




The effect of element characteristics on bainite transformation start temperature using a machine learning approach

Yangni Liu¹, Tingping Hou^{1,*} , Zhuang Yan¹, Tao Yu², Junwen Duan¹, Yuhui Xiao¹, and Kaiming Wu^{1,*}

¹The State Key Laboratory of Refractories and Metallurgy, Hubei Province Key Laboratory of Systems Science in Metallurgical Process, International Research Institute for Steel Technology, Collaborative Innovation Center for Advanced Steels, Wuhan University of Science and Technology, Wuhan 430081, People's Republic of China

²School of Optoelectronic Materials and Technology, Jiangnan University, Wuhan 430056, People's Republic of China

Received: 28 August 2022

Accepted: 28 November 2022

Published online:

1 January 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

ABSTRACT

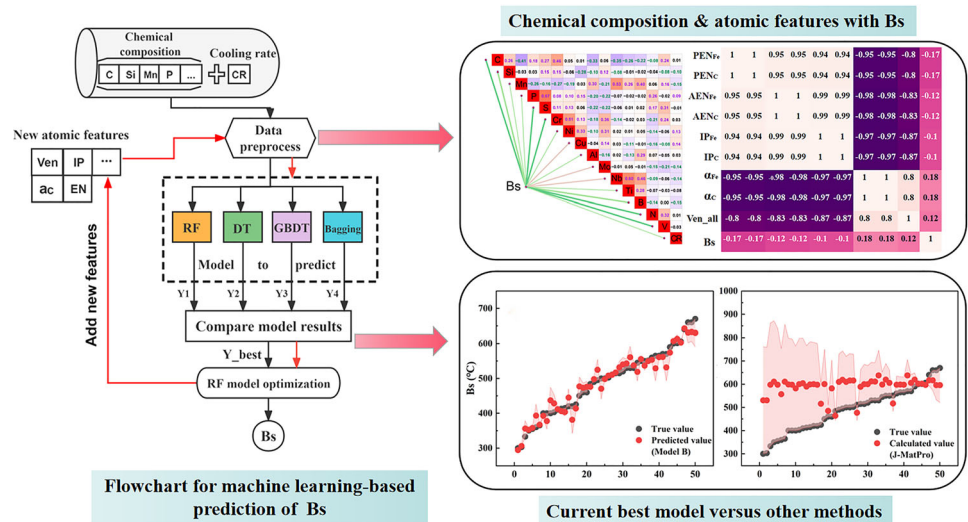
Bainite transformation start temperature (B_s) is an important index to measure the properties of bainitic steel. Based on the experimental results of bainite transformation behavior, the atomic scale characteristics are introduced, and the influence and prediction of different component content on B_s are analyzed by machine learning algorithm. The results show that B_s decreased significantly with the increase in C content (0–0.6wt.%) and Si content (0–0.2 wt.%), while the tendency remains almost unchanged when the Si content is greater than 0.2 wt.%. Furthermore, according to the analysis of atomic scale features, B_s has the strongest dependence on the number of valence electrons and the radius change rate relative to iron. The combination with the above two atomic scale features show the best model performance. The relationship between these two features and B_s is positively proportional, and B_s rises with the increase in their values. Extracting the valuable information about the relationship between B_s and element characteristics from the collected experimental data is of great significance to provide theoretical foundation of possible direction for the advances of designing the excellent properties in steels.

Handling Editor: P. Nash.

Address correspondence to E-mail: houtingping@wust.edu.cn; wukaiming@wust.edu.cn

<https://doi.org/10.1007/s10853-022-08035-5>

GRAPHICAL ABSTRACT



Introduction

Bainite steels are in high demand in many application areas owing to their unique properties, such as high tensile strength, impact toughness and weldability, with ultimate tensile strength up to 2.2 GPa and toughness values up to $130 \text{ MPam}^{1/2}$ [1, 2]. Bainite steels derive their strength mainly from fine bainite plates, which acquire sufficient ductility due to the presence of the ductile phase austenite. Therefore, bainitic steels are often used to manufacture pipelines for gas and oil transportation [3, 4]. The transformation behavior of bainite has been studied by many scholars, among which the common methods for measuring the phase transformation temperature are thermal expansion [5], thermal analysis [6] and metallographic methods [7]. In addition, subcooled austenite continuous cooling transition curve (CCT) about phase transition and phase transition points (bainite transformation start temperature, martensite transformation start temperature) can be obtained by using MUCG 83[8] and J-MatPro software [9].

The transformation kinetics and morphological characteristics of bainite are closely related to the elemental mass fraction, cooling rate and atomic scale

characteristics [10]. Previous studies of Fe-9Ni-C alloys found that the B_s value decreases with increasing carbon content which delays the onset of transformation [11]. As the transformation temperature changes, the microstructure of the early transformation changes and the morphological characteristics of bainite is transformed [12]. The addition of silicon delays the transformation kinetics of bainite. It was found that as the silicon content increased in the range of 1.0 to 2.0 wt%, resulting in more film-like residual austenite and less carbide in bainitic steels, the strength and total elongation of bainitic steels increased [13, 14]. The continuous cooling rate also affects the transformation properties of low-carbon bainite steels. Previous studies have shown that the formation of leaf-shaped bainite during the continuous cooling phase accelerates the subsequent bainite transformation [15]. As the cooling rate increases (1–30 °C/S), the bainite transformation start temperature increases and then decreases [16]. In addition, the valence electrons and radii of each atom are different due to the chemical elements involved in the material from different periods [17]. Previous studies found that the variation of the transition temperature of NiTi-based alloys is correlated with the number of valence

electrons per atom (ev/a). An increment in the number of valence electrons is accompanied by a tendency for the transition temperature to decline when $ev/a < 6.8$ or $ev/a > 7.2$. It is also found that the transition temperature lagging has a relatively wide range of values, which is considered to be related to the atomic radius. During the transformation, an increase in atomic size may lead to more energy dissipation, thus increasing the lag [18].

Previously, Bs was estimated by combining thermal simulation experiments with microstructure, which was inaccurate and limited [12]. So far, most of the bainite phase transformation-related researches have been focusing on experiments clarifying the influence of the phase field on microstructural aspects. The lack of a large amount of experimental data on Bs and related studies on machine learning considerably hinders the further study on revealing influential mechanisms of the bainite phase transformation. So simulation of the data needed is of significance for both the machine learning related studies in materials science and the valuable application of general methods of physics. Powerful data analysis and processing tools in machine learning can significantly reduce errors caused by inaccuracies in experimental operations, exclude noisy data, and provide important information [19]. The method can automatically adjust the weight of each factor according to the target value of the model and perform combined learning.

Machine learning has been used in the field of steel materials [20–25]. For example, Van Bohemen et al. [20] extracted model parameters from the best fit of published time–temperature transformed data and used the model to describe the start curve of bainite formation and predict bainite transformation kinetics. Moreover, the martensite transition temperature of the steel was predicted based on the thermodynamic model, including factors such as chemical composition, austenite grain size, and driving force [21]. Wang et al. [22] used an artificial neural network model to predict CCT diagrams for a class of steels. In addition, Zhang et al. [23] used a Gaussian process regression model based on the alloying elements of a steel to predict martensite transition temperature. They analyzed the intrinsic link between alloying elements and phase transformation point of martensite. Bs is closely related to specific alloying elements and process parameters [24, 25]. Therefore, it is crucial to investigate the intrinsic correlation between important

alloying elements and Bs by collecting relevant elemental content data for bainite-containing steels.

In this study, Bs is predicted by machine learning algorithm, Pearson and Spearman correlation coefficient, the random forest feature importance method using alloying elements (C, Si, etc.) and cooling rate (CR) as input features. The relationship between the alloying elements Si, C and Bs was analyzed. The dataset was divided based on the concentration of element C, and the model's performance was evaluated for low and medium carbon, respectively. Meanwhile, considering the model's generalization capability, new atomic scale features based on alloying elements and cooling rates were added to enable the model to predict Bs better. Finally, the model is validated by randomly selecting some data.

Methodology

In the present work, the random forest algorithm (RF) [26] is used as the dominant machine learning method which is based on the idea of integration learning in Fig. 1. First, the training sets (labeled by 'Decision Tree-1, 2, ..., N') are generated by bootstrapping, and a decision tree is constructed for each training set (green and blue points). Then, when the node has to find suitable features for further splitting, it randomly selects some features and finds the optimal solution (green dots) in Fig. 1. Since the algorithm uses the bag method, it actually obtains the information of the samples and the corresponding features which avoids overfitting [27, 28]. Finally, the results of all trees (labeled by 'Result-1, 2, ..., N') are

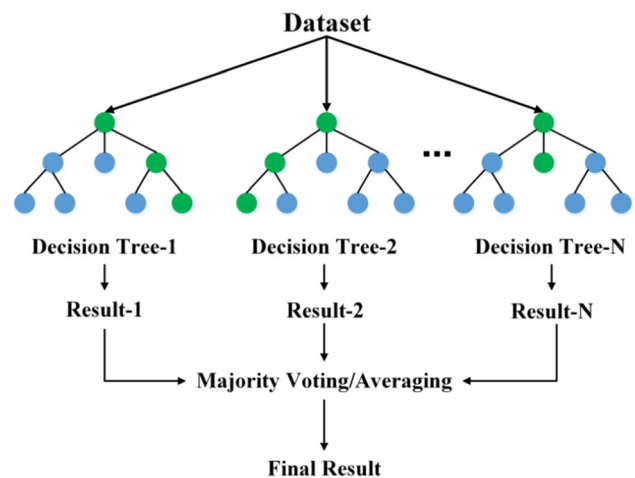


Figure 1 Structure diagram of Random Forest algorithm.

evaluated together, and the final prediction is obtained by voting or averaging (labeled by ‘Majority Voting/Averaging’) [29].

In machine learning modeling, it is necessary to use several methods to compare and select the optimal model as the actual prediction model [30]. Therefore, the final dataset (see Table. S4 in Supplementary Note 1) has been divided into training and testing sets in the ratio of 4:1, four different algorithms (RF, DT (Decision Tree), GBDT (Gradient Boosting Decision Tree) and Bagging) were adopted to build the model. The steps of this study include data collection, data processing, feature analysis, model building and selection, which can be simplified by a visual flowchart as shown in Fig. 2. The black arrow represents the chemical composition and cooling rate as the input characteristics. The prediction scores of Y1, Y2, Y3 and Y4 models are obtained. The prediction scores correspond to the values of three evaluation indexes for model evaluation, and finally, the optimal model is selected for Bs prediction. The red arrow represents the integration of atomic scale features on the basis of the original features and repeats the above steps to obtain a new Bs.

Dataset

In this study, 738 samples were collected from experimental results of previous literature [31–39]. The dataset contains 20 features (chemical composition of

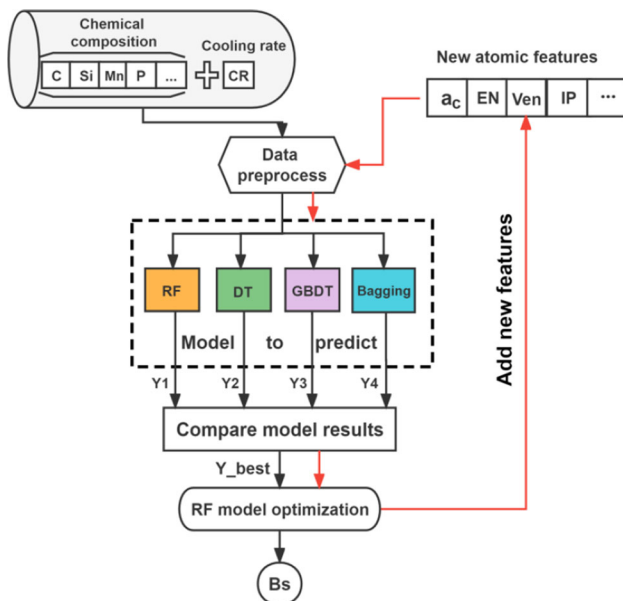


Figure 2 Diagram of machine learning-based alloy design system for the steels with desired phase transformation point.

C, Si, Mn, P, S, Al, etc., and cooling rate) and the target value Bs. The values of unrecorded alloying elements in the samples were set to be zero. All duplicate samples and discrete data from the box plots method [40] have been removed. Moreover, the scatter plot is formed by means of visualization data and used to identify the larger deviation features.

Data processing

Data processing mainly refers to the processing of null values, repeated values, discrete values, etc. (see Tables. S1-3 in Supplementary Note 1).

The Pearson correlation coefficient [41] is used for correlation analysis of the data and is expressed as r to reflect the degree of linear correlation between two variables (X and Y). The value of r is between -1 and 1. The larger absolute value corresponds to the stronger correlation between the variables. The calculation formula is as follows in Eq. (1):

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (1)$$

where n is the number of samples, X_i and Y_i are the observations at point i corresponding to the variables X and Y , and \bar{X} , \bar{Y} are the mean values of the X and Y variables, respectively.

The Spearman correlation coefficient, generally denoted by ρ_s , is used to assess the monotonic relationship and correlation between two continuous variables [42]. When there are no repeated values in the data and the two variables are completely monotonically correlated, the coefficient is 1 or -1. The formula is calculated according to Eq. (2).

$$\rho_s = \frac{\sum_{i=1}^n (R_i - \bar{R})(S_i - \bar{S})}{\left[\sum_{i=1}^n (R_i - \bar{R})^2 \sum_{i=1}^n (S_i - \bar{S})^2 \right]^{\frac{1}{2}}} \quad (2)$$

where n is the total number of samples, R_i and S_i is the ranks of the values for sample i , and \bar{R} and \bar{S} is the average ranks of independent and dependent variables, respectively.

Random forest feature importance [43]

There is an important feature in random forests which exhibits the ability to calculate the importance of individual feature variables. There are many features in the prediction model, and it is desirable to

find the feature variables that are highly correlated with the target value and to guarantee the prediction accuracy by selecting as few features as possible. Therefore, it is necessary to calculate the importance of each feature and rank them.

The extended new atomic features

To improve the generalizability of the model, atomic features are added to the original features, such as atomic radius and valence electrons. In this study, nine new atomic features are constructed, such as electronegativity relative to iron (EN_{Fe}) and carbon (EN_C) atoms, radius change rate relative to iron (α_{Fe}) and carbon (α_C) atoms, the first ionization energy of the relative iron (IP_{Fe}), carbon (IP_C), and the number of valence electrons (Ven_{all}). The electronegativity is further divided into the following parts: Pauling electronegativity relative to iron (PEN_{Fe}) and carbon (PEN_C) atoms, and Allen electronegativity relative to iron (AEN_{Fe}) and carbon (AEN_C) atoms (Calculation formula in Supplementary Note 4) [44].

Model performance evaluation metrics

The performance of the constructed models is evaluated using [30] root-mean-square error (RMSE), mean absolute error (MAE), and R^2 (coefficient of determination) (Eqs. (1)–(3) in Supplementary Note 2).

Empirical formula

The results of the present study are compared with previous formulae. There are four empirical formulas for the calculation of Bs. The following shows formula (3) [45] (Eqs. (4)–(6) in Supplementary Note 3).

$$\begin{aligned} BS = & 720 - 585.63C + 126.6C^2 - 66.34Ni + 6.06Ni^2 \\ & - 0.232Ni^3 - 31.66Cr + 2.17Cr^2 - 91.68Mn \\ & + 7.82Mn^2 - .3378Mn^3 - 42.37Mo + 9.16Co \\ & - 0.1255C0^2 + 0.00284Co^3 - 36.02Cu - 46.15Ru \end{aligned} \quad (3)$$

Results and discussions

The result of the processed data

Based on the above dataset, the outliers of Bs were calculated and removed by adopting the box plots.

Figure 3a shows the calculation results of Bs outliers that the upper and lower limits of Bs are 765 °C and 295 °C, respectively. Three outliers exist beyond the limits. Figure 3b shows the overall distribution of Bs data. Based on the previous experimental and literature data, it can be judged that there are few data of Bs greater than 700 °C. There are only two values in this paper, so this part of the data is classified as abnormal values. Then, the five outliers of Bs (Red pentagrams in Fig. 3b) were removed. Figure 3 (c, d) shows the data distribution of chemical components C, P and Bs, respectively, and the outliers of C and P (Red pentagrams) are also removed. Similarly, the other feature outliers are optimized (see Fig. S1 in Supplementary Note 1). The range of individual feature values for the final dataset is shown in Table 1.

The dependence of Bs on the elements Si and C

Correlation analysis of different chemical components with Bs was performed (Fig. 4). The results showed that the highest correlation between the features was found for the chemical components S and P, and the value of the correlation coefficient r was 0.67. Meanwhile, the elemental features that play an important role on Bs are C, Si, etc. (green line). It indicates that Bs will gradually decrease with the increase in C and Si content. In addition, the features of Cu, Nb, etc., have an positive-going action on the increase in Bs (brown line), which means that Bs will gradually increase with the increase in Cu content. The result indicates that the more C or the less Cu content can reduce Bs, but the dependence of Bs on C was stronger compared to Cu. The correlation results proved that there was no significant linear interaction between the individual elemental characteristics.

To explore the effects of special elements on Bs, the average values of C and Si element contents were taken as representative. When the C concentration is taken in the range of 0–0.6 wt.%, the Bs is 491–586 °C. As shown clearly in Fig. 5a, where the carbon concentration is chosen as the x-axis. The decrease in Bs temperature is inversely proportional to the increase in carbon concentration within the range of 0–0.6 wt.% range. This is because the increase in carbon concentration in austenite would give rise to the decrease in the carbon concentration gradient within the austenite particles prior to

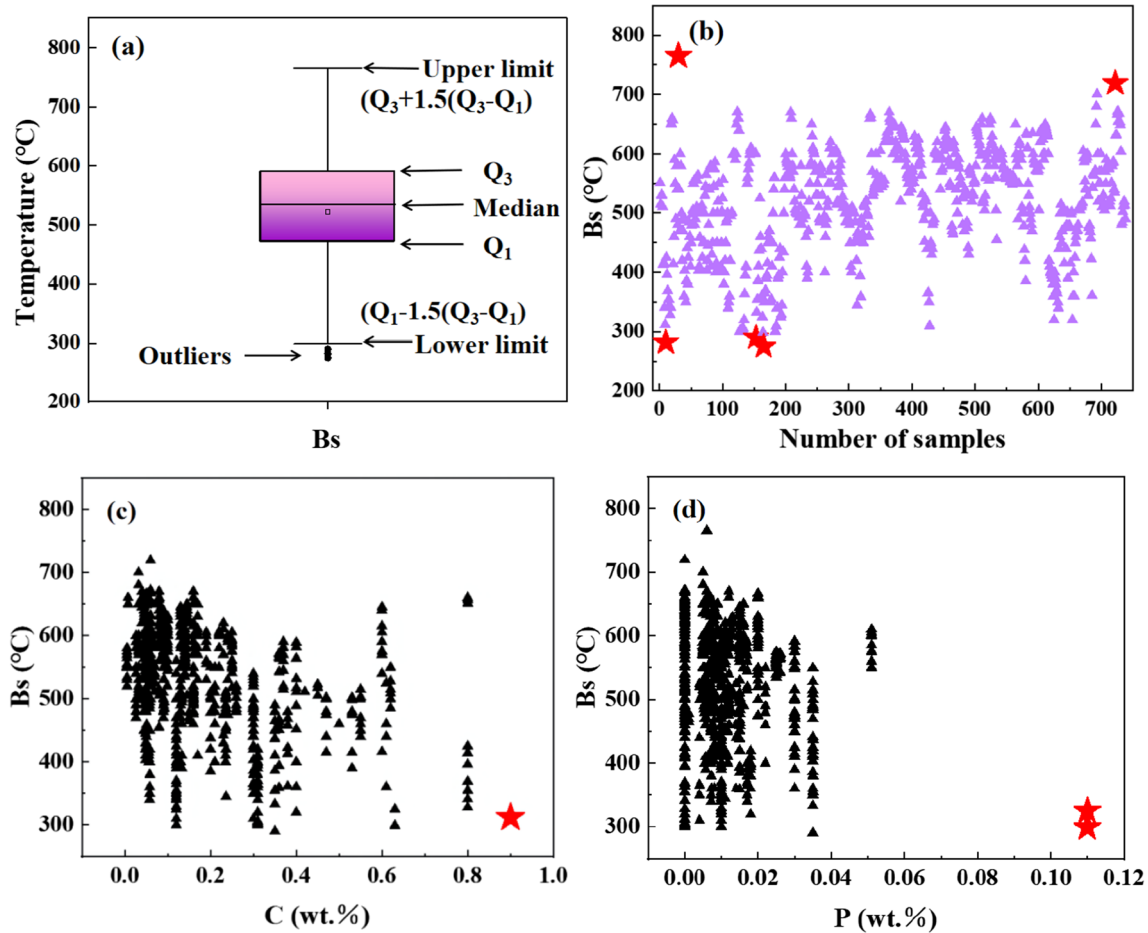


Figure 3 a Outliers of Bs calculated by box diagram. Data distribution, which are b Bs, c chemical element C and d element P. Red pentagram represents the discrete data.

transformation. Correspondingly, the carbon concentration gradient is the effective driving force for bainite growth. Reducing the driving force will certainly increase the time required for the diffusion of carbon atoms from the interface and thus inevitably reduce the bainite growth kinetics [46]. Moreover, more carbon concentration leads to a decrease in the Gibbs energy difference between bainite and austenite, finally triggering the bainite nucleation [47]. When the value of the Si element is taken in the range of 0–1.6 wt.%, the Bs is 463–596 °C. Figure 5b illustrates the relationship between Bs and Si element content. The results demonstrate that Bs decreases significantly with the increase in Si content when the Si content is 0–0.2 wt.%. When the Si content is greater than 0.2 wt.%, Bs decreases gradually and slowly with the increase in Si content. Figure 5b also indicates that if the Si concentration of the steel is more than 0.2 wt.%, the Si has a weak influence on

the bainite reaction. The reason is that Si, as a non-carbide forming element, can inhibit carbide precipitation and can be used as a solid solution element in the steel to stabilize the austenite, thus making the C curve shift to the right and reducing the bainite transformation start temperature [48]. The addition of Si can delay the bainite transformation by affecting the nucleation and growth rate of bainitic ferrite [49].

The results of the model based on the dataset

In this study, different algorithms were used for modeling, and the prediction results are listed in Table 2. The results show that the model has the highest prediction accuracy of 90.5% when modeled RF algorithm, followed by the GBDT algorithm prediction rate of 90.1%. Thus, the above two algorithms are more suitable for predicting the phase transition

Table 1 Range of compositions (wt.%), cooling rate (°C/s) in the dataset

Features	Minimum	Maximum	Average
C	0.01	0.62	0.18
Si	0.00	1.80	0.31
Mn	0.40	2.20	1.40
P	0.00	0.05	0.01
S	0.00	0.05	0.01
Cr	0.00	2.00	0.27
Ni	0.00	2.06	0.16
Cu	0.00	0.36	0.03
Al	0.00	0.05	0.01
Mo	0.00	0.90	0.09
Nb	0.00	0.15	0.02
Ti	0.00	0.16	0.02
B	0.00	0.003	0.001
N	0.00	0.020	0.005
V	0.00	0.18	0.03
CR	0.01	100	13.61

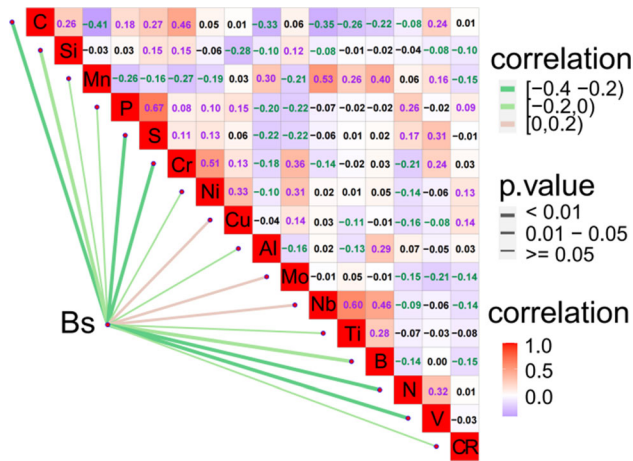


Figure 4 Correlation thermodynamic diagram. The squares represent correlations between features, with red and purple squares representing forward and reverse, respectively. Lines represent correlations between features and Bs. Green and brown lines represent enhancement or weakening, respectively.

point. But through the results in Table 2, it is also clearly found that the prediction accuracy of the four algorithms has little difference (the difference in prediction rate is about 3%). This further indicates that the feasibility of the four algorithms is strong. In addition, the parameters of the RF algorithm are optimized in this paper. The results show that the optimal values of the model parameters are $n_estimators = 171$, $max_depth = 17$, $min_samples_leaf =$

1, $min_samples_split = 2$. Therefore, this set of parameter conditions is chosen to build the final Bs prediction model.

The results of the model based on low-carbon and medium-carbon dataset

Based on the importance of element C to Bs, the entire dataset was divided into low carbon ($Wc \leq 0.25\%$) and medium carbon ($0.25\% \leq Wc \leq 0.6\%$). Figure 6 examines the degree of deviation of the predicted values from the actual values for the low- and medium-carbon data and draws the corresponding scatter plots. Figure 6 also shows the prediction results (R^2 , RMSE, and MAE) for the low- and medium-carbon data based on random forest. Figure 6a and c shows that the errors RMSE = 14.7311 and MAE = 9.0564 for the low-carbon training set are larger than those of 9.1569 and 2.6751 for the medium-carbon training set. Figure 6(b) and d shows that $R^2 = 0.8907$ for the low-carbon test set is smaller than the value of 0.9653 for the medium-carbon test set, which indicates that the model in this paper performs better for the Bs prediction results for the medium-carbon data.

Addition of new features

The correlation analysis of the new features is performed as shown in Fig. 7. Analysis indicates that α_{Fe} , α_C , and Ven_all have a facilitative effect on the elevation of Bs, while PEN_{Fe} and PEN_C have the opposite effect. In addition, IP_{Fe} and IP_C had the same correlation coefficient of -0.1 with Bs, indicating that the dependence of Bs on them was similar. Figure 7 further shows that there is a strong correlation between the new features associated with the atomic parameters. The correlation coefficients between PEN_{Fe} , PEN_C , AEN_{Fe} , AEN_C , IP_{Fe} , IP_C , α_{Fe} and α_C were all above 0.9. Therefore, some of the highly correlated features can be considered for removal without affecting the model performance.

Figure 8 shows the results of ranking the importance of new features on Bs. Note that the importance values of the features are normalized. The results show that Ven_all ranks first in terms of feature importance. The feature Ven_all is related to the number of valence electrons of the element, taking into account the effect of electronic stability on the bainite transformation. In addition, PEN_{Fe} , PEN_C , α_{Fe} ,

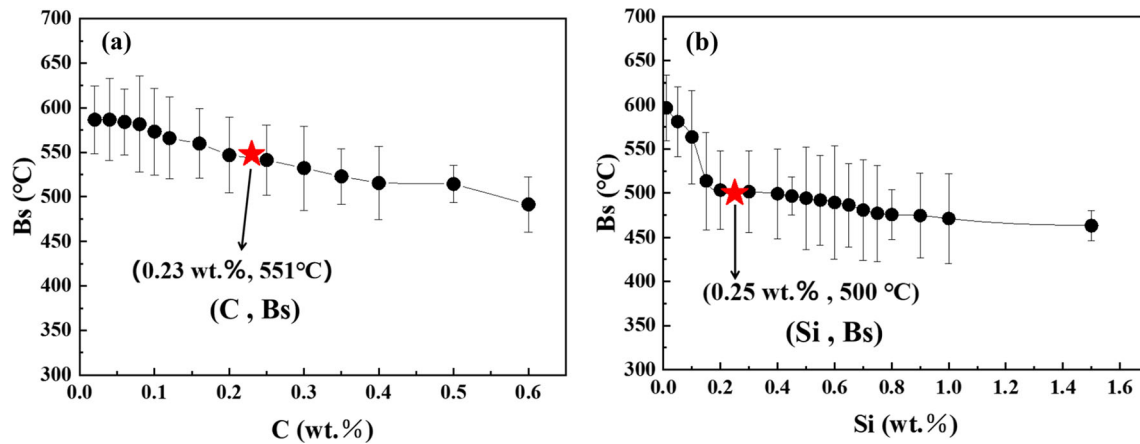


Figure 5 Relationship between **a** C, **b** Si and Bs. The square represents the calculated value. The error bar represents the calculated error value of this point. The pentagrams correspond to the experimental data of C [50] and Si [51], respectively.

Table 2 Model results according to R^2 , RMSE and MAE

Algorithms	R^2	RMSE	MAE
RF	0.905	25.12	17.94
GBDT	0.901	25.18	18.26
Bagging	0.893	27.12	19.07
DT	0.878	30.06	20.59

α_C are also ranked high and may include the influence of alloying elements on the stability of iron and carbon. Therefore, adding such atomic features can improve the performance of the trained model.

Figure 9 shows the results when each feature is added individually. The addition of new features, such as PEN_C , IP_C , and Ven_all reduced the MAE values (green dots). In addition, Ven_all also improved the RMSE (red dots) values without significantly worsening the other evaluation metrics (blue dots). Therefore, Ven_all , PEN_C and AEN_{Fe} may be used as additional beneficial features in addition to chemical composition and cooling rate. Figure 10 shows the results after sequentially adding the remaining eight features and Ven_all combinations. The results show that the model with $Ven_all + \alpha_{Fe}$ has the smallest MAE and RMSE errors and perform well on the R^2 index.

In the present study, the model with the original features is denoted as Model A, and the model with the new features added is denoted as Model B. The comparison of the results of Model A and Model B is reported in Table 3. The results indicate only one special case where the Bagging algorithm model A

($R^2 = 0.893$) gives slightly better results than model B ($R^2 = 0.891$). The predictive ability of Model B of the remaining three algorithms exceeded that of Model A. The four models' prediction results are excellent (R^2 is more than 0.85), which indicates that the bainite transformation start temperature can be accurately predicted by the effective regression model based on the existing data. Furthermore, it can be seen that the random forest model has the highest R^2 (0.913) and the smallest RMSE (24.67) and MAE (17.34), denoting that the model has the best performance. Therefore, in selecting the Bs prediction model, this paper first preprocessed the data, analyzed the characteristics, standardized the data, and adopted the random forest method to predict the special phase transformation point.

Modeling validation

To test the model's accuracy, data validation of the prediction results of the random forest model was conducted in the present work. By randomly selecting 50 data, a scatter plot is drawn with the experimental value as the horizontal coordinate and the predicted values as the vertical coordinate (Fig. 11). The results show the predicted values with the original features (Model A) (in Fig. 11a). The model with the new features (Model B) (in Fig. 11(b)) is in good agreement with the experimental values. Compared to Model A, the predicted values of Model B matched the experimental values better. Figure 11 (c, d) reveals that the calculated Bs values deviate significantly from the experimental values. The vast majority of the Bs values calculated using J-MatPro in

Figure 6 Fitting between the predicted value and experimental value. **a** low-carbon training set, **b** low-carbon testing set, **c** medium-carbon training set, **d** medium-carbon testing set.

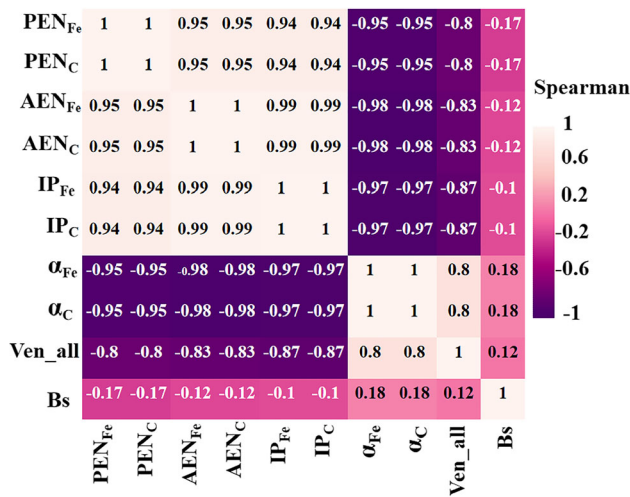
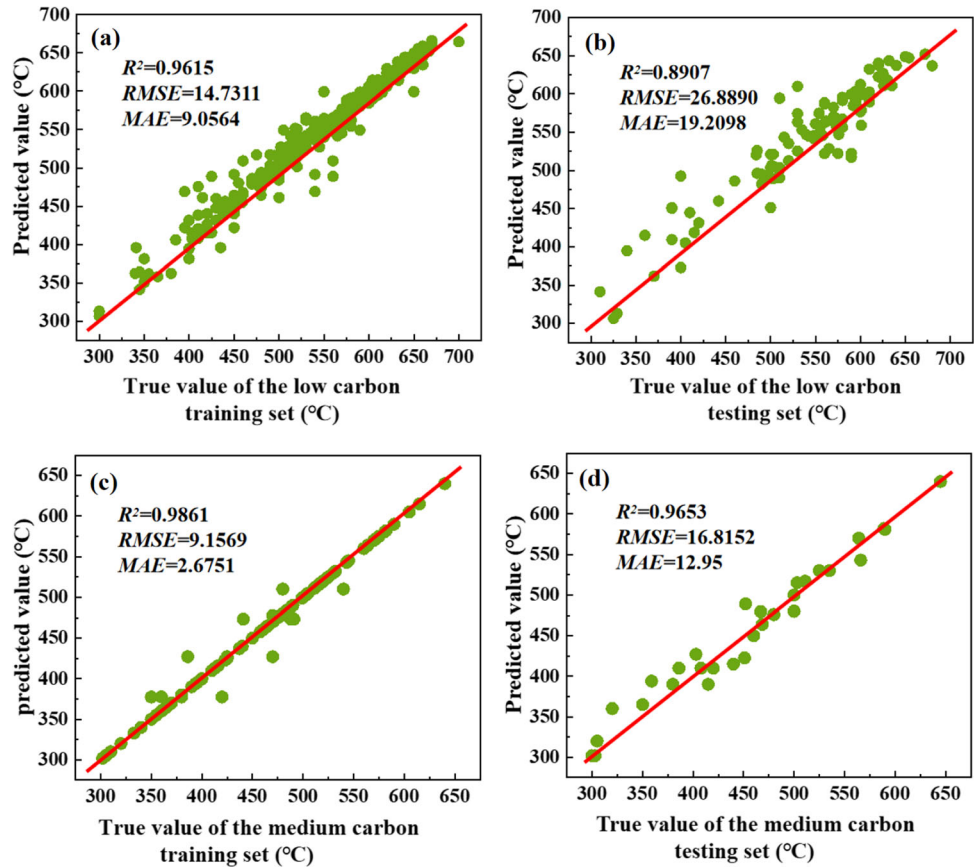


Figure 7 Correlation analysis between the new features and Bs.

Fig. 11c are higher than the experimental values and have larger error values. Figure 11d shows the Bs calculated by the empirical Eq. (1). It can be clearly seen that the data are scattered. There is no clear linear fitting trend with the experimental values. Note that Eq. (1) is the best fit among the four

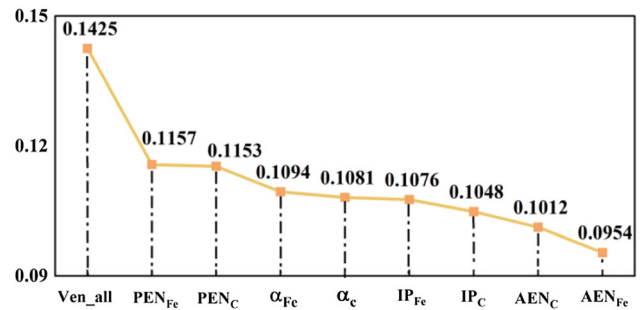


Figure 8 Result of feature importance ranking based on RF.

empirical equations, and the remaining three results are shown in Fig. S2 of Supplementary Note 3.

Conclusion

In this study, the bainite transformation start temperature (Bs) experimental data were collected and preprocessed. The feature has been analyzed using the Pearson correlation coefficient. The influence law of chemical composition Si and C on Bs was further quantified and investigated by means of the random

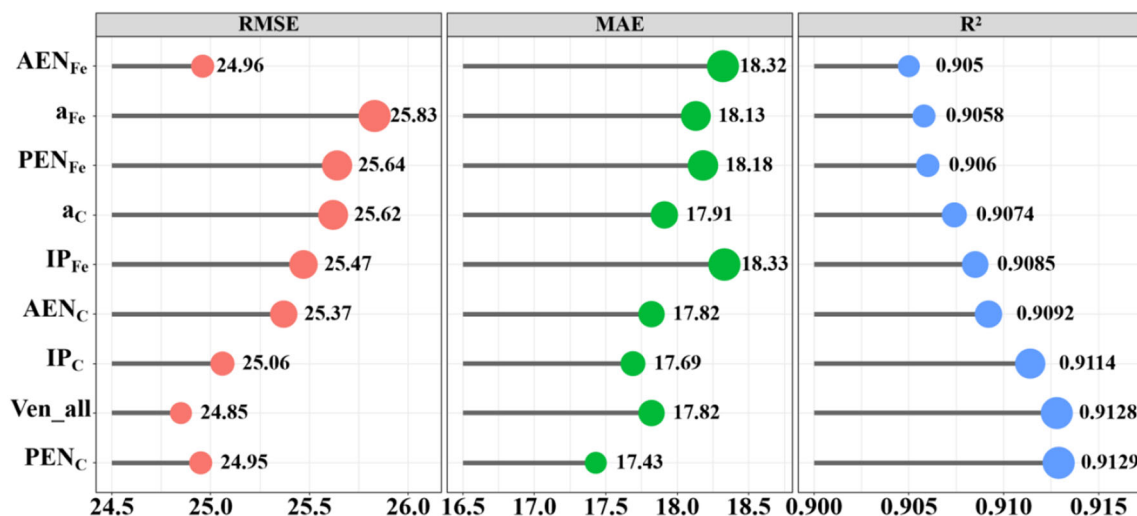


Figure 9 Results of adding new features.

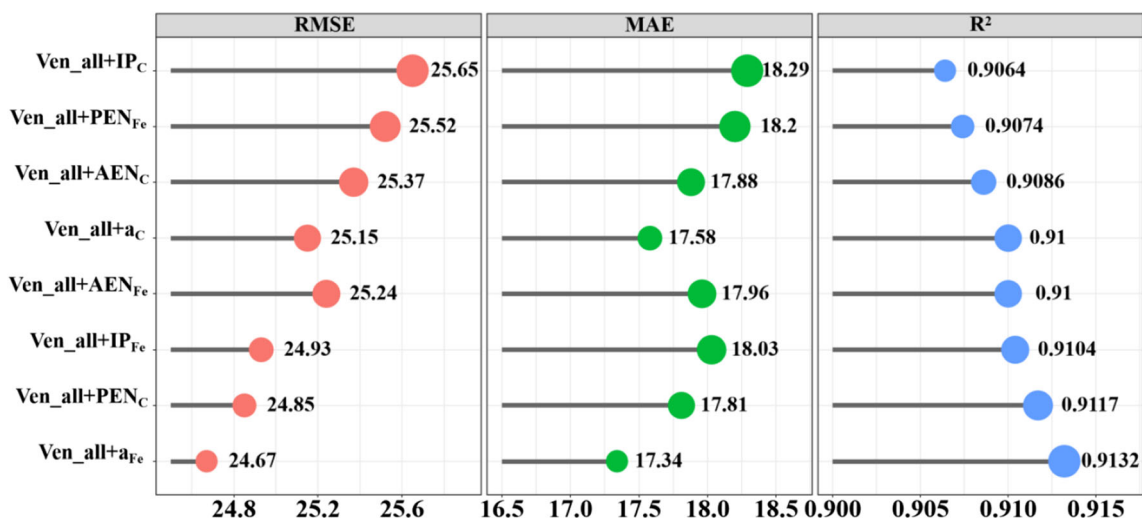


Figure 10 Results of adding the remaining features based on *Ven_{all}*.

Table 3 Comparison of the results of Model A and Model B

Algorithms	Model	hR^2	RMSE	MAE
RF	Model A	0.905	25.12	17.94
	Model B	0.913	24.67	17.34
GBDT	Model A	0.901	25.18	18.26
	Model B	0.905	25.03	18.14
Bagging	Model A	0.893	27.12	19.07
	Model B	0.891	27.48	19.30
DT	Model A	0.878	30.06	20.59
	Model B	0.88	28.59	20.36

forest model. The obtained accuracy could be as high as 91%, with the error within ± 25 °C. Finally, the prediction model of Bs based on the random forest was obtained. In addition, new atomic features were added for prediction, and the results are closer to the experimental values. They can be extended to the case of an unknown new element, which also provides new ideas to study other factors affecting Bs. The following conclusions can be drawn:

- (1) Bs decreases consistently from 586 °C to 491 °C, with increasing C from 0.1 wt.% to 0.6 wt.%. The model prediction results based on the

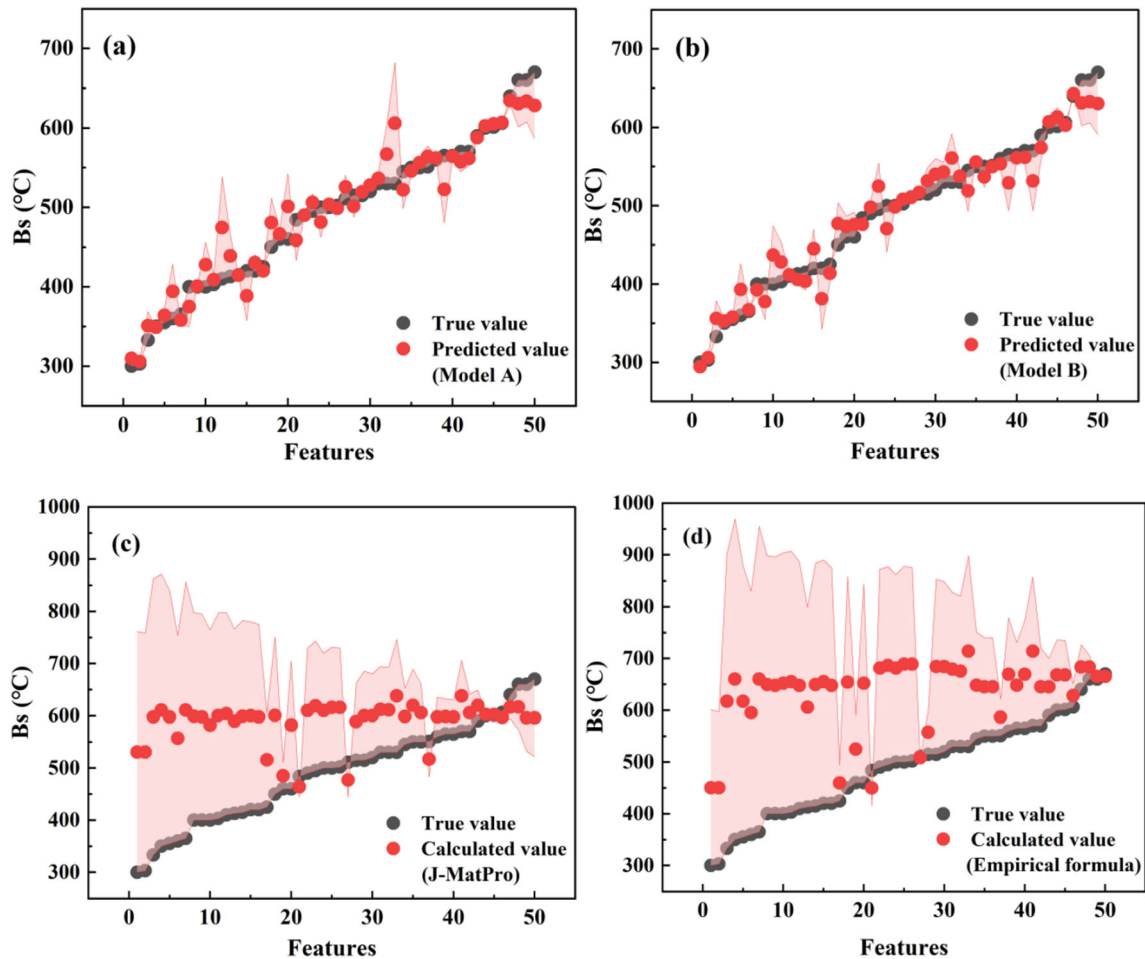


Figure 11 Fitting between the predicted and experimental values of the validation set. **a** Model A, **b** Model B, **c** J-MatPro, **d** Empirical formula. The red shaded area represents the error.

carbon concentration division show that the prediction of medium-carbon steel ($R^2 = 0.9653$) is better than that of low-carbon steel ($R^2 = 0.8907$). It means that the model performs better in predicting B_s for medium-carbon steel.

- (2) B_s temperatures are found to vary significantly with increasing Si. By increasing the silicon elements to 0.2 wt.%, B_s considerably decreases from 596 °C to 513 °C. A turning point is at 0.2 wt.%. While the tendency remains almost unchanged when the Si content is greater than 0.2 wt.%.
- (3) Among the added atomic features, the number of valence electrons ranked first in importance.

In addition, the radius change rate relative to iron and the number of valence electrons both have a similar relationship with B_s in a positive trend. The addition of atomic features improves the performance of the model and enhances the generalization ability of the model.

Acknowledgements

The financial support from the National Natural Science Foundation of China (Grant Nos. 12174296, U1532268 and U20A20279), the Key Research and Development Program of Hubei Province (Grant No. 2021BAA057), Hubei Provincial Colleges and

Universities Excellent Young and Middle-aged Science and Technology Innovation Team Project (Grant No. T201903) and 111 projects (D18018).

Declarations

Conflict of interest The authors declare no competing financial interests.

Supplementary Information: The online version contains supplementary material available at <http://doi.org/10.1007/s10853-022-08035-5>.

References

- [1] Rodrigues PCM, Pereloma EV, Santos DB (2000) Mechanical properties of an HSLA bainitic steel subjected to controlled rolling with accelerated cooling. *Mater Sci Eng, A* 283(1–2):136–143. [https://doi.org/10.1016/S0921-5093\(99\)00795-9](https://doi.org/10.1016/S0921-5093(99)00795-9)
- [2] Yoozbashi MN, Yazdani S, Wang TS (2011) Design of a new nanostructured, high-Si bainitic steel with lower cost production. *Mater Des* 32(6):3248–3253. <https://doi.org/10.1016/j.matdes.2011.02.031>
- [3] Yakubtsov IA, Poruks P, Boyd JD (2008) Microstructure and mechanical properties of bainitic low carbon high strength plate steels. *Mater Sci Eng A* 480(1–2):109–116. <https://doi.org/10.1016/j.msea.2007.06.069>
- [4] Kumar A, Singh A (2021) Mechanical properties of nanostructured bainitic steels. *Materialia* 15:101034. <https://doi.org/10.1016/j.mtla.2021.101034>
- [5] Mao G, Cao R, Chen J (2017) Analysis on bainite transformation in reheated low-carbon bainite weld metals. *Mater Sci Technol* 33(15):1829–1837. <https://doi.org/10.1080/02670836.2017.1325561>
- [6] Kawuloková M, Smetana B, Zlá S, Kalup A, Mazancová E, Váňová P, Rosypalová S (2017) Study of equilibrium and nonequilibrium phase transformations temperatures of steel by thermal analysis methods. *J Therm Anal Calorim* 127(1):423–429. <https://doi.org/10.1007/s10973-016-5780-4>
- [7] Gao J, Li C, Zhang D, Han X (2020) Study on the transformation mechanism of twinning martensite and the growth behavior of variants based on phase-field method. *Steel Res Int* 91(10):2000142. <https://doi.org/10.1002/srin.202000142>
- [8] Yoozbashi MN, Yazdani S (2010) Mechanical properties of nanostructured, low temperature bainitic steel designed using a thermodynamic model. *Mater Sci Eng, A* 527(13–14):3200–3205. <https://doi.org/10.1016/j.msea.2010.01080>
- [9] Zhang CL, Fu HG, Ma SQ, Yi DW, Jian L, Xing ZG, Lei YP (2019) Effect of Mn content on microstructure and properties of wear-resistant bainitic steel. *Mater Res Exp* 6(8):086581. <https://doi.org/10.1088/2053-1591/ab1c8d>
- [10] Wei W, Retzl P, Kozeschnik E, Povoden-Karadeniz E (2021) A semi-physical α - β model on bainite transformation kinetics and carbon partitioning. *Acta Mater* 207:116701. <https://doi.org/10.1016/j.actamat.2021.116701>
- [11] Kawata H, Manabe T, Fujiwara K, Takahashi M (2018) Effect of carbon content on bainite transformation start temperature in middle–high carbon Fe-9Ni-C alloys. *ISIJ Int* 58(1):165–172. <https://doi.org/10.2355/isijinternational.isijint-2017-387>
- [12] Kawata H, Fujiwara K, Takahashi M (2017) Effect of carbon content on bainite transformation start temperature in low carbon Fe-9Ni-C alloys. *ISIJ Int* 57(10):1866–1873. <https://doi.org/10.2355/isijinternational.isijint-2017-239>
- [13] Toji Y, Matsuda H, Raabe D (2016) Effect of Si on the acceleration of bainite transformation by pre-existing martensite. *Acta Mater* 116:250–262. <https://doi.org/10.1016/j.actamat.2016.06.044>
- [14] Zhao J, Jia X, Guo K, Jia NN, Wang YF, Wang YH, Wang TS (2017) Transformation behavior and microstructure feature of large strain ausformed low-temperature bainite in a medium C-Si rich alloy steel. *Mater Sci Eng A* 682:527–534. <https://doi.org/10.1016/j.msea.2016.11.073>
- [15] Chen Y, Chen L, Zhou X, Zhao Y, Zha X, Zhu F (2016) Effect of continuous cooling rate on transformation characteristic in microalloyed low carbon bainite cryogenic pressure vessel steel. *Trans Indian Inst Met* 69(3):817–821. <https://doi.org/10.1007/s12666-015-0564-2>
- [16] Gao B, Tan ZL, Tian Y, Liu YR, Wang R, Gao GH, Wang J, Zhang M (2022) Accelerated isothermal phase transformation and enhanced mechanical properties of railway wheel steel: the significant role of pre-existing bainite. *Steel Res Int* 93(2):2100494. <https://doi.org/10.1002/srin.202100494>
- [17] Zarinejad M, Liu Y (2008) Dependence of transformation temperatures of NiTi-based shape-memory alloys on the number and concentration of valence electrons. *Adv Func Mater* 18(18):2789–2794. <https://doi.org/10.1002/adfm.200701423>
- [18] Xing W, Meng F, Yu R (2017) Strengthening materials by changing the number of valence electrons. *Comput Mater Sci* 129:252–258. <https://doi.org/10.1016/j.commatsci.2016.12.037>
- [19] Guo S, Yu J, Liu X, Wang C, Jiang Q (2019) A predicting model for properties of steel using the industrial big data based on machine learning. *Comput Mater Sci* 160:95–104. <https://doi.org/10.1016/j.commatsci.2018.12.056>

- [20] Van Bohemen S, Morsdorf L (2010) Modeling start curves of bainite formation. *Metall Mater Trans A*(2). <https://doi.org/10.1007/s11661-009-0106-9>
- [21] Van Bohemen S, Sietsma J (2008) Modeling of isothermal bainite formation based on the nucleation kinetics. *Int J Mater Res* 99(7):739–747. <https://doi.org/10.3139/146.101695>
- [22] Wang J, Van Der Wolk PJ, Van Der Zwaag S (2000) On the influence of alloying elements on the bainite reaction in low alloy steels during continuous cooling. *J Mater Sci* 35(17):4393–4404. <https://doi.org/10.1023/A:1004865209116>
- [23] Zhang Y, Xu XJ (2021) Machine learning steel Ms temperature. *Simulation: Transactions of The Society for Modeling and Simulation International* 97(6):383–425
- [24] Lee SJ, Park JS, Lee YK (2008) Effect of austenite grain size on the transformation kinetics of upper and lower bainite in a low-alloy steel. *Scripta Mater* 59(1):87–90. <https://doi.org/10.1016/j.scriptamat.2008.02.036>
- [25] Yakubtsov IA, Boyd JD (2008) Effect of alloying on microstructure and mechanical properties of bainitic high strength plate steels. *Mater Sci Technol* 24(2):221–227. <https://doi.org/10.1179/174328407X243005>
- [26] Zhi YJ, Fu DM, Zhang DW (2019) Prediction and knowledge mining of outdoor atmospheric corrosion rates of low alloy steels based on the random forests approach. *Metals* 9(3):383–383. <https://doi.org/10.3390/met9030383>
- [27] Paul A, Gangopadhyay A, Chintla AR, Mukherjee DP, Das P, Kundu S (2018) Calculation of phase fraction in steel microstructure images using random forest classifier. *Iet Image Process*. <https://doi.org/10.3139/146.101695>
- [28] Li Z, Wen DH, Ma Y, Wang Q, Chen GQ, Zhang RQ et al (2018) Prediction of alloy composition and microhardness by random forest in maraging stainless steels based on a cluster formula. *J Iron Steel Res English Edition* 25(7):7. <https://doi.org/10.1007/s42243-018-0104-5>
- [29] Song SH (2021) Random forest approach in modeling the flow stress of 304 stainless steel during deformation at 700–900 °C. *Materials* 14. <https://doi.org/10.3390/ma14071812>
- [30] Kang MC, Yoo DY, Gupta R (2021) Machine learning-based prediction for compressive and flexural strengths of steel fiber-reinforced concrete. *Constr Build Mater* 266:121117. <https://doi.org/10.1016/j.conbuildmat.2020.121117>
- [31] Halmešová K, Procházka R, Koukolíková M, Džugan J, Konopík P, Bucki T (2022) Extended continuous cooling transformation (CCT) diagrams determination for additive manufacturing deposited steels. *Materials* 15:3076. <https://doi.org/10.3390/ma15093076>
- [32] Lee S, Na H, Kim B, Kim D, Kang C (2013) Effect of niobium on the ferrite continuous-cooling-transformation (CCT) curve of ultrahigh-thickness Cr-Mo steel. *Metall Mater Trans A* 44(6):2523–2532. <https://doi.org/10.1007/s11661-013-1616-z>
- [33] Grajcar A, Zalecki W, Burian W, Kozłowska A (2016) Phase equilibrium and austenite decomposition in advanced high-strength medium-Mn bainitic steels. *Metals* 6(10):248. <https://doi.org/10.3390/met6100248>
- [34] AKrbata M, Krizan D, Eckert M, Kaar S, Dubec A, Ciger R (2022) Austenite decomposition of a lean medium Mn Steel suitable for quenching and partitioning process: comparison of CCT and DCCT diagram and their microstructural changes. *Materials* 15(5):1753. <https://doi.org/10.3390/ma15051753>
- [35] Grajcar A, Zalecki W, Skrzypczyk P, Kilarski A, Kowalski A, Kołodziej S (2014) Dilatometric study of phase transformations in advanced high-strength bainitic steel. *J Therm Anal Calorim* 118(2):739–748. <https://doi.org/10.1007/s10973-014-4054-2>
- [36] Cota AB, Modenesi PJ, Barbosa R, Santos DB (1998) Determination of CCT diagrams by thermal analysis of an HSLA bainitic steel submitted to thermomechanical treatment. *Scripta Mater* 40(2):165–169. [https://doi.org/10.1016/s1359-6462\(98\)00410-2](https://doi.org/10.1016/s1359-6462(98)00410-2)
- [37] Cota AB, Santos DB (2000) Microstructural characterization of bainitic steel submitted to torsion testing and interrupted accelerated cooling. *Mater Charact* 44(3):291–299. [https://doi.org/10.1016/s1044-5803\(99\)00060-1](https://doi.org/10.1016/s1044-5803(99)00060-1)
- [38] Xu FY, Wang YW, Bai BZ, Fang HS (2010) CCT curves of low-carbon Mn-Si steels and development of water-cooled bainitic steels. *J Iron Steel Res Int* 17(3):46–50. [https://doi.org/10.1016/S1006-706X\(10\)60071-4](https://doi.org/10.1016/S1006-706X(10)60071-4)
- [39] Li X, Zhao J, Yang X, Bao JC, Ning BQ (2013) Effects of cooling rate on microstructure and properties of Nb-Ti micro-alloyed steel. *Appl Mech Mater* 341:208–212. <https://doi.org/10.4028/AMM.341-342.208>
- [40] Moeini B, Haack H, Fairley N, Fernandez V, Gengenbach TR, Easton CD, Linford MR (2021) Box plots: A simple graphical tool for visualizing overfitting in peak fitting as demonstrated with X-ray photoelectron spectroscopy data. *J Electron Spectrosc Relat Phenom* 250:147094. <https://doi.org/10.1016/j.elspec.2021.147094>
- [41] Li S, Li S, Liu D, Zou R, Yang Z (2022) Hardness prediction of high entropy alloys with machine learning and material descriptors selection by improved genetic algorithm. *Comput Mater Sci* 205:111185. <https://doi.org/10.1016/j.commat.2022.111185>
- [42] Winter JCF, Gosling SD, Potter J (2016) Comparing the Pearson and Spearman correlation coefficients across

- distributions and sample sizes: a tutorial using simulations and empirical data. *Psychol Methods* 21(3):273. <https://doi.org/10.1037/met0000079>
- [43] Menze BH, Kelm MB, Masuch R, Himmelreich U, Bachert P, Petrich W (2009) A comparison of random forest and its gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bioinf* 10(1):1–16. <https://doi.org/10.1186/1471-2105-10-213>
- [44] Wang JH, Sun S, He YL, Zhang TY (2019) Machine learning prediction of the hardness of tool and mold steels (in Chinese). *Sci Sin Technol* 49:1148–1158. <https://doi.org/10.1360/SST-2019-0060>
- [45] Zhao JC, Notis MR (1995) Continuous cooling transformation kinetics versus isothermal transformation kinetics of steels: a phenomenological rationalization of experimental observations. *Mater Sci Eng* 15(4–5):135–207. [https://doi.org/10.1016/0927-796X\(95\)00183-2](https://doi.org/10.1016/0927-796X(95)00183-2)
- [46] Caballero FG, Miller MK, Babu SS, Garcia-Mateo C (2007) Atomic scale observations of bainite transformation in a high carbon high silicon steel. *Acta Mater* 55(1):381–390. <https://doi.org/10.1016/j.actamat.2006.08.033>
- [47] Speer JG, Edmonds DV, Rizzo FC, Matlock DK (2004) Partitioning of carbon from supersaturated plates of ferrite, with application to steel processing and fundamentals of the bainite transformation. *Curr Opin Solid State Mater Sci* 8(3–4):219–237. <https://doi.org/10.1016/j.cossms.2004.09.003>
- [48] Tian J, Xu G, Jiang Z, Wan X, Hu H, Yuan Q (2019) Transformation behavior and properties of carbide-free bainite steels with different Si contents. *Steel Res Int* 90(3):1800474. <https://doi.org/10.1002/srin.201800474>
- [49] Liu SK, Zhang J (1990) The influence of the Si and Mn concentrations on the kinetics of the bainite transformation in Fe-C-Si-Mn alloys. *Metall Trans A* 21(6):1517–1525. <https://doi.org/10.1007/BF02672566>
- [50] Schindler I, Kawulok R, Opěla P, Kawulok P, Rusz S, Sojka J, Pindor L (2020) Effects of austenitization temperature and pre-deformation on CCT diagrams of 23MnNiCrMo5-3 Steel. *Materials* 13(22):5116. <https://doi.org/10.3390/ma13225116>
- [51] Bräutigam-Matusb K, Altamirano G, Salinas A, Flores A, Goodwin F (2018) Experimental determination of continuous cooling transformation (CCT) diagrams for dual-phase steels from the intercritical temperature range. *Metals* 8(9):674. <https://doi.org/10.3390/met8090674>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.