High-temperature polymer electrolyte membrane fuel cells (HT-PEMFCs) are enticing energy conversion technologies because they use low-cost hydrogen generated from methane and have simple water and heat management. However, proliferation of this technology requires improvement in power density. Here, we show that Machine Learning (ML) tools can help guide activities for improving HT-PEMFC power density because these tools quickly and efficiently explore large search spaces. The ML scheme relied on a 0-D, semi-empirical model of HT-PEMFC polarization behavior and a data analysis framework. Existing data sets underwent support vector regression analysis using a radial basis function kernel. Additionally, the 0-D, semi-empirical HT-PEMFC model was substantiated by polarization data; and synthetic data generated from this model was the subject to dimension reduction and density-based clustering. From these analyses, pathways were revealed to surpass 1 W cm⁻² in HT-PEMFCs with oxygen as the oxidant and CO containing hydrogen.

INTRODUCTION

High-temperature polymer electrolyte membrane fuel cells (HT-PEMFCs), which operate in the temperature range of 120 to 250 °C, offer simpler heat and water management when compared to today’s conventional low-temperature PEMFC variants. The elevated temperature operation of HT-PEMFCs minimizes, or even possibly eliminates, the fuel cell stack radiator, while also eliminating any need for feed gas humidification. Despite the advantages of HT-PEMFCs, they have been maligned over time for vehicle applications due to their low power density, high platinum group metal loadings, unsatisfactory stability, and limited temperature range and water tolerance. Because of these limitations, this fuel cell platform has largely been relegated to stationary power and niche applications.

In 2016, Los Alamos National Laboratory reported a superior polymer electrolyte membrane (PEM) and ionomer binder for HT-PEMFCs based upon a phosphoric acid (H₃PO₄) imbibed polycation. Unlike the standard bearer H₃PO₄ containing polybenzimidazole (PBI) variant, the electrostatic interactions between the tethered cation moiety and phosphate anion in
H$_3$PO$_4$ imbibed polycations permitted the use of wider temperature and humidity ranges without worry of H$_3$PO$_4$ evaporation and leaching. Venugopalan et al. improved upon this concept through the blending of PBI with the polycation to facilitate more H$_3$PO$_4$ uptake and greater ionic conductivity (area specific resistance (ASR) < 0.02 Ω-cm$^2$). The membrane blend was stable at 220 °C and had tolerance to 40% relative humidity at 80 °C while showing remarkably high ionic conductivity over the temperature range of -40 to 240 °C. Notably, both groups attained promising power density values near 0.7 W cm$^{-2}$ with pure oxygen as the oxidant. More recently, power density values as high as 1.2 W cm$^{-2}$ were achieved with a H$_3$PO$_4$/poly(arylene) type of polycation membrane and ionomer binder. The advent of this ion-pair polymer electrolytes for HT-PEMFCs has renewed interest in it for vehicular applications. But, success for this technology still necessitates improved power density, especially with air as the oxidant, while also reducing platinum group metal loadings (e.g., < 6 gPt/vehicle; rated at 90 kW$_{net}$ power).

Perhaps the most significant barrier to better HT-PEMFC power density is the presence of liquid H$_3$PO$_4$ present in the electrode layers with ionomer binder. The phosphate anions in H$_3$PO$_4$ can adsorb to the platinum electrocatalyst surface, blocking reactant sites, and thus causing large activation overpotentials. Furthermore, the presence of liquid acid and hydrocarbon binder stymie delivery of gas reactants to the electrocatalyst surface leading to large concentration overpotentials. These sources of overpotential significantly hamper the power density of HT-PEMFCs and need to be addressed if this platform will ever emerge as a serious competitor for powering light-duty and heavy-duty vehicles (LDVs and HDVs).

Materials science and engineering has played a central role in the evolution of improved electrochemical technologies like fuel cells and batteries. However, the design of new materials is often ‘Edisonian’ and the timeline from discovery to implementation in commercial devices is costly and far too long. The emergence of new Machine Learning (ML) methods across a variety of disciplines (e.g., chemical processes and biology), combined with the ubiquity of powerful computers, has motivated efforts in using these tools for streamlining the timeline from materials discovery to commercialized products. In the context of electrochemical energy storage and conversion, ML has been adapted to capture molecular interactions of materials and relating them to bulk properties so that potential candidates for components and devices can be quickly identified. It is also effective for predicting device performance from experimental operational data. More recently, the implementation of ML approaches coupled with single components modeling is being explored as way to bridge the gap between materials properties prediction and device performance. However, such strategies have not been used for polymer electrolyte membrane based electrochemical systems (e.g., fuel cells and electrolyzers).

In this work, ML and data driven analysis tools were leveraged to identify HT-PEM material properties for achieving high power HT-PEMFCs (∼ 1 W cm$^{-2}$ with oxygen as the oxidant). The approach first starts with a physics-informed, 0-D model that is shown to predict HT-PEMFC polarization behavior data with a few parameters estimated using a Jaya as optimization algorithm. Support vector regression using a radial basis kernel was performed to relate material attributes, such as ion-exchange capacity (IEC), to material properties (H$_3$PO$_4$ uptake and ionic conductivity). This activity informed the ohmic overpotential and parts of the concentration overpotential contributions in the 0-D model that describes HT-PEMFC polarization; thus connecting the materials’ properties prediction to device performance. Verifying the model with existing data sets instilled confidence for generating data from the model for further ML activities using density based clustering via Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) and dimensionality reduction via Uniform Manifold Approximation and Projection (UMAP). These activities revealed that in order to surpass the 1 W cm$^{-2}$ power density objective with air as the oxidant, the oxygen permeability coefficient would need to improve by 2.2x while maintaining facile proton conductivity (∼0.2 S cm$^{-2}$) in the electrode layers. Achieving such a breakthrough could be achieved by new permeable and highly conductive cathode ionomer binders and by applying 187 kPa$_{abs}$ of back pressure on the cell. It is worth mentioning that this pathway for a high-power HT-PEMFC necessitated 218 °C operating temperature – which is possible with acid imibed polycation materials but not plausible with the standard bearer phosphoric acid containing polybenzimidazole materials. Further, the simulated HT-PEMFC with ∼ 1 W cm$^{-2}$ operated with a hydrogen fuel stream containing CO to 12% and had a platinum catalyst.
loading in the electrodes as low as 0.23 mg/cm\(^2\). These attributes are particularly important for addressing fuel cell capital costs, large-scale manufacturing and running the fuel cell on low-cost hydrogen derived from steam reformed methane. In summary, ML and modeling at various scales (e.g., materials properties to device performance) was developed for HT-PEMFCs and was shown to be effective with existing data sets and for identifying pathways to achieve high power density.

RESULTS

For a modeling and data analysis framework to enable the study of materials properties and its effects on the fuel cell performance, it must be implemented in a suitable platform in which the materials models can be embedded. Also, it is desirable that the model implementation allow for the generation of large data sets so that a great number of possible configurations can be explored, with low computational cost. Hence, in this work a Python based, zero-dimensional, steady state, continuum approach was adopted. This model served as a bridge between the materials experimental data and ML predictive models and the system level data analysis so that patterns found using ML can be related to the materials properties in addition to fuel cell configurations and operating conditions. The result of this analysis is a set of clusters that can be associated to the system performance in terms of power density. Furthermore, the distribution of the clusters can be connected to physically meaningful variables, thus informing new experiments and possible designs.

Here, we selected HDBSCAN clustering preceded by UMAP for dimension reduction to find patterns (i.e. clusters) in the data associated with high power density values and to identify the more contributing variables to the clustering. We implemented HDBSCAN with 35 min_samples, and UMAP with 10 n_neighbors, 0.1 min_dist, and 3 n_components. The resulting clusters were further analyzed using the subspace greedy search (SGS) algorithm described by Zhu et al\(^{27}\). The variables identified with the largest contributions to the separation between the classes are the same as those with the higher first-order sensitivity indices from the Sobol's global sensitivity analysis as described in the SI. Furthermore, a set of candidate combinations for the improvement of the fuel cell performance measured as peak power density was found from the cluster analysis.

Figure 1. Schematic representation of the modeling and data analysis approach applied to study the use of ML methods for the fast development of new fuel cell designs.

Figure 1a presents a scheme of the modeling and ML approach for identifying electrochemical material property attributes for improving HT-PEMFC power density. A salient aspect of this Figure is the methods of learning (i.e., a circular loop) that feed into the 0-D, semi-empirical model that describes single-cell HT-PEMFC polarization behavior. The 0-D, semi-empirical model was built off existing HT-PEMFC models that are physics-informed with distributed parameters\(^{28,29,30}\). But, the model also features some semi-empirical expressions\(^{31,32,33}\) to describe concentration polarization overpotential losses that hail from transport phenomena in complex porous structures. The lumped model was selected because it accurately predicts data with few adjustable parameters, and it has relevant descriptors that factor how H\(_3\)PO\(_4\) content in the membrane and CO in the feed gas stream impact HT-PEMFC polarization. The
key assumptions for the model were: continuum level, steady state operation, ideal gas behavior, no gas crossover, no water transport considerations, and uniform distribution of catalyst and ionomer binder and gas transport properties across the electrodes and gas diffusion layers. The SI presents the model equations that include mass balances and reaction rate expressions, parameters for these equations, and descriptions for individual overpotential terms that govern HT-PEMFC polarization. Notably, a new semi-empirical expression was deployed to account for the ionomer binder gas transport (i.e., permeability) and conductivity properties.

**Machine learning embedded materials modeling**

A wide variety of ML methods and implementations exist for data analysis. The selection of the appropriate method must consider aspects like the amount of data available, the purpose of the implementation, and the computational resources. A notable method are the so-called Support Vector Regression Machines or SVRM. The idea of SVRM was first introduced by Drucker et al. 37 as an application of the concept of support vector machines (SVM) first proposed by Boser, Guyon and Vapnik 38. An SVM is an algorithm that looks for the hyperplane that separates two linearly separable data classes maximizing the distance between the hyperplane and the classes. If the data is not linearly separable, then a transformation, known as kernel function, is introduced to map the data into a new space in which it is linearly separable. This feature makes SVMs particularly flexible, as it can be applied to a wide range of datasets. Hence, SVRM was selected for the prediction of bulk material properties (e.g., ionic conductivity) from material characteristics (e.g., ion exchange capacity) under different operating conditions (e.g., temperature) because data was available for bulk HT-PEMs and H_3PO_4 imbibed polycation thin films. Although the Nernst-Planck relationship capture these parameters on ionic conductivity for ideal liquid electrolytes 39, PEMs feature non-ideal behavior and complexity associated with their solid-state structure 40 and concentrated ionic groups. The SVRM was useful for determining the unknown underlying relationship between IEC and temperature and ionic conductivity in the acid imbibed polymer electrolytes.

![Figure 2](image)

*Figure 2. Model for ionic conductivity as a function of temperature and IEC using a support vector regression with a radial basis function kernel. (a-b) polymer electrolyte membrane and (c-d) thin ionomer films that are similar in thickness to the ionomer binder in electrode layers.*

To demonstrate the feasibility of support vector machines for relating modeling material property data that would feed into a device level model, the approach created by Chang et
al.\textsuperscript{41} was adopted. The SI provides a detailed description of the method and thin film ionomer conductivity data measured on interdigitated electrodes (IDEs), and IEC values of the high-temperature polymer electrolyte materials used as thin films and electrode binders (Tables S4 and S5). Bulk HT-PEM ionic conductivity and IEC data was taken from our previous publication\textsuperscript{11}. Figure 2 conveys that the support vector machine models for both the polymer electrolyte membrane and the ionomer binder (characterized as a thin film) give an accurate prediction for the ionic conductivity as a function of material IEC and environmental temperature. These models, which were informed by data, are embedded into the 0-D, semi-empirical HT-PEMFC device model, so that the polarization behavior of the fuel cell can be predicted accurately from materials’ characteristics.

It is recognized that a limited demonstration was only performed with support vector machine models using material IEC and ionic conductivity. Other materials, and their attributes, are also important – e.g., electrocatalysts’ electrochemically active surface area and reactivity (e.g., exchange current and symmetry factor) and reactant gas permeability/diffusion in catalyst layers. Our future work will look to factor in these parameters and considerations into the polarization model. Attaining reaction kinetic and gas transport data will reduce the number of estimated parameters in the fuel cell model and will assist in realizing a more robust and rigorous prediction of HT-PEMFC polarization behavior. Here, ionic conductivity of the electrode binder and HT-PEM were examined as a demonstration of the SVRM effectiveness and the availability of data.

**Model validation**

To assist ML and model validation, both material property data and HT-PEMFC polarization data were included. Material attributes and property data were analyzed using a SVRM in the previous section. The model parameters were estimated using the Jaya optimization algorithm\textsuperscript{34}. Figure 3 presents the polarization curve from the model predictions and the polarization data. Table S6 shows the membrane electrode assembly (MEA) properties and gas reactant flow rates for polarization data presented in Figure 3. Table S7 presents the predicted parameters from the Jaya parameter estimation method for the HT-PEMFC model that successfully described the experimental polarization behavior from our previous work\textsuperscript{11}. It is worth mentioning that this is the first time HT-PEMFC performance has been modeled and substantiated with H\textsubscript{3}PO\textsubscript{4} imbibed polycations and polycation/PBI blends. These materials enable operation at temperatures above 180 °C for the HT-PEMFC – which currently cannot be operated or considered with H\textsubscript{3}PO\textsubscript{4} containing polybenzimidazole (PBI).

Figure 3 also shows the individual overpotential terms as a function of current density from the HT-PEMFC model. These overpotential terms govern the polarization of the HT-PEMFC. Conveying the individual overpotential terms against the HT-PEMFC polarization curve makes it possible to pinpoint the greatest sources of resistances that hamper fuel cell power density. As expected, the activation overpotentials, which arise from electrode kinetics – mainly the sluggish oxygen reduction reaction, cause the largest source of polarization in the cell when extracting up to ~ 2 A cm\textsuperscript{-2} of current density. Beyond this value, the cell approaches the limiting current and concentration polarization overpotential dominate cell performance. Operating the cell at higher temperatures (220 °C as opposed to 200 °C) promoted reaction kinetics and gas permeability resulting in a reduction in activation and concentration overpotentials. As such, the cell displayed less polarization and a 58% improvement in peak power density.
Figure 3. Model predicted and experimental polarization curves and the activation, ohmic, and concentration overpotential terms. Shaded bands represent model variability measure as $f_0$ value from the parameter global sensitivity analysis.

The 7 model parameters (see Table S7) to describe HT-PEMFC polarization were estimated using the Jaya optimization algorithm. The initial values were those reported from the literature for the previously reported data and defined by inspection for the case of $K_{cat}$. As a population based heuristic parameter estimation method, Jaya will yield different results for various runs. Therefore, estimation of the parameter in the model required running the algorithm for 10 iterations and 100 runs to attain a measure of the variability of the algorithm (see Figure S3). The standard error for the estimation was 0.018 V which, when combined with the convergence and stability of the parameter estimation algorithm, provides strong confidence on the model performance. We also used Sobol’s global sensitivity analysis to identify which parameters were the most influential on the variability of the model (see Figure S1 and Figure S2). These results provide information about the physical significance of the parameters involved in the model. The charge transfer coefficient for both the cathode and the anode, as well as the empirical parameter for the acid uptake in the concentration overpotential, were shown to be the most influential on polarization behavior in the low current density regime.

Design space exploration using Machine Learning based visualization

Figure 4. Clustering visualization for the synthetic data: (a) clusters generated by HDBSCAN with preprocessing using UMAP (cluster 1 in green; cluster 2 in cyan; cluster 3 in pink; cluster 4 in blue; cluster 5 in yellow), (b) clusters with power density superimposed as color map.
When selecting an appropriate ML implementation to inform the design of physical system several considerations must be made. Firstly, due to the explorative nature of the task, no assumptions can be made regarding the shape of the clusters. In addition to this, the lack on a priori knowledge about the data and how it should look makes it difficult to establish the number of clusters needed. To deal with these issues, a density-based clustering (DBC) approach was adopted. DBC finds clusters of any shape, as opposed centroid based clustering which assumes a shape (i.e. a sphere of equal variance) for the clusters. Also, in DBC, data points that are in sparse regions are not required to be assigned to a cluster and are instead identified as noise. Therefore, DBC gives us two advantages: (1) we remove the spherical equal variance assumption, allowing for different shapes in our clusters, and (2) we are now able to detect and deal with noise. However, the resolution parameter (namely the radius of the region that is being considered as dense or sparse) is hard to get, especially in higher dimensional data that we cannot easily visualize. This challenge can be overcome by introducing a hierarchical component. Hierarchical clustering finds nested relationships amongst data instead of flat partitions in the data. This characteristic allows for a more complex clustering with higher resolution by further diving clusters until a stopping criterion is met. Thus, combining HC and DB we get a clustering algorithm that makes no assumptions about the shape of the clusters and uses hierarchical clustering to improve the resolution eliminating the need of finding a correct resolution parameter.

In HDBSCAN, as introduced by Campello et al.\textsuperscript{24}, the density is estimated by finding the radius of the region for which the said region is considered dense using mutual reachability as distance measure. The stopping criterion for the hierarchical approach is how long a cluster persist in time before it splits into two clusters or vanishes into noise. The computational cost is, however, is $O(n^2)$ which means it does not scale well. In this work, we use the implementation proposed by McInnes and Healy\textsuperscript{25} in which Dual Three Boruvka’s for Euclidian Minimum Spanning Trees algorithm is used to enhance performance, bringing the computational cost to $O(N\log N)$.

Since HDBSCAN in based on density, we can expect that the performance would decrease with highly sparse data. Also, HDBSCAN performance suffers with high dimensional data. Therefore, using a dimension reduction pre-processing could enhance the clustering output. UMAP is a neighbor graphs-based dimension reduction technique, meaning that a graph is first build for the high dimensional data and then it is embedded in a low dimensional space using a force directed layout using cross entropy to measure the distance between the high dimensional graph and the low dimensional graph. The resulting low dimensional embedding is denser, providing a better starting point for HDBSCAN\textsuperscript{26}.

With an experimentally substantiated semi-empirical HT-PEMFC model for ion-pair polymer electrolytes, it was possible to explore different configurations using a stochastic generator. To produce such configurations 12 input variables (hydrogen stoichiometric ratio, $S_{H2}$; oxygen stoichiometric ratio, $S_{O2}$; temperature, $T$; pressure, $P$; membrane ion exchange capacity, $\text{IEC}_{\text{mem}}$; ionomer binder ion exchange capacity, $\text{IEC}_{\text{io}}$; membrane thickness, $\delta_{\text{mem}}$; ionomer binder thickness, $\delta_{\text{io}}$; carbon monoxide to hydrogen ration in feed stream, $\text{CO}/H_2$; and platinum loading, $L_i$) were allowed to take different values around the base case (see Table S1) in a uniform distribution within a given interval (e.g., ± base value×40%). For example, a given configurations would include a lower catalyst loading, $L_i$, and higher temperature and pressure while another configuration would have the same catalyst loading but with a smaller membrane thickness, $\delta_{\text{mem}}$. The learning agent was then exposed to a variety of scenarios from the generated simulations from which it can then draw and identify patterns and then evaluate the likelihood of new configurations that performs well. The clusters of data can then be associated with desired objectives, such as higher values of peak power density, and then the variables that exert greater influence on the distance between clusters can be revealed and guide the factors and material properties that lead to improved HT-PEMFC power density. To illustrate this concept, Figure 4 shows the visualization for the clustering analysis using HDBSCAN with UMAP as pre-processing. Three major regions are seen in the cluster visualization: a region of low power density (cluster 1 in Figure 4a, and deep blue in Figure 4b), a high power density region in cluster 5, and a transition region in between. Cluster 5 contains configurations that achieve higher power density values. Furthermore, applying the SGS analysis (see Table S2 and Table S3) the variables $S_{O2}$, $T$, and $\text{IEC}_{\text{io}}$ were identified as the most influential in the formation of the clusters. The variables designated by the clustering...
and SGS analyses as the more influential over the peak power density are the same as those with the higher first-order index in the modified Sobol’s global sensitivity analysis - which serves as cross validation for the UMAP-HDBSCAN clustering and SGS analysis approach. With this, not only are a set of candidate configurations (values for design and operating variables) attained, but it also provides insights as to which variables should be prioritized for enhancing HT-PEM performance and efficiency.

**DISCUSSION**

The ML based modeling and analysis framework presented here enables fast identification of materials properties and device operating parameters that enhance HT-PEMFC performance. It is also worth noting that this approach could work even in the absence of any model since the clustering analysis can be performed using only experimental data. The hierarchical structure of the modeling strategy gives the modeler a flexible and powerful tool, in which the ML model are informed from basic physical knowledge of the HT-PEMFC. The model is also conducive for upgrading with additional descriptors as more data is generated from experiments and/or made available in the literature and publicly accessible databases. Here, the data was primarily confined to the ionic conductivity and IEC of the HT-PEM and electrode binders and HT-PEMFC polarization. Future work will look to incorporate electrode reactivity and reactant/product species transport rates. Additionally, this work deployed density based clustering through the HDBSCAN combined with UMAP implementation and these tools allowed for the extraction of the underlying distribution of the data, that is, an accurate description of the polarization behavior of the system. From the cluster analysis, cluster 5 is the one with more combinations that are likely to surpass the 1 W cm\(^{-2}\) goal with air as the oxidant. Although some of the candidate combinations will be deemed as unfeasible either because of practicality concerns, high cost or materials with desired properties cannot be realized, the search space for new configurations was significantly reduced, reducing the amount of experimental permutations to be performed; and thus, accelerating the development of HT-PEMFC technology.

**EXPERIMENTAL PROCEDURES**

**Resource Availability**

**Lead Contact**

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Christopher G. Arges ([carges@lsu.edu](mailto:carges@lsu.edu)).

**Materials Availability**

This study did not generate new unique reagents.

**Data and Code Availability**

The datasets and code generated during this study are available at GitHub: [https://github.com/lbrice1/FCSDAT.git](https://github.com/lbrice1/FCSDAT.git)

**Materials synthesis and characterization**

The SI provides the details about the materials synthesis for QPPSf. The data for the bulk HT-PEM ionic conductivity and corresponding H\(_3\)PO\(_4\) uptake were taken from a previous publication by Venugopalan et al\(^{11}\). The HT-PEMFC polarization data was also taken from the same publication.

The ML and data analysis performed in this work was augmented with new experimental data that examined thin film ionic conductivity of H\(_3\)PO\(_4\) imbibed polycations (i.e., quaternary benzyl pyridinium poly(arylene ether sulfone) (QPPSf)). QPPSf of varying IEC were synthesized based as described in the literature\(^{42,43}\). The SI provides a succinct synthesis procedure. Interdigitated electrodes (IDEs) used for measuring thin film ionic conductivity were manufactured using the procedure by Arges et al.\(^{44,45}\). Thin films of QPPSf were deposited on the IDEs by spin coating a 1 wt% solution of QPPSf dissolved in n-methyl-2-pyrrolidone (NMP) at 4000 rpm for 45 seconds. The IDE with thin film QPPSf were then heated at 120 °C in nitrogen atmosphere to remove the excess NMP from the samples. H\(_3\)PO\(_4\) was imbibed into the thin films by placing a drop of 85wt% H3PO4 on the thin films for 10 minutes. The excess acid was removed by blot drying the IDEs carefully. The electrode pads of the IDE substrate were scraped away using a cotton Q-tip to make electrical connections. The H\(_3\)PO\(_4\) imbibed
Polycation thin film resistance was determined using electrochemical impedance spectroscopy (EIS). The frequency range was set to 100,000 to 1 Hz with an oscillatory amplitude of 0.0001 mA. The equation and IDE dimensions used to calculate in-plane ionic conductivity of the thin films are given in equation 1 and the descriptions below.

\[
\kappa = \frac{1}{R} \frac{d}{[N(N-1)]^l} \tag{1}
\]

Here:
- \( \kappa \): in-plane ionic conductivity
- \( R \): in-plane ionic resistance
- \( d \): spacing between teeth on IDE (100 \( \mu \text{m} \))
- \( l \): length of teeth on IDE (4500 \( \mu \text{m} \))
- \( t \): Polycation film thickness on IDE substrate
- \( N \): number of teeth on IDE substrate (22)

Thickness of the ionomer film coated on IDEs were measure using Ellipsometry (RC2 Ellipsometer). The thickness values of IDEs with QPPSf of different IEC ranged from 10 – 14 nm.

A brief procedure for IEC calculation of pristine QPPSf (i.e., not containing H$_3$PO$_4$) is provided in the SI section. The H$_3$PO$_4$ acid uptake of the new QPPSf ionomers synthesized here and targeted as electrode binders were determined by first preparing these materials as binders for gas diffusion electrodes. The gas diffusion electrodes were fabricated by spray painting catalyst ink on Toray carbon paper. The procedure for catalyst ink preparation was followed from our previous work. A brief description of the procedure as follows: 0.2 g of carbon support catalyst (37% Pt in high surface area carbon, Tanaka Kikinzoku International) with 1.715 g of QPPSf ionomer solution dissolved in 5.5 g of reagent alcohol are blended together and sonicated for 30 minutes to provide a completely dispersed catalyst ink solution. The prepared GDEs were dried at 80 °C to remove residual solvent. The catalyst loading was maintained as 0.5 mgPt cm$^{-2}$. The weight of the non-imbibed GDEs were recorded gravimetrically followed by immersing them in 85 wt% H$_3$PO$_4$ for 10 minutes for acid imbibing. After the acid imbibing step, excess liquid acid was carefully removed from the GDE surface via blot drying. Then, the acid imbibed GDEs were weighed to determine the acid uptake in the electrodes.

**Subspace Greedy Search algorithm**

The SGS algorithm seeks the combination of the most contributing subspace of variables which causes the separation between the clusters. To do this, SGS searches from the lowest dimension and moves to higher dimensions until no possible subspaces are left. A testing score, defined as the k-distance between clusters for each subspace, allows for comparison between different subspaces at the same dimension. The subspaces with the highest scores are kept for further testing and the rest are discard. Further details about the SGS implementation can be found. This analysis can be performed to explore the relationships between all the clusters and specially the clusters of interest, in this case, the clusters with higher power density. Based on which variables appear more frequently in the contributing subspaces for different clusters, the variable classification can be drawn for the underlying distribution. (SI provides more details about the methodology).

**Parameter Estimation**

When implementing a parameter estimation method, the stochastic or deterministic nature of the method, the complexity of the algorithm, the accuracy, and the convergence speed must be considered. Typically, this will lead to a trade-off between accuracy on one hand and stability, complexity, and convergence speed on the other hand. To address this trade-off, a systematic reduction of the number of parameters to be estimated was enacted by using a global sensitivity analysis. This enhances the performance of a population based parameter estimation method. The Jaya implementation for our model is described in the SI and is based off a literature precedent.

**SUPPLEMENTAL INFORMATION**
REFERENCES

1. Rosli, R., Sulong, A., Daud, W., Zulkifley, M., Husaini, T., Rosli, M., Haque, M. (2017). A review of high temperature proton exchange membrane fuel cell (HT-PEMFC) systems. International Journal of Hydrogen Energy, 42, 9293-9314.

2. Xiao, L., Scanlon, E., Ramanathan, L., Choe, E., Rogers, D., Apple, T., & Benicewicz, B. (2005). High-Temperature Polybenzimidazole Fuel Cell Membranes via a Sol-Gel Process. Chemistry of Materials, 17(21), 5328-5333.

3. Lee, K.-S., Maurya, S., Kim, Y., Kreller, C., Wilson, M., Larsen, D., Mukundan, R. (2018). Intermediate temperature fuel cells via an ion-pair coordinated polymer electrolyte. Energy & Environmental Science, 11(4), 979-987.

4. Lee, A., Choe, Y.-K., Matanovic, I., & Kim, Y. (2019). The energetics of phosphoric acid interactions reveals a new acid loss mechanism. J. Mater. Chem. A, 7, 9867-9876.

5. Quatarone, E., Angioni, S., & Mustarelli, P. (2017). Polymer and composite membranes of proton-conducting, high-temperature fuel cells: a critical review. Materials, 10.

6. Li, Q., Jensen, J., Savinell, R., & Bjerrum, N. (2009). High temperature proton exchange membranes based on polybenzimidazoles for fuel cells. Progress in Polymer Science, 34(5), 449-477.

7. Lee, K.-S., Spendelow, J., Choe, Y.-K., Fujimoto, C., & Kim, Y. (2016). An operationally flexible fuel cell based on quaternary ammonium-biphosphate ion pairs. Nature Energy, 1(9), 16120.

8. Wainerwright, J., Wang, J., Weng, D., Savinell, R., & Litt, M. (1995). Acid-doped polybenzimidazoles: a new polymer electrolyte. Journal of the Electrochemical Society, 142(7), 1211-L-123.

9. Pingitore, A., Molleo, M., Schmidt, T., Benicewicz, B., Lipman, T., & Webber, A. (2019). Polybenzimidazole fuel cell technology: Theory, performance, and application. In Fuel Cells and Hydrogen Production: A Volume in the Encyclopedia of Sustainability Science and Technology (pp. 477-514). New York, N.Y.: Springer New York.

10. Chaichi, A., Venugopalan, G., Devireddy, R., Arges, C., & Gartia, M. R. (2020). A solid-state and flexible supercapacitor that operates across a wide temperature range. ACS Appl. Energy Mater., 3(63), 5693-5704.

11. Venugopalan, G., Chang, K., Nijoka, J., Livingston, S., & Geise, G. M. (2020). Stable and highly conductive polycation-polybenzimidazole membrane blends for intermediate temperature polymer membrane fuel cells. ACS Appl. Energy Mater., 3(7), 573-585.

12. Bain, H., Peng, H., Xiang, Y., Zhang, J., Wang, H., Lu, S., & Lin, Z. (2019). Poly(arylene piperidine)s with phosphoric acid doping as high temperature polymer electrolyte membrane for durable, high-performance fuel cells. Journal of Power Sources, 443, 227219.

13. Kongkanand, A., & Mathias, M. (2016). The priority and challenge of high-performance proton-exchange fuel cells. Journal of Physical Chemistry Letters, 7(7), 1127-1137.

14. Hu, Y., Jiang, Y., Jensen, J., Cleemann, L., & Li, Q. (2018). Catalyst evaluation for oxygen reduction reaction in concentrated phosphoric acid at elevated temperatures. Journal of Power Sources, 375, 77-81.

15. Mamlouk, M., & Scott, K. (2010). The effect of electrode parameters on performance of a phosphoric acid doped PBI membrane fuel cell. International Journal of Hydrogen Energy, 35(2), 784-793.

16. Correa-Baena, J.-P., Hippalgaonkar, K., van Duren, J., Jaffer, S., Chandrasekhar, V. R., Stevanovic, V., Bounassi, T. (2018). Accelerating materials development via automation, machine learning, and high-performance computing. Joule, 2(August 15), 1410-1420.

17. Wang, Y., Seo, B., Wang, B., Zamel, N., Jiao, K., & Cordobes Adroher, X. (2020). Fundamentals, materials, and machine learning of polymer electrolyte membrane fuel cell technology. Energy and AI, 1, 100014.

18. Ong, S. P. (2019). Accelerating materials science with high-throughput computations and machine learning. Computational Materials Science, 161, 143-150.

19. Blanco, D. E., Lee, B., & Modestino, M. A. (2019). Optimizing organic electrolysintes through controlled voltage doping and artificial intelligence. PNAS, 116(36), 17683-17689.

20. Gu, G. H., Noh, J., Kim, I., & Jung, Y. (2019). Machine learning for renewable energy materials. J. Mater. Chem. A, 7, 17096-17177.

21. Bao, J., Wang, C., Xu, Z., & Koeppel, B. J. (2019). Physics-informed machine learning with applications to solid oxide fuel cell system modeling and optimization. Richland, Washington: Pacific Northwest National Lab.

22. Wang, B., Xie, B., Xuan, J., & Kui, J. (2020). AI-based optimization of PEM fuel cell catalyst layers for maximum power density via data-driven surrogate modeling. Energy Conversion and Management, 205, 112460.

23. Cao, B., Adutwum, L. A., Oliynyk, A. O., Lubera, E. J., Olsen, B. C., Mar, A., & Buriak, J. M. (2018). How to optimize materials and devices via design of experiments and machine learning: demonstration using organic photovoltaics. ACS Nano, 12, 7434-7444.

24. Campello, R. J., Moulavi, D., & Zimek, A. (2015). Hierarchical density estimates for data clustering, visualization, and outlier detection. ACM Transactions on Knowledge Discovery from Data, 10(1), Article 5.

25. McInnes, L., & Healy, J. (2017). Accelerated Hierarchical Density Clustering. 2017 IEEE International Conference on Data Mining Workshops. New Orleans, LA: IEEE.

26. McInnes, L., & Healy, J. (2018). UMAP: uniform manifold approximation and projection for dimension reduction. ArXiv, abs/1802.03426.

27. Zhu, W., Sun, W., & Romagnoli, J. (2018). Adaptive k-nearest-neighbor method for process monitoring. Industrial & Engineering Chemistry Research, 57(7), 2574-2586.
28. Yin, C., Gao, Y., Li, T., Xie, G., Li, K., & Tang, H. (2020). Study of internal multi-parameter distributions of proton exchange membrane fuel cell with segmented cell device and coupled three-dimensional model. Renewable Energy, 147, 650-662.

29. Kregar, A., Tavcar, G., Kravos, A., & Katrasnik, T. (2020). Predictive system-level modeling framework for transient operation and cathode platinum degradation of high temperature proton exchange membrane fuel cells. Applied Energy, 263, 114547.

30. Oh, K., Jeong, G., Cho, E., Kim, W., & Ju, H. (2014). A CO poisoning model for high-temperature proton exchange membrane fuel cells comprising phosphoric acid-doped polybenzimidazole membranes. International Journal of Hydrogen Energy, 39(36), 21915-21926.

31. Sousa, T., Mamlouk, M., & Scott, K. (2010). An isothermal model of a laboratory intermediate temperature fuel cell using PBI phosphoric acid membranes. Chemical Engineering Science, 65(8), 2513-2530.

32. Cheddie, D. F., & Munroe, N. D. (2007). A two-phase model of a intermediate temperature PEM fuel cell. International Journal of Hydrogen Energy, 32(7), 832-841.

33. Scott, K., & Mamlouk, M. (2009). A cell voltage equation for an intermediate temperature proton exchange membrane fuel cell. International Journal of Hydrogen Energy, 34(22), 9195-9202.

34. Rao, R. V. (2016). Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems. International Journal of Engineering Computations, 7, 19-34.

35. Abdin, Z., Webb, C. J., & Gray, E. (2016). PEM fuel cell model and simulation in Matlab-Simulink based on physical parameters. Energy, 116, 1131-1144.

36. Corrêa, J. M., Farret, F. A., Canha, L. N., & Simões, M. G. (2004). An Electrochemical-based fuel-cell model suitable for electrical engineering automation approach. IEEE Transactions on Industrial Electronics, 51(5), 1103-1113.

37. Drucker, H., Burges, C., Kaufman, L., Smola, A., & Vapnik, V. (1997). Support Vector Regression Machines. Adv. Neural. Inf. Process. Syst., 9, 155-161.

38. Boser, B., Guyon, I., & Vapnik, V. (1992). A training algorithm of optimal margin classifiers. Proceedings of the Fifth Annual Workshop on Computational Learning Theory. Pittsburg.

39. Chang, C., & Ling, C. (2011). LIBSVM: A library for support vector machines. ACM Trans. Intell. Syst. Technol., 2(3), Article 27.

40. Sata, T. (2004). Ion Exchange Membranes. (Royal Society of Chemistry).

41. Grew, K. N., Chu, D., Chiu, W. K. S. (2010). Ionic Equilibrium and Transport in the Alkaline Anion Exchange Membrane. Journal of the Electrochemical Society, 157, B1024-B1032.

42. Arges, C., Wang, L., Parrondo, J., & Ramani, V. (2013). Best practices for investigating anion exchange membrane suitability for alkaline electrochemical devices: case study using quaternary ammonium poly (2, 6-dimethyl 1, 4-phenylene) oxide anion exchange membranes. J. Electrochem. Soc., 160, F1258.

43. Arges, C., & Ramani, V. (2011). Two-dimensional NMR spectroscopy reveals cation-triggered backbone degradation in polysulfone-based anion exchange membranes. Proc. Natl. Acad. Sci., 110, 2490-2495.

44. Arges, C., Li, K., Zhang, L., Kambe, Y., Wu, G.-P., Lwoya, B., . . . Kumar, R. (2019). Ionic conductivity and counterion condensation in nanoconfined polycation and polyanion brushes prepared from block copolymer templates. Mol. Syst. Des. Eng., 4, 365-378.

45. Su, Z., Kole, S., Harden, L., Palakkal, V., Kim, C., Nair, G., Renner, J. (2019). Peptide-modified electrode surfaces for promoting anion exchange ionomer microphase separation and ionic conductivity. ACS Mater. Lett., 1, 467-475.

46. Xu, S., Wang, Y., & Zhi, W. (2019). Parameter estimation of proton exchange membrane fuel cells using eagle strategy based on JAYA algorithm and Nelder-Mead simplex method. Energy, 173, 457-467.
Current intensity (A/cm²)

Voltage (V)

Activation overpotentials

Concentration overpotentials

Ohmic overpotentials

Model (200 °C, 0% CO)
Model (220 °C, 0% CO)
Model (220 °C, 25% CO)

Overpotential (200 °C 0% CO)
Overpotential (220 °C 0% CO)
Overpotential (220 °C 25% CO)
Machine Learning for Guiding High-Temperature PEM Fuel Cells with Greater Power Density

Luis A. Briceno-Mena, Gokul Venugopalan, José A. Romagnoli*, and Christopher G. Arges*

Cain Department of Chemical Engineering, Louisiana State University, Baton Rouge, LA 70803

*corresponding authors: carges@lsu.edu, jose@lsu.edu

Nomenclature

| Symbol | Units | Description |
|--------|-------|-------------|
| $A$    | $cm^2$ | Cross section area of the MEA |
| $\frac{(CO)}{(H_2)}_{Fuel}$ | Dimensionless | Ration of CO to H$_2$ in fuel stream |
| $D_{H_2}^{CL}$ | $cm^2/s$ | Hydrogen diffusivity in the catalyst layer |
| $D_{O_2}^{CL}$ | $cm^2/s$ | Oxygen diffusivity in the catalyst layer |
| $E$    | $V$    | Cell potential |
| $E_{oc}$ | $V$ | Open circuit overpotential |
| $E_{H_2}^A$ | $J/mol$ | Activation energy for the H$_2$ |
| $E_{O_2}^A$ | $J/mol$ | Activation energy for the O$_2$ |
| $F$    | $C/mol$ | Faraday constant |
| $I$    | $A$    | Current |
| $N_{H_2}$ | $mol/s$ | Hydrogen molar flowrate |
| $N_{H_2O}^{cat}$ | $mol/s$ | Water molar flowrate at the cathode |
| $N_{CO}$ | $mol/s$ | Carbon monoxide molar flowrate |
| $P$    | $atm$  | Pressure |
| $P_{H_2}$ | $atm$ | Hydrogen partial pressure |
| $P_{O_2}$ | $atm$ | Oxygen partial pressure |
| $P_{H_2}^0$ | $atm$ | Hydrogen reference partial pressure |
| $P_{O_2}^0$ | $atm$ | Oxygen reference partial pressure |
| $P_{an}$ | $atm$ | Pressure at the anode |
| $P_{cat}$ | $atm$ | Pressure at the cathode |
| $R$    | $J/mol K$ | Universal gas constant |
| $S_{H_2}$ | Dimensionless | Hydrogen stoichiometry ratio |
| $S_{O_2}$ | Dimensionless | Oxygen stoichiometry ratio |
| $T$    | $K$    | Temperature |
| $T_{an}$ | $K$ | Temperature at the anode |
| $T_{cat}$ | $K$ | Temperature at the cathode |
| $T_{ref}^0$ | $K$ | Ref. temp. for associated to ref. exchange current density |
1. Model description

1.1. Model Approach and Assumptions

- Lumped model (i.e., 0-D model)
- Steady-state behavior
- Ideal gas behavior
- Electrode kinetics can be described by the Butler-Volmer equation
- Ohmic resistances from the membrane can be described by Ohm’s law
- Concentration overpotential is a semi-empirical model and transport through porous media is ignored
- O₂ and H₂ solubility in PA imbibed ionomer binders in the electrodes can be described by Henry’s Law. Henry’s constant for hydrogen in this system is assumed to be 4x greater than oxygen.
- Parasitic reactions (e.g., O₂ reduction to hydrogen peroxide and carbon/electrocatalyst corrosion) are not considered.

1.2. Reactions and Overall Mass Balances

The present model includes the chemical reactions for hydrogen oxidation in the anode (R1) and oxygen reduction in the cathode (R2).

\[ H_2 \rightarrow 2H^+ + 2e^- \quad \text{(R1)} \]

\[ \frac{1}{2}O_2 + 2e^- + 2H^+ \rightarrow H_2O \quad \text{(R2)} \]

From the anode and cathode reactions the following species balances can be written:

\[ \frac{dN_{H_2}}{dt} = (S_{H_2} - 1) \frac{I}{zh_{H_2}F} \quad \text{(E1)} \]

\[ \frac{dN_{O_2}}{dt} = (S_{O_2} - 1) \frac{I}{zo_2F} \quad \text{(E2)} \]

\[ \frac{dN_{H_2O}^{cat}}{dt} = \frac{I}{zh_{H_2O}F} \quad \text{(E3)} \]

1.3. Polarization Behavior and Overpotential Losses

The total cell voltage is calculated as:

\[ E(i) = E_{oc} - \eta_{act} - \eta_{con} - \eta_{ohm} \quad \text{(E4)} \]

The theoretical open circuit voltage is calculated as a function of temperature and the species partial pressures as \(^1\):

\[ E_{oc} = 1.23 - 0.9 \times 10^{-3} (T - 298.15) + \frac{RT}{2F} \ln \left( \frac{P_{H_2}}{P_{O_2}} \right) \quad \text{(E5)} \]
From the Butler-Volmer equation\(^2\), the activation overpotential, including the effect of CO, can be expressed as:

\[
\eta_{\text{act}} = \frac{RT}{\alpha_{\text{an}}F} \ln \left( \frac{i}{i_0} \right)^2 + \frac{RT}{\alpha_{\text{cat}}F} \ln \left( \frac{i}{i_0} \right)
\]  \hspace{1cm} (E6)

Where \(\theta_{\text{CO}}\) is the CO coverage calculated as\(^3\):

\[
\theta_{\text{CO}} = 19.9\exp \left( -7.69 \times 10^{-3} T_{\text{an}} \right) + 0.085\ln \left( \frac{y_{\text{CO}}}{y_{\text{H}_2}} \right)
\]  \hspace{1cm} (E7)

For the exchange currents the following expressions are used, introducing a correction for the phosphoric acid (PA) doping\(^4\)\(^5\):

\[
i_{\text{an}}^0 = a_c L_C \left( \frac{p_{\text{H}_2} H_{\text{CL}}^0}{c_{\text{H}_2}^0} \right)^{0.5} \exp \left( -\frac{E_{\text{H}_2}^0}{R} \left( \frac{1}{T_{\text{an}}} - \frac{1}{T_{\text{ref}}} \right) \right) i_0 \text{ an} \exp \left( -\gamma (1 - m_{\text{IO}}) \right)
\]  \hspace{1cm} (E8)

\[
i_{\text{cat}}^0 = a_c L_C \left( \frac{p_{\text{O}_2} H_{\text{CL}}^0}{c_{\text{O}_2}^0} \right) \exp \left( -\frac{E_{\text{O}_2}^0}{R} \left( \frac{1}{T_{\text{cat}}} - \frac{1}{T_{\text{ref}}} \right) \right) i_0 \text{ cat} \exp \left( -\gamma (1 - m_{\text{IO}}) \right)
\]  \hspace{1cm} (E9)

\(m_{\text{IO}}\), the acid mass fraction is calculated from the IEC as:

\[
m_{\text{IO}} = 0.0902 IEC_{\text{IO}} + 0.0352
\]  \hspace{1cm} (E10)

In the above equations, \(\gamma\) is an empirical parameter that helps account how the PA content alters the exchange current density values.

Cheddie et al. proposed an empirical expression for the calculation of Henry’s constant for the oxygen in PA\(^5\). In this model we approximate the Henry’s constant for oxygen as a function of temperature and PA mass fraction in the ionomer binder:

\[
H_{\text{CL}}^{\text{O}_2} = (1.0 \times 10^{-6}) \exp \left( -\frac{900m_{\text{IO}}^2 - 1500m_{\text{IO}} + 650}{T} \right)
\]  \hspace{1cm} (E11)

Hydrogen mass transport is not considered as a limiting factor. Therefore, for the hydrogen, the Henry’s constant is simply assumed to be \(H_{\text{CL}}^{\text{H}_2} = 4H_{\text{CL}}^{\text{O}_2}\)\(^5\)\(^6\).

The concentration overpotential is determined from a semi-empirical equation as proposed by Corrêa et al.\(^7\):

\[
\eta_{\text{con}} = -B \left( \ln \left( 1 - \frac{i}{i_{\text{an}}^0} \right) + \ln \left( 1 - \frac{i}{i_{\text{cat}}^0} \right) \right)
\]  \hspace{1cm} (E12)
The limiting current, \( i_{\text{lim}}^{\text{cat}} \), is affected by the diffusivity in the CL and the partial pressure and the CL thickness \(^{[8]}\). In this work, an empirical parameter, \( K_{\text{cat}} \), is used as a constant for predicting the liming current in the cathode as a function of partial pressure, diffusion, temperature, and ionomer binder thickness:

\[
\frac{i_{\text{lim}}^{\text{cat}}}{T\delta_{\text{CL}}} = \frac{K_{\text{cat}}D_{CL}^{O_2}P_{O_2}}{T}\delta_{\text{CL}}
\]

(E13)

Considering the effect of carbon monoxide coverage for the anode, a similar expression is used but modified to include carbon monoxide coverage

\[
\frac{i_{\text{lim}}^{\text{cat}}}{T\delta_{\text{CL}}} = \frac{K_{\text{an}}D_{CL}^{CL}P_{H_2}}{T\delta_{\text{CL}}}(1 - \theta_{CO})^2
\]

(E14)

It is assumed that

\[
K_{\text{an}} = 0.5K_{\text{cat}}
\]

(E15)

The oxygen diffusivity in PA is approximated from the PA mass fraction in the ionomer binder as \(^{[5]}\):

\[
D_{CL}^{O_2} = (1.0 \times 10^{-6})\exp \left( -\frac{4500 m^2}{T} - 10000 m_{o} + 4010 \right)
\]

(E16)

For the hydrogen diffusivity, it is assumed to be \( D_{CL}^{H_2} = 2D_{CL}^{O_2} \) \(^{[5][6]}\).

The ohmic overpotential accounts for the ionic resistance of both the membrane and the ionomer binder, while all other sources of ohmic losses (e.g., electrical conduction through contacts) are considered negligible \(^{[9]}\). The resistances are considered to be in series so that the global resistance is additive combination of the individual membrane and ionomer resistances:

\[
\eta_{\text{ohm}} = \frac{i}{A\left(\kappa_{\text{mem}} + \kappa_{io}\right)}
\]

(E17)

Where \( \kappa_{\text{mem}} \) and \( \kappa_{io} \) are the ionic conductivities of the membrane and the ionomer binder, which are predicted from a machine learning model informed from experimental data as described below.

2. Support vector regression
Consider the set of training points, \(\{(x_1, z_1), ..., (x_l, z_l)\}\), where \(x_i \in \mathbb{R}^n\) is the feature vector (e.g., the materials characteristics, environmental conditions) and \(z_i \in \mathbb{R}^1\) is the target output (e.g., bulk properties). Under the given parameters \(C > 0\) and \(\epsilon > 0\), the standard form of support vector regression is:

\[
\min_{w, b, \xi, \xi^*} \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_i + C \sum_{i=1}^{l} \xi_i^* 
\]

Subject to

\[
w^T \phi(x_i) + b - z_i \leq \epsilon + \xi_i,
\]

\[
z_i - w^T \phi(x_i) - b \leq \epsilon + \xi_i^*,
\]

\[\xi_i, \xi_i^* \geq 0, i = 1, ..., l.\]

Here \(\phi(x_i) = \exp(-\gamma \|x - x_i\|^2)\) with \(\gamma = \frac{1}{n \text{var}(x)}\) is the kernel function, and \(w\) is the vector of weights.

Usually, due to the high dimensionality of the \(w\), the following dual problem is solved

\[
\min_{\alpha, \alpha^*} \frac{1}{2}(\alpha - \alpha^*)^T Q (\alpha - \alpha^*) + \epsilon \sum_{i=1}^{l} (\alpha_i - \alpha_i^*) + \sum_{i=1}^{l} \sigma_i (\alpha_i - \alpha_i^*) 
\]

Subject to

\[e^T (\alpha - \alpha^*) = 0,\]

\[0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, ..., l,\]

Where

\[Q_{ij} = K(x_i, x_j) \equiv \phi(x_i)^T \phi(x_j)\]

After solving the dual problem, the approximate function is

\[z = \sum_{i=1}^{l} (-\alpha + \alpha_i^*) K(x_i, x) + b\]
3. Global sensitivity analysis

Modified Sobol’s Global Sensitivity Analysis

The Sobol’s method for global sensitivity analysis (GSA) was first introduced in [10]. This method proposes the series expansion of a function \( G = g(z_1, ..., z_J) \) into orthogonal terms of increasing dimensions with mutually independent input variables:

\[
G = g_0 + \sum_{j=1}^{J} g_j(z_j) + \sum_{1 \leq j < b \leq J} g_{jb}(z_j, z_b) + \ldots + g_{1,2,\ldots,J}(z_1, ..., z_J)
\]  
(E24)

Where the index \( j \) denotes the variable of interest, \( b \) another variable, and \( J \) is the total number of evaluated variables. Each term in \( G \) has quadratic integrability over the domain of existence, where \( g_0 \) is a constant. The decomposition of the variance of \( G \) is given by

\[
V(G) = \sum_{j=1}^{J} v_j + \sum_{1 \leq j < b \leq J} v_{jb} + \ldots + v_{1,2,\ldots,J}
\]  
(E25)

Where \( v_j, v_{jb}, v_{1,2,\ldots,J} \) are the individual variances of \( g_j(z_j) \), \( g_{jb}(z_j, z_b) \), and \( g_{1,2,\ldots,J}(z_1, ..., z_J) \) respectively.

The calculated variances can be used to compute the sensitivity indices as

\[
\hat{S}_j = \frac{v_j}{\hat{V}}
\]  
(E26)

\[
\hat{S}_{Tj} = 1 - \frac{\hat{V}_j}{\hat{V}}
\]  
(E27)

Where \( \hat{V}_j \) is the sum of all variances terms that exclude \( z_j \)

This sensitivity indices provide a way to classify variables according to their individual contribution to the changes in the overall variance of the function. Furthermore, the sensitivity indices can indicate whether the model is additive (interactions are present) or not. If \( \hat{S}_j < \hat{S}_{Tj} \), then the model is nonadditive and if \( \hat{S}_j = \hat{S}_{Tj} \) the model is additive [11].

The standard Sobol’s method has been subject of several modifications to improve its performance [12][13]. In this work, we adopted the method proposed by Salas et al. in which a third sampling matrix is used as a contingency for the case of unfeasible solutions [14].

A description of the GSA algorithm is shown on the next page. For this implementation two objective functions were used, one for a parameter sensitivity analysis, \( G_1 \), and another for variable sensitivity analysis, \( G_2 \):
Input: Given \( J, D \in \mathbb{R}; z \in \mathbb{R}^J; u; G = g(z) \)

Output: \( \hat{S}_j, \hat{S}_{Tj} \), \( j = 1, ..., J \)

Initialization
1: Let generate \( D \times J \) matrices \( M_1, M_2, \) and \( M_3 \) with entries \( z_d, j \) varying randomly \( z_j \pm u \% \)

Sampling and resampling
2: for \( m \leftarrow [1, 2] \) do
3: for \( d \leftarrow [1, ..., D] \) do
4: \( g_m[d] \leftarrow g(M_m[d, :] \) \( \{M_m[d, :] \) is the \( d \)-th row of matrix \( M_m \) and \( g_m[d] \) is the \( d \)-th entry of the vector \( g_m \})
5: if \( (g_m[d] \) is not a feasible value) then
6: while \( (g_m[d] \) not a feasible value) do
7: Use the first not used row \( k \) of \( M_3 \)
8: \( g_m[d] \leftarrow g(M_3[k, :]) \)
9: end while
10: \( M_m[d, :] \leftarrow M_3[k, :] \)
11: end if
12: end for

Output varying the input variables
13: \( \hat{g}_0 \leftarrow \frac{1}{2D} \sum_{d=1}^{D} (g_{1,d} + g_{2,d}) \)
14: for \( j \leftarrow [1, ..., J] \) do
15: \( N_j \leftarrow M_2[:, j] \) \( \{N_j[:, j] \) and \( M_1[:, j] \) are the \( j \)-th columns of the matrices \( N_j \) and \( M_1 \) respectively\)
16: end for

\( G_1 = \sqrt{\frac{\sum_{k=1}^{K} (E - E_{exp})^2}{k(k-1)}} \) \hspace{1cm} (E28)

\( G_2 = \max \left( \frac{E_f}{A} \right) \) \hspace{1cm} (E29)
Figure S1. First-order and total sensitivity indices for the parameter sensitivity analysis.

Figure S2. First-order and total sensitivity indices for the variable sensitivity analysis.

4. **Jaya for parameter estimation**

The objective function for the Jaya optimization the minimization of the standard error as:
\[
\min \left( F = \sqrt{\frac{\sum_{k=1}^{k\ell} (\bar{E}_k - \bar{E}_k)^2}{k(k-1)}} \right)
\]  
(E30)

Using the same sampling method from the sensitivity analysis, a solution population is obtained. The best and worst solutions are selected, and the parameters of the solutions in the current population are updated such that:

\[
M_1'[d,j] = M_1[d,j] + a_1 (g_{\text{best}}[j] - |M_1'[d,j]|) - a_2 (g_{\text{worst}}[j] - |M_1'[d,j]|)
\]

(E31)

Where \( M_1'[d,j] \) is the updated \( j \)-th parameter of the \( d \)-th solution, \( g_{\text{best}}[j] \) and \( g_{\text{worst}}[j] \) are the \( j \)-th parameter of the best and worst solution respectively, and \( a_1 \) and \( a_2 \) are two random uniformly distributed numbers in \((0,r)\). The value \( r \) can be considered as a hyperparameter. The objective function is evaluated again for the updated solutions, then is compared against the current best solution and if it is better the solution is kept, otherwise is dropped. The algorithm finishes when the new best solution satisfies the stopping criterion or when the maximum number of iterations is reached. The pseudocode for the algorithm is provided on the next page.

Figure S3. Convergence for the Jaya algorithm over 10 iterations for 100 runs for 3, 5, and 7 parameters selected as ranked from the parameter global sensitivity analysis.
\textbf{Input:} Given $M_1[d, :], d = 1, \ldots, D; k, tol; it_{\text{max}}; r_1, r_2 \in (0, 1); F = f(z), \hat{S}_j, j = 1, \ldots, J$

\textbf{Output:} $z[j], j = 1, \ldots, J$

\textbf{Initialization}

1: $g_{\text{best}} \leftarrow \text{init} > tol$
2: while $g_{\text{best}} > tol \text{ or } it < it_{\text{max}}$ do
3: \hspace{1em} for $d \leftarrow [1, \ldots, D]$ do
4: \hspace{2em} $g[d] = f(M_1[d, :])$
5: \hspace{1em} end for
6: $g_{\text{best}} = \min (g)$
7: $g_{\text{worst}} = \max (g)$
8: Using the $k$ parameters with highest $\hat{S}_j$
9: \hspace{1em} for $j \leftarrow [1, \ldots, k]$ do
10: \hspace{2em} for $d \leftarrow [1, \ldots, D]$ do
11: \hspace{3em} Update solutions using (E31)
12: \hspace{2em} end for
13: \hspace{1em} end for
14: $it \leftarrow it + 1$
15: end while
16: return $M_1'[d, j]$
5. Clustering analysis

SGS comparison

Table S1. Subspaces identified as responsible for cluster separation between cluster \(i\) (columns) and cluster \(j\) (rows).

| Cluster | From | 1 | 2 | 3 | 4 | 5 |
|---------|------|---|---|---|---|---|
| 1       | \(T, \text{Power}\) | \(CO/H_2, IEC_{io}\) | \(S_{O_2}, L_c, IEC, IEC_{io}\) | \(IEC_{io}, S_{O_2}, CO/H_2, \text{Power}\) |
| 2       | \(T, S_{O_2}, IE, C_{\text{mem}}, \text{Power}\) | \(S_{O_2}, T, IEC_{io}, CO/H_2\) | \(T, L_c, S_{O_2}, IE, C_{io}, \text{Power}\) | \(T, \text{Power}\) |
| 3       | \(CO/H_2, IE, C_{\text{mem}}, \delta_{\text{mem}}, \text{Power}\) | \(T, \text{Power}\) | \(S_{O_2}, IEC_{\text{mem}}, L_c\) | \(IEC_{io}, S_{O_2}, \text{Power}\) |
| 4       | \(S_{O_2}, IE, C_{\text{mem}}, CO/H_2\) | \(T, \text{Power}\) | \(S_{O_2}, CO/H_2, S_{H_2}\) | \(IEC_{io}, P\) |
| 5       | \(S_{O_2}, IE, C_{\text{mem}}, CO/H_2\) | \(T, \text{Power}\) | \(IEC_{io}, S_{O_2}, CO/H_2, \text{Power}\) | \(IEC_{io}, L_c, S_{O_2}, IEC, \text{Power}\) |

Frequency of variables showing as the most important

Table S2. Overall cluster comparison

| Variable | Frequency |
|----------|-----------|
| \(S_{O_2}\) | 12        |
| \(IEC_{io}\) | 9         |
| \(T\) | 8         |
| \(CO/H_2\) | 7         |
| \(IEC_{\text{mem}}\) | 6         |
| \(L_c\) | 2         |
| \(S_{H_2}\) | 1         |
| \(P\) | 1         |
| \(\delta_{\text{mem}}\) | 1         |
Table S3. Cluster comparison for cluster 5 against the rest

| Variable   | Frequency |
|------------|-----------|
| $\text{IEC}^{\text{lo}}$ | 3         |
| $S_{O_2}$  | 2         |
| $T$        | 1         |
| $CO/H_2$   | 1         |
| $P$        | 1         |
| $\text{IEC}_{\text{mem}}$ | 0         |
| $L_c$      | 0         |
| $S_{H_2}$  | 0         |
| $\delta_{\text{mem}}$ | 0         |

6. Materials synthesis and characterization

*Synthesis of polycation ionomers (i.e., quaternary benzyl pyridinium poly(arylene ether sulfone) (QPPSf))*

A sample reaction is as follows: 20 g of Udel® poly(arylene ether sulfone) (PSF) was dissolved in 1000 mL of chloroform in a round bottom flask. After PSF was completely dissolved, 13.6 of paraformaldehyde and 30 mL of chlorotrimethylsilane was added. The flask was sealed with rubber septum and stocked with nitrogen. 1050 $\mu$L of SnCl$_4$, the Lewis acid catalyst was added to the sealed flask. The extent of chloromethylation was monitored during the reaction until desired degree of functionalization (DF) values of chloromethylated groups was achieved. The DF values were continuously monitored by $^1$H NMR. The reaction solutions, collected either at the end of the reaction or during the reaction, were cooled down to room temperature and poured into methanol (5:1 volume ratio of methanol to reaction solution) to precipitate the polymer. The solid collected was re-dissolved in chloroform and then precipitated in methanol (5:1 volume ratio) to remove impurities. The solid was collected via vacuum filtration and then vacuum dried at room temperature. CMPSf of different DF values (0.7 DF, 0.9 DF, 1.1 DF, 1.26 DF, and 1.5 DF) were synthesized for thin film conductivity, thin film thickness, IEC, and electrode binder acid uptake experiments. Polycation ionomers (i.e., QPPSf) from CMPSf were synthesized by dissolving CMPSf in NMP (5wt% solution). For converting chloromethylated groups to quaternary pyridinium chloride groups, 3 to 1 mole ratio of base reagent (pyridine) to chloromethylated sites was added. The reaction was run at 60 °C for 24 h. QPPSf solutions of different IEC was synthesized using different DFs of CMPSf.
**IEC of polycation ionomers and QPPSf-PBI membranes**

A 5 wt% of QPPSf in NMP of different IECs synthesized were drop casted on to a levelled glass plates and dried at 60 °C for 24 h. The membranes were then peeled off from the glass plates. IEC of QPPSf was determined using $^1$H NMR. First, the degree of functionalization (DF) of CMPSf was calculated using $^1$H NMR.

$$\text{DF} = \frac{2 \times \text{Area}_{\text{CH}_2\text{Cl substituent} (4.5 \text{ ppm})}}{\text{Area}_{\text{PSf substituent} (7.8 \text{ ppm})}}$$  \hspace{1cm} (E32)

For calculating the ion-exchange capacity of QPPSf ionomers, QPPSf membranes were dissolved in D-DMSO (15 mg in 1 mL of d-DMSO). This solution was used to perform the $^1$H NMR of QPPSf.

Conversion of chloromethylated sites to cation sites:

$$\text{Conversion} = \frac{\text{Area}_{\text{cation substituent} (9.2 \text{ ppm})}}{\text{Ratio} \times \text{DF} \times \text{Area}_{\text{PSf substituent} (1.8 \text{ ppm})}}$$ \hspace{1cm} (E33)

$$\text{Ratio} = \frac{\text{# of protons for cation substituent}}{\text{# of protons for PSf substituent}}$$ \hspace{1cm} (E34)

Determining theoretical IEC of polycations

$$\text{IEC} \left( \frac{\text{mmol}}{g} \right) = \frac{\text{DF} \times 1000}{(\text{MW}_{\text{PSf, monomer}} + \text{DF} \times \text{MW}_{\text{cation}})} \times \text{Conversion}$$  \hspace{1cm} (E35)

Where

$$\text{MW}_{\text{cation}} = (\text{MW}_{\text{cation free base conjugate}} + \text{MW}_{\text{counteranion}} + \text{MW}_{\text{CH}_2 - 1})$$ \hspace{1cm} (E36)
Table S4. Thin film QPPSf in-plane ionic conductivity from 25 °C to 200 °C with different IEC values measured on IDEs

| Temperature (°C) | QPPSf IEC (mmol equiv⁻¹) |
|-----------------|--------------------------|
|                 | 0.9  | 1.24 | 1.45 | 1.72 | 1.96 |
| 25              | 39.4 ± 2.1  | 44.5 ± 2.3  | 45.5 ± 1.7  | 47.8 ± 0.3  | 48.4 ± 1.6  |
| 40              | 40.8 ± 1.6  | 48.1 ± 5.3  | 47.4 ± 4.7  | 48.5 ± 0.3  | 50.7 ± 1.1  |
| 60              | 42.5 ± 0.8  | 47 ± 1.8    | 47.9 ± 1.2  | 49.4 ± 0.3  | 53.3 ± 0.3  |
| 80              | 43.5 ± 0.4  | 48.2 ± 2    | 48.1 ± 1.4  | 50.3 ± 0.2  | 54.6 ± 0.4  |
| 100             | 45.9 ± 0.9  | 49.5 ± 1.4  | 49.3 ± 0.8  | 53.9 ± 2.2  | 57.9 ± 0.4  |
| 120             | 47.9 ± 0.6  | 56.7 ± 2.1  | 54.6 ± 1.5  | 56.8 ± 1.2  | 58.7 ± 0.1  |
| 140             | 49.3 ± 1.7  | 59.6 ± 2.8  | 57.6 ± 2.2  | 66.8 ± 2.1  | 64.8 ± 1.2  |
| 160             | 51.2 ± 1.6  | 62.8 ± 1.9  | 64.7 ± 1.3  | 73.5 ± 1    | 74.3 ± 1.1  |
| 180             | 54.6 ± 1.5  | 66 ± 1.7    | 69.3 ± 1.1  | 77.1 ± 0.9  | 79.8 ± 1    |
| 200             | 56.2 ± 0.9  | 71.3 ± 0.9  | 71.9 ± 0.9  | 80.5 ± 0.6  | 72.2 ± 0.4  |
Table S5. Physical properties of QPPSf ionomer binder materials.

| CMPSf (DF) | QPPSf IEC (mmol equiv⁻¹) | Acid uptake (%) of the electrodes with ionomer binders | Acid mass fraction |
|------------|--------------------------|--------------------------------------------------------|-------------------|
| 0.7        | 0.97                     | 17.1 ± 0.25                                            | 0.12 ± 0.002      |
| 0.9        | 1.24                     | 23.2 ± 3.3                                             | 0.14 ± 0.008      |
| 1.1        | 1.45                     | 29.4 ± 1.5                                             | 0.17 ± 0.003      |
| 1.26       | 1.72                     | 26.6 ± 2.2                                             | 0.18 ± 0.002      |
| 1.5        | 1.96                     | 33.5 ± 1.4                                             | 0.20 ± 0.006      |

7. Model validation

Table S6. MEA properties and gas reactant flow rate feeds for polarization data in Figure 3 of the main manuscript [15].

| Variable | Description                                      | Value | Units         |
|----------|--------------------------------------------------|-------|---------------|
| $A$      | Geometric area of the MEA                        | 5     | $cm^2$        |
| $a_C$    | Catalyst specific area for each electrode       | 45    | $m^2/g_{Pt}$  |
| $L_C$    | Catalyst loading for each electrode             | 0.5   | $mg_{Pt}/cm^2$|
| $S_{H_2}$| Hydrogen stoichiometry ratio                     | 1.2   | Dimensionless |
| $S_{O_2}$| Oxygen stoichiometry ratio                       | 2.2   | Dimensionless |
| $\delta_{mem}$ | Membrane thickness                              | 50    | $\mu m$      |
| $\delta_{io}$ | Ionomer binder thickness                         | 10    | $nm$          |
| $F_{in}^{O_2}$ | Oxygen flow rate                                 | 0.2   | SLPm          |
| $F_{in}^{H_2/CO}$ | Hydrogen or hydrogen-carbon monoxide flow rate  | 0.2   | SLPm          |
| Parameter   | Units               | Description                                      | Value         | Reported value |
|-------------|---------------------|--------------------------------------------------|---------------|----------------|
| $i_{0 \text{ an}}$ | $A/cm^2$          | Reference exchange current density for the anode | 0.144         | 0.144          |
| $i_{0 \text{ ref}}$ | $A/cm^2$          | Reference exchange current density for the cathode | $2.63 \times 10^{-8}$ | $2.63 \times 10^{-8}$ |
| $K_{\text{cat}}$ | $KAs/cm^3 atm$ | Empiric parameter for the limiting current at the cathode side | $1.99 \times 10^7$ | N.A. |
| $\alpha_{\text{an}}$ | Dimensionless     | Charge transfer coefficient for the hydrogen at the anode | 1.02          | $0 - 1$ |
| $\alpha_{\text{cat}}$ | Dimensionless     | Charge transfer coefficient for the cathode      | 0.97          | $0 - 2$ |
| $\gamma$   | Dimensionless      | Empiric parameter for $H_3P$ $O_4$ effect on mass transfer | 4.95          | 4.16          |
| $B$        | $V$                | Empiric parameter for concentration overpotential | 0.090         | 0.016          |

References

[1] Rasheed, R. K., Liao, Q., Caizhi, Z., & Chan, S. H. (2017). A review on modelling of high temperature proton exchange membrane fuel cells (HT-PEMFCs). International Journal of Hydrogen Energy, 41(5), 3142-3165.

[2] Mann, R., Amphlett, J., Peppley, B., & Thurgood, C. (2006). Application of Butler-Volmer equations in the modelling of activation polarization for PEM fuel cells. Journal of Power Sources, 161(2), 775-781.

[3] Dhar, H. P., Christner, L. G., & Kush, A. K. (1987). Nature of CO adsorption during H2 oxidation on relation to modeling CO poisoning of a fuel cell anode. Journal of the Electrochemical Society, 134(12), 3021-3026.

[4] Barbir, F. (2013). Chapter Three - Fuel Cell Electrochemistry. In F. Barbir, PEM Fuel Cells (2nd ed., pp. 33-72). San Diego, CA: Academic Press.

[5] Cheddie, D. F., & Munroe, N. D. (2007). A two-phase model of a in intermediate temperature PEM fuel cell. International Journal of Hydrogen Energy, 32(7), 832-841.
[6] Bernardi, D. M., & Vebrugge, M. W. (1992). A mathematical model of the solid-polymer-electrolyte fuel cell. J. Electrochem. Soc., 139(9), 2477-2491.

[7] Corrêa, J. M., Farret, F. A., Canha, L. N., & Simões, M. G. (2004). An Electrochemical-based fuel-cell model suitable for electrical engineering automation approach. IEEE Transactions on Industrial Electronics, 51(5), 1103-1113.

[8] He, W., Wang, B., & Dickerson, J. H. (2012). Overall concentration polarization and limiting current density of fuel cells with nanostructured electrodes. Nano Energy, 1, 828-832.

[9] Abdin, Z., Webb, C. J., & Gray, E. (2016). PEM fuel cell model and simulation in Matlab-Simulink based on physical parameters. Energy, 116, 1131-1144.

[10] Sobol', I. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo Estimates. Mathematics and Computers in Simulation, 55, 271-280.

[11] Consenza, A., Mannina, G., Vanrolleghem, P. A., & Neumann, M. B. (2014). Variance-based sensitivity analysis for wastewater treatment plant modelling. Science of The Total Environment, 470-471(1), 1068-1077.

[12] Saltelli, A. (2002). Making the best use of model evaluations to compute sensitivity indices. Computer Physics Communications, 145(2), 280-297.

[13] Wu, Q.-L., Courmede, P.-H., & Mathieu, A. (2012). An efficient computational method for global sensitivity analysis and its application to tree growth modelling. Reliability Engineering and System Safety, 107, 35-43.

[14] Salas, S. D., Brandao, A. L., Soares, J. B., & Romagnoli, J. A. (2019). Data-driven estimation of significant kinetic parameters applied to the synthesis of polyolefins. Processes, 7(5), 309.

[15] Venugopalan, G., Chang, K., Nijoka, J., Livingston, S., & Geise, G. M. (2020). Stable and highly conductive polycation-polybenzimidazole membrane blends for intermediate temperature polymer membrane fuel cells. ACS Appl. Energy Mater., 573-585, 3.
[16] Sousa, T., Mamlouk, M., & Scott, K. (2010). An isothermal model of a laboratory intermediate temperature fuel cell using PBI phosphoric acid membranes. Chemical Engineering Science, 65(8), 2513-2530.