

# Subspace Method for Identification and Control of Blast Furnace Ironmaking Process

Zeng Jiu-sun, Liu Xiang-guan, Gao Chuan-hou, Luo Shi-hua

**Abstract**—Process control of blast furnace ironmaking poses a great challenge because of its great complexity. As a main indicator of thermal state in blast furnace, silicon content in hot metal must be maintained at an appropriate level. In the present work, silicon content is taken as the output variable and the subspace identification method is used to identify the model between input and output variables. The identified model is then used for prediction so that future information of silicon content can be obtained. With the predictions of silicon content, predictive control of the ironmaking process becomes feasible. Other practical issues like dimension reduction of input variables and data preprocessing are also discussed.

## I. INTRODUCTION

**B**LAST furnace ironmaking is a critical process in the steel making industry. A blast furnace is a type of metallurgical furnace used for smelting to produce molten iron. In a blast furnace, fuel and ore are continuously dumped through the top of furnace, while air (or pure oxygen) is blown into the bottom of the chamber, so that the chemical reactions take place throughout the furnace as the material moves downward. The end products are usually molten iron and slag phases tapped from the bottom, and flue gases exiting from the top of the furnace. Although the efficiency of blast furnaces is constantly evolving, the automatic control of blast furnace ironmaking remains a problem. It is partly due to the complexity of the whole process and partly the extreme temperature. The temperature in the blast furnace is as high as 1600 °C and constantly changing, making it difficult to maintain at a stable level. As a main indicator of thermal state in blast furnace and the most important index of pig iron quality, silicon content in hot metal must be kept at an appropriate level to facilitate the production of high quality pig iron and stable running of the ironmaking process. Therefore, to better control the process, we need an accurate

predictive model whose prediction is consistent with the observed silicon content in pig iron. Much work has been done on this issue; among them are stochastic models like regressive models [1], neural network models [2], [3], [4] and deterministic models like mechanism mathematical models [5], [6]. It seems these models work quite well in some situations while not so well in other conditions. As it is common knowledge that large-scale industrial processes are influenced by all kinds of noise, models combined of stochastic and deterministic may be more robust. For this reason, the current work adopts combined deterministic-stochastic subspace identification algorithm to establish the predictive model.

## II. SUBSPACE IDENTIFICATION METHOD

System identification is the process of constructing mathematical models of dynamical system using measured input and output data. Until now, numerous time-domain and frequency-domain algorithms have been well developed and a class of identification method—subspace identification has received much attention [7]. Compared to other iterative methods like prediction error method (PEM), this is a fast and numerically reliable way [8]. During the recent 10 years, subspace identification has enjoyed great development both in theory and practice. Subspace identification methods have been applied to various industrial processes from chemical engineering to robot arm control. Three basic subspace identification methods are established, including N4SID [9], MOESP [10] and CVA [11]. The first and third algorithms have been seen more applications in practical industrial systems, and we choose the N4SID for implement in our application here.

Consider the following combined deterministic-stochastic linear time-invariant system:

$$x_{k+1} = Ax_k + Bu_k + \omega_k \quad (2.1)$$

$$y_k = Cx_k + Du_k + v_k, \quad (2.2)$$

$$\text{with } \mathbf{E}\left[\begin{pmatrix} \omega_p \\ v_p \end{pmatrix} \begin{pmatrix} \omega_q^T & v_q^T \end{pmatrix}\right] = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} \delta_{pq} \quad (2.3)$$

where  $A \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{n \times m}$ ,  $C \in \mathbb{R}^{l \times n}$ ,  $D \in \mathbb{R}^{l \times m}$ ,  $S \in \mathbb{R}^{n \times l}$ ,  $R \in \mathbb{R}^{l \times l}$ . The input vectors  $u_k \in \mathbb{R}^{m \times 1}$  and output vectors  $y_k \in \mathbb{R}^{l \times 1}$  are measured while  $\omega_k \in \mathbb{R}^{n \times 1}$  and  $v_k \in \mathbb{R}^{l \times 1}$  are unmeasurable, Gaussian distributed, zero

Manuscript received September 8, 2007. This work was supported in part by the Chinese Department of Science and Technology under Grant 2006A610032.

Zeng Jiu-sun is with the Institute of System Optimum Technique of Zhejiang University, Hangzhou, 310027 China (phone: 086-1396-8006-225; fax: 0571-87967475; E-mail: zjs1020@hotmail.com).

Liu Xiang-guan is also with the Institute of System Optimum Technique of Zhejiang University, Hangzhou, 310027 China. (E-mail: xgliu@zju.edu.cn).

Gao Chuan-hou is also with the Institute of System Optimum Technique of Zhejiang University, Hangzhou, 310027 China. (E-mail: gaochou@zju.edu.cn)

Luo Shi-hua is with the School of Information Management of Jiangxi University of Finance and Economics, Nanchang 331305, China (E-mail: superman100f@yahoo.com.cn)

mean, white noise vector sequences. Here  $E$  is the expected value and  $\delta$  is the Kronecker delta. The identification problem can then be stated as follows:

Given  $N$  input and output measurements and the fact that the two sequences are generated by the unknown combined deterministic-stochastic model (2.1)-(2.3), find the order of system and the system matrices  $A, B, C, D, Q, R, S$ .

Let the input space denote by  $U$  and the output space denoted by  $y(t)$ . Consider the orthogonal projection of  $y(t)$  onto  $U$  such that

$$y_d(t) = \tilde{E}\{y(t)|U\} \quad (2.4)$$

Where  $\tilde{E}$  denotes the orthogonal projection, then  $y_d(t)$  is called the deterministic component of  $y(t)$ . Also, the complementary projection  $y_s(t) = y(t) - \tilde{E}\{y(t)|U\}$  is called the stochastic component of  $y(t)$ .

The approach of N4SID is then used to identify the model. First a state sequence is determined from the projection of input-output data; this projection retains all the information (deterministic and stochastic) in the past that is useful to predict the future. Then the state space matrices are determined from the state sequence and with regression we then get other system matrices. A detailed description of algorithm can be found in [12].

For the convenience of prediction, we design a Kalman filter for the system to estimate the state variables if the system is observable so that the system can be rewritten as:

$$x_{k+1} = Ax_k + Bu_k + Ke_k \quad (2.5)$$

$$y_k = Cx_k + Du_k + e_k \quad (2.6)$$

Where  $K$  is the steady state Kalman gain that can be obtained from an algebraic Ricatti equation [13], innovation  $e_k$  is white noise and independent of past input and output data.

Now a state space model for the blast furnace ironmaking process has been constructed. We need to further develop it into a predictive system for online prediction. The state transition model for the  $k$ th time period to the next can be described as:

$$\begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix} = \begin{bmatrix} A & 0 \\ CA & 0 \end{bmatrix} \begin{bmatrix} x_k \\ y_k \end{bmatrix} + \begin{bmatrix} B \\ CB \end{bmatrix} u_k + \begin{bmatrix} 0 \\ D \end{bmatrix} u_{k+1} \quad (2.7)$$

With the equation, one can get the predictions of the future output from the past input variables, so that a predictive model is constructed.

### III. PRACTICAL CONSIDERING

#### A. Selection of input variables

The blast furnace of our concern has an internal volume of 2500m<sup>3</sup>, together with the supporting equipment; a complex operating process is constructed. Due to the big volume of

blast furnace and the complexity, there are numerous process variables that affect the quality and thermal state in blast furnace. We take silicon content as the output variable but selection of useful process variables needs more consideration. Inclusion of too many parameters will lead to a model of enormous size and complexity. One solution is to select variables and sample points that are well correlated with the output variable, which may reduce the dimension of the identified model effectively. However, such selection would be difficult without prior engineering knowledge and requires a complete residual analysis to test which combination of variables and sample points is the best.

An alternative is to use some dimension reduction technique. In the mean-square error sense, Principal Component Analysis (PCA) is the best linear dimension reduction technique. Through PCA, one can reduce the huge number of process variables into a few variables that represents the major variance. Instead of the using principal components (PCs) as the new variables, here we use another scheme of PCA [14]. This method uses the information in PCs to find important variables in the original dataset. As before, one first calculates the PCs then studies the scree plot to determine the number  $k$  of important variables to keep. Next, one considers the eigenvector corresponding to the smallest eigenvalue (the least important PC) and discards the variable that has the largest (absolute value) coefficient in the vector. Then one considers the eigenvector corresponding to the second smallest eigenvalue and discards the variable contributing the largest (absolute value) coefficient to that eigenvector among the variables not discarded earlier. The process is repeated until only  $k$  variables remain.

The reason for using such a scheme of PCA, on one hand, is that there is much redundant information and most variables remain relatively unchanged, on the other hand for the convenience of coming control. Among the process variables just a few ones are frequently used by operators to control the process. By performing PCA we select these important variables out and take them as the input process variables and other variables are considered as disturbances. Therefore the process can be represented by a model with disturbances (combined deterministic-stochastic model).

#### B. Appropriate level of silicon content

For the purpose of process control, it is essential to use the future information of silicon content for process control. To achieve this, we need to propose an appropriate level of silicon content so that when prediction of future silicon content is available, relative control measures can be taken. In ironmaking process, the main control objectives are larger output, lower energy consumption and better quality. Thus it is important to estimate at which level of silicon content can these objectives be achieved, e.g., the range of silicon content that gives the largest output and consumes the least energy. We then take the measured temperature of hot metal ( $\mu$ ) as an instrument variable to get the range of silicon content ( $Si$ ) with the largest hourly output of hot metal ( $\theta$ ). After removing outliers,  $Si$  is then divided into several intervals

and so is  $\mu$ . So some combinations (called states here) are formed. We then calculate the number of datasets that fall into these states and the mean value of  $\theta$  in each state. The number of datasets here is 500. Datasets with silicon content more than 0.9 are deleted since those datasets are regarded to be measured when production conditions are poor. Fig. 1 gives the number of datasets that fall into each state. It can be seen from Fig. 1 that most datasets fall into 9 states, e.g., the states when  $Si$  is between 0.4 and 0.7 and  $\mu$  is between 1500 and 1530.

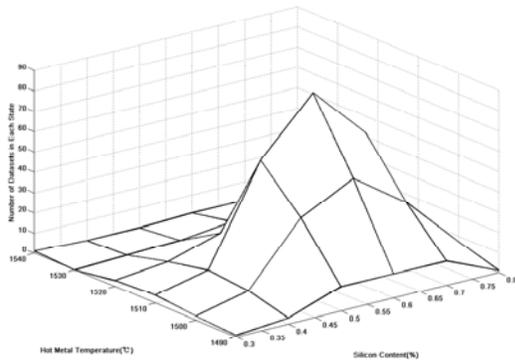


Fig. 1. Number of datasets that fall into each state.

From Fig. 2 it can be seen that there is little difference on average hourly output between each state. Bearing in mind

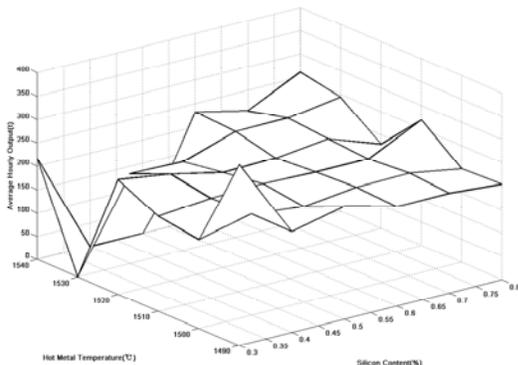


Fig. 2. Average hourly output of hot metal for different combinations of silicon content and hot metal temperature

the fact that the ironmaking process consumes less energy when silicon content and hot metal temperature are lower we can get the best state of production. Here the best state of production is when  $Si$  is between 0.4-0.5 and  $\mu$  is between 1500 and 1510. In this state, good output together with less energy consumption is achieved. In practical production, keeping the silicon content in the selected state becomes one of our control objectives.

### C. Predictive control method

In blast furnace ironmaking, a sustainable process with high output, low cost and good quality is most crucial for production. The ideal situation is just to take small

adjustments of input variables and keep silicon content at an appropriate level- the level that gives higher output and consume less energy. It can be expressed as the following equations:

$$\int_{t_0}^t F(\Delta u_1, \Delta u_2, \dots, \Delta u_n) d\tau = y(t) \in [y_0, y_1] \quad (3.1)$$

$$\min_{\Delta u_1, \dots, \Delta u_n} \left\{ \sum_{k=1}^N (\tilde{y}_k - y_k)^T \Lambda_y (\tilde{y}_k - y_k) + \sum_{k=1}^N \Delta u_k \Lambda_u \Delta u_k \right\} \quad (3.2)$$

where  $\Delta$  is the difference operator and  $\Delta u_k$  is the incremental change of the input variable  $u_j$ .  $y$  is the output variable (here  $y$  is silicon content),  $\tilde{y}_k$  is the predicted value of  $y_k$ .  $\Lambda_y$  and  $\Lambda_u$  are chosen to assign relative importance to the elements in  $y_k$  and  $\Delta u_k$  respectively. For example, quality control may be given a higher emphasis by giving a higher weighting of  $\Lambda_y$ . In practice, operators are apt to adjust a few of the input variables such as temperature of blast, quantity of blast and so on, which makes the operation relatively convenient.

## IV. SIMULATION RESULTS

Time series data used here is collected from blast furnace No. 1 in Baotou Iron and Steel Corporation with the volume of 2500m<sup>3</sup>. The data, with the size of 500 datasets, is divided into training set and validation set. The training set (with 400 datasets) is used to identify the model and the validation set (with 100 datasets) is used to assess the performance of the model.

Through PCA we select out 7 process variables from 35

TABLE I  
SELECTED INPUT VARIABLES

No	Variable	Mean	Standard Deviation
1	quantity of blast	3538.8	72.2
2	temperature of blast	1095.8	18.4
3	pressure of blast	284.8	7.3
4	gas permeability	3155.7	199.3
5	quantity of coal powder	14.2	4.4
6	percent of CO <sub>2</sub> in top gas	27.6	6.0
7	pressure of top gas	175.7	41.2

variables as input variables of the subspace model. Table 1 describes the 7 input variables from No. 6 blast furnace in

Baotou Iron and Steel Corporation in China. As is shown in table 1, the input variables have widely varying mean value and standard deviation, it is necessary to tailor the data to make the calculation easier by ensuring zero mean and unit standard deviation for each input variable. In the present work all the variables have been scaled prior to model

building by the following relation:  $\tilde{x}_n = \frac{x_n - \mu(x)}{\sigma(x)}$ , where

$\mu(x)$  and  $\sigma(x)$  are mean and standard deviation of all data for the variable  $x$ .

After the preprocessing procedures, the subspace framework is then used to identify the model. Fig. 3 gives the

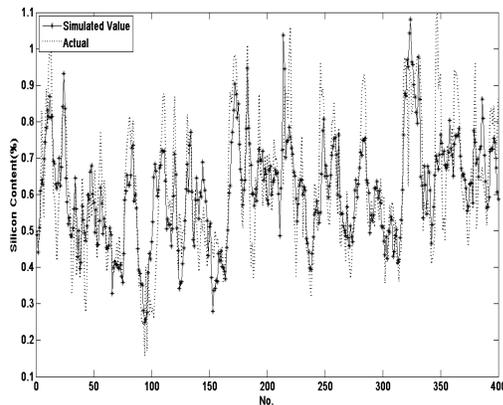


Fig. 3. Simulation result with comparison to actual value of silicon content.

comparison of simulated and actual value of silicon content. With identified model, we can make predictions of future silicon content. The validation set is used to test the performance of the identified model. The results of prediction are shown in Fig. 4.

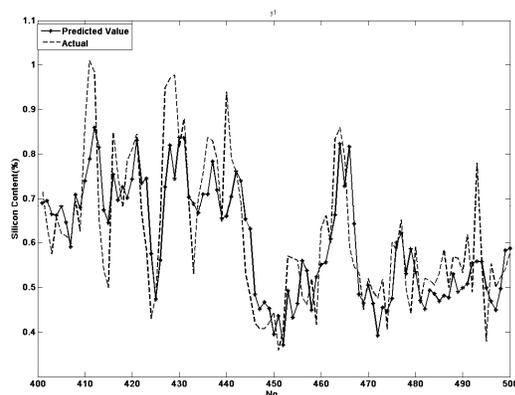


Fig. 4. Prediction result with comparison to actual value of silicon content.

It can be seen from Fig. 4 that predictions are in good agreement with actual values. With the accurate prediction of future information of silicon content, we can take measures in advance to keep the production smooth. For example, if our prediction of silicon content is out of the appropriate level we

obtained in 3.2, we can take adjustments of input variables according to the predictive control method in 3.3.

## V. CONCLUSION

In this work, the subspace identification method is used to construct predictive model for ironmaking process. By selecting out the key input variables using a new scheme of PCA, an easier predictive control method is proposed. The method is illustrated by first obtaining the appropriate level of silicon content and with the desired level of silicon content as one of the control objectives, we propose an optimization-based predictive control method that is easy to implement. Empirical results show that the model has achieved good predictive performance, it is reasonable to believe a smooth production can be realized. For operators of blast furnace, closed-loop operation is a long-term target; further work should be done on the development of closed-loop identification method for the ironmaking process.

## ACKNOWLEDGMENT

The authors thank the Baotou Steel Corporation in China for providing field data and expert knowledge, and special thanks are paid to Yang Ke-ping with Annpop Analytic Software Corporation for suggestions and encouragement. The 3<sup>rd</sup> author would like to thank sponsor and financial support from Zhejiang Provincial Natural Science Foundation of China under Grant No.Y107110, Research Fund for the Doctoral Program of Higher Education of China (for new teachers) under Grant No.20070335161. The 4<sup>th</sup> author would like to thank sponsor and financial support from the Office of Education, Jiangxi Province under Grant No. GJJ08358.

## REFERENCES

- [1] Matias Waller and Henrik Saxon, "On the development of predictive models with application to a metallurgical process," *Industrial & Engineering Chemistry Research*, vol. 39, no. 4, pp. 982-988, Apr. 2000.
- [2] Jian Chen, "A predictive system for blast furnaces by integrating a neural network with qualitative analysis," *Engineering Application of Artificial Intelligence*, vol 14, no. 1, pp. 77-85, Jan. 2001.
- [3] Xiaojing Hao, Peng Zheng and Zhi Xie, "A predictive modeling for blast furnace by integrating neural network with partial least regression," in *Conf. Rec. 2004 IEEE Int Conf. Industrial Technology*, pp. 1162-1167.
- [4] Juan Jimenez, Javier Mochon, Jesus Sainz de Ayala and Faustino Obeso, "Blast furnace hot metal temperature prediction through networks-based models," *ISIJ International*, vol 44, no. 3, pp. 573-580, Mar. 2004.
- [5] Xiang-guan Liu, Jiu-sun Zeng and Min Zhao, "Mathematical model and its hybrid dynamic mechanism in intelligent control of ironmaking," *Journal of Iron and Steel Research International*, vol 14, no. 1, pp. 7-11, Jan. 2007.
- [6] Jose Adilson de Castro, Hiroshi Nogami and Jun-ichiro Yagi, "Three dimensional multiphase mathematical modeling of the blast furnace based on the multifluid model," *ISIJ International*, vol 42, no. 11, pp. 44-52, Nov. 2002.
- [7] S. Joe Qin, "An overview of subspace identification," *Computers and Chemical Engineering*, vol. 30, no. 10-12, pp. 1502-1513, Sep. 2006.

- [8] Wouter Favoreel, Bart De Moor and Peter Van Overschee, "Subspace state space system identification for industrial process," *Journal of Process Control*, vol. 10, no. 2-3, pp. 149-155, Apr. 2000.
- [9] P. Van Overschee, B. De Moor, "N4SID: subspace algorithms for the identification of combined deterministic and stochastic systems," *Automatica*, vol. 30, no. 1, pp. 75-93, Jan. 1994.
- [10] M. Verhaegen, P. Dewilde, "Subspace identification, part 1: the output-error state space model identification class of algorithms," *International Journal of Control*, vol 56, no.5, pp. 1187-1210, May 1992.
- [11] W.E. Larimore, "Canonical variate analysis in identification, filtering and adaptive control," in *Proceeding of the 29<sup>th</sup> Conference on Decision and Control*, pp.596-604.
- [12] M. Jansson, B. Wahlberg, "A linear regression approach to state-space subspace system identification," *Signal Processing*, vol. 52, no. 2, pp. 103-129, Jul. 1996.
- [13] Tohur Katayama, *Subspace Methods for System Identification*. London: Springer-Verlag, 2007, ch. 5.
- [14] I. K. Fodor, "A survey of dimension reduction techniques," UCRL-ID-148494, LLNL, 2002.