Study on AlloYing Yield in Deoxidization and AlloYing of Iron and Steel

Chenyang Lin1, Shuaijun Liu1 and Linli Zhang1*

1 College of Applied Science and Technology, Hainan University, Danzhou, Hainan, 571737, China

*Corresponding author’s email: 583316703@qq.com.

Abstract. With the development of the iron and steel industry, the automation of iron and steel smelting batching scheme has become the research focus of the present. In this paper, Grey correlation model is established and the ranking of the factors influencing the yield of alloy is obtained. By establishing the stepwise regression model, the predicted value of alloying yield is got. The predicted value and the actual value of alloying yield are compared by MATLAB, which proves that the predicted value of the model is accurate and the accuracy is high. The model and conclusions are of guiding significance to the production and smelting of alloy steel.

1. Introduction

With the rapid development of the global economy in nowadays, the demand for steel is increasing dramatically as the basic material of various industries. Along with the rapid development of artificial intelligence, the pursuit of intelligent iron and steel smelting to achieve the optimization of proportioning scheme, the prediction of alloying yield and the optimization algorithm research have become the hotspot at this stage.

At present, the commonly used methods of alloying yield include BP Neural Network [1-2], Genetic Algorithm [3-4], and Support Vector machine [5], get the yield prediction combining with computer simulation technology. In order to improve the accuracy of the prediction of the alloying element yield, this paper establishes a prediction model based on the grey correlation degree and the multiple regression model. Through the MATLAB to solve the historical data, the prediction value and the actual value are basically the same, and the model has high practicability. The effective prediction value of alloy yield provides a reference value for the ingredients of alloy yield in iron and steel industry, which can realize the optimization of batching scheme in iron and steel smelting.

2. Analysis of the yield of C and Mn elements based on the method of grey correlation method

In the process of alloy steel smelting, the yield of alloy elements is affected by many factors, and the degree of correlation is also different. Through the historical data of alloy steel HRB400B smelting, the alloy yield of C and Mn elements is calculated. The grey correlation model was established to analyze the relevant factors, such as converter end point temperature, converter end point C, net weight of molten steel, vanadium nitrogen alloy (imported), vanadium iron (FeV50-B), calcium aluminosilicate, silicon manganese surface (silicon manganese slag), petroleum coke carburizer, manganese silicon alloy (qualified block), silicon carbide (55%), and the correlation coefficient of the factors affecting the alloy yield have been obtained.
2.1 The calculation method of alloying element yield
In the smelting process of alloy steel, one or several different kinds of alloys should be added to the molten steel to make it meet the requirements of the component specifications of the finished alloy steel. In the process of “deoxidation and alloying” of molten steel, some of the alloy ingredients are used for deoxidation of molten steel and converted into deoxidation products, while the other part is absorbed by steel and plays a role in deoxidation and alloying of steel [6]. Alloy yield refers to the ratio of the weight of the alloy element absorbed by the molten steel during deoxidization and the sum of the total mass of the added element. The calculation formula is as follows:

$$\eta = \frac{G_n \times (\eta_1 - \eta_2)}{\sum G_n \times \eta_0} \times 100\%.$$  

(1)

Note:
- \(\eta_1\) and \(\eta_2\) refer to the percentage of an element in the mass of alloy steel;
- \(\eta_0\) refers to the percentage of weight of alloy material;
- \(G_n\) refers to the weight of an alloy material added to this set of data;
- \(\sum G_n\) refers to the sum of the mass of an element contained in each alloy.

2.2 Analysis of the main factors influencing the yield rate by the method of grey correlation degree

2.2.1 Standardized data. Initialize the data of the alloying yield calculated, and exclude the data groups that are not practical, which includes the data group with the yield of alloying element greater than 1, and the data group with zero alloy material (If such alloy material is not added, the correlation degree is recorded as zero). The data sheet of \(C, Mn\) element yield and alloy material of alloy steel HRB400B is shown in Table 1 and Table 2.

| Table 1. Data sheet of yield of element \(C\) in alloy steel HRB400B. |
|---|
| Furnace NO | Converter end point temperature | Converter end point C | Net weight of molten steel | Vanadium nitrogen alloy (imported) | Si\% | FeV\%B | Calcium alumino-silicate | Silicon manganese surface (silicon manganese slag) | Petroleum coke carburizer | Manganese alloy FeMn65Al18 (qualified block) | Silicon carbide (55%) | C element yield% |
| 7A66689 | 1658 | 0.00043 | 74500 | 5 | 41 | 75 | 220 | 85 | 130 | 132 | 0.98495 |
| 7A66686 | 1662 | 0.00005 | 81350 | 5 | 42 | 75 | 220 | 85 | 140 | 132 | 0.976802 |
| 7A66687 | 1660 | 0.00042 | 73850 | 5 | 40 | 50 | 200 | 85 | 150 | 132 | 0.914899 |
| 7A66682 | 1651 | 0.00041 | 70100 | 5 | 40 | 50 | 200 | 90 | 130 | 132 | 0.849099 |
| 7A66681 | 1679 | 0.00041 | 70100 | 4 | 38 | 50 | 200 | 90 | 130 | 132 | 0.928694 |
| 7A66680 | 1618 | 0.00054 | 67400 | 7 | 44 | 50 | 200 | 102 | 145 | 132 | 0.715869 |
| 7A66689 | 1640 | 0.00067 | 64700 | 4 | 40 | 25 | 200 | 75 | 120 | 132 | 0.818724 |
| 7A66688 | 1662 | 0.00067 | 69800 | 7 | 44 | 25 | 200 | 75 | 143 | 132 | 0.815898 |
| 7A66684 | 1674 | 0.00070 | 74800 | 11 | 32 | 25 | 200 | 64 | 124 | 132 | 0.899193 |

| Table 2. Data sheet of yield of element \(Mn\) in alloy steel HRB400B. |
|---|
| Furnace NO | Converter end point temperature | Converter end point C | Net weight of molten steel | Vanadium nitrogen alloy (imported) | Si\% | FeV\%B | Calcium alumino-silicate | Silicon manganese surface (silicon manganese slag) | Petroleum coke carburizer | Manganese alloy FeMn65Al18 (qualified block) | Silicon carbide (55%) | Mn element yield% |
| 7A66689 | 1658 | 0.00011 | 74500 | 5 | 41 | 75 | 220 | 85 | 130 | 132 | 0.879983 | 0.12 |
| 7A66688 | 1662 | 0.00016 | 81350 | 5 | 42 | 75 | 220 | 85 | 140 | 132 | 0.859754 | 0.09754 |
| 7A66687 | 1660 | 0.00013 | 73850 | 5 | 40 | 50 | 200 | 85 | 150 | 132 | 0.75803493 |
| 7A66682 | 1651 | 0.00011 | 70100 | 5 | 40 | 50 | 200 | 90 | 130 | 132 | 0.79064255 |
| 7A66681 | 1679 | 0.00013 | 70100 | 4 | 38 | 50 | 200 | 90 | 130 | 132 | 0.86580277 |
| 7A66680 | 1618 | 0.00012 | 67400 | 7 | 44 | 50 | 200 | 102 | 145 | 132 | 0.73207375 |
| 7A66689 | 1643 | 0.00011 | 64700 | 4 | 40 | 25 | 200 | 75 | 120 | 132 | 0.78216949 |
| 7A66688 | 1662 | 0.00012 | 69800 | 7 | 44 | 25 | 200 | 75 | 143 | 132 | 0.767444 |
| 7A66684 | 1674 | 0.00008 | 74800 | 11 | 32 | 25 | 200 | 64 | 120 | 132 | 0.94687797 |

2
2.2.2 Grey correlation analysis of the yield of C element in alloy steel HRB400B. The data in table 1 is regarded as the data matrix of grey correlation degree about the yield of C element in alloy steel HRB400B, assuming the yield of C element in a given sequence:

\[ y_c = \{y_1, y_2, y_3, \ldots, y_n\} \]  

As a reference sequence, various alloy materials added to the alloy steel HRB400B are used as a comparison sequence:

\[ x_m = \{x_{m1}, x_{m2}, x_{m3}, \ldots, x_{mi}\} \]  

called

\[ \xi(k) = \frac{\min \min |y_c(n) - x(n)| + \rho \max \max |y_c(n) - x(n)|}{|y_c(k) - x(m)| + \rho \max \max |y_c(n) - x(n)|} \]  

is the correlation coefficient of the comparison sequence \( x_m \) to the reference sequence \( y_c \) at \( k \); Among them, \( \rho \in [0, 1] \) is the resolution coefficient, \( \min \min |y_c(n) - x(n)| \) and \( \max \max |y_c(n) - x(n)| \) are the two-stage minimum difference and the two-stage maximum difference [7].

In general, the larger the resolution coefficient \( \rho \) is, the larger the resolution will be; the smaller the resolution coefficient \( \rho \) is, the smaller the resolution will be. A correlation factor corresponds to the correlation degree of the yield of an element, and the information is too scattered for comparison. In order to reflect the correlation index between a certain set of data and the collection rate of elements, this paper concentrates the values of each group of correlation degree at each time into an average value, and processes the scattered values in a centralized way. The following relationships are obtained:

\[ r_i = \frac{1}{n} \sum_{k=1}^{n} \xi(k) \]  

is the correlation degree of sequence \( x_m \) to reference sequence \( y_c \).

2.2.3 Calculating the correlation coefficient of grey correlation degree model by MATLAB. Take the data in table 1 as the text data, select the element yield of C as the reference sequence, and the addition amount of various alloy materials as the comparison sequence. Here \( \rho = 0.5 \). The gray correlation degree of the data is calculated by MATLAB. In order to facilitate the analysis of the influence of various alloy materials, the correlation coefficient is sorted. The operation result is shown in figure 1.

![Figure 1. The result chart of grey correlation degree of C element yield in alloy steel HRB400B.](image)

Note: the number 1 represents Converter end point temperature, 2 represents Converter end point C, 3 represents Net weight of molten steel, 4 represents Vanadium nitrogen alloy (imported), 5 represents FeV50-B, 6 represents Calcium aluminosilicate, 7 represents Silicon manganese surface (silicon manganese slag), 8 represents Petroleum coke carburizer, 9 represents Manganese silicon alloy (qualified block) and 10 represents Silicon carbide (55%).

It can be seen from figure 1 that among the factors affecting the yield of element C in alloy steel HRB400B, silicon manganese surface (silicon manganese slag) is the primary factor, followed by the
The net weight of molten steel of alloy steel, followed by manganese silicon alloy (qualified block), petroleum coke carburizer, ferrovanadium (FeV50-B), converter end point temperature, silicon carbide (55%), vanadium nitrogen alloy (imported), Calcium aluminosilicate, converter end point C. In order to control the alloying yield of C element to reach the qualified index, the addition amount of silicon manganese surface (silicon manganese slag), manganese silicon alloy and petroleum coke carburizer should be strictly controlled.

The similar method is adopted, and the grey correlation coefficient of Mn yield of alloy steel HRB400B is calculated by MATLAB, as shown in table 3.

Table 3. Calculation of correlation degree of Mn recovery in alloy steel HRB400B.

| correlation coefficient | $r_{M1}$ | $r_{M2}$ | $r_{M3}$ | $r_{M4}$ | $r_{M5}$ |
|------------------------|---------|---------|---------|---------|---------|
| value                  | 0.9399  | 0.8017  | 0.9407  | 0.6776  | 0.8470  |

| correlation coefficient | $r_{M6}$ | $r_{M7}$ | $r_{M8}$ | $r_{M9}$ | $r_{M10}$ |
|------------------------|---------|---------|---------|---------|----------|
| value                  | 0.6039  | 0.8766  | 0.8383  | 0.9098  | 0.8802   |

Table 3 shows that the main factors affecting the Mn yield of alloy steel HRB400B are the net weight of molten steel and the converter end point temperature. The main alloy materials affecting the yield are manganese silicon alloy (qualified block), silicon carbide (55%), silicon manganese surface (silicon manganese slag).

The above analysis shows that in the process of deoxidization and alloying of alloy steel, the above factors should be strictly controlled to improve the yield efficiency of alloy elements.

3. Prediction of recovery of C and Mn by stepwise regression analysis model

In this paper, the ranking of factors affecting the yield of C and Mn elements has been obtained by means of grey correlation analysis. In order to make variables provide significant additional explanatory information for the yield of elements, so as to improve the fitting degree of prediction model, the factors affecting the yield of elements are screened by means of stepwise regression analysis. Considering the complexity of the problem and the diversity of data, only HRB400B steel is analyzed.

3.1 Establishment of prediction model of stepwise regression analysis

The several factors influencing the recovery of C and Mn elements are set as independent variable $x_1, x_2, ..., x_m$, and the recovery of elements is set as $y$. The recovery of elements after deoxidization and alloying of molten steel is calculated, this means that the multiple regression model of $y$ and $x_1, x_2, ..., x_m$ is determined, that is, the problem is solved by stepwise regression analysis.

3.2 Solution to the prediction model of stepwise regression analysis

When the stepwise regression analysis method is used to solve the model, the independent variable should be selected to enter the model first. After entering the model, the partial $F$, $R^2$ and $R_j^2$ test should be carried out for the independent variable to determine whether the independent variable has significant influence on the dependent variable $y$ and whether it is necessary to enter the model. Repeat the above operations until the variables outside the model can not pass the partial $F$, $R^2$ and $R_j^2$ tests, and finally determine the regression coefficient between the influencing factors and the yield, and the algorithm is terminated [8].

Because of the phenomenon of data missing and data error when collecting smelting data in reality, we use Excel to screen the data of fifteen groups of variables and eliminate useless data, see table 4:

| Furnace number | 1 | 2 | 3 | 4 | 5 | ... | 149 |
|---------------|---|---|---|---|---|-----|-----|
| Converter end point temperature | 1644 | 1543 | 1674 | 1800 | 1660 | ... | 1660 |
| Converter end point C | 0.00065 | 0.00077 | 5E-04 | 0.0004 | 0.0002 | ... | 0.00042 |
3.2.1 Variable selection. In a multiple regression model, there are many factors that affect the dependent variables, so it is necessary to use partial F test to judge whether these variables are necessary to enter the model.

When $F_j > F_α$, it means that the introduction of $x_j$ into the whole model has a significant impact on $y$, and $x_j$ is necessary to enter the model; when $F_j \leq F_α$, it means that the deletion of $x_j$ in the whole model has little impact on $y$, and $x_j$ removes the model. There are also some indicators to measure the degree of correlation between $y$ and $x_1, x_2, \ldots, x_m$, such as the complex decision coefficient $R^2$ of the whole model. The closer $R^2$ is to 1, the greater the degree of correlation between $y$ and $x_1, x_2, \ldots, x_m$ is. Generally, $R^2$ is greater than 0.64 before the relationship is established. In MATLAB, $R^2$ and $F_j$ are often given at the same time. The closer they are, the more accurate the model is.

First, all 15 groups of variables are introduced into the model. In the MATLAB, execute the stepwise command with the data in table 4 to get the step-by-step regression interactive screen as shown in figure 2. Where blue indicates the variables in the model and red indicates the variables removed from the model.

![Figure 2. Interaction picture of the influence of HRB400B carbon yield of stepwise regression analysis model.](image-url)

Then a variable which has the least influence on the element yield is removed from the model, and the statistical calculation results are obtained. According to the statistical calculation results, the criteria of partial $F$ test and the criteria of $R^2$ and $F_j$ test, the data are tested and analyzed to see whether the model changes better. Repeat this step until the variable cannot be eliminated.
After screening, we found that $x_1, x_2, x_3, x_4, x_5, x_6$, namely converter end point Mn, FeV55N11-A, Vanadium Nitrogen Alloy (imported), FeV50-B, Calcium Aluminosilicate, and FeAl30Si25 have no significant effect on the element yield, and the statistical results after eliminating these independent variables are shown in figure 3.

![Figure 3](image-url)

Figure 3. The interaction picture of the influence of HRB400B carbon yield of stepwise regression analysis model.

From the new statistical results, it can be seen that although the size of residual standard deviation $s$ (RMSE), complex judgment coefficient $R^2$ (R-square) and adjusted complex judgment coefficient $\bar{R}^2$ (Adj R-sq.) have little change, but $R^2$ is close to $\bar{R}^2$, and the value is close to 1, $F$ increases significantly, $p < 0.01$, so the stepwise regression model after screening is more reliable.

After selection, the variables of stepwise regression model are $x_1, x_2, x_3, x_4, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}$, i.e. converter end point temperature, converter end point $C$, steel net weight, vanadium nitrogen alloy (imported), ferrovanadium (FeV50-B), silicon aluminum calcium, silicon manganese surface (silicon manganese slag), petroleum coke carburization, manganese silicon alloy (qualified block), manganese silicon alloy (qualified block), silicon carbide (55%), silicon calcium carbide carbon deoxidizer.

### 3.2.2 Determine regression coefficient.

The model can predict the element yield on line according to the type and quantity of the alloy input, which is of great significance for the realization of computer automatic batching. For the variables in the model, the least square method is used to determine the regression coefficient between variables and yield, so as to synthesize a multiple regression model. In the MATLAB environment, using the historical data of the early smelting of low alloy steel, the multiple regression model of carbon yield of the final alloy steel HRB400B is obtained as follows:

$$
 y = 1.0439 + 1.654 \times 10^{-5} x_1 + 478.885 x_2 + 1.201 \times 10^{-3} x_4 + 1.192 \times 10^{-4} x_{10} + 0.056 x_{11} + 0.134 \times 10^{-5} x_{12} - 1.018 \times 10^{-4} x_{13} - 0.0013 x_{14} - 0.0016 x_{15}.
$$

(6)

According to the model, the $C$ element yield of HRB400B steel can be predicted. The same method is used to predict the recovery of Mn.

### 3.3 Error analysis and model improvement

Based on the fact that there are other constraints in reality, some non-human factors have an impact on the yield of alloy, so there are errors in the prediction [9]. The error and accuracy of the predicted values of $C$ and Mn are calculated:

- Prediction accuracy $= 1 - (\text{absolute error/predicted element yield}) \times 100\%$.
- Absolute value of error $= \text{actual value of element yield} - \text{predicted value of element yield}$.
The results show that the maximum error values of HRB400B carbon and manganese are 0.1211 and 0.0788 respectively, and the prediction accuracy of HRB400B carbon and manganese are 96.57% and 97.95% respectively.

It can be seen that the error of HRB400B carbon yield is larger than that of manganese yield. Figure 4 is the line chart of the comparison between the actual value and the predicted value of the carbon yield of HRB400B.

![Figure 4. Line chart of comparison between actual value and predicted value of carbon yield of HRB400B.](image)

In MATLAB, the regression model is analyzed with the command of regress, and the residual and its confidence interval graph are obtained with the command of rcoplot(r, rint), which is used to test whether the data is abnormal. When the residual confidence intervals of the data contain zero, the error value of the model is smaller and the result is more accurate. The point that the residual confidence interval does not contain zero point is the abnormal point, which should be eliminated. Through regression analysis of the model obtained by HRB400B carbon yield, running the corresponding program in MATLAB environment, we can get the interactive picture as shown in figure 5.

![Figure 5. Residual distribution of HRB400B carbon recovery data.](image)

By observing the residual distribution of HRB400B carbon recovery data, it can be seen that in HRB400B carbon recovery data, the residual confidence interval of the 7th, 60th, 62nd, 77th, 102nd and 133th groups of data does not contain zero, which is abnormal data. After the outliers are eliminated and calculated by stepwise regression analysis, a more reliable multiple regression model can be obtained, thus improving the prediction accuracy of the element yield.

The yield of HRB400B carbon element is predicted by stepwise regression analysis, and the error value and average prediction accuracy are calculated. The data of HRB400B manganese element are processed by the same method. After the improvement of the model, the average prediction accuracy of
carbon and manganese elements in HRB400B is 96.97% and 98.49%, and the accuracy of the two elements is increased by 0.4% and 0.54% respectively.

The prediction model mainly considers the influence of converter end point temperature, converter end point C, net weight of molten steel, vanadium nitrogen alloy (imported), vanadium iron (FeV50-B), calcium aluminosilicate, silicon manganese surface (silicon manganese slag), petroleum coke carburizer, manganese silicon alloy, silicon carbide (55%), silicon calcium carbon Deoxidizer on the alloying yield. The model considers many factors and has high prediction accuracy. In actual production, it is of great significance for online prediction of alloying yield, planning of alloy quantity and type, and minimizing the production cost of alloy steel [10]. Applying it to the computer digital automatic batching system can realize automatic deoxidization and alloying, which will bring new batching mode to iron and steel smelting, greatly improve production efficiency and save production cost.

4. Conclusion
At present, in the process of smelting alloy steel, most iron and steel enterprises only rely on the experience of alchemists to judge the amount of alloy materials, which is a serious waste of alloy materials, and the quality of alloy steel is uneven. The deoxidization and alloying in steel-making process is an important process in the iron and steel smelting process. The key step to increase the output of iron and steel smelting industry is to establish mathematical model of deoxidization and alloying link of historical data and accurately predict the alloying yield.

According to the historical data of HRB400B smelting, the alloying yield of C and Mn elements is obtained by the calculation formula of the alloying element yield. The gray correlation degree model is established, and the relevant data is input into MATLAB for analysis, and the influence ranking of different sub factors on the yield of C and Mn elements is obtained.

Among the factors affecting the C element yield of alloy steel HRB400B, silicon manganese surface(silicon manganese slag) is the primary factor, followed by the net weight of molten steel of alloy steel, and then is manganese silicon alloy (qualified block), petroleum coke carburizer, ferrovanadium (FeV50-B), converter end point temperature, silicon carbide (55%), vanadium nitrogen alloy (imported), Calcium aluminosilicate, converter end point C. The above ranking shows that in order to control the alloying yield of C element to reach the qualified index, the addition amount of silicon manganese surface (silicon manganese slag), manganese silicon alloy and petroleum coke carburizer should be strictly controlled.

Through the establishment of multiple linear regression model to predict the alloying yield, and then calculate the amount of alloy added in the process of deoxidization and alloying. The model can accurately predict the yield of alloy elements in accordance with the current furnace conditions. When simulating the historical production data of steel HRB400B, the maximum error values of the yield of carbon and manganese elements are 0.1211 and 0.0788 respectively, and the prediction accuracy of the yield of alloy elements are 96.57% and 97.95% respectively. Compared with the method of directly fitting the relationship between the end-point condition and the amount of alloy added by the Genetic Algorithm, BP Neural Network and Support Vector machine, this method makes full use of the important parameter of multiple regression, improves the prediction accuracy of the model, and proves the practicability of this method in iron and steel smelting.

References
[1] Li, M.L., Fang, Z.Y., Guo, W.Z. (2019) Study on the yield of alloy in deoxidation alloying of molten steel. J. Chinese & Foreign Entrepreneurs, 21: 87.
[2] Li, T.G., Chen, Y., Zheng, W., Ma, Z.Q., Sun, J.P., Mao, Y. (2019) Prediction model of alloy yield based on BP neural network. J. Shanxi Metallurgy, 42(03): 15-16.
[3] Liu, F.R., Yang, L., He, Y.Z., Liu, Y.Y., Ma, X. (2019) Prediction of alloy yield and modeling and solution of optimization of auxiliary recipe. J. Scientific and Technological Innovation, 21: 138-139.
[4] Lv, X.W., Bai, C.G., Qiu, Q.B., Ouyang, Q., Huang, Y.M. (2007) Study on integrated optimization of sintering ingredients based on genetic algorithm. J. Iron & Steel, 04: 12-15+20.

[5] Han, M., Xu, Q., Zhao, Y., Lin, D., Yang, X.L. (2010) Calculation of alloy addition in converter steelmaking based on yield prediction model. J. Steelmaking, 26(01): 44-47.

[6] Gong, W., Jiang, Z.H., Zheng, W., et al. (2002) Control model of alloy composition during converter smelting. J. Journal of Northeastern University (Natural Science), 23(12): 1155-1157.

[7] Chen, Y.L. (2006) Batching control system under blast furnace trough based on PLC. D. Chongqing University.

[8] Wu, S.L., Han, H.L., Xu, S.B., Niu, B. (2007) Study on mathematical model of blast furnace optimization batching. J. Iron & Steel, 9: 19-23.

[9] Bao, Y.P., Zhang, C.J., Wang, M. (2018) Research status and prospect of alloy reduction in steelmaking process. J. Chinese Journal of Engineering, 40(09): 1017-1026.

[10] Xu, M.H. (2015) Optimization and system development of blast furnace ironmaking ingredients. D. Beijing University of Chemical Technology.