Preprints, 10th IFAC International Symposium on
Advanced Control of Chemical Processes
Shenyang, Liaoning, China, July 25-27, 2018

Fault Detection using Empirical Mode Decomposition based PCA and CUSUM with Application to the Tennessee Eastman Process

Yuncheng Du*, Dongping Du**

**Department of Chemical & Biomolecular Engineering, Clarkson University, Potsdam, NY 13699 USA (Tel: 315-268-2284; e-mail: ydu@clarkson.edu)*
***Department of Industrial, Manufacturing, and System Engineering, Texas Tech University, Lubbock, TX 79409 (Tel: 806-834-7338; e-mail:dongping.du@ttu.edu)*

Abstract: In this work, a new algorithm is developed to identify stochastic faults in the Tennessee Eastman (TE) process, which integrates Ensemble Empirical Mode Decomposition (EEMD), Principal Component Analysis (PCA), Cumulative Sum (CUSUM), and half-normal probability plot to detect three particular faults that could not be properly detected with previously reported techniques. This algorithm includes three steps: measurements pre-filtering, sensitivity analysis, and fault detection. Measured variables are first decomposed into different scales using the EEMD-based PCA for extracting fault signatures, from which a subset of variables that are sensitive to faults are selected with the half-normal probability plot. Based on the specific variables, CUSUM-based statistics are further used for improved fault detection. The algorithm can successfully identify three particular faults in the TE process with small time delay.

Keywords: Process monitoring and control, process data analytics, stochastic faults, sensitivity analysis

1. INTRODUCTION

An important aspect for safe operation in chemical processes is the early detection of abnormal events and malfunctions that are defined as faults (Gerlter, 1998). For a detectable fault, the fault detection and diagnosis (FDD) algorithms can provide symptomatic finger-prints to identify the root cause of the anomalous behaviour. Many methods have been developed, which can be generally classified into three groups (Isermann, 2005): (i) Analytical methods that are solely based on first-principle models of process (Du, et al., 2015; Wong & Lee, 2010); (ii) surrogate (empirical) modelling methods such as multivariate statistical analysis that use the historical process data (Kim, et al., 2016; Chiang, et al., 2015); and (iii) Semi-empirical algorithms that combine first-principles' and surrogate models (Feng, et al., 2016; Jiang, et al., 2016).

Each of these modelling techniques has its own advantages and disadvantages. It is recognized that surrogate models are easier to formulate, however, first-principles models have superior extrapolation ability (Isermann, 2006). This work focuses on the development of a surrogate model. Since data in chemical processes often exhibit high correlation and cross-correlation among variables, multivariate statistical analysis (MVSA) such as principal component analysis (PCA) are used to reduce model complexity, thus leading to improved accuracy (Chiang, et al., 2015). The main drawback of the MVSA is the interpretability, since it depends on subspaces that typically involve linear transformations of the original physical states. To improve the interpretability, a half-normal probability plot is used here to identify variables that are sensitive to faults.

Chemical processes can be operated at different scales (Misra, et al., 2002). For instance, stochastic perturbations may exhibit various energy spectrum and measurement noise may have different frequency ranges. A multiscale interpretation of measurements at different scales can provide signatures for improved FDD. Wavelet transform is one of the most popular multiresolution analysis tools (Sheriff, et al., 2017). However, wavelet transformation may provide inaccurate results since it uses a linear and non-adaptive transformation with fixed wavelet functions. Ensemble empirical mode decomposition (EEMD), as an alternative, is a self-adaptive method (Wu & Huang, 2009), which can be directly derived from data. Thus, the EEMD is more suitable than the wavelet transformation.

Uncertainty is a key challenge for FDD since empirical model based FDD greatly relies on models and data that are often not perfect. Generally, the impact of uncertainty on FDD is not specified in reported studies, which may lead to a loss of accuracy (Du & Du, 2018). Also, it is difficult to identify and diagnose faults using models calibrated with data that is affected by uncertainty such as noise, since fault fingerprints can be camouflaged by noise. To enhance FDD performance, EEMD is combined with PCA as a *pre-filtering* tool in this work to extract fault signatures that can be further used to infer the occurrence of faults. A salient feature of the EEMD-based PCA model is that the extracted fault signatures are more compact, thus adding robustness to FDD algorithms.

Specially, this paper studies the application of EEMD-based PCA and the half-normal probability plot to detect three faults that were found unobservable by other techniques (Ghosh, et al., 2011; Bernal-de-Lazaro, et al., 2016; Ding, et al., 2009; Lee, et al., 2004; Hsu, et al., 2010; Shang, et al., 2017; Lee, et al., 2006; Chiang, et al., 2015). This paper is organized as follows. Section 2 presents the background and the principal methodology used in this paper. The proposed fault detection algorithm is described in Section 3, followed by results and discussion in Section 4 and a brief conclusion in Section 5.

2. THEORETICAL BACKGROUND AND PROBLEM FORMULATION

2.1 Formulation of faults

Let assume the inputs in the Tennessee Eastman (TE) process can be defined as **g**, i.e., $\mathbf{g}=\{g_1, g_2,\ldots, g_n\}$, where n is the total number of inputs. For safe operation and consistent product quality, each input should be operated around a specific mean value with perturbations. For example, Fig. 1 (a) shows an input consisting of perturbations superimposed on two mean values, whereas Fig. 1 (b) shows the dynamics of a measured variable resulting from the input changes in Fig. 1 (a).

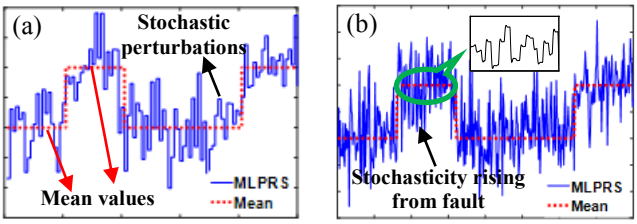

Fig. 1. Illustration of Fault and measured quantity

In this work, one of the mean values in Fig. 1 (a) is considered as a normal operating mode and another represents a fault. The objective is to detect the switch in the mean values of inputs. It is assumed that an input (fault) g_i in $\mathbf{g} = \{g_1, g_2,\ldots, g_n\}$ can be defined mathematically as:

$$g_{i,j} = \bar{g}_{i,j} + \Delta g_i \tag{1}$$

, where $i = 1, \ldots, n$, $j = 1, \ldots, n_i$, $\{\bar{g}_{i,j}\}$ is the mean values for the i^{th} fault g_i, n_i represents the total number of mean values of g_i, and $\{\Delta g_i\}$ are perturbations around each mean value of the i^{th} fault, which are assumed to be time invariant.

2.2 Processing monitoring with PCA

Principal Component Analysis (PCA) is a linear statistical approach to analyse covariance structure of multi-dimensional data $\boldsymbol{X} \in \boldsymbol{R}^{m\times k}$, where k is the number of measured quantities and m is the total number of measurements for each quantity. Using the PCA, it is possible to extract most of the variability in the k variables with a lower-dimensional variable space p ($p \ll k$). Note that, p represents the total number of the principal components retained to capture the majority of variability in $\boldsymbol{X}$. The optimal selection of the principal components and the calculation of T^2 and Q statistics proceed as follows.

(***i***) The covariance matrix C of $\boldsymbol{X}$ can be calculated as:

$$C = \frac{1}{m\text{-}1}\boldsymbol{X}^T\boldsymbol{X} \tag{2}$$

(***ii***) The eigenvalues $\boldsymbol{\lambda}=\{\lambda_i\}$ (i=1,⋯,k) in C can be computed and rearranged in a decreasing order as:

$$det\,(C - \boldsymbol{\lambda I}) = 0 \tag{3}$$

$$\Lambda=\begin{bmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_k\end{bmatrix} \tag{4}$$

, where $\{\lambda_i\}$ are eigenvalues stored in a decreasing order in Λ.

(***iii***) The corresponding eigenvectors $\{e_i\}$ of each eigenvalue λ_i of C can be computed as:

$$Ce_i = \lambda_i e_i \tag{5}$$

Note that an eigenvector matrix can be further defined as: $V = [e_1,\cdots,e_k]$ with all the eigenvectors obtained from (5).

(***iv***) The variance, ϵ_i, captured with each principal component can be approximated using its eigenvalue as:

$$\epsilon_i = \lambda_i / \textstyle\sum_{j=1}^{k}\lambda_j \cdot 100\% \tag{6}$$

(***v***) The total number of principal components p to be retained is decided with a variance threshold ϵ_t. A transformation matrix V_p used to project measurements onto the principal components is generated with eigenvectors as: $V_p = [e_1,\cdots,e_p]$.

(***vi***) Using V_p, $\boldsymbol{X}$ can be projected onto a domain defined by principal components p and a matrix $\boldsymbol{E}$ of residuals as:

$$\boldsymbol{X} = V_p^T\boldsymbol{X} + \boldsymbol{E} \tag{7}$$

From (7), the T^2 and Q statistics can be calculated and used for pre-filtering as explained below.

(***vii***) Using the first p principal components, the T^2 statistic and its corresponding limit T^2_{lim} can be calculated as:

$$T^2 = (\boldsymbol{x}_i - \boldsymbol{m}_c)^T C^{-1}(\boldsymbol{x}_i - \boldsymbol{m}_c) \tag{8}$$

$$T^2_{lim} = \frac{p(m-1)(m+1)}{m(m-p)}F_{p,(m-p),\alpha} \tag{9}$$

, where $\boldsymbol{x}_i$ is the i^{th} row in $\boldsymbol{X}$, $\boldsymbol{m}_c$ is the mean value vector of each column in $\boldsymbol{X}$, and $F_{p,(m-p),\alpha}$ is a F-distribution with p and (m-p) degrees of freedom and a level of significance of α.

(***viii***) The T^2 only explains the steady state correlations, but the residual matrix $\boldsymbol{E}$ can explain the rest of variances that are not captured in principal components. Thus, Q statistics and its corresponding limit Q_{lim} as:

$$Q = r^T r, \text{ where } r = (I - V_p V_p^T)\boldsymbol{x}_i \tag{10}$$

$$Q_{lim} = \theta_1\left[\frac{c_\alpha h_0\sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0(h_0-1)}{\theta_1^2}\right]^{1/h_0} \tag{11}$$

, where r is the residual, c_α is the confidence limit for the (1-α) percentile in a standard normal distribution, h_0 is defined as h_0=1-(2$\theta_1\theta_3$)/3θ_2^2. Each value of $\{\theta_f\}$ (f=1, 2, 3) is computed with eigenvalues as: $\theta_f = \sum_{i=p+1}^{k}\lambda_i^f$. The T^2 and Q statistics obtained from PCA models are used to extract fault features.

2.3 Ensemble empirical mode decomposition

For each variable $\boldsymbol{x}_j$ (j=1,⋯, k) in $\boldsymbol{X}$ with data collected over time domain [0, t_f], Empirical mode decomposition (EMD) decomposes $\boldsymbol{x}_j$ into different scales (or IMFs) as:

$$\boldsymbol{x}_j = \textstyle\sum_{i=1}^{K} c_{j,i}(t) + r_j(t) \tag{12}$$

, where $0 \le t \le t_f$, $c_{j,i}$ is the i^{th} IMF of the j^{th} variable, K is the total number of IMFs, and r_j is the residual of EMD. Note that $\{c_{j,i}\}$ are orthogonal to each other and include different frequencies varying from high to low, but the performance of EMD is jeopardized as mode mixing appears. To overcome the drawback, a noise-assisted method, i.e., ensemble empirical mode decomposition (EEMD) was developed (Wu & Huang, 2009), which can reduce the possibility of undue mode mixing. The EEMD algorithm proceeds as follows.

(***i***) Synthetic measurements of $\boldsymbol{x}_j$, i.e., $\boldsymbol{y}_j$, is generated by adding numerical white noise $n(t)$ with a magnitude of σ to the original measurements of $\boldsymbol{x}_j$, which gives:

$$\boldsymbol{y}_j = \boldsymbol{x}_j + \sigma n(t) \tag{13}$$

(***ii***) The conventional EMD defined in (12) is applied to $\boldsymbol{y}_j$ in (13) to calculate the IMFs $\{c_{j,i}\}$ of $\boldsymbol{y}_j$.

(***iii***) Steps *i* and *ii* will be repeated for L times with different realizations of noise but the same σ, thus an ensemble of IMFs can be obtained and defined as:

$$\left[\{\boldsymbol{c}_{j,i}^{1}(t)\}, \{\boldsymbol{c}_{j,i}^{2}(t)\}, \dots, \{\boldsymbol{c}_{j,i}^{L}(t)\}\right] \tag{14}$$

, where $j \in [1, k]$ represents the j^{th} measured variable in $\boldsymbol{X}$, and $i \in [1, K]$ denotes the i^{th} IMF of decomposition.

(***iv***) The final decomposition results can be calculated with the ensemble means of the corresponding IMFs as:

$$\bar{\boldsymbol{c}}_{j,i}(t) = \frac{1}{L}\sum_{q=1}^{L} \boldsymbol{c}_{j,i}^{q}(t) \tag{15}$$

Results from (15) can be used to reconstruct $\boldsymbol{x}_j$ by combining the EEMD with PCA in order to extract fault signatures.

2.4 Half-normal probability plot

The half-normal probability plot is a graphical tool to estimate which variables are affected by variation in faults significantly. The key is to use a normal curve as the reference distribution against which the significance of effect is tested. This can be calculated as:

$$\left[\Phi^{-1}\left(0.5 + \frac{0.5[i-0.5]}{k}\right), \varepsilon_{p_i}\right] \tag{16}$$

, where $i=1,\cdots,k$ is the i^{th} variable $\boldsymbol{x}_j$ in $\boldsymbol{X}$, Φ^{-1} is the cumulative distribution function of a standard normal distribution, $\{\varepsilon_{p_i}\}$ are organized in an increasing order and can be shown against the coordinates based on the half-normal diagram.

2.4 Cumulative sum (CUSUM) control chart

CUSUM is an efficient tool to detect small shifts, since it can accumulate information (Hawkins & Olwell, 1998). The location CUSUM is used, which calculate two statistics as:

$$C_i^{+} = max\left[0, C_{i-1}^{+} + x_i - (\mu_{i,C} + h)\right] \tag{17}$$

$$C_i^{-} = max\left[0, C_{i-1}^{-} + (\mu_{i,C} - h) - x_i\right] \tag{18}$$

, where h, $\mu_{i,C}$, C_i^{+}, and C_i^{-} are the slack variable, the *in-control* mean value, and the upper and the lower CUSUM statistics, respectively. As seen in (16) and (17), the upper and the lower CUSUM statistics can account for accumulated summations of any small deviations in measurements. The summations can be adaptively corrected with the slack variable and then compared to zero through a *max* operation. Using the CUSUM statistics, a process can be identified as *out-of-control* when either one of them exceeds a threshold ε.

2.5 Tennessee Eastman process

The Tennessee Eastman (TE) process is used as a benchmark process to compare FDD algorithms. The TE process includes five major units: a product condenser, a reactor, a product stripper, a recycle compressor, and a vapour/liquid separator (Downs & Vogel, 1993). There are 41 measured variables, 12 manipulated variables, and 20 input disturbances that can be considered as faults.

Many FDD methods have been reported for the TE process (Chiang, et al., 2015; Sheriff, et al., 2017; Ghosh, et al., 2011; Ding, et al., 2009; Shang, et al., 2017; Lee, et al., 2006). They demonstrated different capabilities for identifying the majority of faults, but most of the previous reported algorithms have consistently failed to detect three particular faults as described in Table 1, i.e., (i) IDV(3) or a step change in the D feed temperature, (ii) IDV(9) or a random variation in the D feed temperature, and (iii) IDV(15) or a stiction in the condenser cooling water valve.

Table 1. Three faults in TE process

Fault #	*Characteristics*	*Size*
IDV (3)	Step-wise changes	5%→10%
IDV (9)	Random variations	5%→10%
IDV (15)	Valve stiction	2%→7%

These faults may have economics or operational impacts. Thus, the incapability to identify there three faults motivates the use of the EEMD-based PCA in combination with the half-normal plot and CUSUM for improved FDD in this work.

3. FAULT DETECTION ALGORITHM

The proposed monitoring approach involves pre-filtering with EEMD-based PCA, sensitivity analysis using the half-normal probability plot, and fault detection with CUSUM statistics.

3.1 Pre-filtering with EEMD-based PCA models

The FDD in this work combines the EEMD-based multiscale modelling with a PCA-based multivariate statistical analysis. The EEMD-based PCA proceeds as a *pre-filtering* tool, which involves four consecutive steps as below.

Step i: For each measured variable $\boldsymbol{x}_j$ ($j=1,\cdots, k$) in $\boldsymbol{X}$, EEMD is used to obtain a family of IMFs $\{\bar{c}_{j,i}\}$, where i ($i=1, \dots, K$) represents the i^{th} IMF (scale) in the decomposition.

Step ii: The i^{th} IMF of each variable $\boldsymbol{x}_j$ is sorted in series to build a new multi-dimensional dataset $\boldsymbol{X}_i^{'}$ as:

$$\boldsymbol{X}_i^{'} = [\,\bar{c}_{1,i} \quad \cdots \quad \bar{c}_{j,i} \quad \cdots \quad \bar{c}_{k,i}] \tag{19}$$

This will produce a family of datasets, $\{\boldsymbol{X}_i^{'}\}$, where $i=1, \dots, K$. Then, the PCA can be applied to each dataset $\{\boldsymbol{X}_i^{'}\}$, which will produce multiple PCA models.

Step iii: For a specified variance threshold ϵ_t and a confidence level α, the T^2 and Q statistics limits of a PCA model at each scale, $\boldsymbol{X}_i^{'}$, are calculated with (9) and (11). These limits are defined as: $\{T^2_{i,\,lim}\}$ and $\{Q_{i,\,lim}\}$. Based on (8) and (10), the T^2 and Q statistics for each set of data in $\{\boldsymbol{X}_i^{'}\}$ are calculated, i.e., $\{T_i^2(t)\}$ and $\{Q_i(t)\}$, which are further normalized with respect to their corresponding limits as:

$$T_{j,\,norm}(t) = T_i^2(t) \,/\, T^2_{i,\,lim} \tag{20}$$

$$Q_{j,\,norm}(t) = Q_i(t) / Q_{i,\,lim} \tag{21}$$

The T^2 or Q statistics at each time t in (19) and (20) can be used to decide the contribution of the normalized T^2 or Q statistics in the reconstruction step below.

Step iv: The j^{th} measured quantity $\boldsymbol{x}_j$ can be reconstructed as:

$$\tilde{\boldsymbol{x}}_j = \sum_{i=1}^{K} \gamma_i(t)\bar{c}_{j,\,i}(t) \tag{22}$$

, where $\bar{c}_{j,\,i}$ is the i^{th} IMF (scale) of the j^{th} variable calculated from (15), and the weighting factor $\gamma_i(t)$ can be defined as:

$$\gamma_i(t) = \begin{cases} 1 & \text{if } TQ_{lim}(t) \geq 1 \\ \left[TQ_{lim}(t)\right]^{\beta} & \text{otherwise} \end{cases} \tag{23}$$

, where $TQ_{lim}(t)$ is described as:

$$TQ_{lim}(t) = T_{j,\,norm}(t) \text{ or } TQ_{lim}(t) = Q_{j,\,norm}(t) \tag{24}$$

The reconstruction of $\boldsymbol{x}_j$ in $\boldsymbol{X}$ uses its corresponding IMFs in each scale when the $\{T_i^2(t)\}$ and/or $\{Q_i(t)\}$ statistics exceed the control limits. To retain fault signatures in $\boldsymbol{X}$, β is chosen to be larger than 1. In this way, the reconstructed variable $\tilde{\boldsymbol{x}}_j$ will only contain features relevant to dynamic changes.

3.2 Fault detection with half-normal plot and CUSUM

Models generated with PCA are not interpretable since they rely on subspaces that do not have any physical explanation. To improve the interpretability, a half-normal probability plot is used to identify variables that are sensitive to faults.

Based on the half-normal probability plot, the CUSUM control chart is used for fault detection. When the CUSUMs statistics exceed a control limit, fault is detected, otherwise, the process will be identified as a normal operating mode. Considering the integrating nature, CUSUM statistics may require some time before a fault can be detected, especially when the faults lead to small variations in process dynamics as studied in this work. Thus, the fault detection delay $ARL_{\text{o.c.}}$, i.e., *out-of-control* average run length, (Bin Shams, et al., 2010), is studied.

4. RESULTS AND DISCUSSION

4.1 Previous attempts for three faults in Table 1

Comparative studies with multivariate techniques for detecting faults in the TE process were previously reported (Yin, et al., 2012). However, faults IDV(3), IDV(9), and IDV(15) were excluded, since either they cannot be identified or the detection rate was found to be very low. In this section, we will first demonstrate the performance of the conventional PCA method for three faults in Table 1 in terms of the fault detectability.

Fig. 2 shows the fault detection results using the T^2 statistics without the *pre-filtering* step proposed in this work. For faults IDV(3), IDV(9) and IDV(15), ten samples of inputs were simulated in each fault profile. The simulation time of each sample was 24 *h* and the sampling time was set to 6 *min*. The first five samples represent a normal operating mode, while the rest samples represent a faulty operating mode. That is, faults were introduced after 120 *h* (i.e.,1200 samples) of a normal operating mode. The settings of the normal and faulty modes follow the descriptions as explained in Table 1, and stochastic perturbations were added to each specific mean value. Also, measurements used for fault detection were contaminated by white noise of a zero mean and a standard deviation of 0.1.

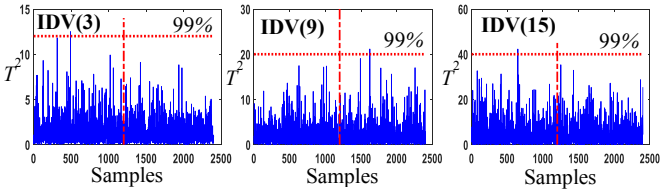

Fig. 2. Illustration of fault detection with PCA

In Fig. 2, a vertical line represents the onset of a fault, whereas a horizontal line denotes a 99% control limit of the T^2. When the T^2 statistic exceeds the control limits after an occurrence of a fault, the fault is considered as detected. However, as seen in Fig. 2, the T^2 fails to surpass the limits after the onsite of any of the faults, which means that these faults cannot be detected with the conventional PCA. The limited capability of previous methods to identify these faults motivates the development of new methods in this work. The Q statistics also fail to detect these faults and the result is not given for brevity.

4.2 Pre-filtering with EEMD-based PCA

The EEMD-based PCA is used as a *pre-filtering* tool to extract fault signatures. For clarify, a dataset of both normal and faulty operating modes with measurements collected over 240 *h* was generated. Since stochastic perturbations are superimposed on the mean values, ten samples were used for simulations. The simulation time of each sample was 24 *h* and the sampling time was 6 *min*, i.e., 2400 samples were used to generate EEMD-based PCA models. To demonstrate the efficiency of the *pre-filtering*, Fig. 3 shows the results obtained with the EEMD-based PCA, in which fault IDV(9) was introduced after 1200 samples (i.e., 120 *h*) of a normal operating. To illustrate the difficulty associated with the detection of IDV(9), manipulated variable XMV(10) is used. This is a critical variable that reflects the reactor cooling water flow in the TE process, which should be adjusted to eliminate any changes in the reactor temperature to remain the conversion level at a desired level.

As seen in Fig. 3 (a), changes in XMV(10) are camouflaged by noise. For an inset in Fig. 3 (a), Fig. 3 (b) shows the results with the *pre-filtering*. As seen, the fault signatures can be extracted by comparing Figs. 3 (a) and (b). As expected, fault IDV(9) can increase the variability in its correlated measured variables as shown in Fig. 3 (b), since it denotes an increase of

random variation around the mean value of fault. Also, as seen in Fig. 3 (b), changes in XMV(10) is insignificant in the absence and in the presence of IDV(9). Similar *pre-filtering* results can be obtained for other faults.

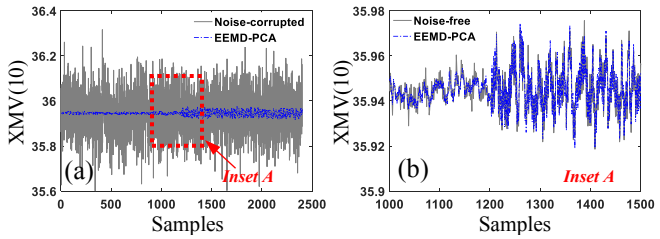

Fig. 3. Illustration of pre-filtering to extract fault features

4.3 Sensitivity analysis with half-normal probability plot

Since the multivariate analysis cannot provide a desired fault detection results for three faults in Table 1 and each fault may have different effect on the measured quantities, we propose to use the half-normal probability plot to find a univariate that is sensitive to a specific fault. The relationship between the faults and their corresponding manipulated variables that can be used for fault detection in shown in Fig. 4. Note that the half-normal probability analysis can be also applied to measured variables. As shown in Fig. 4, the manipulated variable with the highest sensitivity will be used to identify each fault. For example, it was observed that IDV(3), i.e., small step-wise changes in feed concentration, can induce changes in the manipulated variable XMV(10). Since IDV(3) affects the steady state in the reactor, and the reaction is highly exothermic, the manipulated variable XMV(10) should be adjusted to eliminate changes in the mean of the steady state reactor temperature.

4.4 Fault detection with CUSUM

Since the CUSUM based statistics is especially suitable for detecting small changes in the process mean, it is applied to manipulated variables that can be significantly affected by faults in Table 1. Note that the application of CUSUM to measured variables is not given for brevity. Fig. 5 shows the FDD results for three faults in this work.

For all simulations in Fig. 5, the faults were introduced after 1200 samples, i.e., after 24 *h* of normal operations. The set-up of simulation follows the description given in Table 1. The horizontal line in each subplot means the CUSUM statistical threshold, while the vertical line denotes the onsite of faults. As seen, the proposed method can successfully detect these faults that were not observable in previous reported works.

4.5 Comparison of $ARL_{o.c}$

The algorithm resembles the technique developed in the work by (Bin Shams, et al., 2010) in terms of the incorporation of the CUSUM. However, there are significant differences and extensions. The faults in this work are defined as perturbations superimposed on mean values of process inputs, while the previous work only dealt with deterministic faults such as step-wise changes. In addition, the *pre-filtering* in this work can efficiently extract fault signatures and improve the detection efficiency. For comparison, Table 2 shows the results in terms of the *out-of-control* average run length ($ARL_{o.c}$), which is the required detection time before a fault can be identified due to the integrating nature of the CUSUM.

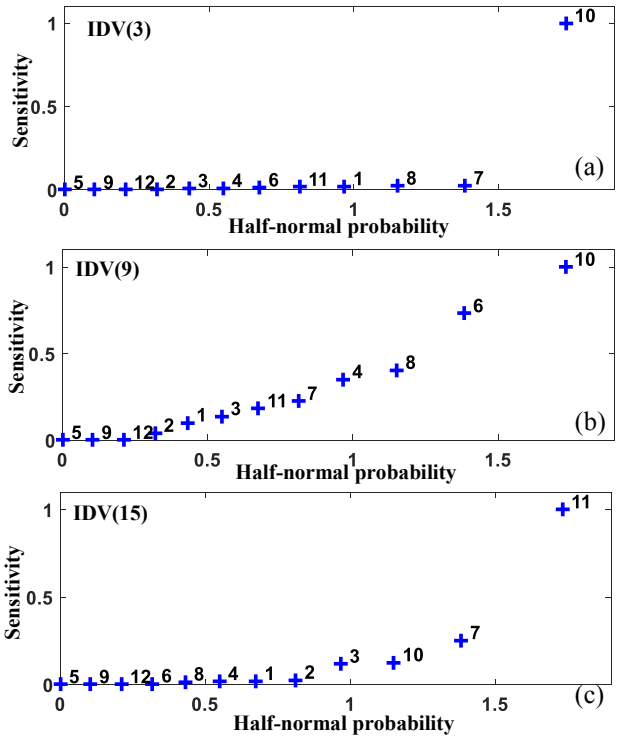

Fig. 4. Illustration of the effect of faults on manipulated variables (XMV) in the TE process

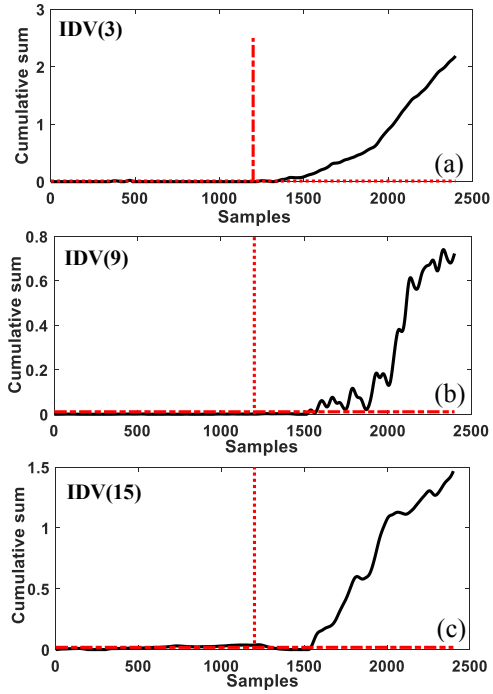

Fig. 5. Fault detection results of three particular faults

As seen in Table 2, there is a longer detection delay for these faults in Table 1 with the previously reported technique. For example, the $ARL_{o.c}$ of IDV(3) was approximately 114.8 *h*,

while only 3.9 *h* is required using the method developed in this work. Note that the slow detection is preferable, as compared to no detection at all in previously reported works.

Table 2. Evaluation of the control performance

$ARL_{o.c}$ (hour)	*previous*	*current*
IDV(*3*)	114.8	3.9
IDV(*9*)	142.2	32.7
IDV(*15*)	65.3	30.3

5. CONCLUSIONS

A new FDD algorithm using the multiscale and multivariate analysis is developed in this current work. The efficiency of the algorithm is demonstrated with a subset of faults in the Tennessee Eastman (TE) process that have been consistently found unobservable or undistinguishable with other previously reported techniques. A *pre-filtering* tool, using the Ensemble Empirical Mode Decomposition (EEMD) and the Principal Component Analysis (PCA), can efficiently extract fault signatures and reduce the effect of noise on the fault detection. The half-normal probability plot can identify variables that are sensitive to faults. The algorithm can successfully identify three particular faults in the TE process with small time delay. It is important to note that the measurement noise is assumed to small in this work, enhancements of the *pre-filtering* step may be required for FDD when the measurement noise is big.

ACKNOWLEDGMENTS

National Science Foundation (NSF-CMMI-1727487 and NSF-CMMI-1727487) is acknowledged for the financial support.

REFERENCES

Bernal-de-Lazaro, J. M. et al., 2016. Enhanced dynamic approach to improve the detection of small magnitude faults. *Chemical Engineering Science,* Volume 146, pp. 166-179.

Bin Shams, M., Budman, H. & Duever, T., 2010. *Fault detection using CUSUM based techniques with application to the Tennessee Eastman process.* Leuven, Belgium, 9th International Symposium on Dynamics and Control of Process Systems (DYCOPS) .

Chiang, L. H. et al., 2015. Diagnosis of multiple and unknown faults using the causal map and multivariate statistics. *Journal of Process Control,* Volume 28, pp. 27-39.

Ding, S. X. et al., 2009. Subspace method aided-driven design of fault detection and isolation system. *Journal of Process Control,* Volume 19, pp. 1496-1510.

Downs, J. J. & Vogel, E. F., 1993. A plant-wide industrial process control problem. *Computers and Chemical Engineering,* 17(3), pp. 245-255.

Du, Y. & Du, D., 2018. Fault detection and diagnosis using empirical mode decomposition based principal component analysis. *Computers & Chemical Engineering,* 115(12), pp. 1-21.

Du, Y., Duever, T. & Budman, H., 2015. Fault detection and diagnosis with parametric uncertainty using generalized polynomial chaos. *Computers and Chemical Engineering,* Volume 76, pp. 63-75.

Feng, Z., Qin, J. S. & Liang, M., 2016. Time-frequency analysis based on Vold-Kalman filter and higher order energy separation for fault diagnosis of wind turbine planetary gearbox under nonstationary conditions. *Renewable Energy,* Volume 85, pp. 45-56.

Gerlter, J., 1998. *Fault detection and diagnosis in engineering systems.* NJ, USA: Taylor & Francis.

Ghosh, K., Ng, Y. S. & Srinivasan, R., 2011. Evaluation of decision fusing strategies for effective collaboration among heterogeneous fault diagnostic methods. *Computers and Chemical Engineering,* Volume 35, pp. 342-355.

Hawkins, D. M. & Olwell, D. H., 1998. *Cumulative sum charts and charting for quality.* New York: Springer.

Hsu, C.-C., Chen, M.-C. & Chen, L.-S., 2010. A novel process monitoring approach with dynamic independent component analysis. *Control Engineering Practice,* Volume 18, pp. 242-253.

Isermann, R., 2005. Model based fault detection and diagnosis - status and applications. *Annual reviews in control,* Volume 29, pp. 71-85.

Isermann, R., 2006. *Fault diagnosis systems: An introduction from fault detection to fault tolerance.* Berlin, Germany: Springer.

Jiang, Q., Huang, B., Ding, S. X. & Yan, X., 2016. Bayesian fault diagnosis with asynchronous measurements and its application in networked distributed monitoring. *IEEE Transactions on Industrial Electronics,* 63(10), pp. 6316-6324.

Kim, Y. et al., 2016. Robust leak detection and its localization using interval estimation for water distribution network. *Computers and Chemical Engineering,* 92(2), pp. 1-17.

Lee, J. M., Qin, S. J. & Lee, I. B., 2006. Fault detection and diagnosis based on modified independent component analysis. *The American Journal of Chemical Engineering,* 52(10), pp. 3501-3514.

Lee, J. M., Yoo, C. & Lee, I. B., 2004. Statistical monitoring of dynamic processes based on dynamic independent component analysis. *Chemical Engineering Science,* Volume 59, pp. 2995-3006.

Misra, M., Yue, H. H., Qin, S. J. & Ling, C., 2002. Multivariate process monitoring and fault diagnosis by multiscale PCA. *Computers and Chemical Engineering,* Volume 26, pp. 1281-1293.

Shang, J., Chen, M., Ji, H. & Zhou, D., 2017. Recursive transformed component statistical analysis for incipient fault detection. *Automatica,* Volume 80, pp. 313-327.

Sheriff, M. Z. et al., 2017. Fault detection using multiscale PCA-based moving window GLRT. *Journal of Process Control,* Volume 54, pp. 47-64.

Wong, W. C. & Lee, J. H., 2010. Fault detection and diagnosis using hidden markov disturbance models. *Industrial and Engineering Chemistry Research,* 49(17), pp. 7901-7908.

Wu, Z. & Huang, N. E., 2009. Ensemble empirical mode decomposition: a noise assisted data analysis method. *Advances in Adaptive Data Analysis,* Volume 1, pp. 1-41.

Yin, S. et al., 2012. A comparison study of basic data driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process. *Journal of Process Control,* 22(9), pp. 1567-1581.