Machine-learning enabled thermodynamic model for the design of new rare-earth compounds

P. Singh\textsuperscript{a,c}, T. Del Rose\textsuperscript{a,b}, G. Vazquez\textsuperscript{c}, R. Arroyave\textsuperscript{d}, Y. Mudryk\textsuperscript{a}

\textsuperscript{a}Ames Laboratory, U.S. Department of Energy, Iowa State University, Ames, Iowa 50011 USA
\textsuperscript{b}Materials Science & Engineering, Iowa State University, Ames, Iowa 50011 USA
\textsuperscript{c}Department of Materials Science and Engineering, Texas A&M University, College Station, TX 77843, USA
\textsuperscript{d}Department of Mechanical Engineering, Texas A&M University, College Station, TX 77843, USA

Abstract

We employ a descriptor based machine-learning approach to assess the effect of chemical alloying on formation enthalpy of rare-earth intermetallics. Application of machine-learning approaches in rare-earth intermetallic design have been sparse due to limited availability of reliable datasets. In this work, we developed an ‘in-house’ rare-earth database with more than 600+ compounds, each entry was populated with formation enthalpy and related atomic features using high-throughput density-functional theory (DFT). The SISSO (sure independence screening and sparsifying operator) based machine-learning method with meaningful atomic features was used for training and testing the formation enthalpies of rare earth compounds. The complex lattice function coupled with the machine-learning model was used to explore the effect of transition metal alloying on the energy stability of Ce based cubic Laves phases (MgCu\textsubscript{2} type). The SISSO predictions show good agreement with high-fidelity DFT calculations and X-ray powder diffraction measurements. Our study provides quantitative guidance for compositional considerations within a machine-learning model and discovering new metastable materials. The electronic-structure of Ce-Fe-Cu based compound was also analyzed in-depth to understand the electronic origin of phase stability. The interpretable analytical models in combination with density-functional theory and experiments provide a fast and reliable design guide for discovering technologically useful materials.

Keywords: Rare-earths, Thermodynamic stability, Density-functional theory, Machine Learning, X-Ray powder diffraction

Introduction and Background

Rare earths find uses in many applications due to their vast span of distinctive physical and chemical properties. Well-known applications include wind turbines, hybrid and electric vehicles, solid state lighting, mobile devices, lasers and optical fibers, fuel cracking and other catalysts, all of which owe most of their core functionalities to 4f-electron compounds \[1\]-\[13\]. A number of future technologies such as magnetic cooling \[14, 15\] rely on rare earths for both magnetocaloric materials and magnetic field sources. Recently, a global “push” for a greener, hydrogen economy brings attention towards solid-state hydrogen liquefaction technology, where rare earth alloys (in particular RT\textsubscript{2} compounds where T = Al, Ni, Co) have potential to play a vital role \[16, 19\]. However, the discovery of rare earth compounds is often built upon serendipitous findings, as can be exemplified by hard rare-earth magnets \[20-22\]. While such encounters will remain a part of future scientific progress in, there is a growing consensus that predictive science enabled by recent developments in theory and machine learning will lead the way for sustainable innovation \[23\].

*Corresponding authors email: p Singh84@ameslab.gov
The lanthanides (mostly trivalent in nature with the exception of Eu and Yb, where a divalent state could be more stable, and Ce, which is often tetravalent) and their compounds have always been a fertile ground for informatics and prediction due to their peculiar chemistry. The concept of the "fraternal fifteen", imagined by K.A. Gschneidner, Jr., presented lanthanides as consecutive houses on the same street but with different number of "kids" or 4f electrons [24]. When moving through the lanthanides, the "houses" (sub-orbitals, e.g., 4f) are filled with varying numbers of kids (f-electrons), while the house’s exterior (outer shells) remain relatively unchanged ([Xe]4f<sup>n</sup>5d<sup>1</sup>6s<sup>2</sup> or, in some cases, [Xe]4f<sup>n</sup>5d<sup>1</sup>6s<sup>2</sup>)t. A famous example constitutes a generalized phase diagram for the trivalent R<sub>R</sub> electronic structure of lanthanide alloys, which predicts crystal structure and melting temperature of compositions with any specific z value [25]. A modified version of this diagram was constructed to predict phase boundaries of intra-lanthanide alloys at pressures up to 25 GPa [26]. This chemical similarity together with a well-established evolution of atomic radii, known as lanthanide contraction, and established rules of some physical behaviors (e.g., de Gennes rule for magnetism) provide a fertile ground for systematic generalization of respective physical and chemical behaviors, suitable for predictive science.

Machine learning models have proven to be useful in identifying crystallographic information of ternary equiatomic rare earth compounds [27,28]. These ML efforts, however, are often limited in their ability to describe true thermodynamic behavior of new rare earth compounds or compositions. Nonetheless, despite the importance of rare-earths based materials and their solid background, modern machine learning and artificial intelligence tools for the targeted investigation of rare earths remain sparse due to a lack of reliable databases with sufficient number of entries required for model training. Naturally, this emphasizes the need of models that can interpolate data from sparse databases. The SISSO (Sure Independence Screening and Parsifying Operator) is one such machine-learning method [29,30] that returns very accurate predictions with limited information [31,34]. The readiness of SISSO-based analytical models using regression and classification makes the model interpretation easier compared to other machine learning methods that face the non-trivial question on the representation of material design space. The SISSO models return the best analytical descriptors trained with limited information for a target property, e.g., thermodynamic stability or formation enthalpy, from a vast feature space constructed from an operator set (e.g., ‘+’, ‘−’, ‘exp’, ‘log’ etc.) and an initial primary feature set (e.g., atomic-size, valence-electron count, Allen electronegativity etc.) [29].

In this work, we present a rapid assessment of alloying effects on the thermodynamic stability of REX<sub>3</sub> (RE=Rare-earth; X=transition-metals) type rare-earth intermetallics using machine-learning (SISSO) trained inexpensive analytical models. The machine-learning model was trained over the density-functional theory (DFT) generated in-house database with 600+ entries. During the training of the model, a total of 26 material descriptors were shortlisted from an extensive pool of 89 features (see supplement) based on their fundamental relevance and statistical correlation with formation enthalpy (see Fig. S1). The Ce-Fe-Cu rare-earth compounds, in this work, were chosen for the model validation. The potential application of Ce-based compounds as high performance magnets make them an ideal candidate for phase stability and electronic-structure analysis. The SISSO trained three-dimensional (3D) descriptor was used to analyze the alloying effect on the phase stability of RE(TM<sub>1</sub>,TM<sub>2</sub>)<sub>2</sub> compounds with cubic structure (MgCu<sub>2</sub> type), where RE=Ln; TM<sub>1</sub>=Mn/Fe/Cu/Co; and TM<sub>2</sub>=Pd/Pt/Re/Rh, and z is the variation in transition metal composition. The phase stability trends are validated through very careful synthesis of Ce-Fe-Cu based rare-earth compounds, i.e., Ce(Fe,Cu)<sub>1−z</sub>_z at z(Fe)= 25, 50 and 75 at.%Fe, followed by detailed X-ray powder diffraction analysis. A comprehensive analysis of chemical alloying on the electronic-structure (the band-structure and Fermi-surface) of Ce-Fe-Cu compounds in cubic (MgCu<sub>2</sub>) and orthorhombic (KHg<sub>2</sub>) phase is also presented to understand the quantum mechanical origin of phase stability. The change in band-structure with Cu strongly correlates with the predicted phase stability. Our work provides a logical next step in the informed discovery of new rare earth compounds by integrating machine learning models with high throughput DFT calculations. This further emphasizes the need of computationally inexpensive and interpretable models for the accelerated discovery of complex alloys.

Methods

Database generation

Structure considerations for database generation: Ames laboratory internal Rare-Earth Information Center (RIC) database, containing over 100,000 references to rare earth materials published before the early 2000s, was scraped for compositional information. Of the resulting 59,000+ potential compositions, select rare earth material classes (REX<sub>2</sub>
were done using 2 high convergence criterion for force (10−6 eV/Å and 10−6 eV). Each crystal structure was fully relaxed (volume count). The generalized gradient approximation of Perdew, Burke and Ernzerhof (PBE) was employed in all calculations [37]. The DFT+U [38] is an alternate scheme for materials with localized d− and f−electrons. However, the non transferability of parameter U across compounds and its arbitrary nature, where the choice of U can strongly influence the observables [39], makes the high-throughput use of DFT+U approach impossible. Franchini et al. [40] have shown that DFT+U method introduces an average (minimum) relative error of about 4% in calculation of heat of formation in binary-manganese oxides, where a single U value was used. Additionally, Söderling et al. [41] also established the effectiveness of GGA functionals in describing the properties of rare-earth metals. The hybrid functionals are less relevant to this discussion for a couple of reasons: (i) most hybrid-functionals are one-shot calculations (no self-consistency), and (ii) they are computationally very demanding, therefore, not useful for machine-learning approaches. Furthermore, the rationale to choose GGA (PBE) over LDA, meta-GGA [42, 43] or hybrid functionals [44] was inspired by the work of Giese and York [45] which highlights the advantage of GGA functionals. Therefore, PBE (GGA) was used in all our calculations [37].

Based on energy and k-mesh convergence tests, we set up high-energy plane-wave cutoff of 520 eV both in relaxation and charge self-consistency. Each crystal structure was fully relaxed (volume + atomic-positions) with high convergence criterion for force (10−3 eV/Å) and energy (10−6 eV). Full relaxation and charge self-consistency were done using 2×2×2 to 8×8×8 depending on total number of atoms per cell (the inverse relation between k-space and r-space helps to choose smaller k-mesh for larger cells and vis-a-vis). The Γ-centered Monkhorst-Pack k-mesh was used for Brillouin zone integration during structural-optimization and charge self-consistency [46]. The shift away from Γ was originally meant to reduce the size of the k−mesh. In metallic compounds, Γ−centered sampling was chosen as it is often a better choice for its faster-convergence than the shifted grid [47].

Formation enthalpy estimate: The formation enthalpy (E_form) of rare-earth based REX2 systems was estimated using the formula E_{form}^{REX_2} = \sum E_i^\text{REX_2} - \sum N_i E_i where E_{total}^\text{REX_2} is the total energy per unit cell of the intermetallic rare-earth compounds per unit cell, N_i and E_i are the number of atoms and ground state total energy of element ‘i’ (e.g., Fe in body-centered cubic, and/or Ce in face-centered cubic).

**SISSO- model training and accuracy**

The SISSO-based ML approach combines regression with compressed sensing to identify high-order analytical models, for example, 1D, 2D, 3D, 4D or even higher dimensional interpretable descriptors [29, 30]. The SISSO models work in a hierarchical order where this first creates a new feature space of varying complexity through algebraic operations (+, −, ×, e^x, e^−x, (−)^1, (−)^2, (−)^3, √, log) on primary feature sets (variables). This allows us to create sufficiently large feature space. All features generated through the combination of the primary features and operators were added recursively to the feature set. The sparse regression method then filters out significant features from an originally large feature space based on their frequency in descriptor generation. We have performed the SISSO training and testing using 5- and 10-fold cross-validation.

The SISSO models were trained on all lanthanides (from La to Lu except Pm) and transition-metals (VIIA-XIIA), as well as IIA, IIA (B-Ga), IVA (C-Sn), VA (N-Te), VIA (O-Te), VIIA (F-I) group elements. The use of such diverse dataset makes the model generalized enough to make predictions of any rare-earth based REX2 compounds. Since the group IVB (Ti, Zr, Hf), VB (V, Nb, Ta), and VIB (Cr, Mo, W) transition metals were not included in the model training, so the confidence in prediction may be low for REX2 compounds with X=IVB, VB, VIB group elements.

**Weighted sublattice description**

In general, the featurization of ML-based models involves composition based vectors for application to disorder materials. Evidently, the number of features grows with increasing alloy complexity, therefore, it would be more
convenient if we separate the alloy features by the role of constituent elements. We used the idea of sublattice
description to distinguish the lattice sites by rare earth (RE) and transition metals (X) in REX. An average value was
assigned to each Wyckoff site, i.e., 8a (RE) and 16d (X), where the transition-metals are twice the amount of rare-
earths. A weight feature was created (e.g., ‘+ plus’ and ‘+plus-2’; ‘− minus’ and ‘− minus-2’) to calculate the contribution of
both the sites according to their weights. An average over the lattice-cell was calculated as an weighted average and
weighted standard deviation over the whole cell (see supplemental Eq. S1-S3).

Experimental details

Three Ce(Fe,Cu) samples with z (Fe)=25, 50, 75 at.% were prepared to verify the mixability of Fe and Cu in the
Ce(Fe,Cu) solid solution. The samples were arc-melted using the high-purity metals: Fe and Cu were at least
99.95 wt.% pure; Ce (at least 99.9 wt.% pure) was provided by Materials Preparation Center of Ames Laboratory. The
samples were wrapped in Ta foil, sealed in evacuated quartz ampoules back-filled with He gas, and annealed at 800°C
for 2 weeks. The samples were slowly cooled in the furnace and analyzed using X-ray powder diffraction (XRPD) and
electron microscopy (SEM/EDX). The SEM analysis revealed strong surface oxidation, however, the EDX confirmed
that the chemical composition of the prepared samples matches the nominal one. The phase analyses of the XRPD
data was performed using Rietica software.

Results and Discussion

Formation enthalpy - training, model-generation, cross-validation, and feature analysis

Training, testing and descriptor analysis: Machine learning and data analytics can accelerate materials design and
discovery through the use of descriptors [29, 30]. Ouyang et al. [29, 30] has shown the usefulness of symbolic
regression in SISSO that allow us to develop analytical descriptors for target properties [57–59]. Despite descriptors
becoming a standard approach for material discovery, literature remains sparse on databases containing thermody-
namic stability information (e.g., formation enthalpy; E_form) of rare-earth based inorganic crystalline solids.

The E_form is a fundamental material property and an important indicator of thermodynamic stability. The ‘in-
house’ Ames Lab Rare-Earth Information Center (RIC) database (‘Ames Lab RIC 2.0’) was used to train the SISSO-
based formation enthalpy descriptor. An analytical 3 dimensional (3D) formation enthalpy descriptor from SISSO
model training was presented in Eq. 1

\[
E_{\text{form}} = -158.157 \times 10^{-3} \times \log \left[ r_c \times \rho_{\text{avg}} \times \chi_{\text{diff}}^{\text{Pauling}} \right] + 18.585 \times 10^{-3} \times \rho_{\text{avg}} \times \rho_{\text{diff}} \chi_{\text{avg}}^2 (\text{1})
\]

\[
+ 3.28 \times 10^{-3} \times (\text{group}_{\text{avg}} - \text{group}_{\text{diff}}) \times (r_{\text{Zunger}} - r_{\text{M miracle}}),
\]

where, \(r_c\) is the covalent-radii (Å), \(\rho_{\text{avg}}\), and \(\text{group}_{\text{avg}}\) are averaged with stoichiometrically weighted mean and \(\rho_{\text{diff}},\r_{\text{diff}}\); \(\chi_{\text{diff}}\), and \(\text{group}_{\text{diff}}\) are stoichiometrically weighted harmonic mean, see Eq. S1-S3 in the supplement. The
‘in-house’ rare-earth database was randomly divided (80:20 ratio) for training and testing of the SISSO based machine-
learning model. The Eq. 1 will allow the high-throughput determination of thermodynamic stability of the specific
crystal phase and compositions that could be synthesized experimentally. The analytically constructed descriptor in
Eq. 1 also conserves the unit of feature sets within SISSO [29].

The sensitivity of 3D descriptors on training data was tested by performing 5 (10)-fold cross validation of Boufounos
et al. [60], where ‘Ames Lab RIC 2.0’ was equally and randomly divided into 5 (10) sets. Following this, one out
of 5 (10) set was left out as a test set and remaining were included in the training. The cross validation was iterated 5
(10) times so that each part was used as a test once during the training. The cross-validation often helps to determine
overfitting or underfitting of the models. Notably, dimensionality (either too large or too small) is the only source of
overfitting in SISSO [29]. In 3D vs 4D descriptor comparison, we found no visible improvements in model accuracy
as shown by root-mean squared error (RMSE; 0.17 eV-atom\(^{-1}\) (3D) vs 0.16 eV-atom\(^{-1}\) (4D)) and coefficient of
determination (R\(^2\); 0.79 (3D) vs 0.82 (4D)) in Fig. 1 (the error analysis see Eq. S4-S5). Therefore, in this work, the
3D descriptor was was used throughout for thermodynamic stability analysis of REX compounds. The observation
about dimensionality was found in good agreement with Ouyang et al. [29], where the disadvantages of using low or
higher order descriptors were clearly outlined. The computational cost is also a critical issue to keep in mind while choosing model dimensionality; therefore, it becomes critical to limit unnecessary features, which is only possible with the choice of optimal descriptor dimension.

Figure 1: Predicted (descriptor in Eq. 1) vs actual (DFT) formation enthalpies from (a,b) 3D, and (c,d) 4D descriptor.

Going back to the interpretation of formation-energy descriptor, a strong correlation among features such as electronegativity, density and atomic-radii with $E_{\text{form}}$ in Eq. 1 makes sense. The atomic stability, also discussed in literature [61], could be rationalized based on atomic radii and electronegativity, which has previously shown usefulness in describing bond character and compound formation [62, 63]. For example, the electronic distribution around a nucleus in an atom determines the atomic-size, i.e., electrons closer to nuclei make a tightly bound system that eventually enhances the electronegativity [64]. The increased screening by inner electrons moving down the group in the periodic table decreases the attraction force on the valence electrons, further balancing out the bonding with the nucleus. The change in electron screening reduces the electronegativity and increases the atomic-radii of the elements.

Feature analysis: The feature complexity was found to increase with increasing model dimensionality as shown by the large feature sets in higher order feature spaces in Fig. 2. The Sure Independence Screening (SIS) creates subspaces of strongly correlated features for the property of interest, in our case formation-enthalpy, from an exponentially large feature set of cardinality $\approx 10^{10}$. The correlated features in the subspace were operated on by an $l_0$-norm regularized minimization Sparsifying Operator (SO ($l_0$)) to help find the best descriptor. The $l_0$-norm of a vector is the number of its non-zero components. The number of terms in the descriptor space depends on dimensionality of the subset space and can be mathematically represented by n-combination (where n represents the dimensionality, i.e., 1D, 2D, and 3D), for example, the 3D descriptor chooses 3 set of features from an exponentially large feature space, i.e., \[ \binom{3000}{3} \approx 4.5 \times 10^9. \]

The frequency of features in the 3D descriptor space for a five-fold cross-validation maximizes at 10—see Fig. 2. The most frequent features are those generated through the stoichiometrically-weighted mean (e.g., $\rho_{\text{avg}}$), stoichiometrically-weighted harmonic mean (e.g., $r_{\text{diff}}$), and stoichiometrically-weighted mean difference (e.g., $r_{\text{red}}$) [31]. Here, we stick to a 3D descriptor space as changes in features may cause over-fitting in top performing higher dimensional descriptors [29]. The feature frequency in Fig. 2 was found near constant for 3D descriptors (4-to-10 features) and performs well both in training and cross-validation. The features in the final descriptor trained
on the full dataset also follow the same trend in choosing the most recurring features in cross-validation. Notably, the 3D descriptor in Eq. 1 shows higher confidence in top features in Fig. 2 such as covalent radii ($r_c$), average density ($\rho_{avg}$), Pauling electronegativity ($\chi_{Pauling}$), group average ($group_{avg}$), and group difference ($group_{diff}$). We also found that the higher frequency of selected features appearing in each run improves the accuracy of the model and makes the descriptor acceptable for phase stability prediction.

As pointed out by Schmidt et al. [65], ideally, descriptors should be uncorrelated, as an abundant number of correlated features can hinder both accuracy and efficiency of the model. This suggests that one needs further feature selection to circumvent the negative effect of dimensionality [66], simplify the models, and improve their interpretability as well as training efficiency [65]. As discussed by Ouyang et al. [29], the analytical nature of the SISSO model helps to circumvent the effect of dimensionality.

Thermodynamic stability assessment of MgCu$_2$ type Ce compounds using the E$_{form}$ descriptor

Background and crystal symmetry of cubic rare-earth Laves phases (MgCu$_2$ type)

Of the 1,200+ Laves phases known today, about 60% contain rare earths, and nearly ~200 of those are binaries. Their excellent chemical and physical properties give them a unique advantage for application as functional materials [67]. The RE(TM)$_2$ Laves phase compounds are the closely related intermetallics that crystallizes in three different structures such as MgCu$_2$ (C15), MgZn$_2$ (C14), and MgNi$_2$ (C36). Recently, some of the non rare-earth C14 Laves phases such as (Hf$_{1-x}$Ta$_x$)Fe$_2$, (Sc$_{1-x}$Ti$_x$)Fe$_2$ and (Zr$_{1-x}$Nb$_x$)Fe$_2$ were found with interesting compositional and temperature dependence of their magnetic properties [68-70]. Similar effects were found in rare-earths such as Ce$_{1-x}$(Ru$_x$Fe$_{1-x}$)$_2$ with C15 crystal phase [71, 72]. Despite the intriguing nature of rare-earth based compounds, the alloying effects on the thermodynamic stability was not systematically assessed either with machine-learning or high-throughput DFT methods.

Here, we explored MgCu$_2$-type Laves phases as model systems as they are the most representative among the rare earth compounds. [67]. The MgCu$_2$ crystal phase has face-centered cubic symmetry with $Fd\bar{3}m$ (#227) space-group. The cubic MgCu$_2$ phase is alternatively represented by its common Strukturbericht designation C15. The Laves phases contains two ($Z=2$) formula units (f.u.) in primitive and eight ($Z=8$) f.u. in conventional cells. The Cu-site forms a tetrahedron at Wyckoff position 16c [3/8, 3/8, 3/8] and Mg-site arranged on diamond cubic lattice.
with Wyckoff position 8b [0, 0, 0], totaling 24 atoms per conventional unit cell as shown in Fig. 3. Similar to the pure metals, the rare-earth elements are arranged in a ABCABC packing sequence in the cubic MgCu$_2$-type structure.

Figure 3: (Color Online) Crystal structure of the MgCu$_2$-type structure. The large (red) atoms represent the Rare Earth (Mg-site), and the smaller (cyan) atoms represent the transition metal (Cu-site).

Thermodynamic stability is a primary fundamental quantity that must be evaluated to determine synthesizability. The main idea behind the exploration of the thermodynamic phase space of rare-earths compounds is to find out the rule of mixing that governs phase selection. Cerium-based binary and ternary intermetallics have been exhaustively explored in the past for their electronic properties [73], and large magnetic response [74-77]. Ce-based alloys are fundamentally interesting for their rich physics [78-82] that can further be tuned using chemical alloying, which may also help improve the physical properties. The polymorphism in Ce is also well established [83], where different stable phases $\alpha, \alpha', \beta, \gamma, \delta$ have been reported at different thermodynamic conditions [84, 85]. Other noted properties, e.g., superconductivity [86, 87], heavy-fermion behavior [88, 89], and complex magnetic properties [90-92] due to varying oxidation states of Ce (nonmagnetic Ce$^{4+}$ ([Xe]$4f^0$), magnetic Ce$^{3+}$ ([Xe]$4f^1$)) make Cerium based system an interesting playground for new discoveries [93].

Effects of chemical mixing of 3d and 4d/5d transition metals on phase stability

The thermodynamic stability analysis using experiments and direct Ab-initio (DFT) approaches is a time-consuming process [94]. Here, we used a descriptor based analytical model as presented in Eq. 1 for the high-throughput assessment of the phase stability of disordered rare-earths with MgCu$_2$ (cubic) crystal phase. We extensively explored the effect of transition metal alloying on the phase stability of pseudo-binary Ce(TM$_1$,TM$_2$,..)$_2$ compounds, where the 3d (TM$_1$=Co/Fe/Mn/Ni) and 4d/5d (TM$_2$=Pd/Pt/Re/Rh) transition metals are alloyed at 16c Wycoff site in Fig. 3 to understand their effect on thermodynamic stability. In Fig. 3, we show the effect of Pd/Pt/Re/Rh alloying with Co on the phase stability of Ce-Co pseudo-binaries. Chemical mixing of Pd or Pt with Co in Fig. 3 shows improvement in the phase stability of Ce-Co-Pd/Pt compound. However, the linear trend suggests that Co and Pd do not want to mix together rather they may prefer to phase separate but still coexist as two phase region of CeCo$_2$ and CePd$_2$. Similarly, the positive slope of phase stability in Fig. 3 for Co-rich Co-Pt compound shows un-mixing of Pt at Co, i.e., Co energetically does not prefer to mix with Pt at transition metal site in Fig. 3.
Figure 4: (Color Online) Thermodynamic stability of cerium based pseudo-binary compounds (a-a) Ce(Co$_z$TM$_{2(1-z)}$)$_2$, (b-b) Ce(Fe$_z$TM$_{2(1-z)}$)$_2$, (c-c) Ce(Mn$_z$TM$_{2(1-z)}$)$_2$, and (d-d) Ce(Ni$_z$TM$_{2(1-z)}$)$_2$, where TM$_2$=Pd/Pt/Re/Rh; and $z$ is elemental composition in atomic-percent.
In Fig. 4b, the strongly reduced (weak formation enthalpy) phase stability of Ce-Co compounds on alloying Co with Re agrees well with Co-Re binary phase diagram, where hexagonal phase becomes more stable at higher at.% Re. Meanwhile, no experimental reports on cubic CeRe₂ further affirms the validity of our predictions [95]. The phase stability analysis indicates that CeRe₂ is very weakly stable, i.e., even if the compound forms it is either metastable or dynamically unstable in ambient conditions. Therefore, it needs to be looked with respect to other competing phases such as hexagonal MgNi₂ crystal type. Notably, the ML predicted formation enthalpies for CeRe₂/Ce(Re)₂/CeCo₂ (-0.09/-0.085/-0.17 eV-atom⁻¹) show reasonable agreement with direct DFT calculations (-0.06/-0.11/-0.26 eV-atom⁻¹). The jump in formation enthalpy at intermediate compositions (near 50at.% Re) of Ce-Co-Re compound in Fig. 4b is suggestive of structural transformation. Similar jump was also observed in volume near 50at.% Re, which indicates towards possible phase change. Moreover, the existence of mixed valency in Ce, e.g., trivalent (+3) and tetravalent (+4), can also lead to the phase change in Ce-Fe compounds, for example, smaller Ce³⁺ ion radii compared to Ce⁴⁺ makes it easier to form the CeFe₂ phase [96]. Mixed valence state, and the fact that the ground state structure of CeFe₂ is not cubic, increases the challenge of accounting it into ML models, therefore, for simplicity, not considered.

In Fig. 4b, we show the effect of alloying Rh-4d (Y) on the phase stability of Ce(Re/Co/TM2₁₋₋₂), where Rh-4d increases the alloy phase stability. Experiments also show that CeRh₂ stabilizes in cubic phase [97], which further confirms the robustness of our predictions.

In Fig. 4b, we show the effect of Pd/Pt/Re/Rh alloying on Fe site in CeFe₂. A monotonic change was observed in phase stability with increasing at.% Fe. The Fe shows large solubility range from min(67 at.%) for Ce(Fe,Re)₁₋₋₂ in Fig. 4b, and max (80 at.%) for Ce(Fe/Pd)₁₋₋₂ in Fig. 4b. However, (near)linear or positive slope in formation enthalpy indicates that Fe does not mix and completely phase separates with other alloying elements. Given the stability of CeFe₂ in cubic phase, we can understand phase stability in Fig. 4b in two ways - (i) Ce-Fe-Y may form different stable layered structure, or (ii) Ce-Fe and Ce-Y will phase separate and coexist together. On the other hand, weakly positive E_form near Fe-rich region indicates towards possible thermal stabilization of pure binary CeFe₂ compound. Notably, the ML predicted E_form for CeFe₂ (-0.04 eV-atom⁻¹) shows reasonably good agreement DFT calculations (0.06 eV-atom⁻¹).

The phase stability of Ce-Mn based rare-earth compound alloyed with Pd/Pt/Re/Rh at Mn-site is shown in Fig. 4b-c. The Ce-Mn system is more interesting in the sense that CeMn₂ does not exist. The large positive change in phase stability of Ce(Mn₁/TM₂₁₋₋₂) (TM₂=Pd/Pt/Re/Rh) on increasing Mn concentration in Fig. 4b-c shows a poor solubility of Mn, which is clear signature of phase change beyond 20-40 Mn at.%. The ML predicted phase stability for CeMn₂ (-0.19 eV-atom⁻¹) is in good agreement with DFT (+0.12 eV-atom⁻¹), which further affirms the thermodynamic instability of CeMn₂. Going back to our discussion of energy instability in CeRe₂, this result confirms metastability or dynamically instability of Ce-Re-Mn compounds in MgCu₂ phase.

The alloying effect of Ni with 4d⁵/5d (TM₂=Pd/Pt/Re/Rh) transition metals on the phase stability of Ce(Ni,TM₂₁₋₋₂) compounds was also analyzed in Fig. 4d-d due to special interest towards RE-Ni compounds for their magnetocaloric properties and hydrogen absorption capacity [98, 99]. Experimentally, CeNi₂ is well known to crystallize in cubic Laves phase with lattice parameter of 7.194 Å [100]. Our model in Eq. also predicts strong stability of CeNi₂ with E_form of -0.18 eV-atom⁻¹ as shown in Fig. 4d-d. The predicted stability was found in good agreement with the DFT (-0.24 eV-atom⁻¹) and experimentally observed cubic MgCu₂ crystal structure [100].

In Fig. 4d, the chemical alloying of Ni with Pd in CeNi₂ shows a weak effect on phase stability up to 40 at.% Ni, while a sharp increase in stability was found in Ni-poor region. Model predictions show a very high stability for CePd₂ (-0.29 eV-atom⁻¹), which is in good agreement with the DFT (-0.24 eV-atom⁻¹). Although CePd₂ was never reported experimentally, our calculations suggest its possible existence in MgCu₂ crystal phase. Our model also predict the stability of CePt₂ as shown in Fig. 4d₁, which is in agreement with experimental observation of cubic MgCu₂ crystal structure [101]. However, the near linear trend in E_form suggests that Ni and Pt either mix very weakly or phase-separate.

In Fig. 4b₂, the phase stability analysis of CeRe₂ shows very interesting behavior, where alloying with Re weakens the phase stability of CeNi₂ (from 100-70 at.%Ni or 0-70 at.% Re), while Re-rich (Ni-poor) region shows increased phase stability. A weakly unstable region for 20-52 at.%Ni is also interesting because of transition from energetically ‘stable’ to ‘unstable’ region. The discrete phase stability in Ce-Ni-Re is suggestive of possible phase transformation, which needs more careful exploration both from theory and experiments as binary Ce-Re compounds are yet to be realized.

The Ce-Rh based compounds, on the other hand, are more interesting because of their critical physical behaviors.
such as Kondo-effect and Fermi-surface effects [102]. Therefore, we also assessed the effect of Rh alloying on thermodynamic stability of CeNi$_2$ in Fig. 4d. The model predicts improved stability of CeRh$_2$ (-0.60 eV-atom$^{-1}$) with respect to CeNi$_2$, which is in good agreement with DFT ($E_{\text{form}}$ (CeRh$_2$) = -0.72 eV-atom$^{-1}$). The experimental realization of CeRh$_2$ [103] further affirms the robustness of our prediction. However, the linear trends for Ce-Ni-Rh in $E_{\text{form}}$ again suggests that Ni and Rh may not want to mix or mix very weakly. However, strong energy stability suggests that both the phases, i.e., CeNi$_2$ and CeRh$_2$, may co-exist together.

**Effect of chemical mixing of 3d transition metals on phase stability**

![Graph showing phase stability](image)

Figure 5: (Color Online) (a-c) Chemical alloying of chemically dissimilar atoms on phase stability of cerium based pseudo-binary compounds Ce(Fe$_z$TM$_2$$_{1-z}$)$_2$, where TM$_2$ = Co/Cu/Zn. (a$_1$-c$_1$) Total density of states at stoicheometric compositions, i.e., $z$ (Fe) = 0, 25, 50, 75, and 100 at.%. The solubility range of Co/Cu/Zn at Fe is marked by arrow.

Low-cost rare earths such as Ce are becoming a popular topic for future applied research due to the criticality of currently available high-performance permanent magnets [104]. Therefore, it would be useful to explore the thermodynamic stability of Ce based compounds, which contain earth-abundant elements and may form the base for future hard magnetic materials. While the MgCu$_2$ type structure is isotropic and unlikely to support hard ferromagnetism, it is useful to see how alloying by common 3d metals can create lattice distortions producing uniaxial symmetry. In Fig. 5a-c, we mixed chemically dissimilar 3d transition elements with similar atomic-sizes together to understand the effects of electron doping (change in valence-electron count) on thermodynamic stability of Ce(TM$_1$,TM$_2$$_{1-z}$)$_2$, where TM$_1$ = Fe/Mn; and TM$_2$ = Co, Cu, and Zn. X = Fe was chosen for its robust magnetic behavior, while Mn was chosen for its intriguing and complex magnetic character. In Fig. 5b-c and Fig. 5a-c$_1$, we analyze the effect of chemical mixing of TM$_1$ = Fe/Mn with TM$_2$ = Co/Cu/Zn on phase-stability and electronic-structure of Ce(TM$_1$,TM$_2$$_{1-z}$)$_2$. The CeFe$_2$ is stable in MgCu$_2$-type phase at room temperature, however, in its ground state it shows a rhombohedral distortion [105]. Whereas, no reports were found that show the existence of cubic structure in Mn based Laves phases. Thus, the explored scenarios are largely hypothetical, allowing us to look into thermodynamic stability of experimentally unavailable rare earth phases using machine learning models.

The ML-predicted phase stability in Fig 5a-c indicates a Fe solubility limit of 62/37/77 at.% in Ce(Fe$_z$TM$_2$$_{1-z}$)$_2$, when chemically mixed with Co/Cu/Zn alloying elements, respectively. This indicates that Fe in Ce(Fe$_z$TM$_2$$_{1-z}$)$_2$ is weakly stable (or unstable) in cubic phase beyond 62/37/77 at.%Fe. The mixing of Co/Cu/Zn at Fe is significant as
each of the alloying elements stabilizes the pure CeFe\textsubscript{2} phase as shown in Fig 5a-c. An important characteristic was observed in Fig. 5c, where the formation enthalpy curve is nearly straight line. This can be attributed to chemical inactivity of the nearly half-filled Fe-3\textit{d} (6e\textsuperscript{s}) and completely filled Zn-3\textit{d} (10e\textsuperscript{s}) shells. The predicted trends in stability of the end compounds, i.e., CeFe\textsubscript{2} (DFT=4 meV-atom\textsuperscript{-1}; ML=20 meV-atom\textsuperscript{-1}) and CeTM\textsubscript{2} (TM=Co/Zn; DFT=(\approx-0.200)/-0.250 meV-atom\textsuperscript{-1}; ML=0.152/-0.190 meV-atom\textsuperscript{-1}), compares well with the DFT numbers. For CeCu\textsubscript{2}, the ML model predicts formation enthalpy of -0.135 eV-atom\textsuperscript{-1} in cubic phase, which is thermodynamically less stable than the orthorhombic (-0.175 eV-atom\textsuperscript{-1}) phase. The ML predictions are in agreement with both DFT and experiments.

To shed more light on trends in energy stability, in Fig. 5a-c, a zoomed-region in total DOS near the Fermi-level of Ce(Fe\textsubscript{z}TM\textsubscript{2}1-z)\textsubscript{2} (TM=Co/Cu/Zn) is shown. The stability at low to intermediate (0-50 at.% Fe) concentration in Fig. 5a is attributed to well-structured DOS peaks just below the Fermi-level, which subsequently disappears on further alloying. Furthermore, Co and Fe bands show no overlap, employing weak or no hybridization. Similarly, the Fe solubility range in Fig. 5b drops drastically to near 30 Fe at.% with Cu alloying whereas it increases to 77 at.% for Zn alloying in Fig. 5c. Despite increasing stability due to Fe/Zn alloying, the Fe/Zn DOS shows no hybridization, i.e., Zn-3\textit{d} and Fe-3\textit{d} are energetically far-apart.

The underlying magnetic structure also plays an important role on thermodynamic stability. As pointed out by Eriksson et al. [106] in CeFe\textsubscript{2}, it is the itinerant nature of electrons in the Ce-4\textit{f} and Fe-3\textit{d} states that are coupled anti-ferromagnetically (AFM). The AFM structure stabilizes CeFe\textsubscript{2} in the cubic phase. The hybridization between Ce-4\textit{f} states with Fe-3\textit{d} itinerant electronic systems can be controlled by compositional tuning. To contrast between itinerant and localized electrons, we choose CeCu\textsubscript{2} as an example where the alloy stabilizes in orthorhombic phase, where the Ce moments are anti-ferromagnetically arranged [107]. This is different from CeFe\textsubscript{2}, where Ce and Fe moments are anti-parallel, which suggests that itinerant electrons hybridize differently than localized 4\textit{f} electrons in CeCu\textsubscript{2}. The stability of given phases can also be correlated to the valence-electron count (VEC) of Fe-3\textit{d} (VEC=8) and Cu-3\textit{d} (VEC=11). This electronic dissimilarity between Fe and Cu explains the stability of CeFe\textsubscript{2} in cubic phase and CeCu\textsubscript{2} in orthorhombic phase. The relation of phase stability with electron-count has also been explained by Gschneidner et al. [67] using the idea of an electron-per-atom ratio.

Figure 6: (Color Online) (a, c) Effect of alloying on the phase stability of cerium based pseudo-binary compound Ce(Mn\textsubscript{z}TM\textsubscript{2}1-z)\textsubscript{2} with chemically distinct atoms (with similar atomic-radius), where TM=Co/Cu/Zn. (a-c) Total DOS at stoichiometric compositions, i.e., z (Mn)=0, 25, 50, 75, and 100 at.%. The solubility range of Co/Cu/Zn at Mn is marked by arrow.
Thermodynamic stability of Ce(Mn, TM2)2, where Mn is alloyed with TM2 = Co/Cu/Zn, is shown in Fig 6a-c. The solubility range for Mn is marked by a dashed arrow. We observed that the predicted formation enthalpy remains positive for most Mn-rich composition range. The solubility limit of Mn in Ce(Mn, TM2)2 is significantly reduced (23/25/30 Mn at.%) compared to the Fe-based compounds (62/37/77 at.%), when alloyed with Co/Cu/Zn (TM2). Intuitively, the phase stability predictions in Fig. 6c for Mn mixed with Zn show linear (straight-line) change, which suggests that Zn does not want to mix with Mn. Notably, the reason for this trend can also be understood from individual electronic configurations of Mn, which has half-filled 3d bands (i.e., 5e−'s), while Zn remains chemically-inactive due to its completely filled d-bands (10e−'s in 3d), i.e., alloying of Mn with Zn have weak or no chemical activity. Also, the positive E_form for CeMn2 agrees with experiments, as this alloy is unstable in the MgCu2 crystal structure. On the other hand, the CeFe2 is a well known phase synthesized experimentally in MgCu2 structure, however, the weak stability of CeFe2 both in ML and DFT predictions can be attributed to relatively large density of states peak (in orange at 100 at.%Fe; and cyan 75 at.%Fe in Fig. 5) near the Fermi-level. It is worth mentioning that the low-T and room-temperature structures of CeFe2 are different, and the structural transition is possibly driven by magnetism. On the other hand, in almost all cases with Fe-poor regions, the increased stability is attributed to a pseudo-gap at E_Fermi in the total DOS. Similar behavior was also found in more complex alloy systems [108].

**Experimental validation of the machine-learning predicted phase stability of Ce-Fe-Cu compounds**

In Fig. 7a, we plot DFT calculated phase energy difference (∆E; eV-atom−1) for Ce(FezCu1−z)2 comparing cubic (MgCu2) and orthorhombic (KHg2) phases with respect to the end points (CeFe2 and CeCu2). The CeFe2 is known to stabilize in cubic phase, whereas CeCu2 in orthorhombic phase. The ML predicted phase stability of cubic and orthorhombic phases of CeCu2 is -0.137 eV-atom−1 and -0.175 eV-atom−1, respectively, which is in agreement with the experiments. Similarly, DFT-calculated ∆E_form in Fig. 7a shows that cubic phase is stable at 0at.%Cu (100at.%Fe) while orthorhombic phase is preferable at 100at.%Cu, in agreement with experiments [72, 86]. Phase stability in Fig. 7a shows a crossover from cubic to orthorhombic phase at 80at.%Cu. Importantly, the thermodynamic stability could be controlled using compositional variation at transition metal site in CeTM2, TM = Fe or Cu. A detailed analysis of phase stability was performed through band-structure, charge-density and magnetization density in supplemental Fig. S4& S5.

![Figure 7](image_url)

To validate our predictions, we synthesize the Ce-Fe-Cu compounds at x (Fe) = 25, 50 and 75 at.%. Our XRPD data in Fig. 7a-d further confirms our predictions that CeFe2 and CeCu2 do not form a solid solution rather they energetically prefer to exist as a mixed cubic (CeFe2) and orthorhombic (CeCu2) intermetallic phase. The ratios of CeFe2 and CeCu2 in the prepared alloys, as determined by Rietveld refinement, agree with the nominal Fe/Cu ratios, indicating that in studied compositions the alloying simply form two-phase mixture avoiding chemical interaction. The Cu atomic position in the CeCu2 compound (KHg2 structure) does not take any Fe and the lattice parameters
of this phase do no change significantly (i.e. the change is <0.01 Å). The behavior of CeFe$_2$ is more interesting since both EDX and XRPD refinement indicate that a small amount of Cu (EDS suggests ∼3%) is likely to mix with Fe in the transition metal position. This is in agreement with calculations that predict lower energy with Cu addition. However, the lattice parameter a of prepared alloys, obtained in the range of 7.308±0.003 (z(Fe)=25) to 7.314±0.003 Å (z(Fe)=75), is higher than that reported for pure CeFe$_2$ (typically reported as ∼7.30 Å). This is in odds with the fact that atomic radii of Cu is slightly smaller or nearly equal to that of Fe. One may speculate that Cu presence slightly decreases Ce valence correspondingly increasing its atomic radii. A number of interesting behaviors has already been reported in lightly doped CeFe$_2$ compounds, so, perhaps, a closer look is warranted at the Ce(Fe$_{1-z}$Cu$_z$)$_2$ system, where z<0.03, as well. In regard to this work, the experimental data clearly support the theoretical results of restricted mixing of Fe and Cu atoms in the Ce(Fe$_{1-z}$Cu$_z$)$_2$ system.

Electronic-structure origin of the phase stability of Ce-Fe-Cu compounds and correlation with machine-learning descriptor

The appearance of atomic-features in the $E_{form}$ descriptor in Eq. [1] in the context of phase stability has already been discussed in the feature analysis section. However, it will also be critical to understand the electronic-structure origin of phase stability and its correlation with the $E_{form}$ descriptor. Therefore, we analyzed the DFT calculated band-structure of Ce-Fe-Cu compounds in Fig. 8 & 9. In Fig. 8 we show the spin-polarized band-structure of Ce(Fe$_{1-z}$Cu$_z$)$_2$ at x=0, 50, 100at.%Cu. The CeFe$_2$ and Ce(Fe$_{0.50}$Cu$_{0.50}$)$_2$ band-structures are shown along high-symmetry directions of cubic phase, while the bands-structure of CeCu$_2$ is shown along high-symmetry directions of orthorhombic phase in the energy-range ($E-E_{Fermi}$) -0.1 to 0.1 eV. The highlighted bands in the band-structure are Ce$^{−4f}$ states that are localized near the Fermi-level both in the cubic (Fig. 8a-d) and orthorhombic (Fig. 8e,f) phases. The DFT calculated magnetic moment-contribution of each species at 0 at.%Cu in the cubic phase is (Ce$=-0.68$, Fe$=1.98$) μ$_B$, whereas magnetic moment of individual species changes to (Ce(1)$=0.55$(near Cu), Ce(2)$=0.61$ (near Fe); Fe (1)$=-1.61$, Fe(2)$=-1.42; Cu$=0.03$) μ$_B$ at 50 at.%Cu. The change in magnetic behavior shows the effect of change in neighboring environment of Ce and Fe sites on alloying with Cu. Notably, we found that both Ce and Fe moments drop with increasing at.%Cu, and the moments drop to zero in the cubic phase where Fe$=Cu$, i.e., Ce$=0$; Cu$=0$ μ$_B$. The reduced magnetic character is directly correlated with the decreased thermodynamic stability of the Ce-Fe-Cu compound. Moreover, the magnetic moment at the Ce/Fe sites in the cubic phase were flipped by adding Cu at the Fe-site in CeFe$_2$. The increasing Cu concentration increases the cell volume in the cubic phase, which was in agreement with the experimental data shown above and found to be responsible for decrease in both magnetic character and phase stability of Ce-Fe-Cu alloy. As can be seen from the band-structure in Fig. 8e,f, the magnetic character changes from strong-ferromagnetic (CeFe$_2$) to non-magnetic (CeCu$_2$) in the cubic phase.

![Figure 8](image.png)

*Figure 8: (Color Online) The spin-polarized band-structure of (a,b) CeFe$_2$, (c,d) Ce(Fe$_{0.50}$Cu$_{0.50}$)$_2$, and (e,f) CeCu$_2$ in cubic phase of Ce-Fe-Cu based rare-earths. The top- and bottom-panels represent the up-spin and down-spin, respectively.*

The energetically stable orthorhombic phase at 100at.%Cu in Fig. 7 has a finite Ce$^{−4f}$ moment of 0.60 μ$_B$, however, Ce$^{−5d}$ has lost its magnetic behavior in absence of Fe$^{−3d}$ states (see comparative plot of cubic and orthorhombic...
phases at 100 at.% in supplement Fig. S2). The decrease in phase stability of the cubic phase in Fig. 7 strongly correlates with increasing at.% Cu, as shown in the total DOS plot in Fig. 5b. The electronic density at $E_{\text{Fermi}}$ was found to increase with increasing at.% Cu, which is evident from the presence of a ‘spaghetti’ configuration of bands in the spin-polarized band-structure plot at 50 at.% Cu in Fig. 8c,d. The behavior of Ce–4f states changes from wavy (represent more hybridization) to much flatter representing reduced hybridization between Ce–Cu due to filled Cu–3d states at 50 at.% Cu.

Microscopically, the delocalization degree of valence electrons can significantly impact the magnetic behavior. Therefore, the loss of magnetic character in Cu doped CeFe$_2$ could also be connected to the difference in electronegativity and atomic-radii of Fe ($\chi_1=1.80$, $r_1=1.40$ Å) with respect to Cu ($\chi_2=1.85$, $r_2=1.35$ Å), especially since higher electronegativity correlates with lower radii due to stronger attraction between the nucleus and valence electrons. Also, Li et al. [109] has shown that electronegativity and atomic-size difference play an important role in influencing the magnetic behavior by modifying attractive interaction between valence electrons of atoms. Looking back to Eq. 1 and correlating that to electronic-structure analysis, the appearance of electronegativity and atomic-radii in $E_{\text{form}}$ descriptor becomes more understandable. While training or cross-validation of the SISSO model was provided with no biases, still the final $E_{\text{form}}$ descriptor picks important physical quantities such as atomic-radii, electronegativity which directly connects to phase stability. This discussion is also consistent with our electronic-structure analysis, which further signifies the importance of using the physics-based descriptors for materials design.

We show the (001) projected Fermi-Surface in the cubic phase of Ce(Fe$_z$Cu$_{1-z}$)$_2$ at 0, 50, and 100 at.% Cu in Fig. 9a-f. Our aim was to understand stability in terms of change in complexity of the Fermi-surface. Is there any relation? In Fig. 9a,b, the (001) Fermi-surface cross-section of CeFe$_2$ shows hole-pockets at $\Gamma$ and K points, and electron-pockets at the X-point in both up-and down-spin channels. Down-spin-channel has multiplet of hole-pockets mainly arising from Fe-3d states while hole pockets in the up-spin channel are mainly from Ce–4f. A similar characteristic was found in band plot in Fig. 8a,b. We found that Cu alloying drastically changed the band-structure in Fig. 8c,d, therefore, it would be interesting to see how Cu affects the Fermi-surface. The (001) Fermi-surface cross-section is shown in Fig. 9c,d, where two hole-pockets were found in the up-spin channel while four hole-pockets in the down-spin channel at $\Gamma$-point. The hole-pocket got much bigger at X-point in down-spin channel in Fig. 9d, while hole pocket in up-spin channel at X-point disappears. This characteristic was seen in band-structure due to change in magnetization direction CeFe$_2$ with Cu alloying. In Fig. 9e,f, CeCu$_2$ becomes non-magnetic where X and $\Gamma$ show much smaller hole-pockets (one each) in both up-and down-spin channels. On the other hand, a Fermi-surface cross-section along (001) shows a slightly enlarged electron-pocket at K-point. This is in agreement with the fact that Cu alloying adds excess electrons. The reduction of hole-pockets on Cu alloying indicates filling of all bonding...
states, which drastically reduces the hybridization of Ce with transition metals and reduces stability of Ce-Fe-Cu pseudo-binary alloys in cubic phase.

**Generalizability of the SISSO based analytical models:** Regarding generalizability of the model and its ability to go beyond known chemistries, we want to emphasize that none of the predicted disordered cases in Fig.4-6 were included in the training or test data. Our calculations show that analytical models are useful for predictive design of disordered alloys. Furthermore, *ab-initio* methods struggle with this because the design of non-stoichiometric disordered supercells is non-intuitive and computationally expensive. In the case where one of the end points is known to not exist, a solubility limit can still be obtained (e.g., solubility limit of Fe in Cu if CeFe$_2$ did not form). Our model predictions for Ce-Fe-Cu shows that Fe-Cu do not want to mix, which is independently verified by DFT calculations and experiments in Fig.7. These findings are promising as they tentatively suggest the model is capable of making safe extrapolations.

We also want to point out that SISSO may not be the only or even best modeling framework that could be used to carry out the proposed task. We note that a priori, it is not always immediately clear when a model is sufficiently simple and yet expressive enough to carry out robust predictions. We highlight, for example, the work by Bartel et al. on the use of SISSO to predict finite temperature corrections to the Gibbs energies of ordered compounds: at the end, the SISSO model is relatively simple (a posteriori), but by carrying out the SISSO methodology the authors were able to explore (order) billion-dimensional feature spaces, arriving at the conclusion that the best model indeed was highly parsimonious, and yet, highly predictive [31].

**Conclusion**

Physical descriptor-based analytical machine-learning models are very useful for accelerated identification of new inorganic compounds with varying chemistry. However, the absence of sufficient density-functional theory or experiment generated entries into the rare-earth databases makes the execution of high-quality machine learning extremely difficult, given the fact that generating accurate DFT or experimental database for rare-earths is very tricky as well as time and resource consuming. In this work, we generated nearly 600+ data points for REX$_2$ type rare-earth compounds. The SISSO models were trained over a sparse ‘in-house’ database that provides three dimensional (3D) analytical descriptor. We demonstrated that computationally inexpensive descriptors can be incorporated to predict the equilibrium phases of rare-earth based compounds purely on a thermodynamics basis, i.e., using formation enthalpies. We present a detailed phase stability analysis of Ce based rare-earth compounds with cubic Laves phase and provide a quantitative guidance for compositional considerations in realizing new metastable materials with unconventional chemistries. We validated our ML and DFT predictions of phase stability through very careful synthesis of Ce(Fe$_{z}$Cu$_{1-z}$)$_2$ compounds with 25, 50 and 75 at.%Fe(z). Our experiments show that Fe and Cu do not form solid solution rather they stabilize as mixed intermetallic phases, i.e., CeFe$_2$ and CeCu$_2$, which is in good agreement with model predictions. The low computational cost and reasonable accuracy of interpretable machine-learning models presented in this work will enable the high-throughput analysis of all possible crystal phases and remove external biases in phase selection. Despite the successful use of limited database and predicting the stability of rare-earth compounds and mixed compositions, there are a number of other challenges that need to be addressed, for example, (i) temperature, and (ii) pressure dependence of ground state structure and its impact on thermodynamic stability. While extracting the complete picture using standard DFT calculations would be prohibitively challenging, we believe that ML methods described here will enable such evaluations and bring the science closer to the desired "materials from PC" land. Our approach will be useful in discovering new and complex rare-earth compounds with new functionalities.

Furthermore, more insights are probably needed into the extrapolation of the SISSO descriptors and the chosen featurization for the complete generalization of descriptors beyond training and test data. However, this requires a separate, yet a detailed investigation to ascertain the ability of this model to extrapolate, with reasonable accuracy, beyond rare-earth and transition metal based alloying chemistries. Notably, simple predictive approaches, such as a second-order polynomial, require unknown coefficients with no physical meaning would still require the utilization of either regression or other machine learning approaches. This suggests that there is no good way to get around the machine learning aspect. Notably, SISSO provides an interpretable analytical model that makes its simpler to use for predictive design, while other regression techniques are mostly a black box. At the same time, there is clearly a need
for more insightful investigation into the extrapolation of the SISSO descriptors for arbitrary alloying chemistries as machine-learning are primarily data sensitive.

Acknowledgements

P.S. and Y.M. acknowledge the insightful discussions with Dr. Vitalij K. Pecharsky and Dr. Duane D. Johnson of Ames Laboratory at Iowa State University. This work was supported by Laboratory Directed Research and Development Program (LDRD) program at Ames Laboratory. Work at Ames Laboratory was supported by the U.S. Department of Energy (DOE), Office of Science, Basic Energy Sciences, Materials Science & Engineering Division. Ames Laboratory operated by Iowa State University for the U.S. DOE under contract DE-AC02-07CH11358. GV and RA acknowledge the support of QNRF under Project No. NPRP11S-1203-170056. DFT calculations were carried out at the Texas A&M High-Performance Research Computing (HPRC) Facility.

Data availability

The authors declare that the data supporting the findings of this study are available within the paper and supplement. Also, the data that support the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request.

References

References

[1] A. R. Jha, “Rare earth materials: properties and application,” CRC Press (Taylor & Francis), U.K. (2014).
[2] H. Kim, S. Alam, N. R. Neelameggham, H. Oosterhof, T. Ouchi, X. Guan, “Rare earth metal technology,” Springer (2017).
[3] D. D. Imholte, R. T. Nguyen, A. Vedantam, M. Brown, A. Iyer, B. J. Smith, J. W. Collins, C. G. Anderson, B. O’Kelley, “An assessment of U.S. rare earth availability for supporting U.S. wind energy growth targets,” Energy policy 113, 294 (2018).
[4] J. Li, K. Peng, P. Wang, N. Zhang, K. Feng, D. Guan, J. Meng, W. Wei, and Q. Yang, “Critical rare-earth elements mismatch global wind-power ambitions,” One Earth, 3, 116 (2020)
[5] T. Sakai, M. Matsuoka and C. Iwakura, “Rare earth intermetallics for metal-hydrogen batteries,” in: K.A. Gschneidner, Jr. and L.R. Eyring, Eds., Handbook on the Physics and Chemistry of Rare Earths (Elsevier, Amsterdam, 1995), 21, 133.
[6] E. Alonso, T. Wallington, A. Sherman, M. Everson, F. Field, R. Roth, and R. Kirchain, “An assessment of the rare earth element content of conventional and electric vehicles,” SAE Intern. J. Mater. Manufactur. 5, 473-477 (2012).
[7] S.V. Eliseeva, J.-C. Büntzli, “Rare earths: jewels for functional materials of the future,” New J. Chem. 35, 1165 (2011).
[8] L.D. Carlos, R.A.S. Ferreira, V. Z. Bermudes, B. Julian-Lopez and P. Escribano, “Progress on lanthanide-based organic-inorganic hybrid phosphors,” Chem. Soc. Rev. 40, 536 (2011).
[9] E.T.C. Vogt and B.M. Weckhuysen, “Fluid catalytic cracking: recent developments on the grand old lady of zeolite catalysis,” Chem. Soc. Rev. 44, 7342 (2015).
[10] S. Dey and G.C. Dhal, “Cerium catalysts applications in carbon monoxide oxidations,” Mat. Sci. Energy Techn. 3, 6 (2020).
[11] E. Snitzer, “Rare earth fiber lasers,” J. Less-Comm. Met. 148, 45 (1989).
[12] J. G. Büntzli and S. V. Eliseeva, “Lanthanide NIR luminescence for telecommunications, bioanalyses and solar energy conversion,” J. Rare Earths 28, 824 (2010).
[13] R. K. Jyothi, T. Thenpalli, J. W. Ahn, P. K. Parthi, K. W. Chung, and J-Y Lee, “Review of rare earth elements recovery from secondary resources for clean energy technologies: Grand opportunities to create wealth from waste,” J. Cleaner Production 187, 7342 (2018).
[14] Y. Madryk, V.K. Pecharsky, Materials for solid state cooling, in: R. Pottgen, T. Justel, and C.A. Strassert, Eds., Rare Earth Chemistry, De Gruyter (2020) pp. 487-510
[15] X. Moya, and N. D. Mathur, “Caloric materials for cooling and heating,” Science 370, 797 (2020).
[16] L. Park, and S. Jeong, “Development of the active magnetic regenerative refrigerator operating between 77 K and 20 K with the conduction cooled high temperature superconducting magnet”, Cryogenics 88, 106-115 (2017).
[17] K.A. Gschneidner, Jr., H. Takeya, J.O. Moorman, and V.K. Pecharsky, “(Dy0.5Er0.5)Al2: A large magnetocaloric effect material for low temperature magnetic refrigeration”, Appl. Phys. Lett. 64, 253 (1994).
[18] T. Numazawa, K. Kamiya, T. Uaki, and K. Matsumoto, “Magnetic refrigerator for hydrogen liquefaction”, Cryogenics 62, 185-192 (2014).
[19] P. B. de Castro et al., “Machine-learning-guided discovery of the gigantic magnetocaloric effect in HoB2 near the hydrogen liquefaction temperature”, NPG Asia Materials 12, 35 (2020).
[20] G. Hoffer and K. Strnat, “Magnetocrystalline anisotropy of YCo5 and Y2Co17,” IEEE Trans. Magn. 2, 487 (1962).
[21] K. Strnat, G. Hoffer, J. Olson, W. Ostertag, and J.J. Becker, “A family of new cobalt-base permanent magnet materials,” J. Appl. Phys. 38, 1001 (1967).
[22] J.J. Croat, J.F. Herbst, R.W. Lee, and F.E. Pinkerton, “Pr-Fe and Nd-Fe-based materials - a new class of high-performance permanent magnets,” J. Appl. Phys. 55, 2078 (1984).
[62] L. Pauling, “The Nature of the Chemical Bond”, Cornell Univ. Press, Ithaca, New York, 1960.

[63] M. Rahm, R. Hoffmann, and N.W. Ashcroft, “Atomic and Ionic Radii of Elements 1-96”, Chem. Eur. J. 22, 14625-14632 (2016).

[64] L.C. Allen, “Electronegativity Is the Average One-Electron Energy of the Valence-Shell Electrons in Ground-State Free Atoms”, J. Am. Chem. Soc. 111, 9093-9014 (1989).

[65] J. Schmidt, M.R.G. Marques, S. Botti, et al. “Recent advances and applications of machine learning in solid-state materials science,” npj Comput Mater 5, 83 (2019).

[66] R.E. Bellman, “Adaptive Control Processes: A Guided Tour,” (Princeton University Press, Princeton, NJ, 2015).

[67] K.A. Gschneider, Jr., and V.K. Pecharsky, “Binary rare earth Laves phases – an overview”, Z. Kristallogr. 221, 375-381 (2006).

[68] Y. Nishihara and Y.Yamaguchi, “Magnetic Phase Transitions in Itinerant Electron Magnets (Hf1−x,Inx)Fe2”, J. Phys. Soc. Jpn. 52, 3630 (1983).

[69] Y. Nishihara, and Y. Yamaguchi, “Coexistence of Ferro- and Antiferromagnetism in the Itinerant Electron System (Sc1−x,Inx)Fe2”, J. Phys. Soc. Jpn. 53, 2201 (1984).

[70] Y. Yamada, Y. Kitaoaka, K. Asayama, and A. Sakata, “NMR Study of Coexistence of Ferro- and Antiferromagnetism in the Itinerant Electron System (Zr1−x,Nb)xFe2”, J. Phys. Soc. Jpn. 53, 3198 (1984).

[71] G. Hilscher, “Onset of coexistence in concentrated ternary alloys II: Laves phase compounds A(Fe1−x,B)x2 (A = Y, Zr; B = Mn, Co and Al)”, J. Magn. Magn. Mater. 27, 1 (1982).

[72] S.B. Roy, G.K. Perkins, M.K. Chattopadhyay, A.K. Nigam, K.J.S. Sokhey, P. Chaddah, A.D. Caplin, and L.F. Cohen. “First Order Magnetic Transition in Doped CeFe3 Alloys: Phase Coexistence and Metastability”, Phys. Rev. Lett. 92, 147203 (2004).

[73] C.K. Barman, P. Singh, D.D. Johnson, and A. Alam, “Revealing the Nature of Antiferroquadrupolar Ordering in Cerium Hexaboride: CeB6”, Phys. Rev. Lett. 122, 076401 (2019).

[74] M.T. Johnson, P.J.H. Bloemen, F.J. a Den Broeder, and J.J. De Vries, “Magnetic anisotropy in metallic multilayers,” Reports Prog. Phys. 26, 121 (1996).

[75] Y. Yamada, Y. Nakazawa, and M. Ishikawa, “Gap Formation in a Valence Fluctuation System of CeNiSn”, Jpn. J. Appl. Phys. 30, 547 (1991).

[76] T. Schmidt, D. Johrendt, C.P. Sebastian, R. Pöttgen, K. Latka, and R. Knies, “Structure, Chemical Bonding and 119Sn Mössbauer Spectroscopy of LaRhSn and CeRhSn”, Z Naturforsch B Chem Sci 60b, 1036 (2005).

[77] B. Chevalier, C.P. Sebastian, and R. Pöttgen, “Hydrogenation of the intermediate valence ternary stannides CeRhSn and CeSn3”, Solid State Sci 8, 1000 (2006).

[78] M. Lenkewitz, S. Corsepius, and G.R. Stewart, “Specific heat and magnetic susceptibility of CeAuSn and CeAgSn”, J. Alloys Compd 241, 174 (1997).

[79] M.T. Johnson, P.J.H. Bloemen, F.J. a Den Broeder, and J.J. De Vries, “Magnetic anisotropy in metallic multilayers,” Reports Prog. Phys. 59, 1409 (1996).

[80] T. Ueno, J. Sinha, N. Inami, Y. Takeiichi, S. Mitani, K. Oto, and M. Hayashi, “Enhanced orbital magnetic moments in magnetic heterostructures with interface perpendicular magnetic anisotropy,” Sci. Rep. 5, 14858 (2015).

[81] M. Bode, M. Heide, K. von Bergmann, P. Ferriani, S. Heinze, G. Bihlmayer, A. Kubetzka, O. Pietzsch, S. Blugel, and R. Wiesendanger, “Chiral magnetic order at surfaces driven by inversion asymmetry,” Nature 447, 190 (2007).

[82] L.M. Miron, K. Garello, G. Gaudin, P.-J. Zermatten, M. V Costache, S. Auffret, S. Bandiera, B. Rodmacq, A. Schuhl, and P. Gambardella, “Perpendicular switching of a single ferromagnetic layer induced by in-plane current injection,” Nature 476, 189 (2011).

[83] G. Yu, P. Upadhyaya, Y. Fan, J.G. Alzate, W. Jiang, K.L. Wong, S. Takei, S.A. Bender, L-T. Chang, Y. Jiang, M. Lang, J. Tang, Y. Wang, Y. Tserkovnyak, P.K. Amiri, and K.L. Wang, “Switching of perpendicular magnetization by spin-orbit torques in the absence of external magnetic fields,” Nat. Nanotechnol. 9, 548 (2014).

[84] S.F. Matar, “Review on cerium intermetallic compounds: A bird’s eye outlook through DFT”, Progress in Solid State Chemistry 41, 55-85 (2013).

[85] B. Johansson, “The transition in cerium is a Mott Transition,” Philos. Mag. 30, 469-482 (1973).

[86] E. King, J. A. Lee, I. R. Harris and T. F. Smith, “New Phase Boundary for Cerium,” Phys. Rev. B 1, 1380 (1970).

[87] G. Seyfarth, A. S. Ruetschi, K. Sengupta, A. Georges, D. Jaccard, S. Watanabe and K. Miyake, “Heavy fermion superconductor CeCu2Si2 under high pressure: Multithermal the valence crossover,” Phys. Rev. B 85, 205105 (2012).

[88] H.L. Luo, M.B. Maple, I.K. Harris, and T.F. Smith, “Superconductivity of the phase CeCo2 at zero and high pressure,” Phys. Lett. A 27, 519-520 (1968).

[89] Y. Omuki, and T. Komatsubara, “Heavy fermion state CeCu6,” J. Magn. Magn. Mater. 63, 281-288 (1987).

[90] M. B. Tang, H. Y. Bai, W. H. Wang, D. Bogdanov, K. Winzer, K. Samwer, and T. Egami, “Heavy-fermion behavior in cerium-based metallic glasses,” Phys. Rev. B 75, 172202 (2007).

[91] S. V. Streitsov, E. Gull, A. O. Shirikov, M. Troyer, V. I. Anisimov, and P. Werner, “Magnetic susceptibility of cerium: An LDA+DMFT study,” Phys. Rev. B 85, 195109 (2012).

[92] F.M. Mathur, S.R. Julian, I.R. Walker, D.M. Freye, R.K.W. Haselwimmer, and G.G. Lonzarich, “Magnetically mediated superconductivity in heavy fermion compounds,” Nature 394, 39-43 (1998).

[93] A. P. Murani, S. J. Levett and J. W. Taylor, “Magnetic Form Factor of Constrained Cerium: Towards Understanding the Magnetism of Cerium,” Phys. Rev. Lett. 95, 256403 (2005).

[94] X. Sun, Y. Lei, R. Zhou, B. Qu, D. Li, B. Zhang, and X.C. Zeng, “New phases of 3d-transition metal-cerium binary compounds: an extensive structural search,” RSC Adv. 7, 40486-40498 (2017).

[95] M. I. Ivanov, V. V. Berezutskii, M. O. Shevchenko, et al., “Interaction in europium-containing alloy systems,” Dop. NAN Ukrainy, No. 8, 90-99 (1993).

[96] B. Gorr, V. Trindade, S. Burk, H.-J. Christ, M. Klauke, D. Mukherji, and J. Rosler, “Oxidation Behaviour of Model Cobalt-Rhenium Alloys During Short-Term Exposure to Laboratory Air at Elevated Temperature”, Oxidation of Metals volume 39, 40486-40498 (2011).

[97] S. V. Streitsov, E. Gull, A. O. Shirikov, M. Troyer, V. I. Anisimov, and P. Werner, “Magnetic susceptibility of cerium: An LDA+DMFT study,” Phys. Rev. B 85, 195109 (2012).

[98] F.M. Mathur, S.R. Julian, I.R. Walker, D.M. Freye, R.K.W. Haselwimmer, and G.G. Lonzarich, “Magnetically mediated superconductivity in heavy fermion compounds,” Nature 394, 39-43 (1998).

[99] A. P. Murani, S. J. Levett and J. W. Taylor, “Magnetic Form Factor of Constrained Cerium: Towards Understanding the Magnetism of Cerium,” Phys. Rev. Lett. 95, 256403 (2005).

[100] X. Sun, Y. Lei, R. Zhou, B. Qu, D. Li, B. Zhang, and X.C. Zeng, “New phases of 3d-transition metal-cerium binary compounds: an extensive structural search,” RSC Adv. 7, 40486-40498 (2017).

[101] M. I. Ivanov, V. V. Berezutskii, M. O. Shevchenko, et al., “Interaction in europium-containing alloy systems,” Dop. NAN Ukrainy, No. 8, 90-99 (1993).
[98] H. Oesterreicher, J. Clinton and H. Bittner, “Hydrides of La-Ni compounds,” Mater. Res. Bull. 11, 1241 (1976).
[99] K.H. Buschow, and H.H. Van Mal, “Phase relations and hydrogen absorption in the lanthanum-nickel system,” J. Less-Common Met., 29, 203 (1972).
[100] W. E. Wallace, T. V. Volkmann and R. S. Craig, “Magnetic Characteristics of Tb-Y and Ho-Y Solid Solution,” J. Phys. Chem. Solids, 31, 2185 (1970).
[101] P. Villars, and K. Cenzual, “CePr2.00 (CePr2, T = 13 K) Crystal Structure: Datasheet from “PAULING FILE Multinaries Edition – 2012” (Springer-Verlag Berlin Heidelberg).
[102] M. Higuchi, H. Yamagami, and A. Hasegawa, “Fermi Surface of CeRh2,” J. Phys. Soc. Jpn. 63, 3014-3025 (1994).
[103] H. Sugawara, T. Yamazaki, J. Itoh, M. Takashita, T. Ebihara, N. Kimura, P. Svoboda, R. Settai, Y. Omuki, H. Sato, S. Uji, and H. Aoki, “Single Crystal Growth and Electrical Properties of CeRh2 and CeIr2,” J. Phys. Soc. Jpn. 63, 1502-1507 (1994).
[104] X.Q. Zhou, S.Y. Liu, X.K. Lv, “Effects of La substitution on the microstructures and properties of sintered NdFeB magnets,” J. Magn. Mater. Dev. 32 (17) (2013).
[105] A. Haldar, K.G. Suresh, A.K. Nigam, “Martensitic features in Si doped CeFe2 revealed by magnetization and transport study”, Intermetallics 18, 1772-1778 (2010).
[106] O. Eriksson, L. Nordström, M. S. S. Brooks, and B. Johansson, “4f-Band Magnetism in CeFe2”, Phys. Rev. Lett. 60, 2523 (1988).
[107] V. Nunez, R. Trump, P.J. Brownt, T. Chattopadhyay, M. Loewenhaupt, and F. Tassett, “Antiferromagnetism of the Kondo lattice compound CeCu, studied by neutron polarimetry”, J. Phys.: Condens. Matter. 4, 1115-1120 (1992).
[108] P. Singh, A.V. Smirnov, and D.D.Johnson, “Ta-Nb-Mo-W refractory high-entropy alloys: Anomalous ordering behavior and its intriguing electronic origin,” Phys. Rev. Mater. 2, 055004 (2018).
[109] K. Li, C. Kang, D. Xue, “Electronegativity calculation of bulk modulus and band gap of ternary ZnO-based alloys”, Mater. Res. Bull. 47, 2902 (2012).