Application of PCA Based Process Monitoring Method to Ironmaking Process*

Tongshuai Zhang, Hao Ye
Tsinghua National Laboratory for Information Science and Technology (TNList), Department of Automation
Tsinghua University
Beijing, China
zts11@mails.tsinghua.du.cn, haoye@tsinghua.edu.cn

Wei Wang
School of Automation Science and Electrical Engineering
Beihang University
Beijing, China
w.wang@buaa.edu.cn

***Abstract*—It is quite challenging to monitor an ironmaking process because of its special characteristics such as frequent fluctuations and lack of direct measurements. To tackle these issues, a two-stage PCA based monitoring method was proposed in our previous work. However, only one type of operating anomaly was considered and the historical data of one accident was utilized. To further evaluate the performance of the two-stage PCA based method, four different anomaly types and 25 corresponding historical datasets collected from three real blast furnaces are tested in this paper. The results demonstrate good potential of our proposed method for anomaly detection in ironmaking process.**

Keywords—principal component analysis*; *process monitoring; fault detection; blast furnace; ironmaking process

I. Introduction

Ironmaking processes, which produce molten iron through complex physical and chemical reactions, are an important part of modern iron and steel industry [1]. It is known that one of the main reactors in the ironmaking process is the blast furnace. As shown in Fig. 1 [1], the inputs of a blast furnace can normally be divided into two parts. The first part containing iron-bearing materials (such as iron ores, sinters, and pellets), cokes and flux are dumped into the blast furnace from its top. The second part containing hot dry air, enriching oxygen, moisture, fuels (such as tar or pulverized coal) are blasted into the furnace from the bottom. Similarly, the outputs containing liquid molten iron and slag flows out of the furnace from the bottom, whereas the coal gas is collected from the top of the furnace. Therefore, with one stream downward and the other upward, complex reactions take place in the blast furnace [2, 3]. Apparently, the prompt detection of potential anomaly is vitally important to keep a blast furnace working normally and steadily. Nevertheless, monitoring the status of a blast furnace is a rather difficult task due to the lack of its accurate mathematical model and direct measurements of its internal states.

Most of the early process monitoring methods applied to blast furnaces are based on expert systems [4, 5, 6, 7]. To achieve desired performance, comprehensive rules and a priori knowledge of ironmaking process are required by these methods. In recent years, several data-driven fault diagnosis methods including Support Vector Machines (SVMs) [8, 9, 10, 11], neural network [12, 13] and state space model assisted approaches [14] have been presented. It is worth noting that sufficient historical data in faulty cases are necessary in most of these methods, but usually they are unavailable in a real ironmaking process.

In contrast to the aforementioned methods, the applications of multivariate statistical process control (MSPC) based monitoring methods to ironmaking processes are very limited, although they have been widely used in the chemical processes [15, 16]. Moreover, most of the currently available works mainly focused on the prediction of silicon content [17, 18, 19, 20] or process modelling problems [21, 22, 23].

[24] and [25] are the representative results on ironmaking process monitoring based on principal component analysis (PCA). Motivated by these two works and the successful applications of PCA to other process industries, we proposed a two-stage PCA based method for ironmaking process monitoring in [26] where the strong disturbances caused by hot blast stoves switching can be dealt with. Note that such disturbances were not considered in [24] and [25].

* The work was supported by National Natural Science Foundation of China under grant 61290324 and 61203068.
Hao Ye is corresponding author (haoye@tsinghua.edu.cn)

Since there was only one type of operating anomaly with its corresponding historical data that was utilized to validate the effectiveness of the proposed two-stage PCA based method in [26], four different anomaly types and 25 corresponding historical data sets collected from three real blast furnaces are tested in this paper. To the best of our knowledge, this is the first work which employs so large scale field data for the validation of PCA based monitoring methods in ironmaking process.

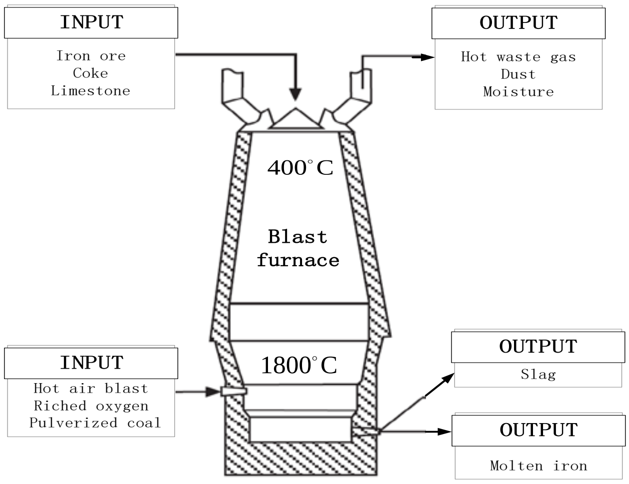

Figure 1. The inputs and outputs of a blast furnace

II. A Brief Introduction to Two-Stage PCA Based Monitoring Method for Ironmaking Process

A. PCA Based Monitoring

Let X denote a data matrix for L observations of p variables, then X can be decomposed as [27],

$$X = TP^T = \sum_{i=1}^{p} t_i p_i^T = \sum_{i=1}^{s} t_i p_i^T + E \tag{1}$$

where t_i, p_i, ($i = 1, \cdots, p$), and E denote the score vector, the loading vector and the residual matrix, respectively. s can be determined by

$$\sum_{i=1}^{s} \sigma_i^2 / \sum_{i=1}^{p} \sigma_i^2 \times 100\% > 95\% \tag{2}$$

where σ_i is the i^{th} largest singular value of X.

Hotelling's T^2 statistic and SPE statistic have been widely used by PCA based monitoring methods, which are respectively defined as [15]

$$T_k^2 = x_k^T P_s \Sigma_s^{-2} P_s^T x_k \tag{3}$$

and

$$SPE_k = r_k^T r_k, \quad \text{with} \quad r_k = x_k - \hat{x}_k, \ \hat{x}_k = P_s P_s^T x_k \tag{4}$$

where x_k denotes a new-coming observation at sampling instant k, P_s is a matrix composed of the first s columns of P (i.e. p_i for $1 \le i \le s$), Σ_s is a diagonal matrix composed of the s largest singular values of X. For a given confidence level α, the corresponding thresholds for T^2 and SPE statistics can be calculated. Please refer to [27] for more details.

B. A Brief Introduction to Two-stage PCA Based Monitoring Method for Ironmaking Process

Blast stoves switching causes severe disturbances, which will affect both the modelling and the monitoring procedure if a standard PCA based method is applied to ironmaking process monitoring. To deal with this issue, a two-sage PCA based method was proposed in [26].

Considering the following two facts, i.e.

No direct measurements are available to indicate the stoves switching;

T^2 statistic is sensitive to the switching disturbances because it reflects the systematic variations of a process.

The first-stage PCA designed in [26] is aimed to achieve effective identification, location and removal of the switching disturbances based on a hypothesis testing framework for T^2 statistic. If the T^2 statistic exceeds the threshold determined by current variation level under hypothesis testing framework, its corresponding observation would be identified as a sample of a peak-like disturbance. And then the start and end of this disturbance will also be identified. And all the identified disturbances will be removed to construct a new training set without disturbances.

Similar to the traditional PCA based monitoring method [15], the second-stage PCA in [26] is responsible for offline process modelling and online monitoring. Additionally, a distribution of the widths of disturbances identified in the first stage can be generated, according to which a time-delay windows is used to distinguish disturbances from anomalies. Therefore, false alarms can be greatly reduced at a cost of delay in alarm generation. And such a delay lasts for only several minutes which is negligible with comparison to the long settling time of the ironmaking process. Note that as opposed to [26], where 'AND' logic of T^2 statistic and SPE statistic for alarm generation was adopted for monitoring, 'OR' logic is executed instead in this paper, which means as long as one of the two statistics goes beyond the controlling limit, the online monitoring model alarms. The reason for such a change is that T^2 statistic and SPE statistic are sensitive to different types of anomalies. Therefore, 'OR' logic is more appropriate to assure the monitoring sensitivity.

In short, the first stage PCA is aimed to construct a new 'clean' training dataset with the stove switching disturbances identified and removed according to the variation of T^2 statistic. The second stage PCA is aimed to build up a process model for monitoring with an alarm logic different from traditional PCA based methods due to its additional procedure to distinguish the switching disturbances and anomalies. Interested readers may refer to [26] for detailed algorithm.

III. CASE STUDY

C. Description of the Dataset

In this paper, the dataset to be tested were collected from No. 2, No. 3 and No. 5 blast furnaces, whose volumes are $2650m^3$, $2000m^3$ and $1500m^3$ respectively, in Liuzhou Iron&Steel Co. Ltd.

Table I shows the main monitored variables in the ironmaking process. Note that though most variables are common in three blast furnaces. Some variables are only available in certain furnace(s). Please refer to the footnotes of Table I for detail.

TABLE I. VARIABLES LIST OF THE DATASET

No.	Variable	No.	Variable
1	Oxygen enrichment rate (%)	18	Cold blast pressure(1) (MPa)
2	Blast furnace permeability index	19	Cold blast pressure(2)[2] (MPa)
3	CO volume[1] (%)	20	Total pressure drop (kPa)
4	H_2 volume[1] (%)	21	Hot blast pressure(1) (MPa)
5	CO_2 volume[1] (%)	22	Hot blast pressure(2)[2] (MPa)
6	Blast velocity at tuyere of blast furnace (m/s)	23	Actual blast velocity (m/s)
7	Enriching oxygen flow (m^3/h)	24	Cold blast temperature (°C)
8	Cold blast flow (10^4 m^3/h)	25	Hot blast temperature (°C)
9	Blast momentum (KJ)	26	Top temperature (1) (°C)
10	Blast furnace bosh gas volume (m^3)	27	Top temperature (2) (°C)
11	BF bosh gas index	28	Top temperature (3) (°C)
12	Theoretical combustion temperature (°C)	29	Top temperature (4) (°C)
13	Blast furnace top gas pressure(1) (kPa)	30	Downcomer temperature[2] (°C)
14	Blast furnace top gas pressure (2) (kPa)	31	Drag coefficient
15	Blast furnace top gas pressure (3) (kPa)	32	Coal injection set value (t/h)
16	Blast furnace top gas pressure (4)[2] (kPa)	33	Actual coal injection rate (t/h)
17	Enriching oxygen pres-sure (MPa)	34	Actual coal injection in last hour (t)

1. Variables only available in No. 3 and No. 5 blast furnaces.

2. Variables only available in No.2 blast furnace.

Hanging, slipping and channeling are the three most common anomalies in ironmaking process. In addition, cold hot furnace condition is also an important concern of blast furnace operators. From 2012 to 2014, there are totally 25 anomalies of No. 2, 3 and 5 blast furnaces confirmed and recorded by the operators, in which all the aforementioned four anomaly types are included. Table II provides the number of times for different types of anomalies occurred in each blast furnace.

TABLE II. SUMMATION OF THE ANOMALIES

Anomalies	Blast Furnace		
	No. 2	***No. 3***	***No. 5***
Hanging	11	5	3
Channeling	1	2	1
Slipping	0	1	0
Cold furnace condition	0	1	0

D. Experimental Results

Due to the page limit, the detailed results after each step of the proposed two-stage PCA based method will not be given. Instead, the final results are summarized in Table III, where the second column shows the blast furnace in which the anomalies occurred, the third column lists the type of each anomaly. The time intervals of the training dataset and testing dataset are provided in the fourth and fifth columns, respectively. The sixth and seventh columns present the statistic(s) using which correct fault alarms are generated ahead of the operators' recorded time, and the corresponding lead time. Note that the sampling interval is 10 seconds, the thresholds of T^2 statistic and SPE statistic are calculated under the confidence level $\alpha = 1\times10^{-4}$.

As shown in Table III, most of the anomalies are detected ahead of the operators' records. More precisely, among the 25 anomalies, PCA method is not applicable to only No. 9 anomaly. The main reason is that this type of anomaly only occurred in the reblowing process which is a transient process. Moreover, 20 of the rest cases can be successfully detected ahead of the operator's records. Thus the correct anomaly detection rate is 83.3% (=20/24).

As shown in the last column of Table III, the leading time lies in two main intervals: i.e. 1~2 hours and 2~3 hours. As the abnormalities in ironmaking process will evolve gradually, detecting the abnormalities in early stage can assist the operators to adjust the working status of the blast furnace in time to avoid further deterioration.

In addition, it can be seen that the T^2 and SPE statistics have different sensitivities to different anomalies. All of the 19 hanging anomalies are detected by T^2 statistic, but only 10 of them are detected by SPE statistic. All the four channeling and one slipping anomalies are detected by SPE statistic, but only two of them are detected by T^2 statistic. Therefore, the different sensitivities may provide valuable information for the classification of different anomalies.

For better understanding, the detailed testing results for No. 1, 14, 19 and 9 anomalies are plotted in Fig. 2~5, respectively, where subfigures (a), (b), and (c) show the original data, the T^2 and SPE statistics, respectively. It is worth noting that both the training data set and the testing data set are adopted to test the online monitoring algorithm of the second-stage PCA.

In addition, instead of choosing a common training set, we choose the normal data not too far before the faulty data as training set. The reason to do so is that the operating point of the ironmaking process often changes, which makes it inappropriate to use one single training set for detecting all anomalies that may corresponds to different operating points of one certain blast furnace. Though it seems impossible to know when a fault will happen in practice, a reasonable update strategy can assure that the monitoring model stays up-to-date.

In Fig. 2, the first 40000 samples constitute the training dataset and the rest samples constitute the testing dataset. Subfigure (a) shows that all the monitored variables with their original amplitudes. Thus certain variables appear overlapped at the bottom as their amplitudes are relatively small. The blue, green and red curves in subfigures (b) and (c) represent normal status, disturbances and abnormal status judged by the monitoring algorithm respectively. The grey dashed lines in subfigures (b) and (c) represents the thresholds of T^2 and SPE statistics. The T^2 alarmed at about the 55580th instant, while the operators found the hanging anomaly and performed recovery actions at 63015th instant and the SPE statistic alarmed at 63470th instant. Therefore, the T^2 statistic alarmed 7435 instants (approximately 20.7 hours) earlier than the operators' monitoring, whereas SPE statistic fails to generate alarm earlier than operators' monitoring. Similarly, Fig. 3~5 show that the T^2 or SPE statistic can detect the anomalies in advance effectively.

It should be mentioned that the reason our method cannot detect the hanging anomaly in Fig. 5 is that the anomaly occurred during the process of reblowing, i.e. a transient process. Thus the static PCA based method is not applicable for such cases. However, as shown in Fig. 5, it is noticed that both SPE and T^2 statistics are first downward and then upward. We believe that the trend may reflect the status of the blast furnace, so the trend how the statistics change may also be useful. Moreover, our proposed method is not desired for detecting anomalies occurred shortly after another anomaly, which is due to the changing of the operating point or the dynamic transient process.

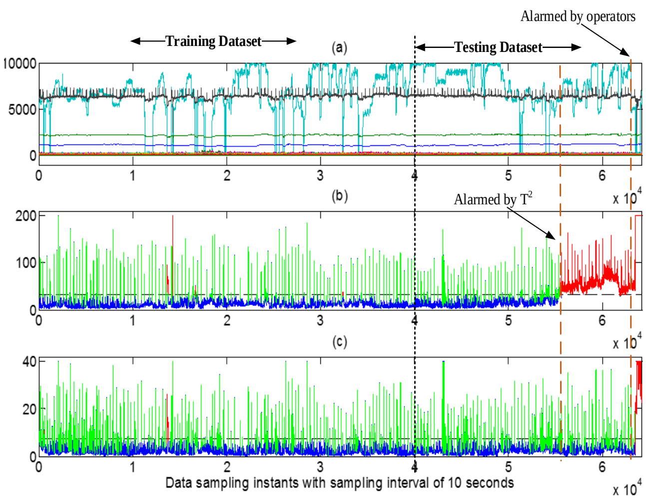

Figure 2. The first hanging of No. 2 blast furnace on December 28th, 2013

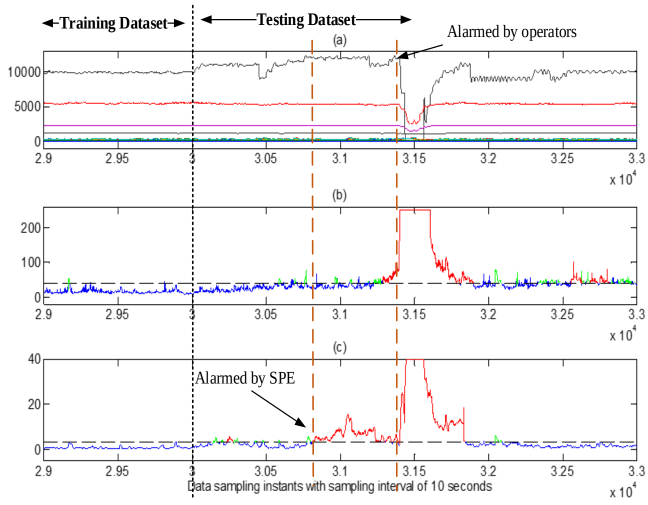

Figure 3. Slipping of No. 3 blast furnace on March 25th, 2012

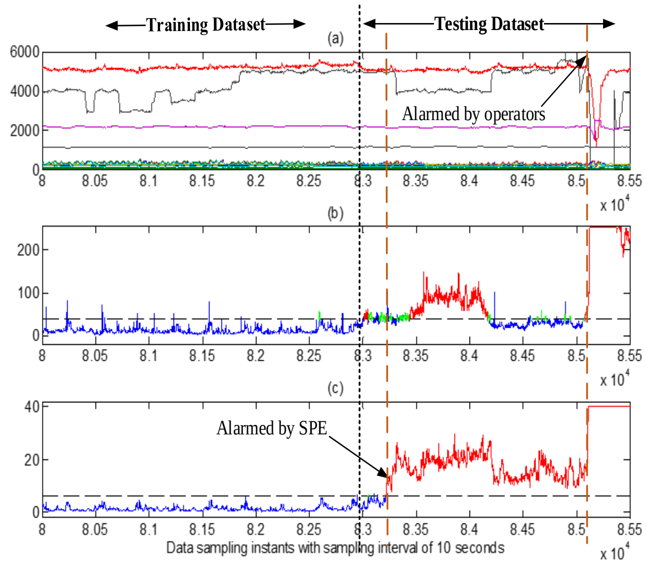

Figure 4. Channeling of No. 3 blast furnace on February 13th, 2013

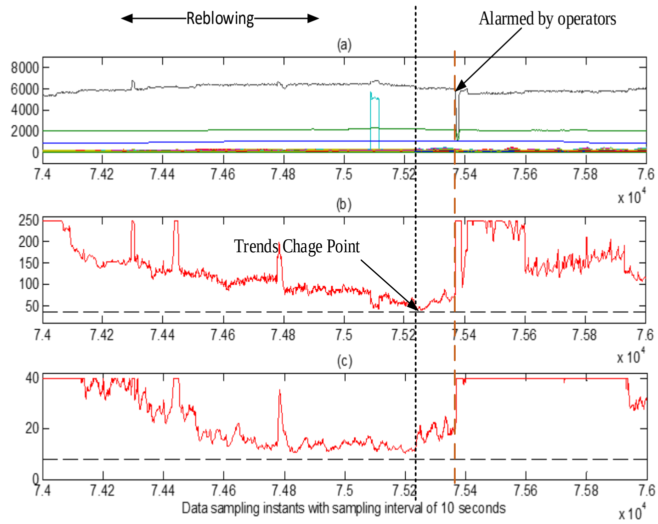

Figure 5. Hanging after reblowing of No. 2 blast furnace on January 21st, 2014

TABLE III. SUMMARY OF SIMULATION RESULTS

Fault No.	Blast Furnace	Anomaly Type	Training Dataset	Testing Dataset	Alarm Statistics	Lead Time
1	No. 2	Hanging	12/21 00:20~12/25 15:46	12/25 15:46 ~12:28 09:40	T^2	20.7 hours
2		Hanging	12/28 13:18~12/28 17:52	12/28 17:52~12/29 00:26	T^2	3.1 hours
3		Channeling	12/30 18:10~12/30 23:44	12/30 23:44~12/31 05:18	SPE	3 hours
4		Hanging	01/01 00:00~01/04 11:27	01/04 11:27~01/05 20:50	T^2 & SPE	1.6 hours
5		Hanging	01/01 00:00~01/04 11:27	01/07 22:54~01/08 04:28	T^2 & SPE	Not advance
6		Hanging	01/17 20:09~01/18 07:17	01/18 07:17~01/18 21:13	T^2 & SPE	3.6 hours
7		Hanging	01/19 16:42~01/19 22:16	01/19 22:16~01/20 09:23	T^2	1.1 hours
8		Hanging	01/19 16:42~01/19 22:16	01/19 22:16~01/20 09:23	T^2	3.0 hours
9		Hanging	01/12 12:21~01/15 06:56	01/21 02:05~01/21 07:39	T^2 & SPE	Not applicable
10		Hanging	01/21 13:12~01/21 18:46	01/21 18:46~01/22 05:54	T^2	7.5 hours
11		Hanging	01/22 06:00~01/22 17:08	01/22 17:08~01/22 22:36	T^2 & SPE	Not advance
12		Hanging	01/24 13:32~01/24 19:06	01/24 19:06~01/25 06:14	T^2	3.4 hours
13	No. 3	Hanging	03/21 20:49~03/22 16:25	03/22 16:26~03/23 13:25	T^2 & SPE	9.3 hours
14		Slipping	03/21 20:49~03/22 16:25	03/25 08:51~03/25 17:15	SPE	1.6 hours
15		Channeling	03/29 08:07~03/30 03:43	03/30 03:43~03/31 10:32	T^2 & SPE	Not advance
16		Hanging	11/12 23:33~11/14 17:32	11/14 17:32~11/16 11:27	T^2	19.6 hours
17		Hanging	01/14 07:24~01/15 14:07	01/15 14:07~01/16 04:05	T^2	1.4 hours
18		Hanging	01/22 05:40~01/25 00:37	01/25 00:37~01/25 09:00	T^2	1.3 hours
19		Channeling	02/12 23:23~02/13 08:22	02/13 08:22~02/13 15:50	SPE	4.9 hours
20		Cold furnace condition	02/15 00:00~02/19 23:57	02/19 23:57~02/26 03:11	T^2 & SPE	1.3 hours
21		Hanging	03/02 09:24~03/06 03:43	03/06 03:43~03/06 14:17	T^2 & SPE	1.5 hours
22	No. 5	Hanging	05/11 00:00~05/12 18:03	05/12 18:03~05/13 19:17	T^2 & SPE	9.4 hours
23		Channeling	06/06 17:06~06/08 02:51	06/08 02:51~06/09 12:30	T^2 & SPE	36 minutes
24		Hanging	10/13 06:53~ 10/13 18:35	10/13 18:35~10/14 01:33	T^2 & SPE	Not advance
25		Hanging	11/12 10:50~11/12 16:24	11/12 16:24~1/12 23:22	T^2 & SPE	3.3 hours

IV. CONCLUSION

In this paper, 25 anomalies corresponding to four different anomaly types of ironmaking process which were collected from three real blast furnaces, are utilized to test the performance of our proposed two-stage PCA based method for ironmaking process monitoring. The results show a good potential of the PCA based methods for anomaly detection in ironmaking process. Furthermore, some interesting topics remain open which could be considered as future works. For example,

- It is important to investigate the diagnostic and classification schemes of the operating anomalies, combining the fault information provided by the two-stage PCA based monitoring method.
- As shown in Table III, the training set is selected not too far before the anomaly to achieve desired performance. Thus a reasonable update strategy needs to be developed for the monitoring model adapted to the change of the process.
- The training sets should not contain abnormal data and ensure that the monitoring model is sufficiently excited. However, such requirements cannot be satisfied in some cases. Thus it is necessary to study how to classify and cluster the process data corresponding to different operating points.

REFERENCES

[1] J.G. Peacey and W.G. Davenport, The Iron Blast Furnace: Theory and Practice: Pergamon Press, 1979.

[2] R.J. Fruehan and A.S. Foundation, The Making, Shaping, and Treating of Steel: Ironmaking volume: AISE Steel Foundation, 1999.

[3] M. Geerdes, H. Toxopeus, C. van der Vliet, R. Chaigneau, and T. Vander, Modern Blast Furnace Ironmaking: An Introduction: IOS Press, 2009.

[4] S. Amano, T. Takarabe, T. Nakamori, H. Oda, M. Taira, S. Watanabe, et al., "Expert system for blast furnace operation at Kimitsu works," ISIJ Int., vol. 30, pp. 105-110, 1990.

[5] J. Dorn, "Guest Editor's Introduction: Expert Systems in the Steel Industry," IEEE Intell. Syst., vol. 11, pp. 18-21, 1996.

[6] H. Lu, B. Gao, L. Zhao, H. Guo, and T. Yang, "Neural Network Expert system of forecasting blast furnace operational conditions," J. Univ. Sci. Technol. B., vol. 3, p. 013, 2002.

[7] A. Klinger, T. Kronberger, M. Schaler, B. Schürz, and K. Stohl, Expert systems controlling the iron making process in closed loop operation: INTECH Open Access Publisher, 2010.

[8] A. Wang, L. Zhang, and N. Gao, "Fault diagnosis of blast furnace based on DAGSVM," in Intelligent Systems, 2006 3rd International IEEE Conference on. IEEE, 2006, pp. 571-574.

[9] A. Wang, L. Zhang, N. Gao, and M. Wang, "Fault diagnosis of blast furnace based on SVMs," in Intelligent Control and Automation vol. 2, 2006. The sixth World Congress on. IEEE, 2006, pp. 5615-5618.

[10] L. M. Liu, A. N. Wang, M. Sha, and F. Y. Zhao, "Multi-class classification methods of cost-conscious LS-SVM for fault diagnosis of blast furnace," J. Iron Steel Res. Int., vol. 18, pp. 17-33, 2011.

[11] L. M. Liu, A. N. Wang, M. Sha, X. Y. Sun, and Y. L. Li, "Optional SVM for fault diagnosis of blast furnace with imbalanced data," ISIJ Int., vol. 51, pp. 1474-1479, 2011.

[12] J. Chen, "A predictive system for blast furnaces by integrating a neural network with qualitative analysis," Eng. Appl. Artif. Intel., vol. 14, pp. 77-85, 2001.

[13] L. Pan and X. Hui, "Blast furnace fault diagnosis based on nonlinear fuzzy neural identification," Adv. Mat. Res., vol. 204, pp. 1254-1257, 2011.

[14] J. Cai, J. Zeng, and S. Luo, "A State space model for monitoring of the dynamic blast furnace system," ISIJ Int., vol. 52, pp. 2194-2199, 2012.

[15] J.F. Macgregor and T. Kourti, "Statistical process control of multivariate processes," Control Eng. Pract., vol. 3, pp. 403-414, March 1995.

[16] S.Joe Qin, "Statistical process monitoring: basics and beyond," J. Chemometr., vol. 17, pp. 480-502, 2003.

[17] V.R. Radhakrishnan and A.R. Mohamed, "Neural networks for the identification and control of blast furnace hot metal quality," J. Process Contr., vol. 10, pp. 509-524, 2000.

[18] T. Bhattacharya, "Prediction of silicon content in blast furnace hot metal using partial least squares (PLS)," ISIJ Int., vol. 45, pp. 1943-1945, 2005.

[19] W. Wang, "Application of BP network and principal component analysis to forecasting the silicon content in blast furnace hot metal," in Intelligent Information Technology Application, 2008. IITA'08. Second International Symposium on. 2008, pp. 42-46.

[20] W. Wang and J. Ma, "Principal component regression approach for forecasting silicon content in hot metal," in Natural Computation, vol. 2, 2009. Fifth International Conference on. IEEE, 2009, pp. 590-593.

[21] X. Liu and W. Wang, "Using multivariate grey model and principal component analysis to modeling the blast furnace," in Multimedia Technology, 2011 International Conference on. IEEE, 2011, pp. 2789-2792.

[22] J. Zeng, X. Liu, C. Gao, and S. Luo, "Subspace method for identification and control of blast furnace ironmaking process," in American Control Conference, 2008. IEEE, 2008, pp. 2481-2485.

[23] D. Noskievičová, "Statistical analysis of the blast furnace process output parameter using arima control chart with proposed methodology of control limits setting," Metalurgija, vol. 48, pp. 281-284, 2009.

[24] F.I. Gamero, J. Colomer, J. Meléndez, and P. Warren, "Predicting aerodynamic instabilities in a blast furnace," Eng. Appl. Artif. Intel., vol. 19, pp. 103-111, 2006.

[25] E. Vanhatalo, "Multivariate process monitoring of an experimental blast furnace," Qual. Reliab. Eng. Int., vol. 26, pp. 495-508, 2009.

[26] T. Zhang, H. Ye, W. Wang, and H. Zhang, "Fault diagnosis for blast furnace ironmaking process based on two-stage PCA," ISIJ Int., vol. 54, pp. 2334-2341, 2014.

[27] I. Jolliffe, Principal Component Analysis: Springer Science & Business Media, 2002.