

AI-Assisted performance evaluation and health monitoring of PEM fuel cells: an experimental study of component-level aging, LSTM-based degradation forecasting and random forest fault diagnosis

Dr. Praful Nandankar^{1*} and Dr. Prashantkumar V. Dhawas²

Proton-exchange-membrane fuel cells (PEMFCs) are at the core of clean hydrogen energy systems, but their long-term operation is constrained by coupled ageing of the membrane-electrode assembly and operational faults such as flooding, membrane drying, fuel starvation, and catalyst-layer degradation. This paper presents an experimental and data-driven study of PEMFC stack behaviour during 1500 hours of continuous operation using two complementary artificial-intelligence models: a Long Short-Term Memory (LSTM) network for voltage-degradation forecasting and remaining-useful-life (RUL) estimation, and a Random Forest (RF) classifier for fault diagnosis. Polarization and power-density characteristics were measured at 0 h, 500 h, 1000 h and 1500 h, while six operating parameters—cell voltage, stack temperature, relative humidity, pressure drop, current ripple and charge-transfer resistance—were recorded for fault-signature extraction. The revised analysis explicitly relates the observed stack-level degradation to the ageing pathways of the carbon paper, membrane, catalyst and ionomer and discusses their cumulative effects. A component-screening database was organized by mapping each key component and ageing mechanism to measurable sensor signatures and AI-learning variables; direct component-level chemical or microscopic measurements were not part of the present experiment and are therefore treated as a limitation rather than as synthetic data. The LSTM model predicted the voltage-degradation trajectory with RMSE of 6.85 mV, MAE of 4.51 mV and R^2 of 0.970 and predicted RUL with R^2 of 0.960. The RF classifier achieved 97.3% accuracy, 97.4% macro-precision, 97.3% macro-recall and 97.3% F1-score on five operating-condition classes. Results demonstrate that combining component-aware degradation interpretation with sequence-learning prognostics and ensemble fault classification provides a practical health-monitoring framework for PEMFC systems.

Hydrogen is increasingly recognized as an important energy carrier for decarbonization because of its high energy density and zero direct carbon emissions at the point of use^[1, 2]. Proton-exchange-membrane fuel cells (PEMFCs) are widely used in transportation, distributed generation and backup power because of their high efficiency, low operating temperature and rapid start-up capability. However, long-term durability remains a major challenge. Ageing does not originate from a single mechanism: the carbon paper can lose hydrophobicity and transport capability, the membrane can undergo chemical and mechanical degradation, catalyst activity can decrease through dissolution, agglomeration and support corrosion, and ionomer degradation or redistribution can reduce proton transport and triple-phase-boundary effectiveness. These mechanisms

interact with water, thermal and electrochemical operating conditions, producing cumulative losses in voltage and power. Faults such as flooding, membrane drying and fuel starvation can further accelerate these ageing processes^[3]. Therefore, reliable health monitoring requires both temporal performance tracking and interpretation of the component-level mechanisms that contribute to the observed stack response^[4,5].

An AI-assisted health-monitoring framework for PEMFC systems is presented using experimental data collected over 1500 h of continuous operation. An LSTM model is developed to estimate stack-voltage degradation and predict RUL, while an RF classifier identifies five operating conditions: normal operation, flooding, membrane drying, fuel starvation and catalyst degradation. In addition, a component-level screening framework is introduced to organize the known ageing pathways of carbon paper, membrane, catalyst and ionomer and to connect these pathways with the measured sensor signatures used for AI learning. The study therefore distinguishes between directly measured stack-level evidence and component-level mechanistic interpretation. The

¹ Assistant Professor, Electrical Engineering Department, Government College of Engineering, Nagpur, India

² Assistant Professor, Electrical Engineering Department, Government College of Engineering, Chandrapur, India

* Corresponding author, E-mail: pppful@gmail.com

Received 1 May 2026; Accepted 22 June 2026; Published online 29 June 2026

main contributions are: (i) experimental characterization of PEMFC degradation over 1500 h; (ii) component-wise interpretation of the cumulative ageing mechanisms affecting the membrane-electrode assembly and gas-diffusion layer; (iii) development of an LSTM-based degradation forecasting and RUL estimation model; (iv) implementation of an RF-based fault diagnosis framework; and (v) integration of component-aware screening, prognostics and diagnostics into a single AI-assisted health-monitoring architecture.

RELATED WORK

Conventional state-of-health estimation of proton-exchange-membrane fuel cells (PEMFCs) has been based on physics-based degradation models such as semi-empirical voltage-decay and electrochemical impedance models^[6, 7]. Nevertheless, although these techniques provide useful physical insights, they require extensive parameter calibration and are often not robust to different operating conditions and ageing profiles, thus limiting their scope for real-time health monitoring^[8–9].

Recent developments in Artificial Intelligence (AI) and Machine Learning (ML) have enabled data-driven approaches that learn complex degradation patterns directly from operational data^[10, 11]. LSTM networks are suitable for modelling temporal degradation trends and predicting remaining useful life (RUL), whereas Random Forest classifiers are effective for robust and interpretable fault diagnosis using multivariate sensor data^[12, 13]. However, AI models can become difficult to interpret when component-level ageing information is not explicitly organized. A useful screening database should therefore connect each key PEMFC component with its dominant ageing mechanisms, expected electrochemical or transport consequences, measurable sensor signatures and the corresponding AI-learning role^[14–16]. The present study develops this linkage at the screening level while retaining the experimentally measured variables as the actual model inputs.

OVERVIEW OF SYSTEM

PEMFC testing stack

The experimental study was performed on a proton-exchange-membrane fuel cell (PEMFC) stack consisting of several single cells (20 cells) with an active electrode area of 25 cm². The stack was operated using a programmable fuel-cell test station that could independently control key operating parameters like hydrogen and air stoichiometry, stack temperature, back pressure and reactant relative humidity. These controls enabled stable operating conditions to be preserved throughout the experimental campaign and enabled a systematic assessment of the fuel-cell performance under both normal and abnormal operating scenarios. To study the long-term degradation behaviour, the PEMFC stack was run under galvanostatic conditions for 1500 h. The performance degradation was periodically assessed by performing polarization (voltage–current) measurements at four ageing milestones: 0 h (beginning of life), 500 h, 1000 h, and 1500 h. The current density was varied from 0.01 A/cm² to 1.2 A/cm² in each polarization test, so that the evolution of activation, ohmic and

concentration polarization losses could be studied during the ageing period.

Sensor instruments

During the experimental campaign, a complete data-acquisition system was used to continuously monitor the operating condition of the PEMFC stack. Six process variables were recorded at a sampling frequency of 1 Hz: stack-average cell voltage, stack temperature, cathode relative humidity, pressure drop across the cathode flow field, current ripple measured from the programmable electronic load, and charge-transfer resistance obtained from periodic electrochemical impedance spectroscopy (EIS) measurements. The stack-average cell voltage is not a measurement of a single representative cell, but the arithmetic mean of the measurements from all cells in the stack by a cell-voltage-monitoring (CVM) module; this lumped indicator was used as it remains sensitive to homogenous ageing trends while reducing cell-to-cell measurement noise. These variables were chosen because they together contain information about the electrochemical performance and the operating condition of the fuel-cell stack. Parameters such as stack voltage and charge-transfer resistance are indicators of the slow ageing of the membrane-electrode assembly while temperature, humidity, pressure drop and current ripple are sensitive to operational anomalies like flooding, membrane drying and fuel starvation. Therefore, the recorded sensor data offered a comprehensive feature set for the degradation prediction by the Long Short-Term Memory (LSTM) network and the fault diagnosis by the Random Forest (RF) classifier.

Conditions of induced faults

Moreover, four representative PEMFC fault conditions were intentionally applied in addition to the normal operation with carefully controlled operating protocols to develop a labelled dataset for fault diagnosis. Flooding due to increased water accumulation in the cathode flow channels limited oxygen transport to the catalyst layer through over-humidification (the cathode reactant relative humidity was increased to 100% RH (full saturation) while the air stoichiometric ratio was reduced to 1.8. Membrane drying was induced by a decrease of the humidification level of the reactant, which led to membrane dehydration and an increase of the ionic resistance (the reactant relative humidity was reduced to 20% RH). Fuel starvation was created by reducing the hydrogen stoichiometric ratio to below the required operating level (specifically to a hydrogen stoichiometric ratio of $\lambda_{H_2} \approx 1.1$). This caused a limitation in the availability of reactant fuel and increased the risk of cell voltage reversal. Repeated potential-cycling experiments led to accelerated degradation of catalysts, which resulted in catalyst dissolution and loss of electrochemically active surface area (potential cycling was performed between approximately 0.6 V and 1.0 V for 8,000 cycles). Multivariate sensor measurements were continuously recorded for each operating condition and labelled appropriately. The resulting dataset contains samples from normal operation as well as the 4 induced fault conditions, providing a representative benchmark for training, validating and testing the Random Forest based fault-diagnosis model. This experimentally generated dataset also made it possible for the proposed AI framework to discriminate between gradual performance de-

gradation and sudden operational faults, thus enabling the full health monitoring of the PEMFC system.

Component-Level aging mechanisms and cumulative effects

PEMFC degradation is a coupled process involving the gas-diffusion layer, membrane, catalyst layer and ionomer. The carbon paper may experience loss of hydrophobicity, pore-structure changes and transport degradation, which can alter liquid-water removal and reactant access. The polymer membrane can undergo chemical and mechanical degradation, leading to thinning, reduced proton-conduction performance and increased gas crossover. The catalyst layer may suffer platinum dissolution, particle growth or agglomeration, electrochemically active surface-area loss and carbon-support corrosion. The ionomer can degrade chemically or redistribute within the catalyst layer, reducing proton transport and weakening the effective triple-phase boundary. These mechanisms are not independent. Changes in water distribution can simultaneously affect the carbon paper, membrane and ionomer, while catalyst-layer degradation can increase charge-transfer resistance and intensify voltage losses. Consequently, the measured stack voltage, relative humidity, pressure drop and charge-transfer resistance represent combined effects rather than isolated signatures of a single component.

The 1500 h experiment provides direct evidence of cumulative stack-level degradation through the downward shift of the polarization curves and the reduction in peak power density from 0.51 W/cm² at beginning of life to 0.31 W/cm² at 1500 h. Because destructive post-mortem characterization of

individual components was not performed in the reported experiment, the component-level mechanisms are used as a mechanistic interpretation of the measured electrochemical response rather than as independently quantified component-degradation rates. This distinction is important for AI learning: the models are trained only on measured electrical, thermal, humidity, pressure, ripple and impedance-derived variables, while the component mechanisms provide a physically meaningful screening layer for interpreting the learned patterns.

Key-Component screening database and AI learning

To address the need for a structured database for key-component screening and AI learning, the experimental records were organized into a component-aware screening matrix. Each record is associated with operating time, measured sensor variables, operating-condition label and the component-level ageing mechanism most consistent with the observed response. The screening matrix does not introduce artificial component measurements; instead, it provides a transparent mapping between known ageing mechanisms and the variables actually available to the LSTM and RF models. The LSTM uses the chronological stack-voltage series to learn cumulative degradation trends, whereas the RF classifier uses the six measured features to distinguish normal operation and the four induced fault conditions. Component-level tags are therefore used for interpretation and screening, not as unmeasured inputs to the models. The key-component aging screening matrix and AI-learning linkage is illustrated in Table 1.

Table 1. Key-component aging screening matrix and AI-learning linkage

Key component	Representative ageing mechanism	Expected cumulative effect	Measured/related screening variables	AI-learning role
Carbon paper	Loss of hydrophobicity; pore/transport degradation	Water-management and mass-transport losses	Relative humidity, pressure drop, cell voltage, current ripple	Interpretation of flooding and transport signatures
Membrane	Chemical/mechanical degradation; thinning; dehydration	Higher resistance, gas crossover and voltage loss	Cell voltage, temperature, relative humidity, charge-transfer resistance	Interpretation of long-term voltage degradation and drying
Catalyst	Dissolution/agglomeration; support corrosion; active-area loss	Higher charge-transfer losses and lower power capability	Charge-transfer resistance, cell voltage, current ripple	Links catalyst degradation to fault-class features
Ionomer	Chemical degradation/redistribution; reduced proton transport	Reduced ionic conductivity and triple-phase-boundary effectiveness	Cell voltage, relative humidity, charge-transfer resistance	Component-aware interpretation of degradation and drying

AI METHODOLOGY

Remaining useful life estimation and degradation forecasting with LSTM

Long Short-Term Memory (LSTM) network was used to predict the temporal degradation behaviour of the PEMFC stack based on the recorded stack-average cell voltage. Before training, the voltage time series was normalized to improve numerical stability and speed up the network convergence. The whole dataset acquired during the 1500 h durability test was divided in a chronological order, where the first 70% of the data were used for training, while the rest 30% were used for testing. The data were split strictly chronologically, without shuffling, so that the test set was fully representative of the last 450 h of the 1500 h durability test (i.e. the most aged part), and the model was therefore tested on previously

unseen, later stage degradation behaviour, not on interpolated or randomly sampled operating points. A sliding-window approach is applied to generate sequential input–output samples, which enables the network to effectively learn the long-term temporal dependencies and degradation patterns from the historical voltage measurements.

The suggested network architecture is composed of two stacked LSTM layers and a fully connected output layer which estimates the future stack voltage. The network parameters were optimized using the Adam optimizer with the mean squared error (MSE) loss function that minimizes the difference between the predicted and experimentally measured voltage trajectories. Because of the gated memory structure, the LSTM network is capable of well capturing the nonlinear degradation dynamics of the PEMFC stack and accurately forecasting the future voltage evolution of the PEMFC stack

under continuous operation conditions. The remaining useful life (RUL) of the PEMFC stack was estimated by extrapolating the predicted voltage trajectory to the pre-defined end-of-life (EOL) threshold of 0.55 V, which corresponds to an approximate 10% voltage degradation relative to the beginning-of-life performance. This value is consistent with end-of-life criteria widely used in the PEMFC durability literature, where a 10–20% relative decline in stack voltage (or equivalent decline in rated power) from beginning of life performance is often used to define practical end of life, including durability targets set for automotive PEMFC stacks. The 0.55 V threshold used here corresponds to the lower (10%) end of this range. The remaining useful life (RUL) was defined as the time difference between the current operating condition and the predicted end-of-life (EOL) point. This method enables prediction of the remaining operational lifetime in a timely manner, providing valuable information for predictive maintenance scheduling, health management and enhancing the reliability and availability of PEMFC systems.

Random forest based fault diagnosis

A Random Forest (RF) classifier was developed to identify the PEMFC stack operating condition based on multivariate sensor measurements. The classifier was trained with six input features, which are, stack average cell voltage, stack temperature, cathode relative humidity, cathode pressure drops, current ripple and charge transfer resistance. These parameters are chosen because they represent both electrochemical degradation and anomalous operating behaviour related to common PEMFC faults. The component-screening matrix described in Section III-E was used to organize the feature-to-mechanism interpretation, while the RF model was trained only with the six measured variables. The RF model consisted of 250 decision trees with a maximum tree depth of nine to balance the classification accuracy and model complexity and reduce the chance of overfitting. The labelled dataset was split into 75% training and 25% testing sets, preserving the class distribution of the five operating conditions: normal operation, flooding, membrane drying, fuel starvation and catalyst degradation. During training, each decision tree was created using random samples and random features, and the final operating condition was determined by majority voting of the decision trees. Then, feature importance analysis was performed to determine the sensor variables that contributed most to fault discrimination and model interpretability.

Data screening and AI-learning protocol

Before AI learning, the recorded data were screened according to operating time, operating condition and measurement availability. The 1500 h durability dataset was retained as a chronological series for LSTM learning, with the first 70% used for training and the final 30% used for testing. For RF diagnosis, labelled samples from normal operation, flooding, membrane drying, fuel starvation and catalyst degradation were organized with the six measured features. The component-aware screening matrix was then used to check whether each feature and class had a plausible physical association with the relevant ageing or fault mechanism. This procedure reduces the risk of assigning a purely statistical meaning to the learned patterns and makes the AI-learning database traceable to physical PEMFC behaviour.

Evaluation of performance

The performance of the LSTM degradation forecasting model was evaluated using three common regression metrics, root mean square error (RMSE), mean absolute error (MAE), and the coefficient of determination (R^2). The magnitude of the prediction errors is quantified by the root mean squared error (RMSE) and mean absolute error (MAE), and the degree of agreement between the predicted and experimentally measured voltage trajectories is quantified by the R^2 value. Together, these metrics provide a comprehensive evaluation of the model's forecasting accuracy and ability to capture the long-term degradation behaviour of the PEMFC stack. The overall classification accuracy, macro-averaged precision, macro-averaged recall, and macro-averaged F1-score were used to measure the classification performance of the Random Forest fault diagnosis model for all fault categories. Also, confusion matrix was used to assess the classification results under each operating condition, and to check possible misclassifications between similar fault modes. All of these evaluation metrics together give a complete evaluation on the prognostic and diagnostic performance of the proposed AI-assisted PEMFC health monitoring framework and confirm its effectiveness for accurate degradation forecasting, remaining useful life (RUL) prediction and real-time fault diagnosis.

EXPERIMENTAL RESULTS AND DISCUSSION

Polarization and power-density characteristics

Fig. 1 presents the polarization and power-density curves measured at four ageing points during the 1500 h durability test. The open circuit voltage of the beginning-of-life (BOL) curve is approximately 0.94 V, and its peak power density is 0.51 W/cm² at a current density of about 1.1 A/cm². The polarization curves gradually shift downwards with increasing operating time with a corresponding decrease in peak power density to 0.31 W/cm² at 1500 h, corresponding to an overall power-density loss of about 39%. This degradation is consistent with the combined effects of membrane ageing, catalyst degradation, increasing ohmic resistance and increasing mass transport limitations. At higher current densities, voltage decay is more pronounced, which indicates a gradual increase in internal resistance and a decrease in reactant transport capability due to ageing induced electrochemical losses. This

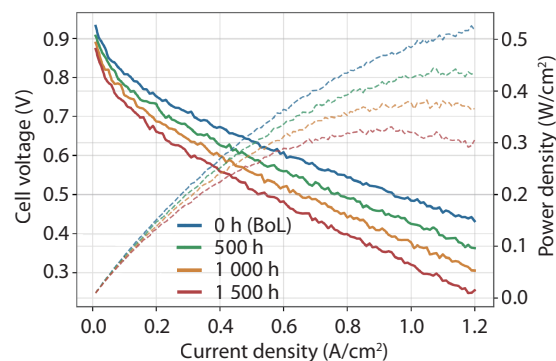


Fig. 1 Polarization and power-density curves at 0 h, 500 h, 1000 h and 1500 h of operation

causes a gradual decrease in the maximum power output with the operating time, which corresponds to the gradual degradation of the performance of the PEMFC stack. The observed clear and consistent degradation trend over the 1500 h test reveals the long-term aging behavior of the fuel cell and lays a strong experimental foundation for further developing the AI-based degradation forecasting and RUL prediction models.

Component aging pattern and cumulative stack-level effect

The measured performance trend is consistent with cumulative ageing of multiple PEMFC components rather than a single isolated failure mechanism. The reduction in peak power density from 0.51 W/cm² to 0.31 W/cm² over 1500 h indicates a substantial loss in stack performance. At the component level, carbon-paper degradation can increase water-retention and mass-transport limitations; membrane ageing can increase ohmic losses and reduce effective proton transport; catalyst ageing can increase charge-transfer resistance and reduce electrochemically active surface area; and ionomer degradation can weaken proton conduction within the catalyst layer. These effects can reinforce one another. For example, poorer water transport can promote either local flooding or membrane dehydration, while catalyst and ionomer degradation can increase electrochemical losses and accelerate the voltage decline. The observed stack response should therefore be interpreted as the cumulative outcome of interacting ageing mechanisms.

LSTM voltage-degradation forecasting

Fig. 2 shows the comparison of the measured voltage trajectory for the stack average with the one predicted by the LSTM model for the last 30% of the degradation test (about 450 h). The predicted voltage agrees well with the measured trend over the entire forecasting horizon, accurately capturing the slow decrease of the stack voltage as the ageing progresses. The model also captures the non-linear degradation behaviour, including the accelerated voltage decay rate at later operational stages, demonstrating its ability to learn complex temporal degradation patterns. For quantitative results, the RMSE, MAE and R² value of LSTM model were 6.85 mV, 4.51 mV and 0.970 respectively, indicating a good match of the predicted voltage profile and the experimental one. The robustness and generalization ability of the proposed forecasting model are confirmed by the low prediction errors and the high coefficient of determination. These results

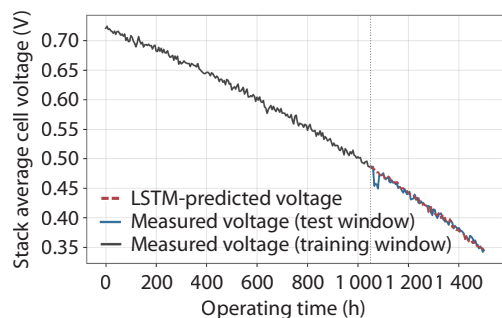


Fig. 2 LSTM-based voltage-degradation forecasting over the test horizon.

demonstrate that the LSTM architecture can learn the temporal degradation characteristics from historical voltage measurements without requiring an explicit physics-based degradation model. Thus, the proposed approach offers a reliable and computationally efficient framework for continuous PEMFC health monitoring, long-term degradation prediction and accurate remaining useful life (RUL) estimation, supporting predictive maintenance and improving the operational reliability of fuel-cell systems.

Remaining-Useful-Life estimation

Fig. 3 presents a parity plot of the predicted and actual remaining useful life (RUL) for 60 evaluation samples, obtained from the test-horizon forecasts. Most of the prediction points are very close to the ideal 45° reference line and within the $\pm 15\%$ error band, which indicates the high accuracy of the proposed model in estimating the remaining life in the wide variety of degradation states. The high agreement between the predicted RUL and the actual RUL validates the robustness and generalization capability of the proposed forecasting framework. The LSTM-based RUL prediction model achieved an RMSE of 78.3 h, MAE of 57.8 h and R² of 0.960, showing excellent prognostic performance. The higher the RUL values, the slightly larger the dispersion, which is expected due to the accumulation of uncertainty over longer prediction horizons, a common feature of multi-step time series prediction. Although the uncertainty is increased, the prediction errors are still acceptable, indicating the stability and reliability of the proposed model. The obtained results show that the proposed LSTM framework is capable of predicting the remaining useful life of PEMFC stacks with a high degree of accuracy, which would allow timely planning of maintenance activities, reduce the risk of unexpected failures and facilitate condition-based maintenance strategies to improve the reliability and operational efficiency of the system.

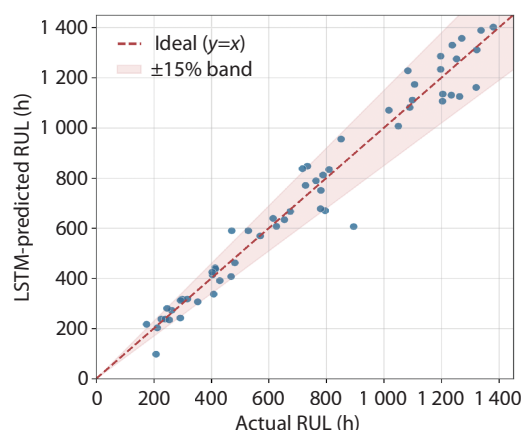


Fig. 3 Remaining-useful-life prediction parity plot (LSTM-based extrapolation to end-of-life voltage threshold)

Random forest fault classification

The performance of the Random Forest (RF) classifier is summarized in Table II. The corresponding confusion matrix for the held-out test dataset is shown in Fig. 4. The proposed classifier achieved an overall classification accuracy of 97.3% and correctly classified most of the PEMFC operating conditions and fault categories. There were only few misclassifica-

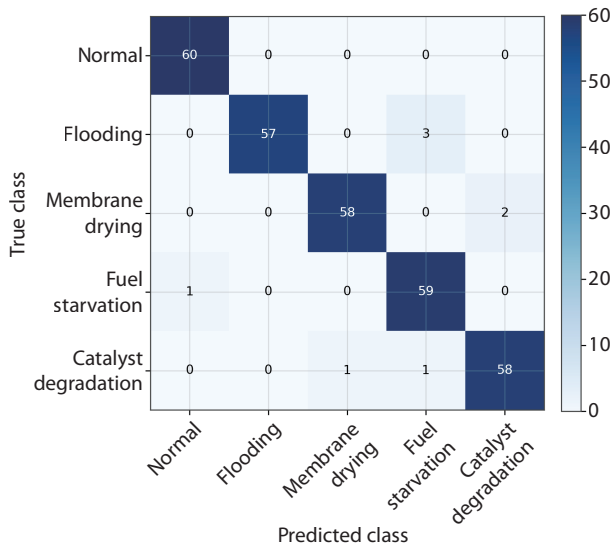


Fig. 4 Confusion matrix for the Random Forest fault classifier on the held-out test set.

tions, mostly between flooding and fuel starvation, and between membrane drying and catalyst degradation. These misclassifications show that the associated fault pairs have partially overlapping electrical and electrochemical signatures under some operating conditions and are therefore more difficult to differentiate. The RF classifier's input features' relative importance is shown in Fig. 5. The most influential features among all measured variables were found to be charge-transfer resistance and relative humidity, followed by cell voltage and pressure drop. This ranking is in accordance

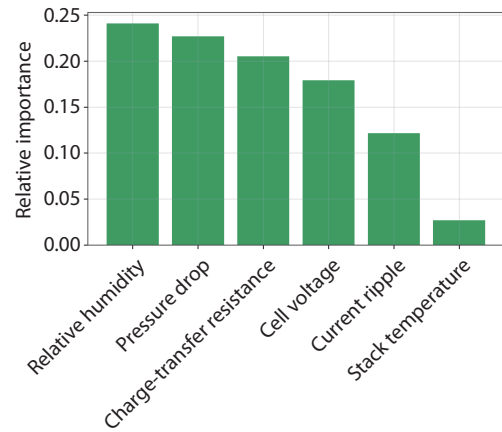


Fig. 5 Random Forest feature-importance ranking for fault diagnosis.

with the underlying electrochemical behaviour of PEMFCs, since the charge-transfer resistance is very sensitive to the catalyst activity and the membrane condition, while the relative humidity is very influential on the membrane hydration and water management. The important contribution of these features demonstrates that they are effective for discrimination of different fault modes. In conclusion, the successful classification outcomes and the meaningful feature importance findings demonstrate that the proposed RF-based framework offers a robust and interpretable way for real-time fault diagnosis and prognostics of PEMFCs, thereby enabling reliable condition monitoring and predictive health management in real-world fuel-cell applications. The performance of AI models are summarised in Table 2.

Table 2. Summary of AI Model Performance

Model	Metric	Value
LSTM (voltage forecast)	RMSE	6.85 mV
LSTM (voltage forecast)	MAE	4.51 mV
LSTM (voltage forecast)	R ²	0.970
LSTM (RUL estimation)	RMSE	78.3 h
LSTM (RUL estimation)	MAE	57.8 h
LSTM (RUL estimation)	R ²	0.960
Random Forest (fault diagnosis)	Accuracy	97.3%
Random Forest (fault diagnosis)	Macro-precision	97.4%
Random Forest (fault diagnosis)	Macro-recall	97.3%
Random Forest (fault diagnosis)	F1-score	97.3%

Fig. 6 presents a comparative overview of the performance metrics for the proposed LSTM and Random Forest (RF) models. The LSTM model shows good accuracy of degradation prediction with low prediction errors and high coefficient of determination. The RF classifier achieves good accuracy of fault classification under various operating conditions. Both models are good, suggesting that they are effective in addressing complementary aspects of PEMFC health monitoring. The collective results verify that the proposed AI-based framework can attain accurate degradation prediction and reliable fault diagnosis simultaneously, thus, enabling comprehensive condition monitoring of PEMFC systems. This integration enables predictive maintenance, improves operational reliability and helps to extend the service life of clean hydrogen fuel-cell systems.

The experimental results show that the health monitoring assisted by AI can significantly improve the conventional diagnostics of PEMFC. The high R² values of LSTM model for both voltage forecasting and RUL estimation indicate that voltage time-series data alone have sufficient information to perform accurate medium-term degradation prediction and hence less reliance on expensive and invasive diagnostic techniques such as in-situ electrochemical impedance spectroscopy (EIS). Similarly, the RF classifier can reach excellent fault-diagnosis performance with only six measurable operating parameters, which makes it suitable for real-time implementation into embedded fuel-cell management systems with limited computational resources. However, it is important to mention that the charge-transfer resistance among the six input features is acquired from periodic broadband EIS

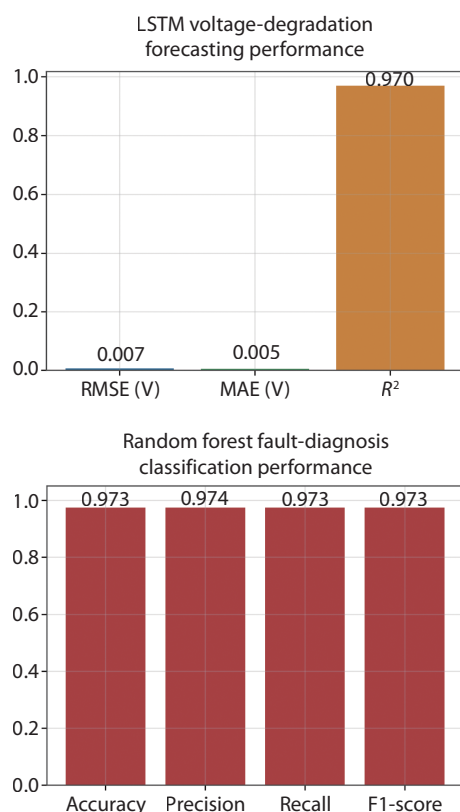


Fig. 6 Consolidated performance summary of the LSTM degradation-forecasting model and the Random Forest fault-diagnosis model.

measurements, which are typically not available in real-time at high sampling rates in most field-deployed PEMFC systems because of the added hardware complexity and measurement time needed for conventional EIS interrogation. However, recent developments in online and hybrid-frequency EIS techniques have dramatically reduced the time of acquisition and computational load required to estimate an equivalent charge-transfer resistance, thus enabling near-real-time updates suitable for periodic, rather than continuous, fault-diagnosis cycles^[17]. The proposed RF classifier can therefore be used in practice in a hybrid sensing scheme where the fast-sampled features (voltage, temperature, humidity, pressure drop and current ripple) are logged continuously, while the charge-transfer-resistance feature is refreshed at a lower update rate using an onboard EIS module. The sensitivity of the classifier performance to a lower update rate for this single feature will be investigated in future work.

The feature importance analysis also shows that the charge transfer-resistance and the relative humidity are the main indicators in the fault identification, providing important insights into the physical factors affecting the degradation and abnormal operation of PEMFCs. A small percentage of misclassification is observed between flooding and fuel starvation and between membrane drying and catalyst degradation. These confusions are physically plausible: both flooding and fuel starvation can generate a transient cell-voltage drop with accompanying increased current ripple at a given current density, since a lack of reactant availability at the catalyst

layer, due to either liquid-water blockage or a decreased hydrogen stoichiometric ratio, generate similar concentration-polarisation signatures; the differentiation between the two is therefore reliant upon secondary cues of pressure drop and relative humidity, which are less pronounced under partial-load conditions. Again, the combination of membrane drying and catalyst degradation increases both charge-transfer and ohmic resistance, so when the catalyst-layer thinning is still moderate, the resulting impedance signature partially overlaps with that of a mildly under-humidified membrane, making the two conditions more difficult to distinguish from instantaneous sensor snapshots alone. Overall classification accuracy is found to be high. These results show that the combination of degradation forecasting and fault diagnosis provides a comprehensive health-monitoring framework, which can enable predictive maintenance, reduce unexpected failures and improve the reliability and operational lifetime of PEMFC systems. Future work can further improve diagnostic performance with additional sensing modalities such as humidity distribution, impedance spectra or thermal imaging along with advanced deep-learning architectures for better fault separation and long-term prognostics.

Study limitation: the present experiment directly measured stack-level electrical, thermal, humidity, pressure and impedance-related variables, but it did not include destructive post-mortem characterization of the carbon paper, membrane, catalyst or ionomer. Accordingly, the component-ageing discussion and screening database are mechanistic and traceability-oriented rather than quantitative component-health measurements. Future experimental work should combine the present AI framework with component-level characterization to provide direct labels for each ageing mechanism and to further improve component-specific diagnosis.

CONCLUSION

In this paper, an experimental study of PEMFC stack degradation and fault behaviour over 1500 hours of operation was presented by integrating an LSTM-based voltage degradation forecasting and remaining useful life (RUL) estimation model with a Random Forest (RF)-based fault diagnosis classifier. The revised analysis additionally considered the ageing patterns of the carbon paper, membrane, catalyst and ionomer and their cumulative interaction in producing the observed stack-level performance loss. The experimental results indicated that the LSTM model achieved an RMSE of 6.85 mV and R^2 of 0.970 for voltage forecasting and R^2 of 0.960 for RUL prediction, while the RF classifier obtained 97.3% overall fault-diagnosis accuracy for the five representative operating conditions. A component-aware screening database was organized to connect ageing mechanisms with measurable variables and AI-learning tasks. Importantly, the present study does not claim direct component-level measurements where these were not collected; the screening database is used to improve physical interpretability and traceability of the measured data. The proposed framework therefore combines experimental evidence, component-level mechanistic interpretation, degradation prediction and fault diagnosis for comprehensive PEMFC health monitoring. Future work will focus on component-resolved experimental characterization, dynamic

operating conditions, multi-stack systems and online adaptive learning so that the AI models can be updated using direct evidence of evolving degradation mechanisms.

■ REFERENCES

1. J. Larminie and A. Dicks, Fuel Cell Systems Explained, 2nd ed. , Wiley, USA, 2003.
2. Y. Wang, K. S. Chen, J. Mishler, S. C. Cho, X. C. Adroher, *Appl. Energy*, 2011, 88, 981
3. A. Nagila, S. Mg, F. D. Deepak, R. Kumar, P. V. Nandankar, *Proc. Int. Conf. Edge Comput. Appl.*, 2022, 1, 711
4. Chandramma, P. Prakash, P. Nandankar, H. Roopa, I. Kathir and P. Singh, *AIP Conf. Proc.*, 2023, 2690, 020047.
5. H. Xu, L. Wang, Z. Liu, G. Qiu, H. Zhou, H. Su, *IEEE Trans. Transp. Electr.*, 2026, Early Access.
6. Z. Hua, Z. Zheng, E. Pahon, M. C. Marion-Péra. D. Hissel, *Energy Convers. Manage.*, 2021, 231, 113825
7. Z. Zheng, R. Petrone, M. C. Pera, M. Hissel, C. Becherif, H. Pianese, N. Yousfi-Steiner, M. Sorrentino, *Int. J. Hydrogen Energy*, 2013, 38, 8914
8. R. Petrone, Z. Zheng, D. Hissel, M. C. Pera, C. Pianese, M. Sorrentino, M. Becherif, N. Yousfi-Steiner, *Int. J. Hydrogen Energy*, 2013, 38, 7077
9. K. Chen, S. Laghrouche, A. Djerdir, M. El Bagdouri, A. El Moudni, *Int. J. Hydrogen Energy*, 2019, 44, 1930
10. C. Zhang, S. Zhao, X. Yang, Y. Zhang, *IEEE Trans. Energy Convers.*, 2021, 36, 41.
11. M. A. R. S. Khan, S. Al-Dahidi, *Renew. Sustain. Energy Rev.*, 2022, 167, 112742
12. P. Archana, P. Divyabharathi, S. R. Balaji, N. Kumaresan, P. Veeramani, S. T. Naitik, S. M. Rafi, P. V. Nandankar, G. Manikandan, *Measurement: Sensors*, 2022, 24, 100575
13. Y. Cheng, H. Zhang, L. Wang, *J. Power Sources*, 2021, 482, 228960.
14. B. Dwarakanath, G. Shrivastava, R. Bansal, P. Nandankar, V. Talukdar, M. A. Usmani, *Int. J. Commun. Netw. Inf. Secur.*, 2023, 14, 342.
15. P. Nandankar, M. V. Aware, *Proc. IEEE Int. Conf. Power Electron. , Drives Energy Syst. (PEDES)*, IEEE, India, 2012.
16. A. T. Jambhulkar, A. M. Bhojar, C. M. Shivankar, P. S. Balwad and P. Nandankar, *Int. J. Innovations Eng. Sci.*, 2023, 8, 1.
17. Q. Li, H. Wang, T. Wang, X. Li, Y. Liu, W. Chen, Z. You, *IEEE Trans. Transp. Electr.*, 2025, 11, 2707



©2026 The Authors. *Energy Lab* is published by Lab Academic Press. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.