ORIGINAL RESEARCH Open access

Machine Learning-Based Evaluation of Activation Strategies and Stability in PEM Water Electrolyzers

Santiago Morales[1*], Laura Rojas[2], Camila Vega[1]

Abstract

Pretreatment of proton exchange membrane water electrolyzers is an essential step undertaken before their routine operation. Such treatments support catalyst activation and membrane saturation, securing optimal system function. In the present work, machine learning is applied to study the influence of three separate activation procedures on the cell's performance and stability. The data needed to build the surrogate models were collected from a lab-scale PEM electrolyzer. After the three tested model types were assessed and verified against experimental records, extreme gradient boosting was chosen for parametric analysis. The modeling outputs show that the activation procedures primarily alter the ohmic resistance early in the cell's lifetime. These insights were further supported by a sensitivity analysis conducted using an explainable artificial intelligence method. Additionally, time-series forecasting via data-driven models was employed to estimate cell stability across the different procedures, revealing strong agreement with experimental observations.

Keywords Machine learning, Cell conditioning, PEM water electrolyzer, Cell stability

*Correspondence:
Santiago Morales
santiago.morales@gmail.com
[1] Department of Intelligent Renewable Energy Systems, Faculty of Engineering, University of Buenos Aires, Buenos Aires, Argentina
[2] Department of Smart Energy Analytics, Faculty of Engineering, Pontifical Catholic University of Chile, Santiago, Chile

Introduction

Decarbonizing the energy landscape is essential not only to protect the environment but also to sustain long-term economic health. This pressing need has driven worldwide pursuit of green energy solutions [1]. Within this effort, proton exchange membrane (PEM) water electrolysis has risen as a frontrunner for producing green hydrogen. These electrolyzers deliver highly pure hydrogen, operate at elevated current densities, and feature a modular architecture that facilitates their integration into current transport and power infrastructures [2]. The membrane electrode assembly (MEA) is the heart of such systems, with core responsibilities including facilitating water-splitting reactions, separating product gases and enabling proton and electron transport. As the key electrochemical and transport steps occur inside the MEA, the MEA governs both the cell's efficiency and operational lifespan. Before regular use, the MEA undergoes activation—either by applying a current or by temperature cycling—to boost stability, consistency, and output. These steps [3] also serve to prevent premature degradation.

PEMECs are activated before deployment to raise their performance. The time required may range from less than 1 h to over 20 h, even for matching materials. Grasping why activation lifts the performance of PEM-based electrochemical devices is as intriguing as it is critical. When membranes form, cations such as Fe3+, Ni2+, and Cr2+ can enter and degrade cellular output [4]. They strike the polymer chains, triggering both mechanical and chemical instability. Some additionally bind protons, thereby obstructing and reducing protonic conductivity. The membrane also experiences

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (CC BY-NC-SA 4.0), which permits non-commercial use, distribution, reproduction, and adaptation in any medium or format, provided that appropriate credit is given to the original author(s) and source, a link to the license is provided, and any changes made are indicated. Adapted material must be distributed under the same license.

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

fluctuations in temperature and moisture during processing. These conditions together affect its water uptake, leading to uneven swelling and structural changes that subsequently affect proton conduction during regular operation [5, 6]. A fraction of the catalyst particles may also be masked by oxide films or hold solvent residues from the catalyst paste, which block the pores of the carbon paper. These phenomena cut down the active electrochemical surface area and diminish cell output. Activation procedures [7, 8] therefore aid in optimizing the start-up phase by conditioning the catalyst layer, enhancing water saturation of the membrane, and flushing contaminants from the membrane.

Although conditioning methods for PEMECs are well established [3], relatively few in-depth studies have examined their influence, compared with research on PEMFCs. One preconditioning approach, applying a steady current density for 24 h, was used to promote membrane hydration and stabilize the catalysts' oxidation state [9]. To harmonize testing and enable trustworthy cross-comparison of PEM electrolyzers across five laboratories, the same test cells, materials, and protocols were used. For a given temperature and flow rate, preconditioning persisted until the current density varied by less than 1%. A low maximum deviation in cell voltage was recorded across the laboratories at several operating temperatures. Electrochemical characterization, combined with in situ imaging, showed that conditioning increased the number of reaction sites [10]. Electrochemical impedance spectroscopy (EIS) data indicated drops in both ohmic and overpotential losses after conditioning. More recently, the influence of one ex-situ and two in-situ conditioning procedures was examined to gauge their impact on near-term cell performance [11]. Experiments deliver vital insights into conditioning strategies through electrochemical diagnostics, yet a recent pivot has shifted toward blending data-driven models with measured data. To scale these systems for market use and permit real-time tracking, data-driven models have become an indispensable resource for predictive analytics, control, and optimization, thereby closing the gap between lab experiments and industrial roll-out [12-14].

In the PEMEC sphere, machine learning has largely focused on screening electrocatalysts [15-17]. In contrast to PEMFCs, where considerable work [18] has pursued data-driven models for performance forecasting and system optimization at multiple scales, equivalent endeavors for PEMECs are still in the formative stage [19, 20]. X-ray tomographic images of the anode PTL in a PEMEC served as training inputs for a machine learning model that predicted oxygen content in the layer [21]. K-nearest neighbors (KNN) and decision tree regression models were used to optimize PEMEC design by selecting the optimal flow-field configuration for maximal hydrogen production efficiency [22]. By uniting datasets from their own experiments and the literature, the authors unveiled the first data-driven framework for efficient hardware selection (flow field, catalyst type, membrane type, cell count) for PEM water electrolyzers. They then proposed design parameters for electrolyzers with large-scale production rates ranging from 50 to 300 mL/min. A database comprising 789 experiments from 30 papers across the past decade was assembled to identify the most decisive parameters affecting PEMEC performance [23]. Leveraging ML methods including SHAP, KNN, and Bayesian optimization, the authors identified operating temperature, catalyst loading, ionomer content, anode/cathode support materials, and PTL pore structure as critical. A similar data-driven cell design optimization was carried out using polynomial regression [14]. A close match in hydrogen production rate was found between a custom cell built using the model-predicted parameters and the model's own forecast. More recently, a data-driven methodology was advanced [24] to forecast degradation trajectories in PEMECs using operational data. Ozdemir and Pektezel [25] investigated how cell voltage, temperature, torque, and flow rate shape performance using machine learning algorithms trained on experimental data. They reported that support vector machines outperformed both random forests and multi-layer perceptron algorithms in predicting current density and hydrogen flow rate, with MAEs of 0.006 and 0.0317 on the training and test data, respectively. The authors have also recently examined PEMEC performance using synthetic data generated from numerical simulations [26]. Among the four ML models tested, SVM demonstrated the highest predictive power, yielding RMSE values of 0.0108 and 0.0371 on training and test sets, respectively.

The foregoing survey makes it clear that research at the intersection of PEMEC experiments and machine learning has concentrated overwhelmingly on refining material properties, cell architecture, and operating parameters. A key variable—cell activation, which shapes both membrane hydration and catalyst activation—has yet to feature as a predictor in data-driven studies. Moreover, economic practicality demands the ability to predict electrolyzer stability under different conditioning protocols. Such foresight would guide the selection of a method that delivers not merely the best short-term

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

output but also robust long-term efficiency. The present study sets out to close these gaps. First, we draw on experimental data acquired from a lab-scale PEMEC across three distinct conditioning strategies and construct data-driven models that incorporate conditioning as an input feature. This enables the model to capture the effect of MEA pretreatment on cell performance. Second, we apply SHAP, an explainable AI technique, to discern which pretreatment yields the highest cell performance. Lastly, we model PEMEC stability using a dataset built from over 200 h of operation. Temporal patterns were learned through an LSTM. With its gated structure, the LSTM sidesteps the vanishing and exploding gradient issues typical of conventional recurrent neural networks.

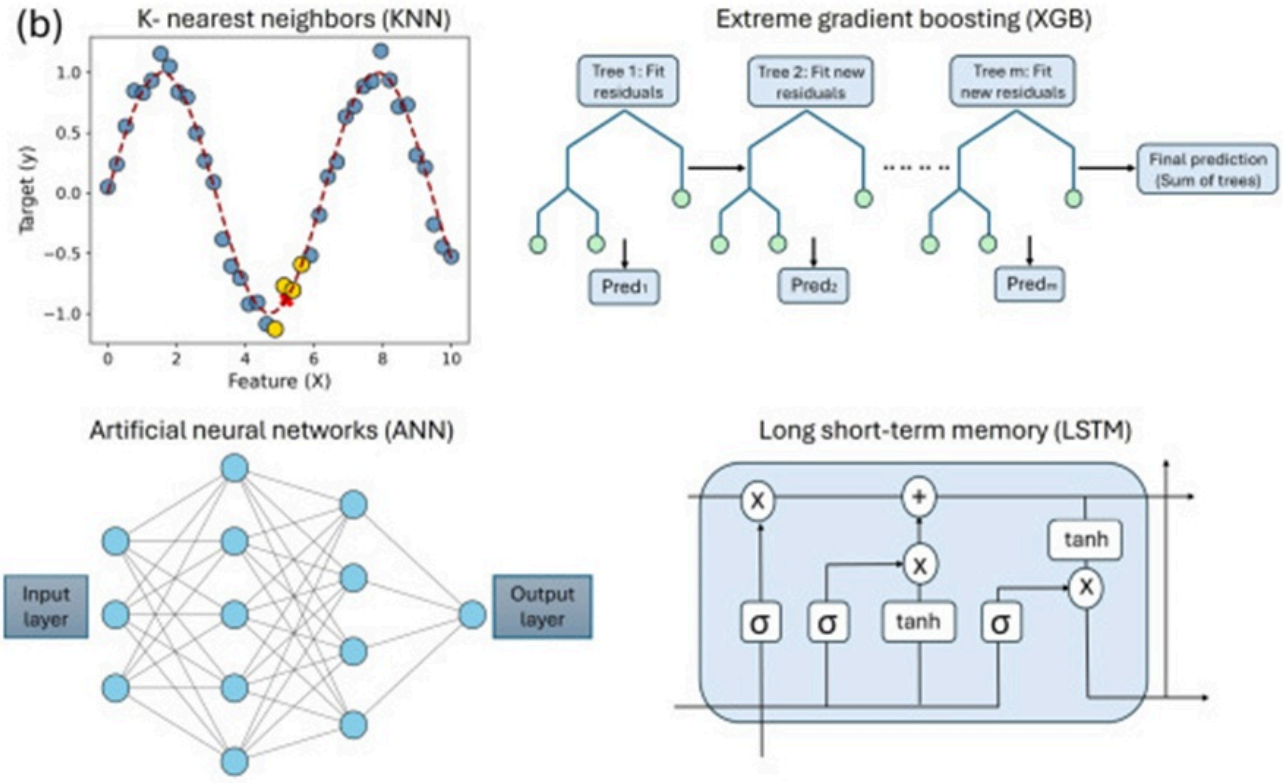
(a)

Mechanical impact (ex-situ)

Thermal impact (in-situ/water (90°C))

Chemical impact (in-situ/acids)

Activation strategy	Description of the activation strategy
Ex-situ	MEA treated before cell assembly (ex-situ) by soaking it in de-ionized water at 60°C for 12 hours.
In-situ/water (90°C)	MEA treated after cell assembly (in-situ) with a supply of de-ionized water at 90°C for 12 hours .
In-situ/acids	MEA treated after cell assembly (in-situ) with: • supply of H_2O_2 for 1 hour • supply of 0.5 mol/L H_2SO_4 for the next1 hour • supply of de-ionized water at 60°C for the next 12 hours.
In-situ/water (60°C)	MEA treated after cell assembly (in-situ) with a supply of de-ionized water at 60°C for 12 hours (reference).



© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

Figure 1. (a) Description of the different MEA activation strategies investigated in this study. The in-situ/water (60 °C) is the reference activation strategy case against which the other pre-treatments methods were compared. (b) Schematic representation of the working principle of the employed machine learning models.

Materials and Methods

Data curation

The dataset was collected from four single cells constructed in-house and fitted with commercial HYDRion MEAs from Ion Power GmbH, featuring a Nafion N115 membrane. The anode catalyst layer contained iridium oxide loaded at 1.0 mg/cm², while the cathode held platinum at 0.3 mg/cm². The MEA's active area spans 4.4 × 4.4 cm², which is larger than the 9 cm² active area used in a recent publication [25]. A 250 µm-thick CURRENTO 2GDL10N-0.25 porous transport layer, coated with 0.25 µm of platinum on both faces, acted as the anode-side PTL. Toray carbon paper (TGP-H-60) served as the cathode PTL. The bipolar plates were machined from titanium grade 2 and featured parallel flow channels. Their anode and cathode surfaces were plated with platinum (2.5 µm) and gold (2.5 µm), respectively. Temperature was controlled by heating each side of the cell with dual heating rods, each controlled by a temperature sensor. The cells ran on test benches supplied with high-purity water at 0.05 L/min. After pretreatment, the operating temperature was held at 60 °C. A more detailed account of the experimental rig and test stations appears in Wolf *et al*. [11].

The input features included applied current density, frequency, and amplitude. An extra feature capturing the pretreatment conditioning type was appended. One-hot encoding [27] converted this categorical conditioning feature into numeric form. The target variables comprised cell voltage along with the real and imaginary impedance components. Min–max scaling was applied to both predictors and outputs. **Table 1** lists the input features and measured output variables, together with the ranges over which the inputs were swept and the outputs recorded. **Figure 1** illustrates each activation strategy. Data points were captured every second across all experiments. From the four datasets, records corresponding to the pretreatment phase were removed. Four ascending current densities—0.5, 1.0, 1.5, and 2.0 A/cm²—were sustained for 12 h apiece. Within each 12 h block at a fixed current density, three polarization curves and three impedance spectroscopy sweeps were collected. To guard against spurious outliers, only readings with a voltage at or above 1.4 V (the open-circuit value) were retained. Entries with missing values were identified and the corresponding rows deleted. After cleaning and merging the data from all four cells, the consolidated database totaled 133,558 records. The database shows high heterogeneity: the categorical pretreatment feature is balanced across the four conditions, yielding a Shannon entropy of 1.997 bits out of a maximum of 2. Similarly, the applied current density displays dispersion among samples, with a standard deviation of 0.62 A/cm² and a coefficient of variation of 1.03. In data-driven modeling, the data are typically split into training and test partitions. The training portion enables ML algorithms to learn input–output mappings; the test portion gauges predictive accuracy on unseen cases. We split the experimental database randomly into 80% for training and 20% for testing. A small holdout set—one polarization curve and accompanying impedance readings at 1.0 A/cm²—was reserved for model validation.

Table 1. Range of measured variables provided to the data-driven model.

Parameters	Type	Unit	Range of values
Current density	Input	A/cm^2	[0.0–2.0]
Frequency	Input	Hz	[0.1–5 × 10^5]
Activation	Input (categorical)	–	ex-situ, in-situ/water(60 °C), in-situ/acids, in-situ/water(90 °C)
Voltage	Output	V	[1.4–1.93]
Impedance (Real)	Output	mΩ	[7.5–20.48]

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

Impedance (Imaginary)	Output	mΩ	[0–3.85]

To forecast electrolyzer stability, an independent database covering 200 h of operation was compiled for each activation procedure. The LSTM model's temporal predictions of cell voltage are derived from this database.

Before constructing machine learning models, it is necessary to determine whether the conditioning strategy type meaningfully affects cell voltage. Significance tests, such as analysis of variance (ANOVA), provide statistical evidence for screening input features, helping keep irrelevant variables and noise out of the models. Given the study's emphasis on the effect of conditioning on performance, we employed ANOVA to test its statistical significance. **Table 2** summarizes the one-way ANOVA on cell voltage, treating conditioning strategy as the categorical factor. The result—a P-value < 0.001 and F(3, 133,554) = 11.81—indicates that MEA conditioning type significantly influences cell voltage. Consequently, it warrants inclusion as an ML input feature. For additional confirmation, we ran pairwise comparisons among conditioning strategies via the Tukey HSD test. **Table 3** reports these comparisons alongside adjusted p-values and confidence intervals. Pairs with adjusted p-values under 0.05 reject the null hypothesis. Based on the adjusted p-values and interval bounds, ex-situ conditioning tends to yield higher voltages than the in-situ approaches. Additionally, the mean voltages for the baseline in-situ case exceed those for the in-situ/water (90 °C) conditioning. These outcomes reinforce the conclusion that conditioning type exerts a statistically significant impact on cell voltage.

Table 2. One-way ANOVA results for the effect of conditioning on cell voltage.

Source	Sum of squares	df	F	p-value
Pre-treatment	0.703	3	11.81	< 0.001
Residual	2649.82	133 554	–	–

Table 3. Tukey HSD pairwise comparisons of voltage (Ewe) for different pre-treatment types. Significant differences (p < 0.05) are indicated.

Group 1	Group 2	Upper CI	Lower CI	p-adj
ex-situ	in-situ/acids	−0.0007	−0.0063	0.007
ex-situ	in-situ/water(60°)	−0.0029	−0.0083	0.000
ex-situ	in-situ/water(90°)	0.0023	−0.0033	0.965
in-situ/acids	in-situ/water(60°)	0.0007	−0.0049	0.223
in-situ/acids	in-situ/water(90°)	0.0059	0.0001	0.042
in-situ/water(60°)	in-situ/water(90°)	0.0079	0.0022	0.000

K-Nearest Neighbors (KNN)

K-nearest neighbors [28] is a non-parametric machine learning method suited for regression tasks. It derives predictions by consulting the target values of the closest training examples. Formally, the KNN prediction is given by:

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

$$\hat{y} = \frac{1}{k} \sum_{x_i \in N_k(x)} y_i \quad (1)$$

Where $\hat{y}$ represents the predicted output, k is the count of nearest neighbors, x_i are points belonging to the neighbor set N_k, x is the query point, and y_i denotes the known target values at x_i. The sole hyperparameter to tune is k, the number of neighbors. The algorithm identifies the k points whose Euclidean distance to x is smallest and outputs the arithmetic mean of their target values from the training set.

Extreme gradient boosting (XGB)

XGBoost (eXtreme Gradient Boosting) is an ensemble method that sequentially constructs decision trees for both regression and classification problems. An individual decision tree operates by repeatedly partitioning the training data into increasingly homogeneous subsets. Each split selects a feature and a threshold that best separates the data. Splitting halts once the tree reaches its maximum depth or a node contains fewer than the minimum allowed samples. XGB proceeds by iteratively fitting a weak decision tree to the negative gradient of the loss function (squared error), which is the core principle of gradient boosting. Each successive tree is fitted to the residuals left by the preceding ensemble.

Artificial neural network (ANN)

Modeled loosely on biological neural systems, artificial neural networks are computational architectures capable of capturing complex regression relationships. The elementary unit is the neuron—a mathematical node that receives inputs and emits an activation signal. Neurons are organized into layers. The input layer passes the predictor variables into the network; the output layer delivers the final predictions. Layers positioned between them are termed hidden layers and are responsible for feature extraction and transformation. Information propagates forward from the input layer, through the hidden layers, to the output layer. Extended descriptions of the XGB and ANN methodologies can be found in Raman *et al*. [19] and Karyofylli *et al*. [20].

Long-short-term memory (LSTM) model for forecasting

LSTM [29] is a specialized class of recurrent neural network designed to overcome the vanishing and exploding gradient problems that afflict conventional RNNs. This is achieved through a gated architecture coupled with a memory cell that governs information retention and deletion. The LSTM unit comprises the following elements:

• Cell state: A pathway that carries long-term temporal information across time steps and influences the unit's output.

• Forget gate: Determines what fraction of the previous long-term memory to keep, using a sigmoid activation:

$$f_t = \sigma(W_f.[h_{t-1}, x_t] + b_f) \quad (2)$$

•

Input gate: Contains a tanh layer that proposes new long-term memory content and a sigmoid layer that decides how much of this proposal to accept and combine with the forget gate's output.

$$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i) \quad (3)$$

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

$$\hat{C}_t = \sigma(W_C.[h_{t-1}, x_t] + b_C) \quad (4)$$

$$C_t = (f_t \times C_{t-1}) + (i_t \times \hat{C}_t) \quad (5)$$

•

Output gate: Controls the final output of the LSTM unit based on the updated cell state.

$$O_t = \sigma(W_o.[h_{t-1,x_t}] + b_o) \quad (6)$$

$$h_t = o_t \times tanh(C_t) \quad (7)$$

In these expressions, W_x and b_x denote the weight matrices and bias vectors associated with the respective gates ($x : f$, i, C, and o). Multiple LSTM units can be stacked into a layer, and successive layers can be arranged to pass information forward. For the stability analysis presented here, we employed a two-layer configuration with 50 and 25 LSTM units, respectively.

Results and Discussion

We begin the presentation of results by gauging the three machine learning models' performance via residual plots, then proceed to validate them using experimental data, and finally conduct a parametric analysis employing the model that excels across all target variables.

Model performance

Residual plots offer a visual means to infer the error spread and patterns arising from ML model predictions. **Figure 2** presents the residual plots for the KNN, XGB, and ANN models. Each was trained to predict three separate target variables: cell voltage (top row), the real impedance component (middle row), and the negative imaginary impedance component (bottom row). **Table 4** lists the search ranges and the final chosen hyperparameter values for the trained models. A residual is defined as the gap between the experimentally measured value and the model's prediction. For instance, the middle-row plots convey how accurately each model reproduces the real impedance component relative to actual experimental records. Whereas cell voltage sheds light on the cell's overall efficiency, impedance measurements help tease apart overlapping phenomena such as electrode kinetics, ohmic resistance, and mass transport losses taking place within the system. Residuals originating from the training and test datasets are distinguished by separate colors. In addition, the root mean square error (RMSE) and the coefficient of determination (R^2) are adopted to quantify model performance, formulated as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2} \quad (8)$$

$$R^2 = 1 - \frac{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2}{\frac{1}{n}\sum_{i=1}^{n}(y_j - \bar{y}_j)^2} \quad (9)$$

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

where n indicates the sample count, y_j stands for the ground truth from experiments, $\hat{y}$ is the model's estimate, and $\overline{y}$ represents the mean of y_j.

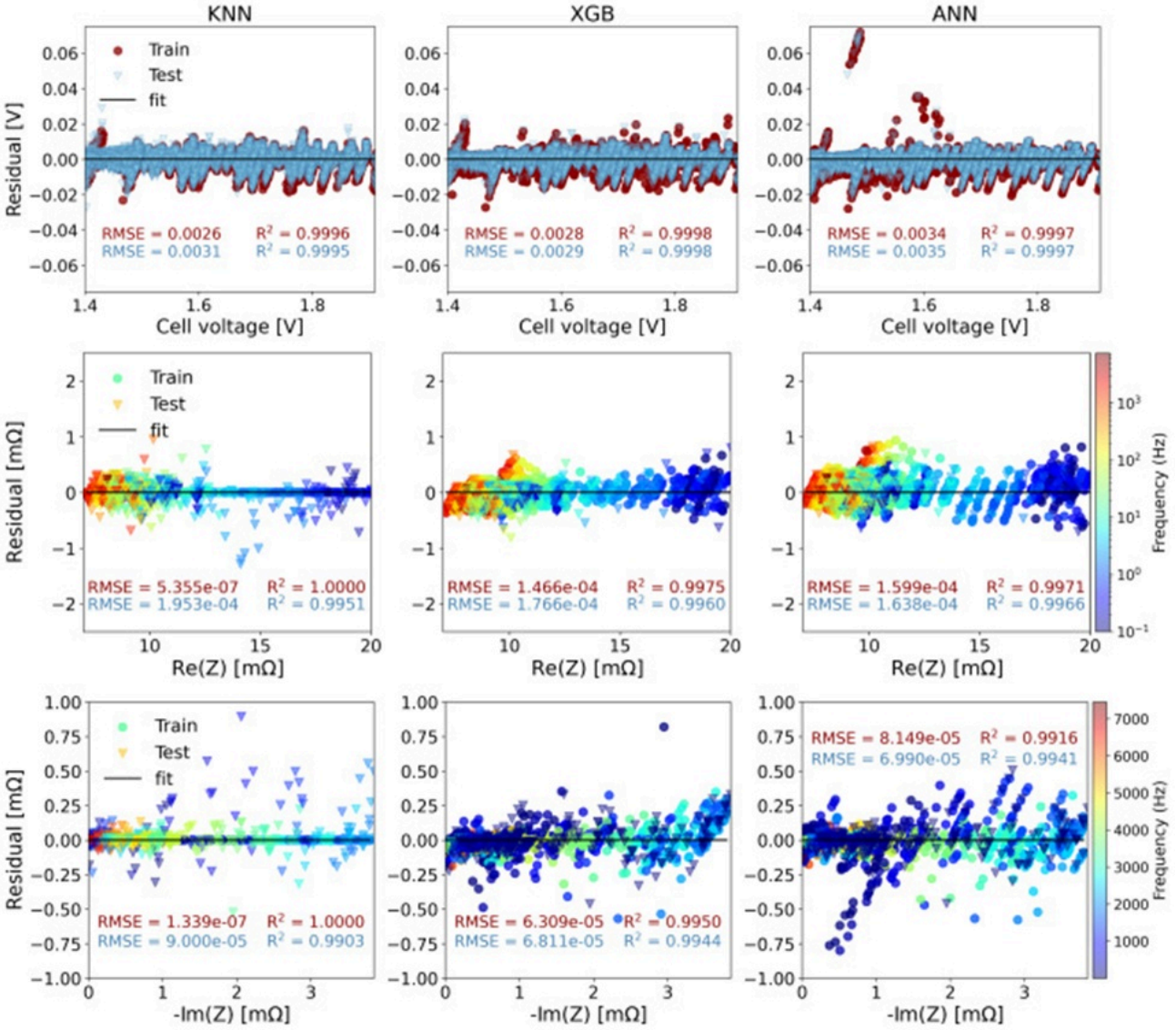


Figure 2. Residual plots showing the comparison between the predicted current densities (top row), real (middle row) and imaginary (bottom row) components of impedance with those obtained from experimental data. The model's performance on the training and test set are shown in each plot by the RMSE and R² values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4. Hyperparameters and search range for each machine learning model.

Model	Hyperparameter	Search range	Selected values (Voltage, Re(Z) and -Im(Z))
KNN	n_neighbors	[3,8,11,14]	3, 8 and 8
XGB	n_estimators	[100–800]	300, 800 and 800
	learning_rate	[0.05–0.20]	0.015, 0.005 and 0.005
ANN	No. of hidden layers	[2–4]	4, 3 and 4

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

No. of hidden units	[20–100]	[100-80-60-30], [60-40-20] and [100-80-60-30]
learning rate	[0.0001–0.01]	0.00025, 0.001 and 0.00015

Inspecting **Figure 2** (top row), the residual spread over the entire cell voltage range looks broadly similar across all three models. For KNN and XGB, both training and test residuals stay within 0.02 V. The ANN model produces a few outliers, chiefly on the test set, where residuals can reach roughly 0.06 V. KNN attains the best training-set result (RMSE = 0.0026), while on unseen test data XGB edges ahead slightly with RMSE = 0.0031. Every model yields R^2 above 0.999, signaling high-quality voltage predictions.

Whereas voltage prediction links directly to applied current density, electrochemical impedance spectroscopy (EIS) delivers a frequency-dependent, complex-valued variable. Impedance is governed by several system attributes—membrane resistance, double-layer capacitance, diffusion—each reflecting distinct internal processes. Capturing these nuances makes impedance inherently more challenging to model. This is apparent from the residual distributions of the real (middle row) and imaginary (bottom row) impedance components in **Figure 2.** KNN yields an almost perfect fit for Re(Z) and Im(Z) on the training set, with RMSE = 5.3×10^{-7} and 1.33×10^{-7} for training and test data, respectively. On the test data, however, its residuals fan out progressively as the applied frequency grows. The same broadening pattern appears in the XGB and ANN models as well; the ANN model's Re(Z) test residuals additionally display a wavy structure. This non-constant residual variance points to heteroskedastic behavior. KNN suffers from inflated variance through overfitting the test data, while ANN generates numerous outliers on Im(Z). Although the test-set RMSE values for both impedance components suggest ANN slightly outdoes XGB, ANN equally exhibits more pronounced heteroskedasticity. For Re(Z), XGB shows tighter residual dispersion than ANN. On the whole, modeling the capacitive (imaginary) side of impedance proves tougher than modeling the resistive side. Hence, among the three contenders, XGB delivers the most well-rounded performance for both voltage and impedance prediction.

The above performance picture rests on a random split of the database into training and testing partitions. Random splitting guarantees that all numerical and categorical feature values are fairly represented in both subsets, offering a window into overall model fit. Yet because MEA activation type serves as an input feature, data leakage becomes a concern if samples tied to a single activation type inhabit both the training and test sets. Data leakage materializes when readings from the same underlying experiment appear in both partitions. To examine this further, **Table 5** sets out R^2 and RMSE metrics for all three models under random versus grouped splitting, restricted to the test data. Grouped splitting ensures every sample belonging to a given activation type is held entirely within the test set, while the training set draws from the remaining activation types. A mild dip in R^2 and RMSE for voltage predictions is evident when switching from random to grouped splitting. This trend is expected, since the models must generalize to a categorical feature value they never encountered during training. A more pronounced fall—around 0.2–0.3 in R^2—appears for Re(Z) predictions under grouped splitting. KNN performs marginally better than XGB and ANN in this scenario. For Im(Z) forecasts, ANN reveals overfitting: its R^2 plunges from 0.994 to 0.714 when the split mode changes. The numbers in **Table 5** suggest that voltage predictions are considerably more resistant to data leakage than impedance predictions.

Table 5. Comparison of methods with performance metrics on Voltage, real (Re(Z)), and imaginary (-Im(Z)) components of impedance on the test data.

Model	Data Partitioning Strategy	−Im(Z) RMSE	−Im(Z) R^2	Re(Z) RMSE	Re(Z) R^2	Voltage RMSE	Voltage R^2
KNN	Random Split	1.15×10^{-4}	0.983	2.11×10^{-4}	0.994	3.0×10^{-3}	0.999
	Group-Based Split	1.91×10^{-4}	0.953	1.3×10^{-3}	0.802	2.3×10^{-2}	0.975
XGB	Random Split	6.81×10^{-5}	0.994	1.76×10^{-4}	0.995	1.84×10^{-2}	0.982

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

	Group-Based Split	1.79×10^{-4}	0.959	1.54×10^{-3}	0.722	2.43×10^{-2}	0.974
ANN	Random Split	6.99×10^{-5}	0.994	1.63×10^{-4}	0.996	3.52×10^{-3}	0.999
	Group-Based Split	4.74×10^{-4}	0.714	1.48×10^{-3}	0.743	2.87×10^{-2}	0.963

Model validation

Residual patterns and summary metrics sketch the models' broad behavior. To sharpen the quantitative picture, we compare their outputs directly with experimental data for (a) thermal and (b) chemical conditioning protocols in **Figure 3**. The comparison relies on polarization curves and Nyquist plots captured at a current density of 1 A/cm². For thermal conditioning, the experimentally recorded polarization curve matches all three models closely, as **Figure 3a** shows. The Nyquist plot, on the other hand, exposes the divergences among the predictions. **Table 6** collects the equivalent circuit fitting parameters. The ANN model strays further from experimental readings than XGB or KNN. Under thermal pretreatment, ANN under-predicts the ohmic resistance and shifts the entire Nyquist trace leftward of the experimental curve. KNN predictions track the impedance data well across the high-to-medium frequency band but diverge notably at low frequencies (Re(Z) ≈ 9.0). XGB captures both the ohmic and charge transfer resistances with fidelity, and its predictions sit close to the experimental points over the full frequency span, though it slightly overestimates the capacitive portion at intermediate frequencies. **Figure 3b** presents the analogous comparison for chemical conditioning. As with the thermal case, the polarization curves from the ML models align well with experimental readings. For impedance, both KNN and XGB maintain good agreement with measured values under chemical conditioning. The ANN model, however, slightly over-predicts the ohmic and charge transfer resistances, shifting the Nyquist trace a touch to the right.

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

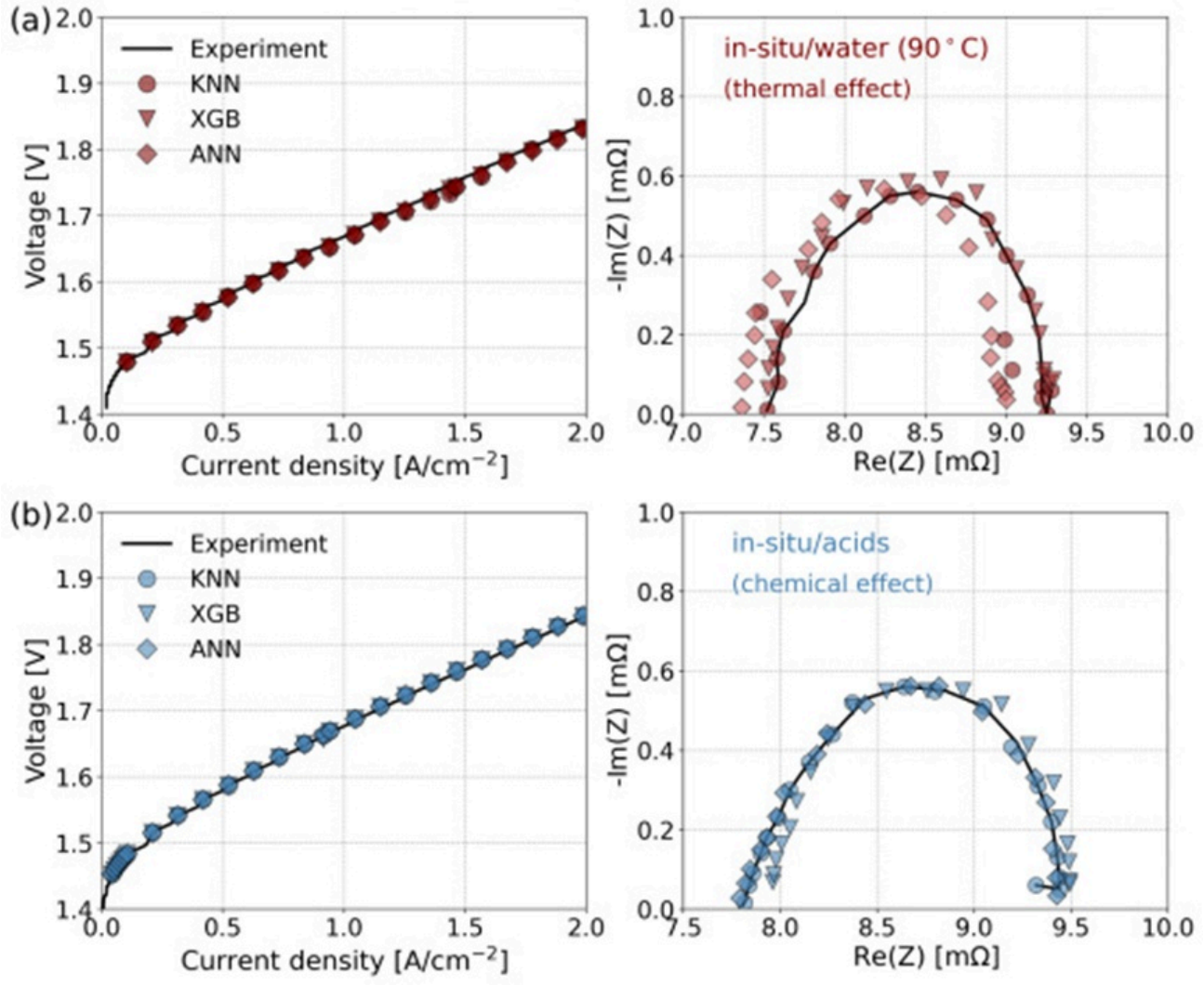


Figure 3. Comparison between experimental measurements and model prediction through polarization curve and Nyquist plot for MEAs pre-treated with (a) deionized water at 90 °C and (b) acidic solutions. The impedance measurements were made at a current density of 1.0 A/cm². Additionally, it is to be noted that as the applied frequency increases, the real component of impedance decreases.

Table 6. Equivalent circuit model fitting parameters for different ML models corresponding to the Nyquist plot in **Figure 3.**

Pre-Treatment Method	Model Type	CPE-P Error (%)	CPE-P Value	CPE-T Error (%)	CPE-T Value	RCT Error (%)	RCT (Ω) Value	Rs Error (%)	Rs (Ω) Value
Thermal Treatment	Experimental Measurement	1.92	0.7439	7.74	0.0076	1.45	1.755	0.20	7.528
	K-Nearest Neighbours (KNN)	3.53	0.7599	14.72	0.0067	2.67	1.719	0.38	7.500
	Extreme Gradient Boosting (XGB)	1.52	0.7182	5.78	0.0095	1.14	1.855	0.16	7.463

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

	Artificial Neural Network (ANN)	2.04	0.7624	8.43	0.0074	1.55	1.676	0.21	7.340
Chemical Treatment	Experimental Measurement	1.51	0.7542	6.19	0.0088	1.19	1.682	0.12	7.832
	K-Nearest Neighbours (KNN)	1.78	0.7538	7.25	0.0091	1.40	1.679	0.15	7.834
	Extreme Gradient Boosting (XGB)	1.87	0.7389	7.56	0.0092	1.46	1.666	0.15	7.921
	Artificial Neural Network (ANN)	1.88	0.7409	7.57	0.0093	1.47	1.695	0.16	7.813

Judging from the impedance residual spreads in **Figure 2,** KNN clings tightly to the training data yet falters on the test set. Although it aligns reasonably with experiments in **Figure 3**, its instance-based learning style drives it to memorize, making its forecasts heavily contingent on local data structure. This manifests as noisy low-frequency impedance predictions for chemical conditioning in **Figure 3b**. Without built-in regularization, KNN remains vulnerable to overfitting. Neural networks, conversely, possess a wealth of trainable parameters that, given sufficient data, can capture intricate input–output relationships. For impedance measurements, where distinct electrochemical processes span a wide frequency range, limited training data may lead ANNs to either overfit or latch onto only the dominant mechanism. These tendencies likely explain ANN's under-prediction in **Figure 3a** and over-prediction in **Figure 3b**. In overall accuracy and faithful reproduction of experimental trends, XGB stands out as the strongest candidate. Accordingly, we select this model to move forward with the parametric analysis.

Effect of activation strategies

We now assess the ex-situ, in-situ/acids, and in-situ/water (90 °C) activation protocols relative to the reference case through electrochemical characterization, drawing on polarization curves and Nyquist plots for the comparison. All findings in this section are generated by the XGB model. Four separate MEA activation protocols were tested in this work to gauge their bearing on PEMEC performance. One of these four pretreatment strategies serves as the reference case against which the remainder are measured. Under the reference protocol, the MEA was hydrated with deionized water at 60 °C for 12 h after being placed inside the cell. We turn next to the influence of the other three activation protocols on cell output relative to this baseline.

Mechanical impact

In the ex-situ protocol, the MEA was hydrated with deionized water at 60 °C for 12 h while still outside the cell (ex-situ) and then fitted into the cell in a wet condition, thereby spotlighting how membrane swelling from hydration shapes performance. **Figure 4a** sets the ex-situ protocol against the reference via polarization curves. At low current densities (0–0.5 A/cm²), the two pretreatment approaches look nearly alike. As current density builds, ex-situ activation yields higher cell voltage readings than the reference. The inset makes clear that this voltage penalty expands, particularly at higher current densities; a gap of 28.5 mV emerges at 2.0 A/cm² between the two protocols. The Nyquist plot in **Figure 4b** captures how current density variations affect impedance spectra under mechanical impact. Several current density values (0.75, 1.25, and 1.75 A/cm²) were intentionally chosen because they were missing from the original training set,

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

purposely demonstrating the model's predictive reach beyond the training database. The ohmic resistance is taken from the high-frequency intercept. A barely perceptible shift of 0.34 mΩ in ohmic resistance accompanies the current density swing from 2.0 to 0.50 A/cm². The charge transfer resistance, taken from the low-frequency intercept, shrinks as current density grows—a consequence of the rising overpotential that intensifies the driving force for the reaction. To spotlight the mechanical impact, **Figure 4c** compares Nyquist plots at 1.75 A/cm². Ohmic resistance values for ex-situ and reference protocols came to 8.79 mΩ and 8.16 mΩ, respectively, while charge transfer resistance increased from 9.14 mΩ to 9.96 mΩ when the MEA was subjected to mechanical impact. Collectively, these numbers point to a performance drop with ex-situ activation.

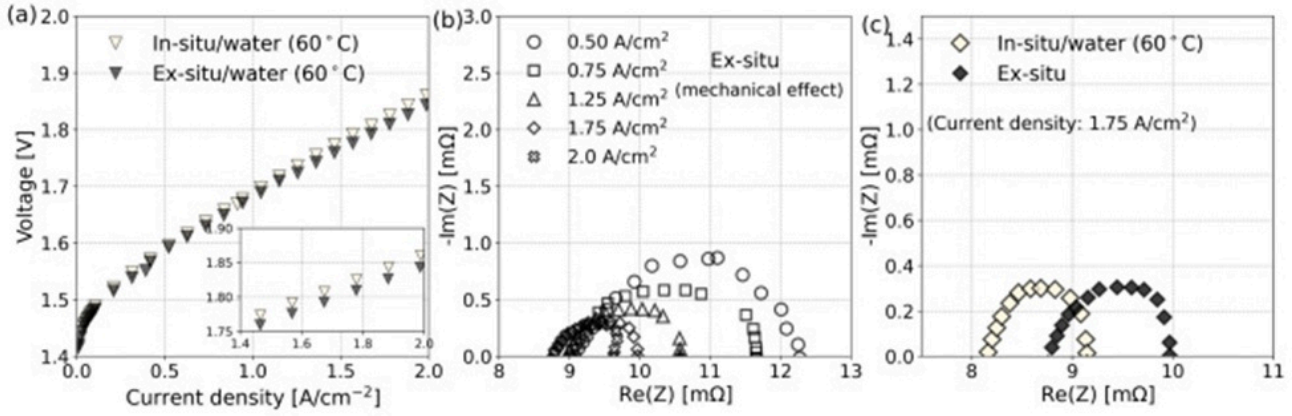


Figure 4. Electrochemical performance comparison between reference and ex-situ activation. (a) Polarization curves highlighting performance reduction due to mechanical impact. (b) Nyquist plots at different current densities (0.5–2.0 A/cm²) for ex-situ activation showcasing the changes in ohmic (RΩ) and charge transfer resistance. (c) Comparison of Nyquist plots between reference and ex-situ activation, illustrating increased RΩ due to mechanical impact.

Chemical impact

For in-situ/acids activation, the MEA was first mounted in the cell (in-situ) and exposed to a steady flow of hydrogen peroxide (H_2O_2) for 1 h. This was followed in sequence by deionized water and 0.5 mol/L sulphuric acid (H_2SO_4), each for 1 h, and finished with a 12 h rinse of deionized water at 60 °C. **Figure 5a** juxtaposes the in-situ/acids protocol with the reference on the polarization curve. Unlike the preceding mechanical case, chemical pretreatment yields a slight edge—shaving off a few millivolts in losses—which the inset of **Figure 5a** further accentuates. **Figure 5b** tracks the response of ohmic and charge transfer resistance to current density for the chemically treated MEA. Ohmic resistance shifts a mere 0.1 mΩ as current density goes from 2.0 to 0.5 A/cm², revealing a narrower spread than in the ex-situ case discussed above. Charge transfer resistance, meanwhile, follows a familiar downward path with increasing current density; a 27.4% change is recorded when moving from 0.5 to 2.0 A/cm². Chemical pretreatment cuts the ohmic resistance, thereby easing ion transport across the membrane. This is borne out by the drop of RΩ from 8.16 mΩ for the reference to 7.75 mΩ for in-situ/acids activation, as shown in **Figure 5c.** Charge transfer resistance likewise falls, from 9.14 mΩ to 8.76 mΩ.

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

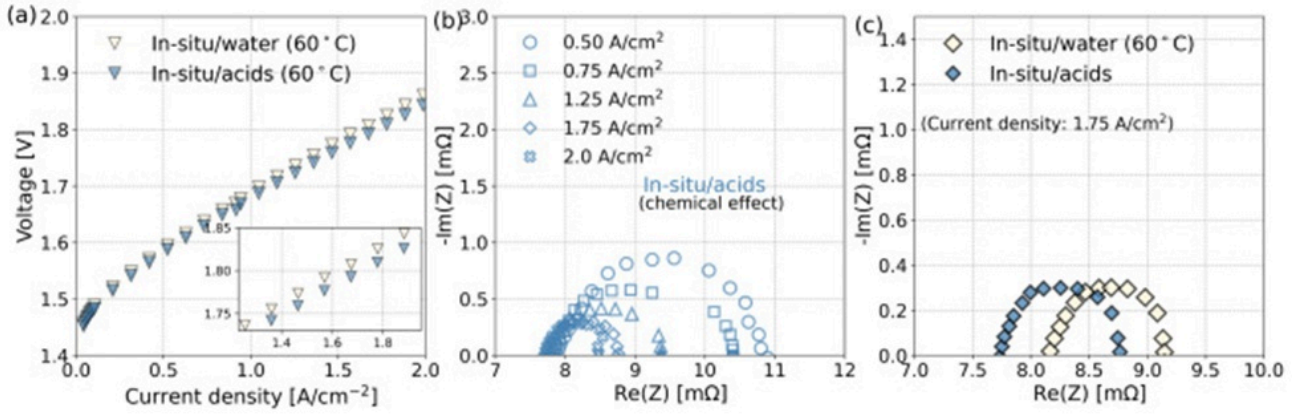


Figure 5. Electrochemical performance comparison between reference and in-situ/acids activation. (a) Polarization curves highlighting performance improvement due to chemical impact. (b) Nyquist plots at different current densities (0.5–2.0 A/cm²) for in-situ/acids activation showcasing the changes in ohmic (RΩ) and charge transfer resistance. (c) Comparison of Nyquist plots between reference and in-situ/acids conditioning, illustrating reduced RΩ due to chemical impact.

Thermal impact

In the thermal protocol, the MEA was mounted in the cell (in-situ) and deionized water at 90 °C was circulated for 12 h. The sole operational difference from the reference is the water temperature during pretreatment—all measurements were taken at 60 °C. As **Figure 6a** illustrates, the MEA pretreated at 90 °C delivers a clear voltage reduction spanning the entire current density range. Unlike the chemical case, the size of this reduction holds fairly steady irrespective of current density: voltage differences at 0.52, 1.54, and 1.95 A/cm² are 14.61 mV, 14.8 mV, and 13.78 mV, respectively. **Figure 6b** shows RΩ staying almost unchanged as current density climbs from 0.5 to 2.0 A/cm², plausibly because pretreatment at 90 °C fully saturates the membrane and thus minimizes the barrier to H^+ migration. Charge transfer resistance declines in the usual logarithmic relationship with overpotential. The boost in ion transport is further underscored in **Figure 6c**, where RΩ drops from 8.16 mΩ (reference) to 7.14 mΩ at 1.75 A/cm².

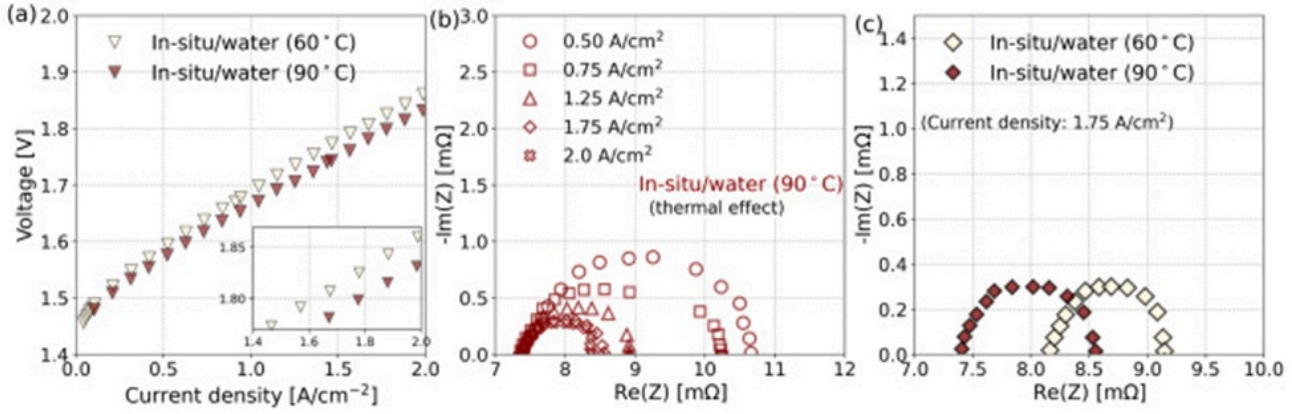


Figure 6. Electrochemical performance comparison between reference and in-situ/water(90 °C) activation. (a) Polarization curves highlighting performance improvement for thermally activated MEA. (b) Nyquist plots at different current densities (0.5–2.0 A/cm²) for in-situ/water(90 °C) activation showcasing the changes in ohmic (RΩ) and charge transfer resistance. (c) Comparison of Nyquist plots between reference and in-situ/water(90 °C) activation, illustrating reduced RΩ for thermal activation.

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

Comparative analysis of conditioning strategies using SHAP

In the earlier section, the electrochemical behavior of the three conditioning methods was assessed relative to a reference condition in which the MEA underwent hydration with deionized water at 60 °C for 12 h. Although the performance deviations associated with each conditioning approach were quantitatively benchmarked against the reference case, a comprehensive comparison among the conditioning methods themselves has not yet been established. In addition, the mechanisms responsible for the observed differences in cell performance under the various conditioning treatments merit further interpretation and discussion. To address these aspects, the present section employs the explainable artificial intelligence technique SHAP.

SHAP [30], derived from principles of game theory, is an interpretability method that assigns contribution scores to input variables influencing the predictions of a machine learning model. The SHAP value quantifies the average effect of a given feature on model outputs when evaluated across every possible combination of input variables. As a result, it provides insights into the relative significance of different features and clarifies why the model places greater emphasis on certain variables than others. The SHAP summary plot presented in **Figure 7a** identifies the extent to which each of the three conditioning strategies influences cell voltage. In the figure, the horizontal axis corresponds to SHAP values, whereas each data point represents an individual experimental observation and its associated SHAP contribution for the relevant conditioning condition. The ordering of the conditioning methods is determined according to their mean SHAP values. The color scale positioned on the left side of the plot indicates feature values. Unlike continuous variables such as current density or frequency, categorical variables do not possess inherent numerical magnitudes. Consequently, these categorical inputs are converted into numerical representations through one-hot encoding. This encoding procedure converts each category into a binary vector, assigning a dedicated column to every category. After applying one-hot encoding, the original pre-treatment-type variable containing four categories is expanded into four separate columns, thereby generating three additional features. Each column represents a distinct pre-treatment condition and can take a binary value of either 0 or 1. For example, for an observation belonging to the ex-situ category, the ex-situ column is assigned a value of 1, while the remaining three columns are assigned values of 0. Through this process, categorical variables are transformed into binary numerical descriptors. Consequently, all four categories, including the reference condition, are displayed separately in **Figure 7a**. A red marker (binary value = 1) signifies that the corresponding conditioning strategy is active for a particular observation, whereas a blue marker (binary value = 0) indicates that the conditioning strategy is not present.

Among the investigated conditioning approaches, the ex-situ treatment exhibits the greatest influence on cell voltage. When this conditioning route is present, its SHAP values are predominantly positive and can reach approximately 0.05. Such behavior indicates a positive association between ex-situ activation and cell voltage, suggesting that this treatment contributes to an undesirable rise in voltage. The next most influential factor is the baseline, or reference, conditioning procedure. Like the ex-situ method, it also demonstrates a positive relationship with cell voltage. Nevertheless, the magnitude of its influence is smaller than that associated with mechanical activation, as reflected by the corresponding SHAP values. The positive SHAP contributions indicate that ex-situ activation tends to elevate cell voltage, thereby increasing voltage losses relative to the baseline condition. In contrast, the remaining two conditioning strategies display the opposite tendency. The SHAP distribution for in-situ/acids activation reveals a negative relationship with cell voltage, accompanied by an average SHAP value lower than that of the reference case. This observation supports the conclusions drawn in the previous sections, demonstrating that chemical activation of MEAs is effective in reducing cell voltage losses. Furthermore, among all activation procedures evaluated, the in-situ/water (90 °C) conditioning method delivers the most favorable overall performance.

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

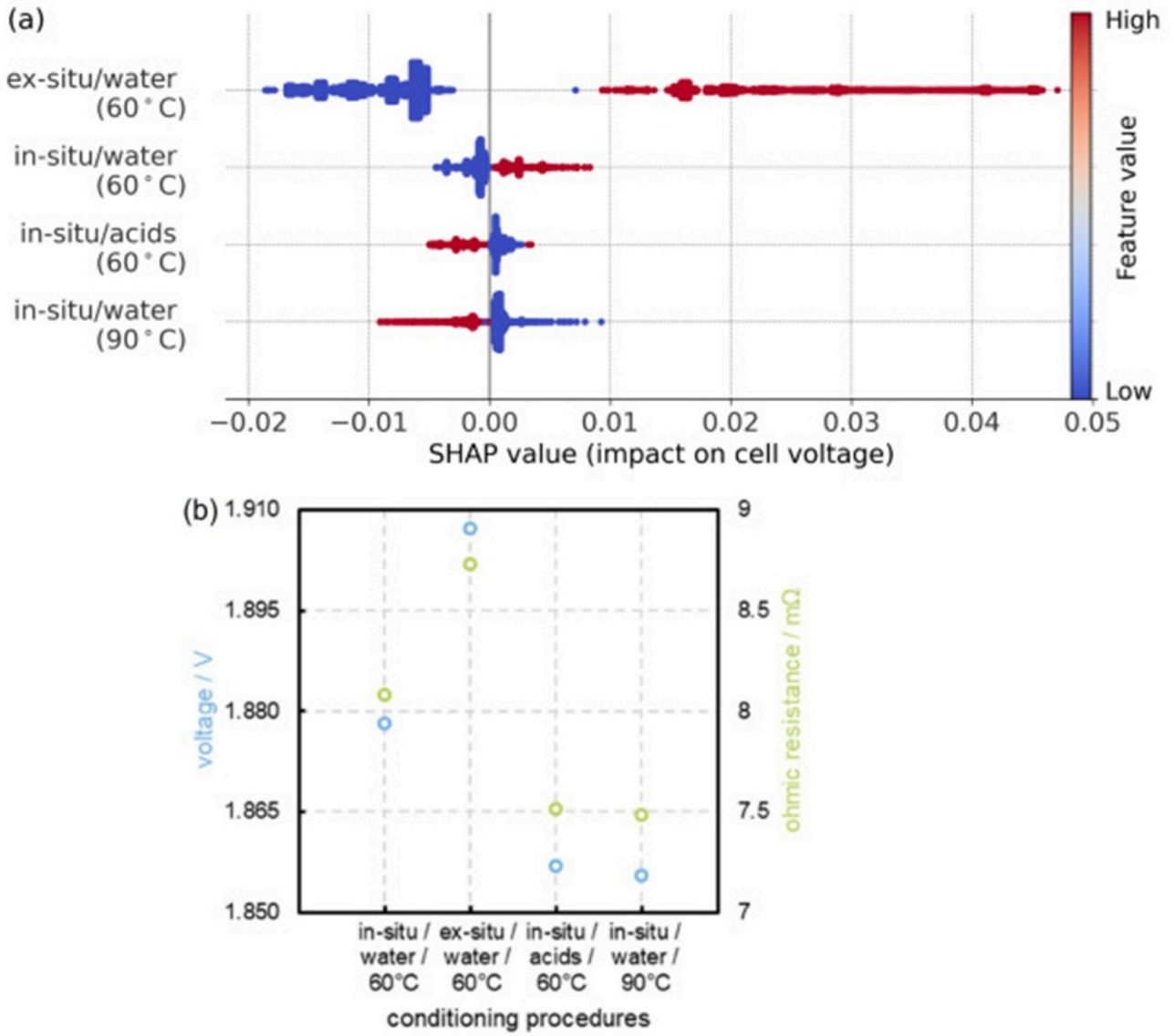


Figure 7. (a) Ranking different MEA conditioning procedures using an explainable methodology (SHAP). The ranking is estimated using the developed XGB data-driven model. (b) Impact of conditioning procedure on cell voltage and ohmic resistance at 2 A/cm² using direct experimental data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Figure adapted with permission from [11].

The preceding analysis demonstrated that ex-situ activation underperforms relative to the reference condition, whereas both in-situ/acids and in-situ/water (90 °C) activation strategies provide measurable improvements. Several factors may account for these differences in electrochemical behavior. In the ex-situ treatment, the membrane is hydrated outside the cell environment for 12 h before assembly. Although both the ex-situ and reference conditions involve exposure to water at an identical temperature, the structural state of the membrane after hydration is expected to be different. Under the reference protocol [11, 31], hydration occurs while the membrane is already integrated within the cell hardware and therefore constrained by the cell geometry. Under such conditions, membrane expansion occurs primarily in the through-plane direction. This mode of swelling promotes tighter interfacial contact between the catalyst layer and the porous PTL. By contrast, during ex-situ activation, the membrane expands without confinement, allowing swelling in both through-plane and in-plane directions. The resulting lateral expansion may adversely affect the catalyst–PTL interface after

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

assembly. Consequently, the ohmic resistance (RΩ) increases, leading to reduced performance compared with the reference case.

The beneficial effects of chemical activation can be explained by the role of the treatment reagents. Hydrogen peroxide is commonly employed to eliminate residual organic contaminants from catalyst layers in commercially produced MEAs [32]. The removal of these impurities lowers interfacial contact resistance. Sulfuric acid treatment contributes through a different mechanism by replacing metal cations present within the membrane with mobile H+ ions. This ion-exchange process increases the proton concentration available for transport, thereby enhancing proton conductivity. The improved performance observed for in-situ/acids activation can therefore be linked to the simultaneous reduction of interfacial losses and enhancement of proton transport within the membrane and ionomer. For the in-situ/water (90 °C) treatment, previous investigations have shown that hydration of Nafion membranes at elevated temperatures increases membrane water content [33]. Greater hydration results in more water molecules surrounding the sulfonic acid groups of the membrane, which facilitates ionic transport. Consequently, membrane conductivity improves, contributing to the superior cell performance obtained through the in-situ/water (90 °C) activation route.

It is also important to recognize the scope and limitations of the SHAP methodology. While SHAP analysis provides a framework for understanding how a machine learning model prioritizes and ranks input features, its conclusions must be assessed against experimental evidence. Such validation is necessary to ensure that the interpretations reflect physical reality rather than model artifacts, thereby increasing confidence in the analytical outcomes. **Figure 7b** compares the conditioning procedures using experimental data reported in [11]. At a current density of 2 A/cm2, the voltage distributions reveal a clear ordering of performance. The ex-situ activation approach produces higher cell voltages than the in-situ treatment conducted at 60 °C, indicating poorer performance. In comparison, the in-situ/acids procedure achieves lower voltages and thus better performance, while the in-situ/water treatment at 90 °C exhibits the most favorable behavior among the tested conditions. The consistency between these experimental observations and the rankings generated by SHAP analysis provides strong support for the validity of the model interpretation.

The findings discussed above highlight the influence of three activation strategies when benchmarked against the reference condition. Each approach offers specific advantages while also presenting certain drawbacks. Ex-situ membrane soaking [11] can increase water uptake and improve the membrane's initial ionic conductivity; however, uncontrolled swelling may introduce errors in catalyst loading and layer integrity [34]. The trends presented in **Figure 4** substantiate the negative consequences of this swelling-induced mechanical effect on overall cell operation. Chemical activation is advantageous because it removes contaminants and residual fabrication-related impurities, which contributes to enhanced electrochemical performance [11]. Nevertheless, additional studies employing spectroscopic characterization methods are required to more rigorously quantify its effectiveness. Thermal activation represents another promising strategy. Its positive influence on performance has been widely documented for PEMFC systems [8] and has more recently been reported for PEMEC applications as well [11]. The predictive results obtained from the present model align closely with these established observations. However, despite its demonstrated short-term benefits, the long-term implications of thermal pre-treatment remain uncertain and should be explored in future investigations.

Effect of pre-treatment on electrolyzer stability using LSTM

Monitoring cell voltage stability at a fixed current density can illuminate long-term system behavior, shedding light on degradation, membrane hydration, and catalyst utilization. A steady voltage trace over extended periods indicates healthy operation. We now examine voltage stability and compare the three pretreatment protocols. **Figure 8** plots voltage stability traces for ex-situ, in-situ/water (90 °C), and in-situ/acids activation. Voltage is recorded at an operating current density of 2 A/cm² over a time span from 58 to 206 h. Before reaching this set point, the cell experienced a succession of lower current densities that were gradually increased. Additionally, roughly every 20 h, EIS and polarization curve measurements were taken to track performance across current levels. All three conditioning approaches yielded stable

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

voltage profiles over the roughly 150 h of observation. For ex-situ activation, cell voltage edged from 1.923 V to 1.921 V between the onset of the first and second break-in procedures at 77.5 h and 101.3 h. A reading of 1.913 V was obtained just before the final break-in procedure commenced at 196.7 h. Across the full operating window, including the intermittent characterization steps, the total voltage decay remained below 10 mV. The other two protocols exhibited comparably steady patterns, underscoring their robustness. Voltage shifts of 5 mV and 17 mV were recorded for in-situ/acids and in-situ/water (90 °C) activation, respectively, between the first and last break-in procedures. In terms of relative standing, ex-situ pretreatment persistently registered the highest cell voltages, consistent with earlier observations. In-situ/acids activation proved the most stable when judged over extended operation. Although the pre-break-in voltage difference between ex-situ and in-situ/water (90 °C) treatments narrowed at certain times—most notably at 77.5 h and 196 h—the chemically treated MEA maintained a nearly flat profile, with voltages of 1.865 V and 1.863 V at those same two instants. LSTM model predictions for all three cases are overlaid in **Figure 8.** The model adheres closely to both training and testing segments, demonstrating its capacity to faithfully capture the system dynamics. Because the chemically activated MEA's voltage barely fluctuates across the full period, the LSTM model attains its lowest RMSE for this case: 0.0058 on training data and 0.0055 on test data. Test-set RMSE values for ex-situ and in-situ/water (90 °C) pretreatments were 0.0062 and 0.0077, respectively.

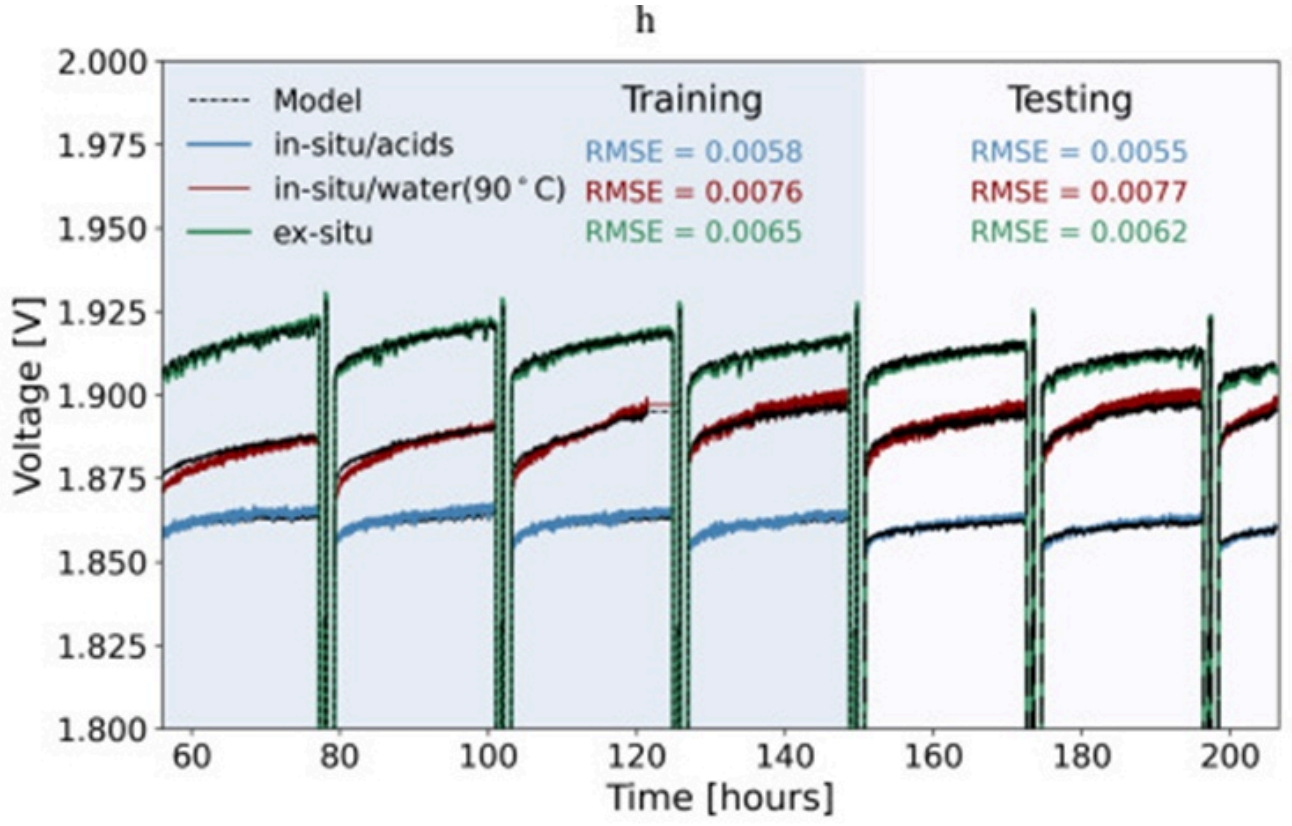


Figure 8. Cell voltage profiles for around 150 h of PEMEC operation at an applied current density of 2 A/cm². The colored lines and the black dashed lines represent the experimental data for the three pre-treatment procedures and model predictions, respectively. The background color grading indicates the data used for training and testing intervals. The performance of the LSTM model is noted through the RSME values. The comparison reveals distinct stability characteristics across the conditioning strategies. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

It is worth underlining that activation strategies can affect short-term and long-term performance in distinct ways. In the preceding section, thermal activation showed a marginal short-run advantage over chemical activation. Yet the traces in **Figure 8** reveal that this hierarchy inverts over longer timescales: chemical activation distinctly outperforms in-situ/water

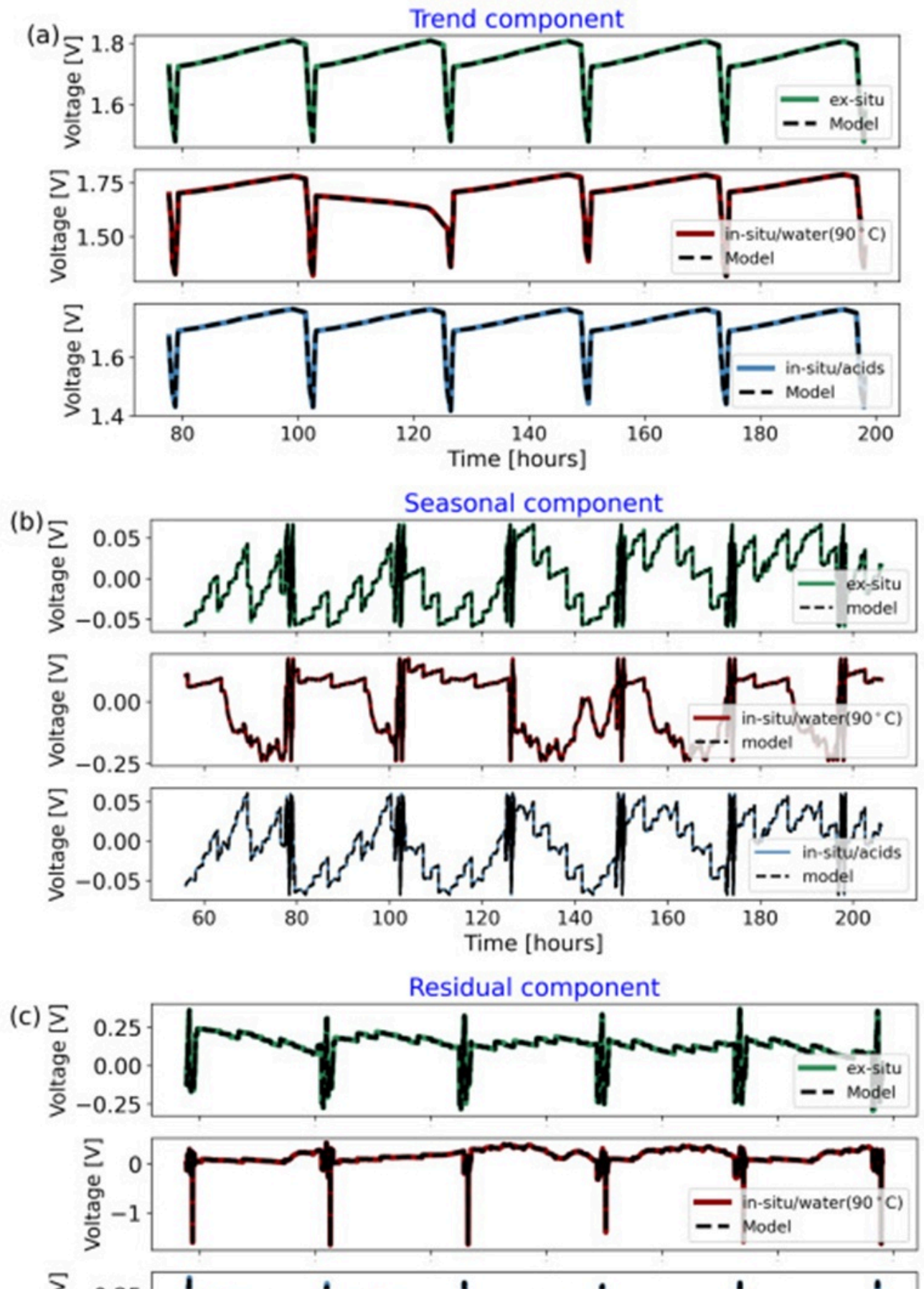
© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

(90 °C) during prolonged operation. While treating the MEA with 90 °C water temporarily elevates its hydration state, ongoing operation with 60 °C water causes that hydration state to gradually fall back. Consequently, cell voltage drifts upward over time and performance slips below that of the chemically activated MEA.

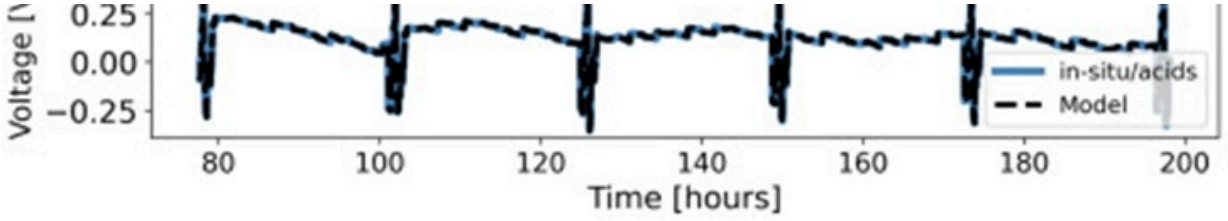
A time series can be broken down into several constituent parts—trend, seasonality, cyclicity, and residual—that combine additively or multiplicatively to describe the data's behavior. The trend reflects the series' long-run trajectory. Periodic, regular swings are captured by the seasonal component, while irregularly timed fluctuations belong to the cyclic component. The residual encapsulates random, unpredictable variation. Traditional time-series techniques [35] zero in on trend and seasonal components, largely because their systematic, repeating signatures lend themselves to extraction by classical statistical tools, which treat residuals as uncorrelated white noise. In practical forecasting, however, capturing at least some portion of the residual is often indispensable.

Figure 9 compares the (a) trend, (b) seasonal, and (c) residual components of experimental data alongside LSTM voltage predictions for the three conditioning protocols, assuming an additive decomposition. Overall, the voltage exhibits a gently rising trend across all three cases, a pattern that recurs after each break-in procedure, and the model tracks this evolution faithfully. The more erratic though periodic seasonal component, displayed in **Figure 9b,** is likewise well reproduced. Notably, the seasonal element for in-situ/water (90 °C) activation exerts a larger influence on the total voltage signal, oscillating with an amplitude near ±0.25 V compared with roughly ±0.05 V for the other two cases. A core strength of the LSTM model is its ability to replicate not only trend and seasonal parts but also the residuals, as **Figure 9c** attests. In every case, the residual component outweighs the seasonal component in its contribution to the voltage time series. The LSTM's capacity to forecast residuals flows from its rejection of the assumptions embedded in conventional methods [35]. Classical approaches regard residuals as purely random and devoid of correlation, yet in reality they frequently contain latent variable interactions, undetected autocorrelations, and nonlinear dependencies. Deep architectures like LSTM, armed with memory cells and gates, can tease out and learn these nonlinear patterns—even recalling past random excursions to guide future predictions. Furthermore, external inputs such as applied current density are fed to the LSTM, sharpening its residual predictions. As a result, these data-driven models can deliver more nuanced time-based voltage forecasts.

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press



© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press



Conclusion

The influence of four pretreatment protocols on PEMEC performance was studied through three machine learning methods—KNN, XGB, and ANN. A PEMEC built around a Nafion N115 membrane with an active area of 4.4 × 4.4 cm² provided the experimental foundation for model construction. Since the investigation centred on conditioning effects, operating parameters were held constant. Model quality was evaluated using RMSE and R^2 metrics. The main takeaways are as follows:

- (1)

All three models predicted cell voltage reliably over the range of current densities and pretreatment types. Divergences appeared for impedance predictions. KNN performed well on training data but overfit the test set. A limited EIS data volume hindered the ANN model's impedance accuracy. XGB delivered the most consistent performance for both voltage and impedance.

- (2)

Ex-situ conditioning prompts membrane swelling in both through-plane and in-plane directions, leading to inferior performance and elevated ohmic resistance compared with the reference. Both in-situ/acids and in-situ/water (90°) activation lowered ohmic resistance and cell voltage relative to the base case.

- (3)

A strong correspondence emerged between model predictions and experimental evidence: SHAP analysis ranked in-situ/water (90°) activation as the most effective short-term strategy, followed by in-situ/acids, the reference, and ex-situ, matching the experimental order reported in [11].

- (4)

LSTM-based time-series forecasts aligned well with measured data. In contrast to conventional forecasting tools, the LSTM proved capable of accurately reproducing the residual component of the time series.

Selecting the right conditioning protocol is vital not only for unlocking optimal cell performance but also for safeguarding electrolyzer longevity. The impact of a given pretreatment is timescale-dependent. Over short observation windows, thermally activated MEAs hold a slight edge over chemically activated ones. Stability analysis, however, demonstrates that in-situ/acids activation outperforms in-situ/water (90°) activation during extended runs. In this landscape, data-driven models serve as powerful tools for identifying the most suitable conditioning protocols, thereby cutting both time and experimental costs.

Nomenclature

ANN = Artificial Neural Network

CL = Catalyst Layer

CPE-T = Constant phase element - Exponent (P)

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

CPE-T = Constant phase element - Magnitude (T)

HER = Hydrogen Evolution Reaction

KNN = K Nearest Neighbors

LSTM = Long-Short Term Memory

MAE = Mean Absolute Error

MAE = Mean absolute error

MAPE = Mean Absolute Percentage Error

MEA = Membrane Electrode Assembly

ML = Machine Learning

MSE = Mean Square Error

OER = Oxygen Evolution Reaction

PEM = Proton Exchange Membrane

PEMEC = Proton Exchange Membrane Electrolytic Cell

PEMFC = Proton Exchange Membrane Fuel Cell

PTL = Porous Transport Layer

R^2 = Coefficient of determination

RCT = Charge transfer resistance

RMSE = Root Mean Square Error

Rs = Solution resistance

SHAP = Shapely Additive Explanations

SR = Split Ratio

SVM = Support Vector Machine

XGB = Extreme Gradient Boosting

Acknowledgements

The authors gratefully acknowledge the financial support by the German Federal Ministry of Education and Research (BMBF) within the H_2Giga project DERIEL (grant number 03HY122C). The authors gratefully acknowledge computing time on the supercomputer JURECA at the Forschungszentrum Jülich under the project no. PAJ2403. The authors would

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

also like to thank Sebastian B. C. Lehmann for his contribution to the graphical designs and Zheng Jiang for data processing which was used in performing equivalent circuit modeling.

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 22 Oct 2021 Revised: 09 Jan 2022 Accepted: 27 Jan 2022
Published online: 15 June 2022

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at https://creativecommons.org/licenses/by-nc-sa/4.0/.

References

1. Rasul MG, Hazrat MA, Sattar MA, Jahirul MI, Shearer MJ. The future of hydrogen: Challenges on production, storage and applications. Energy Convers Manage. 2022;272:116326.

2. Wang Y, Pang Y, Xu H, Martinez A, Chen KS. PEM fuel cell and electrolysis cell technologies and hydrogen infrastructure development–a review. Energy Environ Sci. 2022;15(6):2288-328.

3. Christmann K, Friedrich KA, Zamel N. Activation mechanisms in the catalyst coated membrane of PEM fuel cells. Prog Energy Combust Sci. 2021;85:100924.

4. Zamel N, Li X. Effect of contaminants on polymer electrolyte membrane fuel cells. Prog Energy Combust Sci. 2011;37(3):292-329.

5. Kusoglu A, Dursch TJ, Weber AZ. Nanostructure/swelling relationships of bulk and thin-film PFSA ionomers. Adv Funct Mater. 2016;26(27):4961-75.

6. Xie T, Page KA, Eastman SA. Strain-based temperature memory effect for nafion and its molecular origins. Adv Funct Mater. 2011;21(11):2057-66.

7. Zhiani M, Majidi S, Silva VB, Gharibi H. Comparison of the performance and EIS (electrochemical impedance spectroscopy) response of an activated PEMFC (proton exchange membrane fuel cell) under low and high thermal and pressure stresses.

Energy. 2016;97:560-7.

8. Qi Z, Kaufman A. Activation of low temperature PEM fuel cells. J Power Sources. 2002;111(1):181-4.

9. Siracusano S, Trocino S, Briguglio N, Pantò F, Aricò AS. Analysis of performance degradation during steady-state and load-thermal cycles of proton exchange membrane water electrolysis cells. J Power Sources. 2020;468:228390.

10. Wang W, Li K, Ding L, Yu S, Xie Z, Cullen DA, et al. Exploring the impacts of conditioning on proton exchange membrane electrolyzers by in situ visualization and electrochemistry characterization. ACS Appl Mater Interfaces. 2022;14(7):9002-12.

11. Wolf N, Javed A, Treutlein L, Kungl H, Karl A, Jodat E, et al. Tuning proton exchange membrane electrolytic cell performance by conditioning nafion N115-based membrane electrode assemblies. Electrochem Sci Adv. 2025;e202400038.

12. Chen X, Rex A, Woelke J, Eckert C, Bensmann B, Hanke-Rauschenbach R, et al. Machine learning in proton exchange membrane water electrolysis—A knowledge-integrated framework. Appl Energy. 2024;371:123550.

13. Ding R, Chen Y, Rui Z, Hua K, Wu Y, Li X, et al. Machine learning utilized for the development of proton exchange membrane electrolyzers. J Power Sources. 2023;556:232389.

14. Mohamed A, Ibrahem H, Yang R, Kim K. Optimization of proton exchange membrane electrolyzer cell design using machine learning. Energies. 2022;15(18):6657.

15. Chen L, Tian Y, Hu X, Yao S, Lu Z, Chen S, et al. A universal machine learning framework for electrocatalyst innovation: a case study of discovering alloys for hydrogen evolution reaction. Adv Funct Mater. 2022;32(47):2208418.

16. Timmermann J, Kraushofer F, Resch N, Li P, Wang Y, Mao Z, et al. IrO 2 surface complexions identified through machine learning and surface investigations. Phys Rev Lett. 2020;125(20):206101.

17. Flores RA, Paolucci C, Winther KT, Jain A, Torres JAG, Aykol M, et al. Active learning accelerated discovery of stable iridium oxide polymorphs for the oxygen evolution reaction. Chem Mater. 2020;32(13):5854-63.

18. Ding R, Zhang S, Chen Y, Rui Z, Hua K, Wu Y, et al. Application of machine learning in optimizing proton exchange membrane fuel cells: a review. Energy AI. 2022;9:100170.

19. Raman KA, Hammacher L, Kungl H, Karl A, Jodat E, Eichel RA, et al. Data-driven surrogate modeling for performance prediction and sensitivity analysis of transport properties in proton exchange membrane water electrolyzers. Appl Energy. 2025;386:125529.

20. Karyofylli V, Raman KA, Hammacher L, Danner Y, Kungl H, Karl A, et al. Elucidating parasitic currents in proton-exchange-membrane electrolytic cells via physics-based and data-driven modeling. Electrochem Sci Adv. 2025;e70000.

21. Satjaritanun P, O'Brien M, Kulkarni D, Shimpalee S, Capuano C, Ayers KE, et al. Observation of preferential pathways for oxygen removal through porous transport layers of polymer electrolyte water electrolyzers. iScience. 2020;23(12).

22. Yang R, Mohamed A, Kim K. Optimal design and flow-field pattern selection of proton exchange membrane electrolyzers using artificial intelligence. Energy. 2023;264:126135.

23. Günay ME, Tapan NA, Akkoç G. Analysis and modeling of high-performance polymer electrolyte membrane electrolyzers by machine learning. Int J Hydrog Energy. 2022;47(4):2134-51.

24. Woelke J, Rex A, Eckert C, Bensmann B, Hanke-Rauschenbach R. Predicting future polarization curves from operating data: Machine learning-based investigation of degradation modeling concepts for PEM water electrolysis. Energy AI. 2025;100547.

25. Ozdemir SN, Pektezel O. Performance prediction of experimental PEM electrolyzer using machine learning algorithms. Fuel. 2024;378:132853.

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

26. Ozdemir SN, Pektezel O. A CFD-driven machine learning approach for predicting PEM electrolyzer performance. Fuel. 2026;405:136479.

27. Harris CR, Millman KJ, Van Der Walt SJ, Gommers R, Virtanen P, Cournapeau D, et al. Array programming with NumPy. Nature. 2020;585(7825):357-62.

28. Hastie T, Tibshirani R, Friedman JH, Friedman JH. The elements of statistical learning: data mining, inference, and prediction. Springer; 2009.

29. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput. 1997;9(8):1735-80.

30. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. Nat Mach Intell. 2020;2(1):56-67.

31. Treutlein L, Javed A, Wolf NL, Kungl H, Karl A, Jodat E, et al. Gas crossover in membrane electrolyzers - The impact of MEA conditioning on gas permeability. J Electrochem Soc. 2025;172(6):064507.

32. Napoli L, Lavorante MJ, Franco J, Sanguinetti A, Fasoli H. Effects on nafion® 117 membrane using different strong acids in various concentrations. J New Mater Electrochem Syst. 2013;16(3):151-6.

33. Schuler T, De Bruycker R, Schmidt TJ, Büchi FN. Polymer electrolyte water electrolysis: correlating porous transport layer structural properties and performance: Part I. Tomographic analysis of morphology and topology. J Electrochem Soc. 2019;166(4):F270-F281.

34. Elliot P, Guido B, Shaun MA. Membrane Pretreatment and Cell Conditioning for Proton Exchange Membrane Water Electrolysis. In: ECS meeting abstracts. vol. 2021. 2021. p. 1252.

35. Alfaki MMA, Masih SB. Modeling and forecasting by using time series ARIMA models. Int J Eng Res Technol (IJERT). 2015;4(3):914-8.

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press