

Research Article

Data-Driven Dynamic Modeling for Prediction of Molten Iron Silicon Content Using ELM with Self-Feedback

Ping Zhou,¹ Meng Yuan,¹ Hong Wang,^{1,2} and Tianyou Chai¹

¹State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China

²Control System Center, Manchester University, Manchester M60 1QD, UK

Correspondence should be addressed to Ping Zhou; zhouping@mail.neu.edu.cn

Received 21 August 2014; Revised 17 November 2014; Accepted 18 November 2014

Academic Editor: Jiuwen Cao

Copyright © 2015 Ping Zhou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Silicon content ([Si] for short) of the molten metal is an important index reflecting the product quality and thermal status of the blast furnace (BF) ironmaking process. Since the online detection of [Si] is difficult and larger time delay exists in the offline assay procedure, quality modeling is required to achieve online estimation of [Si]. Focusing on this problem, a data-driven dynamic modeling method is proposed using improved extreme learning machine (ELM) with the help of principle component analysis (PCA). First, data-driven PCA is introduced to pick out the most pivotal variables from multitudinous factors to serve as the secondary variables of modeling. Second, a novel data-driven ELM modeling technology with good generalization performance and nonlinear mapping capability is presented by applying a self-feedback structure on traditional ELM. The feedback outputs at previous time together with input variables at different time constitute a dynamic ELM structure which has a storage ability to tackle data in different time and overcomes the limitation of static modeling of traditional ELM. At last, industrial experiments demonstrate that the proposed method has a better modeling and estimating accuracy as well as a faster learning speed when compared with different modeling methods with different model structures.

1. Introduction

Blast furnace (BF) is a giant countercurrent reactor and heat exchanger in metallurgical industry and is the first step towards the production of steel [1]. During the BF ironmaking system working, the solid raw materials including iron ore and coke are charged layer by layer from the top of the BF, while the compressed air and some auxiliary fuels are introduced through tuyeres just equipped above the hearth for smelting to produce molten iron. The complex chemical reactions and transport phenomena take place in the different zones along the top to the bottom of the BF. Gas-solid, solid-solid, and solid-liquid phases interacted in it and are accompanied with features like high temperature, high pressure, multiphase coupling, and multiphysics field which coexist simultaneously [1–3]. As one of the most complex industrial reactors, the BF has received broad interests both theoretically and experimentally due to its complexity and the key role of iron and steel industry on national economy.

However, it is true that the operation and control of an industrial BF is a serious problem and still relies on the manual operation of foremen experientially [1, 2]. So far, there remain some open problems both in metallurgical fields and in engineering control fields, such as the closed-loop control or operational optimization for the whole BF ironmaking process [4–6].

Undoubtedly, the most crucial obstacle for closed-loop control of BF is that the current regular instruments do not have the ability to feed the need of online measurement for molten iron quality, such as the silicon content ([Si]) in the final hot metal. In the past decades, through continuous efforts and attempts, a great number of models and algorithms have been developed trying to tackle the modeling problem for silicon content prediction. These existing methods include linear model based methods like ARX and ARMAX models [6–8], partial least squares based methods [9], and nonlinear intelligent based methods like artificial neural network (ANN) model [10–12] and support

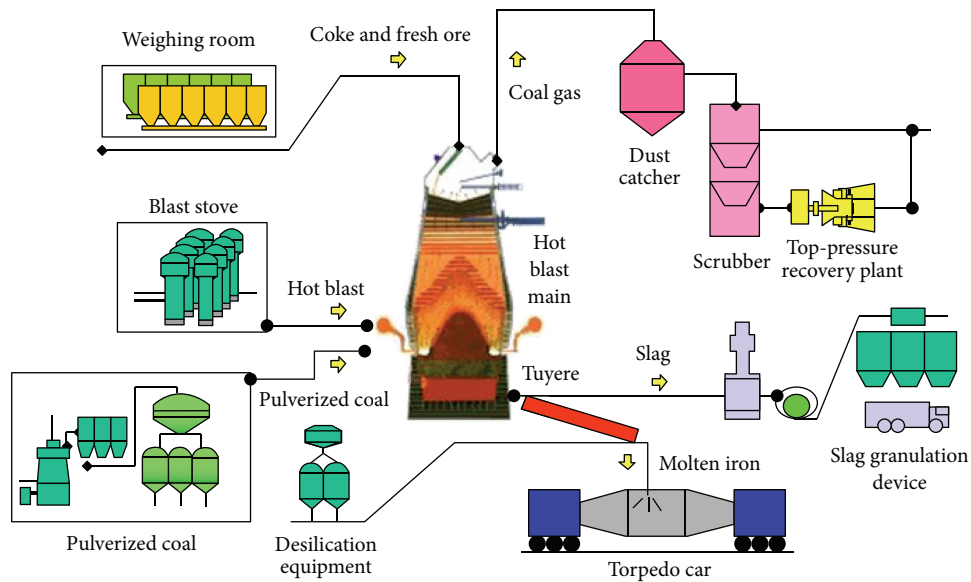


FIGURE 1: Schematic diagram of a typical BF ironmaking process.

vector machine (SVM) model [1, 2, 13, 14]. Though these existing methods have made some achievements in practical application, most of these studies are only focused on the static modeling for [Si] prediction while little attention has been paid to dynamical modeling of this quality parameter.

The BF ironmaking process is a complicated dynamic system with many influential factors and large time lag. To capture the system dynamics, the time series and time delays of the relevant input and output variables should be taken into account during the process modeling. This also means that the existing static prediction models cannot capture the process nonlinear dynamics very well and thus do not provide much accuracy estimation. Therefore, the self-feedback structure which can construct a dynamic system may appear more important for the BF system with serious nonlinear dynamics and large time lag. Moreover, most of the existing prediction models are trained by gradient-based algorithms such as back propagation (BP) algorithm and its variants. It is clear that the learning speed of such intelligent models is insufficiently fast as larger number of training data may be required. Moreover, the BP-like algorithm usually suffers from high computational burden, poor generalization ability, and local optima and overweighting problems [15].

On the other hand, a new machine learning approach that is termed as the extreme learning machine (ELM) has been recently proposed by Huang et al. in [15–18] and verified on a number of benchmark and real-world problems including pattern classification and prediction modeling [16–25]. The ELM and its variants have been considered as a promising learning algorithm in contrast with other algorithms such as BP NN and SVM. This is because ELM has the following advantages: (1) much faster learning speed; (2) higher generalization performance in comparison with BP NN and SVM; and (3) no extra parameters needing to be tuned except the predefined network architecture [15–18, 22–27].

Based on the work of ELM proposed by Huang et al. [15–18], this paper proposed a data-driven dynamic modeling method to predict molten iron silicon content using ELM with the help of principle component analysis (PCA) [28, 29]. In the design procedure of this predictive model, data-driven PCA for reducing the input variables space of ELM has been constructed. Moreover, output self-feedback architecture has been introduced to establish a dynamic ELM model for practical BF dynamic system. This self-feedback structure enables ELM to overcome the static mapping limitation of its feedforward network structure. This improvement can further optimize the application of ELM in the area of dynamic time-series prediction. Lastly, performance of the proposed dynamic ELM based prediction model is compared with other well-known modeling algorithms by industrial experiments on 2[#] BF in Liuzhou Iron & Steel Group Co. of China.

2. Description of BF Ironmaking System

2.1. Process Description. The BF ironmaking is a continuous production process conducted in a closed vertical furnace where materials reduction from iron ore to molten iron takes place every time using carbon coke and gas in high temperature and high pressure environment. Due to the advantages like simple technology, high productivity, and high production efficiency, at present and a long period in the future, the BF smelting will still be the most important way of ironmaking. Indeed, due to the large quantity production, even small improvements of the process can result in considerable profit. Thus the ironmaking BF is regarded as a significant item in the economic development of any country.

Figure 1 is the schematic diagram of a typical BF ironmaking process, which mainly consists of hot blast main, feeding system, air supply system, gas filtration system, slag

treatment system, fuel injection system, and so forth. The inner part of a furnace main is divided into five zones: the throat, the stack, the belly, the bosh, and the hearth from top to bottom. When a BF ironmaking system runs, the solid raw materials consisting of coke and fresh ore are charged layer by layer with definite quantities from the top, while the preheated compressed air, together with pulverized coal, is introduced at the bottom through tuyeres, entering just above the hearth, which is a crucial region of BF where the final molten metal product gathers. The hot air at approximately 1200°C passes upward through the charge and reacts with the descending coke and the supplementary injected oil to generate carbon dioxide, which then changes to CO and H₂ at high temperature. A lot of heat energy is released during this period that can heat up the hearth as high as 2000°C. The generated CO and H₂ further reduce the descending iron ore to form hot metal accumulating in the hearth, and some unreduced impurities (mainly SiO₂) form the slag (mainly CaSiO₃) floating on the hot metal being lighter. The liquid hot metal and slag are periodically tapped out by opening clay-lined tapholes for the subsequent processing. Generally, it will take 6~8 hours for each period of BF ironmaking [30].

2.2. Importance of Modeling for Silicon Content Prediction.

For many countries, such as China, the steel industry is playing an important role in the national economy, and there are thus extensive interests in operational control and optimization of ironmaking BF for saving energy and reducing cost. Generally speaking, control of the BF system often means controlling the hot metal temperature and components, such as silicon content, sulfur content, and phosphorus content in hot metal within acceptable bounds, among which the silicon content is the most important one [31].

For a practical BF production process, silicon content ([Si]) is an important index indicating the chemical heat of molten iron. High silicon content means a large quantity of slag, and this would be easier to wipe off the phosphorus and sulphur in the hot metal. However, excessive silicon content will make cast iron stiff and brittle and even lead to lower yield of metal and easier splashing. In addition, high silicon content will result in a corresponding increase of SiO₂ in the slag, thereby influencing slagging speed of calclime, extending converting time, and intensifying corrosion to furnace lining. From an energy point of view, it would be desired to operate the BF process at low molten metal silicon content, still avoiding the risk of cooling the hearth which may result in chilled hearth. Generally, the content of silicon content should be controlled in 0.5%~0.7%.

Nowadays, it is still an insoluble dilemma to realize the closed-loop control of molten iron quality in ironmaking BF. The main bottleneck is that the direct online measurement on this quality parameter of molten iron is difficult to be realized with the existing conventional measuring means. Moreover, the offline assaying process for this index takes a long lag time, usually more than 1 hour. Therefore, online prediction based molten iron quality modeling must be established. Effective online prediction or estimation for silicon content not only

can offer useful information for operators to judge the inner smelting state and operational condition, but also plays a key role in realizing closed-loop control and operational optimization as well as energy-saving and cost-reducing.

3. Modeling Strategy Using PCA and ELM with Self-Feedback

Data-driven black-box model is a kind of input-output mode. It relies on the development of novel nonlinear signal processing and data analysis technologies along with computer hardware and software technologies and does not require any prior information about the process. The main thought of data-driven model is to approximate the input-output relationships using the strong nonlinear approximation power of some mathematical tools or artificial intelligence technologies, like artificial neural network, fuzzy logic, and support vector machines [32].

The proposed data-driven modeling strategy for silicon content prediction is shown in Figure 2. First, since the BF is a complicated high-dimensional nonlinear dynamic system combined with numerous coupled factors, data-driven PCA technology with a strong ability to handle high-dimensional nonlinear correlated data is introduced to pick a few key factors as the input variables of model so as to reduce the dimension and difficulty for prediction modeling. Then, considering that the BF ironmaking system is a nonlinear system with dynamic time-vary characteristic, the ELM with better nonlinear mapping and fast process capability modeling technology is brought in this paper. In the meantime, output self-feedback structure is put into use on the basis of traditional ELM in this method, and the output variables derived from previous time are fed back to the network input layer. These feedback outputs together with input variables at different time constitute a dynamic ELM structure which has a storage capacity and has the ability to tackle data in different time, thus overcoming the limitation of static modeling of traditional ELM.

Remark 1. As shown in Figure 2, the dynamic ELM based estimation model is developed to achieve the following nonlinear dynamic mapping:

$$y(t) = f_{\text{ELM}} \{X(t), \dots, X(t - k_1), y(t - 1), \dots, y(t - k_0)\}, \quad (1)$$

where $X = [x_1, x_2, \dots, x_n]$ are the values of secondary variables selected by PCA and y is the quality parameter that needs to be estimated. The values of $k_1, k_0 \in Z^+$ are selected according to the time delays and time series of the relevant input and output variables and the sampling frequency of quality parameter y which is generally sampled slower than the process data X significantly.

Remark 2. Note that, in the learning period of the proposed ELM based dynamic estimation model using the training databases, $y(t - 1), \dots, y(t - k_0)$ are the actual (sampling) values of y . After the proposed ELM model is trained and validated well, it will be applied in practice. Since the quality

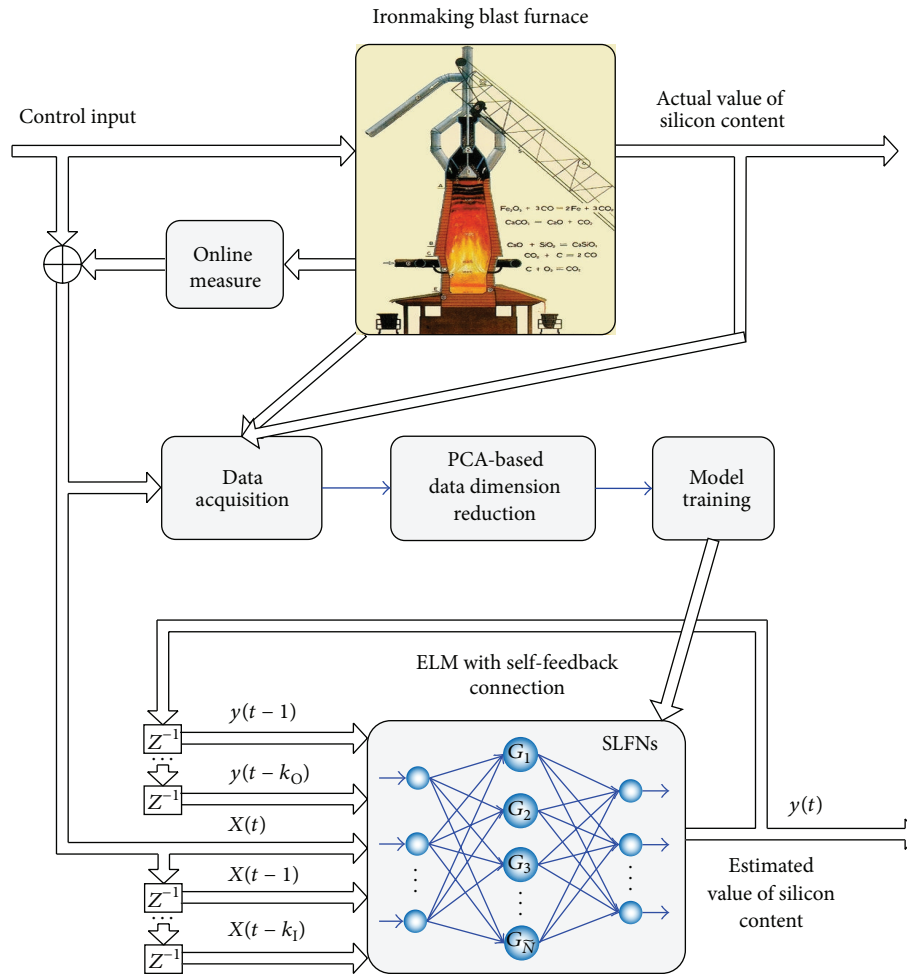


FIGURE 2: Strategy diagram of nonlinear intelligent modeling for silicon content prediction.

parameter y cannot be measured online and the offline assaying process takes a long time which is usually more than one hour, the actual value of y cannot be obtained in real time in practice. Therefore, to achieve the desired dynamic estimation of y , the estimated $\tilde{y}(t-1), \dots, \tilde{y}(t-k_0)$ at past will be used to construct the self-feedback structure for the proposed dynamic ELM based estimation model.

Remark 3. The proposed modeling strategy has two advantages.

- (i) The dynamic property of time series and time delays is considered by introducing the output and inputs in previous time through a self-feedback structure. This self-feedback connection enables ELM to overcome the static mapping limitation of its feedforward network structure. Thus the improved version of ELM can capture the process nonlinear dynamics very well by remembering prior input and output states and using both the prior and current states to calculate new output value.

- (ii) Different from the BP-like modeling algorithm usually suffering from high computational burden, poor generalization ability, and local optima and overweighing problems, the ELM based modeling benefits from much faster learning speed, higher generalization performance, and ease of implantation and use (no extra parameters need to be tuned except the predefined network architecture).

4. Modeling Algorithm

4.1. Selection of Secondary Variables by PCA-Based Dimension Reduction. PCA is a kind of method trying to grasp the main contradiction part in statistical analysis process and analyze the main influencing factors from multiple objects in order to simplify the complex problems. Actually, the principle components conducted by PCA are the combination of column vectors picked by varimax from input matrix. Since correlations and noises always existed in practical industrial data, principle components with a small variance are usually some noisy information. Abandoning this data will not cause

a crucial information loss and can even achieve denoising to some extent.

Data set as shown in the following equation is considered here:

$$u_i = \mathbf{X}v_i, \quad (2)$$

where $\mathbf{X}_{n \times m} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m]$ is the measured n data array on m variables, u_i is the score vector, and v_i is the characteristic unit vector of covariance matrix $\mathbf{X}^T \mathbf{X}$, named load vector. The variance of u_i is λ_i which is also the eigenvalue of $\mathbf{X}^T \mathbf{X}$ and satisfies $\text{Var}(t_i) = \lambda_i, \lambda_1 \geq \dots \geq \lambda_m \geq 0$. PCA is also a procedure used to explain the variance in a single data matrix. The principal component decomposition of \mathbf{X} in (2) can be represented as follows:

$$\mathbf{X} = \mathbf{U}\mathbf{V}^T = \sum_{i=1}^m u_i v_i^T + \mathbf{E}, \quad (3)$$

where $u_i v_i^T$ is the i th principal component and \mathbf{E} is a matrix of residuals. It is to be noted that the score vectors are orthogonal and so are the loading vectors which are of unit length.

Equation (3) indicates that a rank n matrix \mathbf{X} can be decomposed as the sum of n rank 1 principal components. The number of principle components kept in (3) is determined by the total variance. The variance contribution and the total variance of principal component can be represented as follows:

$$\eta_k = \frac{\lambda_k}{\left(\sum_{j=1}^m \lambda_j\right)}, \quad (4)$$

$$C\eta_k = \sum_{i=1}^k \eta_i = \frac{\sum_{i=1}^k \lambda_i}{\sum_{j=1}^m \lambda_j},$$

where η_k is the k th principle component variance contribution and $C\eta_k$ is the total variance of the first k terms. Usually, the total variance varies should be larger than 85%. Only in this case can the data dimension be reduced on the premise of not losing useful information.

After data dimension reduction and noise filtering through PCA, the data measurements are represented as

$$\mathbf{X} \approx \mathbf{U}_k \mathbf{V}_k^T = \sum_{i=1}^k u_i v_i^T, \quad (5)$$

where k is the number of remaining principle components, \mathbf{U}_k is the score vector of the first k terms, and \mathbf{V}_k is the loading vector of the first k terms.

Remark 4. A problem of the PCA-based dimension reduction is that the conducted principle components are comprehensive representation of the original higher-dimension physical variables. However, by computing the *component matrix* which contains the correlations between the principle component and the original physical variable, one can obtain the lower-dimension physical variables which related to the principle components mostly, according to some specific requirements.

4.2. ELM with Self-Feedback Connection. Extreme learning machine (ELM) is an algorithm for single hidden layer feedforward networks (SLFNs) with additive or radial basis function (RBF) hidden nodes whose learning speed can be thousands of times faster than conventional feedforward network learning algorithm like BP algorithm while reaching better approximation performance. In real application, net tends to be used for a finite data set. Huang and Babri prove that a SLFN with at most N hidden nodes and with almost any nonlinear activation function can learn N distinct observations with zero error [18]. And based on the work of [17], the SLFNs (with N hidden neurons) with arbitrary chosen input weights and bias were proved to have the ability to learn N distinct observations with arbitrary small error.

The procedure of the algorithm used here can be summarized as follows: for N arbitrary distinct samples $(\mathbf{X}_i, \mathbf{Y}_i)$, where $\mathbf{X}_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in \mathbf{R}^n$ and $\mathbf{Y}_i = [y_{i1}, y_{i2}, \dots, y_{im}]^T \in \mathbf{R}^m$, the output of a SLFN with \tilde{N} hidden nodes can be represented by

$$f_{\tilde{N}}(\mathbf{X}) = \sum_{i=1}^{\tilde{N}} \beta_i G(\mathbf{a}_i, b_i, \mathbf{X}), \quad \mathbf{X} \in \mathbf{R}^n, \mathbf{a}_i \in \mathbf{R}^n, \quad (6)$$

where \mathbf{a}_i and b_i are the learning parameters of hidden nodes, β_i is the output weight, and $G(\mathbf{a}_i, b_i, \mathbf{X})$ is the output of the i th hidden node with respect to the input data \mathbf{X} .

- (i) For additive hidden node, $G(\mathbf{a}_i, b_i, \mathbf{X})$ is given by $G(\mathbf{a}_i, b_i, \mathbf{X}) = g(\mathbf{a}_i \odot \mathbf{X} + b_i)$, $b_i \in R$, where \mathbf{a}_i is the input weight vector connecting the input layer to the i th hidden node, b_i is the bias of the i th hidden node, and $\mathbf{a}_i \odot \mathbf{X}$ denotes the inner product of vector \mathbf{a}_i and \mathbf{X} in \mathbf{R}^n .
- (ii) For RBF hidden node, $G(\mathbf{a}_i, b_i, \mathbf{X})$ is given by $G(\mathbf{a}_i, b_i, \mathbf{X}) = g(b_i \|\mathbf{X} - \mathbf{a}_i\|)$, $b_i \in R^+$, where \mathbf{a}_i and b_i are the center and impact factor of i th RBF node and R^+ indicates the set of all positive real values.

Remark 5. For the prediction modeling problem considered in this paper, $\mathbf{X}_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T$ are the data of the model inputs variables selected by PCA. $\mathbf{Y}_i = [y_{i1}, y_{i2}, \dots, y_{im}]^T$ are the data of molten iron silicon content that are to be estimated online, which means that $m = 1$. Moreover, the additive hidden node is used in our prediction modeling due to its simple structure.

In supervised batch learning, the learning algorithms use a finite number of input-output samples for training. For N arbitrary distinct samples $(\mathbf{X}_j, \mathbf{Y}_j)$, if a SLFN with \tilde{N} additive hidden nodes can approximate these N samples with zero error, it then implies that there exist \mathbf{a}_j, b_j , and β_j such that

$$f_{\tilde{N}}(\mathbf{X}_j) = \sum_{i=1}^{\tilde{N}} \beta_i G(\mathbf{a}_i, b_i, \mathbf{X}) = \mathbf{Y}_j, \quad j = 1, \dots, N. \quad (7)$$

Equation (7) can be written compactly as

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{Y}, \quad (8)$$

where

$$\mathbf{H}(\mathbf{a}_1, \dots, \mathbf{a}_{\bar{N}}, b_1, \dots, b_{\bar{N}}, \mathbf{x}_1, \dots, \mathbf{x}_N) = \begin{bmatrix} g(\mathbf{a}_1 \odot \mathbf{x}_1 + b_1) & \cdots & g(\mathbf{a}_{\bar{N}} \odot \mathbf{x}_1 + b_{\bar{N}}) \\ \vdots & \ddots & \vdots \\ g(\mathbf{a}_1 \odot \mathbf{x}_N + b_1) & \cdots & g(\mathbf{a}_{\bar{N}} \odot \mathbf{x}_N + b_{\bar{N}}) \end{bmatrix}_{N \times \bar{N}}, \quad (9)$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_{\bar{N}}^T \end{bmatrix}_{\bar{N} \times m}, \quad \mathbf{Y} = \begin{bmatrix} y_1^T \\ \vdots \\ y_N^T \end{bmatrix}_{N \times m}. \quad (10)$$

The purpose of ELM is training the net to find a least-squares solution $\hat{\beta}$ of the linear system $\mathbf{H}\beta = \mathbf{Y}$:

$$\begin{aligned} & \|\mathbf{H}(a_1, \dots, a_{\bar{N}}, b_1, \dots, b_{\bar{N}})\hat{\beta} - \mathbf{Y}\| \\ & = \min_{\beta} \|\mathbf{H}(a_1, \dots, a_{\bar{N}}, b_1, \dots, b_{\bar{N}})\beta - \mathbf{Y}\|. \end{aligned} \quad (11)$$

And the solution of the above linear system can be solved by the inverse of matrix β by the Moore-Penrose method, which is

$$\hat{\beta} = \mathbf{H}^\dagger \mathbf{Y}, \quad (12)$$

where \mathbf{H}^\dagger is the Moore-Penrose generalized inverse of the hidden layer output matrix \mathbf{H} [15–17].

Remark 6. For the simplicity of the paper, the prediction modeling process based on ELM with additive hidden node is summarized as follows: giving a training set $\mathbf{Z} = \{(\mathbf{X}_i, \mathbf{Y}_i) \mid \mathbf{X}_i \in \mathbf{R}^n, \mathbf{Y}_i \in \mathbf{R}^m, i = 1, \dots\}$ for prediction modeling and hidden neuron number \bar{N} , the input weight \mathbf{a}_i and bias b_i can be assigned arbitrarily to calculate the output matrix \mathbf{H} of hidden layer by using (9). After that, the output weight β can be calculated by (12), which is essential for estimating output only based on estimating inputs.

Remark 7. The hidden node number \bar{N} is the only parameter that needs to be predefined in the presented modeling method. In order to achieve optimal approximation ability of training and realize fast convergence aiming at complex industrial data, a proper (maybe optimal) \bar{N} can be determined as the one which results in the lowest validation error through several trainings and validations.

5. Industrial Experiments

5.1. Model Development. In this section, a medium-sized blast furnace (as shown in Figure 3) with the working volume of 2000 m³ in Liuzhou Iron & Steel Group Co. is chosen to perform the validation of the proposed silicon content prediction method. On the foundation of process mechanism and existing monitoring instruments status, measurable parameters influencing silicon content are determined as blast temperature (°C), blast pressure (kPa), oxygen enrichment percentage (%), flow rate of rich oxygen (m³/h), gas permeability (m³/min·kPa), gas volume of bosh (m³/min),



FIGURE 3: The 2nd BF of Liuzhou Iron & Steel Group Co.

TABLE 1: Direct detecting parameters and their instrumentations.

Variable (unit)	Notation	Instrumentation (notation)
Flow rate of cold air (m ³ /min)	q_c	HH-WLB differential pressure flowmeter (FT)
Flow rate of rich oxygen (m ³ /h)	q_o	A+K balance flowmeter (FT)
Blast pressure (kPa)	p_h	DPharp EJA high accuracy pressure transmitter (PT)
Furnace top pressure (kPa)	p_f	DPharp EJA high accuracy pressure transmitter (PT)
Blast temperature (°C)	t_h	Hongguang SBW temperature transmitter (TT)
Blast humidity (g/m ³)	h_c	Air humidity sensor (HT)

bosh gas index (m³/min·m²), blast kinetic energy (kJ/s), blast humidity (g/m³), flow rate of cold air (m³/h), feed blast ratio (wt%), resistance coefficient, volume of coal injection (kg/t), theoretical burning temperature (°C), actual wind speed (m/s), and furnace top pressure (kPa). Figure 4 is a schematic diagram of this actual BF ironmaking system and its measurement system. Through this figure, one can get an intuitive understanding of the detection position distribution of each measurable parameter. In Figure 4, the direct detecting parameters are explained as shown in Table 1, and the indirect detecting variables and their calculation formulas by the direct detecting parameters are listed in Table 2.

Considering the impact of strong correlation between the selected 16 input variables, PCA is used to determine the key input variables that influence the molten iron silicon content mostly. According to (4), the eigenvalue and the variance contribution rate of each component can be calculated as shown in Figure 5. It can be summarized that the cumulative variance contribution rate of the first 6 terms is 98.723%. This means these 6 principal components are sufficient to describe the major variances in the data. Then, by computing the component matrix of principle components, 6 process variables can be determined as the secondary input of the [Si] prediction model. These secondary variables include hot blast pressure x_1 (kPa), hot blast temperature x_2 (°C), oxygen enrichment percentage x_3 (%), volume of coal injection x_4

TABLE 2: Indirect detecting parameters and their calculation formulas.

Variable (unit)	Calculation formulas
Oxygen enrichment percentage (%)	$((0.0163q_o + ((0.21 + (0.29h_c/800)) \times (q_c/60)))/((q_c/60) + (q_o/60)) - (0.21 + (0.29h_c/800))) \times 100$
Gas permeability ($m^3/min \cdot kPa$)	$100q_c/(p_h - p_f)$
Gas volume of bosh (GVB) (m^3/min)	$(1.21q_c/60) + (q_o/30) + (44.8h_cq_c/6000) + (44.8h_cq_o/6000) + ((22.4VCI \times \text{Hydrogen content in coal})/12)$
Bosh gas index ($m^3/(min \cdot m^2)$)	$GVB/78.5398125$
Blast kinetic energy (kJ/s)	$(0.021q_c + ((q_ch_c/60000) + (q_oh_c/60000))/(1 - (h_c/803.6)))/0.2AWS^2/50$
Feed blast ratio (wt%)	$q_c/2000$
Resistance coefficient	$((10000p_h)^2 - 100p_h^2)/GVB^{1.7}$
Theoretical burning temperature ($^{\circ}C$)	$1559 + (0.839t_h) + (4972q_o/q_c) - (6.033h_c) - (3.15VCI \times 1000000/q_c)$
Actual wind speed (AWS) (m/s)	$0.101325(273 + t_h)/(273(0.101325 + p_h)) \times (q_c/3600 \times 4/3.14/30)$

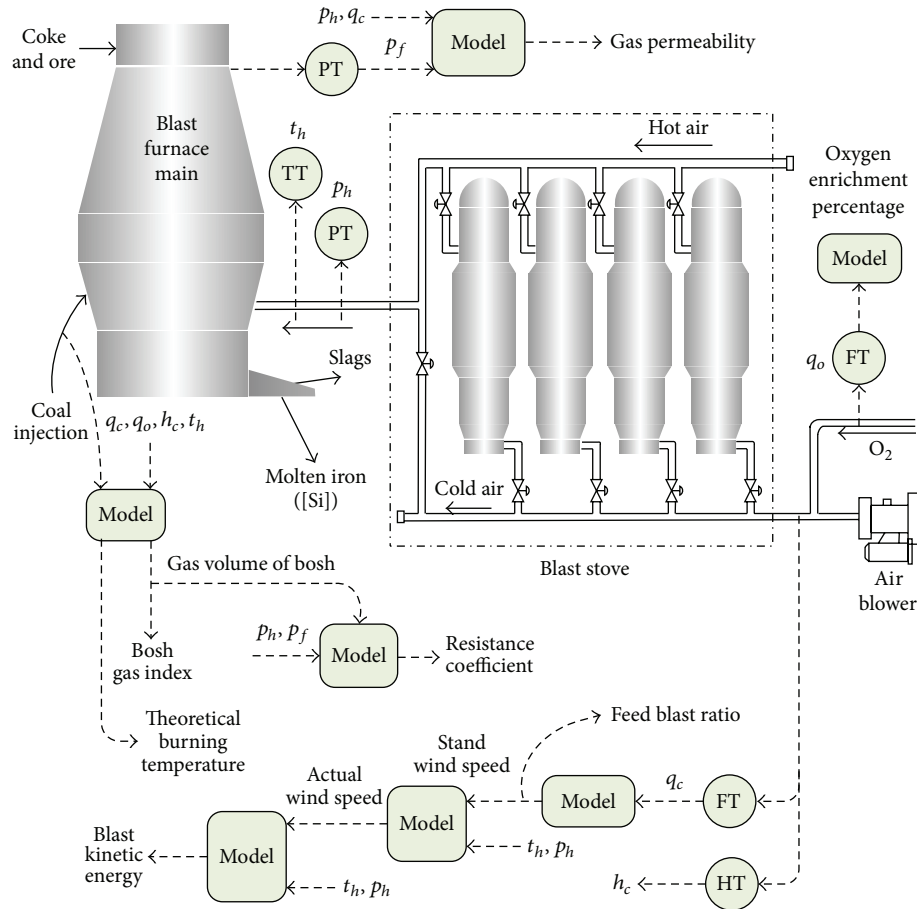


FIGURE 4: Schematic diagram of blast furnace system.

(Kg/t), blast humidity x_5 (g/m³), and gas volume of bosh x_6 (m³/min).

Figure 6 displays the modeling data sets collected from the 9 to 21 October 2013. To better exhibit the performance of ELM based prediction model, optimal learning parameters

need to be made. For the sake of simplicity, we mainly discuss the reason of selecting the optimal number of hidden nodes for the ELM algorithm. The optimal number of hidden units is selected as the one which results in the lowest validation error. Through experiments analysis, the optimal number of

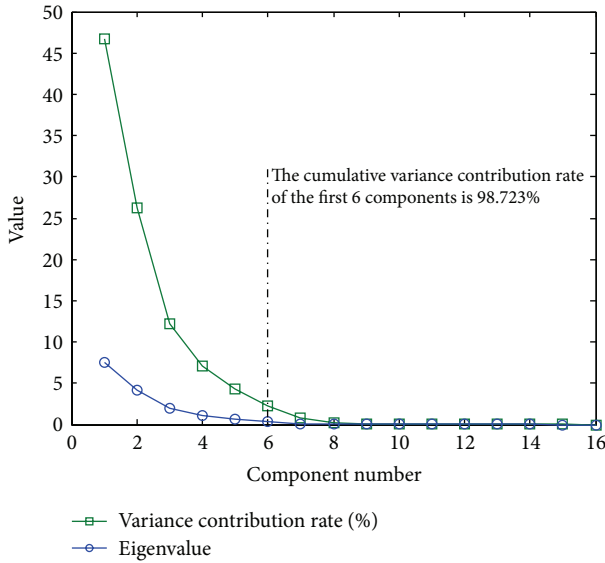


FIGURE 5: Eigenvalue and variance contribution rate of each component.

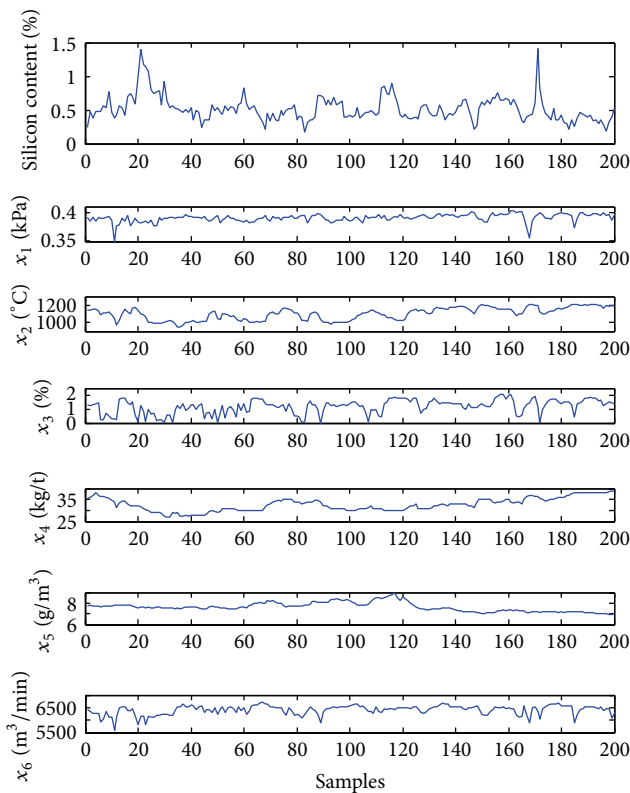


FIGURE 6: Data sets for prediction modeling.

hidden nodes with sigmoidal function is set as $\tilde{N} = 25$. The corresponding modeling result of the developed ELM model with self-feedback structure is shown in Figure 7, where the good modeling accuracy with practical data has been demonstrated.

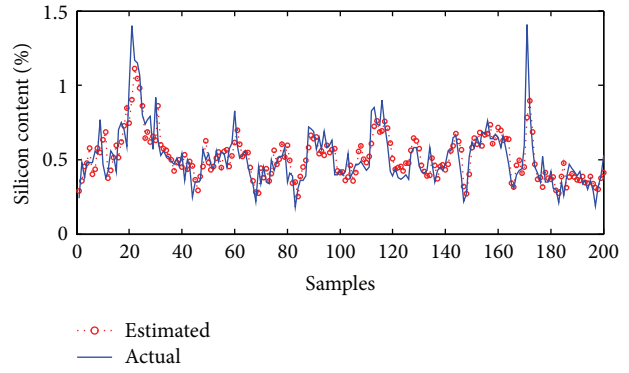


FIGURE 7: Modeling results with the proposed method.

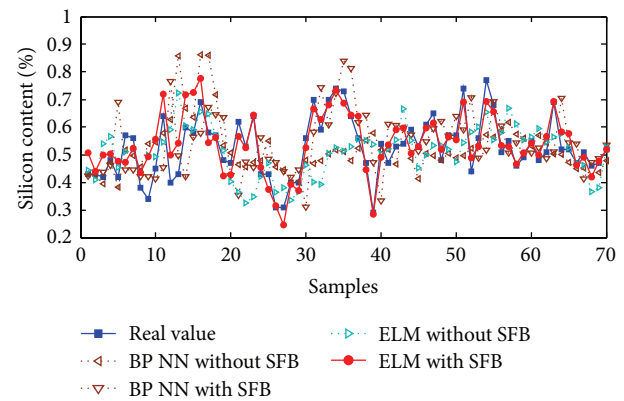


FIGURE 8: Estimation results of molten iron silicon content with different models.

5.2. Experiments Results. The developed ELM based prediction model has been tested on 2[#] blast furnace in Liuzhou Steel of China for quite a long time. Figure 8 shows the estimated results using the proposed modeling method for predicting [Si], where the figure compares the predicted trend with the actual one. Moreover, in order to show the superiority of the proposed method more intuitively, comparisons with various popular prediction models have been made. Here, back propagation neural network without self-feedback (SFB) connection (BP NN for short), BP NN with SFB connection, and traditional ELM without SFB connection have been chosen to conduct the prediction comparison on the same observations. Moreover, the *Levenberg-Marquardt* algorithm is used in traditional BP learning algorithm, which appears to be the fastest method for training moderate-sized feedforward neural networks [16]. From Figure 8, it can be seen that the proposed model has the best estimation performance among all the developed prediction models. For example, it results in the best estimation trend and accuracy, and the shapes of the estimated curve values match the measured ones very well and better than that with other three methods.

It is well known that a good model should have its estimated error autocorrelation close to a white noise. So, in this text, we draw the autocorrelation function of estimating

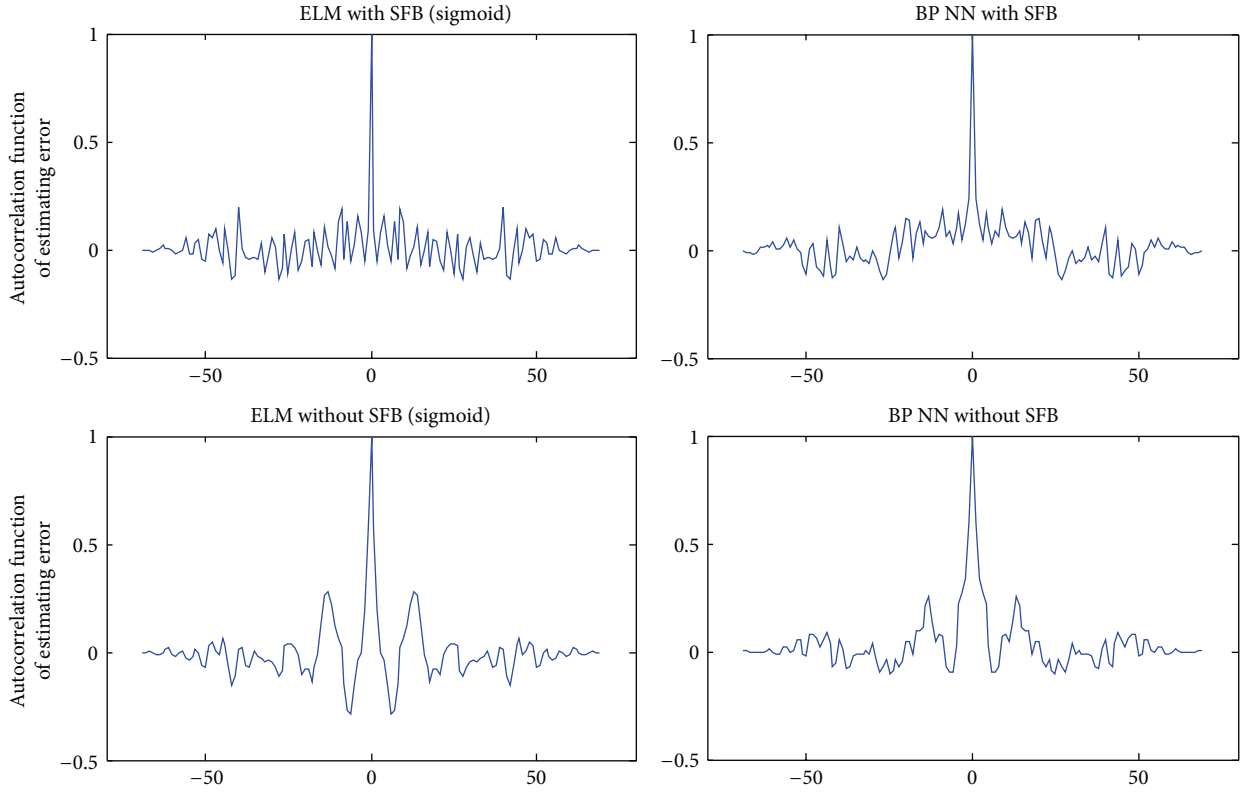


FIGURE 9: Autocorrelation function of estimating error of different models.

error of different models as shown in Figure 9. It can be seen from these figures that the autocorrelation results of algorithm like BP NN without SFB connection and ELM (sigmoid) without SFB connection are much worse than that with a SFB structure, respectively. Although one can obtain that the measuring error autocorrelations of the proposed ELM with SFB connection and BP NN with SFB connection are all satisfactory and close to the shape of the white noise here, the above estimation result (as shown in Figure 8) confirmed the effectiveness and superiority of the proposed method in predicting accuracy.

The estimation and generalization performance of the developed models can be further evaluated quantitatively by calculating the validation accuracy on the testing data set using the standard statistical measures, such as the *root mean squared error* (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{M} \sum_{i=1}^M (\tilde{y}_i - y_i)^2} \quad (13)$$

and the *standard deviation* δ of estimation error $|\tilde{y}_i - y_i|$

$$\delta = \sqrt{\frac{1}{M} \sum_{i=1}^M \left(|\tilde{y}_i - y_i| - \frac{1}{M} \sum_{j=1}^M |\tilde{y}_j - y_j| \right)^2}, \quad (14)$$

where M stands for the number of data points in the time series to be estimated, y_i is the actual value of time series, and \tilde{y}_i is the estimated value at time i by the prediction model.

TABLE 3: Some data statistics of each algorithm.

Algorithm	δ (testing)	RMSE (testing)	Time (seconds)	
			Training	Testing
BP NN without SFB	0.0827	0.1293	0.849955	0.013285
BP NN with SFB	0.0804	0.1272	0.926292	0.013828
ELM without SFB	0.0770	0.1187	0.000678	0.000072
ELM with SFB	0.0355	0.1106	0.000546	0.000089

Table 3 shows the calculated RMSE, δ , and consuming time in training or testing procedure using different methods, respectively. It can be seen from this table that the ELM with feedback connection presented in this paper obtains a much less RMSE and δ than other contrastive model algorithms. Moreover, the proposed model (sigmoid) spent the smallest 0.000546 s CPU time for training and an even faster time 0.000089 s for testing and obtained a very reasonable result. Through the analysis, it can also be seen that, no matter what kinds of algorithm are used, both training and testing results of a model with dynamic feedback are much better than the model without feedback structure when other conditions are the same, which confirm the effectiveness of our developed dynamic self-feedback model structure.

Moreover, the results of practical application indicate that the performance of the developed model is superior to other models and can overcome the problem of “over fitting” excellently. And the gap between every training and

testing is small, which enhances the reliability of the proposed method. The method can also overcome blindness of predefined parameters selection of conventional algorithm; thus convenience is provided for the operators.

6. Conclusions

This paper proposed a data-driven modeling for prediction of molten iron silicon content using PCA and ELM with self-feedback structure. Unlike other methods used for silicon content prediction, the proposed method can predict silicon content more accurately with an extremely fast speed than conventional algorithm, which feed the need for real-time control. Apart from selecting the number of hidden nodes, no other control parameter has to be chosen; thus convenience is provided for the operators. Moreover, the modified ELM with self-feedback structure can overcome the static mapping limitation of traditional ELM and so can cope with dynamic time-series prediction problems very well. Performance of the proposed modified ELM based prediction model is compared with BP algorithm and different model structure on practical industrial data obtained from 2[#] BF in Liuzhou Steel Company of China. The accuracy can basically meet the requirements of actual operation.

Conflict of Interests

The authors have declared that they have no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (61104084, 614730646, 61290323, and 61333007), the Fundamental Research Funds for the Central Universities (N130508002 and N130108001), the IAPI Fundamental Research Funds (2013ZCX02-09), and the III Project (B08015). The authors would like to thank the anonymous reviewers for their constructive comments and the editors for their efforts in editing and polishing the paper. The authors would also like to thank the Ironmaking Plant of Liuzhou Iron & Steel Group Co. in China for providing a lot of experimental support.

References

- [1] L. Jian, C. H. Gao, and Z. H. Xia, "Constructing multiple kernel learning framework for blast furnace automation," *IEEE Transactions on Automation Science and Engineering*, vol. 9, no. 4, pp. 763–777, 2012.
- [2] C. H. Gao, L. Jian, and S. H. Luo, "Modeling of the thermal state change of blast furnace hearth with support vector machines," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 2, pp. 1134–1145, 2012.
- [3] W. Birk, O. Marklund, and A. Medvedev, "Video monitoring of pulverized coal injection in the blast furnace," *IEEE Transactions on Industry Applications*, vol. 38, no. 2, pp. 571–576, 2002.
- [4] H. Saxen, C. H. Gao, and Z. W. Gao, "Data-driven time discrete models for dynamic prediction of the hot metal silicon content in the blast furnace—a review," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 4, pp. 2213–2225, 2013.
- [5] S. Ueda, S. Natsui, H. Nogami, J.-I. Yagi, and T. Ariyama, "Recent progress and future perspective on mathematical modeling of blast furnace," *ISIJ International*, vol. 50, no. 7, pp. 914–923, 2010.
- [6] M. Phadke and S. M. Wu, "Identification of multiinput-multioutput transfer function and noise model of a blast furnace from closed-loop data," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 944–951, 1974.
- [7] M. Castore, G. Gandolfi, S. Palella, and G. Taspardini, "Dynamic model for hot-metal Si prediction in blast-furnace control," in *Proceedings of the Developments Ironmaking Practice*, pp. 152–159, Iron and Steel Institute, London, UK, 1972.
- [8] Y. C. Chao, C. W. Su, and H. P. Huang, "The adaptive autoregressive models for the system dynamics and prediction of blast-furnace," *Chemical Engineering Communications*, vol. 44, pp. 309–330, 1986.
- [9] T. Bhattacharya, "Prediction of silicon content in blast furnace hot metal using partial least squares (PLS)," *ISIJ International*, vol. 45, no. 12, pp. 1943–1945, 2005.
- [10] J. Jiménez, J. Mochón, J. S. de Ayala, and F. Obeso, "Blast furnace hot metal temperature prediction through neural networks-based models," *ISIJ International*, vol. 44, no. 3, pp. 573–580, 2004.
- [11] H. Saxén and F. Pettersson, "Nonlinear prediction of the hot metal silicon content in the blast furnace," *ISIJ International*, vol. 47, no. 12, pp. 1732–1737, 2007.
- [12] J. Chen, "A predictive system for blast furnaces by integrating a neural network with qualitative analysis," *Engineering Applications of Artificial Intelligence*, vol. 14, no. 1, pp. 77–85, 2001.
- [13] X. Tang, L. Zhuang, and C. Jiang, "Prediction of silicon content in hot metal using support vector regression based on chaos particle swarm optimization," *Expert Systems with Applications*, vol. 36, no. 9, pp. 11853–11857, 2009.
- [14] C. Gao, Q. Ge, and L. Jian, "Rule extraction from fuzzy-based blast furnace SVM multiclassifier for decision-making," *IEEE Transactions on Fuzzy Systems*, vol. 22, no. 3, pp. 586–596, 2014.
- [15] N.-Y. Liang, G.-B. Huang, P. Saratchandran, and N. Sundararajan, "A fast and accurate online sequential learning algorithm for feedforward networks," *IEEE Transactions on Neural Networks*, vol. 17, no. 6, pp. 1411–1423, 2006.
- [16] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: a new learning scheme of feedforward neural networks," in *Proceedings of the IEEE International Joint Conference on Neural Networks*, vol. 2, pp. 985–990, IEEE, July 2004.
- [17] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006.
- [18] G.-B. Huang and H. A. Babri, "Upper bounds on the number of hidden neurons in feedforward networks with arbitrary bounded nonlinear activation functions," *IEEE Transactions on Neural Networks*, vol. 9, no. 1, pp. 224–229, 1998.
- [19] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics*, vol. 42, no. 2, pp. 513–529, 2012.
- [20] Y. Lan, Y. C. Soh, and G.-B. Huang, "Ensemble of online sequential extreme learning machine," *Neurocomputing*, vol. 72, no. 13–15, pp. 3391–3395, 2009.

- [21] R. Moreno, F. Corona, A. Lendasse, M. Graña, and L. S. Galvão, "Extreme learning machines for soybean classification in remote sensing hyperspectral images," *Neurocomputing*, vol. 128, pp. 207–216, 2014.
- [22] Y. J. Sun, Y. Yuan, and G. R. Wang, "Extreme learning machine for classification over uncertain data," *Neurocomputing*, vol. 128, pp. 500–506, 2014.
- [23] R. Savitha, S. Suresh, and N. Sundararajan, "Fast learning circular complex-valued extreme learning machine (CC-ELM) for real-valued classification problems," *Information Sciences*, vol. 187, pp. 277–290, 2012.
- [24] Q. Yu, Y. Miche, E. Séverin, and A. Lendasse, "Bankruptcy prediction using extreme learning machine and financial expertise," *Neurocomputing*, vol. 128, pp. 296–302, 2014.
- [25] J. W. Cao, Z. P. Lin, G.-B. Huang, and N. Liu, "Voting based extreme learning machine," *Information Sciences*, vol. 185, no. 1, pp. 66–77, 2012.
- [26] J. Cao and Z. Lin, "Bayesian signal detection with compressed measurements," *Information Sciences*, vol. 289, pp. 241–253, 2014.
- [27] J. W. Cao, T. Chen, and J. Fan, "Fast online learning algorithm for landmark recognition based on BoW framework," in *Proceedings of the 9th IEEE Conference on Industrial Electronics and Applications*, pp. 1163–1168, Hangzhou, China, June 2014.
- [28] J. Zhang, E. Martin, and A. J. Morris, "Fault detection and classification through multivariate statistical techniques," in *Proceedings of the American Control Conference (ACC '95)*, vol. 1, pp. 751–755, June 1995.
- [29] R. P. Good, D. Kost, and G. A. Cherry, "Introducing a unified PCA algorithm for model size reduction," *IEEE Transactions on Semiconductor Manufacturing*, vol. 23, no. 2, pp. 201–209, 2010.
- [30] J. Zhao, W. Wang, Y. Liu, and W. Pedrycz, "A two-stage online prediction method for a blast furnace gas system and its application," *IEEE Transactions on Control Systems Technology*, vol. 19, no. 3, pp. 507–520, 2011.
- [31] L. Jian and C. Gao, "Binary coding SVMs for the multiclass problem of blast furnace system," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 9, pp. 3846–3856, 2013.
- [32] C. Gao, L. Jian, X. Liu, J. Chen, and Y. Sun, "Data-driven modeling based on volterra series for multidimensional blast furnace system," *IEEE Transactions on Neural Networks*, vol. 22, no. 12, pp. 2272–2283, 2011.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

