

# General Screening of Surface Alloys for Catalysis

Matthew M. Montemore,\* Chukwudi F. Nwaokorie, and Gbolade O. Kayode

*Department of Chemical and Biomolecular Engineering, Tulane University, New Orleans,  
LA 70118*

E-mail: mmontemore@tulane.edu

Phone: (504) 314-7120

## Abstract

Intensive research in catalysis has resulted in design parameters for many important catalytic reactions; however, designing new catalysts remains difficult, partly due to the time and expense needed to screen a large number of potential catalytic surfaces. Here, we create a general, efficient model that can be used to screen surface alloys for many reactions without any quantum-based calculations. This model allows the prediction of the adsorption energies of a variety of species (explicitly shown for C, N, O, OH, H, S, K, F) on metal alloy surfaces that include combinations of nearly all of the d-block metals. We find that a few simple structural features, chosen using data-driven techniques and physical understanding, can be used to predict electronic structure properties. These electronic structure properties are then used to predict adsorption energies, which are in turn used to predict catalytic performance. This framework is interpretable and gives insight into how underlying structural features affect adsorption and catalytic performance. We apply the model to screen more than  $10^7$  unique surface sites on approximately  $10^6$  unique surfaces for 7 important reactions. We identify novel surfaces with high predicted catalytic performance, and demonstrate challenges and opportunities in catalyst development using surface alloys. This

work shows the utility of a general, reusable model that can be applied in new contexts without requiring new data to be generated.

## Introduction

Metal alloy surfaces can be high-performing, stable catalysts for a variety of reactions, but the large compositional phase space is difficult to explore, despite significant advances in developing design principles for these systems. Therefore, novel catalysts are often developed by taking existing materials that are known to perform well and modifying them to further improve their performance. Discovering effective catalysts composed of unintuitive combinations of elements generally requires some form of high-throughput screening, which is currently expensive and time-consuming. Developing efficient models can be useful for high-throughput screening, but currently these models must be created from scratch for each reaction of interest, requiring significant time and effort. Hence, a general framework that allows initial screening for many different reactions would greatly streamline the screening process, and would allow significant insight into general trends.

Adsorption energies are the most widely used design parameters in catalyst design. Adsorption energies are often good descriptors for multiple reasons: they control the rate of adsorption/desorption as well as the coverage, and differences in adsorption energies are often related to activation energies through Brønsted-Evans-Polanyi relations.<sup>1</sup> This design paradigm has been successfully used to rationalize catalytic performance<sup>2-4</sup> and to design new catalysts.<sup>3,5-7</sup> However, adsorption energies are difficult to measure and are slow to calculate.

Due to the difficulty of high-throughput screening, significant research has been devoted to developing efficient, parameterized models to predict adsorption energies and hence catalytic performance.<sup>8-15</sup> These models can achieve high accuracy, and have resulted in some successes in screening a particular alloy architecture for a particular reaction. However, past models have generally focused on predicting the adsorption energies of only one or two species, and hence are not applicable to a wide range of reactions. Similarly, they generally focus on a relatively small portion of the vast

combinatorial space of alloy surfaces. To screen for a new reaction or study a new set of surfaces, a new model must be created from scratch, which often requires hundreds or thousands of calculations. Further, a general framework allows understanding of general trends. Previous approaches to constructing a general model have required a DFT calculation for each surface,<sup>16,17</sup> which is much slower than a parameterized model based purely on features that can be extracted from the surface structure.

Indeed, developing general, reusable models is quite difficult in many areas of chemistry and materials science.<sup>18-21</sup> Fitted models are used in a wide range of applications, including molecular dynamics, exchange-correlation functionals, and materials design. These models may be derived from physical theories such as tight-binding,<sup>22</sup> or may be purely based on data-driven approaches such as machine learning.<sup>23</sup> Physically motivated models can have researcher bias, and may miss important features or behaviors. Data-driven approaches are unlikely to generalize beyond the space spanned by the training data and often use low-level features that require a large amount of data. To develop a general, reusable model that is not overly biased by intuition, we used a combined approach of considering data-driven techniques for model selection, as well as our physical understanding. The model decouples the surface from the adsorbate, and this novel architecture allows a significant improvement in generality over previous approaches.

Using this approach, we have developed a general, efficient model that is useful for designing transition metal alloy catalysts for many reactions. The model successfully predicts adsorption energies for a wide variety of species and is applicable to a wide phase space of potential transition metal catalysts. Therefore, the model can be reused in new contexts with no overhead, as there is no need for additional data generation or fitting. This work addresses a major drawback of most predictive models: their lack of reusability. This results in a much more efficient and convenient framework for high-throughput screening for new reactions. In other fields, it has been shown that relatively simple formulas or tables can be superior to expert judgement.<sup>24</sup> Analogously, a general model for adsorption could, in some cases, replace researcher intuition concerning catalytic performance.

# Methods

DFT calculations were performed using the VASP code.<sup>25,26</sup> The PW91 exchange-correlation functional<sup>27</sup> and the projector-augmented wave method<sup>28,29</sup> were used. The plane-wave basis set was cut off at 396 eV, and a  $7 \times 7 \times 1$  k-point mesh was used. All calculations were performed with a  $3 \times 3$  surface cell. Four layers were used, with the bottom two fixed at their bulk positions. Substitutional doping was performed in the top two layers. Adsorption energies are given relative to the gas-phase adsorbate, with the convention that exothermic adsorption is negative. Some calculations reconstructed significantly, which led to unphysical configurations and difficulty in reaching geometric convergence. As done in previous work,<sup>10,30</sup> these calculations were removed.

For each unique atomic environment, we generated a variety of physically motivated features. First, we calculated the electronic coupling between each metal atom and its neighbors, as this has proven useful in previous work predicting electronic structure properties of transition metal surfaces.<sup>9,31</sup> However, instead of assuming a precise functional form for the coupling, we calculated the coupling using a variety of exponents in the expression

$$V_{il,i}^2 = \frac{\sum_j (-r_{d,i})^{\alpha/2} (-r_{d,j})^{\alpha/2}}{d_{i,j}^\beta}, \quad (1)$$

analogous to previous work which fit various parameters for tight-binding models.<sup>32</sup> The index  $i$  represents the central atom under consideration, while  $j$  runs over this atom's neighbors. The values for the d-orbital radii  $r_d$  were taken from previous work, which fit them to band structures,<sup>33</sup> while the values for the atom-atom distances  $d_{i,j}$  were taken from unrelaxed surface geometries, so that the final model is truly predictive without requiring relaxed geometries from DFT. The exponents  $\alpha$  and  $\beta$  were varied from 1 to 9.

In our screening process, rates were predicted by extracting relationships from previous work. Many previous studies have plots of predicted rates as a function of adsorption energies, and we extracted data from these plots. These datasets were interpolated using standard techniques, and were extrapolated where necessary. To account for differences in computational setup, in most cases we used a reference surface

on which the adsorption energy was given in the original work. The adsorption energy difference on this surface between our DFT calculations and the previous work was added to our predicted adsorption energies. Because the results from different functionals are highly correlated,<sup>34</sup> we expect this to be reliable. As an example, the optimal catalyst for the oxygen reduction reaction should bind OH about 0.15 eV less strongly than Pt(111),<sup>35</sup> independent of the computational setup. Costs were calculated by taking the average cost of the atoms in the unit cell, using rough values for the cost per weight of each metal. Identification of symmetry-inequivalent sites and surfaces was done using PyMatGen.<sup>36</sup> Density plots (Figure 4) were created with Datashader.

## Results and Discussion

### Approach and Data

To create the model, we leveraged previous work<sup>17,37</sup> that shows that it is possible to predict the adsorption energy of a variety of species on transition metal alloy surfaces using a few electronic structure parameters: the d-band center  $\epsilon_d$ , the number of p electrons  $n_p$ , the matrix coupling element between the adsorbate and the metal d states  $V_{ad}^2$ , and the idealized filling of the d-band  $f_d$ :

$$E_{\text{ads}} = a_1 + a_2\epsilon_d + a_3n_p + a_4V_{ad}^2 + a_5V_{ad}^2f_d. \quad (2)$$

The value of  $f_d$  is determined by the column of the periodic table that the metal atom is in. For simplicity, in this work we fit Equation (2) separately for each species and for each type of site: top, bridge, and hollow. We did not distinguish between fcc and hcp hollows; fitting these sites separately had little effect on the accuracy.

Therefore, instead of predicting adsorption energies for each species directly from the surface structure, we predicted the electronic structure parameters in Equation (2), and then used these parameters to predict adsorption energies for many species (see Figure 1). Thus, instead of predicting many target variables with complicated relationships to each

other, we only needed to predict a few target variables:  $\epsilon_d$ ,  $n_p$ , and  $V_{ad}^2$ .

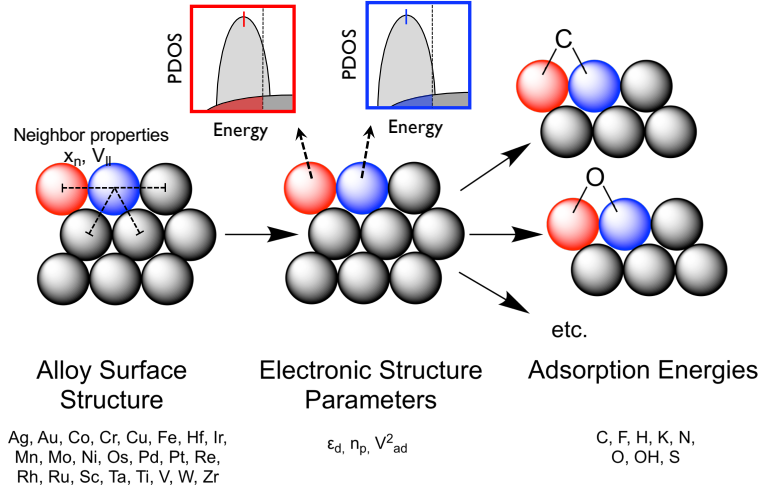


Figure 1: Schematic of overall screening methodology. For each atom in an alloy surface structure, we calculated average atomic properties of its neighbors (e.g.,  $x_n$ ) and calculated the electronic coupling to its neighbors ( $V_{ij}$ ). Using these properties, we predicted the atom’s electronic structure properties, and these electronic structure properties were used to predict adsorption energies for a variety of adsorbed intermediates. These adsorption energies were used to estimate catalytic performance. Because we can predict adsorption energies for many species, the framework can be reused for new reactions with no adjustment.

Our target phase space for this work was fcc(111) and hcp(0001) surfaces with a pure-metal bulk, where atoms in the top two layers are substituted with another metal. This is a large phase space that includes important architectures such as single-atom alloys, near-surface alloys, monolayer overlayers, and monolayer sandwiches. We can also approximate bulk alloys, because the local surface composition and arrangement generally has a larger effect on the adsorption energy than longer-range effects. The method is likely applicable to other structures and architectures as well. Twenty-three metals were included: Cu, Ag, Au, Ni, Pt, Pd, Co, Rh, Ir, Fe, Ru, Os, Mn, Re, Cr, Mo, W, V, Ta, Ti, Zr, Hf, Sc.

To generate data for our electronic structure model, we calculated the projected density of states for atoms in 2273 unique environments on 350 different surfaces. For each of these unique environments, we extracted  $\epsilon_d$  and  $n_p$ . We calculated  $V_{ad}^2$  from the carbon-metal distance for  $\text{CH}_3$  adsorbed in the top site, as in our previous work.<sup>17,37</sup> 335 of these calculations were performed.

We focused on predicting atomic adsorption energies because atomic adsorption

energies are common descriptors for catalysis. We included a set of elements that covers the vast majority of catalytic reactions: C, H, O, N, S, F, K. We also included OH, to show that our framework also applies to intermediates other than atomic species. We calculated 685 adsorption energies total for these species, referenced to the adsorbates in the gas phase.

## Models

As our goal was to maximize the generality of our model, we used a relatively high-bias, low-variance model—a linear model with only a few features—and combined physical insight with data-driven insight. A high-bias, low-variance model is particularly useful for creating more general models with relatively small data sets.<sup>38</sup> This framework is less prone to overfitting and also imbues the model with explanatory power, as differences in adsorption energy can be traced back to surface properties. To choose our features, we examined their predictive power in linear models with varying numbers of terms, and also used regularized regression to examine feature importance. We also carefully considered the physical justification for each feature, as choosing physically motivated features is likely to improve the model’s generality. This is particularly useful when many models have similar accuracy, as is the case here. We fit the electronic structure properties for each metal element separately, and there was some variation in the most accurate feature set between different metals. This created further ambiguity in determining the most important features. Hence, we chose features by hand, considering both data-driven factors and physical motivation, which resulted in a generally applicable model.

For each unique atomic environment in our dataset, we calculated two types of features. First, we calculated a variety of features based on a general form for electronic coupling in a tight-binding model, as a generalization of previous work (see Methods section).<sup>9,31</sup> Additionally, we calculated various property differences between each atom and the mean value for all of its nearest neighbors, again guided by previous work<sup>31</sup> and physical intuition. We denote each of these neighbor properties with a subscript  $n$ . Hence,  $f_d$  is the d-band filling of a surface atom and is used to predict adsorption energies

(Equation (2)), while  $f_{d,n}$  is the average difference in the d-band filling between an atom and its neighbors and is used to predict electronic structure properties.

To select which of these features to use in predicting each electronic structure property, we performed LASSO and ridge regression, fit all possible one- and two-term linear models, and fit a large number of three-term models. This was done for each metal element, each with different coefficients and accuracies. We examined the coefficients in LASSO and ridge regression, where the hyperparameters were tuned using cross validation and accuracy was tested on a separate test set. Three different 20%-80% test-train splits were used, and the results were averaged. Varying the test set size had little impact on the results. We also examined the highest-performing linear models with various numbers of terms, and often checked which features would most improve the test-set errors when added to postulated models. Because of the differences in results across different elements, and because many of the features have a high correlation with other features, no single technique for feature selection was robust and adequate. Further, because of our emphasis on generality, we considered the physical motivation for each potential feature and set of features.

Of the surface parameters we predict, the d-band center varied the most across the dataset and was the most difficult to predict. Based on our data-driven metrics, the neighbor atoms' idealized d-band filling  $f_{d,n}$  was clearly the most important feature. It performed noticeably better than any other feature in one- and two-term models and has the highest weight in ridge regression. It also has strong physical justifications: the filling of neighboring atoms' d-bands will affect electron transfer and hybridization. It is related to the previously developed concept of "diagonal disorder", which explains interactions between transition metal atoms partly in terms of the differences between their d bands.<sup>39</sup> Our data-driven metrics suggest a few other possible features: neighboring atoms' bulk thermal and electrical conductivity, covalent radius, and electronegativity. These are all tabulated quantities for each element. We examined two- and three-term models that include  $f_{d,n}$ , and found that the neighbor atoms' electrical conductivity  $\sigma_n$  consistently gave the best improvement when added to high-performing models, in addition to having



a large coefficient in ridge regression. The conductivity is related to the electronic interactions between an atom and its neighbors, and therefore it is reasonable for the conductivity to affect the d-band. Given  $f_{d,n}$  and  $\sigma_n$  as features, our data-driven metrics suggested that adding an electronic coupling term (Equation (1)) would improve the accuracy. While many of these terms are similar, coupling terms with lower exponents were more accurate. We chose exponent values of  $\alpha = 1.5$  and  $\beta = 3.5$  as these correspond to a derived form for  $V_{pd}^2$ , the coupling between d states and p or s states in neighboring metal atoms.<sup>33</sup> Previous studies have generally used  $V_{dd}^2$ , which we found to be somewhat less accurate. Hence, p-d hybridization—and/or s-d hybridization—may have a larger effect on the d-band center than d-d hybridization. Within uncertainty, our chosen model has the highest accuracy of any three-term model (along with many others). Further, each term in the model has its own data-driven and physical justification.

Somewhat surprisingly, we found that a single-term model for  $n_p$  gave sufficient accuracy. Adding an additional term gave only a slightly smaller test error and nearly identical training error. Within uncertainty,  $V_{pd}^2$  is the feature that gives the lowest test error. The p-d coupling is easy to justify physically, as it will have a strong effect on p-d hybridization and the p band.

The value of  $V_{ad}^2$  generally varies little for a given metal, and thus it is not surprising that a single feature suffices to predict it. Our feature selection process suggested that  $f_{d,n}$  gives the most accurate model, and adding an additional feature does not improve the accuracy. Physically,  $f_{d,n}$  correlates with the neighbor atoms' number of d electrons, which affects the d electrons through hybridization and electron transfer. This in turn affects the coupling between the metal atom's d states and the adsorbate states. The final features in each model and the data-driven and physical justification for each are shown in Table 1.

Linear models with the features in Table 1 gave good accuracy for predicting the electronic structure parameters (see Figure 2a,b,c). The mean absolute error (MAE) and root-mean-squared error (RMSE) of the test sets, averaged over 3 different test-train splits, are shown. Further, when combining these models with Equation (2) to predict

Table 1: Features used in models to predict electronic structure parameters.

Model (Output) <sup>a</sup>	Feature (Input) <sup>b</sup>	Data-driven Justification	Physical Justification
$\epsilon_d$	$f_{d,n}$	Lowest test error in one-term and two-term models; highest rank in ridge regression	d-d hybridization, electron transfer, diagonal disorder
$\epsilon_d$	$\sigma_n$	Lowest test error for two term model with $f_{d,n}$ ; low test error for three term model with $f_{d,n}$ and $V_{pd}^2$	Electronic interaction between neighbors
$\epsilon_d$	$V_{pd}^2$	Lowest test error (within uncertainty) for three term model with $f_{d,n}$ and $V_{pd}^2$	p-d and s-d hybridization
$n_p$	$V_{pd}^2$	Lowest test error in one-term model (within uncertainty); two term models provide little improvement	p-d and s-d hybridization
$V_{ad}^2$	$f_{d,n}$	Lowest test error in one-term model; two term models provide no improvement.	d-d hybridization, electron transfer, diagonal disorder

<sup>a</sup> d-Band center,  $\epsilon_d$ ; number of p electrons,  $n_p$ ; adsorbate-surface matrix coupling element,  $V_{ad}^2$ .

<sup>b</sup> Neighbors' d-band filling,  $f_{d,n}$ ; neighbors' bulk electrical conductivity,  $\sigma_n$ ; p-d coupling to neighbors,  $V_{pd}^2$ .

adsorption energies, we obtain an RMSE of 0.41 eV (see Figure 2). For bridge and hollow sites, the surface parameter values were averaged over the surface atoms in each site. The RMSE varies significantly across adsorbates and sites, ranging from roughly 0.15 to 0.6 eV. Adsorbates that adsorb more strongly and have a larger variation in their adsorption energies have a higher RMSE. However, the root-mean-square percent error (RMSPE) is much more constant, with a range of roughly 6 to 15%, and an overall value of 10%. 75% of the data points have a percentage error less than 10%, and 94% of points have a percentage error less than 20%. Hence, our model is accurate enough to be quite useful in initial screening. While comparing to previous models is difficult due to differing data sets, our errors are in the ranges achieved by other models, but less data is required.<sup>10,12</sup> Error distributions are given in Figure S1. We attribute the relatively low error on a quite heterogeneous dataset to our strategy of decoupling the predictions of the surface properties, which were fit separately for each metal, from the prediction of the adsorption energies, which were fit separately for each adsorbate and site. This addresses the combinatorial challenge of having many metals and many adsorbates, while still allowing each sub-model to be relatively simple.

Constructing linear models with a small number of features has several benefits, as this makes the model easier to interpret and fit. Adding additional features beyond those

in Table 1 did not improve the accuracy significantly, and hence we kept the number of features small in order to preserve these benefits. Further, we tested a few non-linear models, such as kernel ridge regression, and found that the test error was not significantly lower than the linear models. Our analysis suggests that the predictions of the adsorption energies from the electronic structure parameters (Equation (2)) contributes more to the overall error than the predictions of the electronic structure parameters (Table 1).

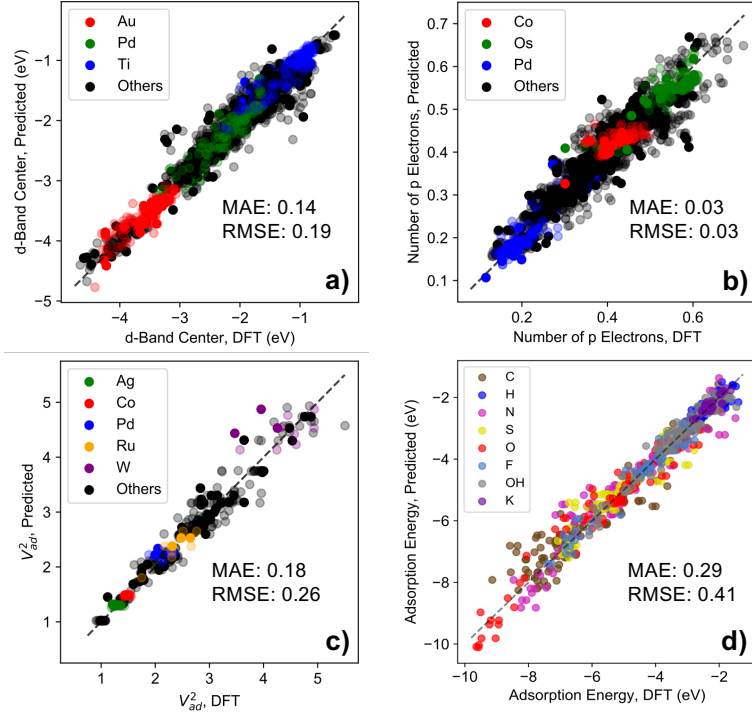


Figure 2: The accuracy of our predictions for a) the d-band center (2273 data points); b) the number of p electrons (2273 data points); and c) the matrix coupling element (335 data points). Test-set data points are opaque, while training-set data points are translucent. Test-set errors, averaged over three test-train splits, are shown. d) The accuracy of our predictions for the adsorption energy from the surface structure, using our predictions for the electronic structure parameters and Equation (2), as shown schematically in Figure 1 (685 data points).

The model sheds insight into the effect of different structural features on electronic structure and adsorption. For metals near the lower right part of the d-block, a high  $f_{d,n}$  leads to a lower  $\epsilon_d$ . This weakens adsorption and is likely due to increased d-d hybridization. The opposite is true for metals near the upper left part of the d-block, where a high  $f_{d,n}$  leads to a higher  $\epsilon_d$ . Hence, metals with mostly empty d bands have little d-d hybridization with metals with nearly full d-bands, due to differences in electronic

energies. In contrast, increasing  $\sigma_n$  raises  $\epsilon_d$  (stronger adsorption) for the lower right part of the d-block but lowers it (weaker adsorption) for the upper left part of the d-block. A higher  $V_{pd}^2$  leads to a higher  $n_p$  for all metals, which strengthens adsorption. For most metals, a higher  $V_{pd}^2$  also leads to a lower  $\epsilon_d$ , which weakens adsorption. In most cases, a higher  $f_{d,n}$  of neighboring atoms leads to a higher  $V_{ad}^2$ , which tends to strengthen adsorptions for adsorbates with no lone pair and/or metals with a low  $f_{d,n}$ . Therefore, the model allows trends in adsorption energies to be traced back to both electronic structure properties and underlying structural features.

To demonstrate the utility of our model, we tested whether it can be used to predict experimental trends. We also compared our adsorption energy predictions to experimental measurements of catalytic activity. To do this requires a volcano plot (or other relationship) that can predict the catalytic performance for a particular reaction and set of experimental conditions based on the adsorption energies of the relevant species. Hence, we developed volcano plots using our model predictions and the experimental data, and then tested how well this scheme predicts catalytic performance.<sup>40–42</sup> As shown in Figure S2, our model accurately captures experimental trends in catalytic performance for pure metals for ammonia synthesis, and for Pt alloys of two different architectures for the oxygen reduction reaction. Our predictions also reproduce previous trends for Pt monolayers on Os<sup>43</sup> and Ru,<sup>44</sup> both of which were shown to be consistent with experiment. More details are given in the Supporting Information.

## Application to a Reaction Pathway

Because our model can predict the adsorption energy of various reaction intermediates, it can be used to generate energetics along a reaction pathway. As a simple example of this, we predicted the energetics of H<sub>2</sub>O decomposition on a variety of surfaces. We took a 3 × 3 surface cell of the close-packed surfaces of all of the fcc and hcp metals included in our fitting, and substitutionally doped into the surface layer each other metal included in our fitting. Up to three dopant atoms were included, in all symmetry-unique configurations. This results in 2612 surfaces. On each of these surfaces, we estimated the

adsorption energies of H, O, and OH across all sites, and took the minimum values. The thermodynamics of this reaction pathway vary quite significantly, and we can quickly predict the behavior of various surfaces for reactions that consume or produce water (see Figure 3).

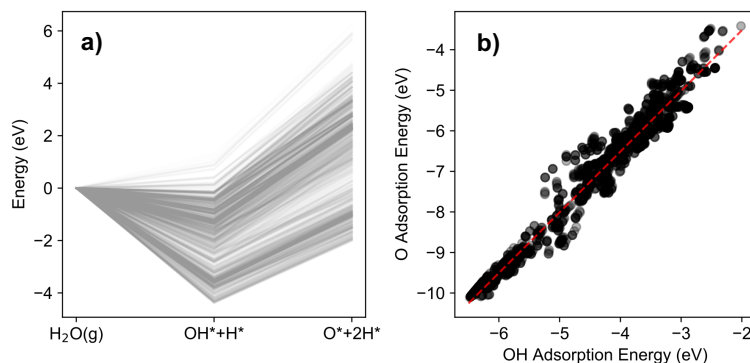


Figure 3: Predicting energetics along a reaction pathway. a) Energetics of  $\text{H}_2\text{O}$  decomposition on surfaces with up to 3 substitutional dopants (2612 surfaces). b) Minimum O adsorption energies as a function of minimum OH adsorption energies on these surfaces. The dashed line is a linear fit.

Because this framework does not assume the validity of scaling relations, which are linear relationships between adsorption energies,<sup>4,45</sup> it can be applied to surfaces that break these relationships.<sup>4,46</sup> Most of the 2612 surfaces examined here closely obey the scaling relation between O and OH, but many do not (see Figure 3). The maximum deviation from the fitted line is 1.2 eV for O and 0.8 eV for OH. Because strongly inhomogeneous surfaces are the most likely to break scaling relations, this gives an estimate for the maximum deviation from this relationship that can be achieved by alloying in close-packed surfaces.

## Application to Catalyst Screening

To understand general opportunities and challenges in catalyst design, we performed simple screenings for several important reactions on a large number of surfaces. We used our model to predict adsorption energies, which were used in conjunction with previously developed relationships between adsorption energies and catalytic activity. We took a  $3 \times 3$  surface cell of the close-packed surface of each fcc and hcp metal in

our set and substitutionally doped each other metal in our set into the top two layers. Up to 9 dopant atoms were included, and all symmetry-inequivalent configurations were included. Each symmetry-inequivalent hollow site on each surface was screened, for a total of approximately  $1.6 \times 10^7$  unique surface sites on  $10^6$  unique surfaces. Using our model, we estimated the adsorption energy of all 8 adsorbates in each of these sites. Then, we extracted relationships from the literature—generally volcano plots—that allow rates to be predicted from adsorption energies (see Figure 4h for an example). As shown in Figure 4, the surface sites generally span many orders of magnitude in predicted activity, as well as several orders of magnitude in estimated cost. Some pure metal surfaces in Figure 4 are shown as points of comparison, using the DFT-calculated adsorption energies in the lowest-energy site.

For two of the reactions, ammonia synthesis and hydrogen evolution, we performed additional DFT calculations of the adsorption energies to confirm that the predictions are reasonable (see Supporting Information). These calculations, which were performed after the model creation process was completed, demonstrate that the model is indeed useful for quickly estimating activity. The RMSPE of these 26 calculations of N and H adsorption energies is 6%, similar to that for the data in Figure 2d. This corresponds to an RMSE of 0.26 eV. Hence, our model is indeed useful for narrowing down a large phase space into a much smaller number of promising candidates.

Our model allows us to quickly identify promising candidates for further study. For example, the map of activity vs. cost for ammonia synthesis (Figure 4a) suggests that there are possibilities for designing catalytic surfaces that are cheaper and at least as active as Ru, the most active monometallic catalyst. For example, some CuFe surfaces are significantly cheaper and predicted to be slightly more active than Ru. Therefore, we calculated the N adsorption energies on two of these CuFe surfaces using DFT and found that the model predictions were accurate, and these surfaces indeed have higher activity than Ru based on the underlying volcano plot (see Figure 4a). We also performed a simple check of the stability of these surfaces. These surfaces have three to four Fe atoms in the surface, surrounding the N atom. We calculated the energy for these Fe atoms

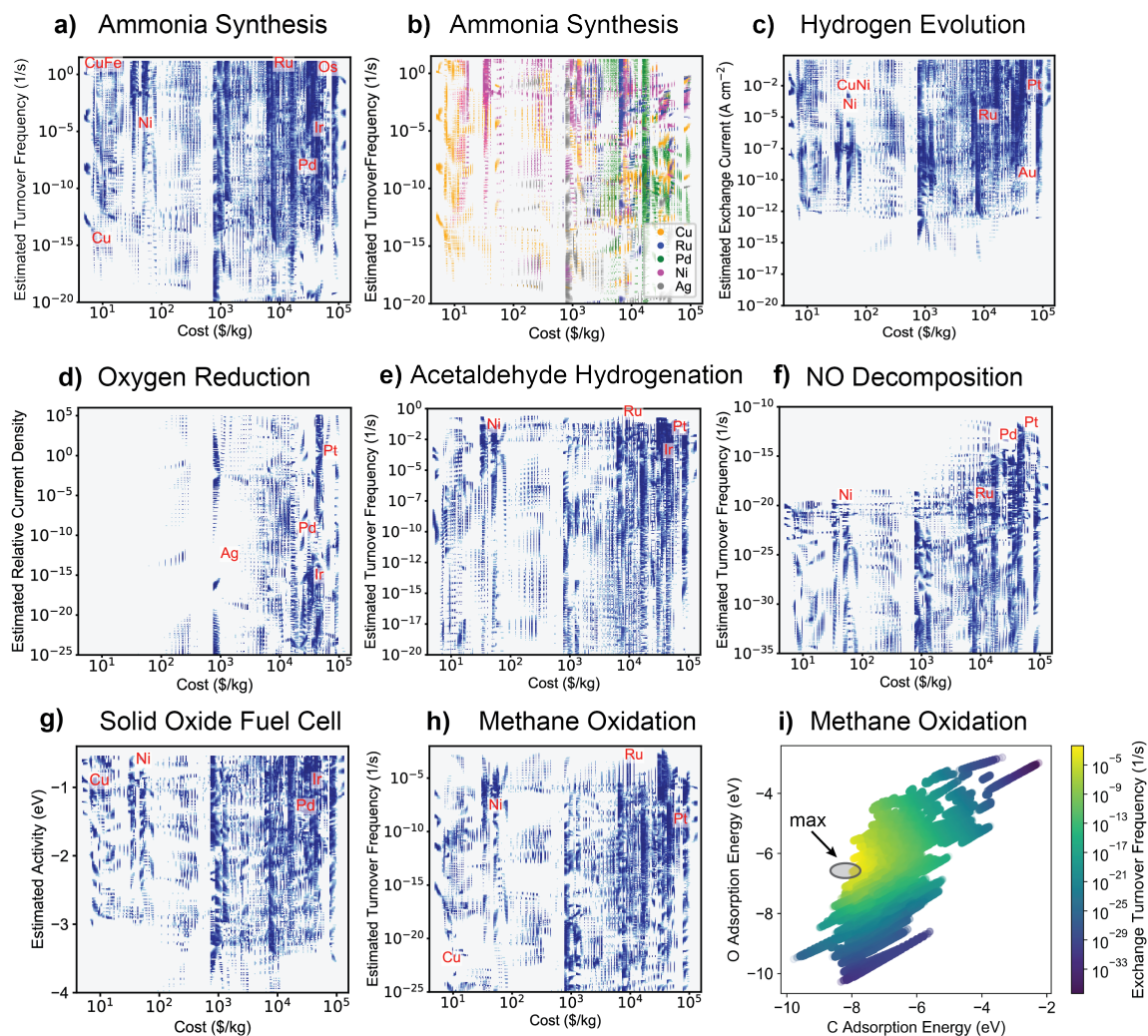


Figure 4: Density plots of screening for estimated catalytic activity and cost for  $1.5 \times 10^7$  surface alloy sites. Dark blue indicates that there are many sites in that region. a) Ammonia synthesis;<sup>47</sup> b) only a subset of surfaces are shown, with those that include particular metals colored separately; c) hydrogen evolution;<sup>48</sup> d) the oxygen reduction reaction (rate defined relative to Pt(111));<sup>35</sup> e) aqueous-phase hydrogenation of acetaldehyde;<sup>49</sup> f) NO decomposition;<sup>50</sup> g) the solid oxide fuel cell anode reaction (hydrogen oxidation), where activity is defined as the reaction energy of the rate determining step;<sup>51</sup> h) methane oxidation in a solid oxide fuel cell.<sup>52</sup> i) Activity for methane oxidation as a function of predicted O and C adsorption energies.<sup>52</sup>

residing in the subsurface, still in the presence of adsorbed N, and found that the Fe atoms are most stable on the surface. This suggests that the Fe surface sites may be stable under reaction conditions. We also identified a few surfaces that are predicted to be cheaper and more active than Pt for hydrogen evolution, such as NiCu (Figure 4c), and confirmed their adsorption energies using DFT. We can also see the range of possible activities and costs for alloys involving a particular metal. For example, there are few surfaces containing Ag that are highly active for ammonia synthesis (Figure 4b).

Screening of a large number of surfaces and sites allows broad insights into catalyst design and suggests where to focus for developing improved catalysts. For example, there are approximately  $10^6$  sites that have an estimated exchange current greater than  $10^{-5}$  A  $\text{cm}^{-2}$  for hydrogen evolution and a cost less than  $10^3$  \$/kg. All of these sites are on surfaces with either Cu or Ag. Hence, out of the phase space we consider here, it may be difficult to find a highly active, low-cost catalyst that does not contain Cu or Ag. For the oxygen reduction reaction, all of the less expensive surfaces (less than roughly \$100/kg) are quite inactive, and have estimated activity below the bottom limit shown in Figure 4d. Hence, it may be difficult to find an effective alloy catalyst for this reaction that is very low in cost. For NO decomposition, there is a clear tradeoff between cost and activity, such that it may be difficult to achieve both high activity and low cost.

For some reactions, such as NO decomposition and hydrogen oxidation in a solid oxide fuel cell, it is difficult to significantly outperform the pure metal surfaces in both cost and activity. Hence, using the design strategy of tuning adsorption energies by alloying, it may be difficult to make significant improvements on catalysts for these particular reactions. In the case of NO decomposition, the maximum rate reached on any of the surface alloys is still significantly lower than the theoretical maximum given by the volcano plot. This is due to the correlation between the adsorption energies of N and O.<sup>17</sup> To illustrate an additional challenge, we show the O vs. C adsorption energies, along with the estimated methane oxidation activity (Figure 4h). Only 20% of our data is shown, but this is representative of the full dataset. The optimal adsorption energies are on the edge of the region of Figure 4h that is accessible by these surface alloys. In the vast majority of cases,



C binding is weaker than optimal, and most sites that bind C with the optimal strength for methane oxidation also bind O too strongly. This suggests that other strategies to stabilize C relative to O may be needed to achieve the optimal catalyst.

This work shows the utility of having a very general model, despite the somewhat lower accuracy as compared to a more specific and complex model. A general model makes high-throughput screening easy for many reactions, and very high model accuracy is not needed as promising candidates can be explicitly checked using DFT. Indeed, we were easily able to identify promising candidates for ammonia synthesis and hydrogen evolution. Further, the large-scale conclusions we draw on the various reactions are insensitive to the error sizes that are present in the model. For example, it is clear that performing the corresponding DFT calculations would not affect our conclusion concerning inexpensive surfaces for the oxygen reduction reaction. For these types of conclusions, the ability to make efficient estimates for a large number of surfaces is more useful than a very low error for a smaller number of surfaces.

Our model also allows the use of a more careful screening methodology, with a negligible additional computational cost. For example, accounting for various sites and the heterogeneity of the surface or using a larger number of adsorption energies as descriptors may somewhat alter the trends shown in Figure 4. Nevertheless, our results show what is possible using high-throughput screening and existing descriptor sets.

## Conclusions

In this work, we have developed a general-purpose, reusable framework for high-throughput screening for catalyst design. We decoupled the prediction of the surface properties for various metals from the prediction of the adsorption energies from the surface properties. This novel, decoupled framework results in a data-efficient model that can be used in new contexts with no additional fitting or data generation. Adding additional species to further increase the generality requires a relatively small amount of additional data. We were able to use the model to quickly identify promising candidates

for highly active sites for ammonia synthesis and hydrogen evolution. As with all efficient models, promising candidates should be checked with DFT prior to experimental synthesis and testing.

This framework may also be useful in related applications, such as corrosion, surface stability, and electrochemistry, where adsorption is important. It can also be used to help detect errors when performing high-throughput calculations. Additionally, it can be used to increase the efficiency of DFT calculations in a variety of situations, by quickly eliminating sites or mechanisms that are very high in energy. In many studies, a large number of DFT calculations are performed, but in the end only a few are used: the lowest energy sites, configurations, or pathways, or the most promising candidates. Having a method to quickly estimate energies can significantly cut down the number of calculations that are needed.

Approximate screening of a large number of surfaces gives insight into what adsorption energy combinations are possible for surface alloys. For methane oxidation, these surface alloy sites barely reach the edge of the volcano-plot maximum, suggesting that other strategies for stabilizing C may be needed to develop the optimal catalyst. For NO decomposition, none of these surface alloy sites approach the volcano plot maximum, and other strategies for stabilizing N and/or destabilizing O are needed for further improvements. Additionally, we gained new insight into the cost-performance possibilities for surface alloys. For example, if a very low-cost catalyst for the oxygen reduction reaction is required for a particular application to be feasible, our results suggest that surface alloys may not be a promising set of materials to focus on. These insights allow design efforts to be focused where they are most likely to be fruitful.

Aside from its practical utility in screening, our framework provides links between alloy surface structure, electronic structure, adsorption, and catalytic properties. Differences in adsorption energies can be traced back to either electronic structure properties or to underlying structural features. This is in contrast to DFT calculations or complex models, which are not directly interpretable.

This work shows how data-driven techniques can be combined with physical insight

and physical models to efficiently generate fitted models with general applicability. In many cases, it is more convenient and efficient to use a pre-built model than to develop a complex, custom built model for each application, despite the somewhat lower accuracy that comes with a more general model. Further, the relatively small number of parameters needed for our framework allows fitting with a relatively small amount of data, which requires less effort and may allow the use of more computationally intensive techniques to generate the underlying data. Because of its generality and easy implementation, this model can act as a consistent, efficient substitute for researcher intuition.

## **Acknowledgements**

We gratefully acknowledge support from Tulane University. C.F.N. was supported by the Louisiana Board of Regents Research Competitiveness Subprogram, contract number LEQSF(2019-22)-RD-A-20. This research was supported in part using computing resources provided by Technology Services at Tulane University, New Orleans, LA. Portions of this research were conducted with high performance computational resources provided by the Louisiana Optical Network Infrastructure (<http://www.loni.org>).

## **Notes**

The authors declare no competing financial interest.

## **Supporting Information**

The supporting information for this work contains error distributions, model validation on trends in experimental data, a test of the model on stepped surfaces, another visualization of the model accuracy, and an example alloy structure. An implementation of the model, called SurfEP, is available at [bitbucket.org/mmmontemore/surfep](http://bitbucket.org/mmmontemore/surfep).

## References

## References

- (1) van Santen, R. A., Neurock, M. and Shetty, S. G. (2010). Reactivity Theory of Transition-Metal Surfaces: A Brønsted–Evans–Polanyi Linear Activation Energy–Free-Energy Analysis. *Chem. Rev.* *110*, 2005–2048.
- (2) Jones, G., Jakobsen, J. G., Shim, S. S., Kleis, J., Andersson, M. P., Rossmeisl, J., Abild-Pedersen, F., Bligaard, T., Helveg, S., Hinnemann, B., Rostrup-Nielsen, J. R., Chorkendorff, I., Sehested, J. and Nørskov, J. K. (2008). First Principles Calculations and Experimental Insight into Methane Steam Reforming over Transition Metal Catalysts. *J. Catal.* *259*, 147–160.
- (3) Nørskov, J. K., Bligaard, T., Rossmeisl, J. and Christensen, C. H. (2009). Towards the computational design of solid catalysts. *Nat. Chem.* *1*, 37–46.
- (4) Montemore, M. M. and Medlin, J. W. (2014). Scaling relations between adsorption energies for computational screening and design of catalysts. *Catal. Sci. Technol.* *4*, 3748–3761.
- (5) Greeley, J., Jaramillo, T. F., Bonde, J., Chorkendorff, I. and Nørskov, J. K. (2006). Computational high-throughput screening of electrocatalytic materials for hydrogen evolution. *Nat. Mater.* *5*, 909–913.
- (6) Studt, F., Abild-Pedersen, F., Bligaard, T., Sørensen, R. Z., Christensen, C. H. and Nørskov, J. K. (2008). Identification of Non-Precious Metal Alloy Catalysts for Selective Hydrogenation of Acetylene. *Science* *320*, 1320–1322.
- (7) Rangan, M., Yung, M. M. and Medlin, J. W. (2011). Experimental and computational investigations of sulfur-resistant bimetallic catalysts for reforming of biomass gasification products. *J. Catal.* *282*, 249–257.

- (8) Xin, H., Holewinski, A. and Linic, S. (2012). Predictive Structure–Reactivity Models for Rapid Screening of Pt-Based Multimetallic Electrocatalysts for the Oxygen Reduction Reaction. *ACS Catal.* *2*, 12–16.
- (9) Inoglu, N. and Kitchin, J. R. (2011). Identification of Sulfur-Tolerant Bimetallic Surfaces Using DFT Parametrized Models and Atomistic Thermodynamics. *ACS Catal.* *1*, 399–407.
- (10) Hoyt, R. A., Montemore, M. M., Fampiou, I., Chen, W., Tritsarlis, G. and Kaxiras, E. (2019). Machine Learning Prediction of H Adsorption Energies on Ag Alloys. *J. Chem. Inf. Model.* *59*, 1357–1365.
- (11) Li, Z., Ma, X. and Xin, H. (2017). Feature engineering of machine-learning chemisorption models for catalyst design. *Catal. Today* *280*, 232–238.
- (12) Tran, K. and Ulissi, Z. W. (2018). Active learning across intermetallics to guide discovery of electrocatalysts for CO<sub>2</sub> reduction and H<sub>2</sub> evolution. *Nat. Catal.* *1*, 696–703.
- (13) Ras, E.-J. J., Louwense, M. J., Mittelmeijer-Hazeleger, M. C. and Rothenberg, G. (2013). Predicting adsorption on metals: Simple yet effective descriptors for surface catalysis. *Phys. Chem. Chem. Phys.* *15*, 4436–4443.
- (14) Andriotis, A. N., Mpourmpakis, G., Broderick, S., Rajan, K., Datta, S., Sunkara, M. and Menon, M. (2014). Informatics guided discovery of surface structure-chemistry relationships in catalytic nanoparticles. *J. Chem. Phys.* *140*, 094705.
- (15) Choksi, T. S., Roling, L. T., Streibel, V. and Abild-Pedersen, F. (2019). Predicting Adsorption Properties of Catalytic Descriptors on Bimetallic Nanoalloys with Site-Specific Precision. *J. Phys. Chem. Lett.* *10*, 1852–1859.
- (16) Andersen, M., Levchenko, S. V., Scheffler, M. and Reuter, K. (2019). Beyond Scaling Relations for the Description of Catalytic Materials. *ACS Catal.* *9*, 2752–2759.

- (17) Montemore, M. M. and Medlin, J. W. (2014). A unified picture of adsorption on transition metals through different atoms. *J. Am. Chem. Soc.* *136*, 9272–9275.
- (18) Medvedev, M. G., Bushmarinov, I. S., Sun, J., Perdew, J. P. and Lyssenko, K. A. (2017). Density functional theory is straying from the path toward the exact functional. *Science* *355*, 49–52.
- (19) Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O. and Walsh, A. (2018). Machine learning for molecular and materials science. *Nature* *559*, 547–555.
- (20) Hansen, K., Biegler, F., Ramakrishnan, R., Pronobis, W., Von Lilienfeld, O. A., Müller, K. R. and Tkatchenko, A. (2015). Machine learning predictions of molecular properties: Accurate many-body potentials and nonlocality in chemical space. *J. Phys. Chem. Lett.* *6*, 2326–2331.
- (21) Senftle, T. P., Hong, S., Islam, M. M., Kylasa, S. B., Zheng, Y., Shin, Y. K., Junkermeier, C., Engel-Herbert, R., Janik, M. J., Aktulga, H. M., Verstraelen, T., Grama, A. and Van Duin, A. C. (2016). The ReaxFF reactive force-field: Development, applications and future directions. *npj Comput. Mater.* *2*.
- (22) Sankey, O. F. and Niklewski, D. J. (1989). Ab initio multicenter tight-binding model for molecular-dynamics simulations and other applications in covalent systems. *Phys. Rev. B* *40*, 3979–3995.
- (23) Behler, J. (2017). First Principles Neural Network Potentials for Reactive Simulations of Large Molecular and Condensed Systems. *Angew. Chemie Int. Ed.* *56*, 12828–12840.
- (24) Grove, W. M. and Meehl, P. E. (1996). Comparative Efficiency of Informal (Subjective, Impressionistic) and Formal (Mechanical, Algorithmic) Prediction Procedures: The Clinical–Statistical Controversy. *Psychol. Public Policy, Law* *2*, 1–30.

- (25) Kresse, G. and Furthmüller, J. (1996). Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* *6*, 15–50.
- (26) Kresse, G. and Hafner, J. (1993). Ab initio molecular dynamics for liquid metals. *Phys. Rev. B* *47*, 558–561.
- (27) Perdew, J. P. and Wang, Y. (1992). Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B* *45*, 13244–13249.
- (28) Blöchl, P. E. (1994). Projector augmented-wave method. *Phys. Rev. B* *50*, 17953–17979.
- (29) Kresse, G. and Joubert, D. (1999). From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* *59*, 1758–1775.
- (30) Ulissi, Z. W., Tang, M. T., Xiao, J., Liu, X., Torelli, D. A., Karamad, M., Cummins, K., Hahn, C., Lewis, N. S., Jaramillo, T. F., Chan, K. and Nørskov, J. K. (2017). Machine-Learning Methods Enable Exhaustive Searches for Active Bimetallic Facets and Reveal Active Site Motifs for CO<sub>2</sub> Reduction. *ACS Catal.* *7*, 6600–6608.
- (31) Xin, H., Holewinski, A., Schweitzer, N., Nikolla, E. and Linic, S. (2012). Electronic Structure Engineering in Heterogeneous Catalysis: Identifying Novel Alloy Catalysts Based on Rapid Screening for Materials with Desired Electronic Properties. *Top. Catal.* *55*, 376–390.
- (32) Fang, S., Kuate Defo, R., Shirodkar, S. N., Lieu, S., Tritsarlis, G. A. and Kaxiras, E. (2015). Ab initio tight-binding Hamiltonian for transition metal dichalcogenides. *Phys. Rev. B - Condens. Matter Mater. Phys.* *92*, 1–15.
- (33) Harrison, W. A. (1989). *Electronic structure and the properties of solids: the physics of the chemical bond*. Dover Publications, Mineola.
- (34) Nayak, S., Bhattacharjee, S., Choi, J. H. and Lee, S. C. (2020). Machine Learning

- and Scaling Laws for Prediction of Accurate Adsorption Energy. *J. Phys. Chem. A* *124*, 247–254.
- (35) Viswanathan, V., Hansen, H. A., Rossmeisl, J. and Nørskov, J. (2012). Universality in Oxygen Reduction Electrocatalysis on Metal Surfaces. *ACS Catal.* *2*, 1654–1660.
- (36) Ong, S. P., Richards, W. D., Jain, A., Hautier, G., Kocher, M., Cholia, S., Gunter, D., Chevrier, V. L., Persson, K. A. and Ceder, G. (2013). Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Comput. Mater. Sci.* *68*, 314–319.
- (37) Montemore, M. M. and Medlin, J. W. (2014). Predicting and Comparing C–M and O–M Bond Strengths for Adsorption on Transition Metal Surfaces. *J. Phys. Chem. C* *118*, 2666–2672.
- (38) Mehta, P., Bukov, M., Wang, C. H., Day, A. G., Richardson, C., Fisher, C. K. and Schwab, D. J. (2019). A high-bias, low-variance introduction to Machine Learning for physicists. *Phys. Rep.* *810*, 1–124.
- (39) Los, J. H., Mottet, C. and Tréglia, G. (2013). Surface segregation trends in transition metal alloys. *Phys. Rev. B - Condens. Matter Mater. Phys.* *88*, 1–11.
- (40) Ganley, J. C., Thomas, F. S., Seebauer, E. G. and Masel, R. I. (2004). A priori catalytic activity correlations: The difficult case of hydrogen production from ammonia. *Catal. Letters* *96*, 117–122.
- (41) Zhang, J., Vukmirovic, M. B., Xu, Y., Mavrikakis, M. and Adzic, R. R. (2005). Controlling the Catalytic Activity of Platinum-Monolayer Electrocatalysts for Oxygen Reduction with Different Substrates. *Angew. Chemie Int. Ed.* *44*, 2132–2135.
- (42) Stamenkovic, V., Mun, B., Arenz, M., Mayrhofer, K., Lucas, C., Wang, G., Ross, P. and Markovic, N. (2007). Trends in electrocatalysis on extended and nanoscale Pt-bimetallic alloy surfaces. *Nat. Mater.* *6*, 241–247.



- (43) Tsai, H. C., Hsieh, Y. C., Yu, T. H., Lee, Y. J., Wu, Y. H., Merinov, B. V., Wu, P. W., Chen, S. Y., Adzic, R. R. and Goddard, W. A. (2015). DFT study of oxygen reduction reaction on Os/Pt core-shell catalysts validated by electrochemical experiment. *ACS Catal.* *5*, 1568–1580.
- (44) Yang, L., Vukmirovic, M. B., Su, D., Sasaki, K., Herron, J. A., Mavrikakis, M., Liao, S. and Adzic, R. R. (2013). Tuning the catalytic activity of Ru@Pt core-shell nanoparticles for the oxygen reduction reaction by varying the shell thickness. *J. Phys. Chem. C* *117*, 1748–1753.
- (45) Abild-Pedersen, F., Greeley, J., Studt, F., Rossmeisl, J., Munter, T. R., Moses, P. G., Skúlason, E., Bligaard, T. and Nørskov, J. K. (2007). Scaling Properties of Adsorption Energies for Hydrogen-Containing Molecules on Transition-Metal Surfaces. *Phys. Rev. Lett.* *99*, 016105.
- (46) Montemore, M. M. and Medlin, J. W. (2013). Site-Specific Scaling Relations for Hydrocarbon Adsorption on Hexagonal Transition Metal Surfaces. *J. Phys. Chem. C* *117*, 20078–20088.
- (47) Jacobsen, C. J. H., Dahl, S., Clausen, B. G. S., Bahn, S., Logadottir, A. and Nørskov, J. K. (2001). Catalyst design by interpolation in the periodic table: Bimetallic ammonia synthesis catalysts. *J. Am. Chem. Soc.* *123*, 8404–8405.
- (48) Nørskov, J. K., Bligaard, T., Logadottir, A., Kitchin, J. R., Chen, J. G., Pandalov, S. and Stimming, U. (2005). Trends in the Exchange Current for Hydrogen Evolution. *J. Electrochem. Soc.* *152*, J23.
- (49) Lee, J., Xu, Y. and Huber, G. W. (2013). High-throughput screening of monometallic catalysts for aqueous-phase hydrogenation of biomass-derived oxygenates. *Appl. Catal. B Environ.* *140-141*, 98–107.
- (50) Falsig, H., Shen, J., Khan, T. S., Guo, W., Jones, G., Dahl, S. and Bligaard, T. (2013). On the Structure Sensitivity of Direct NO Decomposition over Low-Index Transition Metal Facets. *Top. Catal.* *57*, 80–88.

- (51) Rossmeisl, J. and Bessler, W. G. (2008). Trends in catalytic activity for SOFC anode materials. *Solid State Ionics* 178, 1694–1700.
- (52) Kleis, J., Jones, G., Abild-Pedersen, F., Tripkovic, V., Bligaard, T. and Rossmeisl, J. (2009). Trends for Methane Oxidation at Solid Oxide Fuel Cell Conditions. *J. Electrochem. Soc.* 156, B1447–B1456.