in Figure 6.

Next, two groups of FEA simulation experiments were performed on two different initial conditions. The cell discharged at a fixed output voltage in the first group, with a universal initial temperature and SOC, at a 25-s. simulation timestep. The results obtained using the proposed method were compared with those from another simulation under the same setting but instead based on Guo’s method. In the second group, the previous results from Guo’s method in the first group served as the initial condition and started another 25-s. timestep. The simulation results obtained using the proposed method were compared with those from Guo’s method again. Three random locations in subarea block A and three random locations in subarea block B, as shown in Figure 7, were selected on the battery cell to observe the simulation results for potential differences and temperatures. The entire cell results in terms of electrical potential and temperature distributions were also observed.

In both the first and the second group of experiments, for all the six locations within the subarea blocks A and B, the corresponding results by Guo’s method [16] fall within the 95% confidence intervals of our predicted probability distributions without exceptions (cf. Figures 8–11). The first group of experiments compares the entire cell temperature results from the presented method and the one in Guo [16], cf. Figure 9A, the results match exactly. There is no severe unevenness throughout the plates, and the end far from the current collector tab has a relatively lower temperature while the unevenness is growing.

Moreover, the present method is further validated for the electrical potential distributions against Guo’s work [16] for the positive and the negative plates. Comparison of Figure 9B,C illustrates identical behavior, namely the potential is relatively higher near the current collector tab area on the negative plate, while on the positive plate, the potential is relatively lower near the current collector tab area. In the second group of experiments, a comparison of the entire cell temperature results between this paper and Guo’s work [16], cf. Figure 11A reveals an exact match. Similar conclusions are noted for the electrical potentials’ distributions, cf. Figure 11B,C. These observations indicate that the presented method is reliable and of adequate accuracy.

It is noted that the presented work relied on simulating the potentiostatic discharge processes of the battery cell to carry out testing. Additional simulations could be arranged toward various current standards for EV batteries in future work. That helps to fully discover and evaluate the limits of the presented GPR-based data-driven framework before it is applied in an EV battery management system. Some noticeable standards related to the EV battery performance and safety are the “62660” series documents for lithium-ion battery use in EVs by the International Electrotechnical Commission, the “2580” document by UL LLC, and the “USABC” manuals produced and maintained by the United States Council for Automotive Research LLC, along with the United States Advanced Battery Consortium [36–40]. Implementing the framework for real-time battery management systems must consider the data pipelines and communication protocols with the rest of EV systems [41]. The testing in this paper did not cover the battery’s degradation as in work [42,43]. However, because of its data-driven nature, the presented method can easily incorporate empirical data reflecting battery degradation for model enhancement as equivalent circuit models [44] in future work.
5. Conclusions

This paper presents the applications of a GPR-based data-driven framework on a single Li-ion battery cell when the inputs to the GPRs are deterministic. A Li-ion electrode pair is used as a battery cell example, and its thermal and electrical behaviors are solved following the MSMD framework in multiple domains.

The GPRs work in the particle and electrode domains in a battery cell, and the predictions for the different behavior aspects, are not correlated with each other when the inputs are deterministic. Different combinations of the settings in the P2D model compose the training input data pool, and the corresponding P2D model simulation results compose the training target data pool. It is noteworthy that the testing results clearly demonstrate that the GPRs predictions are reliable, accuracy, and stable.

The probabilistic FEA method solves for the temperature and electrode plate potential distribution in the cell domain. PDEs cut functions are defined where variables are not correlated with each other. Furthermore, a combination of Gaussian–Hermite methods with the Maximum Entropy principle is adopted to approximate the solution distributions following the M-DRM procedure. The testing results exhibit that the means of the solution distributions match the results from other methods. Notably, the probability distributions of the temperatures and the plate potential differences are within reasonable ranges.

In this work, the overall GPR data-driven framework is validated against published data. The framework offers accurate and stable predictions for battery cell’s thermal and electrical behaviors during the first timestep in a real-time battery management system where the GPR inputs are deterministic. Future work will examine battery degradation and validation against current EV battery management standards.

Author Contributions: Data curation, C.L. and A.A.P.; Formal analysis, C.L.; Funding acquisition, A.A.P.; Investigation, C.L. and A.A.P.; Methodology, C.L. and A.A.P.; Project administration, A.A.P.; Software, C.L.; Supervision, A.A.P.; Writing—original draft, C.L.; Writing—review & editing, A.A.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kim, G.; Pesaran, A.; Spotnitz, R. A Three-dimensional Thermal Abuse Model for Lithium-ion Cells. J. Power Sources 2007, 170, 476–489. [CrossRef]
2. Srinivasan, V.; Wang, C.-Y. Analysis of Electrochemical and Thermal Behavior of Li-Ion Cells. J. Electrochem. Society 2003, 150, A98–A106. [CrossRef]
3. Barsukov, Y.; Qian, J. Battery Power Management for Portable Devices; Artech House: Norwood, MA, USA, 2013.
4. Pesaran, A.A.; Kim, G.H.; Keyser, M. Integration Issues of Cells into Battery Packs for Plug-in and Hybrid Electric Vehicles. In Proceedings of the 24th International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium & Exhibition, Stavanger, Norway, 13–16 May 2009.
5. Meng, J.; Guo, H.; Niu, C.; Zhao, Y.; Xu, L.; Li, Q.; Mai, L. Advances in Structure and Property Optimizations of Battery Electrode Materials. Joule 2017, 1, 522–547. [CrossRef]
6. Peter, V.d.B.; Frédéric, V.; Joeri, V.M.; Julien, M.; Wout, V.A. SUBAT: An Assessment of Sustainable Battery Technology. J. Power Sources 2006, 162, 913–919.
7. Verbrugge, M.W.; Conell, R.S. Electrochemical and Thermal Characterization of Battery Modules Commensurate with Electric Vehicle Integration. J. Electrochem. Soc. 2002, 149, A45–A53. [CrossRef]
8. Hellweg, B. Microstructural Modeling of Lithium Battery Electrodes; Massachusetts Institute of Technology: Cambridge, MA, USA, 2000.
9. Johnson, V.H.; Pesaran, A.A.; Sack, T. Temperature-Dependent Battery Models for High-Power Lithium-Ion Batteries. In Proceedings of the 17th Annual Electric Vehicle Symposium, Montreal, QC, Canada, 14–17 June 2000.
10. Wang, C.Y.; Srinivasan, V. Computational battery dynamics (CBD)—Electrochemical/thermal coupled modeling and multi-scale modeling. J. Power Sources 2002, 110, 364–376. [CrossRef]
11. He, H.; Xiong, R.; Fan, J. Evaluation of Lithium-Ion Battery Equivalent Circuit Models for State of Charge Estimation by an Experimental Approach. Energies 2011, 4, 582–598. [CrossRef]
12. Hu, X.; Li, S.; Feng, H. A Comparative Study of Equivalent Circuit Models for Li-ion Batteries. J. Power Sources 2012, 198, 359–367. [CrossRef]
13. Newman, J.S. Electrochemical Systems, 2nd ed.; Prentice-Hall: Englewood Cliffs, NJ, USA, 1973.
14. Kim, G.-H.; Smith, K.; Lee, K.-J.; Santhanagopalan, S.; Pesaran, A. Multi-Domain Modeling of Lithium-Ion Batteries Encompassing Multi-Physics in Varied Length Scales. J. Electrochem. Soc. 2011, 158, A955–A969. [CrossRef]
15. Doyle, M.; Fuller, T.F.; Newman, J. Modeling of Galvanostatic Charge and Discharge of the Lithium/Polymer/Insertion Cell. J. Electrochem. Soc. 1993, 140, 1526–1533. [CrossRef]
16. Guo, M.; White, R.E. A distributed thermal model for a Li-ion electrode plate pair. J. Power Sources 2013, 221, 334–344. [CrossRef]
17. Smith, K.A.; Rahn, C.D.; Wang, C.-Y. Control Oriented 1D Electrochemical Model of Lithium Ion Battery. Energy Convers. Manag. 2007, 48, 2565–2578. [CrossRef]
18. Subramanian, V.R.; Boovaragav, V.; Diwakar, V.D. Toward Real-Time Simulation of Physics Based Lithium-Ion Battery Models. Electrochem. Solid-State Lett. 2007, 10, A255–A260. [CrossRef]
19. Ramadesigan, V.; Boovaragavan, V.; Subramanian, V.R. Efficient Reformulation of Solid-Phase Diffusion in Physics-Based Lithium-Ion Battery Models. J. Electrochem. Soc. 2010, 157, A854–A860. [CrossRef]
20. Northrop, P.W.C.; Ramadesigan, V.; De, S.; Subramanian, V.R. Coordinate Transformation, Orthogonal Collocation, Model Reformulation and Simulation of Electrochemical-Thermal Behavior of Lithium-Ion Battery Stacks. J. Electrochem. Soc. 2011, 158, A1461–A1477. [CrossRef]
21. McHutchon, A. Nonlinear Modelling and Control using Gaussian Processes. Ph.D. Thesis, University of Cambridge, Cambridge, UK, 2014.
22. Damianou, A.; Titsias, M.; Lawrence, N. Variational Gaussian Process Dynamical Systems. In Proceedings of the 24th International Conference on Neural Information Processing Systems, Granada, Spain, 12–15 December 2011.
23. Wang, J.; Fleet, D.; Hertzmann, A. Gaussian Process Dynamical Models for Human Motion. IEEE Trans. Pattern Anal. Mach. Intell. 2008, 30, 283–298. [CrossRef]
24. Venables, W.; Ripley, B. Modern Applied Statistics with S-PLUS; Springer: New York, NY, USA, 1999.
25. Hastie, T.; Tibshirani, R.; Friedman, J. The Elements of Statistical Learning; Springer: Berlin, Germany, 2009.
26. Bishop, C. Pattern Recognition and Machine Learning; Springer: Berlin, Germany, 2006.
27. Girard, A. Approximate Methods for Propagation of Uncertainty with Gaussian Process Models; University of Glasgow: Glasgow, UK, 2004.
28. Deisenroth, M.P. Efficient Reinforcement Learning Using Gaussian Processes; KIT Scientific Publishing: Karlsruhe, Germany, 2010.
29. Pearl, J. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference; Morgan Kaufmann: Burlington, MA, USA, 1988.
30. Tunga, M.; Demiralp, M. A Factorized High Dimensional Model Representation on the Nodes of a Finite Hyperprismatic Regular Grid. Appl. Math. Comput. 2005, 164, 865–883. [CrossRef]
31. Balomenos, G. Probabilistic Finite Element Analysis of Structures using the Multiplicative Dimensional Reduction Method; University of Waterloo: Waterloo, ON, Canada, 2015.
32. Jaynes, E.T. Information Theory and Statistical Mechanics. Phys. Rev. 1957, 106, 620–630. [CrossRef]
33. Inverardi, P.L.N.; Tagliani, A. Maximum Entropy Density Estimation from Fractional Moments. Commun. Stat. Theory Methods 2003, 2, 327–345. [CrossRef]
34. Zwilling, C. CRC Standard Mathematical Tables and Formulae, 31st ed.; CRC Press: New York, NY, USA, 2003.
35. Torchio, M.; Magni, L.; Gopaluni, R.B.; Braatz, R.D.; Raimondo, D.M. LIONSIMBA: A Matlab Framework Based on a Finite Volume Model Suitable for Li-Ion Battery Design, Simulation, and Control. J. Electrochem. Soc. 2016, 163, A1192–A1205. [CrossRef]
36. Green, J.M.; Hartman, B.; Glowacki, P.F. A System-based View of the Standards and Certification Landscape for Electric Vehicles. World Electr. Veh. J. 2016, 8, 564–575. [CrossRef]
37. Ruiz, V.; Moretto, P. Standards for the Performance and Durability Assessment of Electric Vehicle Batteries—Possible Performance Criteria for an Ecodesign Regulation; Joint Research Centre: Ispra, Italy, 2018.
38. Castillo, E.C. Standards for electric vehicle batteries and associated testing procedures. In Advances in Battery Technologies for Electric Vehicles; Woodhead Publishing: Sawston, Cambridge, UK, 2015; pp. 469–494.
39. Icdean, C.; Varga, B.; Burnete, N.; Cimerdean, D.; Jurchis, B. Comparison of Different Battery Types for Electric Vehicles. In IOP Conference Series: Materials Science and Engineering; IOP Publishing: Bristol, UK, 2017; Volume 252, p. 012058.
40. Lu, L.; Han, X.; Li, J.; Hua, J.; Ouyang, M. A review on the key issues for lithium-ion battery management in electric vehicles. J. Power Sources 2013, 226, 272–288. [CrossRef]
41. Meissner, E.; Richter, G. Battery Monitoring and Electrical Energy Management: Precondition for Future Vehicle Electric Power Systems. J. Power Sources 2003, 116, 79–98. [CrossRef]
43. Tran, M.-K.; Panchal, S.; Chauhan, V.; Brahmbhatt, N.; Mevawalla, A.; Fraser, R.; Fowler, M. Python-based scikit-learn machine learning models for thermal and electrical performance prediction of high-capacity lithium-ion battery. *Int. J. Energy Res.* 2021, 45, 1–9. [CrossRef]

44. Saleem, K.; Mehran, K.; Ali, Z. Online reduced complexity parameter estimation technique for equivalent circuit model of lithium-ion battery. *Electr. Power Syst. Res.* 2020, 185, 106356. [CrossRef]