The Open Catalyst 2025 (OC25) Dataset and Models for Solid-Liquid Interfaces

Sushree Jagriti Sahoo 1 , Mikael Maraschin 2 , Daniel S. Levine 1 , Zachary Ulissi 1 , C. Lawrence Zitnick 1 , Joel B Varley 4 , Joseph A. Gauthier 2,† , Nitish Govindarajan 3,4† , Muhammed Shuaibi 1,†

¹FAIR at Meta, ²Department of Chemical Engineering, Texas Tech University, Lubbock, TX 79409, USA, ³School of Chemistry, Chemical Engineering and Biotechnology, Nanyang Technological University, 21 Nanyang Link, Singapore 637371, Singapore, ⁴Materials Science Division, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA [†]Co-corresponding Author

Catalysis at solid-liquid interfaces plays a central role in the advancement of energy storage and sustainable chemical production technologies. By enabling accurate, long-time scale simulations, machine learning (ML) models have the potential to accelerate the discovery of (electro)catalysts. While prior Open Catalyst datasets (OC20 and OC22) have advanced the field by providing large-scale density functional theory (DFT) data of adsorbates on surfaces at solid-gas interfaces, they do not capture the critical role of solvent and electrolyte effects at solid-liquid interfaces. To bridge this gap, we introduce the Open Catalyst 2025 (OC25) dataset, consisting of 7,801,261 calculations across 1.511,270 unique explicit solvent environments. OC25 constitutes the largest and most diverse solidliquid interface dataset that is currently available and provides configurational and elemental diversity: spanning 88 elements, commonly used solvents/ions, varying solvent layers, and off-equilibrium sampling. State-of-the-art models trained on the OC25 dataset exhibit energy, force, and solvation energy errors as low as 0.1 eV, 0.015 eV/Å, and 0.04 eV, respectively; significantly lower than than the recently released Universal Models for Atoms (UMA-OC20). Additionally, we discuss the impact of the quality of DFT-calculated forces on model training and performance. The dataset and accompanying baseline models are made openly available for the community. We anticipate the dataset to facilitate large length-scale and long-timescale simulations of catalytic transformations at solid-liquid interfaces, advancing molecular-level insights into functional interfaces and enabling the discovery of next-generation energy storage and conversion technologies.

Date: September 23, 2025

M.S.(mshuaibi@meta.com)

Code: https://github.com/facebookresearch/fairchem Dataset: https://huggingface.co/facebook/OC25 Models: https://huggingface.co/facebook/OC25

1 Introduction

Solid-liquid interfaces are at the heart of several critical technologies, including catalysis, batteries, sensors, and others [3, 4, 9, 10, 13, 15, 17, 20, 22, 27, 28, 31, 41, 45, 47, 49, 54, 59, 66, 67, 71, 75, 76, 79, 80, 82, 83, 85]. In the context of heterogeneous catalysis, interfaces (solid-gas and solid-liquid) are the key zones that enable chemical transformations. In contrast to gas-phase heterogeneous catalysis, thermocatalytic and electrocatalytic processes at solid-liquid interfaces are complicated by the presence of the liquid phase, where solvent and ion effects are intimately coupled with surface reactivity [24, 33, 39, 64, 70, 84]. For example, solvent molecules can stabilize reaction intermediates, thereby impacting adsorption/desorption equilibria, and can act as coreactants. At electrified solid-liquid interfaces, ions result in the formation of an electrical double layer whose structure and properties dictate the thermodynamics and kinetics of charge transfer [44] or may specifically adsorb on the surface and impede catalytic activity [12, 55, 60]. Developing a molecular-level understanding of these solvent and electrolyte effects is therefore essential for rational design of solid-liquid interfaces for

heterogeneous (electro)catalysis. However, the community has so far struggled to achieve an atomic-scale understanding of these complex and dynamic solid-liquid interfaces. A major contributor to this challenge is interference of spectroscopic radiation with the electrolyte, which has complicated unraveling fundamental insights comparable to those that led to the modern physics-informed framework for gas-phase heterogeneous catalyst design [53, 56, 57]. Given the difficulties associated with *operando* spectroscopic measurements of solid-liquid interfaces, computational modeling offers a promising route to develop molecular-level insights into the various physicochemical properties that occur at electrochemical interfaces under reaction conditions.

Although electronic structure methods aimed at simulating electrified interfaces have made significant advances in the past 15 years, [38, 42, 63, 70] fundamental limitations persist - as was discussed in a recent perspective [21]. Briefly, treatment of the electrolyte within density functional theory (DFT) models of the electrical double layer remains a key barrier because of the high computational cost of including explicit electrolyte molecules in models and in performing long time-scale molecular dynamics (MD) simulations of the electrode/electrolyte interface. Three main strategies have emerged to overcome this challenge: (i) the utilization of continuum (implicit) solvation models, which circumvents the need to explicitly model the electrolyte and, to an extent, mitigates the need for dynamics [63], (ii) brute force DFT-based dynamics using explicit solvent, possibly significantly undersampling the solvent configuration space as a result, [29, 61, 81, 86] and (iii) the use of machine learning interatomic potentials (MLIPs) based on DFT calculations (ground-truth) to reduce the computational cost associated with an explicit electrolyte model and perform longer time-scale MD simulations of solid-liquid interfaces [25, 51, 52, 58, 87].

Over the past few years, the development of MLIPs has revolutionized the field of atomistic simulations [6, 7, 14, 16, 18, 19, 46, 62, 78]. What used to be an onerous task that individual research groups would tackle for each system of interest, creating significant redundancy of effort across the community, is increasingly becoming a matter of downloading freely available foundational MLIPs and fine-tuning them for specific system(s) of interest [6, 78]. This has been enabled by the availability of large, open datasets spanning a variety of systems and applications in materials science: bulk materials [5], crystalline organic frameworks [68, 69], molecules [43], and catalysts [11, 73] to train generalized MLIPs. However, the solid-liquid interface, crucial to understanding electrochemical transformation technologies, is largely missing from these open datasets. It should come as no surprise then, that a MLIP that has never been explicitly shown how long-range charges interact, or how a metal surface might behave in the presence of explicit water/solvent molecules, may struggle to adequately model the structure and properties of the electrified double layer crucial to the understanding of electrocatalytic processes. While recent dataset efforts have tried to address this, they are rather limited in both size and chemical diversity, often exploring very specific chemistries and applications [30, 88].

In this work, we introduce the Open Catalyst 2025 (OC25) dataset and models for solid-liquid interfaces from Meta FAIR to broaden the capabilities of generalized foundational MLIPs and to shed light on fundamental aspects of the (electro)chemical transformation technologies that are likely to be critical in our transition to a low carbon economy. The OC25 dataset contains over 7 million single-point DFT calculations sampled from a wide range of transition metal and metal oxide surfaces in contact with several relevant electrolytes, including different cations and anions that have been used in experimental studies. We train several baseline models on the OC25 dataset, demonstrating their utility over existing state-of-the-art models. The dataset, models, and code are all open source and freely available to allow the community to expand upon the work we demonstrate here.

2 OC25 Dataset

The Open Catalyst 2020 and 2022 datasets (OC20/OC22) [11, 73], Catalysis-Hub[77], GASpy[72], and AQCat[65] represent the largest catalyst datasets for training MLIPs. The diversity and scale of these datasets have led to the development of several state-of-the-art ML models for the community [18, 19, 46]. Although models trained on these datasets have demonstrated successful applications for gas-phase heterogeneous catalysis, little has been done for solid-liquid and electrified solid-liquid interfaces. The Open Catalyst 2025 (OC25) dataset hopes to bridge this gap by presenting the largest catalyst solid-liquid interface dataset.

Open Catalyst 2025 (OC25)

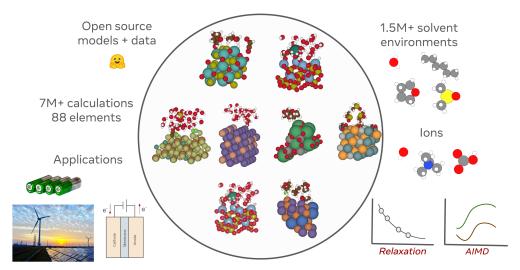


Figure 1 Overview of OC25, including dataset statistics, sampling strategies, relevant applications, and sample snapshots of the dataset.

2.1 Summary

The OC25 dataset is constructed from a combination of DFT relaxations and molecular dynamics simulations of catalyst/solvent/ion structures. Structures consist of a catalyst surface, a solvent, and at least 1 adsorbate molecule. Structures may also contain multiple adsorbates and an ion present to better capture reactive environments and electrocatalytic systems. Surfaces are sampled from unique bulk structures in the Materials Project[32], including oxides. Eight commonly used solvents are sampled across varying solvent depths, with system sizes of 144 atoms on average. Adsorbates are randomly sampled from 98 unique molecules, including OC20 species and additional reactive intermediates. Overall, the OC25 dataset consists of 7,801,261 single-point calculations, spanning 1,511,270 unique systems and 88 unique elements.

OC25 structures correspond to highly off-equilibrium configurations, a property that we have seen to aid in training ML models [5, 11, 18]. To accomplish this, short time-scale MD simulations at high temperature (1000K) were run, minimizing the redundancy in structures that can come from fully relaxed configurations, i.e. OC20 and OC22. The inclusion of off-equilibrium configurations is reflected in the force distributions of OC25 (Figure 2), with forces much higher than those of OC20 and OC22.

2.2 Dataset Generation

The OC25 dataset is generated following a similar pipeline to that of OC20, with the addition of a solvated interface. Structures were created in three stages: (1) adsorbate+surface generation, (2) interface construction, and (3) *ab initio* calculations. All code to generate configurations is provided at https://github.com/facebookresearch/fairchem/.

2.2.1 Adsorbate+Surface generation

We begin by first constructing adsorbate+surface structures in vacuum. A bulk material is randomly sampled from a set of 39,821 materials in Materials Project. For the sampled bulk material, all symmetrically distinct surfaces with miller indices less than or equal to 3 are enumerated and a random surface is selected. The surface is tiled in the xy plane to a length of 8Å. Anywhere between 1-5 adsorbates are then randomly sampled for placement, with a single adsorbate biased 50% of the time and 20% of multiple adsorbate samples being identical molecules. The adsorbates are sampled from the original set of OC20 adsorbates, containing oxygen, hydrogen, C1/C2 molecules, and nitrogen-containing species [11]. This set is extended to also include reactive intermediates from the OC20NEB[74] and OCx24[2], see Appendix C.1. Adsorbate(s) placement is

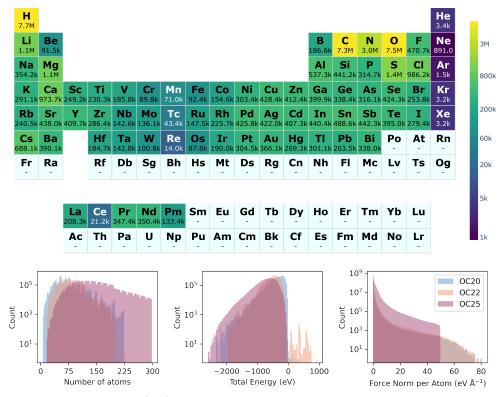


Figure 2 OC25 dataset distribution. (top) OC25 element distribution, with counts corresponding to the number of systems containing an element. (bottom) Distribution of the number of atoms, total energy, and force norm across the OC25, OC20, and OC22 datasets.

performed with the Adsorb-ML workflow[40], randomly placed on sites selected from Delaunay triangulation of surface atoms followed by rotations along the z-axis and wobbles around the x/y-axis. In structures with multiple adsorbates, sites are only considered if their distance to the nearest adsorbate does not result in considerable overlap $(r_{cov} + 0.1\text{Å})$.

2.2.2 Interface construction

Given an adsorbate+surface structure, the solid-liquid interface is constructed by sampling a random solvent and ion combination. Solvents are sampled from a list of eight commonly used solvents (e.g. polar/nonpolar, protic/aprotic). Similarly, ions are sampled from nine cations and anions of varying charges and size. The full list of solvents and ions is illustrated in Figure 3. The surface charge density distribution of the metal interfaces sampled in the dataset are also shown in Figure 3, with values ranging from \sim -80 μ C/cm² to \sim 60 μ C/cm², corresponding to cathodic (reducing) and anodic (oxidizing) conditions, respectively. As can be seen in the ion and surface charge density distributions, a majority of the interfaces have zero surface charge density that corresponds to the condition of the potential of zero charge (PZC).

Given the importance and frequent use of water as a solvent in electrocatalytic applications, we biased our sampling of water, with all other solvents uniformly weighted. An ion is only sampled $\sim 50\%$ of the time. A solvent depth is then sampled between 5-10Å, with more weighting on ≤ 6 Å to limit excessive compute. Given the solvent depth and area of the surface, N solvent molecules are selected, where N is the number of molecules necessary to approximately satisfy the density of the solvent. The resulting solvent+ion box is then randomly packed with Packmol[50] and placed on top of the adsorbate+surface configuration.

To better capture meaningful interactions, for a subset of the dataset, we pre-optimize initial geometries with existing OC20 trained models. EquiformverV2-31M[46] and UMA-S-1[78] were used to relax geometries using loose convergence criteria of maximum per-atom force of 0.5 eV/Å or 50 steps, whichever comes first.

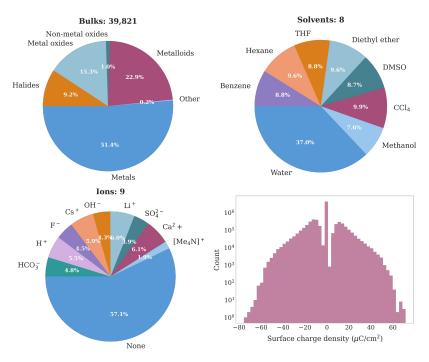


Figure 3 Overview of the bulks, solvents, ions sampled in OC25 and the surface charge distribution (in $\mu C/cm^2$) for the metallic interfaces in the dataset. Adsorbates are sampled from the same set of OC20, with the addition of a few reactive intermediates provided in Appendix C.1.

2.2.3 Ab initio calculations

Configurations are then evaluated with DFT in one of two ways: relaxations or *ab initio* molecular dynamics (AIMD). Structures sampled for relaxations are optimized for only 5 ionic steps. Similarly, short-time scale (10-50 steps) AIMD are performed at constant temperature and volume (NVT) at a temperature of 1000K. We limit simulations to short-time scales to maximize diversity in the dataset.

All DFT calculations were performed with the Vienna Ab Initio simulation Package (VASP)[34–37] v6.3.2. Similar to other large-scale dataset efforts (OC20/OC22), a broad set of settings was selected to balance accuracy and computational costs. Calculations were performed with a revised Perdew-Burke-Ernzerhof (RPBE) functional [26], supplemented with the D3 correction with zero damping to account for the non-local van der Waals dispersion interactions[23], plane wave cutoff energy of 400 eV, and a dipole correction in the z-direction [8]. The k-point mesh was constructed as a function of the cell parameters, similar to OC20, using a reciprocal density of 40. We utilized the non spin-polarized RPBE functional for two reasons: first, the vast majority of the surfaces sampled in this dataset are not magnetic, meaning enabling spin polarization would only add computational expense to dataset generation. Second, for the systems in our database that are magnetic and thus could benefit from spin polarization, ensuring the correct magnetic behavior (ferromagnetism vs antiferromagnetism, etc.) is not trivial, and thus the incremental benefit was deemed to not be worth the additional computational cost. The full set of VASP parameters can be found at https://github.com/facebookresearch/fairchem/.

2.2.4 Force convergence and consistency

The consistency between energy and force (e.g., $F = -\frac{dE}{dx}$) labels in DFT is critical to building reliable datasets for MLIP training. Given a level of theory, a static calculation in DFT codes like VASP are considered complete when the electronic self-consistency loop is converged. Convergence is often defined based on a break condition in total energy (e.g., total energy change between two electronic steps $< \sigma$, "EDIFF" in VASP). For OC25, the electronic termination criteria for the training data was set to 10^{-4} eV, balancing accuracy and computational cost, similar to previous works [5, 11, 73].

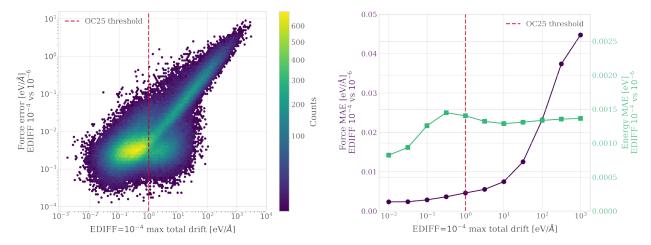


Figure 4 DFT force convergence errors as a function of the total drift in calculations with an electronic termination of 10^{-4} eV ("EDIFF"). Errors are computed against more tightly converged (10^{-6} eV) calculations for a ~ 300 k subset of the dataset. A threshold of 1 eV/Å on the max drift is selected for the OC25 training dataset. All validation and test sets were calculated with the tighter convergence settings.

The net force on any system is expected to be identically zero (i.e., zero acceleration) in the absence of any external fields. However, if the electronic structure calculations are not fully converged, non-zero net forces may be observed. DFT codes often implement routines to correct for these spurious forces, denoted as 'force drift.' For MD calculations in VASP, this drift is calculated and removed from the system during the integration step of MD. However, only the magnitude of the force drift is saved as an output – not the corrected force actually used within the simulation¹. For the validation and test datasets, all calculations were conducted as single points (i.e., non MD-calculations, IBRION $\neq 0$) where the spurious drift was removed using a tighter EDIFF in VASP (10^{-6} eV). This allowed us to directly test the impact of force convergence on final model performance, with potential impact both on the value of OC25 and future dataset efforts and whether training datasets with tighter convergence thresholds are needed.

To determine which samples to include in the final training dataset, we rigorously tested the impact of the net-force drift correction and the "EDIFF" threshold for force convergence on a subset of the dataset. In Figure 4, we show that the magnitude of the drift correction is strongly correlated with the error between the forces determined using more tightly converged calculations (EDIFF= 10^{-6} eV) and the forces determined using training data settings (EDIFF= 10^{-4} eV), and that structures with total drift greater than 10 eV/Å have much larger errors. Energy errors are largely unaffected, with total energy errors as low as 1.5 meV. To ensure reliable forces, we chose a conservative threshold of 1 eV/Å force drift as a quality threshold. The final training dataset was filtered to only include calculations with drifts smaller than this value.

We also investigated the ability of models trained on the unfiltered training data to predict the forces of more tightly converged (EDIFF= 10^{-6} eV) calculations in the validation/test datasets. Surprisingly, we found that models trained on the less tightly converged data were able to predict the forces of the more tightly converged set (see Section 3.2). This suggests that the models are to some degree robust to noise in the training data, and is somewhat at odds with conventional wisdom that any noise in the training data will propagate to reduced model performance.

2.3 OC25 Training, Validation, and Test Splits

The OC25 dataset is divided into training, validation, and test splits to ensure consistent evaluations by the community. Splits are created based on unique bulk-solvent combinations. Of the $\sim 260,000$ unique pairings, $\sim 2.5\%$ are held out for validation and test, respectively. For each data point, DFT total energies and per-atom forces are provided. To test generalizability beyond our defined splits, we generate several explicit out-of-distribution (OOD) splits. Dataset splits are summarized in Table 1.

¹https://www.vasp.at/wiki/index.php/Category:Forces

Table 1 Size of the OC25 train, validation, and test splits.

	Split	Size	Description
Train	All	7,395,512	Training set
Val	Val	203,630	OOD combos
	Test	202,119	OOD combos
	Solvent	11,111	OOD solvents
Test	Ion	$7,\!176$	OOD ions
	Both	6,989	OOD solvents+ions
	Solvation	5,713	$\Delta ilde{E_{solv}}$

2.3.1 Out-of-distribution Splits

Solvents. To asses model performance beyond the solvents used for OC25, a few additional solvents were sampled. These include ethylenecarbonate, acetonitrile, ethanol, and dichloromethane. Bulk materials, adsorbates, and ions used in these configurations remain in distribution.

lons. Ions beyond the ones used in OC25 were also sampled to evaluate generalizability on. These include Cl^- , $\mathrm{PO_4}^{3-}$, Mg^{2+} , and $\mathrm{NO_3}^-$. Bulk materials, adsorbates, and solvents used in these configurations remain in distribution.

Solvents + Ions (Both). Structures with both OOD solvents and ions compose this split. Only bulk materials and adsorbates used in these configurations remain in distribution.

Interfacial solvation. Adsorbate solvation energy is a commonly used property to study the effects of solvents on the interactions of molecules, ions, and complexes in solution on adsorbed reaction intermediates [29]. In the context of catalysis, this often corresponds to the energy difference between the adsorption energy in a solvated environment and in vacuum:

$$\Delta E_{solv} = \Delta E_{ads}^{solv} - \Delta E_{ads}^{vac} \tag{1}$$

As a proxy to this metric, we perform single point DFT evaluations of solvated configurations, and delete the respective regions (e.g. solvent, adsorbate, solvent+adsorbate) to generate the reference configurations. A pseudo-solvation energy, $\Delta \tilde{E_{solv}}$, is calculated based on these static snapshots. We omit any relaxations or molecular dynamics steps that often occur to simplify the task and ensure deterministic MLIP evaluation.

3 Baseline Models and Results

We evaluate OC25 using a set of baseline models that represent state-of-the-art models for catalysis. Baseline models include UMA[78] and eSEN[18], graph neural networks (GNNs) that operate on graphs where atoms are nodes and edges are the interactions between them. We evaluate models of different sizes as well as energy-conserving and direct-force models. We also evaluate the performance of fine-tuning from the latest energy-conserving UMA model.

Baseline results for all splits are evaluated using mean absolute errors (MAE) of energy and forces as primary metrics. Results across the different test sets are provided in Table 2. All evaluation sets are calculated using tighter DFT convergence criteria (EDIFF= 10^{-6} eV) to ensure more accurate force labels.

3.1 Evaluations

Across all splits, energy conserving (cons.) models outperform direct (d.) models in both energy and force predictions. Across model sizes, eSEN-M-d. achieves better results on the Test split compared to both eSEN-S-d. and eSEN-S-cons., while on OOD Solvent and Both splits, eSEN-S-cons. has better energy performance. Generally, models demonstrate competitive performance, with energy and force errors as low as $0.10~{\rm eV}$ and $0.015~{\rm eV/\mathring{A}}$ for eSEN-S-cons.

Table 2 Baseline results across the different **test** splits for different graph neural network models defined in the text. Energy and force mean absolute errors (MAE) are reported in units of eV and eV/ \mathring{A} . Validation results are provided in the Appendix.

			Τe	est	OOD S	Solvent	OOE	Ion	OOD	Both	Solvation
Dataset	Model	# of params	Energy	Forces	Energy	Forces	Energy	Forces	Energy	Forces	Energy
OC25	eSEN-S-d. eSEN-S-cons. eSEN-M-d.	6.3M 6.3M 50.7M	0.138 0.105 0.060	0.020 0.015 0.009	0.351 0.175 0.238	0.047 0.035 0.023	0.216 0.143 0.122	0.035 0.026 0.018	0.389 0.186 0.264	0.052 0.038 0.026	0.060 0.045 0.040
UMA UMA→OC25	UMA-S-1.1 UMA-S-ft	146.6M 146.6M	0.091	$0.064 \\ 0.014$	0.201	$0.101 \\ 0.036$	0.148	$0.090 \\ 0.027$	0.225	$0.108 \\ 0.039$	0.169 0.136

When evaluating UMA (UMA-S-1.1) using the OC20 task, results are worse in both force and solvation energy metrics relative to all eSEN models; a direct comparison of energy metrics is not made due to energy referencing between the levels of theory. UMA fine-tuned on OC25 (UMA-S-ft) achieves results comparable to eSEN-S-cons. for the Test split, though its performance drops on OOD splits. Similar trends are observed for solvation energy, where eSEN-M-d. outperforms all models with an energy MAE of 0.04 eV. Solvation energy errors are generally lower than the corresponding test energy errors across all models, suggesting a potential cancellation of errors when calculating relative properties [1]. UMA models lag behind in solvation energy predictions, with UMA-S-ft reporting a notably higher error of 0.136 eV. Appendix B.3 provides a breakdown of solvation energy error components, with the vacuum term as the main source of error in UMA-S-ft.

3.2 Force convergence

Initially, we trained an eSEN-S-cons. model on the original OC25 dataset (EDIFF= 10^{-4} eV, no drift filtering) and observed considerable force errors in the validation set (Figure 5c), with notable outliers that impact the overall MAE. However, using the same trained model, we then evaluated the same validation set calculated using tighter settings (EDIFF= 10^{-6} eV). The results improved significantly, with force errors dropping from 0.053 eV/Å to 0.016 eV/Å (Figure 5d). Energy errors remained nearly identical, indicating adequate convergence even using the original settings (Figure 5a,b). To our surprise, this suggests that the models are in fact capable of reconciling some numerical noise in the training data, and is somewhat at odds with conventional wisdom that any noise in the training data will propagate to reduced model performance.

Despite these observations, to ensure that the training data are consistent for broader use, we filter the dataset based on a total drift threshold of 1 eV/Å following the results shown in Figure 4. When eSEN-S-cons. is then used to evaluate performance on the original and tighter settings of this filtered set, the gap is much smaller with force errors of $0.0159~{\rm eV/Å}$ and $0.0144~{\rm eV/Å}$, respectively (see Appendix Figure 7). All released validation and test sets are calculated at tighter settings to ensure proper evaluation.

4 Outlook and future directions

OC25 constitutes the largest and most diverse solid-liquid interface dataset available, encompassing a broad spectrum of unique surfaces, solvents, ions, and adsorbate combinations. Despite its extensive coverage of solid-liquid interfaces, OC25 has inherent limitations. Although the systems included are approximately twice as large as those in OC20/OC22, they still fall short of the very large cell sizes needed to rigorously capture bulk-like solvent effects and lower ion concentrations that are often utilized in experiments. The dataset features average ion concentrations of \sim 2M, with a minimum of 0.38M. This is substantially more concentrated than many experiments in electrocatalysis; however, we also note that under reaction conditions and applied potentials, ion concentrations will generally be higher near the (charged) interface compared to the bulk electrolyte. Solvent depths average around 5.6Å, with a maximum of 10Å, corresponding to only a limited number of solvent layers. Although the models in this work may still be able to accurately predict interfacial properties and reactivity, these aspects remain to be tested in future studies. During the generation of this dataset, spin polarization was disabled. As a result, materials for which magnetism is an important consideration may not be accurately represented, and may require additional fine-tuning.

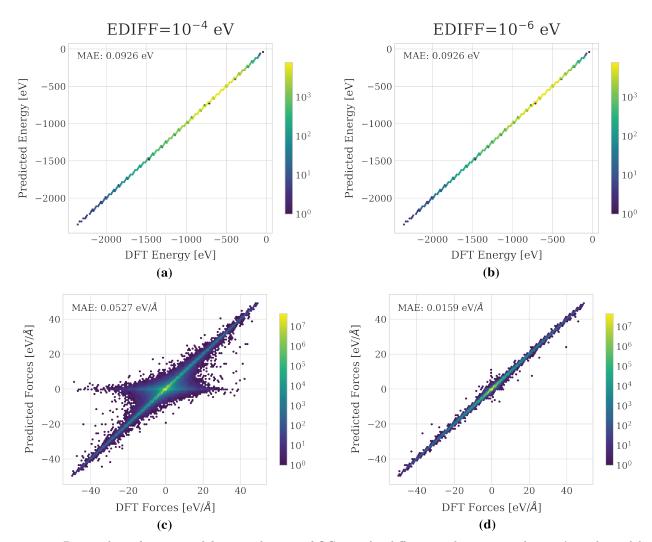


Figure 5 Parity plots of energy and force predictions of OC25 under different evaluation paradigms. A single model is trained on the unfiltered OC25 dataset and evaluated on an identical validation set calculated with the original (EDIFF= 10^{-4} eV) and tighter (EDIFF= 10^{-6} eV) settings. Despite training on the original settings, models are still able to accurately predict the correct force labels.

An important quantity that is currently not explored in this work is the interface workfunction, which can be related to the electrochemical potential measured in experiments. Estimating the workfunction for surface/electrolyte/vacuum interfaces involves evaluating two quantities: the Fermi level and the vacuum potential above the electrolyte. Although interface workfunctions can be estimated by performing single-point DFT evaluations on representative configurations from molecular dynamics simulations using the OC25 models, it would be highly desirable to have ML models that can predict the interface workfunctions directly. Access to the interface workfunction during the simulation can also enable constant potential (grand-canonical) simulations to estimate the thermodynamics and kinetics of electrochemical reactions. Developing models that can accurately predict both the Fermi level and the vacuum potential, and in turn the interfacial workfunction, is an exciting direction for future research in the study of solid-liquid interfaces.

Models trained on OC25 achieve strong predictive accuracy, with energy and force errors reaching as low as 0.105 eV and 0.015 eV/Å, respectively for the eSEN-S-cons model. Out-of-distribution errors are larger for unseen solvents and ions, providing opportunities for improving model generalizability. Novel modeling approaches to better account for long-range interactions and charge can be especially useful for this dataset due to the presence of electrolytes.

The curation of OC25 provided valuable insights into the convergence and consistency of DFT-calculated

force labels and its impacts on MLIP training. In practice, ensuring that the net force across the system is zero is particularly important in training models directly on AIMD data, as some DFT codes (e.g., VASP) do not automatically remove this drift. Additionally, we observe that filtering out calculations where the total force drift exceeds acceptable thresholds can improve data quality. For OC25, this threshold was chosen to be 1 eV/Å, but we encourage future efforts to perform their own analysis. Perhaps most surprisingly, models trained on the unfiltered, less-converged data were able to predict the more tightly converged set of forces very accurately, suggesting that models possess a notable resilience to moderate label noise.

Building on the success of OC20, OC25 introduces explicit modeling of the solid-liquid interface in catalytic systems. Covering an even broader range of chemical diversity, the OC25 dataset, along with its baseline models, hopes to serve as a valuable resource for the broader scientific community.

5 Acknowledgements

N. G. acknowledges support from a startup grant at NTU (award number 024462-00001).

M.M. and J.A.G. gratefully acknowledge support from The Welch Foundation under Grant Number D-2188-20240404.

J.B.V. acknowledges support by the U.S. Department of Energy, Lawrence Livermore National Laboratory (LLNL) under Contract No. DE-AC52-07NA27344.

References

- [1] Kareem Abdelmaqsoud, Muhammed Shuaibi, Adeesh Kolluru, Raffaele Cheula, and John R Kitchin. Investigating the error imbalance of large-scale machine learning potentials in catalysis. *Catalysis Science & Technology*, 14(20): 5899–5908, 2024.
- [2] Jehad Abed, Jiheon Kim, Muhammed Shuaibi, Brook Wander, Boris Duijf, Suhas Mahesh, Hyeonseok Lee, Vahe Gharakhanyan, Sjoerd Hoogland, Erdem Irtem, et al. Open catalyst experiments 2024 (ocx24): Bridging experiments and computational models. arXiv preprint arXiv:2411.11783, 2024.
- [3] Zhongchao Bai, Qian Yao, Mingyue Wang, Weijia Meng, Shixue Dou, Hua kun Liu, and Nana Wang. Low-temperature sodium-ion batteries: Challenges and progress. *Advanced Energy Materials*, 14(17):2303788, 2024.
- [4] Magda H Barecka and Joel W Ager. Towards an accelerated decarbonization of the chemical industry by electrolysis. *Energy Advances*, 2(2):268–279, 2023.
- [5] Luis Barroso-Luque, Muhammed Shuaibi, Xiang Fu, Brandon M Wood, Misko Dzamba, Meng Gao, Ammar Rizvi, C Lawrence Zitnick, and Zachary W Ulissi. Open materials 2024 (omat24) inorganic materials dataset and models. arXiv preprint arXiv:2410.12771, 2024.
- [6] Ilyes Batatia, Dávid Péter Kovács, Gregor N. C. Simm, Christoph Ortner, and Gábor Csányi. Mace: Higher order equivariant message passing neural networks for fast and accurate force fields. arXiv preprint arXiv:2206.07697v2, 2023.
- [7] Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E Smidt, and Boris Kozinsky. E (3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature Communications*, 13(1):2453, 2022.
- [8] Lennart Bengtsson. Dipole correction for surface supercell calculations. Physical Review B, 59(19):12301, 1999.
- [9] Jordi Cabana, Thomas Alaan, George W Crabtree, Po-Wei Huang, Akash Jain, Megan Murphy, Jeanne N'Diaye, Kasinath Ojha, George Agbeworvi, Helen Bergstrom, et al. Ngene 2022: Electrochemistry for decarbonization. ACS Energy Letters, 8(1):740-747, 2022.
- [10] Xiaosheng Cai, Yingying Yue, Zheng Yi, Junfei Liu, Yangping Sheng, and Yuhao Lu. Challenges and industrial perspectives on the development of sodium ion batteries. *Nano Energy*, 129:110052, 2024.
- [11] Lowik Chanussot, Abhishek Das, Siddharth Goyal, Thibaut Lavril, Muhammed Shuaibi, Morgane Riviere, Kevin Tran, Javier Heras-Domingo, Caleb Ho, Weihua Hu, et al. Open catalyst 2020 (oc20) dataset and community challenges. ACS Catalysis, 11(10):6059–6072, 2021.
- [12] Xiaoting Chen, Ian T McCrum, Kathleen A Schwarz, Michael J Janik, and Marc TM Koper. Co-adsorption of cations as the cause of the apparent ph dependence of hydrogen adsorption on a stepped platinum single-crystal electrode. *Angewandte Chemie International Edition*, 56(47):15025–15029, 2017.
- [13] Minju Chung, Joseph H Maalouf, Jason S Adams, Chenyu Jiang, Yuriy Román-Leshkov, and Karthish Manthiram. Direct propylene epoxidation via water activation over pd-pt electrocatalysts. *Science*, 383(6678):49–55, 2024.
- [14] Bowen Deng, Peichen Zhong, KyuJung Jun, Janosh Riebesell, Kevin Han, Christopher J Bartel, and Gerbrand Ceder. Chgnet as a pretrained universal neural network potential for charge-informed atomistic modelling. *Nature Machine Intelligence*, 5(9):1031–1041, 2023.
- [15] Hao Du, Yadong Wang, Yuqiong Kang, Yun Zhao, Yao Tian, Xianshu Wang, Yihong Tan, Zheng Liang, John Wozny, Tao Li, et al. Side reactions/changes in lithium-ion batteries: mechanisms and strategies for creating safer and better batteries. Advanced Materials, 36(29):2401482, 2024.
- [16] Peter Eastman, Pavan Kumar Behara, David L Dotson, Raimondas Galvelis, John E Herr, Josh T Horton, Yuezhi Mao, John D Chodera, Benjamin P Pritchard, Yuanqing Wang, et al. Spice, a dataset of drug-like molecules and peptides for training machine learning potentials. Scientific Data, 10(1):11, 2023.
- [17] Xianbiao Fu, Valerie A Niemann, Yuanyuan Zhou, Shaofeng Li, Ke Zhang, Jakob B Pedersen, Mattia Saccoccio, Suzanne Z Andersen, Kasper Enemark-Rasmussen, Peter Benedek, et al. Calcium-mediated nitrogen reduction for electrochemical ammonia synthesis. *Nature Materials*, 23(1):101–107, 2024.
- [18] Xiang Fu, Brandon M Wood, Luis Barroso-Luque, Daniel S Levine, Meng Gao, Misko Dzamba, and C Lawrence Zitnick. Learning smooth and expressive interatomic potentials for physical property prediction. arXiv preprint arXiv:2502.12147, 2025.

- [19] Johannes Gasteiger, Muhammed Shuaibi, Anuroop Sriram, Stephan Günnemann, Zachary Ulissi, C Lawrence Zitnick, and Abhishek Das. Gemnet-oc: developing graph neural networks for large and diverse molecular simulation datasets. arXiv preprint arXiv:2204.02782, 2022.
- [20] Ashrumochan Gouda, Devendra Sharma, Ashish Kumar, and Venkata Krishnan. Green hydrogen production: From lab scale to pilot scale photocatalysis. In *Towards Sustainable and Green Hydrogen Production by Photocatalysis: Scalability Opportunities and Challenges (Volume 1)*, pages 185–210. ACS Publications, 2024.
- [21] Nitish Govindarajan, Georg Kastlunger, Joseph A. Gauthier, Jun Cheng, Ivo Filot, Arthur Hagopian, Heine Anton Hansen, Jun Huang, Piotr M. Kowalski, Jinwen Liu, Juan M. Lombardi, Mikael Maraschin, Andrew Peterson, Hemanth S. Pillai, Hector Prats, Conor J. Price, René van Roij, Jan Rossmeisl, Ranga Rohit Seemakurthi, Seung-Jae Shin, Audrey Smith, Jia-Xin Zhu, and Katharina Doblhoff-Dier. The intricacies of computational electrochemistry. ACS Energy Letters, 10(9):4277–4288, 2025.
- [22] Ishita Goyal, Nishithan C Kani, Samuel A Olusegun, Sreenivasulu Chinnabattigalla, Rajan R Bhawnani, Ksenija D Glusac, Aayush R Singh, Joseph A Gauthier, and Meenesh R Singh. Metal nitride as a mediator for the electrochemical synthesis of nh₃. ACS Energy Letters, 9(8):4188–4195, 2024.
- [23] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu. *The Journal of Chemical Physics*, 132 (15):154104, 2010.
- [24] Axel Groß and Sung Sakong. Ab initio simulations of water/metal interfaces. *Chemical Reviews*, 122(12): 10746–10776, 2022.
- [25] J. Guo, M. Calegari Andrade, C. Hahn, A. Kulkarni, and N. Govindarajan. Understanding cation and surface charging effects at electrified interfaces using neural network interatomic potentials. *ChemRxiv*, 2025.
- [26] Bjørk Hammer, Lars Bruno Hansen, and Jens Kehlet Nørskov. Improved adsorption energetics within densityfunctional theory using revised perdew-burke-ernzerhof functionals. *Physical Review B*, 59(11):7413, 1999.
- [27] Xi Hao, Weihua Song, Yinghui Wang, Jieling Qin, and Zhenqi Jiang. Recent advancements in electrochemical sensors based on mofs and their derivatives. Small, 21(4):2408624, 2025.
- [28] Deiaa M. Harraz, Kunal M. Lodaya, Bryan Y. Tang, and Yogesh Surendranath. Homogeneous-heterogeneous bifunctionality in pd-catalyzed vinyl acetate synthesis. *Science*, 388(6742):eads7913, 2025.
- [29] Hendrik H Heenen, Joseph A Gauthier, Henrik H Kristoffersen, Thomas Ludwig, and Karen Chan. Solvation at metal/water interfaces: An ab initio molecular dynamics benchmark of common computational approaches. The Journal of Chemical Physics, 152(14), 2020.
- [30] Lukas Hörmann, Wojciech G Stark, and Reinhard J Maurer. Machine learning and data-driven methods in computational surface and interface science. npj Computational Materials, 11(1):196, 2025.
- [31] Haldrian Iriawan, Suzanne Z Andersen, Xilun Zhang, Benjamin M Comer, Jesús Barrio, Ping Chen, Andrew J Medford, Ifan EL Stephens, Ib Chorkendorff, and Yang Shao-Horn. Methods for nitrogen activation by reduction and oxidation. *Nature Reviews Methods Primers*, 1(1):56, 2021.
- [32] Anubhav Jain, Joseph Montoya, Shyam Dwaraknath, Nils ER Zimmermann, John Dagdelen, Matthew Horton, Patrick Huck, Donny Winston, Shreyas Cholia, Shyue Ping Ong, et al. The materials project: Accelerating materials design through theory-driven data and tools. In *Handbook of Materials Modeling: Methods: Theory and Modeling*, pages 1751–1784. Springer, 2020.
- [33] Aidan Klemm, Stephen P Vicchio, Sanchari Bhattacharjee, Eda Cagli, Yensil Park, Muhammad Zeeshan, Ruth Dikki, Harrison Liu, Michelle K Kidder, Rachel B Getman, et al. Impact of hydrogen bonds on co2 binding in eutectic solvents: an experimental and computational study toward sorbent design for co2 capture. ACS Sustainable Chemistry & Engineering, 11(9):3740–3749, 2023.
- [34] Georg Kresse and Jürgen Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science*, 6(1):15–50, 1996.
- [35] Georg Kresse and Jürgen Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Review B*, 54(16):11169–11186, 1996.
- [36] Georg Kresse and Jürgen Hafner. Ab initio molecular-dynamics simulation of the liquid-metal-amorphoussemiconductor transition in germanium. *Physical Review B*, 49(20):14251–14269, 1994.

- [37] Georg Kresse and Daniel Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. Physical Review B, 59(3):1758, 1999.
- [38] Ambarish Kulkarni, Samira Siahrostami, Anjli Patel, and Jens K. Nørskov. Understanding catalytic activity trends in the oxygen reduction reaction. Chemical Reviews, 118(5):2302-2312, 2018.
- [39] Subrata Kumar Kundu, Muhammad Zeeshan, Panuwat Watthaisong, and Andreas Heyden. Liquid phase modeling in porous media: Adsorption of methanol and ethanol in h-mfi in condensed water. *Journal of Chemical Theory and Computation*, 21(12):6121–6134, 2025.
- [40] Janice Lan, Aini Palizhati, Muhammed Shuaibi, Brandon M Wood, Brook Wander, Abhishek Das, Matt Uyttendaele, C Lawrence Zitnick, and Zachary W Ulissi. Adsorbml: A leap in efficiency for adsorption energy calculations using generalizable machine learning potentials. npj Computational Materials, 9(1):172, 2023.
- [41] Nikifar Lazouski, Zachary J Schiffer, Kindle Williams, and Karthish Manthiram. Understanding continuous lithium-mediated electrochemical nitrogen reduction. *Joule*, 3(4):1127–1139, 2019.
- [42] Zachary Levell, Jiabo Le, Saerom Yu, Ruoyu Wang, Sudheesh Ethirajan, Rachita Rana, Ambarish Kulkarni, Joaquin Resasco, Deyu Lu, Jun Cheng, and Yuanyue Liu. Emerging atomistic modeling methods for heterogeneous electrocatalysis. *Chemical Reviews*, 124(14):8620–8656, 2024.
- [43] Daniel S Levine, Muhammed Shuaibi, Evan Walter Clark Spotte-Smith, Michael G Taylor, Muhammad R Hasyim, Kyle Michel, Ilyes Batatia, Gábor Csányi, Misko Dzamba, Peter Eastman, et al. The open molecules 2025 (omol25) dataset, evaluations, and models. arXiv preprint arXiv:2505.08762, 2025.
- [44] Noah B. Lewis, Ryan P. Bisbey, Karl S. Westendorff, Alexander V. Soudackov, and Yogesh Surendranath. A molecular-level mechanistic framework for interfacial proton-coupled electron transfer kinetics. *Nature Chemistry*, 16(3):343–352, 2024.
- [45] Yixuan Li, Liuxiong Luo, Yingqi Kong, Yujia Li, Quansheng Wang, Mingqing Wang, Ying Li, Andrew Davenport, and Bing Li. Recent advances in molecularly imprinted polymer-based electrochemical sensors. *Biosensors and Bioelectronics*, 249:116018, 2024.
- [46] Yi-Lun Liao, Brandon Wood, Abhishek Das, and Tess Smidt. Equiformerv2: Improved equivariant transformer for scaling to higher-degree representations. arXiv preprint arXiv:2306.12059, 2023.
- [47] Yanshuo Liu, Qiang Li, and Kai Wang. Revealing the degradation patterns of lithium-ion batteries from impedance spectroscopy using variational auto-encoders. *Energy Storage Materials*, 69:103394, 2024.
- [48] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. arXiv preprint arXiv:1711.05101, 2017.
- [49] Dharik S Mallapragada, Yury Dvorkin, Miguel A Modestino, Daniel V Esposito, Wilson A Smith, Bri-Mathias Hodge, Michael P Harold, Vincent M Donnelly, Alice Nuz, Casey Bloomquist, et al. Decarbonization of the chemical industry through electrification: Barriers and opportunities. *Joule*, 7(1):23–41, 2023.
- [50] Leandro Martínez, Ricardo Andrade, Ernesto G Birgin, and José Mario Martínez. Packmol: A package for building initial configurations for molecular dynamics simulations. *Journal of Computational Chemistry*, 30(13): 2157–2164, 2009.
- [51] Tristan Maxson and Tibor Szilvási. Transferable water potentials using equivariant neural networks. *The Journal of Physical Chemistry Letters*, 15(14):3740–3747, 2024.
- [52] Tristan Maxson, Ademola Soyemi, Benjamin WJ Chen, and Tibor Szilvási. Enhancing the quality and reliability of machine learning interatomic potentials through better reporting practices. *The Journal of Physical Chemistry C*, 128(16):6524–6537, 2024.
- [53] Andrew J. Medford, Aleksandra Vojvodic, Jens S. Hummelshøj, Johannes Voss, Frank Abild-Pedersen, Felix Studt, Thomas Bligaard, Anders Nilsson, and Jens K. Nørskov. From the sabatier principle to a predictive theory of transition-metal heterogeneous catalysis. *Journal of Catalysis*, 328:36–42, 2015. Special Issue: The Impact of Haldor Topsøe on Catalysis.
- [54] Rui Kai Miao, Ning Wang, Sung-Fu Hung, Wen-Yang Huang, Jinqiang Zhang, Yong Zhao, Pengfei Ou, Sasa Wang, Jonathan P Edwards, Cong Tian, et al. Electrified cement production via anion-mediated electrochemical calcium extraction. ACS Energy Letters, 8(11):4694-4701, 2023.
- [55] Jongsu Noh, Hanjoo Kim, Hyein Park, and Dong Young Chung. The roles of ions in electrochemical interface for electrocatalysis. ACS Catalysis, 15(10):7780–7791, 2025.

- [56] J. K. Nørskov, T. Bligaard, J. Rossmeisl, and C. H. Christensen. Towards the computational design of solid catalysts. *Nature Chemistry*, 1(1):37–46, 2009.
- [57] Jens K Nørskov, Frank Abild-Pedersen, Felix Studt, and Thomas Bligaard. Density functional theory in surface chemistry and catalysis. Proceedings of the National Academy of Sciences, 108(3):937–943, 2011.
- [58] Gbolagade Olajide, Khagendra Baral, Sophia Ezendu, Ademola Soyemi, and Tibor Szilvasi. Application of machine learning interatomic potentials in heterogeneous catalysis. *Journal of Catalysis*, 448:116202, 2025.
- [59] Samuel A Olusegun, Yancun Qi, Nishithan C Kani, Meenesh R Singh, and Joseph A Gauthier. Understanding activity trends in electrochemical dinitrogen oxidation over transition metal oxides. ACS Catalysis, 14(22): 16885–16896, 2024.
- [60] Vincent J Ovalle, Yu-Shen Hsu, Naveen Agrawal, Michael J Janik, and Matthias M Waegele. Correlating hydration free energy and specific adsorption of alkali metal cations during co2 electroreduction on au. *Nature Catalysis*, 5 (7):624–632, 2022.
- [61] Yang Qiu, Debmalya Ray, Litao Yan, Xiaohong Li, Miao Song, Mark H Engelhard, Junming Sun, Mal-Soon Lee, Xin Zhang, Manh-Thuong Nguyen, et al. Proton relay for the rate enhancement of electrochemical hydrogen reactions at heterogeneous interfaces. *Journal of the American Chemical Society*, 145(48):26016–26027, 2023.
- [62] Raghunathan Ramakrishnan, Pavlo O Dral, Matthias Rupp, and O Anatole Von Lilienfeld. Quantum chemistry structures and properties of 134 kilo molecules. *Scientific Data*, 1(1):1–7, 2014.
- [63] Stefan Ringe, Nicolas G. Hörmann, Harald Oberhofer, and Karsten Reuter. Implicit solvation methods for catalysis at electrified interfaces. Chemical Reviews, 122(12):10777–10820, 2022.
- [64] Mohammad Saleheen, Osman Mamun, Anand Mohan Verma, Dia Sahsah, and Andreas Heyden. Understanding the influence of solvents on the pt-catalyzed hydrodeoxygenation of guaiacol. *Journal of Catalysis*, 425:212–232, 2023.
- [65] SandboxAQ. Aqcat25 dataset. https://huggingface.co/datasets/SandboxAQ/aqcat25, 2025.
- [66] Zachary J Schiffer and Karthish Manthiram. Electrification and decarbonization of the chemical industry. *Joule*, 1(1):10–14, 2017.
- [67] Rimmy Singh, Ruchi Gupta, Deepak Bansal, Rachna Bhateria, and Mona Sharma. A review on recent trends and future developments in electrochemical sensing. ACS Omega, 9(7):7336-7356, 2024.
- [68] Anuroop Sriram, Sihoon Choi, Xiaohan Yu, Logan M Brabson, Abhishek Das, Zachary Ulissi, Matt Uyttendaele, Andrew J Medford, and David S Sholl. The open dac 2023 dataset and challenges for sorbent discovery in direct air capture. ACS Central Science, 10(5):923-941, 2024.
- [69] Anuroop Sriram, Logan M Brabson, Xiaohan Yu, Sihoon Choi, Kareem Abdelmaqsoud, Elias Moubarak, Pim de Haan, Sindy Löwe, Johann Brehmer, John R Kitchin, et al. The open dac 2025 dataset for sorbent discovery in direct air capture. arXiv preprint arXiv:2508.03162, 2025.
- [70] Ravishankar Sundararaman, Derek Vigil-Fowler, and Kathleen Schwarz. Improving the accuracy of atomistic simulations of the electrochemical interface. *Chemical Reviews*, 122(12):10651–10674, 2022.
- [71] Meng Tao, Joseph A Azzolini, Ellen B Stechel, Katherine E Ayers, and Thomas I Valdez. Engineering challenges in green hydrogen production systems. *Journal of The Electrochemical Society*, 169(5):054503, 2022.
- [72] Kevin Tran and Zachary W Ulissi. Active learning across intermetallics to guide discovery of electrocatalysts for co2 reduction and h2 evolution. *Nature Catalysis*, 1(9):696–703, 2018.
- [73] Richard Tran, Janice Lan, Muhammed Shuaibi, Brandon M Wood, Siddharth Goyal, Abhishek Das, Javier Heras-Domingo, Adeesh Kolluru, Ammar Rizvi, Nima Shoghi, et al. The open catalyst 2022 (oc22) dataset and challenges for oxide electrocatalysts. ACS Catalysis, 13(5):3066–3084, 2023.
- [74] Brook Wander, Muhammed Shuaibi, John R Kitchin, Zachary W Ulissi, and C Lawrence Zitnick. Cattsunami: Accelerating transition state energy calculations with pretrained graph neural networks. *ACS Catalysis*, 15(7): 5283–5294, 2025.
- [75] Jiangjiang Wang, Gangfeng Wu, Guanghui Feng, Guihua Li, Yiheng Wei, Shoujie Li, Jianing Mao, Xiaohu Liu, Aohui Chen, Yanfang Song, et al. Electrochemical epoxidation of propylene to propylene oxide via halogen-mediated systems. ACS Omega, 8(49):46569–46576, 2023.

- [76] Rui Wang, Lu Wang, Rui Liu, Xiangye Li, Youzhi Wu, and Fen Ran. "fast-charging" anode materials for lithium-ion batteries from perspective of ion diffusion in crystal structure. ACS Nano, 18(4):2611–2648, 2024.
- [77] Kirsten T Winther, Max J Hoffmann, Jacob R Boes, Osman Mamun, Michal Bajdich, and Thomas Bligaard. Catalysis-hub. org, an open electronic structure database for surface reactions. *Scientific Data*, 6(1):75, 2019.
- [78] Brandon M Wood, Misko Dzamba, Xiang Fu, Meng Gao, Muhammed Shuaibi, Luis Barroso-Luque, Kareem Abdelmaqsoud, Vahe Gharakhanyan, John R Kitchin, Daniel S Levine, et al. Uma: A family of universal models for atoms. arXiv preprint arXiv:2506.23971, 2025.
- [79] Simson Wu, Nicholas Salmon, Molly Meng-Jung Li, René Bañares-Alcántara, and Shik Chi Edman Tsang. Energy decarbonization via green h₂ or nh₃? ACS Energy Letters, 7(3):1021–1033, 2022.
- [80] Rong Xia, Sean Overa, and Feng Jiao. Emerging electrochemical processes to decarbonize the chemical industry. JACS Au, 2(5):1054–1070, 2022.
- [81] Saerom Yu, Zachary Levell, Zhou Jiang, Xunhua Zhao, and Yuanyue Liu. What is the rate-limiting step of oxygen reduction reaction on fe-n-c catalysts? *Journal of the American Chemical Society*, 145(46):25352–25356, 2023.
- [82] Fang Zhang, Bijiao He, Yan Xin, Tiancheng Zhu, Yuning Zhang, Shuwei Wang, Weiyi Li, Yang Yang, and Huajun Tian. Emerging chemistry for wide-temperature sodium-ion batteries. Chemical Reviews, 124(8):4778–4821, 2024.
- [83] Peng Zhang, Tuo Wang, and Jinlong Gong. Advances in electrochemical oxidation of olefins to epoxides. CCS Chemistry, 5(5):1028–1042, 2023.
- [84] Xiaohong Zhang, Aditya Savara, and Rachel B Getman. A method for obtaining liquid–solid adsorption rates from molecular dynamics simulations: applied to methanol on pt (111) in h₂o. *Journal of Chemical Theory and Computation*, 16(4):2680–2691, 2020.
- [85] Zishuai Zhang, Aubry SR Williams, Shaoxuan Ren, Benjamin AW Mowbray, Colin TE Parkyn, Yongwook Kim, Tengxiao Ji, and Curtis P Berlinguette. Electrolytic cement clinker precursor production sustained through orthogonalization of ion vectors. Energy & Environmental Science, 18(5):2395–2404, 2025.
- [86] Xunhua Zhao and Yuanyue Liu. Origin of selective production of hydrogen peroxide by electrochemical oxygen reduction. *Journal of the American Chemical Society*, 143(25):9423–9428, 2021.
- [87] Jia-Xin Zhu and Jun Cheng. Machine learning potential for electrochemical interfaces with hybrid representation of dielectric response. *Physical Review Letters*, 135:018003, 2025.
- [88] Yong-Bin Zhuang, Chang Liu, Jia-Xin Zhu, Jin-Yuan Hu, Jia-Bo Le, Jie-Qiong Li, Xiao-Jian Wen, Xue-Ting Fan, Mei Jia, Xiang-Ying Li, et al. An artificial intelligence accelerated ab initio molecular dynamics dataset for electrochemical interfaces. *Scientific Data*, 12(1):997, 2025.

Appendix

A Model training

Baseline model and training hyperparameters followed the same procedures originally proposed in OMol25[43] and UMA[78]. Models were trained with the AdamW optimizer[48], a learning rate of 8e-4, and trained for 40 epochs. Direct models followed a multi-stage scheme, first trained on BF16 and then finetuned at FP32 with a learning rate of 4e-4. A per-atom MAE loss was used for energies and a L2-norm loss for forces, with energy and force coefficients of 10. UMA-S-1.1 was taken directly from the publicly released checkpoints at https://huggingface.co/facebook/UMA. UMA-S-1.1 was finetuned through the "oc20" task head, with all 32 experts available. All models were trained on Nvidia H100 80GB GPU cards. Model hyperparameters are provided in Table 3.

Table 3 Training and model hyperparameters for the baseline models trained in this work.

Hyperparameters	eSEN-S-d./cons.	eSEN-M	UMA-S-ft	
# sphere channels	128	128	128	
lmax	2	4	2	
mmax	2	2	2	
# moe experts	0	0	32	
max neighbors	30/300	30	300	
cutoff radius	6	6	6	
# edge channels	128	128	128	
distance function	gaussian	gaussian	gaussian	
# distance basis	64	128	64	
# layers	4	10	4	
# hidden channels	128	128	128	
learning rate	8e-4	8e-4	4e-4	
# gpus	32/64	32	64	
batch size ($\#$ atoms)	76800	44800	76800	
energy coeff.	10	10	10	
force coeff.	10	10	10	
# of params.	6.3M	50.7M	146.6M	

B Additional results

B.1 Validation

Baseline model results on the validation set are provided in Table 4.

B.2 Error by solvent and ion type

We breakdown model performance as a function of the different solvents and ions. Results are provided for eSEN-S-cons. in Figure 6.

Table 4 Baseline results for the validation split. Energy and force mean absolute errors (MAE) are reported in units of eV and eV/Å.

			Validation	
Dataset	Model	# of params	Energy	Forces
OC25	eSEN-S-d. eSEN-S-cons. eSEN-M-d.	6.3M 6.3M 50.7M	0.138 0.104 0.061	0.020 0.015 0.009
$\begin{array}{c} \hline {\rm UMA} \\ {\rm UMA} {\rightarrow} {\rm OC25} \end{array}$	UMA-S-1.1 UMA-S-ft	146.6M 146.6M	0.093	0.064 0.014

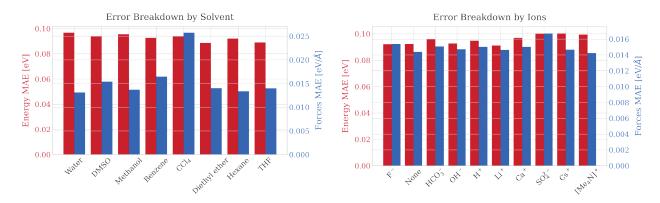


Figure 6 Energy and force mean absolute errors (MAE) broken down for the different solvent and ion types in the validation split. Results reported for the eSEN-S-cons. model.

Table 5 Baseline solvation energy results broken down across the different components. Where, $\Delta \tilde{E}_{solv}$ is the pseudo solvation energy, $\Delta \tilde{E}_{ads}^{solv}$ is the adsorption energy on the solid-liquid interface, and $\Delta \tilde{E}_{ads}^{vac}$ is the adsorption energy in vacuum.

			Energy MAE [eV]		
Dataset	Model	# of params	$\Delta \tilde{E_{solv}}$	$\Delta \tilde{E_{ads}^{solv}}$	$\Delta \tilde{E}^{vac}_{ads}$
OC25	eSEN-S-d. eSEN-S-cons. eSEN-M-d.	6.3M 6.3M 50.7M	0.060 0.045 0.040	0.071 0.057 0.041	0.075 0.057 0.050
$\overline{\text{UMA}}\\ \text{UMA}{\rightarrow}\text{OC25}$	UMA-S-1.1 UMA-S-ft	146.6M 146.6M	0.169 0.136	$0.520 \\ 0.053$	0.407 0.147

B.3 Solvation energy breakdown

For this work, we propose a pseudo solvation energy, $\Delta \tilde{E}_{solv}$, to evaluate baseline models across. Here, we generate a static solid-liquid interface configuration and construct the vacuum and reference configurations directly from this structure (i.e. deleting the solvent to generate the vacuum adsorbate+surface configuration). A break down of the different errors across the terms are provided in Table 5

$$\Delta \tilde{E_{solv}} = \Delta \tilde{E_{ads}}^{solv} - \Delta \tilde{E_{ads}}^{vac} \tag{2}$$

B.4 Force convergence plots on the filtered set

The final OC25 dataset ultimately filtered samples with a total drift <1 eV/Å to ensure the dataset is broadly useful beyond just MLIP training. Figure 7 corresponds to a model trained on the released OC25 dataset (converged with EDIFF= 10^{-4} eV, but with all images with total drift >1 eV/Å removed) and evaluated on the validation set with the same drift filtering. The small gap between force errors at EDIFF= 10^{-4} eV) and (EDIFF= 10^{-6} eV) suggests that the total drift filter does a reasonable job at removing problematic samples in the training set.

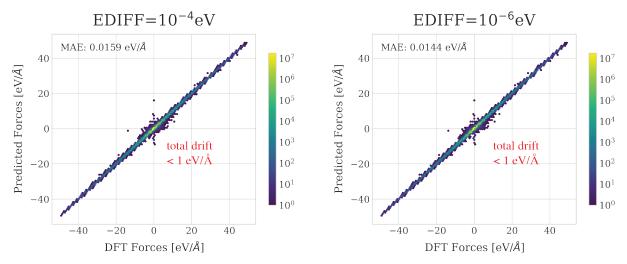


Figure 7 Parity plots of energy and force predictions of OC25 under different evaluation paradigms. A single model is trained on the filtered, released OC25 dataset and evaluated on an identical validation set calculated with the original (EDIFF= 10^{-4} eV) and tighter (EDIFF= 10^{-6} eV) settings.

C Structure generation

C.1 Additional adsorbates

A total of 98 adsorbates were sampled from to create OC25 configurations. These included all of the OC20 adsorbates[11] as well as adsorbates presented in OC20NEB[74] and OCx24[2]. These additional adsorbates are presented in Table 6. We refer readers to the OC20 paper for their full list of adsorbates.

Table 6 Additional adsorbates considered in OC25 alongside the full set of OC20 adsorbates.

Adsorbate class	Adsorbates
O/H Only	*OOH, *H ₂
C_1	*OCHO, *COOH, *OC*O
C_2	CO*COH, *CCOH, *CH ₂ CH ₂ *O, *CHCH ₂ *O, *COHCH ₂ *O, *CH ₂ OH*CH ₂ OH, *OCCHOH, *OCH ₂ CH ₂ OH, *OCH ₂ CH ₂ OH, *OCH ₂ CH ₂ *O, *OCH ₂ CHO, O*C*CO