

# Computational Catalysis and Machine Learning Applications to Water Treatment Technologies

<sup>1</sup>Duo Wang, <sup>2</sup>Ao Xie, <sup>1</sup>Shengcun Ma, <sup>1</sup>Wei Tong, <sup>2</sup>Shiqiang Zou, <sup>1</sup>Anubhav Jain\*

<sup>1</sup> *Energy Storage and Distributed Resources Division, Lawrence Berkeley National Laboratory, Berkeley, California, 94720 USA*

<sup>2</sup> *Department of Civil and Environmental Engineering, Auburn University, Auburn, Alabama, 36849 USA*

\* Corresponding author: [ajain@lbl.gov](mailto:ajain@lbl.gov)

## 1. Abstract

Electrocatalysis can offer new routes to water treatment. For example, electrocatalytic reduction is an emerging technology for treating oxyanions of concern in water. However, identification of highly performant, cost-effective catalysts remains a major barrier to deployment at scale. This article discusses how computational modeling and machine learning can accelerate the search for new catalyst materials. It describes how traditional computational chemistry workflows, now deployed in their basic form for at least two decades, can be expanded in breadth and depth through newly developed machine-learned force fields that have been trained on millions of examples. It also discusses how the theory and machine learning pipeline can effectively integrate with experimental synthesis and characterization platforms to rapidly identify and validate new catalyst chemistries.

## 2. Introduction and opportunities for electrocatalysis in water purification

Ensuring global access to clean water represents a major challenge for the 21st century that requires new water treatment processes. With the rapidly increasing deployment of distributed renewable electricity sources such as solar and wind power, an opportunity exists to develop new cost-effective water treatment processes that are similarly electrified, modular, and distributed.

Electrified water purification harnesses electricity as the thermodynamic driving force to efficiently separate and transform substances of concern from water[1]. Compared to traditional physicochemical water treatment, electrified technologies demand minimal chemical inputs and post-treatment management, thereby reducing operational costs and secondary pollution[2].

Electrified purification methods can utilize both non-Faradaic and Faradaic mechanisms. Processes based on non-Faradaic mechanisms (*i.e.*, without charge transfer), such as electrosorption and electromigration, include techniques such as electrodialysis (ED), electrodeionization (EDI), and capacitive deionization (CDI). These processes leverage electric fields to separate substances of concern and find widespread application in desalination and producing high-quality water[3]. Non-Faradaic water treatment technologies typically operate at lower system voltages ( $< 1.2$  V)[4], resulting in competitive energy consumption (as low as  $0.1$  kWh m<sup>-3</sup> water)[5].

In contrast, electrified water treatment processes centered on Faradaic (*i.e.*, charge transfer) mechanisms, such as electroreduction and electrooxidation, play a crucial role in transforming substances of concern into less harmful, more valuable, or separable forms through redox reactions. These processes include electrolysis, electroplating, electrocoagulation, and electroflotation, which are pivotal in the degradation of organic pollutants, precipitation of heavy metals, and removal of suspended particles[1]. Companies have introduced modular and compact electrocoagulation systems that offer cost-effective industrial wastewater treatment to remove turbidity, metals, oils, bacteria, etc.[6]. Compared to non-Faradaic treatment processes, Faradaic technologies operate within a broader voltage range (up to  $4.0$  V)[7], resulting in a significantly higher energy input than what is predicted by thermodynamics, sometimes reaching up to two orders of magnitude higher than non-Faradaic processes[8].

To reduce the energy input needed for Faradaic water treatment, catalysts have been introduced to mitigate the substantial deviation (*e.g.*, overpotential) from thermodynamics. These catalysts enhance surface affinity, strengthen adsorption, and reduce the activation barrier of target redox reactions; they can also help control side reactions and the final products of the electrochemical process. For example, electrocatalysis can facilitate H<sub>2</sub>O<sub>2</sub> generation as an aquatic disinfectant as well as the transformation of selenate to elemental selenium (*e.g.*, for resource recovery) instead of producing toxic hydrogen selenide[1, 9]. Catalysts are also used to selectively control overpotentials, enabling the redox transformation of recalcitrant aquatic pollutants that surpass the typical electrochemical window of water[1]. For instance, electrooxidation of per- and polyfluorinated alkyl substances (PFAS) can be significantly enhanced by pairing electrodes with high oxygen evolution overpotentials (such as boron-doped diamond, Ti<sub>4</sub>O<sub>7</sub>, SnO<sub>2</sub>, and PbO<sub>2</sub>[10]) with catalysts that facilitate PFAS oxidation.

Although electrocatalysis presents a promising avenue for unlocking Faradaic mechanisms in practical water purification processes, it necessitates concerted efforts to overcome three diverse challenges (**Figure 1**): (i) catalyst materials development, (ii) electrode design, and (iii) system operation. Although careful consideration of all three factors is essential for successful process development[11], the remainder of this article focuses on opportunities for accelerated catalyst materials discovery using theoretical techniques and machine learning (ML).

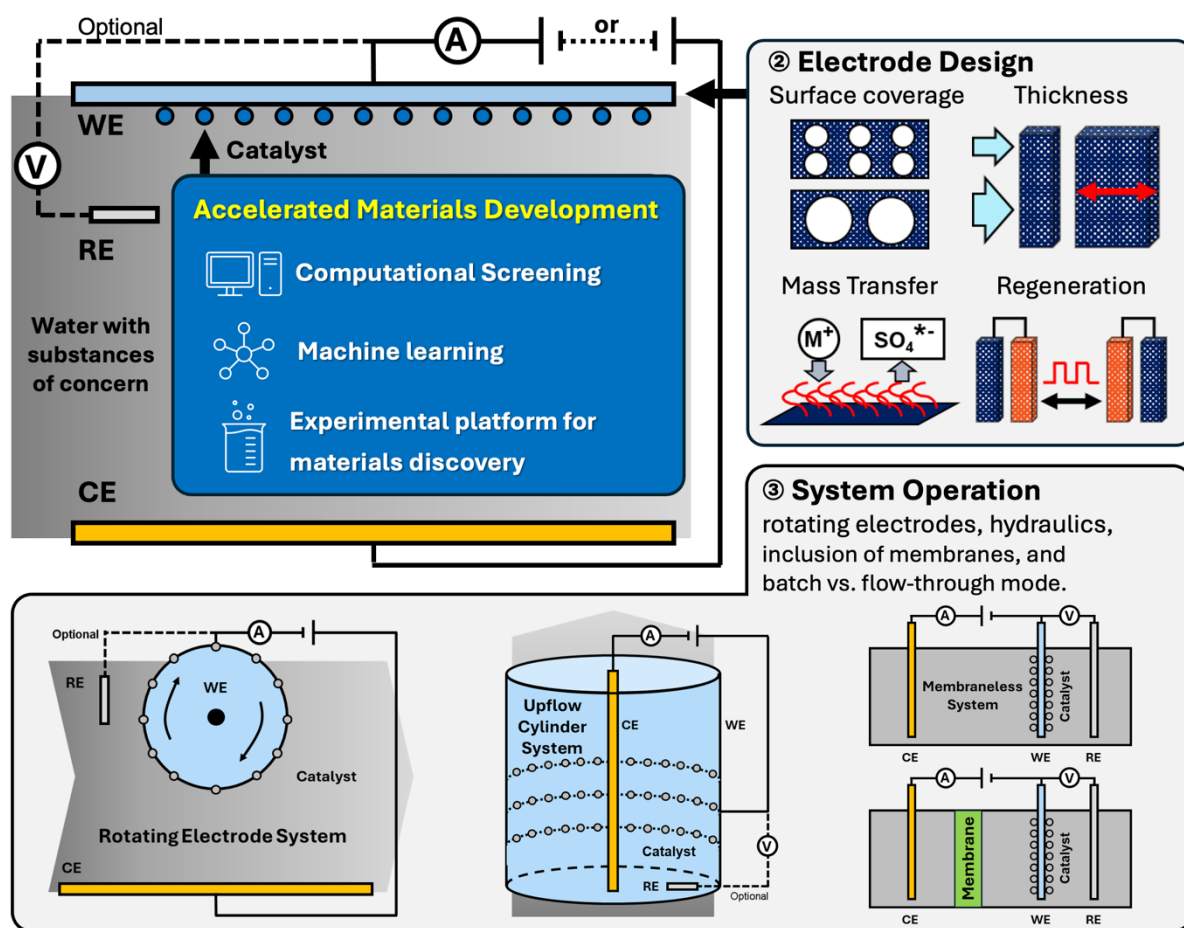


Figure 1 Challenges in realizing electrocatalytic water treatment, including materials (the focus of this article), electrodes, and reactor design. Acronyms used are WE (working electrode), RE (reference electrode), and CE (counter electrode).

### 3. Theoretical screening of catalysts: current practice and limitations

The application of theoretical screening techniques in catalyst design has evolved over multiple decades, expanding across various categories from hydrogen/oxygen evolution reactions, nitrate reduction, to methane/ethanol generation[12-15]. These broad applications highlight the

capability of computational techniques (typically based on density functional theory, or DFT, calculations) in virtual catalyst identification[16].

DFT calculations model reactions at the atomic scale; the resulting calculated reaction parameters (*e.g.*, thermodynamic driving force, activation energy) can be mapped to macroscopic catalyst performance through microkinetic modeling techniques. The overall aim of the modeling is to rapidly simulate which materials (*i.e.*, chemical compositions, surface structures) facilitate the desired surface reactions while avoiding any unintended side reactions. Materials with promising performance in the simulations are prioritized for experimental investigation.

To perform computational catalysis design, the first step is to determine the relevant sequence of surface reactions for a desired electrocatalytic process. Following this, the basic steps of computational catalyst design are outlined in **Figure 2**. First, adsorption energies and reaction barriers are calculated for the targeted surface and reaction species (**Figure 2a**). In general, according to the Sabatier principle, the adsorption energies of reactants on the catalyst surface should neither be too weak nor too strong; the highest catalytic activity tends to occur at adsorption energies close to zero[15, 17] as reflected in so-called "volcano plots" that show clear peaks in catalytic activity as a function of adsorption strength. This balance directly influences the activation energies[18] of reaction pathway steps, which can be precisely calculated using the nudged elastic band (NEB) method within DFT calculations (**Figure 2b**).

Other factors such as pH, applied potential, and surface coverage can significantly impact catalytic efficiency[18]. Such effects can be incorporated into the DFT calculations and used as inputs for microkinetic modeling[19] techniques that use the microscopic information to yield macroscopic reaction rates (**Figure 2d**). With the connection between microscopic parameters and macroscopic performance mapped out, the most promising catalyst compositions and surface

facets can be identified under various conditions via computer simulations. In particular, large databases of materials structures[20] can be used as a basis to generate surface facets and screened computationally for catalytic performance (**Figure 2e**).

Despite the overall successes of theoretical approaches, many simplifications are needed in practice to enable screening of many catalyst candidates at reasonable computing cost. In particular, scaling relationships (in which adsorption energies of complex molecules are modeled using simpler adsorption energies) (**Figure 2c**) and the Brønsted-Evans-Polanyi (BEP) relation (in which activation energies are estimated based on thermodynamic reaction energies) are commonly used approximations when screening materials[15, 19]. These techniques reduce the number and complexity of calculations that need to be performed at the expense of prediction accuracy. Other limitations include constraints on the number of atoms to ~1000 (making it challenging to capture behaviors at higher length scales that are observed in experiments), difficulty of mapping the initial reaction pathways, and high amount of researcher and computer time needed to perform the screening and model secondary effects. As discussed in the next section, machine learning techniques may help mitigate some of these challenges and unlock more rapid and more comprehensive computational screening.

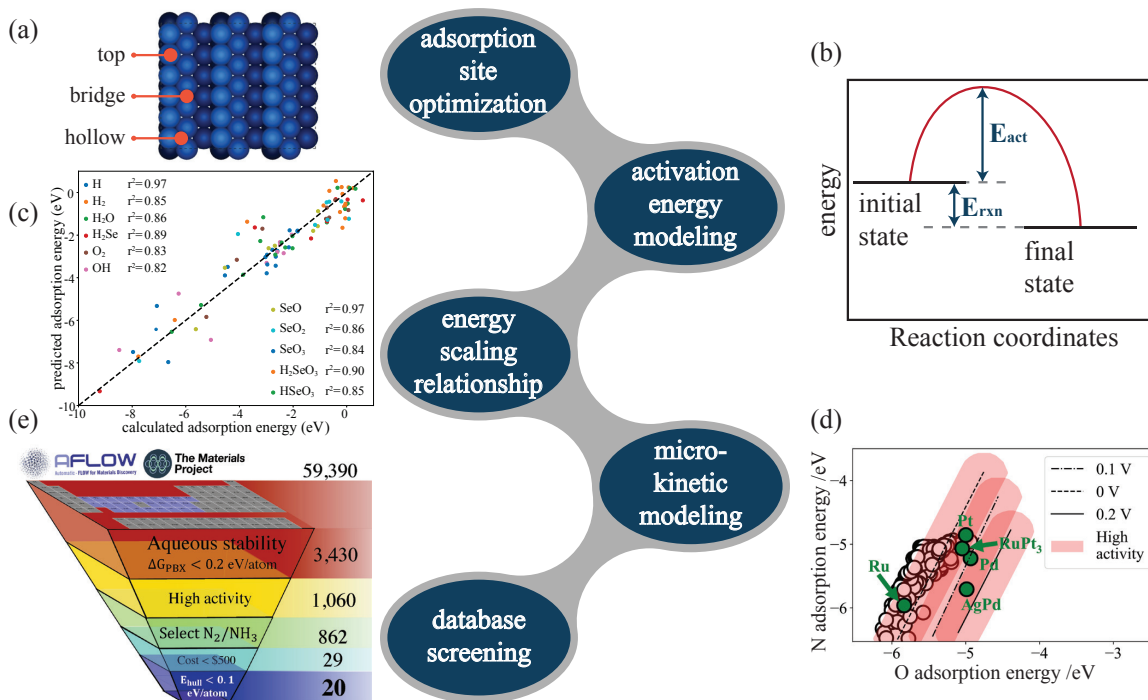


Figure 2: Typical workflow for computational catalysis. (a) Potential adsorption sites on a single materials surface that must be evaluated; (b) activation energy modeling; (c) development of scaling relationships to reduce the complexity of calculations in high-throughput screening; (d) the volcano plot of activity for nitrate removal (figure is obtained from Ref.[14]); and (e) the screening workflow from databases(figure is obtained from Ref.[14]). Copyright 2022 AIP Publishing.

## 4. Opportunities for machine learning to expand capabilities

Machine learning is rapidly transforming many research areas including computational catalysis. Today, machine learning models trained on historical databases of computations can calculate adsorption energies 3 orders of magnitude faster than the equivalent density functional theory procedure while retaining reasonable accuracy[21]. The existence of ML models therefore creates many new opportunities to expand the scope of computational screening in catalyst

identification. Indeed, this transformation is already underway[22] and elements of what is possible are being employed in practice[14].

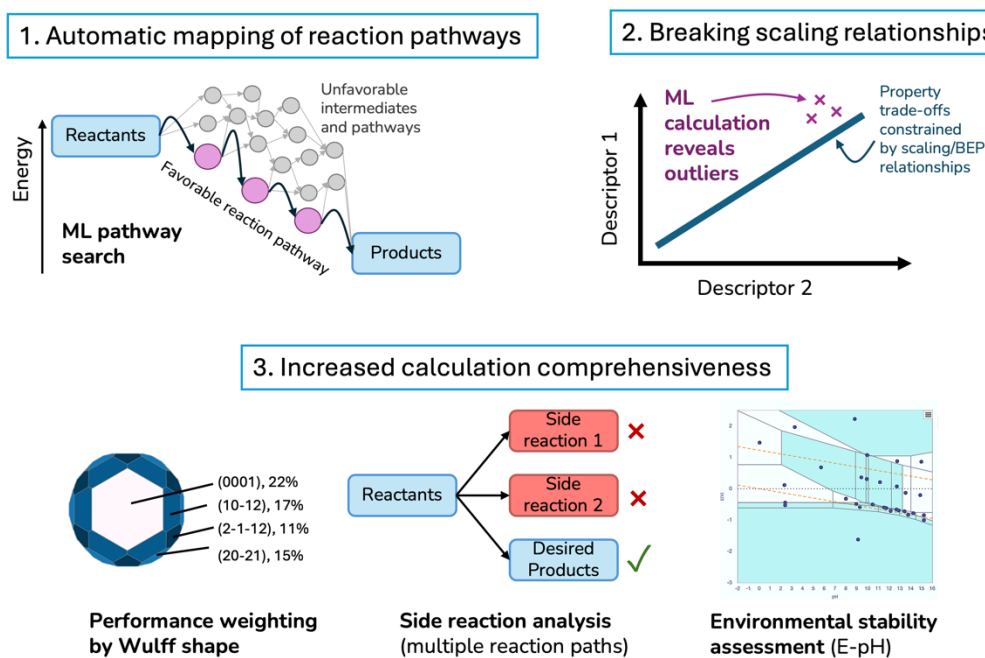


Figure 3 Illustration of 3 areas in which the development of machine learning potentials can improve computational screening workflows.

Today's state-of-the-art machine learned potentials typically take the form of graph neural networks[23], in which atoms of the proposed material are represented as nodes and neighboring atoms are connected by edges. Through successive layers of a neural network, atoms interact with one another through their edges to update their properties. For example, an atom neighboring an oxygen atom may receive a different update than an atom neighboring a palladium atom. As more updates are applied via additional layers, longer range interactions between atoms and their neighbors are captured. At the end of the process, a prediction is made regarding the total energy of the configuration.



Developing accurate and transferable machine learning potentials requires large training data sets from which the neural networks can learn patterns. The largest such data sets today originate from the Open Catalyst Project (OCP) (<https://opencatalystproject.org>). The latest version as of this writing, OC22[24], consists of >62,000 DFT relaxation trajectories made up of nearly 10 million single point calculations that can be used for training. This massive investment of both computer time and researcher labor has enabled the ability to train machine learned potentials with MAEs as low as 0.18 eV as compared to DFT[21]. This makes them sufficient as a "drop-in" replacement for DFT for many practical screening purposes. However, as outlined in **Figure 3**, many other opportunities exist[25] for applying these machine-learned models.

A first opportunity is to automatically map the reaction pathways for a desired overall reaction. Currently, a proposed reaction pathway (*i.e.*, the sequence of intermediates that are formed as various oxidation and reduction reactions occur on a surface) must be mapped in advance to define the calculations that need to be performed. Establishing a reaction pathway can require a mix of expert knowledge, comprehensive characterization, and calculated tests of hypotheses. Furthermore, such pathways may change depending on environment and applied conditions (pH, voltage). Moving forward, machine learning could be applied to automatically enumerate possible reaction pathways and automatically determine thermodynamically favorable pathways via chemical reaction networks[26] or other techniques. Using ML to enumerate and rank possibilities would reduce the need for upfront chemical intuition to solve a catalysis problem. However, a major limitation to this vision today is determining the activation energies of the reactions; data points on activation energies are limited to a few thousand at best. It is possible that more universal ML models, trained further out of equilibrium and near transition state structures using expanded data sets, will help remove this limitation in the future.

A second opportunity for applying ML is to increase accuracy in traditional workflows. As described previously, standard computational screening workflows often skip explicit DFT calculations in favor of simplifications such as scaling relationships to model adsorption energies and BEP relationships to model reaction kinetics. Unlike DFT calculations, these simplifications enforce rigid relationships between simple descriptors and more complex properties, making it difficult to identify the subset of materials that do not follow the usual trends. Indeed, materials whose behavior is not adequately described by fitted descriptor relationships may often be the most interesting for catalysis. Machine learning models may make it possible to remove the use of scaling relationships altogether, replacing them with ML predictions that better model the complex behavior of catalyst materials.

A third opportunity is to extend the complexity and depth of what is calculated. Typically, many aspects of catalytic materials that are possible to calculate are nevertheless skipped due to computational cost and researcher time. This includes calculating Wulff shapes and dominant surfaces (including the effect of high index surfaces), calculating performance as a function of surface coverage, and calculating side reactions with other species. Machine learning potentials offer the opportunity to calculate such behavior on a more regular basis and for more candidates.

Nevertheless, despite the promise afforded by deep learning and neural networks in computational chemistry, challenges remain. Universal ML models still struggle with nonuniform errors (depending on adsorbate and surface chemistry) as well as difficulty generalizing out of distribution[27]. Improving the models will likely require even more training data as well as potential architectural changes to the models. Furthermore, model interpretation remains a challenge for state-of-the-art models. Despite limited interpretations now being possible[28], whole ML models today are largely "black box". Improvements to model interpretability may help

with hypothesis formation, uncovering physical insights that are transferable, understanding model limitations, and improving intuition.

## 5. Theory-Experiment integration: case studies

To illustrate the integration of theory, machine learning, and experiment, we describe advancements in identifying promising electrocatalysts for the removal of the harmful oxyanions in water[29]. In particular, we focus on nitrate pollution; resulting from agricultural and industrial sources, nitrate can cause serious environmental and health problems such as eutrophication and methemoglobinemia[30]. Liu *et al.* previously developed volcano activity plots for transition metal electrocatalysts toward electroreduction of nitrate based on simple thermodynamic descriptors, *e.g.*, binding energy of atomic oxygen and nitrogen[15]. Moreover, they have successfully validated the computational predictions on the activity and selectivity of platinum-ruthenium alloys (i.e., Pt<sub>x</sub>Ru<sub>y</sub>) for nitrate reduction[31], exhibiting 93 – 98% Faradaic efficiencies toward NH<sub>3</sub> production at 0.1 V vs. RHE. The calculated trend corresponded well with subsequent experimental results, showing the catalytic activity of Rh/C electrodes can be enhanced by Sn modification[32]. These studies highlight the kind of synergy that can exist between the experimental and computational results.

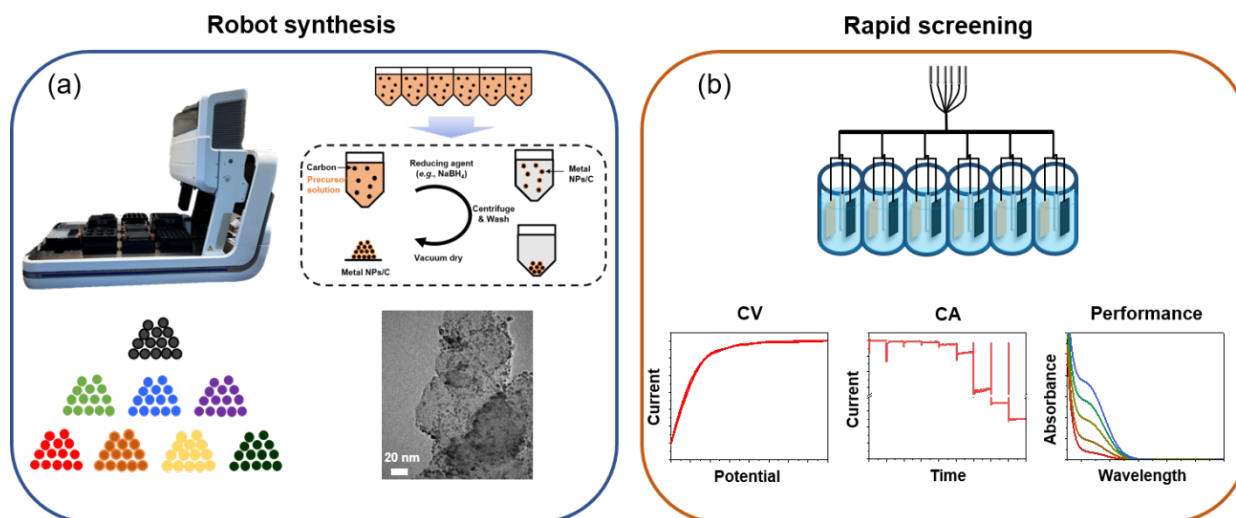


Figure 4. Integrated platform for electrocatalyst synthesis and electrochemical screening. (a) Synthesis of metal NPs on a liquid-handling robot via a chemical reduction method, and (b) screening of electrocatalytic activities via primary CV, secondary CA, and detailed characterization techniques. Adapted with permission from Ref.[33]. Copyright 2024 American Chemical Society.

Following these results and incorporating machine learning to search for a broader search, we identified 20 cost-effective materials out of 59,390 bimetallic alloys for nitrate reduction based on multiple criteria, including stability, nitrogen selectivity, activity, and commercial viability[14]. To rapidly validate the theoretical and ML calculations, we have developed an integrated platform[33] for the synthesis and characterization of electrocatalysts toward electro-reduction of oxyanions (**Figure 4**). For the synthesis of metal nanoparticles (NPs), a chemical reduction reaction was adopted on a versatile liquid handling robot. This robot can handle 6 to 96 reactions at temperatures up to 95 °C with shaking speed up to 2500 rpm. Following synthesis, the electrochemical screening platform comprises two steps: (1) initial cyclic voltammetry (CV)

screening and (2) subsequent chronoamperometry (CA) screening, with test durations ranging from minutes to hours.

Using this platform, we successfully synthesized a number of metal NPs including Rh, In, Pd, Ni, Ag, and Cu[33]. Electrochemical screenings indicate that Cu exhibits the highest electrocatalytic activity for nitrate reduction (71.7% nitrate removal at -1 V) among the four non-noble metals. This trend in the electrocatalytic activity aligns well with the DFT calculations. However, we also observe the discrepancies between the experimental and computational results for the activities of Ag and Ag-Ni (1:3), possibly due to the assumptions made during calculations regarding surface structure and variations in experimental conditions affecting the reaction kinetics and pathways. Thus, while the screening can identify promising directions for further study, they cannot substitute for experimental synthesis and characterization. Even with rapid advancements in theoretical capability and ML, it is important to keep in mind that such screening methods are still intended as a complement to experimental exploration rather than as a fully virtual screening protocol.

## 6. Conclusion

The integration of theoretical calculations, machine learning, and experimentation offers a powerful approach to develop practical, selective, and stable electrocatalysts. By combining machine learning and DFT calculations, researchers can potentially reduce the time needed to perform screening while progressively increasing the fidelity of simulated results. In particular, machine learning holds the promise of improving the mapping reaction pathways, of finding materials that break traditional scaling relationships, and of modeling more complex materials behavior. However, it is important to keep in mind that electrocatalytic processes are extremely

complex and cannot be entirely modeled. Local environmental conditions such as pH, dissolved oxygen/hydrogen concentrations, and temperature at the electrode surface could deviate significantly from the bulk conditions at applied potentials[34, 35], and it is challenging to measure these parameters and model them over extended time scales through computations. Nevertheless, historical successes and new developments in ML offer the possibility of continued extensions of what is possible in virtual catalyst design.

## 7. Acknowledgements

This work was supported by the National Alliance for Water Innovation (NAWI), funded by the U.S. Department of Energy, Energy Efficiency and Renewable Energy Office, Advanced Manufacturing Office under Funding Opportunity Announcement DE-FOA-0001905 under Award Number DE-AC02-05CH11231, and funded in part through contract No. 4600014330 with the California State Department of Water Resources to Lawrence Berkeley National Laboratory.

## 8. References

\*[1] K.C. Zuo, S. Garcia-Segura, G.A. Cerrón-Calle, F.Y. Chen, X.Y. Tian, X.X. Wang, X.C. Huang, H.T. Wang, P.J.J. Alvarez, J. Lou, M. Elimelech, Q.L. Li, Electrified water treatment: fundamentals and roles of electrode materials, *Nature Reviews Materials* 8 (2023) 472-490. <https://doi.org/10.1038/s41578-023-00564-y>.

**This work thoroughly reviewed the principles of electrified water treatment processes and the key role of electrode materials.**

[2] D.B. Gingerich, E. Grol, M.S. Mauter, Fundamental challenges and engineering opportunities in flue gas desulfurization wastewater treatment at coal fired power plants, *Environmental Science-Water Research & Technology* 4 (2018) 909-925. <https://doi.org/10.1039/c8ew00264a>.

[3] M. Soliman, F. Eljack, M.-K. Kazi, F. Almomani, E. Ahmed, Z. El Jack, Treatment technologies for cooling water blowdown: A critical review, *Sustainability* 14 (2021) 376. <https://doi.org/10.3390/su14010376>.

[4] C. Simoes, M. Saakes, D. Brilman, Toward Redox-Free Reverse Electrodialysis with Carbon-Based Slurry Electrodes, *Ind Eng Chem Res* 62 (2023) 1665-1675. <https://doi.org/10.1021/acs.iecr.2c03567>.

[5] L. Gurreri, A. Tamburini, A. Cipollina, G. Micale, Electrodialysis Applications in Wastewater Treatment for Environmental Protection and Resources Recovery: A Systematic Review on Progress and Perspectives, *Membranes (Basel)* 10 (2020) 146.

<https://doi.org/10.3390/membranes10070146>.

[6] WaterTectonics, WaterTecTonics Electrocoagulation,

<https://www.watertectonics.com/approach/packaged-systems/electrocoagulation/2024> (accessed 2024-04-16).

[7] M. Mirabediny, J. Sun, T.T. Yu, B. Akermark, B. Das, N. Kumar, Effective PFAS degradation by electrochemical oxidation methods-recent progress and requirement, *Chemosphere* 321 (2023) 138109. <https://doi.org/10.1016/j.chemosphere.2023.138109>.

[8] M. Veciana, J. Braunig, A. Farhat, M.L. Pype, S. Freguia, G. Carvalho, J. Keller, P. Ledezma, Electrochemical oxidation processes for PFAS removal from contaminated water and wastewater: fundamentals, gaps and opportunities towards practical implementation, *J Hazard Mater* 434 (2022) 128886. <https://doi.org/10.1016/j.jhazmat.2022.128886>.

[9] S.Q. Zou, M.S. Mauter, Direct Electrochemical Pathways for Selenium Reduction in Aqueous Solutions, *Acs Sustainable Chemistry & Engineering* 9 (2021) 2027-2036.

<https://doi.org/10.1021/acssuschemeng.0c06585>.

\*[10] F.Q. Liu, S.T. Jiang, S.J. You, Y.B. Liu, Recent advances in electrochemical decontamination of perfluorinated compounds from water: a review, *Frontiers of Environmental Science & Engineering* 17 (2023) 18. <https://doi.org/10.1007/s11783-023-1618-z>.

**This work comprehensively summarizes the current state of knowledge on electrochemical PFAS treatment, using electro-oxidation, electro-adsorption, electro-coagulation, and coupled technologies.**

[11] N. Singh, B.R. Goldsmith, Role of Electrocatalysis in the Remediation of Water Pollutants, *Acs Catalysis* 10 (2020) 3365-3371. <https://doi.org/10.1021/acscatal.9b04167>.

[12] J. Greeley, M. Mavrikakis, Alloy catalysts designed from first principles, *Nat Mater* 3 (2004) 810-5. <https://doi.org/10.1038/nmat1223>.

\*[13] J. Greeley, Theoretical Heterogeneous Catalysis: Scaling Relationships and Computational Catalyst Design, *Annu Rev Chem Biomol Eng* 7 (2016) 605-35. <https://doi.org/10.1146/annurev-chembioeng-080615-034413>.

**This work provided a review on the scaling relationship using in the surface and catalysis DFT research. The review went through various structures from pure metal to metal oxide surfaces, for various classes of molecules.**

\*[14] R. Tran, D. Wang, R. Kingsbury, A. Palizhati, K.A. Persson, A. Jain, Z.W. Ulissi, Screening of bimetallic electrocatalysts for water purification with machine learning, *J Chem Phys* 157 (2022) 074102. <https://doi.org/10.1063/5.0092948>.

**The authors incorporated machine learning and computation based screening strategy to successfully identify 20 cost-effective materials out of 59390 bimetallic electrocatalysts for NO<sub>3</sub><sup>-</sup> conversions.**

\*[15] J.X. Liu, D. Richards, N. Singh, B.R. Goldsmith, Activity and Selectivity Trends in Electrocatalytic Nitrate Reduction on Transition Metals, *Acs Catalysis* 9 (2019) 7052-7064. <https://doi.org/10.1021/acscatal.9b02179>.

**This work demonstrated a comprehensive work in using DFT calculations to establish energy scaling relationship and BEP relation, for the purpose of high throughput predicting catalyst materials and performed further experimental validation.**

[16] A. Jain, Y. Shin, K.A. Persson, Computational predictions of energy materials using density functional theory, *Nature Reviews Materials* 1 (2016) 1-13.

<https://doi.org/10.1038/natrevmats.2015.4>.

[17] H. Shi, T. Zhou, Computational design of heterogeneous catalysts and gas separation materials for advanced chemical processing, *Frontiers of Chemical Science and Engineering* 15 (2021) 49-59. <https://doi.org/10.1007/s11705-020-1959-0>.

[18] H. Ooka, J. Huang, K.S. Exner, The Sabatier Principle in Electrocatalysis: Basics, Limitations, and Extensions, *Frontiers in Energy Research* 9 (2021) 654460.

<https://doi.org/10.3389/fenrg.2021.654460>.

[19] A.H. Motagamwala, J.A. Dumesic, Microkinetic Modeling: A Tool for Rational Catalyst Design, *Chem Rev* 121 (2021) 1049-1076. <https://doi.org/10.1021/acs.chemrev.0c00394>.

[20] A. Jain, K.A. Persson, G. Ceder, Research Update: The materials genome initiative: Data sharing and the impact of collaborative ab initio databases, *APL Materials* 4 (2016)

<https://doi.org/10.1063/1.4944683>.

[21] J. Lan, A. Palizhati, M. Shuaibi, B.M. Wood, B. Wander, A. Das, M. Uyttendaele, C.L. Zitnick, Z.W. Ulissi, AdsorbML: a leap in efficiency for adsorption energy calculations using generalizable machine learning potentials, *Npj Computational Materials* 9 (2023) 172.

<https://doi.org/10.1038/s41524-023-01121-5>.

[22] H. Mai, T.C. Le, D. Chen, D.A. Winkler, R.A. Caruso, Machine Learning for Electrocatalyst and Photocatalyst Design and Discovery, *Chem Rev* 122 (2022) 13478-13515.

<https://doi.org/10.1021/acs.chemrev.2c00061>.

\*[23] T. Xie, J.C. Grossman, Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties, *Phys Rev Lett* 120 (2018) 145301.

<https://doi.org/10.1103/PhysRevLett.120.145301>.

**This work introduced the use of crystal graph neural networks for predicting materials properties; this representation underpins today's state-of-the-art ML models.**

\*[24] R. Tran, J. Lan, M. Shuaibi, B.M. Wood, S. Goyal, A. Das, J. Heras-Domingo, A. Kolluru, A. Rizvi, N. Shoghi, A. Sriram, F. Therrien, J. Abed, O. Voznyy, E.H. Sargent, Z. Ulissi, C.L. Zitnick, The Open Catalyst 2022 (OC22) Dataset and Challenges for Oxide Electrocatalysts, *Acc Catalysis* 13 (2023) 3066-3084. <https://doi.org/10.1021/acscatal.2c05426>.

**This work describes the most comprehensive training set available today for machine learning potential development.**

[25] F. Göttl, Three Grand Challenges for the Computational Design of Heterogeneous Catalysts, *The Journal of Physical Chemistry C* 126 (2022) 3305-3313.

<https://doi.org/10.1021/acs.jpcc.1c10291>.

[26] J.T. Margraf, H. Jung, C. Scheurer, K. Reuter, Exploring catalytic reaction networks with machine learning, *Nature Catalysis* 6 (2023) 112-121. <https://doi.org/10.1038/s41929-022-00896-y>.

[27] A. Kolluru, M. Shuaibi, A. Palizhati, N. Shoghi, A. Das, B. Wood, C.L. Zitnick, J.R. Kitchin, Z.W. Ulissi, Open challenges in developing generalizable large-scale machine-learning models for catalyst discovery, *ACS Catalysis* 12 (2022) 8572-8581.

<https://doi.org/10.1021/acscatal.2c02291>.

[28] J.A. Esterhuizen, B.R. Goldsmith, S. Linic, Interpretable machine learning for knowledge generation in heterogeneous catalysis, *Nature Catalysis* 5 (2022) 175-184.

<https://doi.org/10.1038/s41929-022-00744-z>.



- [29] Z.X. Wang, D. Richards, N. Singh, Recent discoveries in the reaction mechanism of heterogeneous electrocatalytic nitrate reduction, *Catalysis Science & Technology* 11 (2021) 705-725. <https://doi.org/10.1039/d0cy02025g>.
- [30] X. Zhang, Y. Zhang, P. Shi, Z. Bi, Z. Shan, L. Ren, The deep challenge of nitrate pollution in river water of China, *Sci Total Environ* 770 (2021) 144674. <https://doi.org/10.1016/j.scitotenv.2020.144674>.
- [31] Z.X. Wang, S.D. Young, B.R. Goldsmith, N. Singh, Increasing electrocatalytic nitrate reduction activity by controlling adsorption through PtRu alloying, *Journal of Catalysis* 395 (2021) 143-154. <https://doi.org/10.1016/j.jcat.2020.12.031>.
- [32] W. Siritwatcharapiboon, Y. Kwon, J. Yang, R.L. Chantry, Z. Li, S.L. Horswell, M.T.M. Koper, Promotion Effects of Sn on the Electrocatalytic Reduction of Nitrate at Rh Nanoparticles, *Chemelectrochem* 1 (2014) 172-179. <https://doi.org/10.1002/celec.201300135>.
- [33] S.C. Ma, J. Qian, D. Wang, E.M. Chan, A. Jain, W. Tong, Integrated Electrocatalyst Platform for Electroreduction of Oxyanions in Water, *Acs Applied Energy Materials* (2024) <https://doi.org/10.1021/acsaem.4c00631>.
- [34] J.C. Bui, C. Kim, A.Z. Weber, A.T. Bell, Dynamic Boundary Layer Simulation of Pulsed CO Electrolysis on a Copper Catalyst, *Acs Energy Letters* 6 (2021) 1181-1188. <https://doi.org/10.1021/acsenerylett.1c00364>.
- [35] X. Wang, S. Im, B. Jung, J. Wu, A. Iddya, Q.A. Javier, M. Xiao, S. Ma, S. Lu, B. Jaewon, J. Zhang, Z.J. Ren, C.T. Maravelias, E.M.V. Hoek, D. Jassby, Simple and Low-Cost Electroactive Membranes for Ammonia Recovery, *Environ Sci Technol* 57 (2023) 9405-9415. <https://doi.org/10.1021/acs.est.3c01470>.