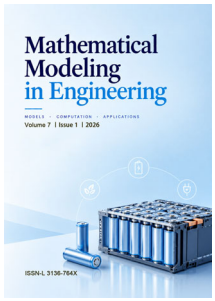
UNIVERSITY OF WEST ATTICA

eISSN: 3136-764X

Mathematical Modeling in Engineering

Vol. 7, No. 3, July, 2026



TabM-LightGBM Model for Electrolyzer Stack Degradation Diagnosis Based on Polarization Curves and Pressure Data

Hugo Lefèvre[a], Louis Bernard[a,*]

[a] *School of ICT, Grenoble INP, Grenoble, 38031, France*

Received 12 May 2026; Revised 31 July 2026; Accepted 07 August 2026; Published 12 August 2026

Abstract

Electrolyzer stack degradation diagnosis is difficult because voltage rise, polarization-curve distortion, gas pressure fluctuation and water-gas transport disturbance can occur simultaneously under variable load and regulation conditions. This paper introduces a TabM-LightGBM model for stack degradation diagnosis based on polarization curves and pressure data. The method first extracts curve-slope, voltage-residual, pressure-variation, operating-state and pressure-reliability features from segmented operating windows. TabM is used to learn multiple tabular representations of feature interactions, while LightGBM performs interpretable degradation classification and maintenance-priority ranking. The Pressure Reliability Index is introduced to reduce interference from transient pressure pulsations during electrochemical ageing and balance-of-plant regulation. A simulated stack monitoring dataset of 6,055 operating windows is constructed for normal, activation-loss, ohmic-growth, mass-transfer-risk and mixed-degradation conditions. The proposed model achieves 93.4% overall accuracy, a macro-F1 score of 0.921 and an expected calibration error of 0.048. Compared with a standard LightGBM baseline, mixed-degradation recall increases by 14.7%, and the false warning rate under pressure disturbance decreases by 11.2%. These results indicate that combining polarization-curve evidence with pressure-feature analysis improves degradation-state recognition, probability calibration and engineering-review reliability. The output reports degradation class, confidence level, dominant evidence and maintenance priority, supporting the distinction among membrane ageing, contact-resistance increase, mass-transfer limitation and uncertain pressure disturbance.

Keywords: Electrolyzer Stack; Degradation Diagnosis; Polarization Curve; Pressure Data; TabM-LightGBM; Maintenance Priority

1. Introduction

Electrolyzer stacks are key equipment for green hydrogen production, and their long-term reliability is directly related to energy conversion efficiency and maintenance cost [1]. Polarization curves provide a structured description of activation loss, ohmic resistance and mass-transfer limitation during stack operation [2]. Pressure data reflect gas-water transport, channel blockage, seal condition and balance-of-plant regulation stability [3]. Recent degradation-diagnosis studies have shown that single-source indicators are insufficient when electrochemical aging and operating disturbance occur together [4]. Therefore, a diagnosis model should connect polarization behavior, pressure response and operating context rather than treating each signal as an isolated alarm.

At present, the diagnostic methods are equivalent-circuit analysis, electrochemical impedance spectroscopy, voltage-residual tracking, pressure-threshold alarming and data-driven classification. The above are feasible; however, they have different deficiencies in online supervision. Equivalent-circuit parameters are more intuitive but need additional measurements. Voltage threshold devices are relatively simple to realise, but they are easily influenced by

Corresponding author: louis.ber@grenoble-inp.fr

https://doi.org/10.68406/mme.2026.vol7iss3nm6:69-80

➢ This is an open access article under the CC BY-NC-SA 4.0 license (https://creativecommons.org/licenses/by/4.0/).

© Authors retain all copyrights.

temperature, current density and water management. Pressure fluctuations are due to transport restrictions; otherwise, they can be caused by temporary regulation of flow. Recently, research on the monitoring of hydrogen-energy equipment degradation has begun to integrate curve-level evidence with dynamic operating data [5]. Research in tabular machine learning has also shown that, with a relatively small amount of data and when feature interpretability is needed, structured features can be effectively used [6]. Many of the test platforms are also recording polarisation and pressure signals at different sampling rates. Generally, the polarisation curve is taken at the end of a planned operating sweep, and a pressure signal is continuously recorded during dynamic operation. If only the two sources are connected at the final decision stage, the model will be unable to determine whether the cause of the warning is electrochemical loss or a disturbance in gas-water transport. Therefore, the diagnosis framework in this paper takes pressure behaviour as the contextual evidence for curve interpretation and does not use an independent alarm path.

This paper develops a TabM-LightGBM model for electrolyzer stack degradation diagnosis based on polarization curves and pressure data. Section 2 reviews degradation mechanisms, pressure-related operating indicators and structured-data diagnosis models. Section 3 presents the proposed method, including feature construction, TabM-LightGBM fusion and degradation probability calculation. Section 4 reports data distribution, diagnostic accuracy, baseline comparison, robustness, interpretability and maintenance-priority outputs. Section 5 summarizes the main findings, engineering value, limitations and future deployment plan. The proposed framework treats diagnosis as a complete process that includes data construction, model prediction, evidence interpretation, confidence assessment and maintenance response.

2. Related Work

2.1. Electrolyzer Stack Degradation and Polarization Curve Diagnosis

Polarisation curve analysis is commonly employed to relate the measured voltage response with electrochemical loss components. The low-current region is mainly due to the activation effect; the middle-current region is ohmic resistance; and the high-current region is limited by mass transport. The above Areas are not entirely independent in practice for stack operations. Membrane dehydration, catalyst-layer degradation, growth of contact resistance and gas accumulation can all increase the voltage. Recently, degradation studies of electrolysers have thus focused on interpreting the shape of the curve rather than using a single operating point [7]. As a result of the engineering conditions, the curve must be interpreted in sections. An increase in voltage at low currents can be caused by catalyst or activation losses, and a nearly parallel upward shift over the middle region is due to resistance growth. At a high current density, the rise in voltage may be caused by a shortage of water supply, local gas accumulation or transport restrictions. The design of the features is to use slope and residual indicators for different current-density ranges to achieve regional interpretation.

Another problem is that the decay is slow and may be hidden by short-term fluctuations in operation. A temporary increase in stack voltage may be caused by a temperature change or gas management; continuous alteration of the curve slope is often associated with material aging. Therefore, the model of maintenance engineers should differentiate between stable degradation evidence and reversible disturbances. The structure of diagnosis should retain the current density, voltage residual, slope change, operating-state tags, etc., and interpret the same voltage increase differently depending on the situation [8]. The Index of Curvature should also be relatively stable over time. The degradation phenomenon should occur within the comparable operating window, and after pressure or temperature stabilisation, no reversible disturbances should be present. Therefore, a single curve snapshot is not used as the input to the model at this time. It is a kind of curve index that adds an operating-state tag to distinguish between short-term fluctuations and extended downturns more clearly.

2.2. Pressure Fluctuation and Gas-Water Transport Indicators

Pressure response provides an extra description of gas-water transport in the stack and auxiliary loop. The two-phase flow instability, channel blockage, risk of gas crossover and balance-of-plant regulation are the causes of abnormal pressure fluctuation. Other Sources of Error for Pressure Signal Include Pump Control, Valve Adjustment, Purging and Measurement Noise. Research on hydrogen systems has shown that changes in pressure should be analysed in light of current-density and voltage behaviour to prevent misdiagnosis [9]. Pressure information is very useful for high-current diagnosis. At a high current density, the generation of gas and transport of water are more pronounced; pressure fluctuations may occur before a substantial increase in voltage. Use the above information correctly; otherwise, the purge operation or temporary valve adjustment will result in a pressure pattern similar to that of a transport failure. Therefore, the new method has added a pressure change, voltage and curve indicator, and is no longer limited to pressure.

Do not use the pressure data as the fault label for degradation diagnosis. A short pressure pulse is a normal change in operation; however, continuous amplification of pressure fluctuations in the high-current area may indicate mass-

transfer limitations. Therefore, a wrong diagnosis will not be made that requires an unnecessary shutdown or delayed maintenance of the electrolyzer stack. Therefore, the method above solves the problem of pressure fluctuation by adding a pressure reliability index to determine whether the pressure channel is stable enough for diagnosis [10]. The two will be used differently. If the pressure fluctuation is due to a temporary operating disturbance, one may re-measure or verify the operating conditions. If there is pressure fluctuation at the same time as a high-current voltage residual and a continuous slope change, the operator should investigate any problems in the water management and gas outlet. Therefore, the diagnosis model needs to return both class probabilities and evidence information.

2.3. Tabular Deep Learning and Gradient Boosting for Industrial Diagnosis

Recently, tabular deep learning has attracted attention in industrial monitoring tasks that are not based on images or long natural-language sequences. Gorishniy et al. [11] proposed TabM-style representation learning to generate multiple feature views and improve interaction modelling for numerical variables. This structure is suitable for analysing electrolysers because curve slopes, residuals, pressure statistics and operating tags are all related through non-linear functions. Feature interaction in traditional tabular modelling is generally handled by manually adding cross-terms or relying on tree splits. TabM provides multiple representations of the same structured data, which is useful for electrolyzer-stack diagnosis because the meaning of pressure variables changes with current density, temperature and curve region. Therefore, a multi-view tabular representation can address the operating-context dependency of a single fixed feature vector more effectively.

LightGBM remains suitable for efficient, stable and explainable modelling of structured industrial data because it supports non-linear feature thresholds and generates feature importance for maintenance inspection. Li et al. [12] showed that LightGBM-based industrial warning models can perform well when features have clear engineering meaning and abnormal samples are limited. In addition, Hollmann et al. [13] demonstrated that tabular foundation models can improve prediction quality on small structured datasets, which supports the combination of learned tabular representations with tree-based decision learning. Therefore, this paper combines TabM representation with LightGBM decision learning to use both learned interactions and interpretable split-based evidence.This hybrid Design is to support degradation classification under the constraint of plant-side monitoring [14]. Confidence estimation is also provided to separate reliable degradation evidence from uncertain operating disturbances in the stack diagnosis [15]. Maintenance priority is then generated as an independent output for engineering review [16].

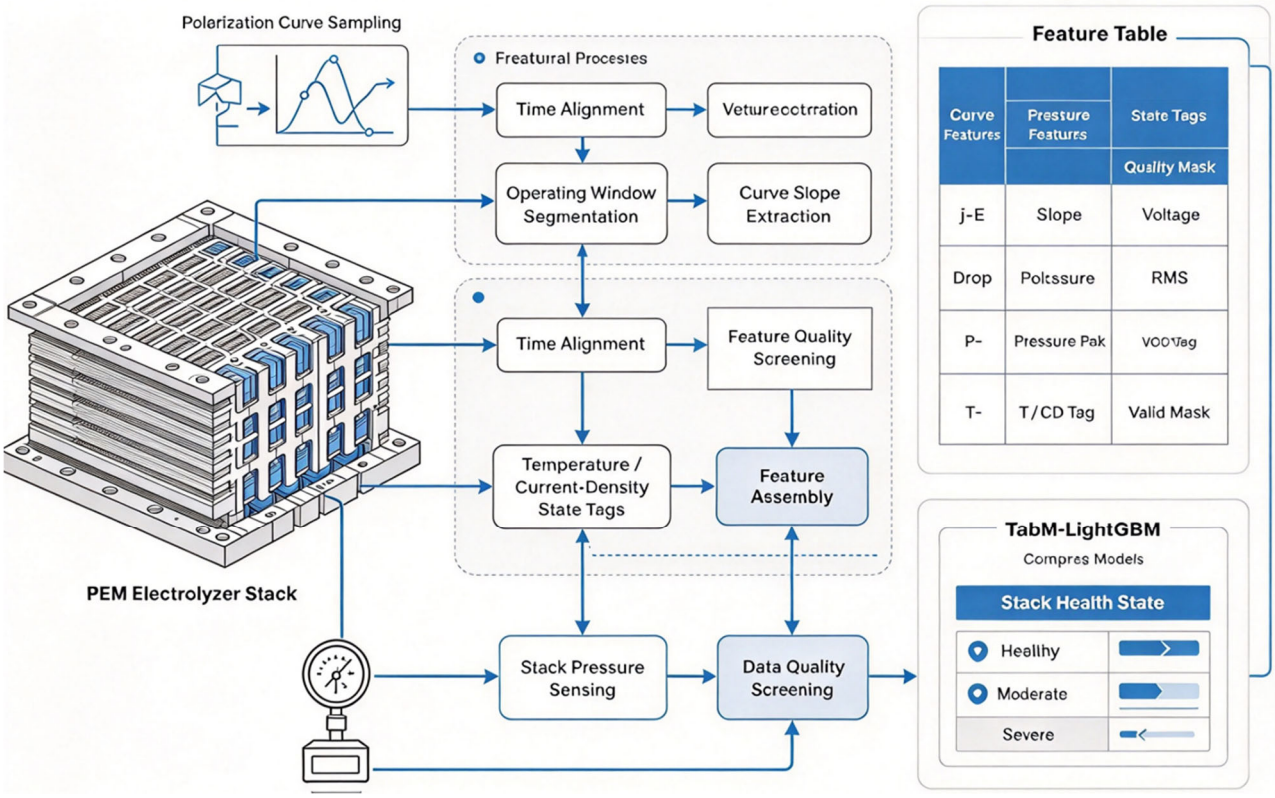


Figure 1. Polarization Curve and Pressure Feature Construction Workflow

3. Material and Method

3.1. Polarization Curve and Pressure Feature Construction

The input data include sampled polarization curves, stack pressure sequences, current density, temperature and operating-state labels. Figure 1 presents the diagnostic data construction process. The figure shows that curve points and pressure windows are aligned by operating time before curve-slope, voltage-residual, pressure-variation and pressure-reliability indicators are extracted. This alignment prevents the model from interpreting a short purge event as permanent electrochemical degradation. It also preserves the physical meaning of low-current activation evidence, middle-current resistance evidence and high-current mass-transfer evidence, so each diagnostic result can be traced back to measurable stack behavior.

The polarization slope is calculated to describe the voltage increase per unit current-density interval:

$$k_j = \frac{V_{j+1} - V_j}{I_{j+1} - I_j} \tag{1}$$

The slope value is then interpreted together with operating pressure because mass-transfer disturbance may change the curve shape without indicating permanent electrochemical aging.

$$P_v = \frac{\mathrm{std}(P_t)}{\mathrm{mean}(P_t)} \tag{2}$$

3.2. TabM-LightGBM Feature Fusion Framework

TabM-LightGBM is designed to combine feature-interaction learning with interpretable tree-based diagnosis. Figure 2 describes the proposed framework. The left part receives polarization-curve and pressure features, the middle part builds multiple TabM representations, and the right part uses LightGBM to output degradation class, confidence and evidence contribution. The model uses structured variables rather than a single voltage or pressure threshold, allowing engineers to review which curve region and pressure condition supported the warning. Confidence is therefore interpreted as evidence reliability under the current operating context, not as authorization for automatic maintenance action.

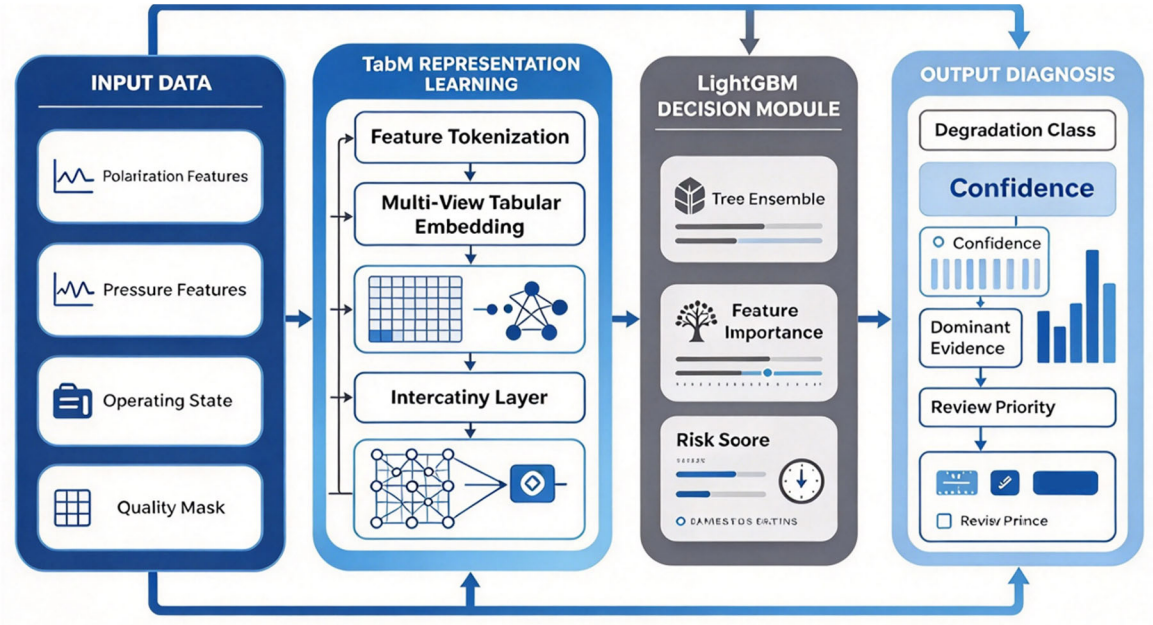


Figure 2. TabM-LightGBM Degradation Diagnosis Framework

The feature vector is first normalized to reduce the influence of scale differences among voltage, current and pressure indicators:

$$z_m = \frac{x_m - u_m}{s_m} \tag{3}$$

After normalization, TabM produces multiple feature views so that interactions among curve slope, pressure variation and operating state can be learned:

$$h_r = \mathrm{Tab}M_r(z) \tag{4}$$

The learned views are aggregated into a fused diagnostic representation before tree-based classification:

$$h = \sum w_r h_r \tag{5}$$

3.3. Degradation Probability and Review Priority

The LightGBM branch estimates the degradation probability from the fused representation. The probability output is used to determine whether the stack is normal, activation-loss dominated, ohmic-growth dominated, mass-transfer limited or affected by mixed degradation.

$$p_c = GBDT_c(h) \tag{6}$$

After the degradation probability is obtained, a review priority score is calculated to support maintenance decisions. This score includes uncertainty, pressure reliability and degradation severity, so that uncertain but severe windows can be sent to manual review.

$$R = a(1 - p_c) + bP_v + cD_c \tag{7}$$

The final output includes degradation class, dominant evidence, confidence level and maintenance priority. A high-confidence ohmic-growth result can be scheduled for resistance-related inspection, while a low-confidence mass-transfer warning may require pressure-channel verification before maintenance action. This output structure is intended to reduce unnecessary shutdowns and improve the traceability of stack degradation diagnosis.

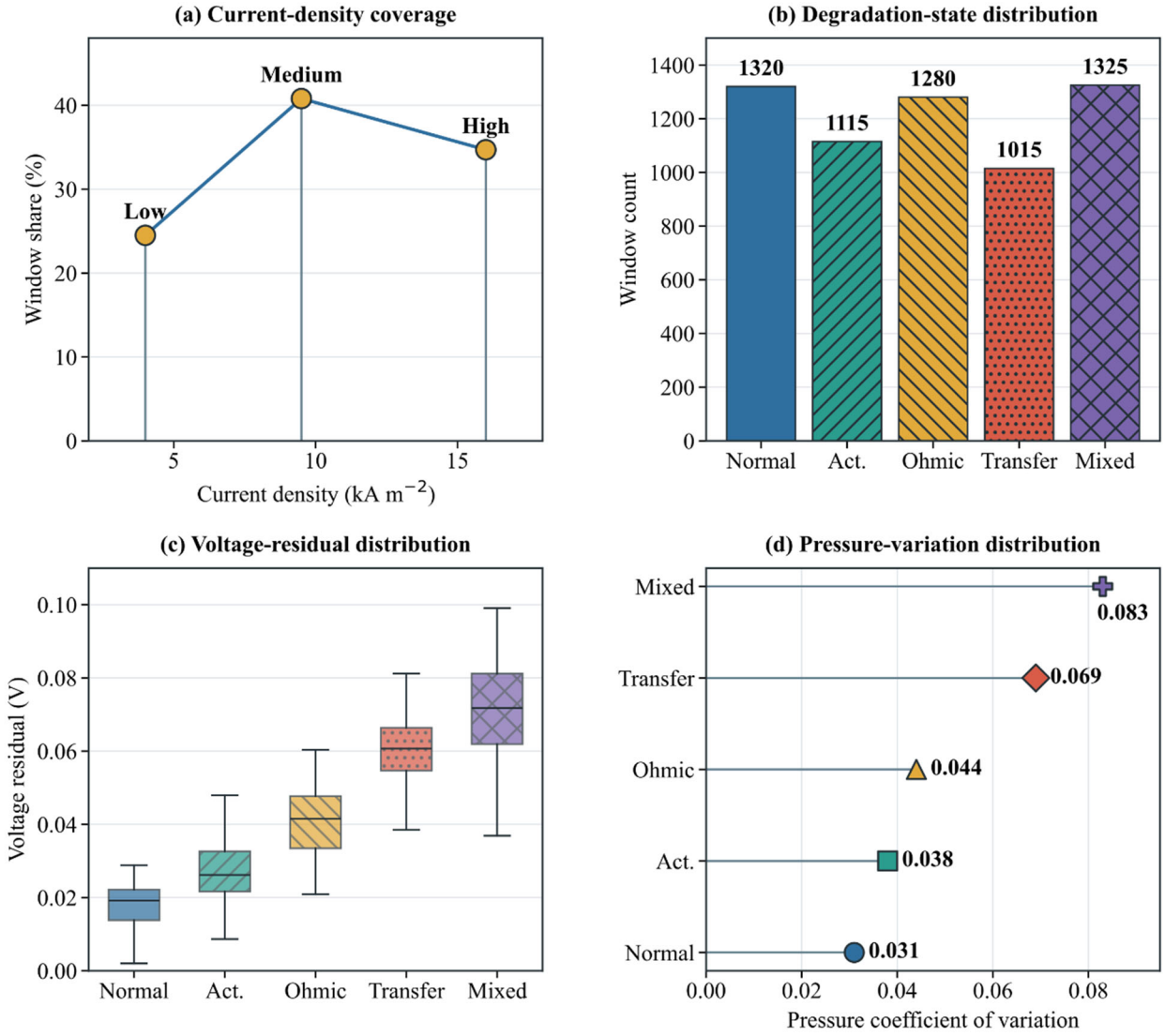


Figure 3. Dataset Structure and Degradation Label Distribution. (a) Current-density coverage. (b) Degradation-state distribution. (c) Voltage-residual distribution. (d) Pressure-variation distribution.

4. Results and Discussion

4.1. Dataset Structure and Degradation Label Distribution

The evaluation set consists of simulated stack monitoring windows for normal operation, activation-loss growth, ohmic-resistance increase, mass-transfer risk and mixed degradation [17]. Figure 3 shows the feature distribution and degradation labels. Figure 3(a) shows current-density coverage across low-, medium- and high-operating regions, which is necessary because different degradation mechanisms become visible in different curve sections. Figure 3(b) presents the degradation-state distribution and indicates that mixed degradation is not artificially overexpanded. Figure 3(c) shows voltage-residual distribution, where ohmic-growth and mixed states have larger residuals. Figure 3(d) shows stronger pressure variation in mass-transfer-risk windows. curve sweeps and dynamic pressure records are jointly constructed so that each degradation label contains electrochemical and transport-related evidence.

According to the distribution results, degradation diagnosis cannot be performed using a single voltage threshold. Some activation-loss windows show moderate voltage increase but a distinct low-current slope change, whereas some pressure-disturbed windows show large pressure fluctuation without long-term voltage drift. If only voltage residuals are used, transient flow disturbance may be interpreted as degradation; if only pressure fluctuation is used, ohmic growth may be ignored. Therefore, the feature system combines curve-level and pressure-level evidence [18]. Macro-F1 and class-wise recall are used to avoid a normal-class bias and to reveal whether the model can identify mixed degradation and mass-transfer risk with sufficient reliability for engineering review.

4.2. Diagnostic Performance and Baseline Comparison

Table 1 shows the operating data statistics of different degradation states. The table is as follows: Window number, Mean voltage, Curve-slope rise, Pressure coefficient of variation and Label ratio. According to the above data, mixed-degradation windows have the highest voltage and pressure fluctuations; activation-loss windows are mainly different in the low-current slope. As shown in the table, the mass-transfer-risk windows are linked to larger pressure fluctuations, but not all of them have a high voltage. The above data show that a multi-feature diagnostic model is required [19]. A compact index of dynamic instability is provided by the pressure coefficient of variation. There is no drop, or we would not know when to be cautious in observing the curve. For instance, a voltage rise at a low-pressure variation is more likely to be caused by electrochemical degradation; a voltage rise at a high-pressure variation may require verification of flow control and gas removal.

Table 1. Operating Data Statistics for Stack Degradation Diagnosis

Degradation State	Windows	Mean Voltage	Slope Rise	Pressure CV	Label Ratio
Normal	1320	1.812 V	0.018	0.031	21.8%
Activation Loss	1115	1.846 V	0.026	0.038	18.4%
Ohmic Growth	1280	1.889 V	0.041	0.044	21.1%
Mass-Transfer Risk	1015	1.935 V	0.057	0.069	16.8%
Mixed Degradation	1325	1.976 V	0.071	0.083	21.9%

The primary diagnostic results are shown in Figure 4. Figure 4(a) presents the confusion matrix, indicating that normal and ohmic-growth states are well separated. Figure 4(b) reports class-wise recall, where mixed-degradation recall is higher than the baseline. Figure 4(c) presents the probability calibration plot and shows that low-confidence samples cluster near state boundaries. Figure 4(d) presents the dominant-evidence distribution, where voltage slope, pressure variation and curve residual contribute differently across degradation classes. Therefore, the model provides both a diagnostic result and supporting evidence [20]. The calibration behavior is important because uncertain pressure-disturbed windows should be routed to review or remeasurement rather than immediately converted into repair commands.

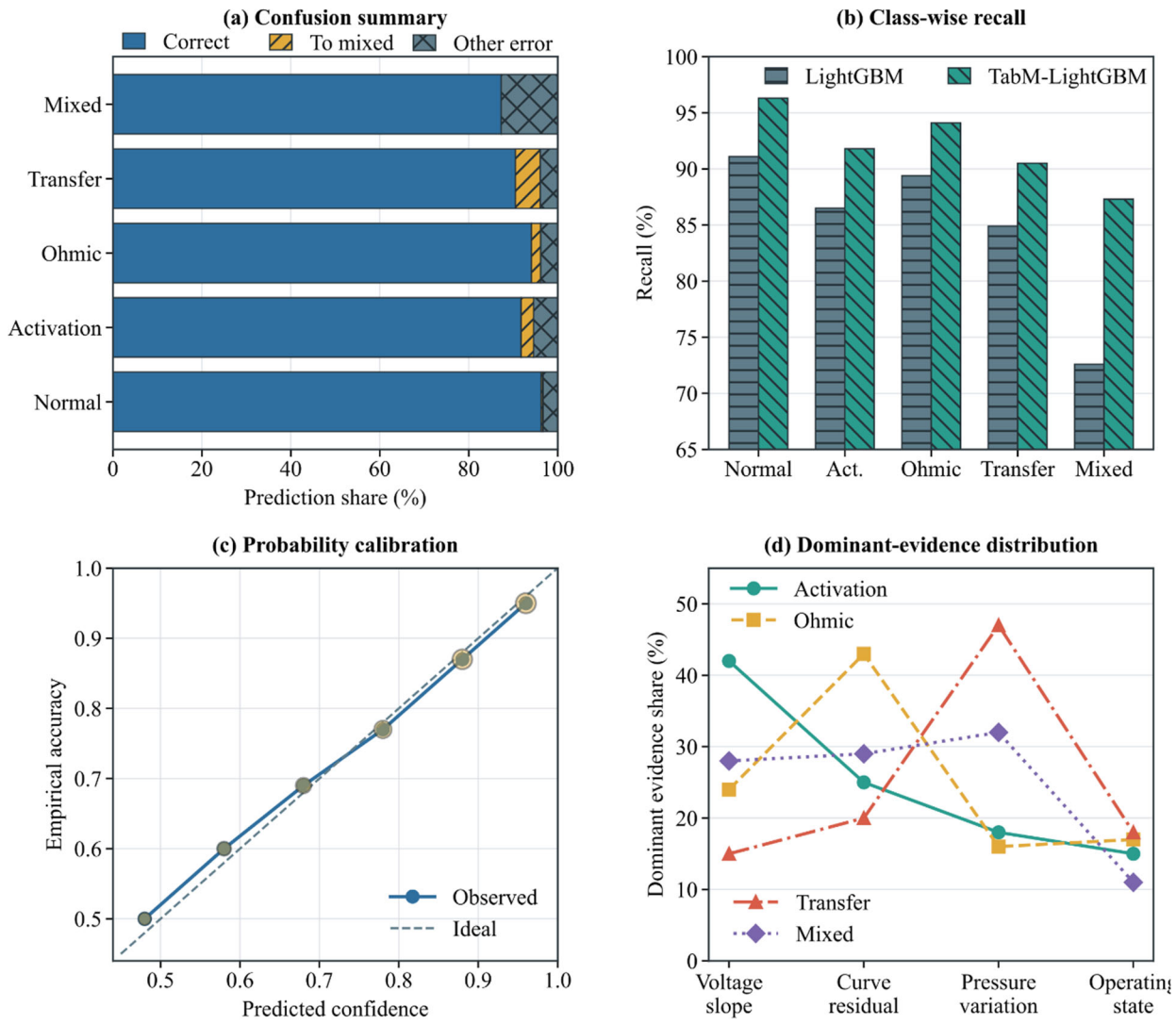


Figure 4. Degradation Diagnosis Results. (a) Confusion matrix. (b) Class-wise recall. (c) Probability calibration. (d) Dominant-evidence distribution.

The improved performance is due to the combination of representation learning and gradient boosting. TabM builds several views of the structured feature set, allowing relationships among polarization slope, voltage residual and pressure variation to be expressed before tree splitting. LightGBM then applies efficient boosting and interpretable split-based decision learning. This hybrid design is more flexible than manual threshold setting and more transparent than a fully opaque neural classifier [21]. For example, the meaning of pressure variation differs between medium-current and high-current regions, and the meaning of voltage residual differs between activation-loss and ohmic-growth classes. TabM represents these interactions, while LightGBM preserves feature-importance evidence for engineering audit.

Figure 5 compares the proposed model with logistic regression, random forest, standard LightGBM and compact multilayer perceptron baselines. Figure 5(a) shows that TabM-LightGBM achieves the best overall accuracy. Figure 5(b) shows higher macro-F1 for rare degradation states, indicating improved balance across classes. Figure 5(c) evaluates high-current windows and mass-transfer risk, where pressure behavior becomes especially informative. Figure 5(d) presents mixed-degradation recall and shows that the proposed method reduces confusion between ohmic growth and pressure-related limitation. The comparison demonstrates that accuracy alone is insufficient for maintenance support because rare mixed states can carry greater engineering risk [22].

4.3. Robustness, Interpretability and Engineering Review

Robustness is tested because electrolyzer operating data may include sensor drift, purge disturbance and load transition. Figure 6 presents the disturbance evaluation. Figure 6(a) reports diagnostic accuracy under voltage noise, pressure noise and combined disturbance. Figure 6(b) reports the missing-feature test, where pressure indicators or curve-slope features are masked during inference. The proposed model degrades gradually rather than collapsing when one evidence source becomes unreliable. This behavior is required for plant operation because diagnosis cannot

assume perfect data quality [23]. Disturbance robustness therefore reflects tolerance to temporary acquisition problems, while missing-feature robustness reflects whether the model can still prioritize severe windows when one source of evidence is unavailable.

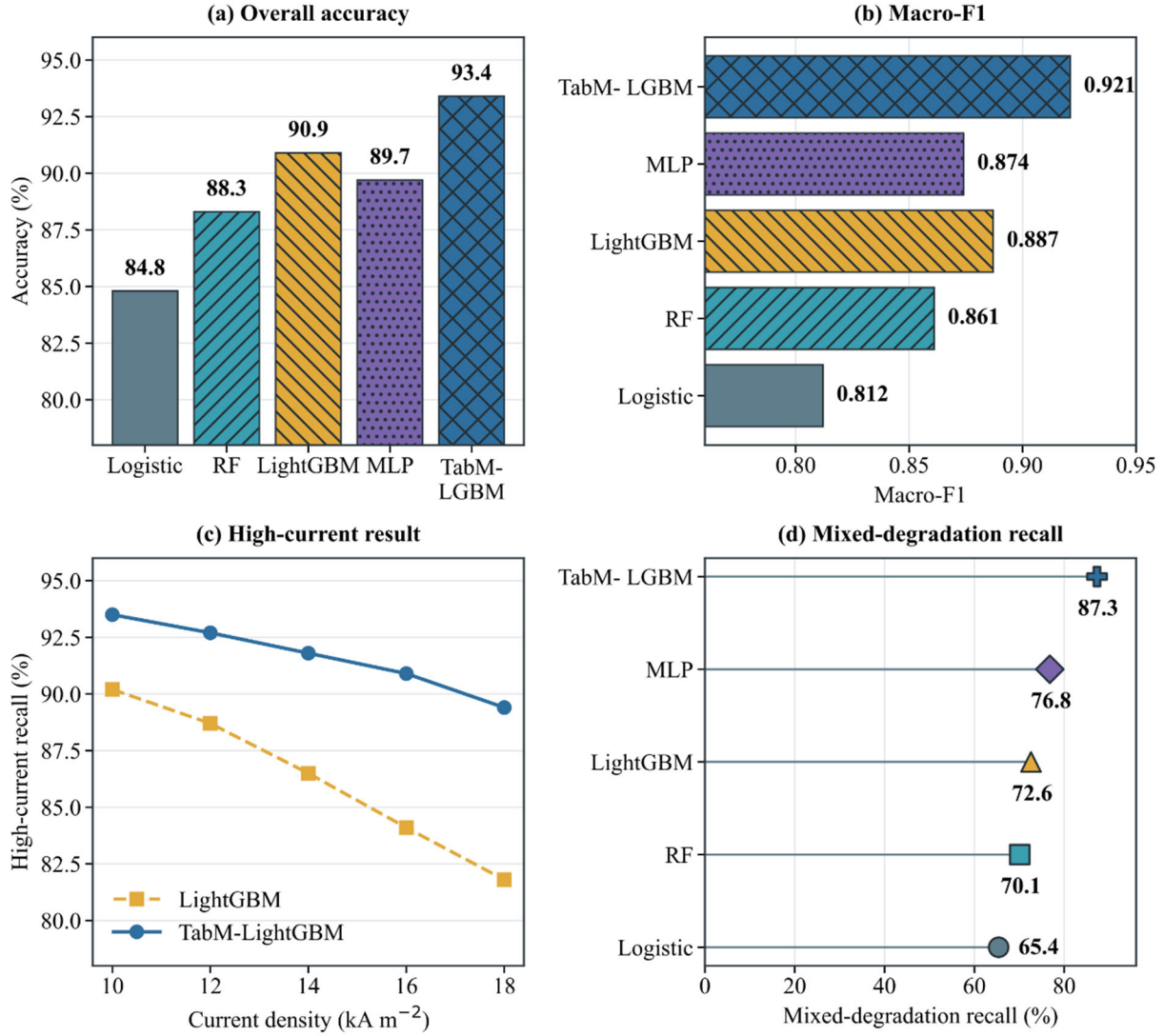


Figure 5. Baseline Model Comparison. (a) Overall accuracy. (b) Macro-F1. (c) High-current result. (d) Mixed-degradation recall.

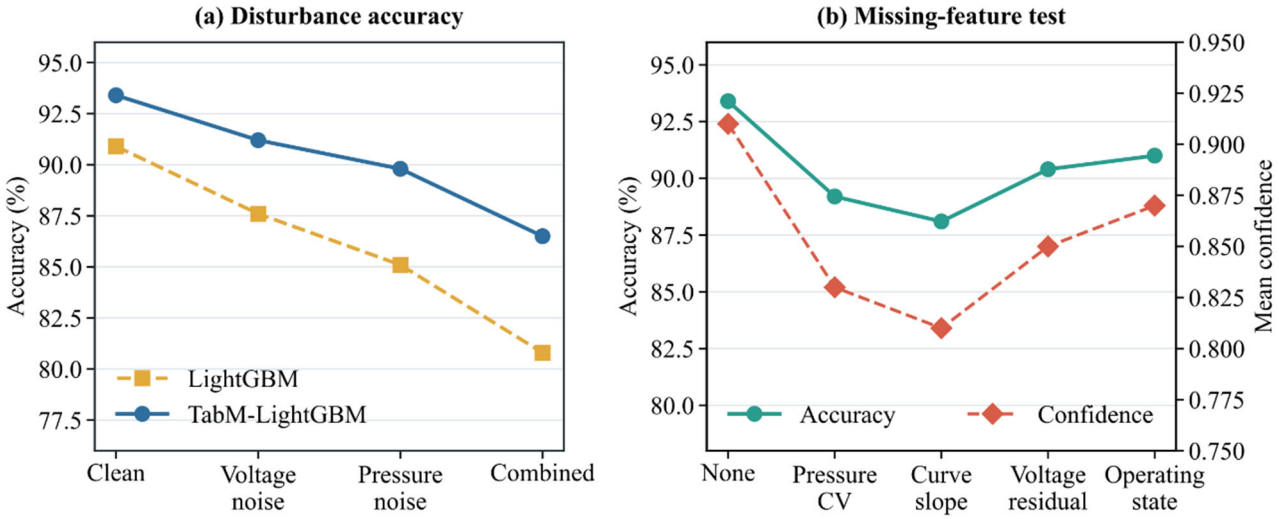


Figure 6. Robustness under Disturbance and Missing Features. (a) Disturbance accuracy. (b) Missing-feature test.

Interpretability is evaluated by examining how feature contributions change across prediction cases. Figure 7 presents the explanation results. Figure 7(a) shows feature importance, where high-current voltage residual, middle-region slope and pressure coefficient of variation are the main contributors. Figure 7(b) shows misclassified samples, most of which occur between mass-transfer risk and mixed degradation. These classes share pressure fluctuation and high-current voltage rise, so the dominant-evidence output should be treated as an audit signal rather than proof of a single electrochemical mechanism. The explanation output can inform whether engineers should inspect the stack, verify gas-water management or repeat the polarization test [24].

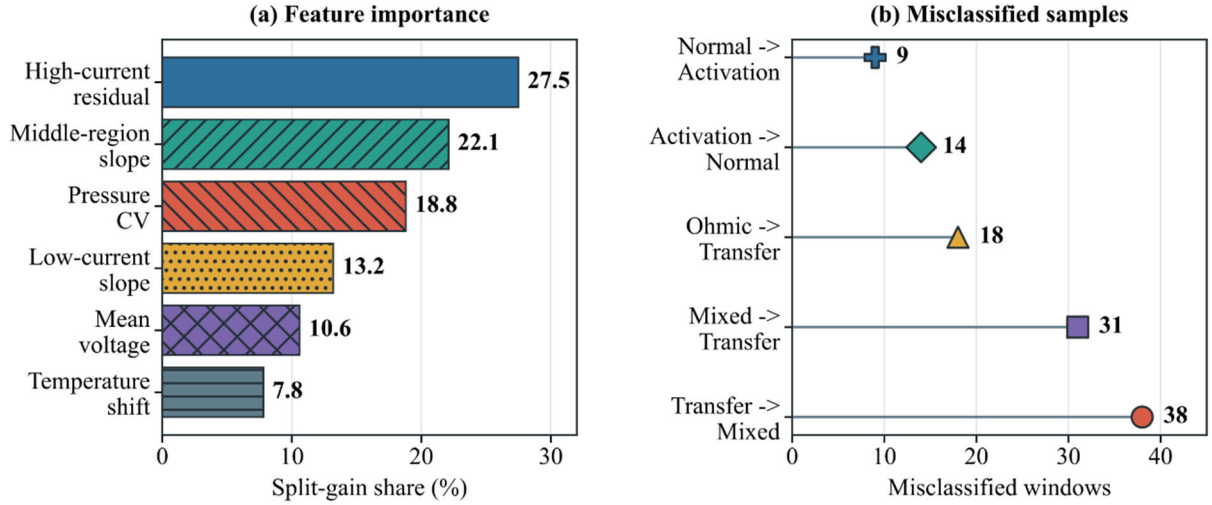


Figure 7. Feature Contribution and Error Analysis. (a) Feature importance. (b) Misclassified samples.

Engineering review outputs are shown in Figure 8. Figure 8(a) shows immediate inspection, scheduled monitoring, remeasurement and low-risk recording groups. Figure 8(b) shows the maintenance-evidence structure, including degradation class, confidence, dominant feature group and pressure reliability. A single class label is insufficient because operators need to know whether a warning is supported by stable electrochemical evidence or by uncertain pressure fluctuation [25]. High-confidence and high-severity cases are routed to immediate inspection, stable moderate-risk cases are routed to scheduled monitoring, uncertain high-severity cases are routed to remeasurement, and normal or weak low-confidence fluctuations are routed to low-risk recording.

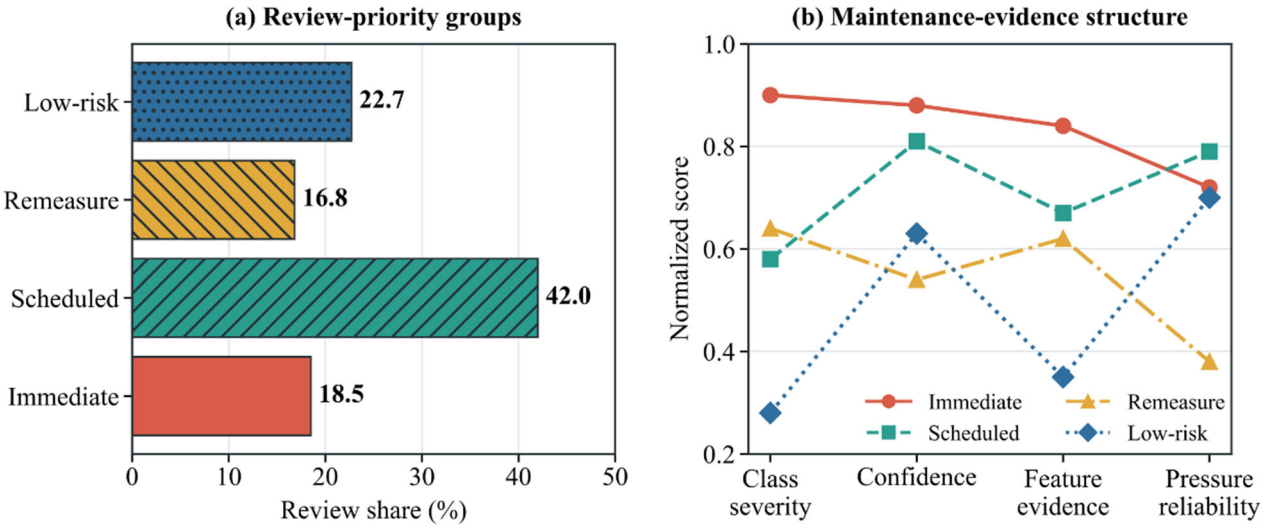


Figure 8. Engineering Review and Maintenance Output. (a) Review-priority groups. (b) Maintenance-evidence structure.

The proposed model can also be used in staged deployment. In the early stage, it can prioritize suspicious windows and reduce the volume of manual examination. After verified maintenance records are collected, plant-specific confidence thresholds can be calibrated. Predictive-maintenance research indicates that model governance should be linked to maintenance feedback rather than treated as a fixed classifier [26]. Thus, the proposed method outputs review priority and evidence descriptors that can be compared with subsequent maintenance records [27]. If many remeasurement cases are later confirmed as normal pressure disturbances, the pressure-reliability threshold should be adjusted; if mixed degradation is repeatedly confirmed after low-confidence warnings, sensitivity to combined curve-pressure evidence should be increased.

Another issue is that degradation states may not remain independent during prolonged operation. Activation loss may be followed by ohmic growth, and pressure instability may contribute to local dry-out or mass-transfer limitation. The dominant state should therefore be understood as an engineering summary rather than a complete electrochemical mechanism [28]. The current study is limited by simulated monitoring windows, simplified pressure features and the absence of all stack designs, membrane materials, catalyst systems and plant-control strategies. Future work should incorporate extended field data, electrochemical impedance characteristics, stack teardown observations and maintenance inspection records [29]. Field validation will improve degradation-threshold calibration and reduce false-positive maintenance decisions [30]. Additional cross-site tests are required before the model can be generalized for stack health evaluation [31].

5. Conclusion

This study proposed a TabM-LightGBM model for electrolyzer stack degradation diagnosis based on polarization curves and pressure data, and the overall conclusion is that curve-region evidence, pressure reliability and calibrated tabular learning should be evaluated jointly for reliable maintenance support. First, the constructed monitoring dataset shows that the proposed model achieves 93.4% overall accuracy, a macro-F1 score of 0.921 and an expected calibration error of 0.048, while improving mixed-degradation recall by 14.7% and reducing the false warning rate under pressure disturbance by 11.2% compared with a standard LightGBM baseline. Second, the comparative results show that TabM representation learning improves the modeling of interactions among voltage residuals, curve slopes and pressure variation, while LightGBM preserves interpretable decision evidence for engineering review. Third, the principal limitation is that the evaluation relies on simulated monitoring windows and simplified pressure features, so long-term field data, EIS, stack disassembly observations and maintenance records are required for deployment validation.

The proposed output separates degradation class, confidence level, dominant evidence and maintenance priority. This structure supports immediate inspection for high-confidence severe degradation, scheduled monitoring for moderate-risk windows, remeasurement for uncertain pressure-affected cases and low-risk recording for weak fluctuations. The Pressure Reliability Index prevents temporary pressure pulsations from being interpreted as confirmed mass-transfer degradation. The model therefore supports maintenance planning while preserving expert review and later audit.

Future work should incorporate long-term stack records from different operating strategies, membrane materials, catalyst systems and balance-of-plant configurations. Field deployment should also include EIS, stack disassembly observations, maintenance work orders and pressure-channel quality records. These data can calibrate thresholds for different stack types and operating hours while preserving traceable evidence for pressure-channel verification, curve remeasurement, trend monitoring and maintenance scheduling.

6. Declarations

6.1. Author Contributions

Hugo Lefèvre contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision, project administration, and funding acquisition. Louis Bernard contributes to validation, analysis, investigation, data collection, draft preparation. All authors have read and agreed with the manuscript before its submission and publication.

6.2. Funding and Acknowledgments

This research received no specific financial support from any funding agency.

6.3. Institutional Review Board Statement

Not applicable.

6.4. Informed Consent Statement

Not applicable.

6.5. Data Availability

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

6.7. Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors utilized DeepL for linguistic polishing, logical organization and draft revision. The generative AI tool was not involved in research design, data collection, data analysis, or the equation of core conclusions. All outputs generated by AI were reviewed, revised and verified manually by the authors. The authors take full responsibility for all content of this manuscript.

7. References

[1] Wang, C. R., Stansberry, J. M., Mukundan, R., Chang, H. M. J., Kulkarni, D., Park, A. M., ... & Zenyuk, I. V. (2025). Proton exchange membrane (PEM) water electrolysis: cell-level considerations for gigawatt-scale deployment. Chemical Reviews, 125(3), 1257.https://doi.org/10.1021/acs.chemrev.3c00904

[2] Wallnöfer-Ogris, E., Schuller, R., Nefedov, A., Bernhardt, M., & Krammer, A. (2024). A review on understanding and identifying degradation mechanisms in PEM water electrolysis cells: Insights for stack application, development, and research. International Journal of Hydrogen Energy, 65, 381-397. https://doi.org/10.1016/j.ijhydene.2024.04.017

[3] Uysal, S. B., Kıstı, M., & Kaya, M. F. (2024). Comprehensive analysis of flow dynamics and pressure variations in PEM water electrolyzers: Experimental and computational insights into gas diffusion electrodes. Fuel, 371, 132092. https://doi.org/10.1016/j.fuel.2024.132092

[4] Dittmar, F., Agarwal, H., & Tübke, J. (2025). Prognostics and health management (PHM) of proton exchange membrane water electrolyzers: A review-based guideline. International Journal of Hydrogen Energy, 106, 806-824. https://doi.org/10.1016/j.ijhydene.2025.01.443

[5] Chen, X., Rex, A., Woelke, J., Eckert, C., Bensmann, B., Hanke-Rauschenbach, R., & Geyer, P. (2024). Machine learning in proton exchange membrane water electrolysis: A knowledge-integrated framework. Applied Energy, 371, 123550. https://doi.org/10.1016/j.apenergy.2024.123550

[6] Borisov, V., Leemann, T., Seßler, K., Haug, J., Pawelczyk, M., & Kasneci, G. (2024). Deep neural networks and tabular data: A survey. IEEE Transactions on Neural Networks and Learning Systems, 35(6), 7499-7519. https://doi.org/10.1109/TNNLS.2022.3229161

[7] Miličić, T., Muthunayakage, K., Vũ, T. H., Ritschel, T. K. S., Živković, L. A., & Vidaković-Koch, T. (2024). The nonlinear frequency response method for the diagnosis of PEM water electrolyzer performance. Chemical Engineering Journal, 496, 153889. https://doi.org/10.1016/j.cej.2024.153889

[8] Xu, B., Ma, W., Wu, W., Wang, Y., Yang, Y., Li, J., Zhu, X., & Liao, Q. (2024). Degradation prediction of PEM water electrolyzer under constant and start-stop loads based on CNN-LSTM. Energy and AI, 18, 100420. https://doi.org/10.1016/j.egyai.2024.100420

[9] Huang, M., Tao, J., Ma, L., Zhuang, X., Wang, H., Tao, H. B., & Zheng, N. (2024). Is it stable for the asymmetric pressure operation of PEM water electrolyzer? Chemical Engineering Journal, 498, 155209. https://doi.org/10.1016/j.cej.2024.155209

[10] Lee, C. Y., Chen, C. H., Chuang, S. M., Dai, C. L., Lai, B. J., Chen, S. Y., & Hsu, J. Y. (2024). Real-time data acquisition inside high-pressure PEM water electrolyzer. Sensors and Actuators A: Physical, 372, 115318. https://doi.org/10.1016/j.sna.2024.115318

[11] Gorishniy, Y., Kotelnikov, A., & Babenko, A. (2025). TabM: Advancing tabular deep learning with parameter-efficient ensembling. The Thirteenth International Conference on Learning Representations. https://openreview.net/forum?id=pIZxEOZCId

[12] Li, S., Jin, N., Dogani, A., Yang, Y., Zhang, M., & Gu, X. (2024). Enhancing LightGBM for industrial fault warning: An innovative hybrid algorithm. Processes, 12(1), 221. https://doi.org/10.3390/pr12010221

[13] Hollmann, N., Müller, S., Purucker, L., Krishnakumar, A., Körfer, M., Hoo, S. B., Schirrmeister, R. T., & Hutter, F. (2025). Accurate predictions on small data with a tabular foundation model. Nature, 637, 319-326. https://doi.org/10.1038/s41586-024-08328-6

[14] Albadwi, A., Selçuklu, S. B., & Kaya, M. F. (2024). Machine learning applications on proton exchange membrane water electrolyzers: A component-level overview. International Journal of Hydrogen Energy, 94, 806-828. https://doi.org/10.1016/j.ijhydene.2024.11.188

[15] Lin, Y. H., & Li, G. H. (2024). Uncertainty-aware fault diagnosis under calibration. IEEE Transactions on Systems, Man, and

Cybernetics: Systems, 54(10), 6469-6481. https://doi.org/10.1109/TSMC.2024.3427345

[16] Urbano, E., Pahon, E., Yousfi-Steiner, N., & Guillou, M. (2024). Accelerated stress testing in proton exchange membrane water electrolysis: Critical review. Journal of Power Sources, 623, 235451. https://doi.org/10.1016/j.jpowsour.2024.235451

[17] Sayed-Ahmed, H., Keller, R., Toldy, A., Müller, M., Froning, D., & Santasalo-Aarnio, A. (2026). Data-driven degradation analysis of proton exchange membrane electrolyzer stacks under dynamic load operation. Energy and AI, 25, 100790. https://doi.org/10.1016/j.egyai.2026.100790

[18] Hayatzadeh, A., Fattahi, M., & Rezaveisi, A. (2024). Machine learning algorithms for operating parameters predictions in proton exchange membrane water electrolyzers: Anode side catalyst. International Journal of Hydrogen Energy, 56, 302-314. https://doi.org/10.1016/j.ijhydene.2023.12.149

[19] Li, H., Jiao, J., Liu, Z., Lin, J., Zhang, T., & Liu, H. (2025). Trustworthy Bayesian deep learning framework for uncertainty quantification and confidence calibration: Application in machinery fault diagnosis. Reliability Engineering & System Safety, 255, 110657. https://doi.org/10.1016/j.ress.2024.110657

[20] Cação, J., Santos, J., & Antunes, M. (2025). Explainable AI for industrial fault diagnosis: A systematic review. Journal of Industrial Information Integration, 47, 100905. https://doi.org/10.1016/j.jii.2025.100905

[21] Gao, F., Zhong, W., Jiang, Q., Peng, X., & Li, Z. (2024). OWFD-UCPM: An open-world fault diagnosis scheme based on uncertainty calibration and prototype management. Knowledge-Based Systems, 286, 111403. https://doi.org/10.1016/j.knosys.2024.111403

[22] Li, T., Chen, J., Liu, T., Sun, C., Zhao, Z., Chen, X., & Yan, R. (2026). Explainable artificial intelligence based intelligent fault diagnosis: A systematic review from applications to insights. Reliability Engineering & System Safety, 267, 111935. https://doi.org/10.1016/j.ress.2025.111935

[23] Vashishtha, G., Chauhan, S., Sehri, M., Zimroz, R., Dumond, P., Kumar, R., & Gupta, M. K. (2025). A roadmap to fault diagnosis of industrial machines via machine learning: A brief review. Measurement, 242, 116216. https://doi.org/10.1016/j.measurement.2024.116216

[24] Chen, G., Yuan, J., Zhang, Y., Zhu, H., Huang, R., Wang, F., & Li, W. (2024). Enhancing reliability through interpretability: A comprehensive survey of interpretable intelligent fault diagnosis in rotating machinery. IEEE Access, 12, 103348-103379. https://doi.org/10.1109/ACCESS.2024.3430010

[25] Hayat, B., Ahmad, S., Shahid, M. A., Khan, A., Islam, M. R., Sayeed, M. S., & Ullah, Y. (2026). A process-guided uncertainty-aware deep learning framework for reliable and interpretable industrial fault diagnosis. PLOS ONE, 21(6), e0349385. https://doi.org/10.1371/journal.pone.0349385

[26] Meitz, L., Senge, J., Wagenhals, T., Schöler, T., Hähner, J., Edinger, J., & Krupitzer, C. (2025). A literature review framework and open research challenges for predictive maintenance in Industry 4.0. Computers & Industrial Engineering, 206, 111193. https://doi.org/10.1016/j.cie.2025.111193

[27] Shaalan, M., Ge, M., Hussain, M., & Rekik, Y. (2024). An explainable predictive maintenance strategy for multi-fault diagnosis of rotating machines using multi-sensor data fusion. Decision Analytics Journal, 10, 100425. https://doi.org/10.1016/j.dajour.2024.100425

[28] Bernhardt, A., Lange, K., Göckeritz, R., Grunwald, E., Hähnel, A., Müller, A., Naumann, V., Rasuli, S., Münchgesang, W., Pollet, B. G., Witte-Bodnar, K., & Ilse, K. (2026). Degradation mechanisms in PEM water electrolysis: Diagnosis and impact. Advanced Materials Technologies, 11(11), e02419. https://doi.org/10.1002/admt.202502419

[29] He, D., Chen, K., Chen, W., Luo, Z., Xiong, Z., Zou, G., Li, G., & Chen, B. (2025). Experimental study on the degradation of proton exchange membrane water electrolyzers under different dynamic loads and flow field configurations. Journal of Power Sources, 641, 236844. https://doi.org/10.1016/j.jpowsour.2025.236844

[30] Javed, A., Wolf, N. L., Meyer, F., Treutlein, L., Kungl, H., Karl, A., Jodat, E., & Eichel, R. A. (2025). Exploring the state-of-operation of proton exchange membrane electrolyzers. International Journal of Hydrogen Energy, 98, 280-294. https://doi.org/10.1016/j.ijhydene.2024.12.055

[31] Dittmar, F., Lickert, T., Smolinka, T., Pinkwart, K., & Tübke, J. (2026). Development and validation of a layer-unspecific physical model for voltage degradation to predict the remaining useful life of proton exchange membrane water electrolyzers. Journal of Power Sources, 668, 239122. https://doi.org/10.1016/j.jpowsour.2025.239122