



OPEN

Putting error bars on density functional theory

Simuck F. Yuk^{1,6}, Irmak Sargin^{2,6}, Noah Meyer³, Jaron T. Krogel⁴, Scott P. Beckman⁵ & Valentino R. Cooper⁴✉

Predicting the error in density functional theory (DFT) calculations due to the choice of exchange–correlation (XC) functional is crucial to the success of DFT, but currently, there are limited options to estimate this a priori. This is particularly important for high-throughput screening of new materials. In this work, the structure and elastic properties of binary and ternary oxides are computed using four XC functionals: LDA, PBE-GGA, PBEsol, and vdW-DF with C09 exchange. To analyze the systemic errors inherent to each XC functional, we employed materials informatics methods to predict the expected errors. The predicted errors were also used to better the DFT-predicted lattice parameters. Our results emphasize the link between the computed errors and the electron density and hybridization errors of a functional. In essence, these results provide “error bars” for choosing a functional for the creation of high-accuracy, high-throughput datasets as well as avenues for the development of XC functionals with enhanced performance, thereby enabling the accelerated discovery and design of new materials.

Advances in computing power and electronic structure methods, in particular density functional theory (DFT), have propelled computation and theory to the forefront of materials design and discovery. Countless efforts now employ data-driven and/or high-throughput computational approaches^{1–8} to isolate individual or groups of materials with desired properties^{3–6,9–11}. For instance, the Materials Genome Initiative has fostered multiple research projects to map out materials’ properties, based on DFT predictions, to explore the physical and chemical properties of known and hypothetical materials^{3,4,12–18}. A striking example is work by Greeley and co-workers which screened over 700 binary alloys and identified BiPt as the best electrocatalyst for the hydrogen evolution reaction, comparable to pure Pt in terms of surface activity¹². Similarly, Hautier et al. explored phosphate chemistry with thousands of compounds for the design of lithium-ion battery cathodes¹³. Emery and co-workers searched through 5,329 ABO_3 perovskite compounds to find thermodynamically-favorable compounds most suitable for a two-step thermochemical water splitting process¹⁷. The combination of high-throughput DFT with materials informatics and machine learning techniques promises even greater search efficiency^{3,19–22}.

Central to the success of DFT are the approximations for many body exchange and correlations (XC) within the framework of the Hohenberg–Sham–Kohn theorems^{23,24}. These approximations attempt to balance accuracy and speed; thereby making DFT an attractive tool for computational materials design. Initial efforts to approximate XC relied on the homogenous electron gas assumption²⁵, in which the energy at a point only depended on the charge density (n) at that point. This local density approximation (LDA)^{23,24,26} has a tendency to over bind leading to errors in calculated physical properties²⁷, including underestimations of lattice parameters^{28–30} and vibrational frequencies^{31,32}. This overbinding is attributed to the LDA exchange which scales as $n^{4/3}$ as opposed to n^2 as expected from Hartree–Fock exact exchange³³.

To improve this approximation, functionals were developed to account for density variations in the electron gas, i.e. including local gradients. These so called generalized gradient approximations (GGA) have a much wider range of forms and thus ushered in an explosion of functionals³⁴. Early derivations include the XC functional of Perdew, Burke and Ernzerhof (PBE), which while improving quantities such as binding energies, tended to overestimate lattice constants³⁵. A more recent incarnation was designed to return to the gradient expansion approximation to give better agreement for solids (PBEsol)³⁶. Unfortunately, GGAs still underestimate band gaps; sometimes predicting semiconductors to be metals³⁷. Hybrid GGAs, such as B3LYP^{38,39}, PBE0⁴⁰ and HSE⁴¹, mix the GGA with non-local exact exchange, resulting in improved descriptions of electronic band gaps as well as covalent, ionic, and hydrogen bonding, but at increased computational cost^{42,43}. Similarly, the non-empirical

¹Department of Chemistry and Life Science, United States Military Academy, West Point, NY 10996, USA. ²Department of Metallurgical and Materials Engineering, Middle East Technical University, 06800 Ankara, Turkey. ³Physics Department, Stanford University, Stanford, CA 94305, USA. ⁴Materials Science and Technology Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA. ⁵School of Mechanical and Materials Engineering, Washington State University, Pullman, WA 99164, USA. ⁶These authors contributed equally: Simuck F. Yuk and Irmak Sargin. ✉email: coopervr@ornl.gov

strongly constrained and appropriately normed (SCAN) meta-GGA provides a significant improvement in accuracy over standard LDA and PBE^{44,45}, but is still quite computationally expensive. A class of van der Waals density functionals (vdW-DF) incorporate long-range vdW interactions allowing for the computation of material functionality in both sparsely and densely packed structures^{45–50}.

It is widely accepted that the main component of the errors in DFT calculations is the exchange–correlation functional approximation and the goal is to choose the most reliable functional^{51,52}. The broad diversity of XC functionals begs the question, for a given composition and structure, what is the best choice of functional? However, one of the major challenges is to determine the reliability and the source of the error for a given functional^{53,54}. The question of DFT calculations' reliability is mainly handled by two different approaches: Bayesian error estimation and statistical analysis. The Bayesian approach is a semi-empirical approach where an ensemble of XC-functionals is used, and the desired quantities are represented in form of a distribution. The spread of the calculations provides an error estimation^{52,53,55–57}. This method has been successfully applied for uncertainty quantification in DFT calculations in a large variety of applications ranging from calculation of physical and surface properties to the refinement of phase diagrams^{53,55,56,58–62}.

In the case of statistical analysis by regression, experimental measurements are expressed as a function of DFT calculation errors. This is an a posteriori analysis to provide material-specific predictions of calculation errors. An example of this method, developed by Lejaeghere et al., employed a linear function⁵¹. Another later effort, used polynomial functions to represent the relationship between calculations and measurements for cubic crystals⁶³. This approach of mapping measurements to calculations and statistically analyzing the error has since been successful for predicting the error in calculations of surface energies and work functions, energetic and elastic quantities, thermal expansion coefficients, and melting temperatures employing DFT-based, semi-empirical methods^{51,52,54,64,65}.

The question of best exchange–correlation functional is particularly relevant for high throughput calculations where one functional may be used to screen a large group of materials. To examine the accuracy of such approaches, here, we employ high-throughput DFT calculations to quantify the accuracy of four XC functionals: LDA, PBE, PBEsol, and vdW-DF with C09 exchange⁴⁶. We explore the properties of 141 binary and ternary oxides encompassing 44 different space groups across all 7 crystal structure types. We examined structures with 4–264 atoms containing 31 different cations. These structures were chosen based on previous work by Hautier and co-workers¹⁴. The lattice parameters, bulk moduli and reaction enthalpies for forming the ternary oxides from the binary oxides were computed and compared against available experimental data, using the Nexus workflow management system⁶⁶. Further details of the methods are given in the Supplementary Information (SI) and the distribution of atomic structures studied is given in Fig. S1. In general, we find that the vdW-DF-C09 and PBEsol functionals have the lowest errors.

In this study, we approached the problem of finding the best functional from the aspect of material-specific calculation errors. We note here that we choose the term error instead of uncertainty as uncertainty is defined as a “non-negative parameter characterizing the dispersion of the quantity values being attributed a measurand”⁶⁷. Based on this, the quantity we are interested in is not a dispersion, but instead, it is a material-specific deviation from the experimental measurements. In addition, we do not consider ensembles of the same functional as in the Bayesian Error Estimation Functionals^{53,55,56}. Instead, we calculate errors, and evaluate them separately for the four different functionals employed in this study. Another difference is that we estimate the error in terms of material-specific parameters using machine learning. Employing materials informatics methods, we quantify the main contributions to the errors and explore their physical origins. Our results emphasize the importance of both the electron density and metal–oxygen bonding/orbital hybridization. We demonstrate the strength of this approach by predicting the lattice constant errors of a distinct dataset of potential perovskites. Overall, our approach provides a means of understanding the expected errors associated with a particular XC, which can lend to greater accuracy for materials screening, along with giving insights that may enable the development of improved XC functionals. As such, the notion of incorporating “error bars” within DFT is to estimate the errors a particular functional would have on a given property with the application of advanced machine learning models and ultimately recommend the functional suitable for a particular class of material.

Results and discussion

DFT-based properties with different XC functionals

Figure 1 shows DFT-optimized lattice constant percent errors relative to experimental values. These errors are good indicators of the over or under-binding caused by the XC. As expected, PBE, on average, overestimates lattice constants while LDA typically underestimates them. Interestingly, vdW-DF-C09 yields similar trends in the lattice constant percent errors as PBEsol; with errors centered around 0%.

The distribution of the lattice constants absolute % errors is depicted in Fig. 2a. The errors of each lattice constant are examined independently. For each XC functional, the *y*-axis indicates the number of lattice constant data points with an absolute percent error less than or equal to the *x*-axis value; the *y*-values asymptote to 100%. The mean absolute relative error (MARE) and standard deviation (SD) are 2.21% and 1.69% for LDA and 1.61% and 1.70% for PBE, respectively. PBEsol and vdW-DF-C09 have significantly lower errors: PBEsol has MARE: 0.79% and SD: 1.35% and vdW-DF-C09 has MARE: 0.97% and SD: 1.57%. Previously, we observed similar accuracy for PBEsol and vdW-DF-C09 for a small dataset of ferroelectric perovskites³⁰. If we were to consider the number of data points with MAREs less than 1%, we see that PBEsol and vdW-DF-C09 would have nearly 80% accuracy, whereas PBE and LDA are dramatically reduced to < 30% and < 20%, respectively.

Intuitively, one would expect XC errors to couple to the chemistry of the compounds¹⁴. The percent error as a function of chemical element is shown in Fig. 3 for the four XC potentials. While no distinctive pattern emerges, for PBEsol and vdW-DF-C09 compounds with lighter chemical elements ($Z < 23$) have lower MAREs < ~ 1%

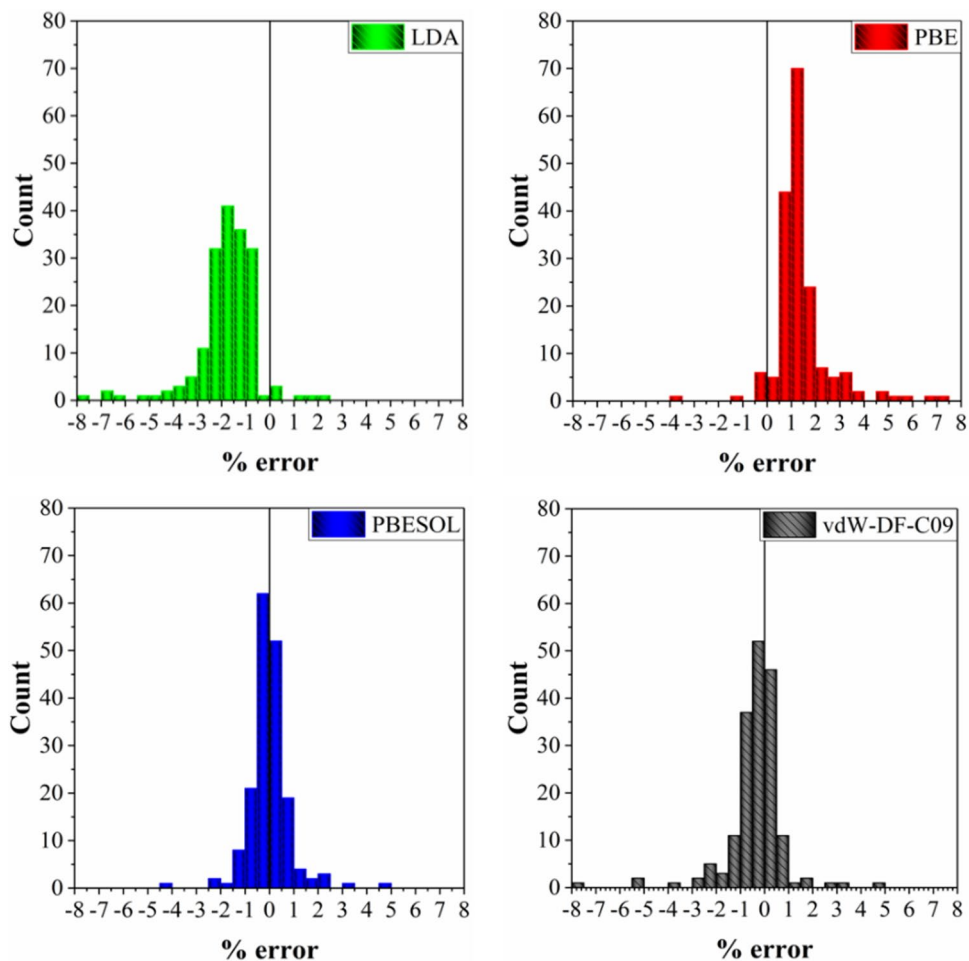


Figure 1. Histogram of DFT-predicted percent errors relative to the experimental lattice parameters using different XC functionals. *a*, *b* and *c* lattice parameters are considered separately for each structure.

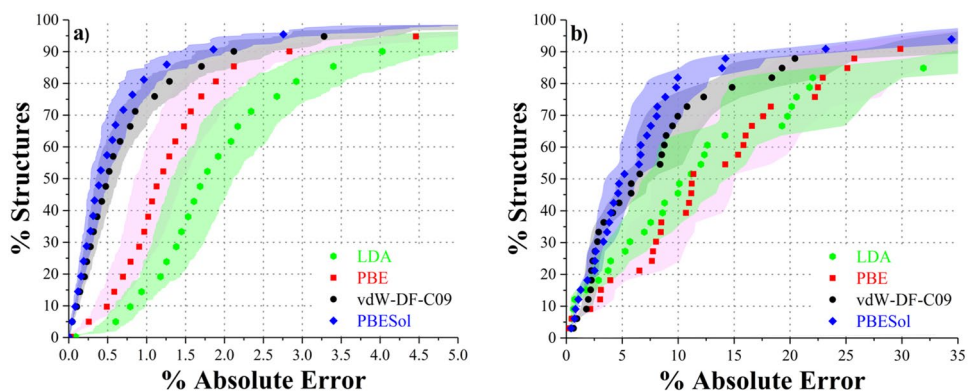


Figure 2. Distribution of DFT-computed percent errors of (a) lattice parameters and (b) bulk moduli for different XCs. *a*, *b* and *c* lattice parameters are considered separately for each structure.

with the greatest errors associated with magnetic elements, specifically Cr, Fe, Ni, and Mo. (All calculations were exclusively spin unpolarized to isolate common XC errors from spin-dependent errors. The results suggest that magnetoelastic couplings may be significant in several compounds.) In fact, a significant reduction in MARE can be achieved for PBESol (MARE = 0.64) and vdW-DF-C09 (MARE = 0.78), if the magnetic ion-containing compounds are excluded (see Table S4).

In addition to lattice parameters, the bulk moduli were computed using the Birch-Murnaghan equation of state as given in Eq. S1. As indicated by Fig. 2b, the predicted bulk moduli MAREs are larger than the lattice

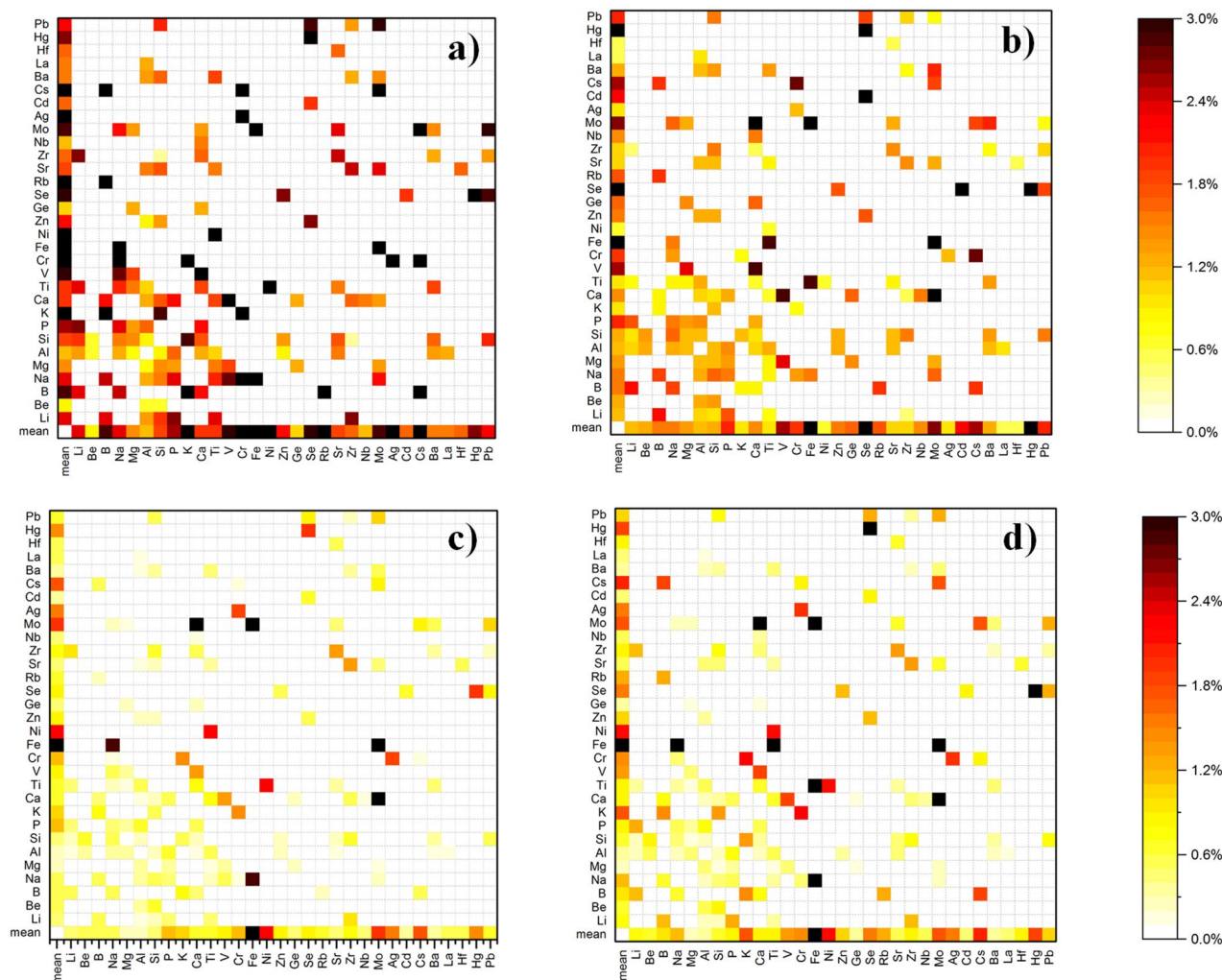


Figure 3. MAREs of DFT-predicted lattice parameters relative to experimental lattice parameters as a function of chemical elements using (a) LDA, (b) PBE, (c) PBEsol, and (d) vdW-DF-C09. For ternary compounds, with a general formula of $A_mB_nO_x$, the species indicated by the y - and x -axes can be on either the A - or the B -site. As a result, the figures are symmetric relative to the diagonal of the plots. Binary compounds are considered in the mean.

parameters for all XC functionals, which is expected. LDA and PBE over and underestimate the bulk moduli respectively, due to the inverted dependence of total energy on volume. LDA yields the largest MARE, 17.93%, and has a SD of 19.82%, followed by PBE with MARE of 15.75% and SD 13.10%. PBEsol and vdW-DF-C09 have lower errors with MARE: 9.49% and SD: 12.79% for PBEsol and MARE: 11.45% and SD: 13.40% for vdW-DF-C09.

Next, we examine computed reaction enthalpies errors for forming ternary oxides from binary oxides, $A_mO_n + B_xO_y \rightarrow A_mB_xO_{(n+y)}$, as shown in Fig. S2. A full list of reaction energies for each XC along with corresponding experimental values are given in Table S3. There is excellent agreement between DFT and experiment for all XCs. This suggests that the cancellation of systematic errors, a hallmark of the success of DFT, is not significantly dependent on the XC. Similar observations were reported for a DFT ($GGA + U$)-high-throughput study of these metal oxides¹⁴.

Together our observations indicate that PBEsol and vdW-DF-C09 are excellent general purpose XCs for examining structural properties, while LDA and PBE fail to capture the structural/mechanical properties of bulk oxides. These present us with a measure of accuracy that may be applied when performing high-throughput calculations.

Materials informatics approach

Exploiting this high-throughput dataset, we employed a materials informatics approach to determine *a priori* the “best” XC for a given metal oxide. (Note: We exclude oxides that have greater than a 5% lattice parameter error for a single lattice parameter for more than two functionals and oxides with greater than a 5% lattice parameter error for multiple lattice parameters. Non-groundstate structures are excluded. This results in 123 data points. We also exclude oxides for which ionic radii value with coordination number of the element do not exist. This left

us with 110 samples to be used for model selection and development. The details are explained in the SI.) Unlike, some other efforts^{68,69} bringing DFT and data-based methods together that seek to predict the magnitude of a physical property, e.g., the lattice parameter, our effort aims to predict the percent error attributed to an XC for a specific property. The philosophy is to provide material-specific “error bars”, i.e., deviations from experimental measurements, for each XC—in this case (as illustrated in Fig. 4) we consider the mean average errors (MAE) and SDs as measures for the range of errors one would expect when studying lattice parameters with each of the functionals. Choosing, beforehand, a “best-fit” XC should help to improve high-throughput dataset accuracy. Here, we focus on the average lattice parameter, $\delta\bar{a}$, as an archetypical physical quantity.

Based on available standard elemental and structure-specific properties: (i) electronegativity of A- and B-site elements (X_A and X_B); (ii) number of valence electrons (n_A and n_B); (iii) atomic numbers (Z_A , Z_B , Z_O), (iv) ionic volumes (V_A , V_B , V_O); (v) DFT ionization percent errors (δIE_A , δIE_B); (vi) number of A-site, B-site, and oxygen atoms (N_A , N_B , N_O); (vii) nominal charge of A- and B-site elements (σ_A and σ_B); and (viii) coordination numbers (CN_A and CN_B), we created the following compound-specific feature set (summarized in Tables S5 and S6) to eliminate inter-feature correlations (see Fig. S4):

- (i) Fractional ionicity, f : This expresses the relative ionicity/covalency of the bonds. For ternary compounds, it is computed using the fractional ionicities of the end member binary compounds. Individual f values are computed from the experimental electronegativities of each binary compound.
- (ii) Charge per valence electron, $\bar{\sigma}^n$: This represents the fraction of valence electrons contributing to the ionic bonding and consequently implies the fraction of electrons contributing to metal–oxygen hybridization. The nominal charge of the A- and B-site ions (σ_A and σ_B) is multiplied by the fractional ionicity of the AO and BO binary oxides f_{AO} and f_{BO} divided by the number of A- and B-valence electrons (n_A and n_B). For ternary oxides, we employ a weighted average value of $\bar{\sigma}^n$ based on the number of A and B-site ions.
- (iii) Valence electrons per atomic number, \bar{n}^Z : This captures electron screening effects and the impact of core or semi-core electrons on the valence electrons. \bar{n}^Z is essentially a weighted average of the number of valence electrons (n_A and n_B). This quantity is mathematically transformed to create Gaussian-like distributions that are compatible with the data-analytics techniques.
- (iv) Pauling electrostatic strength of the metal–oxygen bond, \bar{S} : This is the ionic bond strength. To determine the valence electrons contributing to the ionic bonds, the number of A- and B-valence electrons (n_A and n_B) are multiplied by the fractional ionicity divided by the coordination numbers of the A- and B-site cations (CN_A and CN_B). A weighted average value is used for ternary oxides.
- (v) Mass density, ρ^M : This is the mass to volume ratio and roughly proportional to the average electron density. It is important for highly localized valence states, such as d and f electrons, as well as semi-core states. The mass is calculated by multiplying the atomic weight with the number of A-site, B-site, and oxygen atoms in the stoichiometric compound (N_A , N_B , N_O). The volume is calculated similarly by multiplying the atomic volumes (V_A , V_B , V_O) with the number of A-site, B-site, and oxygen atoms.
- (vi) Oxygen fractional occupation, \bar{N}_O : This is the anion to total ions ratio. It is calculated by dividing the number of oxygen atoms (N_O) by the total number of ions in the stoichiometric compound.

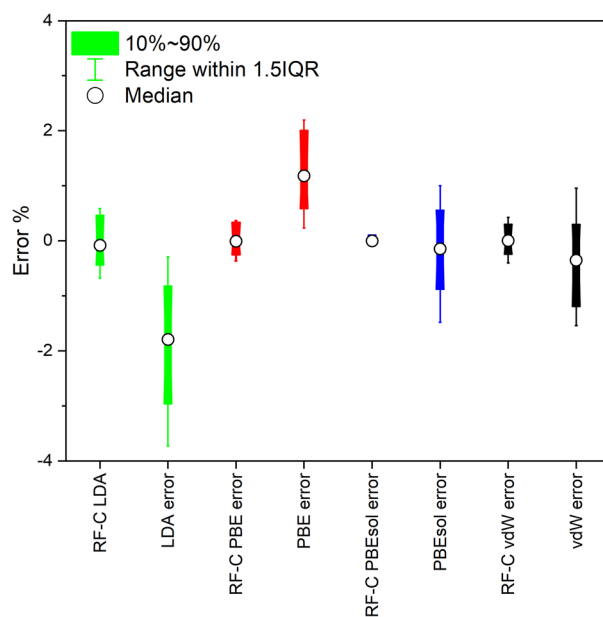


Figure 4. Comparison of the percentage error in average lattice parameter before (XC error) and after RF correction (RF-C XC error).

- (vii) Element specific DFT ionization errors, $\overline{\delta IE}$: These express the difference between DFT computed and experimental first ionization energies for individual elements. This is the only feature that depends on the XC and pseudopotential choice. A weighted average is used for ternary oxides.

To predict $\delta\bar{a}$ for each XC, we trained random forest regression models using the above feature set; this was done with the NumPy, SciPy, and Scikit-learn libraries^{70,71}. Both data before its preparation for machine learning models and feature sets for prediction of errors for all functionals are provided in the Supporting Information. The machine learning codes are also provided in the Supporting Information. The leave-one-out (LOO) cross-validation method was employed with a linear bias correction. The error and the standard deviation for the LOO cross validation is provided in Table S7. MAEs were 0.191, 0.109, 0.115, and 0.139 for LDA, PBE, PBEsol, and vdW-DF-C09, respectively. The predicted versus true error is plotted for each XC in Fig. S3. These results demonstrate the good accuracy of our approach for all XCs. In addition, with an approach similar to the Δ -machine learning approach⁷², we calculated the random forest corrected (RF-C) average lattice parameters and compared them to the experimental average lattice parameters, as shown in Fig. 4. Figure S4 further summarizes the comparison of the RF-CDFT errors and the DFT errors without any correction in greater details. Even though the median of all the RF-C errors is similar, PBEsol shows the smallest variations compared to the experimental values. A similar result was found by Pernot et al. after application of a linear calibration. LDA, PBE, and PBEsol gave similar mean errors, but they differed in terms of width of distribution. LDA has a wider distribution than PBE and PBEsol has the smallest⁶³. Considering linear calibration captures the systematic error, the similarity between their and our results suggest we were able to predict the systematic error by using material-specific features in our ML model.

The key features, i.e., input parameters used in the RF predictions, were further used to identify the most important variables for predicting $\delta\bar{a}$ using the tree interpreter algorithm. This algorithm calculated the contributions of each feature by fitting a linear equation to each sample. The mean and standard deviations of the absolute linear coefficients are given in Table S8. Fitting a different equation to each sample creates a large standard deviation for the linear coefficients and hence the contributions. The contributions summarized below are based on the average values, however, we are confident these average contributions will not change based on different datasets as they are based on the average of linear coefficients fitted to each example separately.

The absolute values of the linearity coefficients were normalized to one. The normalized linearity coefficients are given as a radar plot (see Fig. 5). Equal importance of all features would give vertex values of $1/7$. Deviations from this ideal value allow us to assess specific feature contributions to XC errors; thus, providing insight into the determining factors in the success or failure of each XC.

The contribution of each feature can be summarized as follows:

f : examines the importance of the relative bond ionicity/covalency. In Fig. 5, we see that for both LDA and vdW-DF-C09, f is greater than the 0.14 average value. PBE, on the other hand, is roughly 0.14 while PBEsol exhibits a reduction to ~ 0.10 .

$\bar{\sigma}^n$: Similar to f , the role that ionic bonding and metal–oxygen hybridization plays in determining $\delta\bar{a}$ are explored. In Fig. 5, we observe that it contributes similarly to $\delta\bar{a}$ for all XCs with coefficients that are higher than the average value of $1/7$.

\bar{n}^z : relates to electron screening effects. For LDA and PBE, we observe an average contribution ~ 0.14 (Fig. 5), while vdW-DF-C09 and PBEsol show stronger than average dependence on this feature.

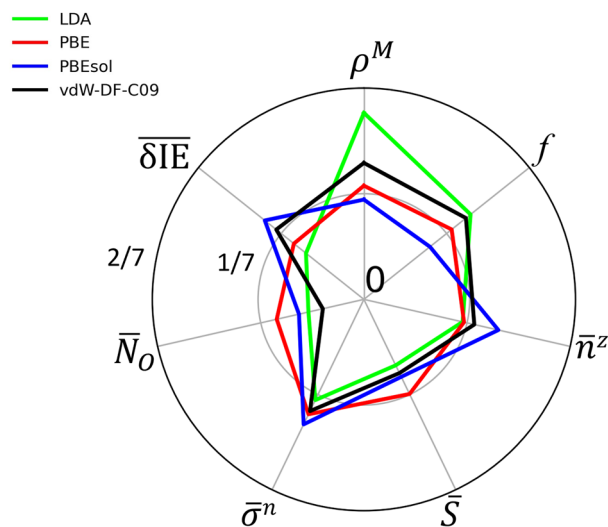


Figure 5. Normalized mean absolute linear dependency coefficients for different XCs obtained from regression analyses. A vertex value of $1/7$ is indicative of equal contributions.

\bar{S} : corresponds to the ionic strength per bond. Here, the predicted errors in PBEsol and vdW-DF-C09 seem to have a roughly average dependence on this feature while, the PBE predicted errors show enhanced correlations and LDA exhibits reduced correlations.

ρ^M : is roughly proportional to the electron density and is viewed as a measure of valence state, i.e. d and f electrons, and semi-core state localization. The relative contribution of this feature to $\delta\bar{a}$ is highest for LDA and vdW-DF-C09; but of limited significance for PBE and PBEsol.

\bar{N}_O : expresses the relative anion:cation ratio. While in general, this feature has the least significance when predicting $\delta\bar{a}$ for all XCs, we observe that removing it from the regression significantly degrades the quality of predictions.

$\delta\bar{I}E$: is the difference between the DFT computed and experimentally available first ionization energies. Unlike other variables, it only depends on the chemical identity of the elements involved in a structure and thus is an indicator of the quality of the pseudopotential as well as the XC. For LDA it is least significant for $\delta\bar{a}$ predictions. Interestingly, even though the same pseudopotential was employed for PBE, PBEsol and vdW-DF-C09, we observe a spread in its relative significance; being least important for PBE and most significant for PBEsol. This perhaps reflects the different philosophies behind XC development.

An analysis of the above data suggests that the largest source of errors is the inability of the functionals to capture and represent non-homogeneities due to limitations in the exchange hole and the short-range expansion of the reduced gradient. It is widely known that LDA exchange holes cannot capture inhomogeneities, therefore resulting in systematic overbinding⁷³. From our analysis, LDA's tendency to overbind is, perhaps, most reflected in the significance of f , $\bar{\sigma}$, \bar{n}^Z and ρ^M in predicting $\delta\bar{a}$. These features are all related to bonding and hybridization; possibly emphasizing the direct relationship between improvements in these quantities and the reduction of XC errors.

Interestingly, the gradient corrected functionals all show significant reductions in ρ^M dependence, largely an indication of the localization or homogeneity of the electrons. For PBE the predicted $\delta\bar{a}$ is roughly equally dependent on all features—perhaps indicative of PBE's poor performance thus requiring a broader set of features to predict the large distribution of errors. On the other hand, vdW-DF-C09 and PBEsol have fewer significant features. PBEsol, for instance, has above average dependencies on $\delta\bar{I}E$, $\bar{\sigma}$, and \bar{n}^Z . While vdW-DF-C09 exhibits significant dependence on $\delta\bar{I}E$ and f , $\bar{\sigma}$, \bar{n}^Z and ρ^M like LDA.

Surprisingly, all XCs have similarly strong dependence on $\bar{\sigma}^n$, which is an indicator of hybridization. To gain further insights into the relationship between $\bar{\sigma}^n$ and hybridization, the hybridization energies for sp , sp^2 , sp^3 , sp^3d , and sp^3d^2 were calculated using a linear sum of the orbital energies following Harrison's prescription⁴² and orbital energies from the literature^{43,44}. In all cases, our calculated energies are linearly dependent on $\bar{\sigma}^n$ as plotted in Fig. S5, typical of hybridization. The orbital and hybridization energies are correlated with the self-interaction energy contributions which rely on the cancellation of errors. The XC functionals compared here are continuous, making it impossible to cancel this energy entirely. A recent study suggested that the error related to the orbital energies may not be directly related to the self-interaction errors but are instead due to the too repulsive exchange response potentials of the LDA which are not fully corrected in the GGA formulation⁷⁴. In either case, the fact that the $\bar{\sigma}^n$ is similar for all XCs is consistent with the fact that typical GGAs and LDAs are incapable of eliminating effects due to poor descriptions of orbital hybridization⁷⁵.

Finally, to demonstrate the power of this approach, the trained regression model was used to predict $\delta\bar{a}$ for a large set of possible ABO_3 perovskites for the four XCs (Fig. 6). As expected, the predicted $\delta\bar{a}$ indicates a systematic under- and over-estimation for LDA and PBE, respectively. Again, PBEsol and vdW-DF-C09 have errors centered around 0, with the vdW-DF-C09 functional having a slightly larger distribution of errors. The distribution of the calculated errors are provided in Supporting Information Fig. S6.

To confirm these predictions, we extracted available data for lattice constant errors for PBE calculations of these oxides from the Materials Project high throughput databases and compared this to experimental data in the Inorganic Crystal Structure Database (ICSD) (see Fig. 7). (N.B. this databases employs VASP with projector augmented wave potentials as opposed to Quantum Espresso (QE) with GBRV pseudopotentials in the training set). To gauge the difference between VASP and QE we include data points from this study and our recent comparative paper³⁰. In general, we find excellent agreement between our machine learning predictions and the VASP and QE calculations; particularly as it relates to non-magnetic materials. The largest deviations occur in three categories:

- 1) Polar structures, e.g. ferroelectric $PbTiO_3$, $AgNbO_3$, and antiferroelectric $PbZrO_3$. We note that structures with large off-center displacements are not well represented in this database. N.B., with the exception of $PbTiO_3$, we typically find better agreement between our QE calculations and the predictions—likely indicative of the differences between QE and VASP.
- 2) Magnetic cations particularly Mo and Mn. Our calculations were explicitly non-magnetic and compounds with these cations were excluded from our training set. Nevertheless, magnetic ions such as Fe, Cr and Ni exhibit remarkably good agreement; possibly an indication of weak spin–lattice couplings within these materials (see Fig. S7).
- 3) Cd-based compounds. We had only two Cd-containing compounds in our training set, neither of which was a perovskite. This suggests the need for additional training data points (see Fig. S7).

Ultimately, these results present strong indications that our approach can reliably predict the errors of a class of materials which may not necessarily be included in our training set. In other words, this allows us to understand the boundaries of predictions for a large dataset using a particular functional. This should allow us to construct much higher accuracy high throughput databases. For example, for the exploration of ABO_3

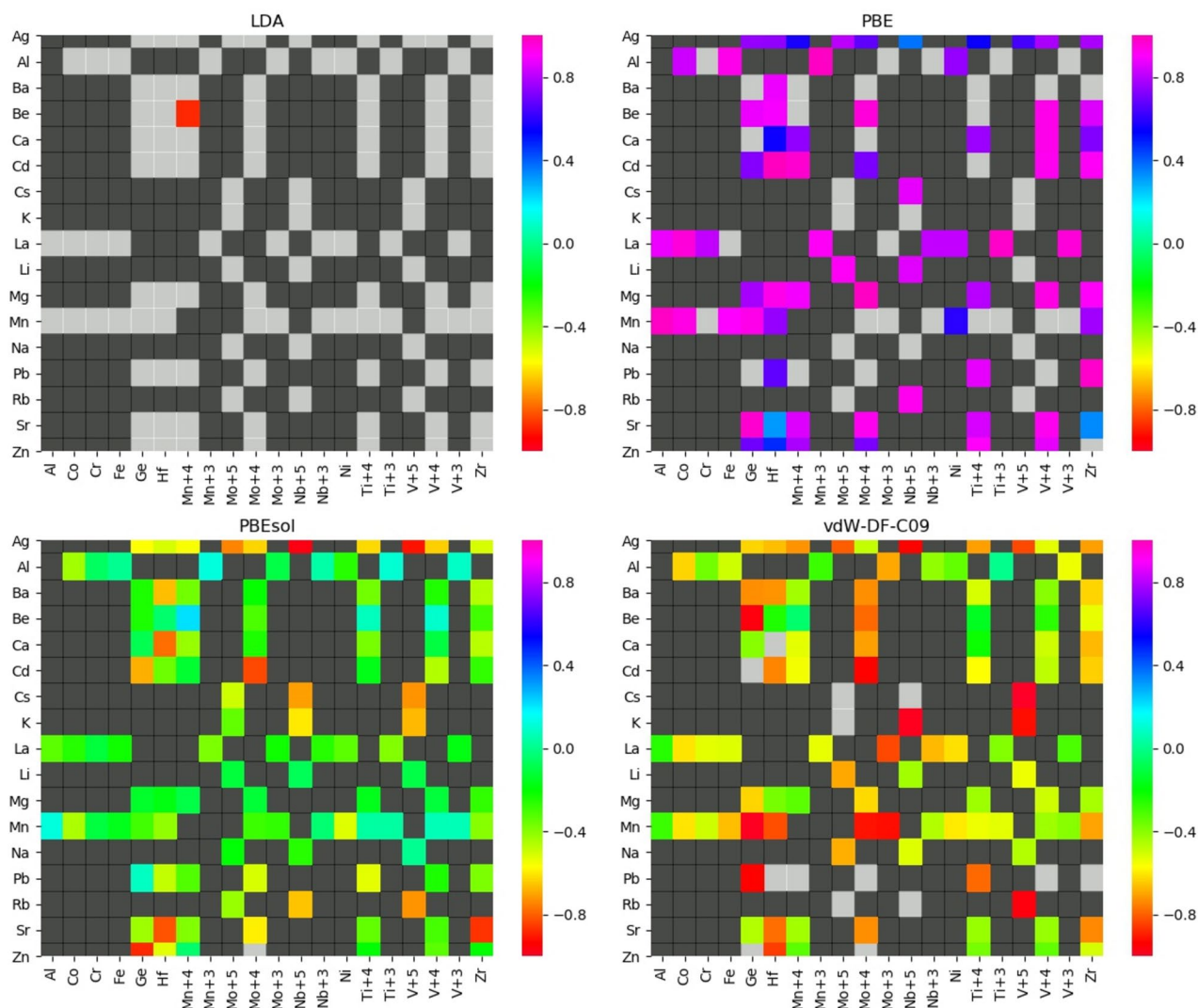


Figure 6. Heat maps showing predicted $\delta\bar{a}$ for selected perovskites. Black masked regions correspond to non-existent A- and B-site combination, and light gray regions show combinations where the absolute error is larger than 1%.

oxides our results indicate that PBEsol and vdW-DF-C09 would be the best choice; at least when predicting lattice constants. Here, we stress that these may not be the optimal XC functionals when exploring other materials properties. For example, we have previously shown that for ABO_3 ferroelectric oxides PBEsol has a tendency to severely overestimate the soft-mode displacements that give rise to polarization and thus, although it gives excellent predictions for lattice constants, would be a poor choice for polar materials.

Conclusion

In conclusion, the impact of XC choice on computed macroscopic properties was investigated for binary and ternary metal oxides. As expected, LDA and PBE lead to under- and overestimations in predicted lattice parameters, respectively. The opposite trend is observed for bulk moduli due to the inverted dependence of total energy to volume. Surprisingly, vdW-DF-C09 exhibits performance comparable to PBEsol. One explanation for this similar behavior, could go back to the philosophy behind the development of the C09 exchange which was to reduce short-range exchange repulsion. To do this, the functional uses the gradient expansion approximation for the enhancement factor $F_x(s) = 1 + \mu s^2$ where $s = |\nabla n|/(2k_F n)$ is the reduced gradient of the density n and $\mu = 0.0864$, which is the value used in PBEsol⁴⁶. Good agreement, within experimental uncertainties (see Table S3), can be seen between DFT-computed and experimental reaction energies regardless of the XC; most likely brought about due to systematic error cancellation. Employing machine learning focused on the prediction of errors, we explore the defining differences in accuracy between different XCs. Most significantly, we find that the representation of electron density and hybridization may be the determining factor behind the accuracy of a functional. This suggests the need to apply more stringent criteria when designing XC functionals; thereby emphasizing the return to the exact density as suggested by Medvedev et al.⁷⁶. Ultimately, vdW-DF-C09 and PBEsol are indicated to be good general purpose XCs for oxides, producing the highest accuracy for high-throughput screening. This

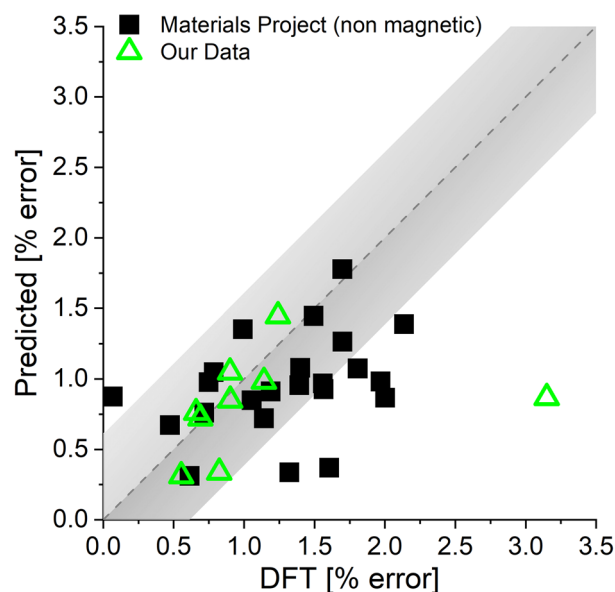


Figure 7. Comparison of $\delta\bar{a}$ computed from DFT vs. experimental values from ICSD. Solid black squares and empty green triangles are Materials project data and our QE data for non-magnetic materials (excluding Cd-based materials), respectively. Our data includes those in this study and our previous work³⁰. Gray regions indicate 1 standard deviation of the computed $\delta\bar{a}$. Predictions for magnetic systems and Cd can be found in the SI.

is illustrated in our application to predicting ABO_3 oxides $\delta\bar{a}$. While further work is needed to understand the deviations for other quantities such as band gaps, elastic, polar or magnetic properties, these results present a meaningful approach, centered on error prediction rather than actual quantities, that may assist in the guidance of choice of XC functional to produce high fidelity datasets and to define routes to creating functionals with improved performances.

Data availability

Data is provided within the manuscript or supplementary information files. Additional data can be obtained through the Constellation data sharing system at <https://doi.ccs.ornl.gov/10.131319/OLCF/2404285>, or by contacting the corresponding author: coopervr@ornl.gov.

Code availability

All DFT calculations were performed using Quantum Espresso (version 6.0), which is an open-source code (see www.quantum-espresso.org/). High-throughput DFT calculations were monitored and conducted using the Nexus workflow management system, which is an open-source code (see <https://www.qmcpack.org/nexus>). Data science models were developed using open source Scikit learn (see <https://scikit-learn.org/stable/>), NumPy (see <https://numpy.org/>), and SciPy (see <https://www.scipy.org/>) libraries. The decision tree interpretation was conducted by tree interpreter algorithm, which is an open-source code (see <https://pypi.org/project/treinterpreter/>).

Received: 10 April 2024; Accepted: 1 August 2024

Published online: 30 August 2024

References

- Morgan, D., Van der Ven, A. & Ceder, G. Li conductivity in Li_xMPO_4 ($M = \text{Mn, Fe, Co, Ni}$) olivine materials. *Electrochem. Solid-State Lett.* **7**, A30–A32 (2004).
- Broderick, S. R., Aourag, H. & Rajan, K. Classification of oxide compounds through data-mining density of states spectra. *J. Am. Ceram. Soc.* **94**, 2974–2980 (2011).
- Armiento, R., Kozinsky, B., Fornari, M. & Ceder, G. Screening for high-performance piezoelectrics using high-throughput density functional theory. *Phys. Rev. B* **84**, 014103 (2011).
- Jain, A. *et al.* A high-throughput infrastructure for density functional theory calculations. *Comput. Mater. Sci.* **50**, 2295–2310 (2011).
- Hautier, G., Fischer, C. C., Jain, A., Mueller, T. & Ceder, G. Finding nature's missing ternary oxide compounds using machine learning and density functional theory. *Chem. Mater.* **22**, 3762–3767 (2010).
- Ong, S. P., Wang, L., Kang, B. & Ceder, G. Li–Fe–P–O₂ phase diagram from first principles calculations. *Chem. Mater.* **20**, 1798–1807 (2008).
- Kang, B. & Ceder, G. Battery materials for ultrafast charging and discharging. *Nature* **458**, 190–193 (2009).
- Sun, W. *et al.* A map of the inorganic ternary metal nitrides. *Nat. Mater.* **18**, 732 (2019).
- Wang, J. *et al.* Epitaxial BiFeO_3 multiferroic thin film heterostructures. *Science* **299**, 1719–1722 (2003).
- Kang, K., Meng, Y. S., Bréger, J., Grey, C. P. & Ceder, G. Electrodes with high power and high capacity for rechargeable lithium batteries. *Science* **311**, 977–980 (2006).

11. Hafner, J., Wolverton, C. & Ceder, G. Toward computational materials design: The impact of density functional theory on materials research. *MRS Bull.* **31**, 659–668 (2006).
12. Greeley, J., Jaramillo, T. F., Bonde, J., Chorkendorff, I. & Nørskov, J. K. Computational high-throughput screening of electrocatalytic materials for hydrogen evolution. *Nat. Mater.* **5**, 909 (2006).
13. Hautier, G. *et al.* Phosphates as lithium-ion battery cathodes: An evaluation based on high-throughput ab initio calculations. *Chem. Mater.* **23**, 3495–3508 (2011).
14. Hautier, G., Ong, S. P., Jain, A., Moore, C. J. & Ceder, G. Accuracy of density functional theory in predicting formation energies of ternary oxides from binary oxides and its implication on phase stability. *Phys. Rev. B* **85**, 155208 (2012).
15. Saal, J. E., Kirklin, S., Aykol, M., Meredig, B. & Wolverton, C. Materials design and discovery with high-throughput density functional theory: The open quantum materials database (OQMD). *JOM* **65**, 1501–1509 (2013).
16. Garrity, K. F., Bennett, J. W., Rabe, K. M. & Vanderbilt, D. Pseudopotentials for high-throughput DFT calculations. *Comput. Mater. Sci.* **81**, 446–452 (2014).
17. Emery, A. A., Saal, J. E., Kirklin, S., Hegde, V. I. & Wolverton, C. High-throughput computational screening of perovskites for thermochemical water splitting applications. *Chem. Mater.* **28**, 5621–5634 (2016).
18. Hinuma, Y., Hayashi, H., Kumagai, Y., Tanaka, I. & Oba, F. Comparison of approximations in density functional theory calculations: Energetics and structure of binary oxides. *Phys. Rev. B* **96**, 094102 (2017).
19. Ward, L., Agrawal, A., Choudhary, A. & Wolverton, C. A general-purpose machine learning framework for predicting properties of inorganic materials. *NPJ Comput. Mater.* **2**, 16028 (2016).
20. Lee, J., Seko, A., Shitara, K., Nakayama, K. & Tanaka, I. Prediction model of band gap for inorganic compounds by combination of density functional theory calculations and machine learning techniques. *Phys. Rev. B* **93**, 115104 (2016).
21. Jain, A., Hautier, G., Ong, S. P. & Persson, K. New opportunities for materials informatics: Resources and data mining techniques for uncovering hidden relationships. *J. Mater. Res.* **31**, 977–994 (2016).
22. Sun, W. *et al.* The thermodynamic scale of inorganic crystalline metastability. *Sci. Adv.* **2**, e1600225 (2016).
23. Hohenberg, P. & Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **136**, B864 (1964).
24. Kohn, W. & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133 (1965).
25. Ceperley, D. M. & Alder, B. J. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.* **45**, 566 (1980).
26. Kohn, W. Nobel Lecture: Electronic structure of matter—wave functions and density functionals. *Rev. Mod. Phys.* **71**, 1253 (1999).
27. Becke, A. D. Perspective: Fifty years of density-functional theory in chemical physics. *J. Chem. Phys.* **140**, 18A301 (2014).
28. Hammer, B., Hansen, L. B. & Nørskov, J. K. Improved adsorption energetics within density-functional theory using revised Perdew–Burke–Ernzerhof functionals. *Phys. Rev. B* **59**, 7413 (1999).
29. Griffin, S. M. & Spaldin, N. A. A density functional theory study of the influence of exchange–correlation functionals on the properties of FeAs. *J. Phys. Condens. Matter* **29**, 215604 (2017).
30. Yuk, S. F. *et al.* Towards an accurate description of perovskite ferroelectrics: Exchange and correlation effects. *Sci. Rep.* **7**, 43482 (2017).
31. Zicovich-Wilson, C. *et al.* Calculation of the vibration frequencies of α -quartz: The effect of Hamiltonian and basis set. *J. Comput. Chem.* **25**, 1873–1881 (2004).
32. De-La-Pierre, M. *et al.* Performance of six functionals (LDA, PBE, PBESOL, B3LYP, PBE0, and WC1LYP) in the simulation of vibrational and dielectric properties of crystalline compounds: The case of forsterite Mg_2SiO_4 . *J. Comput. Chem.* **32**, 1775–1784 (2011).
33. Harris, J. Simplified method for calculating the energy of weakly interacting fragments. *Phys. Rev. B* **31**, 1770 (1985).
34. Burke, K. Perspective on density functional theory. *J. Chem. Phys.* **136**, 150901 (2012).
35. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865 (1996).
36. Perdew, J. P. *et al.* Restoring the density-gradient expansion for exchange in solids and surfaces. *Phys. Rev. Lett.* **100**, 136406 (2008).
37. Heyd, J., Peralta, J. E., Scuseria, G. E. & Martin, R. L. Energy band gaps and lattice parameters evaluated with the Heyd–Scuseria–Ernzerhof screened hybrid functional. *J. Chem. Phys.* **123**, 174101 (2005).
38. Kim, K. & Jordan, K. Comparison of density functional and MP2 calculations on the water monomer and dimer. *J. Phys. Chem.* **98**, 10089–10094 (1994).
39. Stephens, P. J., Devlin, F., Chabalowski, C. & Frisch, M. J. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.* **98**, 11623–11627 (1994).
40. Perdew, J. P., Ernzerhof, M. & Burke, K. Rationale for mixing exact exchange with density functional approximations. *J. Chem. Phys.* **105**, 9982–9985 (1996).
41. Heyd, J., Scuseria, G. E. & Ernzerhof, M. Hybrid functionals based on a screened Coulomb potential. *J. Chem. Phys.* **118**, 8207–8215 (2003).
42. Harrison, W. A. Bond-orbital model and the properties of tetrahedrally coordinated solids. *Phys. Rev. B* **8**, 4487 (1973).
43. Fischer, C. F. Average-energy-of-configuration Hartree–Fock results for the atoms helium to radon charlotte froese fischer. *Atom. Data Nucl. Data Tabl.* **12**, 301–399 (1972).
44. Froese-fischer, C. Erratum: Average-energy-of-configuration Hartree–Fock results for the atoms helium to radon. *At. Data Nucl. Data Tables* **12**, 87 (1973).
45. Dion, M., Rydberg, H., Schröder, E., Langreth, D. C. & Lundqvist, B. I. Van der Waals density functional for general geometries. *Phys. Rev. Lett.* **92**, 246401 (2004).
46. Cooper, V. R. Van der Waals density functional: An appropriate exchange functional. *Phys. Rev. B* **81**, 161104 (2010).
47. Thonhauser, T. *et al.* Van der Waals density functional: Self-consistent potential and the nature of the van der Waals bond. *Phys. Rev. B* **76**, 125112 (2007).
48. Klimeš, J., Bowler, D. R. & Michaelides, A. Van der Waals density functionals applied to solids. *Phys. Rev. B* **83**, 195131 (2011).
49. Berland, K. *et al.* van der Waals density functionals built upon the electron-gas tradition: Facing the challenge of competing interactions. *J. Chem. Phys.* **140**, 18A539 (2014).
50. Berland, K. *et al.* van der Waals forces in density functional theory: a review of the vdW-DF method. *Rep. Prog. Phys.* **78**, 066501 (2015).
51. Lejaeghere, K., Van Speybroeck, V., Van Oost, G. & Cottenier, S. Error estimates for solid-state density-functional theory predictions: An overview by means of the ground-state elemental crystals. *Crit. Rev. Solid State Mater. Sci.* **39**, 1–24 (2014).
52. K. Lejaeghere, in *Uncertainty Quantification in Multiscale Materials Modeling*. (Elsevier, 2020), pp. 41–76.
53. Mortensen, J. J. *et al.* Bayesian error estimation in density-functional theory. *Phys. Rev. Lett.* **95**, 216401 (2005).
54. Proppe, J. & Reiher, M. Reliable estimation of prediction uncertainty for physicochemical property models. *J. Chem. Theory Comput.* **13**, 3297–3317 (2017).
55. Wellendorff, J. *et al.* Density functionals for surface science: Exchange–correlation model development with Bayesian error estimation. *Phys. Rev. B* **85**, 235149 (2012).
56. Wellendorff, J., Lundgaard, K. T., Jacobsen, K. W. & Bligaard, T. mBEEF: An accurate semi-local Bayesian error estimation density functional. *J. Chem. Phys.* **140**, 144107 (2014).
57. Wang, Y., McDowell, D. L. *Uncertainty Quantification in Multiscale Materials Modeling*. (Woodhead Publishing, 2020).
58. Kosłowski, M. & Strachan, A. Uncertainty propagation in a multiscale model of nanocrystalline plasticity. *Reliab. Eng. Syst. Saf.* **96**, 1161–1170 (2011).

59. Pande, V. & Viswanathan, V. Robust high-fidelity DFT study of the lithium-graphite phase diagram. *Phys. Rev. Mater.* **2**, 125401 (2018).
60. Deshpande, S., Kitchin, J. R. & Viswanathan, V. Quantifying uncertainty in activity volcano relationships for oxygen reduction reaction. *ACS Catal.* **6**, 5251–5259 (2016).
61. Pandey, M. & Jacobsen, K. W. Heats of formation of solids with error estimation: The mBEEF functional with and without fitted reference energies. *Phys. Rev. B* **91**, 235201 (2015).
62. Medford, A. J. *et al.* Assessing the reliability of calculated catalytic ammonia synthesis rates. *Science* **345**, 197–200 (2014).
63. Pernot, P., Civalleri, B., Presti, D. & Savin, A. Prediction uncertainty of density functional approximations for properties of crystals with cubic symmetry. *J. Phys. Chem. A* **119**, 5288–5304 (2015).
64. De Waele, S., Lejaeghere, K., Sluydts, M. & Cottenier, S. Error estimates for density-functional theory predictions of surface energy and work function. *Phys. Rev. B* **94**, 235418 (2016).
65. Lejaeghere, K., Jaeken, J., Van Speybroeck, V. & Cottenier, S. Ab initio based thermal property predictions at a low cost: An error analysis. *Phys. Rev. B* **89**, 014304 (2014).
66. Krogel, J. T. Nexus: A modular workflow management system for quantum simulation codes. *Comput. Phys. Commun.* **198**, 154–168 (2016).
67. De Bièvre, P. The 2012 international vocabulary of metrology: “VIM”. *Chem. Int.-Newsmagazine IUPAC* **34**, 26–27 (2012).
68. Jain, D., Chaube, S., Khullar, P., Srinivasan, S. G. & Rai, B. Bulk and surface DFT investigations of inorganic halide perovskites screened using machine learning and materials property databases. *Phys. Chem. Chem. Phys.* **21**, 19423–19436 (2019).
69. Alade, I. O., Olumegbon, I. A. & Bagudu, A. Lattice constant prediction of A₂XY₆ cubic crystals (A = K, Cs, Rb, TI; X = tetravalent cation; Y = F, Cl, Br, I) using computational intelligence approach. *J. Appl. Phys.* **127**, 015303 (2020).
70. T. E. Oliphant, *A Guide to NumPy*. (Trelgol Publishing USA, 2006), vol. 1.
71. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
72. Ramakrishnan, R., Dral, P. O., Rupp, M. & von Lilienfeld, O. A. Big data meets quantum chemistry approximations: The Δ -machine learning approach. *J. Chem. Theory Comput.* **11**, 2087–2096 (2015).
73. Ernzerhof, M., Perdew, J. P. & Burke, K. Coupling-constant dependence of atomization energies. *Int. J. Quantum Chem.* **64**, 285–295 (1997).
74. Gritsenko, O., Mentel, L. & Baerends, E. On the errors of local density (LDA) and generalized gradient (GGA) approximations to the Kohn–Sham potential and orbital energies. *J. Chem. Phys.* **144**, 204114 (2016).
75. R. Jones, in *Computational Nanoscience: Do It Yourself*. (John von Neumann Institute for Computing, 2006), pp. 45–70.
76. Medvedev, M. G., Bushmarinov, I. S., Sun, J., Perdew, J. P. & Lyssenko, K. A. Density functional theory is straying from the path toward the exact functional. *Science* **355**, 49–52 (2017).

Acknowledgements

S.F.Y., J.T.K. and V.R.C. (first principles calculations and high-throughput workflow efforts) were supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division. I.S. and S.P.B. (data analytics) were supported in part by the U.S. Department of Energy under contract 89304017CEM000001. N.M. acknowledges summer support through the ORNL HERE Program, sponsored by the U.S. Department of Energy and administered by the Oak Ridge Institute for Science and Education. We gratefully acknowledge the computational resources provided by the National Energy Research Scientific Computing Center (NERSC), which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Disclaimer

The opinions expressed herein are those of the authors and not necessarily representative of those of the Department of the Army, Department of Defense (DoD), or U.S. Government.

Author contributions

V.R.C. conceived the original idea. S.F.Y. performed the high-throughput DFT calculations. I.S. carried out the data-science methods. J.T.K. developed a high-throughput workflow management module. S.F.Y. and I.S. prepared the figures and tables for the manuscript. S.F.Y. and V.R.C. led the analysis efforts on DFT-predicted results. I.S. and S.P.B. led the data informatics analysis. N.M. and V.R.C. made the comparison efforts on DFT-predicted results. S.F.Y. and I.S. led the writing efforts of manuscript and supplementary materials. All the authors reviewed and edited the manuscript and supplementary materials. This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

Funding

Basic Energy Sciences, DOE, 89304017CEM000001, 89304017CEM000001, HERE program (ORNL).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-69194-w>.

Correspondence and requests for materials should be addressed to V.R.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© UT-Battelle, LLC and Irmak Sargin, Scott P. Beckman 2024