Proc. of the European Safety and Reliability Conference (ESREL 2026)
Edited by José C. Matos, Paulo B. Lourenço, Michael Beer, Daniel Oliveira,
Edoardo Patelli, Hélder S. Sousa, Rui A. Silva and Monica Santamaria
 Published by Research Publishing, Singapore.
doi: 10.3850/ESREL2026061419_esrel26-p27872-cd

Theory-guided data-driven prediction of PEMEC degradation and durability for reliable green hydrogen production

Thomas Waite, Shahin Alipour Bonab, Mohsen Abdolahi, Wenjuan Song, David Flynn, and Mohammad Yazdani-Asrami*

CryoElectric Research Lab, Propulsion, Electrification & Superconductivity group, Autonomous Systems and Connectivity division, James Watt School of Engineering, University of Glasgow, Glasgow G12 8QQ, Scotland, United Kingdom

**Corresponding author's email: mohammad.yazdani-asrami@glasgow.ac.uk*

Accurately predicting the performance degradation of Proton Exchange Membrane Electrolyzer (PEMEC) stacks is critical for optimizing the techno-economic viability of green hydrogen (GH_2) production. Typical approaches for PEMEC degradation modelling choose between physics-based models that are computationally expensive and simplify the interrelated degradation mechanisms, or "black-box" machine learning models that often lack interpretability and generalizability. This work introduces a novel hybrid framework to overcome these limitations: a Theory-Guided Data-Driven (TGDD) model for predicting PEMEC performance degradation. The "theory-guided" component models known current-voltage relationships in electrolyzers, with the data-driven component predicting the (otherwise unknown) exchange current and resistance. This structure makes accurate, generalized, and physically-grounded predictions of PEMEC response curves, enabling simultaneous predictions of operating and reference points to facilitate degradation prognosis in optimal control and state-of-health monitoring. The TGDD model was developed and validated using comprehensive, long-term operational data obtained from PEMECs under various operating conditions. The model demonstrates high accuracy, with an R^2 of more than 94%. Furthermore, the model exhibits enhanced robustness, interpolation, and explainability. Superior interpretability makes the TGDD framework a powerful tool for prognostics and health management in PEMEC systems, enabling future development of intelligent control strategies to mitigate degradation and accelerate the commercialization of GH_2.

Keywords: Control, Degradation, Green Hydrogen, Machine Learning, PEMEC, Prognosis, Safety.

1. Introduction

The global demand to accelerate affordable decarbonization and mitigate climate change has motivated increased investment into Cyber Physical Energy Systems (CPES) (Norbu et al., 2024). CPES seek to optimize the integration of gas, electricity networks, and associated low carbon technologies via digital tools and communication networks that can efficiently and dynamically respond to energy demand (Cali et al., 2025). The integration of Green hydrogen (GH_2) into CPES has been identified as a candidate for decarbonizing the energy demand in hard to abate sectors such as steel manufacturing, aviation, and heavy transport (Zhang et al., 2025). It can also reduce the curtailment of renewable electricity, which in 2025 represented £1.5 billion for the UK alone (Andoni et al., 2025). However, the GH_2 industry faces challenges to growth that are as much technological as they are financial. Especially at the industrial scale, water electrolysis technology is immature in comparison to the conventional (and highly emissive) methods of producing and exploiting H_2 or hydrocarbons. Additionally, the high costs of Polymer Electrolyte Membrane Electrolyzers (PEMECs) and industry-wide uncertainty over system lifetimes and degradation characteristics drive up the price of GH_2 to customers. This creates a negative feedback loop whereby a lack of GH_2 offtake disincentivizes further production (International Energy Agency, 2025). Predicting the degradation of PEMECs is thus important for improving competitiveness of GH_2.

PEMEC degradation is a combination of catalyst agglutination and loss; membrane thinning and thermomechanical damage; and titanium component oxidation and passivation (Khatib et al., 2019; Wallnöfer-Ogris et al., 2024). These processes interact with one another and are variably influenced by the operation, temperature, pressure, and system construction (Makhsoos et al., 2025; Rakousky et al., 2016). In the aggregate, this degradation lowers the power-to-hydrogen efficiency of the PEMEC over time, decreasing its financial viability and remaining useful life.

The complexity and nonlinearity of PEMEC degradation inhibits the development of predictive models for use in control, safety monitoring, and long-term financial forecasts. Numerical models based on physical laws may only capture this complexity to a limited extent, while also being inefficient, difficult to scale, and narrow in scope (Waite and Yazdani-Asrami, 2025). Alternatively, data-driven models can help deal with the nonlinearity of degradation. These approaches learn the complex relationships embedded in a data set to implicitly model the behavior of the underlying system. Studies so far have demonstrated promising results, with high accuracy when predicting electrolyzer cell voltage over long timespans (Alipour Bonab et al., 2025; Hayatzadeh et al., 2024; Xu et al., 2024). However, a major shortcoming of data-driven models is explainability and generalizability (Chen et al., 2024; Polo-Molina et al., 2025). Where numerical approaches explicitly define the relationship between input and output, data-driven models approximate the trends in a dataset as a black-box, with little to no physical basis for their predictions.

This paper proposes a hybrid Theory-Guided Data-Driven (TGDD) modelling framework for application in optimal control, safety monitoring, and maintenance planning. Herein, surrogate machine learning models, trained on experimental data from a wide variety of PEMECs, predict the unknown parameters within physically-grounded semi-empirical models for electrolyzer cell voltage. This approach combines the strengths of both modelling paradigms, enabling accurate cross-domain predictions of PEMEC degradation, reconstruction of polarization curves to simultaneously compare performance at operating and reference currents, and efficient, generalized predictions for heterogeneous PEMEC architectures and operating schema.

2. Methodology

The creation of the TGDD model is divided into three steps: curve fitting and parameter extraction, augmentation and preprocessing, and model training and validation. First, an empirical model for electrolysis cell voltage is fit to polarization data from experimental studies on PEMECs to identify underlying parameters for electrochemical activity and electrical resistance. Second, the fitted parameter values are interpolated over time to improve the model's ability to learn the long-term behavior. In the final step, a data-driven model is trained on the fitted and processed parameter values to predict system degradation within the two domains. **Figure 1** shows the outline of this process, with an illustration of the two parameters predicted by the TGDD model.

2.1. *Description of the Data Set*

A total of 14170 datapoints were collected from 35 experimental publications from the last decade on PEMECs of differing constructions and operating characteristics (Waite et al., 2025).

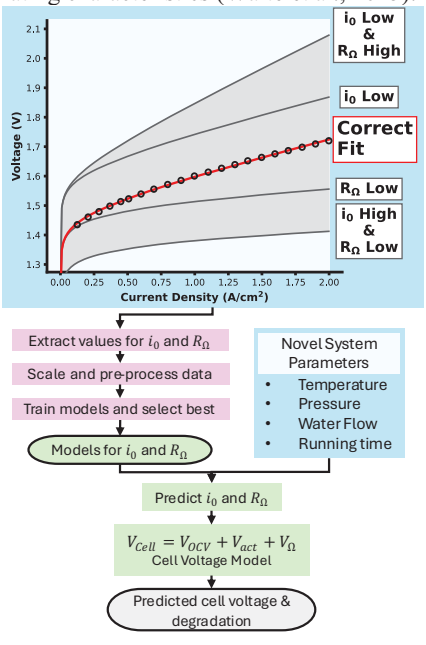


Fig. 1. Illustration of the modelling process

Table 1. Dataset features and values

Feature	Values	Units
Anode material	Ir, Ir-Black, IrO_2, IrO_2/TiO_2, IrOx, IrRu, IrRuOx, $Pt/Ce_{0.5}Zr_{0.5}O_2$, Pt-Black	-
Anode quantity	0.004 – 3.5	mg/cm^2
Cathode material	Pt, Pt/C, Pt/CB, Pt/NOCB, Pt/OCB, Pt-Black	-
Cathode quantity	0.02 - 3	mg/cm^2
Membrane material	Aquivion, Nafion, PFSA	-
Membrane thickness	50 – 178	µm
PTL coating	No, Yes	-
BPP coating	No, Yes	-
Surface area	4 – 600	cm^2
Temperature	20 – 90	°C
Burnt in at start	No, Yes	-
Pressure	1 – 30	Bar
Water flow rate	0.0139 – 50	mL/cm^2/min
Minimum current	0.0-4.0	A/cm^2
Average current	0.1-4.0	A/cm^2
Frequency	0-9700	Hz
Duty	0.2405-1.0	-
Cell voltage	1.326 – 3.094	V

Aiming for a robust, varied dataset, studies were chosen based on shared characteristics, ignoring one-of-a-kind experiments on novel catalysts.

Additionally, to maximize the model scope, the data comes from PEMECs under both steady-state and fluctuating loads. Variable operation is known to have adverse effects on the degradation of PEMECs, so several features have been included to describe dynamic operation. Cell voltages were extracted from 343 polarization curves, and information on the materials, geometries, and experimental variables were recorded from the respective methodological descriptions. **Table 1** provides an outline of the features and value ranges.

The recorded operating characteristics reflect the typical range for PEMECs. Cell temperature is usually about 80°C in experimental literature, but can be lower for industrial electrolyzers or higher to decrease the energy required to drive the reaction. The cathode pressure and water flow rates also demonstrate a reasonably wide range. Finally, a variety of operating points were represented, both static and dynamic and from low to high current densities.

It is important to note that collecting data from a range of experimental sources inherently encodes some inconsistencies. Ideally, a large enough data set will reflect a methodological average, but in practice, experiments vary with regards to duration of polarization measurements, purity of the catalytic coatings and feed water, and cell construction, among other aspects. Respectively, these impact the degree of relaxation of reversible voltage rise, catalyst durability and introduction of impurities, and mechanical stress on the components. While this inconsistency may affect accuracy, it improves the model's cross-experiment generalizability. As such, any data-driven model using experimental data should maximize the quantity of sources.

2.2. *Physical Model*

For a system at standard pressure, the minimum voltage required to be input to initiate electrolysis is the reversible voltage, experimentally defined by the following relationship to temperature (Falcão and Pinto, 2020):

$$V_{rev} = 1.229 - \left(0.9\left(T^{[K]} - 298.15\right)\right) \quad (1)$$

The open-circuit voltage is the adjusted minimum voltage required for the reaction in the presence of a pressure differential. This is described by a Nernst equation (Abomazid et al., 2022):

$$V_{OCV} = V_{rev} + \frac{RT}{2F}\ln\left(\frac{\left(P_{cathode}\sqrt{P_{anode}}\right)}{P_{anode}}\right) \quad (2)$$

where R is the gas constant, F is the Faraday constant, and P_x is the pressure in the anode or cathode chambers.

The electrodes of a PEMEC require a small excess current to activate the catalyst. This relationship is usually described via an approximation of the Butler-Volmer equation. For this research the activation overvoltage is simplified to a combined value representing both the anode and cathode electrodes (Abomazid et al., 2022):

$$V_{act} = \frac{RT}{2\alpha F}\operatorname{asinh}\left(\frac{i}{2i_0}\right) \quad (3)$$

where α is the charge-transfer coefficient (CTC) and i is the applied current density in A/cm^2. The optimal CTC value was determined to be 0.598.

The exchange current density, i_0, is a measure of the activity of the catalyst and is influenced by the particular reaction occurring on the surface, as well as the electrode-specific material choice, temperature, and geometry. The exchange current density of an electrolyzer is generally unknown until a model is fit to polarization data, and the value changes over time as the electrodes of a PEMEC degrade, as seen in **Figure 2**.

The material interfaces within the cell and the impedance of proton flow across the membrane give rise to an electrical resistance. The particular value of the resistance is another unknown parameter depending on the temperature, membrane hydration, material choice, etc. The voltage due to ohmic interactions is described as follows (Falcão and Pinto, 2020):

$$V_\Omega = A_{cell} i R_\Omega \quad (4)$$

where A_{cell} is the active area of the PEMEC in cm^2 and R_Ω is the ohmic resistance of the cell. Like the exchange current density, R_Ω also varies with time as the cell degrades.

The complete model for the PEMEC cell voltage is thus the sum of the component overvoltages:

$$V_{cell} = V_{OCV} + V_{act}\,(i_0) + V_\Omega\,(R_\Omega) \quad (5)$$

To extract values for i_0 and R_Ω, the above empirical model was fit to the collected polarization curves via nonlinear least-square optimization. The current density was maintained below 2.0 A/cm^2, as at higher currents additional mass transfer processes begin to affect the PEMEC performance. The fitting process resulted in 343 pairs of output parameters varying with the collected inputs. The fitting procedure returned a mean goodness-of-fit (R^2) for the whole dataset of 0.998. The results of this curve fitting is reported in **Table 2**.

2.3. *Parameter Extraction*

The changes in i_0 and R_Ω due to regular degradation are very gradual, occurring over hundreds of hours (Rakousky et al., 2016). So linear growth is a common assumption in the literature around PEMEC degradation (Keddar et al., 2022; Krenz et al., 2024; Lu et al., 2023). As such, after extracting the fit parameters, values at the beginnings and ends of measurements were linearly interpolated with 30 equally spaced intervals. The values for i_0 were finally converted to their base-10 logarithm as follows:

$$i_{0,log-transform} = \log_{10}(i_0) \quad (6)$$

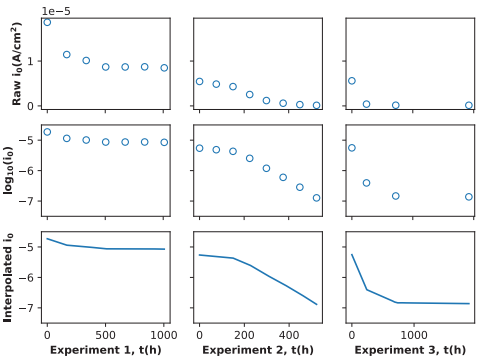


Fig. 2. Exchange current density data before and after processing

Table 2. Extracted Target Parameter Values

Parameter	Range	Units
Exchange current density (i_0)	[7.69e-10, 2.43e-5]	A/cm^2
Ohmic resistance (R_Ω)	[2.56e-4, 3.06e-1]	Ω

The impact of scaling and interpolation on the collected values for i_0 can be seen in **Figure 2**. In the top plot, most values are clustered around zero and at this scale it is difficult to differentiate trends in the data. The second and third plots show how transforming i_0 reveals the parameter's tendency to decay over time. This dramatically improves the ability of a machine learning algorithm to learn the behavior of the system.

2.4. *Data-Driven Model*

In this paper, eXtreme Gradient Boosting regression (XGBoost) was used to model the dataset. XGBoost is a high-performing machine learning ensemble algorithm based on tree boosting, which seeks to identify natural splits in the features of a dataset that maximize the predictive accuracy, successively splitting and narrowing the criteria to form a decision tree (Zhang et al., 2017). A large collection of these decision trees is sequentially generated by minimizing the errors of the prior tree, then the collection votes on the most likely output (Chen and Guestrin, 2016). XGBoost includes additional automated processes for feature selection, subsampling, and sparsity compensation that improve its ability to learn highly nonlinear patterns without overfitting to the training data.

The data sets were split in an 80:20 training/validation ratio, with the validation set used to evaluate the final model accuracy against unseen data. Despite the use of tree-based algorithms, the wide range of feature and target values makes pre-processing necessary. Categorical features were ordinally encoded. Numerical features were processed as follows:

$$x_{transformed} = \frac{x_n - \bar{x}}{\sigma} \quad (7)$$

where x_n is the n'th value of feature x, $\bar{x}$ is the mean of that feature, and σ is the standard deviation.

To maximize accuracy and prevent overfitting, a set of hyperparameters for model complexity and regularization were tuned for each of the cell voltage (V_{cell}), i_0, and R_Ω models. A randomized cross-validation search was carried out over the parameters and ranges detailed in **Table 3**. For each of the 100 randomly selected combinations of hyperparameters, a 5-fold cross validation was carried out on the training set. The highest-performing combination of hyperparameters was then selected and trained on the full 80% training set. Additionally, early stopping was used to ensure the models did not over-fit to the training data. The validation set was here used to check the training progress of the models, stopping early if they began to lose accuracy against the validation set. However, the models were never directly trained on this subset to avoid data leakage.

3. Results and Analysis

Three XGBoost models were developed for this research: one on the raw V_{cell} data, and two on the extracted parameters i_0 and R_Ω. The V_{cell} model represents a typical data-driven approach with minimal pre-processing or feature extraction. The other two predict the TGDD model parameters.

3.1. *Results of Model Training*

Hyperparameter tuning resulted in three high-performing models, with each scoring $R^2 \geq 0.958$ against unseen validation data. All three models were found to score optimally with tree depths at or below the XGBoost default of 5. This indicates the models are fairly conservative.

The parameters for minimum split weight and alpha regularization control how conservative the model is in both structure and output.

Table 3. XGBoost Optimal Hyperparameters

Parameter	Function	Search Range	V_{cell} Model	i_0 Model	R_Ω Model
max depth	Model complexity	[2,15]	4	5	5
minimum child weight	Model complexity	[1,12]	7	7	3
alpha	Regularization	[1e-8, 0.6]	6.2e-3	3.1e-5	1e-8
lambda	Regularization	[1,40]	5.3	29	1

Table 4. Results of Model Training

R^2 Value	V_{cell} Model	i_0 Model	R_Ω Model
Cross Validation	0.99178	0.97085	0.99969
Full Dataset	0.99449	0.97660	0.99989
Training Set	0.99524	0.98111	0.99995
Validation Set	0.99162	0.95822	0.99970
Average polarization curve accuracy	0.99149	0.94239	

In **Table 3**, it can be seen that the optimal R_Ω model has lower values for minimum child weight and alpha regularization than the other two. This implies a lower threshold for tree splits and a lower need to enforce correlation between outputs than in the V_{cell} or i_0 models.

Between the training and testing phases, the V_{cell} and R_Ω models tended to perform slightly better than the i_0 model. Validation set performance was over $R^2 = 0.99162$ for both V_{cell} and R_Ω models, compared to $R^2 = 0.95822$ for the i_0 predictions. Training predictions were more accurate across the board, but followed the same pattern. These scores reflect very good agreement between the predicted values and the dataset. The relatively lower scores for i_0 are likely structural: the values vary across several orders of magnitude, and even after preprocessing the nonlinearity may be harder to represent than with the more stable V_{cell} and R_Ω.

Recombining the parameter models into the full TGDD model reveals a slightly lower average R^2 accuracy than would be expected. To determine this value, R^2 scores were recorded for each of the polarization curves then the average was taken for the complete set of 343 curves. The average R^2 scores for the standard and TGDD models are included in the final row of **Table 4**.

Where the accuracy of the cell voltage model is largely in line with the previous validation results,

the TGDD score of $R^2 = 0.94239$ is markedly lower than that of either the i_0 or R_Ω models. The drop in accuracy compared to the previous results is likely due to the wide range of i_0 values being difficult to accurately represent. Despite the high model accuracy when predicting log-transformed values of i_0, the raw (untransformed) values experience greater relative errors. The i_0^{-1} term in equation 3 increases the impact of these errors, especially for values with a smaller magnitude.

3.2. *Analysis of Model Predictions*

Despite a slightly lower overall predictive accuracy, TGDD modelling has improved explainability over a purely data-driven approach. The model structure means that each output predicts a polarization curve rather than individual voltages, which enables simultaneous prediction of the operating point and a reference point. This is particularly useful for predictive control and state-of-health monitoring, where performance degradation is tracked relative to a static current level. In these cases, explainability and simultaneous predictions outweigh a slight loss in pure predictive accuracy.

Figure 3 demonstrates the TGDD model output against the cell voltage model and the collected data. Here, the cell voltage model slightly outperforms the TGDD model, but the latter more closely matches the contour of the original data. The accuracy of the cell voltage model varies between data points, where the TGDD model sustains a roughly constant level of error. This maintains the ground truth better and improves the quality of interpolation and extrapolation.

Figure 3 also displays the increased explainability of the TGDD approach. The model is additionally able to distinguish the source of predicted voltage rises. As the model provides values for i_0 and R_Ω, it is possible to extract the activation and ohmic overvoltages for any prediction.

Figure 4 illustrates the predicted operating point of a PEMEC run at 3 A/cm^2 for 4000 hours against experimental data (Padgett et al., 2025). The upper half of the plot demonstrates the model's close adherence to the experimental cell voltage data points. The lower half shows the predicted contributions of the overvoltages. The voltage rise from the activation losses is consistently larger than the ohmic overvoltage. However, both the activation and ohmic overvoltages steadily grow over the duration of the test, which corroborates the results from that paper.

This makes sense in the experiment's context. In (Padgett et al., 2025), the authors used a PEMEC with low catalyst loadings (0.41 mg_{Ir}/cm^2 and 0.13 mg_{Pt}/cm^2), and titanium components with noncorrosive platinum coatings. This resulted in durable titanium components with low contact resistances, and relatively weak catalyst coatings. The TGDD model reflects this relationship within the experiment, and could be used to make decisions about the cost—performance trade-off of using more catalytic material or running the electrolyzer at a different temperature.

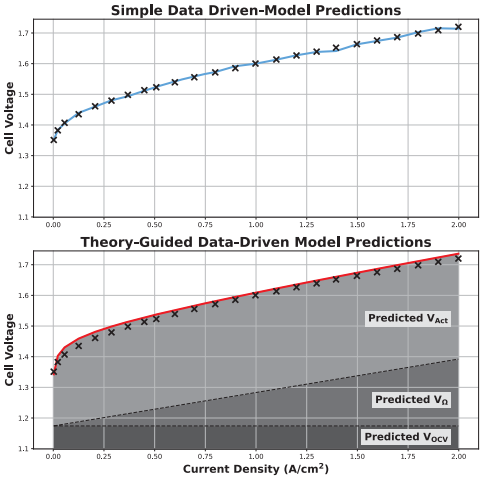


Fig. 3. Comparison of XGBoost modelling on raw cell voltage data vs on electrochemical parameters The TGDD model is slightly less accurate, but maintains the ground truth of the data better than Model 1 and predicts the component overvoltages.

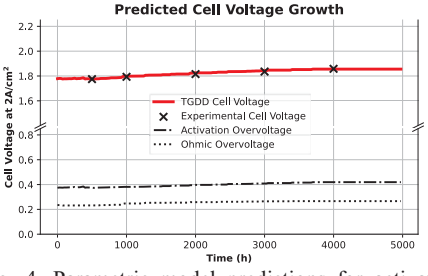
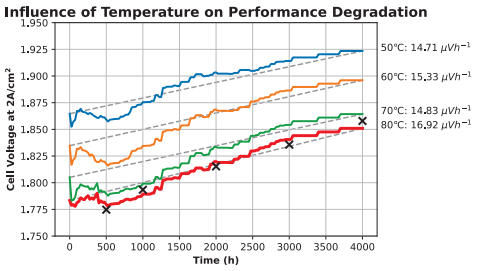
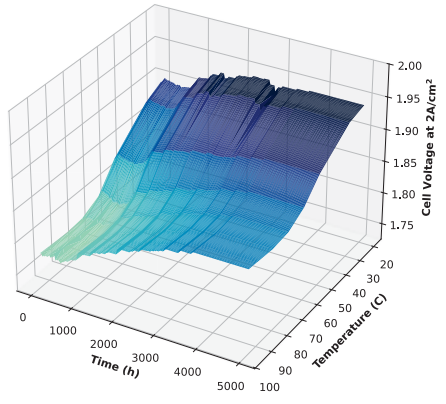


Fig. 4. Parametric model predictions for activation, ohmic, and cell voltages against experimental data



(a) 2D view



(b) 3D view

Fig. 5. Model Predictions of the impact temperature has on long-term degradation

Figure 5 demonstrates the ability of the TGDD model to interpolate within the data set and plausibly generalize PEMEC behavior to provide information for such a design choice. Here, the same setup for **Figure 4** was repeated across a range of common temperatures to evaluate its impact on degradation. Without additional experimental data the accuracy cannot be quantified, but qualitatively, the TGDD model captures a realistic trend: raising the operating temperature is expected to decrease the cell voltage while also increasing the rate of voltage rise (Bazarah et al., 2022; Chandesris et al., 2015). The ability to make such long-term predictions about the magnitude and source of degradation in PEMECs is invaluable for control and prognostics.

3.3. *Assumptions of the TGDD Model*

There were two major simplifications about the performance and degradation of PEMECs. Firstly, R_Ω and i_0 were assumed to vary linearly with time between recorded data points. While this mostly holds true for R_Ω, i_0 can be seen in **Figure 2** to decay quasi-exponentially. Linear interpolation thus slightly overestimates the real value of i_0 between data points, resulting in lower activation overvoltages. In cases where only two polarization curves are recorded with high intermediate degradation, the overestimate could be significant. As such, future TGDD developments should explore more detailed interpolation methods.

Secondly, the theoretical model combines the anode and cathode activation losses and ignores the effects of mass transfer. The former results in a single i_0 value describing the activity of both halves of the cell, meaning activation predictions cannot be precisely isolated. This is not necessarily a major detriment, however, as most degradation occurs in the anode. The latter ignores the diffusion overvoltage that arises at high currents due to an overabundance of protons for a limited amount of reaction sites. As such, the model was only trained on data up to 2.0A/cm^2. Ignoring mass transfer simplifies the model to two parameters and improves fitting consistency but also reduces its ability to represent high-current operation.

4. Conclusions

This research has demonstrated an approach that combines semi-empirical formulae for electrochemical interactions with data-driven modelling of the underlying parameters to create a theory-guided data-driven model (TGDD) model for the performance decay of PEMECs. The integration of domain knowledge with machine-learning dramatically improves the ability of data-driven models to interpolate accurate predictions for the losses of PEMECs and their growth over time. To develop the TGDD model for PEMEC performance degradation, a data set of 14170 entries, split into 343 polarization curves, was first collected from experimental literature recording the cell voltage of varied PEMECs. Electrochemical parameters for catalyst activation and electrical resistance were then extracted from the curves via least-squares fitting, and two XGBoost models were trained on these parameters. This approach resulted in a model with an average coefficient of correlation of R^2 = 0.94239. As a proof-of-concept, TGDD modelling sacrifices some accuracy in predicting cell voltage for

dramatically improved explainability, precision, and interpolation. This increases its potential for degradation-aware control and performance monitoring of PEMECs, where it is useful to compare full electrolyzer response curves.

Acknowledgement

This work is supported in part by U.K. Engineering and Physical Sciences Research Council (EPSRC) under the Grant EP/X038823/1 as part of HI-ACT project.

Data availability statement

All data that support the findings of this study are included within the article.

References

Abomazid, A.M., El-Taweel, N.A., Farag, H.E.Z., 2022. IEEE Trans. Ind. Inform. 18, 5870–5881.

Alipour Bonab, S., Waite, T., Yazdani-Asrami, M., 2025. J. Phys. Energy 8, 015006.

Andoni, M., Couraud, B., Robu, V., Blanche, J., Norbu, S., Chen, S., Kanugrahan, S.P., Flynn, D., 2025. Comparative Techno-economic Assessment of Wind-Powered Green Hydrogen Pathways, in: 2025 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe). Presented at the 2025 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe), pp. 1–5.

Bazarah, A., Majlan, E.H., Husaini, T., Zainoodin, A.M., Alshami, I., Goh, J., Masdar, M.S., 2022. Int. J. Hydrog. Energy 47, 35976–35989.

Cali, U., Lee, A., Hayes, B., Lima, C., Sebastian-Cardenas, D.J., Flynn, D., Kantar, E., Rahimi, F., Thu, K.S., Pasetti, M., Dynge, M.F., Andoni, M., Deveci, M., Kuzlu, M., Alanso, R., Choo, K.-K.R., Mishra, S., Saha, S.S., Norbu, S., Gourisetti, S., Halden, U., Hosseinezhad, V., Robu, V., 2025. Renew. Sustain. Energy Rev. 222, 115845.

Chandesris, M., Médeau, V., Guillet, N., Chelghoum, S., Thoby, D., Fouda-Onana, F., 2015. Int. J. Hydrog. Energy 40, 1353–1366.

Chen, T., Guestrin, C., 2016. XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 785–794.

Chen, X., Rex, A., Woelke, J., Eckert, C., Bensmann, B., Hanke-Rauschenbach, R., Geyer, P., 2024. Appl. Energy 371, 123550.

Falcão, D.S., Pinto, A.M.F.R., 2020. J. Clean. Prod. 261, 121184.

Hayatzadeh, A., Fattahi, M., Rezaveisi, A., 2024. Int. J. Hydrog. Energy 56, 302–314.

International Energy Agency, 2025. Global Hydrogen Review 2025.

Keddar, M., Da Conceicao, M., Doumbia, M.L., 2022. Electrolyzer Degradation-Power Electronics One - Way Interaction Model | Energy Proceedings.

Khatib, F.N., Wilberforce, T., Ijaodola, O., Ogungbemi, E., El-Hassan, Z., Durrant, A., Thompson, J., Olabi, A.G., 2019. Renew. Sustain. Energy Rev. 111, 1–14.

Krenz, T., Rex, A., Helmers, L., Trinke, P., Bensmann, B., Hanke-Rauschenbach, R., 2024. J. Electrochem. Soc. 171, 124501.

Lu, X., Du, B., Zhou, S., Zhu, W., Li, Y., Yang, Y., Xie, C., Zhao, B., Zhang, L., Song, J., Deng, Z., 2023. Int. J. Hydrog. Energy 48, 5850–5872.

Makhsoos, A., Kandidayeni, M., Pollet, B.G., Boulon, L., 2025. J. Power Sources 655, 238003.

Norbu, S., Couraud, B., Challinor, J., Flynn, D., Gibbons, G., Andoni, M., Robu, V., Taha, A., Imran, M., 2024. Cyber-physical energy systems: enabling tailored decarbonization and net-zero pathways for communities, in: IET Powering Net Zero (PNZ 2024). Presented at the IET Powering Net Zero (PNZ 2024), pp. 28–34.

Padgett, E., Yu, H., Blair, S.J., Cullen, D.A., Ahluwalia, R.K., Myers, D.J., Pivovar, B., Alia, S.M., 2025. J. Electrochem. Soc. 172, 054508.

Polo-Molina, A., Portela, J., Rozas, L.A.H., González, R.C., 2025. Modeling Membrane Degradation in PEM Electrolyzers with Physics-Informed Neural Networks.

Rakousky, C., Reimer, U., Wippermann, K., Carmo, M., Lueke, W., Stolten, D., 2016. J. Power Sources 326, 120–128.

Waite, T., Sadeghi, A., Yazdani-Asrami, M., 2025. J. Phys. Energy 7, 035013.

Waite, T., Yazdani-Asrami, M., 2025. J. Phys. Energy 7, 042002.

Wallnöfer-Ogris, E., Grimmer, I., Ranz, M., Höglinger, M., Kartusch, S., Rauh, J., Macherhammer, M.G., Grabner, B., Trattner, A., 2024. Int. J. Hydrog. Energy 65, 381–397.

Xu, B., Ma, W., Wu, W., Wang, Y., Yang, Y., Li, J., Zhu, X., Liao, Q., 2024. Energy AI 18, 100420.

Zhang, H., Si, S., Hsieh, C.-J., 2017. GPU-acceleration for Large-scale Tree Boosting.

Zhang, T., Qadrdan, M., Wu, J., Couraud, B., Stringer, M., Walker, S., Hawkes, A., Allahham, A., Flynn, D., Pudjianto, D., Dodds, P., Strbac, G., 2025. Renew. Sustain. Energy Rev. 208, 114964.