Materials discovery for high-temperature, clean-energy applications using graph neural network models of vacancy defects and free-energy calculations

Matthew Witman,a‡ Anuj Goyal, b‡ Tadashi Ogitsu, c Anthony McDaniel,a‡ Stephan Lany,b‡

We present a graph neural network modeling approach that fully automates the prediction of the DFT-relaxed vacancy formation enthalpy of any crystallographic site from its DFT-relaxed host structure. Applicable to arbitrary structures with an accuracy limited principally by the amount/diversity of the data on which it is trained, this model accelerates the screening of vacancy defects by many orders of magnitude by replacing the DFT supercell relaxations required for each symmetrically unique crystal site. It can thus be used off-the-shelf to rapidly screen 10,000s of crystal structures (which can contain millions of unique defects) from existing databases of DFT-relaxed crystal structures. This modeling approach therefore provides a significant screening and discovery capability for a plethora of applications in which vacancy defects are the primary driver of a material’s utility. For example, by high-throughput screening the Materials Project’s metal oxides, we rapidly “re-discover” and identify new high potential candidate materials for hydrogen generation via solar thermochemical water splitting and energy storage, for CO₂ conversion via reverse water gas shift chemical looping, and for cathodes in solid oxide fuel cells. Thermodynamic modeling on the basis of the high-throughput screening results allows us to connect the predicted defect energies to high temperature process conditions relevant to the different application areas, and we extract the reduction entropies as an additional selection criterion for high-performance materials. Further model development and accumulation of additional training data will only serve to expand the significant utility of this generalizable defect model to solving materials discovery problems in clean energy applications and beyond.

1 Introduction

High-accuracy calculations of vacancy defect formation enthalpies elucidate the primary and critical figure of merit needed to assess a material’s utility across a large variety of applications. These can range anywhere from catalysis (e.g., oxides for water splitting)⁴₅ to degradation resistance in extreme environments (e.g., radiation hardness of transition metal dichalcogenides),⁶,⁷ to neuromorphic computing (e.g., tuning metal-to-insulator transition with oxygen vacancy formation),⁸ to multiferroics (e.g., oxygen vacancy induced magnetic phase transitions).⁹ Density functional theory (DFT) is the method of choice to compute these vacancy formation enthalpies in a high-throughput fashion. However, given the need for supercell construction, atomic force relaxation, and the general presence of multiple non-equivalent atomic sites, the computational effort of defect calculations far exceeds that of the computation of the ideal defect-free material in the primitive cell. Thus, explicit DFT defect calculations exist so far only for a small fraction of the O(100) compounds contained in existing computational databases like the Materials Project (MP),⁹¹² Open Quantum Materials Database (OQMD),¹³ NREL Materials Database (NRELMatDB),¹³¹⁴ Joint Automated Repository for Various Integrated Simulations (JARVIS),¹⁵ and the Quantum Point Defect database (QPDB).¹⁶ Additional, potentially combinatorial, complexity arises from the desire to predict defect behavior in non-ideal materials, e.g., in the presence of atomic site disorder. A successful surrogate modeling approach that avoids the cost requirements of DFT is therefore critical in order to efficiently facilitate materials discovery efforts in these applications.

Previous efforts to predict vacancy formation enthalpies span various methods and material classes within which the models are applicable. Notable examples include modeling vacancy formation enthalpy using a simple hand-derived or machine learning (ML) model based on hypothesized important features and descriptor derived properties to train ML regression models for defect property prediction in semiconductors. For 2D materials consisting of TMDs, hexagonal boron nitride, and other selected wide band gap 2D materials, similar utilization of hand-engineered features and random forests predicted vacancy defect formation enthalpies. These examples provide highly useful models in specific situations by capitalizing on physical intuition regarding important material descriptors. But they do not provide a generalized solution that (1) automatically predicts the vacancy formation enthalpy of any crystal site in any material class and (2) negates the need for manual feature-engineering of material descriptors.

Various deep learning techniques, such as graph or convolutional neural networks, can circumvent such limitations. Rather than requiring the hand-crafting of an input feature vector to describe a training example, the model directly learns a feature representation from iterative convolutions and non-linear transformations of the input data, i.e. a graph representation of the crystal structure. The general success of graph neural networks at performing property prediction on crystal structures is evident from their recent explosion in popularity across a host of materials science applications. In this work, we extend the graph neural network concept for directly predicting vacancy defect formation enthalpies (abbreviated dGNN for short). The only re-
quired input is the non-defected, DFT-relaxed host structure, and
the model output is the predicted relaxed vacancy formation en-
thalpy of any site in the structure. The surrogate model therefore
replaces the need to do an expensive DFT supercell relaxation
for each symmetrically nonequivalent atomic site in the host, of
which there can be a sizable number (up to about 100 in the
present work) in compositionally and structurally complex mate-
rials. Other quantities typically derived from a DFT-relaxed struc-
ture like band gap ($E_g$), effective electron mass ($m^*$), compound
formation enthalpy ($\Delta H_f$), and oxidation state ($s$) can readily be
encoded into the material's graph if available to further improve
prediction accuracy. Existing computational databases containing
10,000s of compounds with potentially up to millions of non-
equivalent defect sites can be screened in an automated fashion
using desktop computational resources with a prediction accuracy
limited mainly by the amount and diversity of the vacancy defect
training data.

As a specific use case, we first focus on the discovery of new so-
lar thermochemical (STCH) water splitting oxides, widely consid-
ered a promising route for renewable hydrogen production[37]–[39]
DFT investigations play a critical role in understanding these ma-
terials and helping to guide experiments,[20]–[25] but can only be ac-
complished on a limited scale. Due to the complexity of the asso-
ciated vacancy defect calculations using DFT, approximately one
year’s work was required to build our training database (~200
host oxide compounds consisting of ~1500 unique defects). Our
highly generalizable dGNN model then extends upon capabilities
of previous work that requires carefully hand-engineered features
in specific crystal systems (e.g., perovskites)[21]–[26] and obtains a
similar expected mean absolute error in oxygen vacancy forma-
tion enthalpy (MAE < 450 meV) in unseen compounds (assum-
ing its cations are represented in the training data). Depending
on the model’s defect predictions, oxide stability, and the string-
gency of these down-selection criteria, we then narrow down a
small number of priority candidates for experimental efforts from
10,000s of possible MP oxides (among which are known STCH oxides we “re-discover” through our screening procedure).

Finally, optimal candidates for other clean energy applications
are highlighted where target ranges for oxygen vacancy formation
enthalpies have been proposed as a primary metric for material
down-selection, such as catalysts for CO$_2$ conversion and cath-
odes for solid oxide fuel cells.[57]–[61] We even extend our analysis
to predict defect density for all materials at finite temperatures,
a capability that can only be achieved due to our method’s abil-
ity to rapidly predict vacancy formation enthalpies of all sites in
the crystal structure (i.e., account for configurational entropy).
Thus materials can be assessed beyond the static picture of single
vacancy defect calculations at 0K for a limited set of materials.
With continued accumulation of training data, this success paves
the way for automated materials discovery in other vacancy de-
fect dependent applications and lays the groundwork for more
complicated machine learning tasks such as correlated vacancy
formation enthalpy or vacancy mobility predictions. All data,
code, and scripts needed to reproduce this study are provided
open source.[19]

2 Results and Discussion

2.1 Developing a diverse database of vacancy defects

We developed an automated workflow for DFT vacancy defect
calculations (Figure 1) as follows: (1) import the stoichiometric
host oxide crystal structures from the ICST[50] that are avail-
able in NRELMatDB; (2) relax the host supercell in ferromagnetic
and different possible anti-ferromagnetic spin configurations; (3)
choose the minimum energy host atomic and magnetic structure;
and (4) perform point defect calculations on the supercell using
the automated defect framework.[21] In the data acquisition step,
we extract host properties (from step 3) and calculate vacancy
formation energies (after step 4), which then supplies input and
target properties, respectively, for training the machine learning
model. Further details on the DFT settings are provided in Section
[51]–[54]. DFT is currently the preferred method for high-
throughput supercell defect calculations. True benchmark calcu-
lations require total-energy methods beyond DFT, such as quan-
tum Monte Carlo or the random phase approximation, which are
currently available only for few defect systems and with restric-
tions in cell size and atomic relaxations.[43]–[50] Given the DFT-
relaxed crystal structure, $\varphi_h$, of the host oxide with total energy
$E_h$ and the DFT-relaxed defected structure, $\varphi_d$, with total energy
$E_d$ (calculated using the standard supercell approach of Ref. [61]),
we compute the enthalpy of defect formation via

$$\Delta H_d = E_d - E_h + \sum_i n_i \mu_{i,\text{ref}}^\varphi.$$  

Here the reference chemical potential of added or removed atoms
($n_i = -1$ and +1, respectively) are taken as the fitted elemental
reference energies (FERE)[14]–[57], $\mu_{i,\text{ref}} = \mu_{i,\text{FERE}}$, which improve the
description of thermochemical properties in DFT calculations.[62]
For an oxygen vacancy, Equation (1), simplifies to $\Delta H_{V_O} = E_{V_O} -$ $E_h + \mu_{O}^\varphi$.

Our final DFT training database consists of ~1500 unique de-
fected oxides from ~200 parent oxides, which span 15 cations (Mg,
Al, Ca, Ti, Mn, Fe, Co, Ni, Sr, Y, Nb, In, Ba, La, Ce), 63 space
groups, and 51 unique stoichiometries. The choice of chemi-
cal space (Figure 2) is motivated by previous literature,[59]–[61] which provides guidance on cations that form stable oxides and
play an active role in tuning oxygen vacancy formation energy via
their redox activity. Previous STCH material searches have targeted perovskite (ABO$_3$) stoichiometry and related structures
systems (cubic, orthorhombic, tetragonal). Our training data is
more diverse and spreads across all seven crystal structure sys-
tems and includes a wide range of stoichiometries to cover dif-
ferent coordination environments and metal ion oxidation states
varying from 2+ to 5+.

2.2 A graph neural network model for vacancy defects

We create a vacancy defect graph neural network (dGNN) sur-
rogate model $f_{\delta\text{GNN}}$, parameterized by weights $\theta$, of the general form

$$\Delta H_d = f_{\delta\text{GNN}} (\varphi_h, i, j, \varphi_d, s; \theta),$$  

Here the reference chemical potential of added or removed atoms
($n_i = -1$ and +1, respectively) are taken as the fitted elemental
reference energies (FERE)[14]–[57], $\mu_{i,\text{ref}} = \mu_{i,\text{FERE}}$, which improve the
description of thermochemical properties in DFT calculations.[62]
For an oxygen vacancy, Equation (1), simplifies to $\Delta H_{V_O} = E_{V_O} -$ $E_h + \mu_{O}^\varphi$.
that will drastically reduce the computational cost associated with Equation (1) when performing high-throughput materials screening. Here, $v_{i} = \{E_g, m^*, \Delta H_f\}$ refers to a set of global input features as derived from the host compound (e.g., the band gap, effective electron mass, and compound formation enthalpy) while $s = \{s_1, s_2, \ldots\}$ refers to site specific input features for each atom (e.g., $s_1$ is the oxidation state of atom 1 in the host structure). Intuitively, for example, we expect oxides with cations in high oxidation states to form O vacancies (i.e., to reduce) more easily than when cations are in lower oxidation states, hence motivating the inclusion of these properties derived from the relaxed host structure.

Requiring only $\mathcal{C}_h$ and the index of the atom to be defected, $i'$, as an input, the model negates the cost of a DFT supercell relaxation for each unique symmetry site when predicting $\Delta H_d$, and executing the ML screening is of negligible computational cost when querying $\mathcal{C}_h$ from existing repositories like MP or NRELMatDB. The inputs for model training (the host’s relaxed POSCAR file, formation enthalpy, bandgap, effective electron mass, and oxidation states) and neutral vacancy formation energies (for both oxygen and cations) are provided in our open-source data repository.

Our dGNN closely follows the original Crystal Graph Convolutional Neural Network (CGCNN) method of Ref. [31], for which we highlight the necessary modifications to predict defect formation enthalpies. The deep learning framework is composed of three major steps.

Crystal embedding. First, $\mathcal{C}_h$ is embedded as a graph with nodes, $V = \{v_0, v_1, \ldots\}$, corresponding to each atom and edges $B = \{b_{ij}\}$ corresponding to bonds that are defined between atoms $i$ and $j$ below a cutoff radius and up to a maximum number of nearest neighbors. CGCNN One-Hot encodes and concatenates a node’s elemental solid properties, $v_v = \text{One-Hot}(\text{Mendelev number, atomic weight, melting temperature, covalent radius, electronegativity, ground state volume per atom, ground state band gap, ground state magnetic moment, and space group number})$ as the initial feature vector, $v_v^{(0)} = v_v$. We additionally concatenate a One-Hot encoding of the site’s oxidation state, as calculated in our DFT defect database, to increase model accuracy

$$v_v^{(0)} = v_v \oplus \text{OneHot}(s_i).$$

Two atoms of the same element type are no longer guaranteed to share an identical encoding due to their dependence on $s_i$ since various elements can assume different oxidation states depending on their local environment. Since One-Hot encoding increases sparsity and dimensionality of the initial node feature vectors and removes any quantitative ordering of a physical property, we also investigate whether a purely continuous encoding strategy can improve performance. This alter-
native approach involves the scaling of each elemental property between [0,1], which can be combined with the element’s ground state electron configuration and valence electrons (e.g., \( \begin{bmatrix} 1s^2 \ 2s^2 \ 2p^6 \ 3s^2 \ 3p^6 \ 4s^2 \ 3d^{10} \ 4p^6 \ 5s^2 \ 4d^6 \ 5p^6 \ \end{bmatrix} \)).\textsuperscript{33} to yield a continuously-scaled elemental representation, \( \mathbf{v}_i \). This vector can further be concatenated with \( \mathbf{s_i} \), scaled between [0,1], and preserves ordering of physical quantities.

CGCNN\textsuperscript{31} also utilized a discretized Gaussian filter applied to the distance between two atoms, \( r_{ij} \), to generate the initial edge features, \( \mathbf{b}_{ij} \), but this representation can be sparse and sensitive to an arbitrary choice of the filter’s standard deviation. A more systematic approach is to use a radial basis set expansion (e.g., similar Behler and Parinello’s G\textsuperscript{30} function\textsuperscript{22}) to generate the initial feature vector according to a set of gaussian widths (\( \eta \)), centers \( (\mathbf{R}_e) \), and a cutoff radius \( (r_c) \):

\[
\mathbf{b}_{ij} = \{ \exp \left[ -\eta (r_{ij} - R_s)^2 / r_c^2 \right] \}. \tag{4}
\]

This simultaneously reduces the sparsity and dimensionality of \( \mathbf{b}_{ij} \) for a moderate basis set size (\( \eta = \{0.5,1.0,1.5\} \), \( R_s = \{1.0,2.0,3.0,4.0,5.0\} \)). Such modifications facilitate distinguishing between very similar crystal structures.

Convolutions. Automated feature extraction is then performed, whereby the feature vector for each node is iteratively updated via convolutions with its neighboring nodes and edges. Mathematically, the output of the \( i \)th convolutional layer can be written as\textsuperscript{31}

\[
v_i^{(l+1)} = g \left( v_i^{(l)} + \sum_j \sigma \left( z_{ij} W_1^{(l)} + b_j^{(l)} \right) \otimes g \left( z_{ij} W_2^{(l)} + b_j^{(l)} \right) \right). \tag{5}
\]

Here \( z_{ij} = \mathbf{v}_i \oplus \mathbf{v}_j \oplus \mathbf{b}_{ij} \) is the concatenation of information from connected nodes in the graph, \( \mathbf{W}_1, \mathbf{b}_1 \) and \( \mathbf{W}_2, \mathbf{b}_2 \) represent weights and biases of different learnable weight matrices (i.e., fully connected neural network layers), \( \sigma \) denotes a sigmoid activation function, \( g \) denotes a softplus activation, and \( \otimes \) denotes element-wise multiplication. This step remains unaffected for the dGNN.

Property prediction. Following \( T \) total convolutions, the original CGCNN method acquires an overall feature vector representing the crystal by pooling all nodes in the structure

\[
\mathbf{v}_c = \text{Pool} \left( v_0^{(T)}, v_1^{(T)}, ..., v_N^{(T)} \right), \tag{6}
\]

which, for example, consists of a summation operation. Thus crystals of arbitrary size are described by a vector of the same dimensionality. The model predicts some final global property by applying one (or more) fully connected layers to \( \mathbf{v}_c \).

However, we are interested in the defect formation enthalpy and therefore isolate the information contained only on the host node/atom to be defected (specified at index \( i' \)) following the \( T \) total convolutions. We therefore replace eq. (6) with

\[
\mathbf{v}_d = g \left( (\mathbf{v}_{i'}^{(T)} \oplus \mathbf{v}_e) \cdot \mathbf{W} + \mathbf{b} \right). \tag{7}
\]

We incorporate the global compound features at this step, \( \mathbf{v}_e = \{ E_g, m^*, \Delta H_f \} \), before applying the subsequent fully connected layer. Additional feed-forward layers may be applied before the final property prediction of \( \Delta H_f \) is then computed through one final output layer. Equation (7) is specifically designed for predicting defect formation enthalpies for the limit of infinite dilution in this work, but could be changed to, for example, expand its applicability to correlated vacancy defects.

All training data, model structures, and hyperparameters used in this study can be found in our Zenodo repository, from which all results can be reproduced when used in conjunction with the dGNN implementation built in Pytorch\textsuperscript{85} provided at \url{https://github.com/mwitman1/cgcnndefect/tree/Paper1} (modified from Ref.\textsuperscript{31}). In brief, due to the small size of our training data set, a dGNN model of minimal complexity is required to facilitate training. In practice the number of trainable parameters often exceeds the number of training examples in deep learning applications and explicit regularization isn’t even always needed to achieve low generalization error.\textsuperscript{85} Nonetheless, our minimal complexity architecture consists of \( T=2 \) convolution steps, \( \mathbf{v}_i \in \mathbb{R}^8 \), and \( \mathbf{v}_d \in \mathbb{R}^{16} \), leading to a GNN with only \( \sim 2,000 \) learnable parameters. After fixing the architecture, the learning rate was adjusted to minimize K-fold test set performance (next section), and over-fitting was minimized via early-stopping using the mean absolute error on a 10% validation set within each train fold.

2.3 Defect GNN validation and performance

We executed three different cross-validation (CV) strategies, demonstrated by the toy examples in Figure 3, to gauge model performance. For defect-wise and compound-wise CV, we utilized \( K = 10 \)-fold cross validation. For each \( k \)-fold, 10% of the training data is held as a validation set for early stopping, and the mean absolute error over all \( n \) defects in the test set is computed, \( \text{MAE}_k = 1/n \sum_i |\Delta H_{d,n} - \Delta H_{d,n}^k| \). The model’s expected prediction error is then estimated across all folds, \( \langle \text{MAE}^k \rangle_k = (1/K) \sum_{k=1}^K \text{MAE}_k^k \). Here \( Y = O \) or \( Y = \) Other filters evaluation of the MAE by a defect’s specific element type to delineate model performance between oxygen and non-oxygen vacancy predictions. The standard deviation in MAE across all \( K \) models, \( \sigma_K(\text{MAE}^k) \), should be small once sufficient data has been collected, i.e., the sampled distribution of training data no longer changes significantly between each fold.

Validating the dGNN with defect-wise CV is a less challenging task, since train and test sets may contain defects from the same material in similar chemical environments (i.e., just above the symmetry tolerance). Validating with compound-wise CV is more challenging, since the test set contains all the defects from a given material (none of which may appear in the training set). This better reflects the practical performance for materials discovery, because one is usually interested predicting for materials for which no DFT vacancy calculations have been performed and thus not a single defect site could exist in training set. Finally, element-wise, leave-one-out CV proves the hardest challenge, where all defects for a compound containing the held-out element are placed in the
Fig. 3 (a) Demonstration of different CV strategies in a toy dataset. In defect-wise CV, different defects from the same compound can appear in both train/validation and test splits, whereas in compound-wise CV, the train/validation and test stratification is performed such that all of compound’s defects can only appear in either train/validation or test splits. The number of compounds is kept constant between folds, leading to small variations in the number of unique defects contained per K-fold test set. In element-wise, leave-one-out CV, all defects in any compound containing the test defects can only appear in either train/validation or test splits. The number of compounds is kept constant between folds, leading to small variations in the number of unique defects contained per K-fold test set. (b) Summary of expected prediction errors and a heuristic uncertainty metric (mean MAE and standard deviation of MAE across K-fold test sets) for different cross-validation (rows) and graph encoding strategies (columns). (c) MAE averaged across the CV test sets (K = 10) models as a function of training data size and the parity plots of the test set predictions with DFT for the “Full” encoding strategy. (d) MAE for the element-wise, leave-one-out CV, separated by test element and encoding strategy. (e) Test set predictions for the $X=\{\text{Y}, \{\text{La}\}, \{\text{Fe}\}, \{\text{Al}\}, \{\text{Co}\}, \{\text{Ti}\}, \{\text{Nb}\}\}$ models.
test set (thus there are X = 15 models and test sets, corresponding to the number of cation types in the training data). At inference time, the ML model is supplied with a graph encoding containing node features that may not be represented by the training data. Figure 3 shows the evaluation metrics for each CV strategy, with the MAE separated by predictions on oxygen and all other vacancy types. Cation vacancies tend to have higher enthalpies, but the mean relative errors between O and non-O defects are quite similar. This CV analysis was repeated for four different encoding strategies, “Element-only” encoding uses $v_x$, while $s_i$ and $v_y$ are empty. Conversely, “DFT-only” encoding keeps $s_i$ and $v_y$ but $v_x$ is empty. The “Full” encoding strategy incorporates all information, while the “Continuous” encoding replaces the one-hot encoding of elemental properties and oxidation state with $v'_x$ and $s'_i$. Three key observations arise from the defect-wise and compound-wise CV. First, our strategy to encode both elemental and DFT data via the “Full” and “Continuous” models provides improved accuracy; nonetheless, “Element-only” and “DFT-only” encoding strategies both lead to reasonable accuracy on their own. Second, the best accuracy for compound-wise CV is obtained with the “Full” model with a $\langle MAE \rangle K < 0.45$ eV. Third, the ability of the model to predict $\Delta H_d^{\text{Other}}$ validates the generality of this approach and its usefulness in other applications requiring predictive modeling beyond oxygen vacancies.

Figure 3 demonstrates the continued decrease in the CV MAE as more defects are added to the training data. While the MAE decreases with defect-wise CV, it starts to plateau, the compound-wise CV still benefits from a log-linear decrease in MAE with increasing data, highlighting the model can still be significantly improved as more training data is acquired. This highlights the need for continued, high-quality DFT defect calculations with automated workflows (see Section S1 for recommended settings to build a larger training dataset consistent with this work). The parity plots correspond to the test set results concatenated across all CV models, with good performance between both oxygen and non-oxygen vacancy defect predictions. Figure 3 shows the results of the element-wise, leave-one-out CV for each of the four different encoding strategies. While this is not a task that one might expect a deep learning approach to succeed at, several elements are well-predicted and close to the target error of 0.5 eV, regardless of encoding type. Yet other elements are poorly predicted for all encoding types. Without any a priori knowledge of which held-out elements are well-predicted, it is not possible to assume that predictions on materials with unseen cation types can achieve $\langle MAE \rangle K < 0.97$ eV (Figure 3). We therefore recommend this current generation of models only be used on compounds containing the fifteen cations spanned by the training database. Nonetheless, Figure 3 demonstrates how the continuous encoding strategy significantly improves property prediction across selected element types. This could be due to the continuous encoding’s preservation of quantitative periodic trends (e.g., electronegativity, mean volume per atom, etc.). In order to expand the quantitative applicability of the model for materials discovery outside of the current cation set, we plan to include a larger chemistry space in the future.

2.4 Comparison with previous modeling efforts

Our cross-validation performance with the current training dataset is comparable to the performance of linear models derived via careful feature engineering to predict neutral oxygen vacancy formation enthalpies. Deml et al. trained a model on 45 binary and ternary oxides and achieved around 0.4 eV MAE on a small test set of 18 oxides that had not been included in model development. Wexler et al. computed SCAN+U vacancy formation enthalpies of 341 ABO$_2$ perovskites and derived a linear model that globally achieved an MAE of 0.7 eV and an MAE of 0.45 eV for the subset (142 materials) with hull energies less than 0.025 eV/atom. Another in depth validation of these studies with ours is presented in Section S3. These linear models are more interpretable than our approach due to the small number of manually derived features, but are less generalizable. For example, they are not capable of predicting neutral cation vacancies, whereas our trained model predicts either. Our model architecture was not designed specific to oxides and could be applied to any material class since it operates generally on any crystal structure, whereas the linear models contain features that can only be calculated if the structure contains O. Figure 3 also demonstrates strong evidence for continued significant model improvement with more training data, which is unclear for the linear models. This pattern of comparable accuracy but increased generalizability also holds true for our method relative to other machine learning efforts for defect predictions, e.g., Frey et al.’s model for transition metal dichalcogenides with MAE=0.67 eV and Cheng et al.’s model for amorphous GeTe. Finally, in a concurrent preprint with ours, Choudary et al. used graph neural networks models to predict total energy of a host structure and with an atom removed to estimate vacancy formation enthalpies, but this neglects the relaxation of the host upon vacancy formation and yields an MAE prediction of 1.5 eV for a single test set, including 2.3 eV for oxides.

2.5 Predictions on known STCH materials

Before utilizing the dGNN model to screen potential candidate materials for STCH water splitting, it is instructive to first validate against additional DFT calculations for experimentally known STCH oxides. The materials behavior for the STCH redox processes can be expressed in terms of reduction enthalpies and entropies. For workable and economic thermodynamic conditions, these considerations lead to a desirable value for the oxygen vacancy defect formation energy in an interval of about [2.3, 4.0] eV (Lieb et al. 2013). Lower formation energies impair the ability to produce hydrogen in the oxidation step, while higher energies prevent significant changes of the O stoichiometry in the reduction step. To predict the defect formation energies for these “unseen” materials, we utilize the expectation across K-fold models, $\langle \Delta H_d^{\text{F}} \rangle K = K^{-1} \sum_{k=1}^{K} \Delta H_d^{\text{F}}$, and the standard deviation, $\sigma(\Delta H_d^{\text{F}}) = \sqrt{K^{-1} \sum_{k=1}^{K} (\Delta H_d^{\text{F}} - \langle \Delta H_d^{\text{F}} \rangle)^2}$, as a heuristic estimate of the uncertainty in the property prediction. As a first test case, we consider the family of BXM oxides (B=Ba; X=Ce,Nb,Pr; M=Mn) which have a higher degree of compositional complexity than the materials in the training...
dataset. Figure 4a compares the $\langle \Delta H_d^{\phi} \rangle_K$ with DFT for BXM oxides for both “Full” and “DFT-only” encoded models. On average, “DFT-only” predictions are only more accurate when structures contain unseen cation types missing from the training set (e.g., Pr), but are generally less accurate if all of a structure’s cation types are in the train set (Figure 2c,d). Nonetheless, both models are quite similar in qualitatively ranking and quantitatively capturing $\Delta H_d^{\phi}$. We additionally tested our model on newly discovered disordered perovskite oxides, $\text{Sr}_{1-x}\text{Ce}_x\text{MnO}_3$ (SCM), which demonstrate suitable STCH performance. In contrast to the BXM family which are line compounds, the SCM family can accommodate a wide range of Ce doping, thereby providing better control over water splitting capabilities by varying Ce concentration. To model SCM alloy structure we employ the special quasirandom structure (SQS) approach and generate two 80-atom supercell structures representing random alloying with differing Ce content. Figure 4b shows the “Full” encoding model predictions on both SQS’s. While the absolute value of the predictions are slightly below the generally accepted optimal range of [2.3, 4.0] eV, the ML model predicts the SQS with higher Ce content to have oxygen vacancy enthalpies closer to the target range, which is experimentally consistent with its improved water splitting capabilities from the increased Ce content.

2.6 High-throughput screening for new STCH materials.

Our “element-only” encoding models can perform consistent predictions on input host crystal structures in which the DFT settings for host relaxations are not identical to those used to create the training set (see Section S2 and Figure S2 for details). Thus we can now efficiently predict $\Delta H_d$ across 10,000s of DFT-relaxed crystal structures included in open source repositories beyond NREL MatDb, such as Materials Project (v2021.03.22). We employ the “Element-only” encoding model at the expense of slightly lower accuracy since Materials Project (MP) data doesn’t necessarily contain all the features needed for the “Full” encoding models. Figure S3 shows that the space of ~35,000 oxides (excluding non-metals) is reduced to about ~2,200 structures by setting a maximum energy above the hull requirement, $E_H < 0.1$ eV/atom, and discarding any materials with cations not present in the training set. From these remaining ~2,200 host oxides, we predict $\langle \Delta H_d \rangle_K$ for the ~48,000 symmetrically unique defect sites. Before proceeding with candidate down-selection, we perform yet another hold-out validation by comparing our MP screening predictions with existing first principles calculations and model predictions that could be easily mined from the literature (see Section S3).

Although we omit host structures with cations outside the training set in this study, adding a small number of training structures in the future (the full search space from Figure 2) could further expand the model’s applicability. Figure S5 shows the predicted $\langle \Delta H_d^{\phi} \rangle_K$ vs. $\Delta H_f$, and, although correlation is evident across the entire enthalpy range, there is little correlation within [2.3, 4.0] eV. Furthermore, a simple model using features derived only from the host composition would clearly be insufficient due to an inability to distinguish individual oxygen vacancies. Within a single structure, these can span a very large range as shown for the min, median, and max predictions for MP structure mp-1247717 ($\text{Ca}_4\text{Mn}_3\text{AlO}_11$).

Assessing an oxide’s STCH potential first requires determining $\Delta H_d^{\phi}$ of all sites and computing the fraction above the minimum threshold of 2.3 eV, denoted $x_{\text{min}}$, and the fraction of defects in the optimal range of [2.3, 4.0] eV, denoted $x_{\text{mg}}$. If the material contains any defects below the target range, the reduced metal oxide cannot readily be regenerated at the oxidation conditions relevant for STCH, and thus we require $x_{\text{min}} = 1$. It may be ideal for all defects to fall within in the target range ($x_{\text{mg}} = 1$) to increase capacity and defect mobility, but these considerations are beyond the scope of this study. In practice, only one defect needs to fall within the targeted range to be considered a promising STCH material ($x_{\text{mg}} > 0$).

Recent total energy calculations for defects in hercynite Fe$_2$O$_3$ have found good agreement between DFT+U, hybrid functional, and the random phase approximation, but uncertainties on the order of a few tenths of an eV should be expected for DFT calculations in transition metal oxides. We therefore extend the defect screening metrics to be uncertainty inclusive ($x_{\text{min,1}}, x_{\text{mg,1}}$), agnostic ($x_{\text{min,2}}, x_{\text{mg,2}}$), or exclusive ($x_{\text{min,3}}, x_{\text{mg,3}}$) for increasingly strict down-selection. Given a host’s $N_s$ symmetry sites and the set of all predictions $(\langle \Delta H_d^{\phi} \rangle_K \equiv H = \langle H_1 \cdots H_{N_s} \rangle)$...
and uncertainties \( \{ \sigma_k(\Delta H^0_k) \} \equiv U = \{ U_1 \ldots U_K \} \), these defect fractions can be calculated subject to a criteria \( C \),

\[
x = \frac{1}{N_X} \sum_{i=1}^{N_X} \begin{cases} 
1 & \text{if } C \\
0 & \text{otherwise}
\end{cases}
\]

which is summarized in Table 1. Figure 5 shows the impact of applying these increasingly strict defect criteria and that a significant number of candidates remain even when using an uncertainty exclusive filter and requiring \( x_{\text{rng}} \to 1 \). See Section S4 for a more detailed discussion on uncertainty.

Down-selection must also consider host oxide stability under STCH relevant conditions. For oxygen chemical potential \( \mu_O = \mu_O^{\text{ref}} + \Delta \mu_O \), typical STCH operating conditions necessitate that the host is stable in a target range \( \Delta \mu_O^{\text{target}} = [-3.0, -2.5] \text{ eV} \).

Given the compound’s energy above the hull in the grand ensemble, \( \phi_U(\Delta \mu_O) \), we define the chemical potential range over which the host stability is below some threshold \( X \),

\[
\Delta \mu_O^{\phi_U \leq X} = [\Delta \mu_O(\phi_U(\Delta \mu_O)) \leq X].
\]

Setting \( X \leq 0.1 \text{ eV/atom} \), for example, helps avoid false negatives during materials’ selection due to the synthesizability of metastable structures or due to uncertainties originating from the specific DFT approach and the convex hull analysis. Our final down-selection criteria requires that the target and stability chemical potential ranges intersect,

\[
\Delta \mu_O^{\phi_U \leq X} \cap \Delta \mu_O^{\text{target}} \neq \emptyset.
\]

The vacancy defect fractions and host stability criteria can be tuned for custom down-selection using our open access data and post-processing scripts.

Figure 5d shows how increasingly stringent \( x_{\text{min}}, x_{\text{rng}}, \) and \( \Delta \mu_O^{\phi_U \leq X} \) criteria can narrow the candidate space from thousands of oxides to just a handful. At each down-selection criteria...

Table 1 Criteria for Equation (8) to determine the defect fractions, \( x_{\text{min}} \) and \( x_{\text{rng}} \), for increasingly strict uncertainty inclusive, agnostic, or exclusive down-selection.

<table>
<thead>
<tr>
<th>Uncertainty</th>
<th>( C ) for ( x_{\text{min}} )</th>
<th>( C ) for ( x_{\text{rng}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inclusive</td>
<td>( H_i + U_i \geq 2.3 ) ( [H_i - U_i, H_i + U_i] \cap [2.3, 4.0] \neq \emptyset )</td>
<td>( H_i \geq 2.3 ) ( [H_i - U_i, H_i + U_i] \in [2.3, 4.0] )</td>
</tr>
<tr>
<td>Agnostic</td>
<td>( H_i &gt; 2.3 ) ( [H_i - U_i, H_i + U_i] \cap [2.3, 4.0] \neq \emptyset )</td>
<td>( H_i &gt; 2.3 ) ( [H_i - U_i, H_i + U_i] \in [2.3, 4.0] )</td>
</tr>
<tr>
<td>Exclusive</td>
<td>( H_i - U_i \geq 2.3 ) ( [H_i - U_i, H_i + U_i] \cap [2.3, 4.0] \neq \emptyset )</td>
<td>( H_i - U_i \geq 2.3 ) ( [H_i - U_i, H_i + U_i] \in [2.3, 4.0] )</td>
</tr>
</tbody>
</table>

Fig. 5 (a) Removing MP oxides containing non-metals, those with \( E_H > 0.1 \text{ eV/atom} \), and those with elements outside the training set yields \(~2.200\) oxides for screening. (b) Screening predictions using the “Element-only” encoding, defect-wise CV models vs. \( \Delta H_f \). (c) The cumulative histogram of compounds that contain a specific fraction of defects predicted to satisfy the criteria \( x_{\text{rng}} \) as defined in Table 1. (d) Utilization of increasingly stringent defect and host compound criteria \( (x_{\text{rng}}, x_{\text{min}}, E_H, \) and \( \Delta \mu_O^{\phi_U \leq X}) \) to narrow the list of potential STCH candidates. At each down-selection stage, we show the number of unique remaining formulas, the number of which overlap with the training data, and an exemplar structure.
we have highlighted one material among many that appear interesting due to relatively low prediction uncertainty, relatively wide stability range, high cation/structure complexity, verified experimental synthesis, and a lack of any STCH-specific experimental investigations that we are aware of. Importantly, this screening “re-disCOVERs” the known STCH material BCM-12R. However, it is eliminated by stricter down-selection criteria since the “Element-only” encoding model under-predicts the oxygen vacancy enthalpies with \{2.0 \pm 0.3, 2.9 \pm 0.4\} eV for the two O sites in BCM, compared to the “Full” encoding predictions, \{2.6 \pm 0.3, 3.8 \pm 0.5\} eV, and DFT predictions, \{2.66, 3.29\} eV. This emphasizes the prudence of considering materials inclusive of their uncertainty, especially since the “Element-only” model necessitated by MP screening has slightly higher CV error. Some oxides with a relatively simple composition satisfying strict down-selection criteria include: Mn$_2$O$_3$, Fe$_2$O$_3$, Fe$_2$O$_5$, Ba$_2$Fe$_2$O$_4$, Mn$_2$CoO$_4$, Mn(FeO$_2$)$_2$, Sr$_2$Mn$_2$O$_5$, Sr$_3$(FeO$_2$)$_2$, Ba(FeO$_2$)$_2$, Ba$_2$In$_2$O$_6$, Fe$_2$NiO$_4$, and Sr$_3$Mn$_2$O$_{$13/4}. All have been experimentally synthesized in the literature. Notably, several have already been investigated in the context STCH performance or other water splitting approaches.

Rediscovering these known STCH materials further validates our approach, and we have now identified many new candidates from which promising STCH materials can be experimentally targeted (see Table S6 for details and the comprehensive list). Raw data for all predicted defect properties and customizable open source scripts for reproducing or modifying the down-selection criteria can be found in the project’s Zenodo repository.

2.7 Beyond STCH: materials discovery across diverse energy applications

Our dGNN methodology and high-throughput database of vacancy predictions can be used to rapidly screen candidate materials in other important clean energy applications. The reverse water-gas shift-chemical looping (RWGS-CL) approach operates in a concept very similar to STCH. H$_2$ gas is used to reduce a metal oxide, after which oxidation with CO$_2$ produces CO for downstream hydrogenation to carbon-based fuels. Like STCH, the process is driven by thermodynamics of oxygen vacancy formation. Ref. 47 strongly correlated the energetics of vacancy formation to other computed properties like O$_2$ surface-adsorbate binding energy to conclude, “It is thus best said that $F_{vac}$ can solely describe the RWGS-CL process and is capable of predicting the CO$_2$ conversion ability of pervoskite oxides.” Based on a previously known CO$_2$-splitting pervoskite (La$_0.75$Sr$_{0.25}$FeO$_3$) for which they computed an average oxygen vacancy formation enthalpy of 3.4 eV, they concluded that candidate materials with similar thermodynamics (e.g., $\Delta H_0^f \in [3.0, 4.0]$ eV) would be highly active. This was confirmed by synthesizing new pervoskites with the desired vacancy thermodynamics and measuring their outstanding activity. By this metric, one can readily use our model to identify optimal candidates for RWGS-CL. Similar to our STCH screening, we “re-disCOVER” experimentally known CO$_2$-splitting oxides, e.g., the La$_2$MnCoO$_6$ system, while discovering new ones (Figure 5).

Oxygen vacancy formation enthalpy has also been correlated with the critical performance metrics for perovskite oxide cathodes in solid oxide fuel cells. Ref. 48 discovered a simple linear relationship for SOFC perovskite cathodes between $\Delta H_0^f$ and a metric for the macroscale oxygen-transfer performance, the area-specific resistance (ASR). The authors noted that successful materials should approximately be targeted with ASR \in [0.02 \Omega \text{ cm}^2, 0.24 \Omega \text{ cm}^2], or between the ASR values in the optimized Ba$_{0.5}$Sr$_{0.5}$Co$_{0.75}$Fe$_{0.25}$O$_3$ (BSCF) and La$_{0.625}$Sr$_{0.375}$Co$_{0.25}$Fe$_{0.75}$O$_3$ (LSCF) systems, respectively. This essentially represents the trade-off between the correlated stability (high defect formation enthalpy) and low operating temperature (low defect formation enthalpy). Using DFT to compute the average $\Delta H_0^f$ in model BSCF and LSCF crystal structures, the authors established guidelines that $\Delta H_0^f \in [0.7, 2.7]$ eV should be targeted.

Once again, we can rapidly target such materials with our approach. BCSF-like BaSr$_2$Fe$_5$(Co$_{0.75}$O$_{2.5}$) (mp-1099936) with $\Delta H_0^f \in [0.7, 1.5]$ and LSCF-like are Sr$_2$LaFe$_2$(CoO$_3$)$_3$ (mp-1218676) with $\Delta H_0^f \in [0.7, 1.5]$ are “re-disCOVERed” in the screening to further validate our approach. Interestingly, the BCSF-like structure is only metastable with $\min (\Delta H_0^f|_{\delta O=0}) = -0.88$ eV, while the LSCF-like is stable with $\min (\Delta H_0^f|_{\delta O=0}) = -0.35$ eV, an observation consistent with the reduced stability of BCSF. For example, some Fe-lacking analogs not discussed in Ref. 48 but computationally investigated elsewhere maintain low predicted vacancy formation enthalpy but improve upon phase stability according to MP phase diagrams. Furthermore, non-simple pervoskite compounds can be identified that display similarly desirable vacancy properties and stability (Figure 6).
2.8 High Temperature Defect Thermodynamics

Thus far we have discussed materials selection purely in terms of zero kelvin predictions of the defect formation enthalpies. However, our approach additionally permits the rapid assessment of defect densities at high temperatures, which is critical in predicting their behavior under realistic process conditions. The defect formation energies $\Delta H_i$ were defined above for the reference energy $\mu^\text{ref}_i$, i.e., the zero temperature limit. At finite temperatures $T$ and partial pressures $p_O$, a corresponding chemical potential $\Delta \mu_i(T, p_O)$ is added to obtain the formation energy under the respective thermodynamic condition (cf. eq. 11). In a given oxide with multiple O sites $i$ and respective reference formation energies $\Delta H_{0,i}$, minimization of the free energy of defect formation yields the dimensionless fractional concentration of O vacancies relative to the nominal O stoichiometry of the respective oxide,

$$[V_o] = \sum \frac{g_i \exp[-(\Delta \mu_{0,i} + \Delta \mu_i)/k_B T]}{1 + \exp[-(\Delta \mu_{0,i} + \Delta \mu_i)/k_B T]},$$

(11)

Here, $k_B$ is the Boltzmann constant, and $g_i = \frac{m_i}{\sum m_i}$ are the normalized degeneracies of the different oxygen sites with their respective multiplicities $m_i$.

To identify oxides which develop a desired degree of O-deficient off-stoichiometry under high-temperature thermodynamic conditions $(T, p_O)$, suitable for different application areas, we numerically invert Equation (11) to solve for the chemical potential $\mu$ for a given target concentration $[V_o]$. At the same time, the respective oxide must be stable under this condition and not decompose into other phases. Here, we include the consideration of a stability threshold $X$ as defined above. For any given temperature, the chemical potential $\Delta \mu_i$ can be translated into the corresponding partial pressure $p_O$ (or vice versa) using the ideal gas law. Note that many "stoichiometric oxides" do not accommodate high levels of defect concentrations, but instead prefer to form a more reduced, ordered phase with lower O content. In this case, there may not be any $(p_O, T)$ conditions for the target $[V_o]$.

On the other hand, oxides that are able to develop a desired level of off-stoichiometry under suitable conditions are considered as potential candidates for functional O-deficient materials in the different application areas.

For a target vacancy concentration of $[V_o] = 1\%$, Figure 7a shows the $p_O$ vs temperature diagram for oxides within $X < 0.05 \text{ eV/atom}$ from the convex hull, using the ML screening of the MP data (cf. Section 2.6). This concentration is generally considered as a demarcation between the dilute, defect-like, and concentrated, alloy like, limits of non-stoichiometric materials, but a similar analysis can be made for any value of $[V_o]$. Desirable process conditions are indicated in Figure 7a for Solid Oxide Fuel Cells (SOFC) and Thermochemical Energy Storage (TCES) using the step reduction for the latter two. A spreadsheet with the oxides falling into the respective regions of interests is included in the SI. All three applications depend crucially on the formation of O vacancies although there are of course other materials considerations that we do not address here. Therefore, the present screening should give valuable insights about potential candidate materials at least for the aspect of O deficient off-stoichiometry.

The list of STCH oxides contains barium, strontium, and lanthanum manganates, which are previously identified classes of oxides for this application but also new suggestions like Ba$_2$Fe$_2$O$_4$. On the other hand, it also contains BaMnO$_3$, which at first sight appears to be a false-positive, because this oxide is known to reduce too easily and therefore be unable to split water. However, it is just one (mp-19267) of 9 different BaMnO$_3$ structures in the MP database, which it is not the ground state. With a ML predicted minimum $V_o$ formation energy of $\Delta H_f = 3.0 \text{ eV}$ it would be a useful water splitter, but the corresponding energy is only $2.2 \text{ eV}$ for the BaMnO$_3$ ground state (mp-1205336) in the MP database, which is too low. Thus, the ML model is consistent with experimental observations, and this example illustrates the tradeoff in choosing the tolerance for the stability criterion. Finally, we note that the list does not contain BCM, an apparent case of a false-negative, resulting from underestimation of the defect energy in the "element-only" encoding used for the MP screening (see Section 2.6). Using the energy from the direct DFT calculation or the "full" encoding ML model, BCM would indeed fall into the STCH process window indicated in Figure 7a.

Our thermodynamic modeling affords direct access to the reduction entropy which is of great benefit to applications that utilize a temperature swing, like STCH and TCES. For example, a large entropy facilitates high $H_2/H_2O$ ratios in the STCH oxidation step. Figure 7b represents the same data as Figure 7a, but showing the O chemical potential $\Delta \mu_O$ as ordinate instead of the $O_2$ partial pressure. The relevant quantity is the differential reduction entropy with change in defect concentration (short-hand $\delta S_{red}$), which equals the slope of the chemical potential,
\[ \Delta S_{\text{red}} = \frac{\partial}{\partial T}(\Delta H_0(T)) \]

We observe in Figure 7 significant variations in \( \Delta S_{\text{red}} \) between the different oxides, which originates from the distribution of defect energies over different O sites. The ideal configurational entropy of 4.6 \( k_B \) for \( V_O \) = 1% is indicated in Figure 7, and the material-specific, numerically determined values are included in the spreadsheet (SI) for the three process windows, giving additional guidance on materials selection over the enthalpy criterion via \( \Delta H_0 \) alone. We further note that additional electronic entropy effects can arise in certain materials, in particular when the O vacancies assume a charged defect state, where the excess electrons either form small polarons or occupy itinerant conduction band states. Such effects are relatively rare in transition metal oxides, where the redox activity is typically dominated by the transition metal ions close to the O defect. They could, however, play a role in the extraordinary behavior of CeO\(_2\) [1–4].

Our high-throughput screening and thermodynamic analysis is a valuable starting point for identifying new potential high-entropy materials.

3 Conclusions

We have developed a powerful, generalized GNN approach for predicting vacancy formation enthalpies of relaxed, defected structures using the relaxed host geometry as input. Therefore, only one DFT relaxation of the host is needed to derive the model inputs, and the model efficiently replaces the computationally intensive supercell calculations with numerous defect relaxations (one per symmetry site) needed to obtain the vacancy formation enthalpies. The model’s applicability is not limited to structures in specific crystal/symmetry classes or elemental compositions, and its accuracy is primarily limited by their representation in the training data. Through careful cross validation, we have thoroughly highlighted the advantages and limitations of the model.

The best model performance was achieved by integrating DFT-computed host compound properties beyond just the relaxed crystal structure into the featurization process (i.e., oxidation states, compound formation enthalpy, band gap, and effective electron mass) to achieve a predicted error below 450 meV for relaxed oxygen vacancy defect formation enthalpy. Nonetheless, models trained only on the crystal structure exhibited just \( \sim 15\% \) higher MAE, since properties like oxidation state and compound formation enthalpy are already indirectly encoded in the crystal structure. \( \Delta H_0 \) of any element/crystal site can be predicted using the same model architecture and learned parameters. Furthermore, so long as two nominally identical materials (relaxed under different DFT settings) have very similar structures, our “element-only” encoding model provides close agreement on the predicted vacancy formation enthalpies because it relies only on the host crystal structure as input. This means that the model is required to screen different databases (e.g., Materials Project) than the models were trained on (e.g., NRELMatDb) as predictions on nominally identical materials provides the same quantitative and qualitative outlook for vacancy formation enthalpies.

While our training database consists of 15 cation elements, we tested an element-wise CV strategy to gauge model performance when predicting vacancies in compounds whose elements were missing from the training set. Finally, since the complexity of the DFT defect relaxations limits the size of the training data that can be collected, we have shown that the model error is still expected to decrease significantly as more data is collected in the future.

We demonstrated the model’s significant utility for novel materials discovery in an exercise of identifying promising candidate oxides in the context of various clean energy applications: solar thermochemical water splitting and energy storage, CO\(_2\) conversion, and SOFC cathodes and electrolytes. We screened structures drawn from a different database (Materials Project) than the source of the training structures (NRELMatDB) using the simplest, “Element-only” graph encoding strategy (i.e., requiring only the host crystal structure as input). Narrowing down the \( \sim 35,000 \) oxides initially queried to as few as \( \sim 10 \) depending on the stringency of down-selection criteria, we identify candidates exhibiting the greatest potential based on predicted vacancy defect enthalpy and host oxide stability criteria, which also “rediscover” known materials from previous experimental literature. This ML strategy therefore efficiently reveals a handful of top candidates from an intractably large space for brute-force DFT or experiments, and will help facilitate the discovery of optimal materials in the future (along with significant potential for more chemistry and structural diversity in the training data). Even stricter down-selection can now additionally be explored based on more detailed properties from first principles calculations that are only tractable across a small number of materials. The final critical contribution of this study is the prediction of defect formation enthalpies across all sites to rapidly estimate defect densities at finite temperatures. By accounting for configurational entropy in high-throughput, we can assess material performance at finite temperatures, rather than relying purely on a zero kelvin picture from individual defect predictions across just a handful of materials.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This material is based upon work supported by the U.S. Department of Energy (DOE), Office of Energy Efficiency and Renewable Energy (EERE), specifically the Hydrogen and Fuel Cell Technologies Office. Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-NA0003525. Part of the work was performed under the auspices of the US Department of Energy by Lawrence Livermore National Laboratory under contract No. DE-AC52-07NA27344. The National Renewable Energy Laboratory (NREL) is operated by the Alliance for Sustainable Energy, LLC, for the DOE under Contract No. DE-AC36-08GO28308. This work used High-Performance Computing resources at NREL, sponsored by DOE-EERE. The views expressed in this article do not necessarily represent the views of the U.S. Department of Energy or the United States Government.
The open-source CGCNN code for training models that directly predict vacancy formation enthalpies has been distributed on the “Paper1” branch at the following GitHub repository (https://github.com/mwitman1/cgcnndefect/tree/Paper1), which is a modified fork of the original open-source CGCNN code (https://github.com/txie-93/cgcnn).

Notes and references

100 D. Xiang, C. Gu, H. Xu and G. Xiao, Small, 2021, 17, 2101524.