Article:
Liu, S., Li, K. and Yu, J. (2026) Adaptive estimation of battery pack state of charge with optical fibre strain measurements. Applied Energy, 407. 127330. ISSN: 0306-2619

https://doi.org/10.1016/j.apenergy.2025.127330

Highlights

Adaptive Estimation of Battery Pack State of Charge with Optical Fibre Strain Measurements

Shiyun Liu,Kang Li,James Yu

- A FBG-assisted battery pack SOC estimation method is proposed to address the challenges due to cell heterogeneity, wiring burden, and computational cost.
- Distributed fibre-optic sensors provide cell-by-cell strain monitoring.
- Strain–charge sensitivity analysis reveals diagnostic peak patterns correlated with ageing and degradation.
- Adaptive GPR-UKF framework dynamically adjusts parameters for robust state estimation.
- 94% RMSE reduction are achieved compared to conventional methods.
- Effectiveness is validated across multiple cell chemistries and operating conditions.

Adaptive Estimation of Battery Pack State of Charge with Optical Fibre Strain Measurements

Shiyun Liu[a], Kang Li[a,*] and James Yu[b]

[a]School of Electronic and Electrical Engineering, University of Leeds, Leeds, LS2 9JT, UK
[b]SP Energy Networks, Glasgow, G2 5AD, UK

ARTICLE INFO

Keywords:
Battery pack
SOC estimation
Gaussian Process
Unscented Kalman Filter
Optical Fibre Strain Measurements

ABSTRACT

Battery packs are critical components in electric vehicles and energy storage systems, yet reliable pack state-of-charge (SOC) estimation remains challenging due to cell-to-cell heterogeneity, dense sensor wiring, and computation that scales with pack size when monitoring all individual cells. This study introduces an optical fibre sensing approach that replaces conventional multi-sensor networks with a more compact set of sensors capable of monitoring entire cell modules while maintaining only pack-level voltage measurements. This allows the development of an innovative strain-charge sensitivity (SCS) analysis methodology that identifies representative cells by capturing subtle mechanical behaviour changes during charging cycles. More specifically, the SCS analysis reveals distinctive peak patterns that correlate with battery ageing states, providing an accurate diagnostic indicator for cell degradation assessment. When the SCS analysis is assisted with a Gaussian Process Regression-based adaptive Unscented Kalman Filter, more accurate and robust battery pack SOC estimation can be achieved under variable operating conditions. Experimental validation using a battery pack composed of two NCR18650 cylindrical cells demonstrates exceptional performance, achieving SOC estimation with 1.28% Mean Absolute Percentage Error under dynamic conditions, significantly outperforming conventional methods. Under static discharging conditions, the adaptive model maintains Root Mean Square Error of 0.46, representing improvements exceeding 67% compared to existing approaches. Additional validation using a $LiFePO_4$ pouch cell pack confirms the effectiveness of the proposed method for different cell technologies, achieving 77.27% improvement in compared to existing methods. This integrated sensing–modelling framework represents a significant advancement in battery-pack state estimation for large-scale applications.

*Corresponding author
k.li1@leeds.ac.uk (K. Li)
ORCID(s): 0000-0001-6657-0522 (K. Li)

Nomenclature

Acronyms

BMS	Battery Management System
BiLSTM	Bidirectional Long Short-Term Memory
CC-CV	Constant Current-Constant Voltage
CDM	Cell Difference Model
CMM	Cell Mean Model
CNN	Convolutional Neural Network
DST	Dynamic Stress Test
DNN	Deep Neural Network
ECM	Equivalent Circuit Model
SEI	Solid-Electrolyte Interphase
EKF	Extended Kalman Filter
ESSs	Energy Storage Systems
EVs	Electric Vehicles
FBG	Fibre Bragg Grating
GPR	Gaussian Process Regression
GRU	Gated Recurrent Units
HPPC	Hybrid Pulse Power Characterization
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MDM	Multiplayer Difference Model
MSE	Mean Squared Error
NIS	Normalized Innovation Squared
OCV	Open Circuit Voltage
OFS	Optical Fibre Sensor
RMSE	Root Mean Square Error
SCS	Strain-Charge sensitivity
SE	Squared Exponential Kernel
TCN	Temporal Convolutional Network
Lin	Linear Kernel
SGW	Savitzky-Golay Filter
SOC	State of Charge
SOH	State of Health
UKF	Unscented Kalman Filter

Symbols

c	Scaling factor for measurement noise covariance adjustment.
c_j	SGW coefficients.
D	Dimensionality of the state vector.
d	The d-th state.
e_k	Residual.
e_ε	Strain measurement error.
e_λ	Wavelength measurement error.
e_V	Voltage measurement error.
E_i	Normalized strain for the i-th cell.
$f(\cdot)$	Latent function mapping inputs to outputs.
$g(\cdot)$	State transition function.
$h(\cdot)$	Measurement function.
i, j	Sample or element indices.
I	Battery pack current (A).
K_T	Thermal response factor.
l	Length-scale parameter for kernel function.
m	Index representing the input dimension.
$\mathcal{L}$	Dataset of N input-output pairs $(\mathbf{x}_i, \mathbf{y}_i)$.
N	Number of samples in the training dataset.
q_{seg}	Charge segments.
Q	Charge (Ah).
Q_a	Available capacity.
Q_n	Total capacity.
(s_{seg})	Strain change segments.
S	Strain sensitivity.
$S_{k,oo}$	Diagonal element of the innovation covariance matrix for the o-th observation.
T	Battery pack average temperature (°C).
$\Delta\lambda$	Wavelength change in the FBG sensor.
ΔT	Change in temperature.
$\mathbf{I}$	Identity matrix with dimension $N \times N$.
$\mathbf{K}_*$	Covariance matrix.
$\mathbf{K}_k$	Kalman gain at time k.
$\mathbf{K}_o$	Covariance matrix for observation model.
$\mathbf{K}_s$	Covariance matrix for the state estimation.
$\mathbf{P}_k$	Covariance matrix at time k.
$\mathbf{Q}$	Process noise covariance matrix.
R^2	Coefficient of Determination.
$\mathbf{R}$	Observation noise covariance matrix.
$\mathbf{R}'_k$	Adjusted observation noise covariance matrix.
$\mathbf{r}_k$	Residual between actual and predicted states.
$\mathbf{S}_k$	Innovation covariance matrix at time k.
$\mathbf{U}_k$	Control input at time k.
$\mathbf{v}_k$	Additive measurement noise at time k.
$\mathbf{w}_k$	Additive process noise at time k.
V	Battery pack terminal voltage (V).
$\mathbf{X}_k$	State vector at time k.
$\mathbf{Y}_k$	Observation vector at time k.
$\mathbf{x}$	Input vector.
$\mathbf{y}$	Output vector.
$W_c^{(d)}$	Weight for covariance calculation.
$W_m^{(d)}$	Weight for mean calculation.

Greek Symbols

α	Spread of sigma points around the mean.
β	Scaling factor for Gaussian distributions.
δ_o	NIS for the o-th observation.
$\Delta\varepsilon$	Change in strain.
ϵ	Noise or error.
ε	Mechanical strain.
η	Coulombic efficiency.
γ	Threshold from chi-square distribution (1 degree of freedom).
$\kappa(\cdot,\cdot)$	Covariance function (kernel).

λ	Wavelength of the FBG sensor.
Λ	Scaling parameter for sigma points.
ρ	Order of polynomial.
σ_{ϵ}	Standard deviation of strain values.
σ_f^2	Signal variance in the kernel function.
σ_{Lin}	Scaling factor for the linear kernel.
σ_{SE}	Scaling factor for the SE kernel.
σ_s^2	Variance of process noise in the state model.
σ_q^2	Variance of the measurement noise.
σ_{SOC}^2	Variance of residuals for SOC.
σ_T^2	Variance of residuals for temperature.
$\sigma_{\mathbf{X}}^2$	Variance of residuals between actual and predicted states.
Θ	Set of optimised hyperparameters.
ω	Half-window size.
ζ	Smoothed value.
$\mathcal{X}_{k\|k-1}^d$	Predicted sigma point.
$\mathcal{Y}_{k\|k-1}^d$	Predicted measurement sigma point.

1. Introduction

Battery pack state-of-charge (SOC) estimation is an essential function of the battery management system (BMS) to optimise performance, maintain cell balancing, and ensure safe and effective charging and discharging operations [1, 2, 3]. However, achieving reliable real-time pack-level SOC estimation remains challenging owing to cell-to-cell heterogeneity, the wiring burden associated with dense sensor deployments, and computation that scales with pack size [4]. While current research has largely focused on state estimation for single cells, these pack-level difficulties have seen limited progress [5, 6].

These challenges also bear directly on the definition of pack SOC, which may be taken as the average across all cells or as boundary values under different operating conditions [7]. Nevertheless, the average often fails to reflect the behaviour of limiting cells, so the cell with the maximum SOC during charging or the minimum SOC during discharging is typically used to represent pack SOC [8]. A more comprehensive approach involves monitoring and estimating the SOC of individual cells [4], which enables detection of cell-level imbalances, improves energy efficiency, and mitigates risks to pack reliability and lifespan.

A common approach for estimating the SOC of a battery pack at the individual cell level involves using both the Cell Mean Model (CMM) to represent overall performance, and the Cell Difference Model (CDM), typically based on simplified equivalent circuit models, to evaluate discrepancies in SOC or internal resistance among individual cells [9, 10, 11, 12, 13, 14, 15]. To further enhance computational efficiency and estimation accuracy, a dual time-scale Extended Kalman Filter (EKF) is applied for sequential state estimation of both models [11, 12, 13]. This approach necessitates the collection of various measurement data, including the terminal voltages of the pack and each cell, as well as the load current.

Furthermore, Jiang et al. [16] proposed a Multiplayer Difference Model (MDM) estimator that operates across multiple timescales to achieve efficient and accurate SOC and capacity estimation for a battery pack. Similarly, Peng et al. [17] developed an Equivalent Circuit Model (ECM) of parallel-connected battery packs, incorporating multiple fuzzy-PI sub-regulators to compensate for variations in resistance and capacity. However, creating a model for each cell remains computationally intensive and requires numerous parameters [18]. In contrast, considering the whole battery pack as a single big cell is much simpler and practically applicable, yet, how to develop and effectively use these simpler models like the ECM is challenging given the inhomogeneity and inconsistence of battery cells within a battery pack during both production and usage [19, 20, 21, 22, 23, 24].

An effective alternative is selecting representative cells to infer the SOC of the entire pack, reducing computational costs [25]. This approach involves two primary steps [4]. The first step is identifying one or several representative cells within the pack [26, 27]. The second step involves applying validated cell-level models to estimate the SOC of these representative cells [28]. The SOC values of these selected representative cells are then extrapolated to estimate the pack SOC. However, the effectiveness of this approach depends on several factors, such as the accuracy of the sensors, the reliability of the observed signals, the robustness of the methods utilised to select the representative cells, and the precision of the models employed to estimate the SOC of the pack [4].

The terminal voltage of individual cells is commonly used to select representative cells within a battery pack, subsequently estimating their SOC using the widely adopted ECM [29, 30]. This strategy presupposes that cells displaying extreme terminal voltages are indicative of extreme SOC values. However, this assumption is not invariably reliable. For instance, a cell exhibiting a high terminal voltage might have greater impedance or could be influenced by voltage measurement noise [31]. Furthermore, significant variations in impedance can lead to elevated terminal voltages, thus complicating the accuracy of SOC estimation [31]

To address the challenges, Wang et al. [31] propose a "M-top/bottom" voltage-based approach that selects a subset of $2M$ representative cells to specifically estimate the maximum and minimum SOC at any given time, which demonstrates the potential in identifying cells with extreme SOC values. Similarly, Chen et al. [32] leveraged the voltage measurements taken at the end of constant current charging phase to determine representative cells exhibiting the highest and lowest voltages, effectively mitigating estimation errors caused by frequent changes in representative cell selection. Building on the direct correspondence between the quasi-Open Circuit Voltage (OCV) and SOC distribution in a battery pack, Chun et al. [33] introduced a method that correlates a battery cell's OCV with its SOC by primarily focusing on the battery cell that has the least remaining capacity. This method reconstructed the OCV operating range to account for factors like ageing which increases the battery's impedance and reduces its capacity. Using

the lowest and highest OCV values of representative cells, a calibration factor was derived to adjust the relationship between the cell's SOC and the overall battery pack's SOC.

In reality, the empirical relationship between OCV and SOC is pivotal for the model-based SOC estimation methods, directly influencing both the performance of the battery model and the accuracy of the estimation results. For instance, the OCV characteristic of a cell has been demonstrated to depend on cell temperature [34]. Accurately tracking the rapidly varying impedance characteristics and precisely modelling the temperature-dependent OCV characteristics can enhance the performance of the battery pack model [30]. In [35], the representative cell of the series branch was selected based on the highest cosine similarity value for internal resistance inconsistencies, thereby facilitating the re-evaluation of the OCV-SOC-Temperature relationships and the SOC estimation for the series-parallel pack.

An alternative method for selecting representative cells in series-connected battery packs involves evaluating individual cell capacities. Due to its significant impact on pack capacity utilization during charging and discharging, the cell with the minimum remaining capacity is considered representative of the entire pack [33, 36]. However, this method faces the challenge of accurately identifying the weakest cell with the lowest capacity before estimating its SOC using the empirical models. To overcome this challenge, Liu et al. [37] proposed a statistics-based approach that uses terminal voltage as a proxy for capacity and employs long-term statistical analysis to identify representative cells based on their historical frequency of exhibiting minimum/maximum voltages. Their improved Vmin + Vmax model method calculates pack SOC through weighted fusion of multiple representative cells rather than relying on a single weakest cell, thereby improving estimation stability and accuracy in real-world applications.

Data-driven methods for estimating battery pack SOC have recently attracted considerable attention for their ability to model input-output relationships using signals from either representative cells or the entire pack's current, voltage, and temperature [7]. For instance, Tang et al. [38] proposed a "leader-follower" approach, where a "leader" cell is selected based on factors like internal resistance and voltage, while "follower" cells adopt the leader's SOC when active balancing hardware is engaged. Manoharan et al. [24] developed a Parallel Artificial Neural Network (PANN) architecture with Bidirectional Long-Short Term Memory (BiLSTM) layers to estimate SOC in parallel cell-connected packs. Unlike the representative cell selection in [38], their model uses pack-level voltage, current, and temperature as inputs under normal conditions, switching to individual cell parameters when cells approach cut-off voltage during charging and discharging. Furthermore, Sun et al. [39] took a more comprehensive approach to address cell inconsistencies by proposing a CNN-BiLSTM-Attention model that incorporates cell inconsistency evolution through three-parameter Weibull distribution modelling, achieving SOC estimation errors below 3% using real-world electric vehicle operational data. Their approach demonstrates how statistical distribution features can be used to significantly improve the estimation accuracy, particularly at lower SOC levels.

Beyond considering the temperature of a representative cell during model development, the temperature distribution across the entire battery pack is crucial in estimating SOC. While higher average temperatures lead to more rapid declines in overall pack capacity and SOC [40], temperature gradients within the pack further exacerbate SOC imbalance and capacity loss, with the cell at the highest temperature ultimately limiting the pack's performance [41, 42]. Thus, for accurate SOC estimation, focusing solely on an individual cell's temperature is insufficient; instead, it is essential to account for factors contributing to SOC imbalance, such as the temperature gradient across the pack.

However, for energy storage systems that contain hundreds or thousands of smaller-sized battery cells, it is impractical to employ conventional temperature or voltage sensors to monitor the characteristics of each individual cell [43, 44]. Further, electrical-based signals are prone to challenges such as signal interference and degradation in harsh operational environments [45]. Optical fibre technology provides an advanced alternative for monitoring the temperature and mechanical properties of batteries, significantly improving the precision of real-time evaluations of battery states [46, 47, 48, 49]. This method capitalizes on the natural mechanical behaviour of lithium-ion batteries, which swell while charging and shrink while discharging [50, 51]. The resulting volume changes, combined with fluctuations in both internal and environmental temperatures, contribute to separator creep [52, 53]. This phenomenon, in turn, accounts for overall strain variations observed on the battery cell's shell [54]. Furthermore, as the battery ages through regular use, certain degradation processes, such as the development of a Solid-Electrolyte Interphase (SEI) layer, induce permanent strain changes on the battery's exterior [52]. This comprehensive information embedded in the strain measurements acquired from the optical fibre sensors (OFSs) has been demonstrated to be particularly useful for effective non-invasive monitoring.

Furthermore, when mounted on battery surfaces, these sensors provide real-time temperature and strain measurements. By combining the OFS sensing with machine learning methods, more accurate SOC estimation can be achieved using non-electrical measurements. For instance, Ee et al. [55] demonstrated that Fibre Bragg Grating (FBG) sensors coupled with Deep Neural Networks (DNNs) deliver comparable performance to conventional electrical signal based approaches for SOC estimation. Peng et al. [56] further confirmed that non-electrical signal based approaches alone can achieve reliable SOC estimation, while FBG measurements serve as auxiliary parameters to enhance estimation accuracy. Li et al. [47] affixed FBG sensors to cell surfaces to capture multi-point strain and temperature variations resulting from charging and discharging cycles. These measurements are then used in a hybrid machine learning framework that

combines convolutional neural networks (CNN) with Gaussian process regression (GPR) methods for state estimation. This approach improves joint estimation accuracy by automatically extracting features from the enriched FBG measurements. Xia et al. [57] achieved remarkable SOC estimation accuracy of up to 99.94% by utilizing DNN models with input signals covering the strain, temperature, voltage, and current measurements. Their work demonstrates that SOC estimation remains feasible using only strain and temperature data obtained from OFSs, which eliminates the need for voltage and current measurements. Sheng et al. [58] investigated multi-position strain measurements to improve the precision of SOC estimation. Their developed deep learning models, including temporal convolutional networks (TCN), CNN, and gated recurrent units (GRU), showed significant improvements in estimation accuracy when using dual strain combinations compared to single strain measurements. Liu et al. [54] also demonstrated that using FBG sensors in a battery pack allows the final pack SOC to be estimated using a strain-based data-driven method in conjunction with a voltage-based equivalent circuit model (ECM). This joint model offers the advantage of reducing the impact caused by corrupted or failed electrical or fibre optical signals, resulting in more accurate and reliable estimation. However, these studies primarily focus on single battery cell, when dealing with the SOC estimation of battery pack comprising many cells, these methods are becoming computationally very intensive if the SOC for each single cell needs to be estimated. This paper addresses this challenge.

This study proposes a novel method combining strain-charge-sensitivity (SCS) analysis and GPR-based adaptive UKF for accurate battery pack SOC estimation through a representative cell approach. Considering the limitations of conventional monitoring systems, this method opens up the opportunities of significantly reducing the number of voltage and temperature sensors in conventional approaches, assisted with the OFSs, thereby significantly reduce the wiring space while enables more accurate real-time state monitoring and estimation of battery cells in a battery pack. Building upon the new OFS assisted sensing approach, the SCS analysis utilises strain and charge signals during the charging of all series-connected cells and identify characteristic peaks in the derivative of strain with respect to charge. By comparing these peak patterns in the SCS profiles, representative cells whose mechanical signatures most accurately reflect the pack's electrochemical behaviour are selected, therefore enables the establishment of a new diagnostic framework for battery-pack management. To estimate SOC profiles, a GPR-based adaptive UKF is developed that integrates GPR state estimation and observation models to dynamically update process and measurement noise covariances in real time. The adaptive framework modulates the Kalman gain, continuously adjusting the relative weighting of sensor measurements and model estimations to account for fluctuations in sensor reliability while preserving estimation accuracy. The method further incorporates temperature effects on SOC behaviour within the model framework. The adaptive UKF leverages GPR-learned thermal-electrochemical relationships for state estimation, with covariance matrices capturing both individual state uncertainties and their interdependencies. Comparative analysis against GPR-UKF with fixed parameters and conventional machine learning approaches demonstrates the superior performance of the proposed new method combining SCS and GPR-based adaptive UKF. The framework proves particularly effective when measurements exhibit fluctuations under dynamic operating conditions, enabling prioritization of reliable information sources for high-fidelity battery state monitoring.

The structure of this paper is as follows. Section 2 provides an overview of the fundamental concepts of GPR and the UKF. Section 3 outlines the methodology, including the framework for the representative cell selection and battery pack state estimation. Section 4 details the experimental setup, followed by data pre-processing procedures in Section 5. Section 6 focuses on performance evaluation metrics and comparative analysis. The results are presented in Section 7, and the paper concludes with a summary of findings in Section 8.

2. Methodology

In series-connected battery configuration, SOC imbalance poses a significant challenge to the battery pack operation and management, complicating the estimation of maximum SOC at the end of charging cycles and minimum SOC during discharge cycles. This imbalance can lead to overcharging or over-discharging of the batteries. Consequently, by limiting the operation of the battery pack with respect to the cell(s) with the lowest available capacity, these safety concerns can be effectively addressed. To efficiently identify representative cells, a method based on strain measurements from OFSs will be introduced in this section.

In order to address the inherent cell-to-cell inconsistency in the pack and to improve the model's capability in representing nonlinear dynamics under uncertainty, prediction and observation models based on GPR are developed and subsequently incorporated into an UKF for SOC and temperature estimation. GPR is a probabilistic machine learning method that provides estimation in the form of a full probability distribution, including mean values and confidence intervals [59]. This non-parametric approach does not require a specific parametric form for the data, allowing for flexibility and adaptability in modelling [60]. By extending multivariate Gaussian distributions to infinite dimensions, GPR can effectively model complex, nonlinear functions, making it particularly useful in situations where the underlying processes are challenging to define or are unknown [47, 61]. UKF estimates the state of a dynamic system by processing sequences of control inputs and observations [62]. Unlike traditional Kalman filters, which rely on linear approximations of state and measurement functions, the UKF employs a deterministic sampling technique known as the sigma-point approach to handle nonlinearity more effectively. This technique involves generating a set of

carefully chosen sample points, known as sigma points (denoted as $\mathcal{X}$), around the mean of the input distribution [63]. These points are selected so that their (possibly weighted) mean and covariance accurately reflect the true mean and covariance of the distribution.

The models and parameters for this GPR-UKF approach are derived from training data using GPR-based non-parametric regression techniques. The UKF employs state-dependent uncertainty estimates, calculated by assessing both measurement noise and the uncertainty associated with the density and quality of the training data. Consequently, the filter automatically increases its uncertainty when it encounters regions with insufficient training data [62].

2.1. Ageing cell selection based on strain-charge sensitivity (SCS) analysis

During operation, battery shells exhibit two types of strain variations. The first type is reversible, as shown in Figure 1. This variation is caused by temperature fluctuations, phase transitions of different electrode materials, and the activities of lithium ions during the charging and discharging processes [51]. In a continuous charge-discharge cycle, with stable environmental conditions, the battery exhibits maximum strain values around the end of charging and minimum strain values at the end of discharging [54, 51]. On the other hand, over the long term, an irreversible strain variation emerges, primarily due to permanent changes in the battery's internal structure, such as the growth of the SEI [52].

Observing mechanical changes over time can improve the evaluation of battery cells' electrochemical performance, although the subtle and minor changes often presents challenges in extracting any meaningful and useful information from real-time strain signals [64, 65, 66]. Capturing the charge variation trend on the dynamic strain signals solves this problem and allows for more accurate estimation of battery states, which are affected by the underlying trends of battery cell service capability. This method is particularly beneficial in detecting subtle charge changes which is especially useful in operating battery packs with cells connected in series, as it is vital to ensure that no battery cell undergoes excessive swelling or thermal stress during charging.

The $\frac{d\varepsilon}{dQ}$ method, also named as strain-charge sensitivity (SCS) analysis, quantifies how the strain in a battery varies with charge accumulation. A higher value of this metric indicates a reduction in the usable capacity of the battery, likely due to irreversible structural changes, such as significant solid material deposition on the electrodes [67]. Such solid deposits hinder ion transport between the electrolyte and electrodes, thereby increasing internal resistance, which in turn leads to greater heat generation during charging and discharging [68]. Additionally, these irreversible structural changes can cause observable variations in surface strain with minimal changes in battery charge [66]. These electrochemical-mechanical relationships often remain unchanged for all cells produced from the same manufacturing batch, allowing the identification of representative cells purely based on the SCS analysis without requiring state

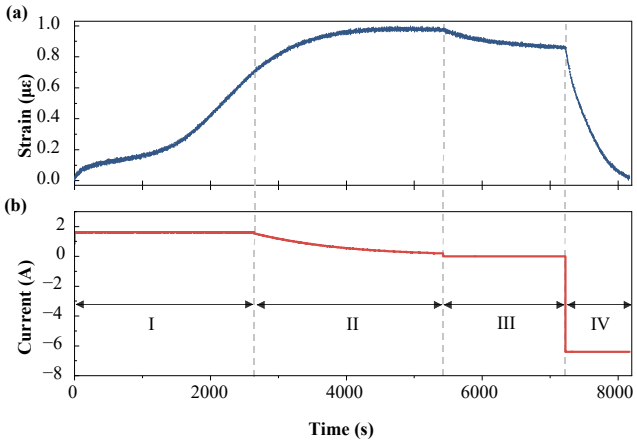

Figure 1: Operational strain variations and current profile over time: (a) Strain; (b) Current, with phases (I) Constant Current Charge, (II) Constant Voltage Charge, (III) Rest, (IV) Constant Current Discharge.

estimation for every single cell in the pack. The reproducible strain patterns provide a solid foundation for identifying cells most accurately representing the pack's overall electrochemical behaviour.

2.2. State prediction and observation models based on GPR framework

In this study, state prediction and observation models for the battery pack are developed using GPR, which are subsequently integrated into the UKF framework. The GPR models predict state transitions and observations of the system while dynamically updating the noise covariances in the UKF.

Considering that the battery pack's average temperature can influence SOC estimation [40], thermal interactions among adjacent cells in the pack's design (to be introduced in Section 3) may cause temperature measurement interference. Employing robust modelling and monitoring strategies is crucial for maintaining reliable SOC tracking.

Therefore, the hidden states of the battery pack, namely the SOC and the average temperature (T), are represented by the two-dimensional vector $\mathbf{X}_k = \left[\mathrm{SOC}_k, T_k\right]$ at each time step k. The control input $\mathbf{U}_k$ corresponds to the battery pack current I_k (positive for charging, negative for discharging). Observations include the terminal voltage (V_k) of the pack and strain (ε_k) from a representative cell, forming the observation vector $\mathbf{Y}_k = \left[\varepsilon_k, V_k\right]$.

2.2.1. GPR-based state prediction model for battery dynamics

In the GPR-based state prediction model, the output $\mathbf{y}_s = \mathbf{X}_k$ of a GPR model indicates the present SOC and the pack average temperature, while the GPR inputs are the previous time step battery states as well as the control input, $\mathbf{x} = \left[\mathbf{X}_{k-1}, \mathbf{U}_{k-1}\right]$. The process noise $\mathbf{w}_{s,k}$ is incorporated into the model, thus the GPR-based prediction model can be

expressed as:

$$\mathbf{y}_s = g(\mathbf{x}) + \mathbf{w}_{s,k} \tag{1}$$

where the function $g(\cdot)$ follows a normal distribution $g(\mathbf{x}) \sim \mathcal{GP}\left(m_s(\mathbf{x}), \mathbf{K}_s\left(\mathbf{x}, \mathbf{x}'\right)\right)$. Here, $m_s(\mathbf{x})$ represents the mean function, and $\mathbf{K}_s$ is the covariance function. The additive process noise $\mathbf{w}_{s,k}$ follows a normal distribution $\mathbf{w}_{s,k} \sim \mathcal{N}(0, \sigma_s^2)$, with σ_s^2 representing the process noise variance.

In the context of battery usage, the current typically exhibits linear or nonlinear variation patterns depending on the charge-discharge protocols [69, 70, 71]. These variations in current also significantly influence the temperature, which often demonstrates strong nonlinear characteristics, especially under high current conditions [72, 73]. Furthermore, the SOC is directly influenced by the charging/discharging current, and significant temperature fluctuations can impact the accuracy of SOC estimation due to the non-linear nature of the electrochemical processes [34, 72]. To accurately model these complex interactions and interdependencies between the battery states and inputs, an appropriate kernel function $\mathbf{K}_s(\mathbf{x}, \mathbf{x}')$ is required.

Specifically, a squared exponential (SE) kernel and a linear (Lin) kernel are employed [61, 74] to define the kernel function $\mathbf{K}_s(\mathbf{x}, \mathbf{x}')$ for modelling $g(\cdot)$ as detailed below:

$$\kappa_{s,m}\left(\mathbf{x}_m, \mathbf{x}'_m\right) = \sigma^2_{\mathrm{SE},s,m} \exp\left(-\frac{\left(\mathbf{x}_m - \mathbf{x}'_m\right)^2}{2l^2_{s,m}}\right) + \sigma^2_{\mathrm{Lin},s,m}\left(\mathbf{x} \cdot \mathbf{x}' + \sigma^2_{0,s,m}\right) \tag{2}$$

$$\mathbf{K}_s\left(\mathbf{x}, \mathbf{x}'\right) = \left(\prod_{m \in \{\mathrm{SOC}, T, I\}} \kappa_{s,m}\left(\mathbf{x}_m, \mathbf{x}'_m\right)\right) \tag{3}$$

where $l_{s,m}$ is the length-scale parameter associated with each input dimension, which determines the sensitivity of the kernel to variations in the input. The constants $\sigma_{\mathrm{SE},s,m}$ and $\sigma_{\mathrm{Lin},s,m}$ serve as scaling factors for the SE kernel and the Lin kernel, respectively. Additionally, $\sigma^2_{0,s,m}$ is introduced as the bias parameter for the Lin kernel, allowing it to capture offsets or shifts in the linear relationships between inputs. This multiple-kernel approach (in Eq. (2)) enables the GPR model to capture both the linear and non-linear dependencies specific to each variable. The multiplicative combination of individual kernels (in Eq. (3)) for SOC, T, and I assumes that the interactions between these variables contribute in a multiplicative manner to the system's output. The set of hyperparameters that will be optimised in the training process for the prediction model can be expressed as:

$$\Theta_s = \left\{ l_{s,m}, \sigma^2_{\mathrm{SE},s,m}, \sigma^2_{\mathrm{Lin},s,m}, \sigma^2_{0,s,m}, \sigma^2_s \mid m \in \{\mathrm{SOC}, T, I\} \right\} \tag{4}$$

The model training is performed to calculate the hyperparameters of the prediction model to maximize the marginal likelihood, following the problem:

$$\max_{\Theta_s} p\left(\mathbf{y}_s \mid \mathbf{x}, \Theta_s\right) \tag{5}$$

Finally, using the trained GPR prediction model, the distribution of the mean output $\hat{\mathbf{y}}_{s*}$ for the new inputs $\mathbf{x}_* = \left[\mathbf{X}_{k-1,*}, \mathbf{U}_{k-1,*}\right]$ can be obtained as follows:

$$\hat{\mathbf{y}}_{s*} = \mathbf{K}_s\left(\mathbf{x}_*, \mathbf{x}\right)\left[\mathbf{K}_s(\mathbf{x}, \mathbf{x}) + \sigma_s^2 \mathbf{I}_s\right]^{-1} \mathbf{y}_s \tag{6}$$

where $\mathbf{I}_s$ denotes the $N \times N$ identity matrix, with N as the number of training samples. $\mathbf{K}_s(\mathbf{x}, \mathbf{x})$ denotes the covariance matrix computed between the training inputs in the state prediction model. The predicted states covariance matrix $\mathbf{K}_{s*}$, which quantifies the uncertainty in the prediction of the test output, is given by:

$$\mathbf{K}_{s*} = \mathbf{K}_s\left(\mathbf{x}_*, \mathbf{x}_*\right) - \mathbf{K}_s\left(\mathbf{x}_*, \mathbf{x}\right)\left[\mathbf{K}_s(\mathbf{x}, \mathbf{x}) + \sigma_s^2 \mathbf{I}_s\right]^{-1} \mathbf{K}_s\left(\mathbf{x}, \mathbf{x}_*\right) \tag{7}$$

2.2.2. GPR-based observation model for battery measurements

The measurement distribution of the general observation model depends only on states [63]. However, in the modelling of battery states and measurements, it is commonly believed that measurements are influenced by control inputs ($\mathbf{U}_k$) as well as the states ($\mathbf{X}_k$) [54, 75, 76].

Accordingly, for the GPR-based observation model, the selected inputs are consistent with those of the state prediction model $\mathbf{x} = \left[\mathbf{X}_{k-1}, \mathbf{U}_{k-1}\right]$, with measurement noise denoted as $\mathbf{v}_{o,k-1}$. The output at the corresponding timestep is $\mathbf{y}_o = \mathbf{Y}_{k-1}$. The model can therefore be expressed as follows:

$$\mathbf{y}_o = h\left(\mathbf{x}\right) + \mathbf{v}_{o,k-1} \tag{8}$$

where the function $h(\cdot)$ follows a normal distribution $h\left(\mathbf{x}\right) \sim \mathcal{GP}\left(m_o(\mathbf{x}), \mathbf{K}_o(\mathbf{x}, \mathbf{x}')\right)$, with mean $m_o(\mathbf{x})$ and covariance function $\mathbf{K}_o$. The additive measurement noise $\mathbf{v}_{o,k-1} \sim \mathcal{N}\left(0, \sigma_o^2\right)$ is normally distributed with variance σ_o^2.

The structure of the covariance function $\mathbf{K}_o(\mathbf{x}, \mathbf{x}')$, provided in Eqs. (9) and (10), uses a hybrid kernel framework incorporating both the SE and Lin kernels.

$$\kappa_{o,m}\left(\mathbf{x}_m, \mathbf{x}'_m\right) = \sigma^2_{\mathrm{SE},o,m} \exp\left(-\frac{\left(\mathbf{x}_m - \mathbf{x}'_m\right)^2}{2l^2_{o,m}}\right) + \sigma^2_{\mathrm{Lin},o,d}\left(\mathbf{x}_m \mathbf{x}'_m + \sigma^2_{0,o,m}\right) \tag{9}$$

$$\mathbf{K}_o\left(\mathbf{x}, \mathbf{x}'\right) = \left(\prod_{m \in \{\mathrm{SOC}, T, I\}} \kappa_{o,m}\left(\mathbf{x}_m, \mathbf{x}'_m\right)\right) \tag{10}$$

The optimised hyperparameters Θ_o for the observation model in the GPR include the length-scale parameter $l_{o,m}$

for each input dimension m, the scaling factors $\sigma_{SE,o,m}$ and $\sigma_{Lin,o,m}$ for the SE and Lin kernels, respectively, and the bias parameter $\sigma^2_{0,o,m}$ for the Lin kernel. Thus, the optimised set of hyperparameters is:

$$\Theta_o = \left\{ l_{o,m}, \sigma^2_{\mathrm{SE},o,m}, \sigma^2_{\mathrm{Lin},o,m}, \sigma^2_{0,o,m}, \sigma^2_o \mid m \in \{\mathrm{SOC}, T, I\} \right\} \tag{11}$$

The hyperparameter optimization problem can be defined as follows:

$$\max_{\Theta_o} p\left(\mathbf{y}_o \mid \mathbf{x}, \Theta_o\right) \tag{12}$$

Therefore, using the GPR observation model, the predictive distribution $\mathbf{y}_{o*}$ for the new input $\mathbf{x}_* = \left[\mathbf{X}_{k-1,*}, \mathbf{U}_{k-1,*}\right]$ is obtained using Eq. (8):

$$\hat{\mathbf{y}}_{o*} = \mathbf{K}_o\left(\mathbf{x}_*, \mathbf{x}\right)\left[\mathbf{K}_o(\mathbf{x}, \mathbf{x}) + \sigma_o^2 \mathbf{I}_o\right]^{-1} \mathbf{y}_o \tag{13}$$

Here, $\mathbf{I}_o$ is an $N \times N$ identity matrix, where N denotes the number of samples in the training set, and $\mathbf{K}_o(\mathbf{x}, \mathbf{x})$ is the covariance matrix obtained from the training inputs in the observation model. The covariance matrix $\mathbf{K}_{o*}$, which quantifies the uncertainty in the prediction of the test output, is given by:

$$\mathbf{K}_{o*} = \mathbf{K}_o\left(\mathbf{x}_*, \mathbf{x}_*\right) - \mathbf{K}_o\left(\mathbf{x}_*, \mathbf{x}\right)\left[\mathbf{K}_o(\mathbf{x}, \mathbf{x}) + \sigma_o^2 \mathbf{I}_o\right]^{-1} \mathbf{K}_o\left(\mathbf{x}, \mathbf{x}_*\right) \tag{14}$$

2.2.3. Integration of GPR-Based prediction and observation models into UKF

When integrating GPR-based state prediction and observation models into the UKF, the deterministic state transition function in the conventional UKF is replaced by the GPR-based state prediction model (Eq. (6)). Likewise, the observation function in the conventional UKF is substituted by the GPR-based observation model (Eq. (13)). This integration enables the GPR models to offer probabilistic representations of both state predictions and measurement uncertainties, thus enabling the UKF to more effectively capture the stochastic behaviour of the battery system dynamics.

The process noise covariance matrix, $\mathbf{Q}$, derives from the variance within the GPR-based state prediction model, which quantifies uncertainty due to process noise during state prediction, as shown in Eq. (3). Similarly, the measurement noise covariance matrix, $\mathbf{R}$, is obtained from the variance in the GPR-based observation model, as given in Eq. (10).

By dynamically updating $\mathbf{Q}_k$ and $\mathbf{R}_k$ based on the variances derived from GPR models, the UKF framework adapts responsively to changing system conditions. Specifically, the GPR-based models generate uncertainty estimates for both the process and the measurements, allowing the UKF to adjust confidence in the state transition and observation models in real-time. This capability ensures that the filter remains responsive to variations in battery behaviour, improving the accuracy of the state estimates.

Furthermore, the Kalman gain, $\mathbf{K}_k$, in the UKF is influenced not only by the prior observation covariance, $\mathbf{P}_k^{\mathbf{Y}^-}$, but also by the prior cross-covariance between the state and the observation [77]. When an observation has high uncertainty while exhibiting strong correlation with the state, simply relying on the GPR-based observation model updated $\mathbf{R}_k$ may not adequately reduce dependence on that observation during state updates. To address the potential influence of uncertain yet correlated observations, an adaptive scaling approach is applied, in which the Normalized Innovation Squared (NIS) δ_o for each observation component is compared to a chi-square threshold [78]. If δ_o exceeds the threshold, indicating an outlier observation, the corresponding elements in $\mathbf{R}_k$ are scaled by a factor c, thereby reducing their impact on the state update.

Table 1 summarises the procedure for estimating the pack's SOC and average temperature, and Figure 2 presents the integrated flowchart of the GPR–UKF adaptive battery-pack state-estimation algorithm.

3. The experimental platform and testing procedures

Due to the strong correlations of charging C-rates and charging protocols with cell shell strain variations during the cycling, the SCS analysis method is naturally a good approach to analyse the cycling characteristics of battery cells. To further establish reliable strain-ageing relationships and validate the proposed framework, concept-proving experiments were conducted on a simple pack unit using the widely adopted CC-CV charging protocol as the primary charging method [68, 71, 79, 80]. The pack working environment also simulate the pack widely working environment that is around 25 and 50% humidity. The proposed state estimation framework was comprehensively validated conducted under both static and dynamic conditions. This section introduces the experimental procedures, covering battery assembly, sensor installation, and testing protocols.

3.1. Battery cell screening and pack assembly

A simple battery pack was constructed using two Panasonic NCR18650B cylindrical lithium-ion cells connected in series, each with a nominal voltage of 3.6 V and a capacity of 3200 mAh. Table 2 provides detailed specifications of the battery pack and its constituent cells.

The first step in the experimental set-up is to select battery cells with closely matching capacity and internal resistance. To achieve this, fresh cells all first underwent a preconditioning process involving three cycles of charging and discharging adopting the constant current-constant voltage (CC-CV) protocol, with voltages ranging from 2.5V to 3.65V, a procedure used to facilitate the battery SEI formation [46]. During the CV charging phase, the current was reduced to 0.02C. A ten-minute relaxation period was introduced between two cycles. Subsequently, the cells underwent capacity testing at C/25 discharge rate and internal resistance assessments through the Hybrid Pulse Power Characterization (HPPC) test. The test data were then used to obtain the capacity and internal resistance of each cell.

Table 1
Summary of the GPR–UKF-based adaptive battery pack state estimation algorithm

GPR-Based Battery Prediction Model:

$$\hat{\mathbf{y}}_{s*} = \mathbf{K}_s\left(\mathbf{x}_*, \mathbf{x}\right)\left[\mathbf{K}_s(\mathbf{x}, \mathbf{x}) + \sigma_s^2\mathbf{I}_s\right]^{-1}\mathbf{y}_s$$

$$\mathbf{K}_{s*} = \mathbf{K}_s\left(\mathbf{x}_*, \mathbf{x}_*\right) - \mathbf{K}_s\left(\mathbf{x}_*, \mathbf{x}\right)\left[\mathbf{K}_s(\mathbf{x}, \mathbf{x}) + \sigma_s^2\mathbf{I}_s\right]^{-1}\mathbf{K}_s\left(\mathbf{x}, \mathbf{x}_*\right)$$

GPR-Based Battery Observation Model:

$$\hat{\mathbf{y}}_{o*} = \mathbf{K}_o\left(\mathbf{x}_*, \mathbf{x}\right)\left[\mathbf{K}_o(\mathbf{x}, \mathbf{x}) + \sigma_o^2\mathbf{I}_o\right]^{-1}\mathbf{y}_o$$

$$\mathbf{K}_{o*} = \mathbf{K}_o\left(\mathbf{x}_*, \mathbf{x}_*\right) - \mathbf{K}_o\left(\mathbf{x}_*, \mathbf{x}\right)\left[\mathbf{K}_o(\mathbf{x}, \mathbf{x}) + \sigma_o^2\mathbf{I}_o\right]^{-1}\mathbf{K}_o\left(\mathbf{x}, \mathbf{x}_*\right)$$

Initialization:

Initial state mean $\hat{\mathbf{X}}_0$, initial covariance $\mathbf{P}_0^{\mathbf{X}}$

State dimension D, scaling parameters α, β

Inputs:

Previous state mean $\hat{\mathbf{X}}_{k-1}$, previous state covariance $\mathbf{P}_{k-1}^{\mathbf{X}}$, previous system input $\mathbf{U}_{k-1}$, current observation $\mathbf{Y}_k$

$\triangleright\ \hat{\mathbf{X}}_{k-1} = \left[\mathrm{SOC}_{k-1}, T_{k-1}\right]$, $\mathbf{U}_{k-1} = I_{k-1}$, $\mathbf{Y}_k = \left[\epsilon_k, V_k\right]$.

Prediction Step:

Compute sigma points for prior state $\mathcal{X}_{k-1}^d$:

$$\mathcal{X}_{k-1}^0 = \hat{\mathbf{X}}_{k-1}$$

$$\mathcal{X}_{k-1}^d = \hat{\mathbf{X}}_{k-1} + \left(\sqrt{(D+\Lambda)\mathbf{P}_{k-1}^{\mathbf{X}}}\right)_d, \quad d = 1, \ldots, D$$

$$\mathcal{X}_{k-1}^d = \hat{\mathbf{X}}_{k-1} - \left(\sqrt{(D+\Lambda)\mathbf{P}_{k-1}^{\mathbf{X}}}\right)_{d-D}, \quad d = D+1, \ldots, 2D$$

$\triangleright$ Scaling factor $\Lambda = \alpha^2(D+\beta) - D$

Compute predicted sigma points for the current state $\mathcal{X}_{k|k-1}^d$:

$$\mathcal{X}_{k|k-1}^d = g\left(\mathcal{X}_{k-1}^d, \mathbf{U}_{k-1}\right)$$

Compute process noise covariance matrix: $\mathbf{Q}_{k-1} = \mathbf{K}_{s*,k-1}$

Compute predicted current state mean and covariance:

$$\hat{\mathbf{X}}_k^- = \sum_{d=0}^{2D} W_m^{(d)}\mathcal{X}_{k|k-1}^d$$

$$\mathbf{P}_k^{\mathbf{X}^-} = \sum_{d=0}^{2D} W_c^{(d)}\left[\mathcal{X}_{k|k-1}^d - \hat{\mathbf{X}}_k^-\right]\left[\mathcal{X}_{k|k-1}^d - \hat{\mathbf{X}}_k^-\right]^T + \mathbf{Q}_{k-1}$$

The weights are defined as:

$$W_m^{(0)} = \frac{\Lambda}{D+\Lambda}$$

$$\triangleright\ W_c^{(0)} = \frac{\Lambda}{D+\Lambda} + \left(1 - \alpha^2 + \beta\right)$$

$$W_m^{(d)} = W_c^{(d)} = \frac{1}{2(D+\Lambda)}, \quad d = 1, \ldots, 2D$$

Update Step:

Compute sigma points for the predicted observation $\mathcal{Y}_{k|k-1}^d$:

$$\mathcal{Y}_{k|k-1}^d = h\left(\mathcal{X}_{k|k-1}^d, \mathbf{U}_k\right)$$

$\triangleright$ $\mathbf{U}_k$ is the system input at the current step.

Compute measurement noise covariance matrix: $\mathbf{R}_k = \mathbf{K}_{o*}$

Compute predicted observation mean and covariance:

$$\hat{\mathbf{Y}}_k^- = \sum_{d=0}^{2D} W_m^{(d)}\mathcal{Y}_{k|k-1}^d$$

$$\mathbf{P}_k^{\mathbf{Y}^-} = \sum_{d=0}^{2D} W_c^{(d)}\left(\mathcal{Y}_{k|k-1}^d - \hat{\mathbf{Y}}_k^-\right)\left(\mathcal{Y}_{k|k-1}^d - \hat{\mathbf{Y}}_k^-\right)^T$$

Compute innovation covariance matrix: $\mathbf{S}_k = \mathbf{P}_k^{\mathbf{Y}^-} + \mathbf{R}_k$

Compute cross-covariance matrix:

$$\mathbf{P}_k^{\mathbf{X}^-\mathbf{Y}^-} = \sum_{d=0}^{2D} W_c^{(d)}\left(\mathcal{X}_{k|k-1}^d - \hat{\mathbf{X}}_k^-\right)\left(\mathcal{Y}_{k|k-1}^d - \hat{\mathbf{Y}}_k^-\right)^T$$

Compute innovation (residual):

$$\mathbf{e}_k = \mathbf{Y}_k - \hat{\mathbf{Y}}_k^-$$

For each observation, compute the NIS:

For $o = 1, 2$: $\delta_o = \frac{\mathbf{e}_{k,o}^2}{S_{k,oo}}$

Compare δ_o to the threshold of chi-square distribution with 1 degree of freedom γ:

If $\delta_o > \gamma$, adjust $\mathbf{R}_k$ to obtain adjusted matrix $\mathbf{R}_k'$:

$\mathbf{R}_{k,oo}' = c \times \mathbf{R}_{k,oo}$, where c is the scaling factor.

Recompute the innovation covariance matrix $\mathbf{S}_k$: $\mathbf{S}_k = \mathbf{P}_k^{\mathbf{Y}^-} + \mathbf{R}_k'$

Compute Kalman gain matrix: $\mathbf{K}_k = \mathbf{P}_k^{\mathbf{X}^-\mathbf{Y}^-}\left(\mathbf{S}_k\right)^{-1}$

Update current state mean and covariance:

$$\hat{\mathbf{X}}_k = \hat{\mathbf{X}}_k^- + \mathbf{K}_k\mathbf{e}_k,\ \mathbf{P}_k^{\mathbf{X}} = \mathbf{P}_k^{\mathbf{X}^-} - \mathbf{K}_k\mathbf{S}_k\mathbf{K}_k^T$$

Outputs:

Current state mean $\hat{\mathbf{X}}_k$, current state covariance $\mathbf{P}_k^{\mathbf{X}}$

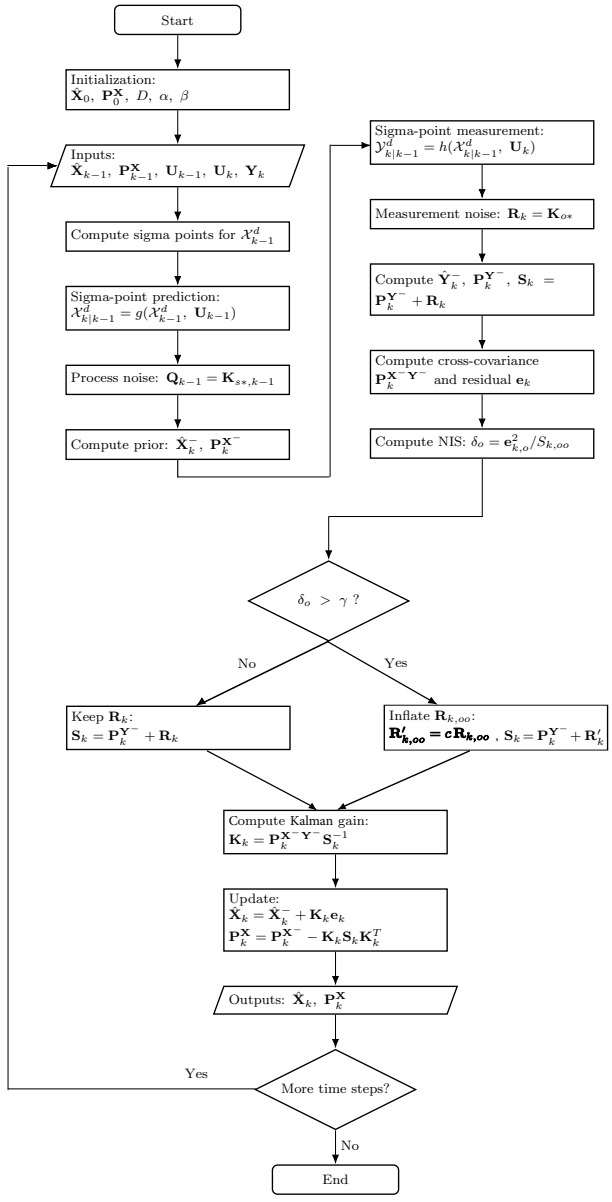

Figure 2: Flowchart of the GPR–UKF-based adaptive battery pack state estimation algorithm.

Furthermore, to compare the mechanical strain of the surface of the cells in the pack with different SOH (capacity), the cells went through a number of cycles in a thermal chamber set at 45°C ambient temperature with 20% humidity. Finally, two cells were selected and assembled into a pack, one with SOH of 96% (Cell 1) and the other with SOH of 99% (Cell 2).

To comply with the practical design considerations in the experimental setup, the cells were arranged side by side with an inter-cell spacing of 10 mm. Each cell was connected to copper test probes for easy disassembly during testing, while the two cells were electrically connected in series via a 10 cm copper wire. The assembled pack was housed in a custom-fabricated 3D-printed polymer enclosure (150 mm × 60 mm × 60 mm) with an open side for thermal management and operational access.

3.2. Layout of fibre optical sensors on the cells

Two FBG sensors with central wavelengths of 1549 nm and 1540 nm, respectively, were affixed onto the surfaces of

Table 2
Battery pack and cell details

Battery or cell properties	Parameters
Nominal Voltage	7.2 V
Nominal Capacity	3200 mAh
Pack Arrangement	2S (2 cells serial-connected)
Inter-cell spacing	10 mm
Pack enclosure	PLA printed (150×60×60 mm)
Cell Nominal Capacity	3200 mAh
Cell Nominal Voltage	3.6 V
Cell cut-off voltage	2.50 V, 4.20 V

the two cells using epoxy. Further, two distinct schemes were examined for laying the FBG sensors [51, 81, 82, 83, 84]. The first approach is to lay the fibre with the FBG sensors at an angle of 30 degrees to the axis of the electrodes as shown in Figure. 3(a). This allows the the optimal capture of the significant wavelength shifts associated to the strain signal trend of the electrodes during the battery's charging and discharging cycles. More specifically, significant wavelength shifts will be observed associated to the increase in overall strain signal strengths during the charging and decrease during discharging, albeit minor fluctuations [51, 81]. Another scheme is to encircle the cylindrical battery cell with the fibre perpendicular to the electrode axis, which allows the capture of the maximal wavelength values. In particular, the signals demonstrated a series of pronounced upward and downward trends throughout a charging/discharging cycle [48, 85]. It is evident that for the later scheme, the strain signals captured by the optical sensors inscribed in the fibre which is laid around the battery's circumference can effectively capture cell-level strain information. Such wavelength data encompass not only the mechanical changes of the electrodes but also the effects from other internal structures, such as separator creep [82, 83]. Besides, another set of FBG sensors (with central wavelengths of 1552 nm and 1555 nm, respectively) for measuring temperatures only were loosely placed at the same location to serve as a reference for monitoring temperature variations. Furthermore, two T-type thermocouples were also affixed to the identified locations on the battery shells to provide reference temperature readings. Thermal paste was applied to all fibres to enhance the thermal conductivity between the battery and the FBGs. The casing of cells was removed to prevent it from being affected by temperature changes, which could lead to inaccurate strain measurements.

3.3. FBG sensor calibration

After the placement of fibre optic sensors on the cells, the FBG sensors were calibrated to facilitate the subsequent thermal-strain signal decoupling from the optical fibre measurements. Initially, the battery was discharged to 0% SOC and placed in a thermal chamber set at 25°C. It was left to rest for four hours, allowing the strain to stabilize and ensuring that the initial strain, ε_0, represented a state prior to lithium-ion insertion. Subsequently, the temperature was varied from 15°C to 50°C with an increment of 5°C in each step, while keeping the humidity constant at 20%. After reaching each set temperature, the sensors were maintained in a constant environment for 2 hours. The average temperature and peak wavelength were then calculated from the data collected in the final 10 minutes at each temperature setting, serving as the representative values for that specific temperature.

3.4. Battery pack testing setup and experimental procedures

The experimental setup for data collection is shown in Figure 3, which includes an optical fibre interrogator, an auxiliary channel for temperature monitoring, a battery tester, and a thermal chamber for environmental control. The constructed pack was placed inside the thermal chamber throughout the experiment, while the room temperature for the experimental facilities was maintained at around 25°C by air conditioning.

A series of tests were then conducted on the battery pack using a NEWARE BTS4000 (60V, 100A) battery tester to control and record the terminal voltage and current of the battery pack. According to the specifications [86], the accuracy for voltage measurements is ±0.1% of the full scale ranging from 0.5 V to 100 V. In this work, the cut-off voltages for the battery pack were set at 8.2V during charging and 5.6V during discharging.

In addition, an auxiliary channel was employed to measure the temperature of each cell. This temperature data is essential for decoupling the fibre strain and temperature signals during the fibre calibration process. The tests were conducted in a BINDER (LIT MK 240) thermal chamber, which maintained ambient temperature and humidity at constant 25°C and 20% relative humidity, respectively, to ensure stable electrochemical conditions and consistent strain patterns for reliable SCS analysis. The host computer with the installed BTS-8.0 software was used to program the test procedure and store the test data. The Micron-Optics SM-130 interrogator was used to measure the wavelength shifts of the FBG sensors. The maximum wavelength stability of the Optical Sensing Interrogator is 5 pm, and the wavelength repeatability is specified as 1 pm [87]. The sampling frequency was set to 1 Hz for all tests performed in the study.

The battery pack was initially tested using a CC-CV charging protocol (0.5C charge up to 8.2 V, with a current cut-off at 0.0625C), followed by a 30-minute rest period and subsequent discharge at 2C down to 5.6 V. Each test phase consisted of 15 continuous cycles to minimise errors in the SOC calculation from the Coulomb counting method [44]. Following this, the battery pack was disassembled, and each cell's SOH was individually assessed by measuring its charging capacity at a C/25 rate and comparing it with its initial capacity. Before reassembly, all cells were discharged to 0% SOC at a C/25 rate.

To evaluate the developed model under complex operating conditions, dynamic performance tests were conducted. The Dynamic Stress Test (DST) was performed with the

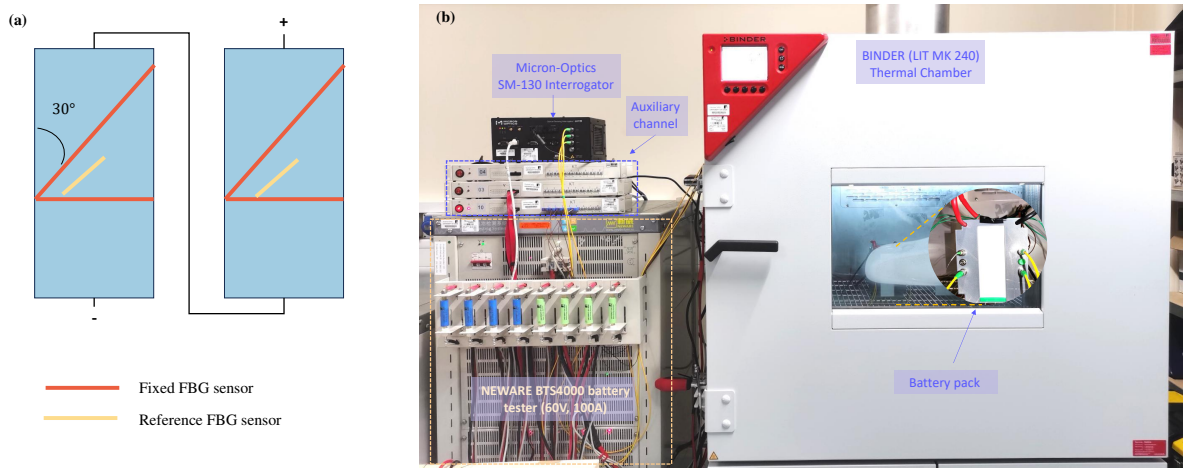

Figure 3: Schematic and experimental setup used in this work: (a) Schematic of the FBG sensor configuration; (b) Experimental setup for data collection.

battery pack placed inside the thermal chamber, ensuring controlled ambient conditions, whereas the Federal Urban Driving Schedule (FUDS) protocol was carried out under the laboratory's air-conditioned environment, where the room temperature was maintained between 23 and 25°C. After the battery pack was fully charged using a CC–CV profile, it was discharged at a constant current of 0.5 C to approximately 10% SOC. Dynamic cycling was then applied until the pack reached the lower cut-off voltage.

4. Data pre-processing

Due to the inherent inconsistencies and noise present in the raw data collected from the battery testing system, pre-processing is indispensable. This process involves addressing missing values, synchronizing time-domain measurements between electrical signals and optical fibre data, decoupling thermal-strain signals from FBG measurements, as well as applying data smoothing and normalization to enhance the accuracy and consistency of the data.

4.1. Synchronization of electrical and optical fibre measurements

The experimental data comprises two distinct datasets. The first consists of electrical measurements, including pack current, terminal voltage, and cell surface temperature, collected using the NEWARE BTS400 battery tester. The second contains fibre peak wavelength data obtained through the Micron-Optics SM-130 Interrogator. Initially, each dataset was individually processed to remove duplicate records and fill missing data using forward-filling techniques based on timestamps, ensuring that only a single entry exists for each time point. Subsequently, by identifying the overlapping time range of the two datasets, a common time index was established to synchronize the datasets. Finally, the datasets were aligned to this common time index through re-indexing, ensuring temporal consistency across all measurements.

4.2. Battery pack SOC calculation

The battery pack SOC is defined as the ratio of the pack's remaining available capacity (Q_a) to its total capacity (Q_n), where the total capacity of the pack is determined by the cell with the lowest capacity (i.e., the representative cell in the pack). The SOC of the pack at each time step is then computed as:

$$SOC(t) = SOC\left(t_0\right) + \int_{t_0}^{t} \frac{I(t) \cdot \eta}{Q_n} dt \tag{15}$$

where $SOC(t)$ and $SOC\left(t_0\right)$ represent the SOC values at time t and the initial time t_0, respectively. η denotes the coulombic efficiency, which is set to approximately 1 for Li-ion batteries [13].

In this study, the total pack capacity was held constant over 15 consecutive cycles for SOC calculation, due to the negligible change in capacity observed within this cycle range [88].

4.3. Thermal-strain signal decoupling for FBG sensor measurements

Considering that the reflected Bragg shift is sensitive to both mechanical variations and temperature fluctuations, a reference sensor has been employed for decoupling to eliminate the impact of temperature [81, 89].

In the aforementioned experiment, a reference FBG sensor was used to exclusively measure the local temperature:

$$\Delta\lambda_{\text{temp}}^{\text{ref}} = K_{\text{T}}^{\text{temp}} \Delta T \tag{16}$$

where $\Delta\lambda_{\text{temp}}^{\text{ref}}$ represents the wavelength change in the reference FBG sensor, calculated as the difference between the wavelength induced by temperature changes and the initial central wavelength of the FBG sensor. $K_{\text{T}}^{\text{temp}}$ denotes the thermal response factor, and ΔT is the change in temperature.

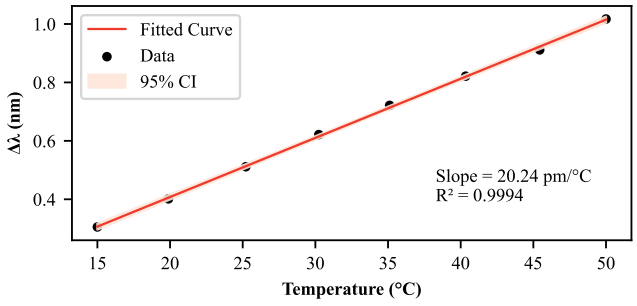

Figure 4: Responses of the wavelength shift (λ_1 in this study) to temperature changes.

The adjacent FBG sensor measures the total variations caused by strain (ε) and temperature, which is given as:

$$\Delta\lambda_{\text{total}} = \Delta\lambda_{\text{strain}} + \Delta\lambda_{\text{temp}}^{\text{adj}} \tag{17}$$

where $\Delta\lambda_{\text{total}}$ is the total change in wavelength measured by the adjacent FBG sensors, $\Delta\lambda_{\text{strain}}$ is the change in wavelength due to strain, and $\Delta\lambda_{\text{temp}}^{\text{adj}}$ is the change in wavelength due to temperature measured by the adjacent FBG sensor.

Defining the strain sensitivity as K_ε, and the thermal response of the adjacent "fixed" FBG sensor as $K_{\text{T}}^{\text{adj}}$, consequently, the equation can be reformulated as:

$$\Delta\lambda_{\text{total}} = K_\varepsilon \Delta\varepsilon + K_{\text{T}}^{\text{adj}} \Delta T \tag{18}$$

This equation illustrates how the total wavelength change ($\Delta\lambda_{\text{total}}$) in the "fixed" FBG sensor is influenced by both the change in mechanical strain ($\Delta\varepsilon$) and temperature variation (ΔT). The thermal response factors ($K_{\text{T}}^{\text{temp}}$ and $K_{\text{T}}^{\text{adj}}$) can be determined through the FBG sensor calibration experiment.

Figure 4 demonstrates a robust linear correlation between the wavelength shift ($\Delta\lambda$) and temperature, exhibiting a thermal sensitivity coefficient of 20.34 pm/°C. The FBG temperature response has been meticulously calibrated across a temperature range of 15 to 50 °C, utilising commercial K-type thermocouples strategically positioned adjacent to each FBG sensor throughout the calibration procedure. The calibration yielded an R^2 value of 99.9%, confirming the exceptional linearity of the temperature-wavelength relationship.

The calibration outcomes for the four FBG sensors are detailed below:

$$\begin{cases} \Delta\lambda_1 = 0.02034 * T \\ \Delta\lambda_2 = 0.02051 * T \\ \Delta\lambda_3 = 0.01004 * T \\ \Delta\lambda_4 = 0.01024 * T \end{cases} \tag{19}$$

where $\Delta\lambda_1, \Delta\lambda_2, \Delta\lambda_3$, and $\Delta\lambda_4$ represent the peak wavelength changes of the four FBG sensors, respectively. T denotes the temperature on the surface of the battery. $\Delta\lambda_1$ and $\Delta\lambda_2$ are influenced by the combined effects of temperature and strain variations on the surfaces of cells 1 and 2, respectively. Meanwhile, $\Delta\lambda_3$ and $\Delta\lambda_4$ act as temperature references for $\Delta\lambda_1$ and $\Delta\lambda_2$, respectively.

Referring to Eqs. (16) and (19), the strain-induced wavelength changes on the surface of the cell can be derived as

$$\begin{cases} \Delta\lambda_1^{\text{strain}} = \Delta\lambda_1 - \frac{0.02034}{0.01004} * \Delta\lambda_3 \\ \Delta\lambda_2^{\text{strain}} = \Delta\lambda_2 - \frac{0.02051}{0.01024} * \Delta\lambda_4 \end{cases} \tag{20}$$

The strain sensitivity of the FBG sensors used in this study is 0.78 pm/$\mu\varepsilon$. Consequently, the strain values can be calculated using Eq. (21):

$$\begin{cases} \varepsilon_1 = 0.78\Delta\lambda_1 - 1.5802 * \Delta\lambda_3 \\ \varepsilon_2 = 0.78\Delta\lambda_2 - 1.5623 * \Delta\lambda_4 \end{cases} \tag{21}$$

where ε_1 and ε_2 represent the strain signals derived from $\Delta\lambda_1$ and $\Delta\lambda_2$, respectively.

4.4. Deciphering $\frac{d\varepsilon}{dQ}$ values

4.4.1. Strain data normalisation

To compare the ageing characteristics of cells, strain values were normalised using Eq. (22):

$$E_i(t) = \frac{\varepsilon_i(t) - \varepsilon_{i,\min}}{\varepsilon_{i,\max} - \varepsilon_{i,\min}} \tag{22}$$

where $E_i(t)$ represents the normalized strain for the i-th cell at time t; $\varepsilon_i(t)$ denotes the original strain of cell i at time t; $\varepsilon_{i,\min}$ is the minimum strain value over the measurement period for cell i; and $\varepsilon_{i,\max}$ is the maximum strain value over the measurement period for cell i. This Min-Max normalisation scales the strain values to a range of [0, 1], where 0 corresponds to the minimum strain and 1 corresponds to the maximum strain for each cell. Here, $i \in \{1, 2\}$ indicates the cell index.

4.4.2. Data preparation for calculating $\frac{d\varepsilon}{dQ}$

A primary limitation of the $\frac{d\varepsilon}{dQ}$ method is its inherent requirement for high-precision equipment, as the strain difference between consecutive data points is typically minimal. Direct calculation of this derivative using strain and total battery charge measurements can lead to significant errors due to noises and inaccuracy being introduced into the measurements [90]. For instance, when utilizing data acquired from a Micron-Optics SM-130 interrogator and a NEWARE BTS4000 battery tester, a direct computation of $\frac{d\varepsilon}{dQ}$ by simply dividing the differential strain with the differential charge often results in excessively noisy data, as illustrated in Figure 5(b) by the yellow curves. To filter out these noises to obtain smooth $\frac{d\varepsilon}{dQ}$ values involves the following steps:

1. Construct an array of 1000 uniformly distributed battery total charge segments (denoted as q_{seg}), which function as a Coulomb counter for the analysis. This array serves to segment total capacity into discrete quantifiable units.

2. Establish a corresponding array for strain change segments $\left(s_{\text{seg}}\right)$, ensuring a one-to-one association with q_{seg}. This mirroring structure facilitates the direct comparison and analysis of battery total charge level and strain data.

3. Process each time-series data point by:

   (a) Identifying the total charge segment closest to the current battery total charge count and recording its index i;

   (b) Computing the change in strain for the subsequent time interval;

   (c) Aggregating changes in strain into the s_{seg} corresponding to index i.

4. After populating the strain change segments $\left(s_{\text{seg}}\right)$, normalize the total strain change in each segment by dividing it by the respective battery total charge change interval (ΔQ). This normalization step adjusts for variations in battery total charge increments and enables a more accurate analysis of strain response relative to battery total charge changes.

5. Apply a smoothing algorithm to data samples in the dataset obtained in step 4 to enhance the accuracy of the results by reducing the impact of measurement noises.

To retain the vital information such as peaks from the results of step 4, a weighted Savitzky-Golay filter (SGW) [91] was applied in step 5. This digital filtering technique uses polynomial fitting for data smoothing by fitting adjacent data samples with a low-degree polynomial via linear least squares. This data smoothing step ensures the preservation of critical data structures and attributes, such as peak heights and widths, while offering a more nuanced approach to balancing noise reduction and data integrity.

The data smoothing step is detailed below. The smoothed value ζ_i' at each point i is computed as a linear combination of the surrounding data points, including itself, weighted by the SGW coefficients. The formula for computing an SGW smoothed point is given below.

$$\zeta_i' = \sum_{j=-\omega}^{\omega} c_j \cdot \zeta_{i+j} \tag{23}$$

where ζ_{i+j} represents the data points within the window centred around sample point i.

The SGW coefficients c_j are determined through an adapted Savitzky-Golay approach combining polynomial least squares fitting with squared Hann window weighting. This enhancement preserves critical signal features whilst effectively suppressing noise. The coefficients are expressed as:

$$c_j = \frac{s_j \cdot h_j^2}{\sum_{k=-\omega}^{\omega} s_k \cdot h_k^2} \tag{24}$$

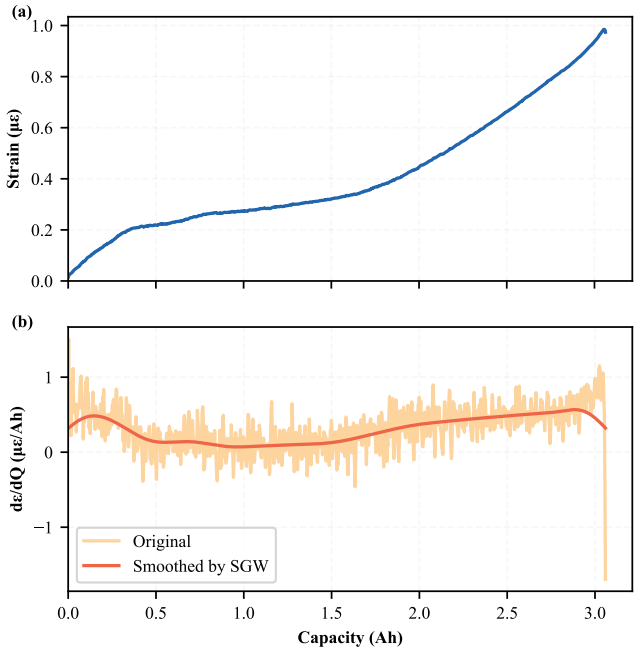

Figure 5: Analysis of strain variations during battery cell charging: (a) Strain variations during the charging process; (b) Derivatives of strain variations with respect to capacity.

where s_j represents the standard Savitzky-Golay coefficients derived from least squares polynomial fitting of order ρ [91], and h_j denotes the Hann window function [92] evaluated at position j within the window. The indices j and k both range from $-\omega$ to $+\omega$, where ω defines the half-window size, yielding a total window size of $2\omega + 1$. The resultant SGW filter minimises the mean squared error of the polynomial fit whilst emphasising central data points through the squared Hann weighting. Throughout this work, the parameters $\omega = 175$ and $\rho = 2$ are adopted unless stated otherwise.

To apply the SGW method for smoothing $\frac{d\varepsilon}{dQ}$ values, the differential quantity $\frac{d\varepsilon}{dQ}$ is first calculated for each pair of consecutive data points. Then, the SGW smoothing is applied to these $\frac{d\varepsilon}{dQ}$ values.

$$\frac{d\varepsilon}{dQ} = \frac{\varepsilon_{i+1} - \varepsilon_i}{Q_{i+1} - Q_i} \tag{25}$$

where ε_i and Q_i are the strain and battery total charge at the i^{th} point, respectively.

After calculating the raw $\frac{d\varepsilon}{dQ}$ values, the SGW filter is applied in step 5:

$$\left(\frac{d\varepsilon}{dQ}\right)_i' = \sum_{j=-\omega}^{\omega} c_j \cdot \left(\frac{d\varepsilon}{dQ}\right)_{i+j} \tag{26}$$

where $(\frac{d\varepsilon}{dQ})_i'$ is the smoothed $\frac{d\varepsilon}{dQ}$ value at point i, providing a cleaner signal that helps identify key features and trends in the data with reduced noise.

Figure 5(a) illustrates the strain variations in a single cell while charging at a C/25 rate, with strain values normalised to a range between 0 and to 1 $\mu\varepsilon$. As the applied current is small, the temperature rise of the battery is minimal (see Figure 9(b)), suggesting that the observed strain measurements are primarily due to the intercalation of lithium ions into the electrode. Figure 5(b) depicts the corresponding derivative of strain with respect to battery total charge level, $\frac{d\varepsilon}{dQ}$. The yellow curve represents the $\frac{d\varepsilon}{dQ}$ results calculated from the dataset shown in Figure 5(a), while the red curve, smoothed by Eq. (26), filters out noise, preserving the main trends and peaks of the data.

5. Performance assessment and comparative analysis

5.1. Comparative analysis

The adaptive GPR–UKF framework for pack-level state estimation is evaluated by comparing it directly with a GPR–UKF that uses fixed process (**Q**) and observation (**R**) covariance matrices.

The process covariance **Q** is determined from the empirical variance of state-prediction residuals generated by the GPR model (see Section 2.2.1). The residual, $\mathbf{r}_k$, is defined as

$$\mathbf{r}_k = \mathbf{X}_k - \hat{\mathbf{X}}_k \tag{27}$$

where $\mathbf{X}_k$ and $\hat{\mathbf{X}}_k$ denote the actual and predicted states, respectively.

The variance of the residuals for each state variable is calculated as

$$\sigma^2_{\mathrm{SOC}} = \frac{1}{N}\sum_{k=1}^{N}\left(r_{\mathrm{SOC},k}\right)^2, \quad \sigma^2_T = \frac{1}{N}\sum_{k=1}^{N}\left(r_{T,k}\right)^2 \tag{28}$$

where N is the total number of samples. Accordingly, the process noise covariance matrix **Q** is constructed as

$$\mathbf{Q} = \begin{bmatrix} \sigma^2_{\mathrm{SOC}} & 0 \\ 0 & \sigma^2_T \end{bmatrix} \tag{29}$$

The observation covariance matrix **R** is determined from instrument specifications (see Section 3.4). The strain measurement error is primarily attributed to the wavelength measurement error introduced by the FBG sensor and the optical fibre interrogator [87, 93]. The strain error is calculated as

$$e_\varepsilon = \frac{e_\lambda}{S} \tag{30}$$

where e_ε represents the strain error, e_λ is the total wavelength error in the measurement, and S denotes the strain sensitivity of the FBG sensors. The voltage measurement error (e_V) is determined based on the accuracy specifications of the measuring equipment [86].

Accordingly, the observation noise covariance matrix **R** is given by

$$\mathbf{R} = \begin{bmatrix} e_V^2 & 0 \\ 0 & e_\varepsilon^2 \end{bmatrix} \tag{31}$$

Additionally, three widely adopted machine-learning regressors are implemented for comparative analysis. The first approach adopts a standard GPR model, utilizing the same hybrid kernel function structure defined in Eq. (3). The second approach is the Extreme Gradient Boosting (XGBoost), a gradient-boosted ensemble of decision trees that iteratively fits CART base learners to residual errors while regularizing tree complexity. XGBoost's strength in modelling non-linear interactions without extensive feature engineering makes it a robust data-driven baseline for battery SOC estimation (algorithmic details and hyper-parameter ranges are provided in Appendix A.1). The third approach employs Long Short-Term Memory (LSTM) networks, which are recurrent neural networks equipped with gated cells to control information flow across time steps [94]. This architecture allows LSTMs to capture long-range temporal dependencies in voltage, current, and temperature sequences, providing another strong baseline for comparison (network architecture and training settings are detailed in Appendix A.2).

5.2. Training and testing procedure

The developed GPR–UKF approach was evaluated using data collected from a pack composed of two NCR18650 cylindrical cells under three distinct conditions: CC discharging, and discharging under DST and FUDS conditions. To further evaluate the approach, a pack dataset collected from a pack containing two $LiFePO_4$ cells was also employed. Due to the $O\left(n^3\right)$ computational complexity of GPR, models were trained and tested separately for each scenario to manage data size efficiently.

For training, Cell 1 was selected as the representative cell in the battery pack, thereby simplifying the model training process. This approach is justified in practical applications, where battery packs typically comprise numerous cells connected in series, making it challenging to utilize all cells for GPR model training [70]. The GPR models were trained with two full cycles per scenario. In this phase, the input for the state prediction model was $[SOC^1_{k-1}, T_{k-1}, I_{k-1}]$, and the output was $[SOC^1_k, T_k]$. For the observation model, the input was $[SOC^1_{k-1}, T_{k-1}, I_{k-1}]$, and the output was $[V_{k-1}, \varepsilon^1_{k-1}]$.

Testing was conducted on both Cell 1 and Cell 2 to assess generalization across different cells. The state prediction model predicted SOC values for both Cell 1 (SOC^1_k) and Cell 2 $\left(SOC^2_k\right)$, based on their respective previous SOC values and the shared average temperature. The observation model predicted the strain for both cells (ε^1_k and ε^2_k), using the actual SOC values for each cell and the common average temperature. The total pack voltage (V_k) remained constant across cells as it represents a pack-level measurement.

For the baseline machine-learning regressors, the training inputs are $[V_k, \varepsilon^1_k, I_k]$ while the outputs are the pack states $[\mathrm{SOC}^1_k, T_k]$, again treating Cell 1 as the representative cell.

5.3. Performance evaluation metrics

Several evaluation metrics are utilised to assess the performance of the proposed GPR-UKF model in estimating battery SOC and average temperature. The selected metrics include the Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Coefficient of Determination (R^2), and Mean Absolute Percentage Error (MAPE). These metrics provide a comprehensive evaluation of model accuracy and robustness. The formulas used are outlined as follows:

a. MAE

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|z_i - \hat{z}_i| \tag{32}$$

b. MSE

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(z_i - \hat{z}_i)^2 \tag{33}$$

c. RMSE

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(z_i - \hat{z}_i)^2} \tag{34}$$

d. R^2

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(z_i - \hat{z}_i)^2}{\sum_{i=1}^{N}(z_i - \bar{z})^2} \tag{35}$$

e. MAPE

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{z_i - \hat{z}_i}{z_i}\right| \times 100\% \tag{36}$$

where N is the number of samples, z_i represents the actual value for the i-th sample, $\hat{z}_i$ represents the predicted value, and $\bar{z}$ is the mean of the actual values.

6. Results and Discussions

In this section, the experimental results used to evaluate the effectiveness of the proposed method for estimating the temperature and SOC of a battery pack are presented.

To investigate strain increment characteristics in the battery cell, the SCS is first analysed at the cell level, followed by an analysis at the pack level by comparing the SCSs of two cells connected in series. Finally, the estimation results and comparisons of different approaches are discussed.

6.1. Experimental results

Figure 6 presents the test results under the CC–CV charging and CC discharging profiles, while Figures 7 and 8 present the results of the battery pack under the DST and FUDS conditions, respectively. These figures illustrate the battery pack current, pack terminal voltage, strain variations, SOC, and shell temperatures for two cells. Specifically, Figure 6 includes data from a test procedure involving 15 cycles,

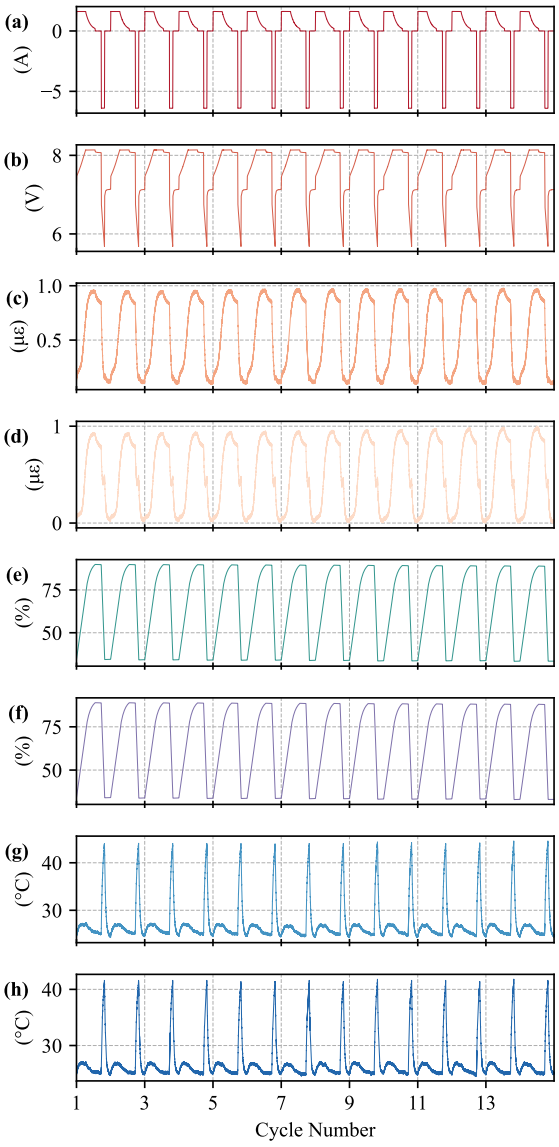

Figure 6: Experimental data from CC-CV charging and CC discharging test over 14 consecutive cycles: (a) Battery pack current; (b) Battery pack terminal voltage; (c) Strain variations in Cell 1; (d) Strain variations in Cell 2; (e) SOC of Cell 1; (f) SOC of Cell 2; (g) Shell temperature of Cell 1; (h) Shell temperature of Cell 2.

as introduced in Section 3.4. The first cycle was excluded from the analysis due to an inconsistent data range, as it began at 0% SOC after pack assembly. The DST and FUDS results start from approximately 90% SOC and continue until the pack reaches its cut-off terminal voltage.

Given the differences in strain scale between cells, strain values were normalized according to Eq. (22). Strain variations increased significantly during CC charging, then decreased in the CV phase and during relaxation after CC discharging, as depicted in Figures 6(c) and 6(d). In addition, over the 14 consecutive cycles, the maximum strain values increased, while the minimum values at the end of each

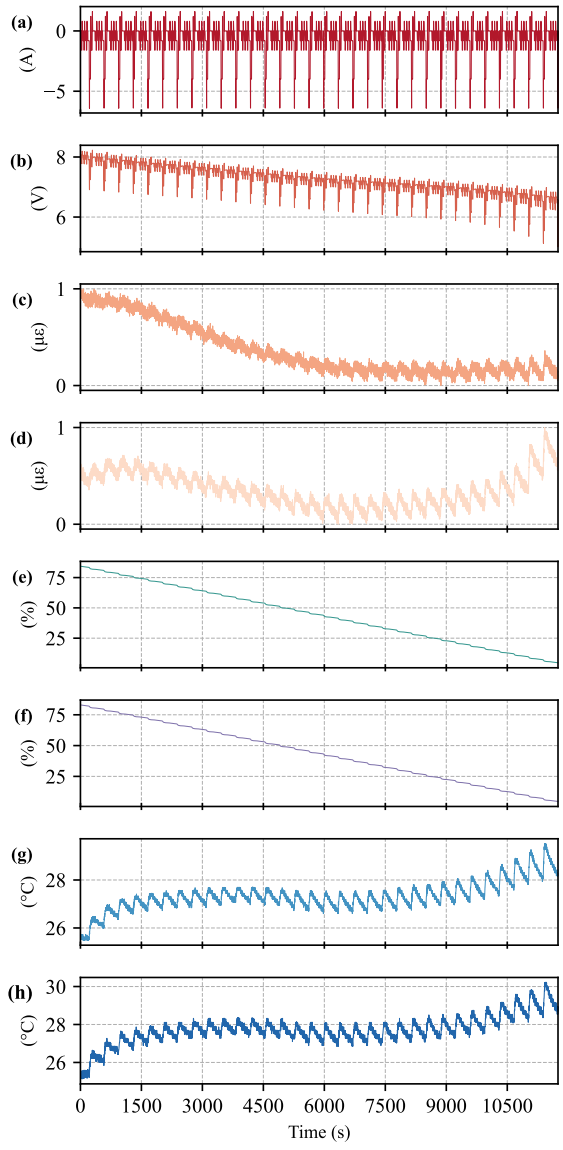

Figure 7: Experimental data under DST test: (a) Battery pack current; (b) Battery pack terminal voltage; (c) Strain variations in Cell 1; (d) Strain variations in Cell 2; (e) SOC of Cell 1; (f) SOC of Cell 2; (g) Shell temperature of Cell 1; (h) Shell temperature of Cell 2.

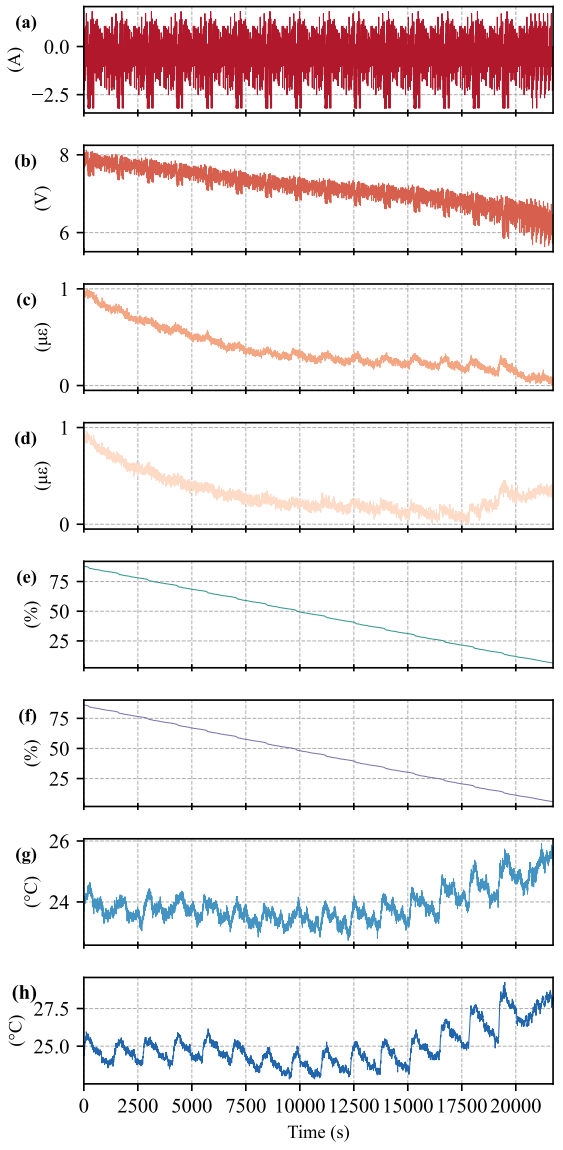

Figure 8: Experimental data under FUDS test: (a) Battery pack current; (b) Battery pack terminal voltage; (c) Strain variations in Cell 1; (d) Strain variations in Cell 2; (e) SOC of Cell 1; (f) SOC of Cell 2; (g) Shell temperature of Cell 1; (h) Shell temperature of Cell 2.

cycle remained nearly unchanged. The shell temperatures of the cells were measured using calibrated FBG sensors. During CC charging, peak temperatures reached approximately 28°C, decreasing in the CV phase and approaching ambient temperature after a 30-minute relaxation period. In contrast, after CC discharging, the peak temperature exceeded 40°C, as shown in Figures 6(g) and 6(h).

In the DST results shown in Figure 7, the strain variations for each 360-second DST cycle (Figures 7(c) and 7(d)) correlated with changes in pack current, voltage, and cell temperatures. However, the overall strain trends differed between cells. For Cell 1 (Figure 7(c)), the strain showed an overall decreasing trend. For Cell 2, the strain initially increased, began to decrease after approximately 1,500 seconds, then displayed another increasing trend after 7,500 seconds, ultimately reaching a maximum at the final DST cycle. Furthermore, starting at around 1,000 seconds, the temperature of Cell 2 (Figure 7(h)) exceeded that of Cell 1 (Figure 7(g)), with a peak above 30°C near the end of the test.

Under the FUDS load, the time-varying current elicits a strongly nonlinear response in the pack. The normalised strain in both cells exhibits an overall decreasing trend, broadly following the net reduction in SOC. The temperatures of the cells remain close to the controlled ambient conditions (approximately 24°C) until the SOC falls below

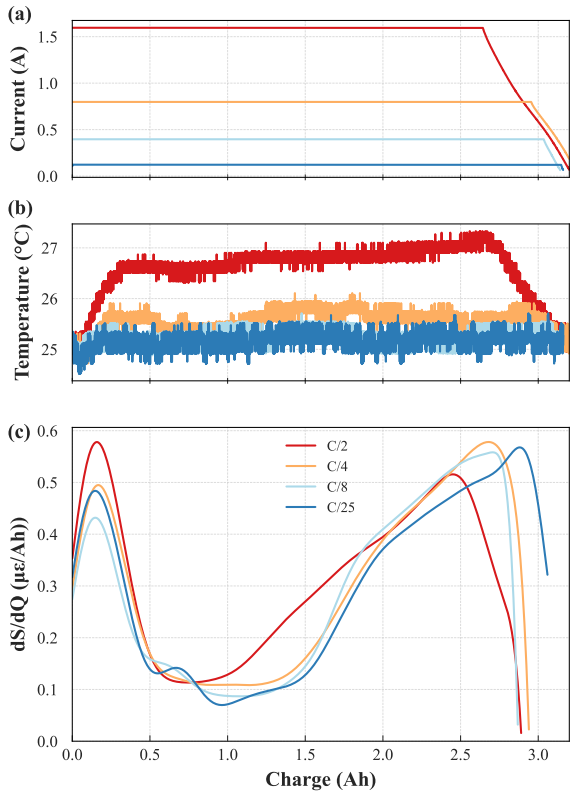

Figure 9: (a) Current profiles under different charging rates; (b) Surface temperature characteristics under different charging rates; (c) SCS analysis under different charging rates for one cell.

30%, after which a more pronounced overall temperature rise is observed, accompanied by noticeable strain fluctuations, particularly in Cell 2.

6.2. Strain-charge sensitivity analysis at cell-level

To investigate the SCS characteristics of the battery cell, experiments were conducted at four different charging rates, ranging from C/25 to C/2, as shown in Figure 9(a). Figure 9(b) presents the shell temperature recorded by calibrated FBG sensors for each charging rate. Following the methodology introduced in Section 4.4, the resulting curves of $\frac{d\varepsilon}{dQ}$ with respect to Q are presented in Figure 9(c).

The results indicate that the charging rate significantly affects the formation and positioning of peaks in the $\frac{d\varepsilon}{dQ}$ curves, though the associated charge values show minimal variation across different C-rates. As charging rates decrease, the number of peaks in the $\frac{d\varepsilon}{dQ}$ curves becomes more pronounced. At high charging rates, the first peak increases in magnitude, and the second peak shifts leftward. Notably, the middle peaks fail to form in the $\frac{d\varepsilon}{dQ}$ curves across charging rates from C/25 to C/2.

At lower charging rates (e.g., C/25), strain evolution is predominantly driven by changes in the graphite electrode [95]. The strain response during phase transitions in graphite is nonlinear, as shown in Figure 5. Initially, strain increases steeply, then slows during the transition, entering a plateau with minimal change before increasing again due to greater volume expansion in later lithiation stages. Thus, the peaks observed in the $\frac{d\varepsilon}{dQ}$ curves are linked to specific phase transitions during different stages of lithiation.

At higher charging rates, however, temperature becomes a critical factor influencing the SCS profiles. This effect is evident in C/2, as shown in Figures 9(a) and (b). Temperature rises substantially during lower SOC levels due to higher internal resistance [96, 97]. The resulting thermal expansion makes the strain response less sensitive to changes originating from the graphite electrode alone. In addition, the combined impact of elevated internal temperature and lithiation leads to a more pronounced first peak in the SCS curve at C/2.

According to Figures 9(a) and (c), the final peak in the SCS profile occurs near the end of the CC charging phase under all four charging rates. At higher charging rates, the cell reaches the cut-off voltage more quickly, transitioning to the CV phase sooner, as shown in Figure 9(a). During the CV phase, the current gradually decreases, leading to reduced temperature effects (Figure 9(b)) and a slower rate of charge increase. The peak values observed in Figure 9(c) occur slightly before the corresponding CV start points, which stems from the application of a smoothing algorithm when deriving the $\frac{d\varepsilon}{dQ}$ values. This discrepancy could be minimised by utilising more advanced smoothing techniques.

Moreover, as the battery ages, internal resistance tends to increase due to factors such as electrolyte decomposition and SEI layer growth. This increased resistance causes the battery to reach the cut-off voltage more quickly during the CC phase, thereby shortening the CC duration. Furthermore, the temperature rise at the end of the CC phase becomes more significant. Consequently, the second peak in the $\frac{d\varepsilon}{dQ}$ curve is expected to shift leftward and/or upward as the battery ages, highlighting its potential as a diagnostic indicator for ageing detection.

6.3. Analysis of SCS under pack level

Figure 10 presents the $\frac{d\varepsilon}{dQ}$ versus Q for two series-connected cells during the C/2 CC-CV charging process. Specifically, Figures 10(a) and 10(b) show results for Cell 1 and Cell 2, respectively, while Figure 10(c) provides a comparative analysis of both cells across varying SOH.

As discussed in Section 6.2, two peaks are evident in Figures 10(a) and 10(b). The first peak, occurring around 0.1 Ah, exhibits distinct behaviours between the two cells. For Cell 1, the first peak reaches approximately 0.35–0.5 $\mu\varepsilon$/Ah, with data from 75 cycles (CC-CV charging, CC discharging) showing no significant trend in its magnitude variations over time. In contrast, for Cell 2, a decreasing trend in the first peak emerges after approximately 20 cycles. The second peak, however, displays a consistent upward trend for both cells as cycling progresses.

Figure 10(c) further compares the two cells, with Cell 1 represented by solid lines and Cell 2 by dashed lines. Curves of the same colour indicate the same cycle numbers for

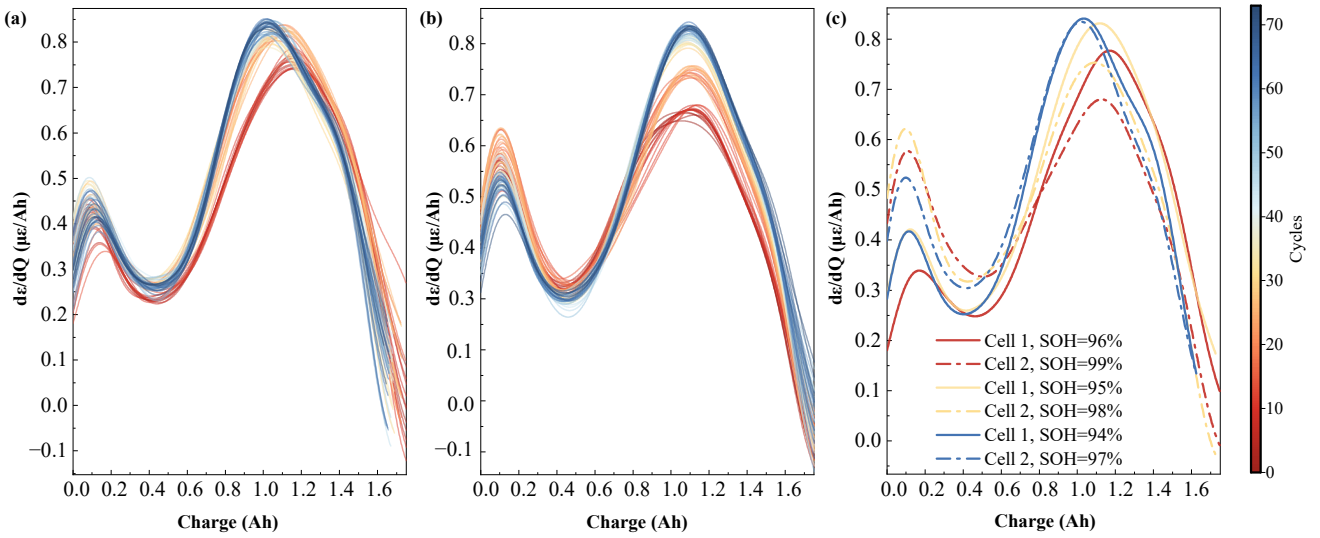

Figure 10: SCS during charging for two series-connected cells at a standard charging rate (C/2): (a) represents Cell 1, (b) represents Cell 2, and (c) compares both cells under varying SOH.

both cells. SOH values were determined by disassembling the battery pack and individually discharging each cell at a C/25 rate after specific cycles. At all measured stages, Cell 1 exhibited lower SOH than Cell 2, indicating more pronounced ageing. It can be observed that under identical cycle numbers and charge conditions, the more aged cell displays a higher second peak value. Additionally, as the number of cycles increases, the charge value corresponding to the second peak shifts to lower values. In the following section, the second peak values from the SCS analysis is used to identify the most aged cell as the representative cell in the pack, as the second peak values provide direct insights into the health condition of cells in the series-connected pack.

6.4. Battery pack state estimation based on the GPR-UKF approach

This section presents a comprehensive evaluation of the GPR-based prediction and observation models applied to a battery pack consisting of two series-connected NCR18650 lithium-ion cells. The assessment under the three distinct conditions follows the procedure of using the SCS approach to identify the representative cell in the pack (as shown in Figure 11), after which the trained models are applied to estimate the pack states. In addition, the proposed GPR-UKF framework is benchmarked against four alternative state estimation methods, enabling a thorough comparison of accuracy and robustness under varied operating conditions. To further demonstrate the generalisability of the proposed approach, the framework is also applied to a pack comprising two series-connected $LiFePO_4$ cells, as detailed in [54]. The resultant state estimation performance is then compared with the findings reported in the referenced study.

6.4.1. Case 1: CC discharging process validation

The optimised hyperparameters for the GPR-based prediction and observation models under the CC discharging

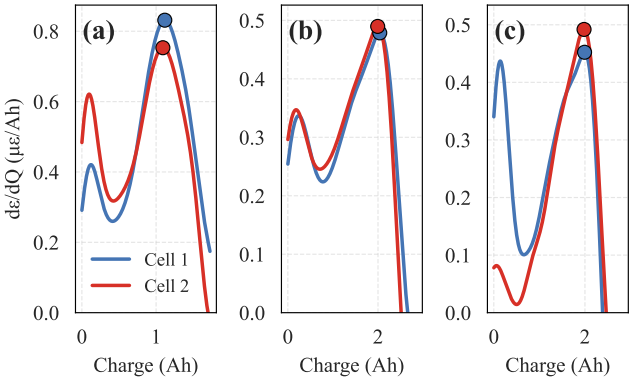

Figure 11: SCS analysis for cells within the pack based on CCCV charging segments under three test protocols: (a) constant–current (CC) cycling, (b) DST, and (c) FUDS. Markers indicate peak SCS values.

Table 3
Optimised Hyperparameters for the GPR-based state prediction model under the CC discharging scenario.

m	$\sigma^2_{SE,s,m}$	$l_{s,m}$	$\sigma^2_{\text{Lin},s,m}$	$\sigma^2_{0,s,m}$	σ^2_s
SOC	2.82×10^{-4}	7.82×10^{8}	1.27×10^{-4}	9.50×10^{4}	
T	2.63×10^{-3}	3.05×10^{5}	1.08×10^{-5}	1.04×10^{1}	1.10×10^{-2}
I	1.54×10^{-5}	2.25×10^{2}	8.49×10^{-3}	8.41×10^{0}	

scenario are listed in Table 3 and Table 4, respectively. Figure 12 illustrates the performance of the GPR-based state prediction and observation models during a CC discharging process lasting approximately 1000 seconds. The insets in each subplot display histograms of the error distributions between the actual and predicted values. Key evaluation metrics for these models are summarised in Table 5.

The calculated matrices $\mathbf{Q}$ and $\mathbf{R}$, derived from the analysis in Section 5.1, are:

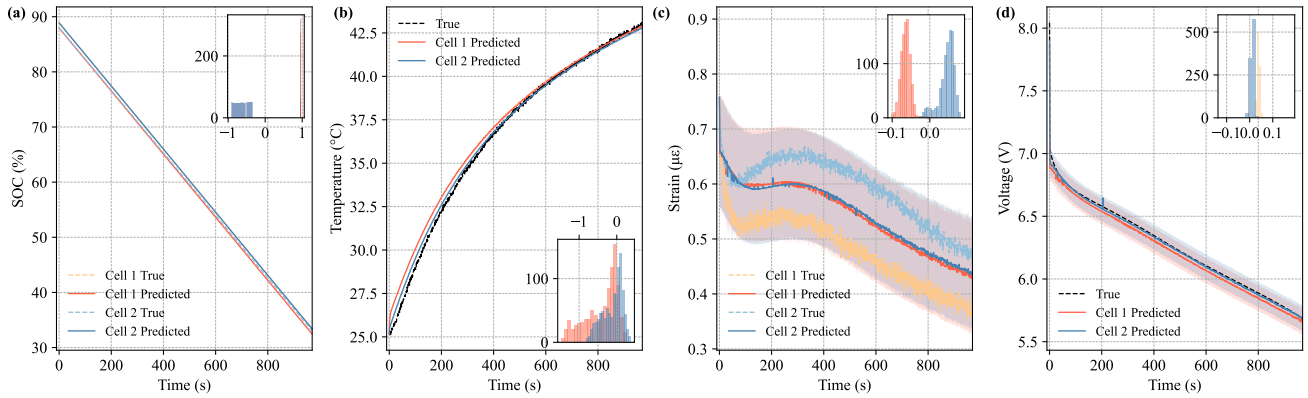

Figure 12: Evaluation of GPR-based prediction and observation models under the CC discharging scenario: (a) Cell SOC; (b) battery pack average temperature; (c) cell strain; (d) battery pack terminal voltage. Insets illustrate the error distributions.

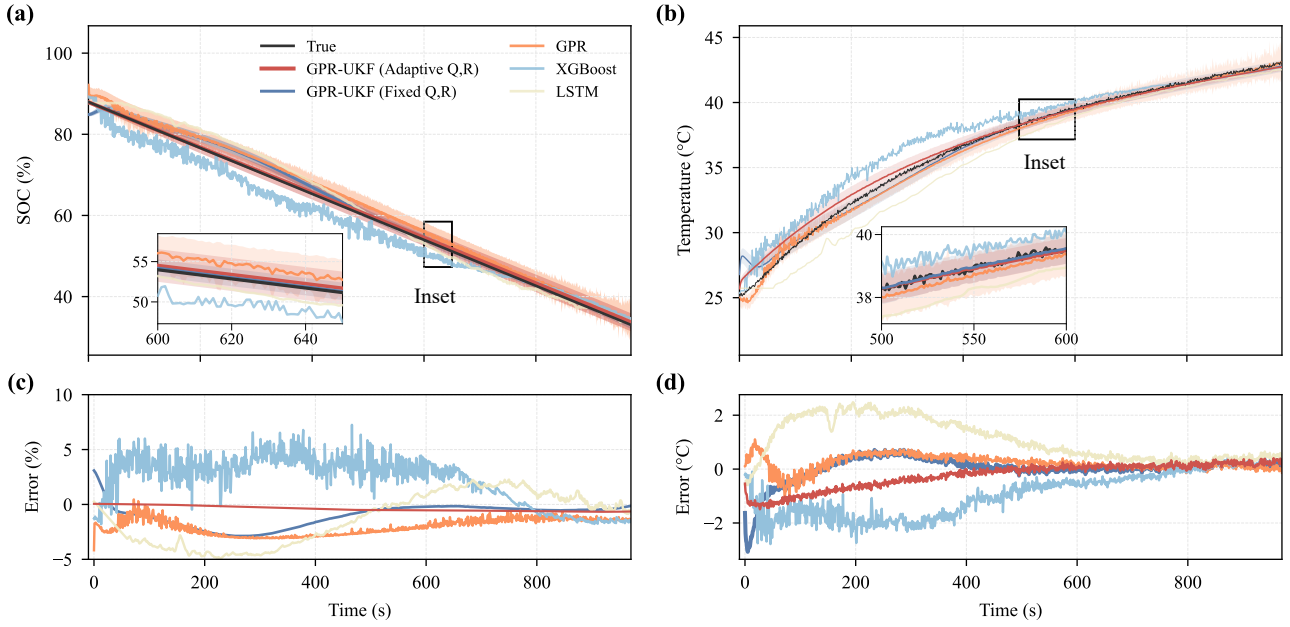

Figure 13: Battery pack SOC and average temperature under the CC discharging scenario: (a) SOC estimation; (b) average temperature estimation; (c) SOC estimation error; (d) average temperature estimation error.

Table 4
Optimised Hyperparameters for the GPR-based observation model under the CC discharging scenario.

m	$\sigma^2_{SE,o,m}$	$l_{o,m}$	$\sigma^2_{\mathrm{Lin},o,m}$	$\sigma_{o,m}$	σ^2_o
SOC	1.77×10^{-5}	1.74	2.69×10^{1}	2.69×10^{1}	1.00×10^{-3}
T	1.36×10^{-4}	2.63×10^{3}	1.96×10^{-3}	1.95×10^{-3}	
I	1.00×10^{-5}	2.04×10^{-4}	2.59×10^{1}	2.58×10^{1}	

$$\mathbf{Q} = \begin{bmatrix} 2.91\times10^{-2} & 0 \\ 0 & 1.79\times10^{-1} \end{bmatrix} \tag{37}$$

$$\mathbf{R} = \begin{bmatrix} 1.00\times10^{-2} & 0 \\ 0 & 2.06\times10^{-2} \end{bmatrix} \tag{38}$$

Table 5
Evaluations for GPR-based states prediction and observation models with Cell1 and Cell2 test data under the CC discharging scenario.

Evaluation	Cell1 Test Data				Cell2 Test Data			
	SOC	T	ε	V	SOC	T	ε	V
MSE	0.1158	0.3355	0.0043	0.0023	0.9425	0.1343	0.0159	0.0022
RMSE	0.3403	0.5792	0.0655	0.0476	0.9708	0.3665	0.1261	0.0470
MAE	0.2944	0.4190	0.0644	0.0393	0.9708	0.3081	0.1240	0.0077
R^2	0.9995	0.9862	−0.1235	0.9800	0.9963	0.9943	−3.1643	0.9808
MAPE (%)	0.6095	1.3273	13.6663	0.6278	1.6822	0.9361	18.5469	0.1149

The GPR-based state prediction model demonstrates both high predictive accuracy and strong generalisation capabilities across both cells, with lower MSE, RMSE, MAE, and MAPE values observed for Cell 2 relative to those predicted using the Cell 1 dataset. Particularly, the

Table 6
Performance metrics for SOC and temperature prediction using five methods in the CC discharging process.

Metric	GPR-UKF (Adaptive Q,R)		GPR-UKF (Fixed Q,R)		GPR		XGBoost		LSTM	
	SOC	T	SOC	T	SOC	T	SOC	T	SOC	T
MSE	**0.2131**	0.2816	2.0463 (89.59%↓)	0.2568 (9.66%↑)	4.6061 (95.37%↓)	**0.1301** (116.45%↑)	9.7226 (97.81%↓)	1.5424 (81.74%↓)	7.0578 (96.98%↓)	1.7420 (83.83%↓)
RMSE	**0.4617**	0.5307	1.4305 (67.72%↓)	0.5067 (4.74%↑)	2.1462 (78.49%↓)	**0.3607** (47.13%↑)	3.1181 (85.19%↓)	1.2419 (57.27%↓)	2.6566 (82.62%↓)	1.3199 (59.79%↓)
MAE	**0.4013**	0.3714	1.0932 (63.29%↓)	0.2927 (26.89%↑)	2.0176 (80.11%↓)	**0.2887** (28.65%↑)	2.7464 (85.39%↓)	0.9965 (62.73%↓)	2.1673 (81.48%↓)	1.0955 (66.10%↓)
R^2	**0.9992**	0.9884	0.9919 (0.74%↑)	0.9895 (0.11%↓)	0.9817 (1.78%↑)	**0.9947** (0.63%↓)	0.9614 (3.93%↑)	0.9366 (5.53%↑)	0.9720 (2.80%↑)	0.9284 (6.46%↑)
MAPE (%)	**0.82**	1.18	1.66 (50.60%↓)	0.94 (25.53%↑)	3.40 (75.88%↓)	**0.86** (37.21%↑)	4.45 (81.57%↓)	2.98 (60.40%↓)	3.33 (75.38%↓)	3.21 (63.24%↓)

coefficient of determination (R^2) for average temperature predictions reaches 0.9943 for Cell 2, outperforming the 0.9862 observed with Cell 1.

However, strain predictions (ε) reveal notable discrepancies across both datasets. As shown in Figure 12(c), error distributions are centred around -0.06 for Cell 1 and approximately 0.5 for Cell 2. Despite relatively low MSE, RMSE, and MAE values in Table 5, the negative R^2 values indicate that the model struggles to capture strain data trends and variance accurately. Conversely, the pack terminal voltage stands out as a consistently well-modelled variable, with R^2 values nearing 0.98 across both test datasets and error distributions tightly centred around zero. This high R^2 value underscores the model's robustness in predicting terminal voltage, supported by minimal error across both cells, as illustrated in Figure 12(d).

In this scenario, SCS analysis was performed using data from the same cycle of the CC-CV charging phase, and Cell 1 was selected as the representative cell. The pack SOC estimation results are shown in Figure 13, with evaluation metrics summarised in Table 6.

In this work, the GPR-UKF with adaptive parameters ($\mathbf{Q}$ and $\mathbf{R}$) serves as the baseline method. Percentage values in parentheses indicate the relative improvement between the baseline GPR-UKF (Adaptive $\mathbf{Q}$, $\mathbf{R}$) method and each comparison method, with the value representing the improvement of the better performing method. For error metrics (MSE, RMSE, MAE, MAPE), a down arrow (↓) indicates that the proposed baseline method achieves better performance (i.e., lower error) than the comparison method, while an up arrow (↑) indicates that the comparison method performs better (i.e., has lower error) than the baseline. For R^2, up arrow (↑) indicates that the baseline achieves a higher R^2 than the comparison method, while a down arrow (↓) means that the comparison method achieves a higher R^2 than the baseline. The best performance (lowest error or highest R^2) among all methods is indicated in bold.

The GPR-UKF approach, incorporating adaptive parameters ($\mathbf{Q}$ and $\mathbf{R}$), demonstrates superior stability and accuracy in SOC estimation compared to the fixed-parameter GPR-UKF approach and three compared machine-learning-based approaches. Specifically, the adaptive GPR-UKF reduces the MSE for SOC estimation to 0.2131, significantly lower than the values observed for the three machine-learning algorithms and that for the fixed-parameter counterpart. For battery temperature estimation, the conventional GPR method distinctly outperforms both GPR-UKF variants, achieving approximately half of the MSE for the GPR-UKF approaches and less than 10% of the MSE values recorded for XGBoost and LSTM methods. However, as shown in Figures 13(c) and (d), prediction errors for all methods progressively converge toward zero, demonstrating each model's ability to refine state estimates iteratively.

Table 7
Optimised Hyperparameters for the GPR-based state prediction model under the DST scenario.

m	$\sigma^2_{SE,s,m}$	$l_{s,m}$	$\sigma^2_{\mathrm{Lin},s,m}$	$\sigma^2_{0,s,m}$	σ^2_s
SOC	3.84×10^{2}	7.73×10^{1}	5.16×10^{-8}	3.33×10^{-1}	1.05×10^{-2}
T	4.39×10^{-5}	2.16×10^{3}	4.39×10^{-2}	5.22	
I	1.00×10^{-5}	7.12×10^{1}	1.00×10^{-2}	1.10×10^{2}	

Table 8
Optimised Hyperparameters for the GPR-Based observation model under the DST scenario.

m	$\sigma^2_{SE,o,m}$	$l_{o,m}$	$\sigma^2_{\mathrm{Lin},o,m}$	$\sigma^2_{0,o,m}$	σ^2_o
SOC	2.54	4.81	0.23	4.62	1.30×10^{-2}
T	2.54	4.81	0.23	4.62	
I	2.54	4.81	0.23	4.62	

Table 9
Evaluation of GPR-based state prediction and observation models using Cell 1 and Cell 2 test data for the DST case

Evaluation	Cell1 Test Data				Cell2 Test Data			
	SOC	T	ε	V	SOC	T	ε	V
MSE	0.4920	0.0709	0.0001	0.0088	0.6759	0.0862	0.0176	0.0030
RMSE	0.7015	0.2663	0.0106	0.0938	0.8222	0.2936	0.1326	0.0549
MAE	0.6092	0.2222	0.0085	0.0333	0.7174	0.2292	0.1241	0.0289
R^2	0.9991	0.8366	0.9806	0.9585	0.9987	0.8058	−5.1795	0.9858
MAPE (%)	1.5909	0.8068	3.2123	0.4695	2.9700	0.8400	83.4000	0.4000

6.4.2. Case 2: DST test validation

The optimised hyperparameters for the GPR-based prediction and observation models under the DST scenario are provided in Table 7 and Table 8. The matrix $\mathbf{Q}$ for this scenario is presented in Eq. (39), while the matrix $\mathbf{R}$, estimated under the same laboratory conditions, is given in Eq. (38).

$$\mathbf{Q} = \begin{bmatrix} 4.58\times10^{-1} & 0 \\ 0 & 6.83\times10^{-2} \end{bmatrix} \tag{39}$$

The performance of the GPR-based models is illustrated in Figure 14, with key evaluation metrics summarised in Table 9. Compared to the CC discharging scenario, the

Table 10
Performance metrics for SOC and temperature prediction with five methods under DST conditions.

Metric	GPR-UKF (Adaptive Q,R)		GPR-UKF (Fixed Q,R)		GPR		XGBoost		LSTM	
	SOC	T	SOC	T	SOC	T	SOC	T	SOC	T
MSE	**0.8341**	0.0629	10.2637 (91.87%↓)	0.2722 (76.89%↓)	328.9966 (99.75%↓)	0.2933 (78.55%↓)	8.0838 (89.68%↓)	**0.0588** (6.97%↑)	29.7694 (97.20%↓)	0.1253 (49.80%↓)
RMSE	**0.9133**	0.2508	3.2037 (71.49%↓)	0.5217 (51.93%↓)	18.1383 (94.96%↓)	0.5416 (53.69%↓)	2.8432 (67.88%↓)	**0.2425** (3.42%↑)	5.4561 (83.26%↓)	0.3540 (29.15%↓)
MAE	**0.6068**	0.2075	2.5669 (76.36%↓)	0.4637 (55.25%↓)	15.5110 (96.09%↓)	0.3712 (44.10%↓)	1.8964 (68.00%↓)	**0.1542** (34.57%↑)	2.9474 (79.41%↓)	0.1996 (3.96%↑)
R^2	**0.9984**	0.8583	0.9802 (1.86%↑)	0.3865 (122.02%↑)	0.3655 (173.20%↑)	0.3389 (153.26%↑)	0.9844 (1.42%↑)	**0.8675** (1.06%↓)	0.9426 (5.92%↑)	0.7175 (19.62%↑)
MAPE (%)	**1.28**	0.76	8.63 (85.17%↓)	1.69 (55.03%↓)	54.96 (97.67%↓)	1.36 (44.12%↓)	10.18 (87.43%↓)	**0.56** (35.71%↑)	22.45 (94.30%↓)	0.71 (7.04%↑)

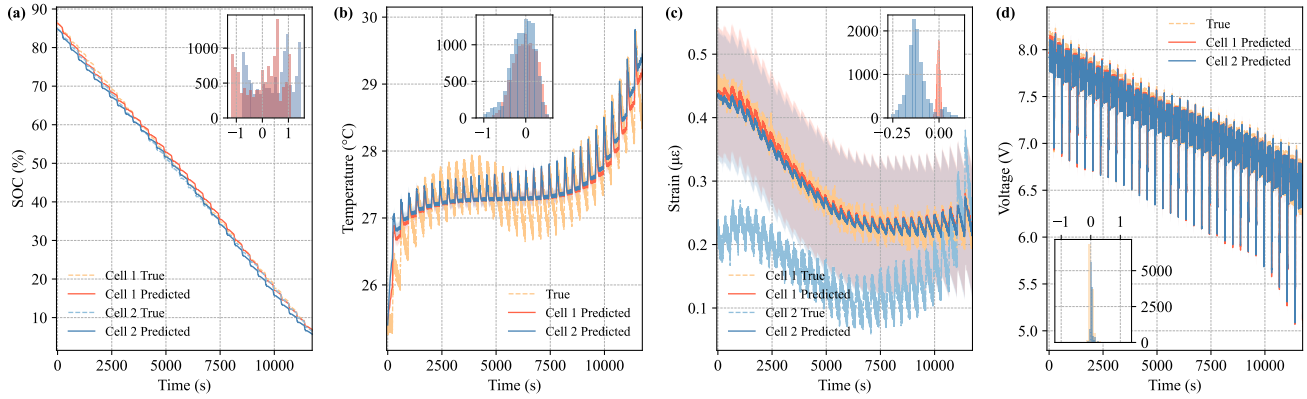

Figure 14: Evaluation of GPR-based prediction and observation models under the DST scenario: (a) Cell SOC; (b) battery pack average temperature; (c) cell strain; (d) battery pack terminal voltage. Insets illustrate the error distributions.

state prediction model under DST conditions exhibits higher errors across both Cell 1 and Cell 2 datasets. For example, the R^2 value for temperature predictions is approximately 0.80 for both cells, considerably lower than the near 0.99 values observed in previous cases.

The observation model effectively captures the variation patterns of ε in the Cell 1 test dataset. However, due to significant discrepancies between the true ε values of Cell 1 and Cell 2, the model trained on Cell 1 data generalizes poorly to Cell 2. As shown in Figure 14(c), Cell 2 predictions exhibit concentrated errors within the range of -0.25 to 0.1$\mu\varepsilon$, resulting in a MAPE of 83.4% and a negative R^2 value of -5.1795, indicating poor predictive capability. Furthermore, predictions for both cells' test datasets also

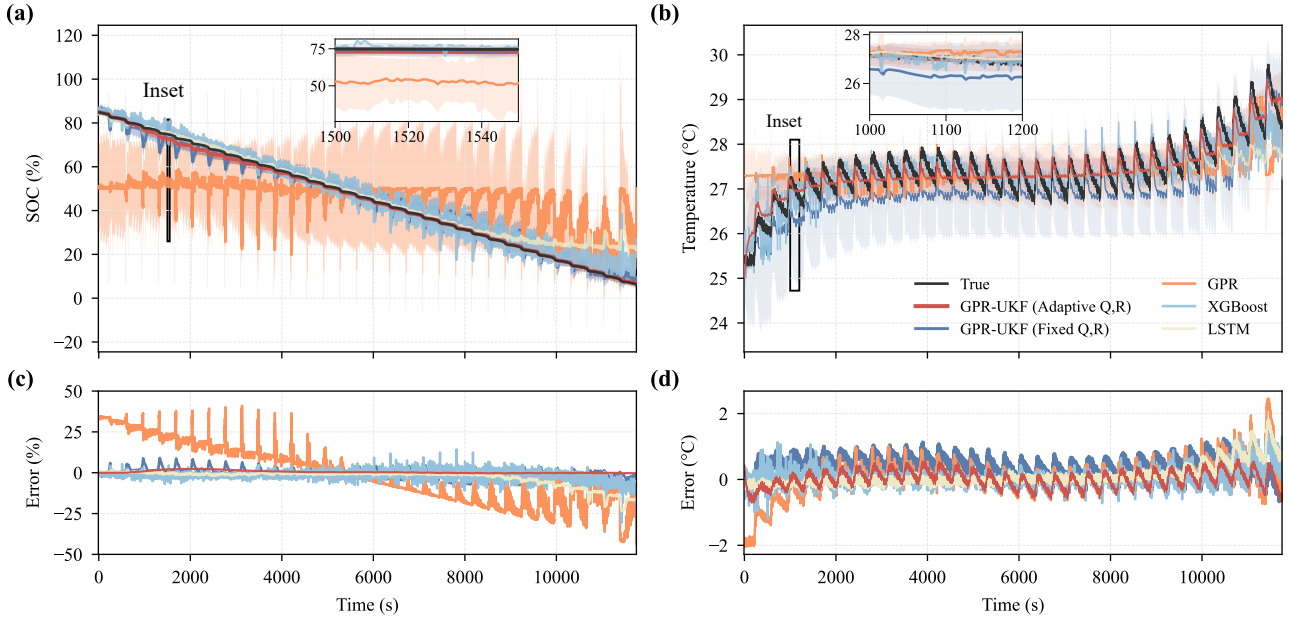

Figure 15: Battery pack SOC and average temperature under the DST scenario: (a) SOC estimation; (b) average temperature estimation; (c) SOC estimation error; (d) average temperature estimation error.

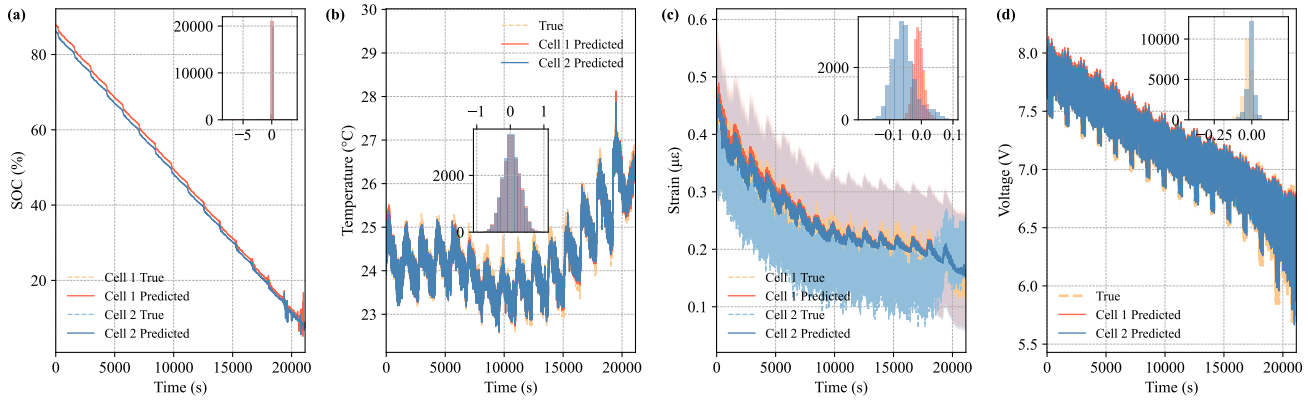

Figure 16: Evaluation of GPR-based prediction and observation models under the FUDS scenario: (a) Cell SOC; (b) battery pack average temperature; (c) cell strain; (d) battery pack terminal voltage. Insets illustrate the error distributions.

exhibit noticeable uncertainty, as illustrated in Figure 14(c). In contrast, the predictions of pack terminal voltage demonstrate high accuracy and reliability with limited uncertainty, as shown in Figure 14(d).

In the cycling dataset used for validation, Cell 2 was selected as the representative cell based on SCS analysis from the same CC-CV charging cycle. The estimation results are presented in Figure 15, and the evaluation metrics are summarised in Table 10.

The GPR-UKF with adaptive parameters demonstrates superior estimation accuracy, as illustrated in Figure 15. In contrast, the GPR-UKF model with fixed parameters exhibits higher estimation errors and greater prediction uncertainties in both SOC and temperature estimations. The standard GPR model under DST conditions struggles to capture SOC and temperature variation patterns effectively, resulting to underfitting. For example, the SOC estimation under DST conditions achieves an MSE of 0.8341 with the adaptive GPR-UKF, compared to 10.2637 for the fixed-parameter approach and 328.9966 for the standard GPR model. On the other hand, the XGBoost prediction curve exhibits pronounced high-frequency fluctuations, and both XGBoost and LSTM predictions substantially diverge from the actual measurements at approximately 10,000 s.

Both XGBoost and the adaptive GPR-UKF method achieve high accuracy in temperature estimation, obtaining MSE values around 0.06 and R^2 exceeding 0.85. A substantial improvement is evident in Table 10 when employing the adaptive GPR-UKF method compared with the standard GPR-based state prediction model and the fixed-parameter GPR-UKF. Moreover, the adaptive GPR-UKF demonstrates superior performance in both SOC and temperature estimation for the battery pack under DST conditions compared other evaluated approaches.

6.4.3. Case 3: FUDS test validation

The optimised hyperparameters for both the GPR-based state and observation models under FUDS conditions are summarised in Table 11 and Table 12, respectively. The corresponding process noise covariance matrix, $\mathbf{Q}$, is given in Eq. (40), while the observation noise covariance, $\mathbf{R}$, remains as defined in Eq. (38).

Table 11
Optimised Hyperparameters for the GPR-based state prediction model under the FUDS scenario.

m	$\sigma^2_{SE,s,m}$	$l_{s,m}$	$\sigma^2_{\text{Lin},s,m}$	$\sigma^2_{0,s,m}$	σ^2_s
SOC	2.44×10^{-1}	2.01×10^{2}	6.53×10^{-3}	3.81×10^{-4}	
T	1.31×10^{-2}	8.27×10^{2}	1.52×10^{-3}	3.37×10^{1}	3.98×10^{-2}
I	5.99×10^{2}	3.74×10^{3}	6.67×10^{-3}	3.02×10^{-4}	

Table 12
Optimised hyperparameters for the GPR-based observation model under the FUDS scenario.

m	$\sigma^2_{SE,o,m}$	$l_{o,m}$	$\sigma^2_{\text{Lin},o,m}$	$\sigma^2_{0,o,m}$	σ^2_o
SOC	2.06×10^{-3}	6.54×10^{-1}	2.77×10^{-2}	2.73×10^{-1}	
T	6.39×10^{-4}	1.13×10^{2}	4.90×10^{-3}	1.88×10^{1}	3.08×10^{-2}
I	1.44×10^{0}	8.89×10^{-5}	1.47×10^{1}	4.23×10^{1}	

Table 13
Evaluation of GPR-based state prediction and observation models using Cell 1 and Cell 2 test data for the FUDS case

Evaluation	Cell1 Test Data				Cell2 Test Data			
	SOC	T	ε	V	SOC	T	ε	V
MSE	0.0137	0.0651	0.0003	0.0017	0.0061	0.0642	0.0044	0.0009
RMSE	0.1170	0.2551	0.0185	0.0412	0.0780	0.2533	0.0663	0.0294
MAE	0.0152	0.2027	0.0150	0.0314	0.0125	0.2013	0.0595	0.0191
R^2	1.0000	0.9140	0.9351	0.9898	1.0000	0.9152	-0.2896	0.9948
MAPE (%)	0.10	0.83	6.78	0.45	0.08	0.83	34.08	0.27

$$\mathbf{Q} = \begin{bmatrix} 0.63 & 0 \\ 0 & 0.27 \end{bmatrix} \tag{40}$$

The performance of the GPR-based models is illustrated in Figure 16, with the corresponding quantitative metrics summarised in Table 13. Despite the FUDS cycle exhibiting stronger nonlinear dynamics and frequent current reversals than DST, the state model trained on Cell 1 generalises well to FUDS and, overall, performs better than under DST.

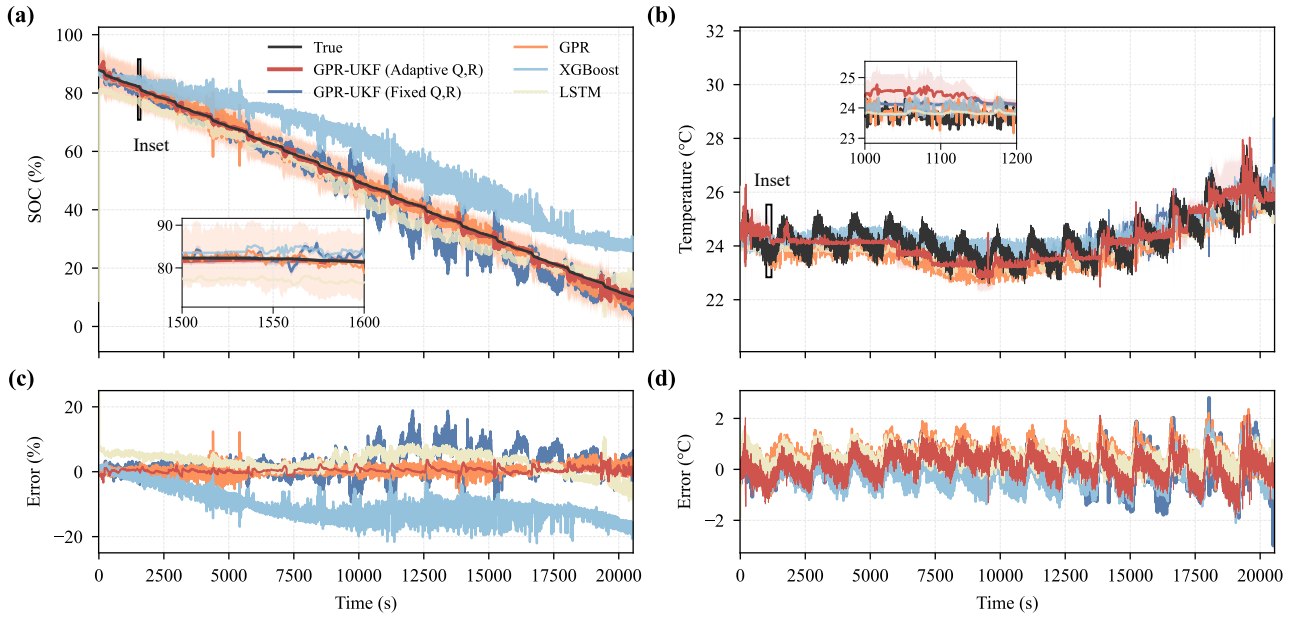

Figure 17: Battery pack SOC and average temperature under the FUDS scenario: (a) SOC estimation; (b) average temperature estimation; (c) SOC estimation error; (d) average temperature estimation error.

Table 14
Performance metrics for SOC and temperature prediction with five methods under FUDS conditions.

Metric	GPR-UKF (Adaptive Q,R)		GPR-UKF (Fixed Q,R)		GPR		XGBoost		LSTM	
	SOC	T	SOC	T	SOC	T	SOC	T	SOC	T
MSE	**0.7359**	0.2663	20.7858 (2724.54%↓)	0.3066 (15.13%↓)	2.0137 (173.64%↓)	0.3239 (21.63%↓)	129.8202 (17541.01%↓)	**0.2100** (26.81%↑)	18.1232 (2362.73%↓)	0.2617 (1.76%↑)
RMSE	**0.8579**	0.5160	4.5591 (431.43%↓)	0.5537 (7.31%↓)	1.4190 (65.40%↓)	0.5691 (10.29%↓)	11.3939 (1228.12%↓)	**0.4582** (12.61%↑)	4.2571 (396.22%↓)	0.5116 (0.86%↑)
MAE	**0.6396**	0.4204	3.2559 (409.05%↓)	0.4501 (7.06%↓)	1.1185 (74.87%↓)	0.4537 (7.92%↓)	10.4496 (1533.77%↓)	**0.3704** (13.50%↑)	3.7136 (480.61%↓)	0.4196 (0.19%↑)
R^2	**0.9985**	0.6078	0.9585 (4.01%↑)	0.5484 (9.77%↑)	0.9960 (0.25%↑)	0.5230 (13.95%↑)	0.7410 (25.79%↑)	**0.6907** (12.00%↓)	0.9638 (3.48%↑)	0.6145 (1.09%↓)
MAPE (%)	**1.9976**	1.7207	10.0855 (404.88%↓)	1.8445 (7.19%↓)	3.4896 (74.69%↓)	1.8528 (7.68%↓)	33.6579 (1584.92%↓)	**1.5237** (12.93%↑)	9.9478 (397.99%↓)	1.7108 (0.58%↑)

When evaluated on Cell 2, the model achieves lower prediction errors than on Cell 1, suggesting that it captures generalisable electrothermal dynamics across cells.

The observation model for strain (ϵ) reveals a distinct behaviour. As shown in Figure 16(c), the predicted strain for Cell 2 follows the trend learnt from Cell 1 but exhibits a systematic offset from the ground truth, as confirmed by the shifted error distribution in the inset. This bias arises from the training data being limited to a single cell.

Following SCS analysis, Cell 2 was chosen as the representative unit for pack-level evaluation. Figure 17 presents the results of pack SOC and temperature estimation under FUDS conditions, with the performance metrics summarised in Table 14. The adaptive GPR–UKF method delivers the most accurate SOC estimation throughout the entire profile, maintaining tight confidence intervals and stable tracking except for moderate increases in residuals during rapid current transitions and more pronounced fluctuations near 20,000 s. This method achieves the lowest overall estimation errors among all approaches (MSE = 0.7359, RMSE = 0.8579, MAE = 0.6396, R^2 = 0.9985, MAPE = 2.00%). In comparison, the standard GPR approach is able to capture the overall SOC trend, while exhibiting greater fluctuations and larger posterior standard deviation bands, particularly during dynamic periods. Both the GPR-UKF (fixed-Q, R) and LSTM methods result in significantly higher SOC errors (MSE = 20.7858 and 18.1232, respectively), while the XGBoost model performs the poorest, with markedly larger deviations (MSE = 129.8202, RMSE = 11.3939, MAE = 10.4496, R^2 = 0.7410, MAPE = 33.66%).

For temperature estimation, XGBoost achieves the lowest prediction errors across all metrics (MSE = 0.2100, RMSE = 0.4582, MAE = 0.3704, R^2 = 0.6907, MAPE = 1.52%), as summarised in Table 14. The GPR-UKF (Adaptive Q,R) demonstrates a higher overall estimation error relative to XGBoost, with an MSE of 0.2663 compared to 0.2100 for XGBoost. Figure 18(b) and (d) further reveal that, in the latter stages of the cycle, the GPR-UKF (Adaptive Q,R) exhibits increasing residual fluctuations and a progressive broadening of the associated standard deviation bands. This pattern is likely attributable to the reduced accuracy of the observation model during these periods, as indicated in Figure 16(c)(d). In contrast, the standard GPR and LSTM methods yield larger errors and greater temporal variability, as indicated by both the numerical results and error profiles.

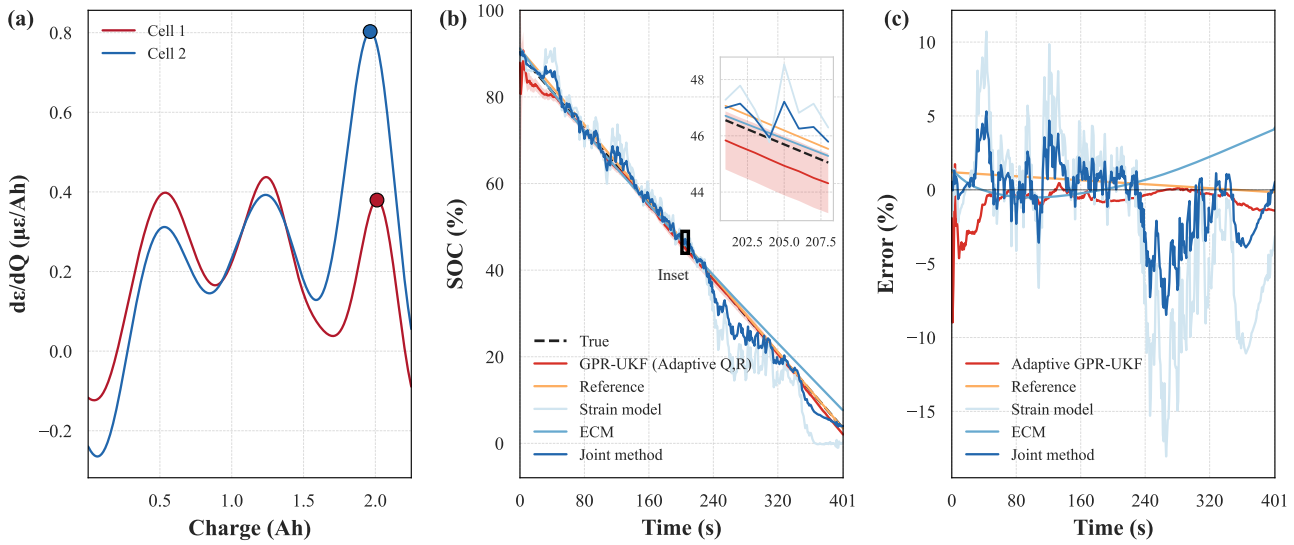

Figure 18: Results for the $LiFePO_4$ battery pack: (a) SCS results using the extracted charging process data; (b) SOC estimations during the discharging process; (c) estimation errors compared with the true values.

Table 15
Comparison of prediction accuracy for different SOC estimation methods compared to the representative cells confirmed by SCS.

Method	MSE	RMSE	MAE	R^2	MAPE (%)
GPR-UKF (Adaptive Q,R)	**1.4683**	**1.2117**	**0.7585**	**0.9977**	**3.1675**
Strain model	36.5278 (95.98%↓)	6.0438 (79.95%↓)	4.5689 (83.40%↓)	0.9416 (5.96%↑)	21.6851 (85.39%↓)
ECM	2.6558 (44.71%↓)	1.6297 (25.65%↓)	1.1452 (33.77%↓)	0.9958 (0.19%↑)	8.6768 (63.49%↓)
Joint method	6.4600 (**77.27%**↓)	2.5416 (52.32%↓)	1.8655 (59.34%↓)	0.9897 (0.81%↑)	7.0693 (55.20%↓)

6.4.4. Case 4: Discharge analysis of a $LiFePO_4$ battery pack.

To further evaluate the proposed framework that integrates the pack terminal voltage with the representative cell strain as the model inputs for the adaptive GPR-UKF approach in battery pack SOC estimation, the experimental data used in [54] have also been adopted in this section. In [54], a $LiFePO_4$ battery pack composed of two pouch cells connected in series was equipped with two FBG sensors on each cell for strain and temperature monitoring. The results from the original study indicated that Cell 1 exhibited a lower initial capacity than Cell 2, reported as 2.54 Ah and 2.58 Ah, respectively. The pack reference SOC was extracted from Cell 1 using the Coulomb counting method. In [54], several approaches, including strain-based methods, ECM-based methods, as well as a joint approach have been compared for pack SOC estimation.

In this work, the representative cell was selected by applying the SCS method to the charging process data, as illustrated in Figure 18(a). The half window size ω was set to 225, and both the polynomial order ρ and the number of charge segments were specified as 2 and 1000, respectively. According to the SCS analysis as illustrated in Figure 18(a), Cell 2 exhibits higher final peak values that appear earlier than those of Cell 1. By examining the real SOC calculations for both cells, it is also evident that Cell 2 maintains a lower SOC than Cell 1 in this operating cycle. Hence, updating the selection of the representative cell using the latest charging process data is both reasonable and effective.

A comparison between the estimation results obtained by the proposed framework and those reported in [54] is shown in Figures 18(b) and (c), where the true SOC is given by the SCS selected representative cell and the Reference SOC is the pack-level real SOC used in [54]. In Table 15, the SOC estimation results produced by all methods for the SCS selected representative cell are summarised. Whereas Table 16 are the comparison results against the Reference SOC reported in [54]. The percentage improvements for the GPR-UKF over each specified method are shown in parentheses.

It is evident that the adaptive framework that integrates battery pack measurements achieves a 77.27% reduction in MSE compared with the joint method, as well as improvements of 95.98% and 44.71% over the strain-based and ECM approaches, respectively. The proposed approach attains 0.7585 MAE and an R^2 value of 0.9977 on the $LiFePO_4$ battery pack discharging test dataset.

By comparing with the reference pack SOC reported in [54], it is evident that the proposed framework achieves superior performance relative to all baseline methods presented in [54]. Figure 18(c) shows that, although GPR-UKF with adaptive **Q** and **R** exhibits a slightly larger initial

Table 16
Comparison of prediction accuracy for different SOC estimation methods in relation to the reference pack SOC from [54].

Method	MSE	RMSE	MAE	R^2	MAPE (%)
GPR-UKF (Adaptive Q,R)	**2.8693**	**1.6939**	**1.2477**	**0.9956**	**3.8793**
Strain model	36.2022 (92.08%↓)	6.0168 (71.84%↓)	4.5040 (72.29%↓)	0.9439 (5.48%↑)	21.6098 (82.05%↓)
ECM	3.0519 (5.98%↓)	1.7470 (3.04%↓)	1.3946 (10.53%↓)	0.9953 (0.03%↑)	9.1854 (57.77%↓)
Joint method	6.4952 (55.82%↓)	2.5486 (33.53%↓)	1.8294 (31.80%↓)	0.9899 (0.58%↑)	6.9939 (44.53%↓)

Table 17
Comparison between voltage and FBG sensors for battery pack monitoring

Parameter	Voltage Sensors	FBG Sensors	Technical Significance
Fundamental Measurement Capabilities			
Cell voltage measurement	✓	×	Critical for operational safety boundaries
Overcharge/overdischarge detection	✓	×	Primary safety mechanism
Temperature monitoring	×	✓	Thermal monitor and safety
Mechanical strain/deformation	×	✓	Indicates internal electrochemical processes
Internal pressure variations	×	✓	Early indicator of degradation mechanisms [51]
Microscale mechanical alterations	×	✓	Reveals ageing processes before electrical manifestation [66]
Fault Detection and Prognostic Capabilities			
Acute Failures			
Internal short circuit	△	✓	FBG provides earlier warning [46]
Thermal runaway	△	✓	FBG offers multiparametric monitoring
Connection failure	✓	△	Voltage sensors superior
Progressive Degradation			
SEI layer growth	×	△	Facilitates preventive diagnostics [98]
Lithium plating	△	✓	FBG enables operando Li plating monitoring [99]
Electrode material degradation	△	△	Complementary approaches
Failure prediction timeframe	Short	Extended	FBG enables predictive intervention [66]
System Implementation Characteristics			
Multiple cell monitoring capability	△	✓	FBG reduces system complexity
EMI susceptibility	High	Immune	FBG increases reliability in noisy environments [100]
Wiring complexity	High	Low	FBG reduces installation and maintenance costs
System scalability	Limited	High	FBG advantageous for large-scale systems
SOC estimation methodology	Voltage-based	Strain-based	Complementary approaches enhance accuracy [47]
Economic and Deployment Considerations			
Technology maturity	High	Medium	Voltage sensors have decades of proven deployment
Initial system cost	Low	High	FBG requires specialized interrogation equipment
Per-unit sensor cost	Low	Medium	Volume production advantages for voltage sensors
Installation complexity	Medium	Low	FBG reduces wiring but requires optical expertise
Maintenance requirements	Medium	Low	FBG has fewer electrical connections to maintain
Replacement/repair cost	Low	High	Voltage sensors use standard electronic components
Total cost of ownership	Medium	Medium	Long-term benefits may offset initial FBG investment

Note: ✓ indicates strong capability, △ indicates limited capability, × indicates absence of capability or disadvantage

estimation error of approximately 0.1%, the performance remains robust with consistently small estimation errors even when the strain-based model and ECM show large deviations, for instance from 240 s to 320 s.

7. Discussions

In this section we summarise and discuss the experimental findings using OFS in place of conventional cell-level voltage and temperature sensors, and we evaluate the state estimation framework under static and dynamic operating conditions.

7.1. Representative cell selection using SCS analysis

The analysis at the cell level where the battery cells are charged with different C-rates, and the results reveal how charging rate and cell temperature critically influence the formation and magnitude of metallurgical phase transition peaks in the strain–charge curves. As shown in Figures 9, lower charging rates accentuate multiple peaks due to more pronounced strain evolution in graphite, while higher charging rates yield larger early peaks driven by both thermal expansion and faster lithiation. This cell-level insight was then extended to the pack level (Figure 10), where two series-connected cells exhibited distinct ageing behaviours, particularly in the positioning and growth of the second peak

in the charge-normalised strain curves over repeated cycling tests. A key observation is that the more aged cell consistently displayed a higher second peak at earlier charging cycles, supporting the use of these profiles for representative cell selection and for indicating capacity fade and thermal effects. Collectively, these results confirm that applying the SCS at both cell and pack levels can offer a comprehensive description of the internal mechanical and electrochemical changes, offering a great diagnostic potential of the SCS approach for battery health monitoring.

Compared to traditional voltage-based sensing and cell selection approaches [29, 30, 31, 32, 35, 37], the FBG-based system combined with SCS analysis presents several noteworthy advantages for battery pack management. While voltage sensors remain indispensable for operational safety, particularly for detecting overcharge and overdischarge events, they are inherently limited in their ability to capture the early mechanical changes that underpin long-term cell degradation. In contrast, FBG sensors offer a direct window into temperature, strain, and internal pressure, enabling earlier detection of subtle degradation mechanisms such as SEI layer growth and lithium plating, often before these manifest electrically. This heightened sensitivity extends the timeframe available for fault prediction and preventive intervention, which is especially valuable for maintaining pack reliability over extended cycles. Additionally, the optical nature of FBG sensors simplifies system wiring and enhances immunity to electromagnetic interference, thereby supporting more scalable and robust monitoring architectures in large battery installations. However, it is important to recognise that FBG-based approaches do not replace the essential role of voltage monitoring in ensuring cell-level safety, but rather complement it by providing richer diagnostic information and enabling a shift towards more predictive and condition-based maintenance strategies. The integration of SCS-guided representative cell selection within an FBG-assisted framework thus marks a significant step forward in advancing battery health diagnostics, while ongoing work is needed to fully address the challenges of operational safety and fault detection in large-scale deployments. Table 17 summarise the comparative analysis of battery pack monitoring technologies.

7.2. Adaptive GPR-UKF for pack level state estimation

Consolidating voltage and strain measurements for the battery pack enables more accurate data-driven estimation of the battery pack SOC and temperature. A set of comparison experiments were conducted, and the proposed the adaptive GPR-UKF, the fixed parameter GPR-UKF, and three commonly utilised machine learning methods were tested under various scenarios.

The comparison results reveal distinct strengths and limitations of these approaches when they were applied under different operating conditions. The open-loop standard GPR model achieved high accuracy under static conditions while degraded in nonlinear scenarios, largely because constraints in the training dataset and the chosen kernel hindered its ability to capture complex dynamic behaviour. The LSTM approach exhibited limited advantages across all test scenarios, attributed to insufficient training data for effective neural network optimization.

A set of quantitative metrics were used to assess the performance of different state estimation approaches across a raneg of test conditions, the results confirm the efficacy of the adaptive GPR-UKF framework. Tables 6, 10, and 14 reveal that the adaptive GPR-UKF approach consistently outperforms alternatives under dynamic operational scenarios. For SOC estimation under DST conditions, the adaptive GPR-UKF achieves an MSE of 0.8341, significantly lower than fixed-parameter GPR-UKF (10.2637), XGBoost (8.0838), and LSTM (29.7694). The most substantial performance differences were observed during the CC discharging phase, where the adaptive model maintains an RMSE of 0.4617 for SOC estimation, representing improvements of 67.72%, 78.49%, 85.19%, and 82.62% compared to fixed-parameter GPR-UKF, standard GPR, XGBoost, and LSTM, respectively. Although individual methods occasionally outperform the adaptive approach in specific conditions, the adaptive GPR-UKF framework demonstrates superior robustness across a range of varying operation conditions, with the dynamic covariance adaptation mechanism being proved to be particularly useful under highly variable load conditions.

To further confirm the effectiveness of the proposed approach in handling different battery packs, another case study was conducted using the dataset from [54], which features a pack of two $LiFePO_4$ pouch cells, different from the first set of experiments where the battery cells are NCR 18650 cylindrical lithium-ion ones. The proposed pack SOC estimation framework was compared with the results reported in [54]. The results demonstrate that updating the representative cell in the pack using SCS during cycling is effective, yielding outstanding SOC estimation when employing the proposed GPR-UKF with adaptive $\mathbf{Q}$ and $\mathbf{R}$. Application of the proposed methodology resulted in a 77.27% improvement in MSE compared to the joint method reported in [54].

The superior performance of the GPR-UKF with adaptive parameters stems from its dynamic parameter-tuning capability, facilitating real-time adjustments of the process and measurement noise covariances ($\mathbf{Q}$ and $\mathbf{R}$). Figure 19 illustrates the evolution of key UKF parameters under the DST scenario. The first column depicts the adaptive GPR-UKF results, including state prediction covariance $\mathbf{P}_k^{\mathbf{X}^-}$ (a1), innovation covariance $\mathbf{S}_k$ (b1), cross-covariance $\mathbf{P}_k^{\mathbf{X}^-\mathbf{Y}^-}$ (c1), and Kalman gain matrix $\mathbf{K}_k$ (d1). For comparison, the second column presents the corresponding parameters for the fixed-parameter GPR-UKF.

In Figure 19(a1), the state prediction covariance indicates higher prior uncertainty for SOC predictions, while temperature prediction covariance remains relatively stable. This reflects greater model confidence in temperature

predictions than in SOC estimates. The innovation covariance $\mathbf{S}_k$, shown in Figure 19(b1), incorporates contributions from predicted observation covariance $\mathbf{P}_k^{\mathbf{Y}^-}$ and measurement noise covariance $\mathbf{R}$. The calculated $\mathbf{S}_k$ for strain (ε) and voltage (V) hovers around 1.0, with near-zero cross-covariance between the two, suggesting that the model treats these measurements as independent.

Figure 19(c1) illustrates the prior cross-covariance matrix between states and observations. A relatively higher positive cross-covariance $\mathbf{P}(SOC, V)$ suggests a positive correlation between SOC and voltage during the tracking process, indicative of the influence of SOC on voltage behaviour as battery load changes. In contrast, SOC exhibits a weaker relationship with strain (ε) under DST conditions. Beyond that, a slight negative cross-covariance can be observed between temperature and voltage, which is consistent with the observed DST behaviour where the overall voltage decreases while temperature generally rises over time.

The Kalman gain matrix $\mathbf{K}_k$ values in Figure 19(d1) highlight the influence of observations on state updates. Due to higher $\mathbf{P}(SOC, SOC)$ values, $\mathbf{K}(SOC, V)$ exhibits a significant weighting, indicating that voltage observations strongly influence SOC updates during the UKF update phase. In contrast, temperature updates rely less on observations, as the model demonstrates greater confidence in its temperature predictions. Voltage observations also exert a corrective effect on temperature updates, with the UKF applying a small inverse correction factor, further showcasing the adaptive nature of the GPR-UKF.

For the fixed-parameter GPR-UKF, reduced performance is linked to the static noise covariance settings. As shown in Figure 19(a2), a high $\mathbf{Q}$ increases prior state covariance $\mathbf{P}_k^{\mathbf{X}^-}$, indicating reduced confidence in state predictions. Simultaneously, a low $\mathbf{R}$ decreases the innovation covariance $\mathbf{S}_k$, leading to an increased Kalman gain $\mathbf{K}_k$ (Figures 19(b2) and 19(d2)). This configuration causes the filter to over-rely on measurements, resulting in greater errors and uncertainty when significant model residuals are present.

8. Conclusion

Battery packs comprising hundreds or thousands of cells arranged in series and/or parallel configurations are vital for the operation of EVs and ESSs. Ensuring their safety and reliability requires real-time monitoring of critical parameters, such as temperature, overcharging, and over-discharging. Traditional approaches that model each individual cell achieve high accuracy but are computationally prohibitive due to the inherent complexity and variability within the pack. Treating the entire battery pack as a single entity reduces computational demands yet fails to capture inconsistencies arising from manufacturing and operational variations.

This study has introduced a novel approach for battery pack state estimation that has successfully addressed the computational challenges of traditional methods by replacing extensive electric sensor networks with a much more compact network of OFSs capable of monitoring entire cells. A key contribution lies in the development of a representative cell methodology that achieves accurate pack state estimation through the selection of representative cells and estimate their states rather than estimating the state of every single cells. The proposed battery pack state estimation approach first employs OFSs to capture strain signals from battery cells, and a SCS methodology is then proposed to identify representative cells by analysing distinctive derivative peaks in strain-charge relationships. The SCS analysis enables early detection of capacity fade and thermal effects by revealing cell ageing based on their characteristic strain-charge behaviours, offering a complementary diagnostic tool to conventional state monitoring approaches using electric signals. Furthermore, the framework incorporates an adaptive GPR-based UKF that dynamically adjusts the process and measurement noise covariances based on state-dependent factors such as training data variability and model reliability. The probabilistic nature of GPR effectively accommodates cell-to-cell heterogeneity and ageing effects, incorporating uncertainty quantification directly into the filtering process to improve estimation reliability under diverse operating conditions.

Experimental results have demonstrated the exceptional SOC estimation accuracy of the proposed method across a range of scenarios and different battery cell chemistries. For the experimental study on a battery pack composed of NCR18650 cylindrical cells, the adaptive GPR-UKF framework has achieved accurate SOC estimation with RMSE of 0.4617 and MAPE of 0.82% under CC discharging conditions, RMSE of 0.9133 and MAPE of 1.28% under dynamic DST conditions, and RMSE of 0.8579 and MAPE of 2.00% under FUDS conditions. Furthermore, under DST conditions, the framework outperforms fixed-parameter GPR-UKF by 71.49%, standard GPR by 94.96%, XGBoost by 67.88%, and LSTM by 83.26% in terms of the RMSE metric. Additional experiments using a battery pack composed of $LiFePO_4$ pouch cells have also demonstrated that the proposed method has maintained high state estimation accuracy with RMSE of 1.2117 and MAPE of 3.17% when the representative cells are identified through SCS analysis, compared to the state estimation results using the reference pack data from the literature with RMSE of 1.6939 and MAPE of 3.88%. The consistently low error rates across different pack configurations and operating conditions confirm the the effectiveness of the proposed framework in a wide range of application scenarios.

The future work will focus on the development of dynamic threshold adaptation strategies based on the SCS analysis results under different operating conditions, enabling more accurate representative cell selection for large-scale battery packs. Furthermore, feature engineering and GPR optimization can be adopted to further improve the computational efficiency in meeting the real-time requirements for grid-level energy storage applications.

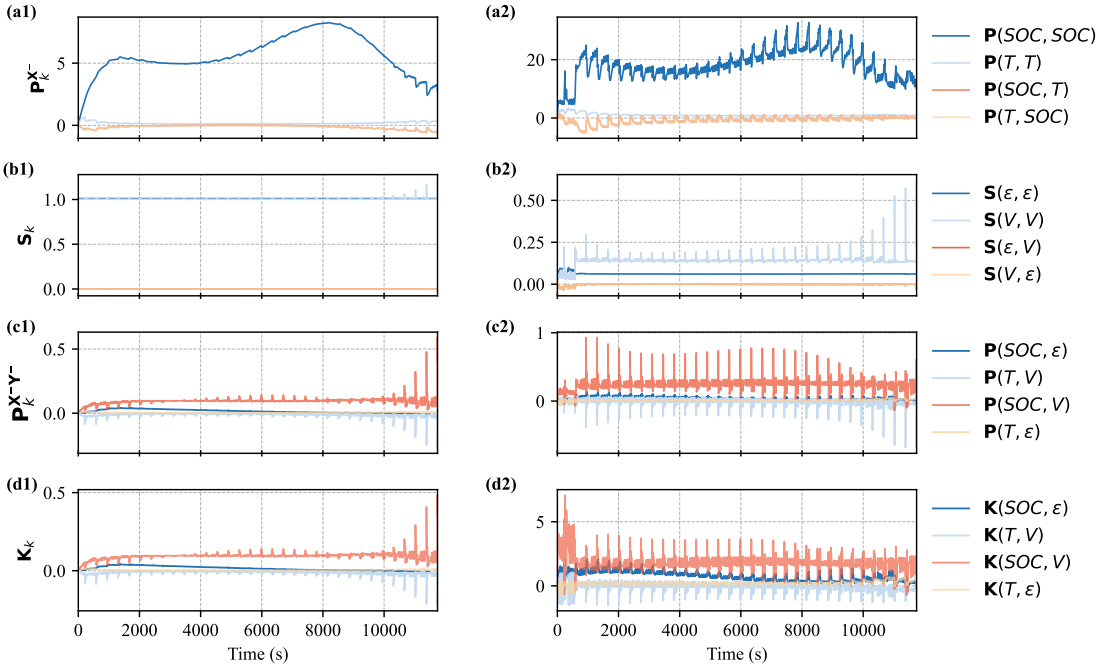

Figure 19: Time evolution of covariance and Kalman gain matrix elements in the DST case. (a) State prediction covariance $\mathbf{P}_k^{\mathbf{X}^-}$, (b) innovation covariance $\mathbf{S}_k$, (c) cross-covariance $\mathbf{P}_k^{\mathbf{X}^-\mathbf{Y}^-}$, and (d) Kalman gain matrix $\mathbf{K}_k$ elements over time.

CRediT authorship contribution statement

Shiyun Liu: Conceptualization, Investigation, Methodology, Visualization, Writing – original draft. **Kang Li:** Conceptualization, Methodology, Supervision, Writing - Review and Editing. **James Yu:** Investigation, Conceptualization.

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 Ofgem funded project 'Resilient and Flexible Multi-Energy Hub Networks for Integrated Green Mobility' and the EPSRC funded project 'Networked Energy Hubs for Accelerating Decarbonization of Transport and Energy Sector (UKRI1779)'

Shiyun Liu would like to thank the University of Leeds and SP Energy Networks for sponsoring his research.

A. Appendix

A.1. XGBoost Implementation Details

The Extreme Gradient Boosting (XGBoost) model was implemented using a multi-output regression framework to simultaneously predict SOC and temperature.

A.1.1. Mathematical Formulation

XGBoost constructs an ensemble of K regression trees, where the final prediction is given by:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(\mathbf{x}_i) \tag{41}$$

where f_k represents the k-th tree and $\mathbf{x}_i = [\varepsilon_i, V_i, I_i]^T$ is the input feature vector consisting of strain, voltage, and current.

The objective function minimized during training is:

$$\mathcal{L} = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{42}$$

where l is the squared error loss function:

$$l(y_i, \hat{y}_i) = (y_i - \hat{y}_i)^2 \tag{43}$$

and $\Omega(f_k)$ is the regularization term:

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 + \alpha \sum_{j=1}^{T} |w_j| \tag{44}$$

where T is the number of leaves in tree f_k, w_j is the weight of leaf j, and γ, λ, and α are regularization parameters.

For multi-output prediction, a multi-output wrapper trains separate XGBoost models internally for each target

Table 18
XGBoost hyperparameters used in the experiments

Parameter	Value
Number of estimators (K)	1000
Maximum tree depth	6
Learning rate (η)	0.1
Subsample ratio	0.8
Column subsample ratio	0.8
L1 regularization (α)	0.1
L2 regularization (λ)	1.0
Random state	42
Number of parallel threads	-1 (all cores)

variable:

$$\begin{aligned}\hat{y}_{SOC} &= \sum_{k=1}^{K} f_k^{(SOC)}(\mathbf{x}) \\ \hat{y}_T &= \sum_{k=1}^{K} f_k^{(T)}(\mathbf{x})\end{aligned} \tag{45}$$

A.1.2. Feature Standardization

Input features are standardized before training:

$$\tilde{x}_j = \frac{x_j - \mu_j}{\sigma_j} \tag{46}$$

where μ_j and σ_j are the mean and standard deviation of feature j.

A.1.3. Hyperparameter Settings

The XGBoost hyperparameters used in the experiments are provided in Table 18.

A.2. LSTM Network Architecture and Training Settings

Long Short-Term Memory (LSTM) networks were implemented to capture temporal dependencies in battery state estimation by processing sequences of voltage, current, and strain measurements.

A.2.1. LSTM Cell Equations

The LSTM cell operations at time step t are defined by:

$$\mathbf{f}_t = \sigma(\mathbf{W}_f \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f) \quad \text{(forget gate)} \tag{47}$$

$$\mathbf{i}_t = \sigma(\mathbf{W}_i \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i) \quad \text{(input gate)} \tag{48}$$

$$\tilde{\mathbf{C}}_t = \tanh(\mathbf{W}_C \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_C) \quad \text{(candidate values)} \tag{49}$$

$$\mathbf{C}_t = \mathbf{f}_t * \mathbf{C}_{t-1} + \mathbf{i}_t * \tilde{\mathbf{C}}_t \quad \text{(cell state)} \tag{50}$$

$$\mathbf{o}_t = \sigma(\mathbf{W}_o \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_o) \quad \text{(output gate)} \tag{51}$$

$$\mathbf{h}_t = \mathbf{o}_t * \tanh(\mathbf{C}_t) \quad \text{(hidden state)} \tag{52}$$

where σ denotes the sigmoid function, $*$ denotes element-wise multiplication, and $\mathbf{W}$ and $\mathbf{b}$ are weight matrices and bias vectors, respectively.

A.2.2. Temporal Sequence Modelling

For sequence modelling with window size $L = 60$, the input sequence is constructed as:

$$\mathbf{X}_{seq} = [\mathbf{x}_{t-L+1}, \mathbf{x}_{t-L+2}, ..., \mathbf{x}_t] \tag{53}$$

where each $\mathbf{x}_i = [\varepsilon_i, V_i, I_i]^T$ represents the feature vector at time step i.

The two-layer LSTM architecture processes the sequence as:

$$\mathbf{H}^{(1)} = \text{LSTM}_1(\mathbf{X}_{seq}) \in \mathbb{R}^{L\times 50} \tag{54}$$

$$\mathbf{H}^{(1)'} = \text{Dropout}(\mathbf{H}^{(1)}, p = 0.2) \tag{55}$$

$$\mathbf{h}^{(2)} = \text{LSTM}_2(\mathbf{H}^{(1)'}) \in \mathbb{R}^{50} \tag{56}$$

$$\mathbf{h}^{(2)'} = \text{Dropout}(\mathbf{h}^{(2)}, p = 0.2) \tag{57}$$

$$\hat{\mathbf{y}} = \mathbf{W}_{out}\mathbf{h}^{(2)'} + \mathbf{b}_{out} \tag{58}$$

where $\hat{\mathbf{y}} = [\hat{y}_{SOC}, \hat{y}_T]^T$ is the output vector containing predictions for SOC and temperature.

A.2.3. Loss Function and Training

The model is trained by minimizing the mean squared error loss:

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N} \|\mathbf{y}_i - \hat{\mathbf{y}}_i\|^2 \tag{59}$$

using the Adam optimizer with default parameters.

A.2.4. Data Preprocessing

Similar to XGBoost, input features and output targets are standardized:

$$\tilde{x}_j = \frac{x_j - \mu_j}{\sigma_j} \tag{60}$$

where μ_j and σ_j are the mean and standard deviation of feature j computed from the training set.

A.2.5. Network Architecture and Configuration

The LSTM architecture and training hyperparameters are provided in Table 19.

References

[1] Yi-Feng Luo, Guan-Jhu Chen, Chun-Liang Liu, Ya-Shuo Chen, and Hua-Sheng Hsieh. An active bidirectional balancer with power distribution control strategy based on state of charge for lithium-ion battery pack. *Applied Energy*, 377:124661, 2025.

[2] Jing Hou, Xin Wang, Tian Gao, and Yan Yang. State of charge estimation of lithium-ion battery pack in electric vehicles based on inconsistency compensation modeling. *IEEE Transactions on Vehicular Technology*, 2025.

[3] Alexis Geslin, Le Xu, Devi Ganapathi, Kevin Moy, William C Chueh, and Simona Onori. Dynamic cycling enhances battery lifetime. *Nature Energy*, 10(2):172–180, 2025.

[4] Shiyun Liu, Kang Li, and James Yu. Battery pack condition monitoring and characteristic state estimation: Challenges, techniques, and future prospectives. *Journal of Energy Storage*, 105:114446, 2025.

Table 19
LSTM architecture and training hyperparameters

Parameter	Value
Architecture	
Number of LSTM layers	2
Hidden units per layer	50
Dropout rate	0.2
Sequence length	60
Input features	3 (strain, voltage, current)
Output features	2 (SOC, temperature)
Training	
Optimizer	Adam
Loss function	Mean squared error
Epochs	100
Batch size	32
Validation split	0.2

[5] Paul Takyi-Aninakwa, Shunli Wang, Guangchen Liu, Carlos Fernandez, Wenbin Kang, and Yingze Song. Deep learning framework designed for high-performance lithium-ion batteries state monitoring. *Renewable and Sustainable Energy Reviews*, 218:115803, 2025.

[6] Tom Rüther, Wesley Hileman, M Scott Trimboli, Gregory L Plett, Matthieu Dubarry, Nikhil Kumar, James Marco, Franz Roehrer, Andreas Jossen, Jan Schöberl, et al. Battery pack states, properties, and characterization techniques beyond cell level. *Cell Reports Physical Science*, 2025.

[7] Chuanxin Fan, Kailong Liu, Chunfei Gu, Quanqing Yu, Rui Wang, Jingyang Fang, and Qiao Peng. Battery pack state of charge estimation towards transportation electrification: Challenges and opportunities. *IEEE Journal of Emerging and Selected Topics in Industrial Electronics*, 2025.

[8] Liang Zhong, Chenbin Zhang, Yao He, and Zonghai Chen. A method for the estimation of the battery pack state of charge based on in-pack cells uniformity analysis. *Applied Energy*, 113:558–564, 2014.

[9] A Xianmin Mu, B Jiahao Zhang, C Guanlin Li, D Zekun Xiao, E Fanpeng Zeng, and F Jing Liu. Estimating soc and soh of energy storage battery pack based on voltage inconsistency using reference-difference model and dual extended kalman filter. *Journal of Energy Storage*, 81:110221, 2024.

[10] Gregory L Plett. Efficient battery pack state estimation using bar-delta filtering. In *EVS24 international battery, hybrid and fuel cell electric vehicle symposium*, pages 1–8, 2009.

[11] Haonan Dong, Wei Huang, and Yixin Zhao. Low complexity state-of-charge estimation for lithium-ion battery pack considering cell inconsistency. *Journal of Power Sources*, 515, 2021. ISSN 03787753. doi: 10.1016/j.jpowsour.2021.230599.

[12] Yuejiu Zheng, Wenkai Gao, Minggao Ouyang, Languang Lu, Long Zhou, and Xuebing Han. State-of-charge inconsistency estimation of lithium-ion battery pack using mean-difference model and extended kalman filter. *Journal of Power Sources*, 383:50–58, 2018. ISSN 03787753. doi: 10.1016/j.jpowsour.2018.02.058.

[13] Haifeng Dai, Xuezhe Wei, Zechang Sun, Jiayuan Wang, and Weijun Gu. Online cell soc estimation of li-ion battery packs using a dual time-scale kalman filtering for ev applications. *Applied Energy*, 95: 227–237, 2012. ISSN 03062619. doi: 10.1016/j.apenergy.2012.02.044.

[14] Yuejiu Zheng, Minggao Ouyang, Languang Lu, Jianqiu Li, Xuebing Han, Liangfei Xu, Hongbin Ma, Thomas A. Dollmeyer, and Vincent Freyermuth. Cell state-of-charge inconsistency estimation for lifepo4 battery pack in hybrid electric vehicles using mean-difference model. *Applied Energy*, 111:571–580, 2013. ISSN 03062619. doi: 10.1016/j.apenergy.2013.05.048.

[15] Cyril Truchot, Matthieu Dubarry, and Bor Yann Liaw. State-of-charge estimation and uncertainty for lithium-ion battery strings. *Applied Energy*, 119:218–227, 2014.

[16] Bo Jiang, Haifeng Dai, and Xuezhe Wei. A cell-to-pack state estimation extension method based on a multilayer difference model for series-connected battery packs. *IEEE Transactions on Transportation Electrification*, 8(2):2037–2049, 2021.

[17] Simin Peng, Yifan Miao, Rui Xiong, Jiawei Bai, Mengzeng Cheng, and Michael Pecht. State of charge estimation for a parallel battery pack jointly by fuzzy-pi model regulator and adaptive unscented kalman filter. *Applied Energy*, 360:122807, 2024.

[18] Shiyun Liu, Jenne Sunku Prasad, Kang Li, and Jihong Wang. Modeling and control of battery and compressed air energy storage systems. In Zhengtao Ding, editor, *Encyclopedia of Systems and Control Engineering (First Edition)*, pages 521–536. Elsevier, Oxford, first edition edition, 2026. ISBN 978-0-443-14080-8. doi: https://doi.org/10.1016/B978-0-443-14081-5.00132-X. URL https://www.sciencedirect.com/science/article/pii/B978044314081500132X.

[19] Jong-Hoon Kim, Jong-Won Shin, Chang-Yoon Jeon, and Bo-Hyung Cho. Screening process of li-ion series battery pack for improved voltage/soc balancing. In *The 2010 International Power Electronics Conference-ECCE ASIA-*, pages 1174–1179. IEEE, 2010.

[20] Jonghoon Kim, Jongwon Shin, Changyoon Chun, and B. H. Cho. Stable configuration of a li-ion series battery pack based on a screening process for improved voltage/soc balancing. *IEEE Transactions on Power Electronics*, 27(1):411–424, 2012. ISSN 0885-8993 1941-0107. doi: 10.1109/tpel.2011.2158553.

[21] Jonghoon Kim, Seongjun Lee, and BH Cho. Discrimination of li-ion batteries based on hamming network using discharging–charging voltage pattern recognition for improved state-of-charge estimation. *Journal of Power Sources*, 196(4):2227–2240, 2011.

[22] Rui Xiong, Fengchun Sun, Xianzhi Gong, and Hongwen He. Adaptive state of charge estimator for lithium-ion cells series battery pack in electric vehicles. *Journal of Power Sources*, 242:699–713, 2013. ISSN 03787753. doi: 10.1016/j.jpowsour.2013.05.071.

[23] Fengxian He, Weixiang Shen, Ajay Kapoor, Damon Honnery, and Daya Dayawansa. State of charge estimation for battery packs using h-infinity observer in underground mine electric vehicles. *Australian Journal of Electrical and Electronics Engineering*, 14(3-4):49–58, 2018. ISSN 1448-837X 2205-362X. doi: 10.1080/1448837x.2018.1451672.

[24] Aaruththiran Manoharan, Denesh Sooriamoorthy, KM Begam, and Vimal Rau Aparow. Electric vehicle battery pack state of charge estimation using parallel artificial neural networks. *Journal of Energy Storage*, 72:108333, 2023.

[25] Xingtao Liu, Wenlong Xia, Siyuan Li, Mingqiang Lin, and Ji Wu. State of charge estimation for lithium-ion battery pack with selected representative cells. *IEEE Transactions on Transportation Electrification*, 2023.

[26] Fulai An, Weige Zhang, Bingxiang Sun, Jiuchun Jiang, and Xinyuan Fan. A novel state-of-energy simplified estimation method for lithium-ion battery pack based on prediction and representative cells. *Journal of Energy Storage*, 63:107083, 2023.

[27] Qiao Wang, Min Ye, Bin Li, Gaoqi Lian, and Yan Li. Co-estimation of state of charge and capacity for battery packs in real electric vehicles with few representative cells and physics-informed machine learning. *Energy*, 306:132520, 2024.

[28] Fei Feng, Xiaosong Hu, Jianfei Liu, Xianke Lin, and Bo Liu. A review of equalization strategies for series battery packs: variables, objectives, and algorithms. *Renewable and Sustainable Energy Reviews*, 116:109464, 2019.

[29] Chaofan Yang, Xueyuan Wang, Qiaohua Fang, Haifeng Dai, Yaqian Cao, and Xuezhe Wei. An online soc and capacity estimation method for aged lithium-ion battery pack considering cell inconsistency. *Journal of Energy Storage*, 29, 2020. ISSN 2352152X. doi: 10.1016/j.est.2020.101250.

[30] Deyang Huang, Ziqiang Chen, Changwen Zheng, and Haibin Li. A model-based state-of-charge estimation method for series-connected lithium-ion battery pack considering fast-varying cell temperature. *Energy*, 185:847–861, 2019. ISSN 03605442. doi: 10.1016/j.energy.2019.07.063.

[31] Weizhong Wang, Pawel Malysz, Deqiang Wang, Ran Gu, Hong Yang, and Ali Emadi. Efficient multi-cell soc estimation for electrified vehicle battery packs. In *2016 IEEE Transportation Electrification Conference and Expo (ITEC)*, pages 1–5. IEEE, 2016.

[32] Yuan Chen, Yanzhong Liu, Yigang He, Zhiqiang Lyu, Yujing Cai, and Siyuan Zhang. Soc estimation of lithium-ion battery pack based on discharge stage division and fusion modeling. *IEEE Transactions on Instrumentation and Measurement*, 2025.

[33] Chang Yoon Chun, Gab-Su Seo, Sung Hyun Yoon, and Bo-Hyung Cho. State-of-charge estimation for lithium-ion battery pack using reconstructed open-circuit-voltage curve. In *2014 International Power Electronics Conference (IPEC-Hiroshima 2014-ECCE ASIA)*, pages 2272–2276. IEEE, 2014.

[34] Cheng Zhang, Kang Li, Jing Deng, and Shiji Song. Improved real-time state-of-charge estimation of lifepo4 battery based on a novel thermoelectric model. *IEEE Transactions on Industrial Electronics*, 64(1):654–663, 2016.

[35] Quanqing Yu, Yukun Huang, Aihua Tang, Chun Wang, and Weixiang Shen. Ocv-soc-temperature relationship construction and state of charge estimation for a series–parallel lithium-ion battery pack. *IEEE Transactions on Intelligent Transportation Systems*, 2023.

[36] Cong-Sheng Huang, Bharat Balagopal, and Mo-Yuen Chow. Estimating battery pack soc using a cell-to-pack gain updating algorithm. In *IECON 2018-44th Annual Conference of the IEEE Industrial Electronics Society*, pages 1807–1812. IEEE, 2018.

[37] Wenxue Liu, Yunhong Che, Jie Han, Zhongwei Deng, Xiaosong Hu, and Ziyou Song. Co-estimation of state-of-charge and capacity for series-connected battery packs based on multi-method fusion and field data. *Journal of Power Sources*, 615:235114, 2024.

[38] Xiaopeng Tang, Yuanqiang Zhou, Furong Gao, and Xin Lai. Joint estimation of state-of-charge and state-of-health for all cells in the battery pack using "leader-follower" strategy. *Etransportation*, 15: 100213, 2023.

[39] Changcheng Sun, Mingyu Gao, Hui Cai, Fei Xu, and Chunxiang Zhu. Data-driven state-of-charge estimation of a lithium-ion battery pack in electric vehicles based on real-world driving data. *Journal of Energy Storage*, 101:113986, 2024.

[40] Partha Pratim Mishra, Aadil Latif, Michael Emmanuel, Ying Shi, Killian McKenna, Kandler Smith, and Adarsh Nagarajan. Analysis of degradation in residential battery energy storage systems for rate-based use-cases. *Applied Energy*, 264:114632, 2020.

[41] Kuan-Cheng Chiu, Chi-Hao Lin, Sheng-Fa Yeh, Yu-Han Lin, Chih-Sheng Huang, and Kuo-Ching Chen. Cycle life analysis of series connected lithium-ion batteries with temperature difference. *Journal of Power Sources*, 263:75–84, 2014.

[42] Naixing Yang, Xiongwen Zhang, Guojun Li, Anjiang Cai, and Yunhua Xu. Effects of temperature differences among cells on the discharging characteristics of lithium-ion battery packs with series/parallel configurations during constant power discharge. *Energy Technology*, 6(6):1067–1079, 2018.

[43] Joachim Schaeffer, Eric Lenz, Duncan Gulla, Martin Z Bazant, Richard D Braatz, and Rolf Findeisen. Gaussian process-based online health monitoring and fault analysis of lithium-ion battery systems from field data. *Cell Reports Physical Science*, 2024.

[44] Jan Figgener, Jonas van Ouwerkerk, David Haberschusz, Jakob Bors, Philipp Woerner, Marc Mennekes, Felix Hildenbrand, Christopher Hecht, Kai-Philipp Kairies, Oliver Wessels, et al. Multi-year field measurements of home storage systems and their use in capacity estimation. *Nature Energy*, pages 1–10, 2024.

[45] Rui Xiong, Wanzhou Sun, Quanqing Yu, and Fengchun Sun. Research progress, challenges and prospects of fault diagnosis on battery system of electric vehicles. *Applied Energy*, 279:115855, 2020.

[46] Wenxin Mei, Zhi Liu, Chengdong Wang, Chuang Wu, Yubin Liu, Pengjie Liu, Xudong Xia, Xiaobin Xue, Xile Han, Jinhua Sun, et al. Operando monitoring of thermal runaway in commercial lithium-ion cells via advanced lab-on-fiber technologies. *Nature communications*, 14(1):5251, 2023.

[47] Yihuan Li, Kang Li, Xuan Liu, Xiang Li, Li Zhang, Bruno Rente, Tong Sun, and Kenneth TV Grattan. A hybrid machine learning framework for joint soc and soh estimation of lithium-ion batteries assisted with fiber sensor measurements. *Applied Energy*, 325: 119787, 2022.

[48] Shiyun Liu and Kang Li. Thermal monitoring of lithium-ion batteries based on machine learning and fibre bragg grating sensors. *Transactions of the Institute of Measurement and Control*, 45(8): 1570–1578, 2023.

[49] C Gervillié-Mouravieff, C Boussard-Plédel, Jiaqiang Huang, C Leau, L Albero Blanquer, M Ben Yahia, M-L Doublet, ST Boles, XH Zhang, JL Adam, et al. Unlocking cell chemistry evolution with operando fibre optic infrared spectroscopy in commercial na (li)-ion batteries. *Nature Energy*, 7(12):1157–1169, 2022.

[50] John Cannarella and Craig B Arnold. State of health and charge measurements in lithium-ion batteries using mechanical stress. *Journal of Power Sources*, 269:7–14, 2014.

[51] Laura Albero Blanquer, Florencia Marchini, Jan Roman Seitz, Nour Daher, Fanny Bétermier, Jiaqiang Huang, Charlotte Gervillié, and Jean-Marie Tarascon. Optical sensors for operando stress monitoring in lithium-based batteries containing solid-state or liquid electrolytes. *Nature Communications*, 13(1):1153, 2022.

[52] John Cannarella. *Coupled mechanical and electrochemical phenomena in lithium-ion batteries*. PhD thesis, Princeton University, 2015.

[53] John Cannarella and Craig B Arnold. Stress evolution and capacity fade in constrained lithium-ion pouch cells. *Journal of Power Sources*, 245:745–751, 2014.

[54] Shiyun Liu, Kang Li, Benjamin Chong, and Ye Chen. State-of-charge estimation of li-ion battery packs based on optic fibre sensor measurements. *Transportation Research Procedia*, 70:388–397, 2023.

[55] Yen-Jie Ee, Kok-Soon Tey, Kok-Sing Lim, Prashant Shrivastava, SBRS Adnan, and Harith Ahmad. Lithium-ion battery state of charge (soc) estimation with non-electrical parameter using uniform fiber bragg grating (fbg). *Journal of Energy Storage*, 40:102704, 2021.

[56] Jun Peng, Shuhai Jia, Shuming Yang, Xilong Kang, Hongqiang Yu, and Yaowen Yang. State estimation of lithium-ion batteries based on strain parameter monitored by fiber bragg grating sensors. *Journal of Energy Storage*, 52:104950, 2022.

[57] Xudong Xia, Wen Wu, Zhencheng Li, Xile Han, Xiaobin Xue, Gaozhi Xiao, and Tuan Guo. State of charge estimation for commercial li-ion battery based on simultaneously strain and temperature monitoring over optical fiber sensors. *IEEE Transactions on Instrumentation and Measurement*, 2024.

[58] Wenjuan Sheng, Junkai Wang, and GD Peng. Enhanced strain assistance for soc estimation of lithium-ion batteries using fbg sensors. *Applied Energy*, 383:125385, 2025.

[59] Preethem S Babu and Indragandhi V. Enhanced soc estimation of lithium ion batteries with realtime data using machine learning algorithms. *Scientific Reports*, 14(1):16036, 2024.

[60] Juan Yan, Kang Li, Er-Wei Bai, Jing Deng, and Aoife M Foley. Hybrid probabilistic wind power forecasting using temporally local gaussian process. *IEEE Transactions on sustainable energy*, 7(1): 87–95, 2015.

[61] David Duvenaud. *Automatic Model Construction with Gaussian Processes*. PhD thesis, University of Cambridge, 2014.

[62] Jonathan Ko, Daniel J Klein, Dieter Fox, and Dirk Haehnel. Gp-ukf: Unscented kalman filters with gaussian process prediction and observation models. In *2007 IEEE/RSJ international conference on intelligent robots and systems*, pages 1901–1907. IEEE, 2007.

[63] Gregory L Plett. *Battery management systems, Volume II: Equivalent-circuit methods*. Artech House, 2015.

[64] Zachary J Schiffer, John Cannarella, and Craig B Arnold. Strain derivatives for practical charge rate characterization of lithium ion electrodes. *Journal of The Electrochemical Society*, 163(3):A427, 2015.

[65] Runlin Wang, Haozhe Zhang, Qiyu Liu, Fu Liu, Xile Han, Xiaoqing Liu, Kaiwei Li, Gaozhi Xiao, Jacques Albert, Xihong Lu, et al. Operando monitoring of ion activities in aqueous batteries with plasmonic fiber-optic sensors. *Nature communications*, 13(1):547, 2022.

[66] Jiaqiang Huang, Laura Albero Blanquer, Julien Bonefacino, Eric R Logan, Daniel Alves Dalla Corte, Charles Delacourt, Betar M Gallant, Steven T Boles, JR Dahn, Hwa-Yaw Tam, et al. Operando decoding of chemical and thermal events in commercial na (li)-ion cells via optical sensors. *Nature energy*, 5(9):674–683, 2020.

[67] Shen Li, Cheng Zhang, Yan Zhao, Gregory J Offer, and Monica Marinescu. Effect of thermal gradients on inhomogeneous degradation in lithium-ion batteries. *Communications Engineering*, 2(1): 74, 2023.

[68] Max Naylor Marlow, Jingyi Chen, and Billy Wu. Degradation in parallel-connected lithium-ion battery packs under thermal gradients. *Communications Engineering*, 3(1):2, 2024.

[69] Cheng Zhang, Kang Li, Lei Pei, and Chunbo Zhu. An integrated approach for real-time model-based state-of-charge estimation of lithium-ion batteries. *Journal of Power Sources*, 283:24–36, 2015.

[70] Antti Aitio and David A Howey. Predicting battery end of life from solar off-grid system field data using machine learning. *Joule*, 5(12): 3204–3220, 2021.

[71] Fujin Wang, Zhi Zhai, Zhibin Zhao, Yi Di, and Xuefeng Chen. Physics-informed neural network for lithium-ion battery degradation stable modeling and prognosis. *Nature Communications*, 15(1): 4332, 2024.

[72] Cheng Zhang, Kang Li, and Jing Deng. Real-time estimation of battery internal temperature based on a simplified thermoelectric model. *Journal of Power Sources*, 302:146–154, 2016.

[73] Shen Li, Anisha N Patel, Cheng Zhang, Tazdin Amietszajew, Niall Kirkaldy, Gregory J Offer, and Monica Marinescu. Internal temperature estimation for lithium-ion batteries through distributed equivalent circuit network model. *Journal of Power Sources*, 611: 234701, 2024.

[74] Arno Solin. *Stochastic Differential Equation Methods for Spatio-Temporal Gaussian Process Regression*. PhD thesis, Aalto University, 2016.

[75] Kwang-Jae Lee, Won-Hyung Lee, and Kwang-Ki K Kim. Battery state-of-charge estimation using data-driven gaussian process kalman filters. *Journal of Energy Storage*, 72:108392, 2023.

[76] Quanqing Yu, Rui Xiong, Cheng Lin, Weixiang Shen, and Junjun Deng. Lithium-ion battery parameters and state-of-charge joint estimation based on h-infinity and unscented kalman filters. *IEEE Transactions on Vehicular Technology*, 66(10):8693–8701, 2017.

[77] Himali Singh, Kumar Vijay Mishra, and Arpan Chattopadhyay. Inverse unscented kalman filter. *IEEE Transactions on Signal Processing*, 72:2692–2709, 2024.

[78] Sarah E Houts, Shandor G Dektor, and Stephen M Rock. A robust framework for failure detection and recovery for terrain-relative navigation. *Unmanned Untethered Submersible Technology*, 2013.

[79] Haining Liu, Ijaz Haider Naqvi, Fajia Li, Chengliang Liu, Neda Shafiei, Yulong Li, and Michael Pecht. An analytical model for the cc-cv charge of li-ion batteries with application to degradation analysis. *Journal of Energy Storage*, 29:101342, 2020.

[80] Zhen Zhang, Xin Gu, Yuhao Zhu, Teng Wang, Yichang Gong, and Yunlong Shang. Data-driven available capacity estimation of lithium-ion batteries based on fragmented charge capacity. *Communications Engineering*, 4(1):32, 2025.

[81] Anurag Ganguli, Bhaskar Saha, Ajay Raghavan, Peter Kiesel, Kyle Arakaki, Andreas Schuh, Julian Schwartz, Alex Hegyi, Lars Wilko Sommer, Alexander Lochbaum, 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. *Journal of Power Sources*, 341:474–482, 2017.

[82] Mustapha Makki, Georges Ayoub, Cheol W Lee, Chulheung Bae, and Xavier Colin. Effect of battery fast cyclic charging on the mechanical and fracture behavior of the lithium-ion battery separator. *Polymer Degradation and Stability*, 216:110469, 2023.

[83] Azadeh Sheidaei, Xinran Xiao, Xiaosong Huang, and Jonathon Hitt. Mechanical behavior of a battery separator in electrolyte solutions. *Journal of Power Sources*, 196(20):8728–8734, 2011.

[84] Charlotte Gervillié-Mouravieff, Wurigumula Bao, Daniel A Steingart, and Ying Shirley Meng. Non-destructive characterization techniques for battery performance and life-cycle assessment. *Nature Reviews Electrical Engineering*, 1(8):547–558, 2024.

[85] Bruno Rente, Matthias Fabian, Miodrag Vidakovic, Xuan Liu, Xiang Li, Kang Li, Tong Sun, and Kenneth TV Grattan. Lithium-ion battery state-of-charge estimator based on fbg-based strain sensor and employing machine learning. *IEEE Sensors Journal*, 21(2): 1453–1460, 2020.

[86] Neware. *BTS4000-60V100A Battery Testing System*. Neware, 2018. URL https://newarebattery.com/wp-content/uploads/2018/specs/BTS4000-60V100A-1CH.pdf. PDF document.

[87] Micron Optics. *Optical Sensing Interrogator | sm130*. Micron Optics, Unknown. URL http://www.micronoptics.ru/uploads/library/documents/Datasheets/Micron%20Optics%20sm130.pdf. PDF document.

[88] Ling Mao, Chuan Yang, Jinbin Zhao, Keqing Qu, and Xiaofang Yu. Joint state-of-charge and state-of-health estimation for lithium-ion batteries based on improved lebesgue sampling and division of aging stage. *Energy Technology*, 11(10):2300567, 2023.

[89] Regis J Van Steenkiste. *Strain and temperature measurement with fiber optic sensors*. CRC Press, 1996.

[90] Andrew Kim. *Unconventional Experimental Setups for Battery Analysis*. PhD thesis, Princeton University, 2019.

[91] Ronald W Schafer. What is a savitzky-golay filter?[lecture notes]. *IEEE Signal processing magazine*, 28(4):111–117, 2011.

[92] KM Muraleedhara Prabhu. *Window functions and their applications in signal processing*. Taylor & Francis, 2014.

[93] Luna Innovations Incorporated. os1200 optical strain gage data sheet. https://lunainc.com/sites/default/files/assets/files/resource-library/LUNA-Data-Sheet-FBG-os1200-final.pdf, 2019. Accessed: October 25, 2023.

[94] Donglei Liu, Shunli Wang, Xiaoxia Li, Yongcun Fan, Carlos Fernandez, and Frede Blaabjerg. A novel extended kalman filter-guided long short-term memory algorithm for power lithium-ion battery state of charge estimation at multiple temperatures. *Energy*, page 137973, 2025.

[95] Kai Li, Yu Huang, Gaoce Han, Wenrong Lyu, Aiqi He, Nini Liu, Yifei Yu, and Yunhui Huang. Real-time electrochemical-strain distribution and state-of-charge mapping via distributed optical fiber for lithium-ion batteries. *Journal of Power Sources*, 624:235526, 2024.

[96] Simone Barcellona, Silvia Colnago, Giovanni Dotelli, Saverio Latorrata, and Luigi Piegari. Aging effect on the variation of li-ion battery resistance as function of temperature and state of charge. *Journal of Energy Storage*, 50:104658, 2022.

[97] Wladislaw Waag, Stefan Käbitz, and Dirk Uwe Sauer. Experimental investigation of the lithium-ion battery impedance characteristic at various conditions and aging states and its influence on the application. *Applied energy*, 102:885–897, 2013.

[98] Ningyue Mao, Haofeng Su, Xiaofan Cui, Ming Han, and Xuan Zhou. In-situ monitoring of li-ion batteries formation and cycling behaviour through optical approach. In *2025 IEEE/AIAA Transportation Electrification Conference and Electric Aircraft Technologies Symposium (ITEC+ EATS)*, pages 1–5. IEEE, 2025.

[99] Xile Han, Hai Zhong, Kaiwei Li, Xiaobin Xue, Wen Wu, Nan Hu, Xihong Lu, Jiaqiang Huang, Gaozhi Xiao, Yaohua Mai, et al. Operando monitoring of dendrite formation in lithium metal batteries via ultrasensitive tilted fiber bragg grating sensors. *Light: Science & Applications*, 13(1):24, 2024.

[100] LHJ Raijmakers, DL Danilov, R-A Eichel, and PHL Notten. A review on various temperature-indication methods for li-ion batteries. *Applied energy*, 240:918–945, 2019.