C. GAO, M. G. SHEN

ISSN 0543-5846
METABK 58(1-2) 29-32 (2019)
UDC – UDK 669.18:669.187:512.9:338.93 =111

END-POINT PREDICTION OF BASIC OXYGEN FURNACE (BOF) STEELMAKING BASED ON IMPROVED TWIN SUPPORT VECTOR REGRESSION

Received – Primljeno: 2018-06-11
Accepted – Prihvaćeno: 2018-09-15
Original Scientific Paper – Izvorni znanstveni rad

In this paper, a novel prediction method for low carbon steel is proposed based on an improved twin support vector regression algorithm. 300 qualified samples are collected by the sublance measurements from the real plant. The simulation results show that the prediction models can achieve a hit rate of 96 % for carbon content within the error bound of 0,005 % and 94 % for temperature within the error bound of 15 °C. The double hit rate reaches to 90 %. It indicates that the proposed method can provide a significant reference for real BOF applications, and also it can be extended to the prediction of other metallurgical industries.

Key words: steel, basic oxygen furnace, twin support vector regression, end-point prediction

INTRODUCTION

In recent years, some significant developments of basic oxygen furnace (BOF) prediction modeling have been achieved [1]-[4]. These achievements are based on the statistical and the intelligent methods. As an intelligent method, Jayadeva proposed a twin support vector machine (TSVM) algorithm in 2007 for the classification applications [5]. In 2010, Peng proposed a twin support vector machine for regression (TSVR) [6], which can be used to establish the prediction model for industrial data. After that, some improved TSVR methods were proposed [7]-[11], the advantages of the algorithm has been proved in these literatures. Therefore, it is a new approach to establish the BOF prediction model.

In this paper, a novel BOF prediction modeling method is firstly proposed based on an improved TSVR algorithm. The proposed model can provide a guiding significance for real BOF production. It is also helpful to other metallurgy preidction and control applications. The improvement of TSVR algorithm enhances the performance and efficiency of the model. The contributions of this work are including: (1) It is the first attempt to predict the end-point information of BOF by using TSVR algorithm. (2) The proposed algorithm is an extension of TSVR algorithm, which is more flexible and accurate to establish a prediction model. (3) The proposed modeling method is also appropriate for the prediction models of other metallurgical applications, such as blast furnace iron making, continuous casting process and so on.

C. Gao (13500422153@163.com), M. G. Shen (Corresponding Author). School of Materials and Metallurgy, University of Science and Technology Liaoning, Anshan, China.

BACKGROUND

Description of BOF steelmaking

According to the requirements of the ingredients, the scrap steel is loaded into the furnace and then poured into the molten iron. The right amount of slagging materials should be added. Then, the oxygen spray gun is inserted into the furnace from the top of the furnace and blows the oxygen so that it can react with the hot metal directly and remove the impurities. After removing most of the sulfur and phosphorus, if the composition and temperature of the molten steel meet the requirements, then stop blowing, raise the spray gun, and prepare the tapping. When the steel is tapping, the furnace body is tilted, the molten steel is injected into the molten ladle from the tapping hole, and the deoxidizer is added to deoxidize and adjust the components. After the steel is qualified, it can be poured into steel castings or ingots, and the ingot can be rolled into various steels.

Recently, the BOF steelmaking process with the sublance measurement is a popular technology in the steel plants. There are two periods for the oxygen blowing process of BOF. In main-blow period, the calculations of the total oxygen volume and the weight of the auxiliary materials can be determined by the static BOF model. However, the relative calculations are not accurate enough due to the compelx process of BOF. In end-blow period, the sublance is adopted to measure the sample information. Because the physical and chemical reactions tend to be stable in this period. Hence, a BOF dynamic model can be used to improve the accuracy of the calculations. The end-point carbon content and temperature can be reached into a satisfactory region.

Improved twin support vector regression algorithm

Based on the concept of TSVR, an improved TSVR algorithm is firstly proposed in this paper. Assume that there exists a sample with an *n*-dimensional vector, and the number of the samples is *l*, which can be expressed as $(\mathbf{x}_1, y_1),\cdots, (\mathbf{x}_l, y_l)$, and then let $\mathbf{A} = [\mathbf{x}_1,\cdots, \mathbf{x}_l]^T \in R^{l\times n}$ be the input data set of training samples, $\mathbf{y} = [y_1,\cdots, y_l]^T \in R^l$ be the corresponding output, and $\mathbf{e} = [1,\cdots, 1]^T$ be the ones vector with appropriate dimensions. Assume $K(.,.)$ denotes a nonlinear kernel function. Let $K(\mathbf{A}, \mathbf{A}^T)$ be the kernel matrix with order *l*. The quadratic programming problems (QPPs) are defined as follows:

$$\begin{aligned}&\min_{\omega_1,b_1,\xi}\frac{1}{2}\left(\mathbf{y}-\left(K\left(\mathbf{A},\mathbf{A}^T\right)\boldsymbol{\omega}_1+b_1\mathbf{e}\right)\right)^T\\&\mathbf{D}\left(y-\left(K\left(\mathbf{A},\mathbf{A}^T\right)\boldsymbol{\omega}_1+b_1\mathbf{e}\right)\right)+\frac{c_1}{2}\left(\boldsymbol{\omega}_1^T\boldsymbol{\omega}_1+b_1^2\right)+c_2\mathbf{e}^T\boldsymbol{\xi}\\&\text{s.t.}:\mathbf{y}-\left(K\left(\mathbf{A},\mathbf{A}^T\right)\boldsymbol{\omega}_1+b_1\mathbf{e}\right)\geq-\varepsilon_1\mathbf{e}-\boldsymbol{\xi},\boldsymbol{\xi}\geq 0\mathbf{e},\end{aligned}\tag{1}$$

and

$$\begin{aligned}&\min_{\omega_2^{\xi},b_2,^*}\frac{1}{2}\left(\mathbf{y}-\left(K\left(\mathbf{A},\mathbf{A}^T\right)\boldsymbol{\omega}_2+b_2\mathbf{e}\right)\right)^T\\&\mathbf{D}\left(\mathbf{y}-\left(K\left(\mathbf{A},\mathbf{A}^T\right)\boldsymbol{\omega}_2+b_2\mathbf{e}\right)\right)+\frac{c_3}{2}\left(\boldsymbol{\omega}_2^T\boldsymbol{\omega}_2+b_1^2\right)+c_4\mathbf{e}^T\boldsymbol{\xi}^*\\&\text{s.t.}:K\left(\mathbf{A},\mathbf{A}^T\right)\boldsymbol{\omega}_2+b_2\mathbf{e}-\mathbf{y}\geq-\varepsilon_2\mathbf{e}-\boldsymbol{\xi}^*,\boldsymbol{\xi}^*\geq 0\mathbf{e},\end{aligned}\tag{2}$$

where $c_1, c_2, c_3, c_4, \varepsilon_1, \varepsilon_2 \geq 0$ are the regulating parameters, $\mathbf{D}$ is a diagonal matrix with the order of $l\times l$ and determined by the following formular:

$$w_i = \exp\left(-\frac{d_i^2}{2\sigma^{*2}}\right), i = 1,2,\cdots,l,\tag{3}$$

where w_i denotes the weight coefficient and σ^* is the width of the Gaussion function. d_i denotes the distance from the training samples to the denoising samples. In the proposed algorithm, the matrix $\mathbf{D}$ can be expressed as $\mathbf{D} = diag\,(w_1, w_2, \cdots, w_l)$.

The first term in the objective function of (1) or (2) is used to minimize the sums of squared Euclidean distances from the estimated function $f_1(\mathbf{x})$ or $f_2(\mathbf{x})$ to the training points, and defined as follows:

$$\begin{cases}f_1(\mathbf{x}) = K\left(\mathbf{x}^T,\mathbf{A}^T\right)\boldsymbol{\omega}_1+b_1,\\ f_2(\mathbf{x}) = K\left(\mathbf{x}^T,\mathbf{A}^T\right)\boldsymbol{\omega}_2+b_2.\end{cases}\tag{4}$$

where $K(\mathbf{x}^T, \mathbf{A}^T) = (K(\mathbf{x}, \mathbf{x}_1), (K(\mathbf{x}, \mathbf{x}_2), \cdots, (K(\mathbf{x}, \mathbf{x}_l))$.

The matrix $\mathbf{D}$ gives different weights for each Euclidean distance. The second term is the regularization term to avoid the over-fitting problem. The third term minimizes the slack vector ξ or ξ^*. To solve the problem in (1) and (2), the traditional TSVR adopts the Lagrange multiplier to obtain the dual problems of (1) and (2). In order to improve the computatonal speed, the problems can be solved in the original solution space instead of dual space by converting the constrained QPPs into a pair of unstrained QPPs. By introducing the plus function, an equivalent unconstrained QPPs can be derived as follows:

$$\begin{aligned}\min_{\mu_1} L_1(\boldsymbol{\mu}_1) &= \frac{1}{2}\left(\mathbf{y}-\mathbf{G}\boldsymbol{\mu}_1\right)^T\mathbf{D}\left(\mathbf{y}-\mathbf{G}\boldsymbol{\mu}_1\right)\\&+\frac{1}{2}c_1\boldsymbol{\mu}_1^T\boldsymbol{\mu}_1+c_2\mathbf{e}^T\left(\mathbf{G}\boldsymbol{\mu}_1-\boldsymbol{\gamma}_1\right)_+,\end{aligned}\tag{5}$$

and

$$\begin{aligned}\min_{\mu_2} L_2(\boldsymbol{\mu}_2) &= \frac{1}{2}\left(\mathbf{y}-\mathbf{G}\boldsymbol{\mu}_2\right)^T\mathbf{D}\left(\mathbf{y}-\mathbf{G}\boldsymbol{\mu}_2\right)\\&+\frac{1}{2}c_3\boldsymbol{\mu}_2^T\boldsymbol{\mu}_2+c_4\mathbf{e}^T\left(\boldsymbol{\gamma}_2-\mathbf{G}\boldsymbol{\mu}_2\right)_+,\end{aligned}\tag{6}$$

where $\boldsymbol{\mu}_1 = \left[\boldsymbol{\omega}_1^T\ b_1\right]^T$, $\boldsymbol{\mu}_2 = \left[\boldsymbol{\omega}_2^T\ b_2\right]^T$, and $\mathbf{G} = \left[K\left(\mathbf{A},\mathbf{A}^T\right)\mathbf{e}\right]$, $\boldsymbol{\gamma}_1 = \mathbf{y}+\varepsilon_1\mathbf{e}$, $\boldsymbol{\gamma}_2 = \mathbf{y}-\varepsilon_2\mathbf{e}$.

From the terms $\frac{1}{2}c_1\boldsymbol{\mu}_1^T\boldsymbol{\mu}_1$ and $\frac{1}{2}c_2\boldsymbol{\mu}_2^T\boldsymbol{\mu}_2$ it is easy to see that the above problems are strongly convex. Therefore, there exists a global unique solution for the problem (5) or (6). However, the plus function in the optimization problem is not differentiable. Hence, a smooth function $s(\cdot)$ can be used to approximate the plus function, which is defined as

$$s(x,\alpha) = \frac{x+\sqrt{x^2+4\alpha^2}}{2}.\tag{7}$$

Then, the problem (4) and (5) are modified as follows:

$$\begin{aligned}\min_{\mu_1} L_1(\boldsymbol{\mu}_1) &= \frac{1}{2}\left(\mathbf{y}-\mathbf{G}\boldsymbol{\mu}_1\right)^T\mathbf{D}\left(\mathbf{y}-\mathbf{G}\boldsymbol{\mu}_1\right)\\&+\frac{1}{2}c_1\boldsymbol{\mu}_1^T\boldsymbol{\mu}_1+c_2\mathbf{e}^T s\left(\mathbf{G}\boldsymbol{\mu}_1-\boldsymbol{\gamma}_1,\alpha\right),\end{aligned}\tag{8}$$

and

$$\begin{aligned}\min_{\mu_2} L_2(\boldsymbol{\mu}_2) &= \frac{1}{2}\left(\mathbf{y}-\mathbf{G}\boldsymbol{\mu}_2\right)^T\mathbf{D}\left(\mathbf{y}-\mathbf{G}\boldsymbol{\mu}_2\right)\\&+\frac{1}{2}c_3\boldsymbol{\mu}_2^T\boldsymbol{\mu}_2+c_4\mathbf{e}^T s\left(\boldsymbol{\gamma}_2-\mathbf{H}\boldsymbol{\zeta}_2,\alpha\right).\end{aligned}\tag{9}$$

The solutions of problem (8) and (9) can be obtained by using Newton-Armijo method. Then, the final regression function can be obtained by

$$f(\mathbf{x}) = f_1(\mathbf{x}) + f_2(\mathbf{x})\tag{10}$$

DYNAMIC PREDICTION MODEL FOR BASIC OXYGEN FURNANCE

In order to control the carbon content and temperature to meet the requirements of high quality steel, the dynamic prediction model is the foundation of the relative control model. By collecting the historical BOF samples, the prediction model for carbon content (C_Model) and for temperature (T_Model) can be estabilshed. Note that the numbers of the input variables of C_Model and T_Model are different. Because the coolant additions only affect the end-point temperature. It means there are three input variables for C_Model and four input variables for T_Model. The output variable

of the relative model is the end-point carbon content or the end-point temperature.

According to the prepared samples of BOF, the regression function $f_{C/T}(\mathbf{x})$ can be given by:

$$f_{C/T}(\mathbf{x}) = f_1(\mathbf{x}) + f_2(\mathbf{x}) \quad (11)$$

where $f_{C/T}(\mathbf{x})$ denotes the estimation function of C_ model or T_model.

The process of C_model and T_model can be described as follows:

Step 1: Initialize the parameters of the prediction model, and normalize the prepared training samples from its original range to the range [-1 1] by *mapminmax* function in Matlab.

Step 2: Denoise the end-point carbon content or the temperature in the training samples by using the wavelet transform algorithm[12]. Then, the denoised samples and can be obtained.

Step 3: By selecting an appropriate parameter σ^* in (3), determine weighted matrix **D**.

Step 4: Select appropriate values of c_1, c_2, c_3, c_4, ε_1, ε_2.

Step 5: Solve the optimization problems in (8) and (9) by Newton-Armijo method and return the optimal vector μ_1 and μ_2.

Step 6: Substitute $\mu_1 = [\omega_1 b_1]^T$ and $\mu_1 = [\omega_2 b_2]^T$ into (10) to obtain the function $f_{C/T}(\mathbf{x})$.

Step 7: Normalize the test samples and substitute them into (10) to obtain the estimated samples, then calculate the relative criteria of the model, which will be described in details later.

Step 8: If the relative criteria are satisfactory, then the C_Model or T_Model is established. Otherwise, return to Steps from 4 to 7.

RESULTS AND DISCUSSION

In order to verify the effectiveness of the proposed model, real heats samples were collected by the sublance from some steel plant in Anshan, P.R. China. The type is low carbon steel of 260 *tons* BOF. Firstly, the samples should be processed because there are a number of abnormal sampels, which means the unexpected information should be deleted, for example, the steel information exceeds the range of the actual value due to the wrong sampling operation. After the preprocessing of the samples, 300 qualified heats samples are obtained. Then, the relative analysis on samples is carried out. Generally, the end-point carbon content and temperature are mainly determined by the carbon content and the temperature measured by the sublance, the end-blow oxygen volume and the end-blow coolant additions. Therefore, other information in the sampels should be removed, including the heat number, the date of the steelmaking and so on. After that, the samples are ready to establish the BOF prediction models (C_Model and T_Model) described before. By considering the real production, the error bound for C_model is set to 0.005 % and which of T_model is set to 15 °C. The performance of each model can be determined by the hit rate within its relative error bound. Normally, a hit rate of 90 % is a good result for real production. Therefore, 200 heats samples are taken as the train samples and 100 heats samples are taken as the test samples. The parameters of the prediction models should be regulated for each run of training 200 samples and testing 100 samples until the evaluation criteria is satisfied. The evaluation criteria are including root mean square error *RMSE*, mean absolute deviation *MAE*, error sum of squares/total sum of squares *SSE / SST*, regression sum of squares/total sum of squares *SSR / SST* and hit rate *HR*. The smaller values of *RMSE* and *SSE / SST*, and the larger values of *SSR / SST* and hit rate are preferred, it shows the model has a better performance. Note that the parameter selection can be determined by the following rules: *SSR / SST* > 0,5 and *HR* >85 with the smallest *SSE / SST*. Especially, the double hit rate *DHR* should be greater than 85 or higher. Based on 300 qualified samples, the parameters of the proposed models are specified and shown in Table 1. c_1, c_2, c_3, c_4, ε_1, ε_2 are described before. The Gaussion function is chosen for the kernel function in this paper, and σ is the width of Gaussion function. σ^* is the parameter in (3) to determine the matrix **D**.

Table 1 **Specified parameters of models**

Prediction Model	c_1	c_2	c_3	c_4	ε_1	ε_2	σ	σ^*
C_Model	0,002	0,002	0,001	0,001	1	1	10	0,3
T_Model	0,003	0,003	0,002	0,002	0,1	0,1	1	0,25

The proposed models can be established by the specified parameters. In order to evaluate the performance of the proposed models, the results are compared with traditional TSVR [6] and *v*-TSVR [7]. Figure 1 shows the performance of the proposed C_model with 50 heats samples, the solid line denotes the trend changes of actual samples, and the dash line denotes the trend changes of predicted samples. It shows that these two sets of samples have a good fitness.The comparison results of the carbon content prediction are listed in Table 2. It is easy to see that the proposed C_Model achieves a best hit rate (*CHR*) of 96 %. The results of *RMSE*, *MAE* and *SSE / SST* are all smaller than those of other two existing methods. Also, the result of *SSR SST* is the second best, and the CPU time is the fastest. It verifies that the proposed C_Model has the best fitting behaviour for carbon content prediction.

Figure 2 shows the performance of the proposed T_ model. The comparisons of the temperature prediction are also listed in Table 2. The T_Model can reach the best hit rate (*THR*) of 94 %. The best results of *RMSE*, *MAE* and *SSE / SST* are obtained. The execution speed of the proposed T_Model is also the fastest. In addition, the double hit rate of the proposed method can reach to 90 %, which is the best result and satisfies the requirements of real applications. Therefore, the proposed T_Model is more efficient to provide a reference for the real applications.

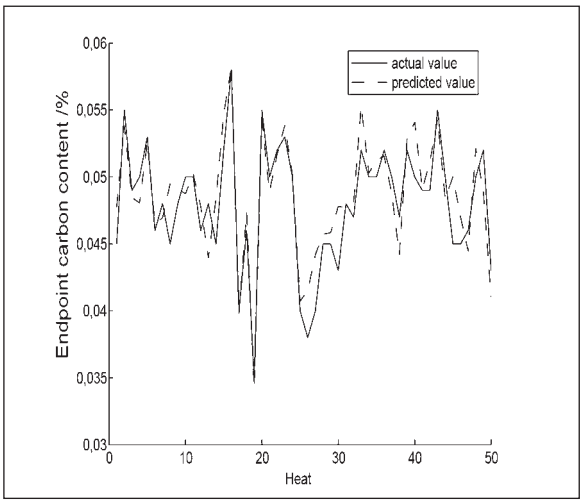

Figure 1 Performance of proposed C_model

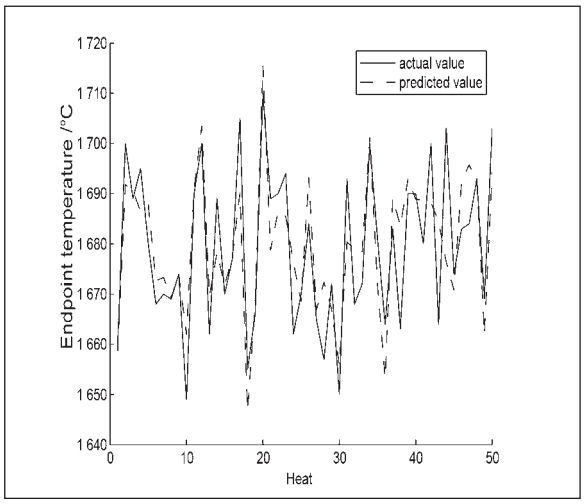

Figure 2 Performance of proposed T_model

Table 2. **Comparisons of three prediction methods**

Model	Criteria	Proposed Method	*v*-TSVR	TSVR
C_model (±0,005 %)	*RMSE*	0,0026	0,0030	0,0032
	MAE	0,0020	0,0025	0,0024
	SSE/SST	0,2189	0,3026	0,3363
	SSR/SST	0,9703	1,2267	0,9418
	CHR, %	96	92	88
	Time, s	1,55	1,75	1,86
T_model (±15 °C)	*RMSE*	8,7359	9,3769	13,2920
	MAE	6,9493	7,4821	10,1886
	SSE/SST	0,3467	0,3995	0,8027
	SSR/SST	0,7257	0,8093	1,0800
	THR, %	94	89	82
	Time, s	1,52	1,79	1,87
	DHR, %	90	82	74

From the above analysis, it can be concluded that the proposed models are effective and feasible, and the hit rate can meet the requirements of the real productions for low carbon steel. For other kinds of steels, the proposed model is still suitable for use. Firstly, the specific samples of the heats should be measured, and the analysis of the influence factors on the specific steel should be carried out. Secondly, the criteria of the model should be selected. Then, the model parameters can be regulated to achieve the best criteria. Finally, the BOF model for the specific kind of steels can be established to guide the real production in the plant.

CONCLUSIONS

In this paper, a BOF prediction method has been proposed based on the improved TSVR algorithm. The simulation results have shown that the proposed models are effective and feasible. The prediction model within the error bound of 0,005 % in carbon content and 15 °C in temperature can achieve a hit rate of 96 % and 94 %, respectively. In addition, the double hit rate of 90 % is the best result by comparing with other existing methods. Therefore, the proposed method can provide a significant reference for real BOF applications. For the further work,, an end-point control model could be established based on the proposed prediction models to realize the control of the end-point carbon content and temperature.

Acknowledgments

This work was supported by Liaoning Province Ministry of Education Scientific Study Project (No. 2017LNQN11).

REFERENCES

[1] M. Brämming, B. Björkman, C. Samuelsson, BOF process control and slopping prediction based on multivariate data analysis, Steel Research International 87(2016)3, 301-310.

[2] M. Han, Z. Cao, Applying input variables selection technique on input weighted support vector machine modeling for BOF endpoint prediction, Engineering Applications of Artificial Intelligence 23(2010)6, 1012-1018.

[3] L. Kong, P. Hodgson, D. Collinson, Modelling the effect of carbon content on hot strength of steels using a modified artificial neural network, ISIJ International 38(1998)10, 1121-1129.

[4] J. Jimenez, J. Mochon, J. Ayala, Ayala, Blast furnace hot metal temperature prediction through neural networks based models, ISIJ International 44(2004)3, 573-580.

[5] Jayadeva, R. Khemchandani, S. Chandra, Twin support vector machines for pattern classification, IEEE Trans. on Pattern Analysis and Machine Intelligence 29(2007)5, 905-910.

[6] X. Peng, TSVR, an efficient Twin support vector machine for regression, Neural Networks, 23(2010)3, 365-372.

[7] R. Rastogi, P. Anand, S. Chandra, A *v*-twin support vector machine based regression with automatic accuracy control, Applied Intelligence 46(2016), 1-14.

[8] M. Tanveer, K. Shubham, M. Aldhaifallah, An efficient implicit regularized Lagrangian twin support vector regression, Applied Intelligence 44(2016)4, 1-18.

[9] Y. Xu, X. Li, X. Pan, Asymmetric *v*-twin support vector regression, Neural Computing & Applications, (2017)2: 1-16.

[10] N. Parastalooi, A. Amiri, P. Aliheydari, Modified twin support vector regression, Neurocomputing 211(2016), 84-97.

[11] Y. Ye, L. Bai, X. Hua, Weighted Lagrange *v*-twin support vector regression, Neurocomputing 197(2016), 53-68.

[12] J. Shen, G. Strang, Asymptotics of Daubechies filters, scaling functions, and wavelets, Applied & Computational Harmonic Analysis 5(1998)3, 312-331.

Note: The responsible translator for English language is L. D. Wang, University of Science and Technology Liaoning, China