Detection of sensor precision degradation by monitoring second-order statistics [*]

Hui Hou[*] Hongquan Ji[*]

[*] *College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao 266590, P. R. China (e-mail: jihq18@sdust.edu.cn)*

Abstract: For industrial processes, there are usually a number of measurement sensors equipped for monitoring and control purposes. In practice, sensors may suffer from the precision degradation phenomenon due to several aspects such as aging and ambient interference. This phenomenon may lead to imprecise or even incorrect control commands and indications, so the corresponding fault detection task is of vital importance. In this paper, inspired by the fact that precision degradation of a sensor can result in the increase of the variable's variance, an algorithm based on second-order statistics analysis is proposed to accomplish the detection task for sensor precision degradation faults. By employing the sliding window technique, second-order statistics of process variables are first extracted. Then, conventional principal component analysis (PCA) is used as a dissimilarity quantification tool, with detection statistics and corresponding control limits established, to perform fault detection. Finally, simulations on a numerical example and the continuous stirred tank reactor (CSTR) benchmark process are performed to illustrate the effectiveness and advantages of the proposed method, in comparison with some existing methods such as PCA, dynamic PCA, and dissimilarity (DISSIM).

Keywords: Fault detection, process monitoring, sensor precision degradation, second-order statistics, variable variance.

1. INTRODUCTION

Over the past decades, data-driven fault detection has received increasingly research attention from both academia and industry (Qin, 2012; Alauddin et al., 2018; Chen et al., 2018; Ji and Zhou, 2020). On one hand, accurate mathematical models for complex modern industrial processes are difficult to acquire, which restricts the use of model-based fault diagnosis methods; on the other hand, the abroad application of distributed control systems and measurement sensors provides a foundation to acquire lots of industry data. Statistical process monitoring (SPM) has developed one of the most booming data-driven techniques (Jiang and Yin, 2018; Deng et al., 2018). Commonly used SPM methods include principal component analysis (PCA), dynamic PCA (DPCA), independent component analysis, as well as their variants (Kourti and MacGregor, 1995; Nomikos and MacGregor, 1995; Ge et al., 2013).

Sensor is one kind of indispensable devices so as to gather information in modern industrial processes. Usually, sensor measurements are utilized for indication or feedback control purposes. Under the circumstance of sensor malfunction or completely failure, the measurement information provided by corresponding sensors will be imprecise or even incorrect. Further, the control performance or production efficiency of industrial processes can be thus negatively influenced. Therefore, it is particularly important to accomplish timely sensor fault detection, in order to ensure that sensors maintain normal operating conditions (Sharifi and Langari, 2017; Ji et al., 2019).

According to the literature, there are four main types of sensor faults, i.e., complete failure, mean bias, mean drift, and precision degradation (Dunia et al., 1996; Wan and Ye, 2012). The complete failure fault can generally be detected easily, as in this case no effective reading will be provided by the sensor. As for mean bias and mean drift faults, they alter the mean of sensor measurement by a constant and gradually changing deviation, respectively. By contrast, the precision degradation fault keeps the measurement mean unvaried but increases the corresponding variable's variance (Dunia et al., 1996).

Most existing sensor fault detection methods aim at detecting the fault type with mean changed (Harmouche et al., 2014), whereas the sensor fault with variance change is relatively rarely considered. Ji et al. (2018) proposed a new sensor precision degradation fault detection and isolation method by applying Kullback-Leibler divergence within the principal component subspace of PCA. Chen et al. (2019) presented a method named cumulative canonical correlation analysis to perform the sensor precision degradation fault detection and discussed this method for both Gaussian and non-Gaussian cases.

Recently, fault detection by monitoring the statistics of process variables has caught more and more attention. A new statistics pattern analysis framework was proposed to solve many challenges related to advanced process monitoring such as nonlinear process dynamics and non-Gaussian distribution (Wang and He, 2010; He and Wang, 2011). Shang et al. (2017) presented a method called recursive transformed component statistical analysis for incipient fault detection in the field of SPM. The aforesaid methods and their variants have exhibited superior monitoring performance for specific problems. In this work, we only consider the sensor precision degradation fault type currently, and inspired by Wang and He (2010) a new method is developed by monitoring second-order statistics. This method is valid to sensor precision degradation and relatively less computational complexity is required.

The remainder of this paper is organized as follows. In Section 2, some preliminaries are provided, including PCA, DPCA, and dissimilarity (DISSIM). Section 3 formulates the sensor precision degradation fault detection problem, and presents the proposed methodology, with a detailed algorithm provided. Simulation studies on a numerical example and the continuous stirred tank reactor (CSTR) process are given in Section 4, followed by some concluding remarks in Section 5.

2. PRELIMINARIES

2.1 PCA

Assume that the industrial data are collected and stacked into a matrix $\mathbf{X} \in \mathbb{R}^{N \times m}$ with N rows of samples and m columns of sensor readings. After the matrix $\mathbf{X}$ is normalized, every column of $\mathbf{X}$ follows a distribution with zero mean and unit variance. Matrix $\mathbf{X}$ is decomposed as follows:

$$\mathbf{X} = \mathbf{T}\mathbf{P}^{\mathrm{T}} + \mathbf{E} \tag{1}$$

$$\mathbf{T} = \mathbf{X}\mathbf{P} \tag{2}$$

where $\mathbf{P}$ is the loading matrix, $\mathbf{T}$ denotes the score matrix, and $\mathbf{E}$ is the residual matrix. The PCA model can be established by eigenvalue decomposition, singular value decomposition or nonlinear iterative partial least squares algorithms. The eigenvalue decomposition method is:

$$\mathbf{S} = \frac{1}{N-1}\mathbf{X}^{\mathrm{T}}\mathbf{X} = \bar{\mathbf{P}}\bar{\mathbf{\Lambda}}\bar{\mathbf{P}}^{\mathrm{T}} \tag{3}$$

where $\bar{\mathbf{\Lambda}} = diag\{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ contains m eigenvalues of $\mathbf{S}$, and $\bar{\mathbf{P}}$ contains corresponding eigenvectors. The number of principal components, A, is calculated by cumulative percent variance criterion (Valle et al., 1999). $\mathbf{\Lambda} = diag\{\lambda_1, \lambda_2, \ldots, \lambda_A\}$ satisfies

$$\mathbf{\Lambda} = \frac{1}{N-1}\mathbf{T}^{\mathrm{T}}\mathbf{T} \tag{4}$$

The PCA method detects fault with T^2 and SPE statistics, and the two statistics are calculated by (5) and (6). If T^2 and SPE statistics are both below their control limits, the process is deemed as normal; otherwise, some anomalies arise.

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

$$\mathrm{SPE} = \left\|(\mathbf{I} - \mathbf{P}\mathbf{P}^{\mathrm{T}})\mathbf{x}\right\|^2 \leq \delta_{\alpha}^2 \tag{6}$$

2.2 DPCA

DPCA considers series correlation of variables and extends PCA method to dynamic systems. The DPCA method uses time lag shift to establish an augmented matrix $\mathbf{X}^l$, i.e.,

$$\mathbf{X}^l = \begin{bmatrix} \mathbf{x}^{\mathrm{T}}(1) & \mathbf{x}^{\mathrm{T}}(2) & \cdots & \mathbf{x}^{\mathrm{T}}(l+1) \\ \mathbf{x}^{\mathrm{T}}(2) & \mathbf{x}^{\mathrm{T}}(3) & \cdots & \mathbf{x}^{\mathrm{T}}(l+2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}^{\mathrm{T}}(N-l) & \mathbf{x}^{\mathrm{T}}(N-l+1) & \cdots & \mathbf{x}^{\mathrm{T}}(N) \end{bmatrix} \tag{7}$$

Where l is the time lag. Parallel analysis method is a common option to determine the value of l. As a practical matter, the value of time lag is usually determined 2. After the augmented data matrix is established, the rest of modeling procedures remain the same as conventional PCA.

2.3 DISSIM

Kano et al. (2002) proposed DISSIM by analyzing the dissimilarity degree of process data. In order to detect the change of operating condition, the DISSIM method defines a dissimilarity index, termed D, to compare the difference between process data distributions.

Consider two data sets $\{\mathbf{X}_i \in \mathbb{R}^{N_i \times m}, \ i = 1, 2\}$. The covariance matrix for $\mathbf{X}_i$ and their mixing covariance matrix are established as follows

$$\mathbf{R}_i = \frac{1}{N_i - 1}\mathbf{X}_i^{\mathrm{T}}\mathbf{X}_i \tag{8}$$

$$\mathbf{R} = \frac{N_1 - 1}{N - 1}\mathbf{R}_1 + \frac{N_2 - 1}{N - 1}\mathbf{R}_2 \tag{9}$$

The orthogonal matrix $\mathbf{P}_o$ is calculated from (10), i.e. through the eigenvalue decomposition on $\mathbf{R}$. Then, the transformation matrix $\mathbf{P}$ is defined by (11)

$$\mathbf{P}_o^{\mathrm{T}}\mathbf{R}\mathbf{P}_o = \mathbf{\Lambda}_c \tag{10}$$

$$\mathbf{P} = \mathbf{P}_o\mathbf{\Lambda}_c^{-1/2} \tag{11}$$

The original data matrix $\mathbf{X}_i$ is transformed to $\mathbf{Y}_i$ via $\mathbf{P}$, and the covariance matrix of $\mathbf{Y}_i$ is denoted as $\mathbf{S}_i$

$$\mathbf{Y}_i = \sqrt{\frac{N_i - 1}{N - 1}}\mathbf{X}_i\mathbf{P} \tag{12}$$

$$\mathbf{S}_i = \frac{1}{N_i - 1}\mathbf{Y}_i^{\mathrm{T}}\mathbf{Y}_i \tag{13}$$

in which $\{\mathbf{S}_1 + \mathbf{S}_2 = \mathbf{I}\}$ holds (Kano et al., 2002). The dissimilarity index D is defined as

$$D = \frac{4}{m}\sum_{p=1}^{m}(\lambda_p - 0.5)^2 \tag{14}$$

where λ_p is pth eigenvalue of matrices $\mathbf{S}_1$ or $\mathbf{S}_2$.

3. METHODOLOGY

For a steady-state process, suppose that there are m sensors equipped, and a measurement sample is denoted as $\mathbf{x} = [x_1, x_2, \ldots, x_m]^{\mathrm{T}}$. These samples are assumed independent and identically distributed ($i.i.d.$) and follow a multi-normal distribution

$$\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{15}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ respectively denote the mean and covariance matrix of $\mathbf{x}$. When a sensor suffers from the precision

degradation, the corresponding variable variance will be increased, which can be modeled as follows. The fault vector $\mathbf{x}_f$ is denoted as

$$\mathbf{x}_f = \mathbf{x}^* + \mathbf{f}_j \qquad (16)$$

where $\mathbf{x}^*$ denotes the fault-free part, $\mathbf{f}_j$ is the fault vector, and they are independent of each other. To be more specific, $\mathbf{f}_j$ is such a vector in which only the jth element is f_j and other elements are zero, implying that the jth sensor is faulty. To imitate the sensor precision degradation, it is supposed that f_j follows a normal distribution with zero mean and a constant variance, i.e. $f_j \sim \mathcal{N}(0, \sigma_j^2)$. For illustration, Fig. 1 shows a sensor measurement sequence with precision degradation fault imposed starting at the 1001th sample. As intuitively observed, since sample 1001, the variance of sensor reading is increased.
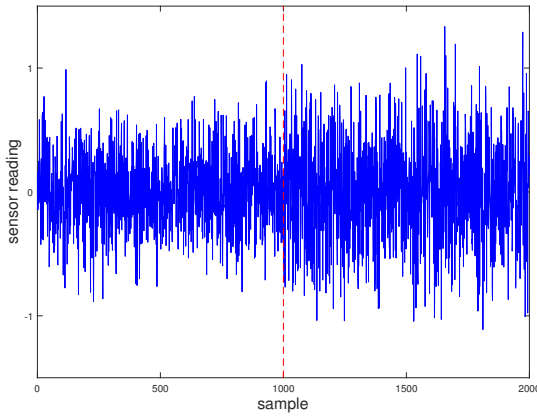

Fig. 1. Illustration of the sensor precision degradation fault

This work aims to achieve successful detection of the sensor precision degradation. As noted above, a distinguish feature for this kind of fault is it can increase the faulty sensor's variance. *Thus, we propose to use a second-order statistics analysis method. Currently, only sample variances are employed for monitoring purposes. If more complex circumstances are considered such as there exist series correlation in some variables, other second-order statistics such as auto-correlation and cross-correlation coefficients can also be incorporated.* The proposed fault detection method can be divided into two parts: the first step is to calculate the variable variances, and the second step is to establish the control limit (offline stage) or perform fault detection (online stage).

3.1 Obtain Variable Variances

The following data matrix $\mathbf{X}^w$ denotes a window of sensor measurement

$$\mathbf{X}^w = \begin{bmatrix} x_1(k-w+1) & x_2(k-w+1) & \dots & x_m(k-w+1) \\ x_1(k-w+2) & x_2(k-w+2) & \dots & x_m(k-w+2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(k) & x_2(k) & \dots & x_m(k) \end{bmatrix} \qquad (17)$$

where w denotes the window width, and k is the current sampling instant. The second-order statistics can

include variance, correlation, auto-correlation, and cross-correlation, and this method only uses the variance statistic currently. For each variable, its sample variance is calculated as follows

$$v_i^2(k) = \frac{1}{w} \sum_{l=0}^{w-1} [x_i(k-l) - \mu_i]^2 \qquad (18)$$

where μ_i is the sample mean of variable x_i. The variances for all variables constitute the following variance sample

$$\boldsymbol{\nu}(k) \triangleq \mathbf{v}^2(k) = [v_1^2(k), v_2^2(k), \dots, v_m^2(k)] \qquad (19)$$

3.2 Fault Detection

In this work, traditional PCA is employed as a tool to quantify the dissimilarities among various variance samples. Two detection statistics, named D_p and D_r are used, with corresponding control limits established. Dp measures the change of variances in the principal component subspace

$$D_p = \boldsymbol{\nu}^{\mathrm{T}} \mathbf{P}_\nu \boldsymbol{\Lambda}_\nu^{-1} \mathbf{P}_\nu^{\mathrm{T}} \boldsymbol{\nu} \qquad (20)$$

D_r measures the change of variances in the residual subspace

$$D_r = \left\| (\mathbf{I} - \mathbf{P}_\nu \mathbf{P}_\nu^{\mathrm{T}}) \boldsymbol{\nu} \right\|^2 \qquad (21)$$

If both D_p and D_r statistics are lower than their corresponding control limits, the process is considered as normal. Otherwise, some faults occur.

According to Wang and He (2010), the central limit theorem states if the samples are independent, their statistics will follow a normal distribution asymptotically. Therefore, in this work the control limits of D_p and D_r can be determined in theory similar to those of T^2 and SPE in PCA. Given the significance level α, the control limit of D_p statistic is computed as

$$T_\nu^2 = \frac{A_\nu(N^2-1)}{N(N-A_\nu)} F_{A_\nu, N-A_\nu; \alpha} \qquad (22)$$

where $F_{A_\nu, N-A_\nu; \alpha}$ denotes the upper α percentile of F distribution with A_ν and $N-A_\nu$ degrees of freedom. A_ν is the number of principal components of PCA applied to the variance sample data matrix. The threshold of D_r can be approximately expressed by

$$\delta_\nu^2 = \theta_1 \left(\frac{c_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + 1 + \frac{\theta_2 h_0(h_0-1)}{\theta_1^2} \right)^{1/h_0} \qquad (23)$$

where $\theta_i = \sum_{j=A_\nu+1}^m \lambda_{\nu,j}^i$, $h_0 = 1 - 2\theta_1\theta_3/3\theta_2^2$, and $\lambda_{\nu,j}$ is the jth eigenvalue of the variance sample covariance. Another method to determine the control limit is the empirical approach based on the calibration data under normal conditions.

In the online detection stage, when a window of measurements are obtained, the variance sample $\boldsymbol{\nu}(k)$ can be obtained at the current instant according to Section 3.1. Then, the D_p and D_r statistics are calculated based on (20) and (21), which are further compared against T_ν^2 and δ_ν^2 to tell whether a fault happens. An algorithm for the fault detection method is summarized in Table 1.

Remark 1. Regarding the proposed method, its detection effect is directly influenced by the fault magnitude and window width. If the fault magnitude is tiny, usually a large window width is required. Nevertheless, if the window width is too large, the detection delay may be increased as well.

Table 1. Fault detection algorithm

| **Off-line design procedure** |
| Input: normal samples; Output: loadings $\mathbf{P}_\nu$, control limits. |
| *S1*: Collect normal process data; |
| *S2*: Slide window and obtain variance sample matrix; |
| *S3*: Apply PCA algorithm; |
| *S4*: Obtain $\mathbf{P}_\nu$ and control limits. |
| **On-line monitoring procedure** |
| Input: online samples;   Output: detection result. |
| *S1*: Obtain online variance sample; |
| *S2*: Calculate D_p and D_r statistics; |
| *S3*: Perform fault detection. |

4. SIMULATIONS

In this section, the effectiveness of the proposed method is illustrated by two simulation studies, including a numerical example and the CSTR process, in comparison with PCA, DPCA and DISSIM.

4.1 A Numerical Example

The process model used here is (Alcala and Qin, 2009)

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} -0.2310 & -0.0816 & -0.2662 \\ -0.3241 & 0.7055 & -0.2158 \\ -0.2170 & -0.3056 & -0.5207 \\ -0.4089 & -0.3442 & -0.4501 \\ -0.6408 & 0.3102 & 0.2372 \\ -0.4655 & -0.4330 & 0.5938 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} + \boldsymbol{\varepsilon} \quad (24)$$

where s_1, s_2, and s_3 are zero-mean Gaussian variables with standard deviation of 1.0, 0.8, and 0.6, respectively. The $\boldsymbol{\varepsilon} \in \mathbb{R}^6$ in (24) also is zero-mean Gaussian variable with a standard deviation of 0.1. One data set containing 2000 normal samples are used for model training. Then, another 2000 independent samples are generated as testing data, where a precision degradation fault is imposed on x_1 starting at sample 1001.
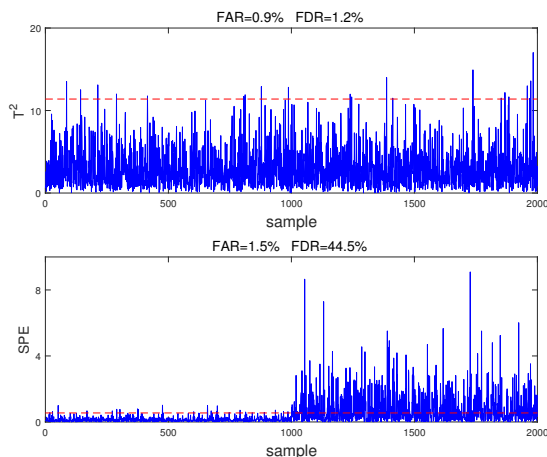

Fig. 2. Fault detection by PCA

Fig. 2 shows the detection results by PCA. False alarm rate (FAR) and fault detection rate (FDR) for the T^2 and SPE statistics are calculated and displayed. As observed, the FDR of SPE is only 44.5%, that is, more than half faulty samples are missed. Therefore, the fault detection result is unsatisfying. Fig. 3 presents the fault detection
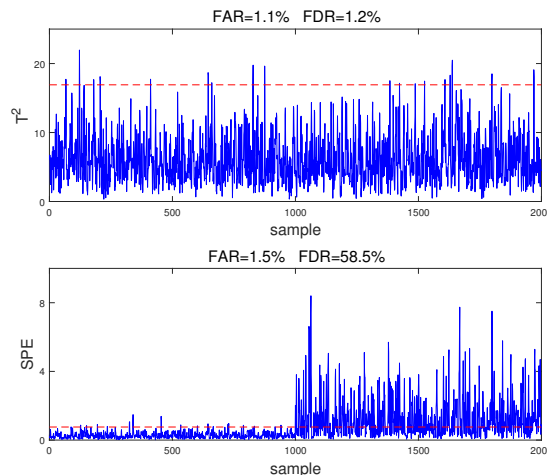

Fig. 3. Fault detection by DPCA

result by DPCA, with time lag $\{l = 3\}$. Compared with PCA, a relatively better detection performance is obtained, however, the FDR is still below 60%.
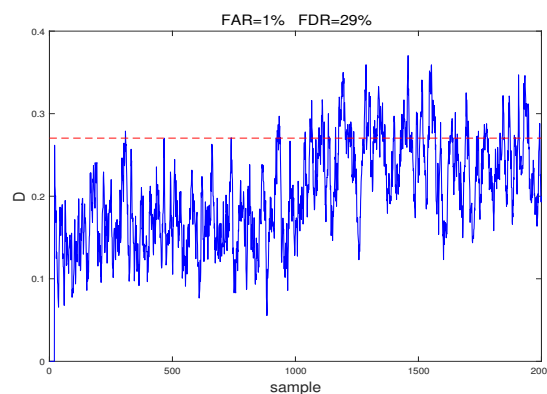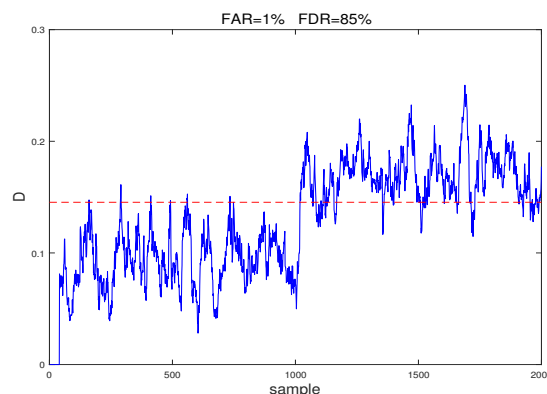

Fig. 4. Fault detection by DISSIM (w=20)

Fig. 5. Fault detection by DISSIM (w=40)

Then we turn to investigate the DISSIM method. Fig. 4 and Fig. 5 show the monitoring results of DISSIM with $w = 20$ and $w = 40$, respectively. It can be observed
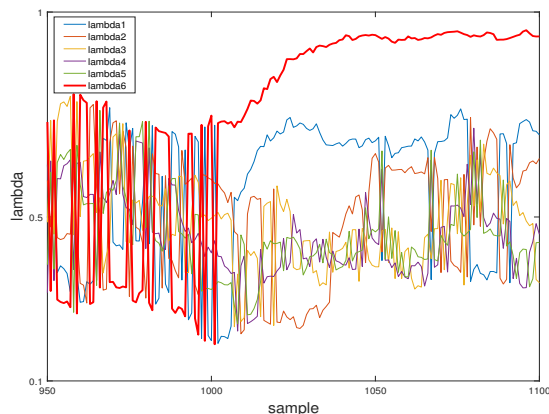
Fig. 6. The eigenvalues of the matrix $\mathbf{S}_i$

that as the window width increases, the fault detection performance of DISSIM is enhanced accordingly. Fig. 6 illustrates the eigenvalues for a period of samples before and after the fault occurrence. As we all know, eigenvectors of the covariance matrix represent the directions of principal components, and eigenvalues are the variances of corresponding scores. Similar to PCA, DISSIM also gets the transformation matrix $\mathbf{Y}$ by projecting the measured variables. The difference is that $\mathbf{Y}$ contains a coefficient, thus the eigenvalues of its covariance matrix are all around 0.5. When the variable x_1 is changed, the maximum and minimum eigenvalues are near one and zero, respectively. According to the definition (14), the change of eigenvalues can be reflected in D. This point has been intuitively demonstrated by Fig. 6. Therefore, the DISSIM with a large window width can detection the sensor precision degradation as well.
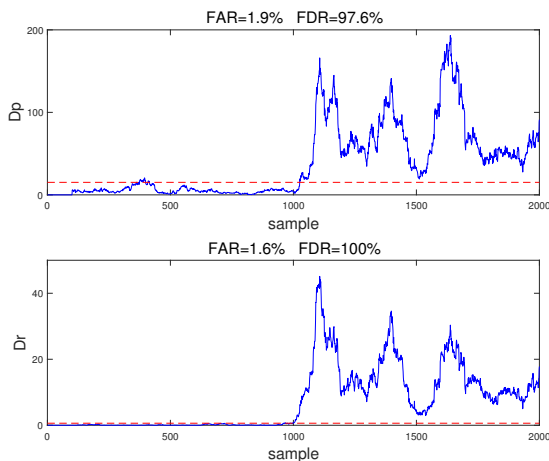

Fig. 7. Fault detection by the proposed method

Fig. 7 presents the fault detection result using the proposed second-order statistics algorithm. For this method, the window width w used in (17) is equal to 100. We can observe from the two sub-figures that both D_p and D_r provide satisfying fault detection performance. Besides, the FARs of these two statistics are approximately equal to the significance level $\alpha = 0.01$, implying that the control limits (22) and (23) are determined reasonably.

4.2 The CSTR Process

The CSTR process (Li et al., 2010) can be described as follows

$$\frac{\mathrm{d}C_A}{\mathrm{d}t} = \frac{q}{V}(C_{Af} - C_A) - k_0 \exp(-\frac{E}{RT})C_A + v_1 \quad (25)$$

$$\begin{aligned}\frac{\mathrm{d}T}{\mathrm{d}t} =& \frac{q}{V}(T_f - T) + \frac{-\Delta H}{\rho C_p}k_0 \exp(-\frac{E}{RT})C_A \\ &+ \frac{UA}{V\rho C_p}(T_c - T) + v_2\end{aligned} \quad (26)$$

where C_A is the outlet concentration of component A, T and T_c are the temperatures of reaction and cooling water, respectively, and q is the flow rate of feed. C_{Af} and T_f are the feed concentration and temperature, respectively. v_1 and v_2 are white Gaussian noises with zero mean, and they are independent. Other variables in (25) and (26) are constants. Four measurements in the process, i.e. C_A, T, T_c, and q are used. Please refer to Li et al. (2010) for more details about the process. A sensor precision degradation fault is imposed on the fourth variable q.
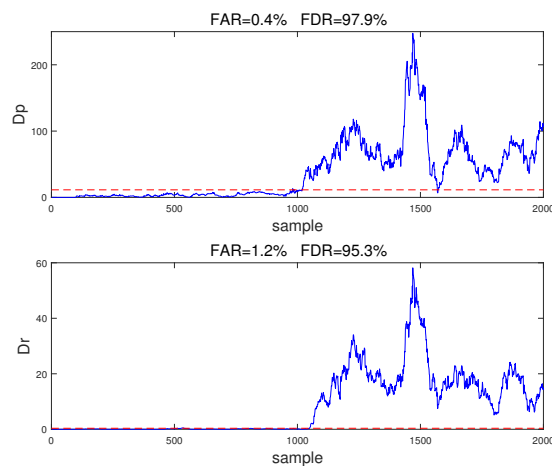

Fig. 8. Fault detection by the proposed method in the CSTR process

Fig. 8 shows the monitoring result by the second-order statistics analysis method. The sensor precision degradation fault is effectively detected, with satisfying FAR and FDR. For comparison purpose, Table 2 shows the fault detection results provided by other methods. Again, for the CSTR process the second-order statistics method is the most efficient.

5. CONCLUSIONS

In this work, the fault detection task for sensor precision degradation within the SPM framework has been involved. Instead of monitoring process variables themselves, the proposed method monitors the second-order statistics of process variables. It has been revealed in the literature that traditional PCA is sensitive to faults which cause mean shift rather than variance change. As analyzed, the characteristic of sensor precision degradation can just be reflected by the deviation (increase) of the faulty variable's variance. Consequently, though traditional PCA and DPCA are not efficient in monitoring the sensor precision

Table 2. Comparison of various fault detection methods

	PCA		DPCA		DISSIM	DISSIM	Proposal	
	T^2	SPE	T^2	SPE	$(w = 20)$	$(w = 40)$	D_p	D_r
FAR	0.9%	1%	1.2%	1.5%	1.1%	1.3%	0.4%	1.2%
FDR	2.7%	31.5%	3.5%	42%	25.5%	76.3%	**97.9%**	**95.3%**

degradation fault, monitoring the second-order statistics by PCA is effective. Besides, it is observed that DISSIM with a large window width can also exhibit satisfying detection performance, because the fault characteristic is indirectly reflected in its detection index. Simulation studies are carried out to demonstrate the effectiveness of the proposed method. In the future, several research directions based on the present work deserves further attention, such as the fault isolation problem especially when the faulty sensor is under closed-loop control.

REFERENCES

Alauddin, M., Khan, F., Imtiaz, S., and Ahmed, S. (2018). A bibliometric review and analysis of data-driven fault detection and diagnosis methods for process systems. *Industrial & Engineering Chemistry Research*, 45(7): 10719–10735.

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

Chen, Z., Ding, S., Peng, T., Yang, C., and Gui, W. (2018). Fault detection for non-Gaussian processes using generalized canonical correlation analysis and randomized algorithms. *IEEE Transactions on Industrial Informatics*, 65(2): 1559–1567.

Chen, Z., Yang, C., Peng, T., Dan, H., Li, C., and Gui, W. (2019). A cumulative canonical correlation analysis-based sensor precision degradation detection method. *IEEE Transactions on Industrial Informatics*, 66(8): 6321–6330.

Deng, X., Tian, X., Chen, S., and Harris, C. J. (2018). Nonlinear process fault diagnosis based on serial principal component analysis. *IEEE Transactions on Neural Networks and Learning Systems*, 29: 560–572.

Dunia, R., Qin, S. J., Edgar, T. F., and McAvoy, T. J. (1996). Identification of faulty sensors using principal component analysis. *AIChE Journal*, 42(10): 2797–2812.

Ge, Z., Song, Z., and Gao, F. (2013). Review of recent research on data-based process monitoring. *Industrial & Engineering Chemistry Research*, 52(10): 3543–3562.

Harmouche, J., Delpha, C., and Diallo, D. (2014). Incipient fault detection and diagnosis based on Kullback-Leibler divergence using principal component analysis: Part I. *Signal Process*, 94: 278–287.

He, Q. P., and Wang, J. (2011). Statistics pattern analysis: a new process monitoring framework and its application to semiconductor batch processes. *AIChE Jorunal*, 57(1): 107–121.

Ji, H., He, X., and Zhou, D. (2018). Diagnosis of sensor precision degradation using Kullback-Leibler divergence. *Canadian Journal of Chemical Engineering*, 96(2): 434–443.

Ji, H., Huang, K., and Zhou, D. (2019). Incipient sensor fault isolation based on augmented Mahalanobis distance. *Control Engineering Practice*, 86: 144–154.

Ji, H., and Zhou, D. (2020). Incipient fault detection of the high-speed train air brake system with a combined index. *Control Engineering Practice*, 100: 104425.

Jiang, Y., and Yin, S. (2018). Recursive total principle component regression based fault detection and its application to vehicular cyber-physical systems. *IEEE Transactions on Industrial Informatics*, 14(4): 1415–1423.

Kano, M., Hasebe, S., Hashimoto, I., and Ohno, H. (2002). Statistical process monitoring based on dissimilarity of process data. *AIChE Journal*, 48(6): 1231–1240.

Kourti, T. and MacGregor, J. (1995). Process analysis, monitoring and diagnosis, using multivariate projection methods. *Chemometrics and Intelligent Laboratory Systems*, 28(1): 3–21.

Li, G., Qin, S. J., Ji, Y., and Zhou, D. (2010). Reconstruction based fault prognosis for continuous processes. *Control Engineering Practice*, 18(10): 1211–1219.

Nomikos, P. and MacGregor, J. (1995). Multivariate SPC charts for monitoring batch processes. *Technometrics*, 37: 41–59.

Qin, S. J. (2012). Survey on data-driven industrial process monitoring and diagnosis. *Annual Reviews in Control*, 36(2): 220–234.

Shang, J., Chen, M., Ji, H., and Zhou, D. (2017). Recursive transformed component statistical analysis for incipient fault detection. *Automatica*, 80: 313–327.

Sharifi, R., and Langari, R. (2017). Nonlinear sensor fault diagnosis using mixture of probabilistic PCA models. *Mechanical Systems and Signal Processing*, 85: 638–650.

Valle, S., Li, W., and Qin S. J. (1999). Selection of the number of principal components: the variance of the reconstruction error criterion with a comparison to other methods. *Industrial & Engineering Chemistry Research*, 38: 4389–4401.

Wan, Y., and Ye, H. (2012). Data-driven diagnosis of sensor precision degradation in the presence of control. *Journal of Process Control*, 22(1): 26–40.

Wang, J., and He, Q. P. (2010). Multivariate statistical process monitoring based on statistics pattern analysis. *Industrial & Engineering Chemistry Research*, 49(17): 7858–7869.