



TECHNICAL ARTICLE

Machine Learning-Guided Exploration of Glass-Forming Ability in Multicomponent Alloys

YI YAO,¹ TIMOTHY SULLIVAN IV,¹ FENG YAN,¹ JIAQI GONG,²
and LIN LI^{1,3} 

1.—Department of Metallurgical and Materials Engineering, The University of Alabama, Tuscaloosa, AL 35487, USA. 2.—Department of Computer Science, The University of Alabama, Tuscaloosa, AL 35487, USA. 3.—e-mail: lin.li@eng.ua.edu

The prediction of glass-forming ability (GFA) in alloy systems is a challenging problem in material science as well as for metallurgical applications. In this study, we build artificial neural network (ANN) models to investigate the GFA of multicomponent alloys, based on the datasets assembled from ternary alloys as well as quinary alloys prepared by magnetron sputtering. Through training the ANN models with different combinations of datasets, we tackle the problem of the influence of the data source on the model performance, especially the generalizability of the models in predicting the GFA in unseen multicomponent alloy systems. The ANN model trained on a combined dataset exhibits the best performance, specifically low root mean square error in leave-one-alloy-system-out validation and high model robustness, for several CoCrFeNi-based multicomponent alloys. To further verify the ANN models, we synthesize CoCrFeNi-Mo metallic thin films by magnetron co-sputtering and characterize the structure and phase information via x-ray diffraction and electron microscopy. The outcomes of our experiments agree reasonably well with the ANN model predictions, indicating that the data-driven machine learning approach can be a useful tool in the future design of multicomponent amorphous alloys.

INTRODUCTION

Metallic glasses (MGs) possess high strength, large elastic limit, and superb corrosion resistance, due to the disordered atomic structures, and thus exhibit great potential for advanced structural and functional applications.^{1–4} However, MGs have a very limited glass-forming ability (GFA) compared to conventional glassy systems (e.g., oxide glasses), significantly hindering their development.^{3,5} Therefore, it is of crucial importance to predict GFA and to identify superior glass-forming alloy systems. Thermodynamically, to form a glass crystallization must be bypassed, and a good GFA implies a small driving force of crystallization.⁶ Accordingly, several thermodynamic models have been developed to

predict the GFA based on the experimental observations and measurements of thermodynamic and physical quantities.^{7–11} Inoue summarized an empirical rule to identify high GFA alloys. He has proposed that an alloy with high GFA should have more than three elements, negative heats of mixing (ΔH_{mix}) of the liquid phases, and a large atomic size difference (δ) (above 12%).⁸ Additionally, based on the nature of glass formation along with metallurgical considerations, it has been found that the liquidus temperature and supercooled liquid region can reflect the stability of glass-forming liquids and crystallization resistance during glass formation.^{6,9,12} For instance, Du designed a parameter $\gamma_m = (2T_x - T_g)/T_l$, where T_x is the onset crystallization temperature, T_g is the glass transition temperature, and T_l is the liquidus temperature.⁹ The parameter, γ_m , has shown a good correlation with the GFA of 39 different MGs, with a statistical correlation factor of 0.931. From a structural perspective, the enhanced atomic size mismatch can

Yi Yao and Timothy Sullivan IV contributed equally to this work.

(Received May 15, 2022; accepted September 22, 2022)

induce large lattice distortion and instability, favoring the glassy phase.^{7,8} The presence of short- to medium-range orders leads to geometric frustration in glass formation, which can also stabilize the disordered structures.¹³ To date, identifying the key factors for GFA is still an open research topic, and the non-equilibrium nature of glass-forming processes may give rise to complex relationships beyond the current understanding. Recently, the rapid development of multicomponent alloys has enabled the MG search in a much broader compositional space.¹⁴ The vast space provides opportunities to find glass-forming candidates, but is difficult to explore by the traditional trial-and-error method.

Data-driven machine learning (ML) technologies have emerged as a new tool to investigate the structure and property relationships in MGs.^{15–38} Learning from the existing data, ML models can predict the properties of unknown alloys, rapidly screening candidates in a vast compositional space, and potentially replacing the trial-and-error method.³⁹ In particular, a variety of ML algorithms have been applied to correlate the glass structure and chemistry with elasticity,²⁹ the propensity for plastic deformation,^{16,17,23,33} and thermal and physical properties.^{18–22,24–32} In regard to GFA, ML models demonstrate success in predicting the glass transition temperature,^{35,36} the critical cooling rate,³⁷ and the critical casting diameter of MGs,^{15,21,26,28} and further identify new glass-forming systems,^{22,31} using random forest,^{25,32} support vector machine (SVM),²⁸ and neural network^{18,31,34} algorithms. Until now, the ML studies of GFA have been mainly focused on binary and ternary alloy systems, and further efforts are required to explore multicomponent alloys using ML algorithms. The exploration of multicomponent alloys in the compositional space features a large variety in selecting chemical elements, and a vast combination of these chemical elements by tuning the composition. The existing datasets of binary and ternary alloys cover a large set of chemical elements to form MGs; whereas the increasing datasets of multicomponent alloys provide refined GFA information for a handful of quaternary or quinary alloy systems. All the data sources are of importance in providing valuable information in training ML models, but how the data sources influence the ML model performance of multicomponent alloys is still in question.

In this study, we build artificial neural network (ANN) models to investigate the GFA of multicomponent alloys based on the datasets assembled from ternary alloys, as well as quinary alloys prepared by magnetron sputtering. Through training the ANN models with different combinations of datasets, we tackle the problem of the influence of the data source on the model performance, especially the generalizability of the models in predicting the GFA in unseen multicomponent alloy systems. The ANN model trained on a combined dataset exhibits the best performance, specifically low root mean square

error in leave-one-alloy-system-out validation and high model robustness, for several CoCrFeNi-based multicomponent alloys. To further verify our ML models, we synthesized CoCrFeNi-Mo metallic thin films by magnetron co-sputtering and have characterized the structural and phase information via x-ray diffraction and electron microscopy. The outcomes of our experiments agree reasonably well with the ANN model predictions, indicating that the data-driven ML approach can be a useful tool in the future design of multicomponent MGs.

METHODS

Datasets

Our data were collected from the following three sources: (1) the handbook *Nonequilibrium Phase Diagrams of Ternary Amorphous Alloys*,⁴⁰ (2) Ren and Ward’s high-throughput sputtering experiments for ternary alloys,³² and (3) recent high-throughput sputtering experiments for quinary alloys by Kube et al.⁴¹ In total, the dataset consisted of 7114 alloy compositions in 44 alloy systems. As detailed in Fig. 1 and Table I, 5421 (76.2%) of these alloys were ternary alloys, which mix 3 elements out of 20 different elements (see Fig. 2). The data came primarily from the handbook and Ren and Ward’s high-throughput sputtering experiments.^{32,40,42} The remaining 1693 alloys (23.8%) were quinary, mixing 5 elements out of 7 different elements (see Fig. 2).⁴³ Refer to online supplementary material for the distributions of elements and alloy systems in the dataset, which are provided in supplementary Figs. S14 and S15. To train the ML model, each data point was labeled, based on its structures: amorphous, indicating the alloy could form MG during the sputtering, or crystal, indicating it could

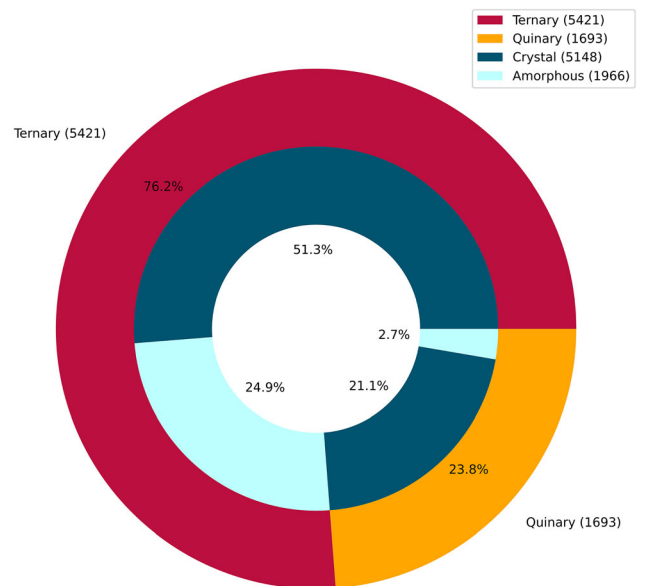


Fig. 1. Distribution of amorphous (light blue) and crystalline (dark blue) phases in the Ternary (red) and Quinary (orange) datasets (Color figure online).

Table I. The information of the datasets

	Data points	Elements	Alloy systems	Amorphous to crystal ratio
Ternary	5421	20	33	1:2.06
Quinary	1693	7	11	1:7.82
Combined	7114	20	44	1:2.62

Fig. 2. The elements in the datasets. The elements denoted in blue only exist in the Ternary dataset, and the elements denoted in red exist in both Ternary and Quinary datasets (Color figure online).

not. Among the collected data, 1966 alloys (27.6%) were identified to have an amorphous structure and 5148 (72.4%) had a crystal structure. Specifically, it was observed that the distribution of the crystal and amorphous structures was imbalanced in the quinary dataset, where 88.7% was crystal and only 11.3% was amorphous. Therefore, we performed a random undersampling test with the synthetic training data. The result showed no discrepancy in performance between the balanced and raw datasets (see supplementary Figs. S13 and S12). Therefore, we accepted this slightly skewed dataset to reduce the risk of information loss from resampling. Notably, the ternary alloy dataset was rich in chemical information, consisting of 20 chemical elements; while the quinary dataset was rich in physical information, with phase and structural properties resulting from 5 element interactions. To explore the influence of the data source on the model performance, the ML models were trained on the “Ternary”, “Quinary”, and “Combined” datasets. The “Combined” dataset represented the combination of the ternary and quinary alloy datasets.

Data Featurization

Thirteen features, as listed in Table II, have been selected to describe the alloys and serve as the input of the ML models. According to the existing studies on the GFA of MGs, these features captured the basic physical, and/or chemical properties of alloys (e.g., valance electron concentration, Pauling’s electronegativity), which distinguish them from each other and were able to relate the elemental properties to the alloy performance. In addition, average atomic size, atomic size difference, average heat of mixing, and standard deviation of the heat of

mixing were chosen as features. These features were used to capture the elemental, thermal, and elastic properties of the alloys considered relevant for glass formation.^{3,7-9,11}

The probability density distributions of the 13 features for the three datasets are displayed in Fig. 3, where the dashed and solid lines represent the crystal (CR) and amorphous (AM) structures, respectively. These plots provide direct visualization of the separation between the crystal and amorphous data in one dimension for a selected feature that is embedded in a higher-dimensional feature space. The blue, green, and red lines represent the Ternary, Quinary, and Combined datasets, respectively. Notably, the distribution of feature values for the Quinary dataset is narrow and concentrated in a small range. This can be attributed to the fact that only 7 different elements exist in the Quinary dataset, compared to 20 in the Ternary dataset. Moreover, the features that measure the standard deviations of the properties (e.g., σ_{VEC} , σ_{PE}) exhibit sharp distributions for the Quinary dataset, reflecting the chemical similarity of the elements that it contains. Additionally, the peaks of the ΔS_{mix} distributions are located at large values for the Quinary dataset, which results from the increased number of components in the system (refer to the definition of ΔS_{mix} in Table II). Overall, the Combined dataset generally preserves the feature space distribution of the Ternary dataset, and is complemented by the details of the Quinary data.

Machine Learning Model

ANN models were employed to train the datasets for GFA prediction, and the architecture of the ANN model is illustrated in Fig. 4. It had a feed-forward structure with one input layer (13 normalized features), two hidden layers with 30 and 15 neurons, and one output layer (1 neuron). The number of hidden layers and the number of neurons per layer were determined through a search of the hyperparameter space to simultaneously minimize the model’s error and its complexity. We opted for a brute force search method to tune the hyperparameters over other possibilities, such as Bayesian optimization, due to the relatively small model size and hyperparameter search space.³⁸ The details of the hyperparameter explorations are provided in the supplementary Tables S1 and S2. Firstly, the values of all the 13 features were normalized and served as the values of input layer neurons. The

Table II. The features and their formulae used in the ANN models

Features	Formula
Average valance electron concentration (VEC)	$\text{VEC} = \sum_{i=1}^n (c_i \text{VEC}_i)$
Standard deviation of VEC	$\sigma_{\text{VEC}} = \sqrt{\sum_{i=1}^n c_i (\text{VEC}_i - \text{VEC})^2}$
Configurational entropy of mixing	$\Delta S_{\text{mix}} = -R \sum_{i=1}^n (c_i \ln c_i)$
Average Pauling's electronegativity (PE)	$\text{PE} = \sum_{i=1}^n (c_i \text{PE}_i)$
Standard deviation of PE	$\sigma_{\text{PE}} = \sqrt{\sum_{i=1}^n c_i (\text{PE}_i - \text{PE})^2}$
Average atomic radius (R)	$R = \sum_{i=1}^n (c_i r_i)$
Atomic radius difference	$\delta = \sqrt{\sum_{i=1}^n c_i \left(1 - \frac{r_i}{R}\right)^2}$
Average melting point	$T_m = \sum_{i=1}^n c_i (T_m)_i$
Standard deviation of melting point	$\sigma_{T_m} = \sqrt{\sum_{i=1}^n c_i ((T_m)_i - T_m)^2}$
Average heat of mixing	$H_{\text{mix}} = \sum_{i=1}^n 4\Delta H_{ij}^{\text{mix}} c_i c_j$
Standard deviation of heat of mixing	$\sigma_{H_{\text{mix}}} = \sqrt{\sum_{i=1}^n c_i c_j \left(\Delta H_{ij}^{\text{mix}} - H_{\text{mix}}\right)^2}$
Average bulk modulus (B)	$B = \sum_{i=1}^n (c_i B_i)$
Standard deviation of bulk modulus	$\sigma_{\text{Bulk}} = \sqrt{\sum_{i=1}^n c_i (B_i - B)^2}$

inputs then went through each of the hidden layers, being transformed by a sigmoid function in each layer. Lastly, the output neuron received the information transmitted from the hidden layer and generated a value between 0 and 1. Thereby, the output value measured the probability to form the amorphous structures and was treated as GFA. In the training dataset, label 0 was given to crystal structure and label 1 was given to amorphous structure. Two metrics were applied to measure the model performance. One was the root mean squared error (RMSE), and the other was the accuracy. RMSE was used as the training metric for hyperparameter tuning. To calculate accuracy, a binary classification was applied to the output values. When the output value was larger than or equal to the threshold of 0.5, the alloy was classified as amorphous; otherwise, the alloy was classified as crystal. The accuracy was the correct prediction percentage, i.e., the ratio of the correct prediction to the total number of samples.

Three ANN models were trained based on the three datasets, Ternary, Quinary, and Combined

models. We evaluated model performance using a ten-fold cross-validation (CV) scheme, and a leave-one-alloy-system-out validation approach. Notably, for the leave-one-alloy-system-out validation, we left the data (alloys) in one alloy system completely out of the training dataset, and that data only served as the validation dataset. Since the leave-one-alloy-system-out validation evaluated the model performance on an entire unseen alloy system, it provided a better measure of the model's generalizability to the new, unseen alloy systems.

Experimental Validation

We employed magnetron co-sputtering to fabricate multicomponent thin films for model validation. The compositions and structure of the thin films were determined by x-ray energy dispersive spectrometer (EDS), x-ray diffraction (XRD), and high-resolution transmission electron microscopy (HRTEM). The details of the experiments are provided in the supplementary material.

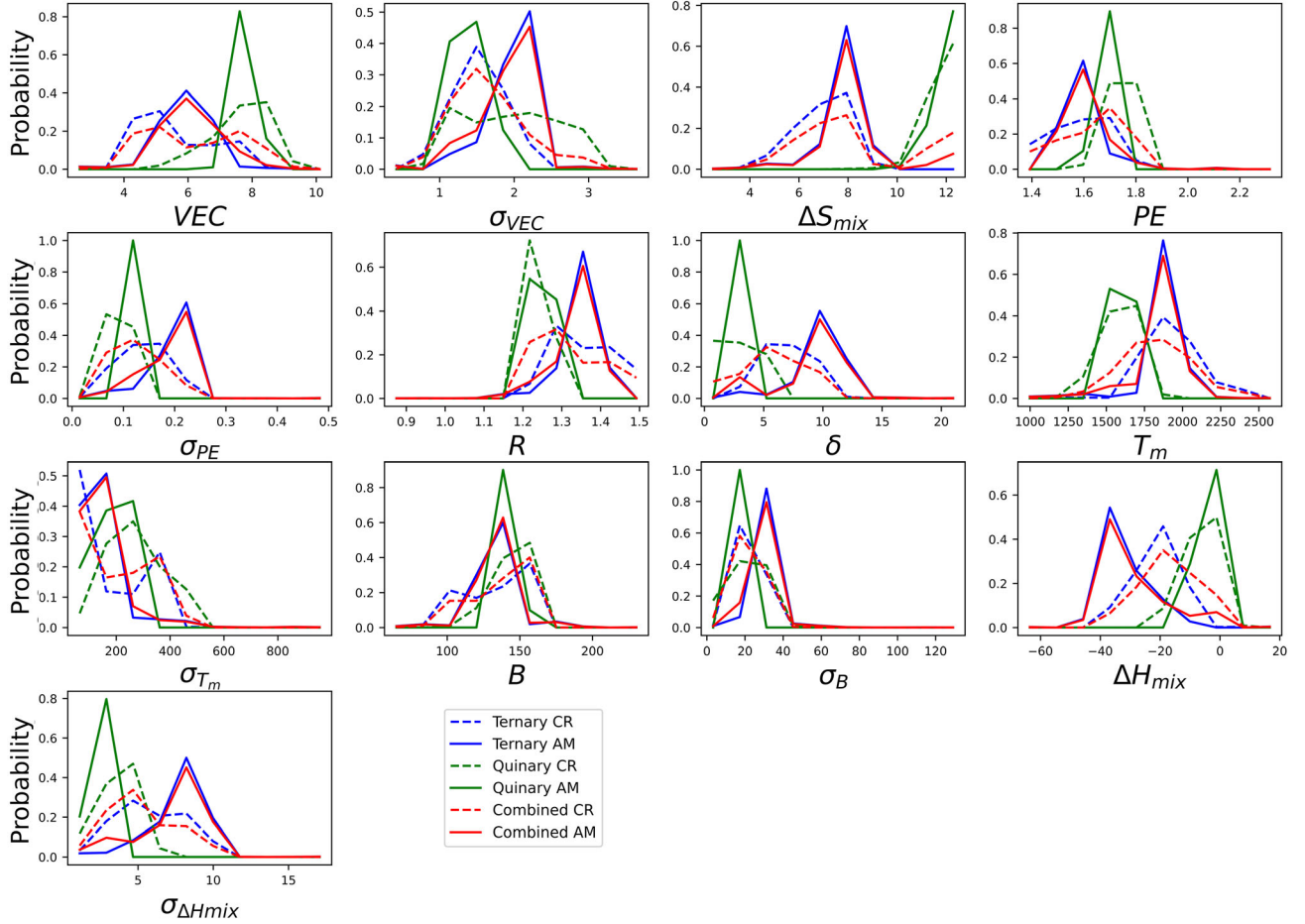


Fig. 3. Probability density distributions of the 13 features. The dashed and solid lines represent the distribution of the *CR* (crystal) and *AM* (amorphous) data points, respectively. The blue, green, and red lines represent the Ternary, Quinary, and Combined datasets, respectively (Color figure online).

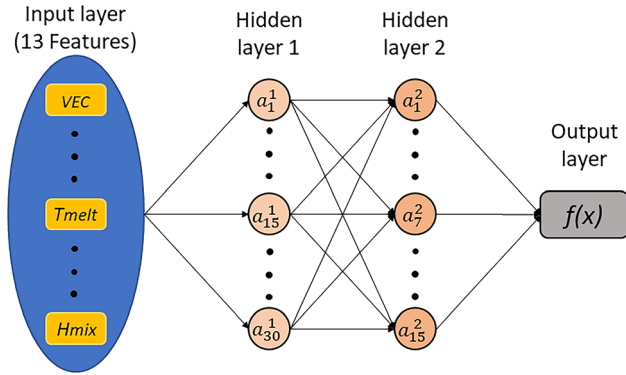


Fig. 4. Schematic of the ANN model. The ANN model has two hidden layers with 30 and 15 neurons, respectively. The output layer generates a value between 0 (crystal) and 1 (amorphous).

RESULTS AND DISCUSSION

ANN Model Performance

The performance of the ANN models trained on three datasets in the ten-fold CV is summarized in Table III. All the models show a good performance with accuracies of 95.2%, 98.6%, and 96.1%, and

Table III. Ten-fold cross-validation (CV) results

Models	RMSE	Accuracy
Ternary	0.189 ± 0.009	$95.2 \pm 0.62\%$
Quinary	0.094 ± 0.024	$98.6 \pm 0.80\%$
Combined	0.173 ± 0.018	$96.1 \pm 0.83\%$

RMSE of 0.189, 0.094, and 0.173 for the Ternary, Quinary, and Combined ANN models, respectively. The limitation of ten-fold CV is that it cannot be used to evaluate the generalizability of the model to explore new alloy systems. In the ten-fold CV scheme, the data are randomly split into training and validation sets. Therefore, the training and validation sets will include data from the same alloy system. In other words, upon training the model, it sees data in all the alloy systems. Therefore, the model metrics obtained by ten-fold CV may not be reliable to evaluate the ML model performance on unseen alloy systems.

The model performance in the leave-one-alloy-system-out validation is shown in Fig. 5 and

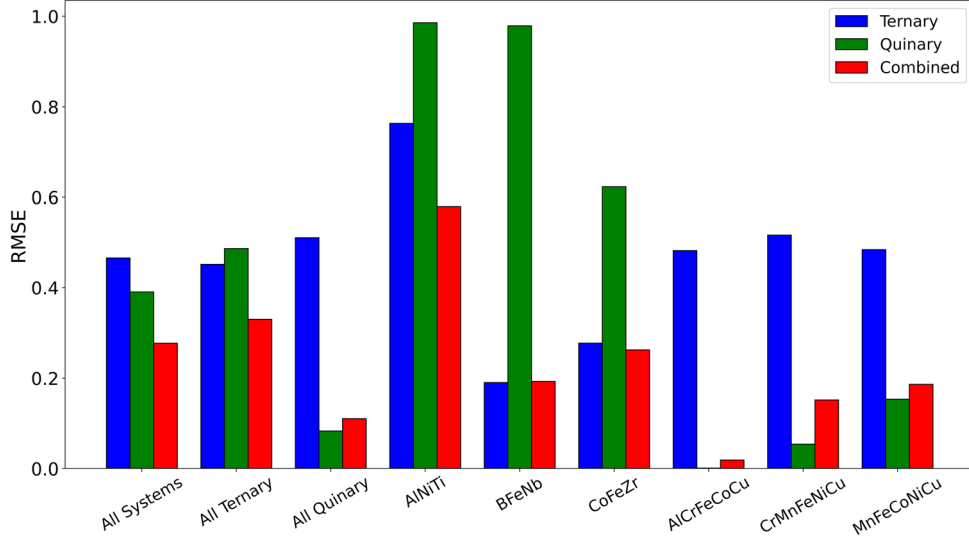


Fig. 5. Root mean squared error (RMSE) for each model when tested with leave-one-alloy-system-out validation for the Combined, Ternary, and Quinary datasets, and the selected alloy systems.

Table IV. The model performance in leave-one-alloy-system-out validation

Models	Validation dataset	RMSE	Accuracy (%)
Ternary	Ternary	0.452	71.2
	Quinary	0.511	52.0
	Combined	0.466	66.7
Quinary	Ternary	0.487	65.6
	Quinary	0.083	97.2
	Combined	0.391	73.1
Combined	Ternary	0.330	85.6
	Quinary	0.110	96.6
	Combined	0.278	88.3

Table IV. The Combined model shows the best performance on the Combined dataset with an overall accuracy of 88.3% and RMSE of 0.278, when compared to the Ternary and Quinary models with accuracies of 66.7% and 73.1%, and RMSE of 0.466 and 0.391, respectively. Exploring the model performance on the ternary systems, the Combined model has the lowest RMSE = 0.330, followed by the Ternary model with RMSE = 0.452 in the leave-one-alloy-system-out validation. For quinary alloy systems, the Quinary model has the lowest RMSE = 0.083, whereas the Combined model ranks second with a slightly higher RMSE = 0.110. The Quinary model behaves the worst in the ternary alloys, especially for the alloy systems containing the chemical elements that are not included in the quinary dataset, such as RMSE = 1.0 for the AlNiTi system, and RMSE = 1.0 for the BFeNb system. While the Ternary model has the largest RMSE (= 0.511) for quinary datasets, and the errors are consistent across different quinary systems, e.g., RMSE = 0.482 for AlCrFeCoCu and RMSE = 0.516 for CrMnFeNiCu, the Combined model represents the most balanced one, and will be employed to

further explore the GFA in the multicomponent alloy systems. The detailed performances of the three models in the leave-one-alloy-system-out validation are in supplementary Table S3 and Figs. S1 ~ S11.

Essentially, the supervised ML model is a mapping function between the input features and the output values.⁴⁴ In this study, the ML models correlate the 13 features of alloys to the GFA, and the density distributions in the 13-feature space determine the ML model performance. Thereby, the differences in model performance in the leave-one-alloy-system-out validation can be attributed to the differences in data distribution in the feature space for the three datasets. As shown in Fig. 3, the Ternary and Quinary datasets are distinct in the feature space distribution, and such distinctions account for the worst performance of the Ternary model in the Quinary dataset and the Quinary model in the Ternary dataset. On the other hand, the Combined model slightly outperforms the Ternary models on the Ternary dataset with additional information from the Quinary dataset. As shown in Fig. 5, the Combined model has a lower

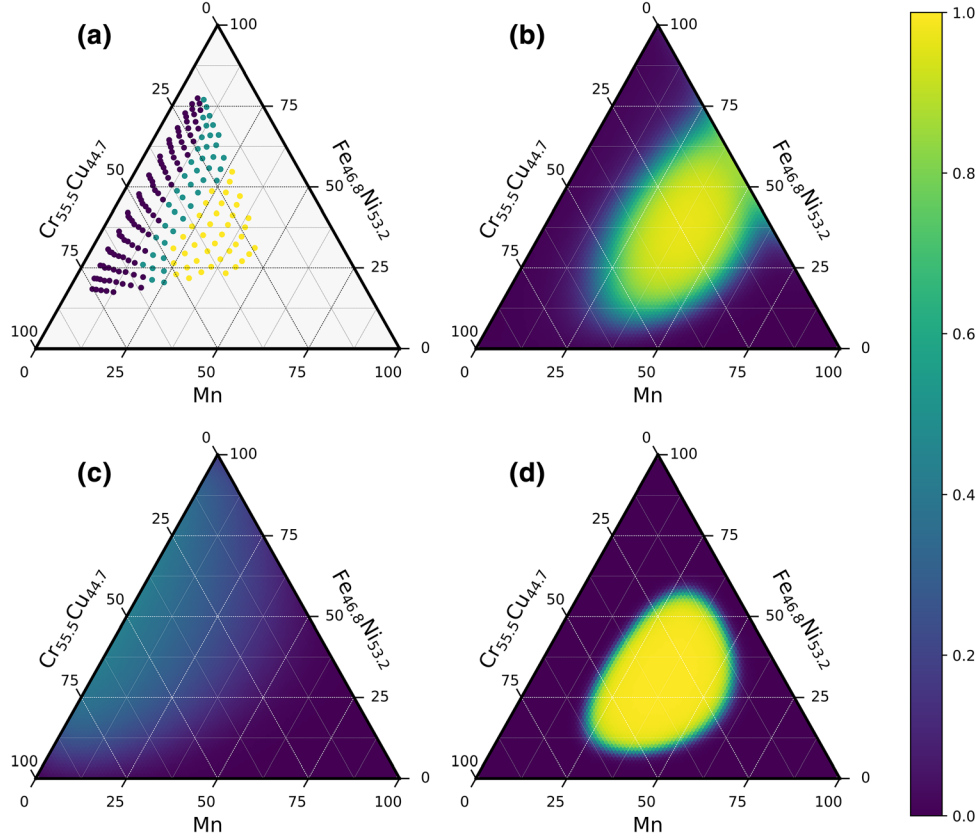


Fig. 6. GFA in the CrMnFeNiCu system. (a) Experimental data with crystal as purple, amorphous as yellow, and crystal-amorphous dual-phase alloys as green. The GFA predictions of models on this system for the (b) Combined, (c) Ternary, and (d) Quinary models (Color figure online).

RMSE than that of the Ternary model in the AlNiTi ternary alloy systems. As Al and Ni are in the Quinary dataset, the addition of these data points provides extra information about Al and Ni properties, leading to better performance. For the Quinary alloy systems, the Combined model shows a much better performance than that of the Ternary model, but is a little worse than that of the Quinary model. Considering the narrow distribution of the quinary data in the feature space, it is reasonable to suspect that the Quinary model is overfitting to the quinary alloy systems. We further examine the robustness of the Quinary and Combined models by calculating:⁴⁵

$$\rho_{adv}(\hat{k}) = E_x \frac{\Delta(x; \hat{k})}{\|x\|_2} \quad (1)$$

where $\Delta(x; \hat{k}) := \min_r \|r\|_2$ subject to $\hat{k}(x+r) \neq \hat{k}(x)$.

Where E_x is the expectation, x is a vector of the features of a data point, $\hat{k}$ is the model, and r is the perturbation. The robustness of the Quinary and Combined models are 23.91 and 86,285.74, respectively. A larger value of the robustness represents a bigger perturbation needed to make the model do an opposite prediction. The significantly enhanced robustness of the Combined model demonstrates its improved noise resistance, suggesting it is a

reliable model for exploring multicomponent alloy systems. Additionally, we have collected 74 more binary and quaternary sputtering data points as a new test dataset. The Combined model performs the best with RMSE = 0.448 and an accuracy of 82.67% compared to RMSE = 0.501 and 0.738 for the Ternary and Quinary models, respectively (see supplementary Table S4).

Figure 6 compares the (a) ground true and (b–d) predicted GFA by the three ANN models in the CrMnFeNiCu alloy systems. The yellow, blue and purple points in Fig. 6a represent the amorphous structure, dual-phase structure (mixture of crystal and amorphous structures), and crystal structure, respectively, based on the experimental results. The Ternary model shows a huge prediction error with the ground truth, as shown in Fig. 6c. It predicts that all the alloys in this system have a low GFA (i.e., less than 0.5), and cannot capture the amorphous structure in the center of the alloy system. It seems that the predicted results by the Quinary model have the best agreement with the experimental data. However, the predicted high-GFA is confined to the area where the alloys show an amorphous structure and the boundary between the high-GFA and low-GFA regions is very sharp. According to Fig. 6d, the Quinary model fails to

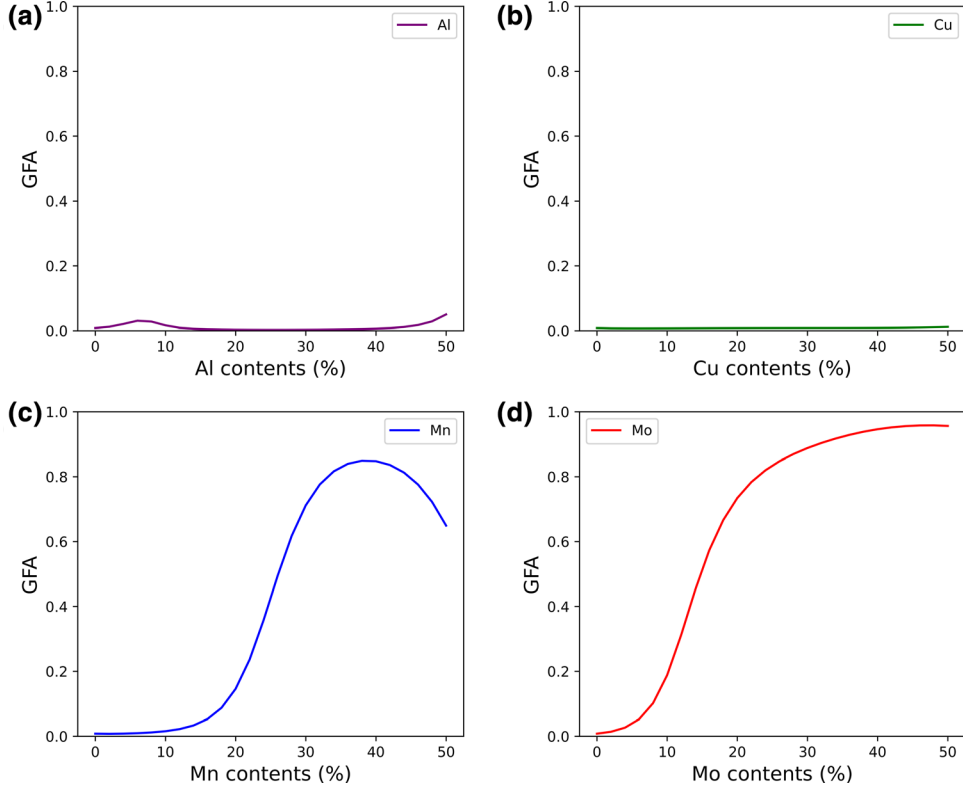


Fig. 7. Predicted GFA by the Combined model for the CoCrFeNi- M systems, where M represents Al, Cu, Mn, and Mo for (a), (b), (c), and (d), respectively.

identify the experimentally shown dual-phase alloy region, denoted as the green dots in Fig. 6a. It is noteworthy that the dual-phase data is not included in the training dataset, the failure of prediction further indicates that the Quinary model overfits the Quinary data. Upon infusing the ternary data, the Combined model shows a slower transition from a high-GFA to a low-GFA region (a broader boundary) in the compositional space, corresponding to the existence of the dual-phase alloys, as illustrated in Fig. 6b.

Prediction of GFA for Multicomponent Alloys

We employed the Combined ANN model to explore multicomponent alloys with high GFA. As a start, we focus on the CoCrFeNi-based alloy system and explore its derivatives by introducing the fifth element. The predicted GFA of four representative alloy systems are displayed in Fig. 7. The model predicts the GFA of CoCrFeNi to be around 0, which is consistent with experimental observations in the literature.⁴⁶ The addition of Al or Cu does not enhance GFA even up to 50% of alloying, but the addition of Mn and Mo does increase the GFA of CoCrFeNi. More specifically, the small addition of Mo leads to a dramatic increase in the GFA, with the maximum value around 0.95.

We further validate the Combined ANN model prediction by fabricating CoCrFeNi-Mo alloys using magnetron co-sputtering with CoCrFeNi target and

pure Mo target, as shown in Fig. 8a. The chemical compositions of the CoCrFeNi-Mo thin film alloys are measured by EDS (Fig. 8b), where the Fe, Co, Cr, and Ni elements are nearly identical, and the Mo contents are ~ 0 at.%, 6 at.%, 13 at.%, 17 at.%, and 23 at.%. We denote these alloys according to the Mo contents, as Mo-0, Mo-6, Mo-13, Mo-17, and Mo-23. The XRD patterns of these alloys are shown in Fig. 8c. Mo-0 exhibits a small (111) peak at around 43° , confirming a crystal FCC structure. With the addition of Mo, the diffraction peaks in the XRD patterns disappear, indicating the structure transition from a crystal to an amorphous structure. Figure 8d compares the Combined model prediction versus experimental observations as a function of Mo content. Evidently, with increasing Mo content, the CoCrFeNi-Mo thin films undergo a transition from a crystalline to an amorphous structure, generally agreeing with the model prediction. As shown in the inset HRTEM of Fig. 8d, a mixture of ordered and disordered atomic arrangements is observed for Mo-13, and the selective area electron diffraction pattern further confirms the dual-phase structure with nanocrystalline diffraction rings and a diffuse halo of amorphous structure. The predicted GFA of Mo-13 is 0.381, reasonably matching the experimental observation of dual-phase structures.

To evaluate the influence of datasets on the CoCrFeNi-Mo prediction, the Combined dataset is visualized by the 2-D projection of the high-

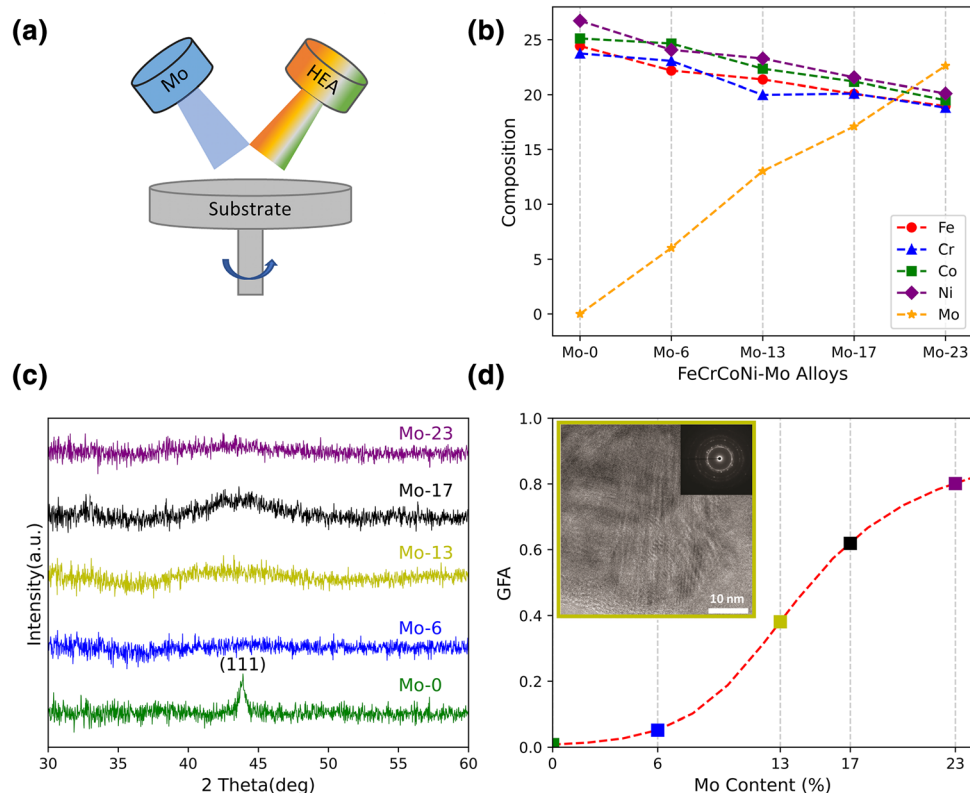


Fig. 8. Experimental validation for the CoCrFeNi-Mo system. (a) Schematic of magnetron co-sputtering with CoCrFeNi target and pure Mo target. (b) The alloy compositions of the five sputtered samples were determined by EDS. (c) XRD results of the sputtered samples with the variation of Mo contents (d) The Combined model prediction versus experimental observations as a function of Mo content. The inset is an HRTEM image of the Mo-13 sample.

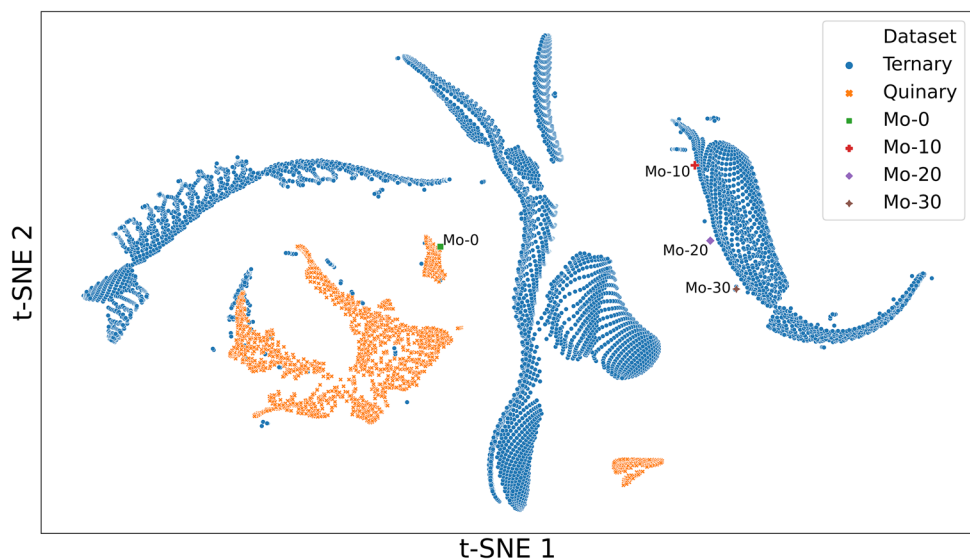


Fig. 9. t-SNE plot of the CoCrFeNi-Mo multicomponent alloys in the Combined dataset. This scatter plot shows a 2D projection of the high-dimensional composition space that is achieved via t-SNE. The Ternary dataset is in blue, and the Quinary dataset is in orange. The composition data correspond to the five CoCrFeNi-Mo alloys are denoted in various symbols (Color figure online).

dimensional feature space, as illustrated in Fig. 9. We use the ML algorithm called t-distributed stochastic neighbor embedding (t-SNE) to represent all composition data from both Ternary (in orange) and Quinary (in blue) datasets in two dimensions.⁴⁷

This technique helps quantify the similarity between points and visualize high-dimensional data, reducing the 13-dimensional space given by the set of possible alloying elements to a 2-D scatter plot. Notably, the Ternary dataset shows a much

broadier distribution than the Quinary dataset, which is consistent with the observation in the feature space (Fig. 3). The CoCrFeNi alloy lies near the cluster of the Quinary dataset, whereas the CoCrFeNi-Mo alloys lie near the cluster of the Ternary dataset. In the Combined dataset, the Quinary dataset provides the information on CoCrFeNi, and the Ternary dataset provides the information on the Mo addition. The newly designed compositions fall in or near the regions spanned by the existing GFA dataset. Also, a large separation between the Mo-0 and the other Mo-contained samples (i.e., Mo-10, Mo-20, and Mo-30) can be observed in the t-SNE plot, which is also accompanied by a large increase in the predicted GFA (Fig. 8d). Such an increase in the predicted GFA may come from the imbalanced data distribution in the Mo-contained alloys (see supplementary Fig. S14), where, among the 25 Mo-contained alloys, 24 of them have amorphous structures. The biased amorphous/crystal distribution could lead to the increase of predicted GFA in the CoCrFeNi-Mo HEAs. However, it is noteworthy that a slight increase in the GFA from Mo-0 and Mo-6 is observed in Fig. 8d, indicating that the model is not completely dominated by the imbalanced Mo data distribution. So far, our model has been used to explore the vicinity of the existing data regions in the t-SNE plot. The generalizability of the model to the unexplored regions (e.g., blank area in the t-SNE plot) will require further experiments, providing more data for model training and validation.

CONCLUSION

We have developed ANN models based on different datasets to investigate the GFA of multicomponent alloys. The ANN model, trained on a combined dataset including both ternary and quinary alloys, exhibits the best performance, well balancing the model accuracy and generalizability, evident by the low RMSEs in leave-one-alloy-system-out validation and high model robustness. Using the ANN model trained on the combined dataset, we explore the GFA in the CoCrFeNi-based alloy systems and identify a CoCrFeNi-Mo alloy system that can have high GFA. To verify the ML model prediction, we perform magnetron co-sputtering to synthesize CoCrFeNi-Mo thin films with Mo content varying from 0 at.% to 23 at.%. A transition from a crystalline to an amorphous structure in the CoCrFeNi-Mo thin films is observed, agreeing nicely with the model prediction. Furthermore, the t-SNE plot reveals the newly identified alloy compositions are in the vicinity of the clusters of the existing datasets, suggesting that the discovery of new alloy systems requires more high-fidelity data to expand the compositional space via extensive research into data collection and generation by a high-throughput means.

SUPPLEMENTARY INFORMATION

The online version contains supplementary material available at <https://doi.org/10.1007/s11837-022-05549-w>.

ACKNOWLEDGEMENTS

This work is supported by the U.S. National Science Foundation (CMMI-1727875), and the National Aeronautics and Space Administration (NASA), Alabama EPSCoR (contract# 80NSSC21M0176). Y.Y and L.L also acknowledge the additional financial support from U.S. National Science Foundation under contract No. (DMR-2104656). This work used the computing resource provided by Extreme Science and Engineering Discovery Environment (XSEDE-MAT210034).

CONFLICT OF INTEREST

All the authors declare that they have no conflicts of interest.

REFERENCES

1. C. Schuh, T. Huftnagel, and U. Ramamurty, *Acta Mater.* 55, 4067–4109. <https://doi.org/10.1016/j.actamat.2007.01.052> (2007).
2. W. Wang, *Prog. Mater. Sci.* 52, 540–596. <https://doi.org/10.1016/j.pmatsci.2006.07.003> (2007).
3. W.H. Wang, C. Dong, and C.H. Shek, *Mater. Sci. Eng. R. Rep.* 44, 45–89. <https://doi.org/10.1016/j.mser.2004.03.001> (2004).
4. B. Zberg, P.J. Uggowitzer, and J.F. Löffler, *Nat. Mater.* 8, 887–891. <https://doi.org/10.1038/nmat2542> (2009).
5. W.L. Johnson, *MRS Bull.* 24, 42–56. <https://doi.org/10.1557/S0883769400053252> (1999).
6. D. Turnbull, *Contemp. Phys.* 10, 473–488. <https://doi.org/10.1080/00107516908204405> (2006).
7. X. Yang and Y. Zhang, *Mater. Chem. Phys.* 132, 233–238. <https://doi.org/10.1016/j.matchemphys.2011.11.021> (2012).
8. A. Inoue, *Acta Mater.* 48, 279–306. [https://doi.org/10.1016/S1359-6454\(99\)00300-6](https://doi.org/10.1016/S1359-6454(99)00300-6) (2000).
9. X.H. Du, J.C. Huang, C.T. Liu, and Z.P. Lu, *J. Appl. Phys.* <https://doi.org/10.1063/1.2718286> (2007).
10. B.S. Dong, S.X. Zhou, D.R. Li, C.W. Lu, F. Guo, X.J. Ni, and Z.C. Lu, *Prog. Natl. Sci. Mater. Int.* 21, 164–172. [https://doi.org/10.1016/S1002-0071\(12\)60051-3](https://doi.org/10.1016/S1002-0071(12)60051-3) (2011).
11. R. Busch, J. Schroers, and W.H. Wang, *MRS Bull.* 32, 620–623. <https://doi.org/10.1557/mrs2007.122> (2011).
12. Z.P. Lu and C.T. Liu, *Phys. Rev. Lett.* 91, 115505. <https://doi.org/10.1103/PhysRevLett.91.115505> (2003).
13. Z.W. Wu, M.Z. Li, W.H. Wang, and K.X. Liu, *Nat. Commun.* 6, 6035. <https://doi.org/10.1038/ncomms7035> (2015).
14. O.N. Senkov, J.D. Miller, D.B. Miracle, and C. Woodward, *Nat. Commun.* 6, 6529. <https://doi.org/10.1038/ncomms7529> (2015).
15. B. Deng and Y. Zhang, *Chem. Phys.* <https://doi.org/10.1016/j.chemphys.2020.110898> (2020).
16. Z. Fan, J. Ding, and E. Ma, *Mater. Today* 40, 48–62. <https://doi.org/10.1016/j.mattod.2020.05.021> (2020).
17. Z. Fan and E. Ma, *Nat. Commun.* 12, 1506. <https://doi.org/10.1038/s41467-021-21806-z> (2021).
18. X. Liu, X. Li, Q. He, D. Liang, Z. Zhou, J. Ma, Y. Yang, and J. Shen, *Acta Mater.* 201, 182–190. <https://doi.org/10.1016/j.actamat.2020.09.081> (2020).

19. J.C. Mauro, A. Tandia, K.D. Vargheese, Y.Z. Mauro, and M.M. Smedskjaer, *Chem. Mater.* 28, 4267–4277. <https://doi.org/10.1021/acs.chemmater.6b01054> (2016).
20. L. Peng, Z. Long, and M. Zhao, *Comput. Mater. Sci.* <https://doi.org/10.1016/j.commatsci.2021.110480> (2021).
21. L.E. Schultz, B. Afflerbach, I. Szlufarska, and D. Morgan, *Comput. Mater. Sci.* <https://doi.org/10.1016/j.commatsci.2021.110877> (2022).
22. Y.T. Sun, H.Y. Bai, M.Z. Li, and W.H. Wang, *J. Phys. Chem. Lett.* 8, 3434–3439. <https://doi.org/10.1021/acs.jpcclett.7b01046> (2017).
23. Q. Wang, J. Ding, L. Zhang, E. Podryabinkin, A. Shapeev, and E. Ma, *npj Comput. Mater.* <https://doi.org/10.1038/s41524-020-00467-4> (2020).
24. Q. Wang and L. Zhang, *Nat. Commun.* 12, 5359. <https://doi.org/10.1038/s41467-021-25490-x> (2021).
25. L. Ward, A. Agrawal, A. Choudhary, and C. Wolverton, *npj Comput. Mater.* <https://doi.org/10.1038/npjcompumats.2016.28> (2016).
26. L. Ward, S.C. O’Keeffe, J. Stevick, G.R. Jelbert, M. Aykol, and C. Wolverton, *Acta Mater.* 159, 102–111. <https://doi.org/10.1016/j.actamat.2018.08.002> (2018).
27. W.-P. Wu, D. Şopu, X. Yuan, O. Adjaoud, K.K. Song, and J. Eckert, *J. Non-Cryst. Solids*. <https://doi.org/10.1016/j.jnoncrysol.2021.120676> (2021).
28. J. Xiong, S.-Q. Shi, and T.-Y. Zhang, *Comput. Mater. Sci.* <https://doi.org/10.1016/j.commatsci.2021.110362> (2021).
29. J. Xiong, T.-Y. Zhang, and S.-Q. Shi, *MRS Commun.* 9, 576–585. <https://doi.org/10.1557/mrc.2019.44> (2019).
30. Z. Zhou, Y. Shang, and Y. Yang, *J. Mater. Inform.* <https://doi.org/10.20517/jmi.2021.12> (2022).
31. Z. Zhou, Y. Zhou, Q. He, Z. Ding, F. Li, and Y. Yang, *npj Comput. Mater.* <https://doi.org/10.1038/s41524-019-0265-1> (2019).
32. F. Ren, L. Ward, T. Williams, K.J. Laws, C. Wolverton, J. Hattrick-Simpers, and A. Mehta, *Sci. Adv.* 4, eaq1566. <https://doi.org/10.1126/sciadv.aq1566> (2018).
33. L. Tian, Y. Fan, L. Li, and N. Mousseau, *Scripta Mater.* 186, 185–189. <https://doi.org/10.1016/j.scriptamat.2020.05.038> (2020).
34. Y. Yao, T. Sullivan, F. Yan, J. Gong, and L. Li, *Scripta Mater.* 209, 114366. <https://doi.org/10.1016/j.scriptamat.2021.114366> (2022).
35. Y. Zhang and X. Xu, *Polym. Chem.* 12, 843–851. <https://doi.org/10.1039/D0PY01581D> (2021).
36. E. Alcobaça, S.M. Mastellini, T. Botari, B.A. Pimentel, D.R. Cassar, A.C.P.L.F. de Carvalho, and E.D. Zanotto, *Acta Mater.* 188, 92–100. <https://doi.org/10.1016/j.actamat.2020.01.047> (2020).
37. B.T. Afflerbach, C. Francis, L.E. Schultz, J. Spethson, V. Meschke, E. Strand, L. Ward, J.H. Perepezko, D. Thoma, P.M. Voyles, I. Szlufarska, and D. Morgan, *Chem. Mater.* 34, 2945–2954. <https://doi.org/10.1021/acs.chemmater.1c03542> (2022).
38. P.W. Guan and V. Viswanathan, <https://arxiv.org/abs/2010.14048> (2020). <https://doi.org/10.48550/arXiv.2010.14048>.
39. R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, and C. Kim, *npj Comput. Mater.* 3, 54. <https://doi.org/10.1038/s41524-017-0056-5> (2017).
40. Y. Kawazoe, J.Z. Yu, A.P. Tsai, and T. Masumoto, *Nonequilibrium Phase Diagrams of Ternary Amorphous Alloys* (Springer, Berlin, 1997).
41. S.A. Kube, S. Sohn, D. Uhl, A. Datye, A. Mehta, and J. Schroers, *Acta Mater.* 166, 677–686. <https://doi.org/10.1016/j.actamat.2019.01.023> (2019).
42. H. Joress, B.L. DeCost, S. Sarker, T.M. Braun, S. Jilani, R. Smith, L. Ward, K.J. Laws, A. Mehta, and J.R. Hattrick-Simpers, *ACS Comb. Sci.* 22, 330–338. <https://doi.org/10.1021/acscombsci.9b00215> (2020).
43. S.A. Kube and J. Schroers, *Scripta Mater.* 186, 392–400. <https://doi.org/10.1016/j.scriptamat.2020.05.049> (2020).
44. S.J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach* (Pearson, 2021).
45. S. Moosavi-Dezfooli, A. Fawzi, and P. Frossard, In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2016), pp 2574–2582. <https://doi.org/10.1109/CVPR.2016.282>.
46. X.B. Feng, W. Fu, J.Y. Zhang, J.T. Zhao, J. Li, K. Wu, G. Liu, and J. Sun, *Scripta Mater.* 139, 71–76. <https://doi.org/10.1016/j.scriptamat.2017.06.009> (2017).
47. L. Van der Maaten and G. Hinton, *J. Mach. Learn. Res.* 9. <https://www.jmlr.org/papers/v9/vandermaaten08a.html> (2008).

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