Journal of Energy Storage 148 (2026) 120194

Contents lists available at ScienceDirect

Journal of Energy Storage

journal homepage: www.elsevier.com/locate/est

Research papers

Digital twin-based prediction of battery parameters from limited initial data using an optimised time-series multi-layer perceptron

Sajad Najafi-Shad [a,b], Davide Domenico Sciortino [a,b,*], Shahaboddin Resalati [a], Paul Henshall [a,b], Denise Morrey [a,b]

[a] Oxford Brookes University, United Kingdom
[b] Centre for Batteries, Electric Vehicles and Electronics (CBEVE), United Kingdom

ARTICLE INFO

Keywords:
Digital twin
Machine learning
Anomaly detection
Genetic algorithm

ABSTRACT

Digital twin (DT) technology is increasingly leveraged for real-time monitoring and predictive modelling of battery systems. However, existing machine learning (ML)-based approaches for battery parameter estimation often rely on large historical datasets, limiting their applicability during the early stages of operation. In this paper, we propose a novel hybrid framework that enables accurate battery parameter prediction and early anomaly detection with minimal initial data. The approach integrates a time-series multi-layer perceptron (TS-MLP) model optimised by a genetic algorithm (GA) to dynamically select input intervals that minimise prediction error. To enhance anomaly detection, an optimised one-class support vector machine (OC-SVM) is trained using both real and synthetically generated anomalies via Gaussian process regression (GPR), ensuring robust performance from early cycles. The proposed method is validated on NASA's 18650 lithium-ion battery dataset. Results show that the model accurately predicts next-cycle cell temperature with a minimum R^2 of 0.9875 and a maximum RMSE of 0.34 °C. This framework provides a reliable and data-efficient solution for early-stage battery diagnostics in DT environments.

1. Introduction

The requirement for uninterrupted real-time observation and prediction, together with an increase in available computational power, have driven the development of Digital Twin (DT) technology [1]. A DT is a virtual representation of a physical asset, system, or process, designed to accurately represent its real-world counterpart. While digital models are employed to simulate, analyse, and optimise system performance, DTs are updated with real-time data and offer additional capabilities, including real-time monitoring, predictive maintenance, and live performance optimisation [2].

Recent advancements in ML methods [3,4], AI-Driven Internet of Things (IoT) [5,6] and DT integration, have facilitated battery systems modelling. These developments offer innovative prospects for improving the accuracy of lithium-ion battery parameter prediction [7]. Accurate prediction of battery parameters such as temperature, SoH, and SoC plays a crucial role in the maintenance, optimisation, and performance enhancement of battery systems. As the battery ages, its State of Health (SoH) is affected by factors such as loss of active material (LAM), loss of lithium inventory (LLI), Solid Electrolyte Interface (SEI) growth, and particle cracking [8]. Consequently, these variations affect the accuracy of initial Remaining Useful Life (RUL) prediction, highlighting the significance of DT approach. Throughout the battery lifecycle, the effects of ageing are difficult to observe in current and voltage measurements [9]. In contrast, battery temperature has a direct effect on battery capacity [9]. Therefore, accurately modelling battery temperature enhances the prediction of RUL and SoH, and is crucial for improving both operational efficiency and battery lifespan [10]. Battery modelling techniques can be classified into three groups:

- Physics-based models (white box): These models, such as the P2D model and the Single Particle Model (SPM), describe the battery cell as a system of porous electrodes and a separator, capturing ion diffusion and electrode reactions through coupled partial differential equations. While these models offer high accuracy in parameter estimation and degradation prediction, they suffer from being complex and computationally expensive. These limitations make them difficult to implement in real-time DT applications [11–13].
- Semi-empirical models (grey box): These models describe battery cell behaviour using electrical networks of resistors and capacitors that mimic the dynamic voltage–current relationship and capture

* Corresponding author at: Oxford Brookes University, United Kingdom.
E-mail address: dsciortino@brookes.ac.uk (D.D. Sciortino).

https://doi.org/10.1016/j.est.2025.120194
Received 9 October 2025; Received in revised form 18 December 2025; Accepted 26 December 2025
Available online 6 January 2026

the approximate ohmic and polarisation effects. However, as batteries age, these models lose accuracy in capturing the non-linear behaviour of the batteries [14–17].

- Data-driven models (black box): These models use Machine Learning (ML) to estimate battery parameters, either directly from raw data or derived features. These models enable highly accurate, adaptable, and interpretable parameter estimation, thanks to their capability of characterising complex and nonlinear ageing behaviours without relying on detailed electrochemical knowledge [18]. Examples include GPR [19], SVM [20], ANN [21], Convolutional Neural Network Convolutional Neural Network (CNN) [22], and Long Short-Term Memory (LSTM) [23].

While semi-empirical and physics-based models are well-developed, data-driven models have recently gained more attention in battery parameter prediction studies [7,9,13,24–26]. Data-driven models typically involve three key stages: *(i)* dataset acquisition, *(ii)* feature selection, and *(iii)* ML deployment. To streamline the otherwise time-intensive process of data collection, open-source datasets such as those provided by NASA [27–29] and the MIT-Stanford repository [28,30,31] are frequently used.

Recent studies have applied various ML techniques, including LSTM, GPR, Support Vector Machine (SVM), and Artificial Neural Network (ANN), to predict battery parameters. These models are mainly supported by feature selection methods as a preprocessing step to enhance the prediction accuracy. Studies show that LSTM networks can capture temporal dependencies and long-term correlations, making them particularly effective for estimating battery parameters. Qu et al. [7] have employed a LSTM to virtually full-discharge Li-ion batteries from partial discharge data, allowing accurate estimation of the actual capacity of batteries. Deng et al. [24] have proposed the Pearson Correlation Coefficient (PCC) method to select input features from the discharge capacity curve. Then, an LSTM method has been employed to recognise battery degradation patterns along with transfer learning methods for SoH estimation, which results in improved generalisation of SoH estimation. Deng et al. [9] have proposed PCC and Grey Relational Grade (GRG) to evaluate the input features. Then, an LSTM-based Seq2Seq model is proposed along with GPR to predict capacity trajectories, enabling accurate diagnosis of degradation with a Root Mean Square Error (RMSE) of 2.75%.

GPR has been demonstrated to provide probabilistic predictions and achieves high accuracy in battery parameter estimation, including State of Charge (SoC), SoH, and capacity, with reported errors below 2% [13,32,33]. He et al. [32] have derived health indicators from the variation coefficient of partial charging curves. An optimised GPR was then employed to estimate battery SoH, achieving an estimation RMSE of 0.85. Aitio et al. [13] have employed GPR to generate battery health trajectories and predicted the battery End of Life (EoL) by combining the proposed health trajectories with stress factors extracted from the raw data.

SVM and ANN are also widely used for battery parameter prediction due to their flexibility in modelling nonlinear relationships. In [26] PCC method is deployed to select input features, including Constant Current (CC) and Constant Voltage (CV) mode durations. Then, a modified ant lion optimisation method is employed to select the optimum values of the Support Vector Regression (SVR) parameters to utilise for the SoH prediction, resulting in improved SoH estimation accuracy compared to typical SVR. Xiong et al. [25] have introduced a hybrid feature selection strategy containing PCC, LASSO regression, neighbourhood component analysis, and the ReliefF algorithm to select the most correlated parameters to the SoH. Then, SVM, ANN, and GPR were deployed to predict the SoH of the Lithium Iron Phosphate (LFP) and Nickel-Cobalt-Aluminum (NCA) batteries, achieving prediction RMSE of 0.8% for NCA batteries and 0.6% for LFP batteries.

Multi-Layer Perceptron (MLP) is also employed for parameter estimation as one of the ML approaches, generally offering less complexity and more computational efficiency compared to deep learning models such as LSTMs or CNNs [34]. This simplicity can result in faster training times and reduced computational resource requirements. This feature makes them a practical choice for short-term predictions like DT, where live predictions are essential [35]. In the context of predicting the parameters of lithium-ion batteries, studies have demonstrated that MLP-based time series architectures can bring higher prediction accuracy while capturing both local and global temporal patterns [36,37].

One of the major challenges in battery systems is the reliable detection of anomalies, which typically are infrequent, unpredictable, and often underrepresented in historical data [38]. In the context of battery parameters, anomalies often appear as spikes or deviations, often caused by factors such as cell ageing, internal faults, or environmental conditions [39]. Data-driven anomaly detection methods can be categorised into unsupervised learning models such as Isolation Forest (IF) [40] or OC-SVM [41] and supervised learning methods including SVM [42], and Random Forest (RF) [43]. Jin et al. [40] have employed IF model as an unsupervised learning method to detect anomalies, enabling more accurate anomaly detection and SoH prediction compared to traditional methods. Zhang et al. [42] have employed SVM to diagnose battery faults when dealing with time-varying battery datasets and achieved up to 92.5% diagnostic accuracy. Jeon et al. [41] have utilised One-Class Support Vector Machine (OC-SVM) to analyse the battery features and categorise battery anomalies to identify the degraded Li-ion battery modules, which enabled the detection of battery anomaly modules using initial 10% discharge data of 60 cycles.

ML methods can also be implemented within a DT framework to improve the accuracy of battery modelling [44,45,47–49]. Jafari et al. [45] have proposed a DT-based SoH prediction method combined with the Extreme Gradient Boosting (XGBoost) model and Extended Kalman Filter (EKF), utilising the NASA dataset. Similarly, in [49], a DT-based ANN and wavelet analysis are used to accurately replicate SoC and SoH, facilitating accurate characterisation of the battery's internal impedance. Further evidence is provided by Deng et al. [48], who integrated a DT, an Equivalent Circuit Model (ECM), and a CNN to generate synthetic data representing different stages of battery ageing. Moreover, physics-based DT models can be employed to increase the accuracy of fault detection algorithms [44]. Beyond prediction and fault detection, DT has also been shown to facilitate battery Battery Management Systems (BMS) operation [50–52]. Table 1 presents the summary of the key studies discussed above.

While the aforementioned studies have proposed novel strategies for utilising DT and ML in estimating battery parameters, the following aspects appear to be overlooked:

- Most existing Machine Learning (ML) frameworks are typically trained, validated, and tested on complete battery life-cycle datasets. Therefore, these models function accurately under controlled offline settings but struggle when deployed in real-time conditions where data arrive sequentially and the model must continuously adapt to the battery's behaviour. This mismatch between offline training and online operation limits their reliability, particularly during the early cycles when only sparse data are available.
- While the concept of a Digital Twin (DT) enables real-time estimation of battery parameters through live data acquisition, extending this capability to long-term parameter prediction remains an unsolved challenge and requires further investigation. DT models must predict future battery behaviour based on limited initial information while accounting for nonlinear ageing dynamics and cycle-to-cycle thermal variations. Current DT implementations rarely address this requirement, leaving a gap in achieving fully predictive DT systems.

Table 1
Summary of key studies on battery modelling, state estimation, prognostics, and DT frameworks.

Study	Methodology	Task/Contribution
[8] Michelini et al	Capacity fade, Internal resistance rise	Evaluation of ageing mechanisms for second-life applications
[9] Deng et al	PCC + Seq2Seq LSTM + GPR	Capacity fade trajectory prediction
[10] Xia et al	Incremental capacity and differential voltage feature extraction	SoH/RUL estimation using ageing-sensitive IC/DV indicators
[11] Trimboli & Plett	Review of physics-based battery models (SPM, P2D, thermal)	Evaluation of modelling strengths and limitations for BMS design
[13] Aitio & Howey	GPR trained on field data with stress factors (T, DoD, irradiance)	EoL forecasting and degradation-rate estimation for off-grid Li-ion systems
[15] Wen et al	Review of numerical, electrochemical and ML modelling methods	Comparative analysis of battery-modelling methodologies
[16] Li et al	ECM development + parameter identification + validation	Improved ECM accuracy under varying operational conditions
[18] Acurio et al	Tikhonov regularisation + Huber regression	Noise-robust SoH estimation under low-SNR
[19] Buchanan & Crawford	CNN + GPR for probabilistic outputs	Probabilistic SoH estimation with uncertainty quantification
[20] Swain et al	RF + SVM with optimised features	RUL prediction using historical cycling data
[21] Li et al	Multi-task NN (SoH + RUL)	Joint estimation of SoH and RUL
[22] Sherkatghanad et al	CNN + BiLSTM + attention	High-accuracy real-time SoC prediction
[23] She et al	MSCC charge-pattern analysis + voltage/polarisation indicators	SoH estimation from real EV charging signatures
[24] Deng et al	PCC-based feature selection + LSTM + transfer learning	Cross-domain SoH prediction across heterogeneous datasets
[25] Xiong et al	PCC + LASSO + NCA + ReliefF + ML models (SVM/NN/GPR)	High-accuracy multi-chemistry SoH prediction (LFP/NCA)
[26] Li et al	PCC + optimised SVR using ALO	Improved SoH prediction accuracy using optimised non-linear regression
[29] Kim et al	PF enhanced using fuzzy-rule uncertainty reduction	Improved RUL prediction with reduced estimation uncertainty
[30] Ansari et al	Hybrid co-estimation (ML + optimisation)	Joint estimation of SoH and RUL
[31] Dai et al	Ridge Regression for SoH + GRU for forecasting	High-accuracy SoH trend prediction
[32] He et al	Partial-charge feature extraction + optimised GPR	High-accuracy SoH estimation
[33] Selvaraj & Vairavasundaram	ML model with Bayesian hyperparameter optimisation	Improved SoC estimation accuracy
[36] Ye et al	IIP-Mixer deep model (patch mixing)	Multiscale RUL prediction with enhanced degradation-pattern learning
[40] Jin et al	Blockchain-secured data + IF anomaly detection	Detection and removal of anomalous battery data to improve SoH/SoC/RUL estimation reliability
[41] Jeon et al	DCAE + OC-SVM applied to IC/DV features	Early detection of instabilities, weak cells, and emerging faults
[42] Zhang et al	Online LS-SVM classifier with adaptive updating	Real-time battery fault classification
[43] Liu et al	RF classifier with minimal sensing	Online localisation of internal short-circuit faults
[44] Yuan et al	Physics-based DT simulating electrochemical/thermal ISC	Synthetic ISC dataset generation for fault diagnostics
[45] Jafari & Byun	DT integrating XGBoost + EKF	Real-time SoH estimation and drift correction using BMS data
[46] Dang et al	ODE-informed neural network	Physics-consistent SoC estimation
[47] Saba et al	DT integrated with ITS data	Real-time EV battery performance monitoring
[48] Deng et al	DT + ECM simulation + CNN + domain adaptation	Fast SoH estimation with minimal labelled data
[49] Di Fonso et al	Wavelet decomposition + NN reconstruction in DT framework	High-fidelity reconstruction of SoC/SoH from laboratory data

1.1. Contribution

This study aims to address the mentioned shortcomings by implementing a Time-Series Multi-Layer Perceptron (TS-MLP) and a GA optimisation within a DT framework. Cell surface temperature is accurately predicted using simulated live data with only a limited number of cycles. Although derived from offline measurements, the data are streamed in real time by sending values at one-second intervals. A live dataset is created from the 18 650 battery cell [27]. After a predetermined number of charge–discharge cycles, a time series of cell temperature values is generated. Then, a GA is employed to identify the most relevant reference points within the temperature time series, which serve as input intervals to the prediction model. Once the MLP model is trained using the optimised intervals, the reference points are shifted one cycle forward, enabling the prediction of future cycle values. As training expands with each cycle, the model's prediction accuracy improves. The model is also equipped with an OC-SVM anomaly detection method. Two sets of synthetic anomalies are generated using GPR. The proposed DT-based TS-MLP method is validated using the NASA dataset [27]. The contributions of this paper can be summarised as follows:

- While the conventional DT–ML frameworks [45,47–49] rely on offline training using complete battery life-cycle data, the proposed approach enables live, cycle-by-cycle prediction and anomaly detection after only a few initial cycles.
- The integration of a Time-Series Multi-Layer Perceptron (TS-MLP) optimised via a Genetic Algorithm (GA) makes the approach computationally efficient, straightforward to implement, and practical. By combining the simplicity of MLP with the ability to capture time-dependent patterns, TS-MLP allows a simple model to understand time series data, making it a highly practical solution for real-time implementation.
- Development of a battery anomaly detection approach employing OC-SVM with GPR-based synthetic anomaly generation improves the performance of the unsupervised OC-SVM in anomaly detection, making it more accurate and reliable compared to prior studies [41,42].

2. Methodology

In this study, cell surface temperature is selected as the key parameter for the DT, as it exhibits significant variation both within individual test cycles and over the battery's lifetime. This parameter is also crucial for monitoring the cell's state and understanding the thermal behaviour [53]. To simulate the data transmission between the real counterpart and DT, the data are input into the model on a cycle-by-cycle basis. Once the cycle data are collected and fed into the model, the GA identifies the optimal reference points within the stored data. The TS-MLP is then trained using the intervals generated based on the selected reference points and tested on the most recent input cycle. Subsequently, the reference points are shifted one cycle forward and the model predicts the temperature profile of the next

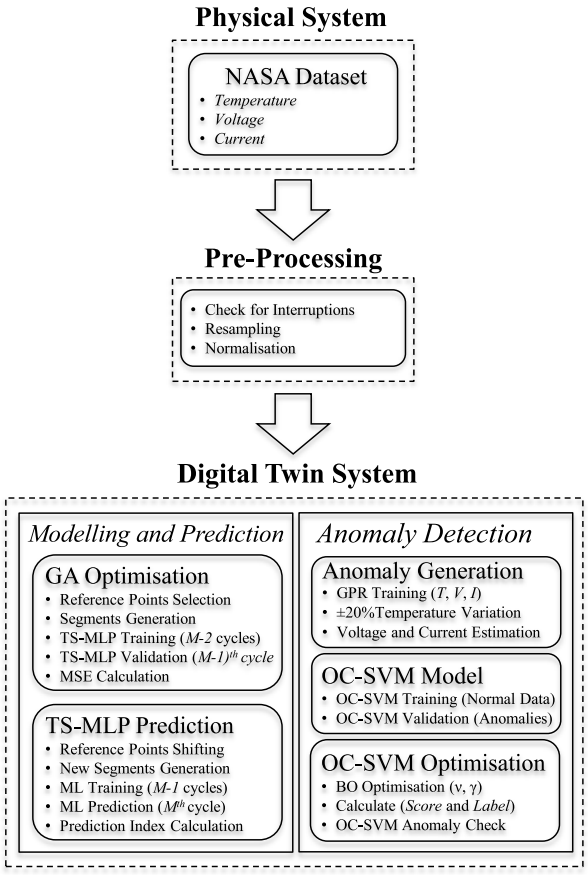

Fig. 1. Conceptual map of the proposed battery digital twin framework.

cycle. With this strategy, the model can begin making predictions after only a few cycles, and its accuracy improves as more cycles are incorporated into the training dataset. To enhance the model's robustness against anomalies, an OC-SVM is trained using the stored data and validated on synthetic anomalies generated by the GPR. The trained model is then applied to assess each incoming sample in real time for potential anomaly detection. Fig. 1 represents the conceptual map of the proposed method.

2.1. ML implementation

The TS-MLP uses a time series of cell surface temperature values to predict future temperatures. The time series is represented as a sequence of scalar values:

$$x_{(t)} = [x_0, x_1, x_2, \dots, x_m] \tag{1}$$

$$x_{(h)} = [x_{m+1}, x_{m+2}, \dots, x_{m+h}] \tag{2}$$

where $x_{(t)}$ is composed of temperature values up to the current time step m, and $x_{(h)}$ represents the future values over a prediction horizon h, corresponding to one future cycle. Fig. 2 illustrates the TS-MLP inputs, the validation period, the prediction horizon, and the reference points. The TS-MLP reference points are denoted as $\alpha, \beta, \gamma, \delta, \omega$ with the corresponding inputs $a_\alpha, a_\beta, a_\gamma, a_\delta, a_\omega$.

In particular, the inputs are formed by shifting backward by the length of the reference points and then selecting a range of h time steps backward. Therefore, selecting appropriate reference points can adjust the input ranges accordingly. In this study, the number of reference points is set to five, which corresponds to five model inputs. To predict future values, a mapping procedure is employed to select

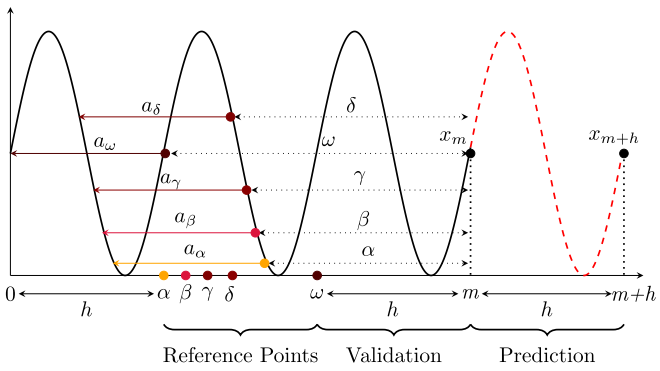

Fig. 2. Schematic illustration of ranges for the DT input, validation, and prediction ranges.

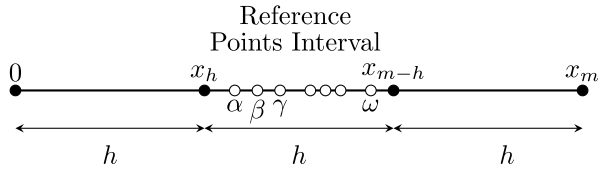

Fig. 3. Selection intervals for the proposed TS-MLP model for the first prediction.

reference points within the reference range. The following constraints are considered:

- Predicting a full cycle of h data points requires the model to be provided with at least h points per input, selected in reverse chronological order starting from the reference points backwards in time. The reference points should be selected from diverse points within a full cycle of length h. Therefore, the span between the earliest and latest reference points should cover at least a full cycle:

$$\omega - \alpha \geq h \tag{3}$$

- To perform a reliable validation, a minimum of h consecutive data points must be available. Hence, the interval between the last data point m and the nearest reference point α must be at least equal to the prediction length:

$$m - \alpha \geq h \tag{4}$$

- Since the input length must match the prediction horizon, the earliest input, ranging from 0 to m - ω must include at least h elements:

$$m - \omega \geq h \tag{5}$$

In this study, each cycle refers to a complete charge–discharge process of the battery cell. The first three cycles were used for model initial training and validation. If a complete cycle is considered as the prediction horizon, the first prediction corresponds to cycle 4. In this case, the most recent cycle, cycle 3, is used for validation. Reference points are selected from cycle 2 and subsequently the training intervals are selected from cycles 1 and 2. The selection intervals of the reference points for the first prediction are shown in Fig. 3, indicating that the minimum value for m is $3h$.

For the second prediction, which targets cycle 5, cycle 4 is considered for validation, reference points are taken from cycles 2 and 3. Therefore, the training dataset is expanded to include data from the beginning to cycle 3. The following predicted cycles correspond to subsequent charge–discharge sequences predicted by the DT-based

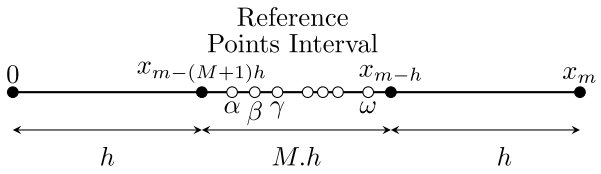

Fig. 4. Selection intervals for the proposed TS-MLP model for the Mth prediction.

model. As the prediction advances through subsequent cycles, the training dataset incrementally incorporates more data, thereby improving the prediction accuracy.

A static learning rate was used during TS-MLP training, as the lightweight architecture and limited cycle-level data allow stable convergence without the need for dynamic learning-rate scheduling. Fig. 4 illustrates the selection intervals for reference points for the Mth prediction.

The interval selected as the validation period, starting from the first point after the nearest reference point and including h subsequent samples, can be expressed as:

$$V = \{m - \alpha + k \mid k = 1, 2, \dots, h\} \tag{6}$$

The mapping procedure selects a subset of consecutive data points according to V:

$$X_i = [x_{m-i-h+k} \mid k = 1, 2, \dots, h\ ] \tag{7}$$

$$i \in \{\alpha, \beta, \gamma, \dots, \omega\}$$

where X_i represents the input sequence for the ith reference point, including h consecutive samples ending at x_{m-i}. The future values of the proposed time series are obtained as a weighted aggregation of predictions derived from each input sequence:

$$x_{(t+h)} = \sum_{j=1}^{s_n} w_j \cdot f_{(X_j)} \tag{8}$$

where $f_{(X_i)}$ is the prediction function applied to segment s_i, $x_{(t+h)}$ is the predicted values for the next h data points, and w_i is the weight assigned to each segment, and is defined as:

$$w_i = g(X_j, c) \tag{9}$$

where g is a relevance function that evaluates the proximity of segment a_i to a context vector c, representing the current state of the time series. Eq. (10) illustrates how the model generates the inputs from past data points to predict the h future values. Once the model has been created and tested, the reference points are shifted forward by one cycle to predict the next cycle.

$$[\text{input}] \Big| [\text{output}] = \begin{bmatrix} x_{m-\alpha-h+1} & x_{m-\alpha-h+2} & x_{m-\alpha-h+3} & \cdots & x_{m-\alpha} \\ x_{m-\beta-h+1} & x_{m-\beta-h+2} & x_{m-\beta-h+3} & \cdots & x_{m-\beta} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ x_{m-\omega-h+1} & x_{m-\omega-h+2} & x_{m-\omega-h+3} & \cdots & x_{m-\omega} \end{bmatrix} \Bigg| \begin{bmatrix} x_{m-h+1} \\ x_{m-h+2} \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{bmatrix} \tag{10}$$

2.1.1. TS-MLP architecture

Each battery cycle is represented by five temperature segments, where each segment contains 1000 uniformly resampled points. These five segments are concatenated to form the TS-MLP input vector $\mathbf{x} \in \mathbb{R}^{5000}$. For compact illustration, the network diagram treats each segment as a single structured input channel. The prediction target is the temperature profile of the subsequent cycle, represented as a 1000-dimensional vector $\mathbf{y} \in \mathbb{R}^{1000}$. The architecture of the proposed TS-MLP used in this study is illustrated in Fig. 5.

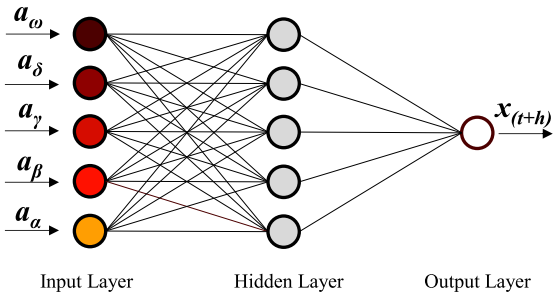

Fig. 5. TS-MLP architecture with five 1000-sample input segments, and a 1000-sample output layer.

Table 2
Summary of the NASA battery dataset experimental configuration.

Parameter	Description
Cell type	18 650 lithium-ion
Rated capacity	2 Ah
Ambient temperature	24 °C
Charge CC mode	1.5
Charge cut-off voltage	4.2 V
Charge CV mode	4.2 V
Charge cut-off current	20 mA
Discharge CC mode	At 2 A
Discharge cut-off voltage	2.7 V
EoL	30% capacity fade
Number of cycles	167
Operational modes	Charge, Discharge, Impedance

The TS-MLP comprises a single hidden layer with five *tansig* neurons, followed by a 1000-neuron *purelin* output layer. The inference cost includes one matrix–vector multiplication of size 5×5000 and one of size 1000×5, resulting in approximately 3×10^4 floating-point operations. This lightweight design facilitates real-time cycle-level inference on microcontroller-based BMS hardware.

2.2. Dataset acquisition

This study uses the publicly available NASA Prognostics Center of Excellence (PCoE) lithium-ion battery ageing dataset [27]. From this repository, cell B0005, an 18 650 cylindrical NCA Li-ion battery is utilised to develop the DT model. The rated capacity of the proposed cell is 2 Ah and the experiment was conducted at 24 °C ambient temperature under charge, discharge, and impedance operational modes. The operational procedures for cell B0005 are as follows:

- Charge:
  1- CC charging at 1.5 A until the cell voltage reaches 4.2 V
  2- CV charging at 4.2 V until the current decays to 20 mA
- Discharge:
  CC discharge at 2 A until the cell voltage reaches a 2.7 V
- Impedance:
  EIS was performed using a frequency sweep between 0.1 Hz and 5 kHz to capture ageing-related changes in internal resistance.

Test cycling continued until the end-of-life (EoL) criterion was reached, defined as a 30% reduction in rated capacity (from 2 Ah to 1.4 Ah). For cell B0005, a total of 167 charge–discharge cycles were recorded. Table 2 summarises the NASA battery dataset used in this study.

NASA datasets consist of multiple charge–discharge cycles, each containing a varying number of data points due to the differences in measurement duration and sampling rate. To ensure consistency, all time series were rescaled to a fixed length of 1000 points using MATLAB interpolation. This normalisation step ensures uniform data

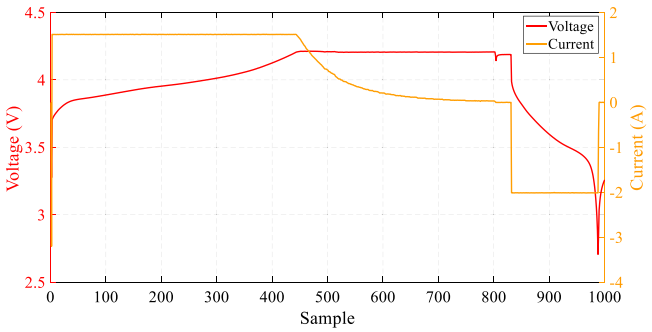

Fig. 6. Voltage and current profiles of the battery for Cycle 1.

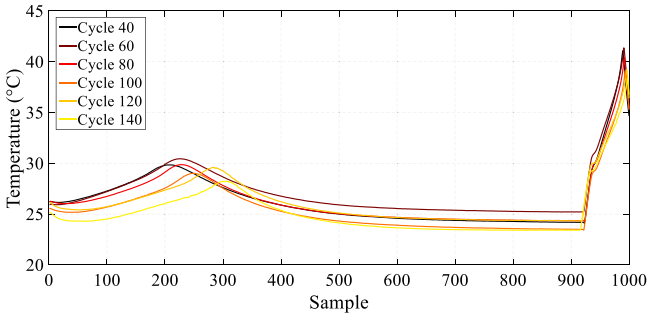

Fig. 7. Temperature profile of the battery at selected ageing cycles (40, 60, 80, 100, 120, and 140).

representation while avoiding any missing samples. For the first prediction, the training interval includes samples 1 to 2000, the validation interval covers samples 2001 to 3000, and the prediction is performed on samples 3001 to 4000. Similarly, for the second prediction, the training interval is extended from samples 1 to 3000, the validation interval spans samples 3001 to 4000, and the prediction interval is from sample 4001 to sample 5000.

Fig. 6 presents the voltage and current profiles from the NASA dataset employed in this study, illustrating the consistent duty cycle applied across all measured cycles. Fig. 7 shows that, although the voltage-current profile remains essentially identical throughout the battery's lifetime, the thermal behaviour gradually changes as the cell ages. This evolving temperature response highlights the need for a cycle-wise temperature prediction strategy within the proposed framework. Therefore, the cell surface temperature is selected as the key parameter for the proposed DT. To simulate the real-time data transmission from the physical counterpart to the DT, the data are fed into the model cycle by cycle. Each cycle consists of a time series of temperature samples, followed by a defined delay, representing the time required to receive the next cycle. The imported temperature time series, together with the stored cycles, is then processed by the GA to select the optimised reference points. A static learning rate was used during TS-MLP training, as the lightweight architecture and limited cycle-level data allow stable and consistent convergence without requiring any form of dynamic learning-rate scheduling.

2.3. Model optimisation

Once a full cycle is stored, the GA begins to identify the optimum reference points. The cost function of the GA is defined as the prediction error of the TS-MLP on the test data, with the model trained using the GA-suggested intervals.

2.3.1. Parent selection

In the proposed GA, parents are selected using the Roulette Wheel Selection (RWS), which provides a probabilistic strategy to prioritise candidates with higher fitness while maintaining population diversity [54]. Accordingly, normalised exponential probability functions are generated to rank candidates. The selection probability of candidate *i*, *Pi*, is given by:

$$P_i = \frac{e^{\left(\frac{-Bm_i}{W}\right)}}{\sum_{j=1}^{n} e^{\left(\frac{-Bm_j}{W}\right)}} \tag{11}$$

where B is the selection pressure, m_i is the Mean Absolute Error (MAE) of set i, n denotes the number of sets, and W is the worst MAE among the selected points. The cumulative probability range C_i is then created as:

$$C_i = \sum_{j=1}^{n} P_j \tag{12}$$

A random number r is then selected from a uniform distribution between 0 and 1. Candidate i is selected if r falls between cumulative probability interval, i.e., C_{i-1} to C_i. This approach ensures that candidates with higher selection probabilities are more likely to be selected, while preserving the population diversity.

2.3.2. Optimisation crossover

To keep the GA diverse, single-point, double-point, and uniform crossover methods are simultaneously utilised to generate offspring [46]. For each crossover, one of the proposed methods is randomly selected to generate offspring from the chosen parent sets.

2.3.3. Optimisation mutation

Mutation is applied by selecting a set of reference points randomly and replacing a part of it with other valid points.

Once parents, offspring, and mutants are generated, they are sorted with respect to their MAE, and the top k sets are selected as the population for the next GA iteration. Whenever the MAE of the top set is less than a threshold (0.01 in this study), it is reported as the optimal selection of reference points, and the trained model is subsequently used to predict next cycle values. Otherwise, the GA updates the reference points and repeats the process.

2.4. Anomaly detection

Since the proposed TS-MLP model is trained exclusively on nominal historical data, it is unable to recognise or respond to abnormal operating conditions. To address this limitation, a detailed anomaly detection framework is introduced in this section. The framework integrates an OC-SVM with a synthetic anomaly generation process and a Bayesian optimisation stage. This enables the DT to reliably distinguish abnormal patterns despite the absence of labelled fault data and the limited availability of nominal samples.

2.4.1. OC-SVM method

OC-SVM is a powerful tool ML for anomaly detection, which models normal data by mapping it into a high-dimensional feature space and defining a decision boundary. Points outside this boundary are considered anomalies. The OC-SVM model produces two outputs:

- *label*: a discrete value that serves as the anomaly indicator, taking +1 for nominal and −1 for anomalous samples.
- *score*: a continuous value representing the distance from the decision boundary, taking negative values for anomalies, positive values for nominal samples.

The OC-SVM problem can be formulated as follows:

$$min_{\mathbf{w},\rho,\nu} \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{\nu m}\sum_{i=1}^{m}\xi_i \tag{13}$$

$$\mathbf{w}\cdot\phi(\mathbf{x}_i) \geq \rho - \xi_i, \quad \forall i = 1,\ldots,m$$

$$\xi_i \geq 0, \quad \forall i = 1, \ldots, m$$

where $\mathbf{w}$ is the weight vector, ρ is the offset of the decision boundary, $\phi(\mathbf{x}_i)$ represents the feature mapping function, ξ_i are slack variables that allow margin violations, ν is the trade-off parameter that regulates the balance between the percentage of outliers and the margin width, and m is the number of training samples.

However, since OC-SCM is an unsupervised model, its accuracy becomes limited when only a small amount of nominal data is available. In such cases, the model cannot form a reliable boundary for distinguishing normal and abnormal behaviour. Therefore, it does not correctly classify all possible anomalies as abnormal. This motivates the need for synthetic anomaly generation and parameter optimisation.

2.4.2. Synthetic anomaly generation

The NASA dataset employed in this study contains only nominal data. Therefore, the anomaly detection model has no exposure to real anomalous samples. To overcome this limitation, a set of synthetic anomaly is generated. These synthetic samples are only intended to define a meaningful abnormal boundary for the proposed OC-SVM model, not to represent a specific real fault.

Several studies have reported a ±20% temperature variation threshold as an indicator of abnormal battery behaviour, associated with degradation [55–57]. Such variations create deviations clearly larger than normal fluctuations and sensor noise, while still being physically realistic.

To generated the synthetic anomalies, a GPR-based mapping was first established to model the relationship between temperature, voltage, and current within each cycle. Once the mapping is created, it is to estimate the corresponding voltage and current values for ±20% of temperature variations, ensuring that all three variables remain physically correlated. Using this mapping, two synthetic datasets are produced for each cycle, corresponding to the ±20% temperature deviations. These represent synthetic anomalies to be used for validating and optimising the anomaly detection method.

2.4.3. Parameter optimisation

These generated samples constitute the synthetic anomalies of each cycle and were treated as abnormal data for evaluating the OC-SVM. Because OC-SVM operates in an unsupervised manner, its initial performance on these anomaly samples is limited, and it does not correctly classify all of them as abnormal. To overcome this limitation, we employed Bayesian Optimisation (BO) to tune the two key OC-SVM parameters, ν which controls the proportion of anomalies and kernel scale (γ) which determines the influence of individual data points in the feature space. The optimisation process continued until all synthetic anomaly samples were assigned negative score values and reliably recognised as anomalies. Once this condition was satisfied, the OC-SVM model was considered fully optimised and ready for deployment.

The following steps are taken to employ the anomaly detection method:

- Train OC-SVM model using normal data from the first cycle.
- Update the model's support vectors using data from subsequent training cycles.
- Update support vectors employing data shifted by ±0.1% of the time frame as acceptable noise to enhance the accuracy and robustness of the model.
- Validate the trained model using the generated anomaly sets, where the model's *label* should be −1 for all the generated anomalies. Therefore, model parameters are optimised to ensure accurate anomaly detection.

2.4.4. OC-SVM evaluation

To evaluate the optimised OC-SVM in a more realistic setting, a total of 50 perturbed points were injected into each 1000-sample test sequence. Among these, 20 points were treated as true anomalies, generated by introducing ±20% deviations in the temperature signal using the trained GPR model, while the remaining 30 points represented acceptable fluctuations created using ±10% variations that should not be classified as anomalous behaviour. These perturbed samples were randomly distributed across the cycle to avoid any positional bias. The optimised OC-SVM was then applied to the augmented sequence to assess its ability to discriminate true anomalies from acceptable fluctuations.

Algorithm 1 DT framework - TS-MLP Prediction with OC-SVM anomaly detection

```
Initialise training dataset D_train ← ∅
Initialise OC-SVM parameters (ν, γ)
for each incoming cycle k do
    Read cycle k from NASA dataset
    Normalise cycle k to 1000 points
    if k ≥ 3 then
        T ← D_train
        V ← cycle k
        mse_val ← 0.01
        while mse_val > 0.01 do
            Select (α, β, γ, δ, ω)
            Train TS-MLP on T
            Validate TS-MLP on V
            Update mse_val
        end while
        Predict cycle k + 1 temperature
    end if
    Generate anomalies for cycle k
    Update OC-SVM with anomalies
    Flag ← Off
    for each sample x in cycle k + 1 do
        Compute OC-SVM label
        if label = −1 then
            Flag ← On
        end if
    end for
    Add cycle k to D_train
end for
```

To quantify the anomaly detection performance, standard classification metrics were employed, including precision, recall, and the F1-score, together with a confusion matrix. The confusion matrix summarises the classification outcomes in terms of true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN), where TP denotes correctly detected anomalies, FP denotes nominal samples incorrectly flagged as anomalous, FN denotes missed anomalies, and TN denotes nominal samples correctly classified as normal. Based on these quantities, the metrics are defined as:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{14}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{15}$$

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{16}$$

These metrics provide a comprehensive measure of the trade-off between correctly identifying anomalous events and avoiding false alarms. The numerical values obtained for the proposed framework are reported and discussed in Section 3.3.

Fig. 8 illustrates the overall strategy to create the DT, predict the parameters, and perform anomaly detection.

2.4.5. Model evaluation

The prediction accuracy across various cycles is evaluated using Mean Absolute Percentage Error (MAPE), Mean Square Error (MSE), RMSE, and R^2. MAPE quantifies the average relative error, while MSE and RMSE represent the magnitude of prediction errors. The R^2 metric indicates how well the model's predictions match the actual SoH. These metrics are defined as follows:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - y_i'}{y_i} \right| \times 100\% \tag{17}$$

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} \left(y_i - y_i' \right)^2 \tag{18}$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(y_i - y_i' \right)^2} \tag{19}$$

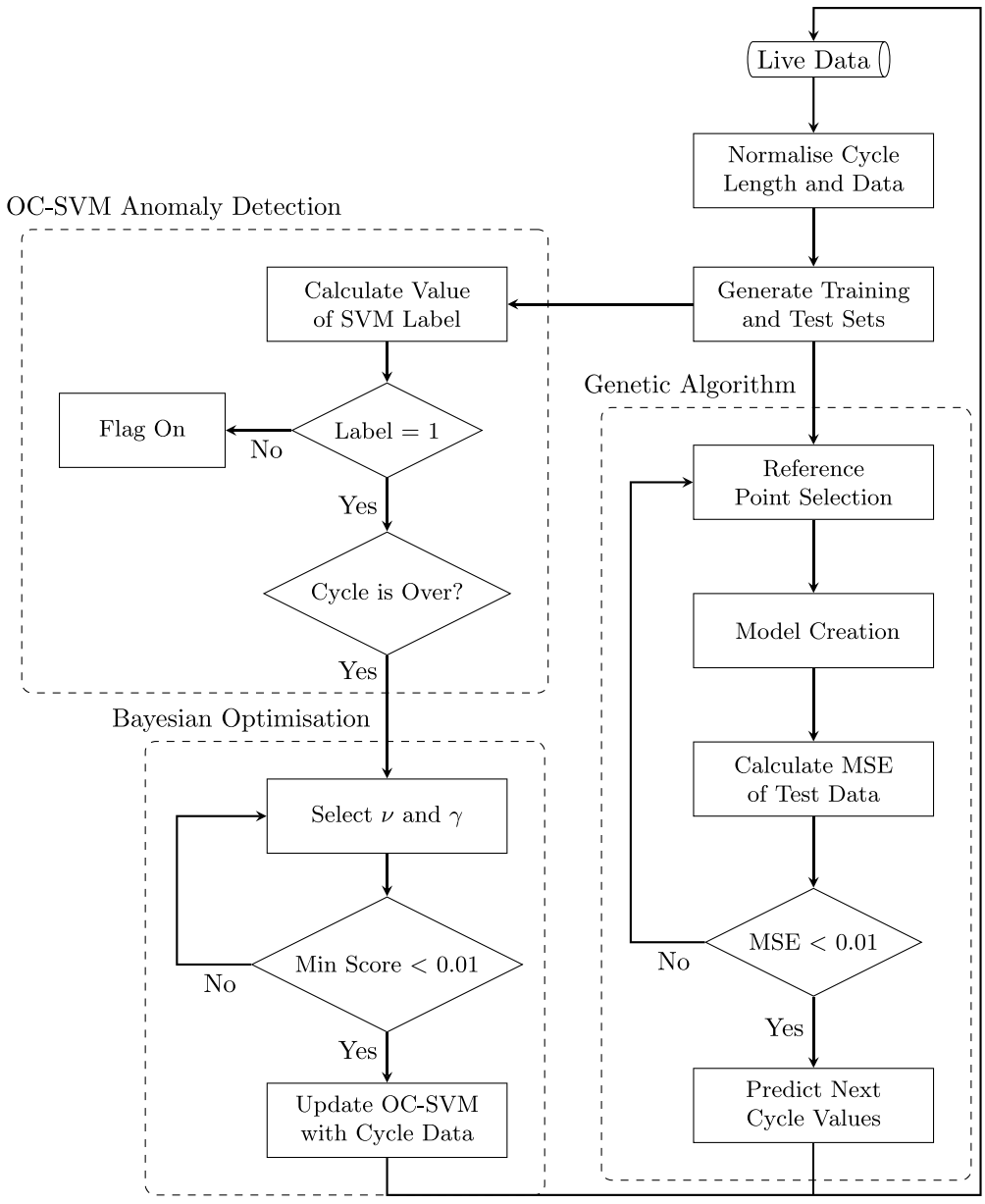

Fig. 8. Overall flowchart of the DT system integrating battery parameter prediction and anomaly detection.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(y_i - y_i'\right)^2}{\sum_{i=1}^{n} \left(y_i - \bar{y}\right)^2} \tag{20}$$

where y_i and y are respectively represent the actual and predicted values of ith sample, n is the number of samples.

2.5. Real BMS integration

The proposed framework is designed with practical BMS deployment as a key consideration. The proposed TS-MLP model requires only five scalar inputs, resulting in a low memory usage suitable for microcontroller-based BMS hardware that typically provides 64–512 kB of RAM. Also, since the MLP architecture is simple, real-time execution remains computationally efficient. The GA optimisation is executed only once per charge–discharge cycle rather than at high frequency, allowing it to run either on the embedded processor or to be offloaded to a supervisory controller. Furthermore, the required measurements, including cell temperature, voltage, and current are typical signals available in commercial BMS units. The cycle level processing used in this study is fully compatible with typical BMS sampling rates (1–10 Hz). Therefore, the proposed DT-based method aligns well with the constraints of real BMS hardware in terms of memory, computational load, and data availability.

It should be noted that the NASA battery datasets used in this study are offline measurements. Therefore, the proposed framework represents a simulation-stage DT, in which cycle-level updating and verification are performed using real measured cycles rather than real-time hardware interaction.

3. Results

In this study, the Digital Twin framework is validated using real experimental cycles from the NASA 18 650 dataset. The predictive component is evaluated on held-out real cycles, demonstrating that the DT can accurately reproduce the thermal behaviour of the physical cell. The diagnostic layer is assessed through anomaly-injected thermal traces, where ±10% fluctuations represent acceptable variations and ±20% deviations represent true anomalies. Finally, the DT performs

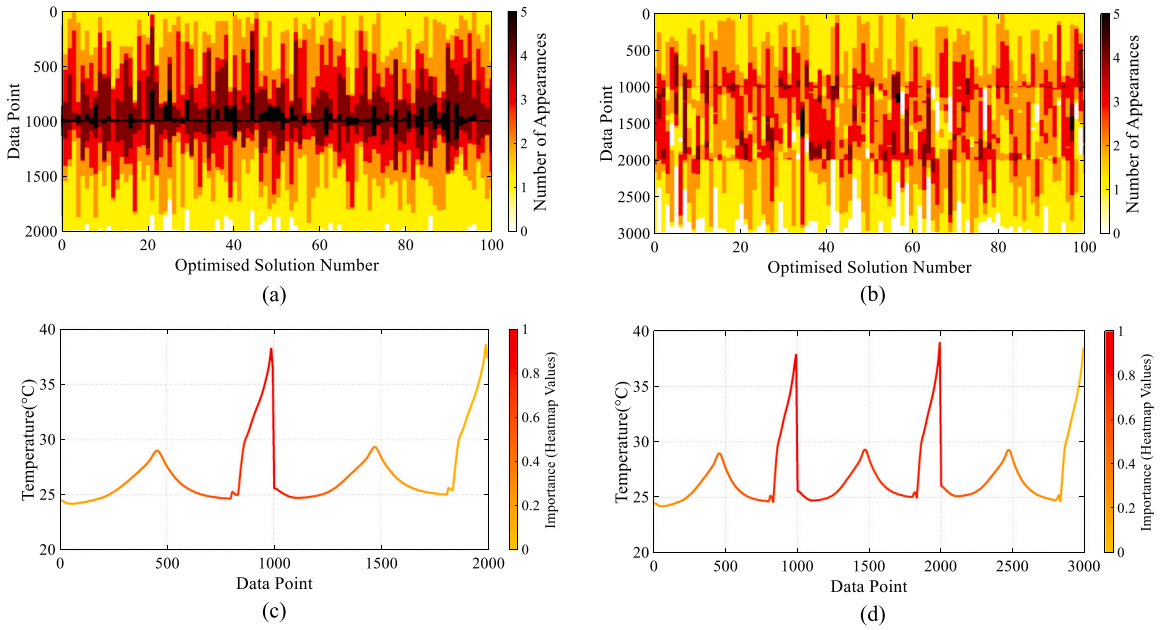

Fig. 9. Reference points optimisation performed with the GA. Heatmaps of optimally selected data intervals across 100 independent runs for the first (a) and second (b) prediction. Composite heatmap showing the frequency of interval selection, overlaid on the original signal for the first (b) and the second (c) prediction.

cycle-by-cycle synchronisation with real measured data, allowing the virtual model to track the battery's long-term evolution despite the absence of real-time sensor feedback.

3.1. Optimisation of reference points

This section reports the results obtained from optimising the reference points. Fig. 9(a) and (b) show heatmaps of selected intervals derived from 100 runs of the GA optimisation for the first and second predictions, respectively. Each column indicates which intervals of the training dataset were selected as inputs for the TS-MLP model. Regions of higher intensity represent intervals selected more frequently, indicating their greater influence during training and, consequently, on the prediction process.

In Fig. 9(a), a cluster around the temperature peak of cycle 1 (data point 900–1100) is highlighted. Fig. 9(b) presents clusters of selected points around the end of cycle 1 and the end of cycle 2 (800–1000 and 1500–2000), with only limited selection from the charging interval of cycles 1 and after cycle 2. As a result, the optimisation process consistently selects temperature fluctuations caused by discharge as inputs across all five intervals, reflecting the most dynamic thermal behaviour. Consequently, these regions are more frequently selected for training, contributing to improved prediction accuracy.

Fig. 9(c) and (d) show the selected training data points superimposed on the normalised heatmaps. Fig. 9(c) shows that the optimiser predominantly selected points within the discharging phase of the first cycle, followed by regions of lower temperatures corresponding to the subsequent charging intervals. In Fig. 9(d), the model selected temperatures from the discharging phases of the first two cycles, as well as the intermediate charging phase. As additional cycles are provided, the model shifts its focus to more recent cycles, as they contain the most up-to-date information to predict future behaviour.

To illustrate the feature importance within the training dataset, the weighted importance maps of the selected training samples are shown in Fig. 10 for the first prediction (cycle 4) and in Fig. 11 for the second prediction (cycle 5). In each optimisation run, the GA selects five reference features from the training data; therefore, across 100 optimisation runs, a total of 500 selections are accumulated. The weighted importance maps visualise the frequency with which each temperature position is selected and directly reflect the behaviour observed in Fig. 9(c)–(d).

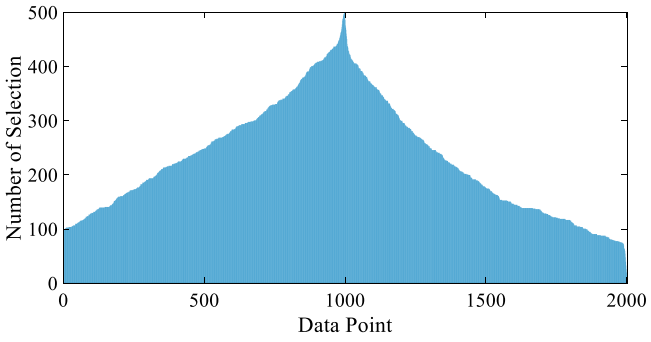

Fig. 10. Histogram of selection frequency of candidate training points for the first prediction across 100 optimisation runs.

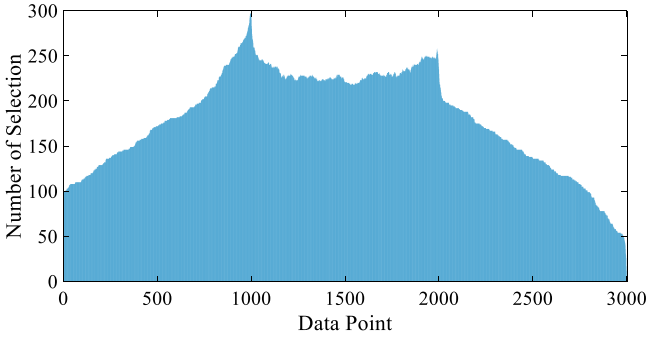

Fig. 11. Histogram of selection frequency of candidate training points for the second prediction across 100 optimisation runs.

For the first prediction, Fig. 10 shows that the highest selection frequencies cluster around the peak-temperature region of the first cycle, indicating that this region carries the strongest predictive information. For the second prediction, Fig. 11 highlights multiple highly

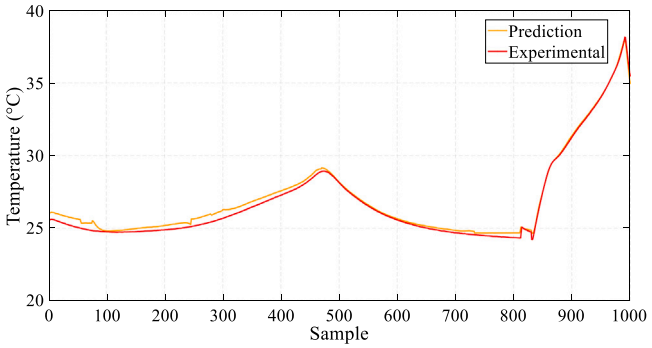

Fig. 12. Prediction of cycle 4 temperature using the proposed DT model.

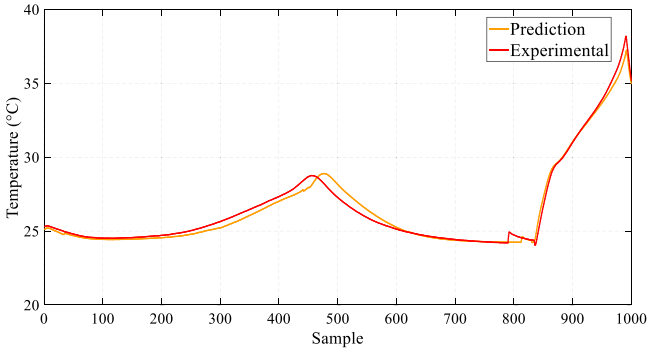

Fig. 13. Prediction of cycle 5 temperature using the proposed DT model.

informative regions, including the peaks of both the first and second cycles as well as the transition interval between them.

These maps therefore provide an explicit and interpretable representation of feature importance, confirming that the GA consistently selects the thermally active and high-gradient regions that contribute most strongly to next-cycle temperature prediction.

3.2. Temperature prediction

Figs. 12 and 13 compare the predicted and experimental temperature profiles for cycles 4 and 5, respectively. In Fig. 12, the temperature prediction of cycle 4 shows a satisfactory level of accuracy across both the charging and discharging phases. As the number of training cycles increases, Fig. 13 demonstrates improved accuracy in predicting cycle 5, with only a slight peak shift and a corresponding underestimation in the region following the charging phase.

To evaluate the generalisation capability of the proposed framework, an additional validation was performed using NASA Battery B0033, which was cycled under harsher discharge conditions (4 A discharge to a 2.0 V cut-off). Fig. 14 presents the comparison between the predicted and measured temperature profiles for four representative future-cycle predictions, including early-, mid-, and late-life stages (cycles 5, 15, 55, and 105). In all cases, the proposed GA–TS-MLP model successfully tracks the real thermal behaviour of the cell, confirming the robustness of the method under different operating conditions.

Table 3 presents the DT-based cycle-by-cycle temperature prediction accuracy for the first four cycles. As DT progresses through more cycles, the training dataset grows. Therefore, the selected inputs can capture long-term variations, which results in improving the prediction accuracy. From cycles 4 to 7, the effect on R^2 is negligible; however, both MAPE and MSE decrease significantly by more than 60%, while RMSE shows a 47% reduction.

Fig. 15 and Table 4 evaluates the performance of the proposed TS-MLP methods while the reference points are selected through two

Table 3
Prediction performance for different target cycles.

Prediction target cycle	Coefficient of determination	Prediction error		
	R^2	MAPE	MSE	RMSE
Cycle 4	0.9913	0.8825	0.0770	0.2775
Cycle 5	0.9906	0.8126	0.0848	0.2913
Cycle 6	0.9974	0.3674	0.0227	0.1664
Cycle 7	0.9976	0.3071	0.0213	0.1461

Table 4
Comparison between GA Optimisation and Random Selection methods across different cycles.

Cycle	Method	RMSE	MSE	MAPE
4	GA Optimisation	0.2775	0.0770	0.8825
	Random Selection	0.6795	0.4616	2.2522
5	GA Optimisation	0.2913	0.0848	0.8126
	Random Selection	0.7819	0.6114	2.4722
6	GA Optimisation	0.1664	0.0227	0.3674
	Random Selection	0.9713	0.9434	2.5399
7	GA Optimisation	0.1461	0.0213	0.3071
	Random Selection	0.7463	0.5569	2.1038

different methods: 1-GA Optimisation 2-Random Selection. The comparison evaluates RMSE, MSE, and MAPE across cycles 4–7. Table 4 demonstrates that the GA-based approach consistently outperforms the random selection method in all three indicators. Averaged over cycles 4–7, the GA-based optimisation achieves a 72% reduction in RMSE, 92% reduction in MSE, and 75% reduction in MAPE compared to the random selection method.

Moreover, the trend analysis based on Fig. 15 shows that:

- The GA-based optimisation demonstrates clear superiority, achieving a 47% reduction in RMSE from 0.2775 °C to 0.1461 °C, a 72% reduction in MSE from 0.0770 to 0.0213, and a 65% reduction in MAPE from 0.8825% to 0.3071%.
- In contrast, the random selection method exhibited an increase of 10% in RMSE from 0.6795 to 0.7463, an increase of 20% in MSE from 0.46 to 0.55, and only a minor 7% reduction in MAPE from 2.25% to 2.10%.

In addition, the trend analysis across consecutive cycles shows that in the GA-based method, all error indicators decrease steadily, reflecting continuous learning and improved prediction stability. Conversely, the random selection approach exhibits fluctuating or increasing error trends, indicating unstable behaviour and lack of convergence.

To highlight the advantages of DT-based ML approach, it is compared against a conventional ML, with a 100-cycle horizon. Both the DT-based TS-MLP and conventional TS-MLP were initially trained on the same dataset. While the conventional TS-MLP model remained static, the DT-based TS-MLP was continuously updated with live data cycle by cycle. Fig. 16 shows the RMSE for both methods. As the cycle number increases, the prediction error of the conventional method gradually increases, whereas the proposed TS-MLP model showed a consistent level of error below the 0.5 °C threshold.

To provide a clearer assessment of the relative performance of the proposed approach, a baseline comparison was conducted using representative temperature-prediction models reported in the literature. Although only a limited number of studies have focused specifically on temperature prediction for NASA battery datasets and even fewer have performed prediction based solely on early-cycle information the extracted RMSE values offer a meaningful point of reference. As summarised in Table 5, the proposed DT-MLP achieves a substantially lower mean RMSE (0.2919 °C) compared with the GPR+RNN, GRU, and LSTM-based baselines, demonstrating the improved prediction accuracy and efficiency of the proposed framework.

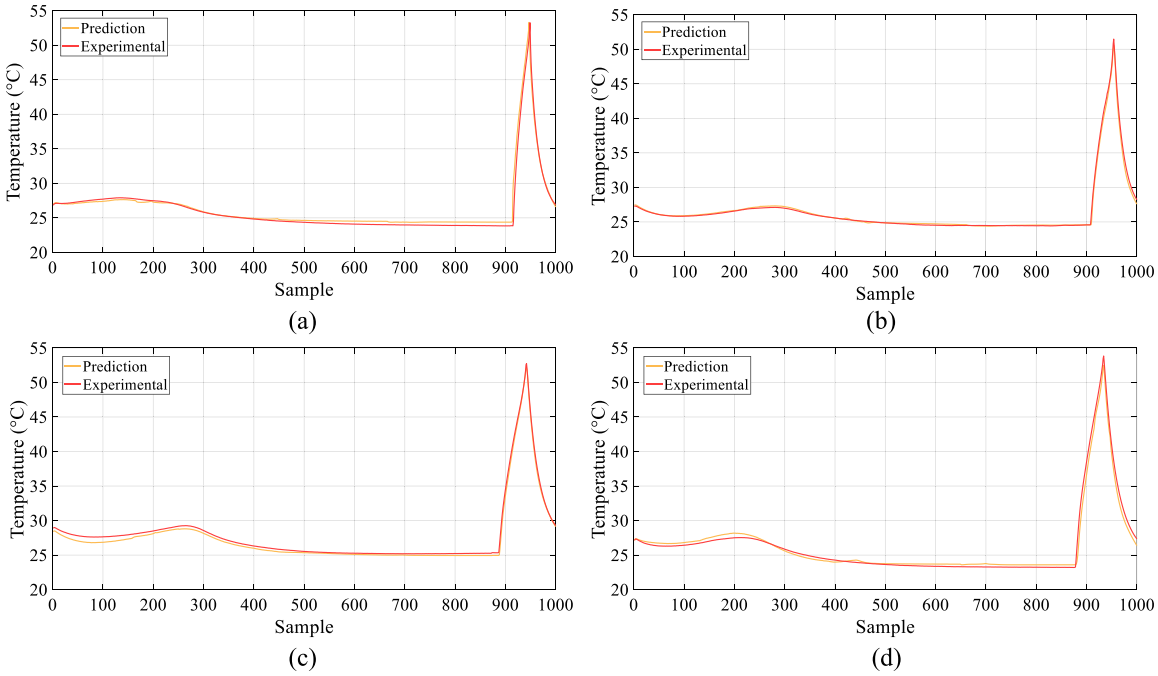

Fig. 14. Temperature prediction on NASA Battery B0033 dataset. (a) Prediction of cycle 5, (b) prediction of cycle 15, (c) prediction of cycle 55, and (d) prediction of cycle 105.

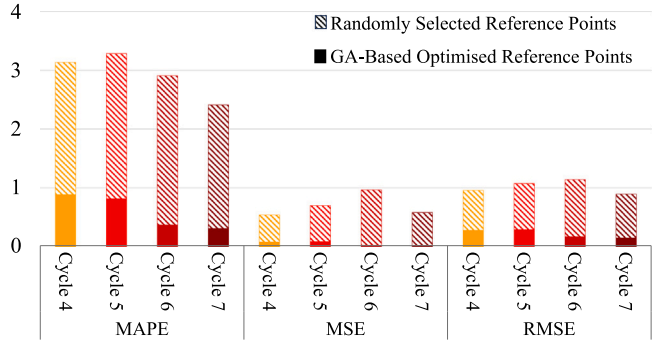

Fig. 15. Accuracy comparison between the proposed DT and randomly selected reference points.

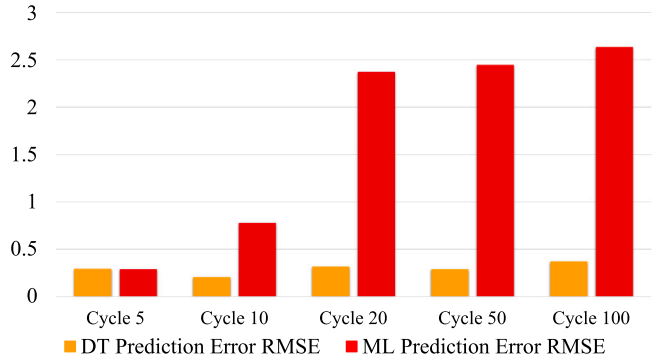

Fig. 16. RMSE comparison between the DT-based and conventional TS-MLP.

3.3. Anomaly detection

Figs. 17 and 18 illustrate the voltage and current profiles of the first cycle, estimated by the GPR based on ±20% temperature variation, respectively. These predicted boundaries reflect the expected deviations

Table 5
Comparison of temperature-prediction RMSE with baseline studies.

Author(s)	Method	RMSE (°C)
Yao et al. [58]	GPR + RNN	0.7411
YuHeng et al. [59]	GRU	0.7976
YuHeng et al. [59]	LSTM	0.9806
Faraji-Niri et al. [60]	CNN + biLSTM	0.5500
This study	**DT-MLP**	**0.2919**

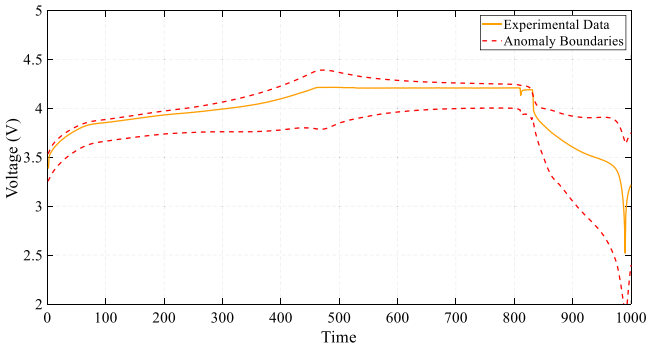

Fig. 17. Voltage profile of the first cycle with anomaly boundaries estimated using the GPR approach.

in voltage and current resulting from significant temperature fluctuations, and serve as a synthetic reference for evaluating anomalous behaviour in subsequent cycles.

Fig. 19 presents the OC-SVM *score* values for the two synthetic anomaly sets generated under ±20% temperature variation. The OC-SVM parameters were optimised using Bayesian Optimisation (BO), which adjusts the kernel scale and the ν parameter to achieve an appropriate decision boundary. After optimisation, all samples in both sets exhibit negative *score* values, confirming that the model accurately classifies them as anomalies. These results confirm that the proposed OC-SVM framework, combined with BO parameter tuning, provides a

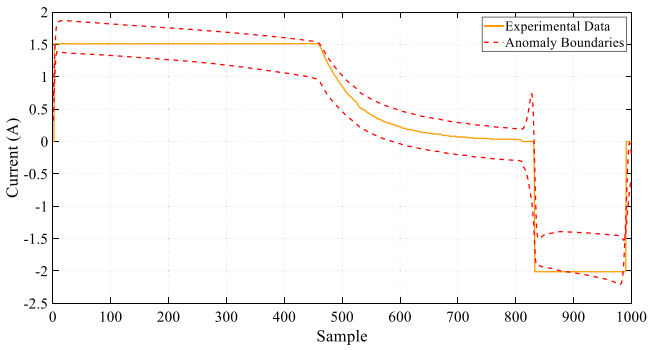

Fig. 18. Current profile of the first cycle with anomaly boundaries estimated using the GPR approach.

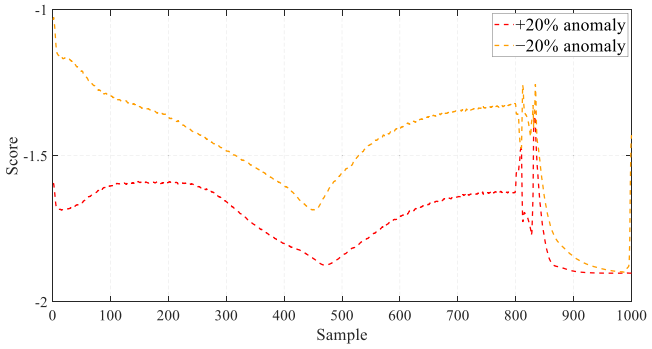

Fig. 19. Score values of the two sets of generated anomalies.

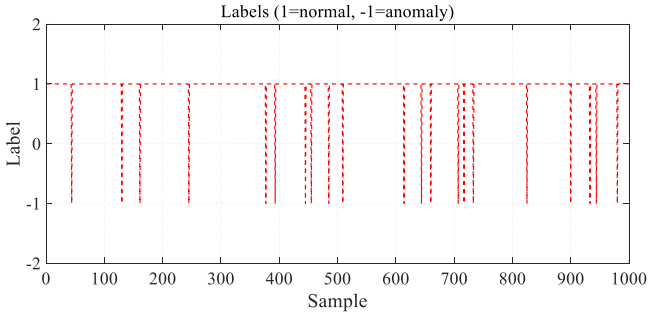

Fig. 20. One-Class SVM anomaly detection method.

reliable anomaly detection mechanism when real fault annotations are not available.

The optimised OC-SVM was evaluated on an augmented 1000-sample test sequence containing 50 perturbed points. Among these, 20 points corresponded to true anomalies generated using ±20% deviations in the temperature signal, while the remaining 30 points represented acceptable fluctuations created using ±10% variations. Fig. 20 illustrates the anomaly detection results of the OC-SVM model for the proposed noisy signal.

As Fig. 20 shows, the model successfully detected all 20 true anomalies, and only one out of the 30 acceptable fluctuations was incorrectly classified as anomalous. The resulting confusion matrix is:

$$\begin{bmatrix} TN = 979 & FP = 1 \\ FN = 0 & TP = 20 \end{bmatrix}$$

This classification outcome corresponds to a precision of 0.952, a recall of 1.000, and an overall F1-score of 0.976. These values indicate that the proposed framework retains perfect sensitivity to true anomalous events while maintaining a very low false-alarm rate, thereby demonstrating the robustness and reliability of the optimised OC-SVM under noisy operating conditions.

4. Conclusion

This study presented a novel approach for predicting battery parameters using a live DT-based framework. A TS-MLP model was employed to accurately estimate cell surface temperatures after only a few initial cycles, thereby overcoming the need for large, time-consuming training datasets. Input intervals for the TS-MLP were optimised through a GA, which selected the most informative intervals to minimise prediction error. Once the test prediction error was below 1%, the TS-MLP was used to predict parameter values for subsequent cycles. In addition, an OC-SVM optimised through BO was employed to detect synthetic anomalies generated using a GPR-based approach, enhancing prediction reliability. The methodology was trained and validated using the NASA open-source dataset for a 18 650 cylindrical NCA cell.

Key conclusions from the analysis are as follows:

- The optimisation consistently selected input points during the discharge phase, ensuring that the most informative regions are used for training.
- After seven initial cycles, the prediction error RMSE decreased by nearly 50% compared to the first prediction (cycle 4).
- GA-optimised reference points outperformed random selection, often reducing prediction errors by an order of magnitude.
- The DT-based TS-MLP maintained prediction errors below 0.5 °C over 100 cycles, outperforming the conventional TS-MLP.
- The optimised OC-SVM successfully identified and correctly labelled all 10 randomly generated anomalies.

The method can be extended to predict additional parameters, with future work focusing on estimating the cell's SoH. Furthermore, an experimental dataset containing real labelled anomalies will be generated for a similar cell to validate the proposed methodology. Another future extension of this work is to conduct an economic impact assessment of the proposed DT-based framework, including potential cost savings from early anomaly detection and reduced maintenance requirements.

A limitation of the current approach is that, as the training dataset expands, the process becomes increasingly time-consuming and computationally inefficient. Future work could explore optimising the number of training cycles once the desired prediction accuracy is achieved, thereby improving computational efficiency as the dataset grows.

List of Abbreviations

ANN	ANNArtificial Neural Network
BMS	Battery Management Systems
BO	Bayesian Optimisation
CC	Constant Current
CNN	Convolutional Neural Network
CV	Constant Voltage
DT	Digital Twin
ECM	Equivalent Circuit Model
EKF	Extended Kalman Filter
EoL	End of Life
GA	Genetic Algorithm
GPR	Gaussian Process Regression
GRG	Grey Relational Grade
IF	Isolation Forest
IoT	Internet of Things
LAM	loss of active material
LFP	Lithium Iron Phosphate
LLI	loss of lithium inventory
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error

MAPE	Mean Absolute Percentage Error
ML	Machine Learning
MLP	Multi-Layer Perceptron
MSE	Mean Square Error
NCA	Nickel-Cobalt-Aluminum
OC-SVM	One-Class Support Vector Machine
PCC	Pearson Correlation Coefficient
RF	Random Forest
RMSE	Root Mean Square Error
RUL	Remaining Useful Life
RWS	Roulette Wheel Selection
SEI	Solid Electrolyte Interface
SoC	State of Charge
SoH	State of Health
SPM	Single Particle Model
SVM	Support Vector Machine
SVR	Support Vector Regression
TS-MLP	Time-Series Multi-Layer Perceptron
XGBoost	Extreme Gradient Boosting

CRediT authorship contribution statement

Sajad Najafi-Shad: Writing – original draft, Visualisation, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualisation. **Davide Domenico Sciortino:** Writing – review & editing, Supervision, Methodology, Conceptualisation. **Shahaboddin Resalati:** Supervision. **Paul Henshall:** Supervision. **Denise Morrey:** Supervision, Project administration.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT for linguistic editing, including grammar correction, stylistic refinement, and exploring alternative formulations informed by existing literature. Generative AI was not used to conduct data analysis or derive results. After using this tool/service, the authors reviewed and edited the content as needed and takes full responsibility for the content of the published article.

Funding

This research received no external funding.

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.

Acknowledgements

The authors wish to thank the Centre for Batteries, Electric Vehicles and Electronics (CBEVE) group at Oxford Brookes University, United Kingdom for their support throughout this research.

Data availability

Data will be made available on request.

References

[1] L. Li, B. Lei, C. Mao, Digital twin in smart manufacturing, J. Ind. Inf. Integr. 26 (2022) 100289.
[2] A. Fuller, Z. Fan, C. Day, C. Barlow, Digital twin: Enabling technologies, challenges and open research, IEEE Access 8 (2020) 108952–108971.
[3] R.S. Nuvvula, P.P. Kumar, S.R. Ahammed, R.S.S. Reddy, B.H. Babu, A. Ali, et al., Distributed reinforcement learning framework for autonomous optimization of grid-scale energy storage systems in renewable energy integration, in: 2024 12th International Conference on Smart Grid, IcSmartGrid, IEEE, 2024, pp. 290–297.
[4] P.P. Kumar, R.S. Nuvvula, S.A. Shezan, B. JM, S.R. Ahammed, A. Ali, et al., Intelligent energy management system for microgrids using reinforcement learning, in: 2024 12th International Conference on Smart Grid, IcSmartGrid, IEEE, 2024, pp. 329–335.
[5] C. Gbaja, Next-generation edge computing: Leveraging ai-driven iot for autonomous, real-time decision making and cyber security, J. Artif. Intell. Gen. Sci. (JAIGS) (ISSN: 3006-4023) 5 (1) (2024) 357–371.
[6] M. Mollashahi, P. Jafari, M. Mehrjoo, A multi-stream deep neural network to predict the energy consumption of smart home appliances, Int. J. Comput. Intell. Appl. 23 (02) (2024) 2450003.
[7] X. Qu, Y. Song, D. Liu, X. Cui, Y. Peng, Lithium-ion battery performance degradation evaluation in dynamic operating conditions based on a digital twin model, Microelectron. Reliab. 114 (2020) 113857.
[8] E. Michelini, P. Höchele, C. Ellersdorfer, J. Moser, Impact of prolonged electrochemical cycling on health indicators of aged lithium-ion batteries for a second-life use, IEEE Access (2024).
[9] Z. Deng, L. Xu, H. Liu, X. Hu, Z. Duan, Y. Xu, Prognostics of battery capacity based on charging data and data-driven methods for on-road vehicles, Appl. Energy 339 (2023) 120954.
[10] F. Xia, K. Wang, J. Chen, State of health and remaining useful life prediction of lithium-ion batteries based on a disturbance-free incremental capacity and differential voltage analysis method, J. Energy Storage 64 (2023) 107161.
[11] M.S. Trimboli, G.L. Plett, Perspectives on employing physics-based battery management for electric vehicles, in: 2024 International Conference on Electrical, Computer and Energy Technologies, ICECET, IEEE, 2024, pp. 1–6.
[12] M. Ohrelius, R. Lindstrom, G. Lindbergh, Simplified physics-based battery model for stationary energy-storage applications, in: Electrochemical Society Meeting Abstracts 245, The Electrochemical Society, Inc., 2024, 249–249.
[13] A. Aitio, D.A. Howey, Predicting battery end of life from solar off-grid system field data using machine learning, Joule 5 (12) (2021) 3204–3220.
[14] M.A. Mohamed, T.F. Yu, G. Ramsden, J. Marco, T. Grandjean, Advancements in parameter estimation techniques for 1RC and 2RC equivalent circuit models of lithium-ion batteries: A comprehensive review, J. Energy Storage 113 (2025) 115581.
[15] J. Wen, L. Wang, X. He, Navigating the intricacies: A critical review of numerical modeling in battery research and design, J. Power Sources 628 (2025) 235902.
[16] F. Li, Z. Li, Y. Zhang, G. Xu, X. Wang, H. Zhang, Analysis and verification of equivalent circuit model of soft-pack lithium batteries, Energies 18 (3) (2025) 510.
[17] S. Najafi-Shad, S.M. Barakati, Intelligent mppt and dc-link voltage control in the hybrid pv-dfig system, in: 2020 28th Iranian Conference on Electrical Engineering, ICEE, IEEE, 2020, pp. 1–5.
[18] B.A.A. Acurio, D.E.C. Barragán, J.C. Rodríguez, F. Grijalva, L.C. Pereira da Silva, Robust data-driven state of health estimation of lithium-ion batteries based on reconstructed signals, Energies 18 (10) (2025) 2459.
[19] S. Buchanan, C. Crawford, Probabilistic lithium-ion battery state-of-health prediction using convolutional neural networks and Gaussian process regression, J. Energy Storage 76 (2024) 109799.
[20] D. Swain, M. Kumar, A. Nour, K. Patel, A. Bhatt, B. Acharya, A. Bostani, Remaining useful life predictor for EV batteries using machine learning, IEEE Access (2024).
[21] P. Li, X. Wu, R. Grosu, J. Hou, M. Ilolov, S. Xiang, Applying neural network to health estimation and lifetime prediction of lithium-ion batteries, IEEE Trans. Transp. Electrification (2024).
[22] Z. Sherkatghanad, A. Ghazanfari, V. Makarenkov, A self-attention-based CNN-Bi-LSTM model for accurate state-of-charge estimation of lithium-ion batteries, J. Energy Storage 88 (2024) 111524.
[23] C. She, Y. Shen, G. Bin, J. Hong, Y. Peng, Accurate state of health estimation of battery system based on multi-stage constant current charging and behavior analysis in real-world electric vehicles, IEEE Trans. Transp. Electrification (2024).
[24] Z. Deng, X. Lin, J. Cai, X. Hu, Battery health estimation with degradation pattern recognition and transfer learning, J. Power Sources 525 (2022) 231027.
[25] R. Xiong, Y. Sun, C. Wang, J. Tian, X. Chen, H. Li, Q. Zhang, A data-driven method for extracting aging features to accurately predict the battery health, Energy Storage Mater. 57 (2023) 460–470.
[26] Q. Li, D. Li, K. Zhao, L. Wang, K. Wang, State of health estimation of lithium-ion battery based on improved ant lion optimization and support vector regression, J. Energy Storage 50 (2022) 104215.
[27] B. Saha, Battery data set, in: NASA AMES Progmostics Data Repository, 2007.

[28] A. Tiane, C. Okar, H. Chaoui, Fault-tolerant framework for state-of-health prediction of lithium batteries, IEEE Trans. Power Electron. 37 (12) (2022) 15528–15537.
[29] G. Kim, J.J. Jung, D.K. Kim, M. Koo, G.J. Nalepa, S. Nowaczyk, Fuzzy particle filtering based approach for battery RUL prediction with uncertainty reduction strategies, Expert Syst. 42 (4) (2025) e70027.
[30] S. Ansari, M. Hannan, A. Ayob, M.G. Abdolrasol, M.A. Dar, Co-estimation of state of health and remaining useful life for lithium-ion batteries using a hybrid optimized framework, J. Energy Storage 105 (2025) 114711.
[31] Z. Dai, A. Li, W. Sun, S. Zhang, H. Zhou, R. Rao, Q. Luo, Estimation and prediction method of lithium battery state of health based on ridge regression and gated recurrent unit, IET Energy Syst. Integr. (2024).
[32] Y. He, W. Bai, L. Wang, H. Wu, M. Ding, SOH estimation for lithium-ion batteries: An improved GPR optimization method based on the developed feature extraction, J. Energy Storage 83 (2024) 110678.
[33] V. Selvaraj, I. Vairavasundaram, A Bayesian optimized machine learning approach for accurate state of charge estimation of lithium ion batteries used for electric vehicle application, J. Energy Storage 86 (2024) 111321.
[34] Y. Jiang, Y. Xu, Y. Yang, F. Yang, P. Wang, C. Li, F. Zhuang, H. Xiong, TriMLP: A foundational MLP-like architecture for sequential recommendation, ACM Trans. Inf. Syst. 42 (6) (2024) 1–34.
[35] S. Ryu, Y. Yu, Quantile-mixer: A novel deep learning approach for probabilistic short-term load forecasting, IEEE Trans. Smart Grid (2023).
[36] G. Ye, L. Feng, J. Guo, Y. Chen, IIP-mixer: Intra–Inter-patch mixing architecture for battery remaining useful life prediction, Energies 17 (14) (2024) 3553.
[37] L. Dang, J. Yang, M. Liu, B. Chen, Differential equation-informed neural networks for state of charge estimation, IEEE Trans. Instrum. Meas. (2023).
[38] S. Kwak, J. Jeong, H. Lee, W. Kim, D. Seo, W. Yun, W. Lee, J. Shin, Few-shot anomaly detection via personalization, IEEE Access (2024).
[39] V. Mariani, G. Adinolfi, A. Buonanno, R. Ciavarella, A. Ricca, V. Sorrentino, G. Graditi, M. Valenti, A survey on anomalies and faults that may impact the reliability of renewable-based power systems, Sustain. 16 (14) (2024) 6042.
[40] R. Jin, B. Wei, Y. Luo, T. Ren, R. Wu, Blockchain-based data collection with efficient anomaly detection for estimating battery state-of-health, IEEE Sens. J. 21 (12) (2021) 13455–13465.
[41] J. Jeon, H. Cheon, B. Jung, H. Kim, ProADD: Proactive battery anomaly dual detection leveraging denoising convolutional autoencoder and incremental voltage analysis, Appl. Energy 373 (2024) 123757.
[42] T. Zhang, R. Li, Y. Zhou, Battery fault diagnosis method based on online least squares support vector machine, Energies 16 (21) (2023) 7273.
[43] H. Liu, S. Hao, T. Han, F. Zhou, G. Li, Random forest-based online detection and location of internal short circuits in lithium battery energy storage systems with limited number of sensors, IEEE Trans. Instrum. Meas. 72 (2023) 1–11.
[44] Z. Yuan, Y. Pan, H. Wang, S. Wang, Y. Peng, C. Jin, C. Xu, X. Feng, K. Shen, Y. Zheng, et al., Fault data generation of lithium ion batteries based on digital twin: A case for internal short circuit, J. Energy Storage 64 (2023) 107113.
[45] S. Jafari, Y.-C. Byun, Prediction of the battery state using the digital twin framework based on the battery management system, IEEE Access 10 (2022) 124685–124696.
[46] D.-C. Dang, A. Opris, D. Sudholt, Crossover can guarantee exponential speed-ups in evolutionary multi-objective optimisation, Artificial Intelligence 330 (2024) 104098.
[47] I. Saba, M. Ullah, M. Tariq, Advancing electric vehicle battery analysis with digital twins in intelligent transportation systems, IEEE Trans. Intell. Transp. Syst. (2024).
[48] Z. Deng, L. Xu, H. Liu, X. Hu, B. Wang, J. Zhou, Rapid health estimation of in-service battery packs based on limited labels and domain adaptation, J. Energy Chem. 89 (2024) 345–354.
[49] R.D. Fonso, R. Teodorescu, C. Cecati, P. Bharadwaj, A battery digital twin from laboratory data using wavelet analysis and neural networks, IEEE Trans. Ind. Inform. 20 (4) (2024) 6889–6899.
[50] Y. Wang, R. Xu, C. Zhou, X. Kang, Z. Chen, Digital twin and cloud-side-end collaboration for intelligent battery management system, J. Manuf. Syst. 62 (2022) 124–134.
[51] Z. Tao, Z. Zhao, C. Wang, L. Huang, H. Jie, H. Li, Q. Hao, Y. Zhou, K.Y. See, State of charge estimation of lithium batteries: Review for equivalent circuit model methods, Meas. 236 (2024) 115148.
[52] F. Naseri, S. Gil, C. Barbu, E. Cetkin, G. Yarimca, A. Jensen, P.G. Larsen, C. Gomes, Digital twin of electric vehicle battery systems: Comprehensive review of the use cases, requirements, and platforms, Renew. Sustain. Energy Rev. 179 (2023) 113280.
[53] D. Sciortino, A. Costa, R. Norris, K. Hayatleh, P. Henshall, Determination of the anisotropic thermal conductivity and specific heat capacity of a li-ion cell, 2024, http://dx.doi.org/10.13140/RG.2.2.16535.18089.
[54] J. Zhang, D. Chen, Q. Yang, Y. Wang, D. Liu, S.-W. Jeon, J. Zhang, Proximity ranking-based multimodal differential evolution, Swarm Evol. Comput. 78 (2023) 101277.
[55] R. Xiong, S. Ma, H. Li, F. Sun, J. Li, Toward a safer battery management system: A critical review on diagnosis and prognosis of battery short circuit, Iscience 23 (4) (2020).
[56] E. Jove, J.-L. Casteleiro-Roca, H. Quintián, F. Zayas-Gato, G. Vercelli, J.L. Calvo-Rolle, A one-class classifier based on a hybrid topology to detect faults in power cells, Log. J. IGPL 30 (4) (2022) 679–694.
[57] A.M. Mutua, R. de Fréin, Quantum-enhanced battery anomaly detection in smart transportation systems, Appl. Sci. 15 (17) (2025) 9452.
[58] Q. Yao, D.D.-C. Lu, G. Lei, A surface temperature estimation method for lithium-ion battery using enhanced GRU-RNN, IEEE Trans. Transp. Electrification 9 (1) (2022) 1103–1112.
[59] Y. Jiang, Y. Yu, J. Huang, W. Cai, J. Marco, Li-ion battery temperature estimation based on recurrent neural networks, Sci. China Technol. Sci. 64 (6) (2021) 1335–1344.
[60] M. Faraji-Niri, T.M. Bui, T.F. Yu, J. Marco, Model-based end of discharge temperature prediction for lithium-ion batteries, IFAC-PapersOnLine 53 (2) (2020) 12701–12707.