



Application of fuzzy learning in the research of binary alloys: Revisit and validation



Huiran Zhang^{a,b,c}, Gaofeng Zhou^a, Shengzhou Li^a, Xingyue Fan^a, Zhiting Guo^a, Tao Xu^a, Yan Xu^d, Xue Chen^a, Dongbo Dai^a, Quan Qian^{a,b}

^a School of Computer Engineering and Science, Shanghai University, Shanghai 200444, China

^b Shanghai Institute of Materials Genome, Shanghai 200444, China

^c Shanghai Institute for Advanced Communication and Data Science, Shanghai University, Shanghai 200444, China

^d College of Mathematics and Physics, Shanghai University of Electric Power, Shanghai 200090, China

ARTICLE INFO

Keywords:

Binary alloys
Hume-rothery rules
Fuzzy learning
Fuzzy rules
Support vector machines

ABSTRACT

Hume-Rothery rules, occupying a central space at the heart of metallurgy, are one of the most important rules in materials science. One limitation to the binary alloys is the lack of predictable tools for understanding the relationship between **alloy structure and solid solubility**. In this paper, with Hume-Rothery rules, we demonstrated the proposed method could be used to give quantitative analysis of solid solubility. Then a hybrid Fuzzy-SVM (Support Vector Machines) method is proposed to predict solid solubility of binary alloy systems. In the study of silver and copper alloy, our methods successfully revealed the understandable relationship between structures (atom size, valency, and electronegativity) and properties (solid solubility). A similar fuzzy rule like H-R's 15% rule was found and an estimation to the solid solubility was given.

1. Introduction

Solid solubility is one of the most important property parameters in the theoretical research of metals and alloys. Hume Rothery proposed three empirical rules to explain the formation of solid solutions: (1) the atomic size of the solvent and solute must not differ by more than 15%, (2) the electrochemical nature of the two elements must be similar, and (3) a higher-valent metal is more soluble in a lower-valent metal than vice versa. Besides these, a second quantitative rule was revealed by Darken and Gurry, from other aspect with a 2-D map where the two coordinates are electronegativity and atomic size. They pointed out an ellipse (one axis of atomic size is $r \pm 15\%$, and another axis of electronegativity is $r \pm 0.4$) to encompass most of the solutes. These solutes show significant solubility in the solvent [1–3]. These rules are empirical which rely on researchers' domain knowledge and can't give a quantitative prediction to the solid solubility.

In recent years, the application of machine learning in materials science shows broad development prospects [4–8]. It combines computing technology and material knowledge and also can be applied to the research of solid solubility. It has been successfully applied to discover material [9–13] and predict materials properties [14–16], crystal structures [17–20]. Möller et al. used the kernel-based Support Vector Regression(SVR) method to gain optimal composition of hard-magnetic

phase. They found that the predictions for the optimal compositions are consistent with the DFT calculations performed *a posteriori* (the average ML-DFT difference is below 4%) [21]. As for alloy system, Zhang et al. applied an artificial neural network(ANN) to predict solid solubility of binary alloys [22]. In fact, ANN is often considered as black boxes and it's opaque to researchers [23]. So, it provides little explanatory insight into the influence between the factors and solubility limit in the prediction process. In our recently researches, we use the Support Vector Machine (SVM) to predict solid solubility with a small dataset in order to provide evidence for the Hume-Rothery rules (H-R rules) and find factors which affect the solid solubility [24]. As Hume Rothery and co-workers said, the solubility limit is determined by three factors in general, and the interaction between them makes the results so complex [2]. It is generally believed that the interplay cannot be expressed using machine learning methods.

Fuzzy approaches have the potential of giving accurate estimates for solid solution. Its human-interpretable IF-THEN criteria is helpful to understand the interplay between these factors [25]. In this paper, we use the fuzzy learning methods to show interplay between structures and properties of binary alloys by representing the H-R rules with the "IF-THEN" rules (Fig. 1(a)). Combining the fuzzy learning and SVM, we enter these rules as a new input matrix to the SVM method. In this hybrid Fuzzy-SVM approach, linear, polynomial, sigmoid, and Gaussian

E-mail address: hrzhangsh@shu.edu.cn (H. Zhang).

<https://doi.org/10.1016/j.commatsci.2019.109350>

Received 25 August 2019; Received in revised form 9 October 2019; Accepted 12 October 2019
0927-0256/© 2019 Elsevier B.V. All rights reserved.

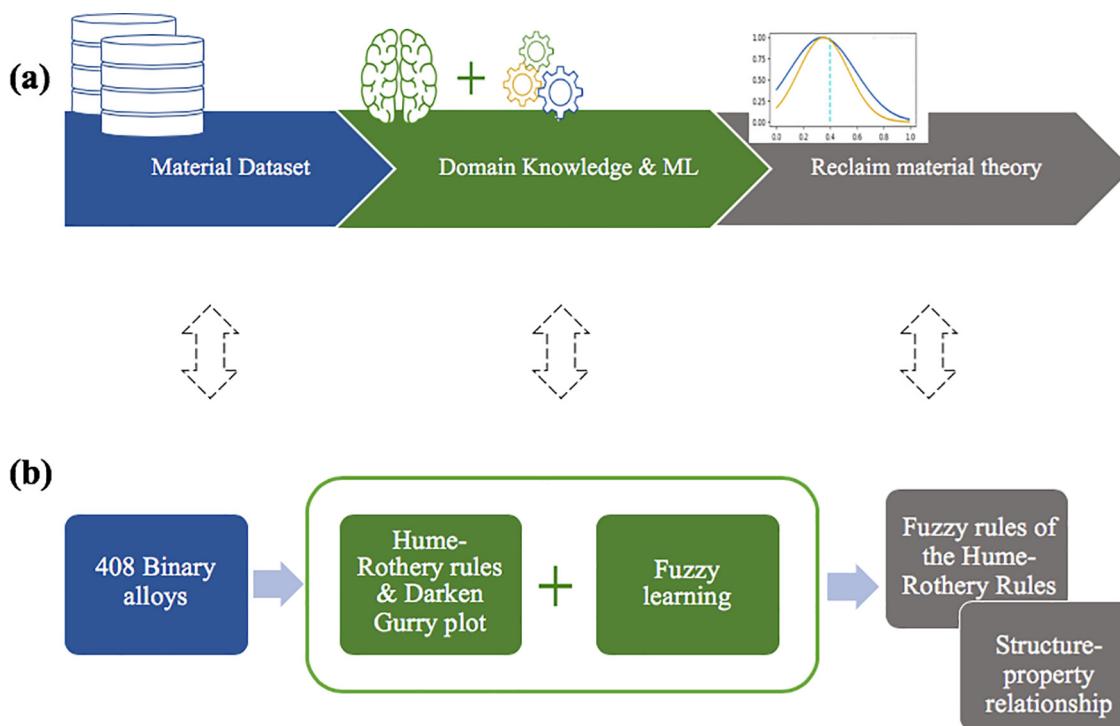


Fig. 1. The flow chart of fuzzy learning method. The input dataset includes 408 binary alloys, 6 initial material features, the targeted property is the solubility limits between solvent and solute. Through combining with binary alloy theories (H-R rules and Darken Gurry plot) and fuzzy learning models to generate fuzzy rules of the H-R rules and structure-property relationship.

kernel functions are employed for the estimation of solid solution.

2. Method

A process which started from material dataset, then combined domain knowledge with machine learning method to revisit material theory is proposed in this paper (Fig. 1(a)). The flow chart of applying fuzzy learning method in binary alloys is given in Fig. 1(b). The input dataset includes 408 silver and copper alloy systems which have six initial material factors and a targeted property. These factors contain atomic size, valence, electronegativity of solvent and solute. The targeted property is the solubility limits between solvent and solute. Through combining with binary alloy theories, three transformed parameters are input into the model. Finally, the fuzzy rules could be obtained by the fuzzy learning method. It could be used to compare with H-R rules to find some situations coincident or different with the H-R rules.

2.1. Fuzzy learning method

A fuzzy model, as a soft computing technique, provides enhanced knowledge in terms of linguistic fuzzy IF-THEN rules that can otherwise be impossible to extract with classical methods [26,27]. Fuzzy learning model based on fuzzy rules are used to identify the potential relationship between input and output variables of a system by fuzzy sets. Generally fuzzy learning models represent relationships between input and output variables with a collection of IF-THEN rules. A fuzzy clustering algorithm called FCM is used to determine the interactive membership values of input-output variables for a given training data set [28]. Let (X_j, Y_j) , $j = 1, \dots, k$ be the set of training data, the steps of fuzzy learning model are given below:

Step 1. An optimal pair (m^*, c^*) is determined with a cluster validity index via iterative search, where m is the rank of fuzziness and c is the size of cluster.

Step 2. Minimize the objective function:

$$\min J(U, c_1, \dots, c_c) = \sum_{j=1}^k \sum_{i=1}^c u_{ij}^m d_{ij}^2 \quad (1)$$

$$\text{s.t. } 0 \leq u_{ij} \leq 1, \forall i, j \quad (2)$$

$$\sum_{i=1}^c u_{ij} = 1, \forall j \quad (3)$$

$$0 \leq \sum_{j=1}^k u_{ij} \leq k, \forall i \quad (4)$$

where J is the objective function, $d_{ij} = c_i - x_j$ is Euclidean distance between the cluster center c_i and the data point x_j .

Step 3. Build fuzzy membership function with Gaussian function.

$$u_{A_k^j}(X_k) = \exp\left(\frac{-(X_k - c_k^j)^2}{\delta_k^j}\right) \quad (5)$$

$$c_k^j = \sum_{i=1}^N u_{ij} x_{ik} / \sum_{i=1}^N u_{ij} \quad (6)$$

$$\delta_k^j = \sqrt{\sum_{i=1}^N u_{ij} (x_{ik} - c_k^j)^2 / \sum_{i=1}^N u_{ij}} \quad (7)$$

Through fuzzy learning, a fuzzy system comprises of a set of IF-THEN fuzzy rules having the follow form:

$$R_i: \text{IF } x_1 \text{ is } A_1^1 \text{ and } \dots \text{ and } x_p \text{ is } A_1^p \text{ Then } y_i = \prod_{j=1}^p A_i^j$$

where $i = 1, \dots, k$, A_i^j ($j = 1, \dots, p$ is the dimension of input vector), are antecedent fuzzy sets, y_i is the output of the i th rule.

2.2. Fuzzy learning with SVM

Another algorithm we usually used is SVM algorithm, which was firstly proposed by Vapnik et al. [29]. Smola and Schölkopf showed that SVR has robust properties according to signal-to-noise ratios and has been used in many materials researches [30]. Considering the potential advantages of both fuzzy and SVM approaches, we propose an SVM approach along with fuzzy regression modeling for the estimation of solid solution using the parameters mentioned by Hume Rothery. In the Fuzzy-SVM approach, linear, polynomial, sigmoid, and Gaussian kernel functions are employed for the estimation of solid solution.

The model mainly includes two parts: training and inference. In the training progress, the parameters of the model are obtained by analyzing training data. After training the model, the prediction accuracy of the model is evaluated through test data in inference stage. The algorithm of Fuzzy-SVM approach is briefly given below:

Step 1. Firstly, using FCM clustering algorithm to determine the cluster centers and membership values:

Step 2. We build a new input matrix through combining the original input and the membership values for each cluster. If $\Gamma = (u_{ij}^l | i = 1, \dots, c^*, j = 1, \dots, k)$, then the membership value can be converted to what we want or need to include in our materials system. There are some examples:

$$X' = [1, \Gamma, X], \text{ or } X'' = [1, \Gamma^2, X], \text{ or } X''' = [1, \Gamma, \Gamma^2, X]$$

Step 3. The complete representation of regression function for a new input matrix is presented as below:

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*)k(x, x_i) + b_i$$

where l is the number of support vectors, α_i and α_i^* are Lagrange multipliers, $k(x, x_i)$ is a kernel function, and b is a bias. In this study, four kernel functions are adopted as the kernel function.

An inference process is proposed to validate the generalization capabilities of the method which determined during the training process. The m^* is a degree of fuzziness and c^* is the number of clusters identified before. For the validation dataset (X_j, Y_j) , $j = 1, \dots, k$ where k is the number of validation data examples. In the Fuzzy-SVM inference mechanism, the estimates of output values are determined by using trained Fuzzy-SVM with specific kernel function.

The inference algorithm proposed for the Fuzzy-SVM method is presented below:

Step 1. The X_j of validation dataset are given as an input to the trained FCM algorithm, the membership values for each input can be calculated.

Step 2. The membership values of inputs obtained in the previous step are combined with the original inputs of the inference data. Thus, the new input matrix, X' for the validation dataset is obtained.

Step 3. The output value of new input matrixes is inferred using trained SVM algorithm with specific kernel function.

3. Result and discussion

3.1. Dataset

The used dataset has also been analyzed by Zhang with ANN [22]. The number of dataset is 408, including eight structural parameters: atomic size, valence, and electronegativity of solvent and solute respectively as well as a property: solubility limits. Here we adopt the first three factors proposed by the H-R rules.

The input parameters of model can be obtained from the rules proposed by the H-R rules. For atom size factor, when the atomic size of

the solvent and solute differ by more than 15%, it is likely to have a low solubility. Hence the difference between solvent and solute will be one of the input parameters. What's more, the 15% rule mentioned by Hume Rothery will also be test by the model. For electronegativity factor, as discussed by Hume Rothery, stable intermetallic compounds are prone to form as the more electronegative is the solute and the more electropositive is the solvent metal, or vice versa [31]. The difference electronegativity of solvent and solute atom will be used which consistent with the Darken Gurry map. And the 0.4 rule mentioned by Darken and Gurry will be test. For valence factor, it seems no quantitative rules, only a relative valence effect discovered by Hume Rothery that a metal of lower valence is more likely to dissolve one of higher valence than vice versa. So, the input parameters of valence will also be the difference of solvent and solute.

Here, we only focus on the metals of the partially soluble (solid solubility between 0 and 100), so that the whole dataset contains 308 examples. The whole dataset is divided into two parts: training dataset and validation dataset, the ratio between them is 7:3. According to the H-R rules, we think that the different factors between solvent and solute are the key to determine the solid solubility. Therefore, the input factors are not all the six factors but the difference between atomic size, electronegativity and valency of solute and solvent.

The input factors are given below:

1. For the atom size factor. The difference of the atomic diameters of solvent and solute divided by the diameter of the solvent atoms is used.
2. For the electronegativity factor. The difference of the electronegativity between the solvent and solute atoms is used.
3. For the valence factor. These are integers and the difference of the valence between the solvent and solute atoms is used.

3.2. Implement of the fuzzy learning method

Fuzzy learning method is applied to binary alloy systems to mine fuzzy rules. For the obtained fuzzy rules, a Gaussian function is applied to visualize them and compare with H-R rules. Firstly, all the extensive solid solutions are put into FCM to generate three clusters and three data centers. Then calculate the variance of each cluster so the two parameters of Gaussian functions are obtained. After getting the fuzzy rules, we can plot them and compare with H-R rules which can revisit and validate H-R rules. Maybe new material theory could be discovered through these comparisons. The rules obtained from the fuzzy learning method were showed in Fig. 2.

Here, the fuzzy rules obtained by fuzzy learning model are visualized by Gaussian functions. Horizontal axis represents the atom size factor, the electronegativity factor and the valence factor discussed in Dataset and vertical axis represents the membership value of solid solubility. After that, the fuzzy rules are compared with the H-R rules. From subgraph of Fig. 2(a1, a2, a3), it can be found that, solid solubilities in the three rules all tend to zero when the value of the atom size factor reaches 15%, which is consistent with H-R rules. On the contrary, the electronegativity's difference between the solute and solvent no more than 0.4 is inconsistent with Darken Gurry plot in Fig. 2(c1, c2, c3), and this situation will be discussed in the next.

As for valence factor, the trend in rule 3 (Fig. 2(b3)) is significantly different from the rest of the rules (Fig. 2(b1), (b2)). The valence factor which H-R rules refer to is the relative valence effect. It is different with the other two rules of H-R rules. A metal with a lower valence is more likely to dissolve in a metal of higher valence. In Fig. 2(b2), the membership value of valence in copper alloy is close to zero which means the valence factor does not affect the solid solubility in this fuzzy rule.

Although H-R rules are simple qualitative rules, but one cannot predict the solid solubility limits accurately without H-R rules. These rules are still useful guidelines for the study of alloy systems [32–35].

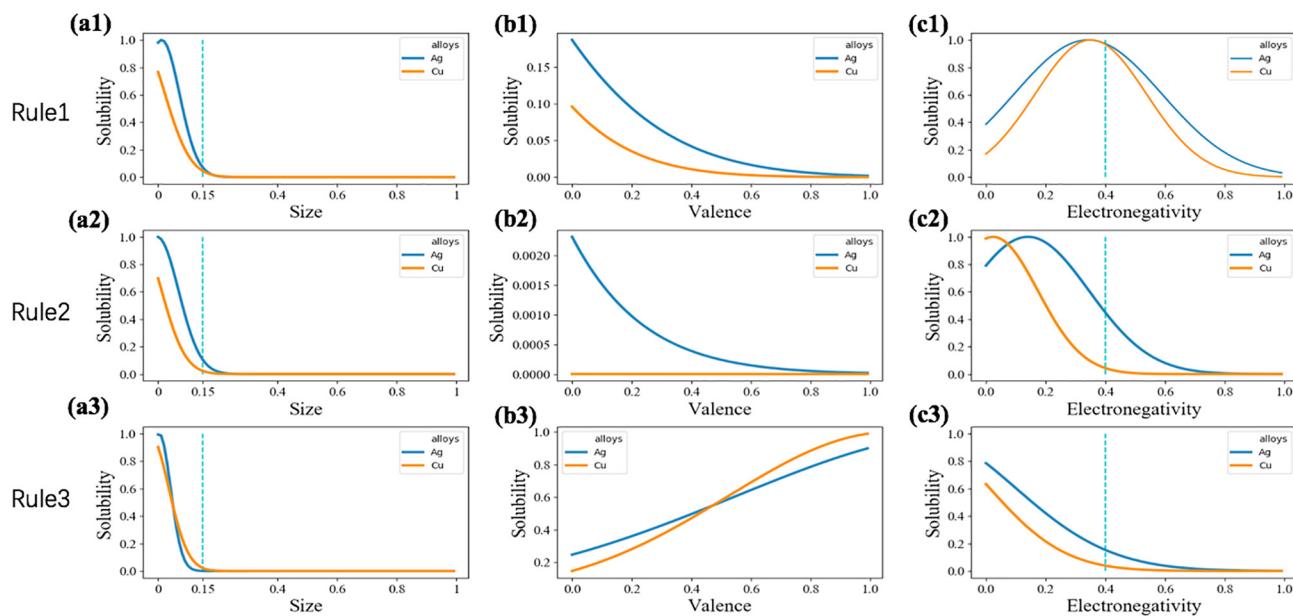


Fig. 2. The rules (a1)-(c3) are obtained from the fuzzy learning model automatically. Horizontal axis represents the atom size factor, the electronegativity factor, the valence factor. Vertical axis represents the membership value of solid solubility. Two types of vertical lines mean the 15% and 0.4 rules which were mentioned by Hume Rothery and Darken Gurry.

Table 1
Statistical error measures applied for method comparison.

Measure	Expression
RMSE	$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y'_i)^2}$
MSE	$\text{MAE} = \left(\frac{1}{n} \sum_{i=1}^n y_i - y'_i \right)$
MaxAE	$\max y_i - y'_i $

Table 2
Comparison of the prediction performance of training phase.

Model type	RMSE	MAE	MaxAE
Fuzzy-SVM-Lin	11.12	6.71	38.80
Fuzzy-SVM-Pol	11.12	6.71	39.37
Fuzzy-SVM-Gauss	2.30*	0.66*	16.00*
Fuzzy-SVM-Sig	12.53	6.17	59.61

*The smallest value.

Table 3
Comparison of the prediction performance of inference phase.

Model type	RMSE	MAE	MaxAE
Fuzzy-SVM-Lin	13.94	7.06	57.92
Fuzzy-SVM-Pol	13.41	6.75	55.48
Fuzzy-SVM-Gauss	8.40*	4.20*	32.74*
Fuzzy-SVM-Sig	8.50	14.18	57.30

*The smallest value.

He et al. used H-R Rules to determine the formation of random solid solution [36]; Tsai et al. verified the availability of the H-R rules for the stability of three groups of stable quasicrystals. And they concluded that the stability of stable quasicrystals and the range of solubility in stable quasicrystals could be explained in terms of the H-R rules [37]; Calvo-Dahlborg et al. used the Hume-Rothery approach in terms of e/a to classify HEAs [38].

The fuzzy rules of electronegative is inconsistent with Darken Gurry

plot in the silver and copper alloy. In Darken Gurry's experiment, the plot indeed doesn't work in silver system. It was attributed to the use of unapproved electronegativity values of silver [39]. However, Gschneidner argued that even if approved electronegativity values of silver were used, the result would still be poor [40]. Yoshiharu Mae found that the atomic radius and electronegativity are not independent factors of each other [41]. In the description of solubility, the atomic radius is effective to the solubility of elements, but the effect of electronegativity is covered by the effect of the atomic radius. This is the reason why the Darken-Gurry plot does not clearly show the solubility of elements in metals. Mizutani et al. examined the H-R stabilization mechanism holds across whole solid solution ranges in series of gamma-brasses except Co-Zn gamma-brasses [42,43].

In total, we applied the fuzzy learning method to mine the fuzzy rules. We compare the fuzzy rules with the H-R rules and indeed find that there is something consistent and inconsistent with the H-R rules. It means that the fuzzy learning method could find structure-property linkage of alloys that Hume Rothery discovered before.

3.3. Implement of the Fuzzy-SVM method

For the purpose of applying the Fuzzy-SVM algorithm, all codes were implemented in python with a python's machine learning library called scikit-learn. The mean absolute error (MSE), the root-mean-square-error (RMSE), the maximum absolute error (MaxAE) is used to evaluate the predictive accuracy of the Fuzzy-SVM model. Let n represents the size of the data set, y and y' represent the actual and predicted values over the dataset. The mathematical expressions of the error measurement for the model estimate are given in Table 1.

Adding the membership value as extra columns to the input matrix to form a new variable. The kernel parameters of SVM are determined by grid search. Four different kernel functions are applied to this model, i.e. the Fuzzy-SVM-Lin, which uses a lineal kernel function; the Fuzzy-SVM-Pol, which uses a Polynomial kernel function; the Fuzzy-SVM-Gauss, which uses a radial basis function; the Fuzzy-SVM-Sig, which uses a sigmoid kernel function. The different types of kernel function used in SVM are given below [44]:

$$\text{Fuzzy - SVM - Lin: Fuzzy - SVM - Lin: } k(x, x') = x, x'$$

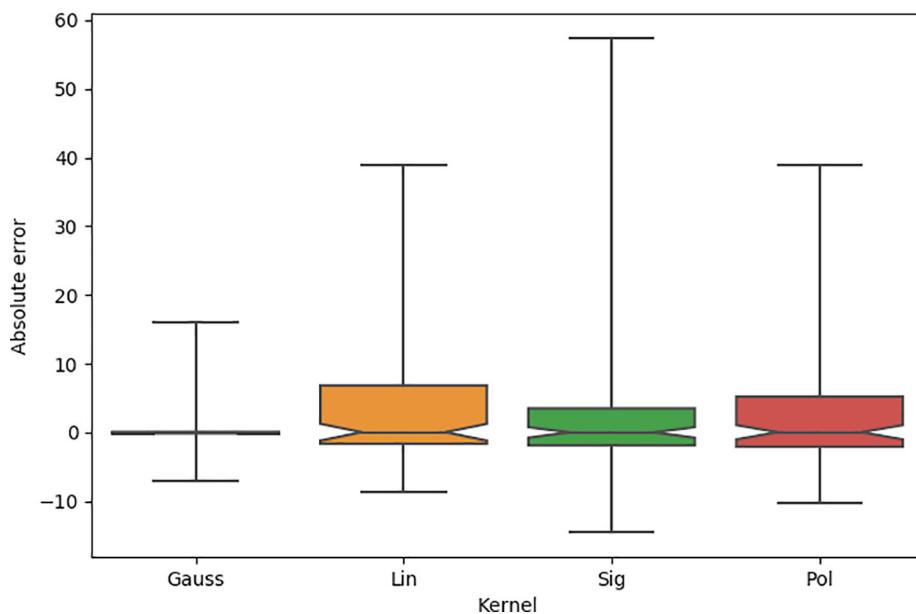


Fig. 3. Notched box plots of real solubility and predicted solubility using Fuzzy-SVM with different kernel functions in training process. The box shows the quartiles of the absolute error while the whiskers extend to show the rest of the distribution. The absolute error distributions of Fuzzy-SVM-Gauss are concentrated around zero and have the lowest range among other methods.

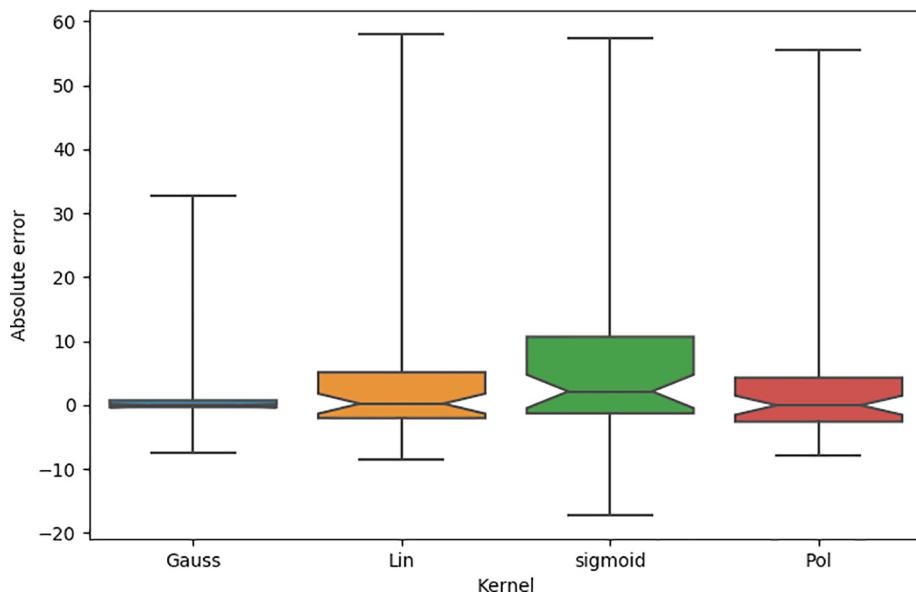


Fig. 4. Notched box plot of real solubility and predicted solubility using Fuzzy-SVM with different kernel functions in inference process. The box shows the quartiles of the absolute error while the whiskers extend to show the rest of the distribution. The Fuzzy-SVM-Gauss is the one that predict the solid solubility with the highest accuracy.

Fuzzy – SVM – Pol: $k(x, x') = (\gamma x, x' + r)^d$

Fuzzy – SVM – Gauss: $k(x, x') = \exp(-\gamma x - x'^2)$

Fuzzy – SVM – Sig: $\tanh(\gamma x, x' + r)$

The Fuzzy-SVM method relies on the selection of hyperparameters. There are two parts hyperparameters:(1) The degree of fuzziness(m) and the number of clusters(c) owing to the FCM algorithm. (2) a proper kernel function and kernel function's parameters owing to SVM algorithm. By testing four different kernel functions and using grid search to determining the kernel function's parameters, the Fuzzy-SVM-Gauss have the minimum statistical error in training and inference process.

The given dataset is divided into two categories: the training set and validation set (7:3), and the ten-fold cross validation will be used in the specific experiments. Firstly, the methods are conducted by training dataset based on Fuzzy-SVM-Lin, Fuzzy-SVM-Pol, Fuzzy-SVM-Gauss, Fuzzy-SVM-Sig. The calculated statistical errors based on Fuzzy-SVM are shown in Table 2. The errors indicated the prediction accuracy of different methods. From Table 2, it can be seen that Fuzzy-SVM with gaussian kernel function's RMSE, MAE and MaxAE outperforms the

other methods. It means that Fuzzy-SVM with gaussian kernel function has better prediction accuracy than others. Once the prediction method is trained, the validation dataset is used to test the generation performance. The validation dataset is given as inference process to the trained method. The generated results are then used to compute statistic error. Table 3 shows the result in inference process and it can be seen that Fuzzy-SVM with gaussian function has a good performance in both training and inference process.

For the predicted value y' obtained in training and inference process, a notched box plot is used to evaluate method's performance. It is observed from Fig. 3 that the absolute error distributions of Fuzzy-SVM-Gauss are concentrated around zero and have the lowest range among these methods. For the training dataset, the Fuzzy-SVM-Gauss method is able to handle the prediction of solid solubility accurately. And we also show the error distributions for the validation dataset in Fig. 4. The Fuzzy-SVM-sigmoid method has the widest error distribution and the Fuzzy-SVM-Gauss is the one that predict the solid solubility with the highest accuracy. From the training and validation dataset, although the Fuzzy-SVM-sigmoid perform good in training process, but as for

validation dataset the accuracy of the method is not well. The Fuzzy-SVM-Gauss performs well in both the training and inference process.

4. Conclusions

In this work, a fuzzy learning method with gaussian function is applied in alloy dataset to uncover fuzzy rules. Then a hybrid fuzzy learning method with SVM is proposed to predict solid solution. The result suggests that the H-R rules' 15% rule is suitable for both silver and copper alloy, but the Darken and Gurry plot is not. Fuzzy-SVM with gaussian kernel could give a relative estimate of solid solubility. In this manuscript, the machine learning method, which based on the fuzzy learning, not only discover the relationship between structure and properties, but also make accurate predictions on properties. It means that the method is effective to find the relationship between the structure and property. The results indicate that machine learning methods can be applied to discover structure-property linkage that was found relying on the materials scientists with domain knowledge previously.

5. Data availability

The raw/processed data required to reproduce these findings cannot be shared at this time as the data also forms part of an ongoing study.

CRediT authorship contribution statement

Huiran Zhang: Conceptualization, Supervision, Formal analysis, Writing - review & editing. **Gaofeng Zhou:** Data curation, Formal analysis, Methodology, Writing - original draft. **Shengzhou Li:** Data curation, Writing - review & editing. **Xingyue Fan:** Writing - review & editing. **Zhitong Guo:** Writing - review & editing. **Tao Xu:** Writing - review & editing. **Yan Xu:** Writing - review & editing. **Xue Chen:** Writing - review & editing. **Dongbo Dai:** Writing - review & editing. **Quan Qian:** Funding acquisition, Project administration.

Acknowledgement

This work was supported by the National Key Research and Development Program of China (No.2018YFB0704400).

References

- [1] J.A. Alonso, S. Simozar, Prediction of solid solubility in alloys, *Phys. Rev. B* 22 (1980) 5583–5589. <https://doi.org/10.1103/PhysRevB.22.5583>.
- [2] W. Hume-Rothery, G.W. Mabbott, K.M.C. Evans, The freezing points, melting points, and solid solubility limits of the alloys of silver, and copper with the elements of the B sub-groups, *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 233 (1934) 1–97.
- [3] W. Hume-rothery, G.F. Lewin, P.W. Reynolds, Peter William Reynolds, The lattice spacings of certain primary solid solutions in silver and copper, *Proc. R. Soc. London. Ser. A - Math. Phys. Sci.* 157 (2006) 167–183, <https://doi.org/10.1098/rspa.1936.0186>.
- [4] R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, C. Kim, Machine learning in materials informatics: recent applications and prospects, *NPJ Comput. Mater.* 3 (2017), <https://doi.org/10.1038/s41524-017-0056-5>.
- [5] S.M. Moosavi, A. Chidambaram, L. Talirz, M. Haranczyk, K.C. Stylianou, B. Smit, Capturing chemical intuition in synthesis of metal-organic frameworks, *Nat. Commun.* 10 (2019) 1–7, <https://doi.org/10.1038/s41467-019-08483-9>.
- [6] K. Gubaev, E.V. Podryabinkin, G.L.W. Hart, A.V. Shapeev, Accelerating high-throughput searches for new alloys with active learning of interatomic potentials, *Comput. Mater. Sci.* 156 148–156 (2019), <https://doi.org/10.1016/j.commatsci.2018.09.031>.
- [7] C. Wolverton, G.R. Jelbert, S.C. O’Keeffe, J. Stevick, M. Aykol, L. Ward, A machine learning approach for engineering bulk metallic glass alloys, *Acta Mater.* 159 (2018) 102–111, <https://doi.org/10.1016/j.actamat.2018.08.002>.
- [8] A.E. Teschendorff, Avoiding common pitfalls in machine learning omic data science, *Nat. Mater.* (2018), <https://doi.org/10.1038/s41563-018-0241-z>.
- [9] P. Raccuglia, K.C. Elbert, P.D.F. Adler, C. Falk, M.B. Wenley, A. Mollo, M. Zeller, S.A. Friedler, J. Schrier, A.J. Norquist, Machine-learning-assisted materials discovery using failed experiments, *Nature* 533 (2016) 73–76, <https://doi.org/10.1038/nature17439>.
- [10] J. Im, S. Lee, T.-W. Ko, H.W. Kim, Y. Hyon, H. Chang, Identifying Pb-free perovskites for solar cells by machine learning, *Npj Comput. Mater.* 5 (2019) 37, <https://doi.org/10.1038/s41524-019-0177-0>.
- [11] A.M. Gopakumar, P.V. Balachandran, D. Xue, J.E. Gubernatis, T. Lookman, Multi-objective optimization for materials discovery via adaptive design, *Sci. Rep.* 8 (2018) 1–12, <https://doi.org/10.1038/s41598-018-21936-3>.
- [12] K. Kim, L. Ward, J. He, A. Krishna, A. Agrawal, C. Wolverton, Machine-learning-accelerated high-throughput materials screening: discovery of novel quaternary Heusler compounds, *Phys. Rev. Mater.* 2 (2018) 1–9, <https://doi.org/10.1103/PhysRevMaterials.2.123801>.
- [13] S. Lu, Q. Zhou, Y. Ouyang, Y. Guo, Q. Li, J. Wang, Accelerated discovery of stable lead-free hybrid organic-inorganic perovskites via machine learning, *Nat. Commun.* 9 (2018) 3405, <https://doi.org/10.1038/s41467-018-05761-w>.
- [14] G. Pilania, X.-Y. Liu, Z. Wang, Data-enabled structure–property mappings for lanthanide-activated inorganic scintillators, *J. Mater. Sci.* 54 (2019) 8361–8380, <https://doi.org/10.1007/s10853-019-03434-7>.
- [15] Ç. Odabaşı, R. Yıldırım, Performance analysis of perovskite solar cells in 2013–2018 using machine-learning tools, *Nano Energy* 56 (2019) 770–791, <https://doi.org/10.1016/J.NANOEN.2018.11.069>.
- [16] Y. Ouyang, Z. Zhang, D. Li, J. Chen, G. Zhang, Emerging theory, materials, and screening methods: new opportunities for promoting thermoelectric performance, *Ann. Phys.* 531 (2019) 1800437, <https://doi.org/10.1002/andp.201800437>.
- [17] J. Lee, A. Seko, K. Shitara, K. Nakayama, I. Tanaka, Prediction model of band gap for inorganic compounds by combination of density functional theory calculations and machine learning techniques, *Phys. Rev. B* 93 (2016) 115104, , <https://doi.org/10.1103/PhysRevB.93.115104>.
- [18] H. Ran Zhang, Y. Zhang, D. Bo Dai, M. Cao, W. Feng Shen, Modelling and optimization of the superconducting transition temperature, *Mater. Des.* 92 (2016) 371–377, <https://doi.org/10.1016/j.matdes.2015.12.081>.
- [19] Y. Zhang, S. Yang, K.U. Leuven, Y.M. Zhang, R. Ubic, D.F. Xue, S. Yang, Predicting the Structural Stability and Formability of ABO₃-Type Perovskite Compounds Using Artificial Neural Networks Improved control of the AM process by in-process monitoring and feedback View project Liquid micro-scale chip cooling solutio, 1 (2012) 1–8. doi:10.1166/mat.2012.1009.
- [20] J.E. Gubernatis, T. Lookman, Machine learning in materials design and discovery: examples from the present and suggestions for the future, *Phys. Rev. Mater.* 2 (2018) 1–15, <https://doi.org/10.1103/PhysRevMaterials.2.120301>.
- [21] D.F. Urban, J.J. Möller, C. Elsässer, G. Krugel, W. Körner, Compositional optimization of hard-magnetic phases with machine-learning models, *Acta Mater.* 153 (2018) 53–61, <https://doi.org/10.1016/j.actamat.2018.03.051>.
- [22] Y.M. Zhang, S. Yang, J.R.G. Evans, Revisiting Hume-Rothery’s Rules with artificial neural networks, *Acta Mater.* 56 (2008) 1094–1105, <https://doi.org/10.1016/j.actamat.2007.10.059>.
- [23] M. Gevrey, I. Dimopoulos, S. Lek, Review and comparison of methods to study the contribution of variables in artificial neural network models, *Ecol. Modell.* 160 (2003) 249–264, [https://doi.org/10.1016/S0304-3880\(02\)00257-0](https://doi.org/10.1016/S0304-3880(02)00257-0).
- [24] S. Li, H. Zhang, D. Dai, G. Ding, X. Wei, Y. Guo, Study on the factors affecting solid solubility in binary alloys: an exploration by Machine Learning, *J. Alloys Compd.* 782 (2019) 110–118, <https://doi.org/10.1016/j.jallcom.2018.12.136>.
- [25] L.-X. Wang, J.M. Mendel, Generating fuzzy rules by learning from examples, *IEEE Trans. Syst. Man. Cybern.* 22 (1992) 1414–1427, <https://doi.org/10.1109/21.199466>.
- [26] L.A. Zadeh, Toward a Generalized Theory of Uncertainty (GTU) - An outline, *2005 IEEE Int. Conf. Granul. Comput.* 2005 (2005) 16. doi: 10.1109/GRC.2005.1547227.
- [27] A. Çelikyilmaz, I. Burhan Türkşen, Fuzzy functions with support vector machines, *Inf. Sci. (Ny)* 177 (2007) 5163–5177, <https://doi.org/10.1016/j.ins.2007.06.022>.
- [28] T.J. Ross, *Fuzzy Logic with Engineering Applications*, John Wiley & Sons Ltd, Chichester, UK, 2010 doi:10.1002/978111994374.
- [29] H. Drucker, C.J.C. Burges, L. Kaufman, A. Smola*, V. Vapnik, Support Vector Regression Machines.
- [30] A.J. Smola, B. Schölkopf, A tutorial on support vector regression, *Stat. Comput.* 14 (2004) 199–222, <https://doi.org/10.1023/B:STCO.0000035301.49549.88>.
- [31] W. Hume-Rothery, Elements of structural metallurgy, 1961.
- [32] G. Ouyang, X. Tan, C.X. Wang, G.W. Yang, Solid solubility limit in alloying nanoparticles, *Nanotechnology* 17 (2006) 4257–4262, <https://doi.org/10.1088/0957-4484/17/16/042>.
- [33] Y.P. Wang, B.S. Li, H.Z. Fu, Solid solution or intermetallics in a high-entropy alloy, *Adv. Eng. Mater.* 11 (2009) 641–644, <https://doi.org/10.1002/adem.200900057>.
- [34] M. Stiehler, J. Rauchhaupt, U. Giegegack, P. Häussler, On modifications of the well-known hume-rothery rules: amorphous alloys as model systems, *J. Non. Cryst. Solids.* 353 (2007) 1886–1891, <https://doi.org/10.1016/j.jnoncrysol.2007.01.052>.
- [35] X. Yang, Y. Zhang, Prediction of high-entropy stabilized solid-solution in multicomponent alloys, *Mater. Chem. Phys.* 132 (2012) 233–238, <https://doi.org/10.1016/j.matchemphys.2011.11.021>.
- [36] Q.F. He, Y.F. Ye, Y. Yang, Formation of random solid solution in multicomponent alloys: from hume-rothery rules to entropic stabilization, *J. Phase Equilibria Diffus.* 38 (2017) 416–425, <https://doi.org/10.1007/s11669-017-0560-9>.
- [37] A.-P. Tsai, A test of hume-rothery rules for stable quasicrystals, *J. Non. Cryst. Solids* 334–335 (2004) 317–322, <https://doi.org/10.1016/j.jnoncrysol.2003.11.065>.
- [38] M. Calvo-Dahlborg, S.G.R. Brown, Hume-Rothery for HEA classification and self-organizing map for phases and properties prediction, *J. Alloys Compd.* 724 (2017) 353–364, <https://doi.org/10.1016/j.jallcom.2017.07.074>.
- [39] Y.M. Zhang, J.R.G. Evans, S. Yang, The prediction of solid solubility of alloys: developments and applications of Hume-Rothery’s rules, *J. Cryst. Phys. Chem.* 1 (2010) 81–97, <https://doi.org/10.1002/chin.201335198>.
- [40] K.A. Gschneidner Jr, LS (Larry) Darken’s contributions to the theory of alloy

- formation and where we are today, 1979.
- [41] Y. Mae, What the darken-gurry plot means about the solubility of elements in metals, *Metall. Mater. Trans. A Phys. Metall. Mater. Sci.* 47 (2016) 6498–6506, <https://doi.org/10.1007/s11661-016-3730-1>.
- [42] U. Mizutani, T. Noritake, T. Ohsuna, T. Takeuchi, Hume-Rothery electron concentration rule across a whole solid solution range in a series of gamma-brasses in Cu-Zn, Cu-Cd, Cu-Al, Cu-Ga, Ni-Zn and Co-Zn alloy systems, *Philos. Mag.* 90 (2010) 1985–2008, <https://doi.org/10.1080/14786430903246320>.
- [43] U. Mizutani, H. Sato, The physics of the hume-rothery electron concentration rule, *Crystals* 7 (2017) 9, <https://doi.org/10.3390/crust7010009>.
- [44] R.G. Brereton, G.R. Lloyd, Support vector machines for classification and regression, *Analyst* 135 (2010) 230–267, <https://doi.org/10.1039/B918972F>.