Fuel Cell Digital Twin for Remaining Useful Lifetime Prediction and Optimisation based on Physics-guided Neural Network

Ming Zhang[c], Amirpiran Amiri[a,*], Tabbi Wilberforce[b], Clyde-Theodore Batista[a], Yuchun Xu[a]

[a] *College of Engineering and Physical Sciences, Aston University, Birmingham, B4 7ET, UK*
[b] *Faculty of Natural, Mathematical and Engineering Sciences, King's College London, London, WC2R 2LS, UK*
[c] *Department of Digital Intelligence, Beijing Bohua Xinzhi Technology Co., Ltd., Beijing, 100101, China*

Abstract

Fuel cells are a critical component in the transition from hydrogen to electricity, with significant potential applications in the transportation sector to facilitate sustainable energy transformation. However, issues related to fuel cell durability present major challenges that hinder their widespread adoption. Digital twin (DT) technology has gained significant attention in recent years for enhancing the lifespan of fuel cells through more effective predictive maintenance. However, developing an effective DT model typically requires extensive data. In this work, we propose a novel fuel cell digital twin (FCDT) framework based on the advanced physics-guided neural network (PGNN) method, which addresses the challenge of massive data dependence while also extending the fuel cell's operational longevity. The PGNN-based approach combines the strengths of physical modeling and data-driven methods, enabling the model to be trained with limited data and providing accurate predictions of the remaining useful lifetime (RUL) using operational parameters as inputs. By interacting with the trained PGNN-based model, the Nelder-Mead algorithm automatically optimises real-time operational parameters, identifying the optimal solution to extend the fuel cell's remaining

*Corresponding author
Email address: a.p.amiri@aston.ac.uk (Amirpiran Amiri)

lifespan. The experiments validate the PGNN-based model's capability to capture complex degradation patterns and provide accurate RUL predictions, even under limited data conditions. Additionally, the optimization process confirms the effectiveness of proposed FCDT framework in refining operating parameters and significantly extending the fuel cell's lifespan.

Keywords: Digital Twin, Fuel Cells, Remaining Useful Lifetime Prediction, Lifetime Optimisation, Physics-guided Neural Network

1. Introduction

Achieving global net-zero emissions and decarbonization has become a central focus of the current energy revolution, driving the shift towards low-carbon and renewable energy sources as the preferred carriers for a more sustainable future. Hydrogen and fuel cell technologies are playing an important role in the transportation sector, offering a clean and efficient solution for reducing emissions and promoting sustainability in vehicles, particularly in heavy-duty and long-distance transport applications [1]. Fuel cells generate electricity through an electrochemical reaction between hydrogen and oxygen, with water as the only byproduct, making them a zero-emission technology that contributes significantly to reducing carbon footprints in energy production. Proton exchange membrane fuel cells (PEMFCs) are highly competitive in transportation applications due to their high power density, quick start-up time, and low operating temperature, and they are already commercially used in most fuel cell vehicles [2]. However, due to the complex structure and material properties of the fuel cell stack, hydrogen-based fuel cells are highly sensitive to operating conditions and environmental factors, leading to irreversible degradation and reduced lifespan [3]. Recent studies show that the average lifetime of PEMFCs is currently around 2000 to 3000 hours, with ongoing research focused on extending this to 6000 to 7000 hours by 2030 [4]. The durability issues of fuel cell stacks currently hinder the large-scale commercial adoption of fuel cell vehicles, as the performance and lifespan of PEMFCs are not yet sufficient to meet the demands of future widespread use.

Constructing an accurate fuel cell degradation model is a potential way to effectively predict the fuel cell's state of health (SoH) and remaining useful life (RUL), facilitating the resolution of fuel cell durability issues. The physis-based degradation modelling method can be fully based on theoreti-

cal principles, or they can combine theoretical principles with experimental data to better capture degradation behavior. For instance, Liu et al. [5] developed a mathematical model using the finite element method to study changes in the catalyst layer (CL) microstructure due to mechanical degradation. Moein-Jahromi et al. [6] introduced a new experiment-based algorithm to evaluate the effect of cyclic loading on CL performance loss, and in a subsequent study [7], they enhanced their model to better predict the aging process of fuel cells under cyclic loading. Zheng et al. [8] developed a model for Pt degradation in the membrane electrode assembly, considering Pt mass loss and particle growth mechanisms. These methods are entirely based on theoretical principles, requiring a deep understanding of the system's degradation mechanisms and precise knowledge of internal parameters. While these models can be beneficial during the fuel cell design stage to improve materials, they are less effective in providing accurate degradation predictions during the system's operational stage. The physical modelling method with experimental data is a effective way to capture degradation behaviour for the FC operation stage. For instance, Zhou et al. [9] developed a degradation prediction model based on a multiphysical aging model combined with a particle filter approach, while Zhang et al. [10] proposed a prognostic model for RUL prediction by integrating multiple data sources, including stack voltage and periodic physical characterization of the fuel cell. Ou et al. [11] proposed a semi-empirical degradation model, based on electrochemical surface area and equivalent resistance, to predict the voltage degradation trend of PEMFCs. Chen et al. [12] used a comprehensive degradation model rooted in thermodynamics and electrode kinetics, coupled with a particle filter algorithm, to predict RUL based on performance degradation data. These prediction methods rely on estimating and updating aging parameters with a physical model to predict voltage. However, they cannot fully model degradation or establish an accurate relationship between operating parameters (such as temperature, pressure, and flow rate) and voltage degradation. This limitation confines the model to steady-state operations with constant operating parameters. When operating conditions change, the model requires recalibration.

On the other hand, data-driven models, especially those utilizing deep learning, have demonstrated significant potential in predicting the RUL of fuel cells. Unlike physics-based models, these approaches do not require detailed information about internal system parameters or precise degradation mechanisms. Instead, they rely on learning degradation patterns directly

from extensive empirical datasets. In this way, the model development process becomes significantly simpler and more flexible, allowing for easier adaptation to varying data and operational conditions. Wang et al. [13] employed a bi-directional long short-term memory network with an attention mechanism (BILSTM-AT) to predict the voltage degradation of the PEMFC stack. Mezzi et al. [14] proposed a prognostic approach for fuel cells based on the Echo State Network, aimed at predicting performance under variable load profiles. Zhang et al. [15] proposed a bi-directional gated recurrent unit (Bi-GRU) neural network for PEMFC power reactor failure prediction in both static and dynamic scenarios. He et al. [16] introduced a novel health indicator (HI) using an auto-encoder, predicting future voltage with a long short-term memory (LSTM) network. Tabbi et al. [17] combined convolutional and recurrent neural networks to predict voltage using multiple operational parameters as input. Chen et al. [18] employed a bidirectional LSTM with Bayesian optimisation for hyperparameter tuning, predicting performance degradation under varying time intervals. Yang et al. [19] integrated an improved grey wolf optimizer (IGWO) and backpropagation (BP) neural network to predict the RUL of PEMFC systems in real-world traffic conditions. To facilitate long-term degradation prediction, Benaggoune et al. [20] investigated a new time-dependent fuel cell performance prediction model based on dilated convolutional neural networks, while Wang et al. [21] proposed a navigation sequence-driven LSTM (NSD-LSTM) for long-term prognostics. To enhance the adaptability of data-driven models, Mohammad et al. [22] proposed a novel WaveNet-GRU deep learning model combined with transfer learning for predicting PEMFC degradation across diverse datasets. Although data-driven models do not require precise physical knowledge or a deep understanding of degradation mechanisms, they are heavily constrained by the need for large and exhaustive datasets to build and validate efficient models. Additionally, these models often struggle to capture or investigate the underlying degradation phenomena associated with the operational parameters of fuel cells, limiting their ability to provide insights into the physical processes driving degradation. They also lack the capacity to integrate real-time operational parameters for fuel cell lifespan extension, making them less effective for real-time optimisation and adaptive control.

The novel physics-guided neural network (PGNN) framework offers significant advantages by combining the strengths of both physics-based and data-driven methods [23]. In this approach, outputs from physics-based

model simulations are integrated with observational data in a hybrid setup, enabling the neural network to leverage both empirical data and underlying physical laws for more accurate predictions. Furthermore, PGNNs embed physical laws directly into the neural network, reducing the dependence on large datasets and allowing for generalization even with limited data. Several successful cases for the application of this modelling approach for various engineering problems have been expressed in the literature. Wang et al. [24] applied the PGNN framework to predict tool wear, accounting for complex cutting conditions and dynamic physical parameters. Xu et al. [25] utilized PGNNs to solve power flow equations, addressing the limitations of data-driven methods that struggle with performance and generalizability due to oversimplified assumptions or neglect of physical laws. Zhang et al. [26] explored a physics-guided fuel-switching strategy for stable combustion in swirl burners. Ye et al. [27] employed the PGNN framework to estimate the state of health of lithium-ion batteries. The potential of PGNN has not yet been fully explored in the context of fuel cell degradation prediction. Developing an accurate degradation prediction model for fuel cells still requires substantial data.

The digital twin (DT) technology represents a groundbreaking approach to managing and optimizing fuel cell systems, specifically addressing the challenge of fuel cell durability. A digital twin is a sophisticated virtual model that mirrors the physical fuel cell, continuously updated with information from embedded sensors, offering a comprehensive view of the fuel cell's performance and operational characteristics. Meraghni et al. [28] proposed a data-driven DT that integrates the system's physical knowledge with a deep transfer learning model, utilizing a stacked denoising autoencoder to continuously update the DT using online measurements. Yue et al. [29] developed a multi-input data-driven model for fuel cell degradation by combining a convolutional neural network and long short-term memory network, using transfer learning to leverage historical data for reliable real-time predictions, especially in the early stages. Zhang et al. [30] proposed a self-adaptive digital twin (SADT) for RUL prediction, utilizing a novel quantile Huber loss function and domain adaptation techniques to enhance adaptability across varying operational scenarios. These DTs have demonstrated strong performance in RUL prediction across various operating conditions. However, their effectiveness often relies on extensive data, which is not always practical or applicable in real-world scenarios.

Therefore, we propose a novel fuel cell digital twin (FCDT) framework

based on the PGNN method to address the fuel cell durability challenge under limited data conditions. Inspired by [31], our FCDT framework offers valuable insights into how various operating parameters influence the system's behavior, enabling timely predictions and optimisation. The digital model is built using the PGNN method, which combines a fuel cell physical model with a deep neural network, enabling accurate RUL prediction under limited data conditions. This model estimates voltage degradation based on real-time operational parameters, including temperature, pressure, flow rate, humidity, and current. The RUL is calculated once a predefined threshold is determined. To facilitate timely optimization, the Nelder-Mead algorithm is employed to optimize real-time operational parameters—such as temperature, pressure, humidity, and current—based on the model's predictions. This real-time optimization mitigates degradation and extends the fuel cell's RUL by identifying the optimal operational parameters. To the best of our knowledge, this is the first attempt to use the PGNN method to establish an FCDT digital model for RUL prediction. We explore how the PGNN-based model can be trained with limited data to perform RUL prediction, reducing data dependence compared to traditional data-driven methods. Additionally, the multivariable optimisation algorithm is used for the first time with the FCDT digital model to optimize real-time operating parameters and extend the fuel cell's RUL.

2. Preliminary

2.1. Fuel Cell Lifetime Challenge

The fuel cell lifetime optimisation problem is defined as maximizing the RUL by adjusting the operating parameters over its service time. The solution consists of two key steps: RUL precise prediction followed by its optimisation for maximum longevity. These can be achieved through three tasks:

1. Prediction Model: the first step is to develop a predictive model capable of forecasting the voltage based on the real-time operating parameters:

$$y_i = F_{\text{pred}}(t_i, x_i) \tag{1}$$

Here, $x_i = [T_i, P_i, F_i, H_i, I_i]$ are the inputs (temperature, pressure, flow rate, humidity, and current) at time step t_i, and y_i is the predicted voltage at the same time step.

2. RUL Calculation: based on the predicted voltage, the RUL can be calculated by determining the time at which the voltage reaches a predefined threshold y_{thre}:

$$t_{\text{EOF}} = F_{\text{pred}}^{-1}(y_{\text{thre}}) \tag{2}$$

The RUL is then computed as:

$$\text{RUL} = t_{\text{EOF}} - t_i \tag{3}$$

3. Optimisation: the final step involves optimising the operating parameters to maximize the predicted RUL. The optimisation problem is formulated as:

$$\max_x f(x) = \text{RUL}_{\text{model}}(x) \tag{4}$$

In this step, optimisation algorithm such as Nelder-Mead are applied to adjust the operating parameters $x_i = [T_i, P_i, F_i, H_i, I_i]$ in real time, extending the fuel cell's remaining lifespan.

2.2. Case Study: PEMFC

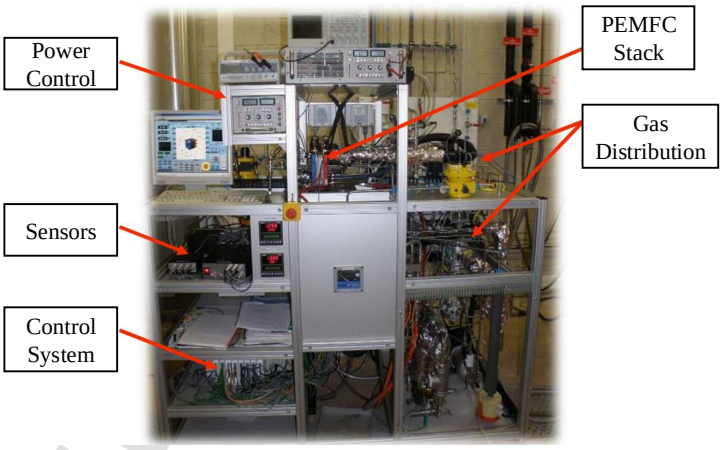

Figure 1: PEMFC test rig in FCLAB.

The PEMFC is an advanced electrochemical device that generates electricity through the electrochemical reaction of hydrogen and oxygen, cogenerating heat and power. The exhaust gas is steam mixed with unreacted

Table 1: Details of operating parameters.

Physical Meaning	**Monitoring Parameters**	**Unit**
Aging time (h)	Time	hour
Single cells voltage	U1 to U5	V
Stack voltage	Utot	V
Current	I	A
Current density	J	A/cm²
Inlet Hygrometry (Air)	HrAIRFC	%
H2		
Inlet and outlet temperature	TinH2, ToutH2	°C
Inlet and outlet Pressure	PinH2, PoutH2	mbar
Inlet and outlet Flow-rate	DinH2, DoutH2	l/mn
Air		
Inlet and Outlet temperature	TinAIR, ToutAIR	°C
Inlet and Outlet Pressure	PinAIR, PoutAIR	mbar
Inlet and Outlet Flow-rate	DinAIR, DoutAIR	l/mn
Water		
Inlet and outlet temperature	TinWat, ToutWat	°C
Flow rate of cooling water	DWAT	l/mn

hydrogen. The PEMFC test rig, developed by the FCLAB Research Federation [32], is shown in Figure 1. The fuel cell system is capable of powering up to 1 kW, with the FC stack consisting of five individual cells. During the degradation aging experiments, the operating temperature is maintained at 55 °C, and the hydrogen fuel pressure is set to approximately 1.5 bar. Each cell has an active area of 100 cm². The PEMFC utilizes commercial membranes, diffusion layers, and machined flow distribution plates. The nominal current density of the cells is 0.70 A/cm², with a maximum current density of 1 A/cm². Given the extended duration of the aging experiments, the system is reset weekly for characterization to ensure reliability, as the fuel cell system may experience operating faults over time. Various physical parameters within the stack are continuously monitored and controlled to ensure precise operating conditions. The detailed operating parameters are presented in Table 1.

3. Method

3.1. Fuel Cell Digital Twin Framework

A digital twin consists of two spaces: the physical space, which includes the real-world system and sensors for measuring operational parameters, and the virtual space, where a digital model of the system is created [33]. Data from the physical space are sent to the virtual space, where the model simulates system behavior and responds to changes in operational parameters. This enables the prediction of future states based on accumulated data, providing insights and control instructions for the physical system. Our proposed FCDT framework is shown in Figure 2 (a). The FCDT digital model is built in the virtual space and is capable of responding to changes in operational conditions. The input to this model consists of the fuel cell system's operating parameters, which are measured by physical sensors. Developing an accurate FCDT digital model is challenging due to the complexity of the fuel cell system's multiple physical parameters and the gradual degradation of the fuel cell stack over time. To address this, we developed an FCDT digital model based on the PGNN method, which significantly reduces dependence on extensive data. Collecting data for the entire degradation period is often costly and, in some cases, impractical in real-world scenarios. Our PGNN-based modelling approach enables the model to be developed using short-term or limited data, greatly enhancing the flexibility and resilience of the digital model. Once the PGNN-based model is established using the historical data, it can evaluate both the historical and current states and accurately predict future states. The optimisation algorithm will work with the established PGNN-based model to automatically adjust the operational parameters. The PGNN-based model will then respond to the predicted future degradation trends. This process allows the optimisation algorithm to identify the optimal operating parameters that will extend the fuel cell's remaining lifespan.

3.2. Physical Model for Estimating Fuel Cell Voltage

In this section, we present the physics-based model used to predict the theoretical voltage of a PEMFC under different operating conditions, based on the Amphlett model [34, 35]. This model calculates the PEMFC voltage by determining the thermodynamic potential, activation overvoltage, and ohmic overvoltage. The process and related equations are as follows:

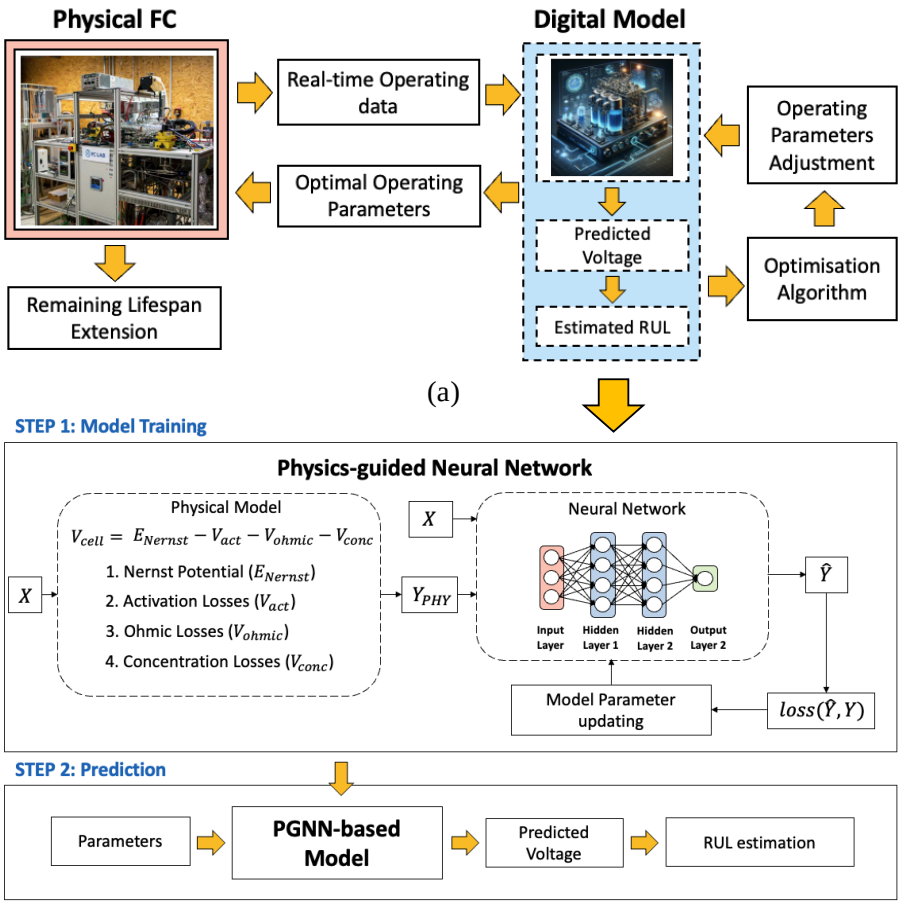

Figure 2: (a) Framework for fuel cell digital twin; (b) PGNN-based model for PEMFC RUL prediction.

1. Thermodynamic Potential (E): The thermodynamic equilibrium potential is calculated using the Nernst equation:

$$E = 1.229 - 0.35 \times 10^{-3}(T - 298.15) + 4.3085 \times 10^{-8} T \ln(p_{H_2} \cdot p_{O_2}) \quad (5)$$

where p_{H_2} and p_{O_2} are the partial pressures of hydrogen and oxygen, respectively.

2. Activation Overvoltage (η_{act}): The activation overvoltage is a function of operating temperature, operating current, and the oxygen concentration at the catalyst reaction sites, given by:

$$\eta_{act} = \alpha_1 + \alpha_2 T + \alpha_3 T \ln(c_{O_2}) + \alpha_4 T \ln(i) \quad (6)$$

where $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ are constant parametric coefficients evaluated from experimental data, T is the operating temperature, c_{O_2} is the oxygen concentration, and i is the operating current.

3. Ohmic Overvoltage (η_{ohm}): The ohmic overvoltage is related to the internal resistancei of the fuel cell and the current, expressed as:

$$\eta_{ohm} = i \cdot R_{internal} \quad (7)$$

where i is the current and $R_{internal}$ is the internal resistance of the fuel cell.

4. Fuel Cell Output Voltage (V): The output voltage of the fuel cell is the thermodynamic potential minus the activation overvoltage and ohmic overvoltage:

$$V = E - \eta_{act} - \eta_{ohm} \quad (8)$$

By combining these equations, the Amphlett model provides a comprehensive description of PEMFC voltage by capturing the effects of temperature, pressure, and current, making it a reliable simulation tool for estimating fuel cell voltage under various operating conditions. In this work, the Amphlett model serves as a physical model, $f_{PHY} : D \to Y$, using theoretical equations to predict voltage under different operating conditions, thus establishing a direct relationship between the operating parameters and the fuel cell's performance.

3.3. Fuel Cell RUL Prediction based on Physics-guided Neural Network

The generic framework of PGNN consists of two key steps: (a) constructing a hybrid combination of physics-based models and neural networks, referred to as hybrid-physics-data (HPD) models, and (b) employing a new quantile Huber loss function in the neural network's learning objective, as described below:

3.3.1. Physics-guided Neural Network Constructing

Consider a predictive learning task where we have a set of input drivers, D, that are related to a target variable Y. One typical approach is to train a neural network, $f_{\mathrm{NN}} : D \rightarrow Y$, on a training dataset, which can then be used to estimate the target variable, $\hat{Y}$. Alternatively, a physics-based numerical model, $f_{\mathrm{PHY}} : D \rightarrow Y$, can be used to simulate the target variable, Y_{PHY}, based on the underlying physical laws governing the relationship between the inputs and the output. However, unlike neural networks, physics-based models often require calibration of their parameters using observational data, which is a resource-intensive and label-expensive process. Additionally, the simulation provided by f_{PHY} may not fully capture the true behaviour of the system, especially when simplifications or approximations in the model lead to discrepancies between the simulated value Y_{PHY} and the real-world data. The goal of HPD modelling is to combine the strengths of both f_{PHY} and f_{NN} by addressing the weaknesses of each. By merging physical insights with data-driven learning, HPD models can more accurately predict the target variable, leveraging the strengths of both physical laws and empirical data while compensating for the deficiencies in each individual method.

The combination of f_{PHY} and f_{NN} is achieved by using the simulated output from the physics-based model, Y_{PHY}, as an additional input to the data-driven model (neural network), alongside the input drivers, D. This results in the following HPD model:

$$f_{\mathrm{HPD}} : X = [D, Y_{\mathrm{PHY}}] \rightarrow Y, \tag{9}$$

which is schematically illustrated in Figure 2 (b). In this framework, if the physics-based model is accurate and Y_{PHY} matches the observed values of Y perfectly, the HPD model will learn to predict $\hat{Y} = Y_{\mathrm{PHY}}$. However, if there are systematic discrepancies (biases) in Y_{PHY}, the HPD model can learn to correct these biases by extracting complex features from the input driver space.

3.3.2. Quantile Huber Loss Function

A quantile Huber loss (QH-loss) function has demonstrated superior performance within the deep learning framework compared to traditional loss functions, such as mean squared error (MSE) and mean absolute error (MAE), in multiple RUL prediction tasks, including rolling bearings [36], lithium-ion batteries [37], and fuel cells [30]. Therefore, we aim to integrate QH-loss into the PGNN framework to improve prediction accuracy.

The QH-loss L_{QH} is defined by the following equation:

$$L_{\mathrm{QH}}(X^s, Y^s) = \sum_{i=1}^{N} \mathbb{E}\left[\rho_{\hat{\tau}_i}^{\kappa}\left(F_Z^{-1}(\hat{\tau}_i) - P(\hat{\tau}_i|x^s)\right)\right], \tag{10}$$

where P represents the predictor, composed of a multi-layer neural network. The quantile distribution output at each quantile $\hat{\tau}_i$ is given by $P(\hat{\tau}_i|x)$. The cumulative distribution function (CDF) for the target distribution $F_Z(z|y^s)$ is obtained by utilizing the Gaussian distribution function centered on the true value y^s, which is mathematically expressed as:

$$F_Z(z|y^s, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{z} \exp\left(-\frac{(t-y^s)^2}{2\sigma^2}\right) dt. \tag{11}$$

3.3.3. PGNN-based Digital Modelling Method

A FCDT digital model can be developed by deploying a PGNN framework. As illustrated in the model training process shown in Figure 2 (b), the physical model generates simulation results, Y_{PHY}, based on the input parameters, X. These inputs, X and Y_{PHY}, are then used to train the neural network to predict the outcome voltage, $\hat{Y}$. Once the training process is complete, the PGNN-based model can estimate the voltage at any given time, thereby enabling an accurate estimation of the RUL value. In this work, we selected the PEMFC as a case study to demonstrate the feasibility and advancements of our proposed method for fuel cell degradation performance prediction.

Regarding the PGNN-based model, the input parameters X include: Time, I, J, T_{inH2}, T_{outH2}, T_{inAIR}, T_{outAIR}, P_{inH2}, P_{outH2}, P_{inAIR}, P_{outAIR}, D_{inH2}, D_{outH2}, D_{inAIR}, D_{outAIR}, D_{WAT}, and HrAIRFC. The output Y of this predictive model is the stack voltage, U_{tot}. The physical model's output Y_{phy} represents the simulated stack voltage, reflecting the designed or expected PEMFC voltage. However, the voltage of the fuel cell degrades over time, and the degradation mechanism is complex, making it challenging for chemical process experts to accurately incorporate it into the physical model. This is where data-driven approaches become valuable. After obtaining the simulated voltage Y_{PHY}, it is combined with the same input parameters and fed into the neural network to predict the actual voltage. The novelty and advantage of our proposed method lie in its ability to train the model effectively under limited data conditions, leveraging the strengths of the PGNN

framework. As a result, the model can predict voltage over any given time span, including both future and historical periods. This flexibility enables the optimisation algorithm to interact with the PGNN-based model effectively, identifying the optimal operating parameters.

In this study, "limited data conditions" refer to training scenarios where only partial degradation trajectories are available, specifically short-term data segments of 100, 200, or 300 hours, representing 20–50% of the full lifetime data. The PGNN-based model is trained using these partial datasets and evaluated on the remaining unseen degradation period, simulating early-life or mid-life deployment scenarios common in real-world fuel cell applications.

3.4. Fuel Cell RUL Optimisation

As for the second key feature of the FCDT framework, we present the fuel cell RUL optimisation process in this section, which extends the RUL of the fuel cell by adjusting the real-time operating parameters. To implement this process, we employ the Nelder-Mead optimisation algorithm [38], which is particularly well-suited for minimizing non-linear objective functions without requiring gradient information. This makes it ideal for optimizing the fuel cell's RUL, where complex, non-linear interactions exist between the fuel cell's operational parameters and its degradation behaviour.

The objective of the optimisation in this work is not to guarantee a global optimum over the entire operating domain, but to identify locally optimal operating adjustments around the current operating point in real time. Nelder–Mead is therefore suitable due to its derivative-free nature, low computational cost, and robustness for local optimisation in non-smooth, model-based objective functions.

The objective function $f(x)$, where x represents operational parameters such as temperature, pressure, flow rate, humidity, and current, is formulated as:

$$\max f(x) = \mathrm{RUL}_{\mathrm{model}}(x) \tag{12}$$

where $\mathrm{RUL}_{\mathrm{model}}(x)$ is derived from PGNN-based predictive model. To ensure feasible solutions, the optimisation includes operational constraints in the form of bounds on the variables:

$$\text{Bounds: } x_{\min} \leq x \leq x_{\max} \tag{13}$$

These bounds represent the safe operating limits, such as allowable temperature and pressure for the fuel cells.

The Nelder-Mead optimisation algorithm follows an iterative workflow that adjusts the operational parameters to extend the fuel cell's remaining lifespan. The iterations are as follows:

1. **Initialization**: Start with an initial simplex in n-dimensional space, which is a set of $n + 1$ vertices that serve as initial guesses for the solution.
2. **Sorting**: At each iteration, sort the vertices of the simplex based on their objective function values such that $f(x_1) \leq f(x_2) \leq \ldots \leq f(x_{n+1})$, where x_1 is the best vertex and x_{n+1} is the worst vertex.
3. **Reflection**: Compute the reflection point x_r using the formula:
$$x_r = \bar{x} + \alpha(\bar{x} - x_{n+1}) \tag{14}$$
where $\bar{x}$ is the centroid of the best n vertices. If $f(x_r)$ is better than $f(x_{n+1})$, replace x_{n+1} with x_r.
4. **Expansion**: If $f(x_r)$ is better than $f(x_1)$, compute the expansion point x_e using the formula:
$$x_e = \bar{x} + \beta(x_r - \bar{x}) \tag{15}$$
If $f(x_e) < f(x_r)$, replace x_{n+1} with x_e; otherwise, replace x_{n+1} with x_r.
5. **Outside Contraction**: If $f(x_r)$ is not better than $f(x_1)$ but better than $f(x_{n+1})$, compute the outside contraction point x_{oc} using the formula:
$$x_{oc} = \bar{x} + \gamma(x_r - \bar{x}) \tag{16}$$
If $f(x_{oc}) \leq f(x_r)$, replace x_{n+1} with x_{oc}; otherwise, proceed to the next step.
6. **Inside Contraction**: If $f(x_r)$ is not better than $f(x_{n+1})$, compute the inside contraction point x_{ic} using the formula:
$$x_{ic} = \bar{x} - \gamma(\bar{x} - x_{n+1}) \tag{17}$$
If $f(x_{ic}) < f(x_{n+1})$, replace x_{n+1} with x_{ic}; otherwise, proceed to the next step.
7. **Shrink**: If none of the above steps replace the worst vertex, shrink the simplex towards the best vertex x_1 using the formula:
$$x_i = x_1 + \delta(x_i - x_1) \tag{18}$$
for $2 \leq i \leq n + 1$.

8. **Iteration**: Repeat the above steps until a stopping criterion is met, such as a maximum number of iterations, a threshold for the change in the objective function value, or a threshold for the change in vertex positions.

The Nelder-Mead algorithm involves several key parameters: α (reflection coefficient), β (expansion coefficient), γ (outside contraction coefficient), and δ (inside contraction coefficient). These parameters control the deformation of the simplex during the optimisation process. The standard settings for these parameters are $\alpha = 1$, $\beta = 2$, $\gamma = 0.5$, and $\delta = 0.5$.

Since the PGNN-based predictive model is nonlinear, the resulting optimisation landscape is generally nonconvex. Consequently, the optimisation is expected to converge to a local optimum. However, for real-time fuel cell operation, local improvements in RUL are practically meaningful and safer than aggressive global exploration, which may violate operational constraints.

4. Results and Discussions

4.1. Physical Model and Degradation Data Preparation

In this study, we utilized the PEMFC physical model and degradation data to validate our proposed FCDT framework. The Amphlett model [34] was selected as the theoretical model for the PEMFC due to its reliability in predicting fuel cell performance. The OPEM (Open Source PEM Cell Simulation Tool) [35] was used as the physical model, $f_{\text{PHY}} : D \rightarrow Y$, which simulates the fuel cell's behavior under various operational conditions. The specific fuel cell parameters used in the physical model are detailed in Table 2.

The degradation data is collected from a PEMFC test rig by the FC lab research federation [32] at constant operation condition. The test rig incorporates five fuel cell stacks, each with a $100cm^2$ active area and a nominal current density of $0.7A/cm^2$. Figure 3 shows the voltage and current profiles of a PEMFC during degradation. The voltage steadily decreases, while current fluctuates, with notable increases around time step 1000 hours. Figure 4 presents histograms of various operating parameters for the PEMFC, including current, voltage, temperature, pressure, and humidity readings. The distributions highlight the variations in each parameter, offering insight into the operational behaviour and variability of the fuel cell system.

Table 2: Amphlett model input parameters

Input	Description	Unit
T	Cell operation temperature	K
PH2	Partial pressure	atm
PO2	Partial pressure	atm
i	Cell operating current	A
A	Active area	cm^2
I	Membrane thickness	cm
JMax	Maximum current density	$A/(cm^2)$
N	Number of single cells	–

Based on the PEMFC system settings, we have determined the input parameters for the physical model as $A = 100$, $I = 0.001$, $J_{\text{Max}} = 1$, and $N = 5$. The cell operating temperature T is set equal to the outlet water temperature, T_{outWAT}, while the partial pressure $PH2$ is equal to the hydrogen inlet pressure $PinH2$, and the partial pressure $PO2$ is equal to the oxygen inlet pressure $PinO2$.

4.2. Implementation Details and Evaluation indicators

Our proposed PGNN-based model consists of four fully connected layers, each with 128 hidden nodes, and uses the ReLU activation function. The default hyperparameter values are outlined in Table 3. All experiments were conducted on a system equipped with an Nvidia GeForce GTX 4060 GPU, an Intel Core i9-13900HX CPU running at 2.20 GHz, and 32 GB of RAM.

Table 3: Default hyperparameters for PGNN

Hyperparameters	Value
Learning rate	0.0003
Batch size	256
Epochs	100
QH-loss factor N_τ	1
QH-loss factor κ	1.0
QH-loss factor σ	0.2

This paper quantitatively evaluates the prediction performance of the

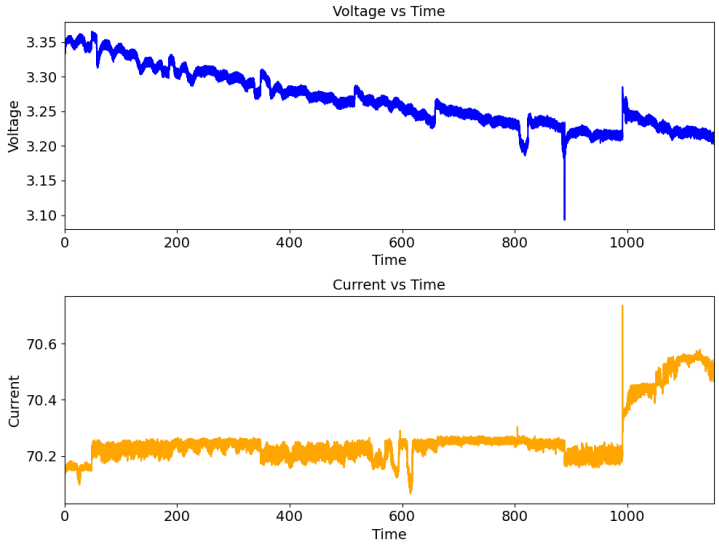

Figure 3: Voltage and Current for the PEMFC degradation data

proposed method using three metrics: R^2 score, mean square error (MSE), and root mean square error (RMSE), defined as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{Q}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{Q}(\bar{y} - y_i)^2} \tag{19}$$

$$MSE = \frac{1}{Q}\sum_{i=1}^{Q}(\hat{y}_i - y_i)^2 \tag{20}$$

$$RMSE = \sqrt{\frac{1}{Q}\sum_{i=1}^{Q}(\hat{y}_i - y_i)^2} \tag{21}$$

where Q is the number of test samples, $\hat{y}_i$ is the predicted voltage, y_i is the true voltage, and $\bar{y}$ is the mean of the true voltage values.

Following [39], we also assess RUL prediction using the error Er_t:

$$Er_t = \frac{r_t - \hat{r}_t}{r_t} \times 100 \tag{22}$$

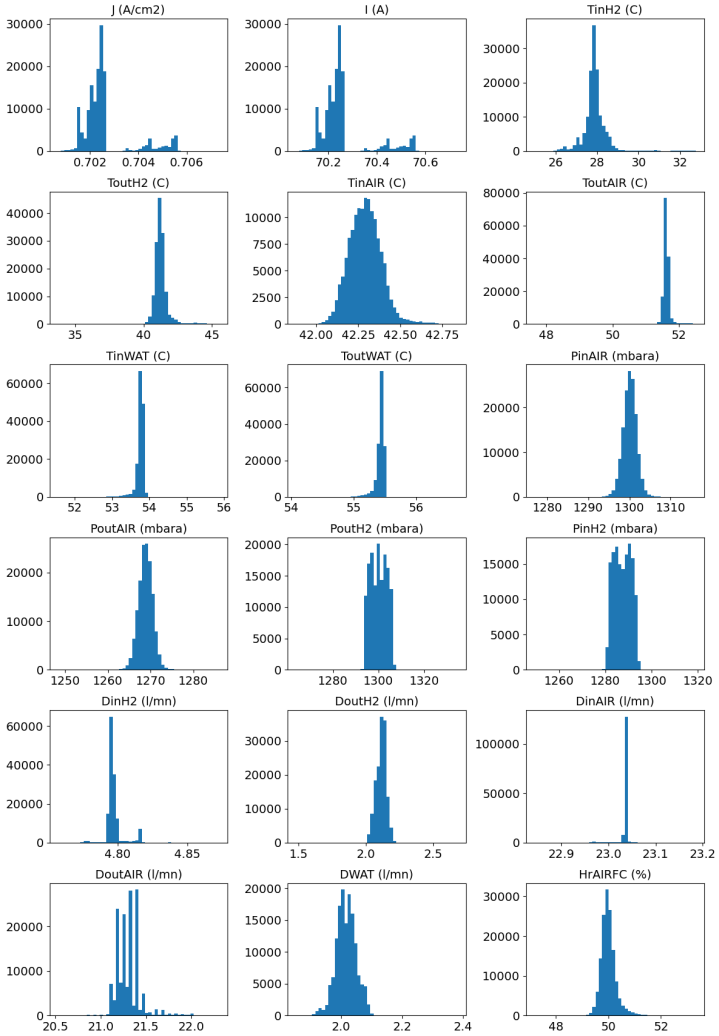

Figure 4: Histograms of Operating Parameters for the PEMFC

where r_t and $\hat{r}_t$ are the actual and predicted RUL values, respectively.

4.3. Results and Analysis for RUL Prediction

Figure 5 presents the results from the training and testing analysis of the PGNN predictive model applied to the PEMFC degradation data. In this experimental setup, the PGNN-based model was trained using data from 0 to 550 hours and then tested for accuracy using data from 550 hours to the end of the test. The result illustrates the voltage behavior over time for both the training and test datasets, alongside the model's predictions. The PGNN-based model demonstrates strong performance in capturing voltage degradation patterns over whole lifetime period, as shown by the close alignment between the predicted and actual test data. The model accurately tracks the gradual decline in voltage, with only minor deviations observed during the later stages of testing. The increased deviation observed after 500 hours is mainly attributed to accelerated degradation effects and occasional operational perturbations, which are underrepresented in early-stage training data. Despite this, the PGNN-based model successfully captures the overall degradation trend and maintains acceptable prediction accuracy.

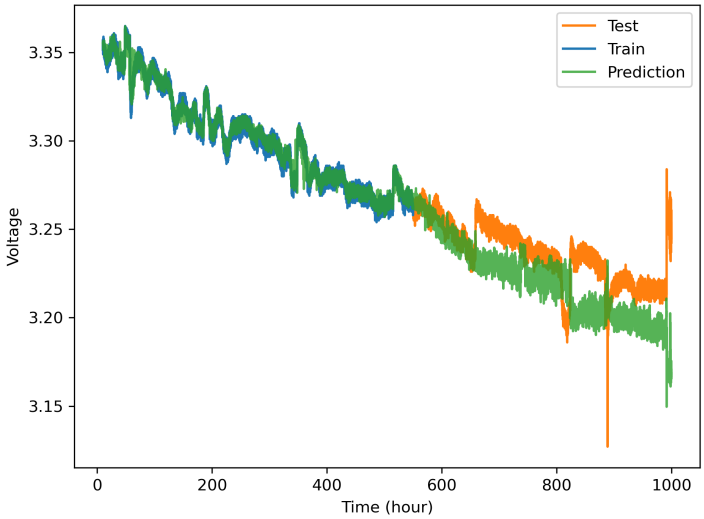

Figure 5: Voltage prediction result.

The detailed quantitative analysis results for the PGNN-based model voltage prediction are shown in Figure 6 and Table 4. The evaluation indicators include R^2, MSE, and RMSE, with the equations provided in (19), (20), and (21). Figure 6 presents these metrics for varying start times, with each bar representing the performance for different starting hours (100, 200, 300, etc.). This means that the PGNN-based model is trained on data from 0 to the starting hour and then tested on data from the starting hour to the end of its lifetime. The R^2 values remain consistently high, indicating that the model effectively captures both the degradation and its variance in the test data. Although the MSE and RMSE values fluctuate slightly, they generally remain low, with a minor peak observed at 200 hours, suggesting that the model performs optimally at later stages. These results demonstrate that the PGNN-based model effectively adapts to the data over time and can provide reliable long-term predictions with different training data periods. The performance fluctuation at 200 hours does not indicate systematic overfitting but reflects transient data variability and sensitivity to the selected training window. The stable performance across longer training windows confirms the robustness of the model.

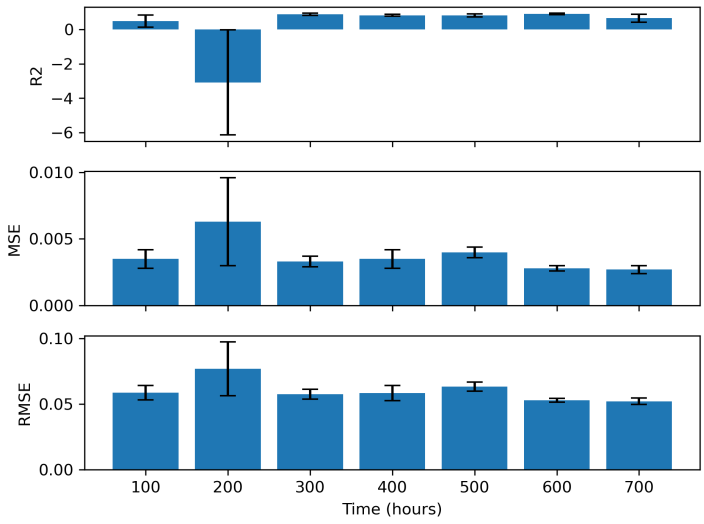

Figure 6: Comparison for different start times.

Table 4: Prediction performance under different start times

Time (h)	R^2	MSE	RMSE
100	0.4814 ± 0.3555	0.0035 ± 0.0007	0.0587 ± 0.0055
200	-3.20 ± 2.80	0.0064 ± 0.0032	0.0790 ± 0.0190
300	0.8815 ± 0.0631	0.0033 ± 0.0004	0.0577 ± 0.0038
400	0.9000 ± 0.0400	0.0034 ± 0.0003	0.0580 ± 0.0030
500	0.9330 ± 0.0149	0.0033 ± 0.0002	0.0571 ± 0.0018
600	0.9500 ± 0.0120	0.0029 ± 0.0002	0.0530 ± 0.0015
700	0.8900 ± 0.0450	0.0035 ± 0.0004	0.0590 ± 0.0032

Table 5 presents the RUL prediction results for the fuel cell test at 550 hours, using the trained PGNN-based model. The table compares the actual RUL with the estimated RUL at various failure thresholds, expressed as percentages of the initial power (P_{init}), along with the error (Er) between the two values. Notably, at the 3.5% threshold of P_{init}, the estimated RUL is significantly higher than the actual RUL, with a large error of -90.06. At the 4.0% threshold, the estimated RUL is much closer to the actual RUL, with a smaller error -6.34, indicating a more accurate prediction. As the threshold increases, the model tends to slightly overestimate the RUL, particularly at the 4.5% and 5.0% thresholds, where the errors are -55.68 and -30.56, respectively. At the highest threshold of 5.5%, both the actual and estimated RUL values are infinite, indicating that the fuel cell did not fail within the observed period. These results highlight that the choice of failure threshold plays a critical role in determining the estimated RUL. The discrepancies observed at lower failure thresholds highlight the sensitivity of RUL estimation to threshold selection, which is a well-known challenge in fuel cell prognostics. The results demonstrate that the proposed model provides reliable RUL estimates within practically relevant threshold ranges (e.g., 4.0–5.0% of Pinit).

While Table 5 presents the quantitative results of the RUL prediction, Figure 7 provides a detailed visualization of the determination process. In this figure, the voltage profiles exhibit a gradual decline, with the predicted voltage generally following the trend of the true voltage. The red dashed lines represent various threshold levels, indicating points where fuel cell failure might be imminent. The vertical dashed line marks the current time at 550 hours, indicating the point of prediction. The close alignment between the predicted and true voltage curves demonstrates that the PGNN-based

Table 5: RUL prediction results.

Failure threshold	Act. RUL	Est. RUL	Er
3.5% of P_{init}	82.20	156.24	-90.06
4.0% of P_{init}	257.78	274.14	-6.34
4.5% of P_{init}	259.52	404.03	-55.68
5.0% of P_{init}	338.40	442.50	-30.56
5.5% of P_{init}	inf	inf	*

model effectively captures the degradation trends, providing reliable RUL predictions.

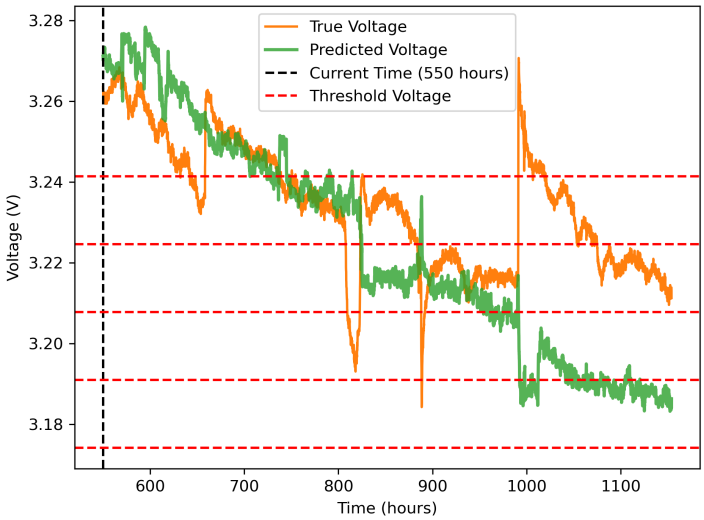

Figure 7: Visualization of RUL Prediction Results at 550 Hours.

4.4. Comprehensive Comparison with Varying Training Dataset Volume

In this section, we present a comprehensive comparison between the proposed PGNN-based model and a set of representative machine learning and deep learning baselines under both sufficient and limited data conditions.

The comparison includes classical machine learning models—Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), K-Nearest Neighbors (KNN), Linear Regression (LR), and a fully connected Neural Network (NN)—as well as recent state-of-the-art deep prognostics models, namely the Non-stationary Transformer (NST) [40], the Temporal Fusion Transformer (TFT) [41], and a hybrid Temporal Convolutional Network with Bidirectional LSTM (TCN-BiLSTM) [42]. These deep learning baselines represent three complementary paradigms for degradation modelling. Specifically, NST is designed to address non-stationary time-series behaviour commonly observed in long-term degradation processes, TFT is a multivariate forecasting architecture capable of modelling complex temporal dependencies and heterogeneous operational inputs, and TCN-BiLSTM combines temporal convolutional feature extraction with recurrent sequence modelling, which has been widely adopted in industrial prognostics applications. All baseline models are trained using the same input variables, data partitioning strategies, and evaluation metrics to ensure a fair and consistent comparison. For all compared models, training and testing are conducted under identical data partitioning strategies, using training data lengths of 100, 300, and 500 hours to evaluate robustness under limited and sufficient data conditions.

Figure 8 illustrates the performance comparison of the proposed PGNN-based model and the benchmark methods under different training dataset sizes. A training dataset of 500 hours is considered the sufficient data condition, under which most methods achieve satisfactory prediction performance, with the exception of SVM. Among all methods, the PGNN-based model consistently achieves the highest prediction accuracy. When the training dataset volume is reduced to 300 hours (approximately 60% of the sufficient dataset), only the LR, NN, and PGNN models maintain positive R^2 values, while the remaining methods exhibit degraded performance, indicating limited generalisation capability. When the training dataset volume is further reduced to 100 hours (20% of the sufficient dataset), representing a severe limited-data condition, the PGNN-based model is the only approach that maintains reliable prediction performance.

The PGNN-based model demonstrates consistently strong performance across all dataset sizes, highlighting its robustness and capability to deliver accurate predictions even under highly limited data conditions. The results further indicate that while modern deep learning models such as NST, TFT, and TCN-BiLSTM achieve competitive prediction accuracy under sufficient data conditions, their performance degrades noticeably as the available train-

ing data decrease. In contrast, the PGNN-based model consistently maintains superior and more stable performance across all data regimes. This behaviour highlights the advantage of incorporating physics-guided information, which significantly enhances data efficiency and robustness in fuel cell degradation prognostics.

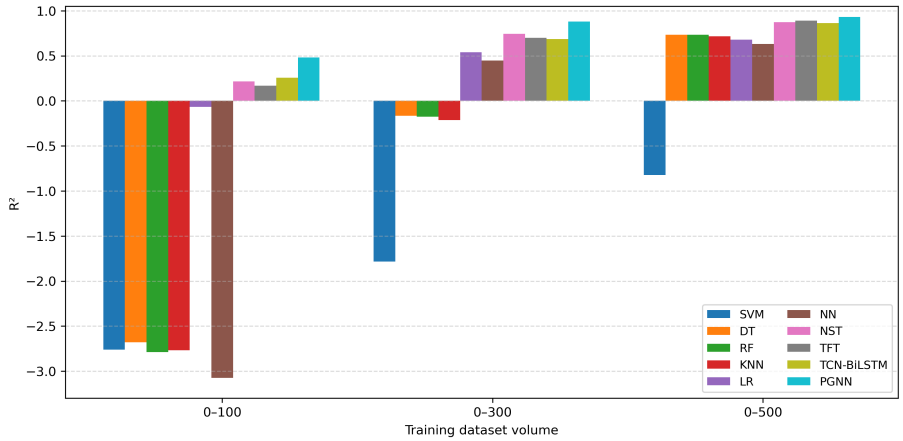

Figure 8: Comparison results with different machine learning methods under different training dataset volume

To further evaluate the performance of our proposed PGNN-based model under limited training data conditions, we tested the model in more complex scenarios. Table 6 provides a detailed analysis of the quantitative evaluation results associated with varying amounts of training data used in the PGNN-based model for predicting fuel cell voltage. The table is organized by time intervals (100, 200, and 300 hours), showing corresponding metrics such as R^2, MSE, and RMSE. The results indicate that as the amount of training data increases, the model's performance improves. At the 100-hour mark, the R^2 values are relatively low, and both MAE and RMSE values are higher, indicating less accurate predictions. However, as more data from the 200- and 300-hour intervals is included, the R^2 values approach 1, while MAE and RMSE decrease, reflecting improved prediction accuracy. This trend suggests that the model benefits significantly from longer training data, enabling it to better capture the underlying degradation patterns of the fuel cell. However, our proposed PGNN-based model is capable of performing well under limited

data conditions, and it can be effectively trained using degradation data from any period, which significantly validates the high flexibility of our method.

4.5. Results and Analysis for RUL Optimisation

The second key step of the FCDT framework is RUL optimisation based on real-time operating parameters by Nelder-Mead optimisation algorithm. Table 7 presents the RUL extension results at 550 hours for various failure thresholds, expressed as percentages of the initial voltage (P_{init}) drop. The actual RUL (Act. RUL) increases as the failure threshold decreases, with the optimized RUL (Opt. RUL) showing significant extension beyond the actual values. At the 3.5% threshold of P_{init}, the estimated RUL is extended by approximately 271.85 hours. As the failure threshold increases to 4.0%, the RUL extension reaches 262.60 hours. At higher thresholds of 4.5% and 5.0%, the RUL is extended by 330.99 and 264.69 hours, respectively. When the threshold reaches 5.5%, the RUL becomes infinite, indicating that no failure occurs within the observed period. These results demonstrate that our proposed approach can significantly extend the fuel cell's lifetime across all failure thresholds. The optimized operating parameters are verified by re-evaluating the PGNN-based model under the optimized conditions and comparing the resulting RUL with the baseline scenario. The consistent extension of RUL across multiple failure thresholds (Table 7) provides indirect validation of the effectiveness of the optimized parameter set.

From a computational perspective, the proposed FCDT framework is well suited for real-time or near–real-time deployment. Once trained, the PGNN-based predictive model requires only lightweight feedforward inference combined with the evaluation of the physics-based model, resulting in low computational overhead. The Nelder–Mead optimisation algorithm is implemented as a local, derivative-free optimiser with a small number of decision variables, which enables fast convergence without the need for gradient computation. These characteristics make the proposed framework suitable for online implementation within digital twin architectures for fuel cell systems.

In practical deployment scenarios, factors such as sensor noise, communication latency, and computational resource constraints must be considered. The proposed framework is compatible with standard industrial computing platforms and can be integrated with existing monitoring systems for periodic or event-triggered optimisation. While detailed hardware-specific

Table 6: Quantitative evaluation results associated with varying training dataset volumes for the proposed PGNN-based model.

		10-100	100-200	200-300	300-400
100 hours	R^2	0.4814 ± 0.3555	-0.8599 ± 0.5284	0.7855 ± 0.1095	0.5775 ± 0.2498
	MSE	0.0035 ± 0.0007	0.0040 ± 0.0005	0.0031 ± 0.0006	0.0047 ± 0.0012
	$RMSE$	0.0587 ± 0.0055	0.0627 ± 0.0041	0.0556 ± 0.0050	0.0680 ± 0.0083
		10-200	100-300	200-400	300-500
200 hours	R^2	-3.0830 ± 3.0640	0.7724 ± 0.1176	0.4438 ± 0.3746	0.7910 ± 0.1116
	MSE	0.0063 ± 0.0033	0.0040 ± 0.0007	0.0044 ± 0.0015	0.0042 ± 0.0005
	$RMSE$	0.0769 ± 0.0205	0.0628 ± 0.0060	0.0650 ± 0.0120	0.0648 ± 0.0038
		10-300	100-400	200-500	300-600
300 hours	R^2	0.8815 ± 0.0631	0.5799 ± 0.3450	0.5864 ± 0.1970	0.8019 ± 0.0417
	MSE	0.0033 ± 0.0004	0.0049 ± 0.0011	0.0047 ± 0.0007	0.0039 ± 0.0004
	$RMSE$	0.0577 ± 0.0038	0.0696 ± 0.0080	0.0680 ± 0.0052	0.0621 ± 0.0031
		400-500	500-600	600-700	*
100 hours	R^2	0.7596 ± 0.1464	0.6643 ± 0.1761	0.1101 ± 0.2328	*
	MSE	0.0030 ± 0.0006	0.0023 ± 0.0001	0.0021 ± 0.0002	*
	$RMSE$	0.0547 ± 0.0055	0.0479 ± 0.0013	0.0455 ± 0.0021	*
		400-600	500-700	*	*
200 hours	R^2	0.7106 ± 0.1345	0.1976 ± 0.3224	*	*
	MSE	0.0023 ± 0.0002	0.0019 ± 0.0001	*	*
	$RMSE$	0.0474 ± 0.0025	0.0438 ± 0.0014	*	*
		400-700	*	*	*
300 hours	R^2	0.8481 ± 0.0490	*	*	*
	MSE	0.0025 ± 0.0002	*	*	*
	$RMSE$	0.0498 ± 0.0020	*	*	*

Table 7: RUL extension results at 550h.

Failure threshold	Act. RUL	Opt. RUL	Ext. Life
3.5% of P_{init}	82.20	354.05	271.85
4.0% of P_{init}	257.78	521.38	262.60
4.5% of P_{init}	269.52	600.51	330.99
5.0% of P_{init}	338.40	603.09	264.69
5.5% of P_{init}	inf	*	*

runtime benchmarking is beyond the scope of this study, it represents an important direction for future system-level implementation and validation.

5. Conclusion and Future Directions

In this paper, we propose a novel FCDT framework enabled by the PGNN-based model and the Nelder-Mead optimisation algorithm. A new PGNN-based modeling method has been developed and tailored to predict the voltage and estimate the RUL by combining the Amphlett model with degradation data. The PGNN-based model then works with the Nelder-Mead optimisation algorithm to optimize the real-time operating parameters collected from the physical PEMFC system, thereby extending the RUL of the PEMFC. The primary conclusions are summarized as follows:

1. The experimental results validate that the proposed PGNN-based model can accurately predict both the voltage and the RUL of the PEMFC. The comparative experiments further demonstrate that the model provides reliable and accurate predictions throughout the entire degradation lifetime of the fuel cell. These findings underscore the model's ability to effectively capture the degradation patterns and predict performance over time, making it a valuable tool for long-term fuel cell prediction.
2. The proposed PGNN-based model has been validated under limited data conditions. Short-term data (from 100, 200, and 300 hours) collected from various points during the degradation period were used to train the PGNN-based model. The comparative results confirm that the model is capable of providing reliable predictions, even when

trained with limited data. This demonstrates the advantage of integrating the PGNN with a physical model, which reduces the model's dependence on extensive experimental data.

3. The experimental results of the RUL optimisation demonstrate that the Nelder-Mead optimisation algorithm effectively works with the PGNN-based model to optimize real-time operating parameters, thus extending the RUL at different failure thresholds. These results indicate that our proposed approach significantly enhances the fuel cell's longevity by identifying optimal operational conditions.

5.1. Future Directions

While the proposed FCDT framework demonstrates strong performance for fuel cell RUL prediction and optimisation under limited data conditions, several promising directions remain for future investigation.

1. Recent advances in industrial soft sensing, including adaptive graph-based learning and physics-aware neural network architectures, offer opportunities to further enhance feature representation while maintaining physical consistency. Integrating such adaptive learning mechanisms with the PGNN framework may improve scalability and generalisation in large-scale industrial systems.
2. In addition, fuel cell degradation is inherently a long-term temporal process. Future work may explore the integration of advanced time-series modelling techniques to better capture long-term degradation dynamics and associated uncertainty, particularly under highly variable operating conditions.
3. Furthermore, while the present study focuses on predictive optimisation performance, incorporating causal inference techniques could provide deeper insight into the underlying mechanisms driving fuel cell degradation. Such approaches may help identify the most influential operating parameters and further strengthen the interpretability and robustness of optimisation strategies in digital twin applications.

Acknowledgment

This work was supported by Innovate UK project DIATOMIC (Digital InnovAtion TransfOrMatIve Change) with grant number 10055175. The authors would like to appreciate the financial support provided through EPSRC IAA 2022-23 Impact Builder Award.

Data Availability Statement

The data supporting the findings of this study are openly available in [IEEE PHM Data Challenge 2014] at [10.25666/DATAUBFC-2021-07-19]. All relevant datasets have been uploaded and can be accessed for reproducibility purposes.

References

[1] K. Kendall, S. Ye, Z. Liu, The hydrogen fuel cell battery: Replacing the combustion engine in heavy vehicles, Engineering 21 (2023) 39–41.

[2] S. Ong, A. Al-Othman, M. Tawalbeh, Emerging technologies in prognostics for fuel cells including direct hydrocarbon fuel cells, Energy 277 (2023) 127721.

[3] L. Vichard, N. Y. Steiner, N. Zerhouni, D. Hissel, Hybrid fuel cell system degradation modeling methods: A comprehensive review, Journal of Power Sources 506 (2021) 230071.

[4] M. Yue, S. Jemei, N. Zerhouni, R. Gouriveau, Proton exchange membrane fuel cell system prognostics and decision-making: Current status and perspectives, Renewable Energy 179 (2021) 2277–2294.

[5] J. Liu, Y. Yin, J. Zhang, T. Zhang, X. Zhang, H. Chen, Mechanical degradation of catalyst layer under accelerated relative humidity cycling in a polymer electrolyte membrane fuel cell, Journal of Power Sources 512 (2021) 230487.

[6] M. Moein-Jahromi, M. Kermani, S. Movahed, Degradation forecast for pemfc cathode-catalysts under cyclic loads, Journal of Power Sources 359 (2017) 611–625.

[7] M. Moein-Jahromi, M. Kermani, Three-dimensional multiphase simulation and multi-objective optimization of pem fuel cells degradation under automotive cyclic loads, Energy Conversion and Management 231 (2021) 113837.

[8] W. Zheng, L. Xu, Z. Hu, Y. Zhao, J. Li, M. Ouyang, Dynamic modeling of pt degradation and mitigation strategies in polymer electrolyte membrane fuel cells, eTransportation 12 (2022) 100171.

[9] D. Zhou, Y. Wu, F. Gao, E. Breaz, A. Ravey, A. Miraoui, Degradation prediction of pem fuel cell stack based on multiphysical aging model with particle filter approach, IEEE Transactions on Industry Applications 53 (4) (2017) 4041–4052.

[10] D. Zhang, P. Baraldi, C. Cadet, N. Yousfi-Steiner, C. Bérenguer, E. Zio, An ensemble of models for integrating dependent sources of information for the prognosis of the remaining useful life of proton exchange membrane fuel cells, Mechanical Systems and Signal Processing 124 (2019) 479–501.

[11] M. Ou, R. Zhang, Z. Shao, B. Li, D. Yang, P. Ming, C. Zhang, A novel approach based on semi-empirical model for degradation prediction of fuel cells, Journal of Power Sources 488 (2021) 229435.

[12] H. Chen, Z. Zhan, P. Jiang, Y. Sun, L. Liao, X. Wan, Q. Du, X. Chen, H. Song, R. Zhu, et al., Whole life cycle performance degradation test and rul prediction research of fuel cell mea, Applied Energy 310 (2022) 118556.

[13] F.-K. Wang, T. Mamo, X.-B. Cheng, Bi-directional long short-term memory recurrent neural network with attention for stack voltage degradation from proton exchange membrane fuel cells, Journal of Power Sources 461 (2020) 228170.

[14] R. Mezzi, N. Yousfi-Steiner, M. C. Péra, D. Hissel, L. Larger, An echo state network for fuel cell lifetime prediction under a dynamic micro-cogeneration load profile, Applied Energy 283 (2021) 116297.

[15] R. Zhang, T. Chen, F. Xiao, J. Luo, Bi-directional gated recurrent unit recurrent neural networks for failure prognosis of proton exchange membrane fuel cells, International Journal of Hydrogen Energy 47 (77) (2022) 33027–33038.

[16] K. He, Z. Liu, Y. Sun, L. Mao, S. Lu, Degradation prediction of proton exchange membrane fuel cell using auto-encoder based health indicator and long short-term memory network, International Journal of Hydrogen Energy 47 (82) (2022) 35055–35067.

[17] T. Wilberforce, A. Alaswad, A. Garcia-Perez, Y. Xu, X. Ma, C. Panchev, Remaining useful life prediction for proton exchange membrane fuel cells

using combined convolutional neural network and recurrent neural network, International Journal of Hydrogen Energy 48 (1) (2023) 291–303.

[18] D. Chen, W. Wu, K. Chang, Y. Li, P. Pei, X. Xu, Performance degradation prediction method of pem fuel cells using bidirectional long short-term memory neural network based on bayesian optimization, Energy 285 (2023) 129469.

[19] J. Yang, L. Wang, B. Zhang, H. Zhang, X. Wu, X. Xu, P. Deng, Y. Peng, Remaining useful life prediction of vehicle-oriented pemfc systems based on igwo-bp neural network under real-world traffic conditions, Energy 291 (2024) 130334.

[20] K. Benaggoune, M. Yue, S. Jemei, N. Zerhouni, A data-driven method for multi-step-ahead prediction and long-term prognostics of proton exchange membrane fuel cell, Applied Energy 313 (2022) 118835.

[21] C. Wang, Z. Li, R. Outbib, M. Dou, D. Zhao, A novel long short-term memory networks-based data-driven prognostic strategy for proton exchange membrane fuel cells, International Journal of Hydrogen Energy 47 (18) (2022) 10395–10408.

[22] M. J. Izadi, P. Hassani, M. Raeesi, P. Ahmadi, A novel wavenet-gru deep learning model for pem fuel cells degradation prediction based on transfer learning, Energy 293 (2024) 130602.

[23] A. Daw, A. Karpatne, W. D. Watkins, J. S. Read, V. Kumar, Physics-guided neural networks (pgnn): An application in lake temperature modeling, in: Knowledge Guided Machine Learning, Chapman and Hall/CRC, 2022, pp. 353–372.

[24] J. Wang, Y. Li, R. Zhao, R. X. Gao, Physics guided neural network for machining tool wear prediction, Journal of Manufacturing Systems 57 (2020) 298–310.

[25] X. Hu, H. Hu, S. Verma, Z.-L. Zhang, Physics-guided deep neural networks for power flow analysis, IEEE Transactions on Power Systems 36 (3) (2020) 2082–2092.

[26] L. Zhang, H. Zhou, Z. Ren, Physics-guided fuel-switching neural networks for stable combustion of low calorific industrial gas, Energy (2024) 131971.

[27] J. Ye, Q. Xie, M. Lin, J. Wu, A method for estimating the state of health of lithium-ion batteries based on physics-informed neural network, Energy 294 (2024) 130828.

[28] S. Meraghni, L. S. Terrissa, M. Yue, J. Ma, S. Jemei, N. Zerhouni, A data-driven digital-twin prognostics method for proton exchange membrane fuel cell remaining useful life prediction, International journal of hydrogen energy 46 (2) (2021) 2555–2564.

[29] M. Yue, K. Benaggoune, J. Meng, D. Diallo, Implementation of an early stage fuel cell degradation prediction digital twin based on transfer learning, IEEE Transactions on Transportation Electrification 9 (2) (2022) 3308–3318.

[30] M. Zhang, A. Amiri, Y. Xu, L. Bastin, T. Clark, Self-adaptive digital twin of fuel cell for remaining useful lifetime prediction, International Journal of Hydrogen Energy 89 (2024) 634–647.

[31] D. Monopoli, C. Semeraro, M. A. Abdelkareem, A. H. Alami, A. G. Olabi, M. Dassisti, How to build a digital twin for operating pem-electrolyser system–a reference approach, Annual Reviews in Control 57 (2024) 100943.

[32] F. Harel, IEEE PHM Data Challenge 2014, FCLAB (2021). doi:10.25666/DATAUBFC-2021-07-19.

[33] Q. Qi, F. Tao, Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison, Ieee Access 6 (2018) 3585–3593.

[34] J. C. Amphlett, R. Baumert, R. F. Mann, B. A. Peppley, P. R. Roberge, T. J. Harris, Performance modeling of the ballard mark iv solid polymer electrolyte fuel cell: Ii. empirical model development, Journal of the Electrochemical Society 142 (1) (1995) 9.

[35] S. Haghighi, K. Askari, S. Hamidi, M. M. Rahimi, OPEM : Open source PEM cell simulation tool, Journal of Open Source Software 3 (27) (2018)

676. doi:10.21105/joss.00676.
URL https://doi.org/10.21105/joss.00676

[36] M. Zhang, D. Wang, N. Amaitik, Y. Xu, A distributional perspective on remaining useful life prediction with deep learning and quantile regression, IEEE Open Journal of Instrumentation and Measurement 1 (2022) 1–13.

[37] Y. Zhang, M. Zhang, C. Liu, Z. Feng, Y. Xu, Reliability enhancement of state of health assessment model of lithium-ion battery considering the uncertainty with quantile distribution of deep features, Reliability Engineering & System Safety 245 (2024) 110002.

[38] F. Gao, L. Han, Implementing the nelder-mead simplex algorithm with adaptive parameters, Computational Optimization and Applications 51 (1) (2012) 259–277.

[39] M. Zhang, N. Amaitik, Z. Wang, Y. Xu, A. Maisuradze, M. Peschl, D. Tzovaras, Predictive maintenance for remanufacturing based on hybrid-driven remaining useful life prediction, Applied Sciences 12 (7) (2022) 3218.

[40] S. Fu, D. Zhang, Y. Xiao, C. Zheng, A non-stationary transformer-based remaining useful life prediction method for proton exchange membrane fuel cells, International Journal of Hydrogen Energy 60 (2024) 1121–1133.

[41] J.-H. Chou, F.-K. Wang, Future prediction on the remaining useful life of proton exchange membrane fuel using temporal fusion transformer model, Renewable Energy 247 (2025) 123019.

[42] W. Yu, F. Wu, Z. Huang, X. Huang, P. Jiang, F. Zhou, X. Wen, Z. Zhan, Pemfc performance degradation prediction model based on temporal convolutional network coupling bi-directional long short-term memory and sparrow search optimization algorithm, International Journal of Hydrogen Energy 205 (2026) 153348.

Declaration of interests

☐ 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.

☒ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Amirpiran Amiri reports financial support was provided by Engineering and Physical Sciences Research Council. Yuchun Xu reports financial support was provided by Innnovate UK. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.