Energy 344 (2026) 139854

Contents lists available at ScienceDirect

Energy

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

Reliable cross-domain lifetime prediction for fuel cells: A time-frequency fusion transfer learning architecture

Hangyu Wu [a,b], Fulin Fan [a,b], Kai Song [a,b,*], Chuanyu Sun [a,b], Yang Li [c], Changjun Xie [d], Xuan Meng [a,b], Jian Mei [a,b], Siew Hwa Chan [e]

[a] School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin, 150001, China
[b] Suzhou Research Institute, Harbin Institute of Technology, Suzhou, 215104, China
[c] School of Electrical Engineering and Automation, Wuhan University, Wuhan, 430072, China
[d] Wuhan Industrial Technology Institute of Artificial Intelligence and New Energy Vehicles, Wuhan, 430119, China
[e] School of Mechanical and Aerospace Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore, 639798, Singapore

ARTICLE INFO

Keywords:
Fuel cell
Domain adaptation technology
Frequency domain analysis
Transfer learning
Remaining useful life

ABSTRACT

Current transfer learning-based methods for predicting the degradation of proton exchange membrane fuel cells (PEMFCs) commonly face insufficient elimination of inter-domain distribution shifts and low knowledge transfer efficiency. Traditional approaches typically extract domain-invariant features by filtering multi-dimensional operational parameters. This process is not only complex but may also introduce interference weakly correlated with voltage degradation, thereby compromising the stability and reliability of cross-domain predictions. To address these challenges, this paper proposes a time-frequency fusion transfer learning (TF-TL) architecture. Unlike existing methods that filter domain-invariant features from multi-dimensional operational parameters, the proposed approach directly aligns domains by leveraging the intrinsic frequency-domain characteristics of voltage signals. The integrated Frequency Domain Adaptation (FDA) technique employs Fast Fourier Transform (FFT) to extract intrinsic low-frequency components from voltage signals. It then aligns features between source and target domains using an enhanced Maximum Mean Discrepancy (MMD) metric. This voltage-centric, frequency-domain alignment avoids dependence on complex auxiliary parameters, thus improving the interpretability and stability of transfer learning. Experimental results show that TF-TL significantly improves generalization across devices and operating conditions, with FDA effectively reducing inter-domain distribution divergence. In multiple transfer tasks, the framework achieves highly accurate and reliable remaining useful life estimation with a relative error below 1.9 %. The proposed architecture provides a theoretical foundation for deploying rapid and reliable online prognostic systems for PEMFCs under complex operating conditions.

1. Introduction

1.1. Background and literature review

Proton exchange membrane fuel cells (PEMFCs), a core enabling technology for the hydrogen economy, are driving the low-carbon transition of transportation powertrains and distributed energy systems owing to their zero-emission nature and high power density [1]. However, under complex operating conditions, PEMFCs experience performance degradation arising from the coupling of various mechanisms, such as electrochemical decay and membrane electrode assembly aging [2]. This difficult-to-predict degradation behavior not only shortens system lifetime but also impedes large-scale commercialization by raising concerns over operational safety and maintenance costs [3]. To address this bottleneck, effective prognostics and health management (PHM) strategies are required to extend the full lifecycle of PEMFCs [4,5].

Data-driven methods can autonomously capture latent degradation patterns from historical data, thereby facilitating the development of mathematical models that describe the relationship between operational parameters and performance decay [6,7]. Compared with traditional mechanistic models, data-driven methods effectively overcome the model mismatch problem caused by complex fuel cell aging mechanisms [8]. These methods exhibit enhanced robustness against noise and

* Corresponding author. School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin, 150001, China.
E-mail address: kaisong@hit.edu.cn (K. Song).

https://doi.org/10.1016/j.energy.2025.139854
Received 12 November 2025; Received in revised form 22 December 2025; Accepted 29 December 2025
Available online 29 December 2025

operational fluctuations, making them a mainstream solution for PEMFC lifetime prediction [9]. In recent years, various data-driven approaches, ranging from deep neural networks to transfer learning frameworks, have been explored to improve degradation modeling accuracy and cross-domain generalization. By comparing Backpropagation Neural Networks (BPNN), Long Short-Term Memory (LSTM) networks, and Gated Recurrent Unit (GRU) networks, Long et al. [10] validated the superiority of GRU in temporal modeling for PEMFC lifespan prediction. Wilberforce et al. [11] proposed a LSTM with Convolutional Neural Network (CNN) fusion architecture that utilizes convolutional neural networks to extract spatial features, significantly improving the prediction accuracy of the Remaining Useful Life (RUL). Zhu et al. [12] developed an Informer-GPR fusion method, which enhances the extraction of key aging features based on a sparse self-attention mechanism, achieving uncertainty quantification and adaptability to dynamic conditions. However, mainstream methods such as LSTM and GRU focus on time-domain degradation trajectories, without explicit representation for frequency-domain features, limiting their capacity to capture multi-physics degradation mechanisms. In addition, modeling time-varying correlations requires large datasets and is vulnerable to overfitting under scarce data conditions, which can significantly increase prediction errors.

The need for improved resolution of complex features has led to the incorporation of frequency-domain analysis for probing PEMFC performance degradation. Sun et al. [13] used Complete Ensemble Empirical Mode Decomposition (CEEMD) to decompose PEMFC voltage into multiple Intrinsic Mode Functions (IMFs) and employed Fast Fourier Transform (FFT) to analyze the potential correlations between these IMFs and other feature parameters, such as current and temperature, from a frequency-domain perspective. Wang et al. [14] extracted frequency-domain health indicators based on the Hilbert transform and built a symbolic GRU model, demonstrating accurate life estimation capability at multiple End-of-Life (EoL) points. Zhu et al. [15] combined TimesNet and Gaussian Process Regression (GPR), transforming time-domain voltage signals into 2D space ones via low-frequency domain analysis, accurately reflecting PEMFC aging characteristics while assessing prediction uncertainty. These studies show that frequency-domain analysis, by decoupling the spectral features of voltage signals, can identify quantitative links between different frequency bands and specific degradation mechanisms, overcoming the ambiguity of time-domain trajectories [16]. However, the accuracy of these predictions is compromised by dataset scarcity.

Transfer learning technology, which leverages knowledge from a source domain to improve and optimize the performance of predictive models in a target domain lacking extensive labeled data, is gradually being applied to the field of PEMFC lifetime prediction [17–19]. However, during transfer, Domain Adaptation (DA) techniques are necessary to eliminate inter-domain discrepancies between the source and target domains; otherwise, distribution shifts can cause the predictive model to fail due to feature mismatch during transfer [20]. Prevailing DA techniques combined with transfer learning are mainly divided into two categories: Feature-based Transfer (FT) and Model-based Transfer (MT) [21]. Zhang et al. [22] used a comprehensive feature evaluation method based on Dynamic Time Warping and Kernel Principal Component Analysis to prioritize key features, establish a domain-invariant feature set, and finally input it into the LSTM model to complete fuel cell transfer prediction. Kebede et al. [23] proposed a transfer learning method based on a Variational Autoencoder and a Bidirectional LSTM model with an Attention Mechanism (BiLSTM-AM) for RUL estimation, which optimizes parameter computation based on the VAE, and falls into the MT category. Wang et al. [24] proposed a novel method using a layer-enhanced quantum LSTM model combined with MT, involving offline training based on source domain data, followed by fine-tuning and transfer to the target stack, achieving accurate RUL estimation. However, MT requires a small amount of target domain data to optimize the initial parameters of the pre-trained model. When the distribution difference between the source and target domains is large, the model cannot adequately decouple domain-specific information, leading to biases in transfer prediction.

1.2. Research gap and contributions

While data-driven methods and transfer learning have advanced PEMFC lifetime prediction, current approaches still face challenges in physical interpretability and robust generalization. Two persistent limitations hinder further progress. First, time-domain and frequency-domain feature processing are often isolated [25]. Typically, frequency analysis is relegated to a mere preprocessing step. This disconnect impedes the synergistic capture of critical interactions between transient voltage dynamics and long-term degradation trends. Second, cross-domain invariant feature screening is often inefficient [26]. It typically depends on exhaustive searches across high-dimensional operational parameters. This not only incurs high computational costs but may also introduce features that are weakly correlated with voltage degradation. This issue becomes acute with large-scale feature inputs, which limits the reliability of predictive models during transfer.

To address the aforementioned issues, this paper proposes a new scheme for PEMFC lifetime prediction based on a developed Frequency Domain Adaptation (FDA) technique. With the FDA, a time-frequency fusion transfer learning (TF-TL) architecture that incorporates frequency-domain FFT is established. This architecture centers on the frequency-domain characteristics of voltage signals, directly extracting domain-invariant features that reflect the essence of degradation, avoiding reliance on complex experimental parameters, and enhancing the universality and interpretability of degradation pattern transfer from a mechanistic perspective. Benefiting from a lightweight structural design, this architecture can be combined with various deep learning models to achieve online generalized prediction across different operating conditions and devices. The specific advantages of the proposed method that contribute to PEMFC lifetime prediction are summarized as follows:

1) The proposed TF-TL architecture for PEMFCs is capable of frequency-domain feature selection and time-domain feature analysis. This architecture can be lightly embedded into various deep learning networks to achieve cross-domain generalized prediction.
2) The proposed FDA technique can achieve domain-invariant characteristics during the transfer process. This technique directly extracts frequency-domain features from the voltage signals to resolve inter-domain discrepancies, thereby avoiding dependence on numerous operational parameters and enhancing the interpretability and stability of generalized predictions.
3) The prediction uncertainty caused by cross-domain differences is effectively and significantly suppressed. Under cross-device and cross-operating-condition scenarios, the TF-TL architecture achieves a high-confidence performance degradation trajectory and accurate RUL estimation.

The remainder of this paper is organized as follows. Section 2 presents the proposed fuel cell lifetime prediction framework. Section 3 describes the data sources utilized in this study. Section 4 validates the efficacy of the FDA in eliminating inter-domain differences and demonstrates the superior performance of the TF-TL architecture. Finally, Section 5 concludes the paper.

2. Methodology

2.1. A transfer learning architecture with time-frequency fusion

Addressing the prevalent issues of inefficient feature screening and weak physical interpretability in existing DA methods, which generally rely on multidimensional operational condition parameters, this paper

proposes a TF-TL architecture centered on the essential frequency-domain characteristics of voltage signals, as depicted in Fig. 1.

Initially, experimental data from both the source and target domains are collected and preprocessed, based on which multiple transfer learning tasks are defined. Subsequently, this paper introduces the TF-TL architecture that incorporates an FDA module. This module analyzes the voltage signals of both the source and target domains via FFT to extract their intrinsic frequency-domain components. Furthermore, based on an improved MMD metric, low-frequency consensus features shared between the two domains are adaptively selected. Meanwhile, Kernel Density Estimation (KDE) is utilized for non-parametric probabilistic modeling of the discrepancies in feature distributions, providing visual and quantitative evaluations for domain alignment.

Unlike traditional methods that rely on external operational condition parameters for feature alignment, the FDA technique directly mines domain-invariant knowledge reflecting the essence of degradation from the frequency-domain representations of voltage signals. This approach enhances the universality and interpretability of the transfer process at the mechanistic level. Finally, the CNN module within the TF-TL architecture captures local degradation patterns, thereby enhancing the capability of baseline models like LSTM to learn domain-invariant degradation laws, which ultimately leads to accurate prediction of the performance degradation trajectory and RUL estimation in the target domain.

2.2. Time-domain and frequency-domain algorithms

2.2.1. Fast Fourier transform

FFT is an efficient algorithm for computing the Discrete Fourier Transform (DFT). Its primary function is to rapidly convert discrete signals from the time domain to the frequency domain, thereby revealing the frequency components of the signal. Direct computation of the DFT is overly complex. However, the FFT method significantly simplifies the calculation process by leveraging the properties of rotation factors in the DFT solution process and incorporating the divide-and-conquer algorithm philosophy [27]. As a result, FFT is widely applied in engineering practices for spectral analysis. The core principles of FFT are as follows:

$$Y[k] = \sum_{n=0}^{N-1} x[n] \cdot e^{-j\frac{2\pi}{N}kn} \quad (k = 0, 1, \ldots, N-1) \tag{1}$$

where $Y[k]$ represents the complex amplitude of the signal at frequency index k, $x[n]$ denotes the uniformly spaced time series after interpolation and resampling, and N signifies the number of sampling points. Regarding the selection of the FFT window length, if the analysis window is too short, low-frequency trends with periods significantly longer than the window are difficult to capture effectively, and their amplitude and phase information can be lost during window sliding. Therefore, this study applies a global FFT transformation to the preprocessed voltage sequence, where the window length is equal to the total number of sampling points N in the sequence. This approach achieves the highest possible frequency resolution for capturing the underlying low-frequency trends.

2.2.2. 1D-CNN

CNN is a deep learning model specifically designed for processing grid-like data. Its core advantage lies in efficiently extracting spatial features through local perception and weight-sharing mechanisms. In the investigation of fuel cell lifespan prediction, CNN can automatically learn local degradation characteristics from multi-dimensional sensor data, including voltage and temperature [28]. Specifically, the convolutional layers capture local degradation patterns, the pooling layers achieve feature dimensionality reduction and enhance translation invariance, while the fully connected layers integrate these features and

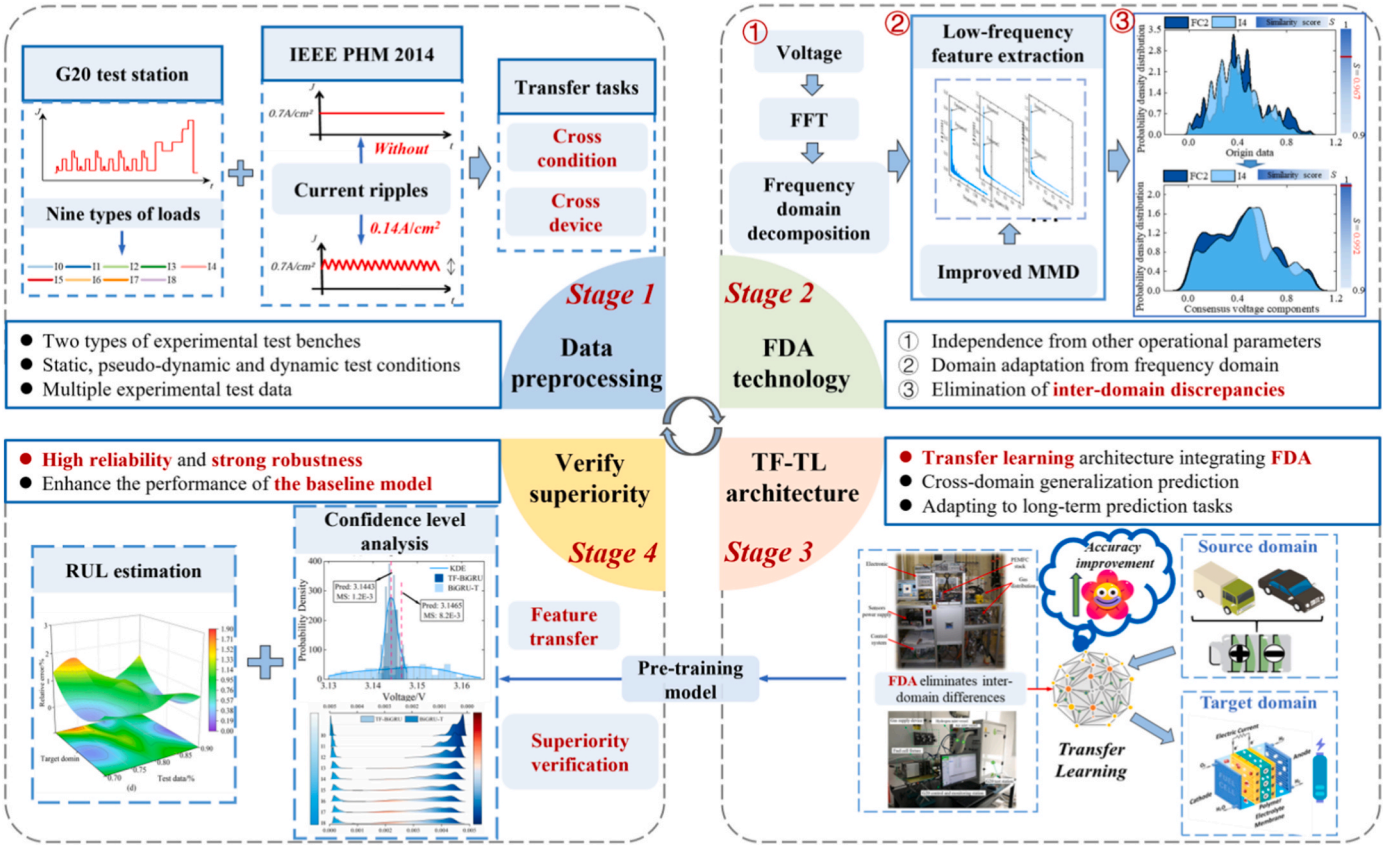

Fig. 1. The research framework of this paper.

map them to the final predictive output. The main structure of CNN is illustrated in Fig. 2.

In the TF-TL architecture, a lightweight 1D-CNN serves as the local feature extractor. This network takes the single-channel time-series signal aligned by the FDA module as input. Its core structure consists of two convolutional blocks: the first block employs 16 filters with a size of 5, followed by a max-pooling layer with a stride of 2. The second block uses 32 filters with a size of 3. All convolutional layers utilize the ReLU activation function. Finally, a global average pooling layer aggregates the feature map along the temporal dimension into a compact 32-dimensional feature vector. This design aims to extract local patterns from the voltage degradation sequence with minimal parameters. The resulting feature vector is then fed into sequence models such as LSTM or GRU for temporal modeling, enabling collaborative cross-domain learning of the degradation patterns.

2.3. Other deep learning models

This section provides an overview of commonly used deep learning models in the field of PEMFC lifespan prediction, including RNNs and their various variants, as well as Temporal Convolutional Networks (TCNs) [29]. These models have been extensively validated in existing research, demonstrating high predictive accuracy and low memory consumption, which aligns well with the rapid response requirements of transfer learning. Based on this, this paper selects to integrate them with the proposed TF-TL architecture to evaluate the transfer universality and performance advantages of the framework under different model conditions. Table 1 summarizes the advantages and disadvantages of the five baseline models.

Table 1
Advantages and disadvantages of baseline models.

Models	Advantages	Disadvantages
LSTM	Effectively captures long-term dependencies and mitigates gradient vanishing	Complex structure, high parameter count, and relatively high computational cost
GRU	Possesses higher computational efficiency and simpler structure than LSTM	Modeling capability for extremely long dependencies may be slightly inferior to LSTM
BiLSTM	Leverages both past and future context, enhancing bidirectional dependency modeling	Higher computational complexity and longer training time than LSTM
BiGRU	Utilizes bidirectional information while typically being more lightweight and faster to train than BiLSTM	Capability to capture extremely long-range dependencies may be slightly inferior to BiLSTM
TCN	Addresses gradient vanishing, supports parallel computation, and offers fast training	May struggle with highly complex temporal patterns and have higher memory requirements

2.4. DA technology based on frequency-domain features

Transfer learning aims to achieve effective knowledge transfer across different domains. Its core objective is to discover generalized feature representations and underlying patterns from the source domain and transfer them to the target domain through an adaptation mechanism, which has a different data distribution. This significantly enhances the predictive capability and generalization performance of the model in the target domain, particularly when labeled data in the target domain is limited. Assume there exist source domain data $D_s = \{(xs\ 1, ys\ 1), (xs\ 2, ys\ 2), \ldots, (xs\ i, ys\ i)\}N_s\ i = 1$ and target domain data $D_t = \{(xt\ 1, yt\ 1), (xt\ 2, yt\ 2), \ldots, (xt\ j, yt\ j)\}N_t\ j = 1$, where $(xs\ i, ys\ i)$ and $(xt\ j, yt\ j)$ are combinations of feature sequences and prediction sequences in the source and target domains, respectively. Each sample consists of a feature sequence and its corresponding label [30].

The traditional mainstream approach of FT involves analyzing and screening common features that are less affected by operating conditions from both the source and target domains. This weakens the data distribution differences caused by different operating conditions, ultimately improving the prediction stability and generalization ability of the model in the target domain. However, aligning parameters that are weakly related to voltage degradation across domains forces the model to learn a large number of non-sensitive features. This not only increases the model training cost but also dilutes the representation weight of key degradation features. Sun et al. [13] employed signal decomposition and FFT analysis to reveal the potential associations between voltage components and experimental parameters from a frequency-domain perspective. The above work validated the feasibility of employing frequency-domain features to extract cross-domain invariant representations.

Unlike the traditional FT domain adaptation process, which relies on other parameters during the experiment, the proposed FDA in this paper extracts essential degradation knowledge from its own frequency-domain characteristics of voltage, avoiding the representation dependency on complex experimental parameters. The specific workflow is shown in Fig. 3.

To effectively mitigate the distribution discrepancy between D_s and D_t, this paper employs the improved MMD as the core metric for domain alignment within the FDA. The improved MMD method optimizes Gaussian kernel parameters adaptively and uses unbiased estimation to enhance robustness in measuring distribution differences. It calculates the median sample distance as the bandwidth and employs unbiased covariance for statistical significance in small samples. This approach captures nonlinear time-series features and addresses the sensitivity of MMD to hyperparameters and instability. Given source domain samples $X_s = \{xs\ i\}n_s\ i = 1$ and target domain samples $X_t = \{xt\ j\}n_t\ j = 1$, the square of the unbiased estimator for the improved MMD is defined as follows:

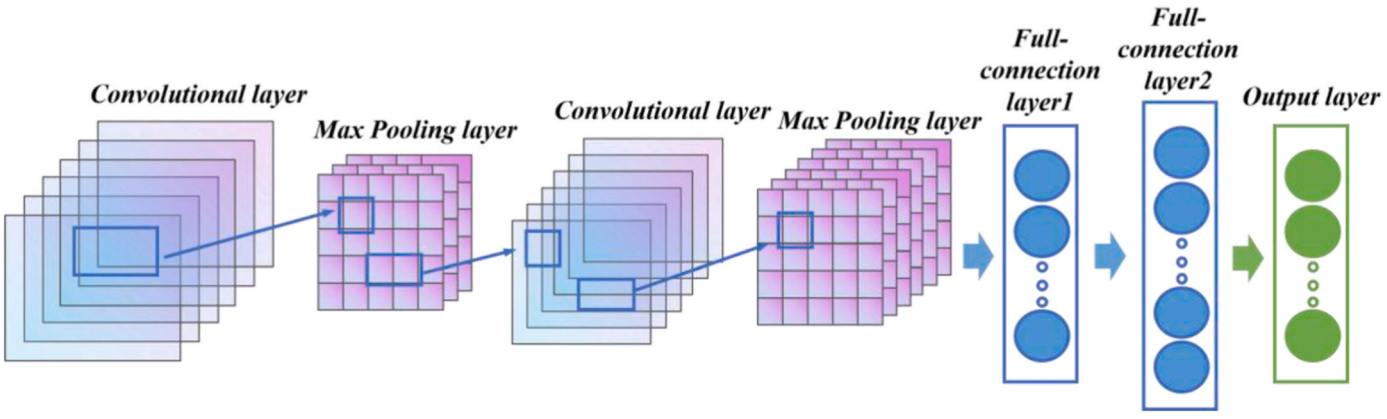

Fig. 2. The structure of CNN.

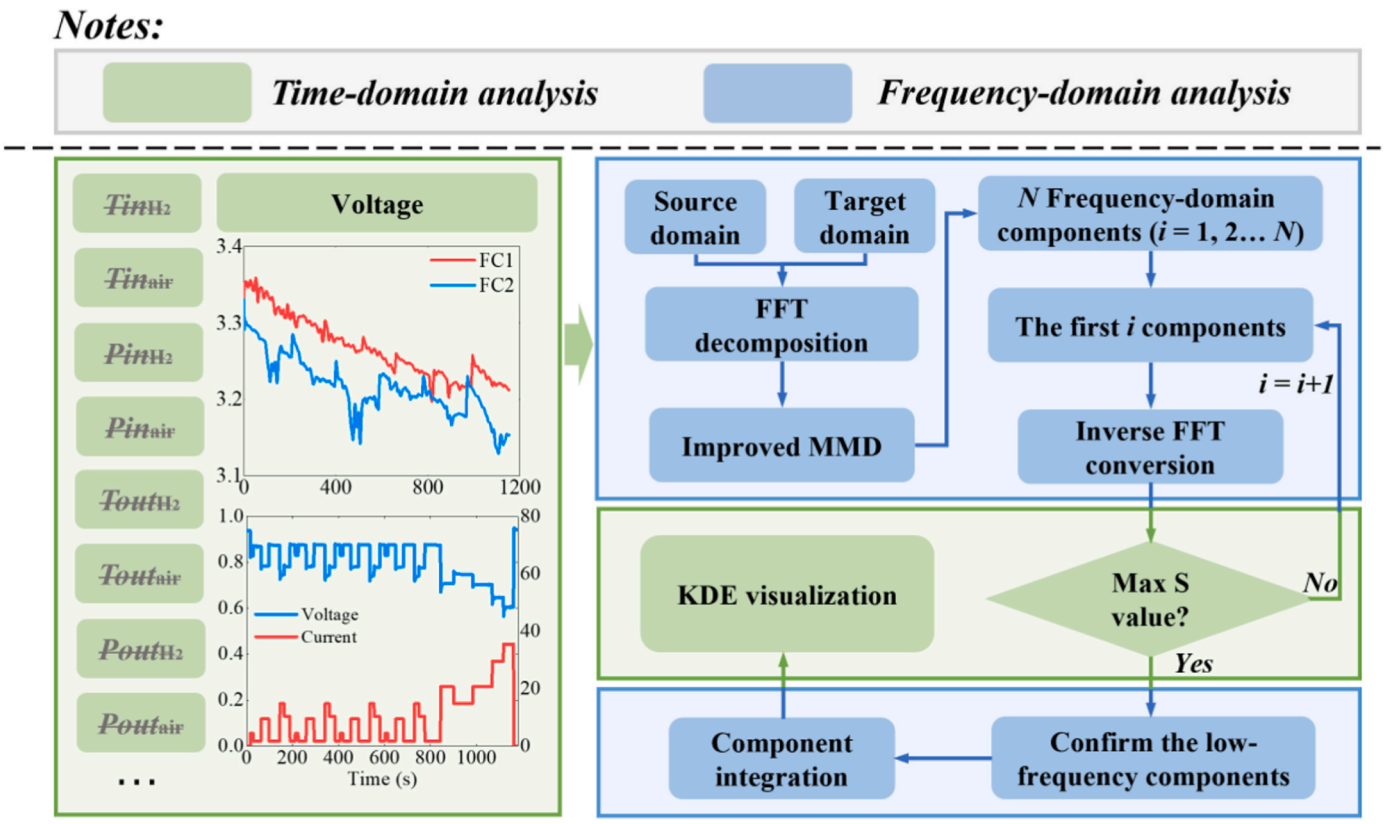

Fig. 3. The workflow of the FDA.

$$MMD^2(X_s,X_t)=\frac{1}{n_s(n_s-1)}\sum_{i=1}^{n_s}\sum_{j\neq i}^{n_s}\kappa\left(x_i^s,x_j^s\right)+\frac{1}{n_t(n_t-1)}\sum_{i=1}^{n_t}\sum_{j\neq i}^{n_t}\kappa\left(x_i^t,x_j^t\right)-\frac{2}{n_sn_t}\sum_{i=1}^{n_s}\sum_{j=i}^{n_t}\kappa\left(x_i^s,x_j^t\right) \tag{2}$$

$$\kappa(x,y)=\exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right) \tag{3}$$

$$\sigma=\text{Median}\left(\{\|x_i-x_j\|:\forall i,j\}\right) \tag{4}$$

where *xi s* and *xt j* are samples, n_s and n_t are sample sizes. σ represents the kernel bandwidth. This study employs a data-driven adaptive strategy to determine σ, setting it to the median of the pairwise Euclidean distances among all samples in the combined set $X_s \cup X_t$, i.e., σ = Median(·). To ensure numerical stability, we define σ = max(σ,ϵ), where ϵ is an extremely small positive constant to prevent division-by-zero errors. This strategy avoids manual hyperparameter tuning and enhances the robustness of the metric across varying data scales. Based on the computed MMD^2 value, a distribution similarity score S is defined to quantify the alignment effect between domains:

$$S=\frac{1}{1+MMD^2(X_s,X_t)} \tag{5}$$

Since $MMD^2 \geq 0$, the value range of S is (0,1]. The closer the source and target domain distributions, the closer the value of S is to 1. Within the FDA, S serves as an intuitive evaluation metric for domain alignment effectiveness. It directly measures the closeness of the low-frequency consensus feature distributions between the source and target domains after FDA processing.

3. Experimental data on PEMFCs

3.1. Static and pseudo-dynamic operating condition data

The aging experimental data for PEMFCs under static and pseudo-dynamic operating conditions were obtained from the publicly available dataset of the IEEE PHM 2014 Data Challenge [31]. The test bench used in the experiments is shown in Fig. 4, and the specific operating parameters during the experiments are listed in Table 2.

The experiments for the two different operating conditions were conducted on two separate fuel cell stacks. Each stack comprised five single cells, with an active area of 100 cm^2 per cell. Stack 1 (FC1) was tested for 1154 h under a static load (Current density = 0.7 A/cm^2), while Stack 2 (FC2) was tested for 1020 h under a pseudo-dynamic load at the same current density (Frequency = 5 kHz, Current ripple = 10 %) [32]. The raw voltage data were preprocessed by sampling at a rate of 3 points per hour and a moving average method was applied for data smoothing. The resulting voltage degradation trends for FC1 and FC2 are shown in Fig. 5.

3.2. Dynamic operating condition data

The dynamic operating condition data were obtained from experiments conducted on a Greenlight 20 (G20) test station [33,34]. The corresponding PEMFC stack is designated as FC3. The experiment comprised 3076 fuel cell dynamic load cycle (FC-DLC) tests, totalling 1008 h. The test bench used for the FC3 durability testing is shown in Fig. 6, and the relevant test parameters are listed in Table 3 [35].

As the durability test for FC3 was designed to simulate real-world vehicle operating conditions, each FC-DLC test involved two distinct driving schedules: “Urban” and “Suburban”. The “Urban” schedule simulates frequent load changes at low vehicle speeds, while the "Suburban" schedule simulates load variations at medium-to-high speeds.

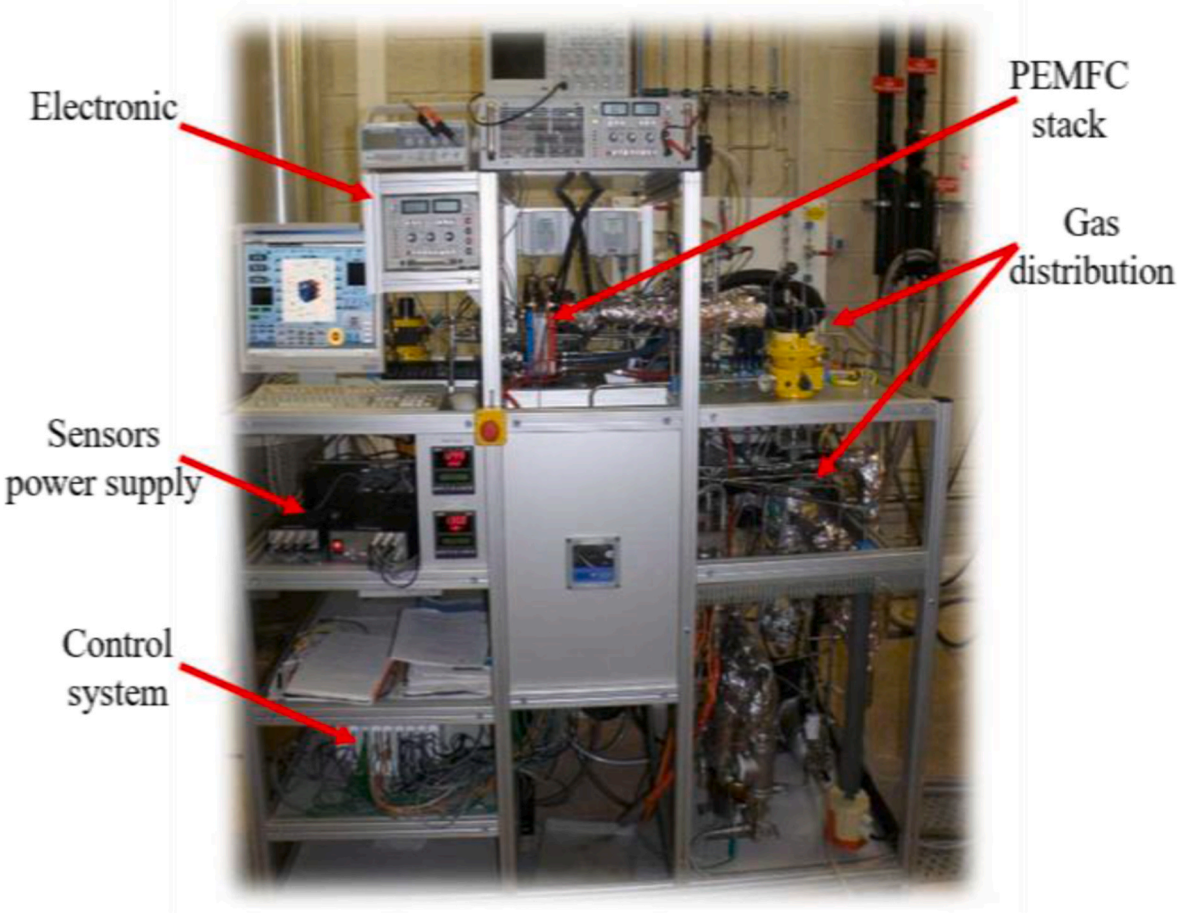

Fig. 4. Aging test bench of FC1 and FC2.

Table 2
Range of controlled physical parameters (FC1 and FC2).

Parameter	Value range
Cooling flow	0~10 L/min
Cooling temperature	20~80 °C
Gas humidification	0~100 % RH
Gas temperature	20~80 °C
Air flow	0~100 L/min
H2 flow	0~30 L/min
Gas pressure	0~2 bar
Fuel cell current	0~300 A

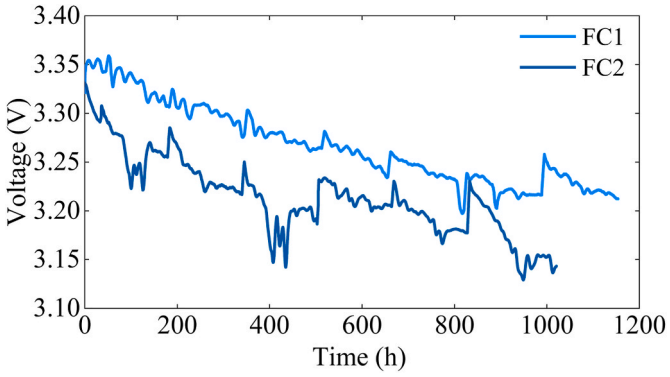

Fig. 5. The voltages of FC1 and FC2 after preprocessing.

Within each dynamic cycle, the "Urban" schedule was repeated four times, followed by one execution of the "Suburban" schedule. This protocol involved nine different load current levels, ranging from 0 % to 100 %. Fig. 7(a) shows the current and voltage profiles during a single FC-DLC test, and Fig. 7(b) presents the voltage variations under the nine categorized load levels.

3.3. Evaluation criteria of prediction performance

To determine the predicting performance, this paper evaluates and compares the forecasting performance of each method based on multiple indicators: Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Coefficient of Determination (R^2), and Relative Error (RE). The smaller the RMSE and MAPE, the higher the prediction accuracy. The closer R^2 is to 1, the better the output value of the model fits the original data. RE is introduced to evaluate the effect of RUL estimation, and the smaller the RE, the better the prediction performance [36].

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{6}$$

$$MAPE = \frac{100}{N}\cdot\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{7}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2} \tag{8}$$

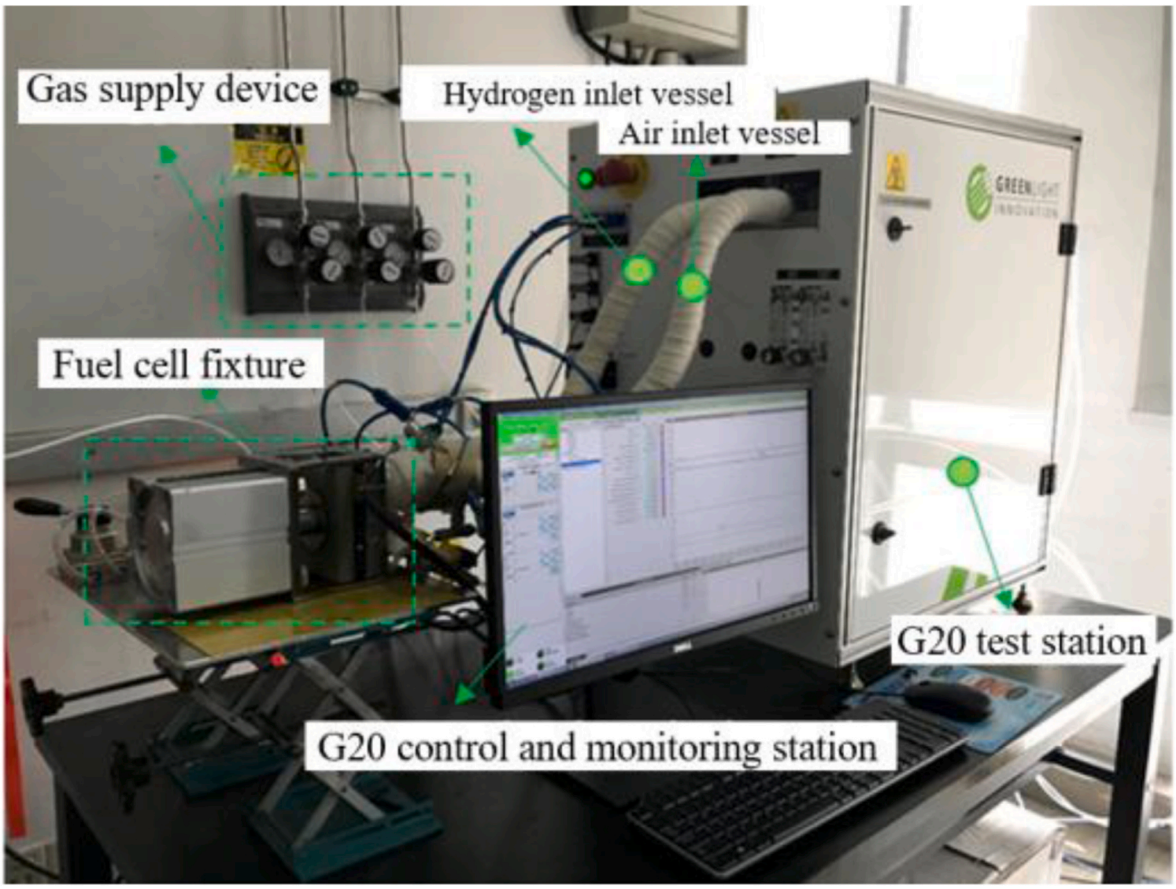

Fig. 6. Aging test bench of FC3.

Table 3
Range of controlled physical parameters (FC3).

Parameter	Value range
Cooling flow	0~10 L/min
Cooling temperature	20~80 °C
Gas humidification	0~100 % RH
Gas temperature	20~80 °C
Air flow	0~100 L/min
H2 flow	0~30 L/min
Gas pressure	0~2 bar
Fuel cell current	0~300 A

$$RE = \frac{\left|RUL_{true} - RUL_{pred}\right|}{RUL_{true}} \cdot 100\% \tag{9}$$

where N is the number of test data, y_i is the real value, and $\hat{y}_i$ is the predicted value. RUL_{true} is the true RUL, and RUL_{pred} is the predicted RUL.

In addition to prediction accuracy, this study also evaluates forecasting reliability using the Mean Standard Deviation (MS) to quantify point-wise uncertainty and the Average Uncertainty (AU) to assess the overall confidence of the entire forecasting process, respectively.

$$MS = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - x_a)^2} \tag{10}$$

$$AU = \frac{1}{N}\sum_{i=1}^{N} MS_i \tag{11}$$

where n is the number of Monte Carlo sampling iterations, x_i is the predicted value from the i-th sampling iteration, and x_a is the mean of the sampled values.

4. Results and discussion

In this section, a transfer pair (A-B) is defined, where A represents the source domain data and B represents the target domain data. Firstly, the effectiveness of the FDA technique within the TF-TL architecture in mitigating inter-domain discrepancies was validated through domain adaptation analysis. Subsequently, the TF-TL architecture was integrated with all baseline models mentioned in Section 2.3 to verify its enhancement of prediction performance. Furthermore, under cross-operating-condition and cross-device scenarios, the superiority of the TF-TL architecture was comprehensively evaluated through analyses of prediction reliability and generalization capability. Finally, the accuracy of TF-TL in assessing the RUL was discussed. The device's processor is the 12th Gen Intel(R) Core(TM) i5-12500H.

4.1. Domain adaptation analysis

This section configures multiple transfer tasks to validate the effectiveness of the FDA technique in eliminating inter-domain discrepancies from the perspective of different source and target domain combinations. Considering that the durability test under dynamic operating conditions involves nine different load levels, the intermediate load I4 is selected as an example for domain adaptation analysis. In addition, static and pseudo-dynamic operating condition data are chosen, forming three transfer pairs for analysis: (FC1–FC2), (FC1–I4), and (FC2–I4).

To verify its advantages in efficiency and robustness, the improved MMD was compared with the standard MMD on identical transfer tasks. For the standard MMD, the bandwidth parameter was optimized

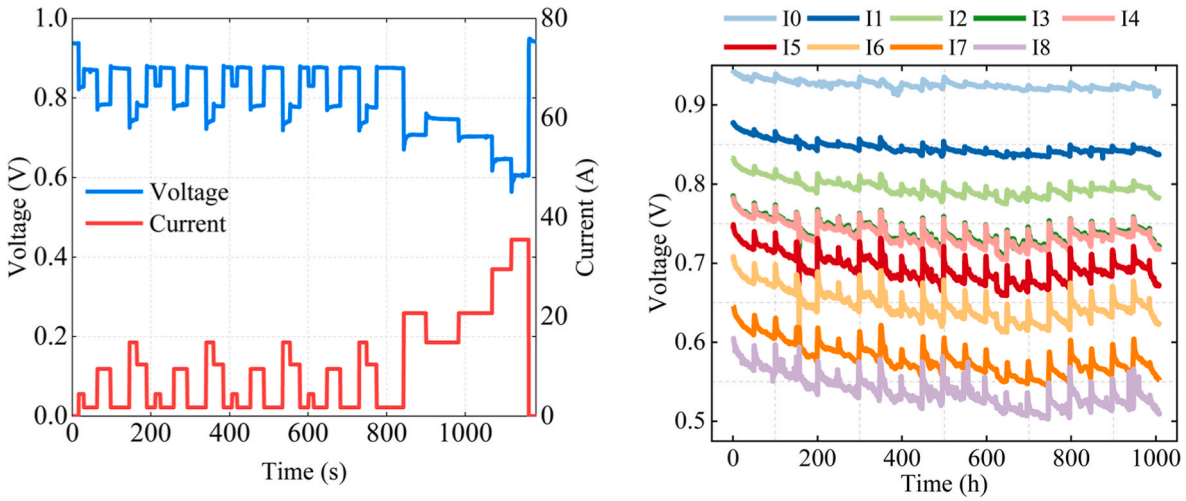

Fig. 7. (a) Current load and voltage values for a single cycle; (b) Voltage under the nine classified load levels.

through a grid search over the set {0.01, 0.1, 1, 10}, selecting the value that yielded the best performance on the source domain validation set. The performance comparison across three transfer pairs is summarized in Table 4. The results show that the improved MMD, via its adaptive bandwidth selection, achieves distribution alignment comparable to the finely-tuned standard MMD. It completely avoids the time-consuming manual tuning and grid search, leading to significantly improved computational speed. This demonstrates that the adaptive strategy effectively mitigates the standard MMD's sensitivity to its bandwidth hyperparameter and enhances the deployment efficiency of the domain adaptation module. This improvement aligns with the TF-TL architecture's objective of a lightweight and practical design.

Fig. 8 visually illustrates the changes in the voltage feature distributions for different transfer pairs before and after FDA processing using KDE. The corresponding quantitative assessment is provided by the introduced distribution similarity score *S*. Although the *S* of original voltage all exceed 0.95, the original voltage distributions across different domains exhibit significant differences due to multiple factors such as operating conditions, equipment specifications, and intrinsic structures. This distribution shift can directly impact the effectiveness of cross-domain knowledge transfer.

After FDA processing, the distribution discrepancies in all three pairs are significantly reduced. The source and target domains achieve distribution alignment while preserving their intrinsic characteristics. Specifically, the similarity *S* for the (FC1-FC2) pair reaches 0.999, nearly eliminating the inter-domain difference, which benefits from their high consistency in hardware specifications and basic operating conditions, with only minor local disturbances. For the (FC1-I4) pair, due to cross-domain differences arising from both operating condition variations and physical stack property differences, FDA does not completely eliminate the distribution shift. However, it effectively aligns the primary peaks and inflection point structures, causing the distribution trends to converge closely. Moreover, the S value for the (FC2-I4) pair is slightly higher than that for (FC1-I4). This phenomenon stems from the inherent connection between the pseudo-dynamic load of FC2 and the dynamic characteristics of the I4 load. The voltage response signals of FC2, subjected to a pseudo-dynamic load containing high-frequency ripple components, inherently encompass certain dynamic disturbance information in their frequency-domain structure. This makes their intrinsic modes closer to those of the voltage signals under the operating condition (I4). Consequently, when extracting low-frequency common features, the FDA technique can more effectively align their shared degradation patterns under dynamic disturbances, resulting in higher distribution similarity.

From a mechanistic perspective, the FDA can capture stable characteristic components of cross-domain voltage signals in the frequency domain and suppress non-stationary components caused by transient operational disturbances and equipment structural differences, thereby achieving the extraction and mapping of domain-invariant features. The aforementioned results not only validate the domain adaptation capability and robustness of FDA in transfer prediction tasks but also demonstrate that in cross-device, cross-condition transfer learning scenarios, this method can significantly reduce the distribution gap while preserving key performance features, laying the foundation for subsequent high-accuracy cross-domain prediction.

Table 4
Comparison of the operational efficiency between Improved MMD and MMD.

Transfer Tasks	Methods	Cost time (s)	*S*
FC1-FC2	MMD	15.3	0.9987
	Improved MMD	1.4	0.9989
FC1-I4	MMD	19.8	0.9853
	Improved MMD	1.1	0.9849
FC2-I4	MMD	17.5	0.9921
	Improved MMD	1.9	0.9923

4.2. Analysis of TF-TL prediction performance

To validate the enhancement effect of the TF-TL architecture on transfer prediction performance, this section systematically evaluates the prediction performance of the deep learning models described in Section 2.3 after integration with TF-TL. The results are compared against those obtained from direct transfer learning of the corresponding models. Here, baseline models employing direct transfer learning are denoted as Model-T, whereas models utilizing the TF-TL architecture are denoted as TF-Model. For the data, the (FC1-FC2) transfer task is adopted. The prediction accuracy of each model with 90 % of the target domain data serves as the basis for horizontal comparison, comprehensively evaluating the effectiveness of TF-TL in enhancing model accuracy.

Table 5 summarizes the error metrics for all prediction models. The baseline Model-T, based on direct transfer learning, achieves moderate accuracy. This performance primarily arises from the retained similarity in voltage distribution between the source (FC1) and target (FC2) domains. However, its relatively low R^2 score indicates limited model fitting capability. In contrast, the TF-Model significantly outperforms Model-T across all metrics. It substantially reduces both RMSE and MAPE while maintaining an R^2 value consistently above 0.95, demonstrating excellent fitting stability and generalization capability. All TF-Model variants demonstrate significantly superior prediction accuracy

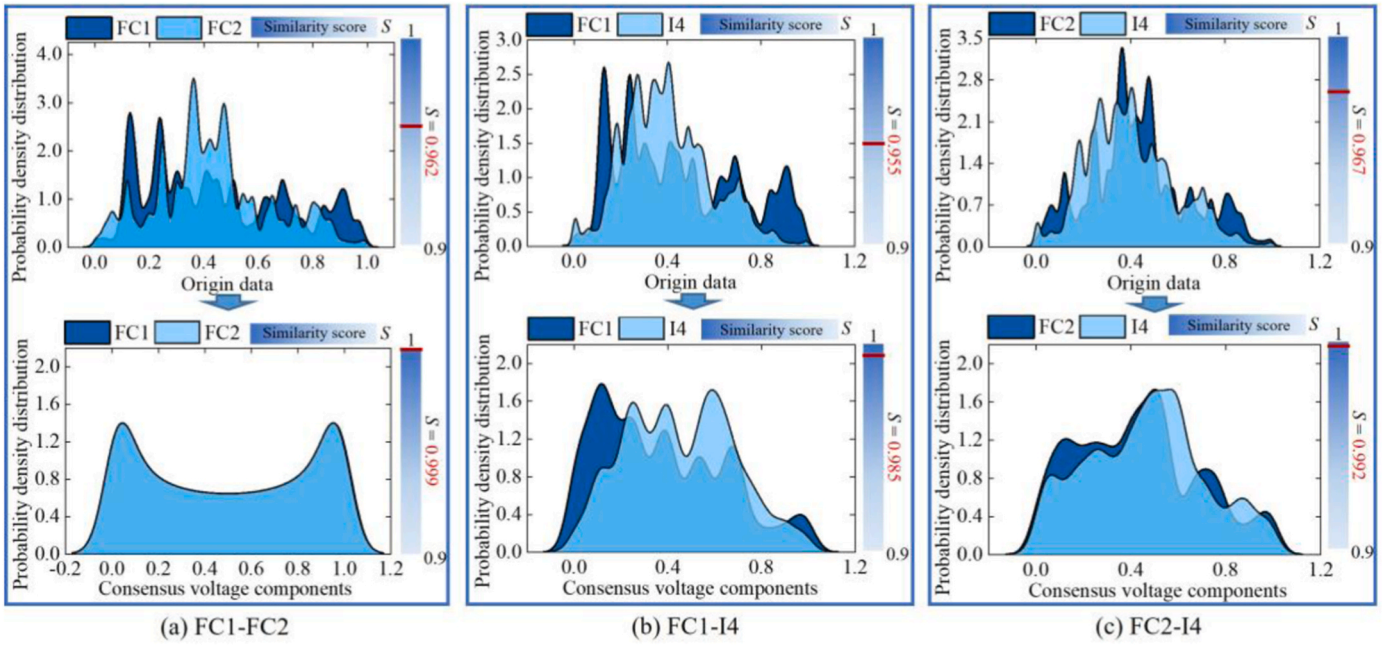

Fig. 8. Probability density distributions.

Table 5
Performance of each prediction model under dynamic conditions.

Model	RMSE	MAPE	R^2	Cost time (s)
LSTM-T	0.0073	0.1843	0.911	72.57
GRU-T	0.0069	0.1826	0.925	65.53
BiLSTM-T	0.0063	0.1713	0.932	134.71
BiGRU-T	0.0057	0.1431	0.941	118.73
TCN-T	0.0065	0.1755	0.935	59.43
TF-LSTM	0.0044	0.1148	0.972	84.24
TF-GRU	0.0041	0.1055	0.982	77.63
TF-BiLSTM	0.0035	0.0849	0.989	150.59
TF-BiGRU	0.0033	0.0657	0.993	114.87
TF-TCN	0.0036	0.0826	0.985	73.18

compared to the Model-T baseline. These results demonstrate that the FFT-based FDA technique within the TF-TL architecture is more effective than the traditional paradigm of screening features from high-dimensional time-domain operational parameters. This approach more effectively captures essential, domain-shared degradation features, thereby providing the model with more discriminative and generalizable inputs.

Computational efficiency is also detailed in Table 4, which lists the total time consumption for each model during training and inference. The results show that the TF-TL architecture moderately increases computational complexity, incurring additional time overhead. However, since PEMFC operational cycles typically span thousands of hours, this added computation time is negligible relative to the system's operational timeline. The substantial improvements in prediction accuracy and cross-domain generalization offered by TF-TL are far more critical. Therefore, the associated temporal cost is justified and acceptable for practical engineering applications.

4.3. Validation of TF-TL superiority

This section comprehensively validates the superiority of TF-TL from the perspectives of prediction reliability and generalization. Based on the comprehensive analysis of baseline model predictions in Section 4.2, the BiGRU model with the smallest error is selected for experimental validation.

4.3.1. Reliability analysis of transfer prediction

This section employs the Monte Carlo (MC) Dropout technique to quantify prediction uncertainty and assess the reliability of the TF-TL model. This method, grounded in Bayesian theory, constructs a probability distribution of prediction outcomes by performing multiple stochastic forward passes with dropout active during inference [37]. Here, 100 forward passes are performed during the prediction phase to systematically analyze the prediction credibility of the BiGRU-T and TF-BiGRU models in the (FC1-FC2) transfer task.

Fig. 9(a) and (b) present the point estimates and interval estimates for the BiGRU-T and TF-BiGRU models, respectively. Long-term prediction results show that compared to BiGRU-T, the prediction curve of TF-BiGRU aligns more closely with the true voltage degradation trajectory of the PEMFC, and its confidence interval is significantly narrower. Local dynamic comparisons (Fig. 9(c) and (d)) further reveal that under the same vertical axis scale, the confidence interval of TF-BiGRU envelopes the dynamic voltage variations much more effectively than that of BiGRU-T, with a more pronounced reduction in interval width. This verifies the superiority of TF-BiGRU not only in point estimate accuracy but also in its precise control over uncertainty.

To further investigate the prediction reliability of TF-TL, a detailed analysis is conducted at four typical time points (Fig. 9(e–h)), representing the initial operation phase, start-stop phase, normal operation phase, and a phase with severe voltage fluctuations of the PEMFC, respectively. The results indicate that at points A-C, TF-BiGRU exhibits sharper, more concentrated peaks in the probability density distribution compared to BiGRU-T, with prediction variance stably maintained in the order of E−4. This robustly validates the strong adaptability of the TF-TL architecture to voltage fluctuations in transfer scenarios. At point D, characterized by severe voltage fluctuations, the variance of TF-BiGRU increases, but it still demonstrates superior prediction accuracy and a more compact uncertainty distribution. In contrast to the bimodal, dispersed distribution of BiGRU-T, TF-BiGRU effectively suppresses excessive dispersion of uncertainty.

The above analysis demonstrates that baseline models equipped with the TF-TL architecture can better handle different fluctuation states of the fuel cell, highlighting the stability of the model in long-term lifecycle transfer predictions. By leveraging its low-frequency feature alignment mechanism, the TF-TL architecture provides high-accuracy and highly

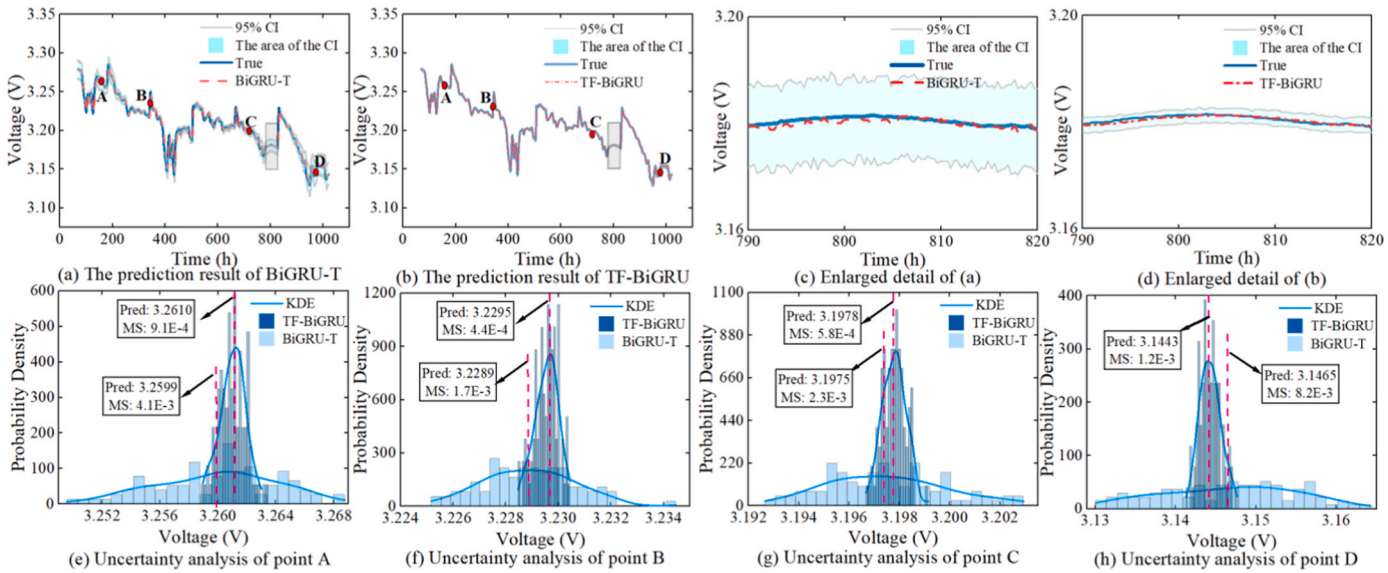

Fig. 9. Reliability analysis results of models.

reliable prediction results for cross-domain, cross-scenario transfer prediction of PEMFCs, offering a reliable quantitative basis for long-term health monitoring and early fault warning of PEMFC systems.

4.3.2. *Generalization analysis of transfer prediction*

Following the validation of TF-TL reliability in Section 4.3.1, this section further discusses its transfer generalization capability. Specifically, static operating condition data FC1 serves as the source domain,

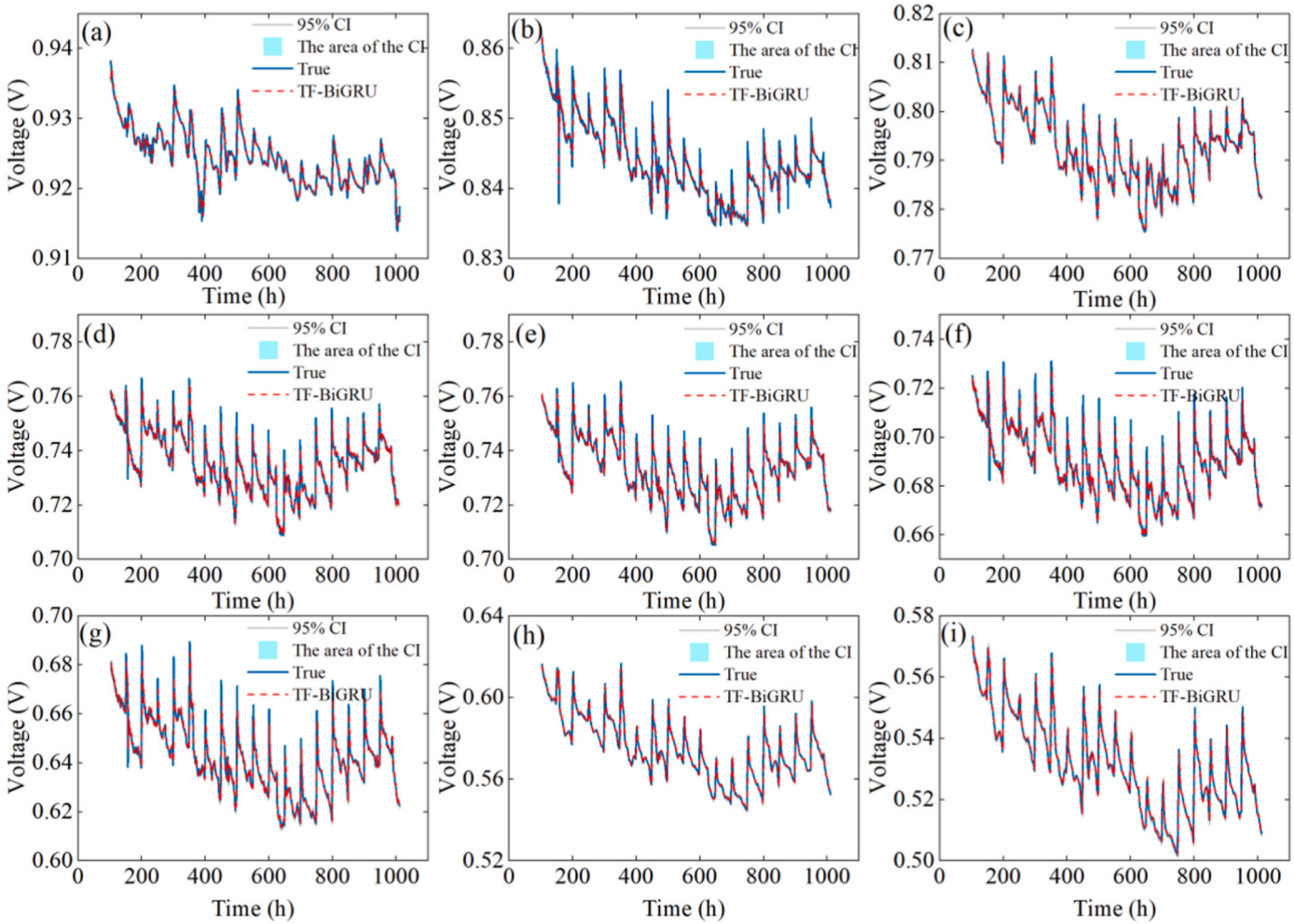

Fig. 10. Generalization validation of TF-BiGRU.

while the nine graded loads from the dynamic operating condition data FC3 serves as the target domains, validating the generalization ability of TF-TL in cross-device and cross-operating-condition scenarios. The testing ratio for the target domain remains consistent with previous sections.

Fig. 10 shows the prediction results of TF-BiGRU for the nine loads under dynamic operating conditions. The overall trends of voltage variation across the nine loads are similar, but significant differences exist in local details. For instance, loads like I1, I3, and I5 exhibit sharp voltage drops and rises around 150h during start-stop events, while load I0 shows frequent minor fluctuations within this interval. Facing loads I0-I8 with higher start-stop frequency and stronger fluctuations, TF-BiGRU accurately captures all voltage mutation points in long-term predictions and maintains close tracking of non-steady-state fluctuations during operation, validating the strong adaptability of the model to complex dynamic conditions.

Furthermore, Table 6 compares the generalization errors of BiGRU-T and TF-BiGRU in the cross-domain prediction task. Under cross-device and cross-condition transfer scenarios, the prediction performance of both models degrades. For BiGRU-T, the R^2 values for loads I3-I6 with severe voltage fluctuations all fall below 0.9, showing a significant decrease compared to its performance ($R^2 = 0.941$) in the same-device transfer task in Section 4.2. This phenomenon is attributed to the intensified domain discrepancy between source and target domains under cross-device and cross-condition scenarios, where the negative transfer effect caused by direct transfer leads to model performance degradation. In contrast, TF-BiGRU demonstrates superior generalization capability, achieving RMSE reductions of 15.1 %–60.4 % and maintaining R^2 consistently above 0.95 across all nine load transfer tasks.

Additionally, this section employs ridge plots to reveal the differences in prediction uncertainty distribution between the TF-BiGRU and BiGRU-T models across the nine load types. As shown in Fig. 11, the horizontal axis represents the uncertainty value range, and the vertical axis represents the different load conditions. The left and bottom axes correspond to TF-BiGRU, while the right and top axes correspond to BiGRU-T. The ridge peaks for TF-BiGRU are generally narrow and tall, and the overall color tone is bluer, indicating its prediction uncertainty is effectively suppressed, leading to a more trustworthy prediction process. In contrast, the ridges for BiGRU-T are wider and flatter, with the I5-I8 load regions appearing orange-red, revealing significantly higher dispersion in its prediction results. Table 7 quantifies the Average Uncertainty (AU) values for the different models, further confirming that as the load dynamics increase, the AU for both models show an upward trend. However, the AU for TF-BiGRU remains consistent in the order of E−4, whereas the AU for BiGRU-T jumps to the E−3 order under high-load conditions.

The above results indicate that the frequency-domain feature alignment mechanism of the TF-TL architecture, based on FDA technology, effectively eliminates inter-domain discrepancies. It maintains stable prediction accuracy while exhibiting high credibility in cross-device and cross-operating-condition scenarios, demonstrating its powerful generalization performance and reliability during the transfer prediction process.

Table 6
The prediction performance of TF-BiGRU under different loads.

Load	TF-BiGRU			BiGRU-T		
	RMSE	MAPE	R^2	RMSE	MAPE	R^2
I0	0.00073	0.05142	0.978	0.00086	0.05937	0.949
I1	0.00105	0.05245	0.972	0.00155	0.08106	0.917
I2	0.00074	0.04268	0.990	0.00187	0.13562	0.935
I3	0.00229	0.12730	0.967	0.00402	0.27331	0.863
I4	0.00216	0.11078	0.964	0.00397	0.27161	0.875
I5	0.00297	0.15632	0.958	0.00518	0.35772	0.855
I6	0.00304	0.15669	0.953	0.00538	0.38760	0.864
I7	0.00159	0.11164	0.989	0.00369	0.34031	0.939
I8	0.00139	0.10955	0.991	0.00332	0.34308	0.942

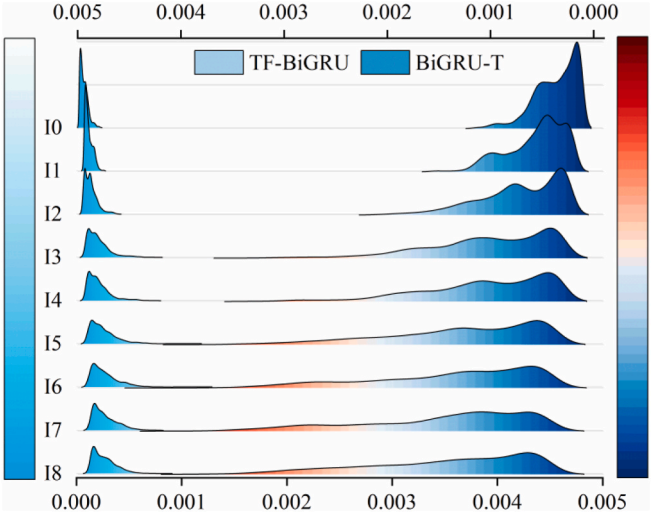

Fig. 11. Uncertainty distribution of predictions under multiple loads.

4.4. RUL estimation of PEMFCs

This section establishes three sets of transfer experiments for FC1, FC2, and FC3, respectively, to analyze the accuracy of the TF-TL architecture in evaluating the RUL under long-term prediction. The offline model is pre-trained based on the stable voltage characteristic components captured in the frequency domain by FDA, and is subsequently employed for the RUL prediction task on the target domain. Considering the nine load groups separated from FC3 exhibit roughly similar trends, the voltage data from load I4 of FC1 and FC2 are selected to evaluate the RUL across the three stacks. The three transfer experiments are (FC2-FC1), (FC1-FC2), and (FC1-I4). Testing ratios of 0.9, 0.85, 0.8, 0.75, and 0.7 for the target domain are set to validate the capability of the TF-TL architecture in long-term RUL estimation.

In addition to the Relative Error (RE) accuracy metric from Section 3.3 for evaluating RUL estimation error, the concept of a Trustworthy Area (TA) is introduced to assess the credibility of the estimation process. The specific formula is as follows:

$$RUL_{true} \times (1 - \alpha_{low}) \leq TA \leq RUL_{true} \times (1 + \alpha_{up}) \tag{12}$$

where RUL_{ture} is the actual RUL. α is an accuracy modifier that defines the lower and upper bounds of the TA, respectively. For fuel cell prognosis, an estimated RUL that is less than the actual RUL is generally more acceptable when a deviation exists. Conversely, an overestimated RUL often leads to delayed maintenance schedules or even premature critical failures. Accordingly, asymmetric bounds are set with $\alpha = 0.1$ for the lower limit and $\alpha = 0.2$ for the upper limit. The estimation process is deemed reliable if the estimated RUL falls within the TA.

For FC1, the initial failure threshold is set to 96 % of its initial voltage (3.345 V), corresponding to 3.2112 V, with the earliest recorded failure time of 893 h. To fully utilize the experimental data of FC2, its failure threshold is determined as 95 % of the initial voltage (3.33 V), equivalent to 3.1635 V, and the earliest failure time is 921 h. Based on the existing analysis of FC3, the voltage at the 620 h sharp voltage drop is adopted as the failure threshold for the I4 load, specifically 0.7258 V [22].

Fig. 12 presents the RUL evaluation results for the three transfer prediction tasks. At a testing ratio of 0.9, the estimated RUL lies above the true RUL, which is attributed to estimation lag caused by the excessively long testing duration. However, at all other testing time points, the estimated RUL consistently falls below the true RUL, and all

Table 7
Average uncertainty under different loads.

	Load	I0	I1	I2	I3	I4	I5	I6	I7	I8
AU(E−4)	TF-BiGRU	0.71	1.15	1.36	2.05	2.08	2.54	2.7	2.69	2.62
	BiGRU-T	3.66	5.55	7.13	10.2	10.5	12.8	13.8	14.5	14.3

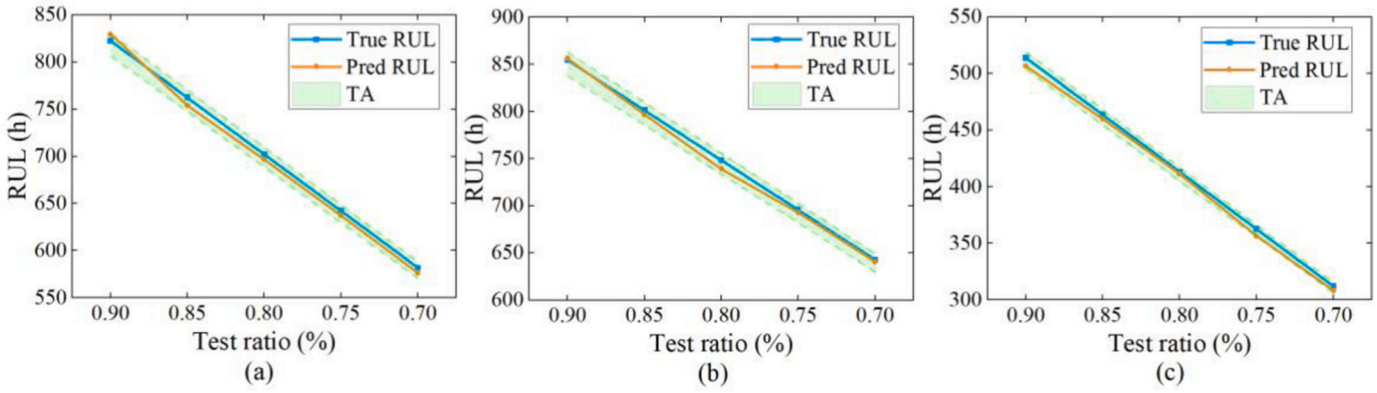

Fig. 12. RUL estimation results for (a) target domain FC1 (b) target domain FC2 (c) target domain I4.

estimated RUL values remain within the TA, indicating a high degree of credibility in the evaluation process of the model optimized by the TF-TL architecture.

Furthermore, Fig. 13 shows a 3D surface distribution of the RE metric during the RUL evaluation, where the X-axis represents the testing ratio, the Y-axis represents the target domains in the three transfer tasks, and the Z-axis represents the magnitude of RE. It can be observed that the maximum RE does not exceed 1.9 %. The results from these multiple transfer tasks demonstrate that TF-TL exhibits accurate RUL evaluation performance across different operating conditions and devices.

4.5. Comparison of superiority over existing methods

To validate the superiority of the proposed TF-TL architecture for RUL estimation, this section conducts a systematic comparison with state-of-the-art methods. The benchmarked approaches include hybrid data-driven/transfer learning methods [23,24], the Self-Adaptive Digital Twin (SADT) [38], and an Improved Physics-Informed Neural Network (Improved PINN) [39,40].

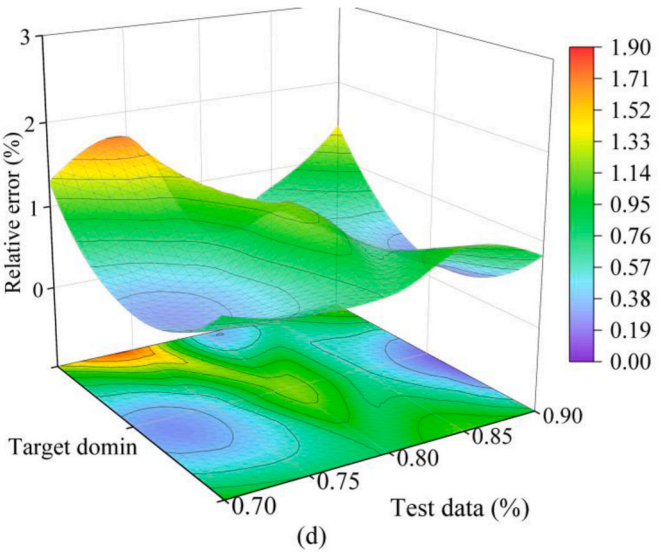

Fig. 13. RE of RUL estimated.

To ensure a fair and consistent comparison, all benchmark methods are evaluated using the best performance reported in their original studies. The RUL of FC2 serves as the unified evaluation standard. Note that the EoL threshold definition varies across studies. To address this variation, we evaluate TF-TL under multiple EoL threshold conditions to ensure an equitable comparison.

Table 8 summarizes the RUL prediction accuracy for each method. The results show that the Improved PINN achieves the highest accuracy, benefiting from its effective integration of mechanistic information. However, it exhibits limitations in tracking the detailed, continuous dynamics of performance degradation. In contrast, the proposed TF-TL method maintains low estimation error across threshold settings and accurately captures detailed degradation trends. These results demonstrate that the TF-TL method offers excellent accuracy, robustness, and stability in predicting the aging trajectory and lifetime of PEMFCs.

5. Conclusion

This study addresses the challenges of distribution shift and high prediction uncertainty in cross-domain degradation prediction for PEMFCs by proposing a TF-TL architecture based on the FDA technology. This architecture leverages the FDA to directly extract low-frequency consensus knowledge from voltage signals, effectively reducing the distribution discrepancy between source and target domains. It enhances various baseline models in a lightweight, embedded manner, significantly improving prediction accuracy and reliability. The main conclusions of this paper are as follows:

Table 8
Comparison of the RUL prediction accuracy.

Method	True RUL (h)	Pred RUL (h)	RE (%)
Proposed method	225	221.53	0.21
	325	321.07	1.21
	425	423.55	0.34
Bi-LSTM-AM [23]	236	232	5.04
LQLSTM [24]	224	221	1.33
SADT [38]	392	371	5.36
	390	369	5.38
SOTA [39]	–	–	2.80
Improved PINN [40]	–	–	0.10

1) The proposed FDA technique significantly eliminates feature distribution differences among static, pseudo-dynamic, and dynamic operating conditions through intrinsic frequency-domain alignment of cross-domain voltage signals, achieving a maximum inter-domain distribution similarity of up to 0.999.
2) The TF-TL architecture can be embedded into different deep learning models, universally enhancing their cross-domain prediction performance. Taking TF-BiGRU as an example, its RMSE decreased by up to 60.4 % compared to the direct transfer method, while its R^2 consistently remained above 0.95. Furthermore, the confidence interval width was significantly narrowed, and the average uncertainty was stably maintained at the E−4 level, demonstrating high accuracy and reliability.
3) In generalization tasks across different operating conditions and devices, the TF-TL architecture exhibits strong adaptability and engineering practical value. In multiple long-term prediction tasks, the relative error of RUL estimation was below 1.9 %, and the results consistently fell within the Trustworthy Area, providing a reliable basis for the online health management of PEMFCs.

The proposed TF-TL architecture exhibits promising potential for deployment in real-world PEMFC online health management systems, primarily due to its lightweight and modular design. However, several practical challenges remain to be addressed in future work. First, the generalization of the model under extreme or unforeseen operating conditions requires further validation. Second, developing an efficient online model update mechanism to adapt to the continuous aging of individual fuel cell stacks is crucial for long-term application. Finally, the micro-level quantitative correlation between the low-frequency characteristics and specific physicochemical degradation mechanisms also merits further investigation.

CRediT authorship contribution statement

Hangyu Wu: Writing – original draft, Methodology, Formal analysis, Conceptualization. **Fulin Fan:** Investigation, Formal analysis. **Kai Song:** Project administration, Funding acquisition. **Chuanyu Sun:** Formal analysis, Data curation. **Yang Li:** Writing – review & editing, Formal analysis. **Changjun Xie:** Supervision, Formal analysis. **Xuan Meng:** Supervision. **Jian Mei:** Supervision. **Siew Hwa Chan:** Supervision, Formal analysis.

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.

Acknowledgement

This work was supported by the Key Research & Development Plan of Heilongjiang Province (Grant No. 2024ZXJ03C06).

Data availability

Data will be made available on request.

References

[1] Meng X, Sun C, Mei J, Tang X, Hasanien HM, Jiang J, et al. Fuel cell life prediction considering the recovery phenomenon of reversible voltage loss. J Power Sources 2025;625:235634. https://doi.org/10.1016/j.jpowsour.2024.235634.
[2] Zhou H, Wu X, Li Y, Fan Z, Chen W, Mao J, et al. Model optimization of a high-power commercial PEMFC system via an improved grey wolf optimization method. Fuel 2024;357:129589. https://doi.org/10.1016/j.fuel.2023.129589.
[3] Yu X, Yang Y, Xie C, Li Y, Zhao B, Zhang L, et al. Remaining-useful-lifetime prediction of proton exchange membrane fuel cell considering model uncertainty quantification on the full-time scale. IEEE Trans Transport Electrification 2024;10:7443–55. https://doi.org/10.1109/TTE.2023.3336324.
[4] Qiu Y, Zeng T, Zhang C, Wang G, Wang Y, Hu Z, et al. Progress and challenges in multi-stack fuel cell system for high power applications: architecture and energy management. Green Energy Intell Transport 2023;2:100068. https://doi.org/10.1016/j.geits.2023.100068.
[5] Liu C, Chen Y, Xu R, Ruan H, Wang C, Li X. Co-optimization of energy management and eco-driving considering fuel cell degradation via improved hierarchical model predictive control. Green Energy Intell Transport 2024;3:100176. https://doi.org/10.1016/j.geits.2024.100176.
[6] Piao S, Huang R, Tsung F. CRULP: reliable RUL estimation inspired by conformal prediction. IEEE Trans Instrum Meas 2025;74:1–11. https://doi.org/10.1109/TIM.2024.3522678.
[7] Li D, Chen J, Huang R, Chen Z, Li W. Sensor-aware CapsNet: towards trustworthy multisensory fusion for remaining useful life prediction. J Manuf Syst 2024;72:26–37. https://doi.org/10.1016/j.jmsy.2023.11.009.
[8] Huang R, Chen Y, Liu C. Revealing state-feature dependencies in dynamic degradation: an Exofeature-Aware transformer for battery state-of-health prediction. IEEE Trans Consum Electron 2025. https://doi.org/10.1109/TCE.2025.3631954. 1–1.
[9] Zuo J, Steiner NY, Li Z, Hissel D. Health management review for fuel cells: focus on action phase. Renew Sustain Energy Rev 2024;201:114613. https://doi.org/10.1016/j.rser.2024.114613.
[10] Long B, Wu K, Li P, Li M. A novel remaining useful life prediction method for hydrogen fuel cells based on the gated recurrent unit neural network. Appl Sci 2022;12:432. https://doi.org/10.3390/app12010432.
[11] Wilberforce T, Alaswad A, A G-P, Xu Y, Ma X, Panchev C. Remaining useful life prediction for proton exchange membrane fuel cells using combined convolutional neural network and recurrent neural network. Int J Hydrogen Energy 2023;48:291–303. https://doi.org/10.1016/j.ijhydene.2022.09.207.
[12] Zhu W. High accuracy and adaptability of PEMFC degradation interval prediction with Informer-GPR under dynamic conditions. 2024.
[13] Sun B, Liu X, Wang J, Wei X, Yuan H, Dai H. Short-term performance degradation prediction of a commercial vehicle fuel cell system based on CNN and LSTM hybrid neural network. Int J Hydrogen Energy 2023;48:8613–28. https://doi.org/10.1016/j.ijhydene.2022.12.005.
[14] Wang C, Li Z, Outbib R, Dou M, Zhao D. Symbolic deep learning based prognostics for dynamic operating proton exchange membrane fuel cells. Appl Energy 2022;305:117918. https://doi.org/10.1016/j.apenergy.2021.117918.
[15] Zhu W, Li Y, Xu Y, Zhang L, Guo B, Xiong R, et al. Interval prediction of fuel cell degradation based on voltage signal frequency characteristics with TimesNet-GPR under dynamic conditions. J Clean Prod 2025;486:144503. https://doi.org/10.1016/j.jclepro.2024.144503.
[16] Shen D, Yang D, Lyu C, Hinds G, Wang L, Bai M. Detection and quantitative diagnosis of micro-short-circuit faults in lithium-ion battery packs considering cell inconsistency. Green Energy Intell Transport 2023;2:100109. https://doi.org/10.1016/j.geits.2023.100109.
[17] Zhu J, Chen N, Shen C. A new data-driven transferable remaining useful life prediction approach for bearing under different working conditions. Mech Syst Signal Process 2020;139:106602. https://doi.org/10.1016/j.ymssp.2019.106602.
[18] Zhou Z, Liu Y, You M, Xiong R, Zhou X. Two-stage aging trajectory prediction of LFP lithium-ion battery based on transfer learning with the cycle life prediction. Green Energy Intell Transport 2022;1:100008. https://doi.org/10.1016/j.geits.2022.100008.
[19] Zhu R, Peng W, Wang D, Huang C-G. Bayesian transfer learning with active querying for intelligent cross-machine fault prognosis under limited data. Mech Syst Signal Process 2023;183:109628. https://doi.org/10.1016/j.ymssp.2022.109628.
[20] Ding Y, Ding P, Zhao X, Cao Y, Jia M. Transfer learning for remaining useful life prediction across operating conditions based on multisource domain adaptation. IEEE ASME Trans Mechatron 2022;27:4143–52. https://doi.org/10.1109/TMECH.2022.3147534.
[21] Li Z, Yuan X, Zhang C, Mao J, Wang Y. Long-term degradation prediction method for proton exchange membrane fuel cells based on hybrid transfer learning. Int J Hydrogen Energy 2025;106:781–9. https://doi.org/10.1016/j.ijhydene.2025.01.405.
[22] Zhang X, Yue J, Huang L, Qiu D, Jiang Y. An enhanced transfer learning based on addernet and deep domain confusion for prognostic of proton exchange membrane fuel cells. Int J Hydrogen Energy 2024;77:795–807. https://doi.org/10.1016/j.ijhydene.2024.06.090.
[23] Kebede GA, Lo S-C, Wang F-K, Chou J-H. Transfer learning-based deep learning models for proton exchange membrane fuel remaining useful life prediction. Fuel 2024;367:131461. https://doi.org/10.1016/j.fuel.2024.131461.
[24] Wang F-K, Kebede GA, Lo S-C, Woldegiorgis BH. An embedding layer-based quantum long short-term memory model with transfer learning for proton exchange membrane fuel stack remaining useful life prediction. Energy 2024;308:133054. https://doi.org/10.1016/j.energy.2024.133054.
[25] Wang C, Dou M, Li Z, Outbib R, Zhao D, Zuo J, et al. Data-driven prognostics based on time-frequency analysis and symbolic recurrent neural network for fuel cells under dynamic load. Reliab Eng Syst Saf 2023;233:109123. https://doi.org/10.1016/j.ress.2023.109123.
[26] Zhang Y, Ma R, Xie R, Feng Z, Liang B, Li Y. A degradation prediction method for PEM fuel cell based on deep temporal feature extraction and transfer learning. IEEE Trans Transp Electrific 2024;10:203–12. https://doi.org/10.1109/TTE.2023.3262588.

[27] Yang W, Chen Z, Zhao H, Li J, Chen S, Shi C. Bidirectional denoising method based on fast fourier transform analysis for TBM field penetration data. Tunn Undergr Space Technol 2025;158:106436. https://doi.org/10.1016/j.tust.2025.106436.
[28] Ji H, Zhang X, Wang T, Yang K, Jiang J, Xing Z. Oil spill area prediction model of submarine pipeline based on BP neural network and convolutional neural network. Process Saf Environ Prot 2025;199:107264. https://doi.org/10.1016/j.psep.2025.107264.
[29] Shi J, Teh J, Lai C-M. Wind power prediction based on improved self-attention mechanism combined with Bi-directional temporal convolutional network. Energy 2025;322:135666. https://doi.org/10.1016/j.energy.2025.135666.
[30] Wei N, Yin C, Yin L, Tan J, Liu J, Wang S, et al. Short-term load forecasting based on WM algorithm and transfer learning model. Appl Energy 2024;353:122087. https://doi.org/10.1016/j.apenergy.2023.122087.
[31] Pan R, Yang D, Wang Y, Chen Z. Performance degradation prediction of proton exchange membrane fuel cell using a hybrid prognostic approach. Int J Hydrogen Energy 2020;45:30994–1008. https://doi.org/10.1016/j.ijhydene.2020.08.082.
[32] Wu H, Wang W, Li Y, Zhu W, Xie C, Gooi HB. Hybrid physics-based and data-driven prognostic for PEM fuel cells considering voltage recovery. IEEE Trans Energy Convers 2023:1–11. https://doi.org/10.1109/TEC.2023.3311460.
[33] Zuo J, Lv H, Zhou D, Xue Q, Jin L, Zhou W, et al. Deep learning based prognostic framework towards proton exchange membrane fuel cell for automotive application. Appl Energy 2021;281:115937. https://doi.org/10.1016/j.apenergy.2020.115937.
[34] Zuo J. Long-term dynamic durability test datasets for single proton exchange membrane fuel cell. Data Brief 2021.
[35] Wu H, Zhang R, Zhu W, Xie C, Li Y, Yang Y, et al. A health index-based approach for fuel cell lifetime estimation. iScience 2024;27:110979. https://doi.org/10.1016/j.isci.2024.110979.
[36] Liu H, Chen J, Hissel D, Su H. Remaining useful life estimation for proton exchange membrane fuel cells using a hybrid method. Appl Energy 2019;237:910–9. https://doi.org/10.1016/j.apenergy.2019.01.023.
[37] Le X-H, Van Binh D, Lee G. Performance and uncertainty analysis in deep learning frameworks for streamflow forecasting via Monte Carlo dropout technique. J Hydrol: Reg Stud 2025;61:102668. https://doi.org/10.1016/j.ejrh.2025.102668.
[38] Zhang M, Amiri A, Xu Y, Bastin L, Clark T. Self-adaptive digital twin of fuel cell for remaining useful lifetime prediction. Int J Hydrogen Energy 2024;89:634–47. https://doi.org/10.1016/j.ijhydene.2024.09.266.
[39] Zhou D, Al-Durra A, Zhang K, Ravey A, Gao F. Online remaining useful lifetime prediction of proton exchange membrane fuel cells using a novel robust methodology. J Power Sources 2018;399:314–28. https://doi.org/10.1016/j.jpowsour.2018.06.098.
[40] Ko T, Kim D, Park J, Lee SH. Physics-informed neural network for long-term prognostics of proton exchange membrane fuel cells. Appl Energy 2025;382:125318. https://doi.org/10.1016/j.apenergy.2025.125318.