POLITECNICO DI TORINO
Repository ISTITUZIONALE

Fusing Driving Behavior Features for Detecting Early Voltage Consistency Anomaly of Battery Pack in Electric Vehicles

Original
Fusing Driving Behavior Features for Detecting Early Voltage Consistency Anomaly of Battery Pack in Electric Vehicles / Li, S., Zhang, H., Misul, D.A., Miretti, F.. - In: IEEE TRANSACTIONS ON TRANSPORTATION ELECTRIFICATION. - ISSN 2332-7782. - (2026). [10.1109/TTE.2025.3642624]

Availability:
This version is available at: 11583/3005830 since: 2025-12-13T15:30:47Z

Publisher:
IEEE

Published
DOI:10.1109/TTE.2025.3642624

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
IEEE postprint/Author's Accepted Manuscript

©2026 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

18 November 2026

Fusing Driving Behavior Features for Detecting Early Voltage Consistency Anomaly of Battery Pack in Electric Vehicles

Shaopeng Li, Hui Zhang, Daniela Anna Misul, *Member, IEEE*, Federico Miretti, *Member, IEEE*

***Abstract*—To prevent battery thermal runaway incidents in electric vehicles (EVs), it is crucial to detect faults on voltage consistency between cells (VCC) in the battery pack by a timely and accurate manner. This paper conducts a naturalistic driving experiment on 20 EVs to collect high-quality vehicle operation data, and proposes an early VCC anomaly detection method that integrates driving behavior features. For four types of micro-running segments of EVs, the model of genetic-algorithm-optimized back propagation neural network (GA-BPNN) is established to estimate the normal voltage variation coefficient between cells (VVCC) during vehicle running. The input features of models include related parameters of ambient temperature, state of charge (SOC), current, vehicle speed, acceleration, and pedal stroke. The GA-BPNN models are validated to exhibit a good robustness. The VVCC anomaly threshold is obtained using methods of Box-Cox transformation, 3σ rule, and box plot. The validation results indicate that the proposed method can effectively identify early abnormalities before VCC exceeds the conventional threshold. The average recall rate for 12 test EVs is 0.8746. This method overcomes the shortcomings of existing methods that cannot automatically adjust anomaly detection standard due to lacking consideration of driving behavior's impact on voltage fluctuation.**

***Index Terms*—Electric vehicles (EVs), Battery safety, Driving behavior, Voltage consistency, Anomaly detection**

I. Introduction

THE rapid development of electric vehicles (EVs) has been accompanied by increasingly severe safety challenges. Incidents of vehicle fires and explosions caused by thermal runaway in battery pack have occurred with concerning frequency [1], [2], [3]. The primary cause of thermal runaway is battery failure, one of the main triggers of which is the inconsistency among individual cells within the battery pack [4], [5]. Differences in initial capacity, internal resistance, and open-circuit voltage among cells can result in discrepancies in current, voltage, and temperature, which may lead to over-voltage, under-voltage, overheating, and even thermal runaway in the battery system over long-term cycling [6], [7], [8]. Excessive inconsistency among cells indicates the presence of excessively aged or defective cells [9], [10]. Accurate evaluation and detection of inconsistency within the battery pack are critical for identifying aging or faulty cells in a timely manner and preventing thermal runaway [11].

A. Literature Review

Terminal voltage is easy to measure and extract in real-time, making it an effective indicator of battery working state. Consequently, many researchers have focused on detecting anomaly of voltage consistency between cells (VCC) for battery safety early-warning in real-world EVs [12], [13]. A common approach involves extracting, estimating, or predicting VCC features such as range, standard deviation (SD), variation coefficient, curve distance, and entropy among all cell voltages. These features are then compared with preset threshold to determine the level of abnormality [14], [15]. For example, Li et al. [16] extracted multiple parameters representing VCC, including the range and variance of metrics such as sample entropy, skewness, and so on. Lu et al. [17] utilized the voltage variation coefficient between cells (VVCC) to evaluate VCC. Wang et al. [18] utilized the average distance between cell voltage curves as a measure of VCC. Huang et al. [19] defined voltage anomaly thresholds and developed a model for voltage prediction, enabling early warning for poor voltage consistency. Qiu et al. [20] detected VCC anomaly by examining if the difference between the maximum theoretical entropy and the measured entropy of all cell voltages in a module or battery pack exceeded a preset threshold. Huang et al. [21] adopted root mean square error as a VCC indicator, and detected voltage anomaly with setting the threshold of local outlier factor to 1.6. Li et al. [22] proposed a VCC evaluation method by Mahalanobis-Taguchi system, and defined thresholds to examine the feature extracted through sample entropy.

However, the above studies have largely overlooked the impact of driving behavior on voltage fluctuations and voltage consistency, and therefore did not adopt differentiated

Manuscript received July 29, 2025. This work was supported in part by the National Key Research and Development Program of China (2023YFC3009700), the National Natural Science Foundation of China (52372341), and the Hubei Provincial Natural Science Foundation (2025AFD745). *(Corresponding author: Hui Zhang; Daniela Anna Misul).*

Shaopeng Li is with the Key Laboratory of Low Altitude Geographic Information and Air Route of Jiangxi Education Institutes, Jiangxi Flight University, Nanchang 330088, Jiangxi, China, and the School of Transportation and Engineering, Jiangxi Flight University, Nanchang 330088, Jiangxi, China (e-mail: spli@jxfu.edu.cn).

Hui Zhang is with the Intelligent Transportation Systems Research Center, Wuhan University of Technology, Wuhan 430063, Hubei, China, and the Engineering Research Center of Transportation Information and Safety, Ministry of Education, Wuhan 430063, Hubei, China (e-mail: zhanghuiits@whut.edu.cn).

Daniela Anna Misul and Federico Miretti are with the Department of Energy (DENERG) and Center for Automotive Research and Sustainable Mobility (CARS@Polito), Politecnico di Torino, 10138 Torino, Italy (e-mail: daniela.misul@polito.it; federico.miretti@polito.it).

anomaly thresholds for multiple driving behavior modes. Zhang et al. [23] observed that VCC varies with changes in driving behavior and environmental conditions. While existing methods are effective in identifying VCC anomaly during advanced fault stages, they struggle to accurately detect early-stage anomaly due to subtle and unstable nature of anomaly features, as well as interference from real-time voltage fluctuation. Soo et al. [24] employed a Transformer based method to predict battery state of health, characterized by battery voltage parameters, for different driving behavior categories, respectively. Therefore, to achieve early anomaly detection, threshold settings should take driving behavior into account. Several researchers have proposed defining voltage anomaly thresholds separately for different operating modes. For instance, Li et al. [16] analyzed statistical features related to skewness, sample entropy, median, etc., for both normal and abnormal batteries. They found that the boundary distinguishing normal and abnormal batteries differs between charging and driving states, and therefore recommended applying distinct methods or thresholds for varying vehicle states. Similarly, Fang et al. [25] set distinct voltage range anomaly thresholds for three scenarios: driving, charging, and starting after static charging. However, these studies only provided differentiated thresholds for various operating modes, without proposing adaptive voltage anomaly detection methods that can dynamically respond to real-time driving states.

To address aforementioned issues, one feasible approach is to estimate or predict the normal VCC under real-time driving behavior. Then, anomalies can be detected if the residual between measured value and model output exceeds a preset threshold [26], [27]. This concept has been employed in several existing studies on voltage anomaly detection. Li et al. [28] predicted the normal voltage during vehicle operation, then performed anomaly detection based on predefined thresholds. Zhang et al. [29] utilized the back propagation neural network (BPNN) to establish instantaneous normal voltage prediction models for different driving behavior modes. However, these studies primarily focus on predicting instantaneous voltage rather than VCC over a given time period. Typically, VCC needs to be calculated using voltage data over a continuous time window. In many studies, the data were obtained from the National Monitoring and Management Center for New Energy Vehicles in China, where the maximum data sampling frequency is 0.1 Hz and several driving behavior signals are not collected [30]. As a result, studies integrating driving behavior for VCC anomaly detection remain scarce.

To establish a driving-behavior-adapted VCC anomaly detection method, our previous work studied the associations between driving behavior and VCC [31]. The key methods of this work are summarized as follows: First, the EVs running process was divided into four types of segments based on the driver's pedal operation behavior, corresponding to four distinct driving behavior modes. Second, for each segment, the VVCC was calculated, and driving behavior parameters containing information on speed, acceleration, and pedal state were extracted. Finally, several algorithms were employed to analyze the influence mechanism, quantitative effect, and importance of driving behavior parameters to VVCC.

B. Contributions of This Work

This study explores the method of early VCC anomaly detection considering driving behavior, the innovations and contributions are highlighted in the following aspects.

1) The method of driving-behavior-integrated early anomaly detection for VCC is proposed. This approach effectively addresses the limitations of existing research that anomaly detection standard cannot adaptively adjust with real-time driving behavior.
2) The GA-BPNN based VVCC estimation models are established for four types of running segments, with input variables including driving behavior parameters. This modeling method is conducive to fully learning the VCC change characteristics at different running states. Meanwhile, the VVCC output by each specific model is better matched with their corresponding single driving behavior mode than by a general model.
3) The practical applicability and accuracy of the proposed method is comprehensively validated using data from 20 EVs. The VVCC estimation models are built using data from four EVs, while their robustness is validated using data from another four EVs. Finally, the effectiveness and accuracy of VCC anomaly detection are verified using data from the remaining 12 EVs.

C. Organization of the Article

The remainder of this paper is organized as follows. Section 2 introduces the naturalistic driving experiment of EVs, the collected data and their preprocessing process. Section 3 describes the methodologies including VCC anomaly detection framework, GA-BPNN based VVCC estimation method, threshold determination approach, etc. Research results and discussion are then presented in Section 4, followed by the conclusions summarized in Section 5.

II. Experiment and Data Acquisition

A. Naturalistic driving experiment

For collecting high-quality vehicular operation data of real-world EVs, a naturalistic driving experiment on 20 EVs is conducted. Specifically, this experiment is performed on 20 electric taxies of same specification, over a period of 10 months from January 28, 2022, to November 28, 2022, at the city of Wuhan, China. During experiment, the EVs operate without any limitations on driving time or routes, and the taxi service has never been influenced by experiment itself. The advanced driving assistant system (ADAS) is customized to collect data (such as battery state, ambient temperature, driving state, etc.) transmitted from controller area network (CAN) of experimental EVs, and data of traffic environment and car following behavior. The connotation and details of the experiment are shown in Fig. 1. The main specifications of the studied EVs are shown in Table I.

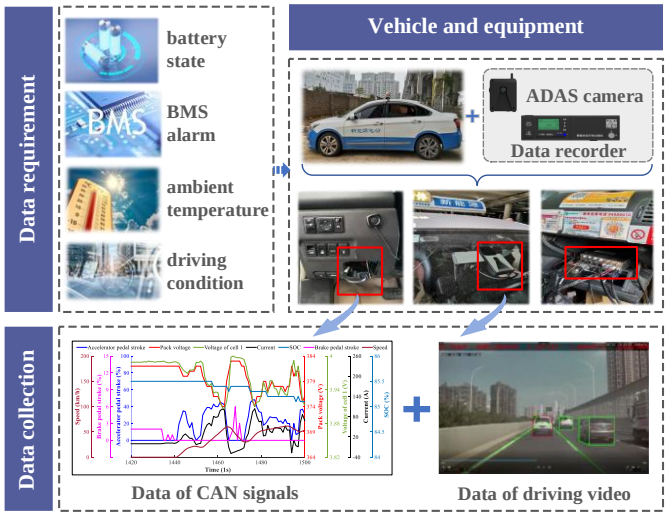


Fig. 1. Core connotation and details of the naturalistic driving experiment

TABLE I
SPECIFICATIONS OF THE EXPERIMENTAL VEHICLES

Specifications	Values
Battery type	Ternary lithium-ion
Number of battery cells	192
Number of battery temperature probes	48
Total energy of battery pack	49 kWh
Rated voltage of battery pack	350.4 V
Nominal voltage of battery cell	3.65 V
Rated motor power	42 kW
Rated motor torque	60 Nm

B. Data collection and preprocessing

During the experiment, data containing over 100 parameters were collected with sampling frequencies ranging from 1 Hz to 100 Hz. This study focuses on 13 key parameters for analysis. Since the raw signals were asynchronous, a resampling procedure was applied to temporally synchronize all parameters at a uniform frequency of 1 Hz. Specifically, a common time axis with a 1-s interval was first established, and then the parameter values were mapped to this timeline using nearest-neighbor interpolation based on their original timestamps. To avoid inadvertently discarding voltage fault samples during outlier handling, the normal range of each parameter was defined using dataset-based statistical thresholds. The descriptions of the 13 parameters used in this study are summarized in Table II.

The sample overview after data preprocessing is provided in Table III. Due to factors such as ADAS malfunctions, vehicle failures, and variations in operational duration, the sample size varies among 20 EVs. The results also show that V4, V8, and V11 each recorded only one BMS alarm, while V13 and V20 did not experience any BMS alarms. The remaining 15 vehicles exhibited alarm counts ranging from 1279 to 892728. Obviously, the differences in alarm frequency across vehicles are significant. Given the reliability of the data, it can be concluded that even for vehicles of the same specification, individual characteristics including differences in operating environment, driving style, usage intensity, aging level, and initial battery condition can lead to large variations in the number and type of BMS alarms recorded. A detailed examination of the alarm records indicates that most BMS alarms in these EVs are related to electrical issues such as sensor and connector faults, whereas only a small proportion are directly associated with battery voltage or temperature anomalies. In particular, for V4, V8, and V11, each alarm corresponded to a transient jump in voltage or current signals that returned to normal in the subsequent frame. These isolated alarms are therefore most likely attributed to momentary sensor glitches rather than data processing errors. It is also important to note that the absence of BMS alarms does not necessarily indicate the absence of voltage or temperature anomalies, as the manufacturer's fault thresholds are conservatively set, allowing minor fluctuations to occur without triggering alarms.

TABLE II
DESCRIPTIONS OF 13 UTILIZED PARAMETERS

Signal	Unit	Frequency	Normal range
Accumulated mileage	km	1 Hz	[Value at beginning, Value at ending]
Accelerator pedal stroke	%	10 Hz	[0, 100]
Brake pedal stroke	%	10 Hz	[0, 100]
Speed	km/h	100 Hz	[0, 150]
Acceleration	m/s^2	10 Hz	[-10, 10]
Ambient temperature	℃	10 Hz	[-10, 60]
Pack voltage	V	50 Hz	[250, 430]
Cell voltage	V	1 Hz	[3.2, 4.5]
Current	A	50 Hz	[-250, 300]
SOC	%	50 Hz	[0, 100]
Probe temperature	℃	1 Hz	[-10, 60]
BMS alarm	/	50 Hz	0 or 1
BMS fault code	/	10 Hz	Set of fault codes

TABLE III
OVERVIEW AFTER DATA PREPROCESSING FOR 20 EVS

EVs No.	Running days	Total samples	Samples with BMS alarm	Initial mileage (km)	Ending mileage (km)
V1	233	5610182	43173	374444.4	466106.2
V2	249	6165498	34904	261363.6	308822.8
V3	236	3363964	25852	251288.4	289585.4
V4	201	4567429	1	363190.7	423139.0
V5	205	4394464	64035	416862.8	500591.7
V6	216	3308534	5037	307120.4	375221.3
V7	164	3751749	1279	289275.9	339988.2
V8	197	4376649	1	343943.7	413229.7
V9	182	4759192	892728	259629.8	337006.3
V10	135	3339187	15043	315936.4	387058.5
V11	140	2367776	1	243100.3	281743.7
V12	209	6674202	17384	307858.4	366060.1
V13	164	1813103	0	215957.4	268063.6
V14	156	3168259	29651	399151.6	439974.0
V15	134	2136208	36320	385780.0	445177.7
V16	150	2784789	12401	292938.7	345616.6
V17	92	2300137	1815	421450.7	503619.3
V18	132	3553767	26001	328588.7	381630.3
V19	90	1956276	1923	325788.6	437635.4
V20	62	1347918	0	234387.0	280090.3

III. METHEDOLOGIES

A. Characterizing method for driving behavior and VCC

Our previous research found the differences in voltage fluctuation patterns at different pedal operations. Thus, the running process of EVs is divided into four segments: segments A with driver's foot on accelerator pedal, segments B with foot idle after completely releasing accelerator pedal, segments C with foot on brake pedal, segments D with foot idle after completely releasing brake pedal. The segment division schematic is shown in Fig. 2. For each segment, driving behavior parameters including acceleration, speed, and pedal status are extracted, as summarized in Table IV. It is noteworthy that segments with a very short duration are filtered out, as they do not contain sufficient information to yield reliable results. Specifically, samples with a duration of less than 5 s are excluded for segments B, C, D, and 10 s for segments A.

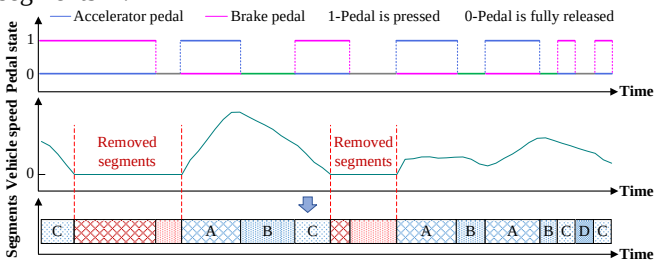


Fig. 2. Running segment division based on pedal status

TABLE IV
EXTRACTED DRIVING BEHAVIOR PARAMETERS

Segments	No.	Driving behavior parameters	Unit
A/B/C/D	1	Maximum speed	km/h
	2	Average speed	km/h
	3	SD of speed	km/h
	4	Maximum positive acceleration	m/s^2
	5	Maximum negative acceleration	m/s^2
	6	SD of acceleration	m/s^2
	7	Average positive acceleration	m/s^2
	8	Average negative acceleration	m/s^2
A	9	Maximum accelerator pedal stroke	%
	10	Average accelerator pedal stroke	%
	11	SD of accelerator pedal stroke	%
C	9	Maximum brake pedal stroke	%
	10	Average brake pedal stroke	%
	11	SD of brake pedal stroke	%

Additionally, for each segment, the VVCC is calculated to measure the VCC in the battery pack during that period. The calculation for VVCC can be expressed as:

$$\bar{V}_j = \frac{1}{n}\sum_{i=1}^{n} V_{ij}\ ,\ j = 1, 2, \ldots, k \tag{1}$$

$$\bar{V} = \frac{1}{k}\sum_{j=1}^{k} \bar{V}_j \tag{2}$$

$$V\max_j = \max\{V_{1j}, V_{2j}, \ldots, V_{nj}\}\ ,\ j = 1, 2, \ldots, k \tag{3}$$

$$V\min_j = \min\{V_{1j}, V_{2j}, \ldots, V_{nj}\}\ ,\ j = 1, 2, \ldots, k \tag{4}$$

$$\delta = \frac{\sqrt{\frac{1}{k}\sum_{j=1}^{k}(V\max_j - V\min_j)^2}}{\bar{V}} \tag{5}$$

where, V_{ij} represents the voltage value of the i cell at the sampling time j; the total number of cells is n (n is 96 in this paper); $\bar{V}_j$ represents average voltage of all cells at sampling time j; $\bar{V}$ is the average value of voltage during the segment; $V\max_j$ and $V\min_j$ represent the maximum and minimum of voltage at the sampling time j; δ is the VVCC during the segment. A larger value of VVCC means a worse VCC.

B. VCC anomaly detection method

We propose a detection method for early VCC anomaly by fusing driving behavior features, which can be summarized as follows. The VVCC estimation models incorporating driving behavior information are developed separately for four types of running segments. If the actual VVCC is significantly higher than the normal value output by the model, it can be determined that the battery pack exhibits a VCC anomaly. Notably, separately modeling each kind of segments could enhance the model's ability to fully learn the mechanisms and characteristics of VCC variations under different driving states, which would effectively improve model accuracy. Additionally, the VVCC output by each model is more closely aligned with their corresponding single driving behavior mode than by a general model. Thus, the proposed estimation method addresses the limitations in existing studies regarding the accuracy and adaptability of estimated values to real-time vehicular running state.

The features exhibit strong correlations with target VVCC, and the modeling utilizes high-quality vehicle operation data. As a result, models with simple network architectures and training processes can achieve a satisfactory performance. Through the preliminary VVCC estimation modeling and testing, the authors find that the GA-BPNN algorithm demonstrates stronger generalization ability compared to decision-tree-based algorithms such as random forest and XGBoost. Therefore, the GA-BPNN algorithm is adopted to construct VVCC estimation models incorporating driving behavior features. The primary framework of the early detection method for VCC anomaly is illustrated in Fig. 3.

The output of the estimation model represents the normal VVCC value for a specific running segment, therefore, the samples used for model training must consist solely of normal data without voltage faults. To ensure data representativeness and reliability, appropriate experimental EVs with fault-free operation data are selected for model development. Table V presents the statistics of segments containing BMS alarms and VCC anomalies for all 20 EVs. It should be noted that certain vehicles, such as V8, V17, and V19, which experienced BMS alarms as shown in Table III, display zero BMS alarms in Table V. This discrepancy arises because those alarm events occurred either during stationary periods or within short-duration segments that were excluded during processing. According to engineering experience, a segment in which the maximum voltage difference within the battery pack exceeds 0.1 V is labeled as a VCC anomaly. As seen in Table V, the VCC anomaly occurrences in segments A of vehicles V20, V19, V13, V17, V3, etc., are significantly

less than those of other EVs. Therefore, data from these vehicles should be prioritized for modeling. While, V14, V12, and V5 should not be used for modeling, due to the relatively more segments with VCC anomaly. In addition, considering the complexity and coupling characteristics of battery faults, data from EVs with BMS alarms should be avoided for model training. Among the 20 EVs, V8, V13, V17, V19, and V20 exhibited no battery fault alarms under the four segments. The EVs V4, V7, and V11 show a very few alarms and VCC anomalies. Hence, normal data from the eight EVs can serve as sample set for modeling.

In summary, the implementation steps for the proposed VCC anomaly detection method can be summarized as follows: First, extract normal data from EVs V8, V13, V17, and V19 for training the GA-BPNN model. Second, verify the robustness and stability of the model using the normal data from vehicles V4, V7, V11, and V20, and calculate the anomaly threshold based on estimation error statistics. Finally, validate the effectiveness and accuracy of the VCC anomaly detection method using data from the remaining 12 vehicles.

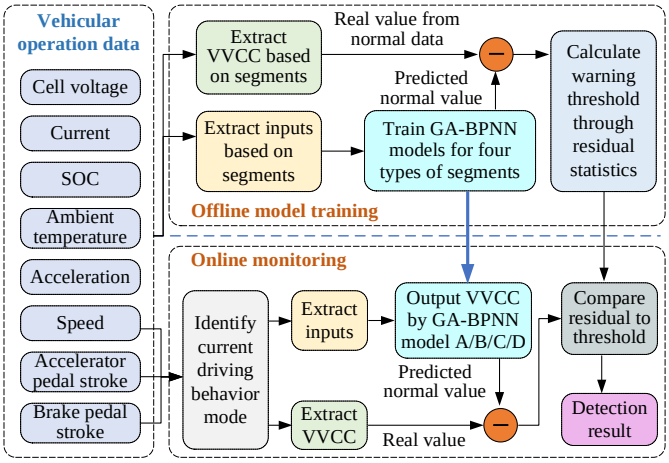


Fig. 3. Early VCC anomaly detection method

C. GA-optimized BPNN model

The GA is a global optimization search algorithm developed by simulating the biological inheritance and evolutionary processes in nature, it is usually used to optimize BPNN models [32], [33], [34]. In this work, we utilize a GA to optimize the BPNN by finding the optimal initial weights and biases, thereby improving the model's performance. The mean squared error (MSE) of VVCC estimation is used to represent the individual fitness; the smaller the MSE, the better the fitness of the individual. Based on individual fitness, the operations of selection, crossover, and mutation are sequentially performed to generate the next generation population. The selection operator used is random uniform selection. The crossover operator is uniform crossover (if a random number is less than 0.5, the gene of the first parent is selected; otherwise, the gene of the second parent is selected). The mutation operator adopted is Gaussian mutation (with an SD of 0.1).

D. Inputs selection for GA-BPNN model

In our previous study [31], a random forest algorithm was employed to establish a regression model linking driving behavior, SOC, and ambient temperature with VVCC, using the data from vehicle V2. The resulting feature importance values quantified the contribution of each variable to VVCC estimation. A summary of these results is presented in Table VI, which illustrates both the increase in model error when a given feature is excluded and the rationale for selecting the corresponding input features in this study.

Firstly, for all the four GA-BPNN models, maximum speed, average speed, SD of speed, average SOC, and average ambient temperature are chosen as input features. Under the four driving patterns, maximum speed, average speed, and SD of speed show strong connections with VVCC, and the importance scores of average SOC are also relatively high. Although ambient temperature exhibits a weak correlation with VVCC, it is included as an input feature because it significantly affects the overall voltage distribution range across different thermal conditions [30], thereby improving the model's robustness under various seasonal and environmental scenarios. In contrast, probe temperature signals are excluded because they merely measure surface temperatures at a few fixed locations within the battery pack and show extremely weak correlation (the Pearson correlation coefficient is 0.002) with voltage variation [29]. Moreover, it is difficult to select or aggregate among 48 probe signals that display substantial spatial temperature differences, which could introduce additional noise into the model. Secondly, three accelerator pedal stroke and brake pedal stroke related parameters are selected as features for model A and C, respectively. Considering the complex fluctuation at segments A, average positive acceleration and maximum positive acceleration are selected as features to enhance the model's fitting capability. Lastly, for segments B, C, and D, average negative acceleration and maximum negative acceleration are selected as features. Although the connection of these two parameters with VVCC is less prominent, it is still significantly stronger than that of maximum positive acceleration, average positive acceleration, and SD of acceleration. Furthermore, current-related parameters are selected as input features in place of voltage-related ones. Including voltage-related parameters could input abnormal voltage information into the model, leading to an output that is not a normal VVCC value, which would hinder VVCC anomaly detection. The input features for the VVCC estimation models are summarized in Table VII.

E. Anomaly threshold determination

To perform VCC anomaly detection, reasonable thresholds need to be defined based on the statistical characteristics of VVCC estimation errors. This paper proposes to derive the threshold using a weighted average of the anomaly boundary values obtained through 3σ rule and box plot. The 3σ rule assumes that the data follows a normal distribution, however, actual samples often deviate from this assumption. To address

this issue, this study refers to previous work [35]. First, apply a Box-Cox transformation to improve the normality of the VVCC estimation residuals. Subsequently, the 3σ rule is used to calculate the anomaly boundary, which is then mapped back to the corresponding value in the original data set. The proposed methods are designed for early-stage anomaly, while BMS alarms from CAN are primarily aimed at significant anomalies that have already occurred. Therefore, the fault detection standards defined by manufacturers are not applicable to this study.

The Box-Cox transformation automatically determines the form and parameters of the transformation function based on the distribution of original data. Data that does not follow a normal distribution will either follow or approximate a normal distribution after transformation. The mathematical formula for Box-Cox transformation [36], [37] is as follows:

$$y^{(\lambda)} = \begin{cases} \dfrac{y^{\lambda}-1}{\lambda}, \lambda \neq 0 \\ \ln y, \lambda = 0 \end{cases}, y > 0 \quad (6)$$

where, y is the original value, λ is the transformation parameter, $y^{(\lambda)}$ represents the transformed value.

To calculate the weighted average of the values obtained from the above two methods, the formula is as follows:

$$Threshold = \alpha \cdot T_{3\sigma} + (1-\alpha) \cdot T_{boxplot} \quad (7)$$

where, $T_{3\sigma}$ and $T_{boxplot}$ are the anomaly boundary derived from the 3σ rule and box plot, respectively, α represents the weight assigned to the former method.

F. Accuracy evaluation for anomaly detection

In anomaly detection and risk identification studies, confusion matrix is commonly used to quantitatively analyze the accuracy of models. The four values in a confusion matrix are: true positive (TP), true negative (TN), false positive (FP), and false negative (FN) [38], [39], [40]. Based on these four categories of samples, various evaluation metrics can be derived to assess model performance, such as recall rate, precision, F1 score, receiver operating characteristic curve (ROC), etc. In this study, recall rate *R* and precision *P* are incorporated, which are defined as follows:

$$R = \frac{TP}{TP+FN} * 100\% \quad (8)$$

$$P = \frac{TP}{TP+FP} * 100\% \quad (9)$$

TABLE V
STATISTICS OF RUNNING SEGMENTS

EVs No.	Segments A			Segments B			Segments C			Segments D		
	Total segments	Segments with BMS alarm	Segments with VCC anomaly	Total segments	Segments with BMS alarm	Segments with VCC anomaly	Total segments	Segments with BMS alarm	Segments with VCC anomaly	Total segments	Segments with BMS alarm	Segments with VCC anomaly
V1	58813	468	19	25401	122	1	27712	216	1	4683	42	0
V2	76034	391	20	18154	92	0	20433	115	0	8050	48	0
V3	35837	291	7	3607	22	1	14389	103	2	1612	11	0
V4	71907	1	25	9295	0	1	2381	0	0	1331	0	0
V5	66944	969	61	7357	69	1	20609	134	1	2684	30	0
V6	56864	99	21	15539	29	2	5015	5	0	1876	0	0
V7	54227	8	28	8278	1	3	5145	2	1	1613	0	0
V8	79312	0	34	9043	0	5	5808	0	0	2105	0	0
V9	75850	11958	29	8157	1573	4	7419	2197	0	2010	451	0
V10	43670	248	11	3287	42	0	7397	19	0	1024	9	0
V11	39838	1	26	2074	0	0	6036	0	0	659	0	0
V12	110214	345	75	7357	16	0	12373	26	3	3502	7	1
V13	29923	0	3	3672	0	0	10408	0	0	1085	0	0
V14	36427	478	182	3577	43	2	2917	52	0	975	15	0
V15	32556	585	18	1856	30	1	3542	42	0	657	8	0
V16	43502	182	9	9047	46	2	7946	60	1	2012	9	0
V17	32912	0	9	2873	0	0	6206	0	0	1096	0	0
V18	34896	268	37	6838	51	0	16613	142	0	3015	14	0
V19	19605	0	2	3742	0	0	4784	0	0	685	0	0
V20	15502	0	1	2056	0	0	5971	0	1	549	0	0

TABLE VI
IMPORTANCE OF FEATURES TO VVCC ESTIMATION

Features	Importance (IncMSE%)			
	Segments A	Segments B	Segments C	Segments D
Average SOC	0.125	0.293	0.596	0.579
Average ambient temperature	0.014	0.011	0.015	0.013
Maximum speed	0.163	0.267	0.119	0.166

Average speed	0.042	0.395	0.209	0.196
SD of speed	0.012	0.015	0.017	0.018
Maximum positive acceleration	0.007	0	0	0.002
Maximum negative acceleration	0.005	0.004	0.008	0.007
SD of acceleration	0.008	0.004	0.006	0.007
Average positive acceleration	0.010	0	0	0.002
Average negative acceleration	0.005	0.010	0.015	0.001
Maximum accelerator/brake pedal stroke	0.086	/	0.003	/
Average accelerator/brake pedal stroke	0.479	/	0.006	/
SD of accelerator/brake pedal stroke	0.045	/	0.005	/

TABLE VII
INPUT FEATURES OF VVCC ESTIMATION MODELS

Segments	Input features of BPNN models			
	Driving behavior		Battery status	Environment
A	Maximum accelerator pedal stroke, average accelerator pedal stroke, SD of accelerator pedal stroke, average positive acceleration, maximum positive acceleration	Average speed, maximum speed, SD of speed	Average SOC, maximum positive current, maximum negative current, average positive current, average negative current	Ambient temperature
B	Maximum brake pedal stroke, average brake pedal stroke, SD of brake pedal stroke, average negative acceleration, maximum negative acceleration			
C	Average negative acceleration, maximum negative acceleration			
D	Average negative acceleration, maximum negative acceleration			

IV. RESULTS AND DISCUSSIONS

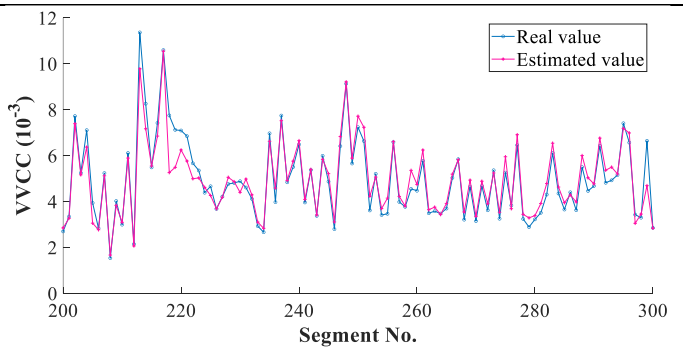
A. Training results of GA-BPNN models

The optimal transfer function, training function, and neurons in hidden layer of BPNN are obtained by 10-fold cross-validation. Then, GA is applied for further optimization. After debugging and validation, the crossover probability is set to 0.6, the mutation probability is the reciprocal of the number of variables, the population size is set to 50, and the maximum generations is set to 30. Table VIII lists model training results before and after optimization by GA. It shows that, the MSE of all models decreases after optimization. Model C shows the most significant reduction in MSE of 49%, while model A has the smallest reduction at 5%.

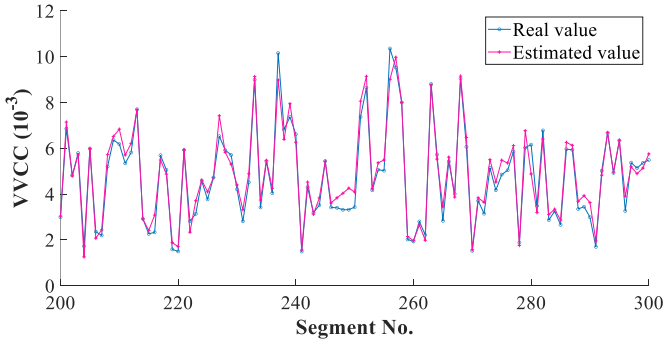
For the visual analysis of test results, we extract segments with indices 200 to 300 from the test set. As shown in Fig. 4, in each subplot, the estimated VVCC is overall very close to the actual value, especially for segments A and B. The VVCC in segments A and B is generally higher, primarily because the vehicle speed is higher and the running state changes more drastically in these segments. This is also the reason why the MSE in segment B is higher than segments C and D, yet its estimation performance is superior. Overall, the test results of four models indicate a good performance.

TABLE VIII
MODEL TRAINING RESULTS BEFORE AND AFTER OPTIMIZATION BY GA

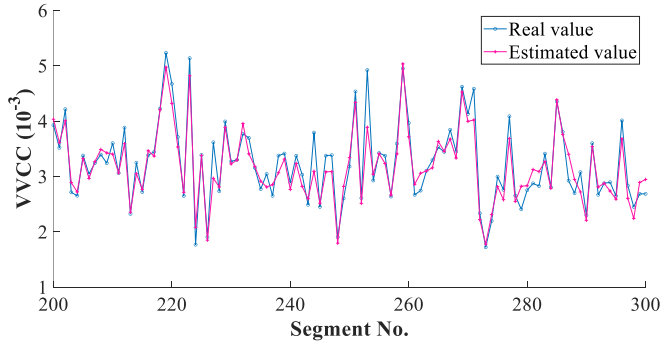
Model	MSE (10^{-7})		MRE	
	Before optimization	After optimization	Before optimization	After optimization
A	2.065	1.957	0.069	0.068
B	4.025	3.572	0.122	0.104
C	4.941	2.496	0.092	0.093
D	3.848	2.933	0.131	0.123



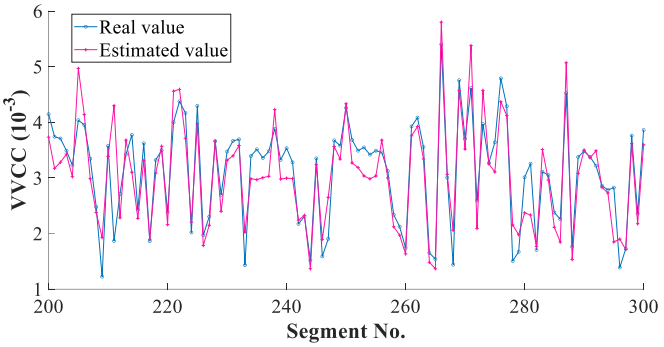
(a)



(b)



(c)



(d)

Fig. 4. Partial results of estimated and real VVCC from test set. (a) Model A. (b) Model B. (c) Model C. (d) Model D

B. Robustness validation of GA-BPNN models

The MSE and MRE obtained from testing the normal operation data from EVs V4, V7, V11, and V20 are shown in Table IX. It can be found that the estimation errors in these EVs are slightly higher than those in the test set composed of V8, V13, V17, and V19, nevertheless, the GA-BPNN models still exhibit a good robustness overall. Among the MRE values for V4, V7 and V20, the minimum is 0.089 and the maximum is 0.288. Notably, the estimation error for V11 is significantly higher than that for other three vehicles, with the MRE for segments D being 0.692. The MRE value of 0.692 does not necessarily indicate poor model accuracy, the main reasons for this are as follows. First, the reference value of vehicle V11's test result is limited. The vehicle with the smallest error in all four segments is V4, while the largest error occurs in vehicle V11. The verification results of V4, V7, and V20 are relatively close. This suggests that although vehicle V11 has relatively few VCC anomaly segments, its VCC is clearly poorer than that of V4, V7, and V20, therefore, greater emphasis should be placed on the validation results of the other three vehicles. Secondly, due to the inherent physical properties of VVCC, including the high precision of voltage measurement and the sensitivity of voltage fluctuations during vehicle operation, the MRE fluctuates significantly.

TABLE IX
ROBUSTNESS VERIFICATION RESULTS OF GA-BPNN MODELS

No.	Model A		Model B		Model C		Model D	
	MSE (10^{-7})	MRE	MSE (10^{-7})	MRE	MSE (10^{-7})	MRE	MSE (10^{-7})	MRE
V4	2.783	0.089	5.519	0.134	13.872	0.198	6.588	0.184
V7	10.083	0.152	16.278	0.207	23.864	0.234	22.176	0.288
V11	8.389	0.213	26.715	0.547	20.239	0.484	21.190	0.692
V20	9.842	0.134	22.124	0.254	22.574	0.261	16.614	0.236

C. Results of anomaly threshold calculation

The optimal values of λ in Box-Cox transformation of VVCC test errors for four segments are: 0.3559, 0.5036, 0.4606, and 0.6045. Since estimated VVCC lower than the measured one is not considered an anomaly, the dataset used for Box-Cox transformation and threshold calculation is composed of those instances where the actual VVCC are greater than the estimated ones. The frequency distributions of test errors before and after transformation, with the determined anomaly boundary, are shown in Fig. 5.

The anomaly boundary calculated by 3σ rule and boxplot method is shown in Table X. It can be observed that the values obtained from the former are consistently higher than those from the latter. Through computation and analysis, it is found that the anomaly detection models perform well when the weight is arbitrarily chosen within the range of zero to one. In the following sections, the effectiveness and accuracy of anomaly detection will be validated and analyzed with a weight value of 0.5. Furthermore, the detection timeliness and accuracy results under different weight values will be discussed in detail to focus on the determination of the optimal threshold.

TABLE X
ANOMALY BOUNDARY OBTAINED BY 3Σ RULE AND BOXPLOT

Segments	3σ rule (10^{-3})	Boxplot (10^{-3})
A	3.703	2.045
B	3.728	2.630
C	4.643	3.161
D	3.737	2.892

D. Effectiveness validation of anomaly detection

This study aims to detect early VCC anomalies, thereby, a key objective is to evaluate the method's capability to identify anomalies before a significant fault occurs (i.e., voltage range over 0.1 V). Fig. 6 illustrates representative analysis segments, showing the estimated VVCC, measured VVCC, and their residual. To ensure clarity, only 30 segments spanning the detected anomalous segment are shown. As Fig. 6(a) shows, from segment No.34169 to No.34178, multiple warnings are triggered, indicating VCC anomalies during this period. In segments No.34170 and No.34178, maximum voltage range exceeds 0.1 V, which signifies the occurrence of fault. Obviously, VCC anomalies are detected in segments No.34169 and No.34177 before the maximum voltage range reaches the threshold of 0.1 V. For the signal curves of segment No.34177, it is evident that the voltage curves of all cells in the 15-second period are significantly scattered and do not show the usual aggregated pattern observed under normal conditions. Thus, for this vehicle, the VCC anomaly before and after these segments mainly reflects an excessively high level of voltage dispersion among the cells. As illustrated in Fig. 6(b), an anomaly warning is triggered at segment No.14821, where cell 9 and 27 show a significant distance from others, with the voltage of cell 9 being the highest among all cells and cell 27 having the lowest, indicating potential over-voltage and under-voltage risks, respectively. For these anomaly instances identified by the models, the maximum cell voltage difference does not exceed 0.1 V at the time the warning is triggered. Furthermore, in segments sequence before and after the warning, the actual VVCC is almost always higher than the estimated one, which implies the VCC remains poor prior to the fault. This result further validates the effectiveness of the proposed method for early anomalies detection.

In the battery pack, early-stage VCC anomalies do not necessarily become more pronounced over time. In the

absence of severe or irreversible damage, such anomalies may remain transient and unstable. Moreover, complex system environment and signal noise can cause subtle anomalies to appear intermittently. It is precisely the subtle and unstable nature of early anomalies that presents challenges for voltage anomaly warning. As demonstrated above, the proposed early-stage VCC anomaly detection method can capture subtle anomalies that are inconsistent with the current driving behavior amidst real-time voltage fluctuations. It is that threshold-based methods and signal processing techniques struggle to achieve, as they are more focused on detecting faults with clearly noticeable abnormal states. In applications, after detecting anomaly segments using this method, it can be further combined with other techniques for fault diagnosis and location.

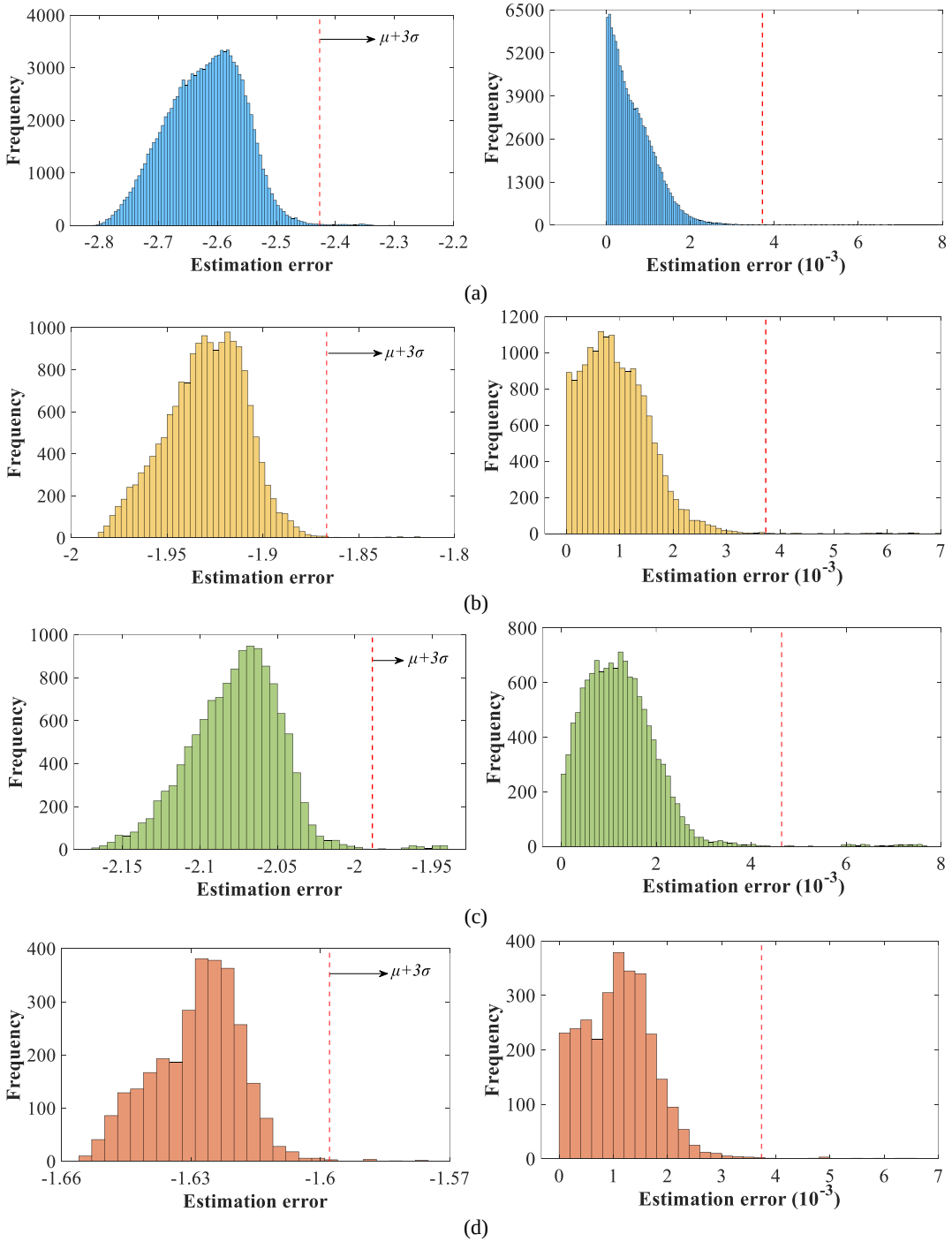


Fig. 5. Frequency distribution and anomaly boundary of VVCC estimation errors after and before Box-Cox transformation. (a) Segments A. (b) Segments B. (c) Segments C. (d) Segments D

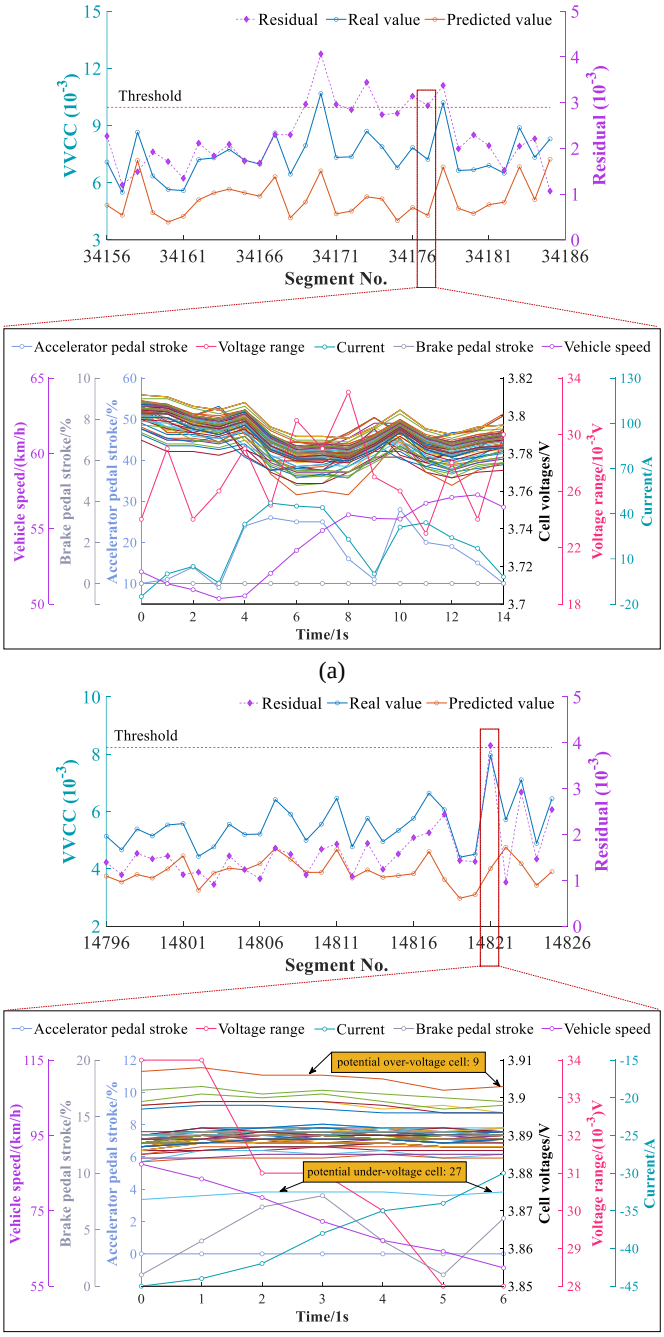


Fig. 6. Examples of VCC anomaly detection. (a) Segment A of V5. (b) Segment C of V1

E. Accuracy validation of anomaly detection

The performance metrics of early-stage VCC anomaly detection on 12 test vehicles are summarized in Table XI. As can be observed, the proposed method exhibits a high accuracy in terms of recall rate. The highest recall rate among the 12 vehicles is achieved by V2, while the lowest one is observed in V10. The recall rates for V2 and V10 are 1.0000 and 0.6364, respectively. The average recall rate across all 12 vehicles is 0.8746, with 10 vehicles having a recall rate above 0.8, and 6 vehicles achieving a recall rate no less than 0.9. Clearly, there is noticeable variation in recall rates among the vehicles, which can be attributed to two main factors: the limited number of fault samples and the complexity of battery system. Although a 10-month experiment involving 20 EVs was conducted, the number of samples containing faults remains relatively small, particularly for severe or long-term faults. This limited data scale may lead to variability in accuracy assessments. In addition, the working states of vehicles and batteries, as well as the fault characteristics of the batteries, are complex. In such a complex system, vehicle differences tend to be amplified, which can affect the robustness of the method. For results of precision, it can be observed that precision values are generally low, which is expected given the early anomaly detection objective of this work. The method aims to detect subtle and unstable early-stage anomalies, battery packs identified with such early signs may not develop obvious faults within the limited 10-month observation period. Therefore, false positives computed against samples with no evident failures are not necessarily erroneous but rather reflect potential early-warning behavior. This distinction explains why precision-based metrics are less appropriate for evaluating early-stage anomaly detection. It should also be emphasized that the proposed model is designed primarily for battery safety state assessment and maintenance guidance, rather than for real-time fault diagnosis or thermal runaway prediction. While false positive results are presented here for completeness, their interpretation must consider the nature of incipient anomalies and the limited duration of experimental observation.

In practice, the trade-off between high recall rate and low precision has implications for operator workload. A high-recall strategy ensures that incipient anomalies are rarely missed, however, it inevitably increases the number of alerts that do not progress into actual faults within the short observation window. To mitigate this issue in real-world deployment, the anomaly scores produced by the model can be further integrated with adaptive threshold, multi-stage verification, or long-term trend monitoring. These mechanisms help filter transient, low-risk signals while preserving the model's high sensitivity to truly emerging faults, thereby balancing safety assurance with practical operator workload.

TABLE XI
RESULTS OF EARLY-STAGE VCC ANOMALY DETECTION

Vehicle No.	TP	FP	FN	TN	Recall rate	Precision
V1	20	1279	1	115126	0.9524	0.0154
V2	20	972	0	121786	1.0000	0.0202
V3	9	49	1	55024	0.9000	0.1552
V5	55	703	8	96590	0.8730	0.0726
V6	20	158	3	79016	0.8696	0.1124
V9	30	863	3	92314	0.9091	0.0336
V10	7	20	4	55141	0.6364	0.2593
V12	71	1598	8	131710	0.8987	0.0425
V14	153	525	31	42935	0.8315	0.2257
V15	15	60	4	38337	0.7895	0.2000
V16	11	143	1	62302	0.9167	0.0714
V18	34	637	3	60646	0.9189	0.0507

In existing studies on the detection of VCC anomalies, most works focus on validating the accuracy of estimation or

prediction models and the effectiveness of anomaly detection, while systematic and large-scale verification of detection accuracy remains limited. This is primarily because real-world vehicle fault data are often confidential and difficult to obtain, and comprehensive accuracy validation requires long-term experiments or complete life-cycle data covering the entire process from anomaly occurrence to fault manifestation. Only a few studies have reported quantitative results such as recall and precision using real vehicle data. For example, Cao et al. [41] developed a deep-learning-based online diagnostic network for lithium-ion batteries, mainly targeting internal short circuit, thermal runaway, over-aging, and electrolyte leakage. Their evaluation with 18.2 million data points from 515 vehicles demonstrated that the method improved the true positive rate by more than 46.5% when the false positive rate was below 0.2. Zhang et al. [42] employed a neural network with interpretable modules on top of an Autoencoder to detect voltage anomalies during charging using real vehicle data, where the precision ranged from 0.2778 to 0.9 and the recall ranged from 0.5 to 0.9 under different hyperparameter configurations. However, the anomaly types and operating conditions in these studies are substantially different from those in this work. The present study focuses on the early detection of VCC anomalies during vehicle driving, rather than fault identification under charging or post-fault conditions, and thus a direct comparison is not applicable. Nonetheless, future research may benefit from benchmarking the proposed method against comparable models once suitable open datasets become available.

F. Discussion for optimal anomaly threshold

Since it is inappropriate to assess model performance by combining TP rate and FP rate, the ROC cannot be used to determine the optimal threshold. Due to the limitation of having no access to vehicle operation data over a much longer period, it is difficult to directly verify if the early anomalies identified by the model will eventually evolve into significant faults. However, the real VCC state of these early anomaly samples identified by the models can be validated through their statistical information. Thus, the optimal threshold is discussed by combining recall rate and the mean VVCC value for the anomaly segments identified by the models.

The recall rate at different thresholds which are computed using various weightings, is shown in Fig. 7. To enhance clarity and avoid potential misunderstanding, the result for precision is also provided, as shown in Fig. 8. Fig. 7 shows that the recall rate of most EVs decreases as the threshold increases, particularly for α greater than 0.5. Once α exceeds 0.5, the high threshold leads to a significant increase in unrecognized anomaly segments, where the maximum voltage range exceeds 0.1 V. For α less than 0.3, the recall rate is very high, with many vehicles showing a result of 1.0. As seen from Fig. 8, the precision remains low across all α values, with the maximum value being below 0.5. The low precision does not indicate poor model performance but rather reflects the nature of early and subtle anomalies, which may not evolve into observable faults within the limited observation period. This result further supports our earlier explanation that computing precision or false positive rates is not appropriate for evaluating early-stage anomaly detection.

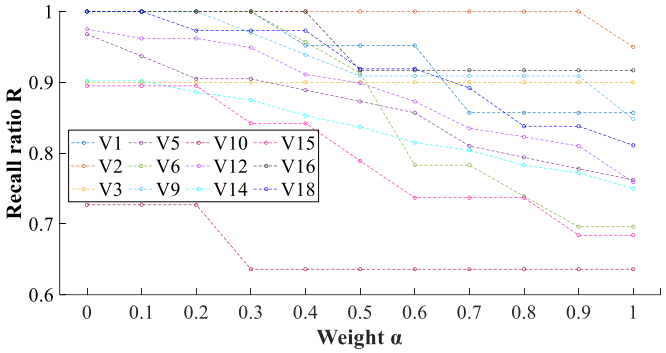


Fig. 7. Recall rate of VCC anomaly detection at different weights

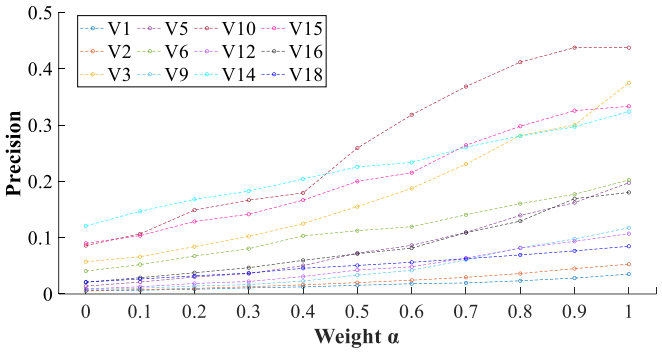


Fig. 8. Precision of VCC anomaly detection at different weights

The mean VVCC of anomaly segments identified by the models using various weightings is presented in Fig. 9. It can be observed that the mean VVCC of the anomaly segments ranges between 0.007 and 0.011. Overall, this value increases with the threshold, however, for certain vehicles such as V3 and V10, it decreases once α reaches 0.7. Lu et al. [17] conducted a VCC evaluation study and verified that, samples with a VVCC in the range of 0.00625 to 0.0125 begin to show a tendency of VCC deteriorating by scores falling below 90. Although the study did not consider the dynamic impact of driving behavior on VVCC, its findings remain relevant for reference. Comparing this value (0.00625) with the results from our model, it is clear that the VCC of identified anomaly samples is in a poor state, further validating the effectiveness of the model in detecting and issuing early warnings for anomalies.

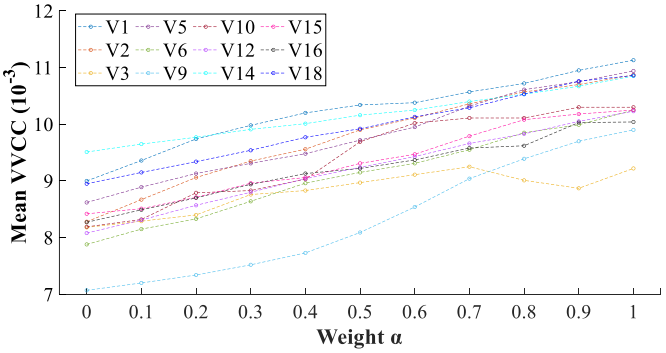


Fig. 9. Mean VVCC of identified anomaly segments at different weights

Due to the complexity of the power battery system and its fault evolution, the definition of VCC faults remains somewhat ambiguous, and there is no complete or universal standard in place. In applications, vehicle manufacturers have established fault alarm standards that balance the timeliness and necessity of the alarm. If both recall rate and anomaly severity are taken into account, the optimal range for weight α in threshold calculation for the experimental vehicles lies between 0.3 and 0.5. For applications, the determination of optimal threshold depends on a comprehensive consideration of anomaly severity and the proactive effect of early warnings. It is also essential to conduct in-depth analysis that integrates the full lifecycle status data of the power battery, fault evolution mechanisms, and characteristics.

V. Conclusion

This paper conducted the research on early VCC anomaly detection by performing naturalistic driving experiments on EVs. A VVCC estimation and anomaly detection method that integrates driving behavior features is proposed. Four GA-BPNN models are established for different running segments to estimate the normal VVCC. The Box-Cox transformation, 3σ rule, and boxplot are used to calculate VVCC anomaly thresholds. The validation results show that the proposed method can effectively detect early anomalies before VCC exceeds conventional thresholds, providing support for the anomaly warning of battery cells. The average recall rate for 12 test vehicles is 0.8746, with 10 vehicles achieving a rate above 0.8 and 6 vehicles achieving a rate of at least 0.9. This method overcomes the limitations of statistics-based and threshold-based methods that fail to account for the impact of driving behavior on voltage fluctuations and enable running-state-adapted anomaly detection. It achieves effective and accurate early warning of minor VCC anomalies. In addition, the proposed GA-BPNN-based framework is computationally efficient and suitable for real-time implementation in the BMS. The computationally intensive training and optimization processes are performed offline, while only the trained lightweight feed-forward BPNN models are deployed in the BMS. For the most complex model (Segment A, with 14 inputs and 80 hidden neurons using ReLU activation), a single forward inference requires approximately 2500 FLOPS, and the total memory footprint is less than 10 KB. As the method operates on a segment basis rather than at every sampling point, the online processing load is further reduced. Considering that typical BMS controllers can achieve tens to hundreds of MFLOPS, the computational demand of the proposed method is orders of magnitude lower, demonstrating its feasibility for embedded implementation.

However, it is worth to point out the limitations of this research. Firstly, the proposed method mainly focuses on identifying subtle anomalies in the early stages of voltage faults and may not be directly applicable to specific fault diagnosis or prediction for future fault. Secondly, the established VVCC estimation models have not accounted for factors such as air conditioning operation, passenger load, and road slope. Finally, the data used in this study is sourced from 20 EVs, with limited variability in terms of vehicle type, geographical location, and other factors.

For future work, publicly available datasets that include detailed fault and thermal runaway information could be used to further develop and verify the method. To build a more robust and universal framework, it is necessary to incorporate data from a broader range of sources, such as different battery types, ambient temperature conditions, charge/discharge states, and driving scenarios. Beyond passenger EVs, the proposed approach could also be extended to large-size electric trucks and stationary energy storage systems. For heavy-duty trucks, additional quantitative features such as vehicle load, road gradient, and module-level variations should be considered to account for the stronger current fluctuations and greater impact of operating conditions on battery consistency. For stationary storage systems, the feature set should instead include charging/discharging power, ramp rate, and grid service modes, with longer time windows to reflect their hour-scale charge-discharge cycles and quasi-steady-state operation characteristics.

References

[1] S. Yang et al., "Comparing different battery thermal management systems for suppressing thermal runaway propagation," *J. Energy Storage*, vol. 101, 2024, Art. no. 114005.

[2] V. S. Hemakumar, V. J. Chakravarthy, S. Surendranath, V. Gundu, M. Ramkumar Prabhu, and S. Hari Chandra Prasad, "A multi-scale modeling approach for predicting and mitigating thermal runaway in electric vehicle batteries," *Therm. Sci. Eng. Prog*., vol. 56, 2024, Art. no. 103029.

[3] S. Kumar and H.-J. Kim, "Recent advances in early warning methods and prediction of thermal runaway events in Li-ion batteries," *J. Ind. Eng. Chem*., vol. 145, pp. 63-74, 2025.

[4] Q. Yu et al., "Unsupervised learning for lithium-ion batteries fault diagnosis and thermal runaway early warning in real-world electric vehicles," *J. Energy Storage*, vol. 109, 2025, Art. no. 115194.

[5] K. Wang et al., "Thermal runaway evolution of a 4S4P lithium-ion battery pack induced by both overcharging and unilateral preheating," *Case Stud. Therm. Eng*., vol. 63, 2024, Art. no. 105324.

[6] J. Hong et al., "Investigation on overcharge-caused thermal runaway of lithium-ion batteries in real-world electric vehicles," Appl. Energy, vol. 321, 2022, Art. no. 119229.

[7] Y. Qin et al., "Temperature consistency–oriented rapid heating strategy combining pulsed operation and external thermal management for lithium-ion batteries," *Appl. Energy*, vol. 335, 2023, Art. no. 120659.

[8] P. Wang, J. Chen, F. Lan, Y. Li, and Y. Feng, "Multiscale feature fusion approach to early fault diagnosis in EV power battery using operational data," *J. Energy Storage*, vol. 98, 2024, Art. no. 112812.

[9] G. Wang, S. Jin, J. Jiao, and J. Xie, "Voltage measurement-based recursive adaptive method for internal short circuit fault diagnosis in lithium-ion battery packs," *Control Eng. Pract*., vol. 145, 2024, Art. no. 105857.

[10] J. Hong, F. Liang, J. Yang, and S. Du, "An exhaustive review of battery faults and diagnostic techniques for real-world electric vehicle safety," *J. Energy Storage*, vol. 99, 2024, Art. no. 113234.

[11] H. Zhao et al., "Voltage fault diagnosis and prognostic of lithium-ion batteries in electric scooters based on hybrid neural network and multiple thresholds," *J. Power Sources*, vol. 618, 2024, Art. no. 235197.

[12] Y. Che, A. Foley, M. El-Gindy, X. Lin, X. Hu, and M. Pecht, "Joint estimation of inconsistency and state of health for series battery packs," *Automotive Innovation*, vol. 4, no. 1, pp. 103-116, 2021.

[13] Z. Zhang, J. Ma, Y. Ma, X. Gong, K. Xiangli, and X. Zhao, "Identification of voltage abnormality in the battery system based on fusion of multiple sparse data observers for real-world electric vehicles," *J. Energy Storage*, vol. 114, 2025, Art. no. 115727.

[14] Q. Liu et al., "Voltage fault diagnosis and misdiagnosis analysis of battery systems using the modified Shannon entropy in real-world electric vehicles," *J. Energy Storage*, vol. 73, 2023, Art. no. 109287.
[15] Y. Xu, X. Ge, R. Guo, and W. Shen, "Recent advances in model-based fault diagnosis for lithium-ion batteries: A comprehensive review," *Renew. Sustain. Energy Rev.*, vol. 207, 2025, Art. no. 114922.
[16] D. Li, J. Deng, Z. Zhang, P. Liu, and Z. Wang, "Multi-dimension statistical analysis and selection of safety-representing features for battery pack in real-world electric vehicles," *Appl. Energy*, vol. 343, 2023, Art. no. 121188.
[17] Y. Lu et al., "A method of cell-to-cell variation evaluation for battery packs in electric vehicles with charging cloud data," *eTransportation*, vol. 6, 2020, Art. no. 100077.
[18] Q. Wang, Z. Wang, L. Zhang, P. Liu, and Z. Zhang, "A novel consistency evaluation method for series-connected battery systems based on real-world operation data," *IEEE Trans. Transport. Electrific.*, vol. 7, no. 2, pp. 437-451, 2021.
[19] Y. Huang, X. Gong, Z. Lin, and L. Xu, "A hybrid data-driven method for voltage state prediction and fault warning of Li-ion batteries," *Case Stud. Therm. Eng.*, vol. 64, 2024, Art. no. 105420.
[20] Y. Qiu, W. Cao, P. Peng, and F. Jiang, "A novel entropy-based fault diagnosis and inconsistency evaluation approach for lithium-ion battery energy storage systems," *J. Energy Storage*, vol. 41, 2021, Art. no. 102852.
[21] P. Huang et al., "Safety risk assessment for automotive battery pack based on deviation and outlier analysis of voltage inconsistency," *J. Clean Prod.*, vol. 466, 2024, Art. no. 142889.
[22] F. Li, Y. Min, Y. Zhang, Y. Zhang, H. Zuo, and F. Bai, "Evaluation method for consistency of lithium-ion battery packs in electric vehicles based on the Mahalanobis-Taguchi system," *J. Energy Storage*, vol. 78, 2024, Art. no. 110045.
[23] X. Zhang, W. Yang, L. Yan, M. B. Kaleem, and W. Liu, "Adaptive internal short-circuit fault detection for lithium-ion batteries of electric vehicles," *J. Energy Storage*, vol. 84, 2024, Art. no. 110874.
[24] Y.-Y. Soo, Y. Wang, H. Xiang, and Z. Chen, "A novel transfer learning model for battery state of health prediction based on driving behavior classification," *J. Energy Storage*, vol. 111, 2025, Art. no. 115409.
[25] W. Fang, H. Chen, and F. Zhou, "Fault diagnosis for cell voltage inconsistency of a battery pack in electric vehicles based on real-world driving data," *Comput. Electr Eng.*, vol. 102, 2022, Art. no. 108095.
[26] L. Zhang, B. Xia, and F. Zhang, "Adaptive fault detection for lithium-ion battery combining physical model-based observer and BiLSTMNN learning approach," *J. Energy Storage*, vol. 91, 2024, Art. no. 112067.
[27] H. Zhao, C. Zhang, L. Xu, C. Liao, L. Wang, and L. Wang, "A deep neural network for multi-fault diagnosis of battery packs based on an incremental voltage measurement topology," *Energy*, vol. 316, 2025, Art. no. 134590.
[28] D. Li, Z. Zhang, P. Liu, Z. Wang, and L. Zhang, "Battery fault diagnosis for electric vehicles based on voltage abnormality by combining the long short-term memory neural network and the equivalent circuit model," *IEEE Trans. Power Electron.*, vol. 36, no. 2, pp. 1303-1315, 2021.
[29] H. Zhang, S. Li, F. Chen, X. Pan, H. Feng, and Y. Sun, "Battery voltage fault diagnosis for electric vehicles considering driving condition variation," *IET Intell. Transp. Syst.*, vol. 18, no. 4, pp. 574-590, 2024.
[30] J. Hong, Z. Wang, and Y. Yao, "Fault prognosis of battery system based on accurate voltage abnormity prognosis using long short-term memory neural networks," *Appl. Energy*, vol. 251, 2019, Art. no. 113381.
[31] S. Li, H. Zhang, N. Ding, M. Acquarone, F. Miretti, and D. A. Misul, "Associations of battery cell voltage consistency with driving behavior of real-world electric vehicles," *Green Energy and Intelligent Transportation*, 2024, Art. no. 100236.
[32] Y. Wu, D. Wu, M. Fei, H. Sørensen, Y. Ren, and J. Mou, "Application of GA-BPNN on estimating the flow rate of a centrifugal pump," *Eng. Appl. Artif. Intell.*, vol. 119, 2023, Art. no. 105738.
[33] L. Bai, Y. An, and Y. Sun, "Measurement of Project Portfolio Benefits With a GA-BP Neural Network Group," *IEEE Trans. Eng. Manage.*, vol. 71, pp. 4737-4749, 2024.
[34] C. Wang, Z. Huang, C. He, X. Lin, C. Li, and J. Huang, "Research on remaining useful life prediction method for lithium-ion battery based on improved GA-ACO-BPNN optimization algorithm," *Sustainable Energy Technol. Assess.*, vol. 73, 2025, Art. no. 104142.
[35] Q. Liu, J. Ma, X. Zhao, K. Zhang, K. Xiangli, and D. Meng, "A novel method for fault diagnosis and type identification of cell voltage inconsistency in electric vehicles using weighted Euclidean distance evaluation and statistical analysis," *Energy*, vol. 293, 2024, Art. no. 130575.
[36] Q. Wang, M. Xie, and F. Yang, "Early battery lifetime prediction based on statistical health features and box-cox transformation," *J. Energy Storage*, vol. 96, 2024, Art. no. 112594.
[37] K. Liu, L. Kang, L. Wan, D. Xie, and J. Li, "Remaining useful life prediction for lithium-ion batteries based on sliding window technique and Box-Cox transformation," *J. Energy Storage*, vol. 74, 2023, Art. no. 109352.
[38] D. Valero-Carreras, J. Alcaraz, and M. Landete, "Comparing two SVM models through different metrics based on the confusion matrix," *Comput. Oper. Res.*, vol. 152, 2023, Art. no. 106131.
[39] Y. Wang, Y. Jia, Y. Tian, and J. Xiao, "Deep reinforcement learning with the confusion-matrix-based dynamic reward function for customer credit scoring," *Expert Syst. Appl.*, vol. 200, 2022, Art. no. 117013.
[40] Y. Zhao et al., "Enhancing battery durable operation: Multi-fault diagnosis and safety evaluation in series-connected lithium-ion battery systems," *Appl. Energy*, vol. 377, 2025, Art. no. 124632.
[41] R. Cao et al., "Model-constrained deep learning for online fault diagnosis in Li-ion batteries over stochastic conditions," *Nat. Commun.*, vol. 16, no. 1, 2025, Art. no. 1651.
[42] X. Zhang, P. Liu, N. Lin, Z. Zhang, and Z. Wang, "A novel battery abnormality detection method using interpretable Autoencoder," *Appl. Energy*, vol. 330, 2023, Art. no. 120312.