Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Data-Driven Fault Diagnosis of Shaft Furnace Roasting Processes Using Reconstruction and Reconstruction-Based Contribution Approaches

Xinglong Lu*, Qiang Liu*, Tianyou Chai* and S. Joe Qin*, **

** State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University Shenyang, China (e-mail: tychai@mail.neu.edu.cn).*
*** Dept. of Chemical Engineering and Materials Science, University of Southern California Los Angeles, CA 90089 USA and The Chinese University of Hong Kong, Shenzhen 2001 Longxiang Blvd., Longgang, Shenzhen, China(e-mail: sqin@usc.edu).*

Abstract: The process faults of shaft furnace roasting processes, e.g. fire-emitting, flame-out, under-reduction, and over-reduction are undesirable for stable operation of the processes. The processes share multiple complexities such as multi-variate and strong correlations, which make it difficult to diagnose the faults using model-based or knowledge-based methods. In this paper, a data-driven fault diagnosis method for shaft furnace roasting processes is presented based on reconstruction and reconstruction-based contribution. The proposed method exploits historical faulty data to derive fault directions to identify ongoing faults with the help of additional explanation from contribution plots. A case study on a simulation system of shaft furnace roasting processes illustrates the effectiveness of the proposed method.

1. INTRODUCTION

Shaft furnace roasting processes are important chemical reduction processes that transform weakly magnetic ore into strongly magnetic one in a high temperature condition. Stable operation of the processes is essential to safety and product quality. But process faults usually occur when there is a mismatched action which cannot meet frequent changes of magnitude, class, and ingredients of the raw ore. Typical process faults are fire-emitting, flame-out, under-reduction and over-reduction. While fire-emitting and flame-out would bring up hazard of equipment damage, under-reduction and over-reduction would degrade ore concentrate to have dissa-tisfactory products. It is necessary to detect and diagnose the process faults in time.

The operating conditions of shaft furnace roasting processes change frequently with complex mechanism. On one hand, it is difficult to model the processes precisely (Wu *et al.*, 2006). On the other hand, process faults are more complex than those caused by sensor or actor failure. Consequently model-based fault diagnosis methods (Zhou & Hu, 2009) are not advisable for diagnosis of shaft furnace roasting processes.

To solve this problem, researchers proposed knowledge-based fault diagnosis methods with artificial intelligence for shaft furnace roasting processes. Yan *et al*. (2008) proposed an intelligent fault prediction system for shaft furnace with case-based reasoning technique by matching an ongoing fault with historical ones in a fault case to get a diagnosis result in the form of probability. Wu *et al.* (2006) and Chai *et al.* (2007) proposed intelligent fault diagnosis systems with rule-based reasoning technique separately, which used process variable observations to reason by rule and then drew diagnosis conclusions. But the methods mentioned above suffer from certain limitations. Firstly diagnosis conclusions are only categorized by results rather than causes. Secondly the accuracy of diagnosis is not reliably guaranteed. Furthermore, with a large number of variables in the process, the correlation among variables is quite complicated because of process coupling and closed loop feedback. It requires extensive prior knowledge and tedious work to establish rules for a fault case, which means considerable cost.

Data-driven fault diagnosis methods have advantages of no need of process model, dimension reduction, easy visua-lization, and ease of use and maintenance. Using multivariate statistical analysis, statistical process monitoring (SPM) has found wide applications in many industrial processes, including chemicals, polymers, and microelectronics manufacturing (Qin, 2003). To take advantage of the convenience of data-driven methods, a data-driven reconstruction and reconstruction-based contribution (RBC) method is therefore applied to the fault diagnosis of shaft furnace roasting processes. With fault directions derived from historical faulty data, it can reconstruct faults to identify fault types. In addition, major contributing fault variables are singled out.

The objective of this paper is to present an application of data-driven fault diagnosis method for an important industrial process, i.e., the shaft furnace roasting process. The paper is organized as follows. Section 2 provides descriptions of the shaft furnace roasting process and the process faults. Section 3 describes the reconstruction and RBC method. Following that, a case study on a simulation system of shaft furnace roasting processes is presented in Section 4. Section 5 discusses conclusions and further work.

2. SHAFT FURNACE ROASTING PROCESS AND PROCESS FAULTS

2.1 Description of Shaft Furnace Roasting Processes

The shaft furnace roasting processes consist of ore feeding, ore preheating, heating, reduction, and cooling & discharge phases, as shown in Fig. 1.

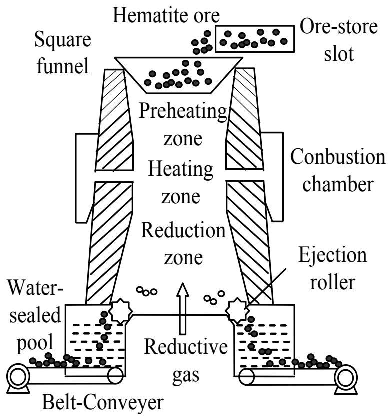

Fig. 1. Illustration of shaft furnace roasting processes

The raw ores of Fe_2O_3 mixed with reductive gas turn to ore concentrate in the reduction zone after being heated in the preheating zone and the heating zone. The following reductions take place when a proper temperature range in the reduction zone is realized.

$$\begin{aligned} &3Fe_2O_3+CO \longrightarrow 2Fe_3O_4+CO_2 \\ &3Fe_2O_3+H_2 \longrightarrow 2Fe_3O_4+H_2O \end{aligned} \tag{1}$$

If the reductions are not completely performed, the output of reaction is a blend of strongly magnetic ore Fe_3O_4 and weakly magnetic one Fe_2O_3. And when reduction overreacts in a higher temperature, following reductions would take place. However resultant FeO is also weakly magnetic.

$$\begin{aligned} &Fe_3O_4+CO \longrightarrow 3FeO+CO_2 \\ &Fe_3O_4+H_2 \longrightarrow 3FeO+H_2O \end{aligned} \tag{2}$$

According to some basic knowledge, the process variables that correlate with faulty operating situations and product quality include the following categories of variables: i) pressures of the heating gas, heating air, and reductive gas; ii) negative pressure inside the furnace; iii) temperatures of the combustion zone, preheating zone, and exhaust gas; iv) flow rates of the heating gas, heating air, and reductive gas, and v) discharging time. In a normal operating situation, control loops can balance the relationship between temperature of the combustion chamber, flow rate of the heating air and discharging time to perform the desired reduction, but also to guarantee normal ranges of negative pressure inside the furnace, temperature of the exhaust gas, and so on. But because of correlations between process variables, a normal situation is at risk of process faults even though variables under closed loop are able to follow desired trajectories. The process faults would alter the correlations between variables. This property provides an option to use data-driven SPM methods to monitor the processes.

2.2 Description of Process Faults

Mismatched set points in control loops would not only degrade ore concentrate, but also adversely lead to process faults. The process faults of shaft furnace roasting processes include the following four main kinds.

a) Fire-emitting: fire emits out of the combustion chamber;

b) Flame-out: flame reaches above the top of the furnace;

c) Under-reduction: raw ores are pushed out of the furnace before fully reduced;

d) Over-reduction: ore concentrate is over reduced before coming out of the furnace.

In a real industrial setting, diagnosis of these faults is mostly based on the operator's observation and experience. It could hardly meet the need of process faults diagnosis. Process faults rarely have well-understood mechanism or patterns of emergence. And the boundary between normal and faulty situations is blurry. Setting up rules for diagnosis requires much process knowledge and trial-and-error rule adjustments. With so many variables, it is complicated to develop a set of rules to describe various faults.

When a fault occurs, the correlation of variables would be broken. The impact on process variables differs from fault to fault. Thus a vector or subspace can be extracted as the direction of every fault, which makes it possible to identify faults (Valle *et al.*, 2001). A data-driven diagnosis method for shaft furnace roasting processes is proposed in the next section.

3. FAULT DIAGNOSIS OF SHAFT FURNACE ROASTING PROCESSES USING RECONSTRUCTION AND RBC APPROACHES

3.1 Fault Diagnosis Strategy of Shaft Furnace Roasting Processes

The use of multivariate statistics for SPM can yield a latent variable model from data. Principal component analysis (PCA) is a basic projection model in multivariate statistics. The adopted method in this paper is based on PCA and applied to a shaft furnace roasting process shown by Chai *et al*. (2011). The strategy of fault diagnosis for shaft furnace roasting processes is shown in Fig. 2. A detailed description for each module will be discussed as following.

PCA model: The model built with historical normal data is the basis of the whole strategy. Fault relevant variables built in the model should be specified beforehand. Seven of variables are used including temperature of the combustion zone, temperature of the exhaust gas, flow rate of the reductive gas, flow rate of the heating gas, negative pressure inside the furnace, heat value of the heating gas, and predictive magnitude of magnetic tube recovery rate.

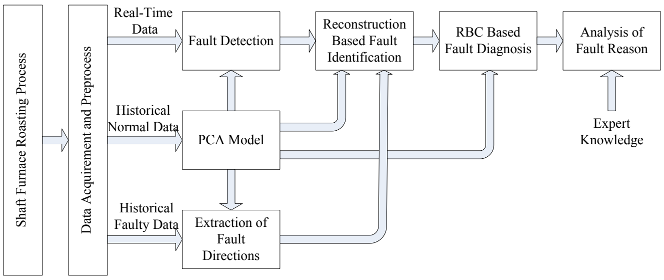

Fig. 2. Strategy of fault diagnosis shaft furnace roasting processes

Fault detection: It is realized by monitoring real-time data of shaft furnace roasting processes with fault detection indices. The PCA model works as a digital template of the shaft furnace in this module. Indices of statistic would show whether or not the process is faulty.

Extraction of fault directions: A dataset of fault directions is derived from historical faulty data of the same variables as used in PCA model. Each direction corresponds to one fault with certain cause. Seven kinds of sensor failure and four kinds of process faults are involved, including fire-emitting, flame-out, under-reduction, and over-reduction.

Reconstruction-based fault identification: It is based on exploitation of real-time data and coordination between the PCA model and the dataset of fault directions. Faults that have happened before can be identified. Fault identification indices will be discussed in details later.

RBC-based fault diagnosis: RBC-based diagnosis can work with PCA model to draw diagnosis results from real-time data. Each one of contribution of variables to the fault is presented for further analysis. When a new fault that has never happened before occurs, reconstruction-based fault diagnosis would fail to diagnose it. After a new fault direction is extracted, the fault can be augmented in the fault library in case the new fault happens again. Fault identification, fault variable diagnosis, and expert knowledge are combined for the root-cause diagnosis. Details about the strategy are shown in Section 3.2.

3.2 Fault Diagnosis Algorithms of Shaft Furnace Roasting Processes

3.2.1 Fault Detection of Shaft Furnace Roasting Processes Based on Principal Component Analysis

Let $\mathbf{x} \in \mathbb{R}^m$ denote a sample vector of m sensors of the shaft furnace roasting processes. Assuming there are N samples of each sensor, a data matrix $\mathbf{X} \in \mathbb{R}^{N\times m}$ is composed of N rows for N samples and m columns for m variables. $\mathbf{X}$ is scaled to zero mean and unit variance. Then the covariance matrix $\mathbf{S}$ of $\mathbf{X}$ can be decomposed by eigendecomposition as

$$\mathbf{S} = \mathbf{P\Lambda P}^{\mathrm{T}} + \tilde{\mathbf{P}}\tilde{\mathbf{\Lambda}}\tilde{\mathbf{P}}^{\mathrm{T}} \tag{3}$$

where $\mathbf{P} \in \mathbb{R}^{m\times l}$ and $\tilde{\mathbf{P}} \in \mathbb{R}^{m\times(m-l)}$ stand for loading matrices of principal components and residual components, l is the number of principal components, and diagonal matrices $\mathbf{\Lambda}$ and $\tilde{\mathbf{\Lambda}}$ respectively contain eigenvalues of the covariance matrix in descending order. Then a new vector $\mathbf{x}$ can be decomposed into two orthogonal subspaces as $\mathbf{x} = \hat{\mathbf{x}} + \tilde{\mathbf{x}}$, where $\hat{\mathbf{x}} = \mathbf{PP}^{\mathrm{T}}\mathbf{x}$ is the projection on principal component subspace and $\tilde{\mathbf{x}} = \tilde{\mathbf{P}}\tilde{\mathbf{P}}^{\mathrm{T}}\mathbf{x}$ is the projection on the residual subspace. The number of components l is determined as Qin & Dunia (2000).

With fault detection indices defined, fault detection can be performed on shaft furnace roasting processes. Qin (2003) presented five kinds of fault detection indices. Among them, SPE and Hotelling's T^2 statistics are most typical and common.

The SPE statistic defined in (4) measures the projection of a sample vector on residual subspace.

$$\mathrm{SPE} \equiv \|\tilde{\mathbf{x}}\|^2 = \mathbf{x}^{\mathrm{T}}\tilde{\mathbf{P}}\tilde{\mathbf{P}}^{\mathrm{T}}\mathbf{x} \tag{4}$$

The process is considered normal if $\mathrm{SPE} \leqslant \delta_\alpha^{\ 2}$, where $\delta_\alpha^{\ 2}$ denotes the upper control limit for SPE with a significance level α.

The Hotelling's T^2 statistic defined in (5) measures variations in principal component subspace.

$$T^2 = \mathbf{x}^{\mathrm{T}}\mathbf{P\Lambda}^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{x} \tag{5}$$

The process is considered normal if $T^2 \leqslant T_\alpha^{\ 2}$, where $T_\alpha^{\ 2}$ denotes the upper control limit for T^2 with a significance level α.

These two control limits are calculated as Alcala & Qin (2009) in this paper.

3.2.2 Fault Diagnosis of Shaft Furnace Roasting Processes Based on Reconstruction

When a process fault occurs, the first step is to detect it. After that, it is necessary to identify the fault for further solution. With historical faulty data and causes available, reconstruction approach can be used to diagnose the faults that have happened before. The detectability, reconstructability, and isolatability of the faults are discussed by Dunia & Qin (1998a) and Dunia & Qin (1998b).

When a certain process fault occurs, samples of sensors in the roasting process need to be captured for further analysis. The same variables as in the PCA model are used here. Let $\mathbf{X}_i \in \mathbb{R}^{N_i\times m}$ denote the faulty data matrix of fault F_i, which consists of N_i rows of N_i samples and m columns of m variables. The work of Valle *et al.* (2001) discussed the relationship between projections of fault direction and faulty data on residual subspace and then provided a method to extract fault direction. We apply SVD on the residual matrix of faulty data $\tilde{\mathbf{X}}_i^{\mathrm{T}}$.

$$\tilde{\mathbf{X}}_i^{\mathrm{T}} = \mathbf{U}_i\mathbf{D}_i\mathbf{V}_i^{\mathrm{T}}$$

The fault direction matrix can be chosen as

$$\tilde{\Xi}_i = \mathbf{U}_i \tag{6}$$

To identify the ongoing fault, it is required to reconstruct sample vectors of the fault with accessible fault directions. The objective of fault reconstruction is to estimate the normal values by eliminating the effect of fault F_i. A reconstructed sample vector $\mathbf{z}_i$ along fault direction Ξ_i is calculated in (7).

$$\mathbf{z}_i = \mathbf{x} - \Xi_i f_i \tag{7}$$

where f_i is the estimation of the magnitude of fault along direction Ξ_i.

Fault reconstruction corrects the effect of a fault, which means it can minimize the fault detection indices of the faulty samples. The reconstructed SPE along direction Ξ_i, i.e. SPE_i becomes

$$\mathrm{SPE}_i = \|\tilde{\mathbf{x}}_i\|^2 = \|\tilde{\mathbf{x}} - \tilde{\Xi}_i f_i\|^2 = \tilde{\mathbf{x}}^{\mathrm{T}}(\mathbf{I} - \tilde{\Xi}_i \tilde{\Xi}_i^{+})\tilde{\mathbf{x}} \tag{8}$$

according to Qin (2012). Then when a process fault takes place, fault detection indices would increase dramatically. If this fault is along the direction Ξ_i, the reconstruction along that direction would correct the effect of the fault in a proper way. So the reconstructed SPE would drop to a relatively normal level. Define a fault identification index

$$\eta_i = \mathrm{SPE}_i / \mathrm{SPE} \tag{9}$$

If Ξ_i is the actual fault, η_i would be close to zero because reconstruction eliminates the effect of the ongoing fault. Then the fault is identified.

3.2.3 Fault Diagnosis of Shaft Furnace Roasting Processes Based on Reconstruction-Based Contribution

Contribution plots are well-known diagnostic tools. It is convenient to be used and requires no prior knowledge. Contribution plots of SPE represent the significance of each variable of SPE, separating fault relevant variables from fault irrelevant ones. The process knowledge on shaft furnace roasting processes is necessary for more convincing explanations of diagnostic conclusion.

If a sample vector x has an abnormal SPE, every single variable has a contribution to it. An investigation into the variables should be carried on, especially into the variables that have a significant contribution. A significant contribution indicates which variables are responsible for an inflated SPE, which means the largest several contributions are likely potential cause of the fault. However the result could be inconclusive and lead to a misdiagnosis because of the smearing effect, which can be avoided by the reconstruction-based contribution.

The reconstruction-based contribution of variable x_i to SPE is used in this work due to the advantage of RBC compared to regular contribution plots. It is defined as Alcala & Qin (2009).

$$\mathrm{RBC}_i^{\mathrm{SPE}} = \mathbf{x}^{\mathrm{T}}\tilde{\mathbf{C}}\Xi_i(\Xi_i^{\mathrm{T}}\tilde{\mathbf{C}}\Xi_i)^{-1}\Xi_i^{\mathrm{T}}\tilde{\mathbf{C}}\mathbf{x} \tag{10}$$

where $\tilde{\mathbf{C}} = \tilde{\mathbf{P}}\tilde{\mathbf{P}}^{\mathrm{T}}$.

4. CASE STUDY ON A SIMULATION SYSTEM OF SHAFT FURNACE ROASTING PROCESSES

In this section, we present a simulation system of shaft furnace roasting processes creating a fine environment approximate to a real shaft furnace field to demonstrate the effectiveness of the proposed method. The simulation system hardware is composed of three parts, the simulated plant, PLC, and fault diagnosis module as shown in Fig. 3.

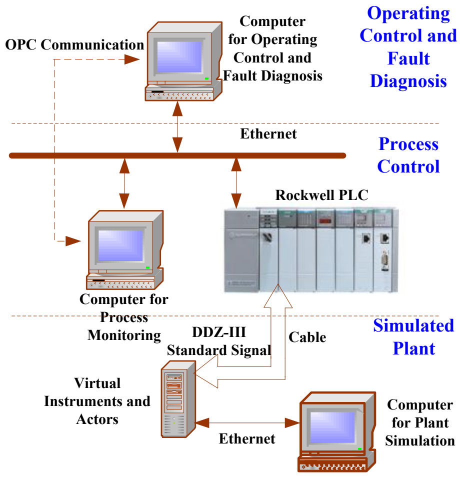

Fig. 3. Structure of the simulation system of shaft furnace roasting processes

The simulated plant is used for simulation of the dynamics of the shaft furnace. PLC achieves process control that a real shaft furnace needs. The fault diagnosis module works as a real-time security for the whole process. There are sensors of pressure, temperature, and flow rate to keep track of variables in the process. The PCA model contains seven fault relevant variables, which are temperature of the combustion zone (TCZ), temperature of the exhaust gas (TEG), flow rate of the reductive gas (FRG), flow rate of the heating gas (FHG), negative pressure inside the furnace (NPF), heat value of the heating gas (HVG), and predictive magnitude of magnetic tube recovery rate (MTRR), as shown in Table 1. As MTRR cannot be collected in real time, a predictive method is adopted from Chai *et al.* (2011).

Table 1. Variables in the PCA model

Variable	No.	Unit
TCZ	1	℃
FRG	2	m^3/h
MTRR	3	%
TEG	4	℃

NPF	5	Kpa
HVG	6	KJ/m^3
FHG	7	m^3/h

A PCA model is derived from normal data. And a fault direction is built using historical faulty data, which consists of seven sensor failures and four process faults, as shown in Table 2.

Table 2. Serial number of faults

Fault	No.
1st sensor failure	1
2nd sensor failure	2
3rd sensor failure	3
4th sensor failure	4
5th sensor failure	5
6th sensor failure	6
7th sensor failure	7
fire-emitting	8
flame-out	9
under-reduction	10
over-reduction	11

In traditional method, every faulty sample corresponds to a fault conclusion. To get an overall conclusion of diagnosis, another fault identification index is defined as follows

$$\sigma_i = N_i / N \tag{11}$$

where N_i stands for the number of samples which draw a conclusion of ith fault, and N stands for the total number of faulty samples.

The sampling period is 1 second in the experiments. We use 600 samples under normal conditions to derive a PCA model for tests. The plots of SPE and T^2 of normal data are shown in Fig. 4. False alarm rate of SPE is 1.0% and false alarm rate of T^2 is 5.17%.

Fault diagnosis results for fire-emitting and under-reduction are shown in Fig. 5 and Fig. 6 respectively. Fig. 5 illustrates that when fire-emitting occurs, SPE and T^2 can both detect it immediately. Both statistics rise visibly above their control limits. The fault identification index σ_i indicates that 100% faulty samples can draw the correct diagnosis conclusion. And RBC plots shows that the 1st and 6th variables, i.e. TCZ and HVG, are responsible for this fault. Because RBCs of these two variables outweigh the others markedly. Fig. 6 can be explained in the same way. 100% faulty samples lead to the correct diagnosis conclusion of under-reduction and the 2nd variable is responsible for this fault because of its RBC outweighing the others significantly.

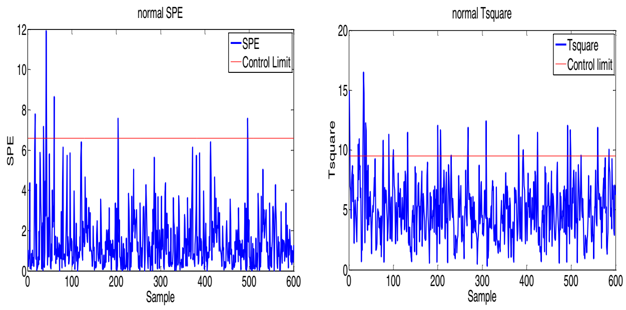

Fig. 4. Plots of SPE and T^2 of normal data

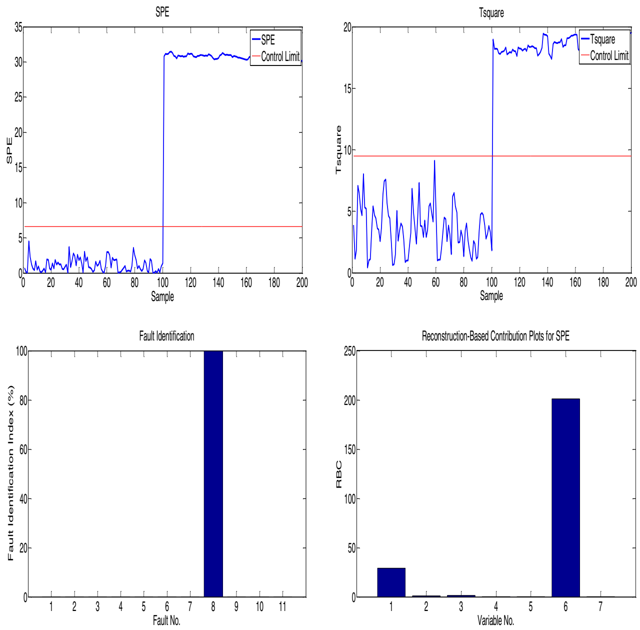

Fig. 5. Diagnosis result for fire-emitting

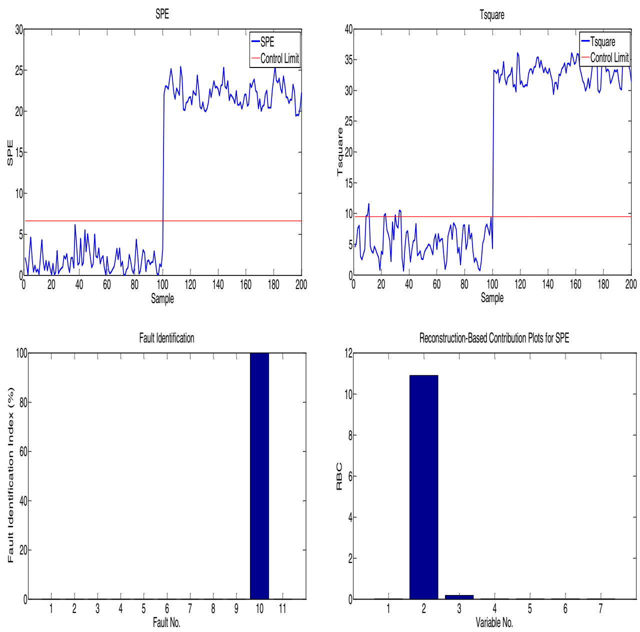

Fig. 6. Diagnosis result for under-reduction

Fault diagnosis results for a new fault are shown in Fig. 7. A new fault means the fault direction is unknown. Fig. 7 presents a confusing diagnosis. The fault is detected by SPE and T^2, but we have got an inconclusive result from fault identification index and RBC plots. 92% of faulty samples indicate the fault is under-reduction while 8% of them indicate the fault is over-reduction. And RBC plots show the fault is not a typical under-reduction fault because the major contributor this time is 6^{th} variable HVG, which means this fault does not share the same cause as under-reduction. When a new fault occurs, the actual cause should be analysed with expert knowledge. And then the direction should be extracted. Thus database of directions of faults could be supplemented in case that this fault would happen again.

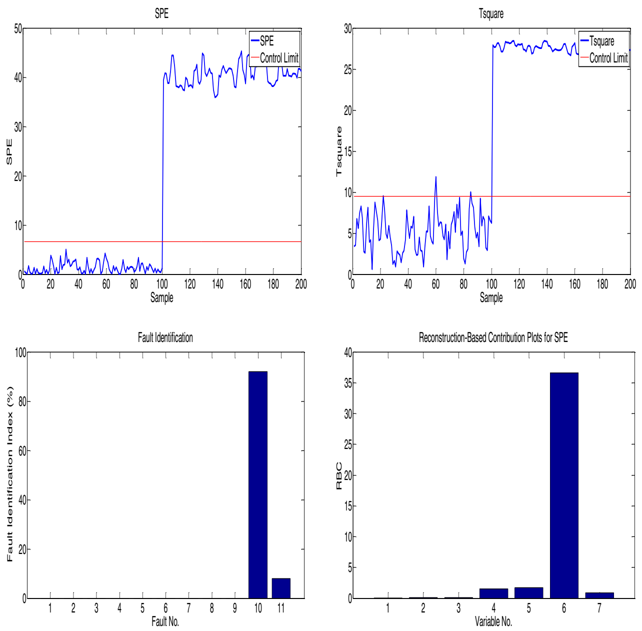

Fig. 7. Diagnosis result for a new fault

5. CONCLUSIONS

In this paper we applied a data-driven diagnosis method for fault diagnosis of shaft furnace roasting processes. The application to the simulation system of shaft furnace roasting process shows that reconstruction and RBC are effective tools for process fault diagnosis of shaft furnaces. When there are much more variables to be monitored in the process, multiblock analysis proposed by Qin *et al.* (2001) and Liu *et al.* (2013) can be effective to interpret the contribution plots. Feedback control would make it difficult to identify the faults because feedback control obscures the source of faults (McNabb & Qin, 2005). Further, operating control in the shaft furnace system is quite complicated. Future work is to find a way to eliminate the feedback control effect in the closed loop system of shaft furnace for accurate diagnosis results.

ACKNOWLEDGEMENTS

This work was supported in part by the Natural Science Foundation of China (61304107, 61020106003, 61290323, 61333007, 61104084, 61203102), the China Postdoctoral Science Foundation funded project (2013M541242), the 111 Project of Ministry of Education of China (B08015), and the IAPI Fundamental Research Funds (2013ZCX02-01).

REFERENCES

Alcala, C. and Qin, S. (2009). Reconstruction-based contribution for process monitoring. *Automatica*, **45(7)**: 1593-1600.

Chai, T., Ding, J. and Wu, F. (2011). Hybrid intelligent control for optimal operation of shaft furnace roasting process. *Control Engineering Practice*, **19(3)**: 264-275.

Chai, T., Wu, F., Ding, J. and Su C. (2007). Intelligent work-situation fault diagnosis and fault-tolerant system for roasting process of shaft furnace. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, **221**: 843-855.

Dunia, R. and Qin, S. (1998a). Joint diagnosis of process and sensor faults using principal component analysis. *Control Engineering Practice*, **6**: 457-469.

Dunia, R. and Qin, S. (1998b). Subspace approach to multidimensional fault identification and reconstruction. *Process Systems Engineering*, **44(8)**: 1813-1831.

Liu, Q., Qin, S. and Chai, T. (2013). Decentralized fault diagnosis of continuous annealing processes based on multilevel PCA. *IEEE Transactions on Automation Science and Engineering*, **10(3)**: 687-698.

McNabb, C. and Qin, S. (2005). Fault diagnosis in the feedback-invariant subspace of closed-loop systems. *Ind. Eng. Chem. Res.*, **44**: 2359-2368.

Qin, S. (2003). Statistical Process monitoring: basics and beyond. *Journal of Chemometrics*, **17**: 480-502.

Qin, S. (2012). Survey on data-driven industrial process monitoring and diagnosis. *Annual Reviews in Control*, **36**: 220–234.

Qin, S. and Dunia, R. (2000). Determining the number of principal components for best reconstruction. *Journal of Process Control*, **10**: 245-250.

Qin, S., Valle, S. and Piovoso, M. (2001). On unifying multiblock analysis with application to decentralized process monitoring. *Journal of Chemometrics*, **15(9)**: 715-742.

Valle, S., Qin, S., Piovoso, M., Bachmann, M. and Mandakoro, N. (2001). Extracting fault subspaces for fault identification of a polyester film process. *Proceedings of the American Control Conference*, **6**: 4466-4471.

Wu, F., Ding, J., Yue, H. and Chai, T. (2006). Intelligent fault diagnosis system for roasting Process of Shaft Furnace. *Journal of Nanjing University of Aeronautics & Astronautics*, **38**: 91-94.

Yan, A., Wang, P. and Zeng, Y. (2008). Intelligent fault prediction system of combustion process in shaft furnace. *CIESC Journal*, **59(7)**: 1768-1771.

Zhou, D. and Hu Y. (2009). Fault diagnosis techniques for dynamic systems. *Acta Automatica Sinica*, **35(6)**: 748-758.