

## Research Article

# Fault Diagnosis Method on Polyvinyl Chloride Polymerization Process Based on Dynamic Kernel Principal Component and Fisher Discriminant Analysis Method

Shu-zhi Gao,<sup>1</sup> Xiao-feng Wu,<sup>1</sup> Gui-cheng Wang,<sup>1</sup> Jie-sheng Wang,<sup>2</sup> and Zi-qing Chai<sup>3</sup>

<sup>1</sup>College of Information and Engineering, Shenyang University of Chemical Technology, Shenyang 110142, China

<sup>2</sup>National Financial Security and System Equipment Engineering Research Center, University of Science & Technology Liaoning, Anshan 114044, China

<sup>3</sup>Beijing Institute of Technology, School of Software, Beijing 100081, China

Correspondence should be addressed to Jie-sheng Wang; wang\_jiesheng@126.com

Received 8 July 2016; Accepted 29 September 2016

Academic Editor: Yaguo Lei

Copyright © 2016 Shu-zhi Gao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In view of the fact that the production process of Polyvinyl chloride (PVC) polymerization has more fault types and its type is complex, a fault diagnosis algorithm based on the hybrid Dynamic Kernel Principal Component Analysis-Fisher Discriminant Analysis (DKPCA-FDA) method is proposed in this paper. Kernel principal component analysis and Dynamic Kernel Principal Component Analysis are used for fault diagnosis of Polyvinyl chloride (PVC) polymerization process, while Fisher Discriminant Analysis (FDA) method was adopted to make failure data for further separation. The simulation results show that the Dynamic Kernel Principal Component Analyses to fault diagnosis of Polyvinyl chloride (PVC) polymerization process have better diagnostic accuracy, the Fisher Discriminant Analysis (FDA) can further realize the fault isolation, and the actual fault in the process of Polyvinyl chloride (PVC) polymerization production can be monitored by Dynamic Kernel Principal Component Analysis.

## 1. Introduction

In today's chemical industry, the polymer production has occupied a very important position. Polyvinyl chloride (PVC) is an important organic synthetic material, and it is also the chemical product which has a variety of uses. PVC resin is a kind of chemical products, the failure mechanism of the production process is complex, and there are serious dynamic and nonlinear problems in the production process; then there is an urgent need to improve the system reliability and safety production, in order to avoid generating fault system; otherwise it will lead to economic losses and even accidents. Therefore, the fault study in production of PVC resin has been an important issue research for experts [1, 2]. At present the principal component analysis is widely applied in chemical process fault diagnosis. Due to the large number of pieces of data in PVC resin production process has serious nonlinear [3] and strong coupling and dynamic problems [4, 5], the rate of false positives and nonresponse rates of the

process by using conventional principal component analysis (PCA) are too high. So in order to make the PCA used in more fields, both the domestic and foreign scholars, respectively, conducted a series of improvements, for example, the kernel principal component analysis (KPCA) and dynamic principal component analysis [6] (DPCA). In 1995, in view of the multi-sensor related timing measurements, a sensor fault detection based on the dynamic principal component analysis (DPCA) method is proposed by Ku et al. [7]; simulation experiments show that this method can effectively detect and identify the fault sensors. DPCA method was applied to the oil state overhaul model by Makis et al.; the simulation results show that DPCA can accurately detect the fault of  $f$  dynamic process [8]. On the nonlinear problem, there are mainly the methods of neural network and principal component analysis, and targeting the problems caused by nonlinear data processing, Kramer proposed nonlinear principal component analysis methods based on the association of the neural network. The simulation results show that it can effectively change multiple

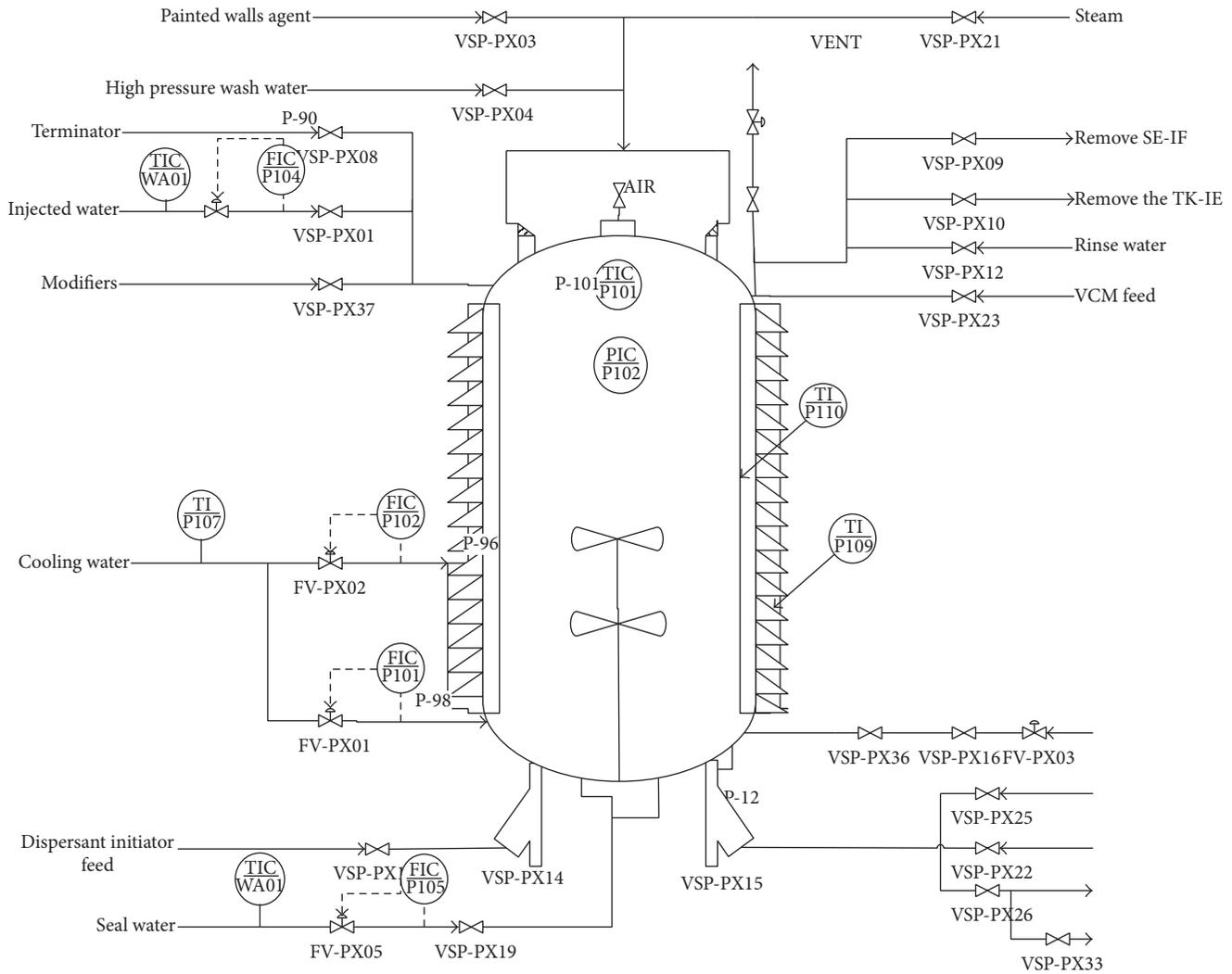


FIGURE 1: Polymerizer process diagram.

variables into a few independent variables [9]. Because the autoassociative neural network is difficult to be trained, Dong and McAvoy proposed a nonlinear PCA method based on element curves and neural network [10]; the results show that the nonlinear function can be expressed as a sum of the single-variable functions. In dealing with nonlinear problems, the introduction of kernel functions (KPCA) method is proposed by Jong-Min et al. [11]. The results show that this method can be carried out for the nonlinear process fault diagnosis. Results show that, in the monitoring process of chemical process, the improved principal component analysis method is better than the traditional principal component analysis.

After analyzing the characteristics of the data for the production of PVC resin, there are serious nonlinear and dynamic data, and this article adopts the method of a compound Dynamic Kernel Principal Component Analysis (DKPCA), using the nonlinear and dynamic process, and, after studying it, there is the defect of the traditional principal component. At the same time this paper also included the fisher discriminant method for further classification of failure

data, to ensure that one quickly finds out the cause of the problem, then for further processing.

## 2. PVC Polymerization Introduction

Taking a unit of the production process of PVC resin as the research object, polymerization process is shown in Figure 1.

The monomer, initiator, dispersion agent, and other materials are fed into the polymerization reactor from inlets. After a series of reactions, the polymers are ultimately generated. In the polymerization process, the heat is ceaselessly released so that the temperature in the polymerization reactor is continuously increased, which will make the reaction more severe and causing material flow imbalances so as to affect the product quality. So the mixing system is added in the polymerization reactor, and the cooling water is timely injected in the jacket and the damper in order to balance the temperature in the polymerization reactor. After the reaction, join the terminators; then the polymerization reaction will be terminated.

From PVC polymerization process, there are many indicators affecting the quality of product and any one variable over its boundaries; it will lead the quality of the product to be out of control. Therefore, perform the real-time monitoring of quality index in the process of polymerization, and then ensure the product quality changed can be diagnosed timely and accurately. After a comprehensive comparative study, we ultimately selected 10 variables identified as the object of study; these 10 variables are as follows: the temperature within the reactor, the reactor pressure, stirring currents, injection into the water flow, seal water flow, jacket water flow, water flow baffles, cooling water inlet temperature, outlet temperature of the water jacket, and the water outlet temperature of the baffle. Monitor the 10 variables, respectively, and ensure the stable operation of the polymerization system.

### 3. Basic Principle of KPCA and DKPCA

**3.1. Basic Principle of KPCA.** PCA is based on data to establish the system statistical model and detect anomalies and failures based on the multivariate statistical techniques. First, take some normal conditions data set  $X_{n \times m}$  ( $n$  is the number of sampling points;  $m$  is the number of variables) to build a statistical model. PCA is made to process data vector which projected onto the main element vector space and residual space. Matrix  $X$  can be expressed as the sum of  $n$  vector cross products; namely,

$$X = TP^T = t_1 p_1^T + t_2 p_2^T + \cdots + t_n p_n^T, \quad (1)$$

where  $T = [t_1, t_2, \dots, t_n]$  is called scoring matrix and  $P = [p_1, p_2, \dots, p_n]$  is called the loading matrix. Each score vector and load vectors are orthogonal, and the load vector length is 1; namely,

$$p_i^T p_j = \begin{cases} 0 & i \neq j, \\ 1 & i = j. \end{cases} \quad (2)$$

Equation (1) is carried out with the right multiplication with  $p_i$  to obtain  $t_i = X p_i$ .

It can be seen that each score vector is equal to the data matrix projection on its load vector, the size of  $t_i$  determines the cover degree of the data matrix projection on the load vectors, the load vector related by maximum score vector represents the maximum change in the direction of the data  $X$ , and it is defined as the first principal component and so on; determine the second main element, the third main element, and the  $n$ th main element. The data  $X$  changes are mainly embodied in the first several principal components, the projection data in the next few load vectors will be small, and it is mainly caused by noise, expressed with matrix  $E$ , and, then, type (1) can be changed for the following form:

$$X = TP^T = t_1 p_1^T + t_2 p_2^T + \cdots + t_k p_k^T + E \quad (k \ll n). \quad (3)$$

Generally adopt the principle of cumulative contribution rate, and determine the number of PCA.  $k$  is the number of selected principal components,  $E$  is the error matrix, ignoring  $E$  usually can serve the purpose of clearing the measurement

noise, equivalent to changing  $n$  dimensional data into  $k$  dimensional PCA and  $n - k$  dimensional residual space  $E$ , it serves dimension reduction purposes. After the principal number is determined, the PCA statistical model is set up through the two subspaces. Commonly the Hotelling  $T^2$  and  $Q$  statistic (the squared prediction error, SPE) are used, which are calculated by the following equation:

$$Q = e_i e_i^T = \bar{X}_i (I - P_k P_k^T) \bar{X}_i^T. \quad (4)$$

The statistic index of control limits is calculated as follows:

$$Q_L = a (b + c z_a)^d, \quad (5)$$

where

$$\begin{aligned} a &= \sum_{i=k+1}^n \lambda_i, \\ b &= 1 + \frac{[\theta_2 h_0 (h_0 - 1)]}{a^2}, \\ c &= \frac{(c_a \sqrt{2\theta_2 h_0})}{a}, \\ d &= \frac{1}{h}, \end{aligned} \quad (6)$$

$$\theta_2 = \sum_{i=k+1}^n \lambda_i^2,$$

$$\theta_3 = \sum_{i=k+1}^n \lambda_i^3,$$

$$h_0 = \frac{(1 - 2a\theta_3)}{(3\theta_2^2)}.$$

$c_a$  is the critical value of normal distribution in the test level and  $\lambda_i$  is the eigenvalue of matrix eigenvalues of original data  $X$ .  $T^2$  statistic is defined as

$$T_i^2 = t_i \lambda^{-1} t_i^T = \bar{X}_i P_k \lambda^{-1} P_k^T \bar{X}_i^T. \quad (7)$$

$T^2$  statistic index of the control limit calculations is calculated as follows:

$$T_{k,m,\alpha}^2 = \frac{k(n-1)}{n-k} F_{k,n-1,\alpha}, \quad (8)$$

where  $\alpha$  is the significance level of test,  $n$  is the data sampling frequency, and  $m$  is a variable number.  $k$  is the number of the principal components and  $F_{\alpha}(k, n - k)$  is the distributed critical value in accordance to the level of test  $\alpha$  and the degrees of freedom  $(k, n - k)$ .

When statistics are within the scope of the control limit, then the system is in a trouble-free state; if the statistics are beyond the control limit, then there is a fault in the system. The basic idea of KPCA is through nonlinear map  $\phi$ , and make the input spaces  $X_1, X_2, \dots, X_n \in R^M$  onto the feature space  $F$ .  $\phi : R^M \rightarrow F$ , and then calculate principal component on the new feature space.

In the feature space  $F$ , covariance matrix can be calculated as follows:

$$C^F = \frac{1}{N} \sum_{i=1}^N \phi(x_i) \phi(x_i)^T. \quad (9)$$

By determining the characteristics of the vector  $C^F$ , we can get the principal component of  $F$ , and the feature vector  $C^F$  is directly related to the input space of PCA.

$$\begin{aligned} \lambda V &= C^F V = \left( \frac{1}{N} \sum_{i=1}^N \phi(x_i) \phi(x_i)^T \right) V \\ &= \frac{1}{N} \sum_{i=1}^N \langle \phi(x_i), V \rangle \phi(x_i), \end{aligned} \quad (10)$$

where  $\langle x, y \rangle$  is representative dot product of  $x$  and  $y$ , and then it can be inferred that, in any condition of  $\lambda \neq 0$ , all solutions of  $V$  can be determined by  $\phi(x_1), \dots, \phi(x_n)$ . So  $\lambda V = C^F V$  is equivalent to

$$\lambda \langle \phi(x_k), V \rangle = \langle \phi(x_k), C^F V \rangle, \quad k = 1, \dots, n. \quad (11)$$

There is a coefficient of  $a_i$  ( $i = 1, \dots, N$ ), where  $V = \sum_{i=1}^N a_i \phi(x_i)$ .

Therefore, combined with the type, it can be obtained that

$$\begin{aligned} \lambda \sum_{i=1}^N a_i \langle \phi(x_k), \phi(x_i) \rangle \\ = \frac{1}{N} \left\langle \phi(x_k), \sum_{i=1}^N \phi(x_j) \right\rangle \langle \phi(x_j), \phi(x_i) \rangle. \end{aligned} \quad (12)$$

At the same time,  $N \times N$  matrix is defined,  $K \in R^{N \times N}$ , and

$$[K]_{ij} = K_{ij} = \langle \phi(x_i), \phi(x_j) \rangle. \quad (13)$$

Therefore, (12) can be simplified as

$$\lambda N a = K a, \quad a = [a_1, \dots, a_N]^T \quad (14)$$

In the application of KPCA, first you have to get the mean centered on high dimensional space. It can be done by using the following formula instead of nuclear matrix  $K$ :

$$\bar{K} = K - I_N K - K I_N + I_N K I_N, \quad (15)$$

where each element of  $I_N$  is equal to  $1/N$  and  $E \in R^{N \times N}$ .

Therefore, the principal  $t$  of vector  $x$  can be acquired through  $\phi(x)$  mapping to the characteristic vector  $V_k$  of  $F$ , where

$$\begin{aligned} t_k &= \langle V_k, \phi(x) \rangle = \sum_{i=1}^N a_i^k \langle \phi(x_i), \phi(X) \rangle \\ &= \sum_{i=1}^N a_i^k k(x_i, X), \quad k = 1, \dots, p. \end{aligned} \quad (16)$$

We found that KPCA is by introducing  $k(x, y) = \langle \phi(x), \phi(y) \rangle$  kernel function, avoiding dot product of the nonlinear mapping, and calculating the feature space. The choice of kernel function is completely determine by  $\phi$  and feature space  $F$ . The most commonly used kernel function is radial basis kernel function:

$$k(x, y) = \exp\left(-\frac{\|x - y\|^2}{\sigma}\right). \quad (17)$$

The kernel function used in this paper is radial basis kernel function.

#### 4. The Principle of Dynamic Kernel Principal Component Analysis

The traditional PCA method to the diagnosis result of system process data with dynamic properties and the nonlinear characteristics is not very ideal, so for such a dynamic nonlinear system one should study a new method; this method must be able to capture the dynamic and nonlinear characteristics of the data at the same time. This is the following Dynamic Kernel Principal Component Analysis (DKPCA) method.

The fault detection principle of Dynamic Kernel Principal Component Analysis is as follows: analyze the system dynamic characteristics at first, the time-series data of the system at normal state were analyzed, and kernel principal component analysis mode of the system under the normal state is constructed. Then a new system data is mapped to the kernel principal component model and, respectively, by principal component scores and the residuals to determine the state of the system.

DKPCA fault diagnosis method is that each of the observation variables is expanded by  $H$  observations in front, the augmented matrix containing the first  $S$  time observations is constructed [12], and augmented matrix is as follows:

$$X(S) = \begin{pmatrix} X_t^T & X_{t-1}^T & \cdots & X_{t-s}^T \\ X_{t-1}^T & X_{t-2}^T & \cdots & X_{t-s-1}^T \\ \vdots & \vdots & \ddots & \vdots \\ X_{t-s-n}^T & X_{t+s-n-1}^T & \cdots & X_{t-n}^T \end{pmatrix}. \quad (18)$$

After the augmented matrix by extending the time sequence, the kernel principal component analysis is used for fault detection. DKPCA flow chart is shown in Figure 2.

#### 5. The Fault Classification Research Based on Fisher Discriminant Method

The FDA's basic idea is to maximize the dispersion between classes at the same time to minimize the dispersion in class, through the optimization objective function to determine the optimal FDA vector FDA, and the vector can represent the direction of the different fault class optimal separation [13–15].

$$\max_{v \neq 0} \frac{v^T S_b v}{v^T S_w v}, \quad (19)$$

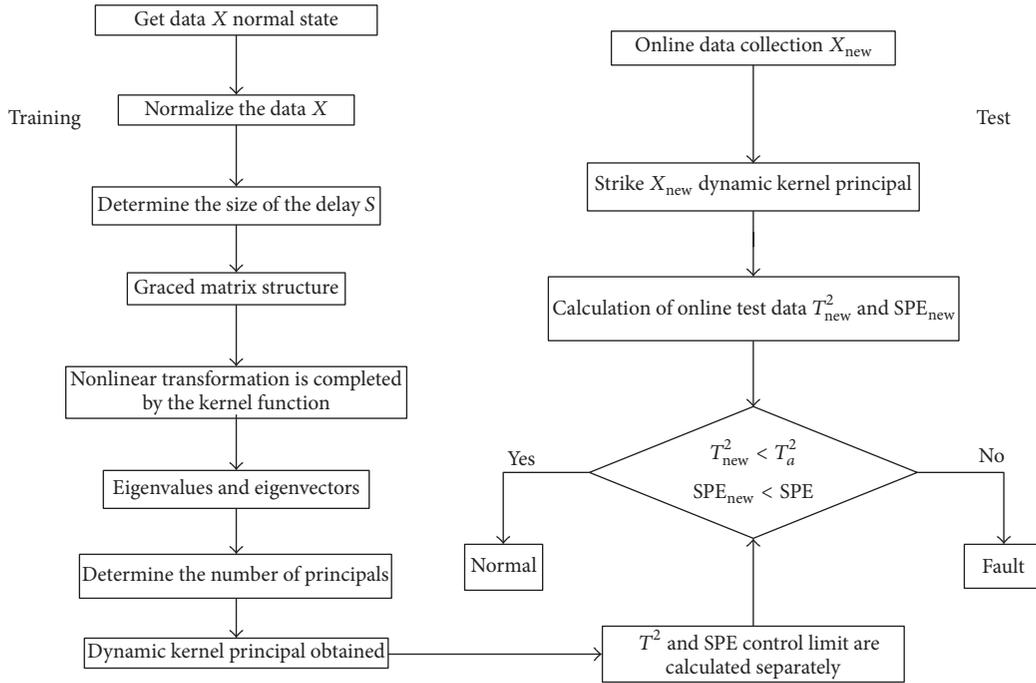


FIGURE 2: Dynamic flow chart of kernel principal component analysis.

where  $S_b$  is the discrete degree matrix between classes,  $S_w$  is the discrete degree matrix in class, and  $v$  is the FDA feature vector [16].

To further classify the fault, the FDA method is adopted for further classification to the failure data detected by DKPCA to ensure that one finds out the cause of the problem quickly, which provides the theoretical basis to make failure countermeasures.

## 6. The Simulation Example Results

First of all, the characteristics of the polymerization process were analyzed, this analysis found that there are 10 process variables affecting the polymerization product quality indicators, 50 sets of data under normal polymerization process are collected and used as the training sample matrix  $X^{50 \times 10}$ , and 200 groups of observed data online are used as test data. Every one minute, sampling occurs one time, in the 51st group failure 1 was introduced, and it is mainly that, because of the rise of temperature, the temperature of the normal production of PVC resin is  $56.5 \pm 0.5$ , as temperatures continue to rise, leading to transformation of resin. In the 125th group failure 2 was introduced, and the stirring electric current is increased by 15%.

According to the above process,  $T^2$  and  $SPE$  are used for fault monitoring of PVC polymerization process by using the traditional principal component analysis, kernel principal component analysis, and the Dynamic Kernel Principal Component Analysis.

- (1) The traditional principal component analysis was adopted, 200 groups of data are to be collected, the

TABLE 1: Diagnostic results.

Diagnostic methods	PCA	KPCA	DKPCA
False alarm rate	17%	8%	1%

sampling interval is 1 min, and the fault detection results are as shown in Figure 3.

- (2) The kernel principal component analysis was adopted, 200 groups of data are to be collected, the sampling interval is 1 min, and the fault detection results are as shown in Figure 4.
- (3) The Dynamic Kernel Principal Component Analysis was adopted, 200 groups of data are to be collected, the sampling interval is 1 min, and the fault detection results are as shown in Figure 5.

In order to more clearly identify which kind of diagnosis methods is more effective, the false positive rates are compared. The lower the false positive rate, the more appropriate this method to be applied in the polymerization kettle fault diagnosis system. Diagnostic results are as Table 1.

It can be clearly seen from the simulation results that there are many cases exceeding the control limits and the false positive rate is quite serious for the traditional principal component analysis method and the kernel principal component analysis method before the 50 sampled points, which will lead to the incorrect judgment of the operators. But, in the fault detection of Dynamic Kernel Principal Component Analysis, there are only two false alarms, and the introduction of fault phase, obvious beyond control limit, thus improves the efficiency of fault diagnosis.

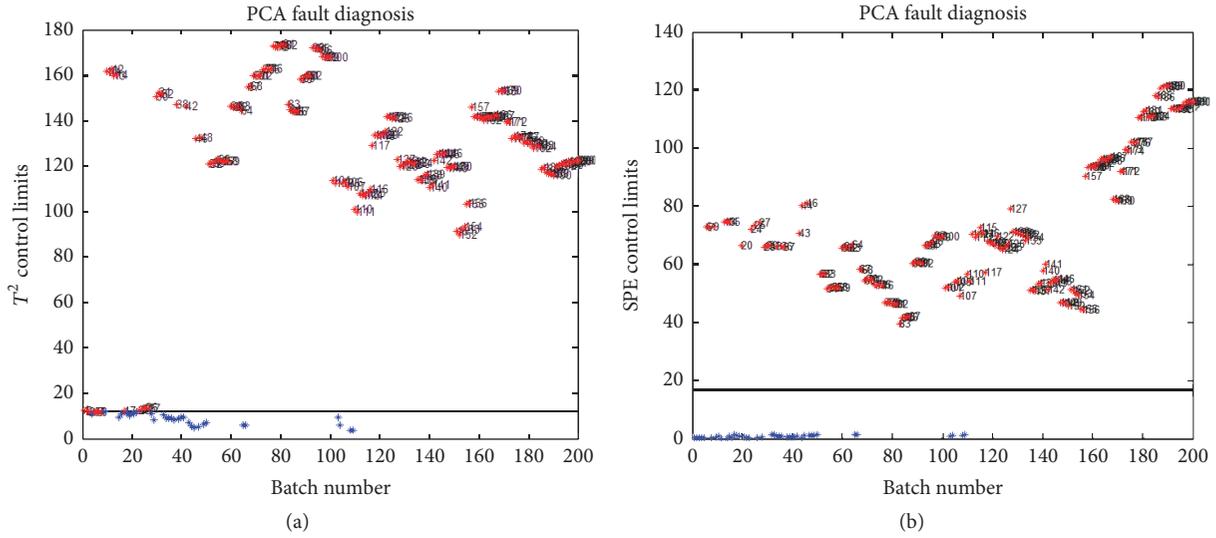


FIGURE 3: Principal component analysis of fault detection results.

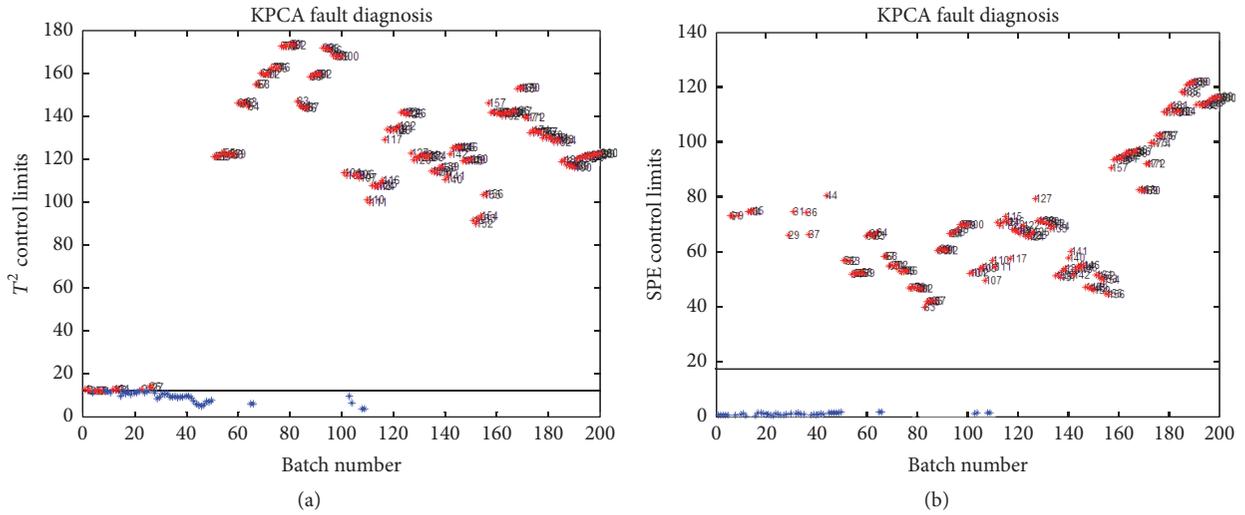


FIGURE 4: Kernel principal component analysis of fault detection results.

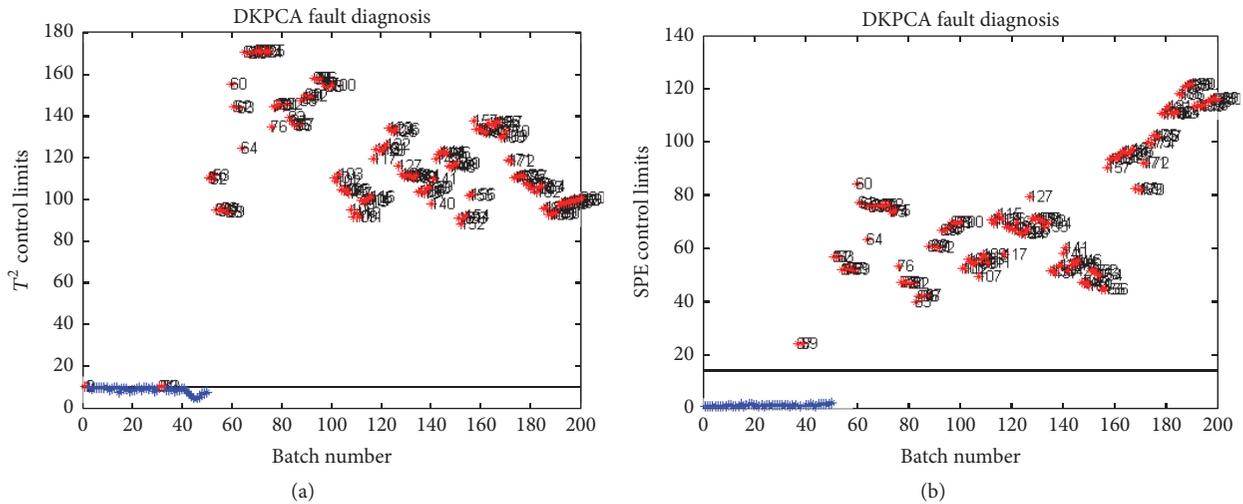


FIGURE 5: Dynamic Kernel Principal Component Analysis of fault detection results.

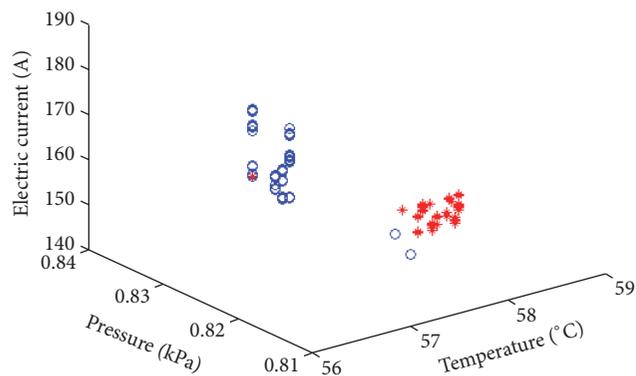


FIGURE 6: Fault classification results of FDA.

In order to further find out the cause of the problem, we need to make further classification of failure data, to ensure rapid recovery production, and now we will extract 100 groups detected by DKPCA. There are two types of failures, respectively, one is resin transformation due to temperature and pressure, and the other failure is caused by the stirring electric current; in order to make better separation of two types of failure, we use the method of judging the FDA, and the results are shown in Figure 6.

By the graph, we can see clearly that \* represents the failure caused by temperature and pressure and  $\circ$  represents the failure caused by stirring electric current increased, fault separation accuracy reached 97%, and it can be used to isolate two faults in the polymerization process.

## 7. Conclusions

Based on the traditional principal component analysis and kernel principal component analysis, the data of dynamic principle were introduced, introducing Dynamic Kernel Principal Component Analysis method, which is used for fault detection to the dynamic and nonlinear strong polymerizing process. At the same time, there are further references to the FDA methods for fault isolation. Simulation results show that the method can real time monitor the change of variables in the polymerization process; the fault of the polymerization process is more sensitive to reduce the probability of false alarm and can quickly find out the cause of the problem. Results indicate that this method can be applied to the fault handling of polymerization, at the same time, and, for the general nonlinear dynamic chemical process, also it has certain applicability.

## Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this article.

## Acknowledgments

This work is partially supported by the Project by National Natural Science Foundation of China (Grant no. 21576127),

the Program for Liaoning Excellent Talents in University (Grant no. LR2014008), the Project by Liaoning Provincial Natural Science Foundation of China (Grant no. 2014020177), the Program for Research Special Foundation of University of Science and Technology of Liaoning (Grant no. 2015TD04), and the Opening Project of National Financial Security and System Equipment Engineering Research Center (Grant no. USTLKFGJ201502).

## References

- [1] S. Z. Gao, X. W. Gao, J. S. Wang, and P. C. Fei, "Rough set-neural network fault diagnosis of polymerization based on improved attribute reduction algorithm of discernibility matrix," *Journal of Chemical Industry and Engineering*, vol. 62, no. 3, pp. 759–765, 2011.
- [2] Y. H. Gao and Z. Yang, "An application of PCA for monitoring and diagnosing fault in a chemical polymeric process," *Journal of Southern Yangtze University*, vol. 4, no. 4, pp. 352–356, 2005.
- [3] G. W. Dou and A. L. Liu, "Fault detection based on kernel principal component analysis," *Chinese Journal of Scientific Instrument*, vol. 30, no. 6, pp. 443–447, 2009.
- [4] H. Song, H. Zhang, and X. Wang, "Multiple faults diagnosis approach for nonlinear system," *Journal of Beijing University of Aeronautics and Astronautics*, vol. 31, no. 11, pp. 1198–1203, 2005.
- [5] S.-H. Jiang, W.-H. Gui, C.-H. Yang, and Z.-H. Tang, "Method based on kernel principal component analysis and support vector machine and its application," *Journal of Central South University (Science and Technology)*, vol. 40, no. 5, pp. 1323–1328, 2009.
- [6] L. Li, J. N. Zhu, and H. B. Shi, "Fault detection of chemical process based on multiscale dynamic kernel principal component analysis, control and instruments in chemical industry," *Control and Instruments in Chemical Industry*, vol. 35, no. 4, pp. 23–26, 2008.
- [7] W. Ku, R. H. Storer, and C. Georgakis, "Disturbance detection and isolation by dynamic principal component analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 30, no. 1, pp. 179–196, 1995.
- [8] V. Makis, J. Wu, and Y. Gao, "An application of DPCA to oil data for CBM modeling," *European Journal of Operational Research*, vol. 174, no. 1, pp. 112–123, 2006.
- [9] M. A. Kramer, "Nonlinear principal component analysis using autoassociative neural networks," *AIChE Journal*, vol. 37, no. 2, pp. 233–243, 1991.
- [10] D. Dong and T. J. McAvoy, "Nonlinear principal component analysis—based on principal curves and neural networks," *Computers & Chemical Engineering*, vol. 20, no. 1, pp. 65–78, 1996.
- [11] L. Jong-Min, C. K. Yoo, S. W. Choi et al., "Nonlinear Process monitoring using kernel Principal component analysis," *Chemical Engineering Science*, vol. 59, 2004.
- [12] H. T. Shi, J. C. Liu, X. D. Ding, and S. Tan, "Fault detection based on hybrid dynamic principal component analysis," *Control Engineering of China*, vol. 19, no. 1, pp. 148–150, 2012.
- [13] Z. Q. Bian and X. G. Zhang, *Pattern Recognition*, Tsinghua University Press, Beijing, China, 1999.
- [14] H. H. Xin, *Process Monitoring based on Fisher Discriminant Analysis*, China University of Petroleum (East China), 2011.

- [15] N. Lv, *Process Monitoring based on Fisher Discriminant Analysis*, Harbin University of Science and Technology, 2009.
- [16] L. H. Chiang, E. L. Russell, and R. D. Braatz, *Fault Detection and Diagnosis in Industrial Systems*, Advanced Textbooks in Control and Signal Processing, Springer, London, UK, 2001.



# Hindawi

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

