2018 European Control Conference (ECC)
June 12-15, 2018. Limassol, Cyprus

Improved Statistical Method Based Exponentially Weighted GLRT Chart and Its Application to Fault Detection*

Raoudha Baklouti[1], Majdi Mansouri[2], Ahmed Ben Hamida[1], Hazem Nounou[3] and Mohamed Nounou[4]

Abstract— **This paper deals with fault detection (FD) of chemical processes. Our previous study [1] has proved the effectiveness of multiscale principal component analysis (MSPCA)-based Moving Window (MW)-Generalized Likelihood Ratio Test (GLRT) to detect faults by maximizing the detection probability for a particular false alarm rate with different values of windows. However, the conventional PCA method is not suitable in nonlinear processes. In fact, this lack affects the monitoring system. To address this problem, we propose, first, to use multistage kernel PCA (MSKPCA) technique to extract the deterministic features and compute the principal components (PCs) in the original space. Second, integrate exponentially weighted moving average (EWMA), that has shown better abilities to reduce the false alarm rates and enhance the FD performances. Therefore, this work focuses on extending MSKPCA, and developing a MSKPCA-based EWMA-GLRT technique in order to improve the FD performance. The performances of the MSKPCA -based EWMA-GLRT are illustrated using Tennessee Eastman benchmark process.**

I. INTRODUCTION

Monitoring applications is of great importance in process engineering in order to obtain a safe operation. Because of a huge amount of data from several variables are used in the process systems. That's way researchers pay great attention to fault detection and diagnosis methods in order to decrease productivity loss and void abnormal event progression.

The widely used and well known multivariate statistical modeling and fault detection techniques are principal component analysis (PCA) [2] and Partial Least Square (PLS) [3], [4] in many applications especially, in chemical processes [5], [6]. By projecting the big data onto a lower-dimensional subspace that contains most of the variance of the original data [5]. However, when the data are highly nonlinear, linear PCA performs poorly [7]. To cope with the problem posed by nonlinear data, the authors in [7] integrates the principal curve algorithm and neural networks. However, their algorithm assumes an approximation of the nonlinear function by a linear combination of many univariate functions. Moreover, in order to compute the principal curves and train the neural network, a nonlinear optimization issue must be solved. To overcome the nonlinear problem, Schölkopf et al. proposed Kernel Principal Component Analysis (KPCA) [8] which better exploits the complicated spatial structure of high-dimensional features. Besides, KPCA is simple and does not entail any nonlinear optimization [9], [10]. On the other hand, process data collected from most chemical processes may contain high level of noise. Thus, this might adversely affect the performance of the conventional PCA monitoring chart in terms of modeling and fault detection. To cope with this issue, Bakshi [11] has developed a multiscale wavelet-based multivariate process monitoring technique through Multiscale Principal Component Analysis (MSPCA), in which, the MSPCA computes the PCA of the wavelet coefficients at each scale. Due to its multiscale nature, MSPCA has been shown to be more appropriate for modeling of data containing contributions from events whose behavior changes over time and frequency [12]. For the modeling task, we propose to use the MSKPCA method. To deal with the detection task, generalized likelihood ratio test (GLRT)-based exponentially weighted moving average (EWMA), applied to the residual computed from MSKPCA, will be developed. The EWMA-GLRT chart is a filtered statistical hypothesis testing fault detection method that takes into consideration the information given by the current and previous data by giving high importance to the more recent data [13], [14].

Therefore, this paper focuses on extending MSKPCA, and developing a MSKPCA-based EWMA-GLRT method in order to enhance the fault detection abilities. The detection performance of the developed MSKPCA-based EWMA-GLRT method is assessed and compared to existing techniques using the Tennessee Eastman Process (TEP).

The rest of the paper is organized as follows. Section 2 is devoted to the KPCA and MSKPCA methods description. Then, Section 3 presents the detection chart EWMA-GLRT. After that, in Section 4 the fault detection performance of MSKPCA -based EWMA-GLRT is studied using the Tennessee Eastman benchmark process. At the end, the conclusions are presented in Section 5.

II. KPCA AND MSKPCA METHODS DESCRIPTION

A. Kernel Principal Component Analysis (KPCA)

Let $X_i \in \mathbb{R}^m$ denotes a i-th sample vector of m number of sensors. Also, assume there are n samples dedicated to each sensor, a data matrix $X \in \mathbb{R}^{N \times m}$ is with each row, displaying a sample. Meanwhile, X matrix is scaled to zero mean for covariance-based PCA and at the same time, to unit variance for correlation-based PCA [15]. The X matrix can be divided into two matrices: a score matrix S and a loading

*Qatar National Research Fund (a member of Qatar Foundation)

[1]Advanced Technologies for Medicine and Signals, National Engineering School of Sfax, Tunisia,

[2]Corresponding author, Electrical and Computer Engineering Program, Texas A&M University at Qatar, Doha, Qatar, Tel:+974.4423.0608, Fax: +974.4423.0065, E-mail: majdi.mansouri@qatar.tamu.edu,

[3]Electrical and Computer Engineering Program, Texas A&M University at Qatar, Doha, Qatar,

[4]Chemical Engineering Program, Texas A&M University at Qatar, Doha, Qatar.

matrix W through singular value decomposition (SVD), so that,

$$X = SW^T, \tag{1}$$

where $S = [s_1\ s_2 \cdots s_m] \in \mathbb{R}^{N\times m}$ is a transformed variables matrix, $s_i \in \mathbb{R}^N$, are the score vectors or PCs, and $W = [w_1\ w_2 \cdots w_m] \in \mathbb{R}^{m\times m}$ is an orthogonal vectors matrix $w_i \in \mathbb{R}^m$ which includes the eigenvectors associated with the covariance matrix of X, i.e., Σ, which is given by,

$$\Sigma = \frac{1}{N-1}X^TX = W\Lambda W^T \ \text{ with } \ WW^T = W^TW = I_N, \tag{2}$$

where $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_m)$ is a diagonal matrix containing the eigenvalues related to the m PCs, $(\lambda_1 > \lambda_2 > \cdots > \lambda_m)$, and I_N is the identity matrix [16].

It must be noted at this point that the PCA model yields same number of PCs as the number of original variables (m). Nevertheless, for collinear process variables, a smaller number of PCs (l) are required so that most of the variations in the data are captures. Most of the times, a small subset of the PCs (which correspond to the maximum eigenvalues) might carry the most of the crucial information in a data set, which simplifies the analysis.

The effectiveness of the PCA model depends on the number of PCs are to be used for PCA. Selecting an appropriate number of PCs introduces a good performance of PCA in terms of processes monitoring. In this study herein, the cumulative percent variance method is utilized to come up with the optimum number of retained PCs. The cumulative percent variance is computed as follows:

$$CPV(l) = \frac{\sum_{i=1}^{l}\lambda_i}{trace(\Sigma)} \times 100. \tag{3}$$

When the number of PCs l is determined, then, the data matrix X is given by,

$$X = SW = [\widehat{S}\ \ \widetilde{S}][\widehat{W}\ \ \widetilde{W}]^T, \tag{4}$$

where $\widehat{S} \in \mathbb{R}^{N\times l}$ and $\widetilde{S} \in \mathbb{R}^{N\times(m-l)}$, are matrices of l retained PCs and the $(m-l)$ ignored PCs, respectively, and the matrices $\widehat{W} \in \mathbb{R}^{m\times l}$ and $\widetilde{W} \in \mathbb{R}^{m\times(m-l)}$ are matrices of l retained eigenvectors and the $(m-l)$ ignored eigenvectors, respectively. Using Equation (4), the following can be written,

$$X = \widehat{S}\widehat{W}^T + \widetilde{S}\widetilde{W}^T = \overbrace{X\widehat{W}\widehat{W}^T}^{\widehat{X}} + \overbrace{X(\mathbf{I}_m - \widehat{W}\widehat{W}^T)}^{R}, \tag{5}$$

where the matrix $\widehat{X}$ represents the modeled variation of X based on first l components, and the matrix R represents the variations corresponding to process noise.

The objective of KPCA, is to compute principal components (PCs) in the map space $\mathbf{F}$. First, we map the nonlinear data into a feature space $\mathbf{F}$ as the following:

$$\Phi : \mathbf{R}^N \to \mathbf{F}. \tag{6}$$

The covariance matrix in the feature space $\mathbf{F}$ is given by:

$$C = \frac{1}{N}\Phi(X_j)\Phi(X_j)^T, \tag{7}$$

It has been shown that, for certain choices of map function Φ, even if feature space $\mathbf{F}$ has arbitrarily large dimensionality, the classical PCA can be performed in $\mathbf{F}$ [17]. To perform PCA in feature space $\mathbf{F}$, we need to compute the eigenvalues $\gamma \geq 0$ and eigenvectors $V \in \mathbf{F}\backslash\{0\}$ satisfying:

$$\gamma V = CV = \frac{1}{N}\Phi(X_j)\Phi(X_j)^TV \tag{8}$$

and assuming that the projected new features have zero mean $\frac{1}{N}\sum_{j=1}^{N}X_j = 0$. Note that all solutions V must lie in the span of $X_1, X_2, \ldots X_N$ as [17]:

$$\gamma(\Phi(X_k).V) = (\Phi(X_k).CV), \quad k = 1, \ldots, N. \tag{9}$$

and there exist coefficients α_k, $k = 1 : N$ so that,

$$V = \sum_{k=1}^{N}\alpha_k\Phi(X_k). \tag{10}$$

By substituting Equations (8) and (10) into Equation (9), we get,

$$N\gamma\alpha = \mathbf{K}\alpha, \tag{11}$$

where, $\mathbf{K}$ is a positive semidefinite kernel matrix, defined as,

$$\mathbf{K_{ij}} =< \Phi(X_i), \Phi^T(X_j) >, \tag{12}$$

and $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_N]$.
The PCs(T) are extracted by projecting the map $\Phi(X)$ onto the eigenvectors V in the feature space $\mathbf{F}$ as:

$$T =< V, \Phi(X) >, \tag{13}$$

and the resulting kernel can be calculated using:

$$Y = \Phi(X)^TV. \tag{14}$$

Two commonly used kernels are the Gaussian and the polynomial [17]:

1) Polynomial: $\mathbf{K}(X,Y) =< X, Y >^d$,
2) Gaussian: $\mathbf{K}(X,Y) = \exp(-\frac{||X-Y||^2}{m})$,

where, d and m are the parameters of Polynomial and Gaussian functions, respectively.

B. Residual Computation

In the section just above we presented how to compute PCs, in the feature space ($Y = \Phi(X)$) using KPCA, which is enough for multivariate fault detection (FD) based techniques. However, when we use univariate based FD techniques such GLRT, EWMA, etc., we need to move to the input space and estimate the approximated pre-images X_i from the KPCA features $Y_i = \Phi(X_i)$ [18]. In the classical PCA, the approximated pre-images X_i can be approximated using eigen-decomposition. However, for KPCA, we define

a projection operator $\mathcal{P}_M$ which maps $\Phi(X)$ to its approximation in the input-space, i.e.,

$$\mathcal{P}_M\Phi(X) = \sum_{i=1}^{N} Y_i V_i. \tag{15}$$

If N is large enough, we get $\mathcal{P}_M\Phi(X) \approx \Phi(X)$. The exact pre-images X approximation is untractable, thus, it is required to approximate Z so that [18].

$$\Phi(Z) \approx \mathcal{P}_M\Phi(X). \tag{16}$$

This can be approximated by minimizing:

$$\rho(Z) = ||\Phi(Z) - \mathcal{P}_M\Phi(X)||^2. \tag{17}$$

Assuming that the kernel is Gaussian, $K(X,Y) = \exp(-||X-Y||^2/2\sigma^2)$, z should satisfy the following equation:

$$z = \frac{\sum_{i=1}^{N}\gamma_i \exp(-||z-X_i||^2/2\sigma^2)X_i}{\sum_{i=1}^{N}\gamma_i \exp(-||z-X_i||^2/2\sigma^2)}. \tag{18}$$

where

$$\gamma_i = \sum_{k=1}^{m} Y_k a_{ki}. \tag{19}$$

We iterate until it converges to a value z_t. The iteration can be represented by the following expression:

$$z_{t+1} = \frac{\sum_{i=1}^{N}\gamma_i \exp(-||z_t-X_i||^2/2\sigma^2)X_i}{\sum_{i=1}^{N}\gamma_i \exp(-||z_t-X_i||^2/2\sigma^2)}. \tag{20}$$

Thus, the residual which will be used for fault detection phase is given by:

$$R = X - Z. \tag{21}$$

C. Multiscale kernel PCA (MSKPCA) method description

Wavelet-based representation is a powerful data-analysis method that provides a separation of deterministic and stochastic features [19]. Given a time domain data set (signal), a scaled signal can be obtained by convoluting the original signal with a low pass filter (h), which is obtained from a scaling basis function of the following form [20]:

$$\Phi_{ij}(t) = \sqrt{2^{-j}}\Phi(2^{-j}t - k), \tag{22}$$

where, k and j are discretized and translation dilation factors. The difference between the original and the approximated signals (detail signal), can be obtained by convoluting the original signal with a high pass filter (g), which is computed from a wavelet basis function of the following form [20]:

$$\Psi_{ij}(t) = \sqrt{2^{-j}}\Psi(2^{-j}t - k). \tag{23}$$

After repeating these approximations, the original signal can be expressed as [20]:

$$\begin{aligned} x(t) &= \sum_{k=1}^{n2^{-J}} a_{Jk}\Phi_{Jk} + \sum_{j=1}^{J}\sum_{k=1}^{n2^{-J}} a_{jk}\Psi_{jk}(t) \\ &= A_J(t) + \sum_{j=1}^{J} D_j(t) \end{aligned} \tag{24}$$

where, n denotes the length of the signal, J represents the maximum possible decomposition depth, $A_J(t) = \sum_{k=1}^{n2^{-J}} a_{Jk}\Phi_{Jk}$ the approximated signal and $D_j(t) = \sum_{j=1}^{J}\sum_{k=1}^{n2^{-J}} a_{jk}\Psi_{jk}(t)$ represents the detailed signal.

By selecting the combination of monitoring process to marry the kernel PCA with wavelet analysis. This combination is expected to benefit from KPCA and de-correlate the cross correlation among variables and from wavelet analysis by capturing the information at multiple scales and de-correlate the auto correlation among the individual variable. Wavelet analysis is first used to analyze the measured data matrix X at multiple scales, resulting in new data matrix $D_1, D_2, \ldots, D_J$, and A_J. Then, KPCA is applied to each new data matrix to extract information in linear feature space for fault detection. Finally, multiscale PCA similarity factor is applied to detect fault pattern. The developed MSKPCA algorithm and its steps are highlighted in Algorithm 1.

Algorithm 1: MSKPCA Algorithm

Input: $n \times m$ data matrix X, Confidence interval α.

- For each column (i.e. process variable) in the data matrix compute the wavelet decomposition;
- For each block (matrix) of scaled and detail coefficients at each scale, the covariance matrix is computed along with the number of principal components, as well as kernel PCA loadings and scores of those wavelet coefficients;
- Once the appropriate number of loadings is selected, wavelet coefficients larger than a certain threshold are selected;
- For all scales together, kernel PCA is carried out by after including only scales with significant events during reconstruction.

Next, we present the detection chart EWMA-GLRT, which is applied with the MSKPCA model for fault detection using Tennessee Eastman benchmark process.

III. EWMA-GLRT FAULT DETECTION CHART

The fault detection step is done using the residuals evaluated from the MSKPCA model. Using the information about the noise distribution of the residuals, a GLRT statistic is computed. To make the decision if a fault is present or not, the test statistic is compared to a computed threshold.
Consider the measurement vector $Y \in \mathbb{R}^N$:

$$Y = \theta + \epsilon, \tag{25}$$

where ϵ denotes the measurement noise assumed to be Gaussian $\mathcal{N}(0, \sigma^2 I_N)$ and θ is an unknown parameter.

GLRT is the detector that corresponds to the highest probability of detection (PD) given a fixed false alarm probability for all values of θ. The two-sided hypothesis test could be formulated as:

$$\begin{cases} \mathcal{H}_0 = \{\theta = \theta_0\}, & \text{(null hypothesis)}, \\ \mathcal{H}_1 = \{\theta = \theta_1\}, & \text{(alternative hypothesis)}. \end{cases} \tag{26}$$

The GLRT parameters θ can be estimated as:

$$\hat{\theta} = \arg\max_{\theta}(p(Y|\theta)), \tag{27}$$

where, p is the marginal density function of Y.

The estimated parameters, $\hat{\theta}$, for both hypothesis, are then substituted into the Likelihood Ratio test $\mathbb{G}(Y)$ provided in the NP-lemma [21] which yields the GLRT,

$$\mathbb{G}(Y) = \frac{p(Y|\hat{\theta}_1)}{p(Y|\hat{\theta}_0)}, \tag{28}$$

The most widely used distribution for random variables is the Gaussian distribution. The residual need to be normalized before the validation or the detection phase. In this case, the normalized residual R is assumed to be Gaussian and the vector $Y \in \mathbb{R}^N$ is formed by one of $\mathcal{N}(0, \sigma^2 I_N)$ or $\mathcal{N}(\theta \neq 0, \sigma^2 I_N)$ and the hypothesis test (26) can be written as,

$$\begin{cases} \mathcal{H}_0 : R \sim \mathcal{N}(0, \sigma^2 I_N) \\ \mathcal{H}_1 : R \sim \mathcal{N}(\mathbf{W}^T\theta, \sigma^2 I_N) \end{cases} \tag{29}$$

where, $\mathbf{W}$ is the $N \times N$ weighting matrix, θ is the mean vector, which is the value of the fault and $\sigma^2 > 0$ is the variance, which is assumed to be known.

The GLRT chart yields the following log-likelihood ratio,

$$\begin{aligned} \mathbb{G}(R) &= 2\log \frac{\sup_{\theta \in \mathbb{R}^N} f_{\mathbf{W}^T\theta}(R)}{f_{\theta=0}(R)} \\ &= 2\log \left\{ \frac{\sup_{\theta} \exp\left\{-\frac{1}{2}\|R - \mathbf{W}^T\theta\|_2^2\right\}}{\exp\left\{-\|R\|_2^2\right\}} \right\} \\ &= \min_{\theta} \|R - \mathbf{W}^T\theta\|_2^2 + \|R\|_2^2, \end{aligned} \tag{30}$$

which is maximized for $\widehat{\theta}$ to obtain the maximum likelihood estimate of θ, which equals $\widehat{\theta} = \mathbf{W}^T\mathbf{W}^{T^{-1}}R$. Then, the GLRT method replaces the unknown parameter, θ, by its maximum likelihood estimate, which gives the following GLRT statistic,

$$\mathbb{G}(R) = \|R - \mathbf{W}^T\widehat{\theta}\|_2^2 + \|R\|_2^2 = R^T\mathbf{W}^T\mathbf{W}^{T^{-1}}R \tag{31}$$

To define the threshold, it is necessary to compute the distribution of the test statistic. Assuming that the normalized residuals follow a Gaussian distribution, i.e.,

$$R \sim \mathcal{N}(\mathbf{W}^T\theta, I_N) \tag{32}$$

where $\theta = 0$ under the null hypothesis $\mathcal{H}_0$, then, the test statistic will follow a Chi-square distribution, i.e.,

$$\mathbb{G}(R) = R^T\mathbf{W}^T\mathbf{W}^{T^{-1}}R \sim \chi^2_{\mu}(\lambda), \tag{33}$$

where $\mu = rank(W^T)$ and $\lambda = (\mathbf{W}^T\theta)^T\mathbf{W}^T\mathbf{W}^{T^{-1}}(\mathbf{W}^T\theta)$.

Assuming that the normalized residuals follow a Gaussian distribution, i.e.,

$$R \sim \mathcal{N}(\mathbf{W}^T\theta, I_N), \tag{34}$$

where $\theta = 0$ under the null hypothesis $\mathcal{H}_0$, then $\mathbb{G}(R) = R^T W^T W^{T^{-1}} R$ will follow a Chi-square distribution i.e.,

$$\mathbb{G}(R) = R^T\mathbf{W}^T\mathbf{W}^{T^{-1}}R \sim \chi^2_{\mu}(\Lambda), \tag{35}$$

where $\mu = rank(\mathbf{W}^T)$ and $\Lambda = (\mathbf{W}^T\theta)^T\mathbf{W}^T\mathbf{W}^{T^{-1}}(\mathbf{W}^T\theta)$. Here, $\mathbf{W}$ is the $N \times N$ identity matrix assumed to be known.

The EWMA-GLRT statistic $\mathbb{EG}$ can be computed as:

$$\mathbb{EG}_k(R) = \lambda\mathbb{G}_k(R) + (1-\lambda)\mathbb{EG}_{k-1}(R) \tag{36}$$

where, λ is the smoothing parameter between 0 and 1.
To define the threshold (α), it is necessary to compute the distribution of the test statistic $\mathbb{EG}$.

Hence, the fault detection decision is given as,

$$\begin{cases} \mathcal{H}_0 & \text{if} \quad \mathbb{EG}(Y) < \alpha \\ \mathcal{H}_1 & \text{else.} \end{cases} \tag{37}$$

If the EWMA-GLRT statistic $\mathbb{EG}$ value is less than the threshold value (α) under null hypothesis, there is no fault and if EWMA-GLRT statistic is higher than threshold value, alternative hypothesis is correct and fault is declared in the process.

Next, the performance of the developed MSKPCA-based EWMA-GLRT technique is assessed and compared to the conventional KPCA methods using Tennessee Eastman Process.

A. Fault Detection and Application to Tennessee Eastman Process

The Tennessee Eastman Process is a well-known benchmark chemical process, which was created by the Eastman Chemical Company to supply a realistic industrial process for evaluating monitoring and process control methods [22]. The test problem is based on an actual chemical process where the components, kinetics, and operating conditions were modified for proprietary reasons. A diagram of the process is contained in Figure 1. The process has five major unit operations: the reactor, the product condenser, a vapor-liquid separator, a recycle compressor and a product stripper. The gaseous reactants A, C, D and E and the inert B are fed to the reactor where the liquid products G and H are formed. The process has 41 measurements and 12 manipulated variables. The training and testing data sets consisted of 53 features and 961 data samples.

Twenty-one testing sets were generated [22] using the preprogrammed faults (Faults 1-21).

To further validate the developed MSKPCA-based EWMA-GLRT fault detection method it would need to be tested on process data from a plant. For each of the testing data sets, the fault was introduced at observation 224, and the fault continued until the end of the data set. The missed detection rates (%) and false alarm rates (%) obtained for all

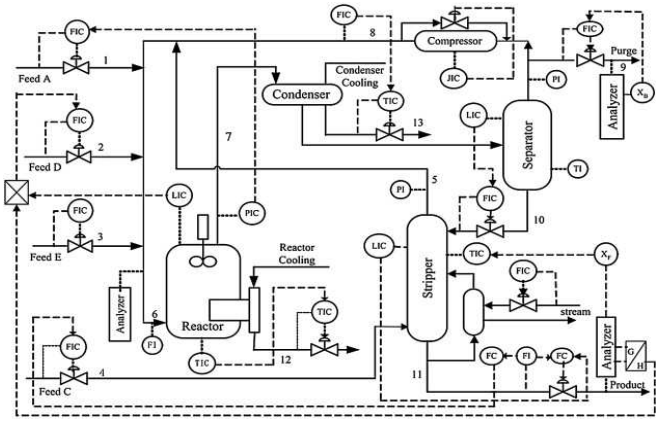

Fig. 1. Tennessee Eastman Process.

faults are summarized in Tables I and II, respectively. We can show from these tables that the MSKPCA-based EWMA-GLRT technique is able to provide significantly lower missed detection rates and false alarm rates, when compared to the classical techniques. Figure 2 presents the monitoring results of fault 12. The MSKPCA-based EWMA-GLRT shows a reduction in false alarm rates and missed detection rates compared to the MSKPCA-based GLRT method.

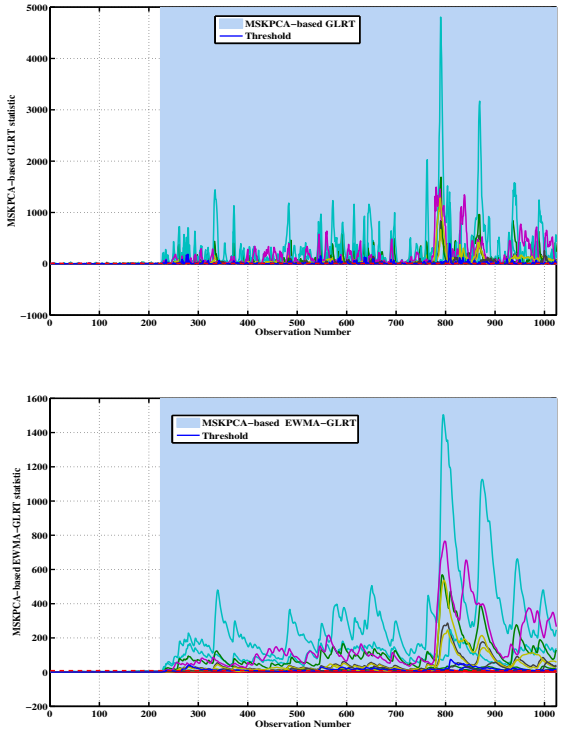

Fig. 2. Monitoring TEP fault 12 using MSKPCA-based GLRT and MSKPCA-based EWMA-GLRT charts.

IV. CONCLUSIONS

In this paper, a multiscale kernel principal component analysis (MSKPCA)-based on exponentially weighted moving average -generalized likelihood ratio test (EWMA-GLRT) is proposed for fault detection of the Tennessee Eastman benchmark process. The fault detection problem was addressed so that the data are first modeled using the MSKPCA method and then the faults are detected using one of the EWMA-GLRT charts. The results demonstrate the effectiveness of the MSKPCA-based EWMA-GLRT technique over the MSKPCA-based GLRT method in terms of lower missed detection and false alarms rates.

TABLE I

SUMMARY OF MISSED DETECTION RATES FOR TEP DATA

Fault Number	MSKPCA-based -GLRT	MSKPCA-based EWMA-GLRT
IDV(1)	0	0.3745
IDV(2)	0.1248	1.7478
IDV(3)	44.4444	60.1748
IDV(4)	56.0549	78.9014
IDV(5)	39.9501	46.5668
IDV(6)	0.1248	0.4994
IDV(7)	25.2185	20.4744
IDV(8)	0.8739	0
IDV(9)	52.0599	73.7828
IDV(10)	4.7441	3.1211
IDV(11)	21.0986	5.1186
IDV(12)	0.9988	0.4994
IDV(13)	2.8714	4.8689
IDV(14)	0	0.2497
IDV(15)	44.9438	65.4182
IDV(16)	17.7278	3.1211
IDV(17)	0.4994	0
IDV(18)	5.3683	5.6180
IDV(19)	29.4632	51.0612
IDV(20)	13.1086	7.2409
IDV(21)	12.9838	17.4782

TABLE II

SUMMARY OF FALSE ALARM RATES FOR TEP DATA

Fault Number	MSKPCA-based -GLRT	MSKPCA-based EWMA-GLRT
IDV(1)	41.2556	14.7982
IDV(2)	34.5291	0
IDV(3)	49.3274	27.3543
IDV(4)	38.5650	12.1076
IDV(5)	38.5650	12.1076
IDV(6)	35.8744	0
IDV(7)	42.6009	3.5874
IDV(8)	49.7758	19.7309
IDV(9)	51.1211	17.9372
IDV(10)	34.0807	0
IDV(11)	38.1166	0.4484
IDV(12)	42.1525	0.8969
IDV(13)	37.6682	0
IDV(14)	42.6009	6.2780
IDV(15)	42.6009	2.6906
IDV(16)	51.5695	27.3543
IDV(17)	60.5381	38.5650
IDV(18)	43.9462	14.3498
IDV(19)	44.3946	19.2825
IDV(20)	38.5650	0
IDV(21)	57.8475	46.1883

ACKNOWLEDGMENT

This work was supported by Qatar National Research Fund (a member of Qatar Foundation) under the NPRP grant NPRP9 − 330 − 2 − 140.

REFERENCES

[1] M. Z. Sheriff, M. Mansouri, M. N. Karim, H. Nounou, and M. Nounou, "Fault detection using multiscale pca-based moving window glrt," *Journal of Process Control*, vol. 54, pp. 47–64, 2017.

[2] M. Mansouri, M. Z. Sheriff, R. Baklouti, M. Nounou, H. Nounou, A. B. Hamida, and N. Karim, "Statistical fault detection of chemical process-comparative studies," *Journal of Chemical Engineering & Process Technology*, vol. 7, no. 1, pp. 282–291, 2016.

[3] C. Botre, M. Mansouri, M. Nounou, H. Nounou, and M. N. Karim, "Kernel pls-based glrt method for fault detection of chemical processes," *Journal of Loss Prevention in the Process Industries*, vol. 43, pp. 212–224, 2016.

[4] C. Botre, M. Mansouri, M. N. Karim, H. Nounou, and M. Nounou, "Multiscale pls-based glrt for fault detection of chemical processes," *Journal of Loss Prevention in the Process Industries*, vol. 46, pp. 143–153, 2017.

[5] B. Wise and N. Gallagher, "The process chemometrics approach to process monitoring and fault detection," *Journal of Process Control*, vol. 6, p. 329348, 1996.

[6] A. Simoglou, E. Martin, and A. Morris, "Multivariate statistical process control in chemicals manufacturing," in *IFAC Conference SAFEPROCESS*, Hull, UK,, 1997, pp. 21–27.

[7] D. Dong and T. McAvoy, "Nonlinear principal component analysis based on principal curves and neural networks," *Computers and Chemical Engineering*, vol. 20, no. 1, pp. 65–78, 1996.

[8] B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural computation*, vol. 10, no. 5, pp. 1299–1319, 1998.

[9] R. Fezai, M. Mansouri, O. Taouali, M. F. Harkat, and N. Bouguila, "Online reduced kernel principal component analysis for process monitoring," *Journal of Process Control*, vol. 61, pp. 1–11, 2018.

[10] M. Mansouri, M. N. Nounou, and H. N. Nounou, "Improved statistical fault detection technique and application to biological phenomena modeled by s-systems," *IEEE transactions on nanobioscience*, vol. 16, no. 6, pp. 504–512, 2017.

[11] B. Bakshi, "Multiscale pca with application to multivariate statistical process monitoring," *AIChE Journal*, vol. 44, pp. 1596–1610, 1998.

[12] M. Mansouri, M. N. Nounou, and H. N. Nounou, "Multiscale kernel pls-based exponentially weighted-glrt and its application to fault detection," *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2017.

[13] M. Mansouri, H. Nounou, M. F. Harkat, and M. Nounou, "Fault detection of chemical processes using improved generalized likelihood ratio test," in *Digital Signal Processing (DSP), 2017 22nd International Conference on*. IEEE, 2017, pp. 1–5.

[14] M.-F. Harkat, M. Mansouri, M. Nounou, and H. Nounou, "Enhanced data validation strategy of air quality monitoring network," *Environmental research*, vol. 160, pp. 183–194, 2018.

[15] M. Tamura and S. Tsujita, "A study on the number of principal components and sensitivity of fault detection using pca," *Computers & Chemical Engineering*, vol. 31, no. 9, pp. 1035–1046, 2007.

[16] J. Jackson and G. Mudholkar, "Control procedures for residuals associated with principal component analysis," *Technometrics*, vol. 21, p. 341349, 1979.

[17] B. Schölkopf, A. Smola, and K.-R. Müller, "Kernel principal component analysis," pp. 583–588, 1997.

[18] S. Romdhani, S. Gong, A. Psarrou *et al.*, "A multi-view nonlinear active shape model using kernel pca." in *BMVC*, vol. 10, 1999, pp. 483–492.

[19] S. Mallat, "A theory of multiresolution signal decomposition: The wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, p. 764, 1989.

[20] M. N. Nounou, H. N. Nounou, N. Meskin, A. Datta, and E. R. Dougherty, "Multiscale denoising of biological data: A comparative analysis," *IEEE/ACM transactions on computational biology and bioinformatics*, vol. 9, no. 5, pp. 1539–1545, 2012.

[21] S. M. Kay, "Fundamentals of statistical signal processing: Detection theory, vol. 2," 1998.

[22] J. J. Downs and E. F. Vogel, "A plant-wide industrial process control problem," *Computers & chemical engineering*, vol. 17, no. 3, pp. 245–255, 1993.