



Phase prediction of Ni-base superalloys via high-throughput experiments and machine learning

Zijun Qin, Zi Wang, Yunqiang Wang, Lina Zhang, Weifu Li, Jin Liu, Zexin Wang, Zihang Li, Jun Pan, Lei Zhao, Feng Liu, Liming Tan, Jianxin Wang, Hua Han, Liang Jiang & Yong Liu

To cite this article: Zijun Qin, Zi Wang, Yunqiang Wang, Lina Zhang, Weifu Li, Jin Liu, Zexin Wang, Zihang Li, Jun Pan, Lei Zhao, Feng Liu, Liming Tan, Jianxin Wang, Hua Han, Liang Jiang & Yong Liu (2021) Phase prediction of Ni-base superalloys via high-throughput experiments and machine learning, Materials Research Letters, 9:1, 32-40, DOI: [10.1080/21663831.2020.1815093](https://doi.org/10.1080/21663831.2020.1815093)

To link to this article: <https://doi.org/10.1080/21663831.2020.1815093>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 27 Sep 2020.



[Submit your article to this journal](#)



Article views: 8149



[View related articles](#)



[View Crossmark data](#)



Citing articles: 49 [View citing articles](#)



ORIGINAL REPORT



Phase prediction of Ni-base superalloys via high-throughput experiments and machine learning

Zijun Qin^{a*}, Zi Wang^{a*}, Yunqiang Wang^b, Lina Zhang^c, Weifu Li^d, Jin Liu^b, Zexin Wang^a, Zihang Li^a, Jun Pan^a, Lei Zhao^e, Feng Liu^a, Liming Tan^f, Jianxin Wang^b, Hua Han^c, Liang Jiang^g and Yong Liu^a

^aThe State Key Laboratory of Powder Metallurgy, Central South University, Changsha, People's Republic of China; ^bSchool of Computer Science and Engineering, Central South University, Changsha, People's Republic of China; ^cNational Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, People's Republic of China; ^dCollege of Science, Huazhong Agricultural University, Wuhan, People's Republic of China; ^eBeijing Key Laboratory of Metal Material Characterization, China Iron & Steel Research Institute, Beijing, People's Republic of China; ^fSchool of Metallurgy and Environment, Central South University, Changsha, People's Republic of China; ^gInstitute for Advanced Studies in Precision Materials, Yantai University, Yantai, People's Republic of China

ABSTRACT

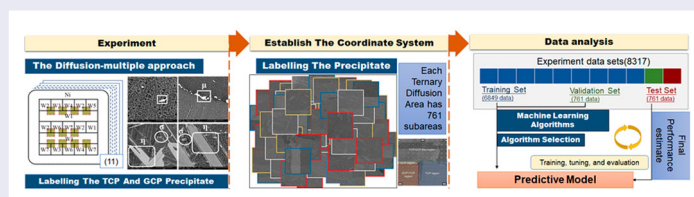
Predicting the phase precipitation of multicomponent alloys, especially the Ni-base superalloys, is a difficult task. In this work, we introduced a dependable and efficient way to establish the relationship between composition and detrimental phases in Ni-base superalloys, by integrating high throughput experiments and machine learning algorithms. 8371 sets of data about composition and phase information were obtained rapidly, and analyzed by machine learning to establish a high-confidence phase prediction model. Compared with the traditional methods, the proposed approach has remarkable advantage in acquiring and analyzing the experimental data, which can also be applied to other multicomponent alloys.

ARTICLE HISTORY

Received 8 June 2020

KEYWORDS

Superalloy; high-throughput experiments; machine learning; diffusion multiple; phase selection



IMPACT STATEMENT

By integrating the high throughput experiments and machine learning algorithms, it is hopeful to facilitate the design of new Ni-base superalloys, and even other multicomponent alloys.

1. Introduction

The multicomponent alloys including Ti-base [1], Ni-base [2,3], Fe-base [4], Al-base [5], and high entropy alloys [6], are composed of five or more metallic elements, which makes it increasingly complex to study the relationships among composition, phases, process, and properties. For example, the Ni-base superalloys, used for turbine blades and disks, generally contain over eight alloying elements, the phase precipitation and mechanical properties are found very sensitive to the composition variation [3,7]. Therefore, there were great efforts on the

composition design of superalloys over the past decades [8,9]. The phases of the Ni-base superalloys generally include γ matrix, γ' phase, topologically close-packed (TCP) phases, and geometrically close-packed (GCP) phases [10,11]. Wherein the TCP and GCP phases are always thought detrimental to the alloys, as their precipitation would deplete the strengthening elements, and facilitate the crack initiation [10,12]. Thereby, the TCP and GCP phases should be avoided during the alloy design, and reliable prediction of these phases becomes indispensable.

CONTACT Feng Liu liufeng@csu.edu.cn The State Key Laboratory of Powder Metallurgy, Central South University, Changsha 410083, People's Republic of China; Liming Tan limingtan@csu.edu.cn School of Metallurgy and Environment, Central South University, Changsha 410083, People's Republic of China

* These authors contributed equally to this work: Zijun Qin, Zi Wang.

Supplemental data for this article can be accessed here. <https://doi.org/10.1080/21663831.2020.1815093>

At the very initial stage, the valence electron concentration was used to predict the formation of the TCP phases in simple binary and ternary systems [13,14]. Thereafter, phase computation, known as PHACOMP [15] and newPHACOMP [16], had been proposed successively. But those methods were unable to accurately predict the precipitation behavior of new generation Ni-base superalloys. So far, the calculation of phase diagram (CALPHAD) is thought a powerful method to bridge the composition and equilibrium phases. As CALPHAD is generally based on the database derived from simple systems within certain composition ranges, the accuracy of TCP and GCP phases prediction in complex systems or outside domains needs further verification.

With the continuous progress in materials design, more statistical methods, especially the machine learning (ML), have been adopted [17,18], providing new hints for exploring multicomponent superalloys. However, as adequate data ranging from tens to over hundreds are needed to train the ML algorithm, enough experimental results are necessary, which is always time-consuming for traditional methods. Hence a high-throughput experiment combining with superalloys diffusion-multiples and automatic composition-phase identification is introduced here. In specific, the diffusion-multiples yield tremendous alloys with composition gradients in large ranges. Aided with advanced microdomain detection technologies, the site dependent composition and phases could be automatically detected. Thus, a reliable experimental database has been established for ML to train phase precipitation models and composition design of superalloys, which can also be broadly applied to other multicomponent alloys.

2. Materials and methods

2.1. Materials and sample preparation

Several superalloys were utilized to assemble a diffusion multiple contain complex compositions, as shown in Table 1. Alloy compositions of W2 to W7 were adjusted based on W1 by substituting Co, W, Mo, Ti, Ta, Nb for Ni respectively. Those alloys were prepared by vacuum induction melting method firstly and homogenized by

heat treatment. Then alloy W1-W7 were arranged and assembled as shown in Figure 1. After hot isostatic pressing at 1180°C/150 MPa for 8 h and two-stage aging process of 1180°C / 1000 h plus 800°C / 2000h, the alloys with different components were closely bonded and fully diffused to form compositional and microstructural gradient.

2.2. Composition identification and microstructure characterization

Micro-XRF (X-ray fluorescence) equipment BRUKER M4 TORNADO was used to automatically and continuously identify the elemental composition at different regions. SEM (scanning electron microscope) was performed by ZEISS SUPRA 55 equipped on an automated photographic platform.

The bulk samples contained 11 diffusion multiples totally, and each diffusion multiple was in a size of $2000 \times 2000 \mu\text{m}$ and divided into 761 sections. Above mentioned division method was shown in Figure 1. Considering the composition changes of the central sections were more significant than these of edge regions, denser data points were collected therein. In specific, we firstly established a rectangular coordinate system based on the actual dimensions of each diffusion multiple, setting the ranges of x and y axes as 0–2000; secondly, set coordinate points in step of 50 nm within $500 \leq x \leq 1500$ and $500 \leq y \leq 1500$, and the square area of $50 \times 50 \text{ nm}$ was the feature region of corresponding center coordinate point. In other areas outside the core domains, the step was set as 100 nm, hence the square area of $100 \times 100 \text{ nm}$ was the feature region of the center point. By marking the existence of TCP and GCP in each area, whole 8371 pieces of data containing alloy composition and these phases were obtained.

2.3. Machinal learning

According to the NFL(No Free Lunch) theorem [19], different machine learning algorithms own respective advantages and disadvantages, the following commonly

Table 1. Measured composition of multicomponent alloys in wt.%.

	Co	W	Mo	Cr	Al	Ti	Ta	Nb	Hf	Ni
W1	13.0	3.0	2.9	12.2	3.03	3.96	3.1	–	0.21	Bal.
W2	27.8	3.0	2.9	12.0	3.05	4.06	3.0	–	0.21	Bal.
W3	13.0	6.1	2.9	11.8	3.05	3.92	3.0	–	0.22	Bal.
W4	13.1	3.0	6.0	12.1	3.05	4.08	3.1	–	0.19	Bal.
W5	13.0	2.9	3.0	12.0	3.07	6.01	3.0	–	0.19	Bal.
W6	13.0	3.0	2.9	12.0	3.10	4.04	8.1	–	0.20	Bal.
W7	13.0	3.0	3.0	11.9	2.98	4.12	3.0	4.0	0.22	Bal.

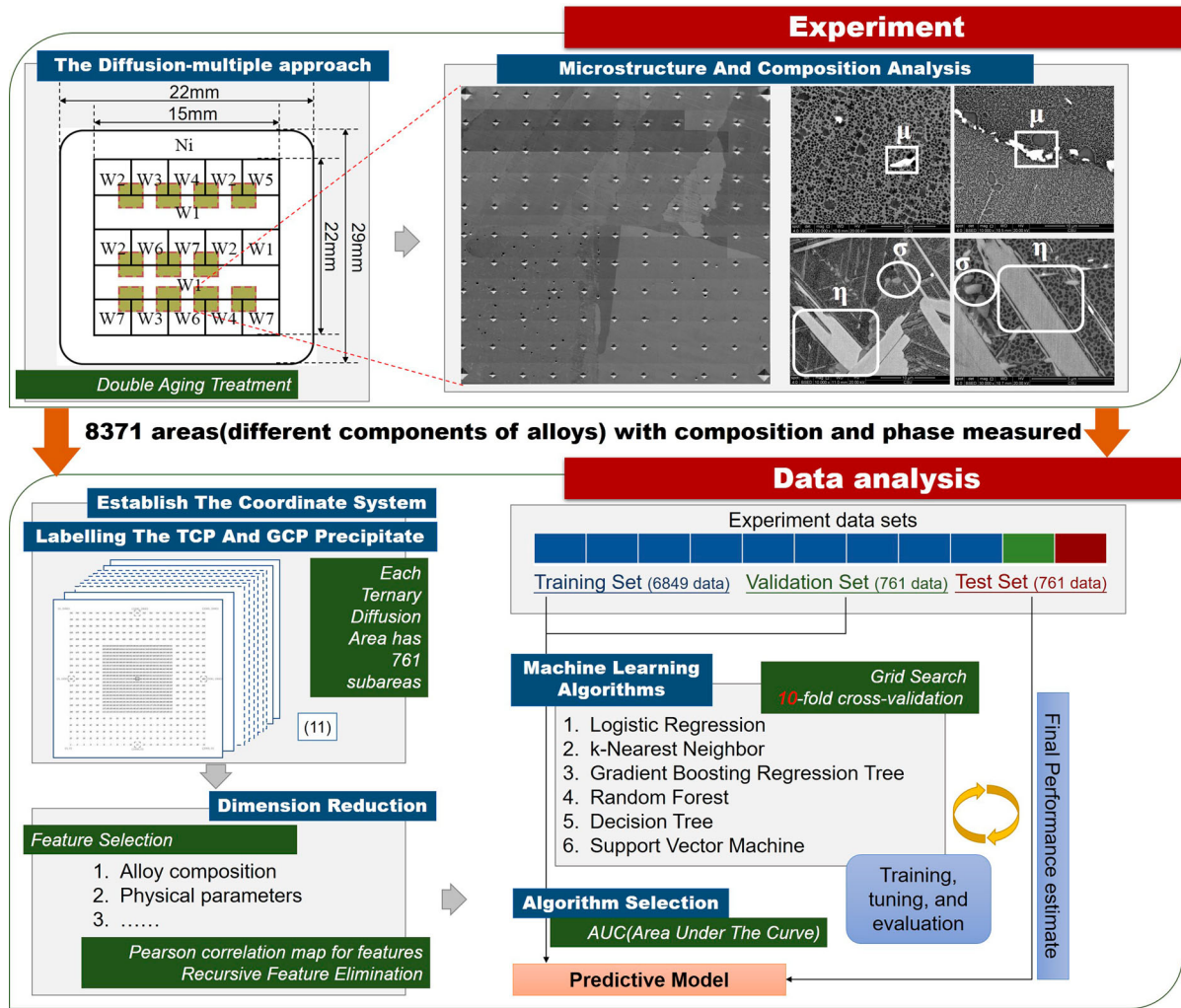


Figure 1. Schematic workflow of the proposed method.

used six different machine learning classification algorithms have been used here: logistic regression (LR), k-nearest neighbor (KNN), gradient boosting regression tree (GBRT), random forest (RF), decision tree (DT), and support vector machine (SVM). The predictability of these algorithms model was judged by comparing area under curve (AUC), as adopted by other authors [20,21]. Before modeling, the features were filtered by Pearson correlation and recursive feature elimination. In order to avoid overfitting of the model and evaluate the feasibility of the models, 761 pieces of data detected from NiX-W-Nb diffusion multiple were picked as a test set, and the remaining 7610 pieces of data were used for training, which were randomly divided into 10 parts, 9 of which were used as the trading set and 1 was used as the validation set. To acquire acceptable accuracy and stability, we applied 10-fold cross validation and a grid-search technique for the hyperparameters selection.

3. Results and discussion

3.1. Phase identification.

The microstructures of seven original alloys, W1-W7, are shown in Figure 2(A-L). The γ , γ' , σ , μ and η phases are the major phases in the diffusion multiple. In specific, the γ matrix is mainly strengthened by solid solution elements. The γ' phase is the $L1_2$ ordered phase based on Ni_3Al , and is the main strengthening phase in Ni-based superalloys. The σ and μ phases are TCP phases, and the η phase is a typical GCP phase. In terms of morphological features, μ and σ phases are mainly in punctate, acicular, irregular, and discontinuous shapes, and the η phase is generally in shapes of large flake and lathy needle. Basically, γ , γ' and GCP in the alloy could be distinguished from the typical morphologies. As both μ and σ phases are harmful and in similar geometric features, these two phases are unified into TCP phases.

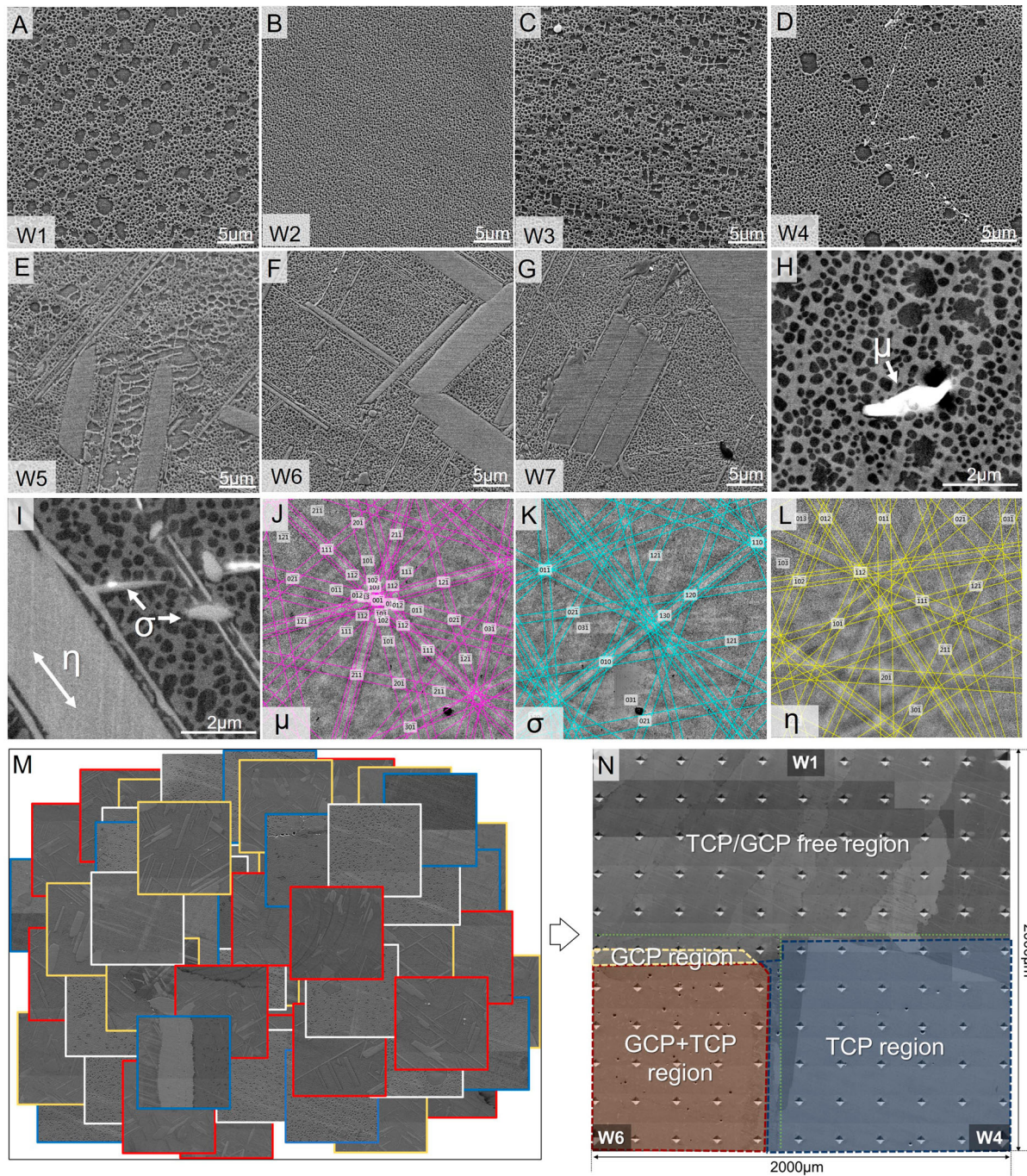


Figure 2. (A-G) SEM images of W1-W7 alloys; (H, I) the magnified images of σ , μ and η phases, respectively; (J, K, L) the electron backscattered diffraction Kikuchi patterns of μ , σ , and η , respectively; (M) labeled and (N) montage SEM images of NiX-6W-4Nb diffusion triple.

In the light of the distributions of TCP and GCP phases, the microstructural images of 761 sections were classified and matched. As shown in Figure 2(M,N), the GCP and TCP phases simultaneously precipitate at the areas marked by red frames; yellow and blue frames correspond to GCP and TCP phase precipitated regions, respectively.

Moreover, the white frame illustrates that there is no harmful phase. Thereby, the distribution of precipitates in the diffusion multiple could be identified directly after the labeled images were integrated into the whole NiX-6W-4Nb diffusion triple. To obtain adequate data source, this labeling method was also applied for the other 10 diffusion triples.

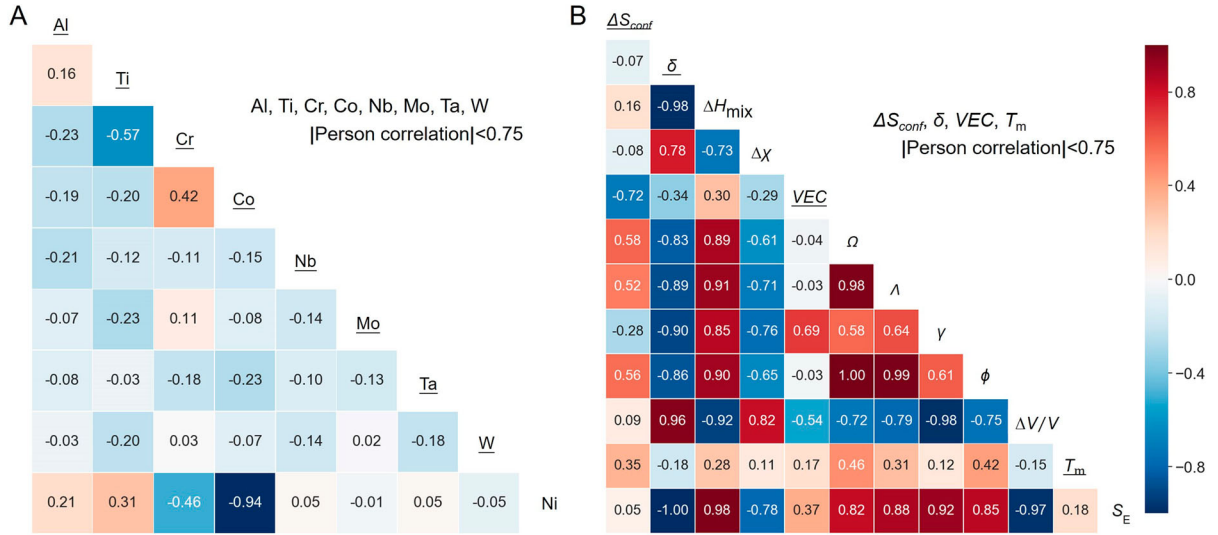


Figure 3. Pearson correlation map for features of (A) alloy composition, and (B) 12 physical parameters.

Table 2. Physic parameters.

	Symbol & Formula	Meaning & Parameter	Ref.
1	$\Delta S_{conf} = -R \sum_{i=1}^n c_i \ln c_i$	Structural entropy. R , gas constant, 8.314 J/(mol·K); c_i , mole fraction of the i th element.	[22]
2	$\delta = 100 \times \sqrt{\sum_{i=1}^n c_i \left(1 - \frac{r_i}{\bar{r}}\right)^2}$	Atomic size difference in multi-component alloys $\bar{r} = \left(\sum_{i=1}^n c_i r_i\right)$, average atomic radius; r_i , atomic radius of the i th element.	[23,24]
3	$\Delta H_{mix} = \sum_{i=1, i \neq j}^n \Omega_{ij} c_i c_j$	Mixing enthalpy of a solid solution $\Omega_{ij} (= 4\Delta H_{AB}^{mix})$, regular melt-interaction parameter between i th and j th elements; ΔH_{AB}^{mix} , mixing enthalpy of binary liquid alloy.	[25]
4	$\Delta \chi = \sqrt{\sum_{i=1}^n c_i (\chi_i - \bar{\chi})^2}$	Electronegativity difference between two alloy components $\bar{\chi} = \left(\sum_{i=1}^N c_i \chi_i\right)$, average electronegativity; χ_i , Pauling electronegativity for i th component.	[26]
5	$VEC = \sum_{i=1}^n c_i (VEC_i)$	Valence electron concentration VEC_i , valence electron concentration of i th.	[27]
6	$T_m = \sum_{i=1}^n c_i (T_m)_i$	Melting point of n -elements alloy, rule of mixtures $(T_m)_i$, melting point of i th.	[28]
7	$\Omega = \frac{T_m \Delta S_{conf}}{ \Delta H_{mix} }$	Geometrical parameter—predicting the solid-solution formation	[28]
8	$\Lambda = \frac{\Delta S_{conf}}{\delta^2}$	Purely geometrical parameter	[29]
9	$\gamma = \omega_S / \omega_L = \left[1 - \frac{[(r_S + \bar{r})^2 - \bar{r}^2] / (r_S + \bar{r})^2}{1 - \sqrt{[(r_L + \bar{r})^2 - \bar{r}^2] / (r_L + \bar{r})^2}} \right]$	The ratio between the solid angles of the smallest and largest atoms $\bar{r} = \left(\sum_{i=1}^N c_i r_i\right)$, average atomic radius; r_L and r_S are the radius of the largest and smallest atoms;	[30]
10	$S_E = S_E(r_i, c_i, \xi)$	The excessive entropy of mixing ξ , the atomic packing fraction of the n -element alloy range from 0.68 (BCC);	[31]
11	$\phi = \frac{S_C - S_H}{ S_E } = \frac{S_C - \Delta H_{mix} / T_m}{ S_E }$	A single dimensionless thermodynamic parameter S_C , the configurational entropy of mixing for an ideal gas.	[31]
12	$\overline{\Delta V / V} = \frac{\sum_{ij} c_i c_j V_i - V_j }{(V_i + V_j) / 2}$	The parameters of a novel two-dimensional structure map $V_i = \frac{4}{3} \pi r_i^3$, atomic percentage of the i th component;	[10]

3.2. Feature set determination

Selecting typical and moderate features is the key for ML, since superabundant features would increase the dimension of the feature space, raise the risk of overfitting, and even destruct the interpretability on results. Therefore, strong related features with absolute value of Pearson correlation coefficient over 0.75 were removed. Firstly, the elemental factors were considered. Because Ni-based superalloys are alloyed with many elements, the interactions among different elements are complicated. As shown in Figure 3(A), even if the strongly related Ni is

removed, the model is still high-dimensional and does not have acceptable generalization ability.

Thereafter, in order to further reduce the dimensions and improve the model interpretability, 12 physics parameters integrating element information were chosen as candidate features, as illustrated in Table 2. Based on Figure 3(B), after removing the significantly related parameters, four physical parameters (ΔS_{conf} , δ , VEC , and T_m) with less correlation can be retained as the features.

Furthermore, recursive feature elimination (RFE) was adopted to reduce feature dimensions. Specifically, RFE

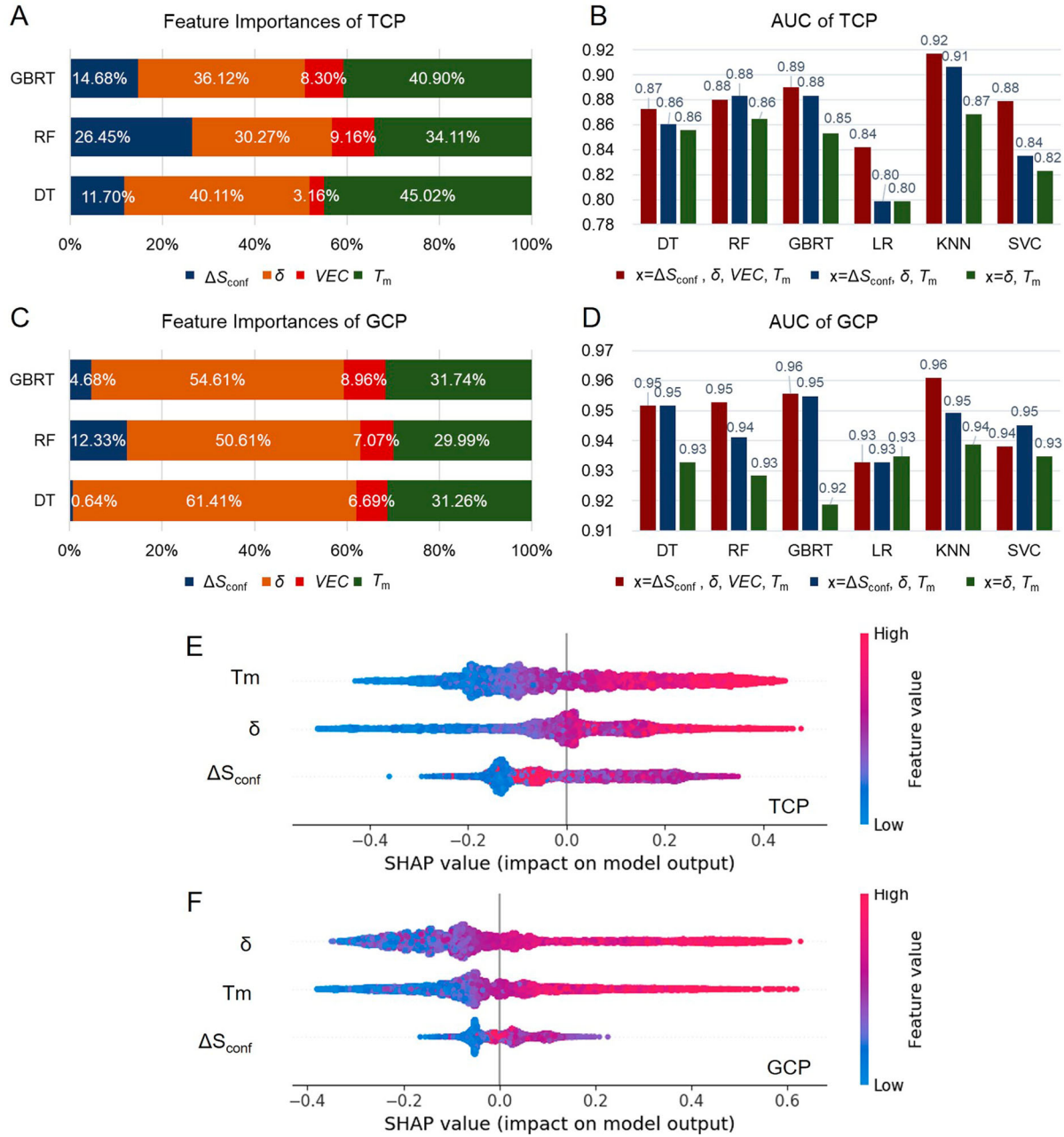


Figure 4. Recursive feature elimination results showing (A) the feature importance and (B) AUC value of TCP phases; (C) the feature importance and (D) AUC value of GCP illustrating predicted by models with different feature sets; the SHAP values of (E) TCP and (F) GCP.

was performed to obtain the importance of each feature through the attribution returned by the algorithm, and remove the weakly weighted features for next round training. According to the feature importance determined by three models of gradient boosting regression tree (GBRT), random forest (RF), and decision tree (DT), as illustrated in Figure 4(A,C), the VEC and ΔS_{conf} have relatively lower feature importance. Moreover, the area under curve (AUC) values of all models are illustrated in Figure 4(B,D). The overall prediction accuracy of these models stayed at high level when ΔS_{conf} , δ , and T_m are used as feature set X. While the model accuracy dropped sharply when both VEC and ΔS_{conf} are removed from the

feature set. Thereafter, the feature X including ΔS_{conf} , δ , and T_m was chosen.

To improve the interpretability of the model, the impact of each eigenvalue on the prediction results was evaluated by SHAP (SHapley Additive exPlanations) [32], as shown in Figure 4(E,F). The dots in the figure represent all data points in the data set. The dots in more reddish color own greater characteristic value, and the bluer the converse. Positive values represent the promotion of harmful phases, and negative values mean the suppression of harmful phases. It was detected that the lower ΔS_{conf} , δ , and T_m were beneficial to suppress the formation of harmful phases. The strengthening

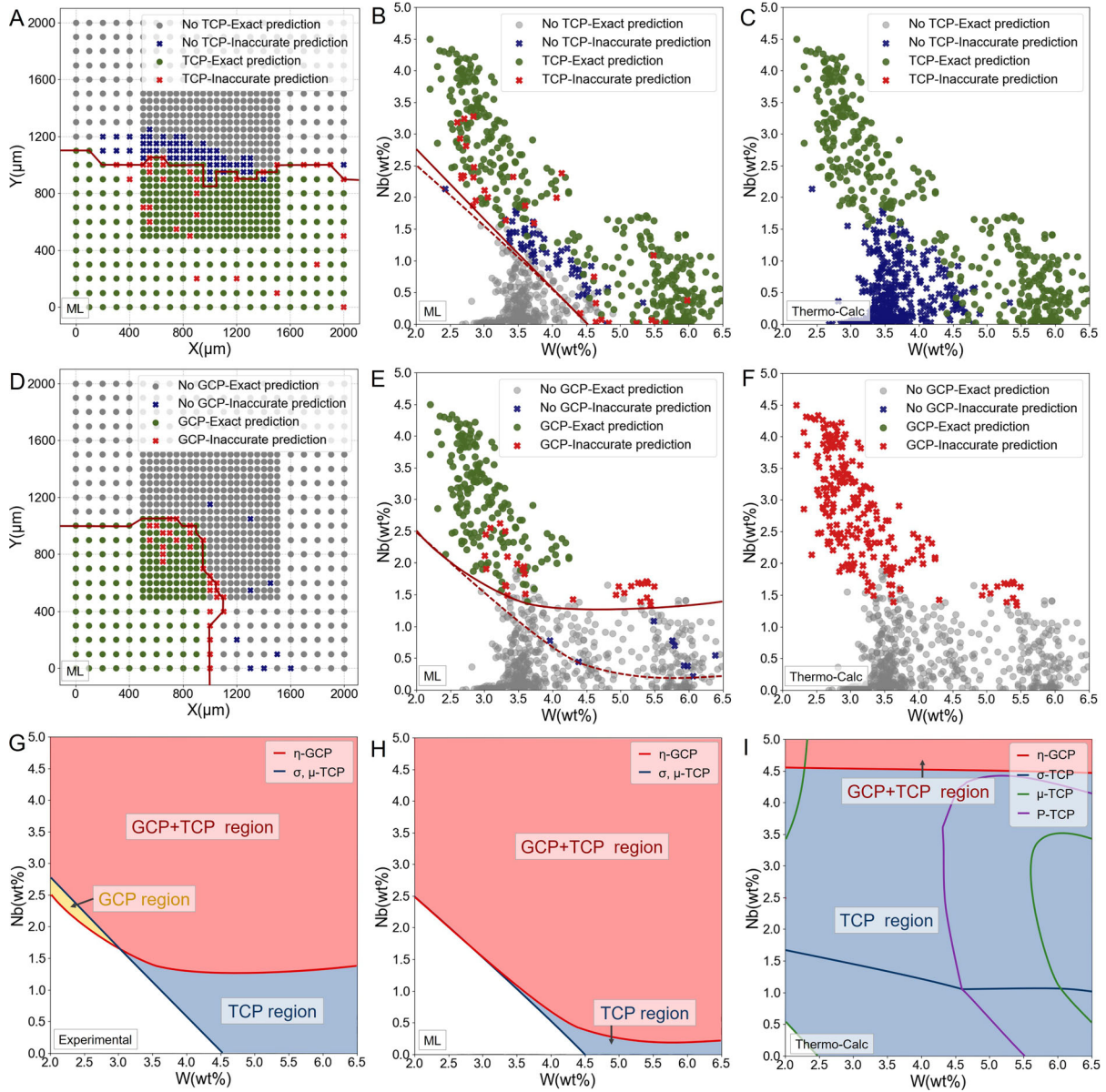


Figure 5. The predicted results showing (A) the distribution and (B) ZPF line of TCP phases by ML and (C) Thermo-Calc; (D) the distribution and (E) ZPF line of GGP phases by ML and (F) Thermo-Calc; and the phase diagram by (G) experimental, (H) ML and (I) Thermo-Calc in NiX-6W-4Nb diffusion triple.

effects of nickel-based superalloys are mainly solid solution strengthening, precipitation strengthening and grain boundary strengthening. The order of atomic size is: $\text{Ni} < \text{Co} < \text{Cr} < \text{Mo} < \text{W} < \text{Al} < \text{Ti} < \text{Nb} < \text{Ta}$, and the element with smaller atomic size difference from the Ni matrix is easier to solute into Ni. For Ni-base superalloys, Co, Cr, Mo and W are the major solid solution strengthening elements. Hence, when the δ is smaller, as the component atomic-size difference is relatively small, the component atoms easily substitute for each other to form solid solution. Meanwhile, the mixing enthalpy values is not negative enough for matrix to form a compound [23]. With low solid solubility in matrix, Al, Ti, Nb and Ta could form intermetallic compounds with Ni, such as γ' - $\text{Ni}_3(\text{Al}, \text{Ti})$, η - $\text{Ni}_3(\text{Al}, \text{Ti})_{0.5}\text{Nb}_{0.5}$ and so on. The γ' phase with low ΔS_{conf} is a long-range ordered face-centered cubic structure, which is the main precipitation strengthening phase. The η phase with high ΔS_{conf} is a close-packed hexagonal structure. Thereby, the overall ΔS_{conf} is smaller, less prone to form harmful phases. Basically, the addition of refractory elements including Cr, W, Ta, Mo and Nb can effectively improve the performance of the alloy, but excessive addition, which would cause the higher T_m , leads to the precipitation of harmful phases.

3.3. Validity of ML model

The algorithm with best general predictability on TCP and GCP phase precipitation was k-nearest neighbor (KNN), which was also selected for the following test. The predicted results of the test set (i.e. NiX-6W-4Nb diffusion triple) were compared with the experimental data to evaluate the feasibility of this ML method. In Figure 5, the green and red points represent the specific area where the harmful phase is accurately and inaccurately predicted respectively. The blue point means that the harmful phase does not precipitate experimentally but is predicted to be appeared, and the gray point manifests that KNN has exactly estimated that the harmful phase does not precipitate there. Additionally, the red line shows the Zero-Phase Fraction (ZPF) line, and the dotted line was the prediction ZPF line.

In general, using the physical parameters ΔS_{conf} , δ and T_m , the KNN model has accurately predicted the precipitation of harmful phases, as the AUC values of TCP and GCP reached 0.88 and 0.92, respectively. Compared with the phase diagram of 800°C calculated by Thermo-calc with thermodynamic database TTNI8, it could be found that the traditional CALPHAD calculation seriously underestimates the precipitation trend of GCP phase and overestimates the precipitation of TCP phase. Thereby, integrating the high throughput experiment and machine learning had better predictability.

Considering the training data has completely excluded NiX-6W-4Nb diffusion triple data, this prediction model had a strong generalization ability, and can be well applied to untrained elementary composition range. This method, can also be extended to predict the phase precipitation in other multicomponent alloys, such as iron steels, Ti alloy, and Al alloys.

4. Conclusions

In conclusion, we established a ML model to predict the relationship between composition and detrimental phases in multicomponent nickel-based superalloys, based on the database obtained from high throughput experiments.

This method used advanced detection equipment to automatically collect the composition and microstructure information of the alloy, and introduced the physic parameters as the feature set to reduces the dimension reasonably. Within six different machine learning algorithms, a trusted model with AUC values of 0.88 and 0.92 for TCP and GCP was finally screened. As a great challenge in alloy design lies in avoiding the precipitation of detrimental phases, this model with sound predictability offers a promising opportunity to find novel Ni-based superalloys within undiscovered composition ranges. Moreover, the method applied in this work can also be extensively used for other multicomponent alloys.

Acknowledgments

The authors acknowledge Prof. Ji-Cheng Zhao from University of Maryland and Prof. Zhanpeng Jin from Central South University for insightful discussions.

Data and materials availability

The data supporting the findings of this work is available in the main text. Raw data are available from the corresponding authors on reasonable request.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by The National Key Research and Development Program of China [grant number 2016YFB0701404]; The Natural Science Foundation of China [grant number 91860105]; The Project founded by China Postdoctoral Science Foundation [grant number 2019M662799].

ORCID

Zijun Qin  <http://orcid.org/0000-0003-0665-5924>

References

- [1] Alabort E, Barba D, Shagiev MR, et al. Alloys-by-design: application to titanium alloys for optimal superplasticity. *Acta Mater.* **2019**;178:275–287.
- [2] Smith T, Esser B, Antolin N, et al. Phase transformation strengthening of high-temperature superalloys. *Nat Commun.* **2016**;7(1):1–7.
- [3] Pollock TM. Alloy design for aircraft engines. *Nat Mater.* **2016**;15(8):809–815.
- [4] Almirall N, Wells P, Yamamoto T, et al. Precipitation and hardening in irradiated low alloy steels with a wide range of Ni and Mn compositions. *Acta Mater.* **2019**;179:119–128.
- [5] Sun W, Zhu Y, Marceau R, et al. Precipitation strengthening of aluminum alloys by room-temperature cyclic plasticity. *Science.* **2019**;363(6430):972–975.
- [6] Ding Q, Zhang Y, Chen X, et al. Tuning element distribution, structure and properties by composition in high-entropy alloys. *Nature.* **2019**;574(7777):223–227.
- [7] Wang Z, Zhang L, Li W, et al. High throughput experiment assisted discovery of new Ni-base superalloys. *Scr Mater.* **2020**;178:134–138.
- [8] Jones N, Christofidou K, Mignanelli P, et al. Influence of elevated Co and Ti levels on polycrystalline powder processed Ni-base superalloy. *Mater Sci Technol.* **2014**;30(15):1853–1861.
- [9] Christofidou KA, Hardy MC, Li H-Y, et al. On the effect of Nb on the microstructure and properties of next generation polycrystalline powder metallurgy Ni-based superalloys. *Metallurgical and Materials Transactions A.* **2018**;49(9):3896–3907.
- [10] Seiser B, Drautz R, Pettifor D. TCP phase predictions in Ni-based superalloys: structure maps revisited. *Acta Mater.* **2011**;59(2):749–763.
- [11] Antonov S, Detrois M, Isheim D, et al. Comparison of thermodynamic database models and APT data for strength modeling in high Nb content γ - γ' Ni-base superalloys. *Mater Des.* **2015**;86:649–655.
- [12] Rae CM, Reed RC. The precipitation of topologically close-packed phases in rhenium-containing superalloys. *Acta Mater.* **2001**;49(19):4113–4125.
- [13] Rideout S, Manly W, Kamen E, et al. Intermediate phases in ternary alloy systems of transition elements. *JOM.* **1951**;3(10):872–876.
- [14] Das D. Intermediate phases in the Mo-Fe-Co, Mo-Fe-Ni, and Mo-Ni-Co ternary systems. *rn.* **1952**;1:2–1.
- [15] Boesch WJ, Slaney JS. Preventing sigma phase embrittlement in nickel base superalloys. *Metal Progress.* **1964**;86(1):109–111.
- [16] Morinaga M, Yukawa N, Adachi H, et al. New PHA-COMP and its applications to alloy design. *Superalloys 1984.* **1984**: 523–532.
- [17] Butler KT, Davies DW, Cartwright H, et al. Machine learning for molecular and materials science. *Nature.* **2018**;559(7715):547–555.
- [18] Balachandran PV, Young J, Lookman T, et al. Learning from data to design functional materials without inversion symmetry. *Nat Commun.* **2017**;8(1):1–13.
- [19] Adam SP, Alexandropoulos S-AN, Pardalos PM, et al. No free lunch theorem: a review. *Approximation and optimization.* Cham, Switzerland: Springer; **2019**. p. 57–82.
- [20] Choubin B, Borji M, Mosavi A, et al. Snow avalanche hazard prediction using machine learning methods. *J Hydrol.* **2019**;577:123929.
- [21] Min X, Yu B, Wang F. Predictive modeling of the hospital readmission risk from patients' claims data using machine learning: a case study on COPD. *Sci Rep.* **2019**;9(1):1–10.
- [22] Morinaga M, Yukawa N, Adachi H, Ezaki H, New PHA-COMP and its applications to alloy design. In: Gell M, et al., editors. *Superalloys 1984.* Warrendale, PA, USA: The Metallurgical Society of AIME; **1984**. p. 523–532.
- [23] Zhang Y, Zhou YJ, Lin JP, et al. Solid-solution phase formation rules for multi-component alloys. *Adv Eng Mater.* **2008**;10(6):534–538.
- [24] Guo S, Hu Q, Ng C, et al. More than entropy in high-entropy alloys: Forming solid solutions or amorphous phase. *Intermetallics.* **2013**;41:96–103.
- [25] Takeuchi A, Inoue A. Classification of bulk metallic glasses by atomic size difference, heat of mixing and period of constituent elements and its application to characterization of the main alloying element. *Mater Trans.* **2005**;46(12):2817–2829.
- [26] Guo S, Ng C, Lu J, et al. Effect of valence electron concentration on stability of fcc or bcc phase in high entropy alloys. *J Appl Phys.* **2011**;109(10):103505.
- [27] Wang S, Xiao X, Xia L, et al. Relationship between the widths of supercooled liquid regions and bond parameters of Mg-based bulk metallic glasses. *J Non Cryst Solids.* **2003**;321(1-2):120–125.
- [28] Yang X, Zhang Y. Prediction of high-entropy stabilized solid-solution in multi-component alloys. *Mater Chem Phys.* **2012**;132(2-3):233–238.
- [29] Singh AK, Kumar N, Dwivedi A, et al. A geometrical parameter for the formation of disordered solid solutions in multi-component alloys. *Intermetallics.* **2014**;53: 112–119.
- [30] Wang Z, Huang Y, Yang Y, et al. Atomic-size effect and solid solubility of multicomponent alloys. *Scr Mater.* **2015**;94:28–31.
- [31] Ye Y, Wang Q, Lu J, et al. Design of high entropy alloys: a single-parameter thermodynamic rule. *Scr Mater.* **2015**;104:53–55.
- [32] Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. *Advances in neural information processing systems*; **2017**.