A hybrid machine learning framework for joint SOC and SOH estimation of lithium-ion batteries assisted with fiber sensor measurements

Yihuan Li [a,b], Kang Li [b,*], Xuan Liu [b], Xiang Li [b], Li Zhang [c], Bruno Rente [d], Tong Sun [d], Kenneth T.V. Grattan [d]

[a] School of Control and Computer Engineering, North China Electric Power University, Beijing, 102206, China
[b] School of Electronic and Electrical Engineering, University of Leeds, LS2 9JT, UK
[c] School of Mechatronics and Automation, Shanghai University, Shanghai 200072, China
[d] School of Science and Technology; City, University of London, London EC1V 0HB, UK

A B S T R A C T

Accurate State of Charge (SOC) and State of Health (SOH) estimation is crucial to ensure safe and reliable operation of battery systems. Considering the intrinsic couplings between SOC and SOH, a joint estimation framework is preferred in real-life applications where batteries degrade over time. Yet, it faces a few challenges such as limited measurements of key parameters such as strain and temperature distributions, difficult extraction of suitable features for modeling, and uncertainties arising from both the measurements and models. To address these challenges, this paper first uses Fiber Bragg Grating (FBG) sensors to obtain more process related signals by attaching them to the cell surface to capture multi-point strain and temperature variation signals due to battery charging/discharging operations. Then a hybrid machine learning framework for joint estimation of SOC and capacity (a key indicator of SOH) is developed, which uses a convolutional neural network combined with the Gaussian Process Regression method to produce both mean and variance information of the state estimates, and the joint estimation accuracy is improved by automatic extraction of useful features from the enriched measurements assisted with FBG sensors. The test results verify that the accuracy and reliability of the SOC estimation can be significantly improved by updating the capacity estimation and utilizing the FBG measurements, achieving up to 85.58% error reduction and 42.7% reduction of the estimation standard deviation.

1. Introduction

To facilitate low-carbon transition of the economy is now a priority in the development agenda of many countries and regions worldwide to mitigate the climate change [1]. One of the key technologies to bridge a carbon neutral future is the battery technology which enables the integration of more renewable energy in the power grid and reduces the greenhouse gas emissions from the transportation sector. The lithium-ion batteries have the merits of long cycle life, high energy density, low self-discharge rate, environmental resilience, and continual decrease of manufacturing costs, which make them overwhelmingly attractive for electric vehicle (EV) and grid energy storage applications [2]. Lithium-ion batteries have to be operated within a proper range in terms of temperature, charging and discharging currents, etc., and violations of the operation conditions will lead to performance degradation, thermal runaway and even explosion, hence strict requirements have been imposed on the safety standards and regulations of battery storage systems [3]. To ensure the operation safety and reliability, and enhance the durability of the battery, an effective and reliable battery management system (BMS) is required for internal states estimation, charge/discharge control, and planned maintenance, etc. [4,5].

The internal state estimation is still an challenging task in BMS, and the two key states, namely state of charge (SOC) and state of health (SOH), have been extensively researched over the years. SOC reflects the real-time remaining capacity of the battery, and has fast time-varying dynamics. SOH reflects the aging or degradation level of the battery, and has slow time-varying dynamics. Further, battery aging will degrade the SOC estimation accuracy, while the capacity is one of the key and widely used indicators of the battery state of health (SOH) to quantitatively assess the battery aging level. Therefore, it is vital for the BMS to accurately estimate the SOC in real-time and calibrate the capacity regularly [3]. Numerous SOC and capacity estimation methods

have been proposed in the literature, which can be roughly categorized into model-based methods and machine learning (ML)-based methods. The performance of the model-based estimation methods is highly dependent on the model accuracy, thus an effective battery model that can well describe the system dynamics is a prerequisite. However, the complexity of batteries internal physicochemical reactions and the uncertainty of external operation environment makes it difficult to build accurate battery models. Therefore, ML techniques have been increasingly employed to estimate the battery SOC and capacity due to their flexibility, reliability, strong adaptability and generalizability.

The main benefit of the ML-based methods is that a priori knowledge of the battery dynamics is no longer required. Besides, various real-world operating conditions can be considered during the model training by adding additional inputs to the model, and this approach is therefore suitable for all types of batteries [6,7]. Techniques such as deep neural networks (DNNs) [8], recurrent neural networks (RNNs) [9], and Gaussian process regression (GPR) [10] can directly map the measured signals (e.g. current, terminal voltage, and surface temperature) to the SOC. In general, the direct measurements are often used as the model inputs to calculate the SOC as the model output. For example, in [11], the battery SOC was estimated using DNNs and the experimental results confirm that the DNN with four hidden layers has the best generalization capability across several drive cycles. In [12], the SOC was estimated using a RNN with gated recurrent unit, which can exploit information of the previous SOCs and measurements and achieve better estimation results than traditional feed-forward neural networks. In [13], a GPR framework was used for SOC estimation under three different ambient temperatures. Compared with the aforementioned ML techniques which only provide the point estimation of SOC, the GPR can not only estimate the SOC using measured quantities, but also quantify the uncertainty of the SOC estimations. The uncertainty quantification can assess the reliability of the estimated results, thus provides more useful information for the decision making in BMS. Besides, methods such as support vector machine (SVM) [14], relevance vector machine (RVM) [15], long short-term memory (LSTM) [16], and convolutional neural network (CNN) [17], just to name a few, have been successfully applied in battery capacity estimation. CNNs have the characteristics of automatic feature extraction and low overfitting risk, and have demonstrated a great potential in battery capacity estimation. In [18], the battery capacity was estimated using a CNN model that combines the concepts of transfer learning and ensemble learning, and the resultant CNN model can be applied to a relatively small dataset while the estimation accuracy and robustness on unseen dataset are also improved. In [19], a CNN-based capacity estimation framework was incorporated with the concepts of transfer learning and network pruning, leading to improved estimation accuracy on small dataset while the size and computational complexity of the model are both greatly reduced.

Most of the existing researches including the aforementioned ones estimate the battery internal states such as SOC and SOH using the traditional externally measured signals such as current, voltage, and surface temperature of a particular location. These external measurements are limited and can hardly reflect the battery electro-chemical-thermal behavior and inhomogeneity of key parameters such as battery thermal distribution profile, hence limiting the state estimation accuracy and leading to over-conservative usage of the battery [20] or potential battery failure [21]. Fiber optic sensors (FOSs) [22], which, in summary, are immune to electromagnetic interference (EMI), robust to corrosive environments, have multiplexing capability with small dimensions, are attractive solutions for distributed battery sensing applications [23]. The latest technology can make it possible to inscribe hundreds or thousands of Fiber Bragg Grating (FBG) sensors into a single optic fiber [24], allowing simultaneous multiple point measurements with much simpler wiring diagram. Moreover, optical fibers are much smaller and lighter than electrical wires and, together with this multiplexing capability, a large number of FBG sensors can be installed for large structures (such as large battery energy storage systems) with much less cable mass and volume. These distinctive features of fiber optic sensors make it possible to develop a very low-cost sensing mechanism for densely instrumenting very large structures, and both the cost and wiring complexity of fiber optic sensing networks can be much lower than fully distributed conventional sensors [25]. Further, the additional measurements allow extraction of a richer set of features for more accurate battery internal state estimation, and are more reliable under strong electromagnetic interference than electric signals. For example, during the charge and discharge processes, the cell electrode volume changes due to Li-ions intercalation/ deintercalation processes in the electrode materials, which manifests as the changes in the strain on the cell surface. The stability and safety of the battery can be affected by these induced strains, particularly in harsh operation conditions, and it may become one of the main reasons leading to potential material failure and other forms of performance degradation if the strains exceed certain level of thresholds [26]. Besides, the cell temperature also changes during the charge and discharge processes due to the electrochemical reactions, resistive heating, and enthalpy changes [21]. Temperature is often monitored to provide early warning of potential thermal hazards, which generally manifest as the thermal runaway and may cause irreversible damage to battery cells when it exceeds certain threshold. These parameters hence can help gain deeper insight into the internal dynamics of the battery, and are regarded as complementary signals to the tradition current and voltage measurements.

FBG sensors are sensitive to strain and temperature and thus their applications in BMS have attracted increased attentions in recent years. To demonstrate the potentials of FBG sensors to aid state estimation in BMS, Sommer et al. [27] attached a pair of FBG sensors externally to lithium-ion pouch cells to monitor intercalation state transition points across various charge/discharge rates. Sharp and repeatable features correlated with intercalation state transitions can be observed which enable better state estimation. Nascimento et al. [28] presented a comparative study of surface temperature monitoring performance between the thermocouples and fiber sensors, and demonstrated that the FBG sensors were better choices for surface temperature monitoring under normal and abuse operating conditions and failure detection. After exploring the relationship between the external FBG sensing signals and diffusion processes [20], Raghavan et al. [23] embedded the FBG sensors inside pouch cells to directly monitor the internal temperature and electrode strain, the results have revealed that batteries with embedded FBG sensors are highly comparable to those without FBG sensors in terms of seal integrity, capacity retention and projected cycle life. Peng et al. [29,30] designed sensitivity-enhanced FBG sensors and mounted them onto the cell surface, the relationship between the strain and SOC/depth of discharge was investigated. Ganguli et al. [31] estimated the SOC and SOH using dynamic time warping and Kalman filtering algorithms based on the internal strain signals obtained from these FBG sensors at different temperature conditions. Similarly, Rente et al. [32] achieved accurate SOC estimation results using dynamic time warping algorithm by correlating the cell surface strain data obtained from FBG sensors with the SOC, the results indicated that installing FBG sensors on the cell surface is a feasible, cost-effective and non-invasive approach for assisting SOC estimation.

Leveraging the latest developments in machine learning approaches and fiber optic sensing technologies in battery condition monitoring, this paper proposes a hybrid machine learning framework for joint estimation of the battery SOC and capacity, the two key internal states for battery management. The main contributions of this paper include: (1) The joint estimation framework takes into account of the intrinsic coupling relationship between SOC and SOH, by using the regularly updated capacity information for SOC estimation, the SOC can be estimated more accurately than siloed estimation methods. (2) The fiber optic sensors are adopted for strain and surface temperature measurements for battery SOC estimation to further improve the accuracy of SOC estimation. (3) GPR is used for battery SOC estimation,

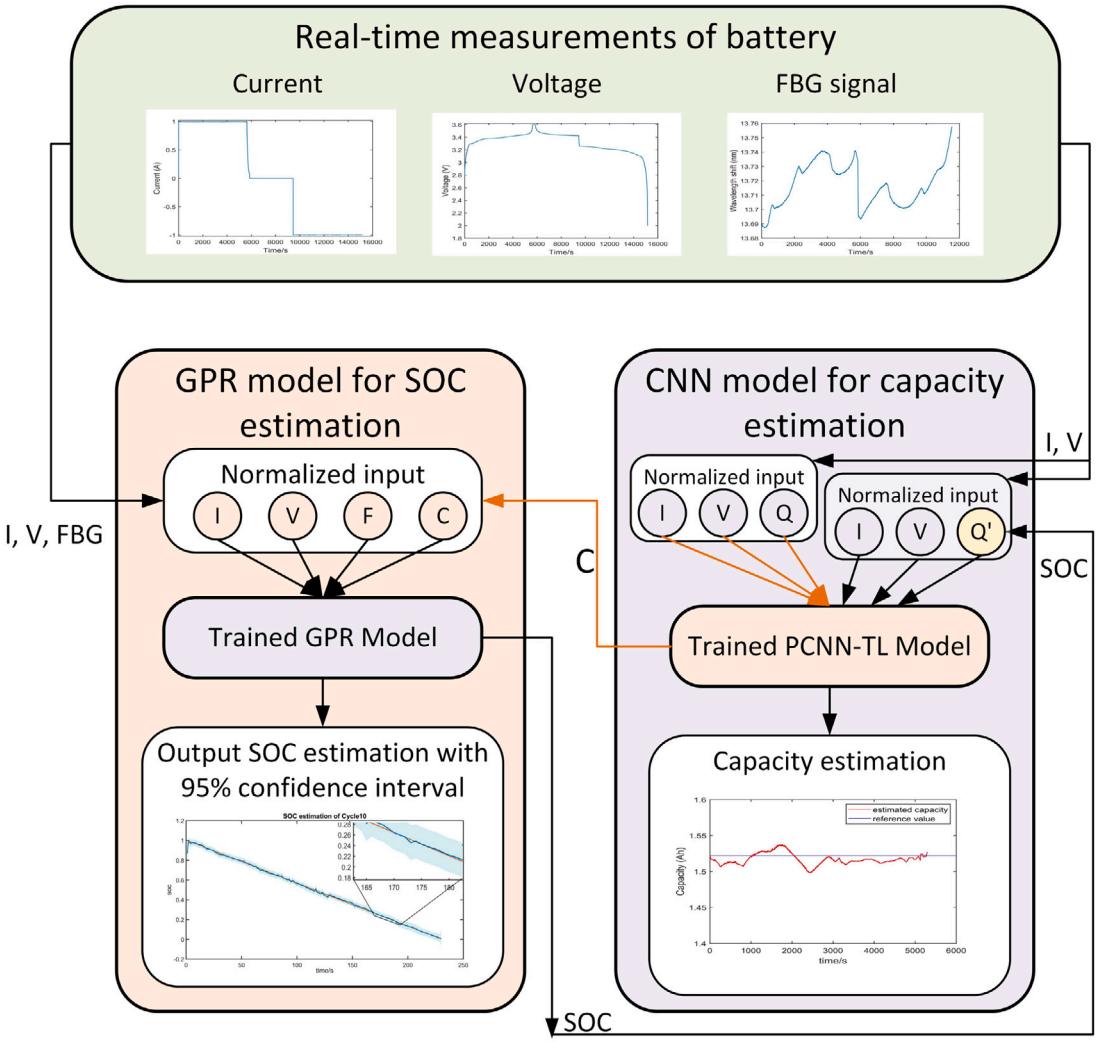

Fig. 1. The flowchart of the proposed battery capacity and SOC joint estimation framework.

which not only provides point estimate of the SOC, but also quantifies the uncertainty of the estimation. This allows up to 85.58% RMSE reduction in SOC estimation and 42.7% reduction in the estimation standard deviation. To the best of our knowledge, this is one of the first attempts to use the GPR method to extract information from FBG measurements for more accurate statistic estimation of the SOC, enabling more trustful management of battery energy storage systems for safer and more reliable battery operations.

The remainder of this paper is organized as follows. Section 2 details the proposed joint estimation framework and introduces the GPR algorithm used for SOC estimation. Section 3 introduces the experimental setup and the new signals measured by FBG sensors. Section 4 presents and discusses the experimental results. Finally, Section 5 concludes the paper.

2. Methodology

In this section, the detailed battery SOC and capacity joint estimation framework is presented. Firstly, the proposed joint estimation framework is introduced. Secondly, a brief overview of the Gaussian Process Regression (GPR) theory is introduced and the implementation procedure of GPR-based battery SOC estimation method is detailed. Further, the CNN-based battery capacity estimation method, which was proposed in our earlier work [19], is briefly introduced.

2.1. The joint estimation framework

Fig. 1 shows the flowchart of the proposed battery capacity and SOC joint estimation framework. The main steps of this framework are introduced as follows: at each time instant, battery current, voltage, and FBG signals are first sampled in real-time, then the current, voltage and the charge capacity are normalized and transformed to a three-dimensional (3-D) image with the size of $15 \times 15 \times 3$ and inputted to the PCNN-TL (pruned convolution neural network with transfer learning) model trained in [19] for online estimation of the battery capacity. Subsequently, the estimated capacity is used in the GPR model to correct the imprecise capacity estimation value, and the updated capacity along with other normalized measurements are used to estimate the SOC online. Finally, the charge capacity calculated by integrating the current with respect to time is replaced by the charge capacity calculated using the estimated SOC, and the new 3-D input generated with the updated charge capacity is fed into the PCNN-TL model to estimate the capacity.

The joint estimation framework is practically more attractive in industrial applications and it can provide more accurate estimates than the traditional state estimation methods since it utilizes the coupling relationships between the capacity and SOC. By online updating the capacity value in the SOC estimation, the impact of battery degradation is taken into account. Therefore the performance of SOC estimation for aged batteries can be further improved.

2.2. SOC estimation

2.2.1. Gaussian process regression

Gaussian process regression (GPR), which is a probabilistic and non-parametric machine learning method, is used for battery SOC estimation in this paper. The GPR method is capable of quantifying the uncertainty of the estimation rather than just provide a point estimate of the SOC, and hence providing more informative outputs than the Kalman Filter (KF) algorithm and its variants. In essence, based on the GPR method, the estimation result of SOC is given in the form of probability distribution, which consists of the mean of the estimation value and confidence intervals.

Let $D = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^{N}$ denote a labeled training dataset with N samples, where $\mathbf{x}_i \in \Re^D$ is a D-dimensional input vector, and $\mathbf{y}_i \in \Re$ is the corresponding output. Suppose that there exists a latent function $f(.)$, to map inputs $\mathbf{x}_i$ to outputs $\mathbf{y}_i$:

$$\mathbf{y}_i = f(\mathbf{x}_i) + \varepsilon_i \tag{1}$$

where $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$ is an independent and identically distributed noise contribution.

In the GPR, the function $f(\mathbf{x})$ is assumed to follow a multivariate Gaussian distribution, and can be described as:

$$f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), K(\mathbf{x}, \mathbf{x})) \tag{2}$$

where $\mathcal{GP}$ denotes a Gaussian process. The mean function $m(\mathbf{x})$ and covariance function $K(\mathbf{x}, \mathbf{x})$, which can fully describe the function $f(\mathbf{x})$, are denoted by:

$$m(\mathbf{x}) = E(f(\mathbf{x})) \tag{3}$$

$$\begin{aligned} K(\mathbf{x}, \mathbf{x}) &= E[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))] \\ &= \begin{bmatrix} \kappa(\mathbf{x}_1, \mathbf{x}_1) & \kappa(\mathbf{x}_1, \mathbf{x}_2) & \dots & \kappa(\mathbf{x}_1, \mathbf{x}_N) \\ \kappa(\mathbf{x}_2, \mathbf{x}_1) & \kappa(\mathbf{x}_2, \mathbf{x}_2) & \dots & \kappa(\mathbf{x}_2, \mathbf{x}_N) \\ \dots & \dots & \dots & \dots \\ \kappa(\mathbf{x}_N, \mathbf{x}_1) & \kappa(\mathbf{x}_N, \mathbf{x}_2) & \dots & \kappa(\mathbf{x}_N, \mathbf{x}_N) \end{bmatrix} \end{aligned} \tag{4}$$

The mean function reflects the expected function value at input $\mathbf{x}$, and the prior mean function is often set to zero in order to avoid expensive posterior computations and hence only the covariance function is inferred [33]. The covariance function $K(\mathbf{x}, \mathbf{x})$, also called the kernel of the Gaussian process, reflects the dependence between the function values at different input points $\mathbf{x}_i$ and $\mathbf{x}_j$. All the assumptions on the properties of the function to be modeled, such as smoothness and periodicity, are reflected in the covariance function. The squared exponential (SE) kernel is common and it is defined as:

$$\kappa(\mathbf{x}_i, \mathbf{x}_i) = \sigma_f^2 exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_i\|^2}{2\lambda^2}\right) \tag{5}$$

where σ_f^2 denotes the signal variance that quantifies the variation of the latent function from its mean, and λ is the characteristic length scale that determines the relative importance of the input variables in estimating the target output.

Based on Eq. (1) and (2), the joint distribution of the training output $\mathbf{y}$ can be expressed as:

$$\mathbf{y} \sim \mathcal{N}(0, K(\mathbf{x}, \mathbf{x}) + \sigma^2 \mathbf{I}) \tag{6}$$

where $\mathbf{I}$ is a $N \times N$ unit matrix. Generally, the unknown hyperparameters $\Theta = (\sigma^2, \sigma_f^2, \lambda)$ of the covariance function need to be optimized in the training process by maximizing the logarithm of the marginal likelihood function of output $\mathbf{y}$. The log marginal likelihood is given by:

$$\log p(\mathbf{y}|\mathbf{x}, \Theta) = -\frac{1}{2}\mathbf{y}^T[K(\mathbf{x}, \mathbf{x}) + \sigma^2\mathbf{I}]^{-1}\mathbf{y} - \frac{1}{2}\log\left|K(\mathbf{x}, \mathbf{x}) + \sigma^2\mathbf{I}\right| - \frac{N}{2}\log 2\pi \tag{7}$$

After obtaining the optimal hyperparameters using the gradient-based method, and given a testing dataset $D_* = \{(\mathbf{x}_{*i}, \mathbf{y}_{*i})\}_{i=1}^{N*}$, the joint multivariate Gaussian distribution of the training output $\mathbf{y}$ and the testing output $\mathbf{y}_*$ can be written as:

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{y}_* \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} K(\mathbf{x}, \mathbf{x}) + \sigma^2\mathbf{I} & K(\mathbf{x}, \mathbf{x}_*) \\ K(\mathbf{x}_*, \mathbf{x}) & K(\mathbf{x}_*, \mathbf{x}_*) \end{bmatrix}\right) \tag{8}$$

where $K(\mathbf{x}, \mathbf{x}_*)$ is the covariance matrix between the testing inputs and the training inputs and $K(\mathbf{x}, \mathbf{x}_*)^T = K(\mathbf{x}_*, \mathbf{x})$, and $K(\mathbf{x}_*, \mathbf{x}_*)$ is the covariance matrix of testing inputs $\mathbf{x}_*$. Then the predictive posterior distribution is derived for the estimation on the new/testing inputs $\mathbf{x}_*$, which can be completely specified by the mean and covariance:

$$\mathbf{y}_*|\mathbf{x}_*, \mathbf{x}, \mathbf{y} \sim \mathcal{N}(\bar{\mathbf{y}}_*, K_*) \tag{9}$$

where the mean $\bar{\mathbf{y}}_*$ of the predictive distribution, which gives the point estimate of the testing output, is given by:

$$\bar{\mathbf{y}}_* = K(\mathbf{x}_*, \mathbf{x})\left[K(\mathbf{x}, \mathbf{x}) + \sigma^2\mathbf{I}\right]^{-1}\mathbf{y} \tag{10}$$

and the covariance matrix K_* provides a measure of uncertainty in the estimate of the test output [34], and it is given by:

$$K_* = K(\mathbf{x}_*, \mathbf{x}_*) - K(\mathbf{x}_*, \mathbf{x})\left[K(\mathbf{x}, \mathbf{x}) + \sigma^2\mathbf{I}\right]^{-1}K(\mathbf{x}, \mathbf{x}_*) \tag{11}$$

2.2.2. GPR-based SOC estimation

In this paper, the GPR is used to estimate the battery SOC for given measurement inputs. As shown in Fig. 2, the input variables to the GPR model are current $I_j(k)$, voltage $V_j(k)$ and wavelength data of the FBG sensors $FBG_j(k)$ at time k in the jth cycle, and the capacity information of the corresponding cycle is updated by the estimation output of the PCNN-TL model (as trained in [19]) to improve the SOC estimation results. Here the wavelength data is obtained from the fiber optic sensors, and two important parameters (i.e. battery surface strain and temperature) can be decoded and extracted from the wavelength for the characterization of the lithiation/delithiation process. The model output is the estimated SOC at time k, denoted by $SOC(k)$.

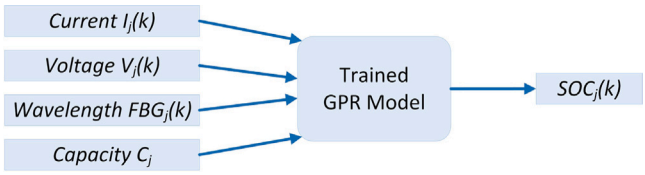

Fig. 2. GPR-based battery SOC estimation.

The GPR-based SOC estimation method mainly consists of two parts, offline training of the model and online estimation of SOC using the trained model. The steps for training a GPR model and then performing SOC estimation are illustrated in Fig. 3, where the blue part on the left represents the offline training process, and the orange part on the right refers to the testing process. The detailed steps can be described as follows:

Training process:

Step 1 — Determine and normalize the training dataset, $D = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^{N}$, where $\mathbf{x}$ contains current, voltage and wavelength measurements as well as the estimated capacity of the corresponding cycle, and $\mathbf{y}$ is the reference value of SOC.

Step 2 — Select a kernel function that can well represent the underlying target function.

Step 3 — Set the initial values for the hyperparameters in the specified kernel function as well as the noise variance.

Step 4 — Optimize the hyperparameters with the training data by maximizing Eq. (7), the log marginal likelihood function.

Testing process:

Step 5 — With the optimal hyperparameters, the GPR model is referred to as the 'trained' model. Then in the testing process, the normalized

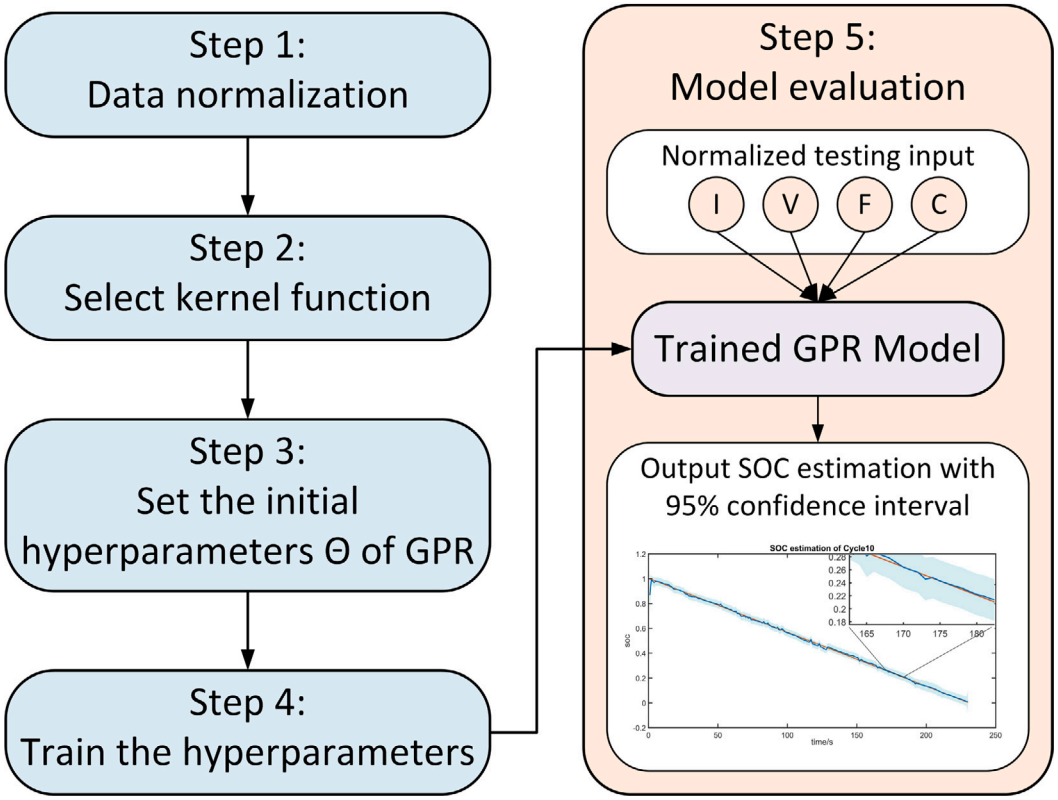

Fig. 3. Flowchart of the GPR-based battery SOC estimation.

testing inputs are fed into the trained GPR model, and the target SOC is outputted in the form of mean and covariance, which provides both the SOC estimation results and the uncertainty measurements.

2.3. CNN-based capacity estimation

The battery aging during its utilization will lead to capacity degradation, which will impact the accuracy of SOC estimation. Considering the importance of timely maintenance and replacement of aged batteries, and the requirement for improved SOC estimation performance, it is meaningful to update the actual capacity in real-time.

In this section, the PCNN-TL model described in our previous work [19] is used for the battery capacity estimation. As shown in Fig. 4, a CNN model that consists of 2 sets of convolutional and max pooling layers, followed by two consecutive convolutional layers, then flattening, and finally two fully-connected layers, was constructed first. Generally, such a model needs to be trained with a large set of labeled training data to optimize its parameters, and will show poor performance if trained on insufficient dataset. Therefore, transfer learning technique was applied to CNN aiming at reducing the required size of datasets by leveraging the knowledge learned from the source task with large dataset to a different but related task with much smaller dataset. As illustrated in Fig. 4, the CNN model was firstly pre-trained on a large source dataset collected from lithium iron phosphate (LFP) cells, each of which was tested for approximately 1000 charge/discharge cycles. Then the learned knowledge (model structure and trained parameters) was transferred to our relatively small target dataset collected from cells that each tested for 30 reference cycles, and the last two convolutional layers and two fully-connected layers were fine-tuned on the target dataset to guarantee model performance. Here the suitable number of fine-tuning layers was determined by trial and error. Finally, a fast recursive algorithm-based network pruning technique was used in the last two fully-connected layers to remove redundant neurons, which can significantly reduce the model size and computational complexity, thus makes it possible to implement the resultant model in the on-board BMS. The resultant model is denoted as PCNN-TL and can achieve fast and accurate capacity estimation on small dataset. More details of this model construction can be found in a previous publication [19].

The input variables of the PCNN-TL model consists of current, voltage and charge capacity of partial charging curves, which are converted into 3 dimensional images with the size of 15 × 15 × 3. While the model output is the maximum available capacity of the discharge cycle which immediately follows the charge cycle that generates the input sample. The capacity estimated by the PCNN-TL model is then used as an input to the GPR model introduced in Section 2.2.2, along with battery current, voltage and FBG wavelength data to update the SOC estimation.

3. Experimental setup

In this work, 4 commercial cylindrical LFP cells with a nominal voltage of 3.2 V and a nominal capacity of 1.6 Ah are used in the experiment. They are tested in parallel using a BTS 4000 battery test system made by NEWARE, and thermocouples with measurement error less than 0.1 °C are attached to measure the cell surface temperature. All cells are tested under a constant temperature of 25 °C, and the battery current, voltage and surface temperature are recorded during the charging and discharging process. High charging current rates are used to accelerate the aging speed of these cells, and a reference cycle is tested every 30 cycles with CC-CV charging and CC discharging process. The CC charging and discharging current is 1 A, with the upper and lower cutoff voltage of 3.6 V and 2.0 V, respectively. The cutoff current of the 3.6 V CV charging process is 75 mA. The sampling frequency for all the equipment used in this experiment was set as 1 Hz. Further, fiber-optic sensors, a promising new sensing technology for battery cell monitoring, are used in this experiment to acquire more process related measurements to improve the SOC estimation accuracy. As shown

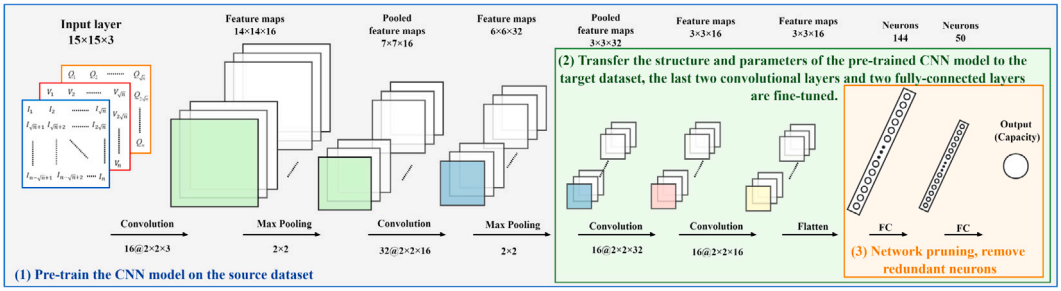

Fig. 4. PCNN-TL model construction.

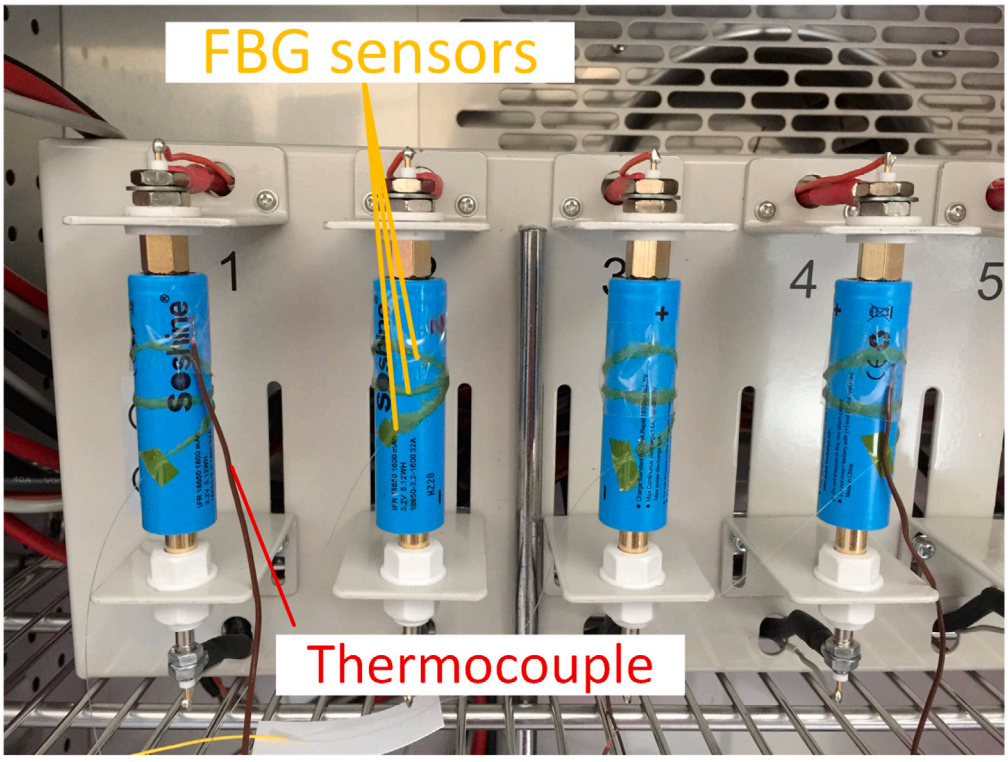

Fig. 5. Cells with FBG sensor integrated.

in Fig. 5, three Fiber Bragg gratings (FBG)-based fiber-optic sensors are directly attached to the surface of each cell without affecting its packaging and integrity. This non-invasive installation approach does not cause potential safety issues and the FBG sensors can be easily mounted on the battery cells.

The FBG sensors are sensitive to strain and temperature variations. These two significant parameters are directly related to the complex processes inside the cells, and the temperature increase and mechanical stress will cause capacity loss and potential risk of batteries. When the battery surface temperature or strain changes, the reflected wavelength changes from the base wavelength λ to λ_s, thus the wavelength shift $\Delta\lambda$ is related to both strain and temperature variations [23,26]. As the three FBG sensors are co-located within a small footprint but have a slightly different orientation, the radial strain signal can be decoupled from the temperature measurement [32]. In this section, the wavelengths that contain both strain and temperature information are directly used for SOC estimation.

The average wavelength shift of the three FBG sensors and the voltage of a cell subject to the CC-CV charging and CC discharging mode is presented in Fig. 6. In Fig. 6, step (1) corresponds to the charging phase with a constant current of 1 A, and step (2) is the 3.6 V constant voltage charging phase, while the final step (3) corresponds to the discharging phase with a constant current of 1 A.

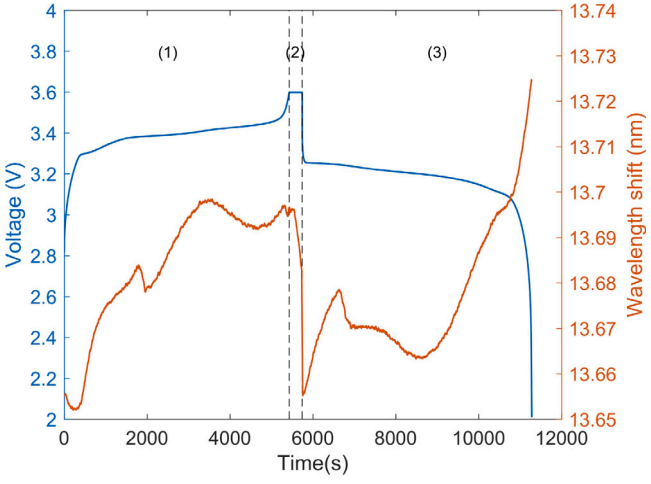

Fig. 6. Wavelength shift of a CC-CV charging and CC discharging cycle.

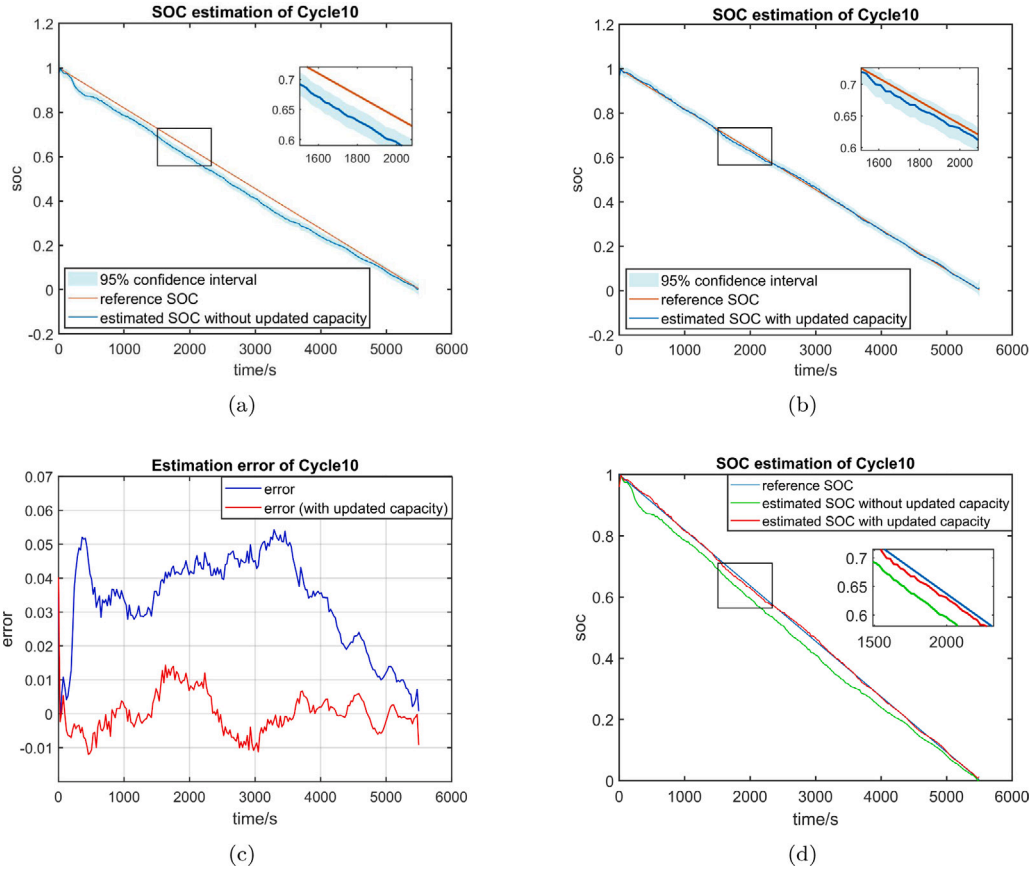

Fig. 7. SOC estimation results of cell 1: (a) SOC estimation results without updated capacity information. (b) SOC estimation results with updated capacity information. (c) SOC estimation error. (d) SOC estimation results with/without updated capacity.

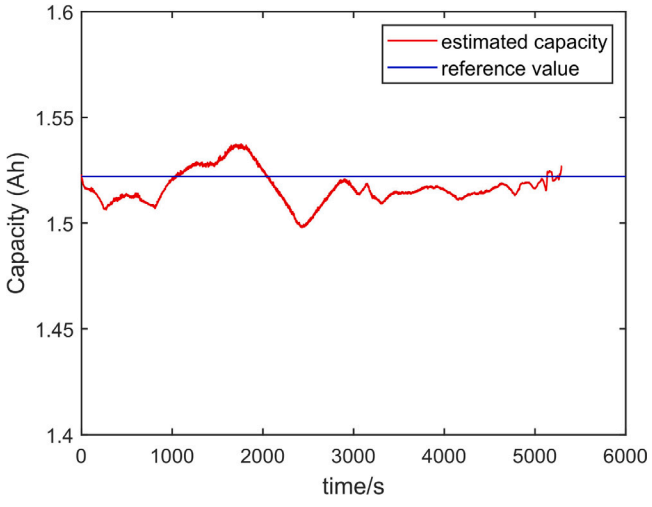

Fig. 8. The estimation results of battery capacity for reference cycle 10 of cell 1.

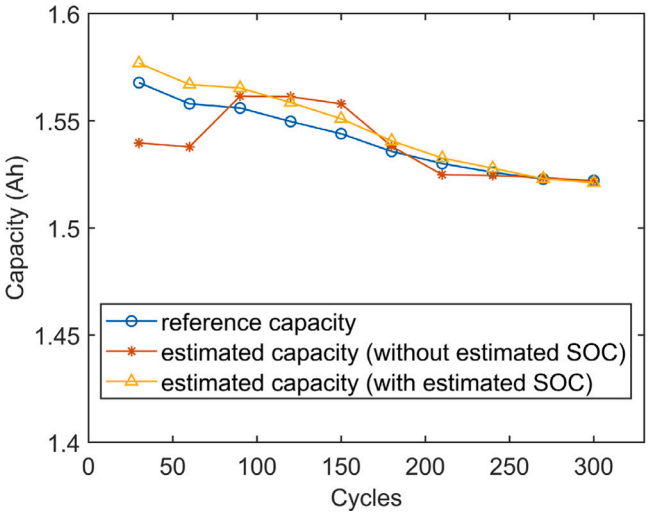

Fig. 9. Capacity estimation results at different reference cycles of cell1.

Since simple functions cannot describe the relationship between SOC and FBG sensor wavelength shift appropriately, therefore, the Gaussian Process Regression algorithm is used to estimate the SOC using this new set of signals.

4. Estimation results and discussions

The performance of the proposed SOC and capacity joint estimation framework is verified on the aforementioned dataset. The root mean-square error (RMSE) is used to evaluate the estimation accuracy and the standard deviation is used to characterize the estimation uncertainty.

The battery capacity is first calculated using the well-trained PCNN-TL model, for which the model inputs are current, voltage, and charge capacity calculated by integrating the current with respect to time. Then the SOC of cell 1 over discharge profiles of reference cycle 10 are estimated and the results are shown in Fig. 7. It is evident that the estimated SOC with updated capacity information (Fig. 7(b)) are much closer to the reference SOC than without it (Fig. 7(a)). This is summarized on Fig. 7(d), where the difference between the curves are

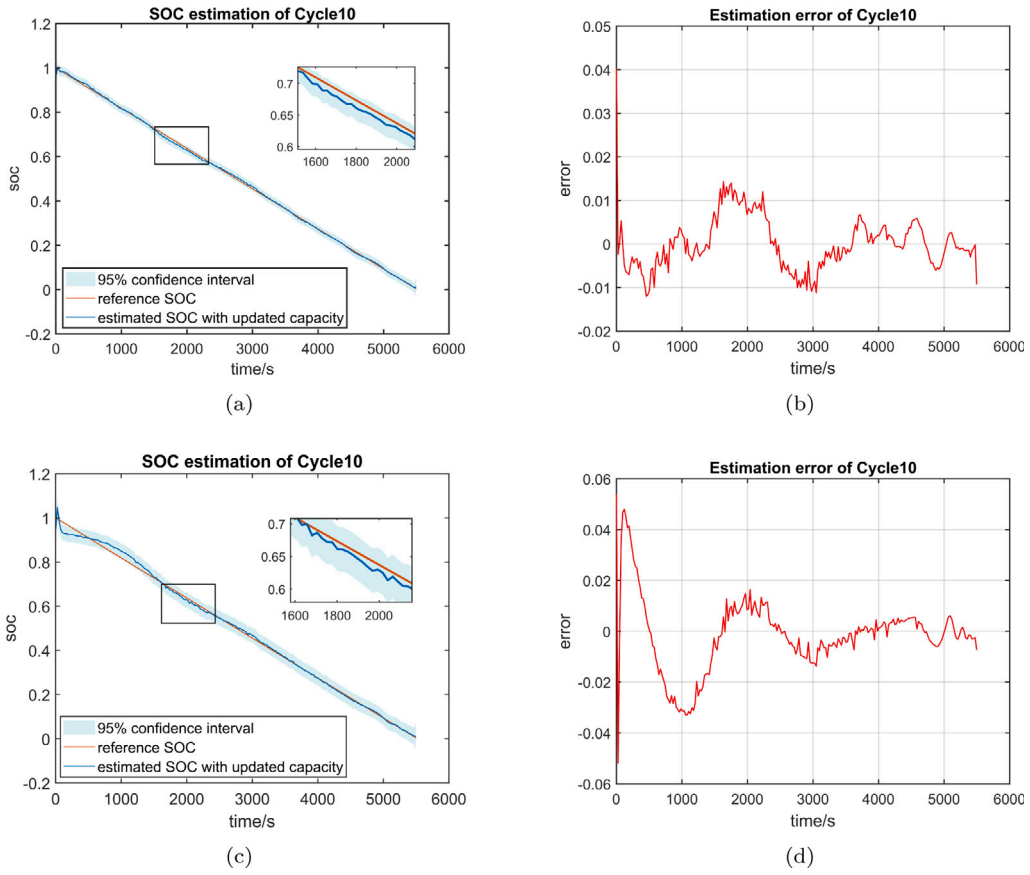

Fig. 10. SOC estimation results of cell 1: (a) SOC estimation results using FBG signals. (b) Error of SOC estimation using FBG signals. (c) SOC estimation results without FBG signals. (d) Error of SOC estimation without FBG signals.

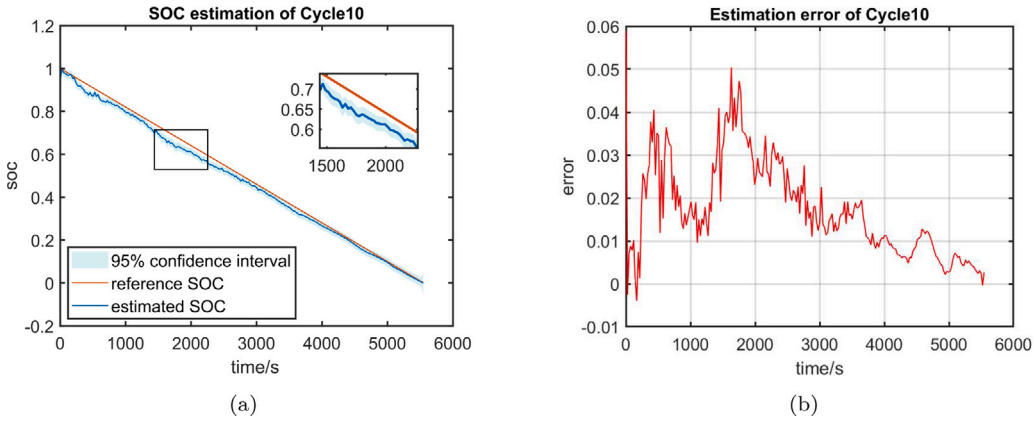

Fig. 11. SOC estimation results of cell 2 using model trained on cell 1: (a) SOC estimation results with updated capacity information. (b) SOC estimation error.

highlighted. As illustrated in Fig. 7(c), though the errors shown in both figures converge to zero, the error of SOC estimation with updated capacity is within 2%, while the error of SOC estimation without updating the capacity is within 6%. Further, the RMSE of the SOC estimation with and without the updated capacity value are 0.62% and 3.59%, respectively. This has clearly demonstrated that accurate capacity estimation is important for SOC estimation. In addition, as shown in Figs. 7(a) and 7(b), the 95% confidence interval in both figures have similar width, and the mean standard deviation of the estimates with and without updated capacity are the same, both are 1.02%. Therefore, it is shown that whether or not the capacity information is updated does not affect the estimation uncertainty.

While estimating the SOC, battery capacity is simultaneously estimated using the well-trained PCNN-TL model, and the estimation results are shown in Fig. 8. The blue and red solid lines represent the reference and estimated capacity value of cycle 10, respectively. As shown in Fig. 8, the estimated capacity is close to the reference value, while its fluctuation is similar to that of the SOC estimation error (red line in Fig. 7(c)), when the error of SOC estimation converges to zero, the estimated capacity also converges to the reference value. Finally, the last 225 consecutive points from the current, voltage and SOC curves of each reference cycle are used to estimate the capacity, and the estimated results for these cycles are illustrated in Fig. 9. The blue line refers to the reference capacity, and the red line represents the capacity estimated using the current, voltage, and charge capacity

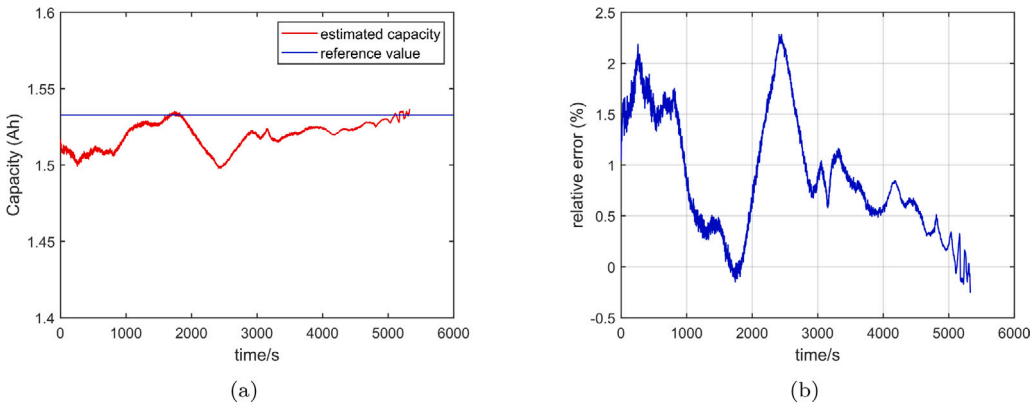

Fig. 12. Capacity estimation results of cell 2 (a)Capacity estimation results using online estimated SOC. (b) The relative error between estimated and reference capacity.

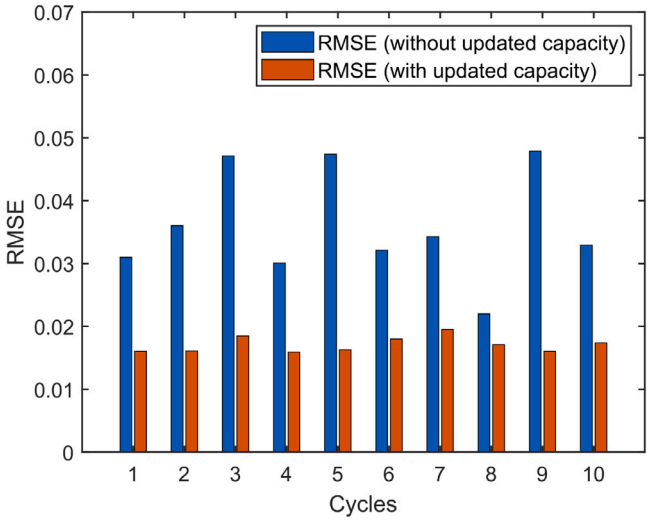

Fig. 13. Bar chart of the SOC estimation RMSE for the discharge profiles of ten reference cycles. The performance of the model trained on cell 1 is evaluated on the data recorded from cell 2.

calculated by integrating the current with respect to time. The RMSE is 0.0126 Ah. While the yellow line refers to the capacity estimated using current, voltage, and the charge capacity calculated from the estimated SOC, and the RMSE is 0.0064 Ah. It is evident that using the estimated SOC information can in turn improve the capacity estimation accuracy.

To investigate the effect of using FBG wavelength signals as model input for the SOC estimation, the data collected from cell 1 was utilized for testing. Two different GPR models were built for SOC estimation, one took current, voltage and capacity as the model input, and the other used current, voltage, FBG wavelength signal, and capacity as the model input. The estimation results on cycle 10 of cell 1 are shown in Fig. 10, and from the enlarged view of Figs. 10(a) and 10(c), it is evident that the shaded blue area is wider when the FBG signal is not fully utilized, which means the estimation uncertainty is higher. Further, Figs. 10(b) and 10(d) confirm that the SOC estimation is more accurate when the FBG signal is utilized as an input to the estimation model. Quantitatively, when using the updated capacity information, the RMSE of the SOC estimation with and without using the FBG signals as input are 0.62% and 1.48%, respectively, and the mean standard deviation of the estimation with and without FBG signals are 1.02% and 1.78%, respectively.

The RMSE and standard deviation of SOC estimation under different input conditions are summarized in Table 1. Two observations can be concluded from Table 1. Firstly, the use of updated capacity can greatly improve the SOC estimation accuracy (reduces the RMSE from 4.3% to 1.48%, or from 3.59% to 0.62%, achieving 65.58% and 82.73% reductions, respectively), but do not affect the estimation uncertainty. Secondly, using FBG signals as input to estimate the SOC can not only reduce the estimation RMSE (the RMSE decrease from 4.3% to 3.59% and from 1.48% to 0.62% achieved 16.51% and 58.11% reductions, respectively), but also reduce the estimation uncertainty (achieving 42.7% reduction on the estimation standard deviation from 1.78% to 1.02%). In summary, using the updated capacity and FBG measurements can reduce the estimation RMSE by up to 85.58% (from 4.3% to 0.62%) and reduce the estimation standard deviation by 42.7%.

To validate the generalization ability of the proposed method, the model trained on cell 1 is directly applied to estimate the SOC of cell 2, and satisfactory SOC estimation results are also achieved. Taking reference cycle 10 of cell 2 as an example, as shown in Fig. 11, the estimation results can still well track the reference values, though the max error is around 5%, it converges to zero at the end of the cycle. A similar trend can be observed from the capacity estimation results shown in Fig. 12, when the SOC estimation error is large, the capacity estimation error is large, while the estimated SOC converged to the reference SOC, the estimated capacity also converged to the reference value. Further, the RMSEs for the discharge profiles of the first ten reference cycles of cell 2 are summarized in Fig. 13. As shown in Fig. 13, the RMSE of SOC estimation is always lower when the capacity is estimated and updated in SOC estimation, which is less than 2% for these ten reference cycles, while the RMSE of the model prediction without using the updated capacity is within 5%, the results again confirm that the proposed joint estimation framework can estimate the SOC more accurately by updating the capacity value.

Finally, the proposed method is also tested on a CC discharge cycle of cell 1 with 1C discharging rate. As shown in Fig. 14, the estimated SOC can track the reference values with the RMSE of 3.48%, and the maximum error around 6%. Table 2 summarizes the RMSE and standard deviation of SOC estimated under different input conditions, the results again confirm that the use of updated capacity can improve the SOC estimation accuracy (reduces the RMSE from 10.42% to 4.89%, or from 7.74% to 3.48%, respectively), and the use of FBG signals not only improves the estimation accuracy (reduces the RMSE from 10.42% to 7.74% and from 4.89% to 3.48%, respectively), but also reduces the estimation uncertainty (achieving up to 45.96% reduction on the estimation standard deviation from 4.33% to 2.34%). Moreover, the capacity estimation results of 30 discharging cycles with 1C current are illustrated in Fig. 15. It is clear that the yellow line is closer to the reference capacity, which indicates that the capacity estimated using the charge capacity calculated by SOC is more accurate than using the charge capacity calculated by current integration, which once again confirms that the estimated SOC can in turn improve the capacity estimation results.

Table 1
SOC estimation results with/without updated capacity and with/without FBG measurements.

Assess	Without FBG signals		With FBG signals	
	Without updated Q	With updated Q	Without updated Q	With updated Q
RMSE	4.3%	1.48%	3.59%	**0.62%**
Mean standard deviation	1.77%	1.78%	1.02%	**1.02%**

Table 2
SOC estimation results on a 1C discharge cycle with/without updated capacity and with/without FBG measurements.

Assess	Without FBG signals		With FBG signals	
	Without updated Q	With updated Q	Without updated Q	With updated Q
RMSE	10.42%	4.89%	7.74%	**3.48%**
Mean standard deviation	4.33%	3.54%	2.34%	**2.31%**

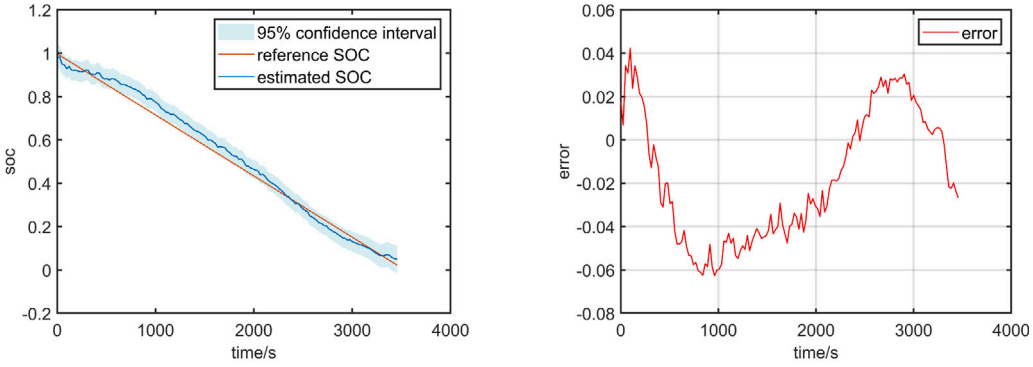

Fig. 14. SOC estimation results on a discharge cycle with 1C current (a) SOC estimation results with updated capacity information. (b) SOC estimation error.

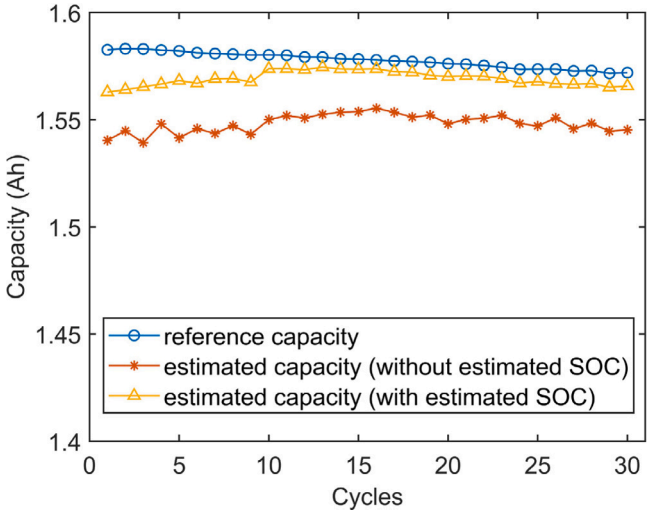

Fig. 15. Capacity estimation results of 30 discharging cycles with 1C current.

In summary, the proposed joint estimation framework has shown to significantly improve the SOC estimation accuracy by updating the imprecise battery capacity in time, and accurate SOC estimation can in turn improve the accuracy of capacity estimation, while traditional SOC estimation methods without capacity calibration cannot eliminate the influence of the erroneous capacity value. Furthermore, the FBG measurements can provide more information on the battery dynamics, therefore, using FBG measurements to assist SOC estimation, the estimation uncertainty can be decreased and estimation accuracy can be improved.

5. Conclusions

This paper has proposed a hybrid machine learning framework to achieve jointly estimation of the battery SOC and capacity, assisted with the instrumentation of FBG sensors to acquire multi-point strain and temperature variation signals due to battery charging/discharging operations. The framework updates the capacity using a well-trained CNN model, and simultaneously estimates the SOC and quantifies the estimation uncertainty using a GPR model. Some conclusions can be drawn below:

- Firstly, the effect of battery aging on SOC estimation is considered, by updating the capacity regularly, the SOC can be estimated more accurately, achieving up to 82.73% reduction on the RMSE. In turn the accurate SOC estimation can further improve the accuracy of the capacity estimation.
- Secondly, a new sensing technology for battery condition monitoring is utilized to enrich the measurements from traditional sensing technologies for batteries. The new sensing technology offers a number of distinctive features, such as multiplexing, lower costs, less complex wiring, and immunity to EMI. The experimental results confirm that the FBG sensing measurements can further improve the SOC estimation accuracy and lower the estimation uncertainty, with up to 58.11% reduction on the RMSE and 42.7% reduction on the estimation standard deviation.
- Thirdly, the GPR algorithm offers a unified framework to incorporate different sensor measurements for simultaneous estimation of the SOC and quantification of the estimation uncertainty, hence the influence of the updated capacity and FBG measurements on the SOC estimation can be assessed systematically.

The experimental results have undeniably verified the performance of the proposed joint estimation framework and that the use of FBG signals is beneficial to SOC estimation. These results confirm that the capacity

estimation is vital for accurate SOC estimation, and the newly introduced FBG signals can further improve the SOC estimation accuracy and reduce the estimation uncertainty. With the mass roll-out of electric vehicles for transport decarbonization and battery storage systems for accepting a large portion of renewable energy into the power grid, continuous, accurate and reliable battery SOC and SOH estimation are not only important in real-life applications for safe, reliable and effective operation and control, it also offers valuable information for whole life-cycle management of batteries, enabling future technological and business innovations to maximize the value chain of batteries. This paper has demonstrated that a holistic approach to integrate novel estimation framework with new sensing technologies such as the FBG sensors can address a number of issues arising from existing siloed approaches in real-life applications, bringing a number of potential tangible benefits which otherwise cannot be achieved by conventional approaches. Our future work will investigate more sophisticated dynamic charging and discharging scenarios under different operation conditions, and explore other application and benefit potentials of the novel sensing technologies when they are combined with battery management schemes and algorithms.

Finally, we would like to emphasize that Fiber Optic Sensors (FOSs) have been chosen as the most suitable means to allow additional practical measurements to be taken. They have been chosen (in addition to conventional, typically electrically-based sensors) due to the inherent advantages that they have for applications such as this battery state monitoring and state estimation. The specific benefits of the use of such FOSs have been discussed by some of the authors in detail in other publications [22] but in summary, FOSs are particularly well suited to these situations, such as when electrical monitoring is unsafe as there is a risk of sparking, or short circuits possible in all electrical systems in harsh environment, thereby offering a significant safety advantage, a key consideration in the design and operation of electric vehicles and battery storage for power grids. Further, FOSs methods such as are used here work well when electrical means of monitoring fail, for example where there is a large amount of electromagnetic noise, and data being read from conventional sensors can be corrupted. In addition, when there are a large number of sensors to be placed (and if it is necessary to place a lot of sensors for a complete monitoring profile of battery systems), FOSs are easy to install and lightweight: the use of an optical network minimizes the weight of cables needed and thus simplifies the set up. Finally, the use of FOS networks can offer a competitive costing option, especially when safety is to the fore in the measurement.

CRediT authorship contribution statement

Yihuan Li: Conceptualization, Methodology, Formal analysis, Software, Writing – original draft, Writing – review & editing. **Kang Li:** Conceptualization, Methodology, Supervision, Resources, Project administration, Funding acquisition, Writing – review & editing. **Xuan Liu:** Validation, Investigation, Data curation, Formal analysis. **Xiang Li:** Investigation. **Li Zhang:** Investigation. **Bruno Rente:** Investigation. **Tong Sun:** Investigation, Writing – review & editing. **Kenneth T.V. Grattan:** Investigation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is partially supported by 'the Fundamental Research Funds for the Central Universities, China (2022MS015)' to Y.H. LI and by SP Energy Networks on the project 'A holistic approach for power system monitoring to support DSO transition' and EPSRC, United Kingdom funded project on 'Creating Resilient Sustainable Microgrids through Hybrid Renewable Energy Systems' under grant EP/R030243/1.

References

[1] Fankhauser S, Smith SM, Allen M, Axelsson K, Hale T, Hepburn C, Kendall JM, Khosla R, Lezaun J, Mitchell-Larson E, et al. The meaning of net zero and how to get it right. Nature Clim Change 2021;1–7.
[2] Gandoman FH, Jaguemont J, Goutam S, Gopalakrishnan R, Firouz Y, Kalogiannis T, Omar N, Van Mierlo J. Concept of reliability and safety assessment of lithium-ion batteries in electric vehicles: Basics, progress, and challenges. Appl Energy 2019;251:113343.
[3] Liu B, Tang X, Gao F. Joint estimation of battery state-of-charge and state-of-health based on a simplified pseudo-two-dimensional model. Electrochim Acta 2020;344:136098.
[4] Liu K, Li K, Peng Q, Zhang C. A brief review on key technologies in the battery management system of electric vehicles. Front Mech Eng 2018;1–18.
[5] Zhang C, Li K, Mcloone S, Yang Z. Battery modelling methods for electric vehicles-A review. In: 2014 European control conference (ECC). IEEE; 2014, p. 2673–8.
[6] Shrivastava P, Soon TK, Idris MYIB, Mekhilef S. Overview of model-based online state-of-charge estimation using Kalman filter family for lithium-ion batteries. Renew Sustain Energy Rev 2019;113:109233.
[7] Tang X, Liu K, Li K, Widanage WD, Kendrick E, Gao F. Recovering large-scale battery aging dataset with machine learning. Patterns 2021;2:100302.
[8] Chemali E, Kollmeyer PJ, Preindl M, Emadi A. State-of-charge estimation of Li-ion batteries using deep neural networks: A machine learning approach. J Power Sources 2018;400:242–55.
[9] Chemali E, Kollmeyer PJ, Preindl M, Ahmed R, Emadi A. Long short-term memory networks for accurate state-of-charge estimation of Li-ion batteries. IEEE Trans Ind Electron 2017;65(8):6730–9.
[10] Sahinoglu GO, Pajovic M, Sahinoglu Z, Wang Y, Orlik PV, Wada T. Battery state-of-charge estimation based on regular/recurrent Gaussian process regression. IEEE Trans Ind Electron 2017;65(5):4311–21.
[11] How DN, Hannan MA, Lipu MSH, Sahari KS, Ker PJ, Muttaqi KM. State-of-charge estimation of Li-ion battery in electric vehicles: A deep neural network approach. IEEE Trans Ind Appl 2020;56(5):5565–74.
[12] Yang F, Li W, Li C, Miao Q. State-of-charge estimation of lithium-ion batteries based on gated recurrent neural network. Energy 2019;175:66–75.
[13] Xiao F, Li C, Fan Y, Yang G, Tang X. State of charge estimation for lithium-ion battery based on Gaussian process regression with deep recurrent kernel. Int J Electr Power Energy Syst 2021;124:106369.
[14] Feng X, Weng C, He X, Han X, Lu L, Ren D, Ouyang M. Online state-of-health estimation for li-ion battery using partial charging segment based on support vector machine. IEEE Trans Veh Technol 2019;68(9):8583–92.
[15] Guo P, Cheng Z, Yang L. A data-driven remaining capacity estimation approach for lithium-ion batteries based on charging health feature extraction. J Power Sources 2019;412:442–50.
[16] You G-W, Park S, Oh D. Diagnosis of electric vehicle batteries using recurrent neural networks. IEEE Trans Ind Electron 2017;64(6):4885–93.
[17] Li Y, Li K, Liu X, Zhang L. Fast battery capacity estimation using convolutional neural networks. Trans Inst Meas Control 2020;0142331220966425.
[18] Shen S, Sadoughi M, Li M, Wang Z, Hu C. Deep convolutional neural networks with ensemble learning and transfer learning for capacity estimation of lithium-ion batteries. Appl Energy 2020;260:114296.
[19] Li Y, Li K, Liu X, Wang Y, Zhang L. Lithium-ion battery capacity estimation — A pruned convolutional neural network approach assisted with transfer learning. Appl Energy 2021;285:116410.
[20] Sommer LW, Kiesel P, Ganguli A, Lochbaum A, Saha B, Schwartz J, Bae C-J, Alamgir M, Raghavan A. Fast and slow ion diffusion processes in lithium ion pouch cells during cycling observed with fiber optic strain sensors. J Power Sources 2015;296:46–52.
[21] Zhang C, Li K, Deng J. Real-time estimation of battery internal temperature based on a simplified thermoelectric model. J Power Sources 2016;302:146–54.
[22] Grattan KTV, Meggitt BT, editors. Optical fiber sensor technology. London, UK: Chapman & Hall; 1995.
[23] Raghavan A, Kiesel P, Sommer LW, Schwartz J, Lochbaum A, Hegyi A, Schuh A, Arakaki K, Saha B, Ganguli A, et al. Embedded fiber-optic sensing for accurate internal monitoring of cell state in advanced battery management systems part 1: Cell embedding method and performance. J Power Sources 2017;341:466–73.
[24] Wang Y, Gong J, Wang DY, Dong B, Bi W, Wang A. A quasi-distributed sensing network with time-division-multiplexed fiber bragg gratings. IEEE Photonics Technol Lett 2011;23(2):70–2.

[25] Su YD, Preger Y, Burroughs H, Sun C, Ohodnicki PR. Fiber optic sensing technologies for battery management systems and energy storage applications. Sensors 2021;21(4):1397.
[26] Nascimento M, Novais S, Ding MS, Ferreira MS, Koch S, Passerini S, Pinto JL. Internal strain and temperature discrimination with optical fiber hybrid sensors in Li-ion batteries. J Power Sources 2019;410:1–9.
[27] Sommer LW, Raghavan A, Kiesel P, Saha B, Schwartz J, Lochbaum A, Ganguli A, Bae C-J, Alamgir M. Monitoring of intercalation stages in lithium-ion cells over charge-discharge cycles with fiber optic sensors. J Electrochem Soc 2015;162(14):A2664.
[28] Nascimento M, Ferreira MS, Pinto JL. Real time thermal monitoring of lithium batteries with fiber sensors and thermocouples: A comparative study. Measurement 2017;111:260–3.
[29] Peng J, Jia S, Jin Y, Xu S, Xu Z. Design and investigation of a sensitivity-enhanced fiber Bragg grating sensor for micro-strain measurement. Sensors Actuators A 2019;285:437–47.
[30] Peng J, Zhou X, Jia S, Jin Y, Xu S, Chen J. High precision strain monitoring for lithium ion batteries based on fiber Bragg grating sensors. J Power Sources 2019;433:226692.
[31] Ganguli A, Saha B, Raghavan A, Kiesel P, Arakaki K, Schuh A, Schwartz J, Hegyi A, Sommer LW, Lochbaum A, et al. Embedded fiber-optic sensing for accurate internal monitoring of cell state in advanced battery management systems part 2: Internal cell signals and utility for state estimation. J Power Sources 2017;341:474–82.
[32] Rente B, Fabian M, Vidakovic M, Liu X, Li X, Li K, Sun T, Grattan KT. Lithium-Ion battery state-of-charge estimator based on FBG-based strain sensor and employing machine learning. IEEE Sens J 2020;21(2):1453–60.
[33] Schulz E, Speekenbrink M, Krause A. A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. J Math Psych 2018;85:1–16.
[34] Ozcan G, Pajovic M, Sahinoglu Z, Wang Y, Orlik PV, Wada T. Online battery state-of-charge estimation based on sparse Gaussian process regression. In: 2016 IEEE power and energy society general meeting (PESGM). IEEE; 2016, p. 1–5.