

Fault detection based on ICA-GLR for non-Gaussian industrial processes

Chun-Chin Hsu^{1*}, Chun-Yuan Cheng¹, Fang-Chih Tien²

¹ Department of Industrial Engineering and Management, Chaoyang University of Technology, Taichung, Taiwan, R.O.C.

² Department of Industrial Engineering and Management, National Taipei University of Technology, Taipei, Taiwan, R.O.C.

ABSTRACT

As the growth of Industry 4.0, online fault detection plays a crucial role in ensuring the manufacturing quality. Generally, the fault detection methods can be classified into model-based and data-driven methods. There are advantages/disadvantages between two methods. In this study, we integrated both methods in order to develop an efficient fault detection method for non-Gaussian industrial processes. The data-driven method, independent component analysis (ICA) is used to extract non-Gaussian information and dimensionality reduction. Meanwhile, the model-based method, generalized likelihood ratio (GLR) test is adopted as the charting statistic. The proposed ICA-GLR method has advantages of 1) detecting a wide range of process changes, 2) estimating the change points and 3) needless prior parameters to be specified by practitioner. The efficiency of the proposed ICA-GLR fault detection method will be verified via implementing one simulated non-Gaussian process and two real manufacturing processes: Tennessee Eastman process and semiconductor manufacturing process. Results demonstrate that the proposed ICA-GLR method has superior fault detectability when compared to traditional methods, such as principal component analysis and ICA.

Keywords: Fault detection, Principal component analysis, Independent component analysis, Generalized likelihood ratio test, Non-Gaussian.

OPEN ACCESS 

Received: December 17, 2020

Revised: February 23, 2021

Accepted: March 23, 2021

Corresponding Author:

Chun-Chin Hsu

cchsu@cyut.edu.tw

 **Copyright:** The Author(s). This is an open access article distributed under the terms of the [Creative Commons Attribution License \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted distribution provided the original author and source are cited.

Publisher:

[Chaoyang University of Technology](https://www.chaoyang.edu.tw/)

ISSN: 1727-2394 (Print)

ISSN: 1727-7841 (Online)

1. INTRODUCTION

Fault detection is becoming incrementally critical in recent years to ensure the product quality in industrial processes. Statistical process control (SPC) is a graphical management tool that can help practitioners visually distinguish between common causes and special causes. Traditionally, the Shewhart \bar{X} chart was used to detect process mean shift. However, the \bar{X} chart is restricted to detect a large process mean shift (i.e. larger than 1.5σ (Montgomery, 2012)). It implies that \bar{X} chart behaviors poorly for a small process mean shift (i.e. smaller than 1.5σ). This is because that \bar{X} chart only takes the latest observation into consideration. Thus, the CUMulative SUM (CUSUM) (Page, 1954) and Exponentially Weighted Moving Average (EWMA) (Robert, 1959) charts were developed to detect small process shifts. However, both CUSUM and EWMA cannot effectively detect large process mean shift. It is intuitional that if a developed chart that can be capable of detecting a wide range of process shift sizes will become an attractive factor for practitioners to apply. Thus, several researchers strived themselves for developing charts in order to meet this goal. One means is to combine several charts. For example, the combination of Shewhart \bar{X} chart and CUSUM chart (Shewhart-CUSUM). The other option is to develop adaptive charts that can pre-predict shift size and use the estimated value to appropriately adjust chart parameters. Although both methods were shown to be effective for detecting a wide range of process shifts, several charting parameters require practitioners to specify before applying control

charts. Recently, the generalized likelihood ratio (GLR) test statistic was utilized to develop control chart. The GLR has been shown to be effective to detect a wide range of process mean shift sizes (Reynolds and Lou, 2010). Furthermore, it is not necessary to adjust control chart parameters during the monitoring period.

As the increasing growth of information technology, it is accessible to collect data simultaneously from several variables. Thus, the data-driven fault detection method is becoming popular to monitor a high-dimensional industrial process. The most well-known data-driven fault detection method is principal component analysis (PCA). PCA can project high dimensional process variables onto a lower dimensional space that contains the most variance of the original data. The original PCA was extended to several methods in order to cope with real process properties, such as dynamic PCA (DPCA) developed by Ku et al. (1995); dynamic PCA with decorrelated residuals (DPCA-DR) developed by Rato and Reis (2013); recursive PCA with a forgetting parameter (RPCA) developed by Li et al. (2000); moving window PCA (MWPCA) developed by Wang et al. (2005). The DPCA and DPCA-DR are intended to monitor autocorrelated industrial processes. The RPCA and MWPCA were introduced to monitor non-stationary industrial processes. Even though the PCA has been successfully implemented in process monitoring, the assumption for extracted PCA components should follow the Gaussian distribution. Martin and Morris (1996) reported that many real industrial processes exhibit PCA extracted variables are rarely conformed to a multivariate Gaussian distribution.

Independent component analysis (ICA) is another dimension reduction tool which can be seen as an extension of PCA. However, it possesses different statistical meaning between PCA and ICA. PCA considers only up to the second ordered statistics (i.e. mean and standard deviation) which intends to decorrelate components, whereas the ICA considers higher ordered statistics which exploits information about cumulants and moments of order greater than two (Lee et al., 2006). Thus, ICA can extract latent variables to be non-Gaussian and mutually independent. Kano et al. (2003) first developed an ICA based SPC monitoring method and showed the superiority over the PCA based SPC. Yoo et al. (2004) proposed a multiway ICA method for batch processing monitoring. Lee et al. (2004) presented a dynamic ICA (DICA) monitoring scheme. In which, ICA was applied to the augmented matrix with time lagged variables. Lee et al. (2006) proposed a modified ICA to relax the drawbacks of original ICA algorithm such as the pre-determination of number of extracted independent components and the pre-determination of the proper order of independent components. Ge and Song (2007) proposed PCA-ICA algorithm to extract Gaussian and non-Gaussian information for fault detection and diagnosis. González and Sánchez (2008) used principal alarms based on ICA to predict the mean shift for process monitoring. Lee et al. (2007) developed a kernel ICA algorithm for monitoring the

multivariate non-linear process. Lu et al. (2008) applied ICA for integrating SPC and engineering process control (EPC). Zhu et al. (2008) showed that ICA outperforms wavelet analysis for cutting force denoising in micro-milling tool condition monitoring.

Early detection of process fault is crucial not only to maintain stable process operation, but also to ensure peoples' safety. Therefore, this study is dedicated to developing an efficient fault detection model for empirical industrial processes. In this study, a fault detection method based on integrating data-driven (i.e. ICA) and model-based (i.e. GLR) methods will be proposed, namely ICA-GLR. The ICA is used to extract non-Gaussian information as well as reducing the data dimensionality. The GLR is used as the charting statistic. The proposed ICA-GLR fault detection method has advantages of 1) being capable of detecting a wide range of process changes, 2) estimating the change points and 3) needless prior parameters to be specified by practitioner. The efficiency of the proposed ICA-GLR will be verified via implementing a simulated non-Gaussian process and two real industrial process data sets, including tennessee eastman process and semiconductor manufacturing process. Results demonstrated the proposed ICA-GLR fault detection method is superior to PCA and ICA based monitoring methods in terms of fault detectability.

The remainder of this paper is organized as follows. A literature review of data-driven and model-based fault detection methods are provided in section 2. The proposed ICA-GLR method is presented in section 3. Three examples are implemented in section 4. Finally, the conclusion is addressed in section 5.

2. LITERATURE REVIEW

In this section, the fault detection based on data-driven methods will be first reviewed which includes principal component analysis (PCA) and independent component analysis (ICA). After that, the model based method, generalized likelihood ratio (GLR) will be provided as well.

2.1 PCA-based Monitoring Method

PCA is a dimensionality reduction technique that can capture the most variability information from the original data set. PCA has been widely used in the fields of pattern recognition, image processing, data compression and process monitoring.

The PCA-based monitoring method involves two phases: the offline model building and online monitoring as described below:

Offline model building

- Step 1. Collect a $p \times n$ data matrix X , for p process variables with n observations. Normalize X to have zero mean and unit variance data matrix Z .

- Step 2. Performing singular value decomposition (SVD) on the covariance matrix $S = \frac{Z' \cdot Z}{n-1}$, then each p -dimensional vector z is transformed into a score vector $y = P'z$, where P is the loading matrix, containing columnwise the vectors of S . Furthermore, $A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$ contains the eigenvalues of S in a descending sequence.
- Step 3. By retaining the first k columns of P and k largest eigenvalues of A , the retained scores can be expressed as $y_k = P'_k z$. The available methods to select the number of retained components include cumulative percentage of variance (CPV), cross-validation and parallel analysis, and they can be referred to Valle et al. (1999); Krzanowski and Kline (1995); Horn (1965), respectively.
- Step 4. Determine the control limits for Hotelling's T^2 and Q (i.e. also referred to squared prediction error (SPE)) charting statistics. The 100(1- α)% control limit for T^2 is expressed as

$$UCL_{T^2} = \frac{k(n^2 - 1)}{n(n - k)} F_{\alpha, k, n-k} \quad (1)$$

And, the 100(1- α)% control limit for Q is defined as

$$UCL_Q = \theta_1 \left(\frac{c_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right)^{1/h_0} \quad (2)$$

where c_α is the confidence interval that corresponds to the 1- α percentile of the normal distribution and $\theta_i = \sum_{j=k+1}^m \lambda_j^2$

with $i = 1, 2, 3$ and $h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2}$.

Online monitoring

Step 1. Normalize the new secured data set, denoted as Z_{new} .

Step 2. Calculate Hotelling's T^2 and Q charting statistics

$$T^2 = z'_{new} P_k A_k^{-1} P'_k z_{new} \quad (3)$$

$$Q = z'_{new} (I - P_k P'_k) z_{new}$$

Step 3. If $T^2 \geq UCL_{T^2}$ or $Q \geq UCL_Q$, it indicates a special cause may exist in the process. Hence, the correction actions should be taken in a bid to bring the process back to the stable circumstance.

2.2 ICA-based Monitoring Method

Collect a $p \times n$ data matrix X , for p process variables with n observations (in contrast to PCA, ICA employs the transposed data matrix). Given S to be the independent component matrix, and A is the unknown mixing matrix.

The relationship between original data matrix and independent components can be expressed as $X = AS$. The objective of ICA is to find a de-mixing matrix W such that the reconstructed signal $\hat{S} = WX$ becomes as independent as possible. Hyvärinen (1999) proposed a fixed-point algorithm for ICA, namely FastICA to secure the de-mixing matrix W :

Step 1. Randomly choose an initial weight vector w_i with unit norm.

Step 2. Let $w_i \leftarrow E\{xg(w^T x)\} - E\{g'(w^T x)\}w$, where g is the first derivative and g' is the second derivative of G in which

$$G(u) = \frac{1}{a_1} \log \cosh(a_1 u), \text{ and } a_1 \text{ is a constant}$$

and $1 \leq a_1 \leq 2$.

Step 3. Normalize $w_i \leftarrow \frac{w_i}{\|w_i\|}$.

Step 4. If w_i has not converged, go back to Step 2.

To divide W into two parts: the dominant part (W_d) and the excluded part (W_e). The T^2 charting statistic is defined as,

$$T^2 = \hat{S}_d^T \hat{S}_d \quad (4)$$

where $\hat{S}_d = W_d X$.

Also, the Q charting statistic can be obtained from:

$$Q = e^T e = (X - \hat{X})^T (X - \hat{X}) \quad (5)$$

where $e(k)$ is the residual at sample k and the predictor $\hat{X} = A\hat{S} = AWX$.

For process monitoring, the kernel density estimation (KDE) is applied to determine the control limits. A univariate kernel estimator with kernel K is defined by

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n k\left\{\frac{x - x_i}{h}\right\} \quad (6)$$

where x is the considered data point, x_i is the observation, h is the smoothing parameter, n is the number of samples and K is the kernel function. There are several kernel functions adopted in the literature in which the Gaussian kernel is the most popular one (Chen et al., 2000; Chen et al., 2004; Silverman, 1986).

2.3 GLR-based Monitoring Method

The Fisher's likelihood ratio test statistic is defined as:

$$\lambda = \frac{L(\theta_0)}{L(\hat{\theta})} \quad (7)$$

where $\hat{\theta}$ is the maximum likelihood estimator (MLE) of θ . Fisher inferred that if θ differs from θ_0 , then the value of the likelihood L when $\theta = \hat{\theta}$ will be larger than when $\theta = \theta_0$. Thus, the rejection region for the test contains values of λ that are smaller than some value λ_R .

Considering n observations $\{x_1, x_2, \dots, x_n\}$ sampled from normally operated process (without fault) and its population follows the normal distribution with mean μ_0 and

variance σ_0^2 . The likelihood function at sample n can be represented as:

$$L(\infty, \mu_0 | x_1, x_2, \dots, x_n) = (2\pi\sigma_0^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma_0^2} \sum_{i=1}^n (x_i - \mu_0)^2\right) \tag{8}$$

the process mean shifts to some value $\mu_1 \neq \mu_0$ at some time τ^* between τ and $\tau + 1$. The likelihood function at sample n can be represented as:

$$L(\tau, \mu_1 | x_1, x_2, \dots, x_n) = (2\pi\sigma_0^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma_0^2} \sum_{i=1}^{\tau} (x_i - \mu_0)^2 - \sum_{i=\tau+1}^n (x_i - \mu_1)^2\right) \tag{9}$$

The log likelihood ratio statistic at sample n is

$$R_n = \ln \frac{\max_{0 \leq \tau \leq n} L(\tau, \mu_1 | x_1, x_2, \dots, x_n)}{L(\infty, \mu_0 | x_1, x_2, \dots, x_n)} = \max_{0 \leq \tau \leq n} \frac{(\hat{\mu}_{1,\tau,n} - \mu_0)}{\sigma_0^2} \left\{ \sum_{i=\tau+1}^n (x_i - \mu_0) - \frac{1}{2} (\hat{\mu}_{1,\tau,n} - \mu_0) \right\} \tag{10}$$

To substitute $\hat{\mu}_{1,\tau,n} = \frac{\sum_{i=\tau+1}^n x_i}{n - \tau}$, the maximum likelihood estimator (MLE) in Equation (9). The R_n can be reduced to

$$R_n = \max_{0 \leq \tau \leq n} \frac{(n - \tau)}{2\sigma_0^2} (\hat{\mu}_{1,\tau,n} - \mu_0) \tag{11}$$

To reduce the computational complexity, only the recent past m samples (i.e. window size) are used to seek the maximum value (Reynolds and Lou, 2010). Let $\hat{\tau}$ be the estimation of the process change-point at which the maximum value has been reached. The GLR charting statistic can be then expressed as:

$$R_{m,n} = \frac{(n - \hat{\tau})}{2\sigma_0^2} (\hat{\mu}_{1,\hat{\tau},n} - \mu_0) \tag{12}$$

A signal is given at sample n if $R_{m,n} > h_{GLR}$.

An alternative way of GLR chart is to calculate the charting statistic $R'_{m,n}$ which has the advantage of showing the difference between $\hat{\mu}_1$ and μ_0 on the plot, that is

$$R'_{m,n} = \sqrt{2R_{m,n}} \sigma_0 = \sqrt{n - \hat{\tau}_n} (\hat{\mu}_{1,\hat{\tau},n} - \mu_0) \tag{13}$$

If $R'_{m,n}$ falls outside the interval of $[-h'_{GLR}\sigma_0, h'_{GLR}\sigma_0] = [-\sqrt{2h_{GLR}}\sigma_0, \sqrt{2h_{GLR}}\sigma_0]$, then a signal will be triggered on the plot.

3. PROPOSED ICA-GLR FAULT DETECTION METHOD

In this section, the proposed ICA-GLR fault detection method will be proposed and the main structure is drawn in Fig. 1. The proposed method constitutes two phases: off-line model building and on-line fault detection. The detailed procedure is presented as follows:

Phase I. Off-line model building

- Step 1. Acquire a $p \times n$ normally operated data matrix \mathbf{X} , for p process variables with n observations.
- Step 2. Standardize the data matrix to have a zero mean and unit standard deviation, expressed as \mathbf{Z} .
- Step 3. Perform FastICA algorithm to obtain the demixing matrix \mathbf{W} . The Matlab code of FastICA can be downloaded from <http://www.cis.hut.fi/projects/ica/fastica/>.
- Step 4. Calculate the euclidean norm (L_2) to each row of \mathbf{W} and then sort the rows by descending order. After that, give a threshold value to select k dominant ICA components by

$$\frac{\sum_{i=1}^k \mathbf{w}_i}{\sum_{i=1}^p \mathbf{w}_i} \geq \text{threshold} \tag{14}$$

where \mathbf{w}_i is the i th sorted row of \mathbf{W} .

- Step 5. Perform the following GLR steps for T^2 and Q , respectively:

1. Given a widow size $m = 400$, which was recommended by Reynolds and Lou (2010).
2. Estimating the mean ($\hat{\mu}_0$) and standard deviation ($\hat{\sigma}_0$) for T^2 and Q , respectively.
3. With $m = 400$, we need to calculate 400 values of $\sqrt{t - \tau_t} (\hat{\mu}_{1,\tau,t} - \mu_0)$ at time t , where $\hat{\mu}_{1,\tau,t}$ represents the estimation of the mean shift.
4. Choosing the maximum absolute values among afore-calculated 400 values to determine the values of $\hat{\tau}_t$, $\hat{\mu}_{1,\tau,t}$ and R'_t .

5. According to the research of Reynolds and Lou (2010), there is a linear relationship between h_{GLR} and the log scale of false alarm rate as shown in Fig. 2. Thus, h_{GLR} can be formulated as

$$h_{GLR} = -0.87 + 1.12 \ln(\text{false alarm rate}) \tag{15}$$

Note that the value of h_{GLR} is the same for T^2 and Q .

6. The control limit for the charting statistic R'_t is $[-\sqrt{2h_{GLR}}\sigma_0, \sqrt{2h_{GLR}}\sigma_0]$. Note that the control limits for T^2 and Q are different due to σ_0 .

Phase II. On-line fault detection

- Step 1. Acquire an on-line new data set.
- Step 2. Perform the same standardization procedure from Phase I.
- Step 3. Calculate ICA-GLR's T^2 and Q charting statistics.
- Step 4. If ICA-GLR's T^2 or Q triggers a signal, then the practitioner should check up the fault and take rectifying measures in order to bring the process back to the stable circumstance. Meanwhile, the estimation of process change point will be provided for helping practitioner diagnose the process fault.

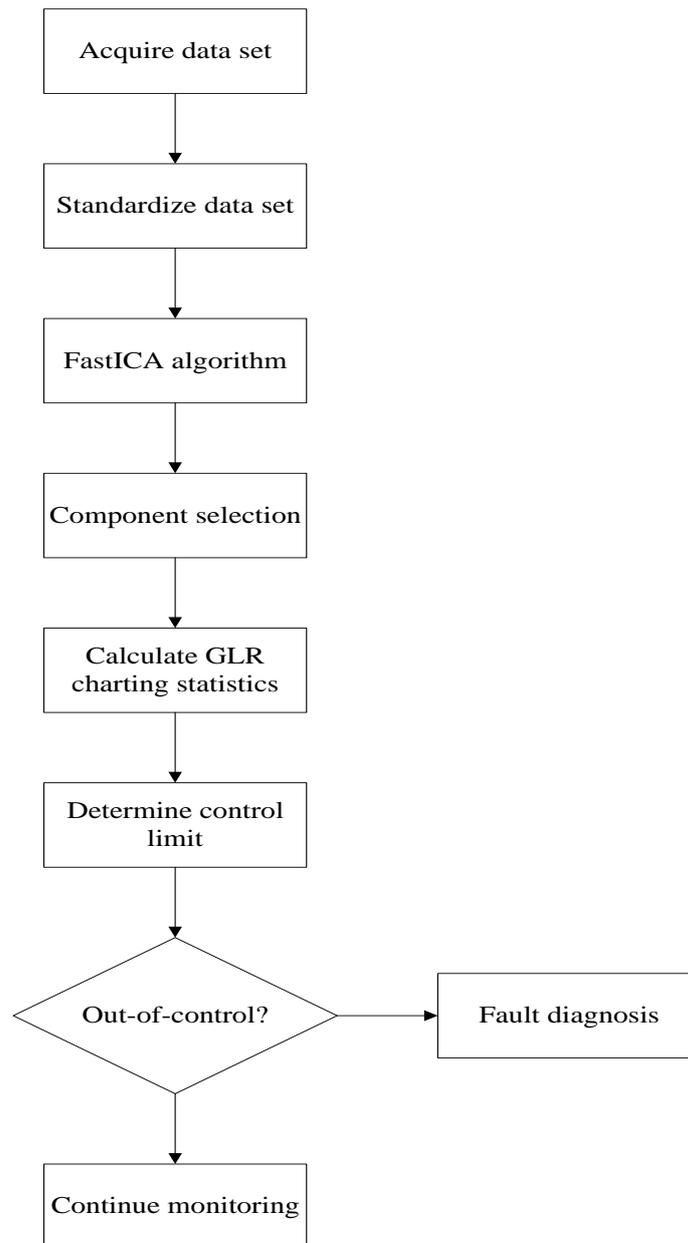


Fig. 1. The main structure of the proposed ICA-GLR fault detection method

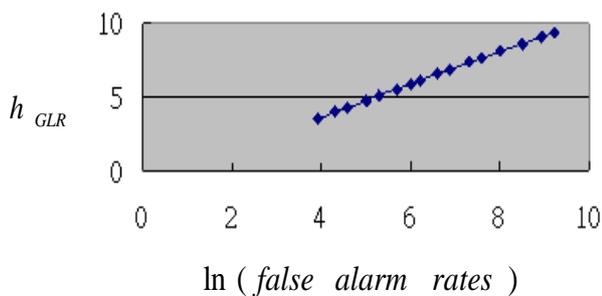


Fig. 2. h_{GLR} vs the log scale of false alarm rate

4. APPLICATIONS

In this section, the efficiency of the proposed ICA-GLR fault detection method will be verified via illustrating three examples: a simulated non-Gaussian process and two real industrial processes, including the tennessee eastman process and semiconductor manufacturing process. Among which, the PCA and ICA based monitoring methods will be used as the benchmark.

4.1 A simulated Non-Gaussian Process

In this section, a multivariate process suggested by Ku et al. (1995) will be given to implement the proposed method.

A dynamic process can be expressed as

$$\mathbf{r}(k) = \begin{bmatrix} 0.118 & -0.191 & 0.287 \\ 0.847 & 0.264 & 0.943 \\ -0.333 & 0.514 & -0.217 \end{bmatrix} \mathbf{r}(k-1) + \begin{bmatrix} 1 & 2 \\ 3 & -4 \\ -2 & 1 \end{bmatrix} \mathbf{u}(k-1)$$

$$\mathbf{g}(k) = \mathbf{r}(k) + \mathbf{v}(k) \tag{16}$$

where \mathbf{g} is the output, \mathbf{r} is the state and \mathbf{v} denotes the input which is assumed to be normal distributed with zero mean and variance of 0.1. The input \mathbf{u} is given by

$$\mathbf{u}(k) = \begin{bmatrix} 0.811 & -0.226 \\ 0.477 & 0.415 \end{bmatrix} \mathbf{u}(k-1) + \begin{bmatrix} 0.193 & 0.689 \\ -0.320 & -0.749 \end{bmatrix} \mathbf{h}(k-1) \tag{17}$$

The input \mathbf{h} is assumed to be uniformly distributed with random vector over interval (-2, 2). The five variables (g_1, g_2, g_3, u_1, u_2) are used to monitor the process.

500 uncontaminated observations are simulated as the historical dataset. Two faults were introduced into the process as follows:

Fault 1 (large shift): a step change of h_1 by 3 is introduced at observation 200 in 500 simulated runs

Fault 2 (small shift): a step change of h_2 by 1.3 is introduced at observation 200 in 500 simulated runs

All data were standardized prior to analysis. The parallel analysis is used to select the number of PCA components and the cumulative percentage of L_2 norm from the first few rows of \mathbf{W} up to 80% criterion is used for determining the number of ICA components. By analyzing the 500 uncontaminated observations, 3 components were retained for PCA and ICA. For a fair comparison, the empirical control limits for each methods were obtained by setting $ARL_0 = 1481.6$, the same setting as Reynolds and Lou (2010) and Reynolds and Stoumbos (2004a,b).

Fig. 3 shows the monitoring results for Fault 1. It demonstrates that PCA fails to detect process fault even if a large process shift. Further analyzing the normality by drawing Q-Q plot and density estimation of the second PCA component (Fig. 4), it shows the data biased from the normal distribution, prompting the poor monitoring result. ICA's T^2 chart can immediately detect the Fault 1 after 200th sample, but the ICA's Q chart does not perform well as ICA's T^2 chart. Related to the proposed ICA-GLR method, both T^2 and Q charts can effectively detect the Fault 1. The ICA-GLR's T^2 chart performs a bit better than Q chart, in which the T^2 chart can instantly detect the fault at sample 200 and the Q chart detects the fault at around sample 220. Fig. 5 shows the estimations of process change point for ICA-GLR charts. It demonstrates that the estimated change points converge at 200 and 220 for T^2 and Q , respectively.

It indicates that the proposed method can accurately predict the process change time which will help practitioner isolate the fault.

Fig. 6 shows the monitoring results for Fault 2. One can find both PCA and ICA based monitoring methods cannot effectively detect the small process change. In contrast, the proposed ICA-GLR's T^2 chart is capable of detecting the small process change after 210th sample, but ICA-GLR's Q fails to detect the change. According to Fig. 7(a), it indicates that the estimated process change converges to 208 under the implementation of ICA-GLR's T^2 chart, but we cannot estimate the change point under the implementation of ICA-GLR's Q chart due to the estimation values change with time (Fig. 7(b)). Through this example, one can find that the proposed ICA-GLR fault detection method can detect a wide range of process changes than traditional monitoring methods.

4.2 Tennessee Eastman Process

The tennessee eastman (TE) process has been widely used in fault detection studies, such as Zhou et al. (2016); Rato and Reis (2013); Ge and Song (2007); Lee et al. (2006); Lee et al. (2007); Lee et al. (2004b); Ku et al. (1995). The plant layout is shown in Fig. 8. The TE process consists five major operation units: reactor, product condenser, recycle compressor, vapor-liquid separator and product stripper. Two liquid products G and H are generated from the gaseous reactants A, B, C, E and the inert B. Two spinoffs, D and F are also produced.

There are a total of 52 variables recorded, including 41 measurements (XMEAS) and 11 manipulated (XMV) variables. A total of 21 faulty modes (Table 1) are introduced in the process for testing the detectability of the monitoring methods. Each dataset was collected at a sample interval of 3 minutes. Each faulty mode contains 960 observations, in which the fault is introduced after 160th observation (8th hour). The dataset can be downloaded from <http://web.mit.edu/braatzgroup>.

The dataset with no fault (IDV(0)), representing data generated under normally operated condition was used to train PCA, ICA and ICA-GLR methods. Before analysis, the dataset is auto-scaled. Parallel analysis is used to determine the number of PCA components and the cumulative percentage of L_2 norm from the first few rows of \mathbf{W} up to 80% criterion is used for determining the number of ICA components. Applying the data set of IDV(0), the number of retained components for PCA and ICA are 11 and 28, respectively. For a fair comparison, the ARL_0 is set to be 1481.6 for securing the empirical control limits for each method.

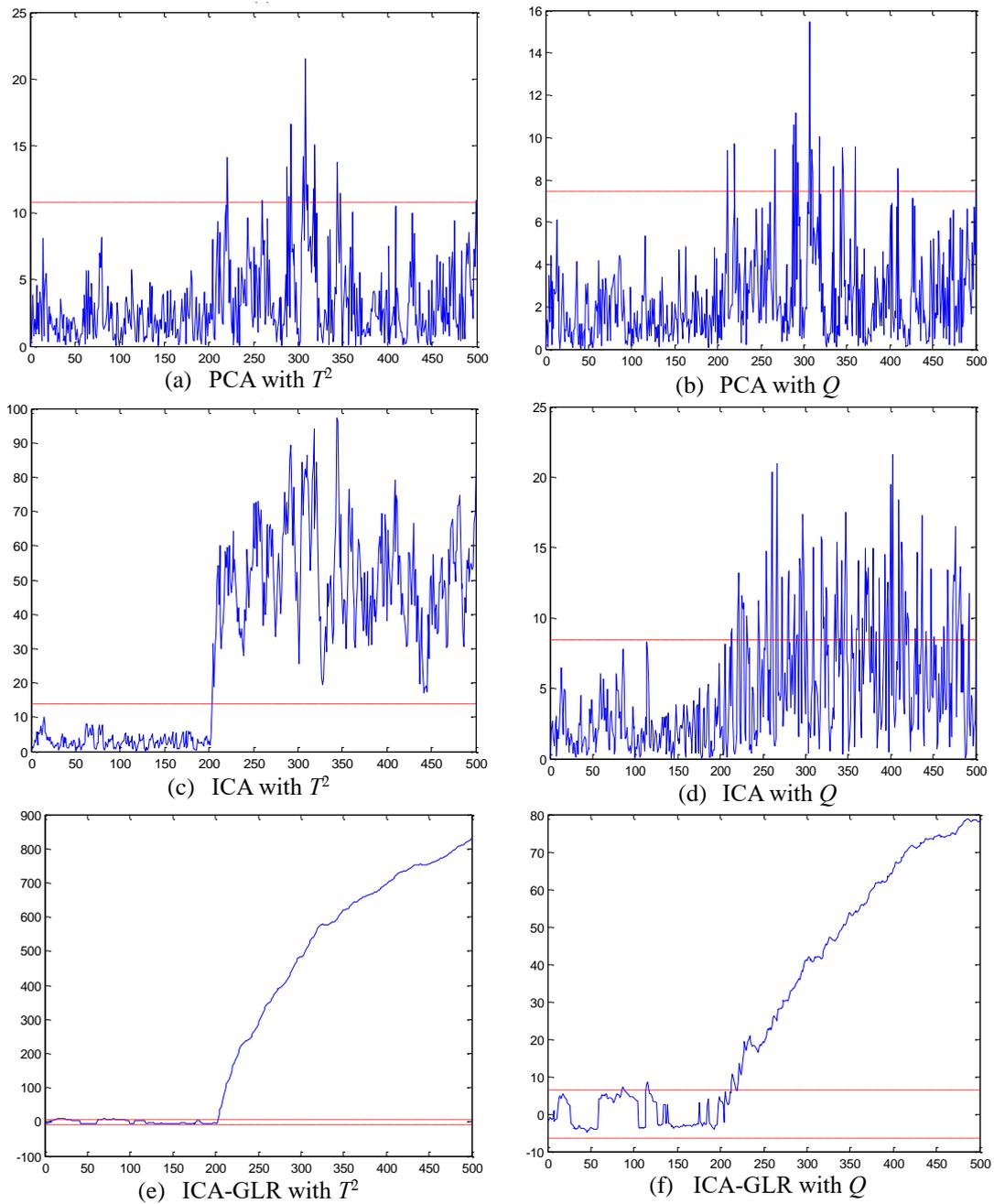


Fig. 3. Monitoring results for Fault 1

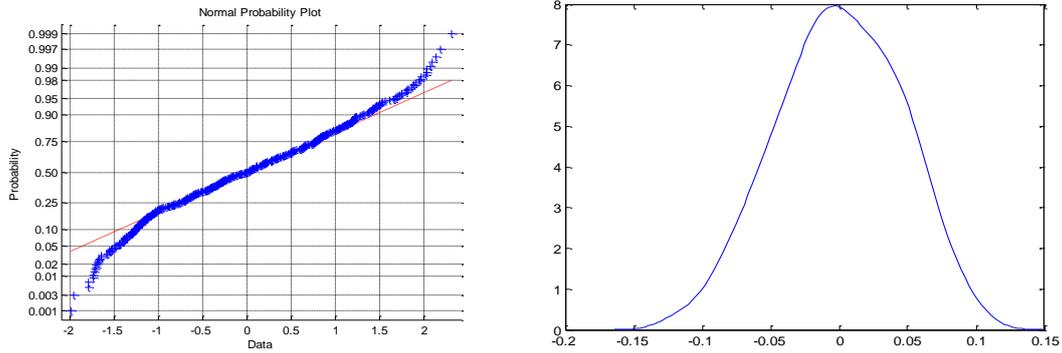


Fig. 4. Q-Q plot (left) and its density estimation plot (right) for the 2nd PCA component

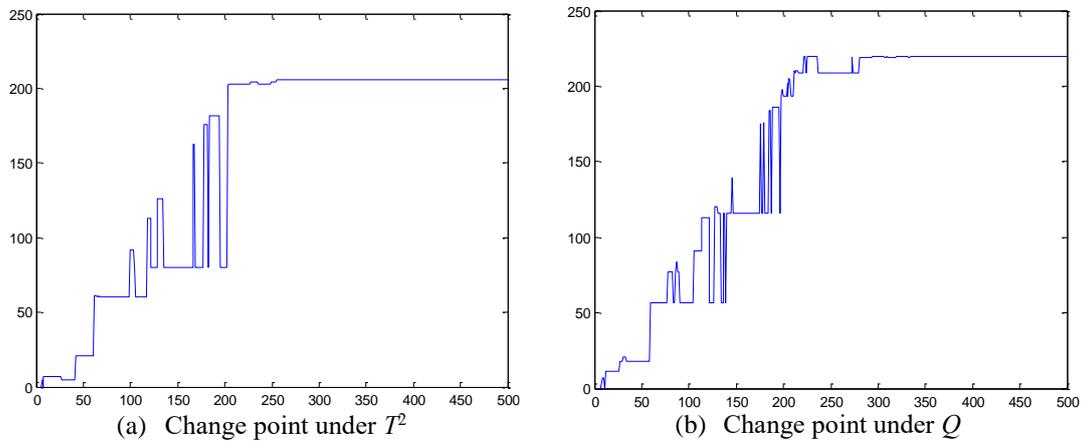


Fig. 5. Estimations of change point for Fault 1

Table 2 shows the fault detection rates for each monitoring method. The faults of number 3, 9 and 15 are excluded from analysis due to all methods are difficult to detect these faults. For an objective comparison between all monitoring methods, the paired-t test was conduct in Table 3. Obviously, the PCA possesses the lowest fault detection rates due to the data set comes from a non-Gaussian process. For ICA implementation, the ICA's Q chart performs better than ICA's T^2 chart, the same result concluded by Rato and

Reis (2013). Generally speaking, the ICA's Q chart can effectively detect faults when the process is non-Gaussian distributed. From Table 2, it show that the proposed ICA-GLR charts produce the highest fault detection rates. Further, Table 3 shows the proposed method significantly outperforming PCA and ICA monitoring methods. The ICA-GLR's T^2 chart performs a little better than ICA-GLR's Q chart, but it is not significantly different in terms of fault detectability.

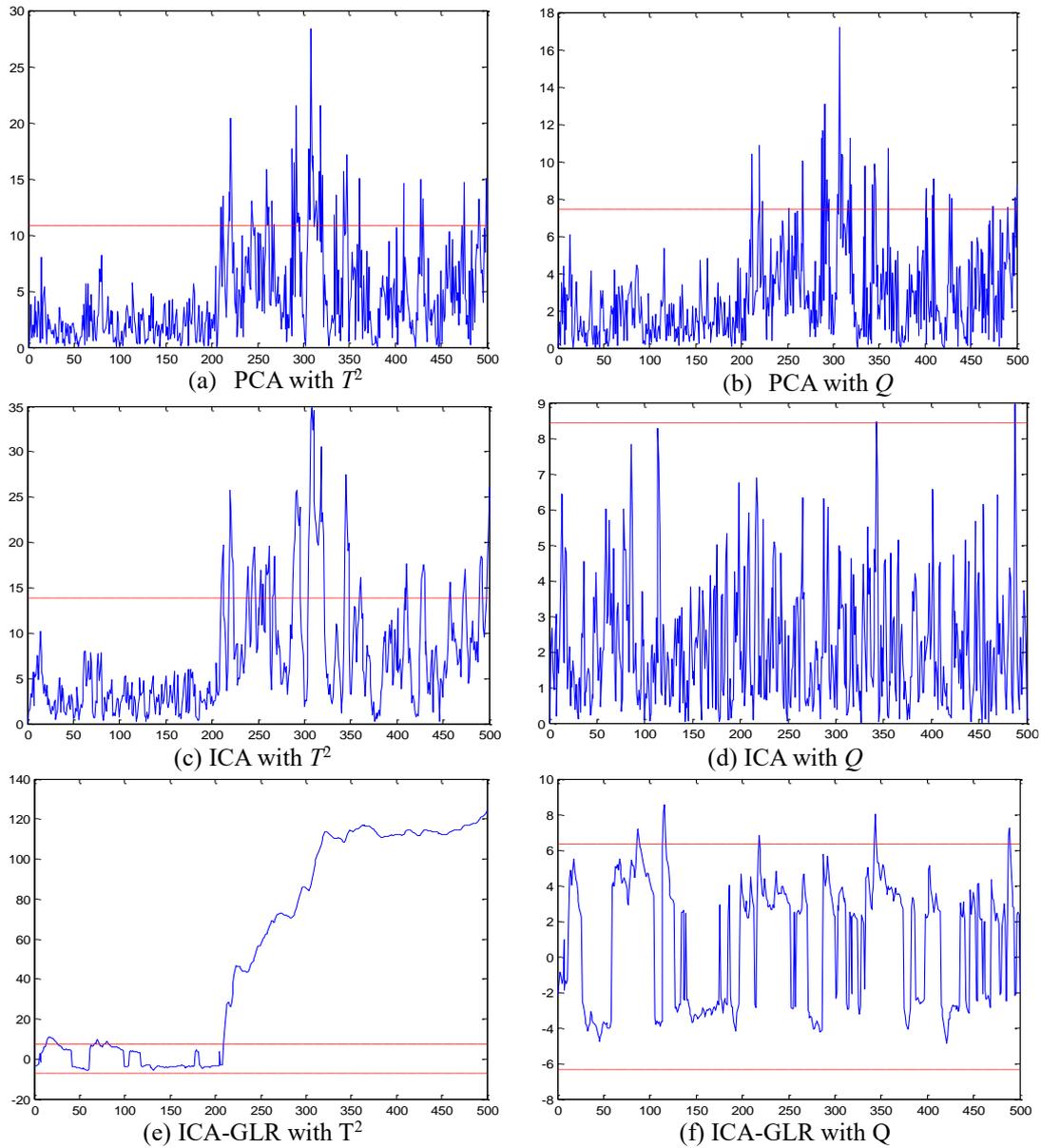


Fig. 6. Monitoring results for Fault 2

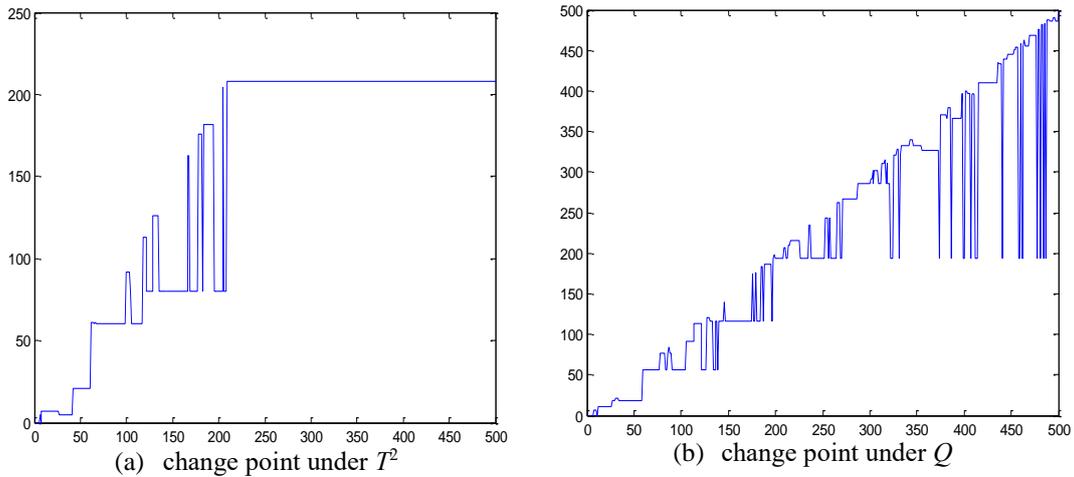


Fig. 7. Estimations of change point for Fault 2

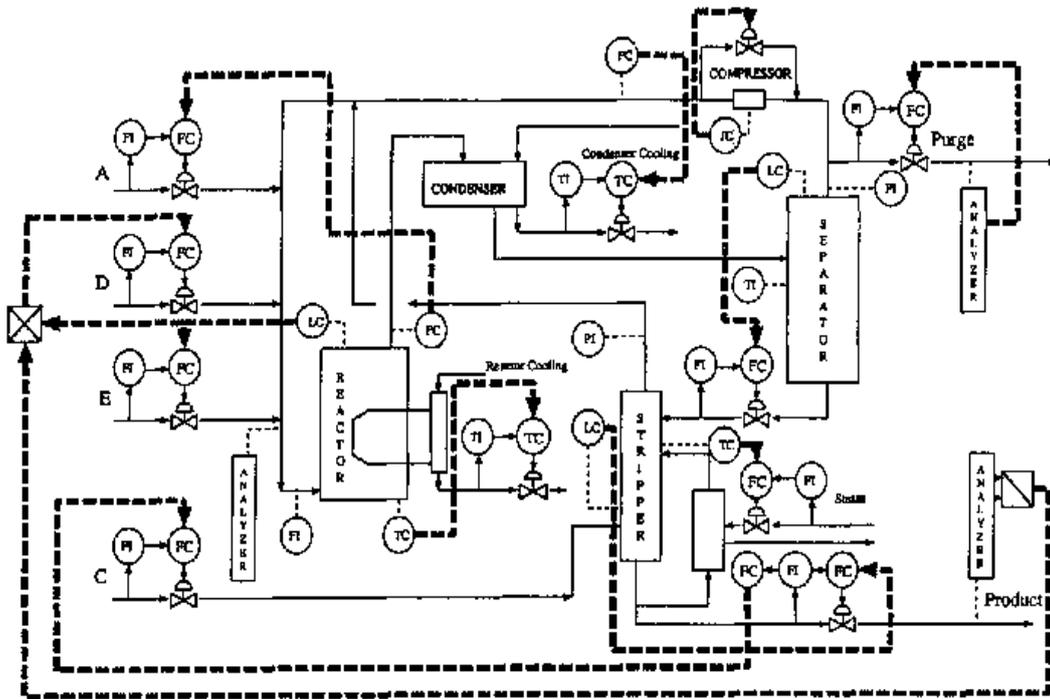


Fig. 8. Layout of TE process (Dong and McAvoy, 1996)

Fig. 9 shows control charts used to monitor Fault 10, a random variation of C feed temperature. The PCA's T^2 and Q charts detect the fault respectively at 232nd sample and 262nd sample, but the data points return within the control limits thereafter. Both ICA's T^2 and Q charts detect the fault at 186th sample and the data points can keep above the

control limit to the end. However, the signal is 26 latter than the real fault induced time, referring to a lag of 78 minutes to indicate the fault. In contrast, the ICA-GLR's T^2 chart can detect the fault at 165th sample, a lag of 15 minutes. Thus, the quick fault response can help practitioner rectify the process more immediately.

Table 1. Faulty modes

No.	State	Faulty type
IDV(1)	A/C Feed Ratio, B Composition Constant (Stream 4)	Step
IDV(2)	B Composition, A/C Ratio Constant (Stream 4)	Step
IDV(3)	D Feed Temperature (Stream 2)	Step
IDV(4)	Reactor Cooling Water Inlet Temperature	Step
IDV(5)	Condenser Cooling Water Inlet Temperature	Step
IDV(6)	A Feed Loss (Stream 1)	Step
IDV(7)	C Header Pressure Loss – Reduced Availability (Stream 4)	Step
IDV(8)	A, B, C Feed Composition (Stream 4)	Random Variation
IDV(9)	D Feed Temperature (Stream 2)	Random Variation
IDV(10)	C Feed Temperature (Stream 4)	Random Variation
IDV(11)	Reactor Cooling Water Inlet Temperature	Random Variation
IDV(12)	Condenser Cooling Water Inlet Temperature	Random Variation
IDV(13)	Reaction Kinetics	Slow Drift
IDV(14)	Reactor Cooling Water Valve	Sticking
IDV(15)	Condenser Cooling Water Valve	Sticking
IDV(16)	Unknown	
IDV(17)	Unknown	
IDV(18)	Unknown	
IDV(19)	Unknown	
IDV(20)	Unknown	
IDV(21)	The valve for stream 4 was fixed at steady state position	Constant position

Table 2. Fault detection rates for TE process

Fault	PCA		ICA		ICA-GLR	
	T^2	Q	T^2	Q	T^2	Q
1	0.9900	0.9938	0.995	0.9913	0.9963	0.9963
2	0.9788	0.9813	0.9813	0.980	0.9925	0.9813
4	0.0037	0.6242	0.0624	0.9638	0.9988	0.9988
5	0.206	0.2197	0.9975	0.9988	0.9988	0.9988
6	0.9875	0.9988	0.9988	0.9988	0.9988	0.9988
7	0.7366	0.9988	0.6542	0.9988	1	0.9988
8	0.9551	0.97	0.9713	0.9663	1	0.9738
10	0.201	0.0849	0.7378	0.7853	0.995	0.975
11	0.0674	0.4894	0.1136	0.7216	0.9863	0.9913
12	0.9513	0.9201	0.9938	0.9963	1	0.9963
13	0.8901	0.9463	0.9463	0.9413	0.9576	0.9551
14	0.6866	0.9988	0.9925	0.9988	0.9975	0.9988
16	0.0612	0.0549	0.7029	0.8215	1	0.9875
17	0.6392	0.8302	0.7253	0.9438	0.9725	0.975
18	0.8826	0.8901	0.8926	0.8964	0.9625	0.9026
19	0.0000	0.0125	0.0362	0.7815	0.9925	0.9975
20	0.2097	0.3171	0.7441	0.8227	0.9938	0.9164
21	0.171	0.3708	0.3933	0.3496	0.8739	0.9263

Table 3. The p -value of paired t test in terms of detection rates

Method A \ Method B		PCA		ICA		ICA-GLR	
		T^2	Q	T^2	Q	T^2	Q
PCA	T^2	--	0.018 (-)	0.008 (-)	0.000 (-)	0.000 (-)	0.000 (-)
	Q	0.018 (+)	--	0.423 (-)	0.006 (-)	0.002 (-)	0.002 (-)
ICA	T^2	0.008 (+)	0.423 (+)	--	0.025 (-)	0.004 (-)	0.006 (-)
	Q	0.000 (+)	0.006 (+)	0.025 (+)	--	0.008 (-)	0.021 (-)
ICA-GLR	T^2	0.000 (+)	0.002 (+)	0.004 (+)	0.008 (+)	--	0.215 (+)
	Q	0.000 (+)	0.002 (+)	0.006 (+)	0.021 (+)	0.215 (-)	--

Note: The “+” indicates the method A generates higher detection rates than method B. Value with bold font indicates statistically significant in mean between two methods (i.e. p -value is lower than 0.05)

4.3 Semiconductor Manufacturing Process

The semiconductor manufacturing includes six main steps: circuit design, wafer fabrication, probing, assembly, test and board assembly as shown in Fig. 10. Within the production cycle, there are several major check points for in house line testing to ensure product functionality. A data set which is publicly available in the UCI Irvine Machine Learning Repository (<http://archive.ics.uci.edu/ml/datasets/SECOM>) contains a total of 1,567 samples with 591 variables. Among them, 1,463 samples are in-control (IC) samples and the remaining 104 samples are out-control (OC) ones.

Before implementing the monitoring methods, the 117 variables from 591 were removed due to the constant values in that variable. Thus, total 474 variables are used for the following analysis and we denoted it as $\{X_1, X_2, \dots, X_{474}\}$. After that, when the variable contains missing values, the mean will fill the missing values as the same operation from Zhang et al. (2016). We use previous 500 samples from IC dataset to be the training data set. Related to testing data set, totaling 304 samples was drawn from the last 200 IC samples along with 104 OC samples. Fig. 11 shows the Q-Q plots and density estimation plots for X_{38} and X_{99} from the training dataset. The figure demonstrates the data set departs from the normal distribution.

The data set is normalized to have zero mean and unit variance and it is denoted as $\{Z_1, Z_2, \dots, Z_{474}\}$. Fig. 12 shows the scatter plots of paired normalized variables (i.e. Z_{15} vs Z_{146} and Z_{148} vs Z_{374}). It shows that it seems an uneasy

work to distinguish OC from IC due to the small process change. For this reason, the control limits for PCA and ICA were secured from tighten setting of $ARL_0 = 200$ in a bid to have a better fault detection chance. However, the ARL_0 for the proposed method is still set to be 1481.6 as recommended by Reynolds and Lou (2010). For component selection, the parallel analysis is used for PCA implementation and the cumulative percentage of L_2 norm from first few rows of W up to 80% criterion is used for ICA implementation. By analyzing the IC dataset, 63 components and 13 components were retained for PCA and ICA methods, respectively.

Fig. 13 shows the monitoring results for each method. It shows both PCA and ICA based monitoring charts fail to trigger the abnormal signal after 200th sample. For PCA, the charting statistics behave random fluctuation pattern within the control limits, indicating PCA cannot perform well when the data is non-Gaussian distributed. The ICA's T^2 shows a process shift pattern, but it cannot provide practitioner with the abnormal signal even the tighter control limit with $ARL_0 = 200$ was given. Furthermore, the ICA's Q abruptly shows a signal at around 260th sample, but it runs back within the control limit. In contrast, the ICA-GLR's both charting statistics immediately emerge signals after 200th sample and kept to the end, providing practitioner the correct process information. In terms of fault detection rate, the ICA-GLR's T^2 possesses 98% and the 99% for ICA-GLR's Q . Fig. 14 shows the estimation of the change point, it shows the estimated process change point is 200 which correctly reflects the real process scenario.

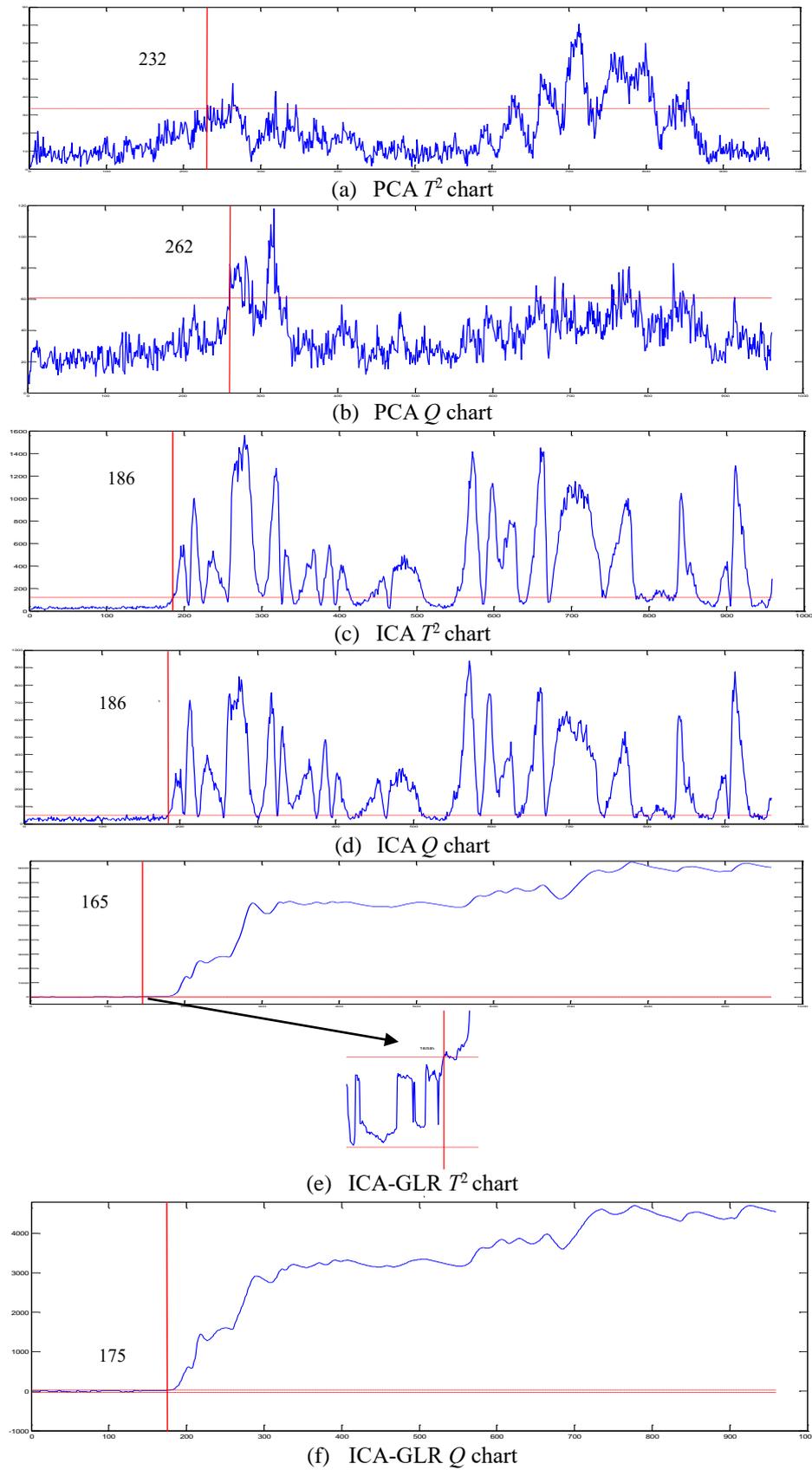


Fig. 9. Monitoring results for Fault 10

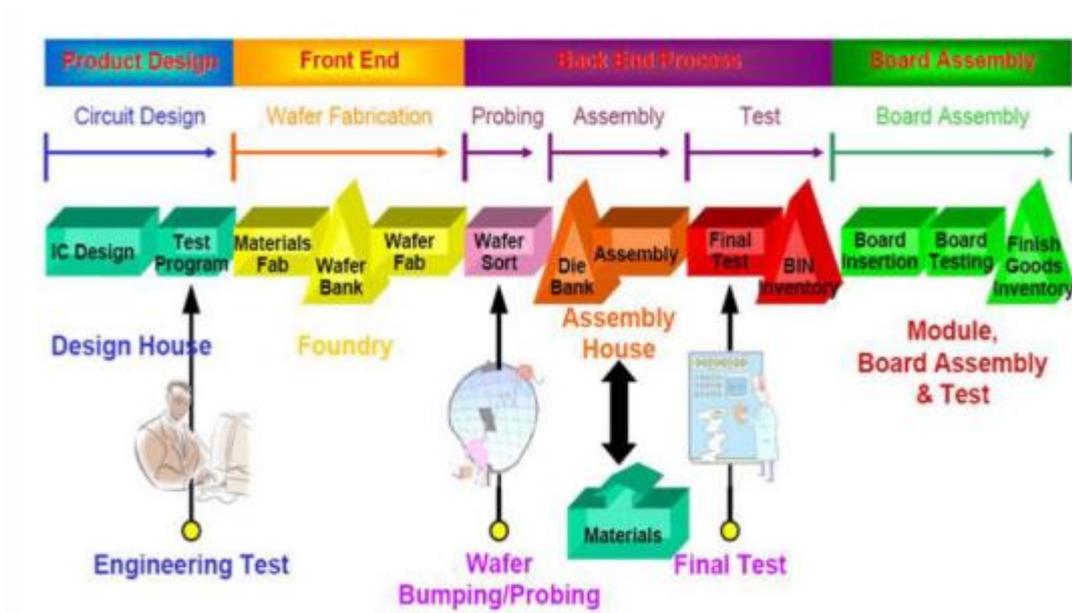


Fig. 10. Semiconductor manufacturing process (Munirathinam and Ramadoss, 2016)

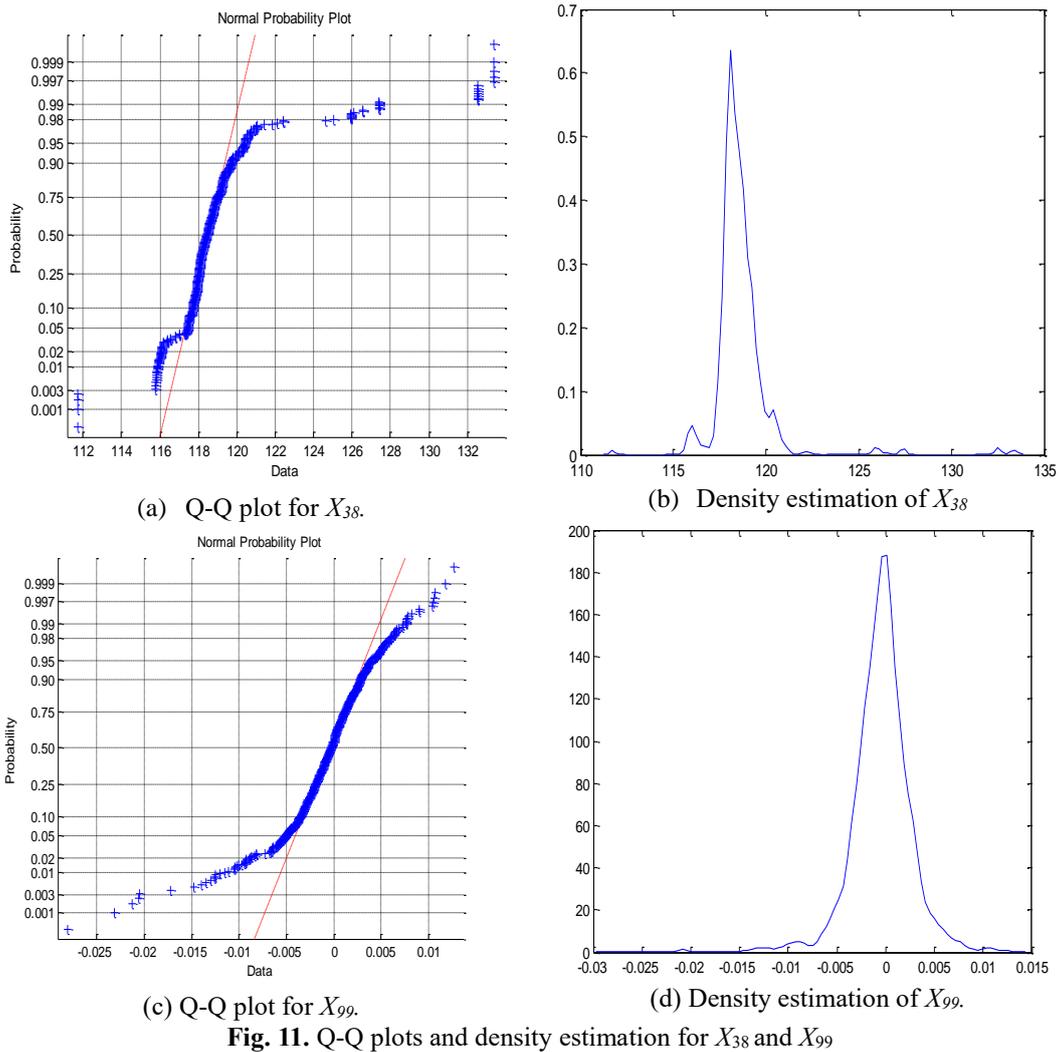


Fig. 11. Q-Q plots and density estimation for X_{38} and X_{99}

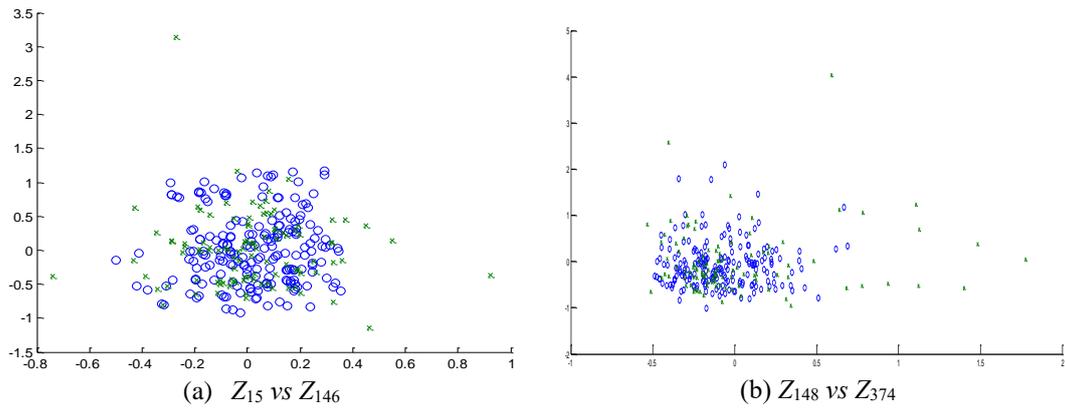


Fig. 12. Scatter plot for Z_{15} vs Z_{146} and Z_{148} vs Z_{374} . The blue circle signifies the IC samples, whereas the green cross represents the OC samples

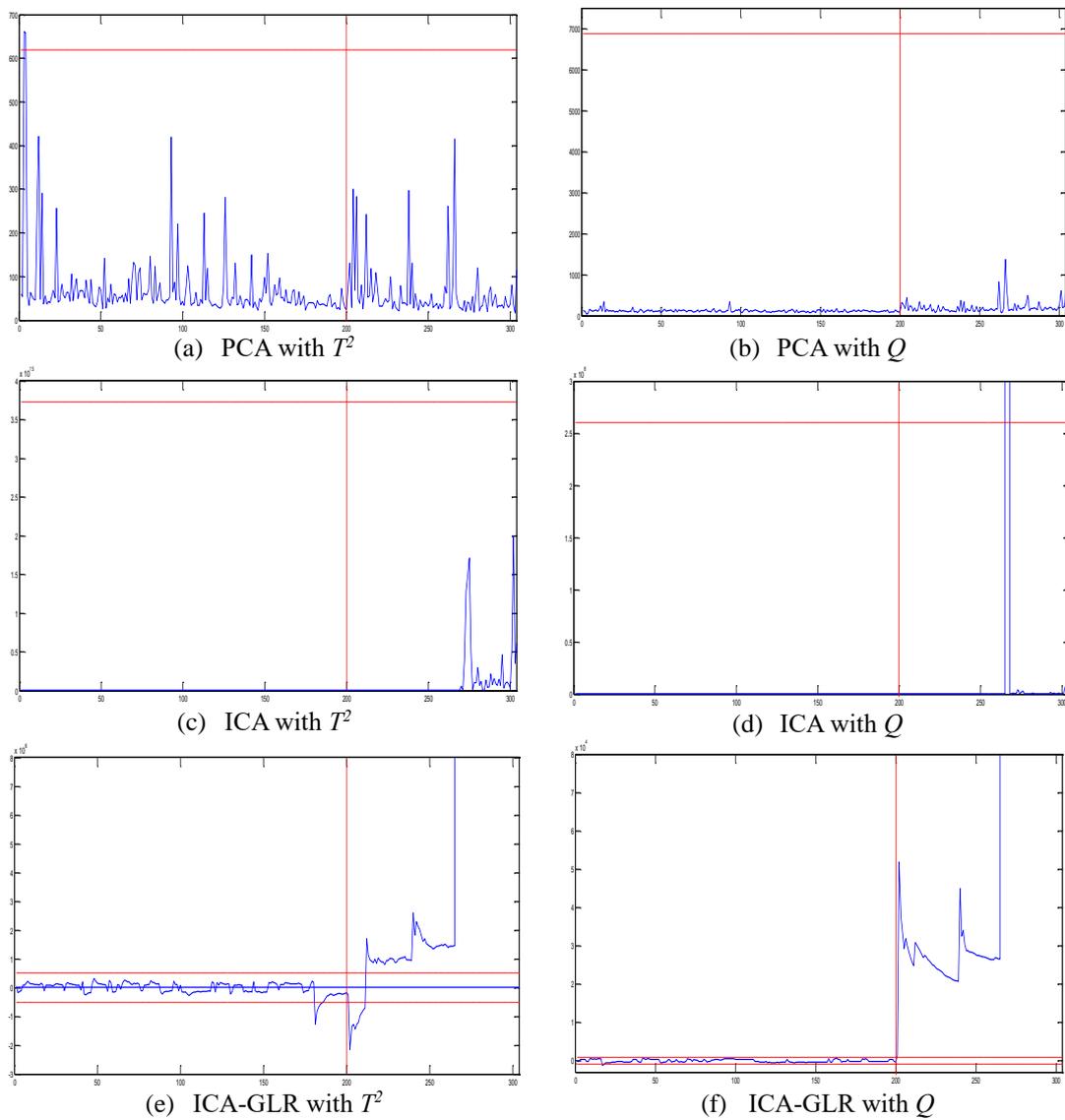


Fig. 13. Monitoring result of semiconductor manufacturing process

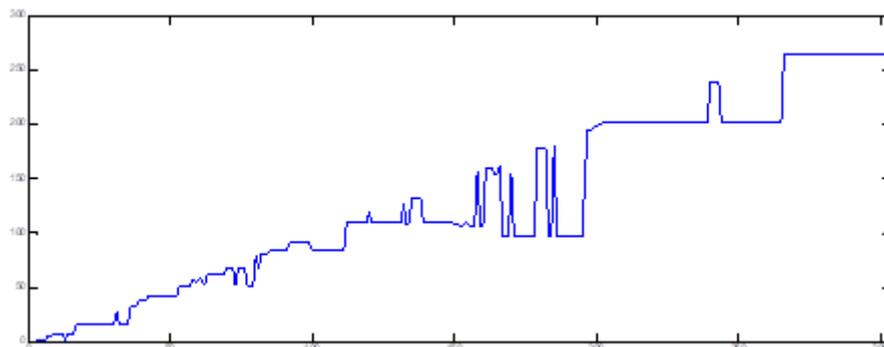


Fig. 14. Estimation of change point of Q charting statistic

5. CONCLUSION

This study proposed an ICA-GLR fault detection method for non-Gaussian industrial process monitoring. The efficiency of the proposed method has been verified via the implementation of three examples. Summarily, the advantages of the proposed method are: 1) it has a superior fault detectability than traditional monitoring methods; 2) it provides the estimation of the process change point which helps practitioners identify the root cause of the process fault; 3) unlike fault detection methods based on machine learning, the proposed method is free from the prior-specified parameters before implementation. Although the proposed method has been shown to have superior performances than traditional methods, the limitations of the ICA-GLR method lie in the lack of consideration of autocorrelation and nonlinearity of process data. Future works can further develop methods to overcome the above-mentioned limitations.

ACKNOWLEDGMENT

This work was supported in part by Ministry of Science and Technology, Taiwan. (Grant No. : MOST 105-2410-H-324 -006 -MY3)

REFERENCES

- Chen, Q., Kruger, U., Leung, A.T.Y. 2004. Regularised kernel density estimation for clustered process data. *Control Engineering Practice*, 12, 267–274.
- Chen, Q., Wynne, R.J., Goulding, P., Sandoz, D. 2000. The application of principal component analysis and kernel density estimation to enhance process monitoring. *Control Engineering Practice*, 8, 531–543.
- Ge, Z., Song, Z. 2007. Process monitoring based on independent component Analysis-Principal Component analysis (ICA-PCA) and similarity factors. *Industrial and Engineering Chemistry Research*, 46, 2054–2063.
- González, I., Sánchez, S. 2008. Principal alarms in multivariate statistical process control using independent component analysis. *International Journal of Production Research*, 46, 6345–6366.
- Horn, J. 1965. A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179–185.
- Hyvärinen, A. 1999. Fast and robust Fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 10, 626–634.
- Kano, M. Tanaka, S., Hasebe, S., Hashimoto, I., Ohno, H. 2003. Monitoring independent components for fault detection. *AIChE Journal*, 49, 969–976.
- Krzanowski, W.J., Kline, P. 1995. Cross-Validation for choosing the number of important components in principal component analysis. *Multivariate Behavioral Research*, 30, 149–165.
- Ku, W., Storer, R., Georgakakis, C. 1995. Disturbance detection and isolation by dynamic principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 30, 179–196.
- Lee, J.M., Qin, S.J., Lee I.B. 2006. Fault detection and diagnosis based on modified independent component analysis. *AIChE Journal*, 52, 3501–3514.
- Lee, J.M., Qin, S.J., Lee, I.B. 2007. Fault detection of Non-linear process using kernel independent component analysis. *The Canadian Journal of Chemical Engineering*, 85, 526–536.
- Lee, J.M., Yoo, C.K., Lee, I.B. 2004. Statistical monitoring of dynamic processes based on dynamic independent component analysis. *Chemical Engineering Science*, 59, 2995–3006.
- Li, W., Yue, H.H., Valle-Cervantes, S., Qin, S.J. 2000. Recursive PCA for adaptive process monitoring. *Journal of Process Control*, 10, 471–486.
- Lu, C.J., Wu, C.M., Keng, C.J., Chiu, C.C. 2008. Integrated application of SPC/EPC/ICA and neural networks. *International Journal of Production Research*, 46, 873–893.
- Martin, E.B., Morris, A.J. 1996. Non-parametric confidence bounds for process performance monitoring charts. *Journal of Process Control*, 6, 349–358.
- Montgomery, D.C. 2012. *Introduction to statistical quality control*, 7th edition, New York, NY: Wiley.

- Munirathinam, S., Ramadoss, B. 2016. Predictive models for equipment fault detection in the semiconductor manufacturing process. *IACSIT International Journal of Engineering and Technology*, 8, 273–285
- Page, E.S. 1954. Continuous inspection schemes. *biometrics*, 41, 100–115.
- Rato, T.J., Reis, M.S. 2013. Fault detection in the Tennessee Eastman benchmark process using principal component analysis based on decorrelated residuals (DPCA-DR). *Chemometrics and Intelligent Laboratory Systems*, 125, 101–108.
- Reynolds, M.R., Jr., Lou, J. 2010. An evaluation of a GLR control chart for monitoring the process mean. *Journal of Quality Technology*, 42, 287–310.
- Reynolds, M.R., Jr., Stumbos, Z.G. 2004a. Control charts and the efficient allocation of sampling resources. *Technometrics*, 46, 200–214.
- Reynolds, M.R., Jr., Stumbos, Z.G. 2004b. Should observations be grouped for effective process monitoring? *Journal of Quality Technology*. 36, 343–366.
- Robert, S.W. 1959. Control chart tests based on geometric moving averages. *Technometrics* 1, 239–250.
- Silverman, B.W., 1986. *Density estimation for statistics and data analysis*, UK: Chapman & Hall.
- Valle, S., Li, W., Qi, S.J. 1999. Selection of the number of principal components: The variance of the reconstruction error criterion with a comparison to other methods. *Industrial and Engineering Chemistry Research*, 38, 4379–4401.
- Wang, X., Kruger, U., Irwin, G.W. 2005. Process monitoring approach using fast moving window PCA. *Industrial & Engineering Chemistry Research*, 44, 5691–5702.
- Yoo, C.K., Lee, J.M., Vanrolleghem, P.A., Lee, I.B. 2004. On-line monitoring of batch processes using multiway independent component analysis. *Chemometrics and Intelligent Laboratory Systems*, 71, 151–163.
- Zhang, C., Chen, N., Zou, C. 2016. Robust multivariate control chart based on Goodness-of-Fit test. *Journal of Quality Technology*. 48, 139–161.
- Zhou, Z., Wen, C., Yang, C. 2016. Fault isolation based on k-Nearest neighbor rule for industrial processes. *IEEE Transactions on Industrial Electronics*, 63, 2578–2586.
- Zhu, K., Hong, G.S., Wong, Y.S., Wang, W. 2008. Cutting force denoising in Micro-milling tool condition monitoring. *International Journal of Production Research*, 46, 4391–4408.