This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

Rules Extraction from Fuzzy-based Blast Furnace SVMs Multiclassifier for Decision-making

Chuanhou Gao, *Senior Member, IEEE*, Qinghuan Ge, and Ling Jian

***Abstract*—Black-box models play an important role on advancing the blast furnace modeling technologies for control purpose. To further enhance their practical applications, this paper is concerned with the transparency and comprehensibility of the blast furnace black-box models. A fuzzy-based SVMs classification algorithm is proposed to perform the tasks of determining the controllable bound from the real data, of reducing feature from extensive candidate inputs and of training the SVMs model parameters. Based on these results, a fuzzy-based blast furnace SVMs 3-class classifier is constructed to serve for the classification problem according to the output lower than its controlled bound, within the controlled bound and higher than the controlled bound. Further, rules extraction is made to achieve the understandability of the constructed SVMs classifier. Through two typical real blast furnace cases, the extracted rules can work well in classifying the hot metal silicon content into low, proper and high range with high transparency as well as encouraging agreements between the predicted values and the real ones. Moreover, there needs very little information on the blast furnace variables when implementing every rule in practice. The extracted rules provide more explicit and direct indication for the blast furnace operators, and thus may serve better for decision-making on the blast furnace control.**

***Index Terms*—Fuzzy, SVMs classifier, rules extraction, blast furnace, silicon content.**

I. INTRODUCTION

BLAST furnace is the upstream unit operation in the route of manufacturing iron and steel with the main purpose producing pig iron, often called hot metal, from iron ore for primary steelmaking. Fig. 1 presents a typical schematic diagram of a blast furnace reactor. When a blast furnace runs, the lump solid raw materials consisting of ore and coke 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, a crucial region of blast furnace in which the final product "hot metal" gathers. The hot air passes upward through the charge and reacts with the descending coke to generate carbon dioxide, which then changes to carbon monoxide 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 carbon monoxide further reduces the descending iron ore to form hot metal accumulating in the hearth, and some impurities, mostly being calcium oxide and silica, form the slag floating on the hot metal being lighter. The liquid hot metal and slag are periodically tapped out through tapholes for the subsequent processing. It will take 6 to 8 h for each period of ironmaking.

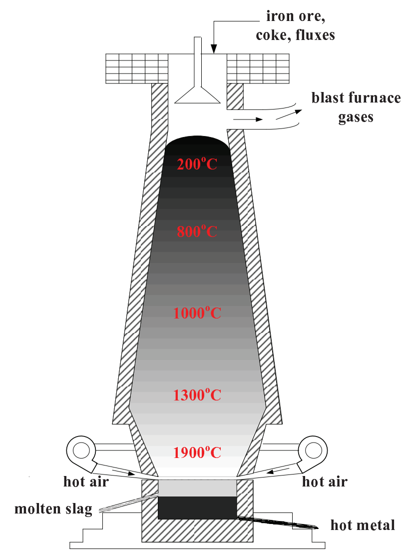

Fig. 1. A schematic diagram of the blast furnace process.

For quite a long time, blast furnace has been regarded as one of the most complex counter-current reactors due to the high number of chemical reactions and transport phenomena occurring between multiphase matters, such as gas-solid, gas-liquid and solid-solid phases interacting. It poses a great challenge to characterize the involved physical and chemical phenomena by quantifying the intricate interactions between blast furnace variables. Ironmaking mechanism-based white-box models [1]–[3] and data-based black-box models [4]–[11] have been extensively developed in the past decades towards the purposes of describing blast furnace process and predicting hot metal component. Based on these studies, great progress has been made in the development of blast furnace modeling techniques. Despite this fact, it should be with caution to use these models for blast furnace control purpose, since the white-box models fail to capture the dynamic disturbances in the ironmaking process, such as unstable slip of furnace charge [12] and aperiodic operations of foremen [13], while the black-box ones can response to these dynamic changes in time but lack of transparency and comprehensibility. Clearly,

Manuscript received April 12, 2012. This work was supported by the National Natural Science Foundation of China under Grant No. 11271326, 61290321 (Major Program), and the Fundamental Research Funds for the Central Universities.

C. H. Gao and Q. H. Ge are with the Department of Mathematics, Zhejiang University, Hangzhou 310027, China (e-mail: gaochou@zju.edu.cn; geqinghuan@126.com).

L. Jian is with the College of Science, China University of Petroleum, Qingdao 266555, China (e-mail: bebetter@upc.edu.cn).

blast furnace white-box models and black-box ones have their own characteristics and complementary advantages. Intuitively, their combination may be more appropriate to blast furnace problem. It is thus expected to develop a blast furnace model that can integrate the advantages of both white-box and black-box models, i.e., being dynamic and also interpretable. An attractive alternative for this purpose is to achieve the understandability of the blast furnace black-box models.

Rules extraction is one kind of important means to enhance the transparency of black-box models, with which the decision may be made for the next step [14], [15]. There are also some studies on rules extraction from a blast furnace system. Martin et al. [16] developed a fuzzy logic based model to predict the hot metal temperature in a Spanish blast furnace. The Takagi-Sugeno-Kang fuzzy system was used to produce the IF-THEN rules with the THEN part as a linear combination of input variables. Based on these IF-THEN rules, the different control actions may be taken to maintain the hot metal temperature at set point. Luo et al. [17] addressed the problem of predicting the hot metal silicon content from two Chinese blast furnaces by a fuzzy system. The fuzzy IF-THEN rules were generated directly from the measured input-output pairs, then these fuzzy rules together with linguistic rules of experienced foremen formed a Fuzzy-Associative-Rules bank with which a fuzzy model was developed finally. Li et al. [18] presented a standardized fuzzy system model to tackle the silicon prediction problem in a small Chinese blast furnace. Like that of Luo et al. [17], a fuzzifier was constructed in [18] to extract some fuzzy rules from the blast furnace input-output pairs. The silicon prediction was performed by a defuzzifier imposed on the extracted fuzzy rules. These studies offer an exciting possibility for describing a blast furnace process with a dynamic and also interpretable model, but still leaving out great potential of improvement. For this reason, the current work is devoted to developing a fuzzy-based blast furnace SVMs multiclassifier from control view of point, and further extracting rules for decision-making on how to control the studied blast furnaces in the next step. In the whole modeling, the fuzzy theory mainly embodies in using the kernel-based fuzzy C-means clustering algorithm [19] to estimate the controlled bound of the silicon content and applying the fuzzy entropy [20] to variables selection.

Fuzzy C-means clustering is a well-known clustering algorithm [21], [22] that takes the weighted mean of all data points as the centroid of a cluster, in which the weight of every data point is the degree of this point belonging to clusters, defined in fuzzy logic. Since it can deal with cases with overlapping clusters, fuzzy C-means clustering has been widely used in the data mining process. However, this algorithm suffers from the presence of noise and outliers in data [23] and also requires significant time for large datasets [24]. To address these problems, some improvement has been made on this algorithm, such as single pass fuzzy C-means for handling large datasets [25] and kernel-based fuzzy C-means for noisy datasets [19], [26]. As real data, blast furnace observations are inevitable to be contaminated by noise, so the kernel-based fuzzy C-means clustering is used to determine the clustering centers of the silicon dataset, and further its controlled bound. The kernel-based fuzzy C-means clustering algorithm integrates the fuzzy C-means clustering algorithm with Mercer kernel function, and can exhibit more strong robustness for pattern recognition problems with outliers or noises [19], [27]. The main improvement is that the kernel-based method computes the membership degree of a input data point (vector) belonging to a cluster in terms of the Euclidian distance between the respective high dimensional feature vectors of this input point and the clustering center, instead of the one between this input data and the clustering center. For this improvement, a nonlinear map is implicitly defined to map the input data space into the high dimensional feature space, like what has been done in the SVMs algorithm. More details about the procedures of the kernel-based fuzzy C-means clustering may be found in [19], [27].

Classification and regression trees (CART) is a kind of frequently-used rules extraction method [28]. Except for being able to extract rules directly from data, CART can be also combined with a certain black-box modeling method, such as SVMs, to improve classification or enhance comprehensibility of the latter. A typical application for the first purpose is to select input features of the SVMs model using CART. Moreira et al. [29] used a CART to select training set based on which the linear SVMs predictor provides the promising results. Huang et al. [30] improved the accuracy of the SVMs model by using hybrid SVM/CART in classifying pathogenic species of bacterial meningitis. Another hybrid use of CART and SVMs is to aim at acquiring comprehensibility of the SVMs model via rules extraction. In this case, a SVMs classifier is constructed firstly; then a reduced training set is collected in the form of support vectors or those points at which the SVMs classifier gives correct classification; finally the CART model trains the reduced training set for rules extraction. Farqud et al. [31] proposed to obtain the reduced training set from support vectors in addressing regression problems using support vector regression. A notable advantage for this hybrid use is that the opaque SVMs model can be made transparency via rules extraction, which can positively affect the studied system acceptance.

The main contribution of this work is to enhance the comprehensibility of the blast furnace black-box SVMs model through CART rules extraction, and provide more direct indication for operators to control the blast furnace hot metal silicon content within proper range. The remainder of this paper is organized as follows: Section II presents the problem formulation. This is followed by a brief review of related methods for classification and rules extraction in Section III. Section IV provides the design of the fuzzy-based blast furnace SVMs multiclassifier and experimental illustration, and rules extraction are made for the studied blast furnaces in Section V. Finally, Section VI concludes this paper.

II. Problem Formulation

Blast furnace control often means to achieve the hot metal temperature and components, such as the silicon content, sulfur content and carbon content within proper bounds. Among these control objectives, the silicon content of the hot metal,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

denoted as z in the context, is the most concerning one since it is closely related to the hot metal quality and also regarded as the chief indicator of the in-furnace thermal status. Typically, a high z means a high cost of coke and excessive generation of heat in the hearth that leads to high furnace temperature, while a low z expresses low furnace temperature, a chilled hearth if the z is too low [6]. From the viewpoint of saving energy, it is better to maintain the z as low as possible when the chilled hearth is avoided.

For the above reasons, the modeling of the z has received much interest in the past decades, but few investigations have been focused on directly tackling the control problem of the z maybe due to facing extreme challenge. Assume that a proper bound of the controlled z is $[z^* - \delta_d, z^* + \delta_u]$, where the z^* is the controlled centriod of the silicon content and $\delta_d + \delta_u$ is the margin of error. These three quantities mainly depend on the quality level of hot metal requested and also differ if they come from different blast furnaces. Generally, the design of a controller is to adjust the blast furnace input variables so that the future z falls in the desired bound. However, the main concern in this work is not to design a controller for the above purpose, but to explore the reasons from the adjustment of some blast furnace input variables that make the z within or beyond the desired bound. The latter may serve as guideline for decision-making to take control actions in the blast furnace process. A feasible solution to address the current concern is to make rules extraction from a SVMs classifier which is performing pattern recognition on the data range of the z, i.e., the following three labels requiring to be recognized: (i) low silicon if $z<z^*-\delta_d$; (ii) proper silicon if $z^*-\delta_d \le z \le z^*+\delta_u$ and (iii) high silicon if $z>z^* + \delta_u$. The rules that say the z proper can be followed while those that say the z low or high should be avoided. Whereas, the rules of saying the z low or high can be conversely utilized to instruct how to control the z within proper bound. In addition, the extracted rules can serve for increasing the comprehensibility of the black-box model, SVMs classifier. To perform the task of extracting rules, it is necessary to understand classification and rules extraction algorithms first.

III. Brief Review of Related Methods

A. Support Vector Machines

SVMs are a kind of kernel-based black-box modeling technique, and were originally introduced for binary classification. The basic idea of SVMs is to design a hyperplane in a high dimensional feature space that can separate two different classes by the greatest extent [32]. This process is undertaken by implicitly defining a high dimensional feature project $\Phi : \mathbb{R}^n \rightarrow \mathcal{F}$ that maps the n-dimensional input patterns $\mathbf{x} \in \mathbb{R}^n$ into the high dimensional feature space $\mathcal{F}$, then finding a linear decision surface $f(\cdot)$ in $\mathcal{F}$ as a binary classifier

$$f(\mathbf{x}) = \mathbf{w}^{\mathrm{T}}\Phi(\mathbf{x}) + b \tag{1}$$

where $\mathbf{w}$ is the normal vector to the hyperplane and $b \in \mathbb{R}$ is the bias term. Mathematically, the classifier of Eq. (1) can be further transformed into the following one

$$f(\mathbf{x}) = \sum_{i=1}^{N} \alpha_i y_i k(\mathbf{x}_i, \mathbf{x}) + b \tag{2}$$

through utilizing kernel trick

$$k(\cdot, *) = \Phi(\cdot)^{\mathrm{T}}\Phi(*) \tag{3}$$

and solving a quadratic programming problem

$$\max_{\alpha_1,\cdots,\alpha_N} \quad \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{N} \alpha_i\alpha_j y_i y_j k\left(\mathbf{x}_i, \mathbf{x}\right) \tag{4}$$

subject to

$$\begin{cases} \sum\limits_{i=1}^{N} \alpha_i y_i = 0 \\ 0 \le \alpha_i \le C, \quad i = 1, \cdots, N \end{cases} \tag{5}$$

Here, $\alpha_1, \cdots, \alpha_N$ are the Lagrange multipliers, $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$ the training data satisfying $\mathbf{x}_i \in \mathbb{R}^n$ and $y_i \in \{-1, 1\}$, $k(\cdot, *)$ the well-known kernel function and C the constant used to penalize the data points crossing the boundaries. The detailed derivative process may be found in [33], [34].

The above binary SVMs classifier can be extended to tackle multiclass classification problem by the following two types of approaches, one is to directly design a multiclass classifier casted as a constrained optimization problem [35], the other is to fuse several binary SVMs into a multiclass classifier. Generally, the former requires to solve a larger optimization problem, so it is thought to be computationally more expensive than the latter [36]. However, the latter decomposes a multiclass problem into multiple independent binary problems so that it cannot capture the correlations between different classes [35]. In practice, there is no criterion saying which means to design a multiclass classifier is better, a proper selection often depends on the specific problem studied.

B. Classification and Regression Trees

CART were pioneered by Breiman et al. [28] to address how to construct a classification or regression model from data based on binary tree structure. Only classification case is considered in this work. There include two main steps to execute the CART algorithm: one is to produce the maximal binary tree by recursive partitioning and the other is to prune this maximal binary tree to reduce the tree complexity but compromised by the model accuracy. During the construction of the maximal binary tree, the Gini diversity index is used to recursively find the most suitable input variable and its corresponding threshold for splitting tree nodes, which is defined in the following.

Definition 3.1 [28]: Given a set S, suppose there are totally m classes that the elements in S can be assigned to, and $p_j (j = 1, \cdots, m)$ are probabilities assigned to each class satisfying $\sum_{j=1}^m p_j = 1$, then the Gini diversity index $G_i(S)$ is defined as

$$G_i(S) = 1 - \sum_{j=1}^{m} p_j^2 \tag{6}$$

TABLE I
A LIST OF CANDIDATE INPUT VARIABLES FROM BLAST FURNACE (A)

Variable name [Unit]	Sym.[#]	Input variable
Basicity of ingredients [wt%]	x_1	$\utilde{q^{0§}}, q^{-1}, q^{-2}, \utilde{q^{-3}}, q^{-4}, q^{-5}$
Blast pressure [kPa]	x_2	$\utilde{q^{0}}, q^{-1}, q^{-2}, q^{-3}, q^{-4}, q^{-5}$
Blast temperature [°C]	x_3	$q^{0}, \utilde{q^{-1}}, q^{-2}, \utilde{q^{-3}}, q^{-4}, q^{-5}$
Blast volume [m^3/min]	x_4	$q^{0}, \utilde{q^{-1}}, q^{-2}, \utilde{q^{-3}}, q^{-4}, q^{-5}$
Coke load of ingredients [wt%]	x_5	$q^{0}, \utilde{q^{-1}}, \utilde{q^{-2}}, q^{-3}, \utilde{q^{-4}}, \utilde{q^{-5}}$
CO percentage in top gas [wt%]	x_6	$q^{0}, \utilde{q^{-1}}, q^{-2}, \utilde{q^{-3}}, \utilde{q^{-4}}, q^{-5}$
CO_2 percentage in top gas [wt%]	x_7	$\utilde{q^{0}}, \utilde{q^{-1}}, q^{-2}, \utilde{q^{-3}}, q^{-4}, q^{-5}$
Feed speed [mm/h]	x_8	$\utilde{q^{0}}, q^{-1}, q^{-2}, \utilde{q^{-3}}, \utilde{q^{-4}}, q^{-5}$
Furnace top pressure [kPa]	x_9	$\utilde{q^{0}}, q^{-1}, \utilde{q^{-2}}, \utilde{q^{-3}}, q^{-4}, \utilde{q^{-5}}$
Furnace top temperature [°C]	x_{10}	$\utilde{q^{0}}, \utilde{q^{-1}}, \utilde{q^{-2}}, q^{-3}, q^{-4}, \utilde{q^{-5}}$
Gas permeability [m^3/min·kPa]	x_{11}	$q^{0}, \utilde{q^{-1}}, q^{-2}, \utilde{q^{-3}}, \utilde{q^{-4}}, \utilde{q^{-5}}$
H_2 percentage in top gas [wt%]	x_{12}	$q^{0}, \utilde{q^{-1}}, q^{-2}, \utilde{q^{-3}}, \utilde{q^{-4}}, q^{-5}$
Oxygen enrichment percentage [wt%]	x_{13}	$\utilde{q^{0}}, q^{-1}, q^{-2}, q^{-3}, q^{-4}, q^{-5}$
Pulverized coal injection [ton]	x_{14}	$\utilde{q^{0}}, q^{-1}, q^{-2}, \utilde{q^{-3}}, \utilde{q^{-4}}, \utilde{q^{-5}}$
Sulfur content [wt%]	x_{15}	$\utilde{q^{-1}}$
Silicon content [wt%]	z	$\utilde{q^{-1}}$

[#] Symbol; [§] represents $q^0 x_1(t)$; variables with wave line are the inputs after feature selection. The marks in Table II have the same meaning.

TABLE II
A LIST OF CANDIDATE INPUT VARIABLES FROM BLAST FURNACE (B)

Variable name [Unit]	Sym.	Input variable
Blast temperature [°C]	x_3	$q^{0}, q^{-1}, q^{-2}, q^{-3}, q^{-4}$
Blast volume [m^3/min]	x_4	$q^{0}, q^{-1}, q^{-2}, q^{-3}, q^{-4}$
Feed speed [mm/h]	x_8	$\utilde{q^{0}}, \utilde{q^{-1}}, q^{-2}, q^{-3}, q^{-4}$
Gas permeability [m^3/min·kPa]	x_{11}	$q^{0}, q^{-1}, \utilde{q^{-2}}, q^{-3}, q^{-4}$
Pulverized coal injection [ton]	x_{14}	$\utilde{q^{0}}, \utilde{q^{-1}}, \utilde{q^{-2}}, \utilde{q^{-3}}, \utilde{q^{-4}}$
Sulfur content [wt%]	x_{15}	q^{-1}
Silicon content [wt%]	z	$\utilde{q^{-1}}$

If the set S is split into S_1 and S_2 under a certain condition C, then the partition Gini diversity index is defined as

$$G_i(S)|_C = \frac{m_1}{m} G_i(S_1) + \frac{m_2}{m} G_i(S_2) \tag{7}$$

where m_1 and m_2 are the numbers of the categories that the elements in S_1 and S_2 can be assigned to, respectively.

From Eq. (6), it is clear that $G_i(S) = 0$ means all samples belonging to one category, so it can act as a measure of purity. For the training set $\mathcal{D}_{tr} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{N}$, the CART algorithm can work as follows [28]:

1) Set the input space $\mathbb{R}^n$ as the root of a binary tree and every input vector $\mathbf{x}_i (i = 1, \cdots, N)$ belongs to this root node.

2) Find the most suitable variable from all input variables and the corresponding threshold as the partition node according to comparing the Gini diversity indexes of all possible partition nodes. The average of two continuous observations of every input variable is thought as a candidate partition node, so there are totally $(N-1) \times n$ possible partition nodes. The optimal partition node gives the lowest Gini diversity index.

3) Two subsamples are produced after the first partition in 2). The same procedure is applied recursively to the two generated subsamples until a certain stopping criterion is satisfied, which yields the maximal binary tree. At this process, Eq. (7) is used to evaluate the partition Gini diversity index.

4) Prune the maximal binary tree by removing some branch nodes without increasing the risk of misclassification, and a set of subtrees are obtained as a result.

5) Pick out the optimal subtree from these candidate subtrees using test sample or cross-validation method.

IV. FUZZY-BASED SVMS CLASSIFICATION ALGORITHM AND BLAST FURNACE EXPERIMENTAL VALIDATION

A. Experimental Data

In the present work, the experimental data is collected from two typical Chinese blast furnaces, one is a medium-sized blast furnace with the inner volume of about 2500 m^3, the other is a pint-sized blast furnace with the volume of about 750 m^3, labeled blast furnace (a) and (b), respectively. The variables closely related to the hot metal silicon content are measured as the candidate inputs for modeling. Tables I and II presents the variables information from these two blast furnaces, in which the last two variables are directly related to the hot metal while the others are related to the blast, the charge or the top gas. As an example, Fig. 2 illustrates the evolution of the hot metal silicon content in the blast furnace (a) and (b). There are totally 800 data points collected with the first 700 points as training set and the remaining 100 points as testing set. The sampling interval is about 1.5 h for the blast furnace (a) while 2 h for the blast furnace (b). Since there is $2-8$ h time delay for the response of the blast furnace outputs to inputs [10], [16], the variables related to the blast, the charge and the top gas with time lags up to the previous 5 taps for the blast furnace (a) and the previous 4 taps for the blast furnace (b) are included as candidate inputs, respectively. However, for the variables related to the hot metal, i.e., the silicon content and sulfur content, only the last silicon content and sulfur content are included for modeling, since too many lag terms of these two variables will strengthen the inertia of silicon models that is not good for capturing some large changes of silicon values [8]. Therefore, there are 86 $(= 14 \times 6 + 2)$ and 27 $(= 5 \times 5 + 2)$ candidate input variables for the blast furnace (a) and (b), respectively. These lagged inputs are also listed in Tables I and II and expressed by the backward shift operator q^{-1}, defined as $q^{-1}x(t) = x(t-1)$. The symbol q^0 indicates $q^0 x(t) = x(t)$.

B. Controlled Bound of the Hot Metal Silicon Content

To design the three-class SVMs classifier mentioned in the Problem Formulation Section, it is first needed to determine the controlled bound of the hot metal silicon content from its historical data. Motivated by Luo et al. [37], data clustering acts as the potential pattern for this task, i.e., estimating $z^* - \delta_d$ and $z^* + \delta_u$. Since z^* is the desired silicon content, it can come directly from a certain clustering center of its historical observations. Whereas, $z^* - \delta_d$ and $z^* + \delta_u$ are the boundaries of the controlled silicon content, so they are taken from the approximate boundaries of clustering centers. From this intuitive analysis, a clustering task with 5 centers is formed to estimate the controlled bound of the hot metal silicon content. The average of the largest clustering center and the second largest one acts as $z^* + \delta_u$, the average of the

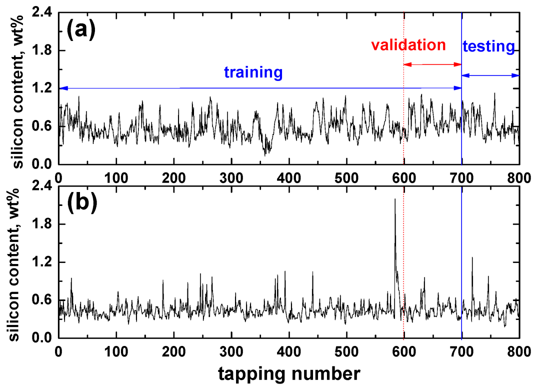

Fig. 2. Evolution of the hot metal silicon content in the blast furnace (a) and (b).

TABLE III
CLUSTERING CENTERS AND CONTROLLED BOUNDS OF THE HOT METAL SILICON CONTENT

Blast furnace	Clustering center		Silicon content: Low $(-1)^\dagger$	Proper $(0)^\dagger$	High $(+1)^\dagger$
(a)	c_α	0.3535	$z<\frac{c_\alpha+c_\beta}{2}=0.4132$	$0.4132\le z\le 0.8251$	$z>\frac{c_\delta+c_\epsilon}{2}=0.8251$
	c_β	0.4728			
	c_γ	0.5920			
	c_δ	0.7272			
	c_ϵ	0.9230			
(b)	c_α	0.3264	$z<\frac{c_\alpha+c_\beta}{2}=0.3736$	$0.3736\le z\le 0.8059$	$z>\frac{c_\delta+c_\epsilon}{2}=0.8059$
	c_β	0.4207			
	c_γ	0.5018			
	c_δ	0.6391			
	c_ϵ	0.9727			

† where -1, 0 and $+1$ indicate the labels for 3-class SVMs classifier.

smallest and the second smallest as $z^*-\delta_d$, and the middlemost one is selected as z^*. This clustering task is also consistent with the one made for identifying the optimal control center of the hot metal silicon content [37].

Applying the kernel-based fuzzy C-means clustering to 5-center clustering task of the silicon dataset can get the results shown in Table III. In the clustering process, the Gaussian kernel

$$k(\mathbf{x}_i,\mathbf{x}_j)=\exp(-\gamma\|\mathbf{x}_i-\mathbf{x}_j\|^2). \quad (8)$$

is used to construct the kenel-based fuzzy C-means clustering algorithm, and the initial centroids are set by randomly selecting 5 points from the samples. Since the final centers are sensitive to the initial centers selected, the results reported in Table III are actually the average of 10 random experimental results. Seen from Table III, the controlled bounds of the hot metal silicon content are $[0.4132, 0.8251]$ and $[0.3736, 0.8059]$ for the blast furnace (a) and (b), respectively. Further counting the numbers of the silicon data points (all samples) within and outside the controlled bounds will yield its distribution in terms of low silicon, proper silicon and high silicon. Shown in Fig. 3 are the results. Clearly, most of the silicon data points of the studied two blast furnaces fall within the respective controlled bound, $574/800=71.75\%$ for the blast furnace (a) and $564/800=70.5\%$ for the blast furnace (b). These results

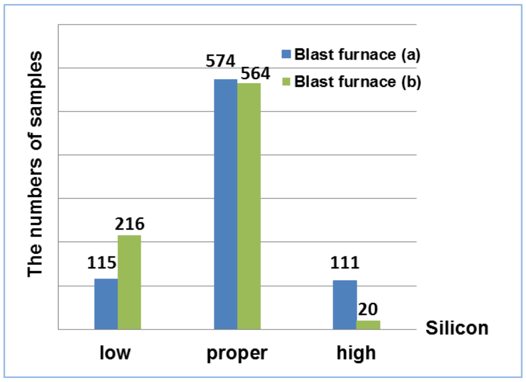

Fig. 3. Distribution of the silicon data points in terms of low silicon, proper silicon and high silicon. The numerals on the top of every column indicate the values of ordinate, such as 115, representing that there are 115 points falling into the range of low silicon for the blast furnace (a), the analogous meaning for other numbers.

conversely render the controlled bound obtained by 5 centers clustering workable.

C. Three Classes Classification of the Hot Metal Silicon Content

With the controlled bound of the hot metal silicon content, it is possible to design a three classes SVMs classifier for the studied blast furnaces. Here, one-against-one method [38] is used to address the current three classes classification problem, in which a binary SVMs classifier is constructed for any two classes among three classes, i.e., 3 binary classifiers needed to be constructed. Like the kernel-based fuzzy C-means clustering, the frequently used Gaussian kernel, Eq. (8), is also selected as the kernel function to construct the three classes SVMs classifier. Therefore, there are kernel parameter γ, penalty factor C, Lagrange multipliers $\alpha_1,\cdots,\alpha_N$ and bias term b needed to be estimated. For this purpose, the grid search for the optimal (γ, C) is firstly made using the training set, i.e., the first 700 points (also including those of other variables listed in Table I or II) in Fig. 2. Since there are very great difference in the magnitude of the candidate inputs, e.g., BT in the magnitude of 10^3 while $[\mathrm{S}]_{N-1}$ in 10^{-2}, the former having a stronger effect on the model parameters than the latter, such data cannot be directly fed into the SVMs models for training. Thus, all the input variables are normalized to the range 0 to 1 through the relationship

$$\bar{x}=\frac{x-x_{\min}}{x_{\max}-x_{\min}} \quad (9)$$

where $\bar{x}$ is the normalized value, x the measured value, $x_{\min}$ the minimum of all measured values of x and $x_{\max}$ the corresponding maximum. With these normalized input variables, the optimal (γ, C) can be searched from the grid set $\{2^{-13},2^{-12.5},\cdots,2^{12.5},2^{13}\}\times\{2^{-13},2^{-12.5},\cdots,2^{12.5},2^{13}\}$ by the ten-fold cross-validation on the previously mentioned training set. This method randomly partitions the training set into 10 subsamples, nine subsamples for training and the remaining one for testing. The cross-validation process is then repeated 10 times, with each of the 10 subsamples used

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

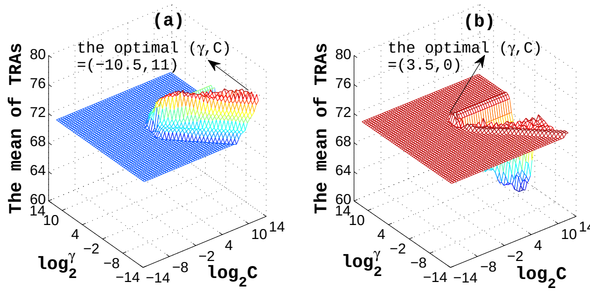

Fig. 4. Ten-fold cross-validation for searching the optimal SVMs model parameters (γ, C) of the blast furnaces (a) and (b). TRAs: the ratio of the numbers of the correct classifications to the size of the training set.

TABLE IV
THREE CLASSES SVMS CLASSIFICATION RESULTS

Blast furnace		Distribution	Silicon content low	Silicon content proper	Silicon content high	TRA‡ (%)	TEA (%)
(a)	86† inputs	True	19	75	6	80.14	79
		prediction	13	86	1		
		correct#	9	70	0		
	42§ inputs	prediction	6	94	0	83.31	79
		correct#	5	74	0		
(b)	27† inputs	True	2	68	30	89.94	68
		prediction	0	100	0		
		correct#	0	68	0		
	9§ inputs	prediction	0	98	2	75.29	70
		correct#	0	68	2		

‡ the training accuracy; † the numbers of original input variables; # the correct prediction; § the numbers of input variables after feature selection.

exactly once as the testing data. The optimal parameters pair corresponds to the one at which the mean testing accuracy (TEA) is the highest. Here, the TEA is defined as the ratio of the number of the correct classifications to the testing set size.

Fig. 4 displayed the grid searching process for (γ, C). There are $53 \times 53 = 2809$ points checked in total. The searching time is about 5993 s and 3462 s for the blast furnaces (a) and (b), respectively, under a Matlab7.1 environment running on a laptop with a 2.20 GHz Intel(R) Core(TM)2 Duo CPU T6670 Processor and a 1.99GB Memory. Clearly, the optimal (γ, C) are $(2^{-10.5}, 2^{11})$ and $(2^{3.5}, 2^{0})$ for the SVMs classifier of the blast furnaces (a) and (b), respectively. Utilizing the optimal parameters pair, it is easy to estimate the Lagrange multipliers $\alpha_1, \cdots, \alpha_N$ and bias term b by solving Eqs. (4) and (5), with all of which the three classes SVMs classifier is further used to perform the classification task on the testing set, i.e., the last 100 data points in Fig. 2. The resutls are shown in Table IV. A look at Table IV may suggest the following information: i) the TRAs (the ratio of the numbers of the correct classifications to the size of the training set) are satisfactory, attaining 80.14% for the blast furnace (a) and 89.94% for (b); ii) the TEAs are encouraging, 79% accuracy rate for the blast furnace (a) superior to that of four classes SVMs classifier [39] but inferior to that of binary SVMs classifier [40], 68% for the blast furnace (b) being the rival of four classes classification accuracy [39]; iii) the predictions that say the silicon content within the proper bound are worth belief, the correctness rate $70/86 = 81.40\%$ for the blast furnace (a) and $68/100 = 68\%$ for the blast furnace (b), while those saying the silicon content outside the proper bound are unreliable; iv) there is large difference between the TRA and the corresponding TEA in the case of the blast furnace (b), more than 20% difference, which implies that overfitting phenomenon is present yet despite using cross-validation to optimize the SVMs models parameters. From these results, the fact is confirmed again that it is more difficult to describe the pint-sized blast furnace than the medium-sized one [41], lower TEA and stronger overfitting phenomenon (meaning worse noise-corrupted) for the former. At the same time, there is still potential to improve the SVMs classifiers by reducing overfitting.

To further reduce overfitting, feature selection is made on the blast furnace input sets for building robust learning model, which also serves for reducing the dimensionality of the underlying SVMs classifier. There are usually two kinds of patterns to perform feature selection, one is forward selection that starts with no variables and works by adding variables one by one until any further addition does not decrease the accuracy, the other is backward selection that starts with all variables and works by removing them one by one until any further removal does not decrease the accuracy. In order to pick out variables more efficiently, variable ranking is firstly made according to a specific criterion, such as F-score, fuzzy entropy, etc. Here, the fuzzy entropy based backward selection is employed for its high efficiency [20]. The larger the fuzzy entropy of a variable is, the less importance this variable is of. During the process of feature selection, the original training set, i.e., containing 700 points, is split into two parts, the first 600 points forms the new training set while the last 100 points acting as the validation set, shown in Fig. 2. Under the condition of the optimal SVMs parameters (γ, C) obtained before feature selection, all original variables are fed into the SVMs model for learning on the new training set and for testing on the validation set, after which the validation set accuracy (VSA), i.e., the ratio of the number of correct classification to the validation set size, will be obtained, 79% for the blast furnace (a) and 69% for (b). These VSAs are denoted as the initial VSAs. Then, the variables are removed one by one according to the fuzzy entropy from the largest to the smallest to observe the VSA. Fig. 5 exhibits the results of the VSAs after removing the input variables in turn, where the numbers with frame are the initial VSAs. If the removal of a variable can make the VSA become higher than the previous maximum VSA, this variable will be deleted. Finally, 42 feature variables (corresponding to 16 blast furnace variables) for the blast furnace (a) and 9 (corresponding to 4 blast furnace variables) for (b) are singled out, respectively, details of which are presented in Table I and II.

These selected feature variables are introduced into the SVMs framework over again, and the original training sets containing 700 points are used to renewedly learn all the SVMs model parameters. The optimal parameters pair (γ, C) become $(2^{-0.5}, 2^{2})$ for the blast furnace (a) and $(2^{2.5}, 2^{0.5})$ for (b) after grid searching. Further, the training and testing results can be obtained from these newly learned SVMs models. Table IV also reports the classification results after feature

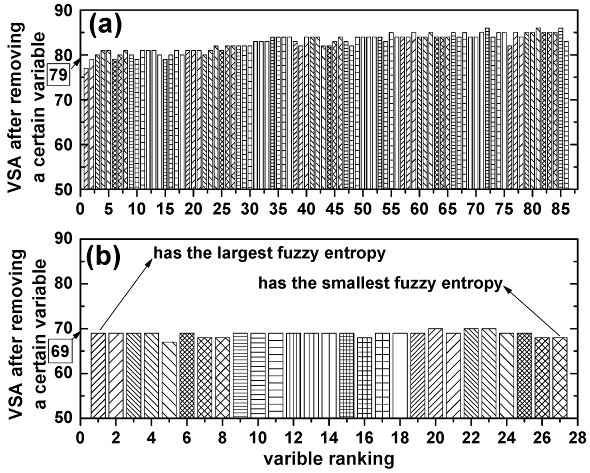

Fig. 5. The validation set accuracy of the blast furnaces (a) and (b) after sequentially removing input variables from the one having the largest fuzzy entropy to the one having the smallest fuzzy entropy.

selection. It is clear that the feature selection has a positive effect on the classification of the studied two blast furnaces. The overfitting phenomenon is greatly reduced in the blast furnace (b) case (decreasing difference between the TRA and the corresponding TEA), and moreover, a little increase of the TEAs is made. For the blast furnace (a), the TRA increases a little but the TEA keeps the same, indicating that the same or better classification results can be obtained with much fewer inputs. Additionally, the models still keep reliability to a certain extent in classifying the silicon content within proper bound, the correctness rate being $74/94 = 78.72\%$ for the blast furnace (a) and $68/98 = 69.39\%$ for (b).

D. Fuzzy-based SVMs Classification Algorithm

From the above discussion, the fuzzy-based SVMs classification algorithm can be summarized as follows:

Algorithm *Fuzzy-based SVMs Classification*

Input: $\{(\mathbf{x}_i, z_i)\}_{i=1}^N$, $\mathbf{x}_i \in \mathbb{R}^n, z_i \in \mathbb{R}$
Output: Three-class SVMs classifier parameters

1: Perform five centers clustering task on series $\{z_i\}_{i=1}^N$ through kernel-based fuzzy C-means clustering algorithm.
2: Denote the five clustering centers from least to greatest by $c_\alpha, c_\beta, c_\gamma, c_\delta, c_\epsilon$.
3: Set $z^* - \delta_d = \frac{c_\alpha + c_\beta}{2}$ and $z^* + \delta_u = \frac{c_\delta + c_\epsilon}{2}$.
4: **for** $j = 1$ to N **do**
5: if $z_j < z^* - \delta_d$, then $y_j = -1$;
6: if $z^* - \delta_d \leq z_j \leq z^* + \delta_u$, then $y_j = 0$;
7: if $z_j > z^* + \delta_u$, then $y_j = 1$;
8: **end for**
9: **for** $l = 1$ to N **do**
10: normalize $\mathbf{x}_l$ to be $\bar{\mathbf{x}}_l$ with every entry in $[0, 1]$
11: **end for**
12: Take $\{(\bar{\mathbf{x}}_i, y_i)\}_{i=1}^N$ as the initial training set for SVMs learning.
13: Make feature selection from $\{\bar{\mathbf{x}}_i\}_{i=1}^N$ by fuzzy entropy based backward selection to generate the new input set $\{\bar{\mathbf{x}}_i'\}_{i=1}^N$ with $\bar{\mathbf{x}}_i' \in \mathbb{R}^{n'}$ and $n' \leq n$.
14: Tune SVMs parameters with grid searching, in which ten fold cross-validation on $\mathcal{D}_{tr} = \{(\bar{\mathbf{x}}_i', y_i)\}_{i=1}^N$ is used.
15: Train SVMs on $\mathcal{D}_{tr} = \{(\bar{\mathbf{x}}_i', y_i)\}_{i=1}^N$ to provide an oracle SVMs mapping a inputs vector to a class label.

V. Rules Extraction from Blast Furnace SVMs Multiclassifier

The fuzzy-based SVMs classifier exhibits encouraging performance in classifying the hot metal silicon content into low, proper and high region. However, it is still unclear how these classification results can serve for decision-making to control the blast furnace system. For this reason, rules extraction is further made from the above black-box SVMs classifier by the CART technique. Like what has been done in the SVMs model, there also includes learning parameters, i.e., producing the maximal binary tree and pruning it, from the training set and evaluating performance through the testing set in the CART algorithm. However, the current training set is composed of those sample points of the original training set after feature selection at which the SVMs model can perform correct classification, while the current testing set is still the original one. With the current training set and according to the procedures of the CART algorithm, the ultimate classification tree can be yielded, as shown in Fig. 6. Clearly, there are 4 rules produced for the blast furnace (a) and 6 rules for the blast furnace (b). Moreover, the rules are all-sided, including those to lead to low silicon, proper silicon and high silicon. This phenomenon conversely renders that the processes during the studied period are persistency of excitation or the gathered data is rich. Further embodying these rules according to the related nodes will obtain the following:

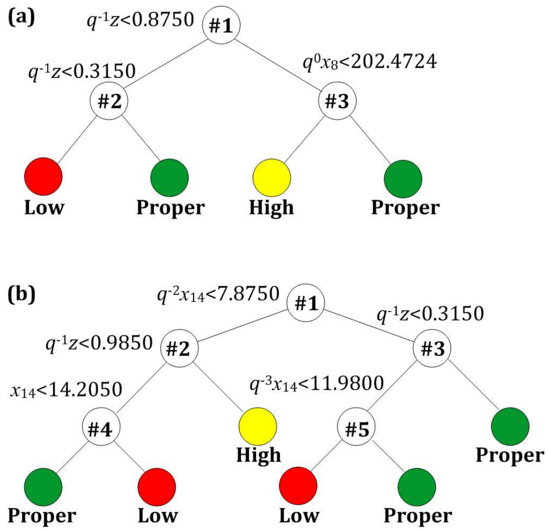

Fig. 6. Binary tree for rules extraction on the data set with correct classifications of the blast furnaces (a) and (b). At each intermediate node, a case goes to the left child node if and only if the condition is satisfied. The predicted class is given beneath each leaf node.

Rules for the Blast Furnace (a)

R_1: **IF** $q^{-1}z<0.3150$, **THEN** low silicon.
R_2: **IF** $q^{-1}z \geq 0.3150$ and $q^{-1}z<0.8750$, **THEN** proper silicon.
R_3: **IF** $q^{-1}z \geq 0.8750$ and $q^0x_8<202.4724$, **THEN** high silicon.
R_4: **IF** $q^{-1}z \geq 0.8750$ and $q^0x_8 \geq 202.4724$, **THEN** proper silicon.

Rules for the Blast Furnace (b)

R_1: **IF** $q^{-2}x_{14}<7.8750$ and $q^{-1}z<0.9850$ and $x_{14}<14.2050$, **THEN** proper silicon.
R_2: **IF** $q^{-2}x_{14}<7.8750$ and $q^{-1}z<0.9850$ and $x_{14} \geq 14.2050$, **THEN** low silicon.
R_3: **IF** $q^{-2}x_{14}<7.8750$ and $q^{-1}z \geq 0.9850$, **THEN** high silicon.
R_4: **IF** $q^{-2}x_{14} \geq 7.8750$ and $q^{-1}z<0.3150$ and $q^{-3}x_{14}<11.9800$, **THEN** low silicon.
R_5: **IF** $q^{-2}x_{14} \geq 7.8750$ and $q^{-1}z<0.3150$ and $q^{-3}x_{14} \geq 11.9800$, **THEN** proper silicon.
R_6: **IF** $q^{-2}x_{14} \geq 7.8750$ and $q^{-1}z \geq 0.3150$, **THEN** proper silicon.

From the above rules, although there refers to 16 blast furnace variables (42 feature variables) in the SVMs classifier for the blast furnace (a), it only needs the information of two blast furnace variables, i.e., feed speed x_8 and the last silicon content $q^{-1}z$, to judge the silicon content low, proper or high, while for the blast furnace (b), only pulverized coal injection x_{14} and the last silicon content are needed. Therefore, the above rules are very simple and convenient for actual usage. At the same time, note that there need two nodes to produce a rule in the case of the blast furnace (a) while three nodes are usually required to yield a rule in the blast furnace (b) case. This result implies that it poses more challenge for controlling the blast furnace (b). Since pint-sized blast furnace generally has stronger noise interference than medium-sized one [41], larger difficulty will be encountered for understanding the dynamics of the former. The difference of the nodes required to produce a rule between the blast furnace (a) and (b) embodies the difference in the difficulties of controlling them. Another notable point is that the rules of the blast furnaces (a) and (b) both require the information on the last silicon content. This phenomenon indicates that the last silicon content is closely related to the current one and also confirms the fact that experienced foremen often utilize the last silicon content to infer the current one and further take possible control actions. In addition, there exists obvious time delay phenomenon between the blast furnace inputs and output, since most of the conditions in the rules are related to the lagged blast furnace variables, which is also consistent with the fact that blast furnace is a system with large delay time. All these results reveal that the above extracted rules are reasonable and attractive to be used for decision-making on the blast furnace control.

Further apply the extracted rules on the testing sets of the blast furnaces (a) and (b) to verify their effectiveness, and the results are evaluated by the following criteria [42] :

i) *Accuracy*, the ability of extracted rules to exactly predict on the testing set, defined as

$$\text{Accuracy} = \frac{Count(\text{correct classifications on the testing set})}{Count(\text{testing set})} \tag{10}$$

where $Count(\cdot)$ is the counting function.

ii) *Fidelity*, the extent to which extracted rules mimic the underlying black-box model, formulated as

$$\text{Fidelity} = \frac{Count(\text{SVMs prediction=rules prediction})}{Count(\text{testing set})} \tag{11}$$

Table V presents the prediction results utilizing the extracted rules. From the accuracy criterion, the extracted rules basically can work as effective as the SVMs classifiers for the studied blast furnaces despite producing a little lower TEAs. Moreover, most of the correct classifications by the SVMs classifiers can be also correctly classified by the extracted rules, high fidelity for both blast furnaces. Namely, the extracted rules can mimic the black-box SVMs model as closely as possible. However, the comprehensibility of the underlying black-box model has been increased greatly through rules extraction, and the explicit information on indicating the silicon to be low, proper or high has been demonstrated by linguistic rules. At this point, the extracted rules can be utilized more directly for decision-making of controlling the blast furnace system than the constructed black-box model. In addition, it should be mentioned that compared with the early fuzzy rules from the blast furnace system [16]–[18], the current ones need less information about the blast furnace variables, so they are more simple and practical.

To compare the performance of the above extracted rules with those extracted from the fuzzy model more visually, the adaptive neuro fuzzy inference system (ANFIS) [43] is constructed in attempt to perform the same classification tasks. A point to note is that the ANFIS can only make the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.

TABLE V
RESULTS OF RULES EXTRACTION FROM THREE CLASSES SVMS CLASSIFIER

Blast furnace	Training data set	Rules number	Accuracy (%)	Fidelity (%)
(a)	correct classification	4	74	93
	original	3	74	93
(b)	correct classification	6	69	96
	original	4	70	79

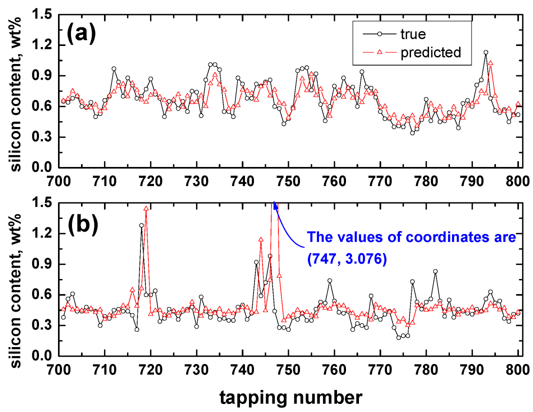

Fig. 7. The numerical prediction of the silicon content through the adaptive neuro fuzzy inference system on the testing sets of the blast furnaces (a) and (b). The arrow points to the invisible point with the values of coordinates $(747, 3.076)$.

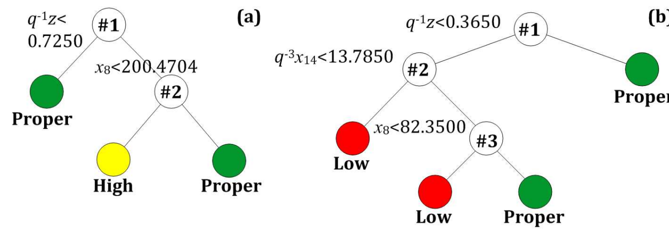

Fig. 8. Binary tree for rules extraction on the original data set of the blast furnaces (a) and (b). At each intermediate node, a case goes to the left child node if and only if the condition is satisfied. The predicted class is given beneath each leaf node.

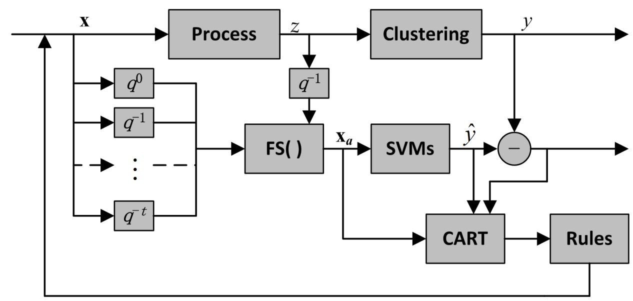

Fig. 9. Block diagram of transparency of the fuzzy-based SVMs black-box models. t, the maximum time delay; FS(), feature selection synthesizer by fuzzy entropy; $\mathbf{x}_a$, the selected inputs; $\hat{y}$, the predicted value of y.

numerical prediction of the silicon content, but cannot give the direct 3-class classification results. It needs to convert the numerical prediction into the classification result according to the silicon divisions in Table III. The binary trees in Fig. 6 render that the CART model only needs the inputs $q^{-1}z$, x_8 to produce the rules in the blast furnace (a) case, and the inputs $q^{-1}z$, x_{14}, $q^{-2}x_{14}$, $q^{-3}x_{14}$ in the blast furnace (b) case. Therefore, a fair comparison can be made if these inputs are fed into the ANFIS for the blast furnaces (a) and (b), respectively. Fig. 7 exhibits the prediction results of these two blast furnaces on their testing sets. Further converting these numerical predictions into the 3-class classification predictions can yield the TEAs to be 76% for the blast furnace (a) and 65% for (b). From the viewpoint of the TEAs, the CART model and the ANFIS have their strong points, the former outperforming the latter in the blast furnace (b) case but the reverse in the blast furnace (a) case. However, the number of fuzzy rules produced by the ANFIS will exponentially increase with the number of inputs, e.g., for the blast furnace (b), there are $5^4 = 625$ (5 is the number of membership functions used) rules produced. The enormous rules number raises higher demand on the computer hardware to balance the model accuracy and efficiency [16].

As a kind of rules extraction method, the CART algorithm can work independent of the SVMs model. Namely, it can be applied directly on the original training set but not those at which the SVMs model has correct classification to produce blast furnace rules. Fig. 8 exhibits the corresponding binary trees for rules extraction and the testing results are evaluated too in Table V. Although some better results are observed, such as a slight increase of accuracy of the blast furnace (b), the extracted rules, as can be seen in Fig. 8, are a little one-sided, no low silicon rule for the blast furnace (a) and no high silicon rule for the blast furnace (b). Therefore, they cannot cover all cases reflecting the silicon range. At this point, it is quite different between the direct use of the CART model and the combination of the SVMs method with the CART algorithm. However, the one-sided rules cannot be chalked up to lack of full excitation on the process or the poor gathered data, since all-sided rules have been extracted from the studied data by combing the SVMs and CART models. The possible reason may be that the CART approach is essentially a kind of white-box modeling methods, which like other white-box models [1]–[3], has difficulty in capturing very few but existent dynamic changes of the blast furnace process. This reason can be further confirmed by the fact that rather low ratio of high silicon or low silicon data points (see Fig. 3) appears in the studied samples. There is much chance that these improper silicon data points result from some unexpected dynamic disturbances [12], [13], so they cannot be described well by pure white-box models. Additionally, it can be noted that the rules numbers extracted from the direct CART models are a little less. In these regards, it will benefit more when applying the CART algorithm to extract rules from the SVMs black-box models of the blast furnace system, through which the advantage (capturing the dynamic disturbances) of the blast furnace SVMs black-box models can be mined fully, and their transparency can be also enhanced.

Resumptively, the kernel-based fuzzy C means clustering, the fuzzy entropy, the SVMs classifier and the CART algorithm are sequentially adopted to classify the hot metal silicon content according to different silicon ranges. A control block diagram in Fig. 9 displays the detailed design principle. Although it is not a novel project to classify the silicon content using the SVMs method or produce fuzzy rules to address the silicon prediction, this work still has large potential for the current concerning blast furnace problems. In our earlier work, the SVMs models were constructed to perform the binary classification task [40] and 4-class task [39] according to the trend change of the silicon content. For the binary classification, the increasing silicon (if $z(t)-z(t-1)>0$) is labeled as one category and the decreasing silicon silicon (if $z(t)-z(t-1)<0$) as the other category, while for the 4-class classification, the increasing (decreasing) silicon is divided into two classes once again according to the increasing (decreasing) extents. These classification modes are quite different from the current one that is classifying silicon according to the silicon range but not the silicon change range. Since control would be the ultimate goal and the blast furnace control often means to control the hot metal temperature and components, such as the silicon content, within acceptable bounds, the current classification mode can serve better for the control purpose. This makes the main difference in comparing the current work with the earlier work [39], [40], but not just confining the difference between 3-class classification and other classification tasks. Another important distinction is that the current work combines the SVMs method with the CART algorithm through which the black-box SVMs model can be enhanced comprehensibility. The extracted rules can account for the output results with detailed and definite input information, which may further serve for the control purpose by linking the output results with control variables. Direct indication on how to adjust inputs can be provided, whereas for the pure black-box SVMs model, the control indication is usually unknown even though good prediction can be obtained. Compared with the direct fuzzy methods, such as ANFIS, the current work can bring into full play the potential of the black-box SVMs models for addressing the blast furnace problems, i.e., capturing the dynamic disturbances, and also avoid the exponential increase of the rules with the inputs. Of course, some challenges have to be encountered in the future, such as integrating a prior knowledge (e.g., the blast furnace process thermodynamics and kinetics) with the current modeling framework, lessening the accuracy difference between the current model and the pure black-box SVMs model, etc.

Notwithstanding obvious difference from our earlier work [39], [40], the contribution made in this work, i.e., using amplitude of input variables as inputs and using CART to improve SVMs classification for rule extraction, are not themselves new. In control applications, both amplitude and difference of variables have been frequently used [44], [45]. As to using CART to improve SVMs classification, there have also been much research on this issue [29]–[31]. At these points, the novelty of the current work seems weak from the methodological viewpoint. However, the purpose of this work is not to develop a new rules extraction method or a new black-box modeling technique, but to address a blast furnace control problem, i.e., to control the hot metal silicon content within a certain range. Since it is not easy to realize the control purpose based on the developed blast furnace models, experience and intuitive of experienced foremen still form the main basis for performing blast furnace control today. For this reason, the current work has tried to address the blast furnace control problem through making rules extraction from blast furnace black-box SVMs model using CART. Therefore, the novelty of this work is to present a new strategy to obtain a possible solution of the blast furnace control issue, for which the control problem is firstly transformed into a 3-class classification problem; then SVMs are used to perform the classification task; finally rules extraction are made from the SVMs black-box model through CART to find the solution of the control problem. Of course, the proposed strategy is not limited to addressing the blast furnace control problem, but can deal with a general control problem analogous to the current one. Also, the used SVMs and CART methods are not unalterable. Other black-box modeling techniques, like neural networks (NN), and other rules extraction methods, like C4.5, can work as well within the framework of the proposed strategy. It will be an interesting attempt to test other combination, such as SVMs and C4.5, NN and CART, etc., in the future investigation.

From the viewpoint of developing blast furnace models, the proposed strategy renders a new modeling thought that can integrate the advantages of blast furnace white-box models and black-box models through enhancing the comprehensibility of the SVMs model. It actually has an effect on advancing blast furnace modeling techniques. The main motivation to pursue the current research is that the operation of a blast furnace is still a serious problem in practice. Although plenty of studies on the modeling and control of a blast furnace system have been carried out in the last decades, experience and intuition of skilled operators are still the main driver to implement blast furnace operation today. This, together with the important impact of blast furnace on the national economy, makes the research on blast furnace modeling and control still very active in the foreseeable future. Another more important reason is that a slight improvement of blast furnace mathematical model may result in considerable profit because of large quantity produced. Hence, even though realization results are not perfect, it is still significant as long as a slight improvement is made.

VI. Conclusion

This paper has developed a fuzzy-based SVMs mutliclassifier to address a 3-class classification problem of the hot metal silicon content from the control viewpoint, for which a fuzzy-based SVMs classification algorithm has been suggested, including the determination of controllable bounds from the real blast furnace data and feature selection from extensive candidate inputs. Then some practicable blast furnace rules are extracted from the constructed SVMs model to instruct the silicon content in low, proper or high range. The encouraging agreements between the predicted values and the real ones reveal the extracted rules can work for the blast furnace system well. Compared with the original SVMs black-box model, the current extracted rules can yield not only the label of classifying the silicon content, but the reason of producing it. Namely, the advantages of black-box models and white-box models have been effectively integrated in the current blast furnace applications. The extracted rules can better serve for making control decision on the blast furnace operation.

REFERENCES

[1] H. Nogami, M. S. Chu, and J. Yagi, "Multi-dimensional transient mathematical simulator of blast furnace process based on multi-fluid and kinetic theories," *Comput. Chem. Eng.*, vol. 29, no. 11-12, pp. 2438-2448, Oct. 2005.

[2] M. S. Chu, X. F. Yang, F. M. Shen, J. Yagi, and H. Nogami, "Numerical simulation of innovative operation of blast furnace based on multi-fluid model," *J. Iron Steel Res. Int.*, vol. 13, no. 6, pp. 8-15, Nov. 2006.

[3] A. Jindal, S. Pujari, P. Sandilya, and S. Ganguly, "A reduced order thermo-chemical model for blast furnace for real time simulation," *Comput. Chem. Eng.*, vol. 31, no. 11, pp.1484-1495, Nov. 2007.

[4] V. R. Radhakrishnan and A. R. Mohamed, "Neural networks for the identification and control of blast furnace hot metal quality, *J. Process Control*, vol. 10, no. 6, pp. 509-524, Dec. 2000.

[5] J. Chen, "A predictive system for blast furnaces by integrating a neural network with qualitative analysis," *Eng. Appl. Artif. Intel.*, vol. 14, no. 1, pp. 77-85, Feb. 2001.

[6] M. Waller and H. Saxén, "Time-varying event-internal trends in predictive modeling methods with applications to ladlewise analyses of hot metal silicon content," *Ind. Eng. Chem. Res.*, vol. 42, no. 1, pp. 85-90, Jan. 2003.

[7] T. Bhattacharya, "Prediction of silicon content in blast furnace hot metal using Partial Least Squares (PLS)," *ISIJ Int.*, vol. 45, no. 12, pp. 1943-1945, Dec. 2005.

[8] H. Sax*é*n, and F. Pettersson, "Nonlinear prediction of the hot metal silicon content in the blast furnace," *ISIJ Int.*, vol. 47, no. 12, pp. 1732-1737, Dec. 2007.

[9] C. H. Gao, J. M. Chen, J. S. Zeng, X. Y. Liu, and Y. X. Sun, "A chaos-based iterated multistep predictor for blast furnace inromaking process," *AIChE J.*, vol. 55, no. 4, pp. 947-962, Apr. 2009.

[10] A. Nurkkala, F. Pettersson, and H. Saxén, "Nonlinear modeling method applied to prediction of hot metal silicon in the ironmaking blast furnace," *Ind. Eng. Chem. Res.*, vol. 50, no. 15, pp. 9236-9248, Aug. 2011.

[11] C. H. Gao, L. Jian, X. Y. Liu, J. M. Chen, Y. X. Sun, "Data-driven modeling based on volterra series for multidimensional blast furnace system," *IEEE Trans. Neural Netw.*, vol . 22, no. 12, pp. 2272-2283, Dec. 2011.

[12] A. Johansson and A. Medvedev, "Detection of incipient clogging in pulverized coal injection lines," *IEEE Trans. Ind. Applicat.*, vol. 36, no. 3, pp. 877-883, May 2000.

[13] T. Miyano, S. Kimoto, H. Shibuta, K. Nakashima, Y. Ikenaga, and K. Aihara, "Time series analysis and prediction on complex dynamical behavior observed in a blast furnace," *Physica D*, vol. 135, no. 3-4, pp. 305-330, Jan. 2000.

[14] R. R. Yager,"Perception-based granular probabilities in risk modeling and decision making," *IEEE Trans. Fuzzy Syst.*, vol. 14, no. 2, pp. 329-339, Apr. 2006.

[15] D. R. Wu and J. M. Mendel, "Linguistic summarization using IF-THEN rules and interval type-2 fuzzy sets," *IEEE Trans. Fuzzy Syst.*, vol. 19, no. 1, pp. 136-151, Feb. 2011.

[16] R. D. Martin, F. Obeso, J. Mochn, R. Barea, and J. Jimnez, "Hot metal temperature prediction in blast furnace using advanced model based on fuzzy logic tools," *Ironmak. Steelmak.*, vol. 34, no. 3, pp. 241-247, May 2007.

[17] S. H. Luo, X. G. Liu, and M. Zhao, "Prediction for silicon content in molten iron using a combined fuzzy-associative-rules bank," in *Proc. Fuzzy Syst. Knowl. Discov.*, Changsha, China, Aug. 2005, pp. 667-676.

[18] Q. H. Li and X. G. Liu, "Fuzzy prediction of silicon content in BF hot metal," *J. Iron Steel Research Int.*, vol. 12, no. 6, pp. 1-4, Nov. 2005.

[19] Z. D. Wu, W. X. Xie, and J. P. Yu, "Fuzzy c-means clustering algorithm based on kernel method," in *Proc. 5th Int. Conf. Comput. Intell. Multimedia Appl.*, Washington, DC, 2003, pp. 49-54.

[20] H. M. Lee, C. M. Chen, J. M. Chen, and Y. L. Jou, "An efficient fuzzy classifier with feature selection based on fuzzy entropy," *IEEE Trans. Syst., Man, Cybern. B*, vol. 31, no. 3, pp. 426-732, June 2001.

[21] J. C. Bezdek, "Cluster validity with fuzzy sets," *J. Cybern.*, vol. 3, pp. 58-71, 1974.

[22] T. C. Havens, J. C. Bezdek, C. Leckie, L. O. Hall, and M Palaniswami, "Fuzzy C-means algorithms for very large data," *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 6, pp. 1130-1146, Dec. 2012.

[23] E. G. Mansoori, "FRBC: a fuzzy rule-based clustering algorithm," *IEEE Trans. Fuzzy Syst.*, vol. 19, no. 5, pp. 960-971, Oct. 2011.

[24] L. O. Hall and D. B. Goldgof, "Convergence of the single-Pass and online fuzzy C-means algorithms," *IEEE Trans. Fuzzy Syst.*, vol. 19, no. 4, pp. 792-794, Aug. 2011.

[25] P. Hore, L. O. Hall, and D. B. Goldgof, "Single pass fuzzy c means," in *Proc. IEEE Int. Fuzzy Syst. Conf.*, Jul. 2007, pp. 1-7.

[26] H. C. Huang, Y. Y. Chuang, and C. S. Chen, "Multiple kernel fuzzy clustering," *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 1, pp. 120-134, Feb. 2012.

[27] X. W. Yang, G. Q. Zhang, J. Lu, and J. Ma, "A kernel fuzzy c-means clustering-based fuzzy support vector machine algorithm for classification problems with outliers or noises," *IEEE Trans. Fuzzy Syst.*, vol. 19, no. 1, pp. 105-115, Feb. 2011.

[28] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, Classification and regression trees. London, U.K.: Chapman & Hall, 1984.

[29] J. M. Moreira, A. M. Jorge, C. Soares, and J. F. De Sousa, "Improving SVM-linear predictions using CART for example selection," in *Proc. 16th Int. Symp. Methodologies Intell. Syst.*, Bari, Italy, 2006, pp. 632-641.

[30] C. Y. Huang, T. H. Tsai, B. C. Wen, C. W. Chung, Y. J. Li, Y. C. Chuang, W. J. Lin, L. L. Li, J. K. Wang, Y. L. Wang, C. H. Lin, and D. W. Wang, "Hybrid SVM/CART classification of pathogenic species of bacterial meningitis with surface-enhanced Raman scattering," in *IEEE Int. Conf. Bioinformatics and Biomedicine*, Hong Kong, China, 2010, pp. 406-409.

[31] M. A. H. Farquad, V. Ravi, and S. B. Raju, " Support vector regression based hybrid rule extraction methods for forecasting," *Expert Syst. Appl.*, vol. 37, no. 8, pp.5577-5589, Aug. 2010.

[32] V. N. Vapnik, *The Nature of Statistical Learning Theoty*. New York: Wiley, 1998.

[33] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines*. Cambridge, U.K.:Cambridge Univ. Press, 2000.

[34] B. Schölkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge, MA: MIT Press, 2002.

[35] K. Crammer and Y. Singer, "On the algorithmic implementation of multiclass kernel-based vector machines," *J. Mach. Learn. Res.*, vol. 2, pp. 265-292, 2001.

[36] C. W. Hsu and C. J. Lin, "A Comparison of methods for multiclass support vector machines," *IEEE Trans. Neural Netw.*, vol. 13, no. 2, pp. 415-425, Mar. 2002.

[37] S. H. Luo, J. Huang, J. S. Zeng, and Q. S. Zhang, "Identification of the optimal control center for blast furnace thermal state based on the fuzzy C-means clustering," *ISIJ Int.*, vol. 51, no. 10, pp. 1668-1673, Oct. 2011.

[38] S. Knerr, L. Personnaz, and G. Dreyfus, "Single-layer learning revisited: A stepwise procedure for building and training a neural network," *In Neurocomputing: Algorithms, Architectures and Applications*, Fogelman-Soulie and Herault, Eds., NATO ASI Series. Springer, 1990.

[39] L. Jian and C. H. Gao, "Binary coding SVMs for the multiclass problem of blast furnace system," *IEEE Trans. Ind. Electron.*, to appear.

[40] C. H. Gao, L. Jian, and S. H. Luo, "Modeling of the thermal state change of blast furnace hearth with support vector machines," *IEEE Trans. Ind. Electron.*, vol. 59, no. 2, pp. 1134-1145, Feb. 2012.

[41] C. H. Gao, Z. M. Zhou, and J. M. Chen, "Assessing the predictability for blast furnace system through nonlinear time series analysis," *Ind. Eng. Chem. Res.*, vol. 47, pp. 3037-3045, 2008.

[42] D. Martens, B. Baesens, and T. V. Gestel, "Decompositional rule extraction from support vector machines by active learning," *IEEE Trans. Knowledge and Data Eng.*, vol. 21, no. 2, pp. 178-191, Feb. 2009.

[43] J.-S. R. Jang, "ANFIS: Adaptive-Network-based Fuzzy Inference Systems," *IEEE Trans. Syst. Man Cybern.*, vol. 23, no. 3, pp. 665-685, May 1993.

[44] Y. P. Pan and J. Wang, "Model predictive control of unknown nonlinear dynamical systems based on recurrent neural networks," *IEEE Trans. Ind. Electron.*," vol. 59, no. 8, pp. 3089-3101, Aug. 2012.

[45] Y. C. Cheng and S. T. Li, "Fuzzy time series forecasting with a probabilistic smoothing hidden markov model," *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 2, pp. 291-304, Apr. 2012.

Chuanhou Gao (M'09 SM'12) received the B.Sc. degrees in Chemical Engineering from Zhejiang University of Technology, China, in 1998, and the Ph.D. degrees in Operational Research and Cybernetics from Zhejiang University, China, in 2004. From June 2004 until May 2006, he was a Postdoctor in the Department of Control Science and Engineering at Zhejiang University.

Since June 2006, he has joined the Department of Mathematics at Zhejiang University, where he is currently a Professor. He was a visiting scholar at Carnegie Mellon University from Oct. 2011 to Oct. 2012. His research interests are in the areas of data-driven modeling, control and optimization, industrial applied mathematics and process control. He is a guest editor of IEEE Transactions on Industrial Informatics, ISIJ International and Journal of Applied Mathematics, and an editor of Metallurgical Industry Automation from May 2013.

Qinghuan Ge received the B.S. degrees in Applied Mathematics from Liaocheng University, China, in 2010, and the M.S. degrees in Operational Research and Cybernetics from Zhejiang University, China, in 2012.

Since July 2012, she has joined Huawei Technologies Co., Ltd. Her research interests are in the areas of machine learning and transparency of black-box modeling techniques.

Ling Jian received the B.Sc. degrees in Applied Mathematics from Qufu Normal University, China, in 2003, M.Sc. degrees in Operational Research and Cybernetics from Zhejiang University, China, in 2006, and Ph.D. degrees in Operational Research and Cybernetics from Dalian University of Technology, China, in 2012.

Since March 2006, he has joined China University of Petroleum, where he is currently an Assistant Professor. His scientific interests includes machine learning and kernel methods.