Received October 29, 2020, accepted November 5, 2020, date of publication November 16, 2020, date of current version November 24, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3037730

Improved Process Monitoring Strategy Using Kantorovich Distance-Independent Component Analysis: An Application to Tennessee Eastman Process

K. RAMAKRISHNA KINI[1] AND MUDDU MADAKYARU[2]

[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

Corresponding author: Muddu Madakyaru (muddu.m@manipal.edu)

ABSTRACT Vowing to the increasing complexity in industrial processes, the need for safety is of highest priority and this has led to development of efficient fault detection (FD) methods. Also, with rapid development of data acquisition systems, process history based methods have gained importance as their dependency is on large volume of sensor data extracted from the process. The industrial data exhibits some degree of non-gaussianity for which Independent Component Analysis (ICA) technique has usually been applied in practice. Recently, a new fault indicator based on Kantorovich Distance (KD) has been proposed which computes distance between two distributions and uses the distance as an indicator of fault. The KD metric has found to provide good monitoring results for data in presence of noise and offers enhanced detection of small magnitude faults. Considering the benefits offered by KD metric, the objective of this work is to amalgamate KD metric with ICA modeling framework to have a fault detection strategy that can improve process monitoring in noisy environment. The proposed ICA-KD FD strategy is illustrated on four processes that includes Modified Continuous Stirred Tank Heater (CSTH), Tennessee Eastman (TE) process and Experimental Distillation Column Process. The simulation results indicate that the proposed FD strategy exhibits improved performance over conventional strategies while monitoring different sensor faults in noisy environment.

INDEX TERMS Process monitoring, fault detection, independent component analysis, Kantorovich distance, small magnitude faults, Tennessee Eastman process, experimental distillation column process, modified continuous stirred tank heater process.

I. INTRODUCTION

In present day process industries, the need for having safe as well as reliable work environment is critical because of continuous and complex day-to-day process operations [1], [2]. With more demand for chemical products in the market, process plants are expanding their working base and this has lead to new process units being added to cater the continuous needs. The addition of new process units has increased the overall complexity which has further increased probability of abnormalities in the process which includes sensor aging, valve being stuck at a constant position, catalyst deactivation, motor/compressor failures etc. Also, since most of the process operations are still monitored by human supervision, the probability of abnormalities being not noticed is large. All this has resulted in many on-board accidents and degraded the product quality which has ultimately lead to process industries incurring heavy financial losses year after year. In order to avoid the above mentioned concerns, it is essential to continuously monitor the chemical plant and this has increased the demand of having fault detection and diagnosis (FDD) systems [3]. The main task of an FDD system is to efficiently capture unacceptable abnormalities in the process and identify its root cause. If efficient FDD systems are installed in chemical process industries, financial losses can be reduced, safety can be guaranteed and product quality can be improved.

The FDD systems are broadly classified into three subgroups: model-based methods, knowledge-based methods and historical data-based methods [4]. The advantages of the

The associate editor coordinating the review of this manuscript and approving it for publication was Dazhong Ma.

three methods have been very well demonstrated in various industrial applications for the last two decades. The model-based methods rely on developing reliable mathematical models of a process through the use of first principles. On the other hand, the knowledge based systems require information of the present process plant and expert plant operator experience. Since modern processes consists of large number of process variables, developing models in both these methods is challenging as well as time-consuming task. The third category of FDD methods is historical data-based methods that require only large amount of process data for capturing information in the process. Due to the significant advancements in the field of online instrumentation and frequent usage of data acquisition systems, large amount of sensor data can be recorded very easily in a short span of time. Since data is available easily, most FDD related industrial applications are based on historical data methods. Though data is available quickly, implementation is no easy task and has its own challenges while monitoring real time applications.

The practical application of historical data based methods is through either uni-variate or multi-variate statistical techniques. The uni-variate methods are very easy to implement but are restricted in monitoring only one variable at a time. Hence, application of uni-variate methods in real-time processes involving multivariate data is not advisable. To deal with multivariate processes, multi-variate statistical based process monitoring techniques have been widely applied in semiconductor, food, chemical, electrical and biology related process in the last few decades [5]. These multi-variate techniques possess ability to analyze high dimensional complex noisy data and extract useful information via few latent variables. Since chemical process industries involve large sensor measurements, using all the measurements for analysis results in poor estimation and there is possibility of noise being retained. Hence, in multi-variate techniques like Principal component analysis (PCA) and Partial Least Squares (PLS), original measured variables are transformed into few immeasurable variables or the latent variables by keeping the information intact [6], [7]. The PCA and PLS multi-variate strategies has been applied in many fault detection problems in last few years [8]–[10]. Despite having good monitoring performance, PCA based methods perform poorly for data which exhibits non-linearity and non-gaussian behavior. The reason is that these methods rely on second order statistical parameters and perform modeling with the assumption that data is gaussian in nature.

The multi-variate method based on Independent Component Analysis (ICA) has received good attention vowing to its ability of extracting information from non-gaussian and non-linear industrial data. ICA disintegrates observed data into latent variables or independent components that are independent and non-gaussian in nature [11], [12]. The ICA method de-correlates the data by reducing higher order statistical dependencies and focuses on making distribution of the disintegrated data as independent as possible. This enables ICA to have more meaningful extraction of information from data in comparison to methods based on PCA or PLS. The ICA based multi-variate method employes higher order statistical dependencies based on either Negentropy during the model development stage. Many extensions of ICA have been developed in the literature. The multi-way ICA strategy was developed for monitoring batch processes and provided better performance in comparison to multi-way PCA [13]. A non-linear process monitoring strategy based on multi-way Kernel ICA further enhanced the monitoring performance for batch processes [14]. The Dynamic ICA strategy was introduced for incorporating dynamics of non-gaussian process data through lagged variables and showed improved monitoring performance [15]. The Weighted ICA strategy used change rate of I^2 statistic for evaluating importance of each independent component which improved monitoring results since the missed detections were reduced [16]. To enhance fault detection where dominant independent components are extracted from multi-variate data and to eliminate the need of initializing de-mixing matrix to a random value, a modified ICA strategy was developed [17], [18]. The adverse effect of outliers in the data was reduced through use of Robust ICA based fault detection strategy where non-gaussian features were extracted through robust whitening and determining of maximum non-gaussian directions [19]. In recent years, artificial intelligence (AI) based methods have been applied for FD problems in different domains of engineering. The convolution neural network (CNN), which is composed hierarchically over many layers to obtain higher level features after each layer, has been applied for fault detection in mechanical domain [20], [21]. The Generative adversarial network (GAN) method has been applied in FD problems when there is imbalance in training and testing data [22]. Recently, the spectral theory of multi-dimensional matrix (STMM) technique has been proposed for event detection in power systems which come with large dimensional data [23]. Most of the AI based approaches for fault detection generally rely on self-learning from very large data sets.

In ICA fault detection strategy and its variant schemes, the monitoring of new process data is carried out with help of three fault detection indicies: I_d^2, I_e^2 and SPE. The I_d^2 monitors the systematic part of ICA model, I_e^2 monitors excluded part of the model and SPE monitors residual part of the model. Recently, a new fault detection metric based on Kantorovich Distance (KD) has been integrated with PCA strategy for enhancing fault detection. The KD metric, which has its roots from optimal mass transport theory, measures distance between two distributions and uses this distance as a measure of fault. It has been observed in the work by [24] that KD metric has provided very good monitoring results in presence of heavy measurement noise and also offered better detection of faults with small magnitude. The KD metric has been employed for change point detection problem where it yielded improved results with lesser false alarms and missed alarms [25]. Considering the attractive advantages offered by KD metric, we propose a novel fault detection strategy in which Kantorovich Distance (KD) metic is amalgamated with

ICA modeling scheme. The performance of conventional fault indicators of PCA as well as ICA strategies usually deteriorate in presence of noise. We feel that the proposed ICA-KD strategy can be very useful in monitoring industrial process data which are rich in noise and also non-gaussian in nature. The contributions of this paper are as follows:

- The Kantorovich Distance (KD) based statistical index is formulated as fault detection index.
- The novel Kantorovich Distance (KD) metric is integrated with non-Gaussian Independent Component Analysis (ICA) modeling framework to have ICA-KD based fault detection strategy. The KD metric is computed sample wise between ICA residuals of training fault free data and new online data using a moving window of fixed length.
- The performance of ICA-KD strategy is compared with conventional PCA and ICA fault detection strategies as well as recently developed PCA-KD strategy with benchmark set-ups and a pilot plant experimental set-up.
- The proposed ICA-KD strategy is able to enhance monitoring performance for non-gaussian data embedded with measurement noise. It is also able to precisely detect faults of smaller magnitude with smooth detection profile.
- The proposed ICA-KD strategy is effective in monitoring different sensor faults with an improved fault detection rate and minimum false alarm rate which clearly shows the robustness of the proposed scheme.

The following section briefly reviews the PCA and ICA techniques used for multi-variate process monitoring along with the conventional fault indicators. In section 3, we present the concept of optimal mass transport and the formulation of Kantorovich distance (KD) metric for discrete distributions. Next, the ICA-KD fault detection strategy is proposed by integrating ICA multi-variate technique with the KD metric in section 4. The section 5 reviews performance of proposed ICA-KD fault detection strategy with help of four case studies: a dynamic multi-variate process, a modified CSTH process, a benchmark Tennessee Eastman Process and an experimental Distillation Column Process. The proposed strategy is compared with PCA based KD strategy as well as conventional fault indicators of PCA and ICA methods. Finally, the paper ends with a brief discussion in the conclusion section.

II. MODEL DEVELOPMENT BASED ON PCA AND ICA

In the current section, a brief overview of PCA and ICA strategies along with their fault indicators are presented.

A. PRINCIPAL COMPONENT ANALYSIS

PCA is a conventional multi-variate method that reduces data dimension with an aim of eliminating cross correlation between the process variables. Consider multi-variate data, $\mathbf{Y} = [\mathbf{y_1}, \mathbf{y_2} \ldots, \mathbf{y_n}]$ where $\in \Re^{n \times m}$ where m and n represents number of variables and observations in $\mathbf{Y}$. After pre-processing $\mathbf{Y}$ to mean of zero and unity variance, singular value decomposition is used to estimate principal components [26]:

$$\mathbf{Y} = \mathbf{T}\mathbf{P}^T \tag{1}$$

where $\mathbf{T} = [\mathbf{t_1}, \mathbf{t_2} \ldots\ldots \mathbf{t_m}]$ is a matrix of PC's and $\mathbf{P} = [\mathbf{p_1}, \mathbf{p_2} \ldots . \mathbf{p_m}]$ is loading vector matrix.

Once the PCA model is developed by using CPV approach for optimum principal components selection, fault indicators T^2 and SPE are employed to check for faults in new process data [27]. The T^2 indicator captures important details from modeled part of PCA and is represented for a newly available data $\mathbf{Y_{new}}$ as:

$$T^2 = \mathbf{Y_{new}}^T \hat{\mathbf{P}} \hat{\mathbf{\Lambda}}^{-1} \hat{\mathbf{P}}^T \mathbf{Y_{new}} \tag{2}$$

where $\hat{\mathbf{\Lambda}}$ and $\hat{\mathbf{P}}$ are matrices corresponding to the retained PC's. The Q or SPE statistics captures information details from residual part of PCA and is represented as:

$$Q = \mathbf{Y_{new}}^T (\mathbf{I} - \hat{\mathbf{P}}\hat{\mathbf{P}}^T) \mathbf{Y_{new}} \tag{3}$$

For $\mathbf{Y_{new}}$, if fault indicators exceed the threshold limits, fault is declared [28].

B. INDEPENDENT COMPONENT ANALYSIS

ICA is a multi-variate technique which uses higher order statistics for extracting non-gaussian latent variables from a multi-variate data. An industrial multi-variate data, $\mathbf{X} = [\mathbf{x_1}, \mathbf{x_2} \ldots, \mathbf{x_n}]^T$ where $\in \Re^{m \times n}$ is represented as linear combination of $k (\leq m)$ unknown independent components (IC's) [29]:

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{E} \tag{4}$$

k represents the IC's, $\mathbf{A} = [\mathbf{a_1} \ldots . \mathbf{a}_k]^T \in \Re^{m \times k}$ is deterministic mixing matrix, $\mathbf{S} = [\mathbf{s_1}, \mathbf{s_2} \ldots . \mathbf{s}_n]^T \in \Re^{k \times n}$ is the matrix with IC's and $\mathbf{E} \in \Re^{m \times n}$ is a residual matrix. The main aim is focused in finding a separating matrix $\mathbf{W}$ such that reconstructed matrix is:

$$\hat{\mathbf{S}} = \mathbf{W}\mathbf{X} \tag{5}$$

The process variables have to be whitened to remove any cross-correlation which is performed by singular value decomposition (SVD) on co-variance of $\mathbf{X}$ and the transformation is expressed by:

$$\mathbf{Z} = \mathbf{Q}\mathbf{X} = \mathbf{B}\mathbf{S} \tag{6}$$

where $\mathbf{Q} = \Lambda^{\frac{-1}{2}} \mathbf{P^T}$ and $\mathbf{B} = \mathbf{Q}\mathbf{A}$ is an orthogonal matrix. From equation (6), the $\mathbf{S}$ can be estimated as follows :

$$\hat{\mathbf{S}} = \mathbf{B^T}\mathbf{Q}\mathbf{X} \tag{7}$$

The equations (5) and (7) yield a relationship between $\mathbf{W}$ and $\mathbf{B}$ which is expressed as:

$$\mathbf{W} = \mathbf{B^T}\mathbf{Q} \tag{8}$$

where the separating matrix $\mathbf{W} = [w_1, w_2 \ldots . w_m]$. Based on approximate form of negentropy, an efficient fixed-point algorithm for ICA has been developed to estimate **B** and the corresponding IC's can be evaluated using equation (8) once **B** is available. Next, the CPV method is employed to determine optimum IC's. The L_2 norm method is used to initially sort rows of separating **W** and then CPV method is applied [30].

The following are three different parts of ICA model: systematic space which represents model with respect to *k* dominant IC's, excluded space representing model with respect to ignored *m-k* IC's and residual space. The ICA based process monitoring consists of three fault detection indicators. The fault indicator I_d^2 monitors systematic part, I_e^2 monitors the excluded part and SPE indicator monitors the residual part of ICA model. The fault indicators at sample 'i' are represented mathematically as [30]:

$$I_d^2(i) = \mathbf{f}_{new\ k}^T(i)\mathbf{f}_{new\ k}(i) \tag{9}$$

$$I_e^2(i) = \mathbf{f}_{new\ m-k}^T(i)\mathbf{f}_{new\ m-k}(i) \tag{10}$$

$$SPE(i) = (\mathbf{X}_{new}(i) - \hat{\mathbf{X}}_{new}(i))^T(\mathbf{X}_{new}(i) - \hat{\mathbf{X}}_{new}(i)) \tag{11}$$

where $\mathbf{f}_{new\ k} = \mathbf{W}_k\mathbf{X}_{new\ k}$, $\mathbf{f}_{new\ m-k} = \mathbf{W}_{m-k}\mathbf{X}_{new\ m-k}$ are computed for a new data $\mathbf{X}_{new}$ with $\mathbf{W}_k$ is a matrix obtained by selection of *k* rows, $\mathbf{W}_{m-k}$ is a matrix obtained by selection of excluded *m-k* rows of the separating matrix **W** and $\mathbf{X}_{new} = \mathbf{Q}^{-1}\mathbf{B}_k\mathbf{f}_{new}(i) = \mathbf{Q}^{-1}\mathbf{B}_k\mathbf{W}_k\mathbf{X}_{new}(i)$. The threshold for ICA fault indicators is computed using Kernel Density Estimation (KDE) technique [15].

III. OPTIMAL MASS TRANSPORT AND KANTOROVICH DISTANCE

Many applications in engineering domain are based on mathematical modeling and extraction of information from the data source. Few applications in this regard are human identification through voice or face in image processing, distinguishing between good and severe tumors in medical applications, detecting abnormalities in process domain and mechanical applications [31]. Recently, methods based on the mathematics of optimal mass transport have received good attention due to their ability in having spatial representation while comparing different data sources. The optimal mass transport problem finds an efficient way of relocating a mass from a source distribution to a destination distribution with respect to a cost function. Optimal mass transport or Earth mover's distance have found many important applications in the field of machine learning, detection of cancers and various image processing problems [32]–[34].

A. KANTOROVICH DISTANCE

Kantorovich Distance (KD) is defined as the mode of transforming a mass of data from one distribution to other relative to a cost function [25], [35]. The advantage with KD metric is that it ensures optimal solution for linear programming problem. Consider two distributions **C** and **D** with supports *e* and *u* respectively. The minimum transfer or transportation distance required for transporting data elements of distribution **C** to distribution **D** is the Kantorovich Distance. If the two distributions are similar, minimum cost for transportation will be very less. If the two distributions are dissimilar, minimum cost for transportation will be large indicating a degree of dissimilarity between the two. Since process monitoring relies on detecting sensor faults based on evaluating the dissimilarity between training and testing data sets, the KD metric suits well in fault detection problems. The cost function for KD metric representation is defined based on the norm:

$$c(e, u) = ||e - u||^p \tag{12}$$

The norm can be selected based on the required application, for example, $||e - u||_1 = \sum_{i=1}^{o} |e_i - u_i|$ or $||e - u||_2 = \sum_{i=1}^{o}(|e_i - u_i|)^2$ or any other form. Using the above represented cost function, KD metric or the p-Wasserstein Distance between two distributions **C** and **D** is represented as:

$$W_p(C, D) = \left(\underset{\gamma \in \Gamma(C,D)}{inf} \int ||e - u||_d^p \gamma(e, u) \right)^{1/p} \tag{13}$$

where $\Gamma(C, D)$ represents set of joint distributions, γ denotes minimum optimal coupling and for the case when p = 1, the Wasserstein distance reduces to Earth mover's distance. The closed form expression for 2-Wasserstein Distance has been developed for gaussian distributions [36]. For two n-dimensional random variables *e* and *u* from two distributions **C** and **D** respectively, the KD metric is expressed as:

$$W_2(C, D) = \left(||\mu_e - \mu_u||^2 2 + Tr\left(\Sigma_e + \Sigma_u - 2(\Sigma_e^{\frac{1}{2}} \Sigma_u \Sigma_e^{\frac{1}{2}})^{\frac{1}{2}}\right) \right)^{\frac{1}{2}} \tag{14}$$

where μ_e and μ_u are the means while Σ_e and Σ_u are the covariance matrices of random variables *e* and *u* respectively. The above closed-form expression holds good for any two distributions in which there is relationship between their covariance matrices. The equation (14) is one way of computing the KD index. An alternate method available for computing the KD index can be found in [24].

B. COMPUTATIONAL IMPLEMENTATION OF KANTOROVICH DISTANCE

Kantorovich Distance is the minimum distance or optimal cost incurred while mapping segments of data between two distributions. When the KD metric is applied on continuous time series data, implementation is possible by considering each observation as one entity or element in both the distributions. However, stacking the data observations into segments and evaluating the segments in both the distributions gives a more smoother transition of KD metric. The computation of KD metric depends on the number of segments in each distribution (*r*) and number of samples in each segment (moving window *j*). Let us consider a study for evaluating the KD metric between two distributions **C** and **D**. The steps involved are:

1) The observations of distribution **C** are divided into *r* segments $C1$, $C_2 \dots C_r$ with each segment having

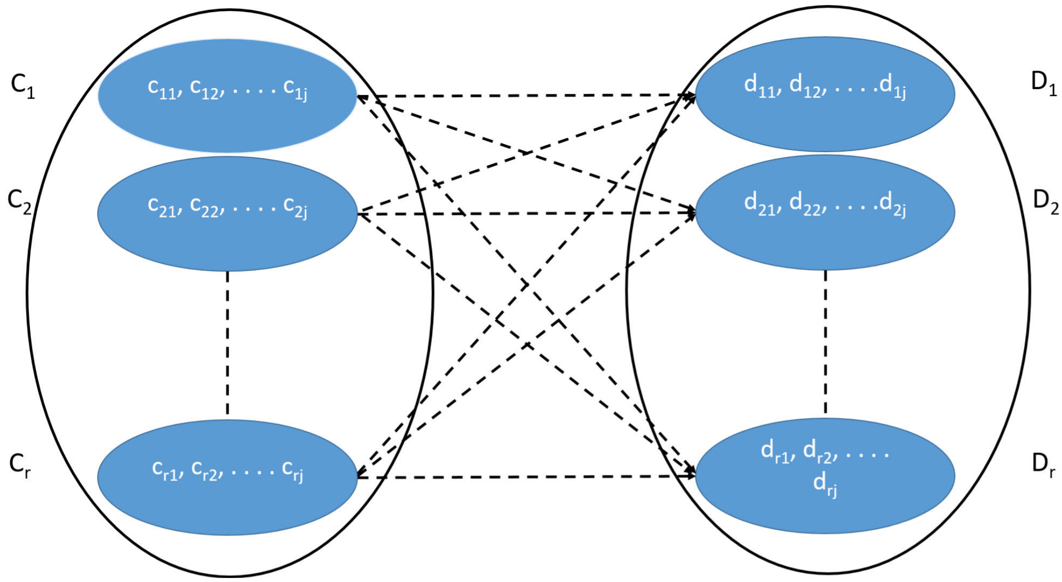

FIGURE 1. Computational implementation of Kantorovich Distance (KD).

j data points i.e $C_1 = [c_{11}, c_{12} \ldots . c_{1j}], C_2 = [c_{21}, c_{22} \ldots . c_{2j}] \ldots C_r = [c_{r1}, c_{r2} \ldots . c_{rj}]$.

2) The observations of distribution **D** are divided into r segments D_1, $D_2 \ldots D_r$ with each segment having j data points i.e $D_1 = [d_{11}, d_{12} \ldots . d_{1j}], D_2 = [d_{21}, d_{22} \ldots . d_{2j}] \ldots D_r = [d_{r1}, d_{r2} \ldots . d_{rj}]$.
3) The KD metric evaluation involves comparing each segment of distribution **C** with each segment of distribution **D** which implies that segment of source distribution C_1 is compared with all segments D_1, D_2, D_3 till D_r of destination distribution. The first sample of $C_1(c_{11})$ is compared with first sample of $D_1(d_{11})$, $D_2(d_{21}) \ldots D_r(d_{r1})$. The second sample of $C_1(c_{21})$ is compared with the second sample of $D_1(d_{12})$, $D_2(d_{22}) \ldots D_r(d_{r2})$ and so on until all the segments are covered.
4) Once comparison of all samples of C_1 with corresponding samples of D_1, D_2, $D_3 \ldots D_r$ is complete, next segment C_2 undergoes comparison with all segments D_1, D_2, $D_3 \ldots D_r$ of distribution **D**. This process will continue until all the segments of distribution **C** are compared with all segments of distribution **D**.
5) The distances of segment comparison between the two distributions is recorded and evaluated by the KD index. If distributions **C** and **D** are relatively similar, the KD index displays a minimum value. However, if the distributions **C** and **D** are differing, the KD index shows relatively larger value suggesting that the distributions are dissimilar. The computational implementation of the KD is presented in Figure 1. The smoothness of the KD metric depends mainly on the ratio of number of observations in each moving window to the total number of segments.

IV. THE PROPOSED ICA-KD BASED FAULT DETECTION STRATEGY

In this section, the ICA strategy is integrated with KD metric to develop a new fault detection scheme which can improve detection of unusual process conditions in presence of noise. Initially, an ICA model is developed from data corresponding to normal operating conditions of the plant and the model is used to detect unusual process conditions in a new data through the KD metric. The residuals generated from ICA model are crucial to determine the existence or non-existence of faults in a process data. Under normal operating conditions, the residuals of a supervised system are zero or close to zero. The residuals deviate away from zero and show a significant value in presence of an abnormality which can be easily differentiated from normal operating mode. For a given data **X**, ICA model is constructed and residuals are computed using the expression:

$$\mathbf{E} = \mathbf{X} - \hat{\mathbf{X}} = \mathbf{Q}^{-1}\mathbf{B_k}\mathbf{W_k}\mathbf{X} \tag{15}$$

where, $\mathbf{B_k}$ and $\mathbf{W_k}$ are the matrices generated by selecting only the dominant k rows of the matrices **B** and **W** respectively. The computation of ICA residuals is very important in the proposed ICA-KD strategy. The residuals are computed using the information from the matrices **B** and **W**. These two matrices are generated in the ICA model development stage using negentropy non-gaussian approximation from the FastICA fixed point algorithm. Hence, it may be noted that ICA residuals are generated using the non-gaussian information from the ICA modeling stage. Once ICA residuals are computed for training and testing data respectively, the KD metric is computed sample wise using closed form expression in equation (14) by utilizing the mean and co-variance of both residual spaces.The proposed ICA-KD fault detection strategy is divided mainly into two parts:

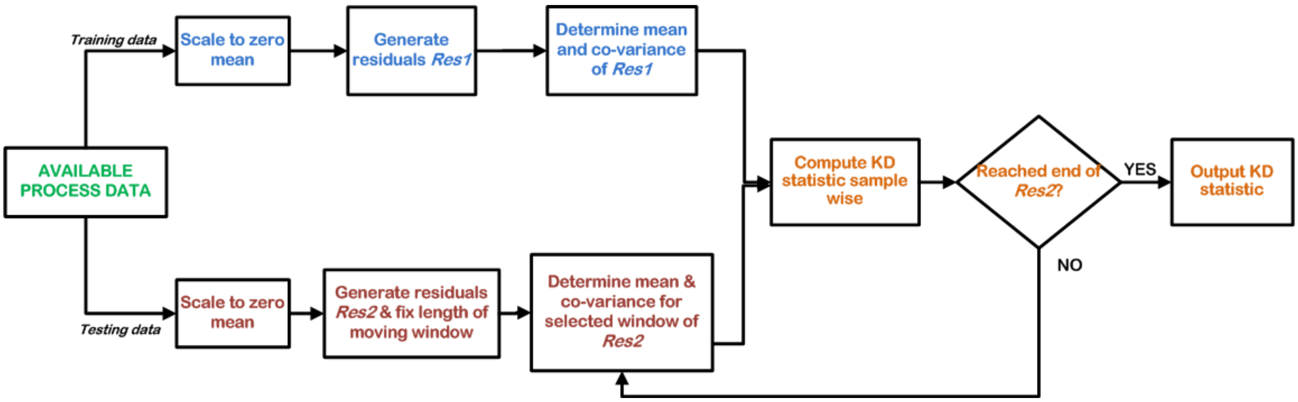

FIGURE 2. Calculation of Kantorovich Distance(KD).

1) Estimating the threshold from data under normal operating conditions (off-line)
2) Determining the fault in a new data using the KD index (on-line)

In fault detection problems, statistical fault indicators are compared with a threshold to take a decision about the fault. If the fault statistic stays within the determined threshold, the supervised system is said to behaving normally. However, the fault indicator statistic exceeding the value of determined threshold indicates the presence of fault in the system. Initially, the data collected from a system exhibiting normal operating conditions is split into two parts - T1 and T2. After scaling to zero mean, ICA models are developed separately for both T1 and T2 which is followed by obtaining residuals - R1 and R2 independently for both the data sets using equation 15. Next, to compute KD metric between the two residual sets, a moving window of **j** samples is employed on R2. Using the equation (14), Kantorovich Distance (KD), KDT is computed between two residuals R1 and R2. Finally, the mean and variance of the generated KDT is employed to generate the threshold which is given as:

$$Thres = \mu_{KDT} + 3\sigma_{KDT} \tag{16}$$

Once threshold is determined, it is used as a reference for determining presence of fault in a new data. The distance between two residuals is computed by KD index which takes a minimum value in absence of fault and shows a relatively large value in presence of fault. Once data under normal process operation is available - T_{norm}, it is scaled to have a mean of zero and an ICA model is constructed. Using this ICA model, residual R_{norm} is generated for T_{norm}. Now, when a new data T_{test} is available, the developed ICA model is used to generate residuals R_{test}. Next, Kantorovich Distance (KD) index KD is computed sample-wise between two residuals, R_{norm} and R_{test} by employing a fixed moving window j. Using the KD metric KD and threshold $Thres$, a decision is taken regarding the fault δ:

$$\delta = \begin{cases} \text{No fault} & \text{if KD} < Thres; \\ \text{fault} & \text{if KD} > Thres; \end{cases} \tag{17}$$

If the value of KD statistic KD is within the threshold $Thres$, the process is declared to be behaving normal. But the KD statistic KD exceeding the threshold $Thres$ indicates presence of fault in the process. The computation of Kantorovich distance statistic based on the two residuals is presented in Figure 2. For a available process data, the offline threshold calculation and on-line fault monitoring using the proposed ICA-KD strategy is presented in Figure 3.

V. CASE STUDIES

In this section, the performance of ICA-KD fault detection strategy is demonstrated through four case studies: (i) a dynamic multi-variate process, (ii) a modified CSTH process and (iii) a bench-mark Tennessee Eastman process(TEP) and (iv) Experimental Distillation Column Process. Three types of sensor faults have been considered in this work, namely, bias fault, intermittent fault and drift fault. The performance of proposed ICA-KD metric is contrasted against PCA-T^2, PCA-SPE, PCA-KD, ICA-I_d^2, ICA-I_e^2 and ICA-SPE fault indicators. In process monitoring problems, the performance of FD strategy is evaluated by two important factors such as False Alarm Rate (FAR) and Fault Detection Rate (FDR). FAR is defined as the ratio of total false alarms and the region of faultless data while FDR is the ratio of number of faults that exceed the control limits and the region of faults. For a good fault detection strategy, the FAR value should be minimum and the FDR value should be maximum.

A. DYNAMIC MULTI-VARIATE PROCESS

Consider a dynamic multi-variate process which is described mathematically [37]:

$$\mathbf{z}_a(i) = \begin{bmatrix} \mathbf{0.118} & -\mathbf{0.191} & \mathbf{0.28} \\ \mathbf{0.847} & \mathbf{0.264} & \mathbf{0.94} \\ -\mathbf{0.333} & \mathbf{0.514} & -\mathbf{0.21} \end{bmatrix} \times \mathbf{z}_a(i-1) + \begin{bmatrix} \mathbf{1} & \mathbf{2} \\ \mathbf{3} & -\mathbf{4} \\ -\mathbf{2} & \mathbf{1} \end{bmatrix} \times \mathbf{u}_a(i-1) \tag{18}$$

$$\mathbf{y}_a(i) = \mathbf{z}_a(i) + \mathbf{v}_a(i) \tag{19}$$

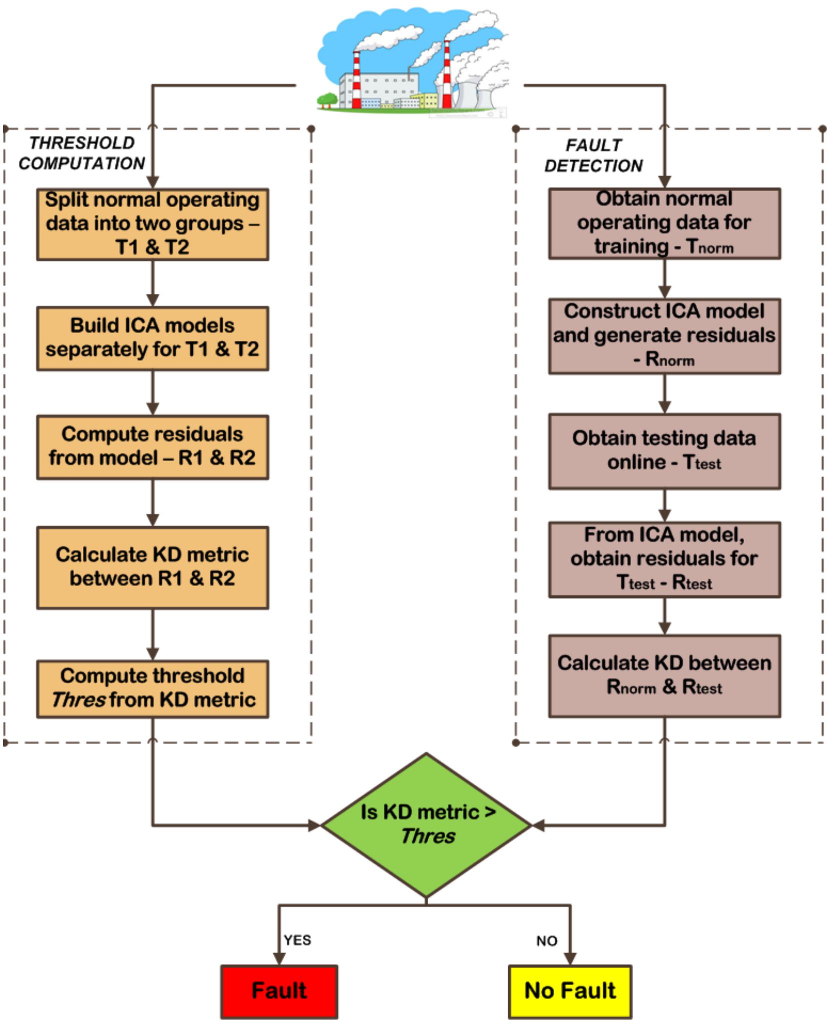

FIGURE 3. Block diagram representation of proposed ICA-KD fault detection strategy.

$$\mathbf{u}_a(i) = \begin{bmatrix} 0.811 & -0.226 \\ 0.477 & 0.415 \end{bmatrix} \times \mathbf{u}_a(i-1) + \begin{bmatrix} 0.193 & 0.689 \\ -0.320 & -0.749 \end{bmatrix} \times \mathbf{w}_a(i-1) \tag{20}$$

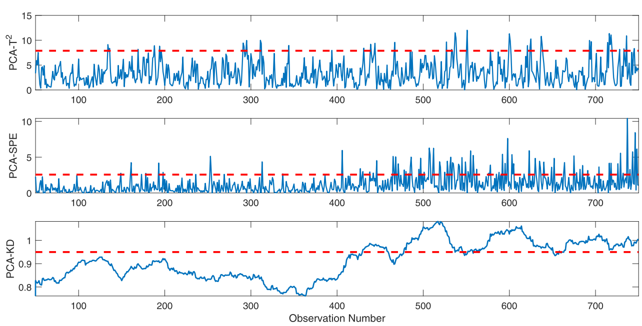

FIGURE 4. Dynamic multi-variate process: Time evolution of PCA fault detection indicies for bias fault.

The data set is generated using the MATLAB simulation. Here, u_a is correlated input, w_a is a random vector where each element has an uniform distribution in range (-1,1), output y_a is z_a plus random vector v_a. The resulting process dynamics is observed and further corrupted with measurement noise having mean zero and unit variance. The dynamic simulator is used to generate 1500 samples with five variables $[\mathbf{y}_{a1}, \mathbf{y}_{a2}, \mathbf{y}_{a3}, \mathbf{u}_{a1}, \mathbf{u}_{a2}]$ and this is split into 750 samples each

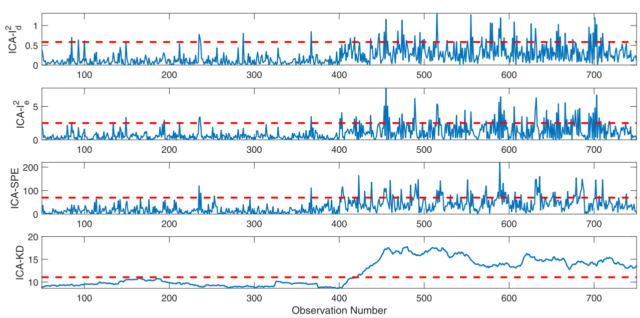

FIGURE 5. Dynamic multi-variate process: Time evolution of ICA fault detection indicies for bias fault.

TABLE 1. Dynamic multi-variate process: Fault detection rates(in percentage).

Faults	PCA-T^2	PCA-SPE	PCA-KD	ICA-I_d^2	ICA-I_e^2	ICA-SPE	ICA-KD
Bias	9.35	25.75	80.86	16.53	18.29	22.14	94.85
Intermittent	7.62	33.14	69.65	10.6	12.8	41.8	93.35
Drift	69	76.43	82.25	68.5	76.86	74.57	88.50

TABLE 2. Dynamic multi-variate process: False alarm rates(in percentage).

Faults	PCA-T^2	PCA-SPE	PCA-KD	ICA-I_d^2	ICA-I_e^2	ICA-SPE	ICA-KD
Bias	3	3.25	0	2.5	2	1	0
Intermittent	2.25	2.84	0	1.16	0.35	0.8	0
Drift	2.86	1.03	0	0.29	0.57	1.71	0

of training and testing data. After training data normalization, the PCA and ICA models are constructed with 4 optimum PC's and 3 IC's being retained respectively. A moving window of 50 is selected for this case study.

The detection capability of proposed ICA based KD fault detection is investigated on three types of sensor faults. First, a bias fault of small magnitude is introduced after sample 400 of the testing data. The monitoring results of PCA and ICA strategy with various fault indicies are presented in Figure 4 and 5 respectively. While T^2 and SPE indicies are unable to detect the fault, the PCA-KD metric shows better response but is unable to clearly detect the fault. The results from Figure 5 indicate that I_d^2, I_e^2 and SPE fault indicies also exhibit incomplete detection of fault. However, the proposed ICA-KD strategy provides a smooth and clear detection of the fault without any missed detections, thus having a clear advantage over other strategies. The superiority of proposed strategy is clearly visible in handling fault of small magnitude.

Next, an intermittent fault is introduced in sampling instants 80 to 170, 320 to 395 and 590 to 675 respectively. Finally, a sensor drift fault is introduced at sampling instant 350 of the testing data. The performance of different fault indicators in monitoring the faults evaluated through FDR & FAR indicies are presented in Table 1 and Table 2 respectively. The FDR and FAR results clearly indicate that ICA-KD strategy over-performs other strategies for dynamic process data which is corrupted with rich measurement noise. The proposed strategy was able to provide smooth transition profile without any missed detections while detecting small magnitude faults.

B. MODIFIED CSTH PROCESS

The modified CSTH process is built with two tanks (i.e., tank-1 and tank-2) connected in such a way that it has interacting effect in-terms of mass and energy. Both the tanks are heated with thyristor power controlled heaters. The cold water enters tank-1 and overflows to tank-2. The fraction of exit flow from tank-2 is recycled back to tank-2. In this study three manipulated inputs are cold water inlet flow rate to the tank-2, heater input of tank-1 and tank-2 respectively. The three measured outputs are temperatures in tank-1, tank-2 and height of the liquid in the tank-2. The schematic of modified CSTH process is shown in Figure 6. The governing dynamic model of modified CSTH process is described through the mathematical equations:

$$V_1\frac{dT_1}{dt} = F_1(u_1)(T_c - T_1) + F_R(u_3)(T_2 - T_1) + \frac{Q_1(u_4)}{\rho C_P} \tag{21}$$

$$\begin{aligned} A_2h_2\frac{dT_2}{dt} &= F_1(u_1)(T_1 - T_2) + F_2(u_2)(T_c - T_2) - F_R(u_3) \\ &\quad \times(T_1-T_2)+\frac{1}{\rho C_P}[Q_2(u_5)-2\Pi r_2 h_2 U(T_2-T_a)] \end{aligned} \tag{22}$$

$$A_2\frac{dh_2}{dt} = F_1(u_1) + F_2(u_2) - F_{out}(h_2) \tag{23}$$

$$F_{out}(h_2) = 10^{-4}\sqrt{0.406h_2^3+0.806h_2^2-0.01798h_2+0.1054} \tag{24}$$

Input and recycle flow rates of tank-1 and tank-2 are given as

$$F_1 = (42379u_1 - 456.85u_1^2 + 8.03684u_1^3) \times 10^{-11} \tag{25}$$

$$F_2 = (196620u_2 - 8796.8u_2^2+190.64u_2^3-1.294u_2^4)\times 10^{-11} \tag{26}$$

$$F_R = 2u_3 \times \frac{1}{3600} \times 10^{-3} \tag{27}$$

Two heat inputs to the tank-1 heater and tank-2 heater are given as

$$Q_1 = 7.9798u_4 + 0.9893u_4^2 - 7.3 \times 10^{-3}u_4^3 \tag{28}$$

$$Q_2 = 104 + 14.44u_5 + 0.95u_5^2 - 8 \times 10^{-3}u_5^3 \tag{29}$$

The data is generated using dynamic simulator from model parameters and steady state nominal operating conditions [38]. Simultaneous perturbation of the inputs using pseudo random binary signals(PRBS) from the idinput function of MATLAB yields the dynamic simulation data. The inflow to the tank-1 and inlet temperature to tank-1 treated as unmeasured disturbances and the measured outputs are corrupted with measurement noise with SNR = 5 to simulate real noisy measurement data. The dimension of data is 2800 × 8 which is split into 1400 samples each of training and testing

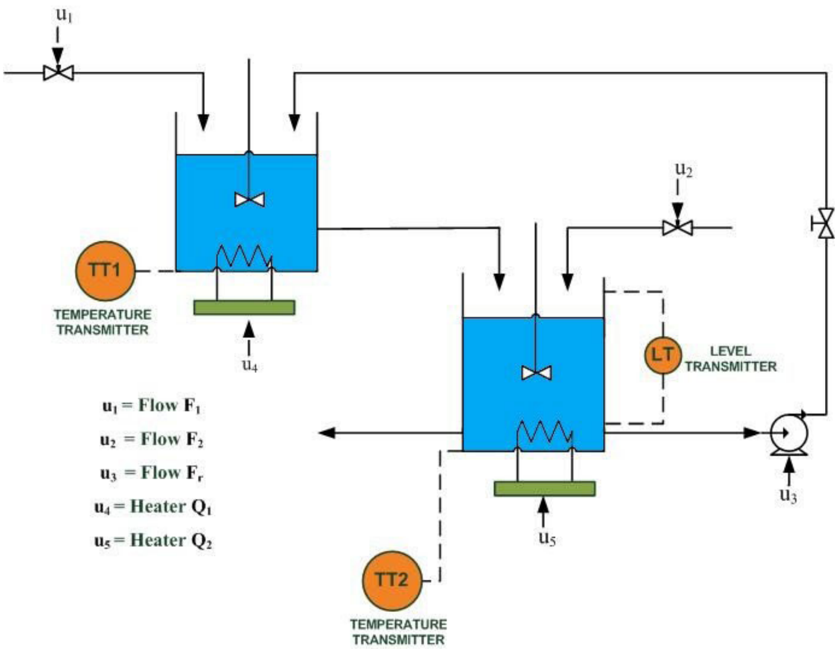

FIGURE 6. Modified CSTH process: A schematic diagram.

data. Post normalization of the training data, PCA and ICA models are constructed with 6 PC's and 6 IC's being retained respectively. A moving window of 50 is selected for this case study. The performance of proposed ICA-KD strategy in monitoring sensor faults for modified CSTH process is described in following section.

This subsection summarizes the results obtained in monitoring sensor faults through proposed ICA-KD fault detection strategy. Three fault scenarios are considered in this section. First, a bias fault of small magnitude is added at sample 350 followed by an intermittent fault which is introduced between samples 300 to 400, 825 to 925 and 1230 to 1330 respectively. Finally, a sensor drift fault is introduced at sampling instant 600 of the testing data. The performance of fault indicators in monitoring these faults evaluated through FDR & FAR matrices are presented in Table 3 and Table 4 respectively. From the tables, it can be inferred that the detection rate of ICA-KD is better in all the three scenarios.

To provide good clarity for the reader, performance of PCA as well as ICA fault indicators in monitoring bias and intermittent faults are presented in detail. The monitoring results of PCA and ICA methods with conventional fault indicators and KD indicies are illustrated in Figure 7 and Figure 8 respectively. As observed in the figures, the conventional PCA-T^2, PCA-SPE, ICA-I_d^2, ICA-I_e^2 and ICA-SPE fail to clearly detect the fault. While PCA-KD technique exhibits an improved response comparatively, it is still unable to clearly detect the fault at few sampling instants. In contrast, the proposed ICA-KD strategy provides enhanced performance with the time evolution providing a smooth detection of fault in the faulty region without any missed detections.

The monitoring results of PCA and ICA in monitoring an intermittent fault are illustrated in Figure 9 and Figure 10. From Figure 9, it is observed that PCA-T^2, PCA-SPE and PCA-KD strategies are unable to detect the fault. The

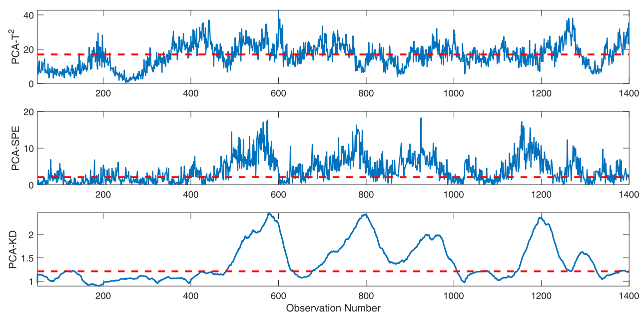

FIGURE 7. Modified CSTH process: Time evolution of PCA fault detection indicies for bias fault.

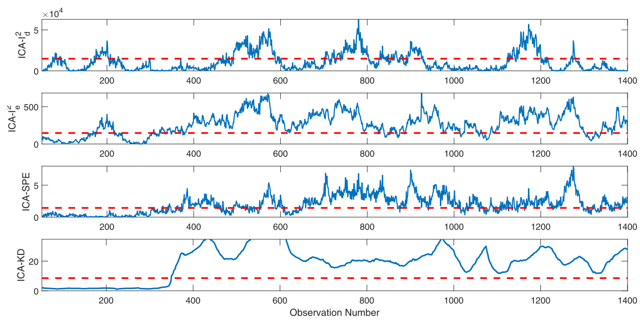

FIGURE 8. Modified CSTH process: Time evolution of ICA fault detection indicies for bias fault.

TABLE 3. Modified CSTH process: Fault detection rates (in percentage).

Faults	PCA-T^2	PCA-SPE	PCA-KD	ICA-I_d^2	ICA-I_e^2	ICA-SPE	ICA-KD
Bias	54	72.55	62.93	31.55	85.27	74	96.36
Intermittent	41	48	50	48	54.33	76.23	88
Drift	89.62	91.28	93.81	92.12	91.25	91.63	94.85

TABLE 4. Modified CSTH process: False alarm rates(in percentage).

Faults	PCA-T^2	PCA-SPE	PCA-KD	ICA-I_d^2	ICA-I_e^2	ICA-SPE	ICA-KD
Bias	9	29.67	4.43	16.33	16	0	0
Intermittent	2.45	19	0	8.55	1.88	0.88	0
Drift	4.33	22.33	0	12.33	21.5	1.05	0

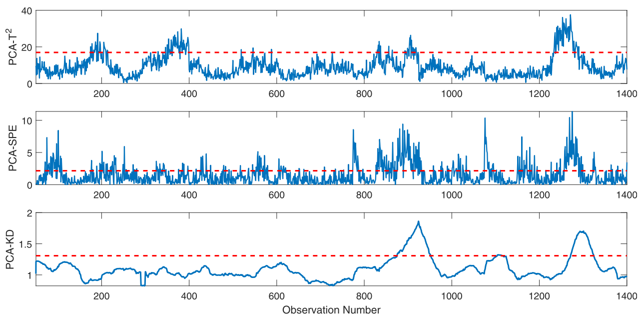

FIGURE 9. Modified CSTH process: Time evolution of PCA fault detection indicies for intermittent fault.

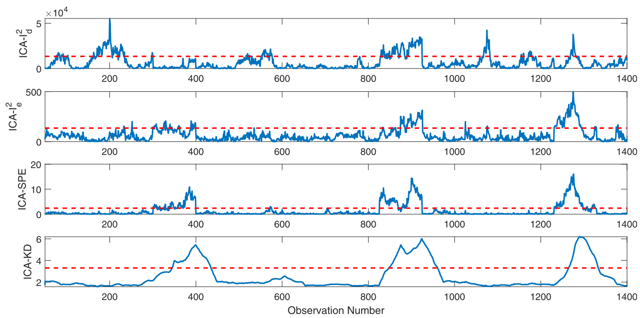

FIGURE 10. Modified CSTH process: Time evolution of ICA fault detection indicies for intermittent fault.

results from Figure 10 indicate that ICA-I_d^2 and ICA-I_e^2 detect the fault with many missed detections in the faulty region. In contrast, the proposed ICA-KD detects the fault clearly without any missed detections, thus providing better results than other strategies. The Modified CSTH data was corrupted with measurement noise and still the proposed ICA-KD was able to deliver better monitoring results as compared to other methods. Also, the ICA-KD method was sensitive to faults of smaller magnitude and hence provided better detection of such faults, thus demonstrating a very good advantage.

C. TENNESSEE EASTMAN PROCESS

In this section, the proposed ICA-KD based FD method is demonstrated on a benchmark Tennessee Eastman (TE) Process. Downs and Vogel developed this bench mark test in Eastman Chemical Company at Tennessee, U.S.A. This process has been considered as a benchmark in control, automation and process related research studies [27], [39], [40]. Through the use of four exothermic and irreversible first order reactions, two products and one by-product are generated by four gaseous reactants. The reactor combined with condenser and stripper of product, vapor liquid separator and recycle compressor comprises of 22 process measurements, 19 composition measurements and 12 manipulated variables. The process comprises of 21 fault scenarios with the training data consisting 500 measurements and testing data 960 measurements respectively. After processing the training data, PCA and ICA models are constructed with 39 PC's and 39 IC's being retained respectively. A moving window of 40 is selected for this case study. Three fault scenarios, namely, IDV(3), IDV(9) and IDV(15) have been excluded in this work since they give very poor FDR values [41].

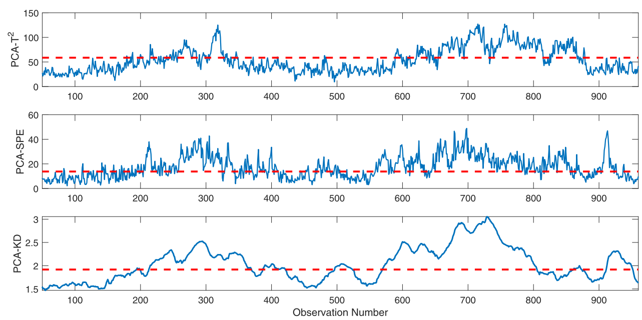

FIGURE 11. TE process: Time evolution of PCA fault detection indicies for IDV(10) fault.

The results of the proposed ICA-KD strategy in monitoring various fault scenarios of TE process are presented here. The FDR & FAR chart for different fault scenarios of TE process are presented in Table 5 and Table 6. Even though PCA-T^2 and PCA-SPE are able to detect the faults with good FDR value, it can be noted that they also tend to have a relatively high FAR value. The performance of PCA-KD strategy is also good and performs as good as conventional fault indicies of PCA. In contrast, it may be observed that the detection rate of

TABLE 5. TE Process: Fault detection rates (in percentage).

Faults	PCA-T^2	PCA-SPE	PCA-KD	ICA-I_d^2	ICA-I_e^2	ICA-SPE	ICA-KD
IDV(1)	98.83	99.15	99.36	98.38	97.62	99	99.5
IDV(2)	98.5	99	99.75	97.78	98	97.62	99.75
IDV(4)	75.56	97.5	99.31	81.63	83.87	45.25	99.5
IDV(5)	28.8	48.5	28.66	99.12	98.88	99.5	99.65
IDV(6)	98.25	99.75	100	100	100	100	100
IDV(7)	99.5	98.88	100	86.38	81.87	84.62	100
IDV(8)	97.26	93.64	95.16	87	86.5	86.5	97.36
IDV(10)	54.36	65.83	66.35	67.13	68.63	71.63	94.58
IDV(11)	64.59	70.32	72.74	49.13	48.13	22	80
IDV(12)	98.75	95	98.6	95.13	94.75	93.63	99.01
IDV(13)	95.39	95.64	95.03	93.37	91.25	86.38	96.69
IDV(14)	100	90.4	100	96.75	95.37	82.75	100
IDV(16)	37.41	66.46	56.33	82.87	83.5	80.13	99.24
IDV(17)	91.15	95.63	98.6	90.75	90.25	83.37	98.81
IDV(18)	90.65	90.25	89.43	89.62	89.88	89.38	90.5
IDV(19)	15.84	61.15	40.89	63.75	61.37	50.5	69.87
IDV(20)	52.49	70.45	90.32	80.63	77.25	63.5	91.08
IDV(21)	49.43	57.61	40.25	48	47.13	41.75	57.53

TABLE 6. TE Process: False alarm rates(in percentage).

Faults	PCA-T^2	PCA-SPE	PCA-KD	ICA-I_d^2	ICA-I_e^2	ICA-SPE	ICA-KD
IDV(1)	7.59	13.29	0	3.75	5	1.25	0
IDV(2)	6.96	6.33	0	3.13	3.75	2.5	0
IDV(4)	3.8	6.33	0	6.25	5	2.5	0
IDV(5)	3.68	6.01	0	4.37	6.25	3.15	0
IDV(6)	0	2.53	0	2.5	3.13	1.25	0
IDV(7)	0	1.9	0	1.87	1.87	1.25	0
IDV(8)	1	5.06	0	5.63	6.25	2.5	0
IDV(10)	2.53	5.06	0	3.1	2.5	0.6	0
IDV(11)	2.27	4.43	0	4.37	5	3.13	0
IDV(12)	3.16	3.8	0	3.75	2.5	0	0
IDV(13)	1.64	1.9	0	3.75	3.13	1.25	0
IDV(14)	2.53	5.7	0	6.25	3.13	4.37	0
IDV(16)	8.86	9.49	0	8.13	7.5	2.5	0
IDV(17)	2.53	5.7	0	6.25	5	0.63	0
IDV(18)	3.16	4.43	0	3.75	2.5	1.25	0
IDV(19)	0	6.33	0	6.25	6.25	0.5	0
IDV(20)	1.87	3.16	0	5.65	1.25	0.6	0
IDV(21)	6.33	7.59	0	10	11.25	4.37	0

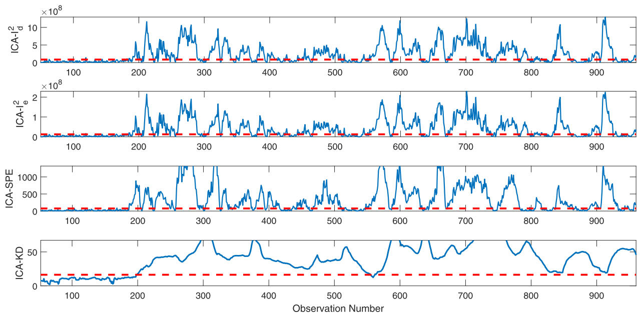

FIGURE 12. TE process: Time evolution of ICA fault detection indicies for IDV(10) fault.

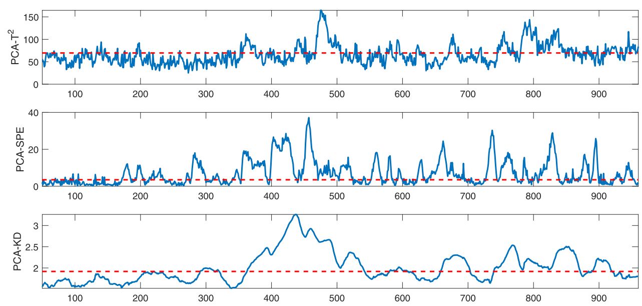

FIGURE 13. TE process: Time evolution of PCA fault detection indicies for IDV(16) fault.

ICA-KD is very good and slightly better in comparison with conventional fault indicators of ICA strategy. The ICA-KD strategy shows good advantage over other FD strategies for IDV(10), IDV(11), IDV(16), IDV(19), IDV(20) and IDV(21) fault scenarios. For providing clarity to the reader, graphical representations of two fault scenarios are presented in detail. The two faults considered here are: IDV(10) which is a ran-

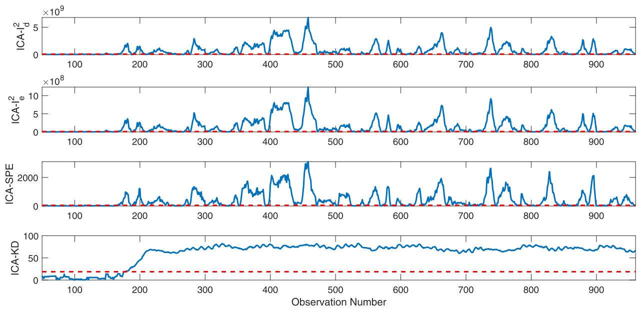

FIGURE 14. TE process: Time evolution of ICA fault detection indicies for IDV(16) fault.

TABLE 7. TE process: False alarm rate for IDV(0) fault-free case.

Fault indicators	FAR
PCA-T^2	5
PCA-SPE	6.35
PCA-KD	4.37
ICA-I_d^2	6.25
ICA-I_e^2	5.73
ICA-SPE	5.1
ICA-KD	0.71

dom variation fault in C feed temperature and IDV(16) which is defined as unknown in the process flowsheet.

The results of PCA and ICA fault detection strategies in monitoring IDV(10) fault is presented in Figure 11 and Figure 12 respectively. Since PCA-T^2, PCA-SPE and PCA-KD detection indicies have many missed detections in the faulty region, their detection performance is not acceptable. On the other hand, the ICA-I_d^2, ICA-I_e^2 and ICA-SPE strategies fare better but are unable to detect the fault precisely. The ICA-KD strategy delivers better results with precise detection of the fault with minimum missed detections. Next, Figure 13 and Figure 14 demonstrates the performance of PCA and ICA in detecting the IDV(16) fault. The conventional PCA indicators and PCA-KD strategy are unable to detect the fault. The performance of ICA-I_d^2, ICA-I_e^2 and ICA-SPE is little better but are incapable of completely detecting the fault. In contrast, the proposed ICA-KD strategy detects the fault neatly without any missed detections as observed in Figure 14. From the results, the FDR value of PCA-KD is little less for IDV (5), IDV (16), IDV (19) and IDV (21) fault scenarios in comparison with PCA-T^2 and PCA-SPE. The performance of KD metric depends on factors like number of optimum principal components selected, and the moving window size. From the selected parameters, the PCA-KD strategy has less performance for only the four fault scenarios of TE process. However, overall it was found that the PCA-KD over performs the conventional fault indicators of PCA strategy. An additional study was carried out to observe the false alarm rate of the proposed ICA-KD strategy on fault-free data (IDV(0)) of 960 observations and the corresponding results are tabulated in Table 7. It may be observed that all the strategies exhibit false alarms, but the ICA-KD strategy has a minimum value of FAR. Thus, from the monitoring results of TE process, it may be concluded that the proposed ICA-KD strategy displays enhanced detection of all faults with smooth detection profile, thus displaying a clear advantage over other methods.

D. EXPERIMENTAL DISTILLATION COLUMN PROCESS

The experimental bubble cap distillation column (DC) data is used to assess the performance of the proposed ICA-KD strategy. The distillation column (DC) is a high energy consuming unit in any chemical process plant and is used for separating components from the mixture of components based on their difference in vapor pressure. Proper monitoring of distillation column process is necessary to avoid any accident and loss of product quality in the industry. The data is generated from the experimental distillation column housed in Chemical Engineering department, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India.

The set-up presented in Figure 15 consists of two bubble caps placed on each tray and is made of stainless steel. The inner details of bubble caps are with weirs of 25mm height and a downcomer of 12 mm diameter. The total number of trays are 10 and spacing between each tray is about 120 mm and therefore the column runs up to a length of 1.5m with a diameter of 100mm. The vapor coming out of distillation column is condensed using shell and tube heat condenser. The liquid is boiled up using a re-boiler which is fitted with 2.0 kW Thyristor heater. The whole column is insulated to avoid any heat loss to the surroundings. The power to the heater is regulated using thyristor controller that accepts 4-20mA input. The reflux flow and feed flow rate can be regulated using rate-controlled metering pumps. To monitor the column temperatures, six resistance temperate detectors (RTD) sensors are fixed at different location of the column. Interfacing is completed between experimental column set-up and the control computer through a data acquisition system. The feed flow and re-boiler heat duty undergo regulation in order to maintain both top as well as bottom product compositions at some desired level. The binary mixture of methanol water system has been applied in this experimental work.

The data is generated from experimental distillation column by perturbing reflux and feed flow sequentially [42]. Once the DC column is at nominal condition, perturbation of feed flow is performed at a magnitude of 50 and reflux flow is kept constant at nominal condition. The column is driven towards nominal conditions and again perturbed using the reflux flow with a magnitude of 40 by maintaining a constant flow rate. The resultant changes in the outputs are recorded. A sampling time of 4 seconds has been considered in this study. The flow rate and feed flow, six temperatures and two outputs with 1024 observations are considered. Next, the generated data is split equally into training and testing data groups. To have a fair comparison, 6 optimum PCs as well as 6 optimum ICs are selected for both PCA and ICA strategies during the model development stage.

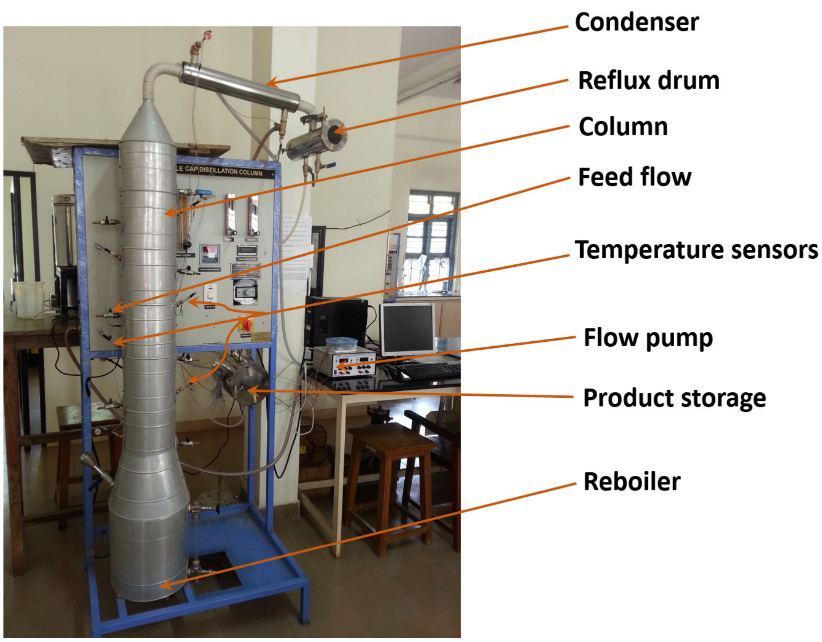

FIGURE 15. A pictorial view of experimental distillation column.

TABLE 8. Experimental DC process: Fault detection rates (in percentage).

Faults	PCA-T^2	PCA-SPE	PCA-KD	ICA-I_d^2	ICA-I_e^2	ICA-SPE	ICA-KD
Bias	25.67	21.7	81.25	78.18	80.94	77.17	96.06
Intermittent	44.64	32.55	94.45	97.64	97.27	97.13	99.31
Drift	81.04	81.28	83.95	83.87	82.49	89.5	94.2

TABLE 9. Experimental DC process: False alarm rates(in percentage).

Faults	PCA-T^2	PCA-SPE	PCA-KD	ICA-I_d^2	ICA-I_e^2	ICA-SPE	ICA-KD
Bias	44.48	20.78	10.67	22.33	28.67	6.67	1.97
Intermittent	28.45	8.48	13.7	14.04	18.14	11.99	2.25
Drift	25.67	18.67	10.67	12.67	8.33	17.53	1.75

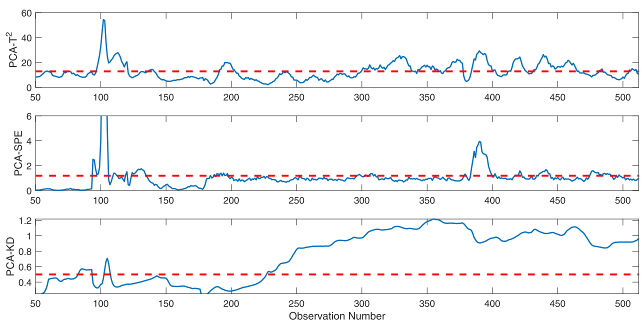

FIGURE 16. Experimental DC process: Time evolution of PCA fault detection indicies for bias fault.

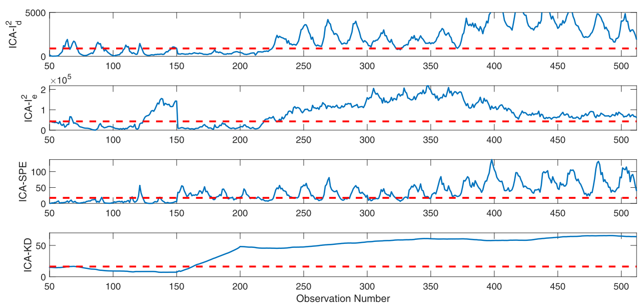

FIGURE 17. Experimental DC process: Time evolution of ICA fault detection indicies for bias fault.

In this case, the monitoring of sensor bias, sensor intermittent and sensor drift faults have been considered. Small magnitude of bias were introduced for bias as well as intermittent faults to assess effectiveness of proposed strategy. The performance of proposed ICA-KD as well as conventional methods evaluated through FDR and FAR is presented in Table 8 and Table 9. It may be observed from the results that the proposed ICA-KD fault detection strategy over performs other methods for all fault scenarios with better FDR value. There are small variations in the generated data due to which there is possibility of false alarms. The proposed ICA-KD strategy exhibits very minimum FAR value in comparison to other methods as seen from results table. The performance of ICA-KD strategy in monitoring a bias fault, which is introduced after sample 150 of testing data, is presented in Figure 16 and Figure 17. It may be observed that conventional indicators of PCA have high false alarms and they are not able to detect the small magnitude of fault. While PCA-KD shows better detection, it detects the fault with a small delay. The ICA based conventional fault indicators partially detect the fault with few false alarms. In contrast, the proposed ICA-KD fault detection strategy exhibits superior detection performance with good FDR value and minimum FAR value.

VI. CONCLUSION

In this paper, we explore the possibility of developing a fault detection strategy using ICA and Kantorovich Distance(KD). The ICA method has been used as modeling framework for fault detection using KD based hypothesis testing. The KD metric evaluates distance between two distributions, i.e, the ICA model residuals of data in fault free operating conditions and ICA residuals of obtained data on the faulty condition. The KD statistic is minimum in absence of fault, but displays significantly larger values when fault is present. To demonstrate the performance of ICA based KD fault detection strategy, four case studies were considered-(i) a dynamic multi-variate process, (ii) a modified CSTH process (iii) a benchmark Tennessee Eastman Process and (iv) Experimental Distillation Column Process. The performance of proposed ICA-KD strategy was compared with PCA-KD, PCA-T^2, PCA-SPE, ICA-I_d^2, ICA-I_e^2 and ICA-SPE strategies. For dynamic multi-variate process as well as modified CSTH process, the data was corrupted with noise and the proposed ICA-KD strategy was able to deliver good monitoring results for noisy data. It was also able to precisely detect faults of small magnitude easily with smooth detection profile. The proposed ICA-KD strategy was also evaluated for its performance on a benchmark TE process which consists of different types of fault scenarios and even in this case, the ICA-KD over-performed other strategies with a good detection rate. Hence, it can be concluded that the proposed ICA-KD clearly emerges as a better choice for formulating FD strategy when compared to recently proposed PCA-KD strategy as well as conventional fault indicators. Due to the advantage of offering good performance in presence of noise, the KD detection metric could be applied for monitoring industrial applications in various fields of engineering.

ACKNOWLEDGMENT

The authors would like to thank Manipal Institute of Technology (MIT), Manipal Academy of Higher Education (MAHE), Manipal, for providing support in carrying out this work.

REFERENCES

[1] P. Peng, Y. Zhang, F. Liu, H. Wang, and H. Zhang, "A robust and sparse process fault detection method based on RSPCA," *IEEE Access*, vol. 7, pp. 133799–133811, 2019.

[2] Z. Chen, Y. Cao, S. X. Ding, K. Zhang, T. Koenings, T. Peng, C. Yang, and W. Gui, "A distributed canonical correlation analysis-based fault detection method for plant-wide process monitoring," *IEEE Trans. Ind. Informat.*, vol. 15, no. 5, pp. 2710–2720, May 2019.

[3] K. R. Kini and M. Madakyaru, "Monitoring multivariate process using improved independent component analysis-generalized likelihood ratio strategy," *IFAC-PapersOnLine*, vol. 53, no. 1, pp. 392–397, Aug. 2020.

[4] N. Md Nor, C. R. Che Hassan, and M. A. Hussain, "A review of data-driven fault detection and diagnosis methods: Applications in chemical process systems," *Rev. Chem. Eng.*, vol. 36, no. 4, pp. 513–553, May 2020.

[5] Y. Zhang, "Enhanced statistical analysis of nonlinear processes using KPCA, KICA and SVM," *Chem. Eng. Sci.*, vol. 64, no. 5, pp. 801–811, Mar. 2009.

[6] M. Alauddin, F. Khan, S. Imtiaz, and S. Ahmed, "A bibliometric review and analysis of data-driven fault detection and diagnosis methods for process systems," *Ind. Eng. Chem. Res.*, vol. 57, no. 32, pp. 10719–10735, Jul. 2018.

[7] B. Wang, X. Yan, and Q. Jiang, "Independent component analysis model utilizing de-mixing information for improved non-Gaussian process monitoring," *Comput. Ind. Eng.*, vol. 94, pp. 188–200, Apr. 2016.

[8] J. Harmouche, C. Delpha, and D. Diallo, "Incipient fault detection and diagnosis based on Kullback–Leibler divergence using principal component analysis: Part I," *Signal Process.*, vol. 94, pp. 278–287, Jan. 2014.

[9] F. Harrou, Y. Sun, and M. Madakyaru, "Kullback-leibler distance-based enhanced detection of incipient anomalies," *J. Loss Prevention Process Industries*, vol. 44, pp. 73–87, Nov. 2016.

[10] A. Bakdi, W. Bounoua, A. Guichi, and S. Mekhilef, "Real-time fault detection in PV systems under MPPT using PMU and high-frequency multi-sensor data through online PCA-KDE-based multivariate KL divergence," *Int. J. Electr. Power Energy Syst.*, vol. 125, Feb. 2021, Art. no. 106457.

[11] M. Aljunaid, H. Shi, and Y. Tao, "Quality-related fault detection based on improved independent component regression for non-Gaussian processes," *IEEE Access*, vol. 7, pp. 158594–158602, 2019.

[12] Z. Li and X. Yan, "Adaptive selective ensemble-independent component analysis models for process monitoring," *Ind. Eng. Chem. Res.*, vol. 57, no. 24, pp. 8240–8252, May 2018.

[13] C. K. Yoo, D. S. Lee, and P. A. Vanrolleghem, "Application of multiway ICA for on-line process monitoring of a sequencing batch reactor," *Water Res.*, vol. 38, no. 7, pp. 1715–1732, Apr. 2004.

[14] X. Tian, X. Zhang, X. Deng, and S. Chen, "Multiway kernel independent component analysis based on feature samples for batch process monitoring," *Neurocomputing*, vol. 72, nos. 7–9, pp. 1584–1596, Mar. 2009.

[15] J.-M. Lee, C. Yoo, and I.-B. Lee, "Statistical monitoring of dynamic processes based on dynamic independent component analysis," *Chem. Eng. Sci.*, vol. 59, no. 14, pp. 2995–3006, Jul. 2004.

[16] Q. Jiang and X. Yan, "Non-Gaussian chemical process monitoring with adaptively weighted independent component analysis and its applications," *J. Process Control*, vol. 23, no. 9, pp. 1320–1331, Oct. 2013.

[17] J.-M. Lee, S. J. Qin, and I.-B. Lee, "Fault detection and diagnosis based on modified independent component analysis," *AIChE J.*, vol. 52, no. 10, pp. 3501–3514, Oct. 2006.

[18] C. Tong, T. Lan, and X. Shi, "Double-layer ensemble monitoring of non-Gaussian processes using modified independent component analysis," *ISA Trans.*, vol. 68, pp. 181–188, May 2017.

[19] L. Cai and X. Tian, "A new fault detection method for non-Gaussian process based on robust independent component analysis," *Process Saf. Environ. Protection*, vol. 92, no. 6, pp. 645–658, Nov. 2014.

[20] O. Janssens, V. Slavkovikja, B. Vervisch, K. Stockman, M. Loccufier, S. Verstockt, R. de Walle, and S. Hoecke, "Convolutional neural network based fault detection for rotating machinery," *J. Sound Vib.*, vol. 377, pp. 331–345, Sep. 2016.

[21] S. Wang, J. Xiang, Y. Zhong, and Y. Zhou, "Convolutional neural network-based hidden Markov models for rolling element bearing fault identification," *Knowl.-Based Syst.*, vol. 144, pp. 65–76, Mar. 2018.

[22] F. Zhou, S. Yang, H. Fujita, D. Chen, and C. Wen, "Deep learning fault diagnosis method based on global optimization GAN for unbalanced data," *Knowl.-Based Syst.*, vol. 187, Jan. 2020, Art. no. 104837.

[23] D. Ma, X. Hu, H. Zhang, Q. Sun, and X. Xie, "A hierarchical event detection method based on spectral theory of multidimensional matrix for power system," *IEEE Trans. Syst., Man, Cybern., Syst.*, early access, Aug. 9, 2019, doi: 10.1109/TSMC.2019.2931316.

[24] S. Kammammettu and Z. Li, "Change point and fault detection using Kantorovich distance," *J. Process Control*, vol. 80, pp. 41–59, Aug. 2019.

[25] B. M. S. Arifin, Z. Li, and S. L. Shah, "Change point detection using the Kantorovich distance algorithm," *IFAC-PapersOnLine*, vol. 51, no. 18, pp. 708–713, 2018.

[26] F. Harrou, M. N. Nounou, H. N. Nounou, and M. Madakyaru, "Statistical fault detection using PCA-based GLR hypothesis testing," *J. Loss Prevention Process Industries*, vol. 26, no. 1, pp. 129–139, Jan. 2013.

[27] T. Lan, C. Tong, H. Yu, and X. Shi, "Statistical monitoring for non-Gaussian processes based on MICA-KDR method," *ISA Trans.*, vol. 94, pp. 164–173, Nov. 2019.

[28] C. Zhang, Q. Guo, and Y. Li, "Fault detection in the tennessee eastman benchmark process using principal component difference based on K-Nearest neighbors," *IEEE Access*, vol. 8, pp. 49999–50009, 2020.

[29] J.-M. Lee, C. Yoo, and I.-B. Lee, "Statistical process monitoring with independent component analysis," *J. Process Control*, vol. 14, no. 5, pp. 467–485, Aug. 2004.

[30] Y. Xu, S.-Q. Shen, Y.-L. He, and Q.-X. Zhu, "A novel hybrid method integrating ICA-PCA with relevant vector machine for multivariate process monitoring," *IEEE Trans. Control Syst. Technol.*, vol. 27, no. 4, pp. 1780–1787, Jul. 2019.

[31] S. Kolouri, S. R. Park, M. Thorpe, D. Slepcev, and G. K. Rohde, "Optimal mass transport: Signal processing and machine-learning applications," *IEEE Signal Process. Mag.*, vol. 34, no. 4, pp. 43–59, Jul. 2017.

[32] J. A. Ozolek, A. B. Tosun, W. Wang, C. Chen, S. Kolouri, S. Basu, H. Huang, and G. K. Rohde, "Accurate diagnosis of thyroid follicular lesions from nuclear morphology using supervised learning," *Med. Image Anal.*, vol. 18, no. 5, pp. 772–780, Jul. 2014.

[33] Y. Rubner, C. Tomasi, and L. J. Guibas, "The earth mover's distance as a metric for image retrieval," *Int. J. Comput. Vis.*, vol. 40, no. 2, pp. 99–121, Nov. 2000.

[34] P. Li, Q. Wang, and L. Zhang, "A novel Earth Mover's distance methodology for image matching with Gaussian mixture models," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 1689–1696.

[35] L. Kantorovich, "On the translocation of masses," *Manag. Sci.*, vol. 5, no. 1, pp. 1–4, Oct. 1958.

[36] A. Takatsu, "Wasserstein geometry of Gaussian measures," *Osaka J. Math.*, vol. 48, no. 48, pp. 1005–1026, 2011.

[37] W. Ku, R. H. Storer, and C. Georgakis, "Disturbance detection and isolation by dynamic principal component analysis," *Chemometric Intell. Lab. Syst.*, vol. 30, no. 1, pp. 179–196, Nov. 1995.

[38] N. F. Thornhill, S. C. Patwardhan, and S. L. Shah, "A continuous stirred tank heater simulation model with applications," *J. Process Control*, vol. 18, nos. 3–4, pp. 347–360, Mar. 2008.

[39] C. Hu, Z. Xu, X. Kong, and J. Luo, "Recursive-CPLS-based quality-relevant and process-relevant fault monitoring with application to the tennessee eastman process," *IEEE Access*, vol. 7, pp. 128746–128757, 2019.

[40] M. Ammiche, A. Kouadri, and A. Bakdi, "A combined monitoring scheme with fuzzy logic filter for plant-wide tennessee eastman process fault detection," *Chem. Eng. Sci.*, vol. 187, pp. 269–279, Sep. 2018.

[41] S. Yin, S. X. Ding, A. Haghani, H. Hao, and P. Zhang, "A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark tennessee eastman process," *J. Process Control*, vol. 22, no. 9, pp. 1567–1581, Oct. 2012.

[42] M. Madakyaru, F. Harrou, and Y. Sun, "Monitoring distillation column systems using improved nonlinear partial least squares-based strategies," *IEEE Sensors J.*, vol. 19, no. 23, pp. 11697–11705, Dec. 2019.

K. RAMAKRISHNA KINI received the B.E. degree in electronics and communication engineering from NMAMIT, NITTE, and the M.Tech. degree in control systems from the Manipal Institute of Technology, Manipal, India. He is currently a Ph.D. Research Scholar and an Assistant Professor with the Instrumentation and Control Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education. His research interests include fault detection and diagnosis, soft sensor modeling, latent variable regression modeling using wavelets, and advanced process control.

MUDDU MADAKYARU received the B.E. degree in chemical engineering from RVCE, the M.Tech. degree in chemical plant design from the National Institute of Technology Karnataka, India, and the Ph.D. degree in process control from the Indian Institute of Technology Bombay, Mumbai, India. He was a Postdoctoral Researcher with Texas A&M University, Doha, Qatar, for four years. He is currently a Professor with the Department of Chemical Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, India. He has published more than 30 articles in peer-reviewed journals and 15 international conference proceedings papers. His research interests include advanced process control, including system identification, fault detection and diagnosis, model predictive control, and latent variable regression modeling using wavelets. He is also a Life Member of the Indian Society for Technical Education and the Indian Society of Systems for Science and Engineering, and a Fellow of the Institution of Engineers, India.