Original Research Article

Research on Optimization of "Deoxidation Alloying" of Molten Steel Based on Linear Programming

Anna Li*, **Dongqing Xu**

School of Finance, Anhui University of Finance and Economics, Bengbu 233030, Anhui Province, China. E-mail: lan0528hh@163.com

Abstract: Aiming at the optimization of the supporting solution for molten steel "deoxidation alloying", the cost of "deoxidation alloying" is minimized from an economic perspective. Using Excel, Eviews and spss software programming, through factor analysis, clustering dimension reduction, principal component analysis Multiple linear regression analysis and linear programming optimization analysis, the author found out the main factors that affected the yield of alloy elements. This paper establishes a multiple linear regression mathematical model that affects the main factors of alloy elements and yield. According to the reference alloy price, the linear programming model is adopted to find the optimal solution of alloy ingredients.

Keywords: Alloy Element Yield; Optimization Scheme; Principal Factor Analysis; Multiple Linear Regression Analysis; Linear Programming Model

China Steel Association expects that China's steel demand will be 890 million tons in 2020, increasing 2% year-on-year. In 2020, supply contraction may be greater than demand contraction, which forms a better support for the steel markets. Optimization research of deoxidation alloying is crucial. For different steel types, different types and amounts of alloys are added at the end of melting to make the elements of the alloy reach the standard. The finished steel meets certain standards and requirements in some physical properties. Reducing production costs is a key factor for iron and steel enterprises to improve their competitiveness. Therefore, the optimization of the deoxidation alloying process has become a key research issue for relevant personnel. This paper intends to establish relevant mathematical models for the deoxidation alloying link through historical data, optimize the number and type of alloys And ensure the maximum production cost of alloys while ensuring that the molten steel meets the standards[1].

1. Related research

Steel is the foundation of industrial production, and economists usually use steel output as one of the important indicators to measure the economic strength of a country. Therefore, the research of smelting steel by relevant personnel is quite rich. Fang Yue and Gao Zhen obtained the influencing factors of the element yield during the "deoxidation alloying" process of iron and steel through principal component analysis[2]. Cheng Ruonan, Wang Ruimei, *et al.* used the Pearson correlation coefficient to study the relationship between different related factors and element yields, and used neural networks to predict element yields, and reduced the cost of alloy ingredients by establishing multivariate linear programming[3].

doi: 10.18282/ims.v3i1.310

2. Research on influencing factors of element yield

2.1 Research ideas

In this paper, based on the historical data of the Mathhorcup D question of steel deoxidation alloying ingredients optimization in 2019, the element yield is calculated by the element yield calculation formula, and the influencing factors are obtained through the principal component analysis.

2.2 Calculation of element yield and analysis of influencing factors

2.2.1 Calculation of element yield

2.2.1.1 Modeling ideas

According to the historical data of Annex I of Mathhorcup D in 2019 and the description of commonly used alloy compositions in Annex II, through the element yield calculation formula: element yield = (mass of alloy elements into steel / total mass of alloy elements added) × 100%.

Among them, the quality of the alloy elements entering the steel is related to the continuous casting positive sample, the converter end point, the molten steel quality and the total alloy mass, that is, the quality of the alloy elements entering the steel = (molten steel quality + alloy total mass) * continuous casting positive sample-converter end point * Quality of molten steel; the total mass of alloy elements added is related to the mass of each alloy and the content of elements in each alloy, for example: the amount of C element added = $\sum(M_{alloy}*W_C)$.

2.2.1.2 Model solution

Using excel software, according to the definition of the yield rate of the two elements C and Mn, the specific historical yield rate of the two elements C and Mn is calculated.

Heat number	C element yield rate	Mn element yield rate	Heat number	C element yield rate	Mn element yield rate
7A06267	0.990459628	0.999948679	7A06555	0.864667944	0.525947201
7A06341	0.897526848	0.999887696	7A06301	0.869129409	0.522311972
7A06274	0.912306765	0.999649762	7A06443	0.886713661	0.5078459
7A06132	0.951261368	0.999347456	7A06296	0.902594353	0.506947012
7A06610	0.932282571	0.999071351	7A06744	0.752114961	0.469346953

Table 1. Partial historical yield rates of C and Mn elements

2.2.2 Influencing factors of element yield

2.2.2.1 Modeling ideas

First of all, data processing is performed to remove 386 samples of the wrong yield rate obtained from the previous question from the 810 yield rate samples; then 424 samples of the reliable yield rate in the last step are re-acquired; SPSS software is adopted to perform factor analysis. Through dimensionality reduction and principal component analysis, the main factors affecting the yield of C and Mn elements can be found out.

2.2.2.2 Establishment of the model

There are n cases, with p indicators observed in each case. There is a strong correlation between the p indicators. In the future, it is easy to study, standardize the sample observation data, so that the average value of the variable after the standard dialect is 0, and the variance is 1. Here, the original variable and the normalized variable vector are represented by X, and F1, F2, ..., Fm (m < p) represents the standardized common factor.

The factor model is:

$$\begin{cases} X_1 a_{11} F1 + a_{12} F_2 + \cdots\cdots + a_{1m} F_m + \xi_1 \\ X_2 = a_{12} F_1 + a_{12} F_2 + \cdots\cdots + a_{1m} F_m + \xi_2 \\ \vdots \\ X_p = a_{p1} F_1 + a_{p2} F_2 + \cdots\cdots + a_{pm} F_m + \xi_p \end{cases}$$

The matrix form of the factor model is: $X = AF + \xi$

Where F is the main factor, ξ is the special factor, and A is the factor loading matrix.

$$\begin{pmatrix} A_{11} & a_{12} & a_{13} \cdots & a_{1n} \\ A_{21} & a_{22} & a_{23} \cdots & a_{2n} \\ \vdots & \vdots & \vdots \quad \vdots & \vdots \\ A_{p1} & a_{p2} & a_{p3} \cdots & a_{pr} \end{pmatrix}$$

After obtaining the factor model, since its public factors may not necessarily reflect the essential characteristics of the problem, it is better to explain the actual meaning of each public factor in the future and reduce the subjectivity of the explanation. The matrix form of the rotated factor model is: $X = A_1F' + \xi$ (F' is the rotation matrix)

Factor score: Factor analysis is to express the variable as a linear combination of public factors. At this time, the public factor can also be expressed as a linear combination of variables, which is the factor score coefficient, as follows:

$F_j = \beta_{j1}X_1 + \beta_{j2}X_2 + \ldots + \beta_{jp}X_p$ $(j = 1,2,\ldots, m)$

Where β_{jp} is the score of the jth common factor on the pth original variable.

2.2.2.3 Solving the model

It can be seen from the matrix of correlation coefficients of the initial variables that the correlation coefficients among multiple variables are larger, and the corresponding significance is generally smaller, and the significance is less than 0.05, which is necessary for factor analysis.

KMO sampling suitability measure		.542
Bartlett sphericity test	Approximate chi-square	429.909
	Degrees of freedom	36
	Distinctiveness	.000

Table 2. KMO and Bartlett inspection table

Table 2 shows the KMO test and the spherical Bartlett test table. It is generally considered that KMO statistics greater than 0.5 are acceptable. In this case, KMO statistics are 0.542, which is acceptable. The significance of Bartlett's test is 0.000, which is less than 0.01, showing that there is a significant correlation between various variables, that is, the null hypothesis that the correlation matrix is rejected as the unit matrix.

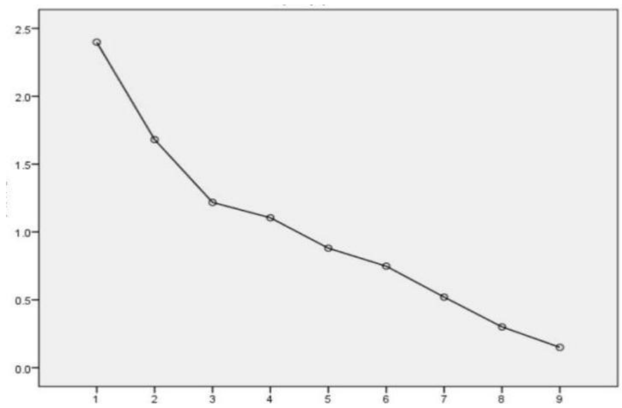

Figure 1. Gravel.

Figure 2 is a gravel chart regarding initial eigenvalues (variance contribution rate). Observation shows that the declining trend of the eigenvalue after the fourth public factor accelerates, so it is more appropriate to select 4 public factors.

	Ingredient			
	1	2	3	4
Converter end temperature	-.141	.356	.323	.444
Converter end point C	-.514	.000	.159	.005
Converter end point Mn	-.110	.458	.026	-.220
Converter end point S	.172	-.045	-.225	.528
Converter end point P	.147	.548	-.115	.089
Converter end point Si	.116	.003	-.057	-.480
Net weight of molten steel	-.137	-.055	.591	-.113
C element yield	.063	.009	.450	.046
Amount of added C element	.424	.017	.057	-.054

Table 3. Matrix table of component score coefficients

Table 3 is the matrix of factor score system, from which the final factor score formula can be known as:

F1 = -0.141 × converter end temperature-0.514 × converter end point C－0.11 × converter end point Mn + 0.172 × converter end point S + 0.147 × converter end point P + 0.116 × converter end point Si －0.137 × molten steel net weight + 0.063 × C Rate + 0.424 × amount of added C element;

F2 = 0.356 × converter end temperature + 0.458 × converter end point Mn－0.045 × converter end point S ＋ 0.548 × converter end point P ＋ 0.003 × converter end point Si－0.055 × molten steel net weight + 0.009 × C element yield + 0.017 × add The amount of C element;

F3 = 0.323 × converter end temperature + 0.159 × converter end C + …… + 0.057 × added amount of C element;

F4 = 0.444 × converter end temperature + 0.005 × converter end C + …… －0.054 × the amount of added C element.

In order to study the total utility, the scores of the four common factors can be weighted and summed, and the weight is the variance contribution rate corresponding to the common factors. In this example, the variance contribution rate is used as the value. The variance contribution rates of the four rotated common factors are 22.262%, 17.652%, 16.754%, and 14.456%, so the total score formula is: ZF = 22.262% × F1 + 17.652 % × F2 + 16.754% × F3 + 14.456% × F4.

The final calculation result is as follows: ZF_C = 0.14975176 × converter end temperature-0.08706502 × converter end point C + 0.0289108 × converter end point Mn + 0.06897842 × converter end point S + 0.12305684 × converter end point P－0.0525851 × converter end point Si + 0.04247332 × molten steel net weight + 0.0976565 × C element yield + 0.09913526 × C element added.

Similarly, the final calculation score formula of Mn is: ZF_{Mn} = 0.19023095 × the amount of added Mn element + 0.22484376 × net weight of molten steel + 0.15732321 × the yield of Mn element + 0.07825331 × converter end point P + 0.09742162 × converter end temperature-0.14924826 × Converter end point S + 0.1405504 × converter end point Mn-0.0064781 × converter end point Si + 0.11916059 × converter end point C

From the results of SPSS analysis, it is concluded that the order of influencing factors affecting the yield of element C of the alloy is: (1) converter end temperature; (2) converter end P; (3) amount of added C element; (4) converter end C; (5) converter end S; (6) Si at the end of the converter; (7) Net weight of molten steel; (8) Mn at the end of the converter.

The order of influence of the factors affecting the yield of alloy Mn elements is as follows: (1) net weight of molten steel; (2) added Mn element; (3) converter end point S; (4) converter end point Mn; (5) converter end point C; (6) converter end point temperature; (7) converter end point P; (8) End point of converter Si.

3. Prediction of element yield and improvement of the model

3.1 Research ideas

Predict the yield of alloying elements C and Mn: find the relationship between the yields of alloying elements C and Mn and their impact factors, using the four impact factors ranked in the top four as variables, using Eviews software for multiple linear regression analysis; use these two models to predict the yield of alloying elements C and Mn.

To further improve the prediction model, Excel's trend line function is used to derive a variety of functional relationships; the F test and t test are adopted to continuously filter out the variables that do not meet the requirements; finally the optimal model is obtained to improve its prediction accuracy .

3.2 Prediction of element yield rate-multiple linear regression model

3.2.1 Model establishment

The yield of alloy element C: $\eta_c = C + a_1 * x_1 + a_2 * x_2 + a_3 * x_3 + a_4 * x_4$

Among them, X1: converter end temperature; X2: converter end point C; X3: converter end point P; X4: added C amount; a1, a2, a3, a4 are coefficients.

Yield of alloying element Mn: $\eta_{Mn} = Mn + a_5 * x_5 + a_6 * x_6 + a_7 * x_7 + a_8 * x_8$

Among them, X5: the amount of Mn added; X6: net weight of molten steel; X7: converter end point S; X8: converter end point Mn; a5, a6, a7, a8 as coefficient.

3.2.2 Solving the model

3.2.2.1 Prediction model of alloy element C

$$\eta_c = 0.571724 + 0.0000143 * x_1 + 63.93576 * x_2 - 109.0404 * x_3 + 0.002485 * x_4$$

S.E=0.056403 means that the average error between the actual observation point of the alloy element C yield and its estimated value is 0.056403, and the error is less, indicating that the model is well established.

Perform the F test and t test at the 5% significance level:

F inspection: F=19.69457>$F_{0.05}(4,419)\approx 2$

t test: $|t(\hat{a}_1)| = |1.377168| > t_{0.025}(419)\approx 0.9$; $|t(\hat{a}_2)| = |1.729304| > t_{0.025}(419)\approx 0.9$;

$|t(\hat{a}_3)| = |-1.790509| > t_{0.025}(419)\approx 0.9$; $|t(\hat{a}_4)| = |6.375110| > t_{0.025}(419)\approx 0.9$;

It shows the end temperature of the converter, the amount of added C element; the end point of the converter C has a significant effect on the yield of alloying element C, while the end point of the converter has no significant effect on the yield of alloying element C.

3.2.2.2 Prediction model of alloy element Mn

$$\eta_{Mn} = 0.365235 + 0.000520 * x_5 + 0.0000017 * x_6 53.676548 * x_7 - 34.23206 * x_8$$

S.E=0.047705 shows that the average error between the actual observation point of the alloy element Mn yield and its estimated value is 0.047705, the error is small, indicating that the model is well established.

Perform the F test and t test at the 5% significance level:

F inspection: F=21.60956>$F_{0.05}(4,419)\approx 2$

t test: $|t(\hat{a}_5)| = |4.957343| > t_{0.025}(419)\approx 0.9$; $|t(\hat{a}_6)| = |0.962281| > t_{0.025}(419)\approx 0.9$;

$|t(\hat{a}_7)| = |-0.819896| < t_{0.025}(419)\approx 0.9$; $|t(\hat{a}_8)| = |-1.957713| > t_{0.025}(419)\approx 0.9$;

It indicates the net weight of molten steel; the added Mn element and converter end Mn have a significant effect on the yield of alloying element Mn, while the converter end S has no significant effect on the yield of alloying element Mn.

3.3 Improvement of the prediction model of element yield[4-8]

3.3.1 Modeling ideas

Use Excel's icon trend line function to derive a variety of functional relationships, use F test and t test to continuously filter the variables that do not meet the requirements, and finally get the optimal model.

3.3.2 Solving the model

$$\eta_c = -58.29167 + 1.57 * 10^{-21} * x_1^6 - 1.3 * 10^{16} * x_2^6 - 8.71 * 10^{18} * x_3^6 - 4.16 * 10^{-11} * x_4^6 + 2.79 * 10^{-8} * x_4^5 - 7.63 * 10^{-6} * x_4^4 + 0.001093 * x_4^3 - 0.086189 * x_4^2 + 3.539288 * x_4$$

$$\eta_{Mn} = 0.803211 + 1.02 * 10^{-19} * x_5^6 + 2.35 * 10^{-16} * x_6^3 - 40.15641 * x_8$$

3.3.3 Comparison of models

3.3.3.1 Prediction model of the yield of alloying

element C

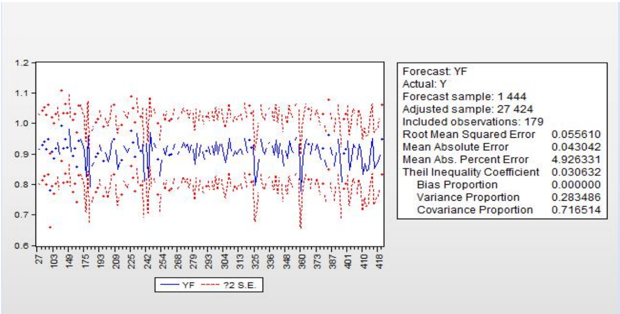

Figure 2. Initial model.

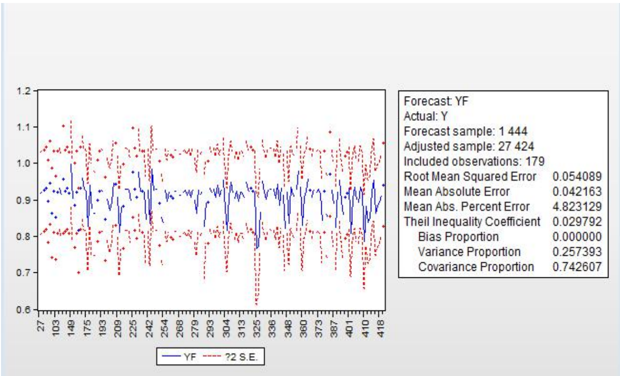

Figure 3. The prediction map of the model after modification.

A. Comparison of regression coefficients:

a. After the modification, the coefficient of X_1 is negative (X_1 represents the end point C of the converter), which is consistent with the expected result. The higher the end point C of the converter, the lower the oxidizability of the molten steel and the lower the element C of the alloy absorbed by the molten steel, otherwise, the higher.

b. The coefficient of X_4 after modification has positive and negative, and X_4 is a polynomial, so it can not be accurately compared. However, the overall prediction accuracy of the model has been improved to a certain extent.

B. Goodness of fit:

a. The R^2 value of the original model is 0.311650, $\overline{R^2}$ = 0.295826, indicating the converter end temperature, the converter end C, the converter end P; the ability of the added C to explain the yield of alloying element C is 31.165%. The modified sample determination coefficient indicates that the interpretation ability is 29.5826%.

b. The R^2 value of the modified model is 0.3488, and $\overline{R^2}$ = 0.314120, indicating the modified variable end temperature of the converter, the end point of the converter C, the end point of the converter P; 34.88%, the modified sample decision coefficient indicates that the interpretation ability is improved to 31.4120%.

C. t test: ｜t（â）｜>$t_{0.025}(419)\approx 0.9$, all meet the conditions.

D. Coefficient estimation error: The coefficient error of the modified model is generally smaller than that of the model before modification, and the model optimization is successful. The overall error S.E = 0.055666 <0.056403, so the overall error is reduced.

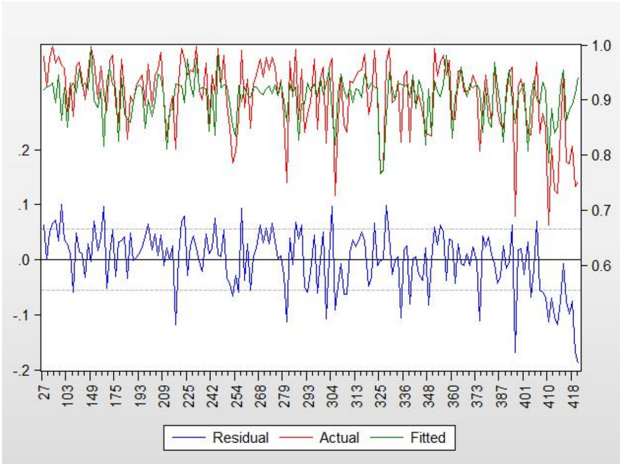

Figure 4. The residual distribution map of the modified model.

E. Observation and analysis of residual distribution:

a. In the residual distribution table, most of the residuals in each period fall within the dotted frame of + ô, and the model fits well.

b. There is no regularity in the overall residuals, indicating that the model fits well. Occasionally large residuals are caused by missing data in the corresponding sample.

4. Research on the optimization of the supporting plan for the "deoxidation alloying" of molten steel

4.1 Research ideas

Because the prices of different alloys are different, their choice directly affects the cost of molten steel deoxidation alloying. According to the prediction results of alloy yield and the prices of different alloys, the optimization calculation of molten steel deoxidation alloying cost is realized, and the alloy batching plan is given. By establishing the objective function of the lowest cost and writing the constraint conditions, the lowest cost is solved by the linear programming model and the optimization principle, etc., and the alloy batching scheme is given[9].

4.2 Optimization research-linear programming model[10-19]

4.2.1 Model establishment

4.2.1.1 Selection of decision variables

Suppose the element to be alloyed is i, and there are n in total. The types of alloys for adjusting C and Mn are m. The addition amount of each alloy is x_1, x_2……x_m is the decision variable.

$X= (x_1,x_2……x_m)^T$

Obviously, the decision variable is non-negative.

The n independent variables of the decision variables constitute the n-dimensional Euclidean space En, that is, any point in the n-dimensional space corresponds to a group of XI (I = 1, 2,…, n) represents an alloy addition scheme.

4.2.1.2 Establish the objective function

Aiming at the lowest cost of alloy ingredients

$$\min Z=\sum_{j=1}^{m} p_j x_j$$

Among them, P_j is the price of the j-th alloy, the unit is yuan/ton; Z is the total cost of the alloy ingredients, the unit is yuan.

4.2.1.3 Determine the constraints

a. Constraints on the target composition of molten steel

$$LL_i \le \frac{\sum_{j=1}^{m} F_{ij}\eta_i x_j + F_{i,0}M_{Molten\,steel}}{\sum_{j=1}^{m} x_j + M_{Molten\,steel}} \le UL_i$$

i = 1, 2,…, n, there are n inequalities, j = 1,2,…, m, there are m inequalities. In the formula, F_{ij} is the content of element i in the j-th alloy, unit%; η_i is the yield of element i, unit%; $F_{i,0}$ is the original content of element i molten steel before alloying, unit%; $M_{Molten\,steel}$ is the quality of molten steel, unit t; LL_i is the lower limit of the composition of element i, unit%; UL_i is the upper limit of the composition of element i, unit%.

b. Non-negative constraints

$X_j \ge 0$ j =1,2 , …, m

4.2.2 Solving the model

According to the classification of different steel types, select the historical normal smelting furnace times, and replace the average value of the number of normal yields of each steel type with a large number of normal furnace times. And take the average molten steel as the molten steel weight.

Use LINGO software to get the following results:

"Objective value: 0.2662077E + 10" indicates that the optimal target value is 0.2662077 * 10 ^ 10 yuan.

"Total solver iterations: 2" means that the simplex method is used to iterate twice to obtain the optimal solution.

"Value" gives the value of each variable in the optimal solution. That is, the amount of silicon manganese surface (silicon manganese slag) is 299853.5 tons, and the amount of petroleum coke carburizer is 105884.1 tons.

"Reduced cost" lists the coefficients of the variable in the row of the discriminant number in the optimal simplex table, indicating the rate of change of the objective function when the variable has a slight change. The reduced cost value of the base variable should be 0. The corresponding Reduced Cost value represents the amount by which the objective function increases when xi increases by one unit (assuming other single non-base variables remain unchanged). In this model, the Reduced Cost value corresponding to x5 and x8 is 0, indicating that when there is a slight change in x5 and x8, the value of the objective function does not change.

"Dual Price" lists the coefficients of the slack variables in the row of the discriminant number in the optimal simplex table, that is, the change rate of the objective function when the constraint conditions change slightly, and each corresponding constraint in the output result has A dual price (shadow price).

5. Conclusion

This paper analyzes the four main factors that affect the yield of C element: (1) converter end temperature; (2) converter end point P; (3) amount of added C element; (4) converter end point C; four main factors that affect Mn element yield are also analyzed: (1) net weight of molten steel; (2) added Mn element; (3) converter end point S; (4) converter end point Mn. So the input of these main factors should be closely controlled during steel-making to make the yield close to the predicted value.Based on optimized logistic regression, the model concludes that the end temperature of the converter, the amount of added C element, and the end point of the converter C have a significant effect on the yield of alloying element C. The steel plant should allocate the investment ratio of each major factor to achieve cost reduction and efficiency increase, thereby achieving the goals of cost saving and long-term development of the steel plant.

References

1. Qian Z, Ge L, Wang W. Research on optimization of converter alloy composition based on grey correlation method (in Chinese). Journal of Fuyang Teachers College (Natural Science Edition) 2020; 37(1): 11-16.
2. Fang Y, Gao Z. Model of influencing factors of element yield based on principal component analysis (in Chinese). China Metal Bulletin 2019; (6): 140-141.
3. Cheng R, Wang R, Ren C, *et al*. Optimization of the "deoxidation alloying" batching scheme for molten steel (in Chinese). Energy Saving 2020; 39(2): 86-87.
4. Ren X, Zhu R, Yao Q, *et al*. Research on the yield of molten steel "deoxidation alloying" based on multiple linear regression (in Chinese). Computer Programming Skills and Maintenance 2020; (3): 23-25+70.
5. Jiang H. Research on the prediction model for the yield of molten steel "deoxidation alloying" (in Chinese). World Nonferrous Metals 2020; (1): 180-181.
6. Zhou S, Dong C, Zhang Z, *et al*. Research on the prediction of deoxidized alloying composition based on fuzzy programming model (in Chinese). Scientific Consulting (Science and Technology Management) 2019; (10): 84.
7. Tan Y, Wu J, Li K. Research on the prediction of metal element yield based on gray neural network model (in Chinese). Information Recording Materials 2019; 20(9): 40-41.
8. Li M, Wu Y, Guo W. Research on the yield of alloy in molten steel deoxidation alloying. Chinese and Foreign Entrepreneurs 2019; (21): 87.
9. Du J, Han C, Cai K. Predictive control for mold level of continuous casting based on least squares support vector machine. Journal of Iron and Steel Research 2007; (8): 28-31.
10. Yang G. Mathematical modeling (in Chinese). Shanghai: Shanghai University of Finance and Economics Press; 2015.
11. Fan H, Feng Z, Yu R. Optimization of the scheme cost of "deoxidized alloying" in steel and water (in

Chinese). Yunnan Chemical Industry 2020; 47(2): 45-46.
12. Liu P, Zhu J, Gao Z, *et al*. Analysis of deoxidation alloying ingredients based on BP neural network (in Chinese). Journal of Qilu University of Technology 2019; 33(5): 74-80.
13. Zhang T, Zhu J, Wang H, *et al*. Research on optimization of batching scheme for molten steel deoxidation alloying based on fuzzy linear programming (in Chinese). Journal of Qinghai University 2019; 37(5): 73-81.
14. Dai Y, Xie W, Chen J, *et al*. Analysis and research of deoxidized alloying ingredients based on gray correlation (in Chinese). Scientific Consulting (Science and Technology Management) 2019; (9): 31.
15. Dong C, Gong X, Xu C, *et al*. Research on the prediction of deoxidation alloying composition based on BP neural network (in Chinese). Scientific Consulting (Science and Technology Management) 2019; (9): 45.
16. Zhou H, Zhang Q, Han X, *et al*. Optimized design of deoxidation alloying scheme for converter smelting (in Chinese). World Nonferrous Metals 2019; (13): 12-13.
17. Liu F, Zhou Y, Tang J, *et al*. Optimization of the "deoxidation alloying" batching scheme for molten steel (in Chinese). Modern Computer 2019; (22): 8-13.
18. Wang M. Research on the optimization of the "deoxidation alloying" batching scheme for molten steel (in Chinese). Think Tank Times 2019; (31): 292+296.
19. Zheng T. Optimization of batching scheme for molten steel "deoxidation alloying" (in Chinese). Think Tank Times 2019; (29): 287+294.