Postprint of K-Nearest Neighbor Fault Diagnosis Strategy Based on Variance Maximization Rotation Transformation

Authors: Zhang Cheng, Guo Qingxiu, Li Yuan

Date: 2018-05-24T00:00:00+00:00

Abstract

To improve the fault detection capability of FD-KNN for latent variables in nonlinear and multimodal processes, a K-nearest neighbor fault detection and diagnosis strategy based on variance-maximizing rotation transformation is proposed. First, a rotation transformation is established through the variance maximization method to transform the original data into a new orthogonal space; next, the FD-KNN method is executed in this orthogonal space for fault detection; finally, a contribution plot-based fault diagnosis strategy is presented by combining the contribution plot method. Through a nonlinear simulation example, the effectiveness of the method for latent variable fault diagnosis is demonstrated; simultaneously, testing is conducted on the Tennessee Eastman Process, a typical nonlinear industrial process, and comparison is made with PCA, FD-KNN, and PC-KNN methods, with experimental results further proving the effectiveness of the method.

Full Text

Varimax Rotation-Based K-Nearest Neighbors Fault Diagnosis Strategy

Zhang Cheng, Guo Qingxiu, Li Yuan†

Research Center for Technical Process Fault Diagnosis & Safety, Shenyang University of Chemical Technology, Shenyang 110142, China

Abstract: To improve the fault detection capability of FD-KNN for latent variables in nonlinear and multimodal processes, this paper proposes a K-nearest neighbor fault detection and diagnosis strategy based on varimax rotation. First, a rotation transformation is established via the varimax method to transform the original data into a new orthogonal space. Next, the FD-KNN method is executed in this orthogonal space for fault detection. Finally, a contribution plot-based fault diagnosis strategy is presented. A nonlinear simulation example demonstrates the effectiveness of the proposed method for latent variable fault diagnosis. Additionally, tests on the Tennessee Eastman process—a typical nonlinear industrial process—show comparative results against PCA, FD-KNN, and PC-KNN methods, further validating the effectiveness of the approach.

Keywords: K-nearest neighbors; varimax rotation; fault detection; fault diagnosis; process control; principal component analysis

0 Introduction

With the rapid development of industrial production, process monitoring and fault diagnosis technologies have become increasingly important for ensuring production safety and improving efficiency. Since distributed control systems can automatically measure and store large amounts of multivariate sampling data, data-driven multivariate statistical process control (MSPC) methods have emerged as a hot research topic [1]. However, modern industrial processes are characterized by mechanization, electrification, automation, and chemical processing, resulting in numerous complex industrial systems with nonlinear and multimodal features. Typical examples include penicillin fermentation processes, semiconductor aluminum etching processes, the Tennessee Eastman (TE) process, polymerization processes, and wastewater treatment processes [2–6].

Principal component analysis (PCA) is a typical data-driven MSPC method that has been widely applied in chemical production processes with promising results [7–9]. PCA decomposes the original space into a principal component subspace (PCS) and a residual subspace (RS) through linear transformation based on variance information, and monitors these two subspaces using Hotelling' s T^2 and squared prediction error (SPE) statistics. The control limits for T^2 and SPE are typically determined under the assumption that process variables follow a multivariate Gaussian distribution, which affects PCA' s fault detection performance in nonlinear processes [10]. To address nonlinear process monitoring, several PCA-based nonlinear methods have been proposed and rapidly developed [11–12]. In recent years, kernel principal component analysis (KPCA) based on kernel theory has been effectively used for nonlinear system process monitoring [13–15].

To capture process data dynamics and perform related application analyses, dynamic PCA (DPCA) was developed by augmenting time series variables to expand the data matrix and incorporate autocorrelation characteristics [16–17]. DPCA extends traditional PCA to dynamic multivariate processes and demonstrates good performance in fault diagnosis of systems with time delays due to its consideration of serial correlation in data. However, methods such as PCA and KPCA are suitable for monitoring single-mode processes; when applied to multimodal processes, they typically exhibit high false alarm rates [18].

To monitor multimodal processes, He et al. proposed a fault detection method based on K-nearest neighbors (FD-KNN) [19]. FD-KNN uses the distance similarity between the current sample and its neighbors to measure the sample status, making variable scaling a primary consideration. FD-KNN is typically combined with Z-score standardization to adjust monitoring variable scales and improve detection performance. However, when samples have high dimensions, FD-KNN requires calculating distances in high-dimensional data, resulting in significant computational time and storage space requirements. To address these limitations, He et al. proposed a combined PCA and FD-KNN detection method (principal component-based K-nearest neighbor rule for fault detection, PC-KNN) [20]. PC-KNN first obtains the principal component subspace using PCA, then executes FD-KNN in this subspace. By calculating inter-sample distances in a low-dimensional principal component subspace, PC-KNN reduces computational complexity and improves efficiency. However, PC-KNN only detects faults occurring in the principal component subspace; when faults occur entirely in the residual subspace, PC-KNN becomes ineffective. In summary, PC-KNN does not demonstrate superior detection performance compared to FD-KNN.

Considering the limitations of PC-KNN and the advantages of FD-KNN in handling nonlinear and multimodal processes, this paper proposes a varimax rotation-based K-nearest neighbor fault detection and diagnosis strategy (Rot-KNN) for typical industrial processes with nonlinear and multimodal characteristics. Varimax rotation is a method used in PCA that maximizes the sum of variances of factor loadings through coordinate transformation while rotating the loading matrix toward a simple structure. Compared to PC-KNN, Rot-KNN increases computational load but enables more comprehensive fault detection without the residual subspace detection deficiency of PC-KNN. Compared to FD-KNN, Rot-KNN improves detection capability for latent variable faults. Compared to PCA, KPCA, and DPCA, Rot-KNN effectively handles process monitoring with nonlinear and multimodal characteristics, improving fault detection rate (FDR) while reducing false alarm rate (FAR).

---

1 FD-KNN

The fundamental principle of FD-KNN is to measure sample differences using the distance between a sample and its nearest neighbors. First, the k nearest neighbors of sample X are identified in the training set. Second, the sum of squared distances between X and its k nearest neighbors is calculated:

$$D_i^2 = \sum_{j=1}^{k} d_{ij}^2$$

where d_{ij} is the Euclidean distance between sample X_i and its j-th nearest neighbor $X_{(j)}$. Finally, the detection control limit D_α^2 is determined based on the non-central chi-square distribution with confidence level α [21]. Since kernel density estimation (KDE) is a typical non-parametric distribution estimation method that has achieved good results [22–26], it can also be applied to determine the control limit. When the D^2 statistic of a test sample exceeds the control limit D_α^2, it is classified as a fault; otherwise, it is normal.

The specific calculation process of FD-KNN is as follows:

1) Offline Modeling a) Find the k nearest neighbor samples for each sample X_i in the training set X; b) Calculate the distance D_i^2 between X_i and its k nearest neighbors using Eq. (1); c) Determine the control limit D_α^2 for the training model.

2) Online Detection a) Find the k nearest neighbor samples for test sample X^* in the training set X; b) Calculate the distance statistic D^{*2} between X^* and its k nearest neighbors using Eq. (1); c) Compare D^{*2} with D_α^2: if $D^{*2} > D_\alpha^2$, X^* is judged as a fault; otherwise, it is normal.

Due to differences in variable scales and units, effective standardization of sampling data is typically required before process monitoring. Z-score is a commonly used data preprocessing method that adjusts variable means and variances to enable comparison of variables with different units on the same scale. Therefore, Z-score is often combined with monitoring methods to improve process monitoring capability, as seen in FD-KNN, PCA, and KPCA.

The combination of Z-score and FD-KNN is effective for monitoring processes with independent variables. Consider a dataset containing two measurement variables x_1 and x_2, where $x_1 \sim N(0, 16)$ and $x_2 \sim N(0, 1)$. Randomly generating 200 training samples and two fault samples F1 and F2 yields the scatter plot shown in Figure 1: see original paper. Fault samples F1 and F2 deviate from the data center along the x_1 and x_2 directions, respectively, with a fault magnitude of 4 times the standard deviation length (e.g., F1(6,4,0), F2(0,4)). Applying Z-score standardization to this dataset produces the result shown in Figure 1: see original paper, where fault samples are separated from training samples, and FD-KNN can accurately identify both faults. The D^2 detection results are shown in [Figure 2: see original paper].

However, when variables are correlated, the Z-score and FD-KNN combination has limitations for detecting subtle faults. For example, when two measurement variables satisfy the relationship $x_1 = \cos(\theta)s - \sin(\theta)t$, $x_2 = \sin(\theta)s + \cos(\theta)t$ with $\theta = \pi/4$, the data distribution is shown in [Figure 3: see original paper]. Applying the Z-score and FD-KNN combination yields the detection results in

[Figure 4: see original paper], where Fault 2 is misclassified as normal. This error occurs because Z-score standardization is applied to observed variables x_1 and x_2, which cannot simultaneously standardize the underlying variables s and t. If data standardization is performed along the s and t directions, both faults can be detected by FD-KNN. For this example, establishing a new coordinate system based on data dispersion and performing standardization along the new directions can improve FD-KNN detection capability.

2.1 Varimax Rotation Transformation

Let random vector $x = [x_1, x_2, \dots, x_n]^T$ contain n random variables with covariance matrix $\Sigma = E(xx^T)$. Let λ_i and p_i be the eigenvalues and eigenvectors of Σ, respectively, with λ_i sorted in descending order. Consider the linear transformation:

$$y = P^T x$$

where $P = [p_1, p_2, \dots, p_n]$ is an orthogonal matrix. The random variables $y = [y_1, y_2, \dots, y_n]^T$ have variances $Var(y_i) = \lambda_i$ and covariances $Cov(y_i, y_j) = 0$ for $i \neq j$, meaning y_i and y_j are uncorrelated. As shown in [1,2], this transformation represents the directions of maximum variability, and its covariance matrix has a simple yet refined diagonal structure.

2.2 Varimax Rotation-Based K-Nearest Neighbor Fault Detection and Diagnosis Strategy

Let the training dataset be $X \in \mathbb{R}^{m \times n}$, where m and n are the numbers of samples and monitored variables, respectively. First, compute the covariance matrix Σ of X and perform eigen-decomposition. Denote the eigenvalues and eigenvectors as λ_i and p_i, respectively, sorted in descending order. Next, apply the rotation transformation $Y = XP$ according to Eq. (2) and standardize Y to obtain the standardized dataset Y_{std}. Finally, apply the FD-KNN method on Y_{std} for fault detection. Performing standardization sequentially along orthogonal directions with large variability before applying FD-KNN can improve fault detection rates.

The Rot-KNN procedure is as follows:

1) Model Building a) Center the training set X; b) Perform eigen-decomposition of the covariance matrix Σ of X; c) Apply rotation transformation $Y = XP$ and standardize Y to obtain the centered matrix Y_{std}; d) Execute the FD-KNN method on Y_{std} and determine the statistical control limit D^2_α using KDE.

2) Online Detection For a test sample x_{te}, perform the following steps: a) Center x_{te} using the mean from step 1a; b) Apply rotation transformation $y_{te} = x_{te}P$ and standardize y_{te} as in step 1c, denoted as $y_{te,std}$; c) Find the k nearest neighbors of $y_{te,std}$ in Y_{std} and calculate the distance statistic D^2_{te}; d) Compare D^2_{te} with D^2_α: if $D^2_{te} > D^2_\alpha$, x_{te} is a fault; otherwise, it is normal.

When a sample is detected as faulty, it is necessary to diagnose the cause. In classical methods such as PCA and KPCA, contribution plots are typically used for fault diagnosis [27-29]. For a fault sample x centered to x_0, after varimax rotation transformation it becomes $y = x_0P$. The statistical index D^2 for sample y can be calculated as:

$$D^2 = \sum_{j=1}^{k} \left(\frac{1}{k} \sum_{h=1}^{n} \lambda_h^{-1} (y_h - y_h^{(j)})^2 \right) = \sum_{h=1}^{n} \lambda_h^{-1} \sum_{j=1}^{k} (y_h - y_h^{(j)})^2$$

where $y_h^{(j)}$ is the h-th variable of the j-th nearest neighbor sample. The contribution of variable y_h to D^2 can be derived as:

$$con_h = \lambda_h^{-1} \sum_{j=1}^{k} (y_h - y_h^{(j)})^2$$

If sample x is detected as faulty and y_h is an out-of-control variable after coordinate rotation, then $con_h > D^2_\alpha$. Since $y_h = \sum_{i=1}^{n} x_i p_{ih}$, Eq. (6) can be rearranged as:

$$con_h = \lambda_h^{-1} \sum_{j=1}^{k} \left(\sum_{i=1}^{n} p_{ih} (x_i - x_i^{(j)}) \right)^2 = \lambda_h^{-1} \sum_{j=1}^{k} \left(\sum_{i=1}^{n} p_{ih} \Delta x_i^{(j)} \right)^2$$

where $\Delta x_i^{(j)} = x_i - x_i^{(j)}$. The contribution of original variable x_i to out-of-control variable y_h is:

$$Con_{i \to h} = \lambda_h^{-1} p_{ih} \sum_{j=1}^{k} (x_i - x_i^{(j)}) \left(\sum_{l=1}^{n} p_{lh} (x_l - x_l^{(j)}) \right)$$

The total contribution of variable x_i is:

$$C_i = \sum_{h \in \mathcal{S}} Con_{i \to h}$$

where $\mathcal{S}$ is the set of out-of-control variables. Variables with larger C_i are typically considered the primary causes of faults.

1) Fault Diagnosis For a fault sample $x = [x_1, x_2, \ldots, x_n]$, the diagnosis procedure is: a) Calculate each variable' s contribution con_h to D^2 using Eq. (6) and determine the number of out-of-control variables s; b) Calculate the contribution $Con_{i\to h}$ of original variable x_i to each out-of-control variable y_h using Eq. (8); c) Calculate the total contribution C_i of variable x_i using Eq. (9).

---

3 Simulation Experiments

3.1 Numerical Example

This example validates the proposed method using a nonlinear system with four observed variables, modeled as:

$$\begin{cases} x_1 = s - 0.2t^2 + 0.3s^2 + e_1 \\ x_2 = t + 0.1s^2 + e_2 \\ x_3 = s + t + e_3 \\ x_4 = 0.5s - 0.3t + e_4 \end{cases}$$

where latent variables s and t follow uniform distributions on intervals $[-50, 50]$ and $[0, 1]$, respectively. Five hundred samples were randomly generated for training, and 100 samples for validation. Two fault classes were created by adjusting initial values of latent variables, with Fault 2 having a larger deviation scale than Fault 1 relative to the uniform distribution of variables. [Figure 5: see original paper] shows the sample distribution, while [Figure 6: see original paper] displays the distribution of each variable.

First, conventional FD-KNN was applied with neighbor number $k = 5$ (optimized via testing). After Z-score standardization, the detection results are shown in [Figure 7: see original paper]. FD-KNN cannot effectively detect Fault 1. As seen in [FIGURES:5] and [6], Fault 1 exhibits a slight shift in the direction of latent variable t, but this shift is not visually apparent in observed variables x_i; in other words, from the perspective of observed variables alone, Fault 1 samples appear normal. Therefore, Z-score standardization along observed variables cannot improve FD-KNN detection performance.

Next, Rot-KNN was applied and compared with FD-KNN. The observed dataset was centered, and varimax rotation transformation was performed to obtain a new dataset. The standardized variable distributions are shown in [Figure 8: see original paper]. From variable y_4, Fault 1 clearly deviates from the normal trajectory. Applying FD-KNN on the rotated dataset yields the detection results in [Figure 9: see original paper], where both fault classes are completely detected. Observation of fault magnitudes shows Fault 2 has a larger scale than Fault 1, consistent with the initial setup.

For comparison, traditional PCA-T^2, PCA-SPE, and PC-KNN methods were also applied. Based on a cumulative variance contribution rate (CPV) of 90%, two principal components were selected, dividing the original space into PCS and RS. [Figure 12: see original paper] shows PCA-T^2 and PCA-SPE results. Combined detection from both subspaces only identifies Fault 2, indicating these methods are effective only for large-scale faults. Fault 1 is missed primarily because T^2 and SPE control limit determination assumes multivariate Gaussian distributions, which is not satisfied here, as shown in [Figure 13: see original paper]. PC-KNN, which executes FD-KNN in the PCS, yields the results in [Figure 14: see original paper]. PC-KNN monitors sample variations only in PCS without considering RS variations, resulting in high false alarm rates when faults occur entirely in RS.

Based on Rot-KNN detection results, diagnosis was performed on two fault samples (the 50th sample from each class). First, Eq. (6) identified the fourth rotated variable y_4 as out-of-control, as shown in [Figure 10: see original paper]. Then, Eq. (7) calculated original variable contributions to y_4, displayed in [Figure 11: see original paper]. Variables x_1 and x_2 show large contributions to y_4 for both faults, indicating their abnormal variations cause the faults. Indeed, from model (10), abnormal changes in latent variable t can be determined through the linear combination of x_1 and x_2 since they satisfy $t = 0.1853x_1 - x_2$.

3.2 TE Process

The TE process is a simulation system proposed by Downs et al. based on a real chemical production process at Tennessee Eastman [30]. The TE simulation system accurately designs nonlinear relationships between unit operations and material/energy balances, as well as vapor-liquid equilibrium laws, enabling simulation of various real faults in chemical processes. Due to severe coupling, high nonlinearity, and open-loop instability, the TE process has been widely used as a benchmark platform for research in process control technology, process monitoring and optimization, and process operation system integration [31]. As shown in [Figure 15: see original paper], the TE process comprises five operating units: a reactor, condenser, recycle compressor, separator, and stripper. It includes four gaseous reactants (A, C, D, E), two liquid products (G, H), and byproduct F and inert gas B. The entire process contains 22 continuous process measurement variables, 19 composition measurement variables, and 11 control variables. As a standard benchmark problem that realistically simulates many typical characteristics of actual complex industrial processes, the TE process has been extensively applied in control, optimization, process monitoring, and fault diagnosis research [32–34].

The TE simulation system can simulate both normal and 21 fault environments with a 3-minute sampling interval [27]. Following [36], 33 variables were selected for fault analysis; 960 normal samples were used for model building; 480 normal samples for model validation; and each fault type contains 960 samples with faults introduced at time 161 and continuing to the end.

As described in [35], faults 3, 9, and 15 have small magnitudes and stable process behavior, making them difficult to detect. PCA-T^2, PCA-SPE, and FD-KNN can promptly and accurately detect faults 1, 2, 6, 7, 8, 12, 13, and 14 with FDR above 95%. Various PCA-based methods, such as KPCA and DPCA, achieve FDR below 85% for faults 5, 10, 16, and 19. This section compares Rot-KNN with PCA-T^2, PCA-SPE, and FD-KNN for faults 5, 10, 14, 16, and 19. Parameter settings are listed in , and FDR/FAR results are shown in . For fault 14, Rot-KNN achieves 100% FDR with lower FAR than other methods. For faults 5, 10, 16, and 19, Rot-KNN achieves FDR above 90%, while other methods remain below 60%. [Figure 16: see original paper] shows the fault detection control charts for Rot-KNN and FD-KNN on faults 5, 10, 16, and 19.

The contribution plot-based diagnosis strategy was applied to these faults, with results shown in [Figure 17: see original paper]. Since all methods effectively detect fault 14, it was first diagnosed to validate the method. Fault 14 is caused by a stuck reactor cooling water valve, which directly affects reactor temperature (x_9), reactor cooling water outlet temperature (x_{21}), and reactor cooling water flow (x_{32}) [29]. [FIGURE:17(a) shows Rot-KNN diagnosis results, where variables x_9, x_{21}, and x_{32} show significant contributions to D^2, confirming the diagnosis validity. Next, faults 5, 10, and 16 (where Rot-KNN shows high detection rates) were diagnosed. Fault 5 is caused by abnormal condenser cooling water inlet temperature, affecting condenser cooling water outlet temperature (x_{33}) and stripper underflow (x_{17}) measurements, as confirmed in [FIGURE:17(b)]. Fault 10 results from abnormal feed C temperature, affecting stripper underflow (x_{17}) and stripper temperature (x_{18}) measurements, shown in [FIGURE:17(c)]. Fault 16 diagnosis in [FIGURE:17(d) indicates it is mainly caused by abnormal variations in stripper-related variables: stripper underflow (x_{17}), stripper temperature (x_{18}), and stripper valve (x_{31}), meaning the fault occurs primarily in the stripper refining process.

---

4 Conclusion

This paper proposes the Rot-KNN method to improve FD-KNN detection rates for latent variable faults through varimax rotation transformation. Simulation and industrial examples demonstrate its effectiveness compared to PCA, FD-KNN, and PC-KNN. While Rot-KNN outperforms FD-KNN and PC-KNN in detection performance, it has higher computational complexity. Reducing computational complexity is a focus for future research. Additionally, determining the neighbor number k remains an open question; this paper uses cross-validation, which is complex. Future work will focus on k determination methods in Rot-KNN.

References

[1] Qin S J. Statistical process monitoring: basics and beyond [J]. Journal of Chemometrics, 2010, 17 (8-9): 480-502.

[2] Fan S K S, Lin S C, Tsai P F. Wafer fault detection and key step identification for semiconductor manufacturing using principal component analysis, AdaBoost and decision tree [J]. Journal of the Chinese Institute of Industrial Engineers, 2016, 33 (3): 151-168.

[3] Ringwood J V, Lynn S, Bacelli G, et al. Estimation and Control in Semiconductor Etch: Practice and Possibilities [J]. IEEE Trans on Semiconductor Manufacturing, 2010, 23 (1): 87-98.

[4] Wang Fan; Tan Shuai; Yang Yawei, et al. Hidden Markov model-based fault detection approach for multimode process [J]. Industrial & Engineering Chemistry Research, 2016, 55 (16).

[5] Jayavani S, Deka H, Varghese T O, et al. Recent development and future trends in coir fiber-reinforced green polymer composites: Review and evaluation [J]. Polymer Composites, 2016, 37 (11): 3296-3309.

[6] Katsoyiannis A, Samara C. Ecotoxicological evaluation of the wastewater treatment process of the sewage treatment plant of Thessaloniki, Greece [J]. Journal of Hazardous Materials, 2007, 141 (3): 614.

[7] Zhang Cheng, Li Yuan. Study on the fault-detection method in batch process based on statistical pattern analysis [J]. Chinese Journal of Scientific Instrument, 2013, 34 (9): 2103-2110.

[8] Shams M A B, Budman H M, Duever T A. Fault detection, identification and diagnosis using CUSUM based PCA [J]. Chemical Engineering Science, 2011, 66 (20): 4488-4498.

[9] Jaffel I, Taouali O, Harkat M F, et al. A fault detection index using principal component analysis and mahalanobis distance [J]. IfacPapersonline, 2015, 48 (21): 1397-1401.

[10] Li Yuan, Yan Yayun, Tang Xiaochu. Multi-stage online product quality prediction based on local model [J]. Journal of System Simulation, 2016, 28 (4): 966-971.

[11] Jia F, Martin E B, Morris A J. Non-linear principal components analysis for process fault detection [J]. Computers & Chemical Engineering, 1998, 22 (12): S851-S854.

[12] Turnip A, Hong K S, Jeong M Y. Real-time feature extraction of P300 component using adaptive nonlinear principal component analysis [J]. BioMedical Engineering OnLine, 2011, 10 (1): 83.

[13] Lee J M, Yoo C K, Sang W C, et al. Nonlinear process monitoring using kernel principal component analysis [J]. Chemical Engineering Science, 2004, 59

(1): 223-234.

[14] Wang Huang, Yao Ma. Fault detection of batch processes based on multivariate functional kernel principal component analysis [J]. Chemometrics & Intelligent Laboratory Systems, 2015, 149: 78-89.

[15] Zhang, Yingwei. Enhanced statistical analysis of nonlinear processes using KPCA, KICA and SVM [J]. Chemical Engineering Science, 2009, 64 (5): 801-811.

[16] Rato T J, Reis M S. Advantage of using decorrelated residuals in dynamic principal component analysis for monitoring large-scale systems [J]. Industrial & Engineering Chemistry Research, 2013, 52 (38): 13685-13698.

[17] Rato T J, Reis M S. Fault detection in the Tennessee Eastman benchmark process using dynamic principal components analysis based on decorrelated residuals (DPCA-DR) [J]. Chemometrics & Intelligent Laboratory Systems, 2013, 125 (7): 101-108.

[18] Wang Guozhu; Liu Jianchang; Zhang Yingwei, et al. A novel multi-mode data processing method and its application in industrial process monitoring [J]. Journal of Chemometrics, 2015, 29 (2): 126-138.

[19] He Q P, Wang Jin. Fault detection using the k-nearest neighbor rule for semiconductor manufacturing processes [J]. IEEE Trans on Semiconductor Manufacturing, 2007, 20 (4): 345-354.

[20] He Q P, Wang Jin. Principal component based k-nearest-neighbor rule for fault detection [C]// Proc of American Control Conference. 2008: 1606-1611.

[21] He Q P, Wang Jin. Large-scale semiconductor process fault detection using a fast pattern recognition-based method [J]. IEEE Trans on Semiconductor Manufacturing, 2010, 23 (2): 194-200.

[22] Ma Hehe; Hu Yi; Shi Hongbo. Fault detection and identification based on the neighborhood standardized local outlier factor method [J]. Industrial & Engineering Chemistry Research, 2013, 52 (6): 2389–2402.

[23] Kano M, Sakata T, Hasebe S. Just-in-time statistical process control for flexible fault management [C]// Proc of Sice Conference. 2010: 1482-1485.

[24] Wang Guozhu; Liu Jianchang; Li Yuan, et al. Fault detection based on diffusion maps and k nearest neighbor diffusion distance of feature space [J]. Journal of Chemical Engineering of Japan, 2015, 48 (9): 756-765.

[25] Karunamuni R J, Alberts T. A generalized reflection method of boundary correction in kernel density estimation [J]. Canadian Journal of Statistics, 2010, 33 (4): 497-509.

[26] Terrell G R, Scott D W. Variable Kernel Density Estimation [J]. Annals of Statistics, 1992, 20 (3): 1236-1265.

[27] Miller P, Swanson R E, Heckler C E. Contribution plots: a missing link in multivariate quality control [J]. Applied Mathematics & Computer Science, 1998, 8, 775.

[28] Westerhuis J A, Gurden S P, Smilde A K. Generalized contribution plots in multivariate statistical process monitoring [J]. Chemometrics & Intelligent Laboratory Systems, 2000, 51 (1): 95-114.

[29] Huang Jian; Yan Xuefeng. Gaussian and non-gaussian double subspace statistical process monitoring based on principal component analysis and independent component analysis [J]. Industrial & Engineering Chemistry Research, 2015, 54, 1015.

[30] Downs J, Vogel E F. A plant-wide industrial process control problem [J]. Computers & Chemical Engineering, 1993, 17 (3): 245-255.

[31] Wu Yongjian, Yuan Decheng, Guo Jinyu. Application of process monitoring method based on hidden markov model in TE process [J]. Journal of Shenyang University of Chemical Technology, 2004, 18 (2): 144-146.

[32] Ge Zhiqiang, Song Zhihuan. Nonlinear probabilistic monitoring based on the gaussian process latent variable model [J]. Industrial & Engineering Chemistry Research, 2010, 49 (10): 4792-4799.

[33] Wang Jin, He Q P. Multivariate statistical process monitoring based on statistics pattern analysis [J]. Industrial & Engineering Chemistry Research, 2010, 49 (17): 7858-7869.

[34] Jia, Qilong; Zhang, Yingwei. Quality-related fault detection approach based on dynamic kernel partial least squares [J]. Chemical Engineering Research & Design, 2016, 106: 242-252.

[35] Guo Xiaoping, Yang Meng, Li Yuan. Fault location method based on improved reconstruction contribution map [J]. Chinese Journal of Scientific Instrument, 2015, 36 (5): 1193-1200.

[36] Chiang L H, Russell E L, Braatz R D. Fault detection and diagnosis in industrial systems [M]. Springer Science & Business Media, 2001.

Note: Figure translations are in progress. See original paper for figures.