

The Open Catalyst 2022 (OC22) Dataset and Challenges for Oxide Electrocatalysts

Richard Tran,[†] Janice Lan,[†] Muhammed Shuaibi,[†] Brandon M. Wood,[†] Siddharth Goyal,[†] Abhishek Das, Javier Heras-Domingo, Adeesh Kolluru, Ammar Rizvi, Nima Shoghi, Anuroop Sriram, Félix Therrien, Jehad Abed, Oleksandr Voznyy, Edward H. Sargent, Zachary Ulissi,^{*} and C. Lawrence Zitnick^{*}



Cite This: *ACS Catal.* 2023, 13, 3066–3084



Read Online

ACCESS |



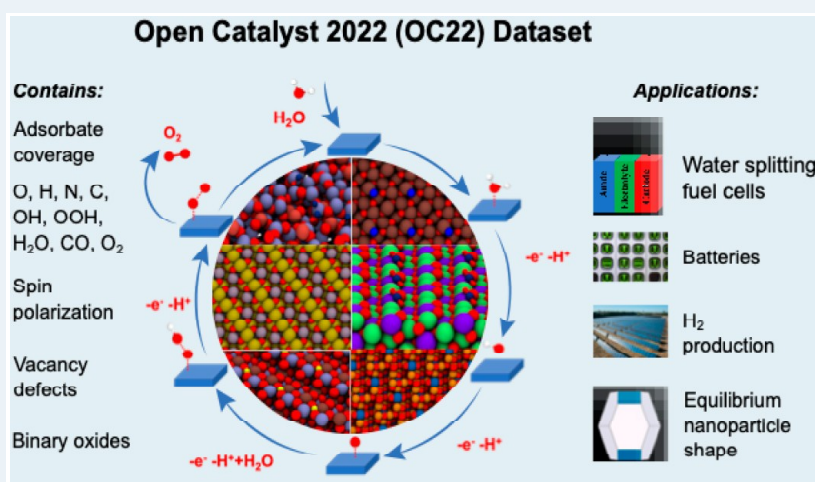
Metrics & More



Article Recommendations



Supporting Information



ABSTRACT: The development of machine learning models for electrocatalysts requires a broad set of training data to enable their use across a wide variety of materials. One class of materials that currently lacks sufficient training data is oxides, which are critical for the development of Oxygen Evolution Reaction (OER) catalysts. To address this, we developed the Open Catalyst 2022 (OC22) dataset, consisting of 62,331 Density Functional Theory (DFT) relaxations (~9,854,504 single point calculations) across a range of oxide materials, coverages, and adsorbates. We define generalized total energy tasks that enable property prediction beyond adsorption energies; we test baseline performance of several graph neural networks; and we provide predefined dataset splits to establish clear benchmarks for future efforts. In the most general task, GemNet-OC sees a ~36% improvement in energy predictions when combining the chemically dissimilar Open Catalyst 2020 Data set (OC20) and OC22 datasets via fine-tuning. Similarly, we achieved a ~19% improvement in total energy predictions on OC20 and a ~9% improvement in force predictions in OC22 when using joint training. We demonstrate the practical utility of a top performing model by capturing literature adsorption energies and important OER scaling relationships. We expect OC22 to provide an important benchmark for models seeking to incorporate intricate long-range electrostatic and magnetic interactions in oxide surfaces. Data set and baseline models are open sourced, and a public leaderboard is available to encourage continued community developments on the total energy tasks and data.

KEYWORDS: catalysis, oxides, renewable energy, datasets, machine learning, graph convolutions, force field

INTRODUCTION

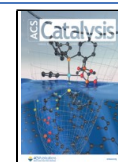
Advances are needed in technologies to produce, store, and use low-carbon-intensity energy. Renewable energy is often produced by intermittent sources (e.g., sunlight, wind, or tides), so efficient grid-scale storage is required to transfer power from times of excess generation to times of excess demand. There are a number of promising storage techniques including the conversion of renewable energy to a chemical form, e.g. water splitting to H₂, or CO₂ conversion to liquid fuels and high-value chemical feedstock. Inorganic oxides are

abundant electrocatalysts that are extensively used in these applications. However, the complex nature of oxide surfaces compared to simpler metals present a number of challenges to

Received: November 3, 2022

Revised: January 31, 2023

Published: February 16, 2023



Open Catalyst 2022 (OC22) Dataset

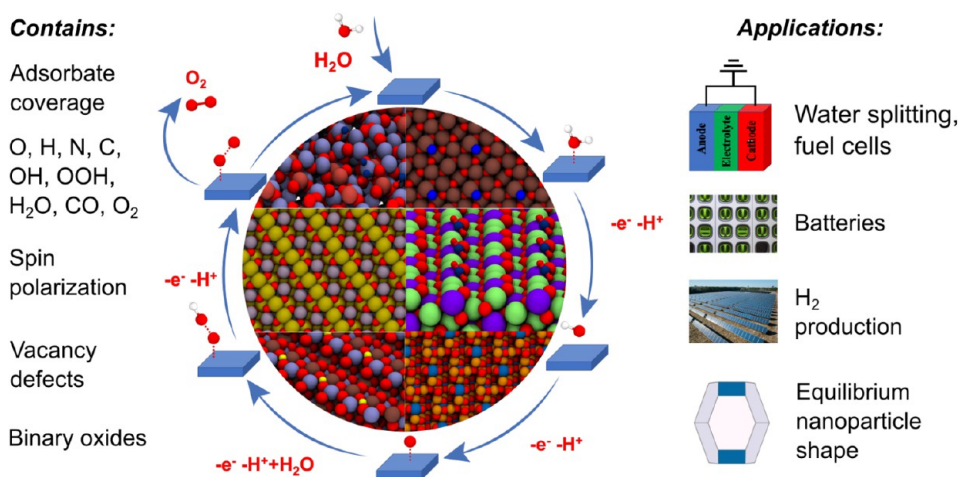
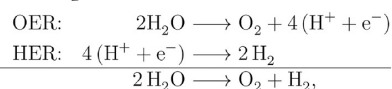


Figure 1. Overview of the contents and impact areas of the OC22 dataset. The water nucleophilic attack mechanism is highlighted for the OER reaction, with H_2O and O_2 as reactants and products, respectively. Inset images are a random sample of the dataset.

catalyst design. Developing generalizable machine learning methods to quickly and accurately predict the activity and stability of oxide catalysts would have a major impact on renewable energy storage and utilization.

As a motivating example of the need and challenges for oxide electrocatalysts, we consider water splitting for the generation of clean H_2 ; an energy-dense fuel that is used in fuel cells or ammonia synthesis. Electrochemical water splitting consists of two coupled half-reactions



which split two water molecules to evolve H_2 and O_2 gas. This process is extremely energy intensive. The OER overpotential is the larger contributor to the inefficiency of this reaction; it is quite complicated due to bond rearrangements and the formation of an O–O bond. Water splitting typically uses very harsh acidic conditions to reduce gas solubility and improve proton conductivity, and for which high performance proton exchange membranes are widely available. Unfortunately, for these conditions there are very few known materials that are stable and active, except extremely expensive metal oxides, such as those using Ir.¹ Currently, there are significant efforts to design complex multicomponent oxide OER catalysts to reduce the cost and improve their activity and stability.^{2,3} Computational chemistry can play a critical role in helping screen, discover, and understand such materials.

Computational methods can be used to predict the activity and stability of a proposed oxide catalyst, but these techniques are significantly more complicated than for metal catalysts and present many additional challenges. First, there are many oxide polymorphs (crystal structures) for any given chemical composition that must be considered to identify the most stable catalyst structure.⁴ Second, the surface of an oxide catalyst is often prone to reconstruction, leaching, doping, and defects.⁵ Third, the environment can lead to a number of possible surface terminations. Fourth, it is difficult to determine a catalyst's active site and there are often multiple competing mechanisms to consider.⁶ To add to these challenges, computational chemistry methods such as the

widely used Generalized Gradient Approximation (GGA) are less accurate for oxide materials due to the strong electron correlation and complicated electronic structure. Large system sizes and the likelihood of long-range electrostatic or magnetic interactions also result in slower convergence. These additional configurational and computational complexities make the creation of datasets and machine learning models for oxides significantly more expensive and challenging, leading to much fewer and smaller datasets than for metal systems (see⁷ for a sample of representative datasets in catalysis).

We address these challenges by generating a large oxide dataset to accelerate the development of machine learning (ML) models for materials design and discovery. The training of accurate and generalizable ML models requires large datasets. For example, the OC20 dataset⁸ (ca. 250 million single-point calculations) considered different adsorbates (small adsorbates, C1/C2 compounds, and N/O-containing intermediates) on top of randomly sampled low Miller index facets of stable materials from the Materials Project,⁹ but did not include metal oxide materials due to the complexities above. The release of the OC20 dataset helped enable rapid advances in the accuracy and generalizability of Graph Neural Network (GNN) models,¹⁰ with decreases of 55+% in the key S2EF metrics in the first two years. Initial baseline models like CGCNN¹¹ and SchNet¹² focused on local environment representations. Key advances since then include invariant angular interactions (DimeNet/DimeNet++^{13,14}), faster and more accurate but nonenergy conserving models (ForceNet¹⁵ and SpinConv¹⁶), and triple/quadruplet interactions (GemNet-dT,¹⁷ GemNet-XL,¹⁸ and GemNet-OC¹⁹). Other approaches include the use of transformers (3D-Graphormer²⁰) and more effective augmentation and learning strategies (Noisy-Nodes²¹). These and further advances are necessary to accurately predict properties of complex structures such as oxide systems.

In this work, we present the Open Catalyst 2022 (OC22) dataset (Figure 1) for the oxygen evolution reaction and oxide electrocatalysts more generally, as well as accompanying tasks and GNN baseline models. OC22 is intended to complement OC20, which did not contain any oxide materials, and further

enable the development of generalizable ML models for catalysis. This dataset spans the configurational complexity for oxide surfaces described above, including varying surface terminations, adsorbate+slab configurations and coverage, and nonstoichiometric substitutions and vacancies. To encompass the additional complexities in this dataset, we also expand on the primary tasks in OC20 to include the DFT total energy as a target. A more general property, DFT total energy offers the ability to address potential applications beyond those that just require simple adsorption energies.

With the creation of new datasets, the question arises of whether the data in them is complementary to other datasets for training ML models (see recent reviews for a perspective of catalysts informatics^{22–24}). This is especially important when consolidating data with a variety of computational methods in anthological dataset collections such as the Catalysis-Hub,²⁵ Catalyst Acquisition by Data Science,²⁶ and the NFDI4Cat consortium.²⁷ For instance, models can be trained jointly using multiple datasets, or transfer learning may be used to train a model on a larger dataset and fine-tuned on a smaller dataset. Recently, the OC20 dataset enabled the catalysis community to use transfer learning to improve model performance²⁸ on other smaller datasets. The small molecules and drug discovery communities have seen success in using transfer learning to transfer between varying levels of electronic structure calculations²⁹ or between related tasks.^{30–32} In this work, we explore the extent OC20 can aid OC22 via transfer learning or by jointly training on both datasets.

We train a variety of leading GNN models on two related proposed community challenges for OC22: (1) predict the DFT total energy and force for a given structure and (2) predict the DFT relaxed total energy given an initial structure. We also evaluate our models' performance on the established task of predicting the relaxed structure given an initial structure. The dataset is split into train/validation/test splits indicative of the situation commonly found in catalysis where the properties of unseen crystal compositions need to be predicted. Splits contain a combination of isolated surfaces (a.k.a slab) and surface with adsorbate (a.k.a adsorbate+slab) systems. All baseline models, data loaders and training scripts for each of these tasks are available at <https://github.com/Open-Catalyst-Project/ocp>. While we focus on a subset of tasks, models capable of solving these tasks on the OC22 dataset will likely be able to address numerous related catalysis problems.

The OC22 Dataset. OC22 is designed to provide a training dataset for constructing generalized models to aid in predicting catalytic reactions on oxide surfaces. To achieve this, we built the dataset in four stages: (1) bulk selection, (2) surface selection, (3) initial structure generation, and (4) structure relaxation. The dataset contains slabs and adsorbate+slabs, 19,142 and 43,189 systems, respectively. This resulted in 9,854,504 single-point calculations, each of which yielded forces and energies that were later partitioned into suitable train, validation, and test validation splits. We prioritized diversity in composition, surface termination, and adsorbate configurations in constructing our dataset to ensure that our models can generalize well. As a result of our emphasis on creating an unbiased and diverse dataset, OC22 structures may not always be the most stable or pertinent for a particular reaction pathway of interest—data still meaningful for building generalizable models. All source code used to generate the adsorbate configurations will be provided in the Open Catalyst

Data set repository at <https://github.com/Open-Catalyst-Project/Open-Catalyst-Dataset>.

Bulk Selection. We begin by confining our set of bulk oxide materials to 4,728 unary (A_xO_y) and binary ($A_xB_yO_z$) metal-oxides from the Materials Project⁹ where A and B are metals. These oxides can be composed of any combination of metals or semimetals listed in the [Supporting Information \(SI\)](#). In our list of 51 metals, Ce and Lu were the only lanthanides considered due to the utility of cerium-based oxide compounds in catalytic reactions^{33,34} and to add additional variety with lutetium-based oxides. For each chemical system, we considered bulk materials with the top five lowest energies above hull with less than 150 atoms to provide equal chemical distribution and diversity in our set of oxides. We note that under this criteria, some materials may exhibit an energy above hull exceeding 0.1 eV/atom (the threshold initially used in OC20). In addition to chemical diversity, we also included materials with a variety of electronic band gaps (E_G). [Table 1](#)

Table 1. Overview of the Chemical, Structural, and Adsorbate Composition of the Entire Dataset of Slabs and Adsorbate+Slabs

Chemical formula	
Unary (A_xO_y)	6,190
Binary ($A_xB_yO_z$)	56,141
Elements sampled	
Alkali	13,541
Alkaline	13, 974
p-block metals	14,029
Metalloids	8,292
Transition metals	48,561
Crystal structures	
Triclinic	6,214
Monoclinic	16,294
Orthorhombic	7,258
Tetragonal (Rutile)	11,550 (4,318)
Trigonal	4,411
Hexagonal	2,680
Cubic	9,606
Band gaps	
$E_G = 0$ eV	1,366
$0 \text{ eV} < E_G < 3.2$ eV	2,591
$E_G > 3.2$ eV	598
Adsorbates	
O	10,816
H	5,298
N	4,000
C	3,905
OH	4,092
OOH	4,424
H ₂ O	4,846
CO	3,994
O ₂	1,814
Calc. with PBE+U: 20,812	

lists the number of metallic ($E_G = 0$ eV), semiconducting ($0 \text{ eV} < E_G < 3.2$ eV), and insulating ($E_G > 3.2$ eV) materials considered in our dataset (all electronic properties were derived from the Materials Project). Many oxides such as TiO₂ are also useful for photocatalysis which typically requires semiconducting properties to allow for photoelectron

excitation. We also considered 173 unary and binary rutile structures.

Our selection criteria for bulk oxides prioritized chemical diversity over stability. We acknowledge that many of the materials we selected are not electrochemically stable which is a prerequisite for viable electrocatalytic materials. Pourbaix analysis have previously demonstrated that only oxides composed of 26 of the 51 elements we considered are relatively stable under aqueous conditions.³⁵

We also ignored the fact that certain chemical systems have a far greater set of distinct bulk structures than others. For instance, the Materials Project database has reported over 300 entries for chemical systems such as Ti–O and Mn–Li–O while no entries were reported for 200 chemical systems (see the SI). Other databases such as the Automatic-Flow³⁶ and Open Quantum Materials Database³⁷ have also made significant efforts in exploring oxides and contain chemical systems unexplored in the Materials Project. However, to ensure all oxides were obtained using a consistent methodology and open source licensing, we extracted entries from the Materials Project only.

Surface Selection. We constructed our dataset by first randomly sampling 4,286 bulk oxides from our original bulk oxide set of 4,728. We limited our dataset to slabs of less than 250 atoms. We construct each slab and adsorbate+slab using the process shown in Figure 2. Given a random oxide selected

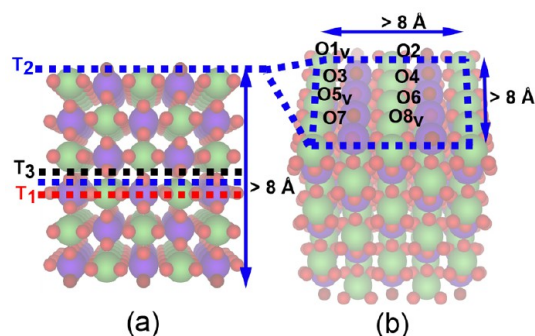


Figure 2. Construction of rutile (110) slabs and adsorbate+slabs. (a) Dashed lines indicate the different possible terminations (T_1 , T_2 , and T_3). The slab is symmetric about T_3 . (b) The T_2 terminated surface with its periodic boundary (blue dashed lines) contains 8 oxygen sites. Random removal of 3 surface oxygen (dark red) creates vacancy defects (transparent).

from our bulk dataset, we enumerate through all possible surface terminations with a maximum Miller index less than or equal to 3. As with Figure 2a all slabs are capped with the same terminating surface regardless of stoichiometry. We randomly select one termination which we replicated to a depth of at least 8 Å and a width in each cross-sectional direction of at least 8 Å.

Next we decorated the surface of the slab with a random number of oxygen vacancies which can act as active sites for reactions such as CO_2 capture³⁸ and OER.^{39,40} To do so, we first identify all existing oxygen lattice sites on the surface as with Figure 2b. We then select a random number of surface oxygen to remove ranging from 0 (no vacancies) to all surface oxygen. We do the same on the other surface to maintain surface symmetry and avoid the manifestation of nonphysical dipole moments which can lead to diverging DFT energies.

The SI provides the chemical space distribution of all slabs and adsorbate+slabs successfully calculated in the dataset. Table 1 summarizes the distribution of elemental composition, crystal structures, bulk band gap, and number of components of the entire dataset of slabs and adsorbate+slabs.

Initial Structure Generation. To construct our adsorbate+slab, we first randomly sample one adsorbate from the set shown in Figure 3. This adsorbate set includes O^* , OH^* , OH_2^* , OOH^* , and O_2^* which are the intermediates in the proposed reaction mechanisms of OER. To expand the possible chemistry of adsorbates on oxides beyond OER, we also included monatomic H^* , O^* , N^* , and C^* , as well as CO^* . Table 1 shows the distribution of the 9 sampled adsorbates across the dataset.

We then determine the coverage of our random adsorbate on our randomly constructed slab. In contrast to the OC20 dataset, here we allow for more than one adsorbate of the same type to bind to the surface. The adsorbate can bind to three types of sites: the surface oxygen, the under-coordinated surface metal, or an oxygen vacancy. The maximum number of adsorbates allowed on the surface is limited by the sum of these three types of sites. However, we also ensure that all adsorbates are always separated by a distance greater than the M–O bond of the host material to avoid adsorbate overcrowding.

In this effort, we implemented specific strategies for placing adsorbates on the aforementioned surface sites as shown in Figure 3. The first row of placement strategies demonstrates that all adsorbates are able to bind to any undercoordinated surface metal at the lattice position of oxygen. This includes lattice positions of vacancies introduced during slab generation. An adsorbate containing oxygen will always bind to the metal via the oxygen atom as shown for OH^* , O_2^* , CO^* , H_2O^* and OOH^* . We also considered intermediates that arise due to formation of oxygen dimers which play a role in one of the possible mechanisms of OER.^{6,41} In this configuration, a pair of monatomic oxygen atoms can adsorb on to adjacent undercoordinated metals to form a dimer of 1.68 Å which is longer than the bond length of O_2^* .

The second row demonstrates how specific molecules that are able to form new molecules with the addition of oxygen can also bind to existing surface oxygen. For example, binding to a surface oxygen with the monatomic adsorbates will form a dimer molecule whereas CO^* and OH^* can bind to form CO_2^* and OOH^* respectively. Incorporating these reactions in the dataset will allow for the exploration of intermediate surface reactions that are only possible on oxides.

Lastly, we also allowed for a 4-fold rotational degree of freedom about the normal of the surface for all adsorbates. We randomly select the degree of rotation for each adsorbate on the surface after identifying the adsorbate sites.

Structure Relaxation. The OC22 dataset uses different computational settings than those used for the OC20 dataset. The OC22 dataset models the exchange-correlation effects with the Perdew–Burke–Ernzerhof (PBE), generalized gradient approximation (GGA)⁴² which is generally accepted for modeling surface reactions on oxides.^{6,43,44} In contrast, the OC20 dataset utilizes the RPBE DFT functional. We also accounted for strong electron correlations in some transition metal oxides by applying the Hubbard U correction in accordance to the suggestions made by the Materials Project.⁹ The last row of Table 1 shows the total number of slabs and adsorbate+slabs calculated using Hubbard U corrections.

Adsorbate-specific placement strategies

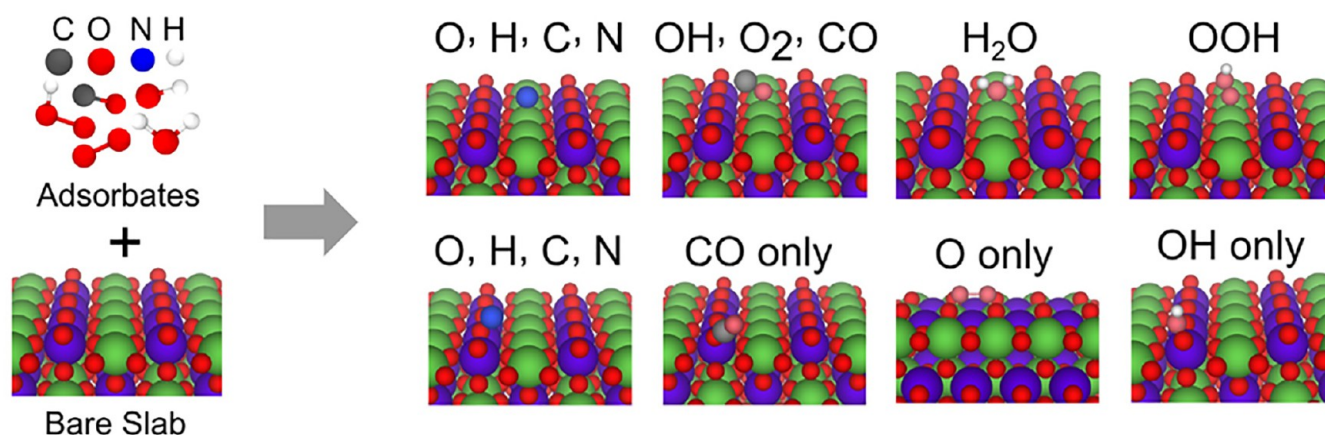


Figure 3. Overview of the adsorbate specific placement strategies. Adsorbates include H^* , O^* , N^* , C^* , OOH^* , OH^* , OH_2^* , O_2^* , and CO^* (left). Adsorbates can either bind to undercoordinated surface metals (first row of strategies) or to surface oxygen to form new intermediates (second row).

Although higher-level theory single-point calculations (e.g., hybrid functionals⁴⁵) are often used to verify the final electronic structure and energy of a surface, they still use a scheme similar to the one here to obtain the optimized structure. Models developed for this dataset will greatly accelerate more accurate workflows by focusing expensive calculations on the most stable and relevant structures.

In contrast to the OC20 dataset, all calculations were performed with spin-polarization to account for the significant spin states in metal oxides. Although some oxide materials exhibit magnetic polymorphism, we only considered one polymorph for each slab with all slabs being initialized with ferromagnetic or nonmagnetic configurations in accordance to the magnetic moments of each metal suggested by Horton et al.⁴⁶ These different magnetic states for a single crystal structure can significantly change thermodynamic properties at the surface. For example, rutile VO_2 has been demonstrated to have several different spin states with nonmagnetic surfaces yielding significantly lower surface energies than ferromagnetic surfaces for the same slab.⁴⁷ For further details regarding the computational settings, we refer the reader to the SI.

We allowed all atoms of the slab and adsorbate+slab to be relaxed. This will not only yield a lower DFT energy, but also allows for more accurate calculations of the surface energy by ensuring both surfaces are relaxed. This is in contrast to the OC20 dataset where only the adsorbates and the surface atoms were relaxed.

Systems that did not converge ionically were set aside for use in alternative tasks. All intermediate structures, energies, and forces are stored for future training and evaluation. The algorithms implemented to produce all input slabs and adsorbate+slabs were constructed with the aid of Python Materials Genomics (pymatgen)⁴⁸ and are available in the Open Catalyst Data set repository (https://github.com/OpenCatalyst-Project/OpenCatalyst-Dataset/tree/OC22_dataset). All calculations are performed using the Vienna ab initio simulation package (VASP).^{49–53} In total, we used over 240 million core-hours to create this dataset.

TASKS

The goal of the OC22 dataset is to efficiently simulate atomic systems with practical relevance to OER and other oxide applications. One approach to screening materials relies on simple descriptors such as adsorption energy and surface energy. These descriptors alongside the Sabatier principle⁵⁴ and surface Pourbaix diagrams⁵⁵ can be used to correlate with more complex outputs like activity and selectivity. Unfortunately, the primary bottleneck to doing so are computationally expensive DFT calculations. Calculations are further exacerbated for OC22 as its systems are larger and more complex than that of OC20. Again, we focus on structure relaxations as they have been a useful means to informing catalyst activity for a broad range of applications.^{56–61} Models developed for OC20 have shown great progress on their proposed tasks.^{10,15–17,19,20} In all of the OC20 tasks, energies were referenced to represent adsorption energy. While advantageous for screening purposes, this referencing, however, implicitly limited models to only studying adsorbate+slab combinations and not any one in isolation. In the context of OER, this is especially problematic as typical discovery pipelines require exploring different coverages and configurations of the surface.^{4,35,62–66} Figure 4 illustrates a typical workflow for OER where studying different surface terminations are necessary before running an adsorption calculation. Here, we propose modified variations of the OC20 tasks that would enable models to study surfaces with and without the presence of an adsorbate.

In all tasks, structures can contain a surface and adsorbate combination or just an isolated surface (a.k.a slab). The surface is defined by a unit cell periodic in all directions with a vacuum layer at least 12 Å. All ground truth targets are computed using DFT.

We briefly summarize the OC20 tasks below. For all tasks, energy is referenced to correspond to adsorption energy. See the original OC20 manuscript for more details.⁸ **Structure to Energy and Forces (S2EF)** takes a given structure and predicts the energy and per-atom forces. **Initial Structure to Relaxed Energy (IS2RE)** takes an initial structure and predicts the relaxed energy. **Initial Structure to Relaxed Structure**

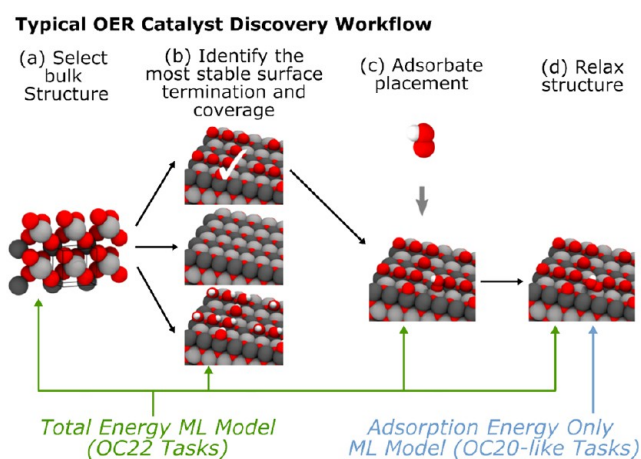


Figure 4. A typical OER workflow, motivating the need for total energy models beyond adsorption energies. Total energy models would allow one to study all parts of this workflow, and not just the final relaxation like adsorption energy models. (a) A bulk structure is selected from material datasets and a surface is created. (b) Surface terminations are enumerated and studied with DFT to identify the most stable termination. Surface Pourbaix diagrams are created and used to make this decision. (c) Only after the most stable termination is identified, an adsorbate is placed. (d) The adsorbate+slab system is relaxed and the referenced adsorption energy is computed.

(IS2RS) takes an initial structure and predicts the relaxed structure. The size of the train and validation splits for each task is listed in Table 2.

In the curation of both OC20 and OC22, slabs and adsorbate+slabs were relaxed in parallel, with adsorbates being placed on unrelaxed slabs. OC20 makes an assumption in computing an adsorption energy such that the corresponding relaxed slab reference is comparable to that of the adsorbate+slab combination. This assumption was feasible given that the majority of the surface was constrained.

Unlike OC20 where surface atoms are constrained, all atoms in OC22 are unconstrained. While this enables the community to study other surface properties like surface energy, the assumption that the relaxed clean surface and adsorbate+slab surface are comparable no longer holds. Computing an adsorption energy in the same manner of OC20 would correspond to an incorrect reference, resulting in an ill-posed, noisy target (see SI for more details). Instead, we modify the OC20 S2EF and IS2RE tasks to target DFT total energy rather than adsorption energy. We use the IS2RS task as is with no modifications.

Structure to Total Energy and Forces (S2EF-Total) takes a given structure and predicts the DFT total energy and per-atoms forces. Compared to S2EF, S2EF-Total differs only in its energy prediction. S2EF takes the DFT total energy and

references it by subtracting off a clean surface and gas phase adsorbate energy. S2EF-Total is only interested in the DFT total energy. The two tasks are related as follows:

$$\hat{E}_{S2EF} = \hat{E}_{S2EF-Total} - E_{slab}^{DFT} - E_{gas}^{DFT} \quad (1)$$

Initial Structure to Total Relaxed Energy (IS2RE-Total) takes a given structure and predicts the relaxed DFT total energy. Similar to S2EF-Total, IS2RE-Total is related to IS2RE as follows:

$$\hat{E}_{IS2RE} = \hat{E}_{IS2RE-Total} - E_{slab}^{DFT} - E_{gas}^{DFT} \quad (2)$$

DFT total energies are not meaningful on their own. Physically relevant properties like adsorption energy include some reference. A model that can predict a DFT total energy, however, gives the flexibility to reference to whatever is desired. Adsorption energy in this context would involve two predictions—one of the adsorbate+slab and one of the clean surface. For OER this is particularly important to identify the most stable surface coverage (or termination). While this problem is also important for OC20, those systems were much less complicated and the proposed adsorption energy tasks are typically sufficient.

Of the proposed tasks, S2EF-Total is the most general and closest to a DFT surrogate. Models trained for this task would enable researchers to study properties derived from isolated surfaces such as surface stability with respect to the bulk energy (surface energy), a necessary and important step in the catalyst discovery pipeline. Total energies also allows us to leverage surface trajectories and their energies for training, data that was previously unusable in OC20 using the specified bare slab energy reference.

■ BASELINE GNN MODELS

A wide range of models for catalyst and molecular applications have been proposed.^{15–21,67} We evaluate our tasks using the latest state of the art models. Additionally, we baseline alternative model architectures including equivariant and (non)energy-conserving models. Code for all baseline models are implemented in PyTorch⁶⁸ and PyTorch Geometric,⁶⁹ and are publicly available in our open source repository at <https://github.com/Open-Catalyst-Project/ocp>.

Graph Neural Networks (GNNs) have continued to grow in popularity as an efficient and accurate architecture for modeling atomic interactions. Unlike descriptor based models,^{70–73} where hand crafted representations are used to describe atomic environments, GNNs learn atomic representations through several message passing steps.⁷⁴ Consistent with related work,^{8,12,13} graphs are constructed with atoms treated as nodes and interactions between atoms as edges. Periodic boundary conditions are accounted for in graph construction

Table 2. Size of Train and Validation Splits^a

Task	Train			ID			OOD		
	Adslabs	Slabs	Total	Adslabs	Slabs	Total	Adslabs	Slabs	Total
S2EF-Total	6,642,168	1,583,125	8,225,293	313,238	81,489	394,727	356,633	94,036	450,669
IS2RE-Total	31,244	14,646	45,890	1,701	923	2,624	1,862	918	2,780
IS2RS	31,244	14,646	45,890	1,701	923	2,624	1,862	918	2,780

^aS2EF-Total structures come from a superset of IS2RE-Total systems, including unrelaxed systems (e.g. 50,810 train systems). Splits are sampled based on catalyst composition, ID for those from the same distribution as training, OOD for unseen catalyst compositions. Splits consist of both adsorbate+slab (adslabs) and slab systems. Validation and test splits are similar in size with exclusive compositions.

consistent with OC20. A cutoff radius is introduced for computational tractability.

We benchmark GNNs that have either performed well on OC20 or other molecular datasets. For *S2EF-Total*, we benchmark a larger sample of models including SchNet,¹² DimeNet++,¹⁴ ForceNet,¹⁵ SpinConv,¹⁶ PaiNN,⁷⁵ GemNet-dT,¹⁷ and GemNet-OC.¹⁹ IS2RS baselines are limited to the top performing models - SpinConv, GemNet-dT, and GemNet-OC. IS2RE-Total baselines include SchNet, PaiNN, DimeNet++, and GemNet-dT. Top performing *S2EF-Total* models were also evaluated for IS2RE-Total via an iterative relaxations approach.⁸

SchNet and DimeNet++ proposed continuous edge filters and directional message passing, respectively. ForceNet and SpinConv proposed architectures with direct force predictions in place of using energy derivatives with respect to atomic positions. PaiNN is an equivariant model with spherical harmonics up to order $l = 1$. We modify PaiNN's original architecture to make direct force predictions as our experiments showed a boost in performance. GemNet-dT incorporates symmetric message passing, scaling factors, equivariant predictions, and several efficient architecture improvements over the similar DimeNet++. GemNet-OC expands on GemNet-dT to efficiently capture quadruplet interactions, the current state of the art model across all tasks for OC20.

Unless otherwise noted, graph edges were computed on-the-fly via a nearest neighbor search for a cutoff radius of 6 Å and a maximum of 50 neighbors per atom. GemNet-OC uses different cutoffs for the type of interaction, e.g. triplets and quadruplets. Initial model sizes were taken directly from corresponding OC20 configurations. To accommodate for the fact OC22 has 16× less data, a light hyperparameter sweep was done for all models, with particular emphasis on learning rates, schedulers, and batch sizes. Effective batch sizes were set to ~192–256 for S2EF and ~4–64 for IS2RE. S2EF models used identical learning rate schedulers to more fairly compare baselines, decaying the learning rate at epochs 2, 3, 4, 5, and 6. IS2RE used a reduce on plateau learning rate scheduler. Full details on model hyperparameters and training configurations can be found in the SI.

All experiments used the following loss function⁸ to balance energy and force predictions:

$$\mathcal{L} = \lambda_E \sum_i |E_i - E_i^{DFT}| + \lambda_F \sum_{i,j} \frac{1}{3N_i} |F_{ij} - F_{ij}^{DFT}|^p \quad (3)$$

where λ_E and λ_F are empirical parameters, E_i is the energy of system i , F_{ij} is the force on the j th atom in system i , N_i is the number of atoms in system i , and p is the norm order. With the exception of GemNet-dT and GemNet-OC which used $p = 2$, all *S2EF-Total* models used $p = 1$. For IS2RE-Total only the energy term is evaluated, i.e. $\lambda_F = 0$. Baseline *S2EF-Total* models were trained with $\lambda_E = 1$ and $\lambda_F = N_{atoms}^{-2}$ to ensure size invariance, as detailed by Batzner et al.^{76,77}

EVALUATION METRICS

All our tasks use the same evaluation metrics proposed by OC20. The only difference is rather than ground truth values being DFT adsorption energies, we use DFT total energies for OC22. We briefly mention the metrics below but refer readers to the OC20 manuscript⁸ for a more detailed description.

S2EF-Total. The *S2EF-Total* task uses the same metrics as the OC20 S2EF task. Metrics include Energy Mean Absolute

Error (MAE), Force MAE, Force cosine, and Energy and Forces within Threshold (EFwT). Ground truth targets correspond to DFT total energy and per-atom forces.

IS2RE-Total. Similarly, IS2RE-Total uses the same metrics as the OC20 IS2RE task. Metrics include Energy MAE and Energy within Threshold (EwT). Ground truth targets correspond to the DFT total energy of the relaxed structure.

IS2RS. IS2RS metrics here are identical to that of OC20. Metrics include Average Distance within Threshold (ADwT), Force below Threshold (FbT), and Average Force below Threshold (AFbT). Ground truth targets are the relaxed structure. DFT is also used to evaluate predicted relaxed structures.

Consistent with OC20, our evaluation metrics still focus on accuracy. Given the complexity of OC22, we are interested in how previously successful models will perform on larger more intricate systems. In addition, we focus on models that are significantly faster than traditional DFT-based techniques. Models that can calculate energy and force estimates in under 10 ms would significantly aid oxide-related research.

ML EXPERIMENTS

The availability of large, diverse datasets like OC20 allows us to explore more interesting experiments alongside the OC22 dataset. In addition to training our baseline models on just OC22, we examine the extent the OC20 dataset and its pretrained models can benefit OC22 performance, and vice versa.

The varied training strategies are summarized in Figure 5. For each task we first study the performance using baseline

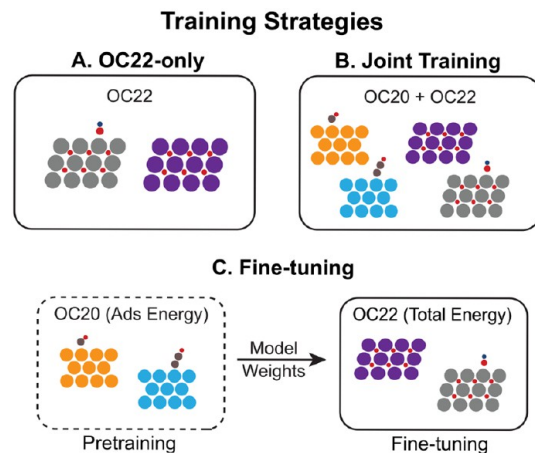


Figure 5. Various training strategies explored in OC22. **A.** The OC22-only strategy involves just using OC22 for the proposed tasks. **B.** Joint training refers to models trained on both OC20 and OC22 simultaneously. **C.** In fine-tuning, pretrained models for OC20 are used as starting points to train on just OC22.

models just trained on OC22 (**OC22-only**). This is the standard strategy when introducing a new dataset. Next, we leverage both OC20 and OC22 via **joint training** (OC20+OC22). In joint training we train a combined dataset of OC20 and OC22 systems. For *S2EF-Total*, we explore combined datasets with different sizes of OC20—2M, 20M, and All. While the OC20 energies were originally expressed as adsorption energy, for these experiments we use the DFT total energy which is also publicly accessible. One of the limitations to joint training is the need to train on a larger combined

Table 3. Predicting Total Energy and Force from a Structure (*S2EF-Total*)^a

		<i>S2EF-Total</i> Test							
Training	Model	Energy MAE [eV] ↓		Force MAE [eV/Å] ↓		Force Cosine ↑		EFwT [%] ↑	
		ID	OOD	ID	OOD	ID	OOD	ID	OOD
OC22-only	Median Baseline	163.4235	160.4547	0.0745	0.0729	0.0018	0.0018	0.00	0.00
	SchNet ¹²	7.9241	7.9248	0.0601	0.0823	0.3633	0.2195	0.00	0.00
	DimeNet++ ^{13,14}	2.0952	2.4751	0.0426	0.0585	0.6062	0.4360	0.00	0.00
	ForceNet ¹⁵			0.0564	0.0619	0.3507	0.2795	0.00	0.00
	SpinConv ¹⁶	0.8364	1.9440	0.0377	0.0631	0.5905	0.4118	0.00	0.00
	PaiNN ⁷⁵	0.9513	2.6300	0.0449	0.0583	0.4852	0.3449	0.00	0.00
	GemNet-dT ¹⁷	0.9385	1.2713	0.0316	0.0405	0.6647	0.5303	0.00	0.00
OC20-2 M + OC22	GemNet-OC ¹⁹	0.3742	0.8287	0.0294	0.0397	0.6913	0.5500	0.02	0.00
	PaiNN ⁷⁵	0.3993	1.5293	0.0475	0.0644	0.4670	0.3201	0.01	0.00
	SpinConv ¹⁶	0.9310	1.7897	0.0361	0.0552	0.6210	0.4641	0.00	0.00
OC20-20 M + OC22	GemNet-OC ¹⁹	0.4207	0.9135	0.0288	0.0373	0.6929	0.5602	0.01	0.00
	PaiNN ⁷⁵	0.3601	1.4538	0.0457	0.0605	0.4795	0.3410	0.01	0.00
	SpinConv ¹⁶	0.9721	1.5337	0.0355	0.0524	0.6008	0.4705	0.01	0.00
OC20-All + OC22	GemNet-OC ¹⁹	0.3105	0.8274	0.0268	0.0366	0.7216	0.5845	0.08	0.01
	SpinConv ¹⁶	1.2968	1.7044	0.0397	0.0467	0.5293	0.4417	0.00	0.00
	GemNet-OC ¹⁹	0.3111	0.6893	0.0269	0.0342	0.7064	0.5859	0.07	0.00
OC20 → OC22	SpinConv ¹⁶	1.1248	1.9661	0.0359	0.0505	0.6016	0.4578	0.00	0.00
	GemNet-dT ¹⁷	0.5715	1.0401	0.0307	0.0406	0.6734	0.5378	0.02	0.00
	GemNet-OC ¹⁹	0.2391	0.9378	0.0297	0.0407	0.6780	0.5364	0.13	0.00
	GemNet-OC-Large ^{b,19}	0.2173	1.0315	0.0272	0.0400	0.7298	0.5776	0.19	0.00

^aResults are shared for the OC22-only, joint, and fine-tuning training strategies. Experiments are evaluated on the test set. ^bFirst fine-tuned energy output weights on OC22 energies, then fine-tuned entire network on OC22 energies+forces.

dataset, which can significantly increase training time. To address this, we additionally explore **fine-tuning** (OC20 → OC22) experiments. In fine-tuning, models are initialized with pretrained weights learned from training on OC20. The pretrained models are then fine-tuned by training on just OC22. While approaches to fine-tuning vary in which portion of the network weights are updated, we limit our experiments to updating all the weights and leave more rigorous strategies as future work for the community.²⁸ For *S2EF-Total*, we experiment with fine-tuning using different fractions of the OC22 dataset. All fine-tuning experiments are performed using public OC20 adsorption-energy model checkpoints found at <https://github.com/Open-Catalyst-Project/ocp/blob/main/MODELS.md>.

Through these experiments we hope to share results that provide insights beyond just performance on OC22. Building a dataset that spans all possible applications, chemical diversity, and level of DFT theory is not computationally feasible. However, as we demonstrate with OC22, by leveraging large datasets, such as OC20, we may be able to train effective models with much smaller datasets for specific domains; even if they contain critical differences like DFT theory and material compositions.

RESULTS

We report results for all baseline models and tasks below. All validation results can be found in the SI.

S2EF-Total. Results on SchNet,¹² DimeNet++,¹⁴ ForceNet,¹⁵ SpinConv,¹⁶ PaiNN,⁷⁵ GemNet-dT,¹⁷ and GemNet-OC¹⁹ are shown in Table 3 (top). All models make energy and per-atom force predictions. SchNet and DimeNet++ make force predictions via a gradient of energy with respect to atomic positions, while all other models make direct force predictions. Across all metrics, GemNet-OC performs the best.

While GemNet-dT also demonstrates competitive force metrics, GemNet-OC significantly outperforms all models on energy based metrics. This may be due to the large receptive field (cutoff = 12 Å) of GemNet-OC better capturing long-range interactions and its unique ability to explicitly capture quadruplet interactions.

Results across the two test subsplits, In Domain (ID) and Out of Domain (OOD), are shown in Table 3. As expected, ID metrics are better than OOD. Unlike OC20 where ID and OOD-based splits had fairly close metrics, OC22 OOD metrics are substantially higher than ID. By definition, OOD contains combinations of material species not seen in the training set, i.e., if Ag–Cu is OOD, then a Ag–Cu only interaction has never been seen during training. This suggests generalization in the context of total energy predictions is more challenging than a referenced adsorption energy. Although physically motivated, the OC20 adsorption energy target can also be thought of as a form of Δ -learning,^{78–80} simplifying the complexity of the problem to learning a correction to some base property. To explore this in the context of OC22, we report results on a per-element linearly fit reference in the SI that helps improve performance. We refrained from making this the base task for OC22 in order to encourage alternative schemes or approaches to target normalization. OC20 results on the proposed tasks are also available in the SI, with similar poor performance suggesting *S2EF-Total* to be a generally more challenging task.

Joint training experiments on OC20 and OC22 are conducted for the top performing models, GemNet-OC, PaiNN, and SpinConv. Table 3 additionally contains results of different sizes of OC20 combined with OC22. To stay consistent with OC22, DFT total energy targets were used for OC20. With the addition of OC20 training data, GemNet-OC saw improvements in both energy and force predictions while PaiNN and SpinConv saw improvements to either only energy or forces, respectively. This suggests that despite the

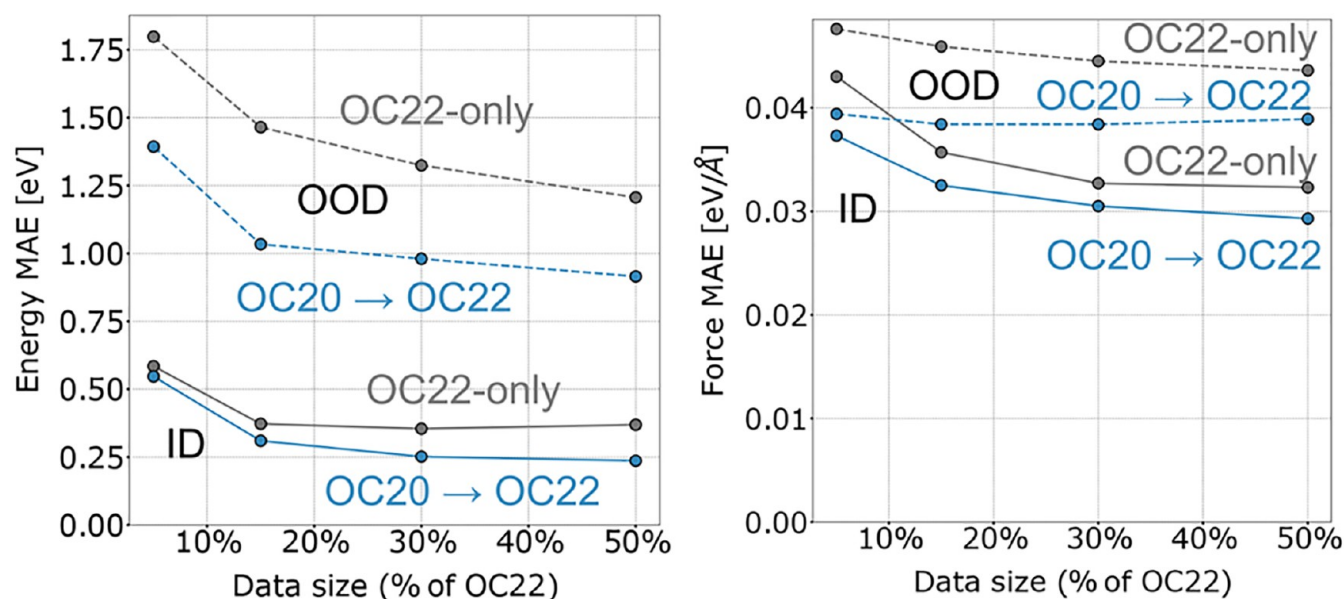


Figure 6. Results of GemNet-OC on *S2EF-Total* across different training data sizes. Two strategies are compared here: OC22-only and fine-tuning. Results are reported for both ID (solid) and OOD (dashed) on the test set.

Table 4. *S2EF-Total* Fine-Tuning Results Trained on Various Fractions of the OC22 Dataset^a

		<i>S2EF-Total</i> Test							
Training	Fraction of OC22	Energy MAE [eV] ↓		Force MAE [eV/Å] ↓		Force Cosine ↑		EFwT [%] ↑	
		ID	OOD	ID	OOD	ID	OOD	ID	OOD
OC22-only	5%	0.5847	1.7983	0.0430	0.0476	0.4967	0.4083	0.00	0.00
	15%	0.3727	1.4651	0.0357	0.0459	0.6142	0.4805	0.01	0.00
	30%	0.3546	1.3242	0.0327	0.0445	0.6589	0.5132	0.04	0.00
	50%	0.3692	1.2064	0.0323	0.0436	0.6566	0.5134	0.02	0.00
	100%	0.3742	0.8287	0.0294	0.0396	0.6913	0.5500	0.02	0.00
OC20 → OC22	0%	487.1210	434.6896	0.3650	0.3617	0.1944	0.1946	0.00	0.00
	5%	0.5469	1.3938	0.0373	0.0394	0.5482	0.4766	0.00	0.00
	15%	0.3103	1.0339	0.0325	0.0384	0.6212	0.5180	0.03	0.00
	30%	0.2519	0.9803	0.0305	0.0384	0.6565	0.5356	0.08	0.00
	50%	0.2365	0.9154	0.0293	0.0389	0.6790	0.5464	0.13	0.01
	100%	0.2391	0.9378	0.0297	0.0407	0.6780	0.5364	0.13	0.00

^aGemNet-OC¹⁹ was used for all experiments. Note, a fraction of 0% for OC22 corresponds to the baseline of directly evaluating a pretrained checkpoint from OC20 on OC22, with no additional training. All experiments are evaluated on the test set.

differences in DFT theory, the additional data is still meaningful in improving model predictions. However, increasing the amount of OC20 data had mixed results. GemNet-OC generally saw improvements across all metrics while SpinConv and PaiNN saw either minor improvements or worse performance. We note that training samples were randomly drawn, i.e., experiments with a larger proportion of OC20 would have seen fewer samples of OC22 during training. The differences in trends could be a result of model data efficiency and capacity. Exploring alternative sampling strategies to joint training could aid models and improve trends further. For our fine-tuning experiments, we evaluate GemNet-OC, SpinConv, and GemNet-dT models. Fine-tuning is performed by first training a model on OC20. This pretrained model is then fine-tuned by training on only OC22. Trained OC20 models are publicly available and were directly obtained from <https://github.com/Open-Catalyst-Project/ocp>. SpinConv saw little improvements on forces and worse performance for energies. GemNet-dT and GemNet-OC saw significant improvements to energy MAE

and minor improvements to force MAE for ID data. For OOD data, GemNet-dT generally saw improvements with fine-tuning while all other models either saw similar or worse results. To drive performance further, we trained GemNet-OC-Large, a larger, more parametrized version of GemNet-OC under a more careful fine-tuning strategy. Here, the energy output weights were first fine-tuned on OC22 energies, afterward, the entire network was fine-tuned on energies and forces. The large variant resulted in improved ID energy and force predictions, with OOD still seeing little or negative impacts. Fine-tuning experiments were extremely delicate and required careful selection of learning rates and hyperparameters, details are highlighted in the SI. While our initial fine-tuning results were generally limited to energy improvements, we hope the future development of more rigorous methods could lead to better performance across all metrics.

A potential benefit accompanying pretraining and fine-tuning is the need for less training data. A model initialized with meaningful weights could simplify the need to learn interactions and representations from scratch by utilizing an

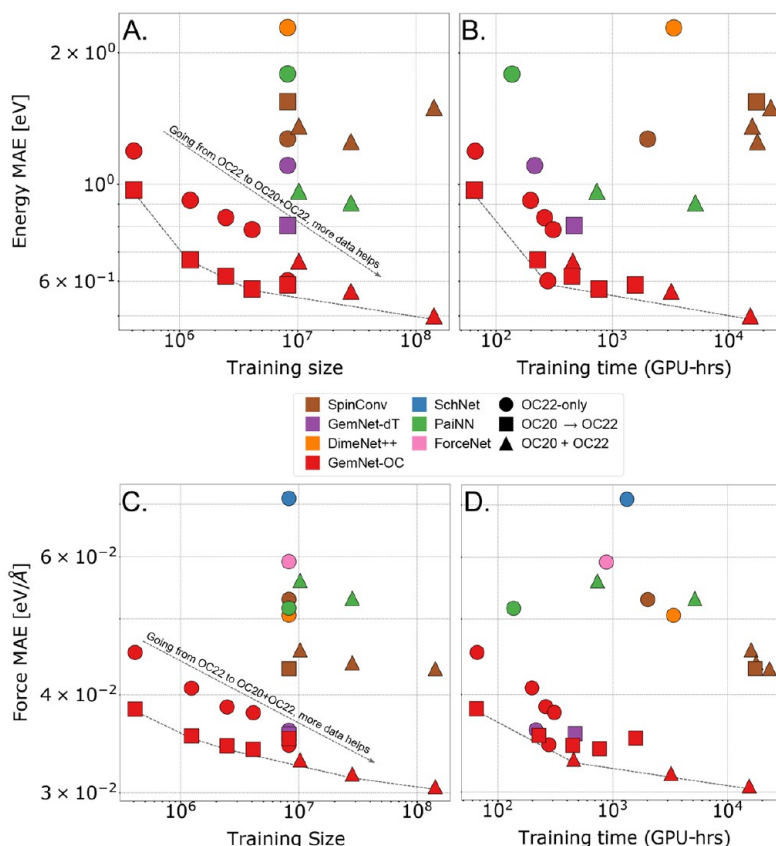


Figure 7. Summary of *S2EF-Total* test results as a function of training size (A,C) and training time (B,D). Models are color coded and the respective training strategy is indicated by different shapes. For fixed dataset sizes, fine-tuning experiments see improvements in both energy and force predictions. Increasing data consistently helps performance when moving from OC22 to OC20+OC22. Pareto fronts are provided for current optimums across training sizes and times. Fine-tuning experiments do not consider the dataset sizes and training times used during pretraining. Results are averaged across both ID and OOD splits.

Table 5. Predicting Total Relaxed Energy from an Initial Structure (IS2RE-Total)^a

IS2RE-Total Test						
Approach	Training	Model	Energy MAE [eV] ↓		EwT [%] ↑	
			ID	OOD	ID	OOD
Direct	OC22-only	Median Baseline	176.2560	171.8536	0.00	0.00
		SchNet	2.0012	4.8468	1.03	0.45
		DimeNet++	1.9600	3.5186	0.65	0.38
		PaiNN	1.7160	3.6835	0.88	0.38
		GemNet-dT	1.6771	3.0837	1.49	0.45
		SchNet	3.0384	4.2996	0.38	0.53
	OC20+OC22	DimeNet++	1.9614	3.4605	1.18	0.42
		PaiNN	1.7330	3.7519	0.76	0.49
		GemNet-dT	2.5234	4.2290	0.80	0.60
		GemNet-OC ^b	1.1527	1.7476	3.66	0.98
	OC20 → OC22	SpinConv	1.7367	2.6668	1.49	0.94
		GemNet-dT	1.8129	2.0439	1.64	0.83
		GemNet-OC	1.3294	1.5841	2.02	1.40
Relaxation	OC20+OC22	SpinConv	2.2959	2.5902	1.26	0.68
		GemNet-OC	1.2007	1.5339	2.63	2.15
		SpinConv	1.7997	2.8884	1.41	0.57
	OC20 → OC22	GemNet-OC	1.1201	1.8490	3.89	1.77
		GemNet-OC-Large	1.2534	2.1154	1.60	0.98

^aResults are shared for the OC22-only, joint, and fine-tuning training strategies. Experiments are evaluated on the test set. ^bGemNet-OC pretrained on OC20+OC22 *S2EF-Total*

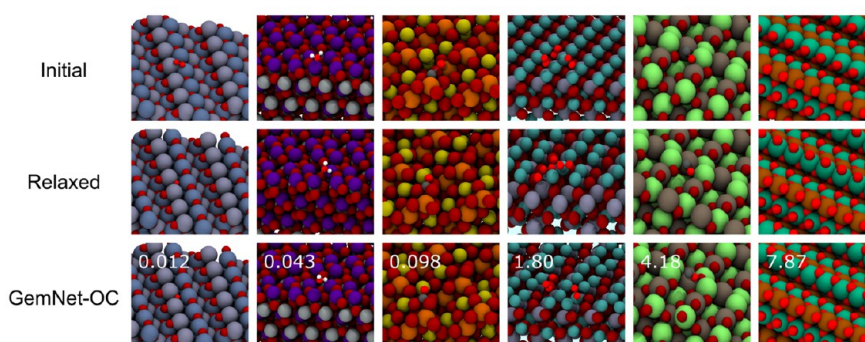


Figure 8. Demonstration of GemNet-OC solving the IS2RS and IS2RE-Total tasks via the relaxation approach. Initial, DFT Relaxed, and the ML predicted relaxed structures are shown for each system. The first three columns were randomly sampled from “successful” cases in which IS2RE-Total energy MAE was less than 0.1 eV, while the latter columns are “failure” cases, with energy MAEs greater than 0.5 eV. Oxygen found in the adsorbate is illustrated with a high contrast red and made smaller to distinguish it from oxygen in the catalyst material.

Table 6. Predicting Relaxed Structures from Initial Structures (IS2RS)^a

		IS2RS Test					
Training	Model	ADwT [%] ↑		FbT [%] ↑		AFbT [%] ↑	
		ID	OOD	ID	OOD	ID	OOD
OC22-only	IS baseline	43.39	45.26	0.00	0.00	0.03	0.10
	SpinConv	51.33	47.08	0.00	0.00	4.08	1.47
	GemNet-dT	57.84	54.17	0.00	0.00	4.16	3.54
OC20+OC22	GemNet-OC	59.47	55.72	0.00	0.00	5.49	4.45
	SpinConv	53.99	52.39	0.00	0.00	2.64	2.38
	GemNet-OC	58.55	58.44	0.00	0.00	8.01	6.58
OC20 → OC22	SpinConv	54.21	51.42	0.08	0.00	6.31	3.24
	GemNet-OC	55.55	50.50	0.08	0.00	9.02	6.59
	GemNet-OC-Large	57.23	54.63	0.00	0.00	10.41	8.09

^aAll models predicted relaxed structures through an iterative relaxation approach. The initial structure was used as a naive baseline (IS baseline). Experiments are evaluated on the test set.

alternative dataset. To explore this, we evaluated the performance of a pretrained GemNet-OC model fine-tuned on various fractions of OC22. As shown in Figure 6, a fine-tuned GemNet-OC consistently outperforms its OC22-only variant across all data sizes for the ID split, with diminishing returns for both strategies around ~50%. On OOD, energy performance continues to improve with data size. In Table 4, we additionally show the performance of a pretrained OC20 GemNet-OC used to directly evaluate OC22 (Fraction = 0%). As expected, energy metrics are extremely poor given the OC20 original target is adsorption energy. Force metrics are also extremely poor, suggesting the fine-tuning performance is not merely a result of a good pretrained model, but an actual transfer of knowledge from the two datasets. Figure 7 illustrates the various models and approaches as a function of training size and time. Notably, we see a strong linear trend in performance with data size. With saturation yet in sight, we expect more joint dataset efforts to continue to aid in performance. While for a fixed dataset size, fine-tuning efforts improved performance, they were often more costly in training time (Figure 7B,D). We anticipate future fine-tuning developments to be not only more accurate, but efficient as well. Similar fine-tuning experiments with OC20 models trained on DFT total energy targets were also performed. Results were consistent with those shared above, suggesting that despite a difference in targets, models are learning a similar underlying representation that is being transferred to OC22.

IS2RE-Total. We explore two approaches for predicting relaxed energies from initial structures—“Direct” and “Relaxation”.⁸ The first directly predicts the relaxed energy with a single call to the model. The relaxation approach uses a *S2EF-Total* model to run a structural relaxation—iteratively predicting forces and updating atomic positions until a relaxed structure and its corresponding energy is reached. While OC20 has shown relaxation based approaches to be superior to direct, they are 200–300× slower, motivating the potential benefits of direct models.

Table 5 presents IS2RE-Total results on both direct and relaxation approaches under the different training scenarios. Whereas OC20 saw relaxation based approaches to consistently perform better, we see mixed results here. The best relaxation-based approach, GemNet-OC, achieves an EwT of 3.89% indicating models have significant room for improvement. For the relaxation approach, fine-tuning consistently outperforms OC22-only. The best direct approach, GemNet-OC, also only achieves an EwT of 3.66%. Here, joint training consistently hurts performance. Following literature efforts,¹⁸ fine-tuning was done from the top performing *S2EF-Total* checkpoint - GemNet-OC OC20-All+OC22. While the best performing ID results come from a direct approach, OOD metrics are considerably better via the relaxation method, indicating their ability to better generalize. We evaluate OC20 IS2RE-Total performance in the SI and observe similar poor results, suggesting IS2RE-Total to be a considerably more challenging variation.

IS2RS. To evaluate the prediction of relaxed structures from initial structures, we select the top performing *S2EF-Total* models GemNet-dT, SpinConv, and GemNet-OC. Similar to OC20, we use these models to run ML driven structure relaxations (Figure 8). Relaxed structures were then evaluated with DFT to determine whether the predicted relaxed structures are valid. Table 6 shows GemNet-OC outperforming all other models across all metrics. Joint training and fine-tuning approaches both improve DFT force based metrics over OC22-only. GemNet-OC-Large fine-tuned achieves the best force metrics. Pursuant to OC20, non-DFT distance based metrics like ADwT struggle to correlate well with the practical DFT metrics.¹⁰ Both FbT and AFbT results indicate the models need significant improvement to achieve the level of accuracy needed for practical applications.

Does OC22 Benefit OC20? Alongside developing more accurate models, exploring augmentation strategies is another opportunity to improve performance on existing datasets like OC20.¹⁰ An interesting question is whether OC22 data may improve model performance on OC20. It has already been shown that the use of auxiliary data such as off-equilibrium MD or rattled data can lead to state-of-the-art results on OC20.¹⁹

To that end, we explore the impact that jointly training with OC22 and OC20 has on OC20 performance. Note OC22 is a significantly smaller and more limited dataset. OC20 contains ~134 M training data points and spans a large swath of materials. OC22 on the other hand is only ~6% of the size of OC20, limited to only oxide materials, and places no constraints on atoms in the systems. Table 7 compares the

Table 7. GemNet-OC Results Trained on Either OC20 or Both OC20+OC22 and Evaluated on OC20 and OC22^a

Training Data	Energy MAE [eV] ↓	Force MAE [eV/Å] ↓	Force Cosine ↑
OC22 evaluation			
OC20	55.9003	0.3842	0.1674
OC20+OC22	0.6606	0.0305	0.6574
OC20 evaluation			
OC20	0.3938	0.0219	0.6505
OC20+OC22	0.3171	0.0231	0.6486

^aResults are averaged across all ID/OOD validation splits. Total energies are used for all dataset targets.

performance of GemNet-OC trained on OC20 and OC20+OC22 as evaluated on both OC20 and OC22 separately. As expected, when trained on only OC20, OC22 metrics are poor—attributed to the lack of oxides in OC20 and the difference in DFT theories. When trained on OC20+OC22, however, we see a significant improvement in energy MAE (~20%). Force based metrics are either no different or slightly worse. Despite the joint dataset containing only a small fraction of OC22, it aided by a margin larger than any of the previous MD or rattled data efforts. Exploring in more detail as to how and why such improvements were observed could aid in systematically curating datasets to further improve OC20 performance.

Adsorption Energy from Total Energy Models. One of the difficulties with assessing the utility of total energy models is that, like DFT, total energy values themselves are meaningless without an appropriate reference. Here we explore the performance of adsorption energy predictions via the total energy models. As previously detailed (see Tasks), calculating

adsorption energies with OC22 data is ill-posed because of the potential inconsistencies resulting from relaxing the clean slab and adsorbate+slab in parallel. To address this, we sampled a subset of systems (~700) from OC22 and reran DFT calculations with a more conventional procedure. Slabs were first relaxed, then adsorbates were placed on the relaxed slab, and finally the adsorbate+slab system was relaxed. Additionally, the bottom layers in the surface were fixed. This new data allowed us to validate the predicted adsorption energy from total energy models.

Total energy models were utilized to calculate adsorption energies using two different schemes. In the first approach, which we will refer to as *Mixed-ML*, only the adsorbate+slab energy (\hat{E}_{sys}) in eq 4 is predicted with ML. Note, because OC22 includes systems with multiple adsorbates, the number of adsorbates, n_{ad} , is included in the calculation. Although this will yield the averaged adsorption energy across several adsorbates, for convenience, we will refer to it as the adsorption energy.

$$E_{\text{ad}}^{\text{Mixed}} = \frac{\hat{E}_{\text{sys}} - E_{\text{slab}}^{\text{DFT}} - E_{\text{gas}}^{\text{DFT}} \cdot n_{\text{ad}}}{n_{\text{ad}}} \quad (4)$$

Calculating the adsorbate+slab energy is the bottleneck in estimating adsorption energies, so replacing these DFT calculations with an ML potential could substantially improve the efficiency of screening new catalysts across a large swath of the chemical space. The second approach, *Full-ML*, predicts the energy of the slab (\hat{E}_{slab}) and the adsorbate+slab (\hat{E}_{sys}), which would improve throughput even further.

$$E_{\text{ad}}^{\text{Full}} = \frac{\hat{E}_{\text{sys}} - \hat{E}_{\text{slab}} - E_{\text{gas}}^{\text{DFT}} \cdot n_{\text{ad}}}{n_{\text{ad}}} \quad (5)$$

The adsorption energy results for three models, GemNet-OC, PaiNN, and SpinConv are reported in Table 8. All models

Table 8. Adsorption Energy Predictions from Total Energy Models on a Subset of the OC22 Validation ID Dataset^a

Training Data	Model	Mixed-ML Ads. Energy MAE [eV] ↓	Full-ML Ads. Energy MAE [eV] ↓
		MAE [eV] ↓	MAE [eV] ↓
OC22	GemNet-OC	0.678	0.767
OC20-All+OC22	GemNet-OC	0.691	0.724
OC22	PaiNN	1.295	0.965
OC20-2M+OC22	PaiNN	0.795	0.825
OC22	SpinConv	1.357	1.001
OC20-2M+OC22	SpinConv	0.984	0.980

^aTwo scenarios are considered, *Mixed-ML*, where only the energy of adsorbate+slab is predicted with ML, and *Full-ML*, where both the slab and adsorbate+slab are predicted with ML.

were trained on the *S2EF-total* task and used to run ML relaxations from the same initial structures as the DFT validation. More details on these calculations can be found in the SI. As expected, the error is higher than predicting adsorption energy directly as was done in OC20, but predicting total energy is a more challenging and general task.

Comparing the *Mixed-ML* and *Full-ML* adsorption energy results allows us to examine whether a cancellation or compounding of errors occurs. In the *Full-ML* approach, the ML potential could have some systemic bias compared to the

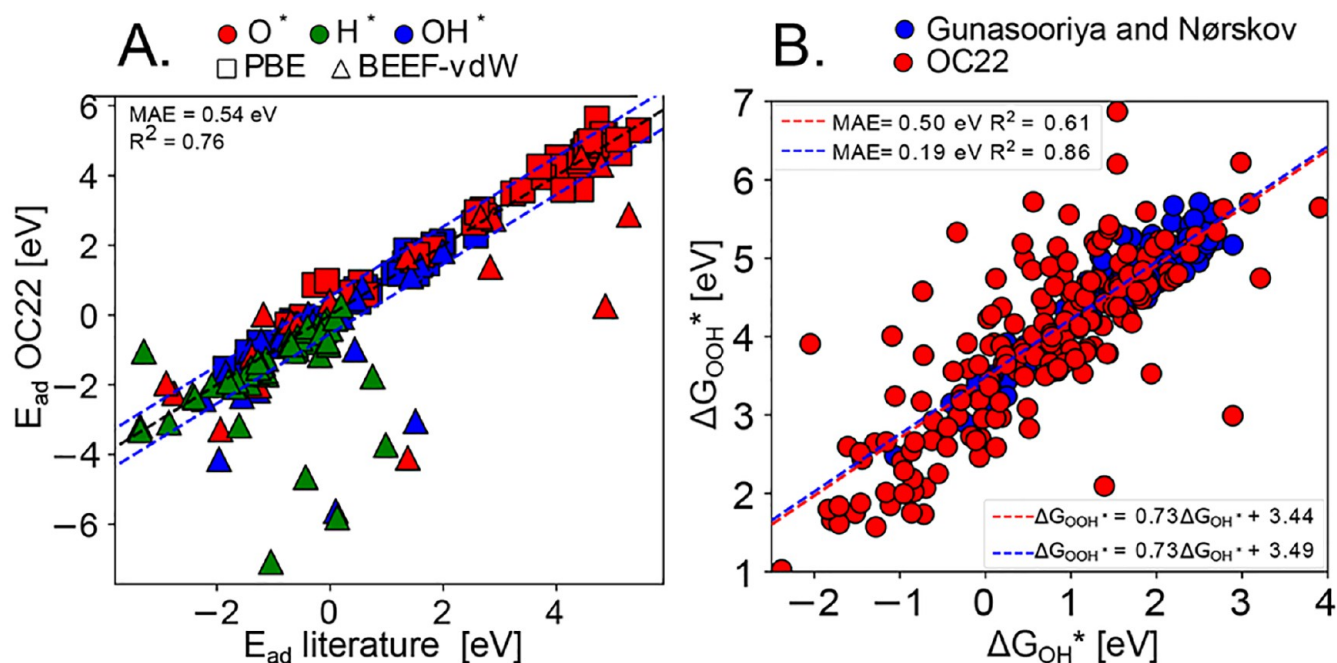


Figure 9. (A) Comparison of OC22 predicted (y-axis) and literature (x-axis) values for the adsorption energies of O^* , H^* , and OH^* across different OOD metal oxide compounds for rutile structures (see SI for a comparison of perovskite structures). A parity line (black-dashed) is provided for reference as well as a line above and below to indicate the mean absolute error (blue-dashed). (B) A comparison of ΔG_{OOH^*} (y-axis) and ΔG_{OH^*} (x-axis) with predicted (red) and literature (blue) data points shown along with their corresponding linear fits (dashed lines). All predictions were performed using the GemNet-OC OC20+OC22 model.

total energy DFT labels, yet still produce accurate adsorption energies if this error is “cancelled out” by subtracting off the ML slab prediction containing a similar bias. Alternatively, because ML is used for both an adsorbate+slab and slab prediction, errors could accumulate in the final prediction. While GemNet-OC does not observe any cancellation of errors, PaiNN and SpinConv see some improvements on the Full-ML approach. However, these trends go away when the training is augmented with OC20 data, with both approaches producing similar results. While adsorption energy prediction from total energy models is still far from chemical accuracy ~ 0.1 eV, we demonstrate the utility of current models compared to literature studies below. Additionally, we envision that the release of the OC22 dataset will lead to rapid modeling improvements as was the case for OC20.

Comparisons with Literature. To demonstrate relevance of our models beyond the OC22 dataset, we use our models to compare predicted adsorption energies and trends with corresponding existing data in the literature for O^* , H^* , and OH^* .

Figure 9A plots our predicted values for the adsorption energies of O^* , H^* , and OH^* against a sample set of literature values^{55,81} for rutile surfaces. The adsorption energies demonstrated a linear correlation of at least 75% with the majority of predicted values within 0.6 eV of the literature. The 1 eV discrepancy of several outliers could be accounted for by several differences in computational parameters between the literature values that utilized BEEF-vdW⁵⁵ and the training data for OC22. These discrepancies include the lack of Hubbard U corrections and the absence of spin-polarization in the literature. Despite the large deviation in some of the BEEF-vdW data points, this comparison demonstrates that a majority

of our predicted adsorption energies for some compositions are agnostic of the functional.

From the adsorption energy, we can obtain scaling relationships which are useful for identifying optimal catalysts across a variety of materials. Predicting these trends with OC22 will demonstrate the value of our model as a viable tool for screening catalysts. Figure 9B shows the scaling relationship between the Gibbs adsorption energy of OH^* and OOH^* calculated at standard conditions ($T = 298.15$ K, $P = 1$ bar) in literature⁶² and predicted data points using OC22 (see Table S11 in the SI for the Gibbs energy corrections). Our linear fitting of the slope (0.73) and intercepts (3.44) for ΔG_{OOH^*} vs ΔG_{OH^*} are consistent within 0.05 eV of the literature slope (0.73) and intercepts (3.49).⁶² We also demonstrated a similar fitting for ΔG_{O^*} vs ΔG_{OH^*} (see SI) albeit with our intercept overestimating by approximately 0.7 eV. Despite the significantly higher MAE and lower R^2 in the predicted relationships, we were still able to obtain a significant linear correlation above 60%.

While adsorption energy plays an important role for catalyst discovery, incorporating vibrational and zero point energy contributions are necessary for Gibbs energy calculations. Gibbs energy of adsorption (ΔG_{ad}) is often necessary for constructing accurate reaction pathways, creating microkinetic models, and determining the overpotential. However, at specific atmospheric conditions, the adsorption energies can generally be shifted by a constant correction value to obtain ΔG_{ad} . To demonstrate this, we plotted a minimal set of data points (15–20) of ΔG_{ad} at standard temperature and pressure against the adsorption energy to obtain a constant shift (ΔG_{corr}) between the two quantities. We can therefore add ΔG_{corr} to the adsorption energy of any system with the same adsorbate to get ΔG_{ad} (see Figure S4). This method of

calculating ΔG_{ad} when applied to predicted adsorption energies circumvents the need for a separate model of ΔG_{ad} .

For more details in regards to literature validation and other scaling relationships, we refer the reader to the SI.

DISCUSSION

There are many challenges to building large datasets and fitting generalizable models in computational catalysis, some of which were recently summarized.¹⁰ All of the challenges described also apply to the OC22 dataset - model performance varies across adsorbates and materials, direct force predictions tend to perform the best despite breaking energy conservation, developing helpful metrics for common tasks like local relaxations is difficult, and choosing the right calculations to improve the performance and generalizability of models is challenging. This work adds to these difficulties by highlighting additional challenges in capturing long-range interactions, developing models that go beyond adsorption energy, and fitting models with multiple datasets and levels of theory.

The performance of baseline models in this work is impressive given the difficulty of predicting the total system energy of complex oxide surfaces, but challenges still remain. The best results on the most general *S2EF-Total* task using a transfer learning approach from OC20 has an energy MAE of 0.24 eV for ID performance and 0.94 eV for OOD performance. Using that same model to predict relaxed total energies yields energy MAEs of 1.12 eV for ID and 1.85 eV for OOD predictions. These results are somewhat more impressive on a per-atom basis as is common for formation energy estimates of materials. However, for predicting experimentally relevant properties like the overpotential for the OER, these results are far from sufficient. We note that the initial baseline models for OC20 were similarly unhelpful for catalyst activity predictions, but rapid contributions from the broader community greatly improved their accuracy and predictive power. We hope that similar progress is seen for the tasks here. We also expect that the current models may already be helpful for certain more limited tasks, such as accelerating future oxide calculations with the use of online fine-tuning.⁸²

The tasks proposed in this work aim to push the community more in the direction of a general purpose potential, rather than separate models for each specific property. As an example, the tasks in OC20 were limited to the prediction of a specific property - the adsorption energy, following the most common approach in the community.^{83–85} This was a reasonable choice as the adsorption energy is a common descriptor for catalytic properties, and the adsorption energy itself was thought to be easier to fit than the DFT total energies. However, defining the tasks in this way meant that resulting models could only predict the adsorption energy and were unhelpful for predicting other surface properties like the surface energy. These limitations are highlighted in oxide catalysis where the stability of various surface terminations is needed. The total energy tasks in this work should encourage models that serve as general DFT surrogates - making predictions on a much wider range of properties.

Future Directions. Long-Range Interactions. The OC22 dataset contains long-range interactions that are likely difficult to capture in existing GNN models. Unlike metal surfaces which have a sea of electrons that can screen interactions, many of the oxides in OC22 are semiconductors with considerable partial charges (especially on the oxygen

atoms). Electrostatics have very long-range effects (energy decaying as $1/r$), and the partial charges can vary from system to system. The interaction of magnetic spins in systems with spin polarization is also long-ranged. This poses a challenge for the GNNs used in this work, which are often developed under the assumption that local interactions dominate. The use of several message passing steps or long-range local cutoffs may allow for these long-range interactions to be captured. There has been considerable effort in developing ML models that include long-range interactions,^{86–88} and we expect those approaches to be very useful in improving predictions for OC22.

Higher-Level Theory. The OC22 dataset also highlights the challenges of requiring multiple levels of theory for varying properties and materials. The OC20 dataset was constructed with the RPBE functional and neglected spin polarization, which represented a good trade-off between accuracy and computational cost for adsorption energies. However, some oxide surfaces require proper selection of Hubbard *U* corrections and can exhibit significant spin polarization. Combining datasets with multiple levels of theory, or upgrading datasets from less accurate to more accurate methods are popular questions in the small molecule community,^{89,90} but applying these ideas to OC20/OC22 will require extending these approaches to large datasets and inorganic materials, and we hope the community rises to this challenge. An obvious future direction is to improve the data quality with far more expensive hybrid functionals on the relaxed structures here.

Magnetic and Charge Effects. While additional information beyond just atom positions and atomic numbers have yet to improve performance on earlier datasets like OC20, it remains an open problem, how to best incorporate additional physics. In the context of OC22, magnetic configurations play an important role in oxide chemistry. Oxides exhibit different magnetic configurations for the same structure. These magnetic polymorphs can lead to different energetic, structural and magnetic trajectories along with different oxidation states during relaxation which can drastically affect chemisorption.^{91–93} Identifying ways to include magnetic moments in future models or training strategies might be an important contribution to improving GNN performance.

We would also like to explore the effect of charge balancing in the future for our oxide systems. In contrast to metallic systems, semiconductors can thermodynamically favor surface reconstruction over electron promotion when dopants, charges or stoichiometric defects are present. Essentially, bonds can be broken or created in such a way that electrons are prevented from getting promoted into the conduction band in a process called self-compensation.^{94,95} As a result, nominally identical initial structures, but with different numbers of electrons (and consequently, Fermi levels), can relax to significantly different final geometries (e.g., adatoms being expelled or dimers formed or broken⁹⁶). This effect can be extremely long-range as it only depends on the total number of shared electrons in the system, i.e., a vacancy introduced on one side of the slab may affect a local geometry on the other side.

Capturing the complexity of such phenomena with ML models would require a more in-depth analysis of a system's Fermi level's and the model's ability to capture long-range interactions. In particular, the differences in forces and in final geometries induced by electronic doping should be explored, as well as the ability of GNNs to differentiate between them

using an additional doping descriptor. To disentangle the effect of oxygen vacancies on Fermi level vs their local effects, a purely charging-based dataset may be prepared. While Fermi level positions are results of DFT calculations, electron counting and band filling,^{95,97} can be leveraged to provide empirically similar information as additional data.

Solvation Effects. The impact of solvation is another area for future research. Although we did not directly model solvation effects in this work, we are able to account for the consequences of partial surface dissolution by incorporating random oxygen vacancies at varying coverages. Oxide surfaces are prone to partial dissolution from solvents which can lead to surface vacancy defects that can modify catalytic properties.⁹⁸ Modeling OER on RuO₂ and IrO₂ in the presence of these defects allowed for previous computational studies to obtain descriptors of overpotential and activity that more closely reflect experimental observations.^{99,100}

Training Strategies. Joint training on both the OC20 and OC22 datasets leads to several unexpected results. Surprisingly, naively fitting on both OC20 and OC22 (much smaller dataset) leads to large accuracy improvements for predicting OC20 energies, as shown in Table 7. In addition, models trained on either OC22 or OC20+OC22 both appear to follow the same log–log scaling for energy MAE (Figure 7). These observations open the door to using a wide array of existing large datasets (NOMAD,¹⁰¹ Materials Project,⁹ OQMD³⁷) that although different, could aid in model development. These ideas can be rationalized if all of these datasets together can help learn more flexible and useful representations, regardless of their specific tasks or details.

Fine-tuning and transfer learning baselines were investigated as potential routes to improve accuracy across both OC20 and OC22 and reduce the computational intensity of training GNNs for these tasks. The most accurate models for both OC20 and OC22 were models trained on both datasets simultaneously, which indicates that a common representation can be learned and shared by both datasets. Surprisingly, the limited fine-tuning experiments in this work did not improve substantially on the accuracy/cost Pareto front (Table 7). However, there are many possible fine-tuning strategies and a large number of variations (e.g., which sections of the GNN to freeze or fit, or leaving this decision to an attention block²⁸), and we expect more progress from the community in this area. These approaches are necessary to encourage the reuse of large models, and to reduce the computational cost of obtaining state-of-the-art models for future small datasets.

Alternative Property Predictions. Models trained on OC22 could predict the total energies for any slab or adsorbate+slab which ultimately allows us to determine any thermodynamic quantity including adsorption energy, surface energy, and reaction energy. Adsorption and reaction energies are useful for identifying viable catalysts. We can also predict the surface energy in order to construct elaborate phase diagrams which can be used to assess the thermodynamic stability of a surface at varying adsorbate coverages. Pourbaix diagrams (applied potential vs pH) are especially important for determining the thermodynamic viability of electrocatalysts. The surface energy can also be used to model the equilibrium crystal structure or Wulff shape. With a predictive model that circumvents DFT calculations, all these applications, which ordinarily require hundreds of DFT calculations, are possible with little to no computational cost.

This dataset will have a broad impact in discovering oxide catalysts for a variety of reaction families and unraveling complex reaction mechanisms in these systems. Oxide materials are likely present in any reaction under strong oxidative conditions, such as the accelerated degradation of long-lived contaminants like PFOA¹⁰² or systematically upgrading chemical building blocks.¹⁰³ Photocatalysis, which directly uses available sunlight to drive chemical reactions also relies heavily on oxides such as TiO₂ due to their desirable optical properties⁸¹ and could benefit from this dataset. One example which is currently computationally expensive to study is the Mars-van Krevelen (MvK) mechanism, which is one of the most common catalytic mechanisms in ionic crystals.^{104,105}

In the MvK, an adsorbate binds to a surface oxygen to form a new intermediate which desorbs to leave behind an oxygen vacancy, which can later be replenished by oxygen atoms from incoming adsorbates. By explicitly including oxygen defects and vacancies in the dataset generation process, we hope the resulting models will be helpful for accelerating these studies. Similar reactions that could benefit from these approaches are CO₂ capture on carbides¹⁰⁶ or nitrate reduction on nitrides.¹⁰⁷

Experimental Outlook. Ultimately, the goal of developing accurate computational models is to inform experimental design and discovery, either through direct quantitative agreement or by providing insight into the underlying physical phenomena. It is important to be cognizant of the fact that the catalyst systems we simulate with DFT are idealized versions of the actual physical systems observed in experiments. For instance, the structure and composition of an oxide catalyst are prone to change during the reaction due to interactions with reaction intermediates and the surrounding medium, which complicates the connection with idealized DFT calculations. One way to enhance this work would be to complement it with experimental validation and auxiliary data from other modeling techniques and experiments like microkinetic models and operando spectroscopy. In spite of these challenges, DFT has proven to be an essential tool for providing atomistic level insights for catalysis and we envision that the improvements made in modeling oxide catalysts as a result of the OC22 dataset, e.g. considering the influence of surface coverages and defects on the oxide structure at a much larger scale than previously possible, will pave the way toward strengthening the connection with experiments and unearthing underlying catalyst design principles.

■ ASSOCIATED CONTENT

Data Availability Statement

The full dataset is provided at <http://opencatalystproject.org> and available in an ASE¹⁰⁸ trajectory or model-ready LMDB format. Baseline models, dataloaders, and trainers are provided in the open source repository <https://github.com/Open-Catalyst-Project/ocp>.

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acscatal.2c05426>.

Details on OC20 S2EF-Total and IS2RE-Total results, results using an alternative reference scheme, a discussion on adsorption energy for OC22, performance on OC22 adsorbate+slabs and slabs, independently, training and hyperparameters for baseline models, full results on the validation splits, additional validation with literature, a description of calculated corrections for

Gibbs adsorption energy, and the Hubbard U corrections used (PDF)

AUTHOR INFORMATION

Corresponding Authors

Zachary Ulissi – Scott Institute for Energy Innovation, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States; Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15217, United States; orcid.org/0000-0002-9401-4918; Email: zulissi@andrew.cmu.edu

C. Lawrence Zitnick – Fundamental AI Research, Meta AI, Menlo Park, California 94025, United States; Email: zitnick@meta.com

Authors

Richard Tran – Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15217, United States; orcid.org/0000-0002-0308-2182

Janice Lan – Fundamental AI Research, Meta AI, Menlo Park, California 94025, United States

Muhammed Shuaibi – Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15217, United States; Fundamental AI Research, Meta AI, Menlo Park, California 94025, United States

Brandon M. Wood – Fundamental AI Research, Meta AI, Menlo Park, California 94025, United States

Siddharth Goyal – Fundamental AI Research, Meta AI, Menlo Park, California 94025, United States

Abhishek Das – Fundamental AI Research, Meta AI, Menlo Park, California 94025, United States

Javier Heras-Domingo – Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15217, United States; orcid.org/0000-0002-4322-3146

Adeesh Kolluru – Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15217, United States

Ammar Rizvi – Fundamental AI Research, Meta AI, Menlo Park, California 94025, United States

Nima Shoghi – Fundamental AI Research, Meta AI, Menlo Park, California 94025, United States

Anuroop Sriram – Fundamental AI Research, Meta AI, Menlo Park, California 94025, United States

Félix Therrien – Department of Electrical and Computer Engineering, University of Toronto, Toronto, Ontario M5S 3G4, Canada; Department of Physical and Environmental Sciences, University of Toronto Scarborough, Scarborough, Ontario M1C 1A4, Canada

Jehad Abed – Department of Electrical and Computer Engineering, University of Toronto, Toronto, Ontario M5S 3G4, Canada; Department of Materials Science and Engineering, University of Toronto, Toronto, Ontario M5S 3G4, Canada; orcid.org/0000-0003-1387-2740

Oleksandr Voznyy – Department of Physical and Environmental Sciences, University of Toronto Scarborough, Scarborough, Ontario M1C 1A4, Canada; orcid.org/0000-0002-8656-5074

Edward H. Sargent – Department of Electrical and Computer Engineering, University of Toronto, Toronto, Ontario M5S 3G4, Canada; orcid.org/0000-0003-0396-6495

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acscatal.2c05426>

Author Contributions

[†](R.T., J.L., M.S., B.M.W., and S.G.) These authors contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors acknowledge Shaama M. Sharada, Samira Siahrostami, Andrew J. Medford, Tiago F. Goncalves, Selin B. Bilgi, and Sophia Kassabian for their assistance in reviewing calculations and helpful discussions. The authors also acknowledge helpful discussions with Aleksandra Vojvodic, John Kitchin, as well as Johannes Gasteiger on modeling choices.

REFERENCES

- (1) Trasatti, S. Electrocatalysis by oxides - Attempt at a unifying approach. *J. Electroanal. Chem.* **1980**, *111*, 125–131.
- (2) James, M. I.; Sun, X. Recent progress on earth abundant electrocatalysts for oxygen evolution reaction (OER) in alkaline medium to achieve efficient water splitting – A review. *J. Power Sources* **2018**, *400*, 31–68.
- (3) Yuan, N.; Jiang, Q.; Li, J.; Tang, J. A review on non-noble metal based electrocatalysis for the oxygen evolution reaction. *Arabian Journal of Chemistry* **2020**, *13*, 4294–4309.
- (4) Flores, R. A.; Paolucci, C.; Winther, K. T.; Jain, A.; Torres, J. A. G.; Aykol, M.; Montoya, J.; Nørskov, J. K.; Bajdich, M.; Bligaard, T. Active learning accelerated discovery of stable iridium oxide polymorphs for the oxygen evolution reaction. *Chem. Mater.* **2020**, *32*, 5854–5863.
- (5) Chen, B. W. J.; Xu, L.; Mavrikakis, M. Computational Methods in Heterogeneous Catalysis. *Chem. Rev.* **2021**, *121*, 1007–1048.
- (6) González, D.; Heras-Domingo, J.; Sodupe, M.; Rodríguez-Santiago, L.; Solans-Monfort, X. Importance of the oxyl character on the IrO₂ surface dependent catalytic activity for the oxygen evolution reaction. *J. Catal.* **2021**, *396*, 192–201.
- (7) Andersen, M.; Reuter, K. Adsorption enthalpies for catalysis modeling through machine-learned descriptors. *Acc. Chem. Res.* **2021**, *54*, 2741–2749.
- (8) Chanussot, L.; Das, A.; Goyal, S.; Lavril, T.; Shuaibi, M.; Riviere, M.; Tran, K.; Heras-Domingo, J.; Ho, C.; Hu, W. others Open catalyst 2020 (OC20) dataset and community challenges. *ACS Catal.* **2021**, *11*, 6059–6072.
- (9) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. The Materials Project: A materials genome approach to accelerating materials innovation. *APL Materials* **2013**, *1*, 011002.
- (10) Kolluru, A.; Shuaibi, M.; Palizhati, A.; Shoghi, N.; Das, A.; Wood, B.; Zitnick, C. L.; Kitchin, J. R.; Ulissi, Z. W. Open Challenges in Developing Generalizable Large-Scale Machine-Learning Models for Catalyst Discovery. *ACS Catal.* **2022**, *12*, 8572–8581.
- (11) Xie, T.; Grossman, J. C. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.* **2018**, *120*, 145301.
- (12) Schütt, K.; Kindermans, P.-J.; Felix, H. E. S.; Chmiela, S.; Tkatchenko, A.; Müller, K.-R. SchNet: A continuous-filter convolutional neural network for modeling quantum interactions. *Advances in Neural Information Processing Systems* **2017**, 991–1001.
- (13) Gasteiger, J.; Groß, J.; Günnemann, S. Directional Message Passing for Molecular Graphs. *International Conference on Learning Representations*; ICLR, 2020.
- (14) Gasteiger, J.; Giri, S.; Margraf, J. T.; Günnemann, S. Fast and Uncertainty-Aware Directional Message Passing for Non-Equilibrium Molecules. *arXiv preprint*, Apr. 5, 2020, ver. 3. DOI: 10.48550/arXiv.2011.14115.

- (15) Hu, W.; Shuaibi, M.; Das, A.; Goyal, S.; Sriram, A.; Leskovec, J.; Parikh, D.; Zitnick, C. L. Forcenet: A graph neural network for large-scale quantum calculations. *arXiv (Machine Learning)*, Mar. 2, 2021, ver. 1. <https://arxiv.org/abs/2103.01436>.
- (16) Shuaibi, M.; Kolluru, A.; Das, A.; Grover, A.; Sriram, A.; Ulissi, Z.; Zitnick, C. L. Rotation invariant graph neural networks using spin convolutions. *arXiv (Machine Learning)*, June 17, 2021, ver. 1. DOI: 10.48550/arXiv.2106.09575.
- (17) Gasteiger, J.; Becker, F.; Günnemann, S. GemNet: Universal directional graph neural networks for molecules. *Advances in Neural Information Processing Systems* 2021, 34, 6790–6802.
- (18) Sriram, A.; Das, A.; Wood, B. M.; Goyal, S.; Zitnick, C. L. Towards training billion parameter graph neural networks for atomic simulations. Published as a conference paper at *International Conference on Learning Representations*. See the following: <https://openreview.net/pdf?id=0jP2n0YFmKG>.
- (19) Gasteiger, J.; Shuaibi, M.; Sriram, A.; Günnemann, S.; Ulissi, Z.; Zitnick, C. L.; Das, A. Developing Graph Neural Networks for Large and Diverse Molecular Simulation Datasets. *Transactions on Machine Learning Research (TMLR)* 2022. See the following: <https://openreview.net/forum?id=u8tvSxm4Bs>.
- (20) Ying, C.; Cai, T.; Luo, S.; Zheng, S.; Ke, G.; He, D.; Shen, Y.; Liu, T.-Y. Do Transformers Really Perform Badly for Graph Representation? *Advances in Neural Information Processing Systems* 2021, 34. See the following: <https://proceedings.neurips.cc/paper/2021/file/f1c1592588411002af340cbaedd6fc33-Paper.pdf>
- (21) Godwin, J.; Schaarschmidt, M.; Gaunt, A. L.; Sanchez-Gonzalez, A.; Rubanova, Y.; Veličković, P.; Kirkpatrick, J.; Battaglia, P. Simple GNN Regularisation for 3D Molecular Property Prediction and Beyond. *International Conference on Learning Representations*. 2021. See the following: <https://openreview.net/pdf?id=1wVvweK3olb>.
- (22) Takahashi, K.; Takahashi, L.; Miyazato, I.; Fujima, J.; Tanaka, Y.; Uno, T.; Satoh, H.; Ohno, K.; Nishida, M.; Hirai, K.; Ohya, J.; Nguyen, T. N.; Nishimura, S.; Taniike, T. The Rise of Catalyst Informatics. *Towards Catalyst Genomics* 2019, 11, 1146–1152.
- (23) Medford, A. J.; Kunz, M. R.; Ewing, S. M.; Borders, T.; Fushimi, R. Extracting knowledge from data through catalysis informatics. *ACS Catal.* 2018, 8, 7403–7429.
- (24) Schlexer Lamoureux, P.; Winther, K. T.; Garrido Torres, J. A.; Streibel, V.; Zhao, M.; Bajdich, M.; Abild-Pedersen, F.; Bligaard, T. Machine learning for computational heterogeneous catalysis. *ChemCatChem* 2019, 11, 3581–3601.
- (25) Winther, K. T.; Hoffmann, M. J.; Boes, J. R.; Mamun, O.; Bajdich, M.; Bligaard, T. Catalysis-Hub.org, an open electronic structure database for surface reactions. *Scientific Data* 2019, 6, 75.
- (26) Fujima, J.; Tanaka, Y.; Miyazato, I.; Takahashi, L.; Takahashi, K. Catalyst Acquisition by Data Science (CADS): A web-based catalyst informatics platform for discovering catalysts. *Reaction Chemistry and Engineering* 2020, 5, 903–911.
- (27) Wulf, C.; Beller, M.; Boenisch, T.; Deutschmann, O.; Hanf, S.; Kockmann, N.; Kraehnert, R.; Oezaslan, M.; Palkovits, S.; Schimpler, S.; Schunk, S. A.; Wagemann, K.; Linke, D. A Unified Research Data Infrastructure for Catalysis Research – Challenges and Concepts. *ChemCatChem* 2021, 13, 3223–3236.
- (28) Kolluru, A.; Shoghi, N.; Shuaibi, M.; Goyal, S.; Das, A.; Zitnick, C. L.; Ulissi, Z. Transfer learning using attentions across atomic systems with graph neural networks (TAAG). *J. Chem. Phys.* 2022, 156, 184702.
- (29) Smith, J. S.; Nebgen, B. T.; Zubatyuk, R.; Lubbers, N.; Devereux, C.; Barros, K.; Tretiak, S.; Isayev, O.; Roitberg, A. E. Approaching coupled cluster accuracy with a general-purpose neural network potential through transfer learning. *Nat. Commun.* 2019, 10, 2903.
- (30) Dai, W.; Jin, O.; Xue, G.-R.; Yang, Q.; Yu, Y. Eigentransfer: a unified framework for transfer learning. *Proceedings of the 26th Annual International Conference on Machine Learning* 2009; pp 193–200. See the following: <http://machinelearning.org/archive/icml2009/papers/141.pdf>
- (31) Rosenbaum, L.; Dörr, A.; Bauer, M. R.; Boeckler, F. M.; Zell, A. Inferring multi-target QSAR models with taxonomy-based multi-task learning. *Journal of Cheminformatics* 2013, 5, 33.
- (32) Turki, T.; Wei, Z.; Wang, J. T. Transfer learning approaches to improve drug sensitivity prediction in multiple myeloma patients. *IEEE Access* 2017, 5, 7381–7393.
- (33) Song, X. Z.; Zhu, W. Y.; Wang, X. F.; Tan, Z. Recent Advances of CeO₂-Based Electrocatalysts for Oxygen and Hydrogen Evolution as well as Nitrogen Reduction. *ChemElectroChem* 2021, 8, 996–1020.
- (34) Dey, S.; Dhal, G. C. Cerium catalysts applications in carbon monoxide oxidations. *Materials Science for Energy Technologies* 2020, 3, 6–24.
- (35) Wang, Z.; Zheng, Y.-R.; Chorkendorff, I.; Nørskov, J. K. Acid-stable oxides for oxygen electrocatalysis. *ACS Energy Letters* 2020, 5, 2905–2908.
- (36) Curtarolo, S.; Setyawan, W.; Wang, S.; Xue, J.; Yang, K.; Taylor, R. H.; Nelson, L. J.; Hart, G. L. W.; Sanvito, S.; Buongiorno-Nardelli, M.; Mingo, N.; Levy, O. AFLOWLIB.ORG: A distributed materials properties repository from high-throughput ab initio calculations. *Comput. Mater. Sci.* 2012, 58, 227–235.
- (37) Saal, J. E.; Kirklin, S.; Aykol, M.; Meredig, B.; Wolverton, C. Materials design and discovery with high-throughput density functional theory: the open quantum materials database (OQMD). *JOM* 2013, 65, 1501–1509.
- (38) Liu, B.; Li, C.; Zhang, G.; Yao, X.; Chuang, S. S.; Li, Z. Oxygen Vacancy Promoting Dimethyl Carbonate Synthesis from CO₂ and Methanol over Zr-Doped CeO₂ Nanorods. *ACS Catal.* 2018, 8, 10446–10456.
- (39) Asnavandi, M.; Yin, Y.; Li, Y.; Sun, C.; Zhao, C. Promoting Oxygen Evolution Reactions through Introduction of Oxygen Vacancies to Benchmark NiFe-OOH Catalysts. *ACS Energy Letters* 2018, 3, 1515–1520.
- (40) Lopes, P. P.; Chung, D. Y.; Rui, X.; Zheng, H.; He, H.; Martins, P. F. B. D.; Strmcnik, D.; Stamenkovic, V. R.; Zapol, P.; Mitchell, J. F.; Klie, R. F.; Markovic, N. M. Dynamically stable active sites from surface evolution of perovskite materials during the oxygen evolution reaction. *J. Am. Chem. Soc.* 2021, 143, 2741–2750.
- (41) Dau, H.; Limberg, C.; Reier, T.; Risch, M.; Roggan, S.; Strasser, P. The Mechanism of Water Oxidation: From Electrolysis via Homogeneous to Biological Catalysis. *ChemCatChem* 2010, 2, 724–761.
- (42) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* 1996, 77, 3865.
- (43) Heras-Domingo, J.; Sodupe, M.; Solans-Monfort, X. Interaction between Ruthenium Oxide Surfaces and Water Molecules. Effect of Surface Morphology and Water Coverage. *J. Phys. Chem. C* 2019, 123, 7786–7798.
- (44) Van Den Bossche, M.; Grönbeck, H. Adsorbate Pairing on Oxide Surfaces: Influence on Reactivity and Dependence on Oxide, Adsorbate Pair, and Density Functional. *J. Phys. Chem. C* 2017, 121, 8390–8398.
- (45) Rousseau, R.; Glezakou, V.-A.; Selloni, A. Theoretical insights into the surface physics and chemistry of redox-active oxides. *Nature Reviews Materials* 2020, 5, 460–475.
- (46) Horton, M. K.; Montoya, J. H.; Liu, M.; Persson, K. A. High-throughput prediction of the ground-state collinear magnetic order of inorganic materials using density functional theory. *npj Computational Materials* 2019, 5, 1–11.
- (47) Wahila, M. J.; Quackenbush, N. F.; Sadowski, J. T.; Krispeneit, J. O.; Flege, J. I.; Tran, R.; Ong, S. P.; Schlueter, C.; Lee, T. L.; Holtz, M. E.; Muller, D. A.; Paik, H.; Schlom, D. G.; Lee, W. C.; Piper, L. F. The breakdown of Mott physics at VO₂ surfaces. *arXiv (Strongly Correlated Electrons)*, Dec. 9, 2020, ver. 1. DOI: 10.48550/arXiv.2012.05306.
- (48) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Comput. Mater. Sci.* 2013, 68, 314–319.

- (49) Kresse, G.; Hafner, J. Ab initio molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium. *Phys. Rev. B* **1994**, *49*, 14251–14269.
- (50) Kresse, G.; Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **1996**, *6*, 15–50.
- (51) Kresse, G.; Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **1996**, *54*, 11169–11186.
- (52) The calculations in this work have been performed using the ab-initio total-energy and molecular-dynamics package VASP (Vienna ab-initio simulation package) developed at the Institut für Materialphysik of the Universität Wien.
- (53) Kresse, G.; Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **1999**, *59*, 1758.
- (54) Kuo, D.-Y.; Paik, H.; Kloppenburg, J.; Faeth, B.; Shen, K. M.; Schlom, D. G.; Hautier, G.; Suntivich, J. Measurements of oxygen electroadsorption energies and oxygen evolution reaction on RuO₂ (110): a discussion of the sabatier principle and its role in electrocatalysis. *J. Am. Chem. Soc.* **2018**, *140*, 17597–17605.
- (55) Dickens, C. F.; Montoya, J. H.; Kulkarni, A. R.; Bajdich, M.; Nørskov, J. K. An Electronic Structure Descriptor for Oxygen Reactivity at Metal and Metal-oxide Surfaces. *Surf. Sci.* **2019**, *681*, 122–129.
- (56) Huang, H.-C.; Li, J.; Zhao, Y.; Chen, J.; Bu, Y.-X.; Cheng, S.-B. Adsorption energy as a promising single-parameter descriptor for single atom catalysis in the oxygen evolution reaction. *Journal of Materials Chemistry A* **2021**, *9*, 6442–6450.
- (57) Nørskov, J. K.; Abild-Pedersen, F.; Studt, F.; Bligaard, T. Density functional theory in surface chemistry and catalysis. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 937–943.
- (58) Bligaard, T.; Nørskov, J. K. Ligand effects in heterogeneous catalysis and electrochemistry. *Electrochim. Acta* **2007**, *52*, 5512–5516.
- (59) Hammer, B.; Nørskov, J. K. *Adv. Catal.*; Elsevier, 2000; Vol. 45; pp 71–129.
- (60) Seh, Z. W.; Kibsgaard, J.; Dickens, C. F.; Chorkendorff, I.; Nørskov, J. K.; Jaramillo, T. F. Combining theory and experiment in electrocatalysis: Insights into materials design. *Science* **2017**, *355*, No. eaad4998.
- (61) Nørskov, J. K.; Bligaard, T.; Logadottir, A.; Bahn, S.; Hansen, L. B.; Bollinger, M.; Bengaard, H.; Hammer, B.; Sljivancanin, Z.; Mavrikakis, M.; et al. Universality in heterogeneous catalysis. *Journal of catalysis* **2002**, *209*, 275–278.
- (62) Gunasooriya, G. K. K.; Nørskov, J. K. Analysis of acid-stable and active oxides for the oxygen evolution reaction. *ACS Energy Letters* **2020**, *5*, 3778–3787.
- (63) Back, S.; Tran, K.; Ulissi, Z. W. Discovery of acid-stable oxygen evolution catalysts: high-throughput computational screening of equimolar bimetallic oxides. *ACS Appl. Mater. Interfaces* **2020**, *12*, 38256–38265.
- (64) Patniboon, T.; Hansen, H. A. Acid-Stable and Active M–N–C Catalysts for the Oxygen Reduction Reaction: The Role of Local Structure. *ACS Catal.* **2021**, *11*, 13102–13118.
- (65) Zagalskaya, A.; Evazzade, I.; Alexandrov, V. Ab initio thermodynamics and kinetics of the lattice oxygen evolution reaction in iridium oxides. *ACS Energy Letters* **2021**, *6*, 1124–1133.
- (66) Vinogradova, O.; Krishnamurthy, D.; Pande, V.; Viswanathan, V. Quantifying confidence in DFT-predicted surface pourbaix diagrams of transition-metal electrode–electrolyte interfaces. *Langmuir* **2018**, *34*, 12259–12269.
- (67) Liu, Y.; Wang, L.; Liu, M.; Zhang, X.; Oztekin, B.; Ji, S. Spherical message passing for 3d graph networks. *arXiv preprint*, Nov. 24, **2021**, ver. 5. DOI: 10.48550/arXiv.2102.05013.
- (68) Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L. others Pytorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing Systems* **2019**, 8026–8037.
- (69) Fey, M.; Lenssen, J. E. Fast graph representation learning with PyTorch Geometric. *arXiv (Machine Learning)*, Apr. 25, **2019**, ver. 3. DOI: 10.48550/arXiv.1903.02428.
- (70) Behler, J. Perspective: Machine learning potentials for atomistic simulations. *J. Chem. Phys.* **2016**, *145*, 170901.
- (71) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98*, 146401.
- (72) Chmiela, S.; Tkatchenko, A.; Sauceda, H. E.; Poltavsky, I.; Schütt, K. T.; Müller, K.-R. Machine learning of accurate energy-conserving molecular force fields. *Science Advances* **2017**, *3*, No. e1603015.
- (73) Bartók, A. P.; Kermode, J.; Bernstein, N.; Csányi, G. Machine Learning a General-Purpose Interatomic Potential for Silicon. *Phys. Rev. X* **2018**, *8*, 041048.
- (74) Gilmer, J.; Schoenholz, S. S.; Riley, P. F.; Vinyals, O.; Dahl, G. E. Neural Message Passing for Quantum Chemistry. *Proceedings of the 34th International Conference on Machine Learning* **2017**, 1263–1272.
- (75) Schütt, K.; Unke, O.; Gastegger, M. Equivariant message passing for the prediction of tensorial properties and molecular spectra. *International Conference on Machine Learning* **2021**, 9377–9388.
- (76) Batzner, S.; Musaelian, A.; Sun, L.; Geiger, M.; Mailoa, J. P.; Kornbluth, M.; Molinari, N.; Smidt, T. E.; Kozinsky, B. E. (3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.* **2022**, *13*, 1–11.
- (77) Musaelian, A.; Batzner, S.; Johansson, A.; Sun, L.; Owen, C. J.; Kornbluth, M.; Kozinsky, B. Learning Local Equivariant Representations for Large-Scale Atomistic Dynamics. *arXiv (Computational Physics)*, Apr. 11, **2022**, ver. 1. DOI: 10.48550/arXiv.2204.05249.
- (78) Shuaibi, M.; Sivakumar, S.; Chen, R. Q.; Ulissi, Z. W. Enabling robust offline active learning for machine learning potentials using simple physics-based priors. *Machine Learning: Science and Technology* **2021**, *2*, 025007.
- (79) Ramakrishnan, R.; Dral, P. O.; Rupp, M.; von Lilienfeld, O. A. Big data meets quantum chemistry approximations: the Δ -machine learning approach. *J. Chem. Theory Comput.* **2015**, *11*, 2087–2096.
- (80) Zhu, J.; Vuong, V. Q.; Sumpter, B. G.; Irle, S. others Artificial neural network correction for density-functional tight-binding molecular dynamics simulations. *MRS Commun.* **2019**, *9*, 867–873.
- (81) Comer, B. M.; Medford, A. J. Analysis of photocatalytic nitrogen fixation on rutile TiO₂ (110). *ACS Sustainable Chem. Eng.* **2018**, *6*, 4648–4660.
- (82) Musielewicz, J.; Wang, X.; Tian, T.; Ulissi, Z. FINETUNA: fine-tuning accelerated molecular simulations. *Machine Learning: Science and Technology* **2022**, *3*, 03LT01.
- (83) Sorescu, D. C.; Thompson, D. L.; Hurley, M. M.; Chabalowski, C. F. First-principles calculations of the adsorption, diffusion, and dissociation of a CO molecule on the Fe(100) surface. *Physical Review B - Condensed Matter and Materials Physics* **2002**, *66*, 354161–354163.
- (84) Sholl, D. S.; Steckel, J. A. *Density Functional Theory*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2009; pp 1–31.
- (85) Wellendorff, J.; Silbaugh, T. L.; Garcia-Pintos, D.; Nørskov, J. K.; Bligaard, T.; Studt, F.; Campbell, C. T. A benchmark database for adsorption bond energies to transition metal surfaces and comparison to selected DFT functionals. *Surf. Sci.* **2015**, *640*, 36–44.
- (86) Ko, T. W.; Finkler, J. A.; Goedecker, S.; Behler, J. General-purpose machine learning potentials capturing nonlocal charge transfer. *Acc. Chem. Res.* **2021**, *54*, 808–817.
- (87) Zubatiuk, T.; Nebgen, B.; Lubbers, N.; Smith, J. S.; Zubatyuk, R.; Zhou, G.; Koh, C.; Barros, K.; Isayev, O.; Tretiak, S. Machine learned Hückel theory: Interfacing physics and deep neural networks. *J. Chem. Phys.* **2021**, *154*, 244108.
- (88) Behler, J.; Csányi, G. Machine learning potentials for extended systems: a perspective. *European Physical Journal B* **2021**, *94*, 1–11.
- (89) Grambow, C. A.; Li, Y.-P.; Green, W. H. Accurate thermochemistry with small data sets: A bond additivity correction

and transfer learning approach. *J. Phys. Chem. A* **2019**, *123*, 5826–5835.

(90) Duan, C.; Chu, D. B.; Nandy, A.; Kulik, H. J. Two Wrongs Can Make a Right: A Transfer Learning Approach for Chemical Discovery with Chemical Accuracy. *arXiv (Chemical Physics)*, Jan. 11, **2022**, ver. 1..

(91) Biz, C.; Fianchini, M.; Polo, V.; Gracia, J. Magnetism and Heterogeneous Catalysis: In Depth on the Quantum Spin-Exchange Interactions in Pt₃M (M= V, Cr, Mn, Fe, Co, Ni, and Y)(111) Alloys. *ACS Appl. Mater. Interfaces* **2020**, *12*, 50484–50494.

(92) Biz, C.; Fianchini, M.; Gracia, J. Strongly Correlated Electrons in Catalysis: Focus on Quantum Exchange. *ACS Catal.* **2021**, *11*, 14249–14261.

(93) Ren, X.; Wu, T.; Sun, Y.; Li, Y.; Xian, G.; Liu, X.; Shen, C.; Gracia, J.; Gao, H.-J.; Yang, H.; Xu, Z. J. others Spin-polarized oxygen evolution reaction under magnetic field. *Nat. Commun.* **2021**, *12*, 1–12.

(94) Zunger, A. Practical doping principles. *Appl. Phys. Lett.* **2003**, *83*, 57–59.

(95) Pashley, M. Electron counting model and its application to island structures on molecular-beam epitaxy grown GaAs (001) and ZnSe (001). *Phys. Rev. B* **1989**, *40*, 10481.

(96) Voznyy, O.; Thon, S.; Ip, A.; Sargent, E. Dynamic trap formation and elimination in colloidal quantum dots. *J. Phys. Chem. Lett.* **2013**, *4*, 987–992.

(97) Voznyy, O.; Zhitomirsky, D.; Stadler, P.; Ning, Z.; Hoogland, S.; Sargent, E. H. A charge-orbital balance picture of doping in colloidal quantum dot solids. *ACS Nano* **2012**, *6*, 8448–8455.

(98) Heras Domingo, J. Modeling of RuO₂ surfaces and nanoparticles. *Their potential use as catalysts for the oxygen evolution reaction*. Ph.D. Thesis, Universitat Autònoma de Barcelona, 2019.

(99) Dickens, C. F.; Nørskov, J. K. A Theoretical Investigation into the Role of Surface Defects for Oxygen Evolution on RuO₂. *J. Phys. Chem. C* **2017**, *121*, 18516–18524.

(100) Zagalskaya, A.; Alexandrov, V. Role of Defects in the Interplay between Adsorbate Evolving and Lattice Oxygen Mechanisms of the Oxygen Evolution Reaction in RuO₂ and IrO₂. *ACS Catal.* **2020**, *10*, 3650–3657.

(101) Draxl, C.; Scheffler, M. NOMAD: The FAIR Concept for Big-Data-Driven Materials Science. *arXiv (Materials Science)*, May 14, **2018**, ver. 1. DOI: 10.48550/arXiv.1805.05039.

(102) Liang, S.; Pierce, R. D., Jr; Lin, H.; Chiang, S.-Y.; Huang, Q. J. Electrochemical oxidation of PFOA and PFOS in concentrated waste streams. *Remediation Journal* **2018**, *28*, 127–134.

(103) Védrine, J. C. Metal oxides in heterogeneous oxidation catalysis: State of the art and challenges for a more sustainable world. *ChemSusChem* **2019**, *12*, 577–588.

(104) Mars, P.; van Krevelen, D. W. Oxidations carried out by means of vanadium oxide catalysts. *Chem. Eng. Sci.* **1954**, *3*, 41–59.

(105) Hinuma, Y.; Toyao, T.; Kamachi, T.; Maeno, Z.; Takakusagi, S.; Furukawa, S.; Takigawa, I.; Shimizu, K. I. Density Functional Theory Calculations of Oxygen Vacancy Formation and Subsequent Molecular Adsorption on Oxide Surfaces. *J. Phys. Chem. C* **2018**, *122*, 29435–29444.

(106) Gracia, J. M.; Prinsloo, F. F.; Niemantsverdriet, J. Mars-van Krevelen-like mechanism of CO hydrogenation on an iron carbide surface. *Catal. Lett.* **2009**, *133*, 257–261.

(107) Abghoui, Y.; Skúlason, E. Electrochemical synthesis of ammonia via Mars-van Krevelen mechanism on the (111) facets of group III–VII transition metal mononitrides. *Catal. Today* **2017**, *286*, 78–84.

(108) Larsen, A. H.; et al. The atomic simulation environment—a Python library for working with atoms. *J. Phys.: Condens. Matter* **2017**, *29*, 273002.

Recommended by ACS

Unveiling the Hidden Energy Profiles of the Oxygen Evolution Reaction via Machine Learning Analyses

Tomohiro Fukushima, Kei Murakoshi, *et al.*

JULY 24, 2023

THE JOURNAL OF PHYSICAL CHEMISTRY LETTERS

READ 

First Step of the Oxygen Reduction Reaction on Au(111): A Computational Study of O₂ Adsorption at the Electrified Metal/Water Interface

Alexandra M. Dudzinski, Karsten Reuter, *et al.*

AUGUST 29, 2023

ACS CATALYSIS

READ 

Acetylene Semi-Hydrogenation on a Perovskite Oxyhydride Surface: Insights from First Principles and Microkinetic Modeling

Kristen Wang Romero, De-en Jiang, *et al.*

JUNE 26, 2023

ACS CATALYSIS

READ 

Electrocatalyst with Dynamic Formation of the Dual-Active Site from the Dual Pathway Observed by *In Situ* Raman Spectroscopy

Chao Jing, Jian-Qiang Wang, *et al.*

AUGUST 05, 2022

ACS CATALYSIS

READ 

Get More Suggestions >