Performance Evaluation of Independent Component Analysis-Based Fault Detection Using Measurements Corrupted with Noise

K. Ramakrishna Kini¹ · Muddu Madakyaru²

Received: 22 June 2020 / Revised: 18 January 2021 / Accepted: 8 February 2021 / Published online: 22 March 2021
© The Author(s) 2021

Abstract

The detection of sensor faults has proven to be easier through data-driven methods which rely on historical data collected from sensors that are placed at various locations in a process plant. Since the distribution of industrial process variables is random and non-Gaussian, the independent component analysis (ICA) method has been better suited for fault detection (FD) problems. Whenever data comes with any level of noise, there is difficulty in separating useful information, which hence degrades the monitoring quality of an FD strategy. In this paper, the robustness of FD strategies is assessed for different noise realizations of sensor data using stochastic simulations. The main objective of this work is to demonstrate that ICA-based FD strategies are more robust for different noise levels in comparison with principal component analysis (PCA). The ICA modeling algorithm is improved to avoid random initialization of a de-mixing orthogonal matrix during computation of independent components. Two case studies are considered for evaluating the robustness of FD strategies: a simulated quadruple tank process and a simulated distillation column process. Comparisons have been carried out between ICA, dynamic ICA, modified ICA and PCA strategies for different sensor noise levels. The simulation results reveal that ICA-based FD strategies over-perform PCA FD strategy in monitoring sensor faults for different levels of noise.

Keywords Independent component analysis · Robust fault detection · Distillation column process · Quadruple tank process · Stochastic simulation

1 Introduction

An active field of research in the process control domain, process monitoring-fault detection (PM-FD) aims to quickly detect sensor faults in chemical process plants (Venkatasubramanian et al. 2003; Ramakrishna Kini and Madakyaru 2020). Until recently, the task of detecting process deviations was dependent on human supervision, which eventually lead to more cost due to the increased complexity in modern process plants. To deal with process deviations or faults, model-based and process history methods have been applied

in practice. The model-based PM-FD methods rely on comparing information of measured variables with an explicit known mathematical model of the process. A few commonly used methods include parity relations, expert systems and Kalman filters (Sanjula and Zukui 2019). In contrast, process history PM-FD schemes rely purely on available historical data. Owing to the tremendous advances in the field of computer technology, acquiring data from sensors has become an easy task. Once large data is available, the current status of the process plant can be studied, and this would aid in diagnosing faults.

Process history methods are further classified into two categories: univariate and multivariate techniques. While univariate techniques monitor a single variable at a time, multivariate methods excel in monitoring multiple variables. In recent years, a few of the commonly applied multivariate methods include principal component analysis (PCA), principal component regression (PCR) and partial least squares (PLS) (Yin et al. 2014). PCA, a classic dimensionality reduction method, is one of the oldest multivariate methods used in fault detection problems (Nor et al. 2020; Alauddin et al.

✉ Muddu Madakyaru
muddu.m@manipal.edu

K. Ramakrishna Kini
kr.kini@manipal.edu

¹ Department of Instrumentation and Control Engineering,
Manipal Institute of Technology, Manipal Academy of Higher
Education, Manipal 576104, India

² Department of Chemical Engineering, Manipal Institute of
Technology, Manipal Academy of Higher Education, Manipal
576104, India

2018). In PCA, large-dimension data is transformed to new dimensional subspace with an aim of retaining the original data information in the transformed variables (Harkat et al. 2020). Despite being the core method in fault detection problems, PCA FD strategy carries with it a few disadvantages. First, modeling in PCA is carried out by considering second-order statistics of mean and variance. Secondly, PCA assumes that the latent variables adhere to Gaussian distribution, which is not often true with industrial processes.

More recently, a PM-FD method based on independent component analysis (ICA) has been proposed for capturing non-Gaussian characteristics of data through higher-order statistical representation. The ICA method, initially developed for blind source separation problems, has progressed very well to find applications in PM-FD problems (Tong et al. 2017; Li and Hongbo 2010). In ICA, large data is decomposed into linear combination of independent components (ICs) that are independent with no orthogonality constraint (Hyvarinen 2013). ICA modeling involves computation of ICs by using an iterative algorithm in which separating the matrix is randomly initialized, and this usually results in different solutions for ICA algorithms (Lee et al. 2006). In this work, the de-mixing orthogonal matrix $\mathbf{V}$ is initialized to identity the matrix instead of a random matrix, as this helps in providing a constant solution. Once the ICA model is developed, it is used to monitor new data through three fault indicators, namely, I_d^2 , I_e^2 and squared prediction error (SPE). Many extensions of conventional ICA technique have been developed in recent years. This includes dynamic ICA (DICA), which considers the dynamics of the data via lagged variables, modified ICA (MICA) (Tong et al. 2017; Yingwei and Yang 2010), where dominant ICs are extracted from data to reduce computational complexity, kernel ICA (KICA) (Tian et al. 2009), which is used to handle both nonlinear and non-Gaussian industrial data, and multi-block ICA where ICA process monitoring is carried out in different blocks (Yingwei and Chi 2012). A multi-way kernel “FastICA” based on feature samples was developed, which was used for diagnosing batch processes, while nonlinear detection using integrated kernel PCA and locality preserving projections (LPPs) was also developed (Lianfang and Xuemin 2014).

Though ICA and its variant methods have been applied successfully in the process monitoring domain, they still face a stiff challenge of extracting information from high-dimensional noisy data available from industrial plants. In this work, the robustness of ICA and PCA fault detection strategies is validated by their ability in monitoring process data in the presence of different levels of noise. To test the robustness of an FD strategy towards different noise realizations of sensor data, stochastic simulations have been carried out. The main aim is to show that the ICA-based models are more robust for different levels of noise as compared to

conventional PCA method. Stochastic simulations involving 1000 runs will be performed for noise with a signal-to-noise ratio (SNR) of 20, 10 and 5. The SNR is defined as the ratio of the variance of the signal to variance of the noise. $\text{SNR} = 20$ corresponds to quality data with less noise, $\text{SNR} = 10$ corresponds to a medium level of noise, while $\text{SNR} = 5$ indicates a high amount of noise in the data. Comparisons have been carried out between ICA-, DICA-, MICA- and PCA-based fault detection strategies for different noise levels.

The organization of this paper is as follows: Section 2 provides a detailed overview of ICA, DICA, MICA and PCA modeling methods along with the fault indicators. Next, robust fault detection based on ICA involving stochastic simulation with different noise realization is presented through a block diagram representation in Sect. 3. To test the effectiveness of different FD strategies against noise realizations, two case studies have been considered: a simulated quadruple tank process (QTP) and a simulated distillation column (DC) process in Sect. 4. The paper terminates with a conclusion section in Sect. 5.

2 Data-Driven Modeling Framework

2.1 Independent Component Analysis

ICA is a multivariate technique with the sole objective of extracting hidden independent components (ICs) from non-Gaussian data. Available training data, $\mathbf{X}_a = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]^T \in \mathbb{R}^{m \times n}$ is represented as a linear mixture of p ($p \leq m$) unknown ICs $s_1, s_2, \dots, s_p$ and a mixing or a separation matrix. The original data and ICs are related which is described as:

$$\mathbf{X}_a = \mathbf{A}\mathbf{S} + \mathbf{E} \quad (1)$$

where n and m represent number of observations and variables, respectively, and $\mathbf{A} = [\mathbf{a}_1 \dots \mathbf{a}_p]^T \in \mathbb{R}^{m \times p}$ is a deterministic mixing matrix, $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n] \in \mathbb{R}^{p \times n}$ is the IC matrix and $\mathbf{E} \in \mathbb{R}^{m \times n}$ is a residual matrix. Since two unknown matrices $\mathbf{S}$ and $\mathbf{A}$ cannot be computed, the aim is focused in finding a separating matrix $\mathbf{W}$ such that the reconstructed matrix is given by:

$$\hat{\mathbf{S}} = \mathbf{W}\mathbf{X}_a \quad (2)$$

The ICA model development stage consists of four steps, namely, centering, whitening, preprocessing and iterative computing. Initially, the data $\mathbf{X}_a$ is preprocessed to have a mean of zero. Next, whitening is performed by singular value decomposition (SVD) on the covariance of $\mathbf{X}_a$ which generates the transformation $\mathbf{Z} = \mathbf{Q}\mathbf{X}_a$ with $\mathbf{Q} = \mathbf{\Lambda}^{-1}\mathbf{U}^T$, where $\mathbf{\Lambda}$ is a matrix involving diagonal eigen values and $\mathbf{U}$ is an eigen vector matrix obtained using the covariance of $\mathbf{X}_a$. The

transformation post whitening stage is:

$$\mathbf{Z} = \mathbf{Q}\mathbf{X}_a = \mathbf{Q}\mathbf{A}\mathbf{S} = \mathbf{V}\mathbf{S} \quad (3)$$

where $\mathbf{V}$ is an orthogonal matrix. Next, the preprocessing step involves developing important relationships between original data and the separating matrix $\mathbf{W}$, which is useful for determining the ICs. From Eq. (3), $\mathbf{S}$ can be estimated as follows:

$$\hat{\mathbf{S}} = \mathbf{V}^T \mathbf{Z} = \mathbf{V}^T \mathbf{Q}\mathbf{X}_a \quad (4)$$

Comparing Eqs. (2) and (4), a relation between $\mathbf{W}$ and $\mathbf{V}$ is obtained, given by:

$$\mathbf{W} = \mathbf{V}^T \mathbf{Q} \quad (5)$$

To avoid random solutions in each iteration, instead of random initialization of $\mathbf{V}$, it is initialized to an identity matrix $\mathbf{I} \in R^{m \times m}$. Next, each column $\mathbf{v}_i$ of $\mathbf{V}$ is updated one by one through an iterative algorithm. The computation of $\mathbf{V}$ requires all ICs to be independent, and criteria for independence is based on non-Gaussianity which is measured through kurtosis and negentropy. Kurtosis, the fourth-order moment, is sensitive to outliers which makes negentropy a better choice. For a variable having density function $f(Y)$, H is represented as:

$$H(Y) = - \int f(Y) \log f(Y) dy. \quad (6)$$

Any difference in entropy between the given distribution and normal distribution having the same mean-variance is computed by negentropy which is mathematically given by:

$$J(Y) = H(Y_{guass}) - H(Y) \quad (7)$$

Y_{guass} is a random Gaussian variable with a similar covariance matrix as that of Y . For estimating negentropy, a simpler approximation was proposed (Aapo and Erkki 2000):

$$J(Y) \approx [E\{G(\hat{Y})\} - E\{G(y_g)\}]^2 \quad (8)$$

where Y and y_g are assumed to have a mean of zero and unity variance, and $G(\cdot)$ is a non-quadratic function which can be evaluated as (Aapo and Erkki 2000):

$$G(k) = \exp(-ck^2/2) \quad (9)$$

where $c \approx 1$. By maximizing the objective function, extraction of independent components takes place sequentially:

$$J(\mathbf{v}) \approx [E\{G(\mathbf{v}^T \mathbf{x}_a)\} - E\{G(y_g)\}]^2 \quad (10)$$

Each element of $\mathbf{V}$ corresponds to one IC and can be estimated by the use of a deflation scheme. By maximizing the sum of the N -unit cost function, one unit cost function is generalized for determining $\mathbf{V}$. The resulting optimization problem can be described as:

$$\max \sum_{i=1}^N J(\mathbf{v}_i) \quad i = 1, 2, \dots, N \quad (11)$$

S.T. $E\{(\mathbf{v}_k^T \mathbf{x}_a)(\mathbf{v}_j^T \mathbf{x}_a)\} = \delta_{jk}$ and $\mathbf{s}_i = \mathbf{v}_i^T \mathbf{x}_a$ where $\mathbf{v}_i$ is the row of $\mathbf{V}$ at the maximum of $J(\mathbf{v}_i)$. This undergoes maximization at certain optima of $E\{g(\mathbf{v}^T \mathbf{x}_a)\}$ under the constraint $E\{(\mathbf{v}^T \mathbf{x}_a)^2\} = \|\mathbf{v}\|^2 = 1$ which is described by:

$$F(\mathbf{v}) = E\{\mathbf{x}_a g(\mathbf{v}^T \mathbf{x}_a)\} - \gamma \mathbf{v} = 0 \quad (12)$$

$\gamma = E\{\mathbf{v}_0^T \mathbf{x}_a g(\mathbf{v}^T \mathbf{x}_a)\}$ and v_0 is the optimum value of $\mathbf{v}$. Solving the above by using the Newton method yields the fixed-point algorithm (Ali and Mahdi 2012):

$$E\{\mathbf{z} g(\mathbf{v}^T \mathbf{z})\} - E\{g'(\mathbf{v}^T \mathbf{z})\} \mathbf{v} \quad (13)$$

The above algorithm extracts only one independent component at a time. This fixed-point algorithm has to be run m times to estimate the remaining ICs. After each iteration, the vectors have to undergo orthogonalization to make sure that they do not converge to the same maxima. Hence, performing the deflationary orthogonalization through the Gram–Schmidt method:

$$\mathbf{v}_{j+1} = \mathbf{v}_{j+1} - \sum_{i=1}^j (\mathbf{v}_{j+1}^T \mathbf{v}_i) \mathbf{v}_i \quad (14)$$

$$\mathbf{v}_{j+1} = \frac{\mathbf{v}_{j+1}}{\|\mathbf{v}_{j+1}\|} \quad (15)$$

where j is the number of previously estimated vectors. Once the ICs are computed using the fixed-point algorithm, the next task is selection of dominant ICs which is needed for two reasons: robust performance and reduction of complexity. A few methods have been commonly preferred for selection of dominant ICs (Lee et al. 2006). However, the cumulative percentage variance (CPV) method has been applied in the current work to determine optimum ICs. First, the L_2 norm is used to sort the rows of separating W from descending to ascending order, $\lambda = \|\mathbf{W}\|_2$. Next, CPV method is applied to determine p -dominant ICs which is expressed as:

$$\text{CPV}(p) = \frac{\sum_{i=1}^p \lambda_i}{\sum_{i=1}^m \lambda_i} * 100 \quad (16)$$

Once a model is developed, determining the prediction capability is crucial for its acceptance. Generally, the prediction

capability of a model is reduced when large noise is present in the data. In this work, the R^2 test has been used to evaluate the goodness of the developed PCA as well as the ICA model. The coefficient of determination, denoted as R^2 , is the ratio of the sum of squares regression to the total sum of squares denoted mathematically as:

$$R^2 = \frac{SSQ}{TSSQ} \quad (17)$$

SSQ is the sum of squares which determines how far the predicted value is from the mean of the original data, and TSSQ determines how the original data points vary from their mean.

Once an ICA model is developed from data exhibiting normal process operation, the future behavior of variables is compared with the present behavior through the fault detection indicators. The ICA model may be represented as:

$$\mathbf{X} = \underbrace{\mathbf{A}_p \mathbf{S}_p}_{\text{Systematic space}} + \underbrace{\mathbf{A}_{m-p} \mathbf{S}_{m-p}}_{\text{Excluded space}} + \underbrace{\mathbf{E}}_{\text{Residual space}} \quad (18)$$

The ICA model consists of three parts: The systematic space represents a model with respect to p optimum ICs, the excluded space represents a model with respect to ignored $m-p$ ICs, and the residual space. To monitor the three parts, three fault indicators are employed namely, I_d^2 to monitor systematic space, I_e^2 to monitor excluded space and SPE to monitor residual space. The reconstructed vectors $\hat{\mathbf{f}}_{\text{new}p} = \mathbf{W}_p \mathbf{X}_{\text{new}p}$ and $\hat{\mathbf{f}}_{\text{new}m-p} = \mathbf{W}_{m-p} \mathbf{X}_{\text{new}m-p}$ are computed for a new data $\mathbf{X}_{\text{new}}$ where $\mathbf{W}_p$ is a matrix corresponding to p rows, and $\mathbf{W}_{m-p}$ corresponds to excluded $m-p$ rows of the separating matrix. The fault detection indicators of ICA at sample i are represented as (Yuan et al. 2019):

$$I_d^2(i) = \hat{\mathbf{f}}_{\text{new}p}^T(i) \hat{\mathbf{f}}_{\text{new}p}(i) \quad (19)$$

$$I_e^2(i) = \hat{\mathbf{f}}_{\text{new}m-p}^T(i) \hat{\mathbf{f}}_{\text{new}m-p}(i) \quad (20)$$

$$SPE(i) = (\mathbf{X}_{\text{new}}(i) - \hat{\mathbf{X}}_{\text{new}}(i))^T (\mathbf{X}_{\text{new}}(i) - \hat{\mathbf{X}}_{\text{new}}(i)) \quad (21)$$

where $\hat{\mathbf{X}}_{\text{new}} = \mathbf{Q}^{-1} \mathbf{B}_p \hat{\mathbf{f}}_{\text{new}}(i) = \mathbf{Q}^{-1} \mathbf{B}_p \mathbf{W}_p \mathbf{X}_{\text{new}}(i)$. Kernel density estimation(KDE) technique is employed for computing the threshold limits for fault indicators (Lee et al. 2004).

2.2 Dynamic ICA

In chemical process plants, variables do not remain at a fixed state, as they have a tendency to move around a nominal operating range, and this will result in variables having strong autocorrelation. The conventional process monitoring methods perform modeling assuming that the present sampling instants are independent from the past sampling instants. However, due to the autocorrelation exhibited by variables,

present sampling instants have dependency on past sampling instants, and this information can be included in the modeling stage through dynamic ICA strategy. In dynamic ICA strategy, the variables are supplemented with past observations to generate an augmented matrix which is represented by (Lee et al. 2004):

$$\mathbf{X}(l) = \begin{bmatrix} \mathbf{x}_t^T & \mathbf{x}_{t-1}^T & \cdots & \mathbf{x}_{t-l}^T \\ \mathbf{x}_{t-1}^T & \mathbf{x}_{t-2}^T & \cdots & \mathbf{x}_{t-l-1}^T \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_{t+l-n}^T & \mathbf{x}_{t+l-n-l}^T & \cdots & \mathbf{x}_{t-n}^T \end{bmatrix} \quad (22)$$

where $\mathbf{x}_t^T$ is an observation vector at sampling instant t , and l denotes the number of lagged measurements. The selection of a proper value for time lag l is crucial for capturing dynamic information. A few methods available in the literature for choosing the right number of lags include a subspace identification criterion and the Akaike information criterion. However, it may be noted that selecting $l = 1$ or $l = 2$ yields good results for dynamic multivariate techniques (Lee et al. 2004). Once a reference dynamic ICA model is constructed using fault-free data, the fault indicators described in Eqs. (19)–(21) are used for detecting faults in a new data set.

2.3 Modified ICA

The modified ICA (MICA) algorithm was proposed as an improvement of the ICA algorithm in order to have structured ordering of the independent components (ICs). In the MICA algorithm, the initial estimate of ICs is done using normalized principal components (PCs) from the PCA model, and then, the FastICA algorithm is utilized to update the retained ICs with same variance being maintained as that of PCs. The starting step of MICA includes applying PCA to extract PCs from data $\mathbf{X}_a = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]^T \in \mathbb{R}^{m \times n}$.

$$\mathbf{X}_a = \mathbf{T} \mathbf{P}^T \quad (23)$$

where $\mathbf{T}$ and $\mathbf{P}$ are score and loading vector matrices, respectively, generated from covariance of $\mathbf{X}_a$. The last few eigenvalues of $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m)$ are close to zero, and they can be left out. The retained eigenvalues or PCs undergo whitening which is expressed as $\mathbf{Z} = \mathbf{Q} \mathbf{X}_a$ with $\mathbf{Q} = \mathbf{\Lambda}^{-1} \mathbf{P}^T$. The whitened components of $\mathbf{Z}$ serve as initial estimates for the ICs. The main objective of MICA is to ensure that elements of $\mathbf{S}$ become as independent as possible, and the MICA model is defined as:

$$\mathbf{S} = \mathbf{C}^T \mathbf{Z} \quad (24)$$

where $\mathbf{C} \in \mathbb{R}^{m \times p}$ such that $\mathbf{C}^T \mathbf{C} = \mathbf{D}$ and $\mathbf{D} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$ considering all the dominant eigenvalues. The variance of each element of $\mathbf{S}$ is the same as the variance of the score elements of PCA since the following condition is satisfied:

$$\frac{\mathbf{S}\mathbf{S}^T}{n-1} = \mathbf{D} \quad (25)$$

This is followed by normalization of $\mathbf{S}$ such that:

$$\mathbf{S}_{\text{no}} = \mathbf{D}^{-\frac{1}{2}} \mathbf{S} = \mathbf{C}_{\text{no}}^T \mathbf{Z} \quad (26)$$

where $E(\mathbf{S}\mathbf{S}^T) = \mathbf{I}$. Hence, the objective of MICA is reduced in determining $\mathbf{C}_{\text{no}}$. The elements of the $\mathbf{C}_{\text{no}}$ matrix are computed by a fixed-point algorithm using the negentropy approximation (Lee et al. 2007). Before the fixed-point algorithm is executed, the following initialization is ensured:

$$\mathbf{C}_{\text{no}} = [\mathbf{I}_m; \mathbf{0}] \quad (27)$$

where $\mathbf{I}_m$ is a m -dimensional identity matrix and $\mathbf{0}$ is an $m \times (p-m)$ zero matrix. After the computation of $\mathbf{C}_{\text{no}}$, $\mathbf{W}$ and $\mathbf{A}$ matrices are computed using:

$$\mathbf{W} = \mathbf{D}^{1/2} \mathbf{C}_{\text{no}}^T \mathbf{Q} \quad (28)$$

$$\mathbf{A} = \mathbf{P} \mathbf{A}^{1/2} \mathbf{C}_{\text{no}} \mathbf{D}^{-1/2} \quad (29)$$

The next step is to determine dominant ICs. As the variance of each IC in $\mathbf{S}$ is the same as that of PCs, any standard criterion for determining optimum PCs in the PCA strategy can be employed. The CPV technique is considered for determining optimum ICs in this study. Two fault indices are defined for the MICA strategy: the T^2 and SPE for modeling systematic and residual parts, respectively. For a new data $\mathbf{X}_{\text{new}}$, they are defined as follows:

$$T^2 = \mathbf{s}^T \mathbf{D}^{-1} \mathbf{s} \quad (30)$$

$$\text{SPE} = (\mathbf{X} - \hat{\mathbf{X}})^T (\mathbf{X} - \hat{\mathbf{X}}) \quad (31)$$

where $\mathbf{s} = \mathbf{W} \mathbf{X}_{\text{new}}$ and $\hat{\mathbf{X}} = \mathbf{A} \mathbf{s}$. If the fault indices exceed a predetermined reference threshold, then a fault is declared (Lee et al. 2006).

2.4 Principal Component Analysis

PCA is the most widely applied technique for capturing correlation between the process data variables. After normalization to zero mean and unity variance, the scaled data $\mathbf{X} \in \mathbb{R}^{n \times m}$ is used to estimate principal components (PCs) through singular value decomposition:

$$\mathbf{X} = \mathbf{T} \mathbf{P}^T \quad (32)$$

where $\mathbf{T}$ and $\mathbf{P}$ are the PC and loading vector matrix, respectively (Harrou et al. 2013). The loading vector matrix is related with covariance of $\mathbf{X}$, which is described mathematically as:

$$\mathbf{\Sigma} = \mathbf{P} \mathbf{\Lambda} \mathbf{P}^T \quad \text{with} \quad \mathbf{P} \mathbf{P}^T = \mathbf{P}^T \mathbf{P} = \mathbf{I}_m \quad (33)$$

where $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m)$ is a matrix comprising eigenvalues of $\mathbf{\Sigma}$ arranged in diagonal format in descending order. Once optimum PCs p are computed via CPV technique, the PCA model can be represented as:

$$\mathbf{X} = \underbrace{\hat{\mathbf{T}} \hat{\mathbf{P}}^T}_{\text{modeled space}} + \underbrace{\tilde{\mathbf{T}} \tilde{\mathbf{P}}^T}_{\text{residual space}} = \hat{\mathbf{X}} + \mathbf{F} \quad (34)$$

The $\hat{\mathbf{X}}$ and $\hat{\mathbf{T}}$ represent model parameters for p retained PCs. The residual matrix $\mathbf{F}$ represent model parameters for $m-p$ ignored PCs.

The T^2 statistics monitors the modeled part of the developed PCA model. It is represented in the following manner for a new data $\mathbf{X}_{\text{new}}$:

$$T^2 = \mathbf{X}_{\text{new}}^T \hat{\mathbf{P}} \hat{\mathbf{\Lambda}}^{-1} \hat{\mathbf{P}}^T \mathbf{X}_{\text{new}} \quad (35)$$

The matrix $\hat{\mathbf{\Lambda}}$ contains diagonal eigenvalues of retained PCs. The Q statistics monitors the residual part of the developed PCA model, and it is expressed as:

$$Q = \mathbf{X}_{\text{new}}^T (\mathbf{I} - \hat{\mathbf{P}} \hat{\mathbf{P}}^T) \mathbf{X}_{\text{new}} \quad (36)$$

For $\mathbf{X}_{\text{new}}$, a fault may be declared if the value of T^2 and Q fault indicators is greater than the threshold (Sanjula and Zukui 2019).

3 ICA-Based Robust Fault Detection Strategy

Once a multivariate model is developed using ICA, evaluating the model for robustness and prediction capability is crucial for its acceptance. The developed multivariate model should be robust enough to perform in harsh and noisy industrial environments. It has been observed that a multivariate FD strategy which performs well in less noisy environments may not be able to replicate the same performance in a noisier environment. Hence, the presence of noise does have its effect and can degrade the fault handling capacity of an FD strategy. In this paper, we present a multivariate FD strategy based on ICA and present the behavior of the developed ICA model when noise is introduced in the underlying data. The aim of this work is to study the effect of different levels of measurement noise on performance of ICA, DICA and

MICA models. After this, similar noise realizations are performed on a PCA model followed by drawing comparisons between PCA- and ICA-based FD methods.

Algorithm 1 : Robust ICA fault detection algorithm

1. For a given training and testing data set, perform scaling to have a mean of zero.
 2. Whiten training data to remove cross-correlation and develop relations between $\mathbf{Q}$, $\mathbf{V}$ and $\mathbf{W}$ to generate $\mathbf{W} = \mathbf{V}^T \mathbf{Q}$.
 3. Initialize $\mathbf{V}$ to identity matrix $\mathbf{I} \in R^{m \times m}$ and calculate the number of ICs using a fixed-point algorithm.
 4. Determine optimum ICs and threshold limits for I_d^2 , I_e^2 fault indicators.
 5. Fix the number of stochastic simulations and add noise with a defined SNR to the testing data.
 6. Perform stochastic simulations and calculate fault indicators I_d^2 , I_e^2 and SPE statistics in each simulation.
 7. Compute the values of the false alarm rate (FAR) and missed detection rate (MDR) for each simulation.
 8. If the developed model possesses lower values of FAR and MDR, declare the developed fault detection strategy as robust.
-

It is unfair to come to a conclusion regarding the robustness of a fault detection scheme based on few rounds of simulations. In the field of system identification and signal processing, when faced with significant uncertainty in a process model involving making a estimation, stochastic simulation has proved to be a good solution. Thus, in this study, stochastic simulations have been carried out with 1000 realizations. The different noise realizations that would be carried out involve introducing noise with a signal-to-noise ratio (SNR) to the data. In the present task, noise with a defined SNR, i.e., SNR = 20, SNR = 10 and SNR = 5, is added to the data by performing stochastic simulation. The data with SNR = 20 corresponds to quality data with less noise, the data with SNR = 10 corresponds to a medium level of noise and SNR = 5 indicates very noisy data. The algorithm of the robust fault detection strategy is described in Algorithm 1.

Figure 1 describes a block diagram representation of the proposed strategy. The available process data is initially split into training and testing groups. The training group data is preprocessed to a zero mean which is then used to develop the ICA model. After determining the optimum ICs, KDE technique is employed to compute the threshold for fault indicators. For the testing data group, stochastic simulation of 1000 runs is performed for noise with an SNR that is fixed at 20, 10 and 5. Next, the fault indicators for ICA are computed and compared with the threshold. During the absence of a fault, the value of the fault indicator will be less than the control limit value. A fault is declared when the fault indicator exceeds the control limit value. Similar noise realizations

are also performed on DICA, MICA and PCA strategies for having comparison. In process monitoring problems, the performance of an FD strategy is assessed using false alarm rate (FAR) and missed detection rate (MDR). The FAR is defined as the ratio of the total number of false alarms to the total number of data in a normal operating range. The MDR is defined as the ratio of the total number of faults which do not exceed the control limits to the total number of data in the faulty range. A fault detection strategy is deemed to be robust if FAR and MDR values are minimum for different levels of noise.

4 Case Studies

In this section, two case studies are considered: simulated quadruple tank process and simulated distillation column (DC) process to demonstrate the effectiveness of ICA-based strategies over PCA-FD strategy.

4.1 Types of Sensor Faults

A fault is defined as a continuous series of events where a variable deviates from its regular operating range. Consider a process having a large number of variables. A measurement $M(t)$ associated with a process variable may be represented as:

$$M(t) = M_n(t) + F(t) \quad (37)$$

where $M_n(t)$ is the true value of the measurement $M(t)$, and $F(t)$ is a possible fault at time instant t . The value of $F(t)$ in the equation would be zero when the process variable is operating normally and it takes a non-zero value in the presence of a fault. A bias fault is recognized by sudden jump of a variable from normal operating value to higher or lower abnormal value. In this fault, measurement shifts from its normal value of $M_n(t)$ to a new value $M_n(t) + F$. This can be represented in time series format as:

$$M(t) = \begin{cases} M_n(t) & \text{if } t < t_a; \\ M_n(t) + F & \text{if } t \geq t_a; \end{cases} \quad (38)$$

where F is a bias term representing the rate of increase (%), and t_a is a time instant where the fault appears. The aging or drift fault occurs due to aging of a sensor and can be represented in time series format as:

$$M(t) = \begin{cases} M_n(t) & \text{if } t < t_a; \\ M_n(t) + V(t - t_a) & \text{if } t \geq t_a; \end{cases} \quad (39)$$

Fig. 1 Representation of the proposed strategy

Fig. 2 A schematic of the quadruple tank process

where $N_n(t)$ is the true value of the measurement, V is the rate of the slow drift behavior and t_a is the time instant where the fault occurs.

4.2 Simulation Study on a Quadruple Tank Process

In the last few years, the quadruple tank system has been applied in many control as well as process problems since it demonstrates very useful multivariable phenomena between the four tanks. In this process, voltages of a pump v_1 and v_2

Table 1 QTP: MDR & FAR for bias fault

Indicators	SNR = 20	SNR = 10	SNR = 5
MDR			
PCA- T^2	0.13	0.33	2.3
PCA-Q	35.92	68.31	71.51
ICA- I_d^2	0	0	0.01
ICA- I_e^2	0	0	0.07
ICA-SPE	0	0.03	0.02
DICA- I_d^2	0	0.06	0.84
DICA- I_e^2	0	0.03	2.92
DICA-SPE	0	0.05	0.56
MICA- T^2	0	0	0.02
MICA-SPE	0	0	1.02
FAR			
PCA- T^2	33.65	32.87	32.70
PCA-Q	0	0	0
ICA- I_d^2	3.83	3.13	5.16
ICA- I_e^2	2.16	2.82	2.63
ICA-SPE	4.09	5.3	5.37
DICA- I_d^2	0.8	1.9	10.57
DICA- I_e^2	2.15	5.8	10.99
DICA-SPE	1.24	4.5	13.34
MICA- T^2	3.24	4.89	5.17
MICA-SPE	1.01	1.77	3.69

are inputs, while the heights of the liquid level h_1 , h_2 , h_3 and h_4 are measured outputs. The equations that govern the quadruple tank process (QTP) presented through Fig. 2 are (Karl 2000):

$$\frac{dh_1}{dt} = \frac{q_1 k_1 V_1}{A_1} + \frac{a_3 \sqrt{2gh_3}}{A_1} - \frac{a_1 \sqrt{2gh_1}}{A_1} \quad (40)$$

Fig. 3 QTP: monitoring results of PCA with bias fault for SNR = 5

Fig. 4 QTP: monitoring results of ICA with bias fault for SNR = 5

$$\frac{dh_2}{dt} = \frac{q_2 k_2 V_2}{A_2} + \frac{a_4 \sqrt{2gh_4}}{A_2} - \frac{a_2 \sqrt{2gh_2}}{A_2} \quad (41)$$

$$\frac{dh_3}{dt} = \frac{(1 - q_2) k_2 V_2}{A_3} - \frac{a_3 \sqrt{2gh_3}}{A_3} \quad (42)$$

$$\frac{dh_4}{dt} = \frac{(1 - q_1) k_1 V_1}{A_4} - \frac{a_4 \sqrt{2gh_4}}{A_4} \quad (43)$$

In the above equations, $a_1 \dots a_4$ represent area of the outlet pipes; $A_1 \dots A_4$ represent the cross-sectional area of the tanks; q_1 and q_2 are valve ratios; k_1 and k_2 denote pump constants; and g is the gravity of the earth.

In this case, dynamic simulations are carried out to generate a data set for which standard nominal operating points are considered (Karl 2000). Perturbation of inputs is done at their normal operation range using a pseudo-random binary signal

in the range of frequency $[0 \ 0.03 \ \omega_n]$, where $\omega_n = \pi/T$ represents Nyquist frequency. The idinput function in the System Identification Toolbox of MATLAB is used to generate 2048 samples of data, which were later split equally into training and testing data. For ICA and modified ICA(MICA) strategies, three optimum ICs are selected, whereas three PCs are selected for the PCA strategy. For the dynamic ICA strategy, a lag of $l = 1$ is considered, and six optimum ICs are selected for the same. The model performance of ICA is assessed using R^2 values which are 0.975, 0.941 and 0.891 for SNR = 20, SNR = 10 and SNR = 5, respectively. A value of $R^2 = 0.975$ suggests that 97% of the total sum of squares in testing data is described by the ICA model and only 3% in the residuals.

Fig. 5 QTP: monitoring results of DICA with bias fault for SNR = 5

Table 2 QTP: MDR & FAR for drift fault

Indicators	SNR = 20	SNR = 10	SNR = 5
<i>MDR</i>			
PCA- T^2	20.57	21.28	22.47
PCA-Q	5.54	6	7.03
ICA- I_d^2	4.5	8.76	9.4
ICA- I_e^2	4	6	6.6
ICA-SPE	4.7	7.65	9.95
DICA- I_d^2	5.72	6.13	8.46
DICA- I_e^2	1.88	2.75	3.13
DICA-SPE	2.72	4.16	5.19
MICA- T^2	3.50	4.46	7.76
MICA-SPE	4.14	6.60	9.56
<i>FAR</i>			
PCA- T^2	3.05	3.27	3.37
PCA-Q	8.46	8.96	9.35
ICA- I_d^2	1.14	1.2	3.13
ICA- I_e^2	2.47	3.71	4.57
ICA-SPE	0.41	4.6	2.62
DICA- I_d^2	8.47	8.2	10.59
DICA- I_e^2	9.59	10.5	12.65
DICA-SPE	7.55	9.29	10.95
MICA- T^2	1.59	3.85	5.16
MICA-SPE	1.16	4.60	6.44

First, a sensor bias fault with a magnitude of +1 is introduced in the tank 1 measurement at sample 175 of the testing data. The fault indicators of different strategies evaluated through FAR as well as MDR for different noise realizations is presented in Table 1. While PCA-Q statistics has no false

Fig. 6 Schematic of a distillation column

alarms, the PCA- T^2 statistics has many more false alarms than the fault indicators of ICA, DICA and MICA methods. The PCA strategy has very high value of MDR with increased noise level in comparison to other strategies. The performance of ICA, MICA and DICA methods is acceptable even though MDR value increases with added levels of noise. For providing clarity to the reader, monitoring plots of PCA, ICA and DICA strategies in monitoring bias fault for the case of SNR = 5 are presented in Figs. 3, 4 and 5, respectively.

In the next case, a fault with a rate of 0.008 resembling a drift sensor fault is introduced in sensor data-3 (i.e tank 3 height measurement) from sampling time instant 150 till the end of testing data. The performance of different fault

Fig. 7 DC process: variation of output signals

Fig. 8 DC process: scatter plot, observed vs. predicted values

indicators evaluated through FAR/MDR values for the cases with SNR = 20, SNR = 10 and SNR = 5 is presented in Table 2. The results demonstrate that MDR values for PCA and ICA fault indicators are almost comparable except for PCA- T^2 which has a high MDR value. The FAR values of all the methods are also in the same range, and the FAR value increases as more noise is added in the data. Thus, it can be concluded that the monitoring performance of ICA, DICA and MICA strategies are slightly better in comparison

to PCA strategy in monitoring bias and drift faults. The ICA and its variant fault detection methods are more robust for different levels of noise in the quadruple tank process case study.

4.3 Distillation Column Simulation Study

A distillation column (DC) is an energy-consuming unit in any chemical process plant and is used for separating compo-

nents from mixtures of components based on the difference in vapor pressure. Proper monitoring of distillation columns is necessary in an industry to avoid any accident and loss of product quality. A schematic of a DC process applied in an industrial setup can be observed in Fig. 6. In this example, the distillation column consists of 32 plates, and 10 resistance temperature detectors (RTDs) are used to monitor temperature at different locations of the column. The flow rates of feed and reflux are used to perturb the distillation column.

The Aspen simulator is used for generating distillation column process data. To begin, flow rates of the feed as well as reflux streams are perturbed from their nominal operating ranges (Harrou et al. 2018; Madakyaru et al. 2013). Once the system has reached a steady-state condition, these perturbations are used to generate data. The input variables consist of 10 temperatures corresponding to measurements at various locations of the column along with flow rates of feed and reflux. The variations of the output (i.e., x_B and x_D) for changes in the input perturbations are presented in Fig. 7. A total data length of 4096 samples is generated with 14 variables, which are then split equally into training and testing data sets.

The training data is used for developing PCA- and ICA-based models which will then be used for monitoring faults in testing data. For ICA and MICA strategies, 8 and 10 ICs are selected, whereas 8 PCs are selected for the PCA strategy. For the DICA strategy, a lag of $l = 1$ is considered and 14 ICs are selected for the same. Figure 8 demonstrates a scatter plot of observed values vs. predicted values for the case SNR = 5, and it is observed that predicted values follow observed values perfectly. The goodness of the model in fitting training data is found using R^2 statistics. The value of R^2 for the cases with SNR = 20, SNR = 10 and SNR = 5 is found to be 0.978, 0.955 and 0.9152, respectively. Since R^2 values reduces as more noise is present in the data, it may be concluded that noise has its effect on the prediction capability of the ICA model.

First, the ability of different FD strategies in handling sensor bias fault is being investigated. To simulate this fault, a bias with a magnitude of 20 is introduced in testing data temperature variable 5 at sampling time instant 750. To obtain a valid conclusion from the simulation results, 1000 stochastic simulations of different noise realizations have been carried out, and the results are tabulated in Table 3. From the table, it is observed that despite having no false alarms, the T^2 and Q statistics have a very high MDR which indicates that they are unable to detect bias fault in the presence of noise. In contrast to PCA, the fault indicators of ICA, DICA and MICA strategies have less FAR and MDR values, thus demonstrating enhanced performance in detecting the bias fault even with increase in noise levels.

Next, the capability of different FD strategies in monitoring sensor drift fault is presented in this section. A slow

Table 3 DC process: MDR & FAR for bias fault

Indicators	SNR = 20	SNR = 10	SNR = 5
<i>MDR</i>			
PCA- T^2	0	5.62	98.22
PCA-Q	87.26	99.35	100
ICA- I_d^2	0	2.76	1.17
ICA- I_e^2	0	0	0
ICA-SPE	0	0	6.53
DICA- I_d^2	0	0	1.52
DICA- I_e^2	0	1.04	3.5
DICA-SPE	0	1.08	6.53
MICA- T^2	0	0.5	1.75
MICA-SPE	0	0	2.06
<i>FAR</i>			
PCA- T^2	0	0	0
PCA-Q	0	0	0
ICA- I_d^2	0.23	0.16	0.2
ICA- I_e^2	0.21	0.13	0.2
ICA-SPE	0.09	0.07	0.02
DICA- I_d^2	2.49	5.48	11.23
DICA- I_e^2	1.32	3.49	5.32
DICA-SPE	1.15	6.70	8.17
MICA- T^2	0	0	0.06
MICA-SPE	0	0	0.09

ramp fault with a rate of 0.015 resembling an aging sensor fault is introduced in temperature variable 7 at sampling time instant 900. Stochastic simulations are carried out to assess the performance of fault indicators in monitoring this fault for different noise realizations. The time evaluation of PCA, ICA and MICA fault indicators can be observed in Figs. 9, 10 and 11, respectively. The results of stochastic simulation evaluated using FAR and MDR are presented in Table 4. It is observed that the PCA fault indicators have zero FAR but very high MDR value, and this clearly suggests that performance is not good in monitoring drift fault in the presence of noise. In contrast, ICA, DICA and MICA strategies demonstrate better detection of drift fault when noise level is increased in the data. Hence, it can be inferred that performance of ICA, DICA and MICA strategies is commendable in handling sensor faults of the DC process in the presence of noise.

5 Conclusion

The aim of this study was to find the effectiveness of ICA-based fault detection strategies against different levels of noise in the measured variables, which was then compared with the PCA-based fault detection strategy. Whenever data

Fig. 9 DC process: monitoring results of PCA with drift fault for SNR = 10

Fig. 10 DC process: monitoring results of ICA with drift fault for SNR = 10

comes with any level of noise, there is difficulty in separating useful information, and this leads to poor model prediction. Hence, to check the robustness of the FD strategies, stochastic simulations were performed for noise with SNR = 20, SNR = 10 and SNR = 5. The data with SNR = 20 corresponded to quality data with less noise, the data with SNR = 10 corresponded to a medium level of noise and SNR = 5 indicated very noisy data. The performance of ICA FD, dynamic ICA FD and modified ICA FD strategies was compared with the PCA-FD strategy through two case studies, i.e., the quadruple tank process and the distillation column process. The FAR and MDR metrics were used to assess the monitoring performances of different strategies. It was observed that the value of MDR as well as FAR became increasingly higher with

increased level of noise in the data for all the FD strategies. The MDR value for the PCA strategy was very high, and it further increased with increased levels of noise. In contrast, the ICA-based FD strategies had much lower values of MDR for different noise levels. Hence, it can be inferred that ICA, dynamic ICA and modified ICA strategies were more effective to deal with sensor faults in comparison to the PCA strategy for different levels of noise. For the future, the proposed strategy can be extended for considering wavelet-based ICA for handling process plants with multi-resolution data.

Funding Open access funding provided by Manipal Academy of Higher Education, Manipal

Fig. 11 DC process: monitoring results of MICA with drift fault for SNR = 10

Table 4 DC process: MDR & FAR for drift fault

Indicators	SNR = 20	SNR = 10	SNR = 5
<i>MDR</i>			
PCA- T^2	16.03	20.48	26.41
PCA-Q	32.61	69.74	80.81
ICA- I_d^2	4.77	6.45	14.89
ICA- I_e^2	13.57	15.13	16.10
ICA-SPE	6.45	11.09	12.04
DICA- I_d^2	5.34	10.75	15.37
DICA- I_e^2	6.51	14.51	20.08
DICA-SPE	6.23	13.87	23.38
MICA- T^2	6.20	7.66	9.45
MICA-SPE	7.14	10.49	12.16
<i>FAR</i>			
PCA- T^2	0	0	0
PCA-Q	0	0	0
ICA- I_d^2	0.22	0.4	0.8
ICA- I_e^2	0.02	0.17	4.37
ICA-SPE	0	0	2.58
DICA- I_d^2	11.37	15.45	14.89
DICA- I_e^2	13.51	16.56	16.10
DICA-SPE	12.13	14.29	12.04
MICA- T^2	0	0	0.06
MICA-SPE	0	0	0.06

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material

in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ajami, A., & Daneshvar, M. (2012). Data driven approach for fault detection and diagnosis of turbine in thermal power plant using independent component analysis (ICA). *Electrical Power and Energy Systems*, 43, 728–735.
- Alaiddin, Md, Khan, F., Imtiaz, S., & Ahmed, S. (2018). A bibliometric review and analysis of data-driven fault detection and diagnosis methods for process systems. *Industrial and Engineering Chemistry Research*, 57(32), 10719–10735.
- Cai, L., & Tian, X. (2014). A new fault detection method for non-Gaussian process based on robust independent component analysis. *Process Safety and Environmental Protection*, 92(6), 645–658.
- Harkat, M.-F., Kouadri, A., Fezai, R., Mansouri, M., Nounou, H., & Nounou, M. (2020). Machine learning-based reduced kernel PCA model for nonlinear chemical process monitoring. *International Journal of Control, Automation and Electrical Systems*, 31, 1196–1209.
- Harrou, F., Nounou, M. N., Nounou, H. N., & Madakyaru, M. (2013). Statistical fault detection using PCA based GLR hypothesis testing. *Journal of Loss Prevention in the Process Industries*, 26(1), 129–139.
- Harrou, F., Sun, Y., Madakyaru, M., & Bouyedou, B. (2018). An improved multivariate chart using partial least squares with continuous ranked probability score. *IEEE Sensors Journal*, 18(16), 6715–6726.
- Hyvarinen, A. (2013). Independent component analysis: Recent advances. *Philosophical Transactions of the Royal Society*, 371(20110534), 1–19.
- Hyvarinen, A., & Oja, E. (2000). Independent component analysis: Algorithm and applications. *Neural Networks*, 13(4–5), 411–430.

- Johansson, K. H. (2000). The quadruple-tank process: A multivariable laboratory process with an adjustable zero. *IEEE Transactions on Control Systems Technology*, 8(3), 456–465.
- Kammammettu, S., & Li, Z. (2019). Change point and fault detection using Kantorovich Distance. *Journal of Process Control*, 80, 41–59.
- Lee, J. M., Joe Qin, S., & Lee, I. B. (2006). Fault detection and diagnosis based on modified independent component analysis. *AIChE Journal*, 52(10), 3501–3514.
- Lee, J. M., Qin, S. J., & Lee, I. B. (2004). Statistical monitoring of dynamic processes based on dynamic independent component analysis. *Chemical Engineering Science*, 59(14), 2995–3006.
- Lee, J. M., Qin, S. J., & Lee, I. B. (2007). Fault detection of non-linear processes using kernel independent component analysis. *The Canadian Journal of Chemical Engineering*, 85(4), 526–536.
- Madakyaru Muddu, Nounou Mohamed N., Nounou Hazem N. (2013). Enhanced modeling of distillation columns using integrated multiscale latent variable regression. In *IEEE symposium on computational intelligence in control and automation (CICA)* (pp. 74–80).
- Nor, N. M., Hasan, C. R. C., & Hussain, M. A. (2020). A review of data-driven fault detection and diagnosis methods: Applications in chemical process systems. *Reviews in Chemical Engineering*, 36(4), 513–553.
- Ramakrishna Kini, K., & Madakyaru, M. (2020). Improved anomaly detection based on integrated multi-scale principal component analysis using wavelets: An application to high dimensional processes. *IFAC-PapersOnLine*, 53(1), 398–403.
- Tian, X., Zhang, X., Deng, X., & Chen, S. (2009). Multiway kernel independent component analysis based on feature samples for batch process monitoring. *Neurocomputing*, 72(7–9), 1584–1596.
- Tong, C., Lan, T., & Shi, X. (2017). Ensemble modified independent component analysis for enhanced non-Gaussian process monitoring. *Control Engineering Practice*, 58, 34–41.
- Venkatasubramanian, V., Rengaswamy, R., Yin, K., & Kavuri, S. N. (2003). A review of process fault detection and diagnosis part 1: Quantitative model based methods. *Computers and Chemical Engineering*, 27(3), 293–311.
- Wang, L., & Shi, H. (2010). Multivariate statistical process monitoring using an improved independent component analysis. *Chemical Engineering Research and Design*, 88(4), 403–414.
- Yin, S., Ding, S. X., Xie, X., & Luo, H. (2014). A review on basic data-driven approaches for industrial process monitoring. *IEEE Transactions on Industrial Electronics*, 61(11), 6418–6428.
- Yuan, X., Shen, S.-Q., He, Y.-L., & Zhu, Q.-X. (2019). A novel hybrid method integrating ICA-PCA with relevant vector machine for multivariate process monitoring. *IEEE Transactions on Control Systems Technology*, 27(4), 1780–1787.
- Zhang, Y., & Ma, C. (2012). Decentralized fault diagnosis using multiblock kernel independent component analysis. *Chemical Engineering Research and Design*, 90(5), 667–676.
- Zhang, Y., & Zhang, Y. (2010). Fault detection of non-Gaussian processes based on modified independent component analysis. *Chemical Engineering Science*, 65(16), 4630–4639.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.