

Redefining the Stability of Water Oxidation Electrocatalysts: Insights from Materials Databases and Machine Learning

*Raul A. Marquez,[†] Erin Elizabeth Oefelein,[‡] Thuy Vy Le,[†] Kenta Kawashima,[†] Lettie A. Smith,[†]
and C. Buddie Mullins^{†,§,||,⊥,▽,*}*

[†] Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712, United States.

[‡] Department of Statistics and Data Sciences, The University of Texas at Austin, Austin, Texas 78712,
United States.

[§] McKetta Department of Chemical Engineering, The University of Texas at Austin, Austin, Texas
78712, United States.

^{||} Texas Materials Institute, The University of Texas at Austin, Austin, Texas 78712, United States.

[⊥] Center for Electrochemistry, The University of Texas at Austin, Austin, Texas 78712, United States.

[▽] H2@UT, The University of Texas at Austin, Austin, Texas 78712, United States.

* Corresponding author: mullins@che.utexas.edu

Number of pages: 17

Number of figures: 18

Additional Information on Machine Learning Models

The machine learning models and libraries discussed in this Perspective are available at the following [GitHub link](#). These resources are intended for community use, allowing researchers to analyze their datasets or categorize materials into specific groups depending on their needs. We advise compiling materials databases from well-controlled experimental conditions and including a comprehensive set of features to improve result accuracy. Additionally, we encourage exploring other tools and models to leverage machine learning in electrocatalysis.

Supporting note: Constructing decision trees using gradient boosting.

The gradient boosting technique, which involves a series of decision trees, each acting as a weak learner, builds each successive tree on the deficiencies of its predecessors to enhance the overall model. However, in our study, the XGBoost model produced only one tree. Three main factors explain this outcome:

First, after extensive testing with various machine learning models, such as Linear Regression and Support Vector Machines using a Linear Kernel, we discovered that the independent and dependent variables in our dataset exhibited a simple linear relationship. In such scenarios, adding more trees would not improve model performance and might lead to overfitting.

Second, we implemented early stopping to prevent overfitting and ensure the model's effectiveness on unseen data. This technique halts the training process when the model's performance on a validation set meets a predefined threshold. In our case, this criterion was satisfied after just one tree, resulting in an XGBoost model composed of this single tree.

Third, we carefully screened our transition metal (TM) X-ide datasets to ensure a fair comparison under homogeneous testing conditions. We removed certain entries because they used different electrolytes (e.g., NaOH) or due to incomplete data provided in the study. Consequently, our datasets became relatively small after this screening. More complex models with multiple trees could overfit and underperform on new, unseen data. Thus, employing a single-tree model promotes simplicity and reduces the risk of overfitting, making it preferable when working with small datasets.

We utilized the *plot_tree* function from the XGBoost library to visualize the resulting trees, shown in **Figures 5a, S13, and S14**, for TM selenides, phosphides, and sulfides, respectively.

Supporting Figures, Tables, and Schemes

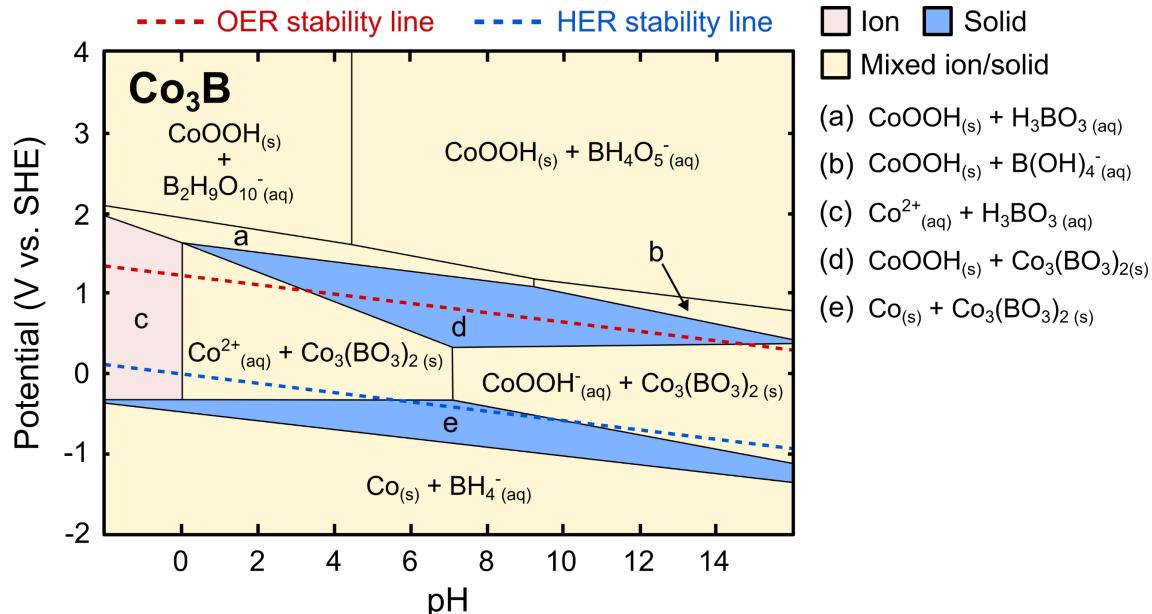


Figure S1. Pourbaix diagram for Co_3B . The plot was adapted from the diagram generated using the Materials Project.^{39–42} Ion concentration: $10^{-5} \text{ mol}\cdot\text{L}^{-1}$.

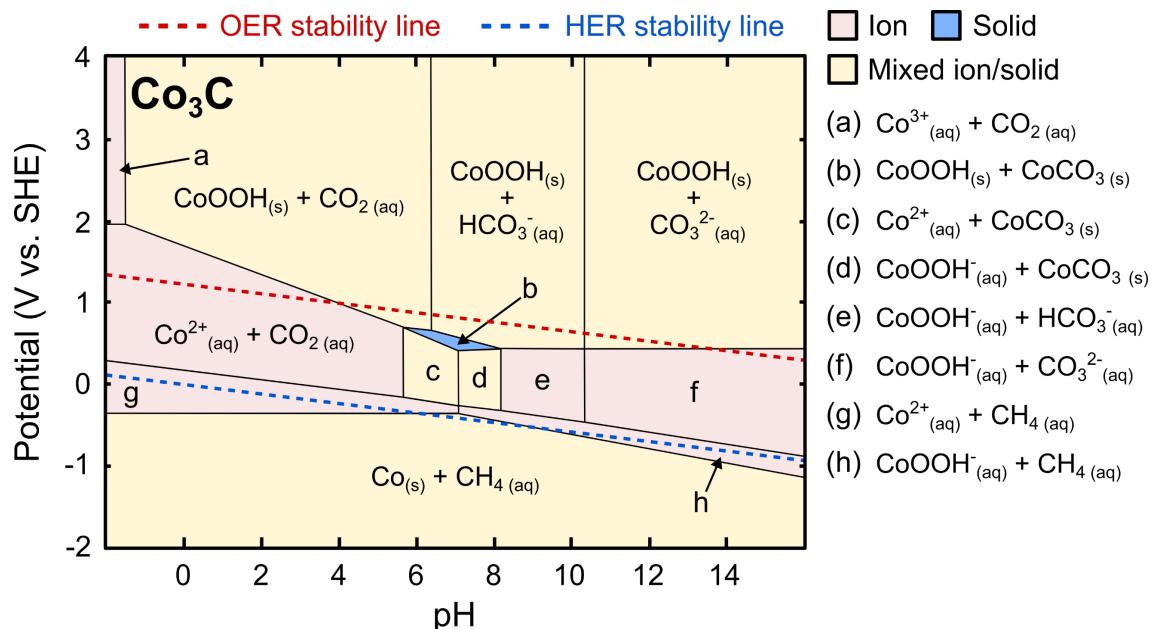


Figure S2. Pourbaix diagram for Co_3C . The plot was adapted from the diagram generated using the Materials Project.^{39–42} Ion concentration: $10^{-5} \text{ mol}\cdot\text{L}^{-1}$.

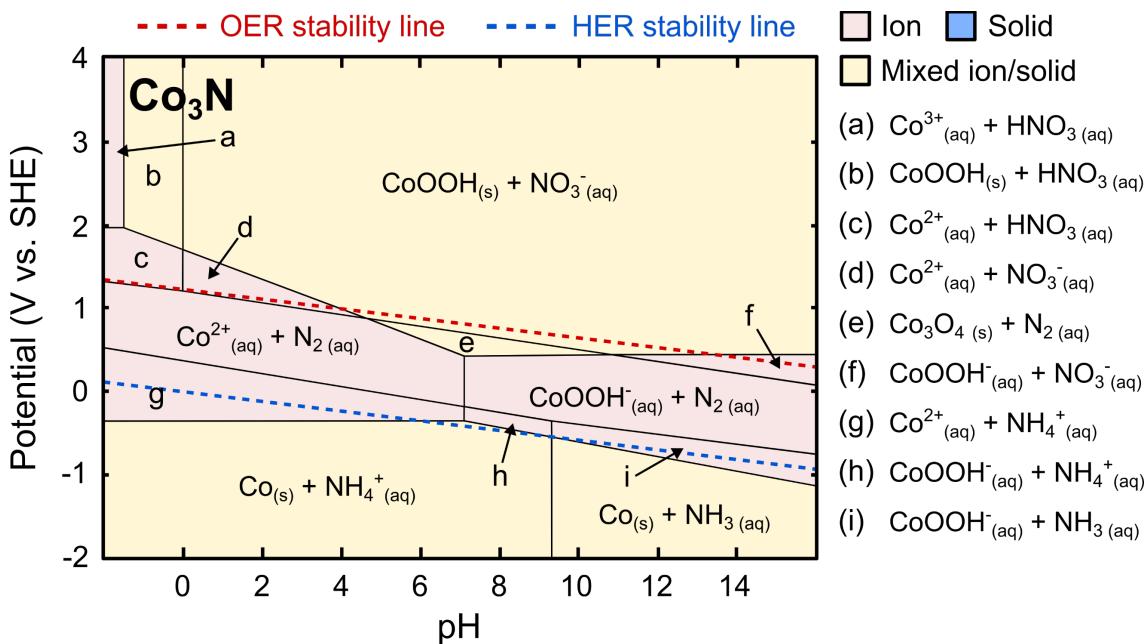


Figure S3. Pourbaix diagram for Co_3N . The plot was adapted from the diagram generated using the Materials Project.^{39–42} Ion concentration: $10^{-5} \text{ mol}\cdot\text{L}^{-1}$.

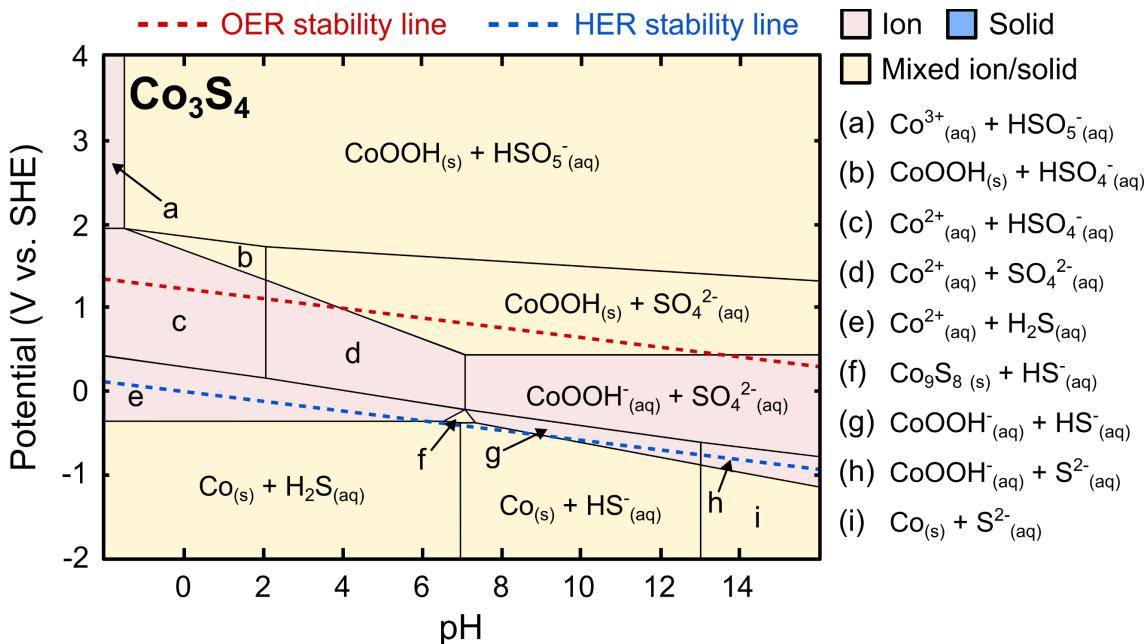


Figure S4. Pourbaix diagram for Co_3S_4 . The plot was adapted from the diagram generated using the Materials Project.^{39–42} Ion concentration: $10^{-5} \text{ mol}\cdot\text{L}^{-1}$.

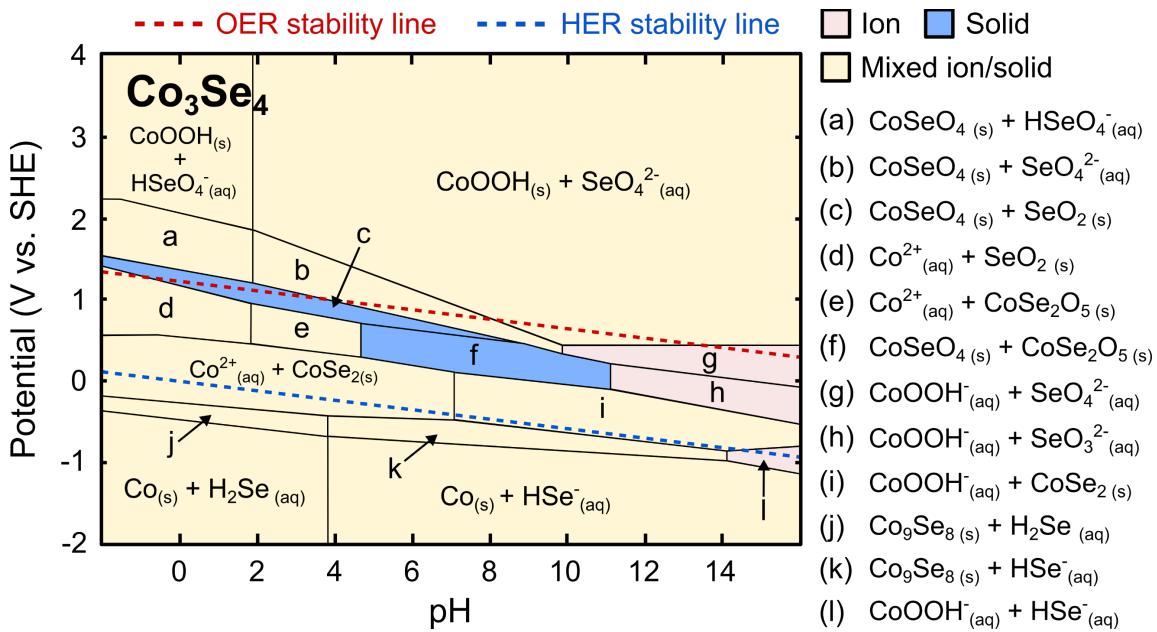


Figure S5. Pourbaix diagram for Co_3Se_4 . The plot was adapted from the diagram generated using the Materials Project.^{39–42} Ion concentration: $10^{-5} \text{ mol}\cdot\text{L}^{-1}$.

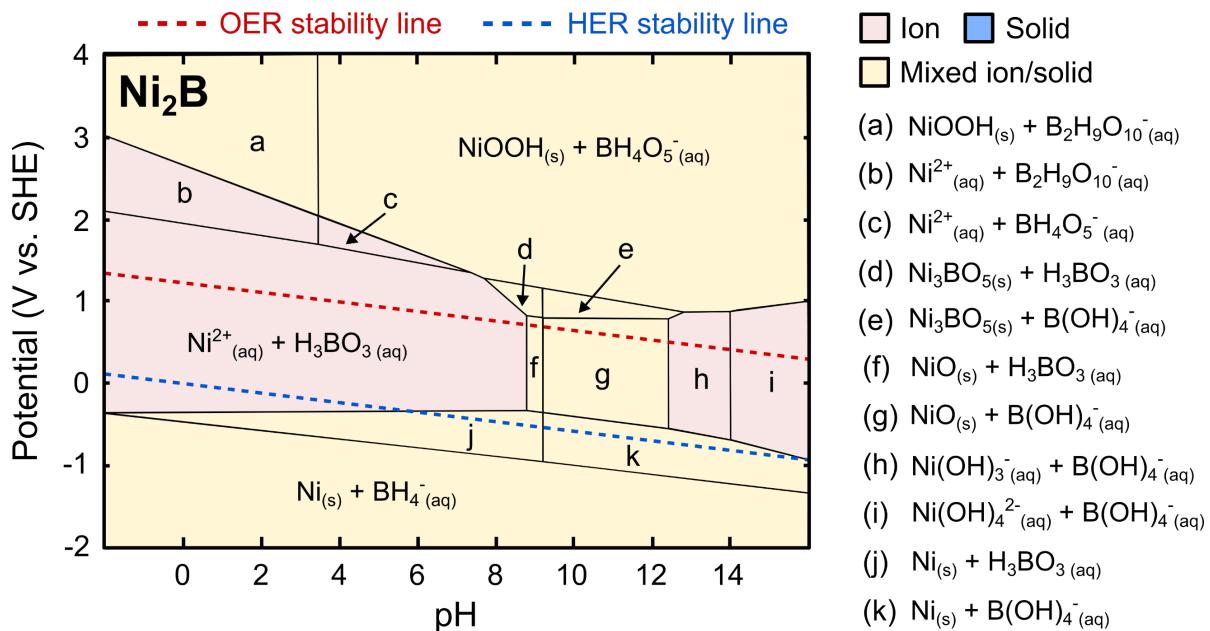


Figure S6. Pourbaix diagram for Ni_2B . The plot was adapted from the diagram generated using the Materials Project.^{39–42} Ion concentration: $10^{-5} \text{ mol}\cdot\text{L}^{-1}$.

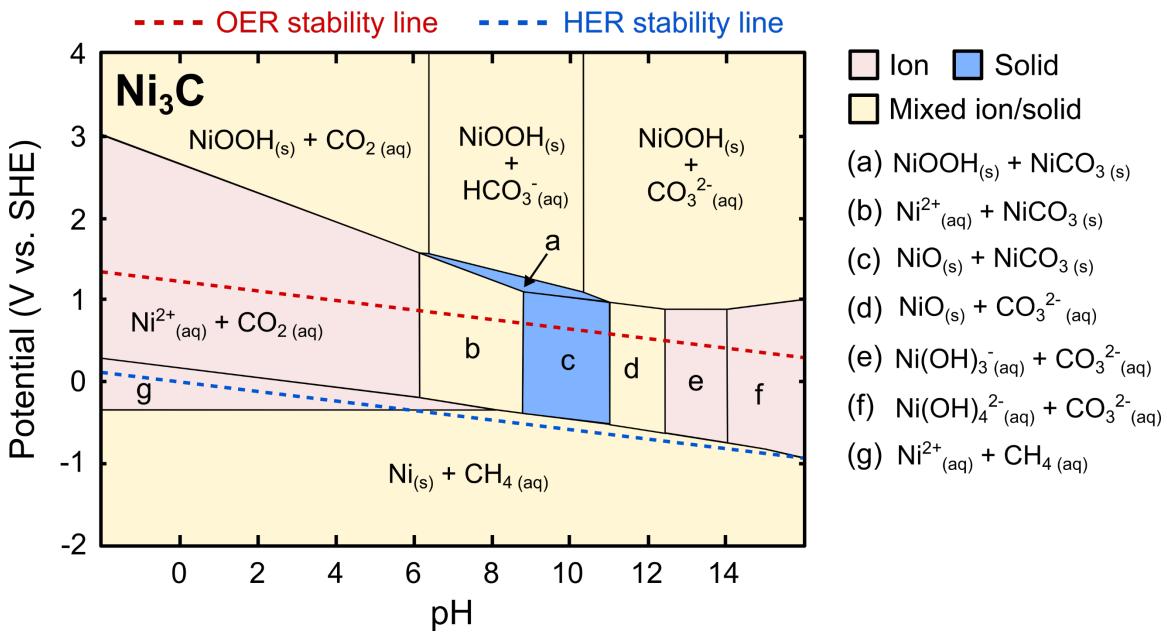


Figure S7. Pourbaix diagram for Ni₃C. The plot was adapted from the diagram generated using the Materials Project.^{39–42} Ion concentration: $10^{-5} \text{ mol}\cdot\text{L}^{-1}$.

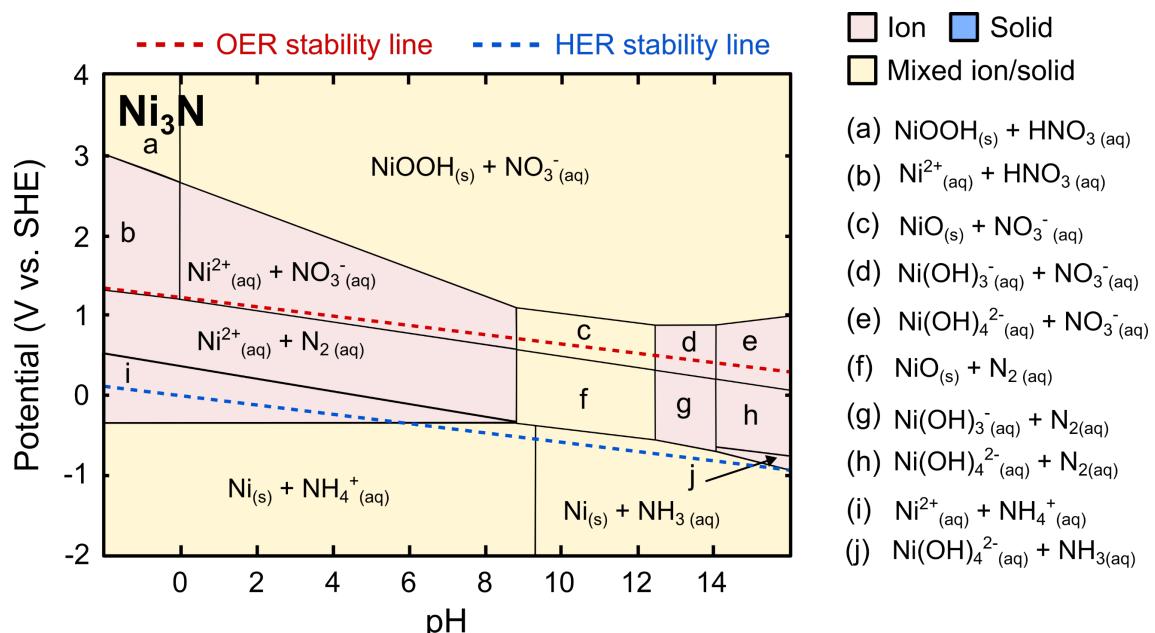


Figure S8. Pourbaix diagram for Ni₃N. The plot was adapted from the diagram generated using the Materials Project.^{39–42} Ion concentration: $10^{-5} \text{ mol}\cdot\text{L}^{-1}$.

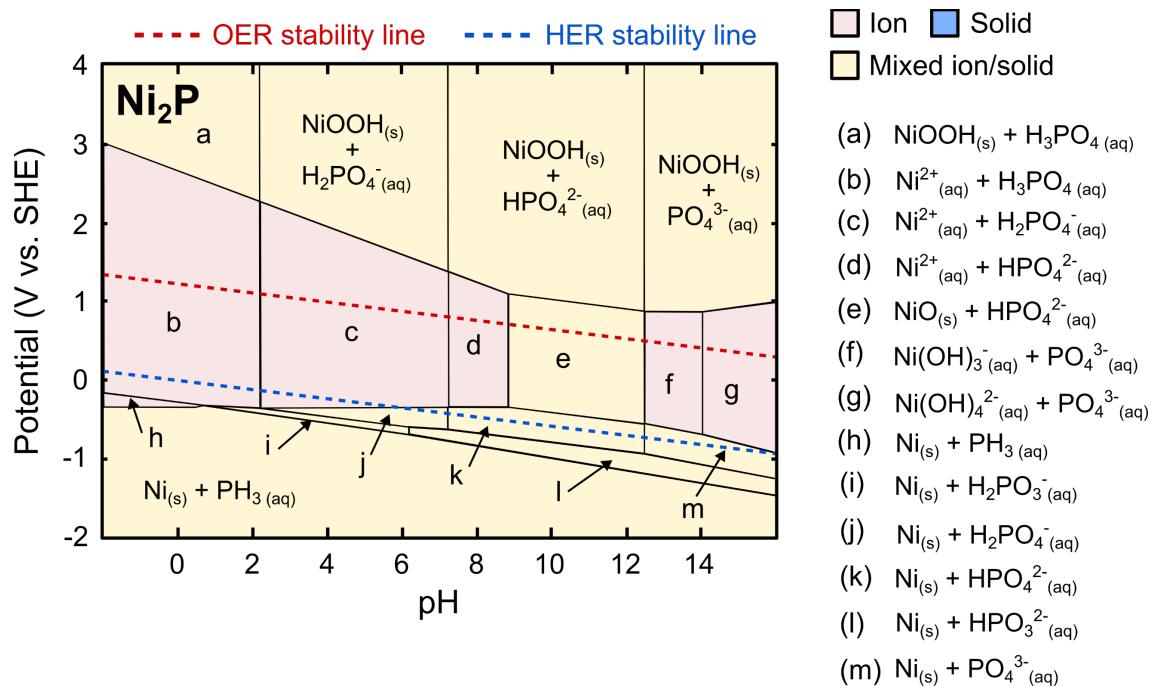


Figure S9. Pourbaix diagram for Ni₂P. The plot was adapted from the diagram generated using the Materials Project.^{39–42} Ion concentration: 10^{-5} mol·L⁻¹.

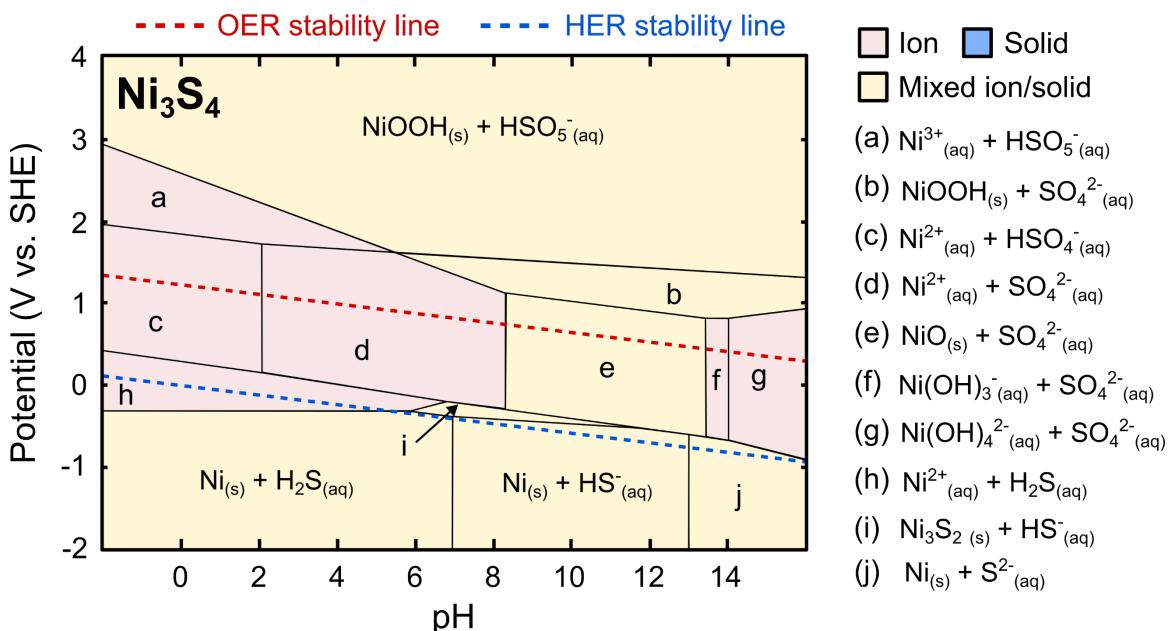


Figure S10. Pourbaix diagram for Ni₃S₄. The plot was adapted from the diagram generated using the Materials Project.^{39–42} Ion concentration: 10^{-5} mol·L⁻¹.

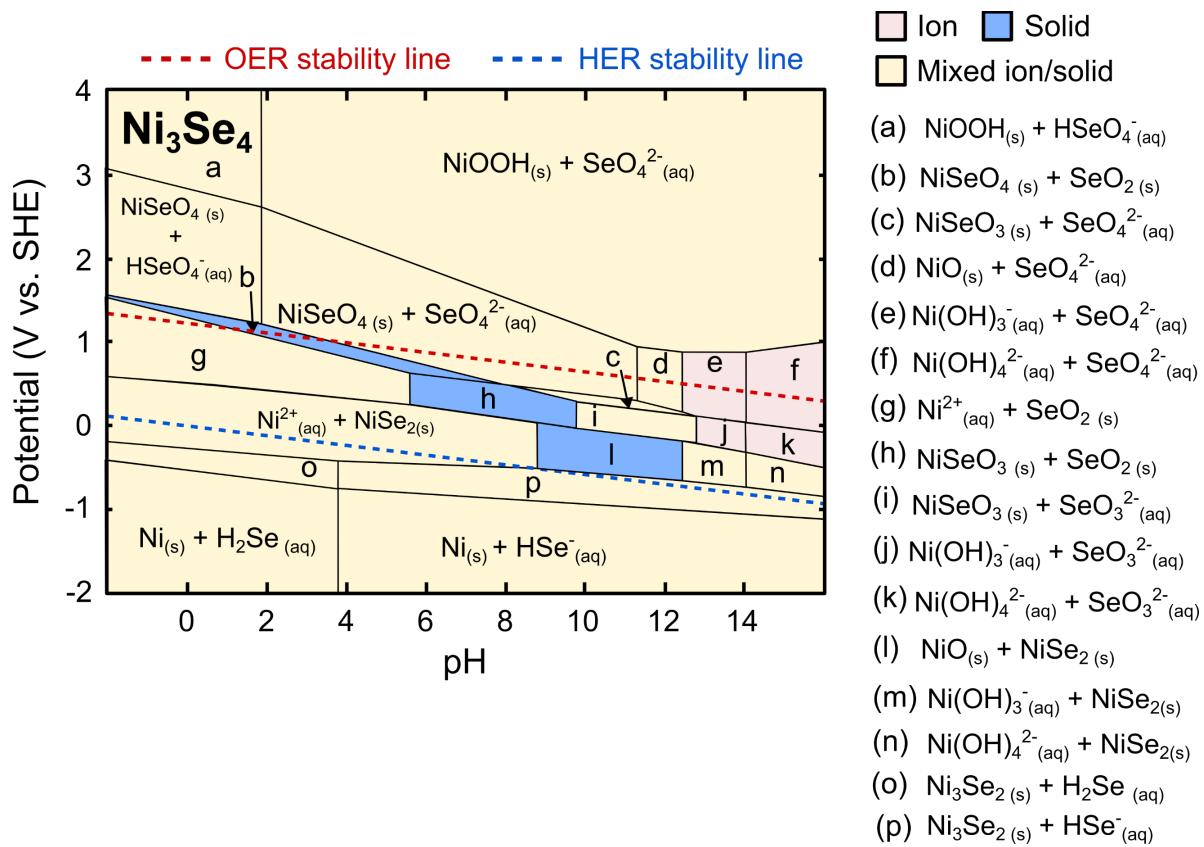


Figure S11. Pourbaix diagram for Ni_3Se_4 . The plot was adapted from the diagram generated using the Materials Project.^{39–42} Ion concentration: $10^{-5} \text{ mol}\cdot\text{L}^{-1}$.

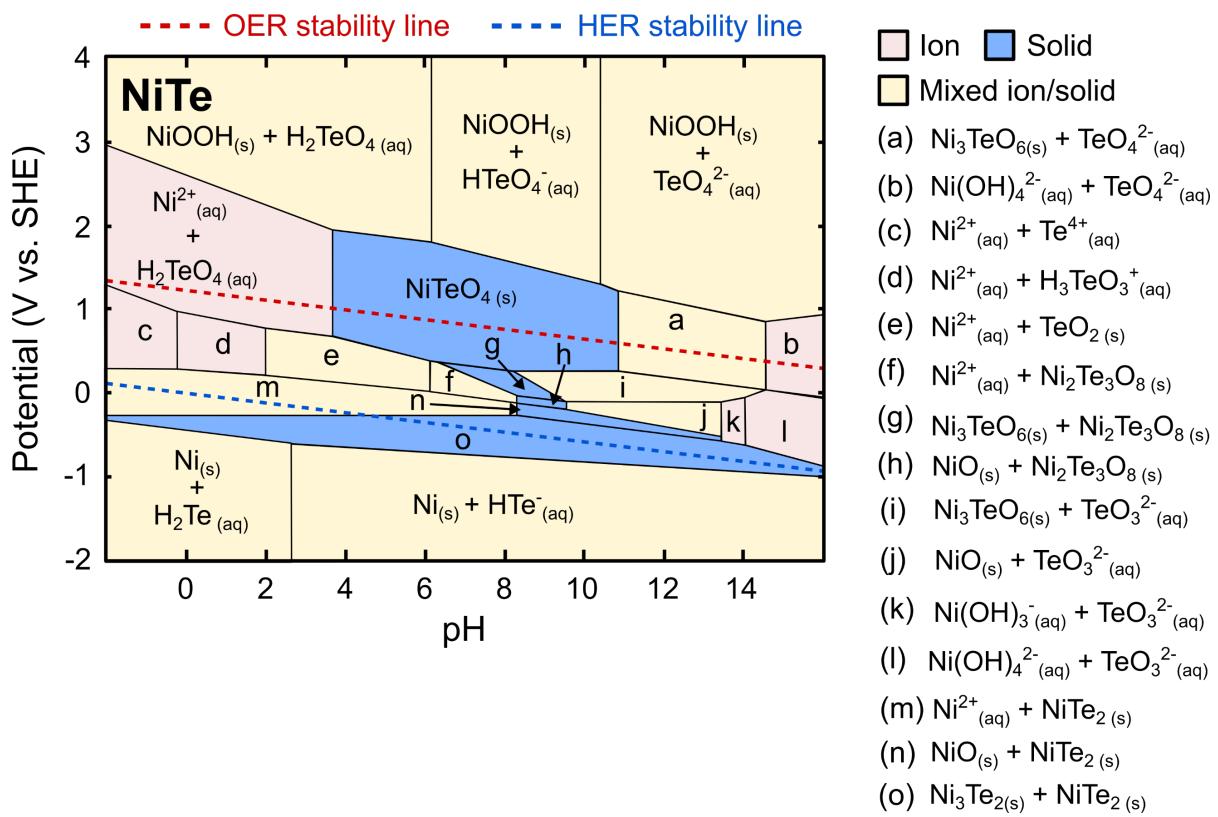


Figure S12. Pourbaix diagram for NiTe. The plot was adapted from the diagram generated using the Materials Project.³⁹⁻⁴² Ion concentration: $10^{-5} \text{ mol}\cdot\text{L}^{-1}$.

P

Transition metal phosphides

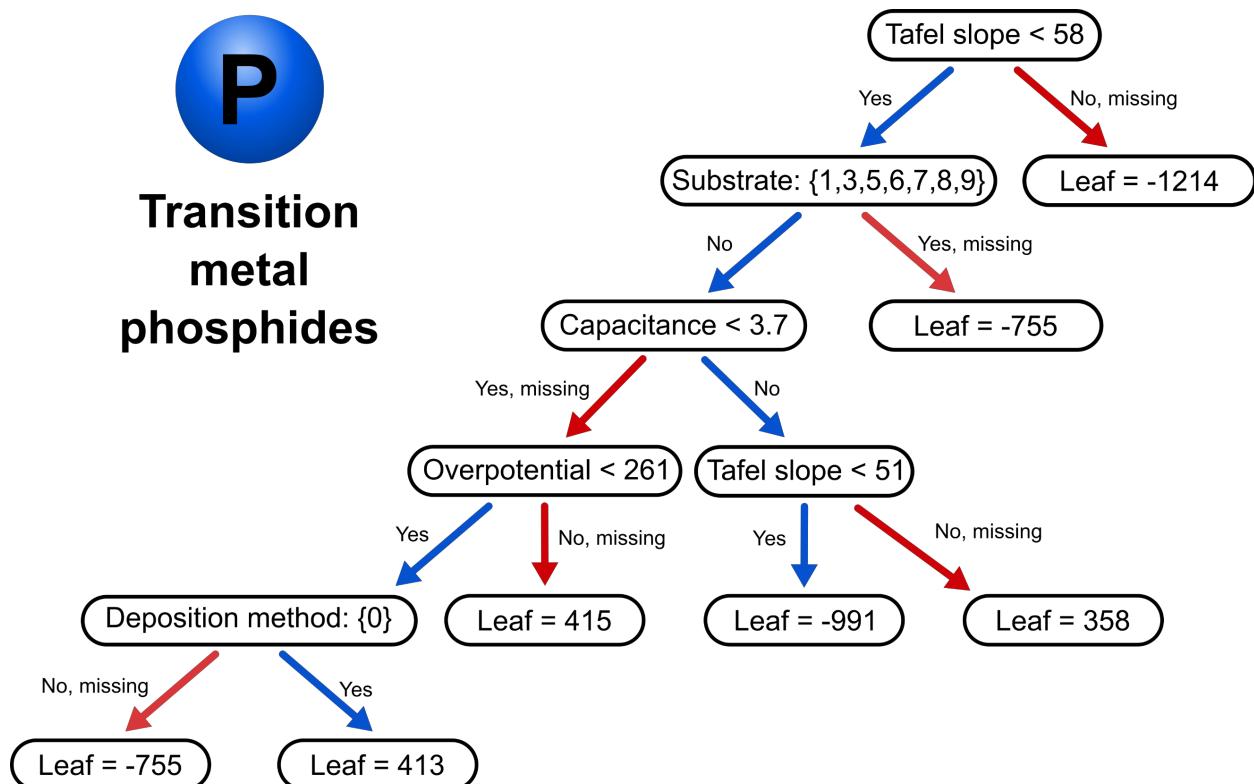


Figure S13. Decision tree for the TM phosphide dataset. A blue arrow denotes a catalyst node with identical features, while a red arrow indicates a precatalyst node. The leaves symbolize the purest nodes for each catalyst and precatalyst subgroup. A total of 126 data points, each representing a TM phosphide material from our compilation, were analyzed.

Supporting note: The following codes apply for the decision trees and force plots of the TM phosphide, sulfide, and selenide databases:

Substrate type: "Glassy carbon": 0, "carbon cloth": 1, "graphite": 2, "Au glass": 3, "carbon foam": 4, "Cu foil": 5, "FeNi foam": 6, "Ni foam": 7, "carbon fiber paper": 8, "Pt sheet": 9, "carbon paper": 10, "NiFe alloy": 11, "Fe foam": 12, "Ti mesh": 13, "Cu mesh": 14.

Deposition method: "Self-supported electrode": 1, "powder-coated electrode": 2.

Initial elemental content: "Element present": 0, "no element present": 1.



Transition metal sulfides

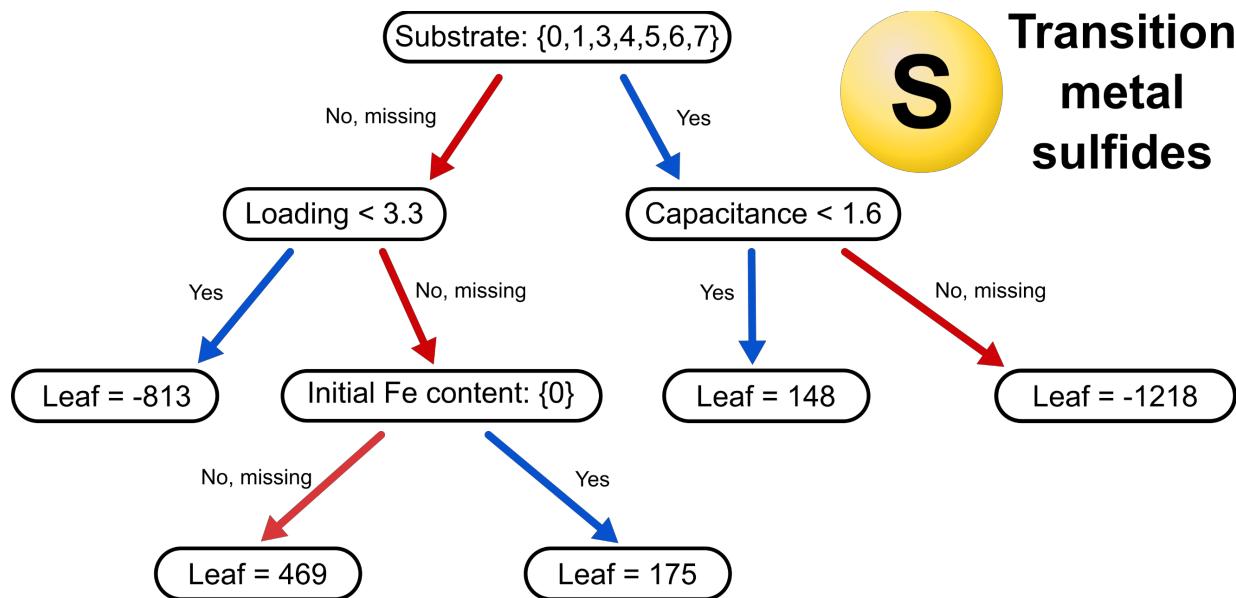


Figure S14. Decision tree for the TM sulfide dataset. A blue arrow denotes a catalyst node with identical features, while a red arrow indicates a precatalyst node. The leaves symbolize the purest nodes for each catalyst and precatalyst subgroup. A total of 104 data points, each representing a TM phosphide material from our compilation, were analyzed.

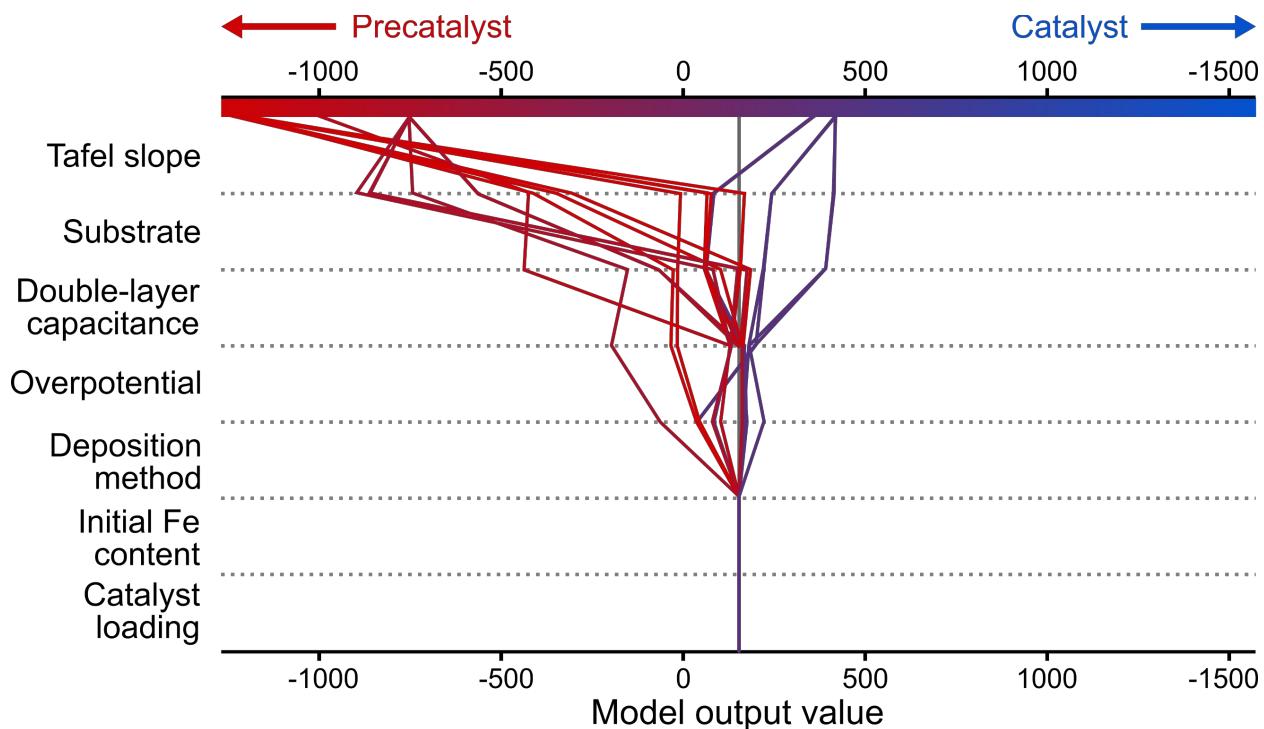


Figure S15. Decision plot for the TM phosphide dataset. Each colored line represents a single TM phosphide electrocatalyst from our compilation of the literature. Blue lines represent a catalyst classification, whereas red lines denote a precatalyst. The features are ranked in descending order based on their importance to the model, with the most influential feature at the top. A total of 126 data points, each representing a TM phosphide material from our compilation, were analyzed.

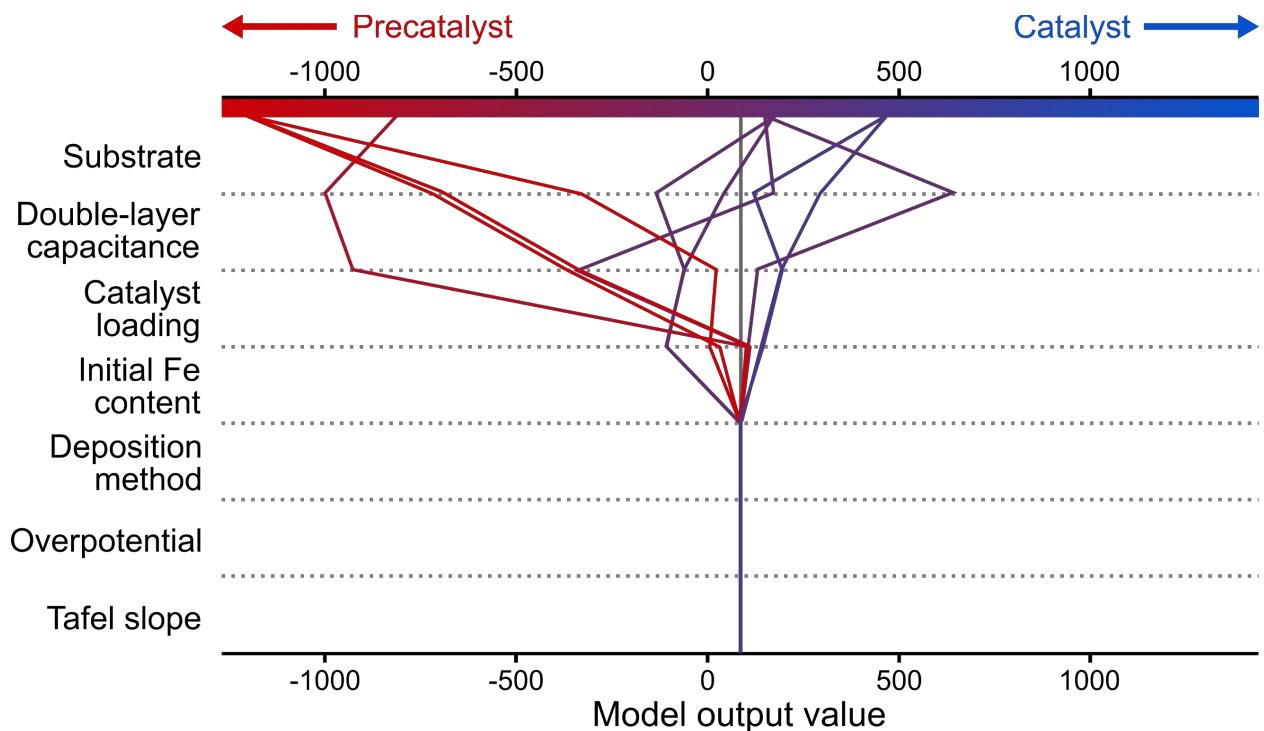


Figure S16. Decision plot for the TM sulfide dataset. Each colored line represents a single TM phosphide electrocatalyst from our compilation of the literature. Blue lines represent a catalyst classification, whereas red lines denote a precatalyst. The features are ranked in descending order based on their importance to the model, with the most influential feature at the top. A total of 126 data points, each representing a TM phosphide material from our compilation, were analyzed.

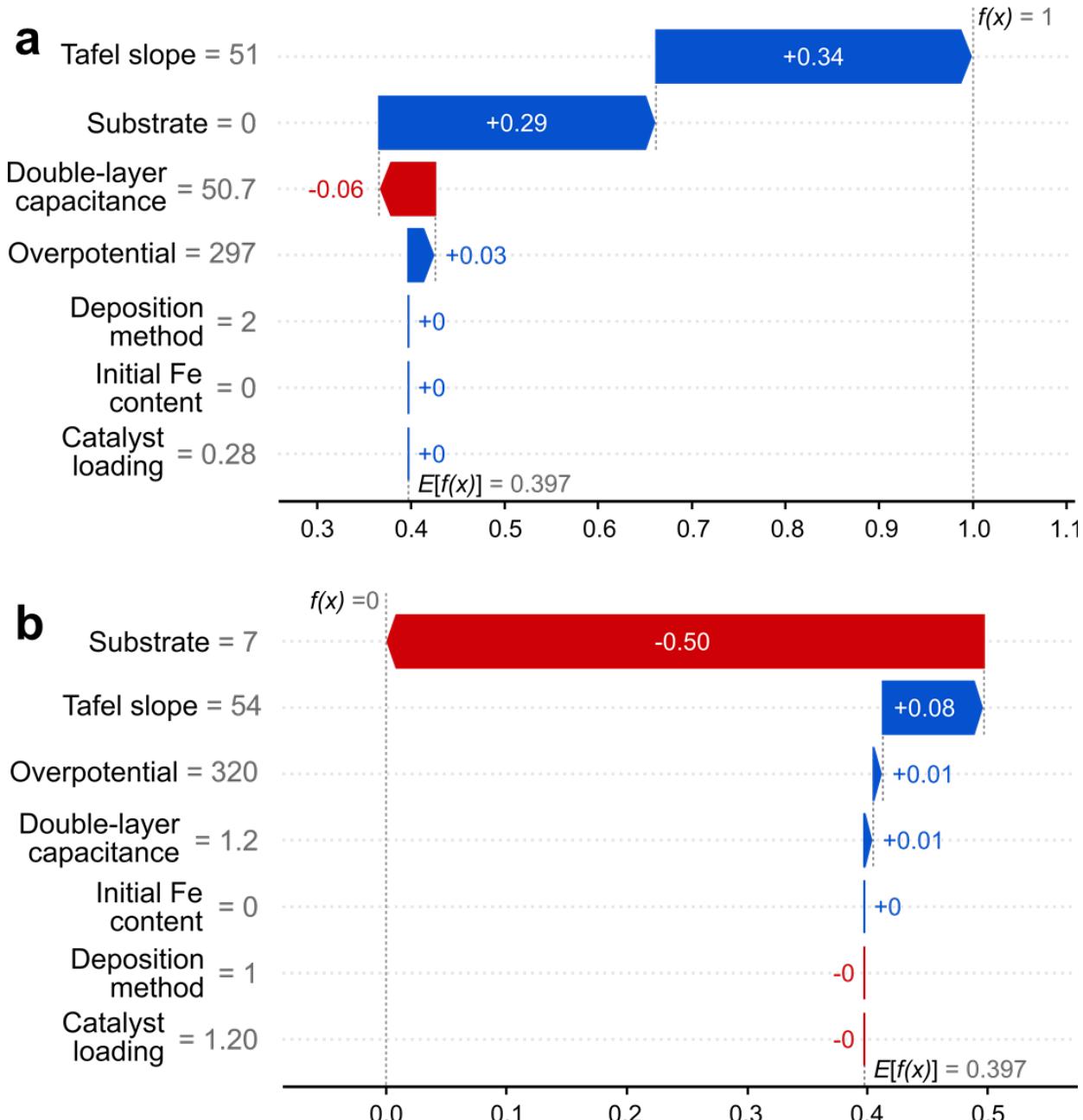


Figure S17. Force plots for a single (a) catalyst and (b) precatalyst from the TM phosphide dataset. The bars represent the contribution of specific features to the classification as either a catalyst (blue) or a precatalyst (red). These features are ranked in descending order of importance to the model, with the most influential feature at the top.

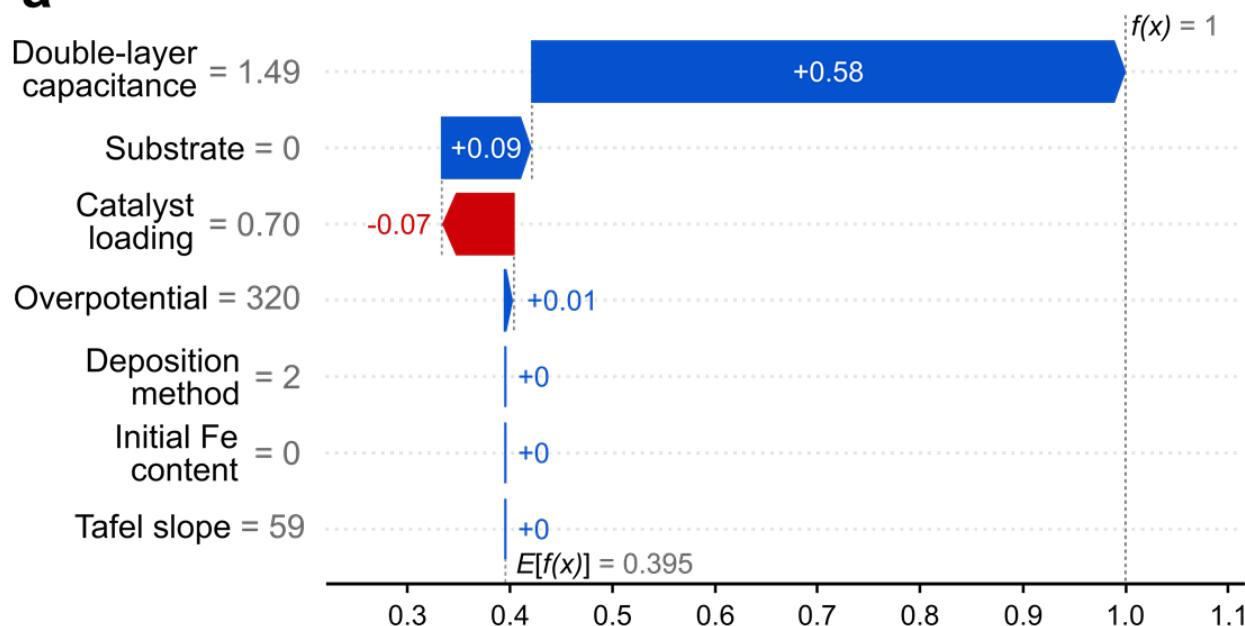
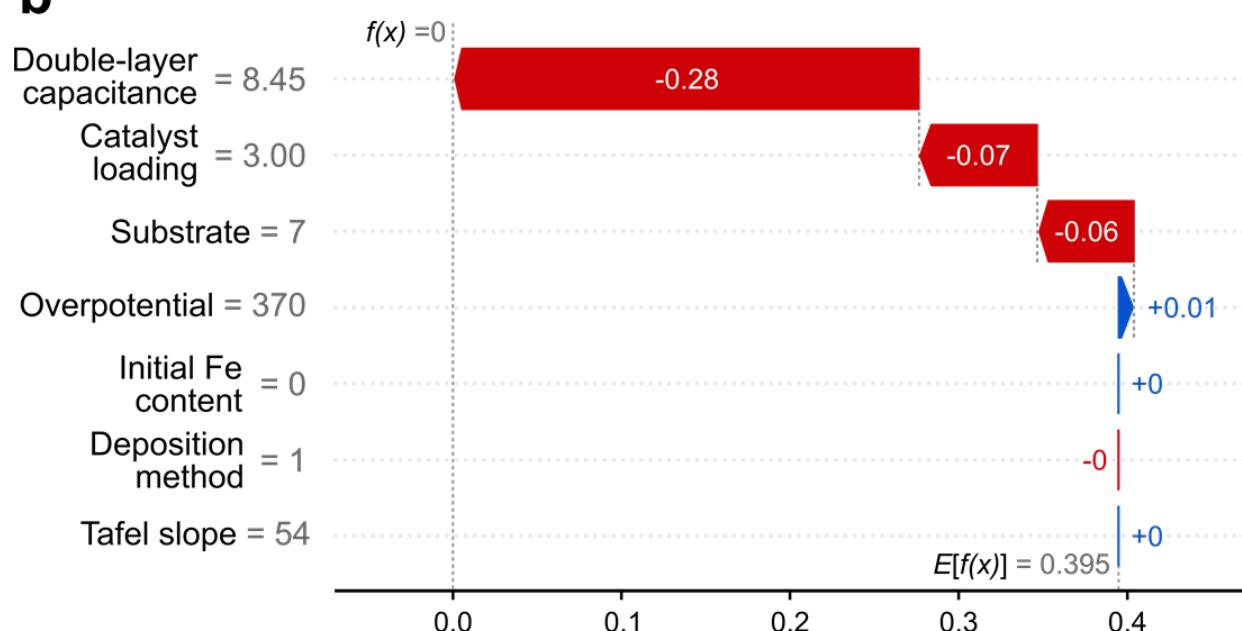
a**b**

Figure S18. Force plots for a single (a) catalyst and (b) precatalyst from the TM sulfide dataset. The bars represent the contribution of a specific feature towards the catalyst (blue) and precatalyst (red) classification. The features are ranked in descending order based on their importance to the model, with the most influential feature at the top.

References

- (1) Márquez, R. A.; Kawashima, K.; Son, Y. J.; Castelino, G.; Miller, N.; Smith, L. A.; Chukwuneke, C. E.; Mullins, C. B. Getting the Basics Right: Preparing Alkaline Electrolytes for Electrochemical Applications. *ACS Energy Lett.* **2023**, *8* (2), 1141–1146. <https://doi.org/10.1021/acsenergylett.2c02847>.
- (2) Hoang, T. T. H.; Gewirth, A. A. High Activity Oxygen Evolution Reaction Catalysts from Additive-Controlled Electrodeposited Ni and NiFe Films. *ACS Catal.* **2016**, *6* (2), 1159–1164. <https://doi.org/10.1021/acscatal.5b02365>.
- (3) Yoon, Y.; Yan, B.; Surendranath, Y. Suppressing Ion Transfer Enables Versatile Measurements of Electrochemical Surface Area for Intrinsic Activity Comparisons. *J. Am. Chem. Soc.* **2018**, *140* (7), 2397–2400. <https://doi.org/10.1021/jacs.7b10966>.
- (4) Morales, D. M.; Risch, M. Seven Steps to Reliable Cyclic Voltammetry Measurements for the Determination of Double Layer Capacitance. *J. Phys. Energy* **2021**, *3* (3), 034013. <https://doi.org/10.1088/2515-7655/abee33>.
- (5) Kawashima, K.; Márquez-Montes, R. A.; Li, H.; Shin, K.; Cao, C. L.; Vo, K. M.; Son, Y. J.; Wygant, B. R.; Chunangad, A.; Youn, D. H.; Henkelman, G.; Ramos-Sánchez, V. H.; Mullins, C. B. Electrochemical Behavior of a Ni₃N OER Precatalyst in Fe-Purified Alkaline Media: The Impact of Self-Oxidation and Fe Incorporation. *Mater. Adv.* **2021**, *2*, 2299–2309. <https://doi.org/10.1039/D1MA00130B>.
- (6) Márquez-Montes, R. A.; Collins-Martínez, V. H.; Pérez-Reyes, I.; Chávez-Flores, D.; Graeve, O. A.; Ramos-Sánchez, V. H. Electrochemical Engineering Assessment of a Novel 3D-Printed Filter-Press Electrochemical Reactor for Multipurpose Laboratory Applications. *ACS Sustain. Chem. Eng.* **2020**, *8* (9), 3896–3905. <https://doi.org/10.1021/acssuschemeng.9b07368>.
- (7) Son, Y. J.; Marquez, R. A.; Kawashima, K.; Smith, L. A.; Chukwuneke, C. E.; Babauta, J.; Mullins, C. B. Navigating iR Compensation: Practical Considerations for Accurate Study of Oxygen Evolution Catalytic Electrodes. *ACS Energy Lett.* **2023**, *8* (10), 4323–4329. <https://doi.org/10.1021/acsenergylett.3c01658>.
- (8) de Groot, M. T. Alkaline Water Electrolysis: With or without Iron in the Electrolyte? *Curr. Opin. Chem. Eng.* **2023**, *42*, 100981. <https://doi.org/10.1016/j.coche.2023.100981>.
- (9) Bernt, M.; Gasteiger, H. A. Influence of Ionomer Content in IrO₂/TiO₂ Electrodes on PEM Water Electrolyzer Performance. *J. Electrochem. Soc.* **2016**, *163* (11), F3179. <https://doi.org/10.1149/2.0231611jes>.
- (10) Bender, G.; Carmo, M.; Smolinka, T.; Gago, A.; Danilovic, N.; Mueller, M.; Ganci, F.; Fallisch, A.; Lettenmeier, P.; Friedrich, K. A.; Ayers, K.; Pivoar, B.; Mergel, J.; Stoltzen, D. Initial Approaches in Benchmarking and Round Robin Testing for Proton Exchange Membrane Water Electrolyzers. *Int. J. Hydrogen Energy* **2019**, *44* (18), 9174–9187. <https://doi.org/10.1016/j.ijhydene.2019.02.074>.
- (11) Bernt, M.; Siebel, A.; Gasteiger, H. A. Analysis of Voltage Losses in PEM Water Electrolyzers with Low Platinum Group Metal Loadings. *J. Electrochem. Soc.* **2018**, *165* (5), F305–F314. <https://doi.org/10.1149/2.0641805jes>.

- (12) Anantharaj, S.; Kundu, S.; Noda, S. "The Fe Effect": A Review Unveiling the Critical Roles of Fe in Enhancing OER Activity of Ni and Co Based Catalysts. *Nano Energy* **2021**, *80*, 105514. <https://doi.org/10.1016/j.nanoen.2020.105514>.
- (13) Trotochaud, L.; Young, S. L.; Ranney, J. K.; Boettcher, S. W. Nickel–Iron Oxyhydroxide Oxygen-Evolution Electrocatalysts: The Role of Intentional and Incidental Iron Incorporation. *J. Am. Chem. Soc.* **2014**, *136* (18), 6744–6753. <https://doi.org/10.1021/ja502379c>.
- (14) Ehlers, J. C.; Feidenhans'l, A. A.; Therkildsen, K. T.; Larrazábal, G. O. Affordable Green Hydrogen from Alkaline Water Electrolysis: Key Research Needs from an Industrial Perspective. *ACS Energy Lett.* **2023**, *1502–1509*. <https://doi.org/10.1021/acsenergylett.2c02897>.
- (15) Son, Y. J.; Kim, S.; Leung, V.; Kawashima, K.; Noh, J.; Kim, K.; Marquez, R. A.; Carrasco-Jaim, O. A.; Smith, L. A.; Celio, H.; Milliron, D. J.; Korgel, B. A.; Mullins, C. B. Effects of Electrochemical Conditioning on Nickel-Based Oxygen Evolution Electrocatalysts. *ACS Catal.* **2022**, *12* (16), 10384–10399. <https://doi.org/10.1021/acscatal.2c01001>.
- (16) Lira Garcia Barros, R.; Kraakman, J. T.; Sebregts, C.; van der Schaaf, J.; de Groot, M. T. Impact of an Electrode-Diaphragm Gap on Diffusive Hydrogen Crossover in Alkaline Water Electrolysis. *Int. J. Hydrogen Energy* **2023**. <https://doi.org/10.1016/j.ijhydene.2023.09.280>.
- (17) Klaus, S.; Cai, Y.; Louie, M. W.; Trotochaud, L.; Bell, A. T. Effects of Fe Electrolyte Impurities on Ni(OH)₂/NiOOH Structure and Oxygen Evolution Activity. *J. Phys. Chem. C* **2015**, *119* (13), 7243–7254. <https://doi.org/10.1021/acs.jpcc.5b00105>.
- (18) Kim, Y.; Jung, S.-M.; Kim, K.-S.; Kim, H.-Y.; Kwon, J.; Lee, J.; Cho, H.-S.; Kim, Y.-T. Cathodic Protection System against a Reverse-Current after Shutdown in Zero-Gap Alkaline Water Electrolysis. *JACS Au* **2022**, *2* (11), 2491–2500. <https://doi.org/10.1021/jacsau.2c00314>.
- (19) Huang, L.-F.; Hutchison, M. J.; Santucci, R. J.; Scully, J. R.; Rondinelli, J. M. Improved Electrochemical Phase Diagrams from Theory and Experiment: The Ni–Water System and Its Complex Compounds. *J. Phys. Chem. C* **2017**, *121* (18), 9782–9789. <https://doi.org/10.1021/acs.jpcc.7b02771>.