Gaussian Process Regression With Automatic Relevance Determination Kernel for Calendar Aging Prediction of Lithium-Ion Batteries

Kailong Liu , *Member, IEEE*, Yi Li, Xiaosong Hu , *Senior Member, IEEE*, Mattin Lucu , *Student Member, IEEE*, and Widanalage Dhammika Widanage , *Member, IEEE*

Abstract—Battery calendar aging prediction is of extreme importance for developing durable electric vehicles. This article derives machine learning-enabled calendar aging prediction for lithium-ion batteries. Specifically, the Gaussian process regression (GPR) technique is employed to capture the underlying mapping among capacity, storage temperature, and state-of-charge. By modifying the isotropic kernel function with an automatic relevance determination (ARD) structure, high relevant input features can be effectively extracted to improve prediction accuracy and robustness. Experimental battery calendar aging data from nine storage cases are utilized for model training, validation, and comparison, which is more meaningful and practical than using the data from a single condition. Illustrative results demonstrate that the proposed GPR model with ARD Matern32 (M32) kernel outperforms other counterparts and can achieve reliable prediction results for all storage cases. Even for the partial-data training test, multistep prediction test, and accelerated aging training test, the proposed ARD-based GPR model is still capable of excavating the useful features, therefore offering good generalization ability and accurate prediction results for calendar aging under various storage conditions. This is the first-known data-driven application that utilizes the GPR with ARD kernel to perform battery calendar aging prognosis.

Index Terms—Battery health, calendar aging prediction, data-driven model, Gaussian process regression, lithium-ion batteries.

I. INTRODUCTION

LITHIUM-ION (Li-ion) batteries are the promising candidates for electric vehicle (EV) applications, owing to their impressive features such as high energy density, high efficiency, and environmental friendliness [1]. However, reliable calendar aging prediction is still a bottleneck for the performance enhancement of EVs. In real automotive applications, Li-ion batteries generally degrade with the calendar and cyclic modes. Considering that more than 75% of battery life is spent under parking mode for EVs [2], calendar aging prediction therefore becomes a prerequisite for battery service life diagnosis.

Calendar aging for most Li-ion batteries is mainly caused by the growth of solid electrolyte interface (SEI) during storage [3]. Specifically, when a battery is stored, the reduction of its electrolyte solvents such as ethylene carbonate would cause the formation of Li-based products, further resulting in the generation of SEI on the anode particle of battery [4]. In such cases, Li-ion battery capacity would decrease over time [5]. The corresponding capacity aging rate is highly dependent on several key factors including the storage temperature and battery state-of-charge (SOC) [6]. Therefore, a key but challenging issue for calendar aging prediction is to simultaneously take these factors into account. It is vital to develop suitable models to capture capacity degradation dynamics under various storage conditions.

Several physics-based models have been reported in the literature to explain battery calendar aging behaviors [7], [8]. Although electrochemical dynamics of batteries during storage have been analyzed in the simulation environment, these models are highly time-consuming and complex to parametrize, making them overly expensive for real-time calendar aging prediction on a long period scale.

To overcome the above challenges, calendar aging prediction approaches based on semiempirical models have been designed. For instance, Schmalstieg *et al.* [9] proposed an Arrhenius-based semiempirical model to capture calendar cell aging. Petit *et al.* [10] developed an empirical capacity loss model to evaluate the effects of SOC and temperature on storage lifetime of Li-ion batteries. In [11], instead of using Arrhenius acceleration model, a semiempirical approach based on the Eyring acceleration model was adopted to predict battery calendar aging, while the SOC drifting was also taken into account. By considering the effects

of temperature and storage conditions, De Hoog *et al.* [12] proposed a semiempirical combined model to estimate the calendar lifetime for a nickel-manganese-cobalt oxide battery. In [13], by taking the initial surface layer caused by cell formation into account, an extended semiempirical model was proposed to improve the calendar aging predictive ability. These referred works belong to open-loop models without strong generalization abilities; in a way, their performance highly depends on the quality of test experiments.

Data-driven models, which are free of assuming any mechanism a priori, are also gaining increasing attention in the battery state-of-health (SOH) estimation and remaining useful life diagnosis [14]. Different intelligent techniques such as support vector regression [15], [16], Bayesian prediction [17], [18], and artificial neural network [19]–[21] have been successfully applied to build data-driven models for battery cyclic aging prediction. On one hand, some review papers have summarized these state-of-the-art applications [22], [23], concluding that several limitations still exist as: 1) data-driven approaches are mainly used to capture battery cyclic aging states but very few attempts have been done for calendar aging diagnosis. 2) Most publications fit the model on aging data obtained under constant operating conditions, ignoring various cases of temperature and SOC. Such models are infeasible for predicting capacity under different conditions. On the other hand, in a previous publication, a critical review on various data-driven models in battery aging domain was presented, in which the Gaussian process regression (GPR) is identified as one of the most powerful techniques. Detailed comparisons for different machine learning techniques are referred to Table 5 of [14] and the corresponding discussions. In fact, beyond the performance of simple structure and computationally acceptable predictions, GPRs enjoy the significant merits of being nonparametric and able to consider the uncertainty of predicted values. Through formulating specific input features, GPR-based models have been applied successfully in both academic and industrial domains [24]–[26]. However, to the best of our knowledge, there is still a lack of researches by using GPR in battery calendar aging prediction domain.

Besides, most existing works just use single conventional kernels to develop their GPR techniques without considering the correlations of multidimensional input variables. In the light of this, it could be a promising way through developing an improved GPR technique with the multidimensional kernel structure to capture the battery capacity degradation dynamics under different temperature and SOC storage conditions.

Based on the above discussions, this article is concerned with machine learning-enabled calendar aging prediction for Li-ion batteries, where both the corresponding storage temperature and battery SOC can be taken into account simultaneously. Several key original contributions are made in this article. First, nine cases of experimental calendar aging data are collected under various storage temperatures and SOC levels over 480 days, constituting a well-rounded database to train and validate the calendar aging model. Second, because the battery calendar aging involving local fluctuations over storage time is a highly nonlinear process, a framework based on the GPR model is proposed to
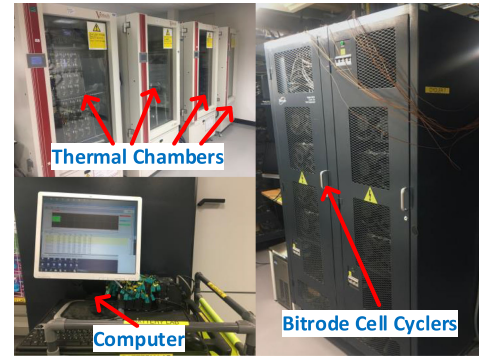

Fig. 1.   Calendar aging test equipment.

efficiently capture the capacity degradation dynamics with reliable confidence ranges. Third, due to the input features involving storage temperature and cell SOC, the isotropic kernel function of GPR is modified with an automatic relevance determination (ARD) structure, which brings the benefits that irrelevant inputs can be removed by fixing large length scales. Meanwhile, various predictors can be formulated to improve prediction accuracy and robustness. Finally, based on our dataset, the prediction performance of our proposed GPR model is investigated in terms of different kernel functions, and compared with a regression calendar-life (RCL) model. This is the first known data-driven application by utilizing GPR with ARD kernel to handle battery calendar aging predictions. Obviously, due to mechanism-free characteristics, the proposed GPR+ARD model can be readily extendable to other battery types for calendar aging prognosis.

The rest of this article is organized as follows. Section II presents the calendar aging experiments and the collected dataset. Section III introduces the developed model framework and several key quantitative metrics, followed by the description of ARD-based GPR model. Section IV analyzes the comparison and verification results. Finally, Section V concludes this article.

II. CALENDAR AGING TEST

Fig. 1 illustrates the equipment used for conducting the battery calendar aging tests under different conditions. The cells were stored in the Votsch VT3050 Thermal Chambers, and operated by the Bitrode MVC 16-100-5 Cell Cyclers. The generated battery data were monitored and stored by a computer. Commercial Panasonic NCRBD batteries from a commercial automotive company were the cells used to study calendar aging characteristics of Li-ion batteries. The battery has a 3Ah nominal capacity, 2.5 V lower cut-off voltage, and 4.2 V upper cut-off voltage. Because the rate of degradation can be minimized through keeping the SOC at a low or medium level and lowering the battery temperature [27], all cells are stored at 10 °C and moderate 50% SOC prior to any tests.

The calendar aging test was performed under various storage temperatures (10 °C, 25 °C, and 45 °C) and SOC levels (20%, 50%, and 90%) for a storage time of 480 days. All batteries were set in the temperature chambers with an open-circuit status
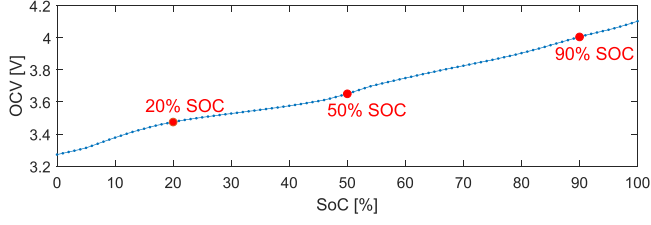
Fig. 2.　OCV–SOC curve for the tested NCRBD battery.

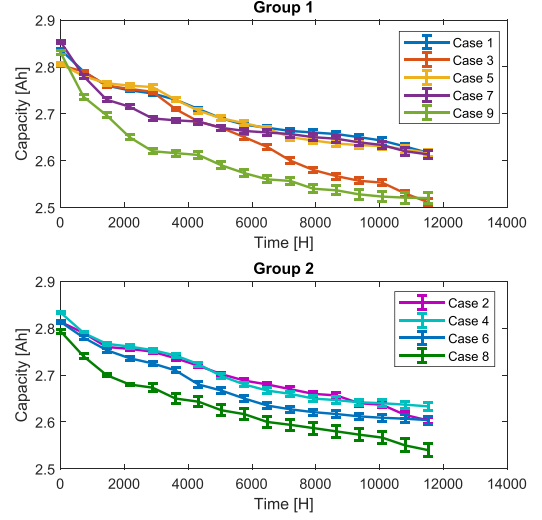Fig. 3.　Calendar aging dataset for different storage conditions.

Fig. 4.　Capacity degradations in calendar aging with standard deviations versus storage hours for different cases.

TABLE I
INITIAL BATTERY CAPACITY VALUES FOR DIFFERENT CASES

Group 1	Case 1	Case 3	Case 5	Case 7	Case 9
C_{ini} [Ah]	2.837	2.802	2.807	2.852	2.830
Group 2	Case 2	Case 4	Case 6	Case 8	
C_{ini} [Ah]	2.812	2.833	2.813	2.794	

during storage. For each storage temperature and SOC, three cells were studied to obtain the average battery capacity and minimize any battery-to-battery discrepancies.

Periodic check-ups were performed every 30 days to obtain the capacity information during storage. For all tests, 1 C-rate is equal to 3 A. At each check-up, the temperature chambers were set to 25 °C. Each cell was then charged by a constant-current constant-voltage (CC-CV) pattern with 1/2 C-rate in the CC phase until the terminal voltage reached 4.2 V, followed by a CV phase until the current dropped below 1/10 C-rate. After resting for 3 h, the cells were discharged by a CC pattern with 1/3 C-rate until the lower cut-off voltage of 2.5 V. The average discharge capacity (over the three cells) was selected as the battery capacity for each condition. Before calibrating the SOC of cells, CC-CV pattern would be implemented again to recharge all batteries to their full-charging states. After another 3 h rest period, the batteries were discharged to their specified SOC setpoints by a well-controlled coulomb counting method. Fig. 2 illustrates the open-circuit voltage (OCV)–SOC curve for our adopted NCRBD battery. The OCV–SOC points at which the batteries were stored are also highlighted with red. Subsequently, each temperature chamber was readjusted to its specified storage condition again.

Following this procedure, the calendar aging dataset that contains nine storage cases was obtained, as shown in Fig. 3. Five cases (Case 1, Case 3, Case 5, Case 7, and Case 9) are labeled as "Group 1" and another four cases (Case 2, Case 4, Case 6, and Case 8) are labeled as "Group 2." Detailed capacity aging curves with the standard deviations versus time for various storage conditions are illustrated in Fig. 4. Several of the capacity fade trends illustrate an initial rapid capacity fade followed by a more linear decrease. This phenomenon is attributed to the presence of excess anode electrode area in comparison to the cathode electrode area. Known as anode "overhang" in literature [28],

[29], an outflow of Li-ions can occur from the active regions of the anode to its excess passive regions, leading to the initial rapid capacity fade. The subsequent linear capacity fade is then attributed to the irreversible capacity fade due to SEI growth.

The initial battery capacities C_{ini} for these cases are all different from each other, as described in Table I. It can be also seen that the initial measured capacities do not start from the nominal capacity of cell, which is a practical and likely scenario to occur.

III. TECHNIQUE

This section elaborates the modeling methodology as well as the corresponding quantitative metrics. Additionally, the fundamentals of GPR technique with ARD kernel are presented, followed by a brief description of a RCL model for comparison purposes.

A. Model Development and Quantitative Metrics

On the basis of the tested calendar aging dataset, machine learning-based techniques can be developed to capture capacity degradation dynamics with various storage conditions. Compared with the existing data-driven models that normally consider just capacity information, an innovative model structure, which also takes both storage temperature term and battery SOC term into account, is developed for calendar aging prediction, as shown in Fig. 5. This model framework can be mainly divided into two parts. For the prediction of next capacity point, output $C_{sto}(t + 1)$ can be predicted after using GPR to learn the underlying mappings among all input terms including the historical
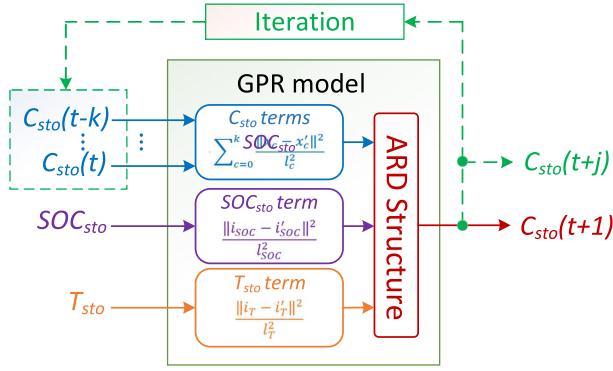
Fig. 5.   Model framework for calendar aging prediction.

capacity data vector $[C_{\text{sto}}(t-k),\ldots,C_{\text{sto}}(t)]$, storage SOC level SOC_{sto}, and temperature T_{sto}. In GPR, these mappings are reflected within the covariance functions. Detailed pseudo-code of this mapping can be found in Section II of [30]. For the multistep prediction, as illustrated in the green dashed line of Fig. 5, an iteration process that uses the previously predicted capacity as the next input to further predict new capacity value is conducted until the jth capacity value is achieved. Here j and k represent the horizons of future and previous calendar capacities, respectively. In order to use our collected dataset and verify the prediction performance of the proposed model, the capacity prediction in this article is conducted in steps of 30 days (720 h) for a total time duration of 480 days (11 520 h). Through a trial-and-error method, the case of "$k = 2$" is selected due to a good trade-off between computational effort and prediction accuracy.

After constructing the proper input–target pairs from the battery calendar aging dataset, the GPR technique is employed to study the potential mapping mechanism, giving rise to the capacity prediction model that considers various storage conditions. Based on our collected dataset as shown in Fig. 3, in order to ensure enough aging information can be learned for pure machine-learning techniques, the dataset from "Group 1" (green cases) that covers all temperatures and SOCs is applied for model training purpose. After training, the dataset from "Group 2" (yellow cases) is used to verify the effectiveness of the proposed model.

Moreover, to evaluate the prediction performance of the data-driven model, several key quantitative metrics are adopted in this article [31]. Here, N is the total number of predicted points, y_j and $\tilde{y}_j$ stand for each actual capacity data and each predicted value, respectively.

1) *Maximum absolute error (MAE):* By defining as (1), MAE is used to illustrate the maximum difference between the predicted and real test values. The larger the MAE values, the poorer the predicted accuracy is [31].

$$\text{MAE} = \max_{1 \le j \le n} |y_j - \tilde{y}_j|. \tag{1}$$

2) *Root-mean-square-error (RMSE):* RMSE is another widely used metric to measure the overall difference between the predicted values and real test values. By

defining as (2), the closer RMSE reaches to 0, the better the prediction accuracy is achieved [31].

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(y_j - \tilde{y}_j\right)^2}. \tag{2}$$

3) *Fit-goodness (R^2):* R^2 is defined by (3) to measure the match quality of a model to the real test data [32]

$$R^2 = 1 - \frac{\sum_{j=1}^{N}\left(y_j - \tilde{y}_j\right)^2}{\sum_{j=1}^{N}\left(y_j - \bar{y}\right)^2} \tag{3}$$

where $\bar{y}$ is the mean of the predicted values. It is evident that as R^2 approaches 1, the corresponding model well describes the variability of the target class.

4) *Calibration score (CS):* By defining as (4), CS reflects the frequency of real data lying within the obtained confidence range [24]

$$\text{CS} = \frac{1}{N}\sum_{j=1}^{N}[|y_j - \tilde{y}_j| < 2\sigma]_I \tag{4}$$

where $[\cdot]_I$ represents the Iverson bracket. For a GPR model, 95.4% is a general confidence range with the interval corresponding to $\pm 2\sigma$ [24]. Therefore, the ideal CS should get close to 0.954: be less or larger than this value indicates that the developed model is overconfident or underconfident, respectively.

B. GPR Technique With ARD Kernel Structure

Derived from the Bayesian framework, the GPR is able to undertake nonparametric regression with the Gaussian process. By defining the mean function $m(i)$ and covariance function $\kappa(i, i')$ of a real process $f(i)$ as

$$\begin{cases} m(i) = \text{E}\left(f(i)\right) \\ \kappa\left(i, i'\right) = \text{E}\left[\left(m(i) - f(i)\right)\left(m(i') - f(i')\right)\right] \end{cases}. \tag{5}$$

The probability distributions of GPR can be specified by [30]

$$f(i) \sim \text{GPR}\left(m(i), \kappa\left(i, i'\right)\right). \tag{6}$$

Supposing that the same Gaussian distribution exists between the training set i and the new dataset i', then the corresponding output y' can be calculated by the conditional distribution as [24], [33]

$$p\left(y' | i, y, i'\right) = N\left(y' | \bar{y}', \text{cov}\left(y'\right)\right) \tag{7}$$

with

$$\begin{cases} \bar{y}' = \kappa(i, i')^T[\kappa(i, i)]^{-1}y \\ \text{cov}\left(y'\right) = \kappa\left(i, i'\right) - \kappa(i, i')^T[\kappa(i, i)]^{-1}\kappa\left(i, i'\right) \end{cases} \tag{8}$$

where y', $\bar{y}'$, and $\text{cov}(y')$ are the GPR posterior prediction, its corresponding mean and covariance, respectively. $N()$ indicates a normal distribution; $\kappa(i, i)$, $\kappa(i', i')$, and $\kappa(i, i')$ are the covariance matrices between just training inputs, just validation inputs, as well as training and validation inputs, respectively; y denotes the training output vector. It should be known that the uncertainty quantification of GPR in this article is actually the confidence boundaries to reflect the "scope compliance"

uncertainty of predicted values. This uncertainty is caused by differences between the modeled context and the application context, which is not the same as the standard deviations of measurements [30].

The performance of GPR is fully determined by its $m(i)$ and $\kappa(i, i^{'})$, indicating that the corresponding kernel function must be selected and learned carefully from the training dataset. Among various kernel types, several simple but effective kernel functions are particularly noteworthy.

Squared-exponential (SE) function is a more widely used kernel function given as [30]

$$\kappa_{\mathrm{SE}}\left(i, i^{'}\right) = \sigma_{\mathrm{SE}}^2 \exp\left[\frac{-\left(i - i^{'}\right)^2}{2l_{\mathrm{SE}}^2}\right] \qquad (9)$$

where σ_{SE} and l_{SE} are hyperparameters to control the amplitude and length scales. To some extent, SE kernel belongs to a stationary-type kernel in that the correlations between different points are purely affected by the term $i - i^{'}$, leading to a smooth distribution. This would be too strict for capacity degradation data with many local fluctuations; therefore, an alternative is the Matern32 (M32) kernel function as [30]

$$\kappa_{M32}\left(i, i^{'}\right) = \sigma_{M32}^2 \left[1 + \frac{\sqrt{3}\left(i - i^{'}\right)}{l_{M32}}\right] \exp\left[-\frac{\sqrt{3}\left(i - i^{'}\right)}{l_{M32}}\right] \qquad (10)$$

where σ_{M32} and l_{M32} represent the hyperparameters to control the function amplitude and smoothness, respectively.

In practice, due to the limited capture ability of the SE function and M32 function, these isotropic kernels would provide unreliable predictive results for nonlinear mapping that involves multidimensional input variables. For the calendar aging model, the inputs should not only contain the capacity terms, but also involve the storage temperature and cell SOC. In order to extract these features and improve accuracy, the isotropic SE and M32 kernels are modified with the ARD structure [34], as denoted by (11) and (12).

$$\begin{cases} \kappa_{\mathrm{SE}\sim\mathrm{ARD}}(i, i^{'}) = \sigma_{\mathrm{SE}}^2 \exp[-\frac{1}{2}r] \\ \kappa_{M32\sim\mathrm{ARD}}(i, i^{'}) = \sigma_{M32}^2(1 + \sqrt{3r}) \exp(-\sqrt{3r}) \end{cases} \qquad (11)$$

with

$$r = \frac{\left\|i_T - i_T^{'}\right\|^2}{l_T^2} + \frac{\left\|i_{\mathrm{SOC}} - i_{\mathrm{SOC}}^{'}\right\|^2}{l_{\mathrm{SOC}}^2} + \sum_{C=0}^{k} \frac{\left\|i_C - i_C^{'}\right\|^2}{l_C^2} \qquad (12)$$

where hyperparameters l_T, l_{SOC}, and l_C determine the relevancies of temperature, SOC, and capacities inputs with respect to the regression results, respectively. Generally, a large value leads to a low relevancy. For GPR, the "learning" implies the optimization of hyperparameters within the covariance function, using the training dataset. In this article, a standard gradient descent optimizer is used to fit the hyperparameters of GPR through maximizing the log marginal likelihood [30]. Here the threshold of gradient descent optimizer is $2 \times e^{-5}$, which is defined by Matlab GPR toolbox.

TABLE II
IDENTIFIED PARAMETERS FOR RCL MODEL

Parameters	α_1	α_2	α_3	α_4	p
Values	0.0006	0.5270	-2.7049	-1.0185	0.52

Therefore, ARD structure can be seen as a powerful tool for input features extraction. By using the ARD kernel for calendar aging prediction, irrelevant input features among capacity, storage temperature, and SOC would be effectively removed by fixing large length scales for them, yielding a sparse and explanatory subset of features. Besides, various predictors with different length scales are generated to improve prediction accuracy and robustness.

C. Regression Calendar-Life Model

In order to demonstrate the effectiveness of the proposed GPR+ARD model, a simplified RCL model is also adopted and compared. This RCL model is actually a typical semiempirical model which has been applied in publication [11]. Generally, the capacity loss ΔQ during storage is expressed as a function of the battery SOC $\mathrm{SOC}_{\mathrm{sto}}$, storage temperature T_{sto}, and storage time duration t as [11]

$$\Delta Q = f\left(\mathrm{SOC}_{\mathrm{sto}}, T_{\mathrm{sto}}, t\right). \qquad (13)$$

Then, (13) can be simplified by assuming decoupling. First, both battery SOC and T_s versus time, implying that the calendar degradation trend is similar over time and can be shaped by a coefficient as [11]

$$\Delta Q = C\left(\mathrm{SOC}_{\mathrm{sto}}, T_{\mathrm{sto}}\right) t^p. \qquad (14)$$

In the literature, the shaping coefficient $C(\mathrm{SOC}_{\mathrm{sto}}, T_{\mathrm{sto}})$ is usually assumed to follow the Arrhenius relation, allowing the $\mathrm{SOC}_{\mathrm{sto}}$ and T_{sto} to be decoupled as [11]

$$C\left(\mathrm{SOC}_{\mathrm{sto}}, T_{\mathrm{sto}}\right) = A(z) \exp\left[-\frac{E_a(SOC_{\mathrm{sto}})}{kT_{\mathrm{sto}}}\right]. \qquad (15)$$

It is noteworthy that the activation energy, denoted E_a in (15), could be approximated by an affine linear dependence of SOC [11], as expressed in (16)

$$\Delta Q = \alpha_1 \exp\left[\alpha_2 \cdot SOC_{\mathrm{sto}}\right] \exp\left[(\alpha_3 \cdot SOC_{\mathrm{sto}} + \alpha_4)/T_{\mathrm{sto}}\right] t^p. \qquad (16)$$

For this RCL model, input parameters include the storage SOC, temperature, and time duration. Following the same datasets as GPR, all data from "Group 1" (green cases) are used to train the RCL model, while all data from "Group 2" (yellow cases) are used to validate the trained RCL model. In this article, five parameters (α_1, α_2, α_3, α_4, and p) require to be identified. An advanced heuristic method named biogeography-based optimization (BBO) is adopted to calculate these parameters by minimizing the RMSE between the predicted values and the real test data through 20 independent runs. More details regarding the BBO technique and the identification procedure can be found in [35]. The corresponding identified parameters are shown in Table II.
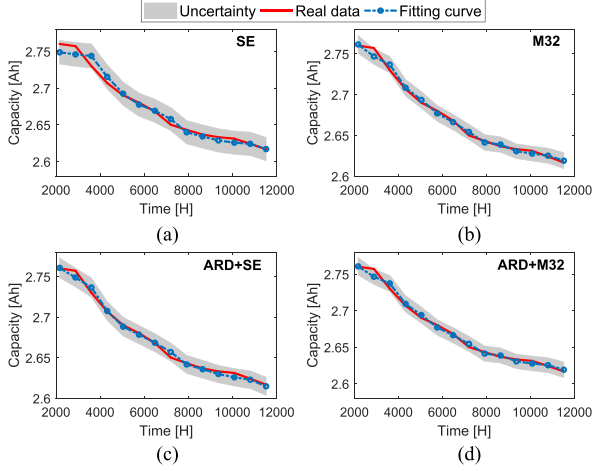
Fig. 6. Training results by using different kernel functions for Case 5 data. (a) SE. (b) M32. (c) ARD+SE. (d) ARD+M32.

Fig. 7. Prediction results by using different kernel functions for Case 6 data. (a) SE. (b) M32. (c) ARD+SE. (d) ARD+M32.

IV. RESULTS AND DISCUSSIONS

A. Performance Comparisons

In this subsection, two comparisons are first conducted to quantify the improvement by using GPR+ARD model for calendar aging prediction.

1) Comparisons of Various Kernel Functions: To evaluate the performance of different kernel functions in the calendar aging prediction domain, four covariance functions, including solo SE, solo M32, SE with ARD kernel (ARD+SE), and M32 with ARD kernel (ARD+M32), are compared with respect to their training and prediction performance. The initial values of all GPR models' hyperparameters are set through using MATLAB GPR toolbox, and defined as follows: for the solo SE and M32 kernels, $\sigma_{SE} = \sigma_{M32} = 0.1$, $l_{SE} = l_{M32} = 2$; for the ARD+SE and ARD+M32 kernels, $\sigma_{SE} = \sigma_{M32} = 0.1$, $l_T = l_{SOC} = 2$, $l_c = 1$. Here, Case 5 with the worst training results and Case 6 with the largest one-step prediction errors are specified for performance comparisons.

Fig. 6 illustrates the training results by using different kernel functions for Case 5 data. It is evident that experimental capacity presents a nonlinear declining trend during storage with 0.5 SOC and 25°C temperature. Although solo SE function can capture the overall degradation trend with the largest confidence range, some points are still mismatched especially at beginning. After modifying SE with the ARD structure, as shown in Fig. 6(c), the training performance can be effectively improved. Here the MAE for ARD+SE case is 0.0091, which is 33.1% less than that at solo SE case. For solo M32 and ARD+M32 kernels, it seems that better training results are obtained for both the cases. After using the ARD structure, the MAE for ARD+M32 case becomes 0.0078, which is 8.3% less than that of the solo M32 case. These satisfactory training results are mainly due to the variable length scales of the ARD structure. We can conclude that with the same calendar aging dataset, the training performance can be improved by using ARD-based kernel functions.

After training, the GPR models with different kernel functions are applied to predict the future capacity in different storage conditions. Fig. 7 and Table III illustrate the prediction results
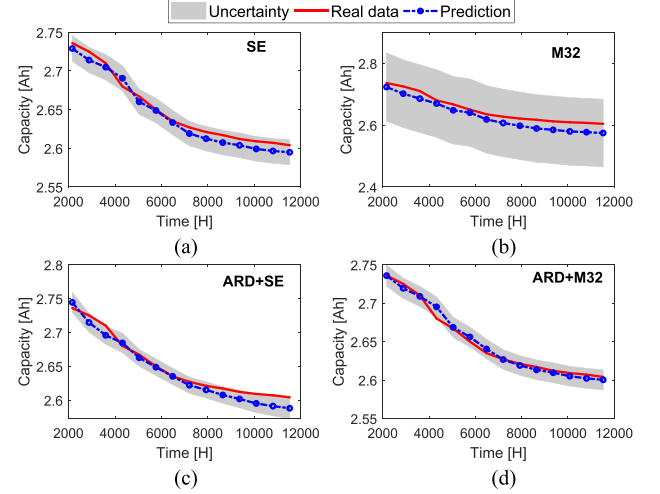
Kernel types.	solo SE	solo M32	ARD+SE	ARD+M32
MAE	0.0161	0.0300	0.0154	**0.0109**
RMSE	0.0100	0.0228	0.0082	**0.0054**
R^2	0.948	0.927	0.957	**0.978**
CS	0.929	1	0.929	**0.954**

and the corresponding quantitative metrics for Case 6 data. It can be observed that by using solo SE kernel, although the obtained 95% confidence range almost covers the overall degradation trend, the mean prediction values still mismatch the real experimental data in most time points (here the CS and RMSE values are 0.929 and 0.0100, respectively). For solo M32 kernel, the prediction values are all lower than the actual values (here MAE, RMSE, and R^2 become the worst ones as 0.0300, 0.0228, and 0.927, respectively). Besides, the corresponding 95% confidence range distributes in a wide region, implying that high uncertainty is achieved in this case. These prediction failures are mainly caused by the overfitting problem, implying the poor generalization ability of solo kernel structure. In comparison, the predicted values get closer to the real capacity data by using the ARD-based SE kernel, indicating the effectiveness of ARD structure. But several mismatch points still exist, especially after 8000 h points for this case, which means the SE kernel cannot capture the overall capacity degradation dynamics. In Fig. 7(d), by using the ARD-based M32 kernel, the capacity trend is well captured as desired. Quantitatively, the RMSE for Case 6 data here is just 0.0054, which is 76.3% and 34.1% less than the solo M32 case and ARD+SE case, respectively. Besides, the 95% confidence range distributes in a narrow region for such a case, indicating a high credibility for the prediction results. This satisfactory performance is caused by the strong feature extraction abilities of ARD and high robustness of M32 kernel. Accordingly, ARD-based M32 kernel is selected for predicting calendar aging in the following studies.

2) Comparisons of Training Results for GPR and RCL Models: Next, in order to further evaluate the effectiveness of
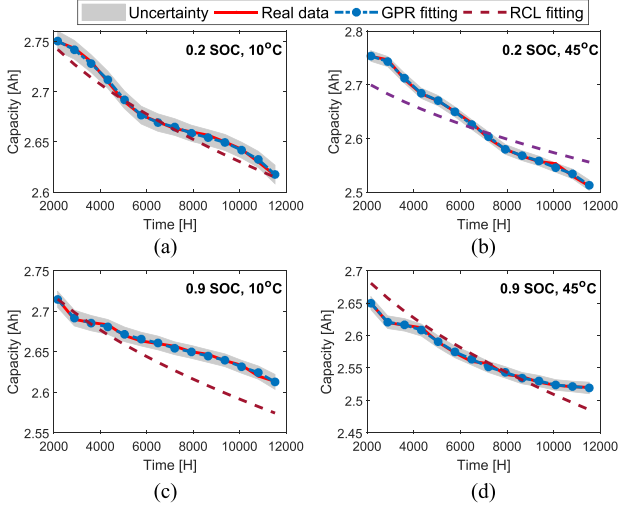
Fig. 8.   Comparisons of training results by using GPR model and RCL model. (a) Case 1. (b) Case 3. (c) Case 7. (d) Case 9.

Fig. 9.   Prediction results by using full data from "Group 1" cases as the training set. (a) Case 2. (b) Case 4. (c) Case 6. (d) Case 8.

TABLE IV
QUANTITATIVE TRAINING METRICS FOR GPR MODEL AND RCL MODEL

Case No.	Case 1	Case 3	Case 7	Case 9
MAE (RCL)	0.0214	0.0432	0.0463	0.0401
MAE (GPR)	0.0042	0.0041	0.0046	0.0029
RMSE (GPR)	0.0012	0.0010	0.0009	0.0008
CS (GPR)	0.954	0.954	0.954	0.954

TABLE V
PREDICTION QUANTITATIVE METRICS AFTER FULL-DATA TRAINING

Case No.	Case 2	Case 4	Case 6	Case 8
MAE	0.0115	0.0107	0.0122	0.0076
RMSE	0.0043	0.0051	0.0052	0.0046
R^2	0.981	0.985	0.978	0.988
CS	0.929	0.929	0.929	0.954

the GPR model for calendar aging prediction, the RCL model is used as a comparison. Fig. 8 and Table IV illustrate the training results and the corresponding quantitative metrics for different storage conditions, respectively. It is worth noticing that the RCL model gives a general trend of capacity aging without the direct uncertainty quantification for the predicted values. Even for the best fitting results obtained under 0.2 SOC and 10 °C storage temperature, the MAE is larger than those of GPR model, respectively. For the remaining three cases of "Group 2" validation, the corresponding fitting results also present large differences with the measured data, implying that this RCL model is inadequate to capture capacity degradation of our dataset case. The main reason that makes RCL model bad would be the unrecorded initial capacity fading for such dataset. In comparison, by using the ARD-based GPR model, both the overall capacity decline trend and local nonlinear fluctuations are well fitted as desired. From Table IV, the MAE and RMSE for all cases by using GPR model are within 0.005 and 0.0012, respectively. Moreover, the CS values are all equal to 0.954, indicating the high training accuracy and good generalization ability by using our proposed GPR+ARD data-driven model.

B. Full-Data Training Results of ARD+M32

Fig. 9 and Table V present the prediction results and the corresponding quantitative metrics for "Group 2" cases after full-data training based on the "Group 1" cases. It can be seen that the predicted capacity values for all cases match the actual data well. The trained ARD-based GPR model captures the overall
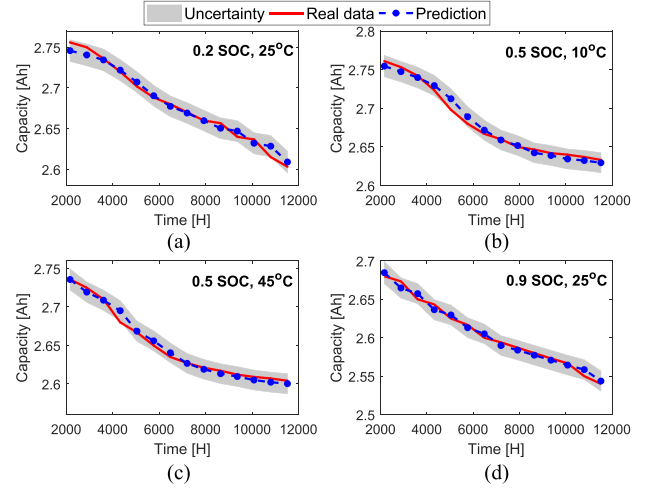
capacity degradation trends well as the RMSE of all predicted samples are less than 0.006. Besides, among all samples, the maximum MAE value is 0.0122, obtained for Case 6 data. This is mainly caused by the insufficient training data as only Case 5 covers 0.5 SOC condition. However, this MAE is still less than 0.5% capacity range, indicating that high accuracy is also achieved for such a case. For Case 8 with 0.9 SOC and 25 °C storage temperature, the corresponding CS value reaches 0.954, which means that the actual results are all covered within the obtained confidence range. Interestingly, CS values for other cases are all 0.929, implying that the corresponding confidence ranges are also reliable. Therefore, it can be concluded that the full-data-trained GPR model with ARD structure is effective and highly accurate for battery calendar aging prediction under various storage conditions.

C. Partial-Data Training Results

For GPR technique, the inclusion of a larger number of relevant data could lead to explain the data better and learn more underlying mapping information, further resulting in more accurate prediction results, and narrower confidence boundaries. However, collecting calendar aging data under various storage conditions is an extremely time-consuming process in real-world applications. In such a case, it is meaningful to develop a reliable model with a satisfactory accuracy level based on partial training data. To evaluate the partial-data training results and the corresponding prediction performance of the proposed ARD-based GPR model, the capacity data before 8000 h of all "Group 1"
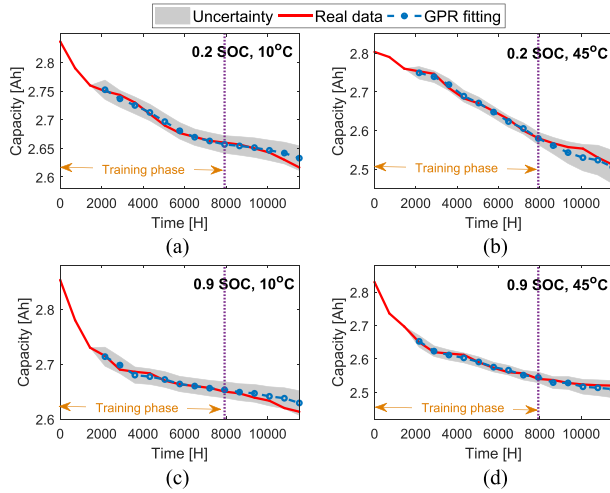
Fig. 10. Training results by using partial data from "Group 1" cases as the training set. (a) Case 1. (b) Case 3. (c) Case 7. (d) Case 9.

Fig. 11. Prediction results by using partial data from "Group 1" cases as the training set. (a) Case 2. (b) Case 4. (c) Case 6. and (d) Case 8.

cases (nearly 7/3 split) are chosen as the dataset for the training part, while the remaining data are employed as the validation set.

Fig. 10 illustrates the results for "Group 1" cases based on the partial-data training. From Fig. 10, it is observed that for various cases with different storage conditions, the capacity values are highly similar to the real data in training phase, indicating that an accurate fitting result is obtained for our GPR model. After 8000 h, apart from Case 9 that still presents the highly similar trend with just 0.0032 MAE, other cases all show more or less differences. Specifically, Case 1 and Case 7 obtain the MAE values of 0.0116 and 0.0154 at 11 520 h, respectively. Case 3 achieves 0.0123 MAE at 10 800 h. It therefore proves that decreasing the training data will result in the information loss of capacity fade in calendar aging, further reducing the extrapolation and generalization performance of the trained model. Even so, by using the GPR model with the ARD structure, all MAE values are still less than 0.5% capacity range, which means that the training results are still reliable. To further evaluate the prediction results of partial-data training, all data from "Group 2" cases are then employed as the validation set.

After training the GPR model based on the partial data from "Group 1" cases, the prediction results for "Group 2" cases are presented in Fig. 11. Moreover, detailed quantitative metrics are examined in Table VI. It is seen that Case 4 presents much higher accuracy in the whole validation process with the smallest values of MAE (0.0112) and RMSE (0.0057), respectively. For Case 2, the corresponding RMSE is 0.0065, indicating that satisfactory overall capacity prediction is also achieved. Here the MAE is 0.0128, caused by a short-period mismatch around 11 000 h. From Fig 11(d), the predicted values present more fluctuations in comparison with those in the full-data training case. Quantitatively, here the MAE and RMSE for Case 8 become 0.0118 (55.3% increase) and 0.0061 (32.6% increase). Even so, the result of Case 8 still presents a satisfactory capacity prediction. In comparison, Case 6 has the worst prediction due to several mismatches occurring after 7000 h. The RMSE for
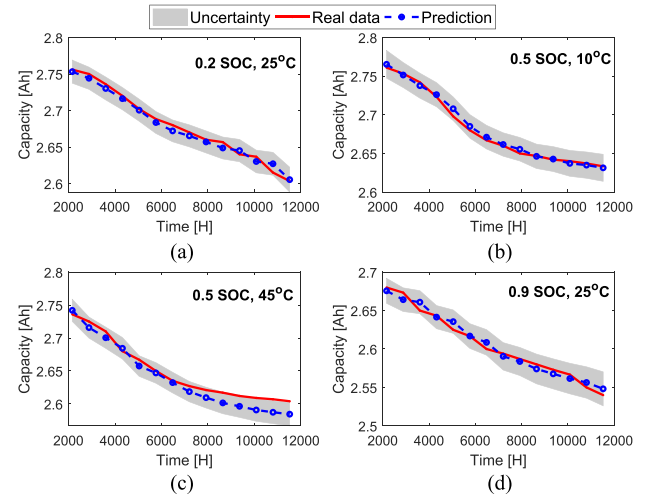
TABLE VI
PREDICTION QUANTITATIVE METRICS AFTER PARTIAL-DATA TRAINING

Case No.	Case 2	Case 4	Case 6	Case 8
MAE	0.0128	0.0112	0.0167	0.0118
RMSE	0.0065	0.0057	0.0084	0.0061
R^2	0.964	0.979	0.959	0.972
CS	0.929	0.929	0.896	0.954

Case 6 reaches 0.0084, which is 61.5% more than that under the full-data training. This result is reasonable due to the decreased capacity characteristics covered by using partial-data training. However, the MAE for Case 6 is still less than 0.6% capacity range (here is 0.0167), which means that the corresponding predicted accuracy is also acceptable. Besides, the CS values for all cases are all larger than 0.896, implying that the confidence levels are reliable. In conclusion, these facts signify that with a suitable partial-data training, the proposed GPR model is also capable of excavating the useful information, therefore providing reliable and accurate prediction results for calendar aging under various conditions.

D. Multistep Calendar Aging Prediction

Multistep calendar aging prediction is more meaningful in real-world applications as it can provide the entire future trend of capacity degradation. To evaluate the multistep prediction performance of our proposed ARD-based GPR model, a multistep prediction test is conducted for all "Group 2" cases in comparison with the RCL model.

In this test, after obtaining a new predicted capacity value by our GPR+ARD model, a recursive process is iteratively conducted to predict future capacity until the last one is achieved. It should be known that due to the structure as illustrated in Fig. 5, ARD+GPR model requires the information of first $k + 1$ historical capacity points ($k = 2$ in the article), while the RCL model just requires the initial one capacity point. Here the comparison between two models is conducted after the $k + 1$th
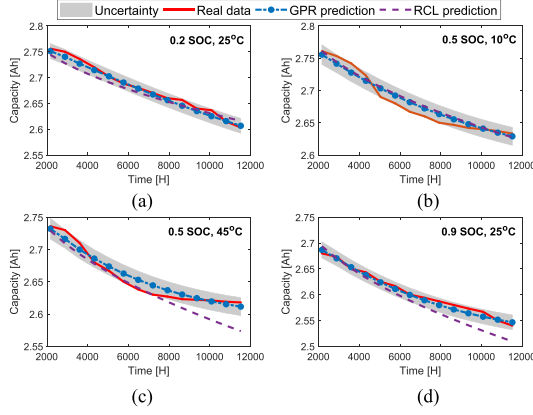
Fig. 12. Multistep calendar aging prediction results of ARD-GPR model and RCL model. (a) Case 2. (b) Case 4. (c) Case 6. (d) Case 8.

TABLE VII
QUANTITATIVE METRICS FOR MULTISTEP PREDICTION

Case No.	Case 2	Case 4	Case 6	Case 8
MAE (RCL)	0.0223	0.0167	0.0472	0.0413
MAE (GPR)	0.0132	0.0148	0.0152	0.0124
RMSE (GPR)	0.0073	0.0102	0.0104	0.0069
R^2 (GPR)	0.961	0.946	0.943	0.970
CS (GPR)	0.954	0.911	0.907	0.954

capacity points. Fig. 12 and Table VII illustrate the multistep prediction results and the corresponding quantitative metrics for both RCL and GPR+ARD models, respectively. From Fig. 12, it can be observed that relatively large predicted mismatches exist for cases by using the RCL model (here the worst MAE reaches 0.0472 for Case 6 at 11 520 h), indicating the poor generalization ability of the RCL model for our dataset. In comparison, by using the GPR+ARD model, although several mismatches occur especially in large local fluctuations of Case 4 and Case 6, the entire capacity decline trends are still captured reliably for all cases. These increased local mismatches are reasonable as the predicted errors are accumulated for multistep conditions. Here the RMSE for Case 4 and Case 6 become 0.0102 and 0.0104, which are nearly twice larger than those of one-step prediction cases. These results are mainly caused by the poor training dataset for 0.5 SOC condition (just Case 5 owns the information of 0.5 SOC). However, the worst MAE is still within 0.7% capacity range (here it is 0.0152 for Case 6), indicating that the corresponding multistep prediction results are acceptable for all cases. Moreover, all the obtained uncertainty ranges cover the local fluctuations. It can be concluded that even for the multistep prediction, the developed GPR+ARD model can capture the overall capacity degradation trends well with an acceptable confidence level.

E. Prediction at New Condition Through Accelerated Aging Data Training

In the real world, batteries experience a wide range of storage temperatures and SOCs. Developing a lifetime model based on converting accelerated aging data to predict new degradation
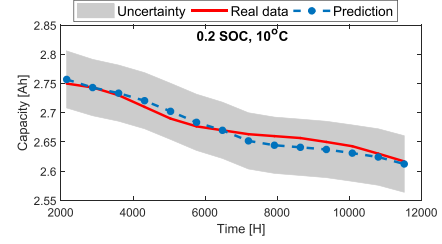

Fig. 13. Prediction result of Case 1 through training based on the accelerated aging data.

TABLE VIII
QUANTITATIVE METRICS AFTER ACCELERATED AGING TRAINING

Case No.	Case 5	Case 6	Case 8	Case 9	**Case 1**
MAE	0.0066	0.0082	0.0074	0.0089	0.0162
RMSE	0.0027	0.0035	0.0031	0.0037	0.0106
R^2	0.989	0.978	0.983	0.975	0.942
CS	0.954	0.954	0.954	0.929	1

case is another promising research topic [32], [36]. To evaluate the corresponding performance of proposed GPR+ARD model, a test regarding the entirely new condition prediction is conducted in this subsection.

In this test, GPR+ARD model is trained through using accelerated aging data. Specifically, the aging data under the relatively high SOCs and temperatures from Case 5, Case 6, Case 8, and Case 9 are utilized for model training. After that, data from Case 1 that represents an entirely new storage condition are used for validating our proposed model. Fig. 13 and Table VIII illustrate the prediction results of Case 1 after accelerated aging data training and the corresponding quantitative metrics, respectively. One obvious observation is that the obtained uncertainty bounds become relatively wider than the tests from previous subsections. This is hardly surprising given that temperature and SOC in Case 1 are both different from the utilized accelerated aging data. In such situations, the covariance values calculated by the kernel function are smaller, leading to the broader confidence boundaries. However, it is clear that these uncertainty bounds still cover the real data. The overall capacity degradation trend of Case 1 can be captured by the predicted capacity values, indicating that the proposed GPR+ARD model also presents effectiveness for such a case.

F. Further Discussions

Due to the lack of exploiting machine learning-based approaches for calendar aging prediction in the existing published work, for the first time, this article focuses especially on the development of the GPR technique with ARD kernel to achieve satisfactory capacity prediction under various storage conditions. In this article, the calendar aging dataset is acquired from an OEM automotive company with some initial degradation due to the reduced begin-of-life (BOL) capacity of the battery. Then, the observed trends would inevitably decrease the prediction performance of the utilized RCL model, while favoring the step-by-step GPR model. Based upon our test results, several useful

observations can be made. 1) In order to take full advantage of the semiempirical model, a well-designed aging test that covers the battery's nominal capacity and considers overhang effects is recommended [28]. 2) In real-world applications, missing data related to any usage and subsequent degradation of a cell is a practical and likely scenario to occur. In such circumstances, our proposed GPR+ARD model outperforms the RCL model with regard to prediction performance and uncertainty quantification. 3) To avoid underfitting problem of pure data-driven technique, a dataset covering enough useful information is suggested in the training phase [14], [30]. Future work could include an effective combination of the proposed GPR technique with battery electrochemical knowledge or electrothermal models, and the performance improvements in research areas such as the conversion of accelerated aging data to predict entirely new degradation cases, and the holistic aging predictions regarding both calendar and cycling modes.

V. CONCLUSION

In this article, effective capacity prognosis under various storage conditions for Li-ion batteries was presented. The GPR technique with ARD kernel was employed to synthesize a data-driven model for battery calendar aging prediction. Based upon the GPR toolbox of MATLAB 2018 with a 2.40 GHz Intel Pentium 4 CPU, our proposed GPR model can be well trained within 10 s. Illustrative results corroborated that the ARD+M32 kernel outperforms other kernels in both training and validation processes (here the MAE and RMSE are less than 0.011 and 0.0055 for all cases). Based upon our measured dataset, such GPR model exhibits improved prediction performance with higher accuracy and better generalization ability. Moreover, the uncertainty level of predicted results can be considered simultaneously. Even for the partial-data training test, multistep prediction test, and accelerated aging training test, the predicted results were satisfactory in terms of the accuracy (here the worst RMSE were less than 0.0105) and the reliable confidence range for various storage conditions. Without any requirements of electrochemical knowledge it is worth noting that the proposed model can be easily extended to other battery types for resilient calendar aging prognosis.

REFERENCES

[1] X. Hu, D. Cao, and B. Egardt, "Condition monitoring in advanced battery management systems: Moving horizon estimation using a reduced electrochemical model," *IEEE/ASME Trans. Mechatronics*, vol. 23, no. 1, pp. 167–178, Feb. 2018.

[2] S. Grolleau *et al.*, "Calendar aging of commercial graphite/LiFePO₄ cell-predicting capacity fade under time dependent storage conditions," *J. Power Sources*, vol. 255, pp. 450–458, 2014.

[3] M. Dubarry, N. Qin, and P. Brooker, "Calendar aging of commercial li-ion cells of different chemistries–A review," *Current Opinion Electrochemistry*, vol. 9, pp. 106–113, 2018.

[4] S. S. Zhang, "A review on electrolyte additives for lithium-ion batteries," *J. Power Sources*, vol. 162, no. 2, pp. 1379–1394, 2006.

[5] L. Su *et al.*, "Path dependence of lithium ion cells aging under storage conditions," *J. Power Sources*, vol. 315, pp. 35–46, 2016.

[6] M. Ecker *et al.*, "Calendar and cycle life study of Li(NiMnCo)O₂-based 18650 lithium-ion batteries," *J. Power Sources*, vol. 248, pp. 839–851, 2014.

[7] H. Ekström and G. Lindbergh, "A model for predicting capacity fade due to SEI formation in a commercial graphite/LiFePO₄ cell," *J. Electrochemical Soc.*, vol. 162, no. 6, pp. A1003–A1007, 2015.

[8] K. L. Gering, "Novel method for evaluation and prediction of capacity loss metrics in li-ion electrochemical cells," *Electrochimica Acta*, vol. 228, pp. 636–651, 2017.

[9] J. Schmalstieg, S. Käbitz, M. Ecker, and D. U. Sauer, "A holistic aging model for Li(NiMnCo)O₂ based 18650 lithium-ion batteries," *J. Power Sources*, vol. 257, pp. 325–334, 2014.

[10] M. Petit, E. Prada, and V. Sauvant-Moynot, "Development of an empirical aging model for Li-ion batteries and application to assess the impact of vehicle-to-grid strategies on battery lifetime," *Appl. Energy*, vol. 172, pp. 398–407, 2016.

[11] E. Redondo-Iglesias, P. Venet, and S. Pelissier, "Eyring acceleration model for predicting calendar ageing of lithium-ion batteries," *J. Energy Storage*, vol. 13, pp. 176–183, 2017.

[12] J. De Hoog *et al.*, "Combined cycling and calendar capacity fade modeling of a nickel-manganese-cobalt oxide cell with real-life profile validation," *Appl. Energy*, vol. 200, pp. 47–61, 2017.

[13] S. L. Hahn, M. Storch, R. Swaminathan, B. Obry, J. Bandlow, and K. P. Birke, "Quantitative validation of calendar aging models for lithium-ion batteries," *J. Power Sources*, vol. 400, pp. 402–414, 2018.

[14] M. Lucu, E. Martinez-Laserna, I. Gandiaga, and H. Camblong, "A critical review on self-adaptive Li-ion battery ageing models," *J. Power Sources*, vol. 401, pp. 85–101, 2018.

[15] D. Yang, Y. Wang, R. Pan, R. Chen, and Z. Chen, "State-of-health estimation for the lithium-ion battery based on support vector regression," *Appl. Energy*, vol. 227, pp. 273–283, 2018.

[16] J. Wei, G. Dong, and Z. Chen, "Remaining useful life prediction and state of health diagnosis for lithium-ion batteries using particle filter and support vector regression," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5634–5643, Jul. 2018.

[17] X. Hu, J. Jiang, D. Cao, and B. Egardt, "Battery health prognosis for electric vehicles using sample entropy and sparse Bayesian predictive modeling," *IEEE Trans. Ind. Electron.*, vol. 63, no. 4, pp. 2645–2656, Apr. 2016.

[18] C. Hu, H. Ye, G. Jain, and C. Schmidt, "Remaining useful life assessment of lithium-ion batteries in implantable medical devices," *J. Power Sources*, vol. 375, pp. 118–130, 2018.

[19] C. Sbarufatti, M. Corbetta, M. Giglio, and F. Cadini, "Adaptive prognosis of lithium-ion batteries based on the combination of particle filters and radial basis function neural networks," *J. Power Sources*, vol. 344, pp. 128–140, 2017.

[20] G.-W. You, S. Park, and D. Oh, "Diagnosis of electric vehicle batteries using recurrent neural networks," *IEEE Trans. Ind. Electron.*, vol. 64, no. 6, pp. 4885–4893, Jun. 2017.

[21] Y. Zhang, R. Xiong, H. He, and M. Pecht, "Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 5695–5705, Jul. 2018.

[22] M. H. Lipu *et al.*, "A review of state of health and remaining useful life estimation methods for lithium-ion battery in electric vehicles: Challenges and recommendations," *J. Cleaner Prod.*, vol. 205, pp. 115–133, 2018.

[23] K. Liu, K. Li, Q. Peng, and C. Zhang, "A brief review on key technologies in the battery management system of electric vehicles," *Frontiers Mech. Eng.*, vol. 14, no. 1, pp. 47–64, 2019.

[24] R. R. Richardson, C. R. Birkl, M. A. Osborne, and D. Howey, "Gaussian process regression for in-situ capacity estimation of lithium-ion batteries," *IEEE Trans. Ind. Informat.*, vol. 15, no. 1, pp. 127–138, Jan. 2019.

[25] G. O. Sahinoglu, M. Pajovic, Z. Sahinoglu, Y. Wang, P. V. Orlik, and T. Wada, "Battery state-of-charge estimation based on regular/recurrent Gaussian process regression," *IEEE Trans. Ind. Electron.*, vol. 65, no. 5, pp. 4311–4321, May 2018.

[26] D. Yang, X. Zhang, R. Pan, Y. Wang, and Z. Chen, "A novel Gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve," *J. Power Sources*, vol. 384, pp. 387–395, 2018.

[27] P. Keil, "Aging of lithium-ion batteries in electric vehicles," Ph.D. dissertation, Inst. Elect. Energy Storage Technol., Technische Universität München, 2017.

[28] M. Lewerenz, G. Fuchs, L. Becker, and D. U. Sauer, "Irreversible calendar aging and quantification of the reversible capacity loss caused by anode overhang," *J. Energy Storage*, vol. 18, pp. 149–159, 2018.

[29] B. Gyenes, D. Stevens, V. Chevrier, and J. Dahn, "Understanding anomalous behavior in coulombic efficiency measurements on Li-ion batteries," *J. Electrochemical Soc.*, vol. 162, no. 3, pp. A278–A283, 2015.

[30] C. K. Williams and C. E. Rasmussen, *Gaussian Processes for Machine Learning*, vol. 2. Cambridge, MA, USA: MIT Press, 2006.

[31] K. Liu, K. Li, Q. Peng, Y. Guo, and L. Zhang, "Data-driven hybrid internal temperature estimation approach for battery thermal management," *Complexity*, vol. 2018, 2018.

[32] M. Ecker *et al.*, "Development of a lifetime prediction model for lithium-ion batteries based on extended accelerated aging test data," *J. Power Sources*, vol. 215, pp. 248–257, 2012.

[33] R. R. Richardson, M. A. Osborne, and D. A. Howey, "Gaussian process regression for forecasting battery state of health," *J. Power Sources*, vol. 357, pp. 209–219, 2017.

[34] J. Zhao, L. Chen, W. Pedrycz, and W. Wang, "Variational inference based automatic relevance determination kernel for embedded feature selection of noisy industrial data," *IEEE Trans. Ind. Electron.*, vol. 66, no. 1, pp. 416–428, Jan. 2019.

[35] K. Liu, C. Zou, K. Li, and T. Wik, "Charging pattern optimization for lithium-ion batteries with an electrothermal-aging model," *IEEE Trans. Ind. Informat.*, vol. 14, no. 12, pp. 5463–5474, Dec. 2018.

[36] D.-I. Stroe, M. Świerczyński, A.-I. Stan, R. Teodorescu, and S. J. Andreasen, "Accelerated lifetime testing methodology for lifetime estimation of lithium-ion batteries used in augmented wind power plants," *IEEE Trans. Ind. Appl.*, vol. 50, no. 6, pp. 4006–4017, Nov./Dec. 2014.

Xiaosong Hu (SM'16) received the Ph.D. degree in automotive engineering from the Beijing Institute of Technology, Beijing, China, in 2012, and the second Ph.D. degree with dissertation in automotive research center from the University of Michigan, Ann Arbor, USA, in 2012.

He is currently a Professor with the State Key Laboratory of Mechanical Transmissions and with the Department of Automotive Engineering, Chongqing University, Chongqing, China. He was a Postdoctoral Researcher with the Department of Civil and Environmental Engineering, University of California, Berkeley, CA, USA, between 2014 and 2015, as well as with the Swedish Hybrid Vehicle Center and the Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden, between 2012 and 2014. He was also a Visiting Postdoctoral Researcher with the Institute for Dynamic Systems and Control, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, in 2014. His research interests include battery management technologies and modeling and controls of electrified vehicles.

Dr. Hu has published more than 100 high-caliber journal/conference papers. He has been a recipient of several prestigious awards/honors, including SAE Ralph Teetor Educational Award in 2019, Emerging Sustainability Leaders Award in 2016, EU Marie Curie Fellowship in 2015, ASME DSCD Energy Systems Best Paper Award in 2015, and Beijing Best Ph.D. Dissertation Award in 2013.

Kailong Liu (M'18) received the B.Eng. degree in electrical engineering and the M.Sc. degree in control theory and control engineering from Shanghai University, China, in 2011 and 2014, respectively, and the Ph.D. degree in electrical engineering from the Energy, Power and Intelligent Control Group, Queen's University Belfast, Belfast, U.K., in 2018.
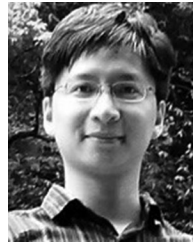
He is a Research Fellow with the Warwick Manufacturing Group, University of Warwick, U.K. He was a Visiting Student Researcher with Tsinghua University and North China Electric Power University, China, in 2016. His research interests include modeling, optimization, and control with applications to electrical/hybrid vehicles, energy storage, and battery management systems.

Dr. Liu was the Student Chair of IEEE QUB student branch and a recipient of awards such as EPSRC Scholarship, Santander International Scholarship, and QUB ESM International Scholarship.

Mattin Lucu (S'19) received the M.Sc. degree in integration of renewable energy sources into the electricity grid from the University of the Basque Country, UPV-EHU, Spain, in 2016, and is currently working toward the Ph.D. degree in control engineering with the IK4-Ikerlan Technology Research Centre, University of the Basque Country, Barrio Sarriena, Spain.

During his graduate studies, he worked as R&D intern successively with the EneR-GEA research group (ESTIA Engineering School, France) in wind turbine emulation and control, and with the EDP in the analysis of photovoltaic power insertion in low-voltage distribution networks. In 2016, he joined the IK4-Ikerlan Technology Research Centre. His research interests include electrochemical energy storage systems and machine learning algorithms applied to their lifetime estimation.

Widanalage Dhammika Widanage received the honour's (first class) degree in electronic and communication engineering and the Ph.D. degree in system identification from the University of Warwick, in 2004 and 2008, respectively.

He is currently an Assistant Professor of Modeling and Energy Storage with the WMG, Warwick University, Coventry, U.K. He leads the battery modeling research with the department and recently secured funding to lead the modeling activity as a PI (for WMG) on the Faraday Multiscale Modeling project, as a Co-Investigator (C0-I) on the EPSRC Prosperity Partnership with Jaguar Land Rover, and a PI and Co-I on four Innovate UK projects (PI and three Co-I). His research interests include system identification theory, applied across several applications including batteries.

Dr. Widanalage is a member of the European Materials Modelling Community Interoperability and Repository Advisory Group (EMMC-IRAG) and supported the compilation of the materials modeling-terminology, classification, and metadata document which is now publicly available to increase the use of material modeling with industrial end-users. He was the recipient of the WMG Early Career Researcher of the Year Award in 2016.

Yi Li received the B.E. degree in chemical engineering from the East China University of Science and Technology, Shanghai, China, in 2008, the M.S. degree in material science from the Technical University of Munich, Munich, Germany, in 2012, and the Ph.D degree in electrical engineering from Vrije Universiteit Brussels, Brussels, Belgium, in 2019.

She is currently a Senior Research Associate with Lancaster University, U.K. Her research interests include lithium-ion battery aging mechanism identification, state of health estimation, and lifetime prediction.