



Data-Intensive Electrocatalysis: A DFT-to-ML Predictive Workflow for CO₂ Reduction and On-Demand H₂O₂ Electrosynthesis

Ikenna Okechukwu^{a*}, Chinwendu Faustina Achilonu^a, Nkiruka Nancy Duru^a, Samson Uchenna Eze^b, Rebecca Chioma Okereke^c, Nneoma Anne Ezealuma^d, Oluwafemi Fikayomi Bamidele^e, Odinakachukwu Nwaehiri^a

^aDepartment of Pure and Industrial Chemistry, University of Nigeria, Nsukka, Nigeria.

^bDepartment of Chemistry Education, Federal College of Education (Technical) Omoku, Nigeria.

^cDepartment of Geology, University of Port Harcourt, Port Harcourt, Nigeria.

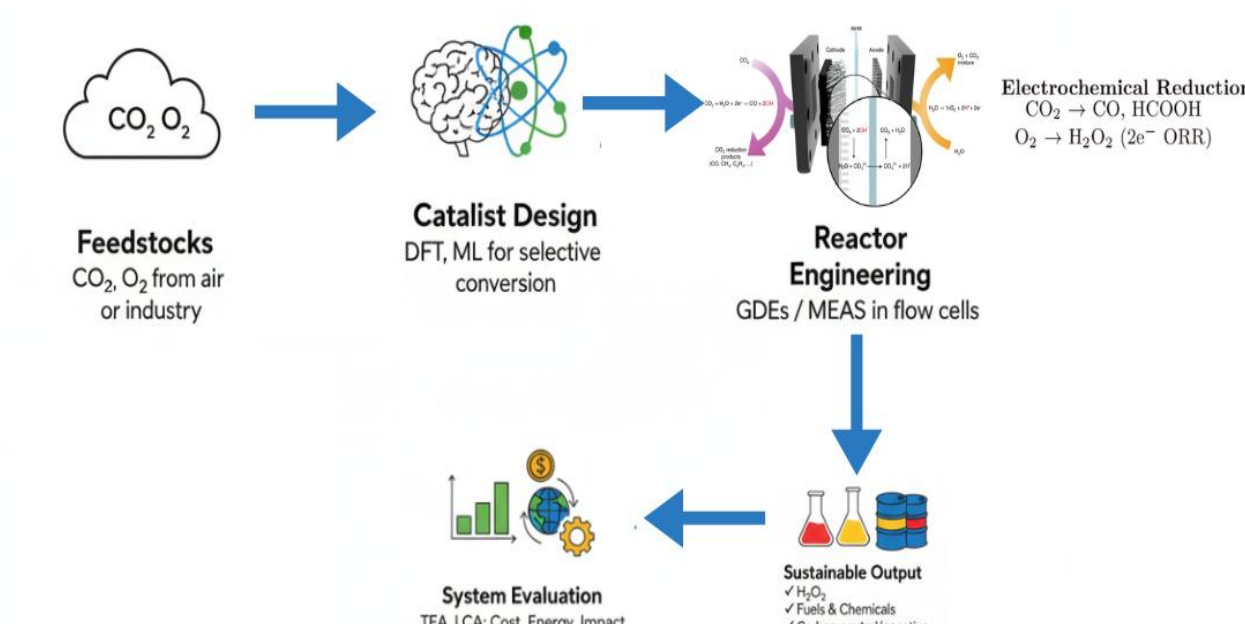
^dDepartment of Geology, Nnamdi Azikiwe University, Awka, Nigeria.

^eDepartment of Chemistry, University of Ibadan, Ibadan, Nigeria.



1. INTRODUCTION

Electrocatalytic CO₂ reduction and on-demand H₂O₂ electrosynthesis offer sustainable routes to fuels and chemicals, but discovering effective catalysts remains a major challenge. While Density Functional Theory (DFT) can accurately predict catalyst properties, it is too slow and expensive for large-scale screening. To accelerate this process, we use Machine Learning (ML) trained on a high-quality DFT database to rapidly predict high-performance catalysts for both reactions. These candidates are then integrated into advanced Gas Diffusion Electrodes (GDEs) to maximize efficiency in practical reactor systems. Finally, a full Techno-Economic Analysis (TEA) and Life Cycle Assessment (LCA) evaluate commercial viability and confirm the environmental sustainability of the technology from cradle to grave.



2. PROBLEM STATEMENT

- How can we accelerate catalyst discovery beyond the slow and costly limits of traditional DFT screening?
- What role do catalyst-layer play in advancing scalable CO₂ and H₂O₂ electrocatalytic systems?
- How can we implement Life Cycle Assessment (LCA) and Techno-Economic Analysis (TEA) for these products to evaluate their production efficiency, processing feasibility, and overall sustainability?

3. DATA PREPARATION

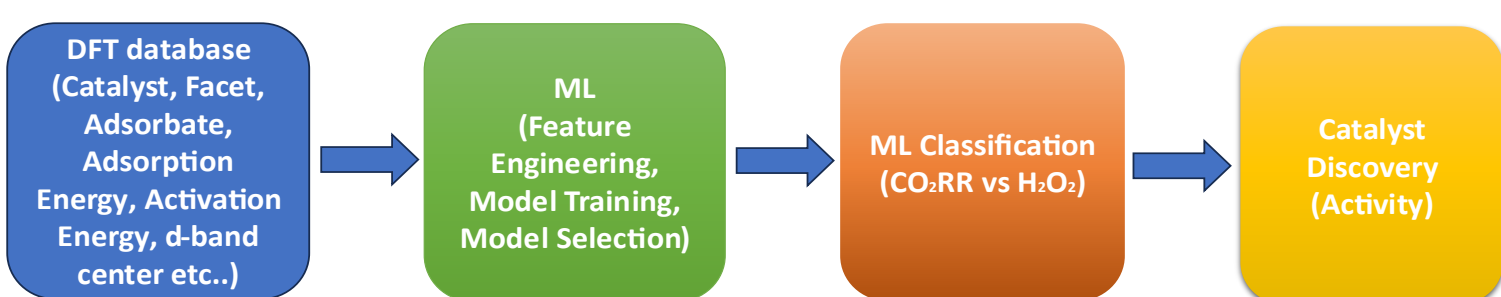


Figure 1: A DFT-ML Workflow for Predicting Electrocatalyst Selectivity.

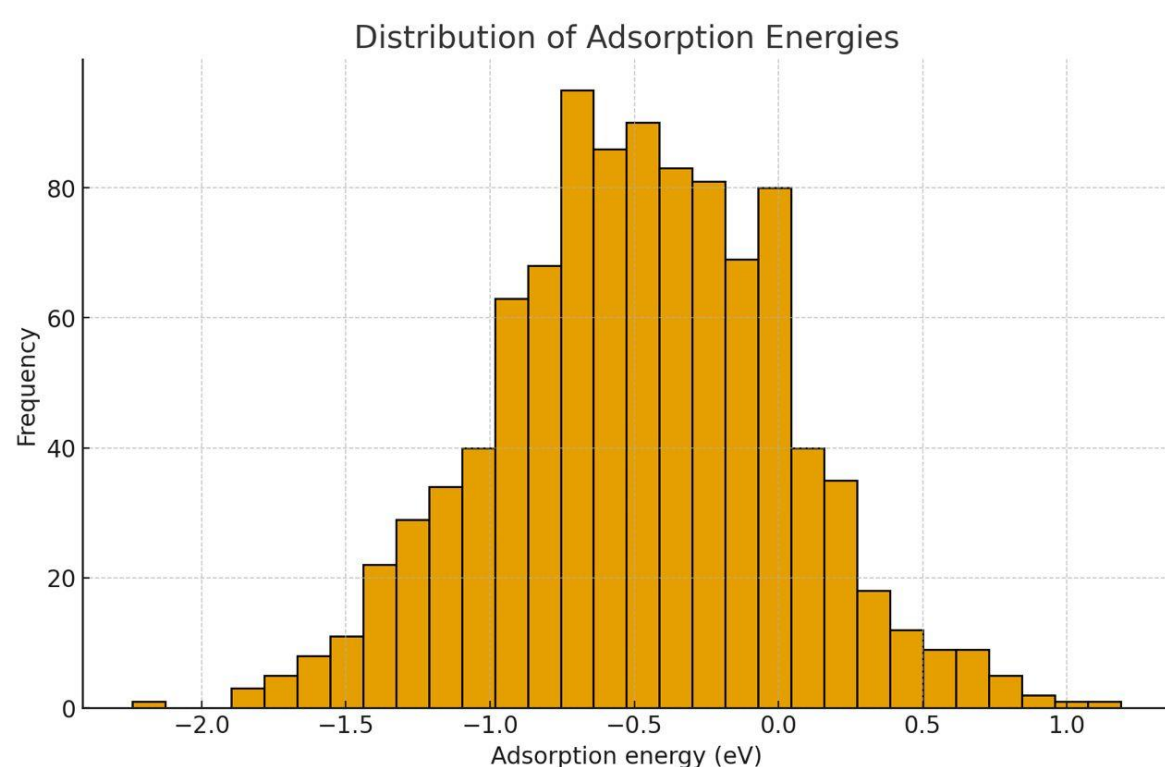


Figure 2: A histogram of the distribution of calculated adsorption energies across the catalyst dataset.

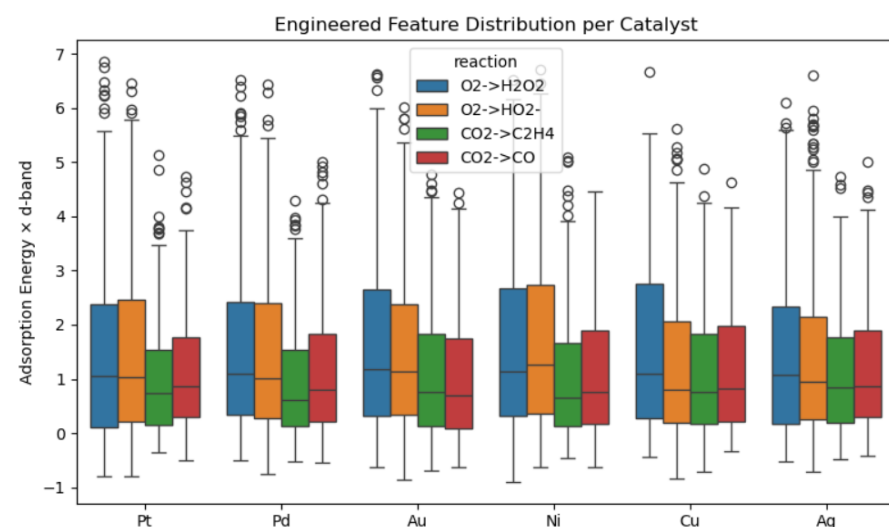


Fig 3: Comparison of a key catalytic descriptor across six metal catalysts.

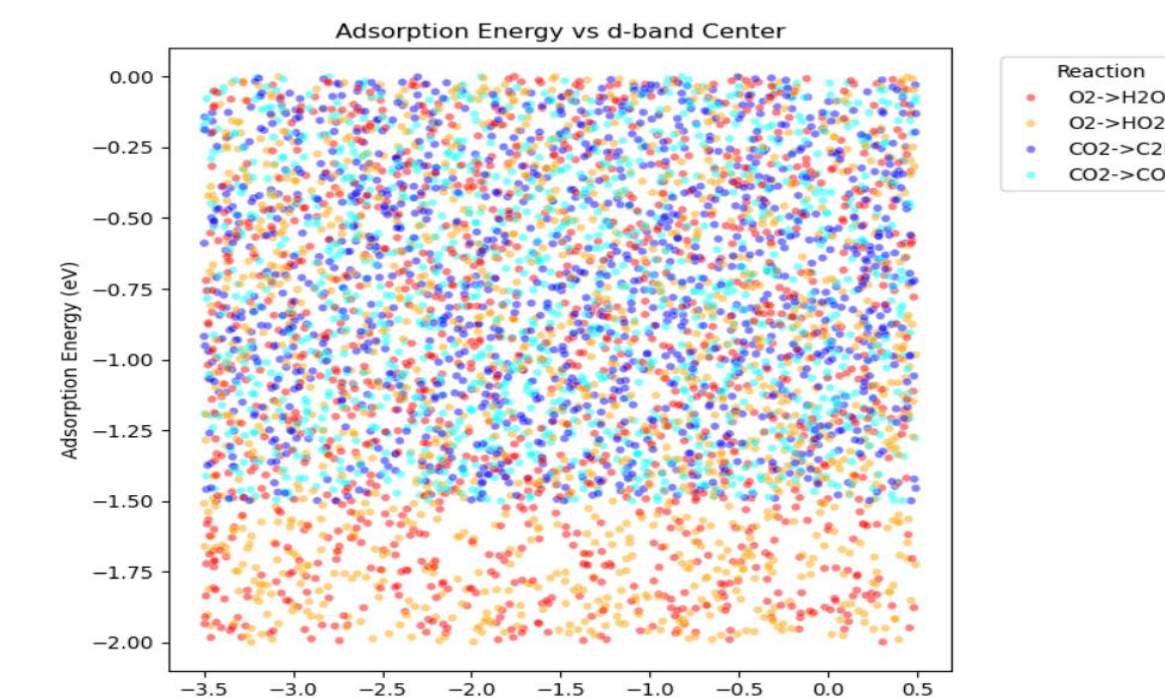


Fig 4: A plot showing the correlation between two key electronic descriptors for ORR and CO₂RR pathways.

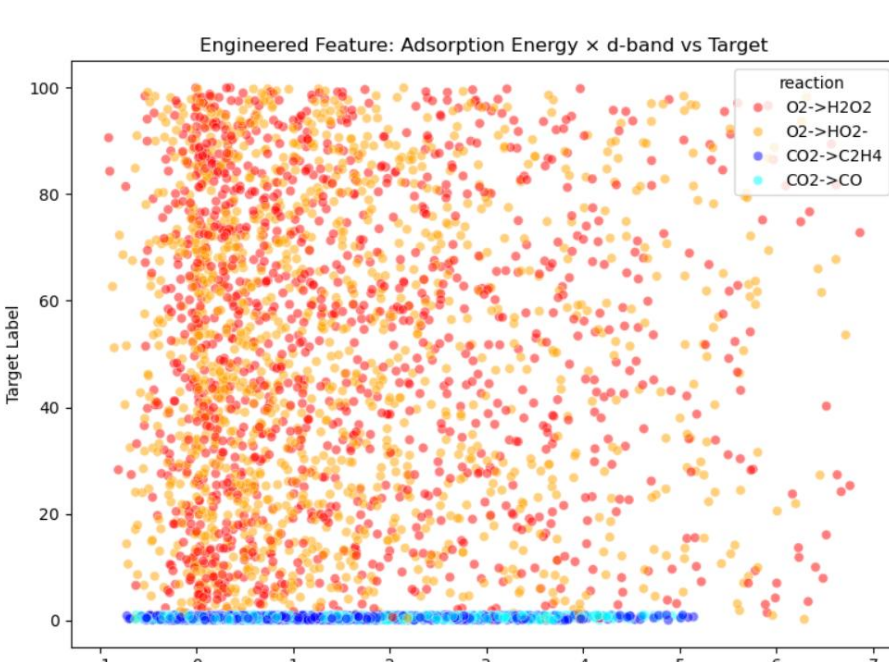


Figure 5: Scatter plot of the target label versus the engineered feature (Adsorption Energy x d-band), colored by reaction pathway.

4. MACHINE LEARNING PREDICTION

The Model Shows a Moderately Strong, Positive Correlation
R² Value: The R² value of 0.668 (for CO₂ Reduction) and 0.647 (for H₂O₂ Electrosynthesis) is the most important metric here. It means the model can successfully explain about 65-67% of the variation in catalyst performance.

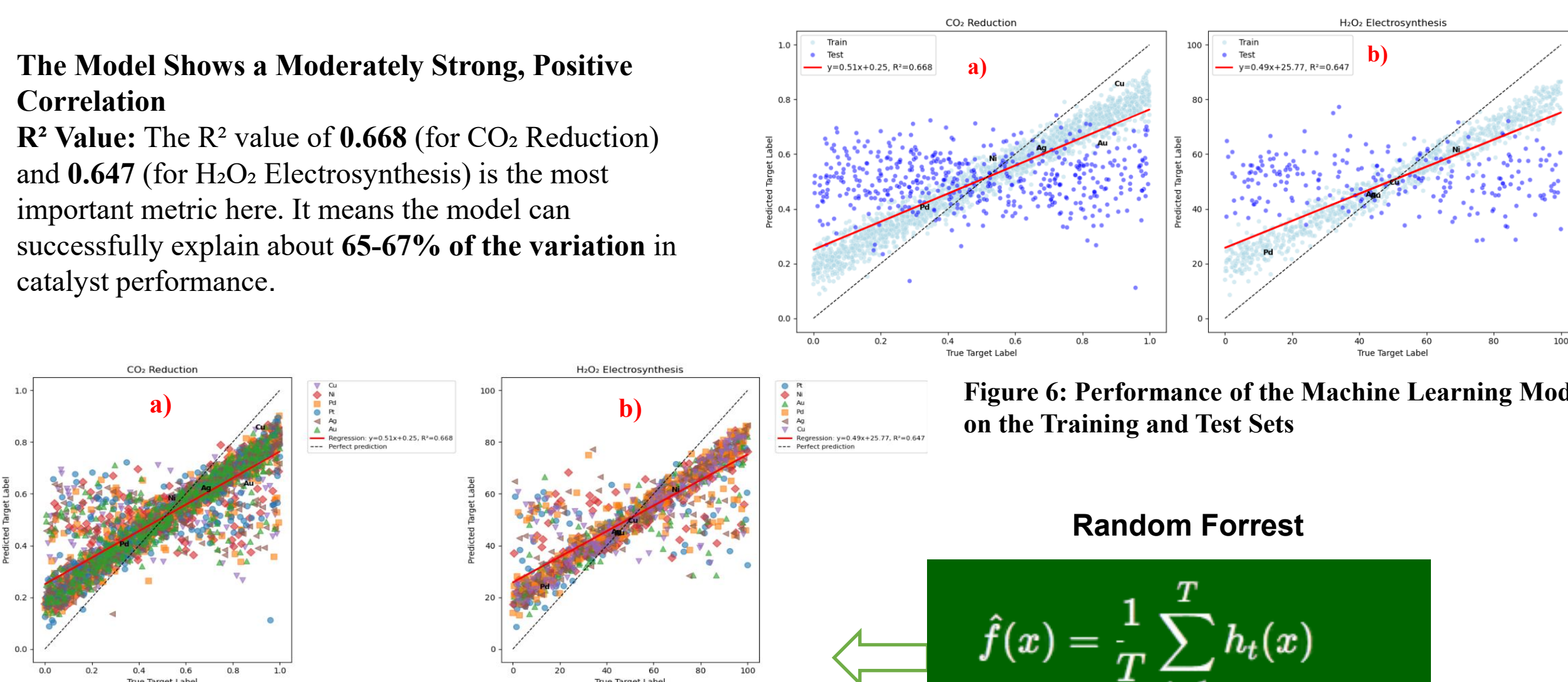


Figure 7: ML model performance for (a) CO₂ Reduction and (b) H₂O₂ Electrosynthesis.

5. ERROR VALIDATION

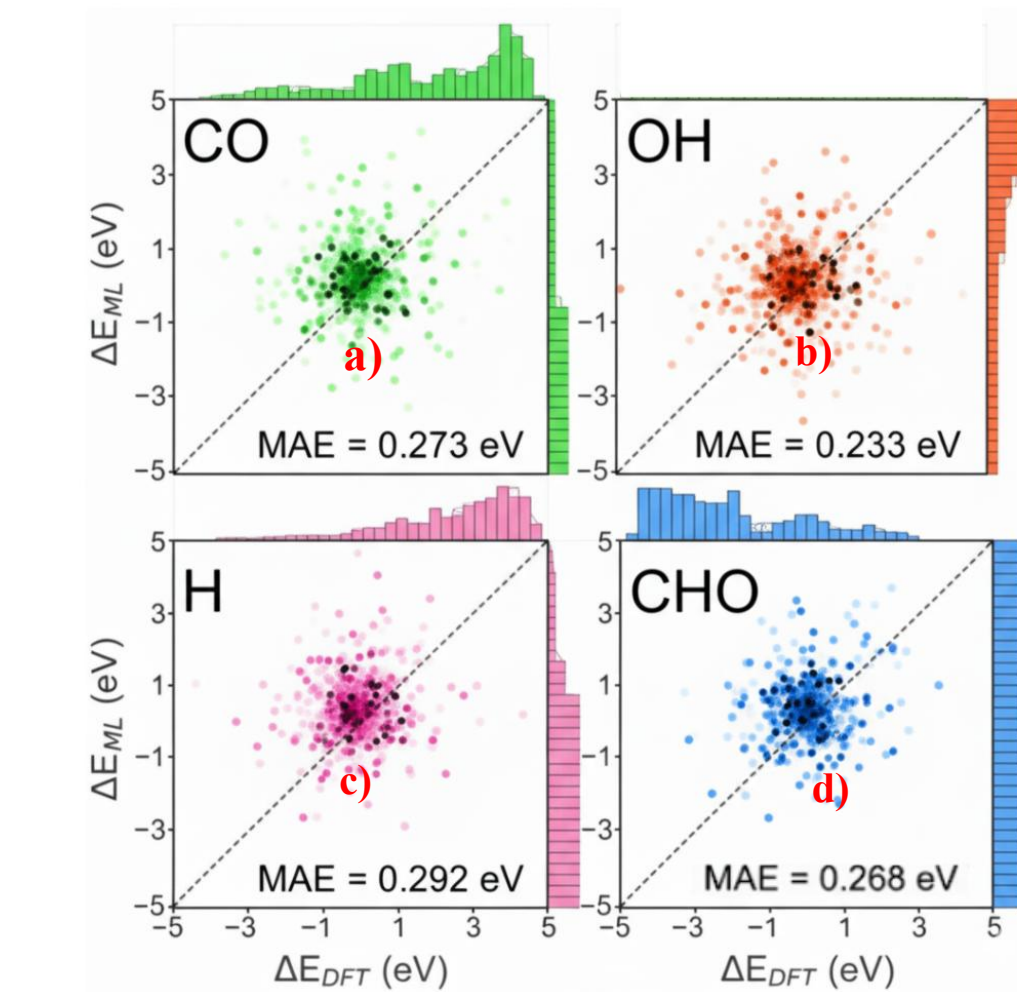


Figure 8: A validation of the ML model showing the Mean Absolute Error (MAE) for key adsorbate energies from the Test Set

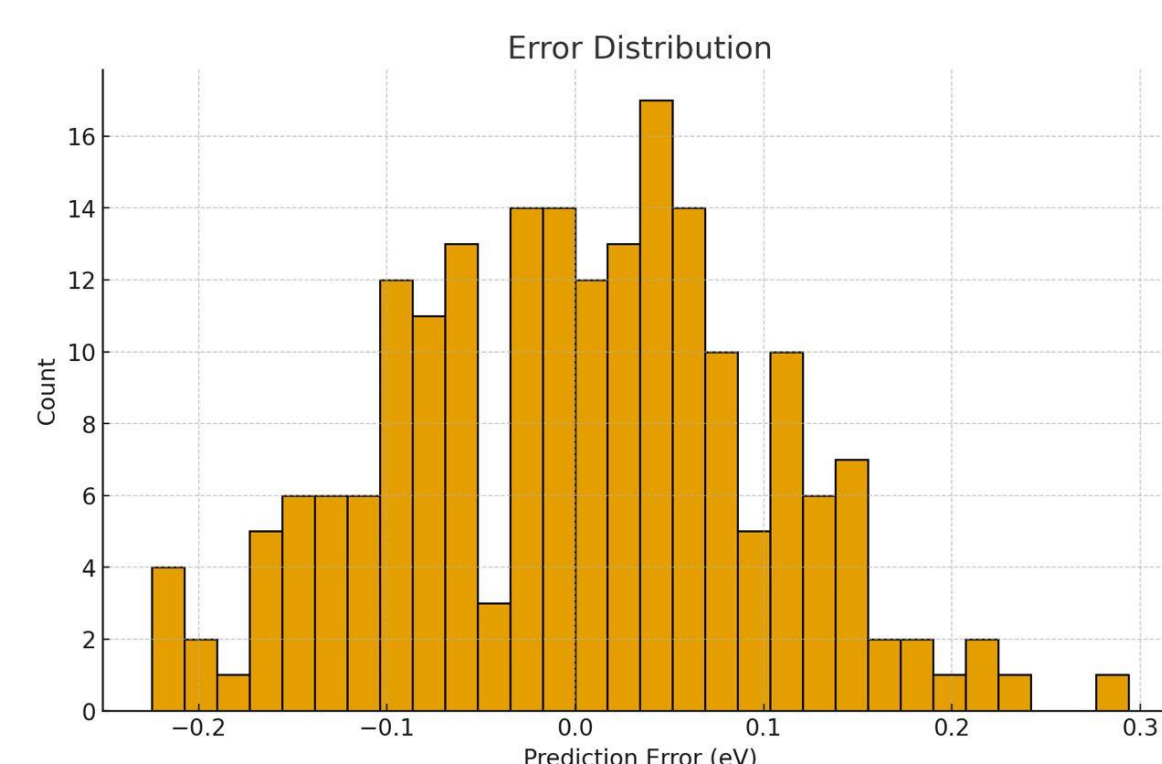


Fig 9: Histogram of the prediction errors from the Machine Learning model

An error of less than 0.3 eV is often considered "chemically accurate" enough for initial screenings.

6. CATALYST SELECTIVITY

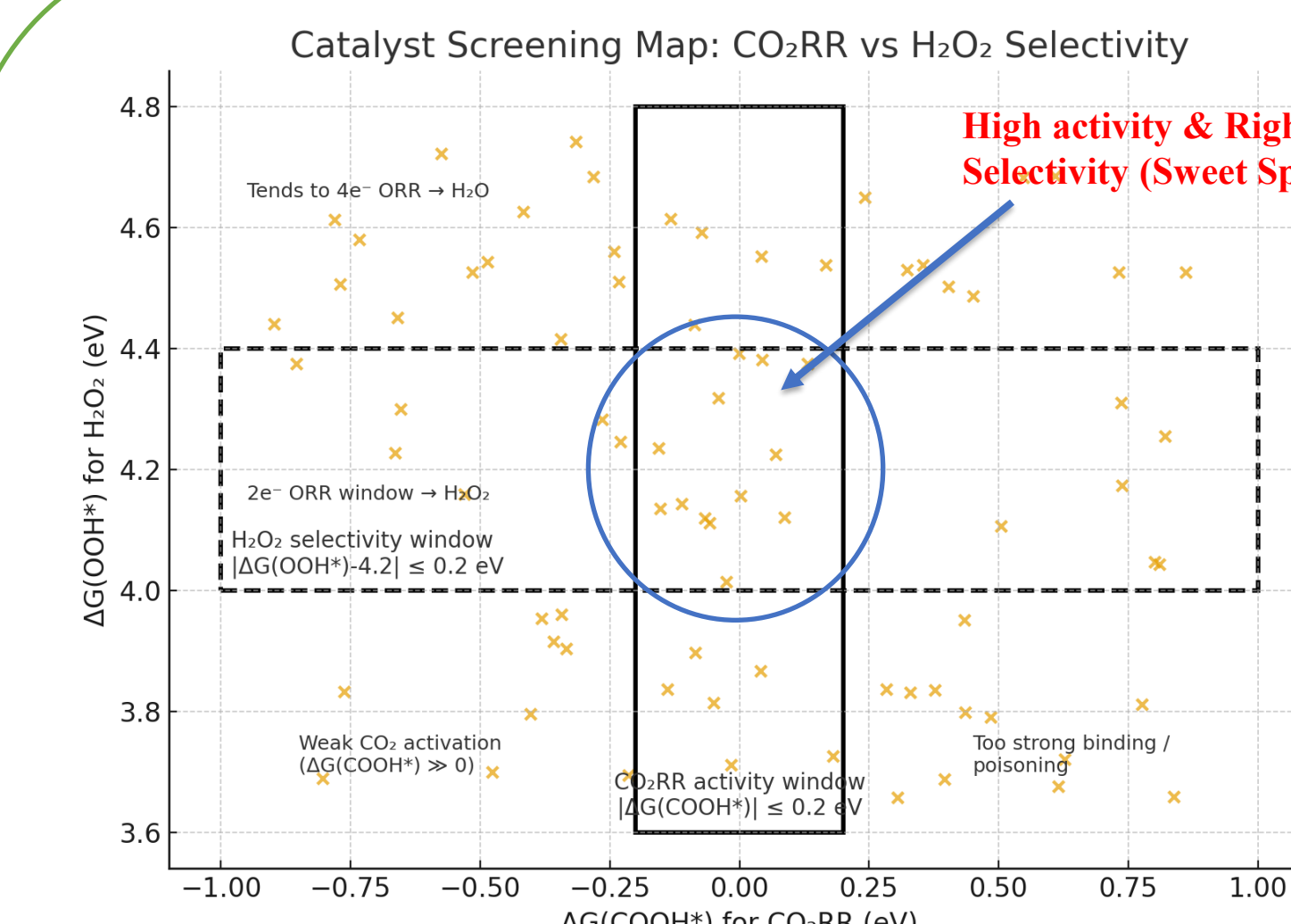


Figure 10: A Catalyst Screening Map Comparing Descriptors for CO₂RR Activity and H₂O₂ Selectivity.

The volcano plot's intersection marks the catalyst's "sweet spot" balancing high activity and selectivity for optimal efficiency.

9. CONCLUSIONS

- Our DFT-to-ML model accelerates catalyst screening (R² ≈ 0.67), identifying candidates like Cu, Au for CO₂ reduction and Ni, Ag for H₂O₂ synthesis. Performance must be validated in practical GDE systems, and ultimate scalability depends on favourable TEA and LCA for commercial and environmental viability.
- Our work advances innovative electrochemical technologies (SDG 9), promotes sustainable resource use (SDG 12), and reduces emissions through cleaner CO₂ conversion and H₂O₂ synthesis (SDG 13).
- Future directions for using machine learning (ML) to predict catalysts from Density Functional Theory (DFT) data will focus on creating more robust, efficient, and reliable predictive models.

7. GDEs MODULES

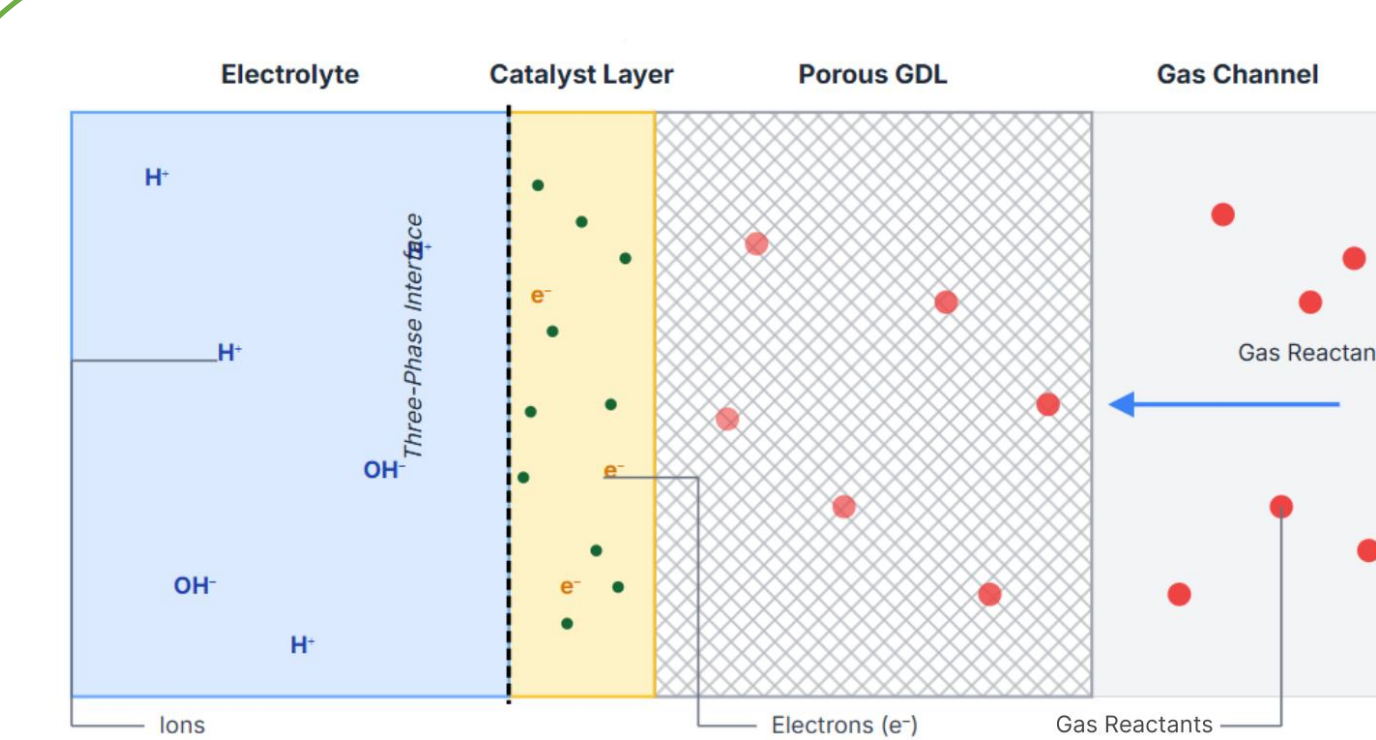


Figure 11: Gas Diffusion Electrode (GDE) Cross-Section

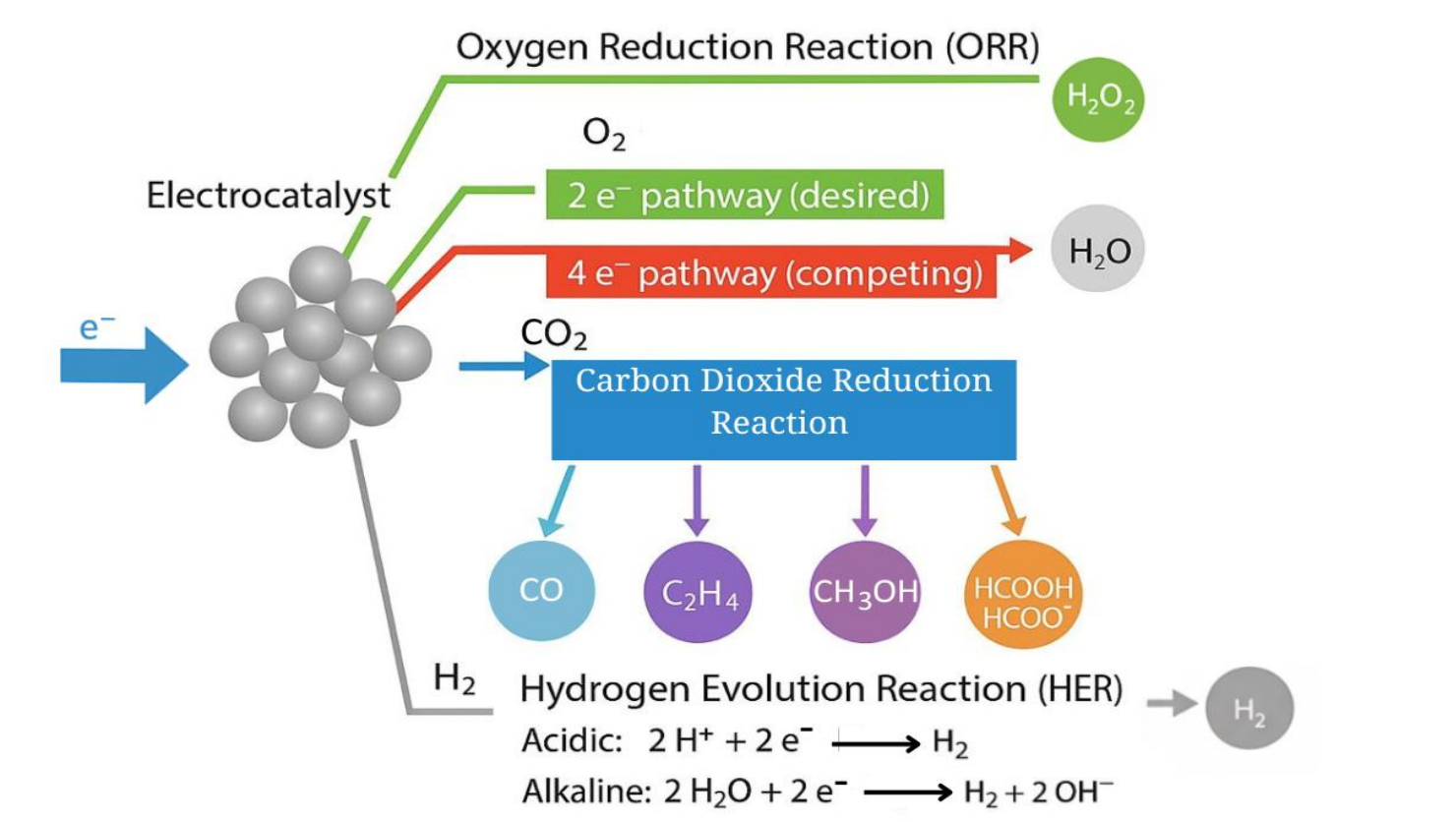


Figure 12: An Illustration of the selectivity challenge between ORR, CO₂RR, and HER on an electrocatalyst.

10. ACKNOWLEDGEMENTS

The authors acknowledge the University of Nigeria, Nsukka, for institutional support and the Hugging Face team for providing access to the DFT dataset used in this study.



8. SYSTEM EVALUATION

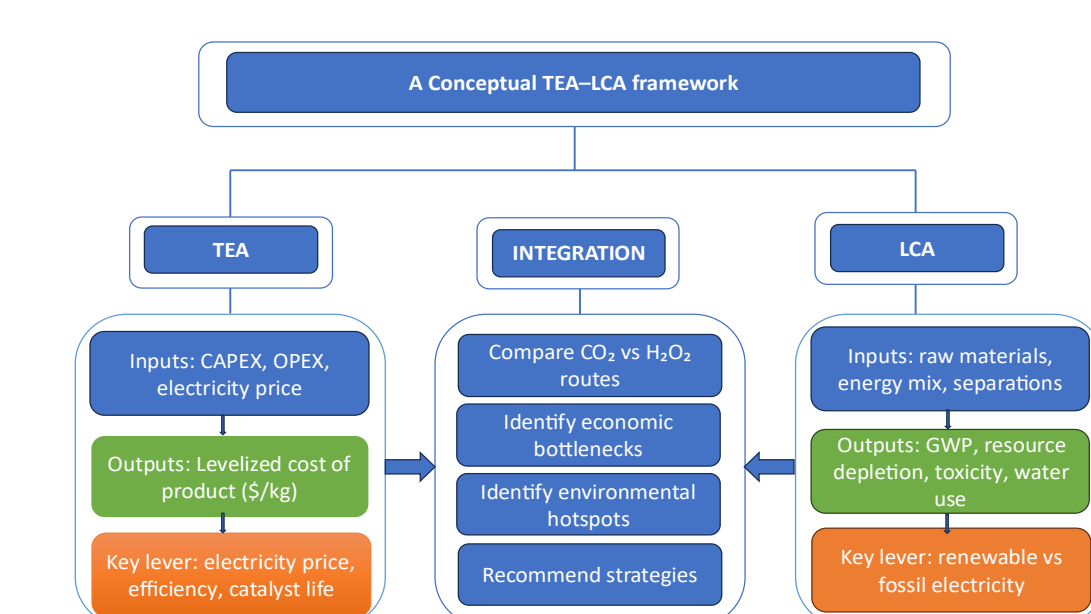


Fig 13: Conceptual Framework for Integrating TEA and LCA

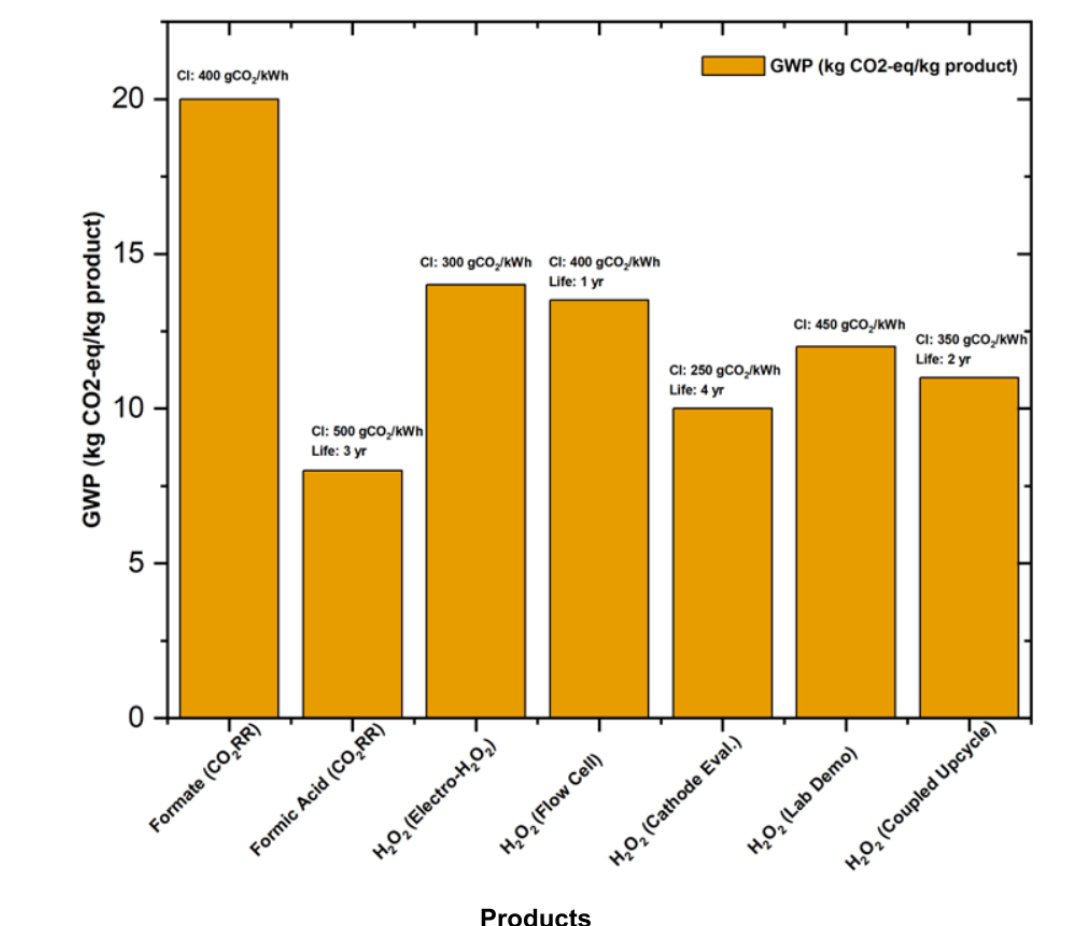


Fig 14: Comparative GWP of electrocatalytic products, annotated with electricity carbon intensity and reactor lifetime.

11. REFERENCES

- Wang, G., Chen, J., Ding, Y., Cai, P., Yi, L., Li, Y., Tu, C., Hou, Y., Wen, Z., Dai, L. (2021). *Electrocatalysis for CO₂ conversion: from fundamentals to value-added products*. Chem Soc Rev. 2021 Apr 21;50(8):4993–5061. doi: <https://doi.org/10.1039/d0cs00071j>
- Zhou, J., Gu, Z., Wang, X., Zhang, X., Ouyang, Y., & Shi, L. (2025). Screening of single-atom catalysts for CO₂ electroreduction to CH₄ using DFT calculations and machine learning. *Journal of Materials Chemistry A*, 13(29), 24073–24083. <https://doi.org/10.1039/D5TA03431K>
- Wang, K., Qu, T., Li, Q., Tan, S., & Chen, X. (2023). Techno-Economic Analysis of Electrocatalytic CO₂ Reduction into Methanol: A Comparative Study between Alkaline Flow Cell and Neutral Membrane Electrode Assembly. *Catalysts*, 13(6), 1005. <https://doi.org/10.3390/catal13061005>
- Li, F., Thevenon, A., Rosas-Hernández, A., Wang, Z., Li, Y., Gabardo, C. M., ... Sargent, E. H. (2021). Molecular Tuning of CO₂-to-Ethylene Conversion. *Nature*, 577(7791), 509–513. <https://doi.org/10.1038/s41586-019-1782-2>