A PREDICTION METHOD OF SILICON CONTENT IN HOT METAL OF BLAST FURNACE

In blast furnace smelting, the silicon content in hot metal can indirectly reflect the blast furnace temperature and measure the quality of hot metal. For more accurate prediction, according to the reduction reaction, the input parameters affecting the silicon content are selected to form a data set. The Weighted Random Forest (WRF) prediction model and the Scaling Coefficient Particle Swarm Optimization (SCPSO) algorithm are proposed. The prediction method based on SCPSO-WRF has higher prediction hit rate and lower mean error than those traditional methods. The results show that the prediction hit rate and mean error of SCPSO-WRF are 89,1 % and 0,0291 respectively. The prediction method has theoretical research and practical application value.

Key words: blast furnace, hot metal, silicon, particle swarm optimization, random forest

INTRODUCTION

Blast furnace smelting is an extremely complex process. There is nonlinearity between the parameters. Silicon content can indirectly reflect the blast furnace temperature [1]. If the silicon content can be predicted, it will have reference significance for operators to control the furnace condition. In order to make the prediction results more accurate, researchers have established various prediction models, including Random Forest (RF) [2], Support Vector Machine (SVM) [3], neural network [4], extreme learning machine [5], etc. In [2], the input parameters are selected based on experience, which lacks metallurgical and statistical analysis. The prediction model is RF, and the grid search is used to optimize the model parameters. In [3], the process of blast furnace smelting is analyzed, and the input parameters are selected according to the grey correlation degree. The prediction model is SVM, and the Particle Swarm Optimization (PSO) algorithm is improved to optimize the model parameters. In [4], the input parameters are selected according to experience and correlation coefficient. The prediction model is the recurrent neural network with a long short-term memory structure, and the grid search is used to optimize the model parameters. In [5], the input parameters are selected according to the correlation coefficient. The prediction model is a multiple kernel extreme learning machine, and the improved grey wolf optimizer is used to optimize the model parameters. There are still some problems, such as high prediction error. There are not only the reasons for the selection of input parameters, but also the factors for the prediction model. To fundamentally improve the prediction hit rate, according to the influence of blast furnace gas on reduction reaction, the input parameters are

selected and a high-quality data set is formed. Then a prediction method based on SCPSO-WRF is proposed. The effect is verified by simulation experiments.

PARAMETER SELECTION

In the task of silicon content prediction, the quality of data set determines the hit rate of the model and has a great impact on the prediction results. Selecting appropriate input parameters as input parameters is the basis of high hit rate, and the selection of input parameters is a very important part. Generally, the correlation coefficient between input parameters and silicon content is calculated, and the parameters with large correlation are selected, hoping to infer the value of silicon content according to these parameters. When the input parameters tend to statistical analysis, the generalization performance of the model will be reduced due to the lack of smelting theoretical support.

In this paper, the input parameters are selected according to metallurgy and statistics. The selected input parameters are supported by both theory and data. The reduction of iron oxide by CO and H_2 in low and medium temperature zone is called indirect reduction, and the reduction of iron oxide by carbon in high temperature zone is called direct reduction. Because the reduction of silicon consumes a lot of heat, silicon is directly reduced by carbon only in the high temperature zone [1]. Among the many factors affecting silicon content, the direct reduction of iron is an important factor. When the smelting conditions of blast furnace are stable, the indirect reduction degree of CO can be calculated according to the content of CO_2 in gas composition. Once the indirect reduction degree is determined, the direct reduction degree is also determined. Therefore, the change of CO_2 content in blast furnace gas also reflects the change of direct reduction degree of iron, and finally affects the degree of silicon reduction [6]. The reduction of iron ore is affected by gas flow conditions.

K. Yang, C. Hu, School of Computer Science and Software Engineering, University of Science and Technology Liaoning, Anshan 114051, China. Corresponding author: Kai. Yang (asyangkai@126.com)

Gas flow conditions include temperature, flow rate, pressure and composition [1]. In a certain range, the increase of gas temperature can accelerate the reduction reaction, the increase of gas flow rate is conducive to diffusion and can promote reduction, and the increase of gas pressure can accelerate reduction [1]. Gas composition includes CO, CO_2, H_2 and N_2. In theory, the increase of indirect reduction can reduce carbon consumption and fuel ratio, so as to reduce the reduction of silicon. According to the above analysis, the selected input parameters are: fuel ratio, gas pressure, gas utilization rate, gas flow rate, gas temperature and gas composition. The correlation coefficient of each input parameter is shown in Table 1.

Table 1 **Correlation coefficients**

Parameters	Correlation coefficients
fuel ratio	0,4457
gas pressure	-0,4788
gas utilization rate	-0,6475
gas flow rate	-0,5843
gas temperature	-0,7327
CO	0,1427
CO_2	-0,6409
H_2	-0,6545
N_2	0,6406

WEIGHTED RANDOM FOREST

In RF, the weight values of all decision trees are the same. The weight values of each decision tree in WRF are different, and these weight values are optimized by SCPSO.

SCPSO is used to optimize three parameters of WRF, which are maximum depth, maximum features number and number of decision trees. The dimension of the solution is three, representing three parameters respectively. The fitness value of the objective function is the prediction hit rate of RF on the test set, and the weight values of all decision trees are the same. After determining the values of the three parameters, SCPSO is used to optimize the weight vector of WRF. The dimension of the solution is the number of decision trees, and each decision tree has a different weight value. The fitness value of the objective function is the prediction hit rate of WRF. The final predicted value $\hat{y}$ of WRF is calculated as follows:

$$\hat{y} = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix} \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix} \quad (1)$$

w_i is the weight of the i-th decision tree, $\hat{y}_i$ is the predictive value of the i-th decision tree.

PARTICLE SWARM OPTIMIZATION

In PSO, $x_i = \left(x_i^1, x_i^2, \cdots, x_i^N \right)$ and $v_i = \left(v_i^1, v_i^2, \cdots, v_i^N \right)$ are the position and velocity of the particle. $p_i = \left(p_i^1, p_i^2, \cdots, p_i^N \right)$ is the optimal individual position, and $g = \left(g^1, g^2, \cdots, g^N \right)$ is the optimal global position [7]. The formulas [8] are shown in (2) and (3).

$$v_i^d (t+1) = \omega v_i^d (t) + c_1 r_1 \left(p_i^d (t) - x_i^d (t) \right) + c_2 r_2 \left(g^d (t) - x_i^d (t) \right) \quad (2)$$

$$x_i^d (t+1) = x_i^d (t) + v_i^d (t+1) \quad (3)$$

i is the i-th particle; d is the d-th dimension; t is the t-th iteration; c_1 and c_2 are local and global learning factors; r_1 and r_2 are random numbers in the range [0, 1]; ω is the inertia weight.

SCALING COEFFICIENT PARTICLE SWARM OPTIMIZATION

In SCPSO, the movement of particles is guided by the elite solutions g_1 and g_2, which are the top two global optimal solutions found. Particles search near g_1 and g_2, which increases the possibility of finding a better solution.

Add two random coefficients in the range [0, 1] to the current position and the optimal position respectively. The mathematical expressions are $r_2g_1 - r_1x_i$ and $r_2g_2 - r_1x_i$, respectively. Because of r_2 and r_1, the value of $r_2g_1 - r_1x_i$ can be smaller or larger than $g-x_i$, making it easier to find a better solution.

A new random number balance mechanism is proposed in this paper. c_1 and c_2 are random numbers in the range [-2,2]. The direction of v_i is controlled by the signs of c_1 and c_2. As shown in Figure 1, x_i is the current position of the particle, g_1 and g_2 are the top two global optimal solutions found, and x_i is the midpoint of $g_1g'_1$ and $g_2g'_2$. In SCPSO, particles may move in multiple directions. The relations between the signs of c_1 and c_2 and the direction of v_i is shown in Table 2.
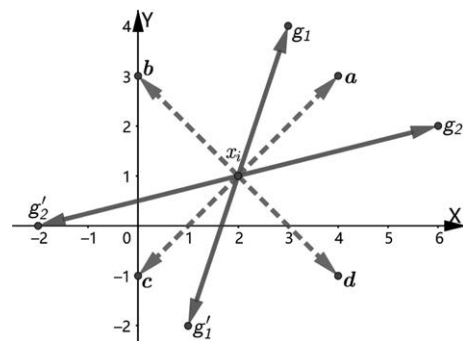

Figure 1 Particle movement graph of SCPSO

Table 2 **Corresponding relations between sign and direction**

c_1	c_2	v_i
+	+	**a**
-	+	**b**
-	-	**c**
+	-	**d**

According to the above improvement of PSO, the formulas of SCPSO are shown in (4) and (5).

$$v_i^d (t+1) = c_1 \left| r_2 * g_2^d (t) - r_1 * x_i^d (t) \right| + c_2 \left| r_2 * g_1^d (t) - r_1 * x_i^d (t) \right| \quad (4)$$

$$x_i^d (t+1) = r_3 * x_i^d (t) + v_i^d (t+1) \quad (5)$$

i is the i-th particle; d is the d-th dimension; t is the t-th iteration; r_1, r_2 and r_3 are random numbers in the range [0,1]; c_1 and c_2 are random numbers in the range [-2,2].

The execution process of SCPSO is as follows:

Step 1: Randomly initialize the position of particles, calculate the fitness values, and update g_1 and g_2.

Step 2: Is the maximum number of iterations reached? If yes, exit the program; If not, go to step 3.

Step 3: Update the velocity according to formula (4).

Step 4: Update the position according to formula (5).

Step 5: Check the out of bounds and randomly initialize the out of bounds position.

Step 6: Calculate the fitness values of the new positions. If the new fitness value is better, update g_1 and g_2.

Step 7: Output the position of g_1 and the fitness value of g_1, and go to step 2.

BENCHMARK FUNCTIONS VALIDATION

The experimental results are analyzed from three aspects: optimal value, average value and worst value. PSO, Grey Wolf Optimizer (GWO) [9] and SCPSO are selected for comparative experiment. In the experiment, the number of iterations is 1000, the population size is 10, the solution dimension is 10, and it runs 20 times. In PSO, c_1 is 0,2, c_2 is 0,3, ω is 1. f_1-f_3 are unimodal functions, f_4-f_6 are multimodal functions, and the details of the benchmark function are shown in Table 3. The experimental results are shown in Table 4. The convergence graphs of some functions is shown in Figures 2 and 3.

Table 3 **Benchmark functions**

Functions	Range	f_{min}
$f_1(x) = \sum_{i=1}^{n} x_i^2$	[-100, 100]	0
$f_2(x) = \sum_{i=1}^{n} \left(\sum_{j=1}^{i} x_j \right)^2$	[-100, 100]	0
$f_3(x) = \sum_{i=1}^{n} |x_i|^{(i+1)}$	[-1, 1]	0
$f_4(x) = \sum_{i=1}^{n} -x_i \sin\left(\sqrt{|x_i|}\right)$	[-500, 500]	-4190
$f_5(x) = \frac{1}{4000} \sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	[-600, 600]	0
$f_6(x) = \left\{ \left[\sum_{i=1}^{n} \sin^2(x_i) \right] - \exp\left(-\sum_{i=1}^{n} x_i^2\right) \right\} \times \exp\left[-\sum_{i=1}^{n} \sin^2\left(\sqrt{|x_i|}\right) \right]$	[-10, 10]	-1

Table 4 **Experimental results**

|	Algorithm	Best	Mean	Worst
f_1	PSO	7,38448E-27	3,23768E-18	4,44644E-17
	GWO	1,1957E-192	1,8028E-172	3,6054E-171
	SCPSO	0	0	0
f_2	PSO	8,47349E-07	0,026957592	0,452895906
	GWO	2,03442E-46	1,29676E-16	2,08431E-15
	SCPSO	0	0	0
f_3	PSO	2,25597E-12	1,99283E-09	2,08237E-08
	GWO	1,2157E-245	3,1785E-213	6,357E-212
	SCPSO	0	0	0
f_4	PSO	-3854,250869	-3362,844277	-2786,653363
	GWO	-4189,621475	-2905,459369	-1567,193305
	SCPSO	-4189,785082	-4175,450608	-3942,954201
f_5	PSO	0,032593858	0,139883743	0,491535418
	GWO	0	0,032449576	0,648991522
	SCPSO	0	0	0
f_6	PSO	1,96831E-29	4,96644E-14	8,97943E-13
	GWO	-1	-0,10070719	0,001263564
	SCPSO	-1	-1	-1
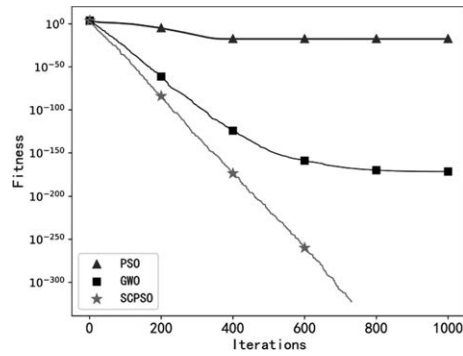
Figure 2 Convergence graph of f_1
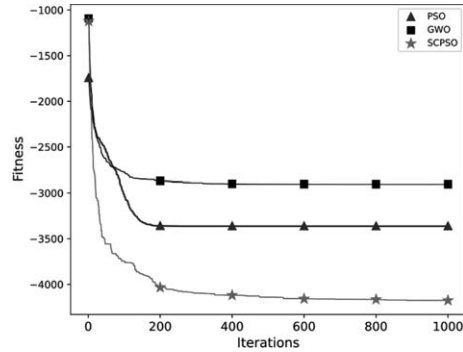
Figure 3 Convergence graph of f_4

It can be seen from Table 4 that SCPSO finds the actual global optimal value on f_1, f_2, f_3, f_5 and f_6. On f_4, the optimal values found by SCPSO is closer to the actual global optimal value than PSO and GWO. It can be seen from Figures 2 and 3, SCPSO can locate the range of optimal values faster than PSO and GWO. It can be concluded from Figures 2 and 3 that SCPSO not only has higher accuracy but also faster speed than PSO and GWO, which basically meets the expectation.

RESULTS AND DISCUSSION

The three parameters are maximum depth, maximum features number and number of decision trees, and their value ranges are [1,20], [1,9] and [1,300]. The optimization results show that the maximum depth is 16, the maximum features number is 4, the number of decision trees is 10, and the weight vector is [0,08930418, 0,10679194, 0,16228444, 0,06976635, 0,06209202, 0,05738346, 0,1694795, 0,14866342, 0,09103784, 0,0399059]. The model is evaluated from two aspects: hit rate and mean absolute error.

(1) Hit rate

$$H_{rate} = \frac{1}{N} \sum_{i=1}^{N} I\left(|y_i - \hat{y}_i| \leq 0.05 \right) \times 100\% \qquad (6)$$

y_i is the true value. $\hat{y}_i$ is the predictive value. I is an indicator function. If the input is true, output 1; otherwise, output 0. N is the number of test samples.

(2) Mean Absolute Error

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |y_i - \hat{y}_i| \qquad (7)$$

In order to verify the performance of the models, five models are compared in simulation experiments. The size of the data set is 1039, the size of the training

Table 5 **Experimental results**

Model	H_{rate}	MAE
SVM	80,76 %	0,03418691684509517
SCPSO-SVM	85,57%	0,03221314153087282
RF	84,29%	0,029747403846153857
SCPSO-RF	86,53%	0,029428927678688356
SCPSO-WRF	89,1%	0,029179851103108095

set is 727, and the size of the test set is 312. The experimental results are shown in Table 5.

In Table 5, the performance of SCPSO-SVM is better than SVM. It can be seen that the models optimized by SCPSO have been improved. SCPSO has good performance in solving extreme values and parameter optimization. The H_{rate} of SCPSO-WRF is the highest 89,1%. The MAE of SCPSO-WRF is about 0,0291. SCPSO-WRF is significantly better than the other four models. The predictive values and MAE of test samples 1-50 are shown in Figures 4 and 5. The prediction errors of some samples are shown in Table 6.
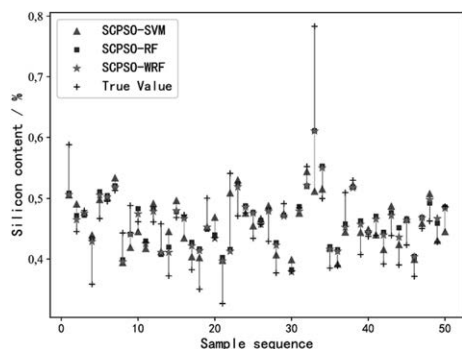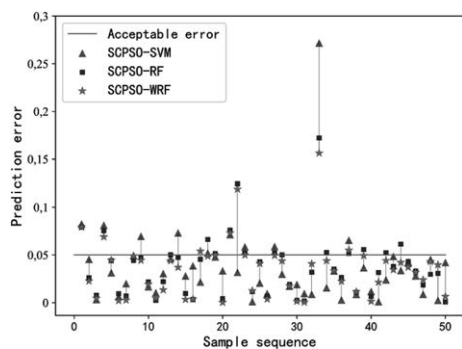

Figure 4 Predictive values of test samples 1-50

Figure 5 MAE of test samples 1-50

Table 6 **Prediction error of some samples**

Sample ID	Error of SCPSO-RF	Error of SCPSO-WRF
39	0,05571	0,049221706
42	0,052285	0,044122776
53	0,050221935	0,043928493
59	0,051466	0,049939097
108	0,06624	0,048783048
134	0,05258	0,043923345
144	0,06133	0,042104011
237	0,05199	0,049958506
279	0,05232125	0,049409093
286	0,050136667	0,043692323

It can be seen from Table 6 that the prediction error of SCPSO-RF is close to 0,05, but greater than 0,05. According to formula (6), these samples are not hit. By

optimizing the weight value of the decision tree, the prediction error of SCPSO-WRF is less than 0,05. SCPSO-WRF prediction method further improves the prediction hit rate and reduces the error. The input parameter selection method, WRF and SCPSO proposed in this paper make the prediction of silicon content more accurate.

CONCLUSION

In this paper, SCPSO and WRF are proposed and verified by experiments. SCPSO is used to optimize the hyper-parameters of WRF. The silicon content prediction model based on SCPSO-WRF is established. The results show that the performance of SCPSO-WRF is better than the other models. SCPSO-WRF can meet the task requirements of silicon content prediction. The establishment of prediction model has theoretical research and practical application value for improving hot metal quality and adjusting blast furnace operation.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Z. Y. Liang, Ironmaking, Metallurgical Industry Press, Beijing, 2009, pp. 130-144
[2]   W. H. Wang, X. G. Liu, X. Y. Liu, Prediction model of silicon content in blast furnace hot metal based on random forests algorithm, Metallurgical Industry Automation 38 (2014) 5, 33-38
[3]   X. Xu, C. C. Hua, Y. G. Tang, et al, Modeling of the hot metal silicon content in blast furnace using support vector machine optimized by an improved particle swarm optimizer, Neural Computing and Applications 27 (2016) 6, 1451–1461
[4]   Z. L. Li, C. J. Yang, W. H. Liu, et al, Research on hot metal Si-content prediction based on LSTM-RNN, Journal of Chemical Industry and Engineering 69 (2018) 3, 992-997
[5]   Y. M. Fang, X. D. Zhao, P. Zhang, et al, Prediction modeling of silicon content in liquid iron based on multiple kernel extreme learning machine and improved grey wolf optimizer, Control Theory & Applications 37 (2020) 7, 1644-1654
[6]   S. R. Na, Analysis of ironmaking calculation, Metallurgical Industry Press, Beijing, 2010, pp. 373-378
[7]   C. Hu, K. Yang, Prediction of silicon content in hot metal based on golden sine particle swarm optimization and random forest, Metalurgija 61 (2022) 2, 325-328
[8]   Y. Liu, L. Ma, H. Z. Zhang, et al, Intelligent optimization algorithm, Shanghai People's Publishing House, Shanghai, 2019, pp. 51, 53 and 57
[9]   S. Mirjalili, S. M. Mirjalili, A. Lewis, Grey Wolf Optimizer, Advances in Engineering Software 69 (2014), 46-61