PAPER • OPEN ACCESS

Enhancing oxygen reduction reaction mass activity in polymer electrolyte fuel cells via molecular machine learning

MACHINE LEARNING
Science and Technology

PAPER

Enhancing oxygen reduction reaction mass activity in polymer electrolyte fuel cells via molecular machine learning

Ekaterina Gracheva[1], **Shin-ichi Yamazaki**[2,*], **Guillaume Lambard**[1,*], **Keitaro Sodeyama**[1], **Tsutomu Ioroi**[2] and **Masafumi Asahi**[2]

[1] Data-driven Materials Design group, Center for Basic Research on Materials (CBRM), National Institute for Materials Science (NIMS), 1-1 Namiki, Tsukuba 305-0044, Ibaraki, Japan
[2] Research Institute of Electrochemical Energy, Department of Energy and Environment, National Institute of Advanced Industrial Science and Technology (AIST), 1-8-31 Midorigaoka, Ikeda 562-8577, Osaka, Japan
[*] Authors to whom any correspondence should be addressed.

E-mail: s-yamazaki@aist.go.jp, lambard.guillaume@nims.go.jp, gracheva.ekaterina@nims.go.jp, sodeyama.keitaro@nims.go.jp, ioroi-t@aist.go.jp and m.asahi@aist.go.jp

OPEN ACCESS

RECEIVED
9 February 2026

REVISED
16 June 2026

ACCEPTED FOR PUBLICATION
18 June 2026

PUBLISHED
3 July 2026

Keywords: materials science, machine learning, chemoinformatics, polymer electrolyte fuel cell (PEFC), oxygen reduction reaction, active learning

abstract

The search for efficient fuel cell additives is a critical challenge for improving their oxygen reduction reaction activity. Here, we present the first successful application of active learning (AL) to guide the experimental discovery of organic molecules that enhance the mass activity (MA) of Pt cathode catalysts in polymer electrolyte fuel cells (PEFCs), a class of fuel cells often referred to in the literature as proton-exchange membrane fuel cells. Using our in-house simplified molecular input line entry system-X molecular characterization tool, we trained neural networks, autonomously tailored for our data, on an initial set of 96 organic compounds, of which 84 passed RDKit validity checks for modeling. Model performance improved over three AL cycles, with the mean absolute error decreasing from 0.202 to 0.165 and the root mean square error decreasing from 0.271 to 0.218. The cycles were guided by predictive ranking, experimental feasibility constraints, and LSTM-based molecule generation. Among the additives identified, melam, a melamine dimer, achieved a 67% improvement in MA over the no-additive baseline. These results demonstrate the feasibility of AL in low-data regimes for PEFC additive discovery and highlight the importance of data consistency, model retraining, and close machine learning–experiment coordination.

1. Introduction

As the global community accelerates efforts to achieve carbon neutrality, hydrogen-based energy systems are gaining significant attention due to their potential to provide clean and efficient power. Polymer electrolyte fuel cells (PEFCs) are a leading clean energy technology due to their high efficiency, low operating temperature, and potential for zero emissions [1–4]. PEFCs are being widely studied for applications ranging from automotive transportation to portable power systems and stationary energy sources. Despite extensive research, several challenges remain before PEFCs can be deployed at a large scale, including issues related to durability, cost, and catalyst performance [5–12].

One key performance metric in PEFCs is the oxygen reduction reaction (ORR) activity at the cathode. Despite decades of study, the ORR on Pt under operating conditions remains the rate-limiting step relative to the hydrogen oxidation at the anode, constraining overall PEFCs efficiency. Modifying the local chemical environment at Pt, via adsorbed organic additives, is a promising route to enhance ORR kinetics.

Prior works have explored various classes of organic molecules, such as ionic liquid [13–18], oleylamine [19], pyrene amine with octylamine [20], tetraazaporphyrin [21, 22], melamine (polymer)

[23–26], and guanamine [27] as activity-enhancing additives for electrocatalysis in PEFCs. These studies have demonstrated that modification of Pt surface with appropriate molecules significantly enhances the ORR activity, and that changes in molecular structure can significantly impact interfacial chemistry and catalytic performance. However, the selection process in these works has typically relied on expert judgment or limited chemical libraries, without systematic data-driven exploration of the broader chemical space. Although the underlying mechanisms are not fully understood, these additives are thought to modulate the interfacial chemistry, alter adsorption behavior, or destabilize reactive intermediates, ultimately improving catalytic activity [28].

Selecting the right additives is, however, a complex problem. The chemical space of potential organic molecules is vast, and traditional trial-and-error experimental screening is both time-consuming and costly. In many cases, materials scientists rely on chemical intuition or structural similarity to known additives to select candidates. This approach, while practical, can easily overlook unconventional or structurally diverse compounds that might exhibit strong performance. As a result, the exploration of additive space remains incomplete, and more efficient, systematic methods are needed to accelerate discovery and reduce experimental overhead.

To the best of our knowledge, no prior study has applied a closed-loop active-learning framework to systematically discover *organic additives* that enhance ORR mass activity (MA) in PEFCs. Such an approach enables efficient navigation of the additive space and offers a foundation for systematic discovery of high-performing, structurally novel candidates.

Molecular machine learning (MML) has emerged as a powerful tool to assist in materials discovery and design, particularly in domains where large datasets are unavailable and each experiment is expensive [29]. In recent years, MML has been successfully applied to many tasks, including catalyst optimization [30, 31], polymer design [32–35], and molecular generation for pharmaceuticals [36, 37]. However, its adoption in PEFC additive screening has been limited [38, 39]. This is partly due to the complexity of the problem, molecular structure-property relationships are often non-linear and require meaningful molecular representations, and partly due to the lack of high-quality experimental data.

To address the challenge of identifying effective organic additives that can enhance the ORR activity of Pt catalysts, we employ an active learning (AL) framework. In AL, a machine learning (ML) model is used not only to learn from existing data but also to guide the selection of new experiments [40–42]. It is an iterative process that aims to improve the model's accuracy while minimising the number of required experiments. At each step, the model selects the most informative or promising candidates to test next, based on an acquisition strategy such as uncertainty or expected improvement. This approach has shown promise in chemistry and materials science, particularly in small-data scenarios where efficient use of experimental resources is critical [43–46].

In the present study, we combine prediction-based candidate ranking, molecular generation, and experimental feasibility filtering within a closed-loop active-learning workflow. This strategy allows the search to focus on candidate additives that are not only predicted to improve ORR MA, but are also realistic for experimental validation.

With this work, we aimed to demonstrate the applicability of AL to PEFC additive discovery. Our primary objective was to improve MA by optimizing the choice of organic additives adsorbed on the Pt surface. We employed the simplified molecular input line entry system (SMILES-X) framework, using LSTM-based molecular representations to support both property prediction and candidate generation.

Through experimental feedback cycles, the integration of AL enabled us to improve model accuracy and identify promising new molecular candidates. Among the top-performing candidates generated by the model and selected for validation, melam, a dimer of melamine, demonstrated a 67% improvement in MA compared to the baseline, pure solvent. This result illustrates the potential of the approach for accelerating materials discovery. More broadly, our results suggest that AL can guide molecular discovery even in low-data regimes and provide a reproducible framework for accelerating fuel cell development. Moreover, the methods presented here are generalizable and can be applied to a wide range of materials science problems involving molecular structure–property relationships.

Our key contributions are as follows:

- We demonstrate a closed-loop active-learning workflow for improving Pt-catalyst MA using organic additives.
- We develop a predictive ML model, trained and refined over multiple experimental feedback cycles.
- We integrate a generative component to propose new candidate molecules, leading to several high-performing suggestions.

- We highlight the importance of data consistency, experimental feedback, and uncertainty-aware exploration in small-data settings.

The remainder of this paper is organized as follows:

Section 2 outlines the ML and experimental procedures. Section 2.1 provides a detailed description of the molecular model, dataset, and learning strategy. Section 2.2 describes the processing of organic compounds and the MA measurements, respectively. In section 3, we present the results across three AL cycles. Section 4 discusses these findings, and section 5 concludes the paper with final remarks and potential future directions.

2. Methods

2.1. MML

To accelerate the discovery of effective organic additives for PEFCs, we employed an AL framework driven by predictive modeling and molecular generation. For both tasks, we used the SMILES-X tool [47], a MML model that learns structure–property relationships directly from SMILES strings, the simplified molecular input line entry system [48], a widely used text-based representation of molecular structures. Although SMILES representations are not unique and can lead to syntactically invalid molecules during generation, these issues were addressed in this work through data augmentation, canonicalization, and post-generation validity checks.

Graph neural networks (GNNs) provide an attractive alternative because they operate directly on molecular graphs and avoid some limitations of text-based molecular representations [49, 50]. However, standard message-passing GNNs propagate information locally along graph edges and may therefore require multiple layers to capture dependencies between distant atoms. This can be problematic in low-data settings, where shallow architectures are often preferred, whereas deeper GNNs may suffer from over-squashing or over-smoothing effects that limit long-range information propagation [51, 52]. In molecular systems, this limitation is particularly relevant for long-range or directional interactions, such as electrostatic effects, which are not explicitly represented in standard message passing [53]. In contrast, the LSTM architecture used in SMILES-X was originally developed to model long-distance dependencies in sequential data and can therefore exploit non-local patterns within SMILES strings [54]. This makes SMILES-X a practical and defensible choice for the present low-data experimental setting. In addition, one of the secondary aims of this study was to evaluate the applicability of our previously developed method to the present PEFC additive discovery problem.

For molecular generation, SMILES-X uses a second LSTM-based model trained to generate SMILES strings. Molecules are constructed token by token, using statistical likelihood and property-based heuristics to select the next character. Notably, SMILES-X reuses the property prediction model during generation to estimate the properties of partially built molecules, guiding the generation toward promising candidates even before the full structure is completed [55].

Our AL loop consisted of three cycles. After training on the initial dataset, the model was used to evaluate a list of organic molecules proposed by the experimental team based on domain knowledge and practical feasibility. These compounds were ranked by the SMILES-X prediction model according to their expected MA, and the top candidates were selected for experimental validation.

In the second cycle (Cycle 2), the model was retrained using the newly acquired experimental data, which improved its predictive accuracy and enabled its use not only for prediction but also for molecular generation. SMILES-X was then used to generate new molecules predicted to act as effective additives. These generated candidates were prioritized based on predicted performance and chemical feasibility, and were considered alongside additional researcher-suggested candidates.

In the third cycle (Cycle 3), the model was retrained using the full set of available data to further improve its reliability. This final step provided a more robust predictor for evaluating additive performance and concluded the AL process.

The training dataset was expanded after each cycle by incorporating newly measured and re-measured compounds, as detailed in section 3.

The AL workflow is summarized as a flowchart in figure 1, showing the progression from Cycle 1 to Cycle 3, including experimental data acquisition, model retraining, candidate evaluation, molecular generation, and feedback into subsequent cycles. The corresponding evolution of the ML model performance is summarized in table 3.

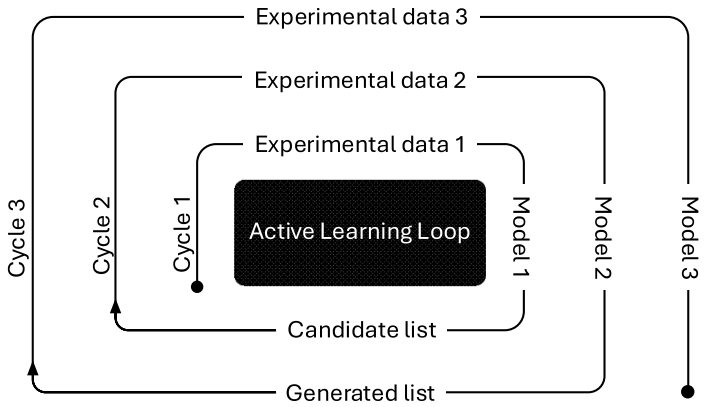

Figure 1. Overview of the active learning workflow combining SMILES-X-based prediction and generation with experimental validation.

Table 1. Hyperparameters selected for each dataset in the three active learning cycles[a].

Cycle	Batch size	Learning rate	Embedding size	LSTM units	Dense layer size
Cycle 1	16	$10^{-3.5}$	2	8	8
Cycle 2	8	$10^{-3.5}$	2	8	256
Cycle 3	8	$10^{-2.5}$	1	16	256

[a]Embedding sizes of 1–2 are sufficient here due to character-level SMILES tokenization and the small-data regime; increasing capacity led to overfitting.

Throughout all cycles, we focused on predicting improvements in MA at 0.9 V relative to a baseline electrolyte. Molecules were encoded into latent vectors using the LSTM encoder of SMILES-X, and a regression model was trained to estimate MA. Model performance was assessed using 10-fold cross-validation, where the data is split into ten parts, and each subset is used once as a test set while the remaining nine serve for training. For each fold, we performed five independent training runs with different random seeds to account for variability due to model initialization and training dynamics. We report the average prediction across these five runs for each test fold, and use the standard deviation across runs to indicate uncertainty, represented as vertical error bars on the prediction plots.

The model architecture and training process were governed by several hyperparameters, including batch size and learning rate, which influence training stability and convergence, as well as embedding size, number of LSTM units, and the size of the time-distributed dense layer, which determine the model's representational capacity. The architecture-related hyperparameters of the SMILES-X model were selected using the zero-shot neural architecture search algorithm epsinas [56]. This procedure was applied separately in each AL cycle to tailor the geometry of the LSTM-based architecture to the available experimental dataset, enabling dataset-specific model design in the low-data regime considered in this work. The selected hyperparameter values for each cycle are listed in table 1.

2.2. Characterization

The ORR activity of Pt catalyst (TEC10E50E, Pt content: 46.5%, Tanaka Kikinzoku Kogyo Co. Ltd) was measured by a method described in [21, 23, 24]. Briefly, electrochemical measurements of a Pt catalyst-modified electrode (evaluation of the ORR activity and electrochemically active surface area (ECSA)) were conducted in the absence of additives. Then, the modified electrode was pulled up from the electrolyte solution (0.1 M HClO4), and the electrode was immersed in the solution of the additives for 10 min. Then, the electrode was retrieved from the solution of the additives, rinsed with purified water, and re-immersed in the electrolyte solution (0.1 M HClO4). The same electrochemical measurements were conducted again. By comparing the electrochemical properties after and before the treatment with additives, effects of the additives on the catalytic activity can be revealed. The ORR MA was calculated at 0.9 V and 0.95 V (*vs* a reversible hydrogen electrode). Relative MA is defined as the ratio of the MA after/before the treatment of the additives, and is discussed as an indicator of the effect of additives.

Specific activity is determined by dividing MA with ECSA. Most of the additives were dissolved in acetone at 0.7 mM. The other compounds were dissolved in water, 0.1 M HClO4, or dimethylsulfoxide. We selected the solvent based on the solubility of the additives. Concentrations of some compounds were changed from 0.7 mM to other values (0.02 mM, 0.1 mM, 0.85 mM, 3 mM, 15 mM, or 20 mM).

3. Results

We first constructed a curated molecular dataset for model training and active-learning-driven candidate selection. For modeling consistency, we selected the fixed-concentration subset (0.7 mM; 68 samples) for Cycle 1 training (table 2). After adding 17 new measurements in Cycle 2 and several re-measured points, the training set grew to 85. Cycle 3 incorporated 12 additional measurements (8 expert-suggested and 4 generated, commercially available), yielding 93 training samples used for the final predictor (figure 5 and table 3).

Each compound was characterized thanks to experimentally measured properties relevant to fuel cell performance. These included MA at 0.9 V and 0.95 V, and ECSA, measured as dimensionless ratios before and after additive adsorption. Initial analysis showed strong correlations between these properties (see appendix, figure A1). Since slightly more data was available for MA at 0.9 V, we selected it as the primary prediction target for model training.

Cycle 1: initial model and candidate evaluation

To build the first model, all candidate compounds were validated using RDKit [57], an open-source cheminformatics toolkit widely used for molecular structure handling. RDKit was used to parse the SMILES strings and ensure that they cORResponded to chemically valid molecules, i.e. structures that could be interpreted without errors and complied with basic valence and bonding rules. Of the 96 compounds considered, 84 passed this validation step and were used for model training. These samples spanned various solvent and concentration combinations. To ensure consistency in the training data, we compared four filtering strategies based on whether solvent type and additive concentration were fixed:

- Use all data regardless of additive concentration and solvent
- Fix additive concentration, disregard solvent type
- Fix solvent, disregard additive concentration
- Fix both concentration and solvent

Table 2 summarizes these dataset sizes. From a chemical standpoint, additive concentration is expected to influence performance more than solvent choice. To confirm that filtering did not adversely affect model accuracy, we trained a separate model for each filtering strategy using the same set of hyperparameters, which were optimized on the full dataset. As shown in figure 2, the differences in performance were minimal. We therefore chose the dataset with fixed concentration (68 samples) to balance size and consistency.

Using this model (Model 1), we evaluated a list of 34 candidate molecules provided by the experimental collaborators. These predictions helped prioritize compounds for experimental validation in the next cycle.

Cycle 2: updated model and candidate generation

Figure 3 compares the predictions of Model 1 with the experimental measurements obtained for the 18 expert-suggested candidate additives. This step served as a prospective evaluation of the initial model on newly proposed compounds, rather than as an assessment of a retrained model on an expanded dataset. The comparison showed moderate predictive agreement: while compound-level deviations remained, the prospective error was comparable to the cross-validation error, indicating that the initial model could provide useful trend-level guidance while still benefiting from additional consistent experimental data. These newly acquired measurements were therefore added to the original dataset to create an expanded training set. In addition, several samples were re-evaluated to improve data consistency and overall data quality.

We retrained the SMILES-X model on this updated data, consisting of 85 data points, to produce Model 2, which showed improved accuracy due to the increased sample size and more targeted data

Table 2. Number of valid RDKit entries under different data filtering strategies. Fixing additive concentration while allowing solvent variation provided a good balance between data consistency and size.

Filtering strategy	Number of samples
No filtering (all data)	84
Fixed concentration only	68
Fixed solvent only	61
Fixed concentration and solvent	48

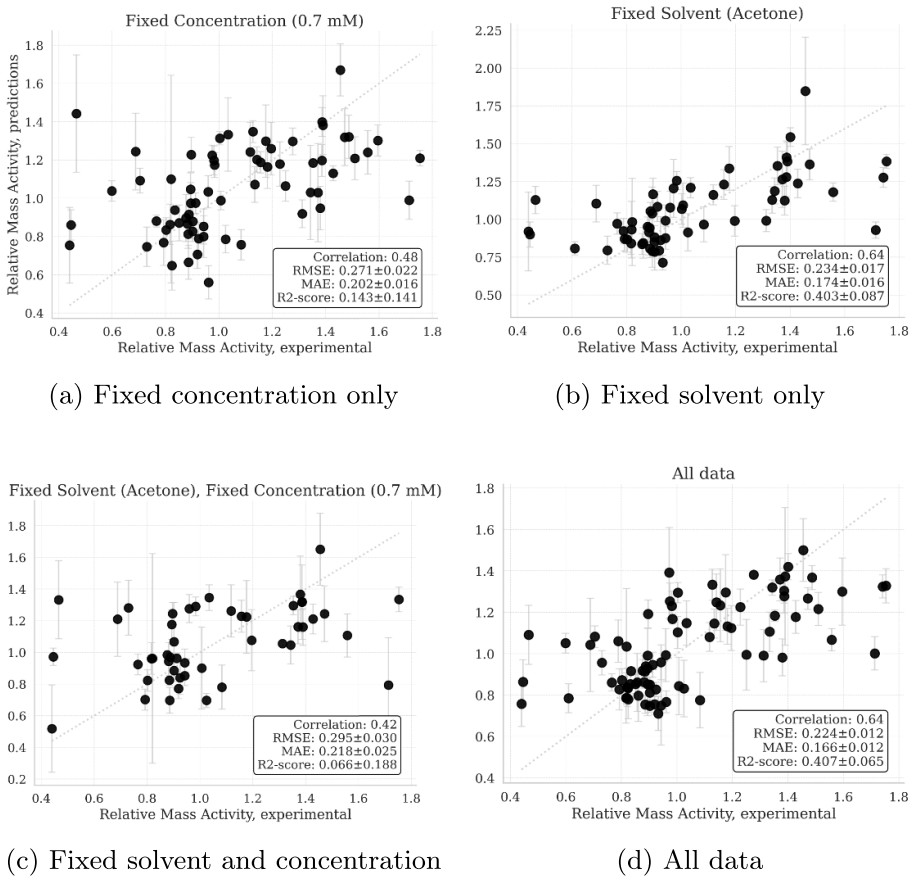

Figure 2. Performance of initial SMILES-X models trained on datasets filtered by different additive concentration and solvent criteria.

(figure 4). Model 2 was also used to guide a generative model trained on the same data. Together, they suggested new candidate molecules predicted to improve MA. Notably, four of these generated molecules were found to be commercially available, allowing direct experimental validation without additional synthesis steps.

Cycle 3: final model training

In the final cycle, eight more compounds from the chemists' candidate list were evaluated, along with the four commercially sourced molecules generated in Cycle 2, bringing the total dataset to 93 data points. These new measurements were added to the training set, and the final predictive model was trained using the full dataset (Model 3). As shown in figure 5, the final model achieves the highest predictive accuracy among all training cycles. Red circles mark the SMILES-X–enerated molecules, all of which exhibit substantial performance gains, with MA improvements exceeding 20% relative to the no-additive baseline. The top-performing compound in this group is melam, a dimer of melamine, which boosts MA by 67%.

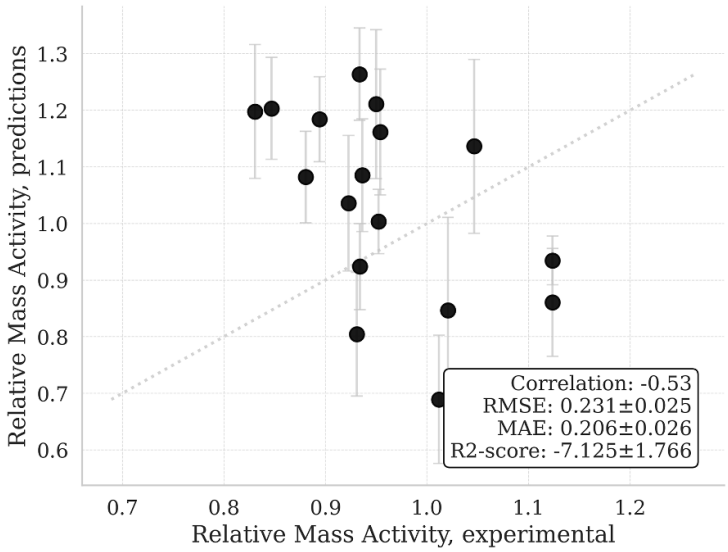

Figure 3. Predicted (gray) and experimental (black) relative mass activities for 17 expert-selected candidates measured after Cycle 1. Predictions come from the fixed-concentration model.

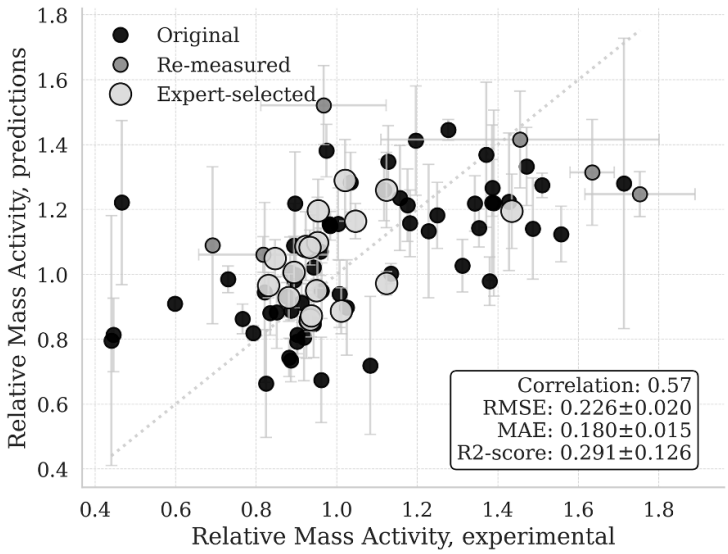

Figure 4. Cross-validation performance of Model 2 trained on the expanded dataset after Cycle 2. Black points represent the original dataset; dark gray points indicate re-measured data; large light-gray circles show results for chemist-suggested compounds.

4. Discussion

The development of next-generation energy materials increasingly depends on the ability to search large molecular design spaces efficiently. In this context, ML and AL offer powerful strategies to accelerate discovery, reduce experimental cost, and help navigate the trade-offs between performance, stability, and manufacturability. While much attention has been paid to models achieving state-of-the-art accuracy, our study highlights a complementary and often overlooked benefit: the ability of even modest, data-limited models to provide meaningful experimental guidance when integrated into a human-in-the-loop workflow.

Our results show that ML models, even when trained on small and noisy datasets, can effectively prioritize experimental targets. In this study, a model trained on only 68 initial measurements was used to select 17 promising additives from an expert-curated list. As shown in figure 3, the model's predictions for these new candidates showed useful trend-level agreement with the experimental outcomes, with test-set RMSE and MAE values comparable to those obtained during cross-validation. Specifically, the model

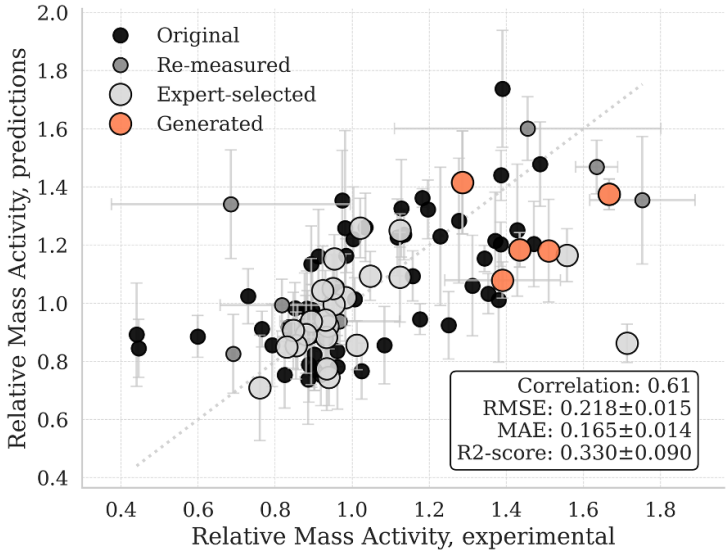

Figure 5. Final SMILES-X model performance using 93 data points. Black points represent the original dataset; dark gray points indicate re-measured data; large light-gray circles show results for chemist-suggested compounds; red circles correspond to generated compounds.

Table 3. Performance of ML models across three active learning cycles. All metrics are calculated by 10-fold cross-validation.

Cycle	Training samples	RMSE	MAE	Pearson r	R^2
Cycle 1	68	0.271	0.202	0.48	0.14
Cycle 2	85	0.226	0.180	0.57	0.29
Cycle 3	93	0.218	0.165	0.61	0.33

trained on the initial 68 samples achieved an out-of-sample RMSE of 0.271, while predictions on the 17 new samples yielded an RMSE of 0.231. This level of agreement indicates that even models trained on limited experimental datasets can detect performance trends, enabling researchers to focus effort on the most promising candidates and thereby reduce screening time and material waste.

Model performance improved steadily across three AL cycles, each consisting of retraining with new experimental data. As summarized in table 3, every cycle added new, informative samples that reduced prediction error and strengthened the model's ability to generalize. This effect is especially pronounced in small-sample settings typical of materials science, where each measurement carries significant informational weight. By Cycle 3, the model trained on the expanded dataset achieved the best overall performance, as reflected by the lowest RMSE and MAE and the highest Pearson correlation reported in table 3. These results confirm that the integration of human expertise, model-guided selection, and iterative feedback can progressively enhance predictive power.

The final model was further used to screen novel molecules generated using SMILES-X. While generative models can produce unrealistic or synthetically inaccessible structures, we focused on commercially available candidates to ensure experimental viability. Four of the top-ranked molecules were tested, all showing improved MA in line with the model's predictions, though in some cases the model slightly overestimated performance.

The strong performance of melam is particularly informative because it suggests that motifs related to melamine-based chemistry remain valuable targets for ORR enhancement, while also pointing to nearby molecular variants that may offer improved stability, cost, or process compatibility. Although melam suffers from stability and cost limitations, its performance demonstrates the model's ability to suggest useful directions for exploration. A second generated candidate, melamine, achieved a 51%. It is approximately three orders of magnitude less expensive than the best-performing compound, Co-tBuTAP (90% MA enhancement [21]), as of August 2025 [58, 59], and contains no transition metals, making it a highly practical alternative. These findings exemplify how even simple generative strategies, when

coupled with informed filtering and predictive models, can yield promising new candidates that balance performance with real-world constraints.

Taken together, these results illustrate the multi-faceted role ML can play in early-stage materials discovery. Beyond maximizing model accuracy, the integration of ML with expert insight and iterative experimental feedback enables more informed decisions, better data quality, and the discovery of alternative compounds with desirable properties. Our workflow shows that progress does not require perfect models, only useful ones, continuously improved by the right experiments. As data grows and tools evolve, such frameworks are likely to become standard practice in the search for new functional molecules.

5. Conclusion

We demonstrated the successful application of AL to optimize fuel cell performance through the screening of organic electrolyte additives. Our results show that ML models, even when trained on limited and noisy data, can meaningfully guide the selection of candidates for experimental validation. By integrating expert knowledge with predictive modeling and molecular generation, we were able to prioritize promising compounds.

Model performance improved steadily with each iteration, and the final model was used to screen new, commercially available molecules. Among the generated candidates, melam (a melamine dimer) achieved a 67% increase in MA relative to the additive-free baseline. Melamine itself also showed a 51% improvement while offering advantages in cost, availability, and solvent compatibility. Notably, melamine has been known to be a good ORR enhancer [23–26], but our ML model was not aware of this when generating the SMILES. These results illustrate the potential for ML-guided discovery to yield practical, sustainable alternatives even without large datasets or state-of-the-art models.

Overall, this work highlights the value of ML tools in small-data regimes. By combining predictive models with experimental iteration, candidate generation, and domain expertise, our approach provides a viable strategy for accelerating materials development in complex chemical systems.

Following the completion of the three AL cycles, the experimental team identified an improved additive application method that leads to significantly enhanced MA. This discovery opens a new direction for optimization, and the data and models developed during the current study are already being reused to inform the next set of experiments under the updated protocol.

In parallel, the existing models are being applied to related experimental setups involving alternative catalyst morphologies. These extensions highlight the broader applicability of the trained models and their underlying representations. Looking forward, the models developed here can be adapted to similar tasks through a combination of strategies. When conditions change between experiments (such as in catalyst structure, measurement setup, or additive application method) domain adaptation techniques become useful. For example, linear regression or other lightweight models can be employed to translate predictions between related domains, leveraging the same molecular features. Alternatively, neural networks can be fine-tuned via transfer learning by retraining only the final layers, allowing the core molecular understanding to be retained while adjusting to new conditions. These approaches extend the utility of ML in data-scarce scientific environments and offer a flexible framework for accelerating discovery across related systems.

Data availability statement

The data cannot be made publicly available upon publication because they contain commercially sensitive information. The data that support the findings of this study are available upon reasonable request from the authors.

Conflict of interests

The authors declare that they have no competing interests.

Funding

This work is supported by New Energy and Industrial Technology Development Organization (NEDO).

Scientific contribution

This work establishes a closed-loop experimental workflow in which predictive modeling, candidate generation, and feasibility-based experimental selection are combined to identify organic additives for improving Pt-catalyst ORR MA under low-data conditions.

Author contributions

Ekaterina Gracheva 0000-0002-9704-5939
Data curation (equal), Investigation (equal), Methodology (equal), Software (equal), Validation (equal), Visualization (equal), Writing – original draft (equal), Writing – review & editing (equal)

Shin-ichi Yamazaki
Conceptualization (equal), Data curation (equal), Formal analysis (equal), Investigation (equal), Validation (equal), Visualization (equal), Writing – original draft (equal), Writing – review & editing (equal)

Guillaume Lambard 0000-0003-0275-4079
Data curation (equal), Investigation (equal), Methodology (equal), Software (equal), Supervision (equal), Validation (equal), Visualization (equal), Writing – review & editing (equal)

Keitaro Sodeyama
Conceptualization (equal), Funding acquisition (equal), Project administration (equal), Supervision (equal), Writing – review & editing (equal)

Tsutomu Ioroi
Conceptualization (equal), Data curation (equal), Formal analysis (equal), Funding acquisition (equal), Investigation (equal), Project administration (equal), Resources (equal), Supervision (equal), Validation (equal), Visualization (equal), Writing – review & editing (equal)

Masafumi Asahi 0000-0002-2122-3073
Conceptualization (equal), Data curation (equal), Formal analysis (equal), Investigation (equal), Methodology (equal), Project administration (equal), Resources (equal), Supervision (equal), Validation (equal), Visualization (equal), Writing – review & editing (equal)

Appendix. Correlation matrix of measured properties

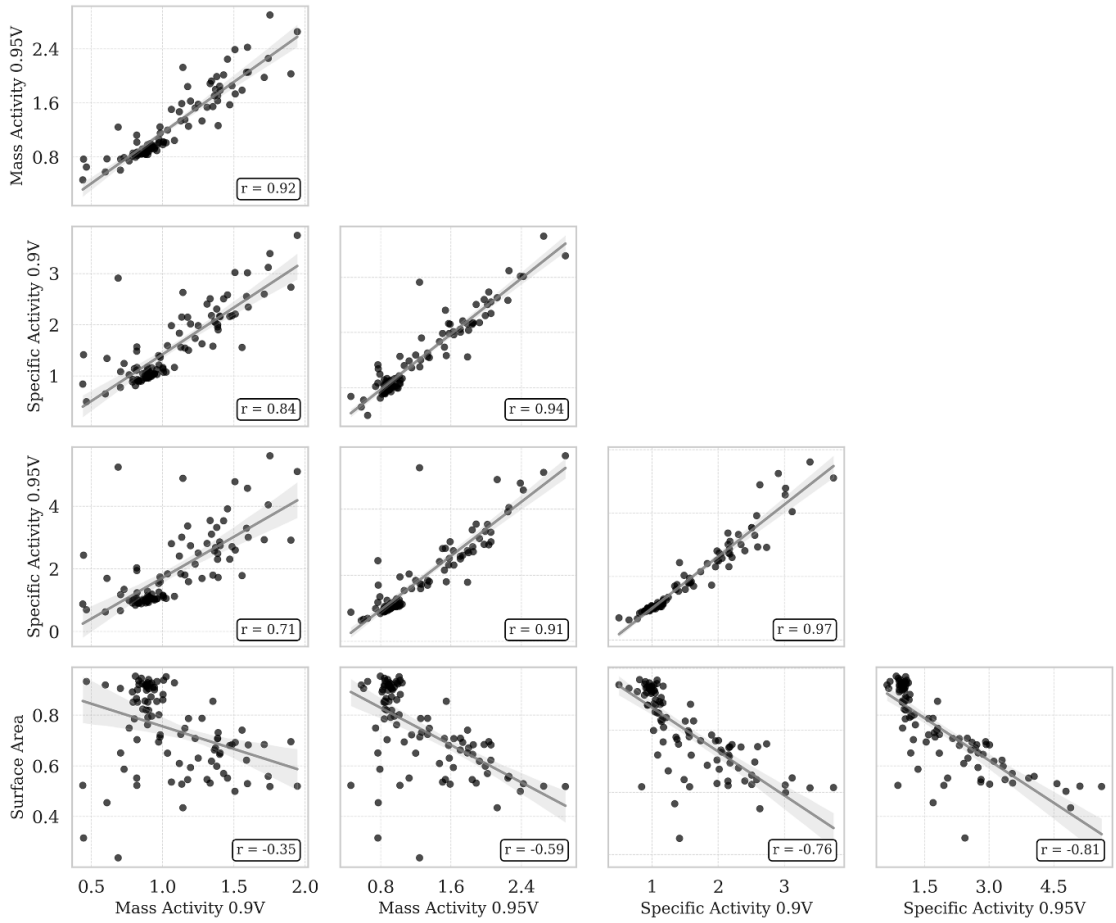

Figure A1. Correlation matrix of measured properties across the initial dataset. Mass activity and Specific activity at 0.9 V, 0.95 V, show strong mutual Correlation, justifying the use of MA at 0.9 V as a primary prediction target.

References

[1] Majlan E, Rohendi D, Daud W, Husaini T and Haque M 2018 Electrode for proton exchange membrane fuel cells: a review *Renew. Sustain. Energy Rev.* **89** 117–34

[2] Gasteiger H A and Marković N M 2009 Just a dream—or future reality? *Science* **324** 48–49

[3] Zeng W, Guan B, Zhuang Z, Chen J, Zhu L, Ma Z, Hu X, Zhu C, Zhao S and Shu K 2025 Comprehensive review on the advances and comparisons of proton exchange membrane fuel cells (pemfcs) and anion exchange membrane fuel cells (afcs): from fundamental principles to key component technologies *Int. J. Hydrog. Energy* **102** 222–46

[4] Vallejo-Cervantes C, Espinoza-Andaluz M and Iranzo A 2025 Technical review of commercial lt-pemfc technologies: performance, applications and challenges *Int. J. Hydrog. Energy* **176** 151480

[5] Liu C-Y and Sung C-C 2012 A review of the performance and analysis of proton exchange membrane fuel cell membrane electrode assemblies *J. Power Sources* **220** 348–53

[6] Guo Y, Pan F, Chen W, Ding Z, Yang D, Li B, Ming P and Zhang C 2021 The controllable design of catalyst inks to enhance pemfc performance: a review *Electrochem. Energy Rev.* **4** 67–100

[7] Borup R L, Kusoglu A, Neyerlin K C, Mukundan R, Ahluwalia R K, Cullen D A, More K L, Weber A Z and Myers D J 2020 Recent developments in catalyst-related pem fuel cell durability *Curr. Opin. Electrochem.* **21** 192–200

[8] Hidayat E F, Zain S M and Noviyanti A R 2025 Advances in proton exchange membrane fuel cell (pemfc) materials: a review of developments from 2021 to 2025 *J. Power Sources* **657** 238124

[9] Zhang F, Zu B, Wang B, Qin Z, Yao J, Wang Z, Fan L and Jiao K 2025 Developing long-durability proton-exchange membrane fuel cells *Joule* **9** 101853

[10] Qasem N A 2024 A recent overview of proton exchange membrane fuel cells: fundamentals, applications and advances *Appl. Therm. Eng.* **252** 123746

[11] Chen S, Liu Y, Fu X and Wang W 2024 Recent advances in morphology control of platinum catalysts toward oxygen reduction reaction *Front. Energy* **18** 330–55

[12] Li J-R, Liu M-X, Liu X, Yu X-H, Li Q-Z, Sun Q, Sun T, Cao S and Hou C-C 2024 The recent progress of oxygen reduction electrocatalysts used at fuel cell level *Small Methods* **8** 2301249

[13] Snyder J, Fujita T, Chen M and Erlebacher J 2010 Oxygen reduction in nanoporous metal–ionic liquid composite electrocatalysts *Nat. Mater.* **9** 904–7

[14] Zhang G-R, Munoz M and Etzold B J 2016 Accelerating oxygen-reduction catalysts through preventing poisoning with non-reactive species by using hydrophobic ionic liquids *Angew. Chem., Int. Edn.* **55** 2257–61
[15] Li Y, Hart J, Profitt L, Intikhab S, Chatterjee S, Taheri M and Snyder J 2019 Sequential capacitive deposition of ionic liquids for conformal thin film coatings on oxygen reduction reaction electrocatalysts *ACS Catal.* **9** 9311–6
[16] Li Y, Intikhab S, Malkani A, Xu B and Snyder J 2020 Ionic liquid additives for the mitigation of nafion specific adsorption on platinum *ACS Catal.* **10** 7691–8
[17] Avid A, Ochoa J L, Huang Y, Liu Y, Atanassov P and Zenyuk I V 2022 Revealing the role of ionic liquids in promoting fuel cell catalysts reactivity and durability *Nat. Commun.* **13** 6349
[18] Wang L, Morales-Collazo O, Brennecke J F and Jia H 2024 Impact of a protic dimeric ionic liquid on pem fuel cell performance and durability *Int. J. Hydrog. Energy* **58** 40–45
[19] Chung Y-H, Chung D Y, Jung N and Sung Y-E 2013 Tailoring the electronic structure of nanoelectrocatalysts induced by a surface-capping organic molecule for the oxygen reduction reaction *J. Phys. Chem. Lett.* **4** 1304–9
[20] Miyabayashi K, Nishihara H and Miyake M 2014 Platinum nanoparticles modified with alkylamine derivatives as an active and stable catalyst for oxygen reduction reaction *Langmuir* **30** 2936–42
[21] Yamazaki S, Asahi M and Ioroi T 2019 Promotion of oxygen reduction on a porphyrazine-modified pt catalyst surface *Electrochim. Acta* **297** 725–34
[22] Yamazaki S, Asahi M, Taguchi N and Ioroi T 2019 Electrochemical analysis of the porphyrazine-induced enhancement of ORR activity of pt catalysts for the development of porphyrazine-adsorbed pt catalysts *J. Electroanal. Chem.* **848** 113321
[23] Asahi M, Yamazaki S, Taguchi N and Ioroi T 2019 Facile approach to enhance oxygen reduction activity by modification of platinum nanoparticles by melamine-formaldehyde polymer *J. Electrochem. Soc.* **166** 498
[24] Yamazaki S, Asahi M, Taguchi N, Ioroi T, Kishimoto Y, Daimon H, Inaba M, Koga K, Kurose Y and Inoue H 2020 Creation of a highly active pt/pd/c core–shell-structured catalyst by synergistic combination of intrinsically high activity and surface decoration with melamine or tetra-(tert-butyl)-tetraazaporphyrin *ACS Catal.* **10** 14567–80
[25] Yamazaki S, Asahi M, Taguchi N and Ioroi T 2022 Sulphonated melamine polymer for enhancing the oxygen reduction reaction activity and stability of a pt catalyst *J. Electroanal. Chem.* **908** 116103
[26] Wada N, Nakamura M and Hoshi N 2020 Structural effects on the oxygen reduction reaction on pt single-crystal electrodes modified with melamine *Electrocatalysis* **11** 275–81
[27] Yamazaki S, Asahi M, Taguchi N and Ioroi T 2023 Triazine compounds for the high enhancement of the ORR activity and durability of Pt/C *Electrochim. Acta* **472** 143417
[28] Jinnouchi R and Minami S 2024 The melamine-driven solvation effect promotes oxygen reduction on a platinum catalyst: machine learning-aided free energy calculations *J. Phys. Chem. Lett.* **16** 265–73
[29] Butler K T, Davies D W, Cartwright H, Isayev O and Walsh A 2018 Machine learning for molecular and materials science *Nature* **559** 547–55
[30] Goldsmith B R, Esterhuizen J, Liu J-X, Bartel C J, Sutton C 2018 Machine learning for heterogeneous catalyst design and discovery
[31] McCullough K, Williams T, Mingle K, Jamshidi P and Lauterbach J 2020 High-throughput experimentation meets artificial intelligence: a new pathway to catalyst discovery *Phys. Chem. Chem. Phys.* **22** 11174–96
[32] Kim C, Batra R, Chen L, Tran H and Ramprasad R 2021 Polymer design using genetic algorithm and machine learning *Comput. Mater. Sci.* **186** 110067
[33] Sharma A, Mukhopadhyay T, Rangappa S M, Siengchin S and Kushvaha V 2022 Advances in computational intelligence of polymer composite materials: machine learning assisted modeling, analysis and design *Arch. Comput. Methods Eng.* **29** 3341–85
[34] Gracheva E, Lambard G, Samitsu S, Sodeyama K and Nakata A 2021 Prediction of the coefficient of linear thermal expansion for the amorphous homopolymers based on chemical structure using machine learning *Sci. Technol. Adv. Mater.* **1** 213–24
[35] Wu S, Kondo Y, Kakimoto M-A, Yang B, Yamada H, Kuwajima I, Lambard G, Hongo K, Xu Y and Shiomi J 2019 Machine-learning-assisted discovery of polymers with high thermal conductivity using a molecular design algorithm *npj Comput. Mater.* **5** 66
[36] Yang X, Wang Y, Byrne R, Schneider G and Yang S 2019 Concepts of artificial intelligence for computer-assisted drug discovery *Chem. Rev.* **119** 10520–94
[37] Jiménez-Luna J, Grisoni F and Schneider G 2020 Drug discovery with explainable artificial intelligence *Nat. Mach. Intell.* **2** 573–84
[38] Ding R, Zhang S, Chen Y, Rui Z, Hua K, Wu Y, Li X, Duan X, Wang X and Li J 2022 Application of machine learning in optimizing proton exchange membrane fuel cells: a review *Energy AI* **9** 100170
[39] Ming W, Sun P, Zhang Z, Qiu W, Du J, Li X, Zhang Y, Zhang G, Liu K and Wang Y 2023 A systematic review of machine learning methods applied to fuel cells in performance evaluation, durability prediction and application monitoring *Int. J. Hydrog. Energy* **48** 5197–228
[40] Settles B 2012 *Active Learning* (*Synthesis Lectures on Artificial Intelligence and Machine Learning*) (Morgan & Claypool Publishers)
[41] Chen C-T and Gu G X 2020 Generative deep neural networks for inverse materials design using backpropagation and active learning *Adv. Sci.* **7** 1902607
[42] Lookman T, Balachandran P V, Xue D and Yuan R 2019 Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design *npj Comput. Mater.* **5** 21
[43] Xu P, Ji X, Li M and Lu W 2023 Small data machine learning in materials science *npj Comput. Mater.* **9** 42
[44] Lambard G, Sasaki T, Sodeyama K, Ohkubo T and Hono K 2022 Optimization of direct extrusion process for Nd-Fe-B magnets using active learning assisted by machine learning and bayesian optimization *Scr. Mater.* **209** 114341
[45] Kim C, Chandrasekaran A, Jha A and Ramprasad R 2019 Active-learning and materials design: the example of high glass transition temperature polymers *Mrs Commun.* **9** 860–6
[46] Ang S J, Wang W, Schwalbe-Koda D, Axelrod S and Gómez-Bombarelli R 2021 Active learning accelerates ab initio molecular dynamics on reactive energy surfaces *Chem* **7** 738–51
[47] Lambard G and Gracheva E 2020 Smiles-x: autonomous molecular compounds characterization for small datasets without descriptors *Mach. Learn.: Sci. Technol.* **1** 025004
[48] Weininger D 1988 Smiles, a chemical language and information system. 1. Introduction to methodology and encoding rules *J. Chem. Inf. Comput. Sci.* **28** 31–36
[49] Wu X, Wang H, Gong Y, Fan D, Ding P, Li Q and Qian Q 2023 Graph neural networks for molecular and materials representation *J. Mater. Inform.* **3** 12
[50] Li L, Zhang Y, Wang G and Xia K 2025 Kolmogorov–arnold graph neural networks for molecular property prediction *Nat. Mach. Intell.* **7** 1346–54

[51] Alon U and Yahav E 2020 On the bottleneck of graph neural networks and its practical implications (arXiv: 2006.05205)
[52] Di Giovanni F, Giusti L, Barbero F, Luise G, Lio P and Bronstein M M 2023 On over-squashing in message passing neural networks: the impact of width, depth and topology *Int. Conf. on Machine Learning* (PMLR) pp 7865–85
[53] Thürlemann M and Riniker S 2023 Anisotropic message passing: graph neural networks with directional and long-range interactions *The 11th Int. Conf. on Learning Representations*
[54] Hochreiter S and Schmidhuber J 1997 Long short-term memory *Neural Comput.* **9** 1735–80
[55] Lambard G SMILES-X: a molecular machine learning tool (available at: https://github.com/Lambard-ML-Team/SMILES-X) (Accessed 03 July 2025)
[56] Gracheva E 2025 Neural architecture search: two constant shared weights initialisations *Artif. Intell. Rev.* **58** 296
[57] RDKit: open-source cheminformatics (available at: www.rdkit.org/)
[58] TCI Chemicals Melamine (CAS No. 108-78-1) - Product T0337 (available at: www.tcichemicals.com/JP/en/p/T0337) (Accessed 20 August 2025)
[59] Sigma-Aldrich 2,7,12,17-Tetra-tert-butyl-5,10,15,20-tetraaza-21H,23H-porphyrin (Dye Content 85) (available at: www.sigmaaldrich.com/JP/ja/product/sial/42836) (Accessed 7 September 2025)