

## Research Article

# Fault Detection and Identification for Nonlinear Process Based on Inertia-Based KEPCA and a New Combined Monitoring Index

Loubna El Fattahi <sup>1</sup> and El Hassan Sbai <sup>2</sup>

<sup>1</sup>Department of Physics, Faculty of Science, Moulay Ismail University, Meknes, Morocco

<sup>2</sup>High School of Technology, Moulay Ismail University, Meknes, Morocco

Correspondence should be addressed to Loubna El Fattahi; [l.elfattahi@edu.umi.ac.ma](mailto:l.elfattahi@edu.umi.ac.ma)

Received 22 April 2021; Revised 31 May 2021; Accepted 22 June 2021; Published 14 July 2021

Academic Editor: Jose R. C. Piqueira

Copyright © 2021 Loubna El Fattahi and El Hassan Sbai. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the present study, we introduce a new approach for the nonlinear monitoring process based on kernel entropy principal component analysis (KEPCA) and the notion of inertia. KEPCA plays double roles. First, it reduces the data in the high-dimensional space. Second, it constructs the model. Before data reduction, KEPCA transforms input data into high-dimensional feature space based on a nonlinear kernel function and automatically determines the number of principal components (PCs) based on the computation of the inertia. The retained PCs express the maximum inertia entropy of data in the feature space. Then, we use the Parzen window estimator to compute the upper control limit (UCL) for inertia-based KEPCA instead of the Gaussian assumption. Our second contribution concerns a new combined index based on the monitoring indices  $T^2$  and SPE in order to simplify the detection task of the fault and prevent any confusion. The proposed approaches have been applied to process fault detection and diagnosis for the well-known benchmark Tennessee Eastman process (TE). Results were performing.

## 1. Introduction

Since faults are integrated into most industrial processes' execution, fault detection and identification is mandatory in monitoring and diagnosing different processes, in order to ensure a riskless functionality and improve the productivity of an industrial process. In most industrial systems, including chemical processes, the collected data are large, extremely correlated, and subject of noise. To this end, a data preprocessing is required to get rid of redundant data and extract only the significant information. This limitation has led to the development of multivariate statistical methods, such as principal component analysis (PCA) [1, 2], recursive PCA [3], entropy PCA [4], kernel principal component analysis (KPCA) [5–7], and modified kernel PCA [8]. Thereby, PCA is a linear projection method that has a weakness in the case of some chemical processes that present nonlinear characteristics; thus, PCA does not handle well fault detection [9]. Recently, the KPCA gained more

attention in different research fields, particularly in the chemical process fault detection and diagnosis [10–12]. KPCA enhances the computation efficiency by selecting only some features from a higher dimensional feature space. Thus, the computational complexity of the kernel matrix is minimised in KPCA [13]. Unless, KPCA is known to perform better result than PCA in monitoring nonlinear industrial processes [14, 15]; however, it is still confusing in determining and selecting automatically the optimal number of features in the feature space. In fact, we need to provide as an input the variance expressed by the maintained PCs. To eliminate this deficiency, we suggest applying our method, inertia-based KPCA combined to the Shannon entropy [16, 17], for the first time to monitor the chemical Tennessee Eastman process. Indeed, by incorporating the concept of entropy to the inertia-based KPCA, we obtain the inertia-based KEPCA, and ensuring an extraction of the PCs with higher valuable information. Accordingly, each sample contributes to the inertia that is considered as a probability,

where Shannon entropy can be applied. Despite the widespread use of the Hotelling'  $T^2$  and the square predicted error (SPE) statistics in monitoring, especially in chemical processes [2, 3, 8–10], they simultaneously detect fault very frequently, which is confusing for process operators conducting the fault analysis. Hence, the second role of our contribution is to cope with this issue by combining the monitoring indices  $T^2$  and SPE statistics into a single index. The combined index is constructed by integrating the information gained from the principal and the residual spaces together. With the application of our proposed new combined index, the detection of the fault is achieved and compared to another index developed in [10]. The simulation benchmark has been carried out on the chemical Tennessee Eastman process. We demonstrate that the proposed inertia-based KEPCA method resolves the shortcoming of the standard KPCA method, in terms of selecting the features automatically and in term of fault detection.

The present work is organized as follows. Section 2 presents the inertia-based KEPCA approach for fault detection and identification; also, the proposed new combined index is discussed. Section 3 presents the application of our approaches: inertia-based KEPCA and the new combined index to monitor the nonlinear Tennessee Eastman process. Section 4 concerns conclusion driven from this study.

## 2. Methods and Algorithms

*2.1. Introduction to the Inertia-Based KEPCA Method.* KEPCA is an unsupervised learning technique, as we documented it in [16]; also, it is an enhancement for the existing linear method EPCA [18]. In fact, it is the combination of the EPCA and the kernel mapping. The input data are indirectly mapped, using a kernel function, into a high-dimensional feature space and then executes linear EPCA in the output features. As a result, the reproducing kernel Hilbert space (RKHS) is constructed using the Gaussian function. Based on the concept of inertia, each mapped sample contributes to both explained and residual inertia, which are considered as a probability distribution. Afterward, the maximum of Shannon entropy was applied to the probability distribution to extract features expected to preserve only the most valuable information among all features. Furthermore, the inertia-based KEPCA is simply calculated away from any nonlinear optimization, and within it, large nonlinearities can be eliminated. The related algorithm of KEPCA is provided in detail in [16]. It is worth mentioning that the inertia-based KEPCA-based monitoring has the same construction steps as the standard KPCA [14, 15], and the main difference lies in performing the entropy to inertia-based PCA instead of only applying the PCA.

The Gaussian function for the data mapping step for the nonlinear process monitoring is given by

$$K_G(x, x') = \exp\left(-\frac{\|x - x'\|^2}{c}\right), \quad (1)$$

where  $c$  is the kernel bandwidth, predetermined by the user, proportional to the number of variables in the process at the hand.

*2.2. Parzen Window Estimator for Probability Density Function (PDF).* Great number of estimators exists in the literature for probability density estimation. Considering the nonparametric estimation methods, we find the Parzen window estimator, kNN estimator, adaptive estimator, and variable estimator that is the combination of both Parzen and kNN estimators and others [19]. Hence, the method of Parzen remains the most largely used one for the estimation of the PDF. In fact, the Parzen window estimator is a practical tool that can extract density distribution empirically and provide a smooth PDF from population under consideration. The estimator's expression is given as follows [20, 21]:

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n G\left(\frac{D(x, X_i)}{h}\right), \quad (2)$$

where  $G$  is the kernel function (Gaussian in our case),  $h$  is the bandwidth, and  $D(x, X_i)$  is the distance separating sample  $x$  and samples  $X_i$ . Though, the choice of the bandwidth  $h$  of the Parzen estimator is awkward. In this regard, in a diversity of research studies have been done mean integrated squared error (MISE) [22, 23], cross-validation, and others [19]. Here, we opt for the AMISE method cited in [24] to select the optimal value that gives effective results in smoothing parameter selection, especially in univariate random data.

*2.3. Fault Detection Metrics.* The fault detection phase is mandatory in monitoring nonlinear processes. Consequently, a variety of indices was developed [25]. The most widely used among them are the  $T^2$  and SPE statistics. They measure the variation within the feature space founded by the multivariate statistical method such as KPCA. Indeed,  $T^2$  refers to the sum of normalized squared scores, while the SPE is the sum of squared errors, in the principal space and residual space, respectively. In the present work, we will use the same expressions elaborated by previous research studies for the  $T^2$  and SPE [2, 3, 8–10] and combine both of them in order to construct one index.

*2.4. Proposed New Combined Index (NCI).* Practically, one index is much easier for process monitoring, rather than two indices. Since  $T^2$  and SPE behave in a complementary way, they can be combined into one index. In this regard, Yue and Qin [26] proposed a combined index based on the  $T^2$  and SPE statistics. Thereafter, Choi et al. [10] suggested another unified index but with a regularized parameter. Accordingly, the main difference between our NCI and the one in [26] lies in the use of the probability density function (PDF) estimated through the Parzen window estimator to derive the upper control limits instead of the Gaussian assumption. This is because the Gaussian assumption-based UCL has been proven to not be valid and may provide false results considering that data can be non-Gaussian [15]. Therefore, our proposed NCI will be as follows:

$$\text{NCI}(x) = \frac{\text{SPE}(x)}{\text{ucl}_{\text{SPE}}} + \frac{T^2(x)}{\text{ucl}_{T^2}}. \quad (3)$$

The expressions  $\text{ucl}_{\text{SPE}}$  and  $\text{ucl}_{T^2}$  are given by integrating the estimated density through the Parzen window estimator (equation (2)) of training samples in the feature space. Let  $\beta$  be the significance level and  $\hat{f}$  the estimated PDF. Thus, the expressions are as follows:

$$\int_{-\infty}^{\text{SPE}_{\beta}} \hat{f}(\text{SPE}) d\text{SPE} = \beta, \quad (4)$$

$$\int_{-\infty}^{T^2_{\beta}} \hat{f}(T^2) dT^2 = \beta.$$

The fault is detected ( $F_k = 1$ ) by the NCI, if the following condition is satisfied:

$$F_k = \text{NCI}_k > \text{NCI}_{\text{UCL}}(\beta). \quad (5)$$

**2.5. Faulty Variable Identification.** Once the fault is detected, it is necessary to know which variables are the root of the cause. In this context, several works have proposed localization methods based on the calculation of the contributions of each variable to the monitoring indices [8, 11]. In these approaches, the variable with the greatest contribution is considered as the one that causes the appearance of the fault. The present work uses the sensitivity analysis method [8, 27] to identify the faulty variables. It is founded on the fact that the parameters causing the problem can be detected as a rate of change in the output variables of the system. The contributions of to the  $i^{\text{th}}$  variable of a given vector  $x_i \in \mathbb{R}^m$ , with  $m$  variables, to each index are

$$\text{Con}(T^2_i) = x_i \frac{\partial T^2}{\partial x_i}, \quad (6)$$

$$\text{Con}(\text{SPE}_i) = x_i \frac{\partial \text{SPE}}{\partial x_i}.$$

Accordingly, the contribution of the same variable to our proposed index will be

$$\text{Con}(\text{NCI}_i) = x_i \left( \frac{1}{\text{ucl}_{T^2}} \frac{\partial T^2}{\partial x_i} + \frac{1}{\text{ucl}_{\text{SPE}}} \frac{\partial \text{SPE}}{\partial x_i} \right). \quad (7)$$

**2.6. Outline and Online Inertia-Based KEPCA Fault Detection Algorithm.** Throughout the offline training phase, the normal operating data are used, where several parameters are determined: our NCI, including  $T^2$  and SPE statistics, the significance level, and the UCL of the NCI. Whereas, the online monitoring is useful to compute the detection indices  $T^2$  and SPE of the further vector of measurement. The process is judged to be under abnormal conditions if it surpasses the UCL of the NCI.

The following lines represent the steps of the outline and online processing of the inertia-based KEPCA fault detection algorithm.

Phase1: offline training.

- Step 1: scale the input data obtained under normal operating conditions (NOC)
- Step 2: choose the type of kernel function for mapping
- Step 3: compute the kernel matrix of the input data under NOC and centre it
- Step 4: obtain eigen-decomposition and sort eigenvalues with their corresponding eigenvectors in the descending order
- Step 5: orthonormalize the eigenvectors
- Step 6: compute the explained and residual inertia of each PCs
- Step 7: compute the sum of Shannon entropy of each inertia
- Step 8: select the optimal number of PCs that corresponds to the maximum of sum of entropy
- Step 9: calculate NCI using the inertia-based KEPCA method
- Step 10: determine UCL of the NCI

Phase 2: online monitoring.

- Step 1: obtain test sample  $x$  and normalize it as in the offline phase
- Step 2: compute the kernel vector of the test sample  $x$  and centre it
- Step 3: select the PCs of the test sample, which have the maximum of Shannon entropy of the inertia
- Step 4: compare NCI of the test sample with its respective upper control limit provided by the model development phase; the process is in control, if NCI is less than its monitoring index. Other way, the process is out-of-control.

### 3. Application: Tennessee Eastman Process

To demonstrate our proposed approach inertia-based KEPCA monitoring, we consider the famous chemical benchmark Tennessee Eastman process (TEP). Figure 1 shows the control structure of the process. It is a non-linear real industrial process described by dynamic properties. It encompasses five main parts: a compressor, a reactor, a condenser, a stripper, and a separator and is adopted as an unavoidable industrial benchmark process in evaluating monitoring approaches [28]. TEP comprises 22 continuous variables (in blue), 12 variables which are manipulated (in orange), and 19 composition measurement (in green). The sampling interval was set at 3 min, and the fault was injected at sample 160. More details and operating conditions are described in [28, 29]. Tables 1 and 2 describe the variables and the studied faults of the TE process, respectively.

**3.1. Fault Detection Rule.** Concerning the fault detection, some works consider that fault is detected once the monitoring index exceeds its control limit at the first time [9, 10], while other researchers define the detection time as the first sample number after eight sequential samples that have exceeded the control limit [8]. Whereas in [16], the fault

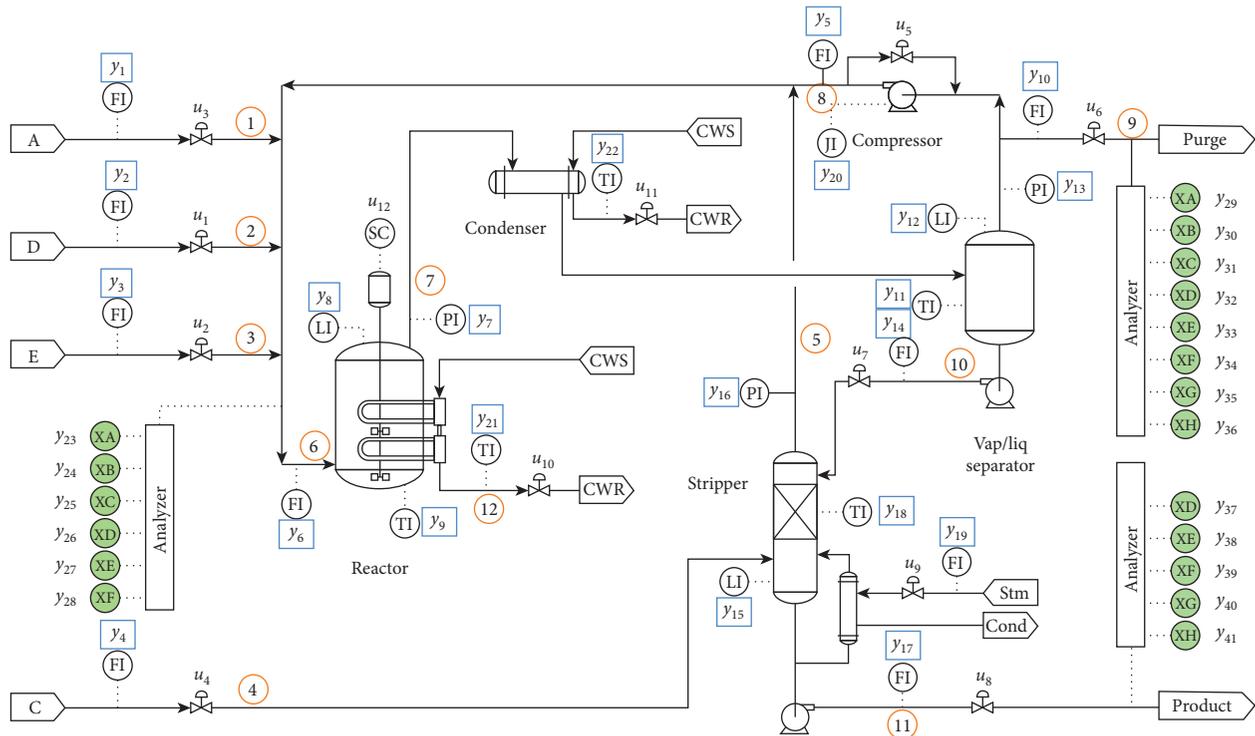


FIGURE 1: Tennessee Eastman process flowchart.

TABLE 1: Variables of the TE process.

No.	Process variable
1	A feed (stream 1)
2	D feed (stream 2)
3	E feed (stream 3)
4	Total feed (stream 4)
5	Recycle flow (stream 8)
6	Reactor feed rate (stream 6)
7	Reactor pressure
8	Reactor level
9	Reactor temperature
10	Purge rate (stream 9)
11	Product separator temperature
12	Product separator level
13	Product separator pressure
14	Product separator under flow (stream 10)
15	Stripper level
16	Stripper pressure
17	Stripper under flow (stream 11)
18	Stripper temperature
19	Stripper steam flow
20	Compressor work
21	Reactor cooling water outlet temperature
22	Separator cooling water outlet temperature
23–28	Components A, B, C, D, E, and F in stream 6
29–36	Components A, B, C, D, E, F, G, and H in stream 9
37–41	Components D, E, F, G, and H in stream 11
42	MV for D feed flow (stream 2)
43	MV for E feed flow (stream 3)
44.	MV for A feed flow (stream 1)
45	MV for total feed flow (stream 4)

TABLE 1: Continued.

No.	Process variable
46	MV for compressor recycle valve
47	MV for purge valve (stream 9)
48	MV for separator pot liquid flow (stream 10)
49	MV for stripper liquid prod flow (stream 11)
50	MV for stripper steam valve
51	MV for reactor cooling water flow
52	MV for condenser cooling water flow

TABLE 2: Faults description for the TE process.

Fault no.	Process variable	Fault type
1	A/C feed ratio, B composition constant (stream 4)	Step
2	B composition, A/C feed ratio constant (stream 4)	Step
3	D feed temperature (stream 2)	Step
4	Reactor cooling water inlet temperature	Step
5	Condenser cooling water inlet temperature	Step
6	A feed loss (stream 1)	Step
7	C header pressure loss-reduced availability (stream 4)	Step
8	A, B, and C feed compositions (stream 4)	Random variation
9	D feed temperature (stream 2)	Random variation
10	C feed temperature (stream 4)	Random variation
11	Reactor cooling water inlet temperature condenser	Random variation
12	Cooling water inlet temperature	Slow shift
13	Reaction kinetics	Sticking
14	Reactor cooling water valve	Sticking
15	Condenser cooling water valve	Sticking
16–20	Unknown	Unknown
21	Valve position constant (stream 4)	Constant position

detection is counted when a monitoring index surmounts the control limit in at least two consecutive occasions. To be more accurate in terms of comparison, we consider in our study, fault detection as successfully accomplished, when the monitoring index surpasses the control limit in at least two times consecutively as in [16]. Then, we evaluate the monitoring performance by the computation time (CT), the missed detection rate (MDR), and false alarm rate (FAR) metrics as in [25] based on the fault detection rule mentioned above.

**3.2. Monitoring Performance Metrics.** To evaluate the fault detection performance, three metrics are generally used: computation time, which is defined as the period between the time when a fault was introduced and the time of its detection. The false alarm rate (FAR) which expresses the percentage of the violated samples that exceeds the detection limit, to the faultless data. It can be computed as follows [25]:

$$\text{FDR} = \frac{\text{Violated samples}}{\text{Faultless data}} \times 100. \quad (8)$$

The missed detection rate (MDR) is provided by the percentage of the faulty data that does not surpass the detection limit to all the faulty data.

$$\text{MDR} = \frac{\text{Missed detection}}{\text{Faulty data}} \times 100. \quad (9)$$

**3.3. Experimental Result.** In this section, we demonstrate the efficiency of our proposed approach inertia-based KEPCA using faults 1 and 14 of the nonlinear TE process. The fault 1 concerns the change in the component A/C feed ratio and B composition constant, while the fault 14 is related to the valve of the reactor cooling water. To construct the model, we used all aforementioned variables apart from the 12<sup>th</sup> of the manipulated variable (the agitation speed of the reactor's stirrer) as it is a constant. Besides, all 500 samples gathered under the NOC were employed as a training set. While, 960 samples were gathered under faulty operating conditions and used as test data, with 52 variables for each. The number of automatically selected PCs is 17, expressing maximum of inertia entropy. The parameter of the Gaussian function, incorporated in model development for mapping, was set as  $c = n\sigma^2W$ , where  $n$  is the number of studied variables,  $\sigma^2$  is the variance of the input data, and  $W$  is a constant depending on the process being examined. For the TE process,  $W$  is equal to 40 as it was set in [15].

In Figure 2, the number of retained PCs is determined automatically. The retained PCs are necessary and sufficient in terms of information content (Shannon entropy). They are estimated to be 17 PCs considering all 52 variables of the TE process. Contrary to the work in [15], the number of retained PCs was indirectly specified through the cumulative percent eigenvalue (CPE), as an input parameter. Besides,

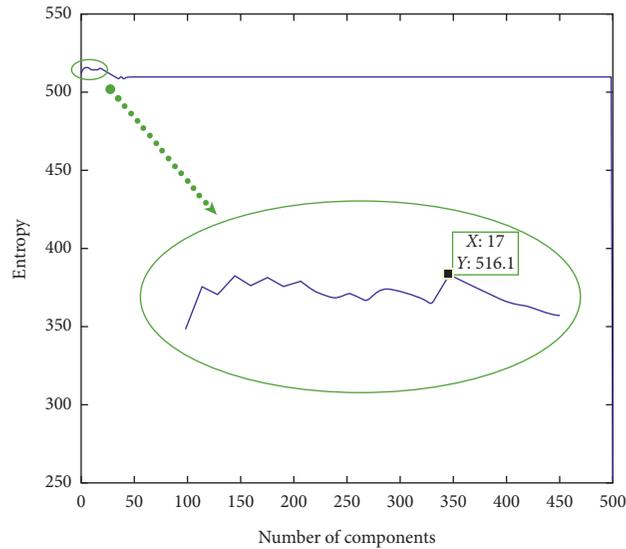


FIGURE 2: The selected optimal number of PCs.

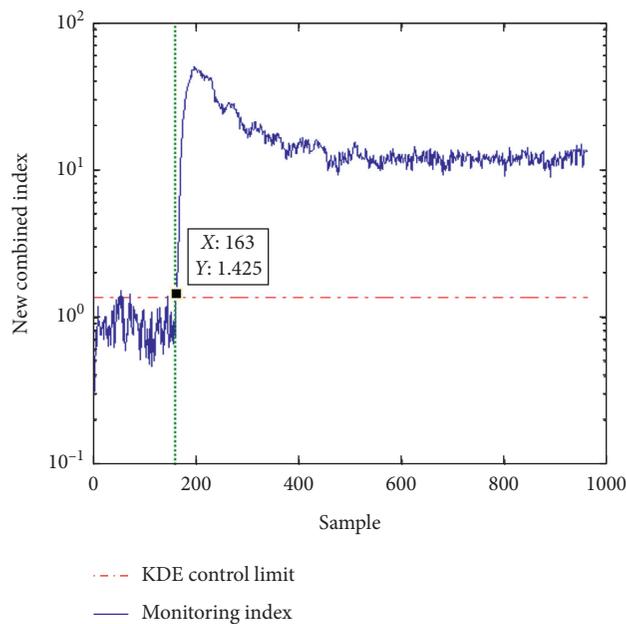


FIGURE 3: Monitoring index for fault 1.

they examine only 33 variables without a clear justification, which may leave out valuable information from data.

Here, it is worth noting that for all studied faults, we obtained the same number of selected PCs, which is 17 PCs.

In Figures 3 and 4, we can see the monitoring index plotted as a solid curve, whereas the horizontal dashed line represents the 99% confidence limit. The vertical dashed line represents the 160<sup>th</sup> sample when the fault was introduced. Our NCI of inertia-based KEPCA model detects quickly fault 1 at 163<sup>th</sup> sample and fault 14 at 161<sup>th</sup> sample. Therefore, the inertia-based KEPCA-based monitoring with our NCI demonstrates the existence of the fault 1 and fault 14 along the fault operation period.

Figure 5 shows the result of variables contribution to the fault 1. It can be seen that the largest contribution comes from variable number 1 and variable number 44. Considering Table 1, they correspond to the A feed and the A feed flow valve variables, respectively. Evidently, these two variables are closely connected to the component A feed. While, Figure 6 reveals the variables' large contribution to the fault 14. We can see that variable numbers 9, 21, and 51 are the cause of this fault. Referring to Table 1, we can find that the variable number 9 concerns reactor temperature, variable number 21 deals with reactor cooling water inlet temperature, and variable number 51 relates to the reactor cooling

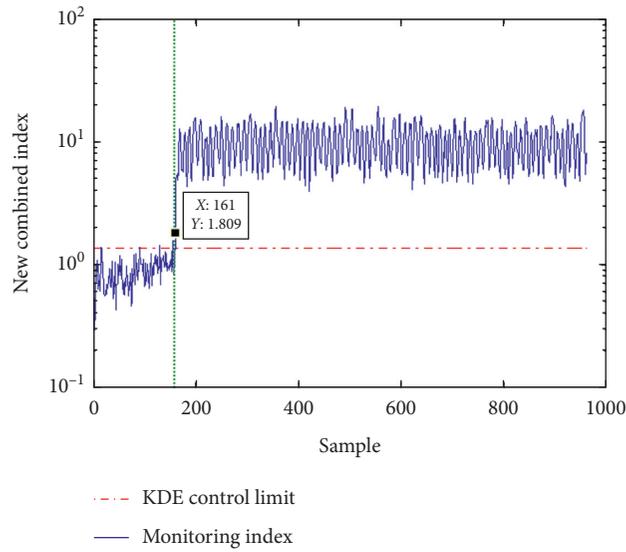


FIGURE 4: Monitoring index for fault 14.

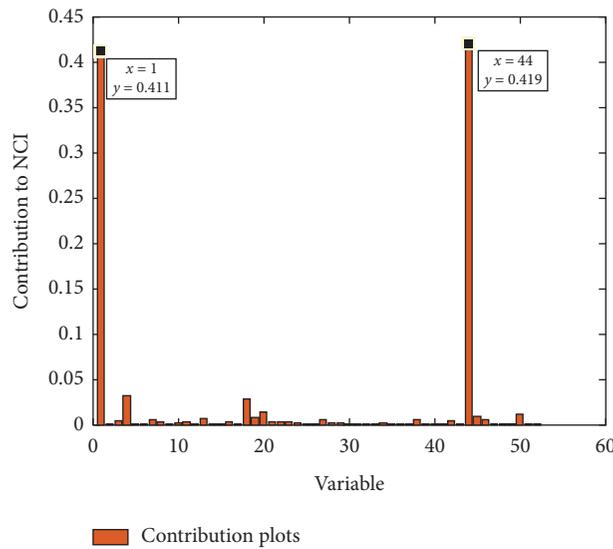


FIGURE 5: NCI-based contribution plot for fault 1.

water flow valve. Obviously, all mentioned variables are connected to fault of the reactor cooling water valve.

Tables 3 and 4 provide the detection rates and computation time, respectively, for all 21 faults using inertia-based KEPCA with our proposed NCI compared to the index proposed by Choi et al. [10] named the unified index. Results on TE show that the choice of the monitoring index

affects the process monitoring result. In term of FDRs, our NCI-Parzen has overall higher rates over its counterpart, the unified index. Besides, comparing our results to the results given in [15] named KPCA-KDE, we can see that our method competes with KEPCA-KDE. Both inertia-based KEPCA and KPCA-KDE provide good results, in terms of the detection rate for faults ID 1, 2, 4, 6, 7, 8, 12, 13, 14, and

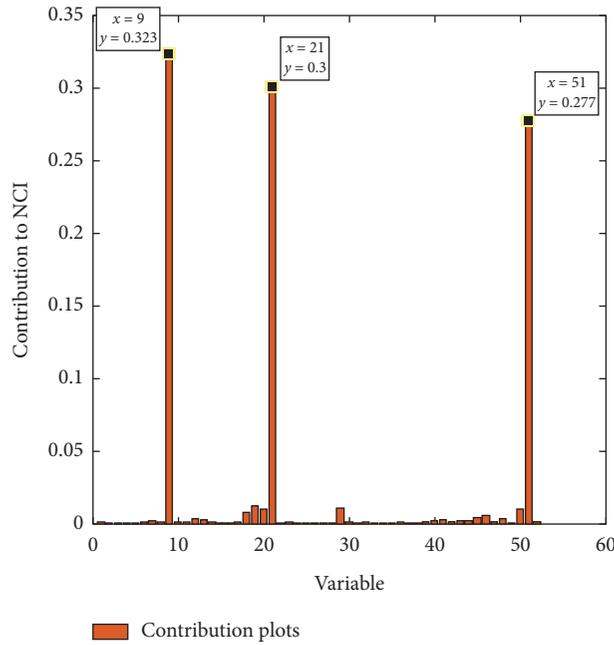


FIGURE 6: NCI-based contribution plot for fault 14.

TABLE 3: Fault detection rate (%) using different indices.

Fault ID	NCI	Unified index	KPCA-KDE
1	99.63	99.63	99.75
2	98.50	98.50	98.63
3	4.24	0.62	1.75
4	96.13	93.26	99.88
5	30.09	24.47	26.88
6	99.88	99.88	99.88
7	99.88	99.88	99.88
8	97.75	97.25	98.00
9	4.74	0.50	2.25
10	57.68	43.57	53.50
11	71.29	61.17	79.88
12	98.88	98.75	97.63
13	95.38	95.26	95.63
14	99.88	99.88	99.75
15	11.74	1.62	2.88
16	48.31	25.34	44.62
17	93.13	89.26	93.50
18	89.89	89.39	—
19	13.73	2.50	13.50
20	60.42	44.94	57.75
21	46.32	40.07	—

17. However, for faults that are proved difficult in detection (faults ID 3, 9, and 15), our method inertia-based KEPCA performs better than KPCA-KDE. Moreover, Table 4 shows that the elapsed time before a fault is detected for our NCI is

either less or equal to the one for the unified index. Whereas, the computation time for the KPCA-KDE and inertia-based KEPCA is still almost the same. Therefore, integrating our NCI to the inertia-based KEPCA technique for monitoring

TABLE 4: Computation time in minute.

Fault ID	NCI	Unified index	KPCA-KDE
1	9	9	6
2	36	36	33
3	129	129	1725
4	9	9	3
5	3	3	3
6	3	3	3
7	3	3	3
8	45	66	48
9	9	18	1725
10	57	81	180
11	18	21	15
12	9	9	42
13	111	114	105
14	3	3	6
15	1725	1737	1725
16	3	45	81
17	72	75	45
18	45	255	24
19	33	249	36
20	243	249	105
21	855	855	—

grants better monitoring results compared to the unified index combined to inertia-based KEPCA assumption.

#### 4. Conclusion

The present study aims to investigate our proposed method inertia-based KEPCA for selecting automatically the sufficient number of PCs for model development. The model was constructed under NOC, and only 17 PCs were automatically estimated to be sufficient and expressed maximum inertia entropy. Hence, this approach can be integrated to the statistical multivariate methods. In addition, our developed new combined index based on  $T^2$  and SPE integrates the information gained from the principal and the residual spaces together. The NCI detects more and earlier fault than the unified index in terms of fault detection rate and computation time. Thus, it can be expanded to the nonlinear process fault detection in order to simplify the detection task and avoid confusion. The UCL of our NCI was derived directly from its PDF estimated using the Parzen window estimator. Overall results show that our approaches improve dimensionality reduction, the fault detection, and identification applied to the nonlinear benchmark Tennessee Eastman process.

#### Data Availability

The Tennessee Eastman data used to support the findings of this study have been deposited in <http://depts.washington.edu/control/LARRY/TE/download.html>.

#### Conflicts of Interest

The authors declare that there are no conflicts of interest.

#### References

- [1] B. De Ketelaere, M. Hubert, and E. Schmitt, "Overview of PCA-based statistical process-monitoring methods for time-dependent, high-dimensional data," *Journal of Quality Technology*, vol. 47, no. 4, pp. 318–335, 2015.
- [2] J. Wang and Q. P. He, "Multivariate statistical process monitoring based on statistics pattern analysis," *Industrial & Engineering Chemistry Research*, vol. 49, no. 17, pp. 7858–7869, 2010.
- [3] W. Li, H. H. Yue, S. Valle-Cervantes, and S. J. Qin, "Recursive PCA for adaptive process monitoring," *Journal of Process Control*, vol. 16, 2000.
- [4] X. Deng and X. Tian, "Entropy principal component analysis and its application to nonlinear chemical process fault diagnosis," *Asia-Pacific Journal of Chemical Engineering*, vol. 9, no. 5, pp. 696–706, 2014.
- [5] 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.
- [6] S. Romdhani, S. Gong, and A. Psarrou, "A multi-view nonlinear active shape model using kernel PCA," in *Proceedings of the British Machine Vision Conference 1999*, pp. 48.1–48.10, Nottingham, UK, 1999.
- [7] X. Zhang, J. Chen, and Q. Gan, "Anomaly detection for aviation safety based on an improved KPCA algorithm," *Journal of Electrical and Computer Engineering*, vol. 2017, Article ID 4890921, 8 pages, 2017.
- [8] X. Deng, X. Tian, and S. Chen, "Modified kernel principal component analysis based on local structure analysis and its application to nonlinear process fault diagnosis," *Chemometrics and Intelligent Laboratory Systems*, vol. 127, pp. 195–209, 2013.
- [9] J. Shang, M. Chen, and H. Zhang, "Fault detection based on augmented kernel Mahalanobis distance for nonlinear dynamic processes," *Computers & Chemical Engineering*, vol. 109, pp. 311–321, 2018.

- [10] S. W. Choi, C. Lee, J.-M. Lee, J. H. Park, and I.-B. Lee, "Fault detection and identification of nonlinear processes based on kernel PCA," *Chemometrics and Intelligent Laboratory Systems*, vol. 75, no. 1, pp. 55–67, 2005.
- [11] J.-H. Cho, J.-M. Lee, S. Wook Choi, D. Lee, and I.-B. Lee, "Fault identification for process monitoring using kernel principal component analysis," *Chemical Engineering Science*, vol. 60, no. 1, pp. 279–288, 2005.
- [12] X. Deng and X. Tian, "Multivariate statistical process monitoring using multi-scale kernel principal component analysis," *IFAC Proceedings Volumes*, vol. 39, no. 13, pp. 108–113, 2006.
- [13] M. Kano, S. Hasebe, I. Hashimoto, and H. Ohno, "A new multivariate statistical process monitoring method using principal component analysis," *Computers & Chemical Engineering*, vol. 25, no. 7-8, pp. 1103–1113, 2001.
- [14] J.-M. Lee, C. Yoo, S. W. Choi, P. A. Vanrolleghem, and I.-B. Lee, "Nonlinear process monitoring using kernel principal component analysis," *Chemical Engineering Science*, vol. 59, no. 1, pp. 223–234, 2004.
- [15] R. T. Samuel and Y. Cao, "Nonlinear process fault detection and identification using kernel PCA and kernel density estimation," *Systems Science & Control Engineering*, vol. 4, no. 1, pp. 165–174, 2016.
- [16] L. E. Fattahi and E. H. Sbai, "Kernel entropy principal component analysis using Parzen estimator," in *Proceedings of the 2018 International Conference on Intelligent Systems and Computer Vision (ISCV)*, pp. 1–8, Fez, Morocco, April 2018.
- [17] L. E. Fattahi and E. H. Sbai, "Clustering using kernel entropy principal component analysis and variable kernel estimator," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 3, 2021.
- [18] K. Slaoui, *Optimisation des méthodes d'analyses factorielles, discriminante, et de classification des données statistiques et dynamiques par le principe de maximum d'entropie*, Ph.D. thesis, Sidi Mohamed Ben Abdellah University, Fez, Morocco, 2000.
- [19] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*, Chapman and Hall/CRC, Boca Raton, FL, USA, 1986.
- [20] M. Rosenblatt, "Remarks on some nonparametric estimates of a density function," *The Annals of Mathematical Statistics*, vol. 27, no. 3, pp. 832–837, 1956.
- [21] E. Parzen, "On estimation of a probability density function and mode," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1065–1076, 1962.
- [22] J. Liang, "Multivariate statistical process monitoring using kernel density estimation," *Developments in Chemical Engineering and Mineral Processing*, vol. 13, no. 1-2, pp. 185–192, 2008.
- [23] R. L. Hyndman, X. Zhang, and M. L. King, "Bandwidth selection for multivariate kernel density estimation using MCMC," in *Proceedings of the Econometric Society 2004 Australasian Meetings*, Perth, Australia, 2004.
- [24] L. Sun and M. Jiang, "Research on the optimized clustering method based on CFSFDP," in *Proceedings of the 2019 34th Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, pp. 28–33, Jinzhou, China, June 2019.
- [25] I. Jaffel, O. Taouali, I. Elaïssi, and H. Messaoud, "Comparative study of PCA approaches for fault detection: application to a chemical reactor," in *Proceedings of the 2013 International Conference on Control, Decision and Information Technologies (CoDIT)*, pp. 57–62, Hammamet, Tunisia, May 2013.
- [26] H. H. Yue and S. J. Qin, "Reconstruction-based fault identification using a combined index," *Industrial & Engineering Chemistry Research*, vol. 40, no. 20, pp. 4403–4414, 2001.
- [27] L. Petzold, S. Li, Y. Cao, and R. Serban, "Sensitivity analysis of differential-algebraic equations and partial differential equations," *Computers & Chemical Engineering*, vol. 30, no. 10–12, pp. 1553–1559, 2006.
- [28] 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.
- [29] T. J. McAvoy and N. Ye, "Base control for the Tennessee Eastman problem," *Computers & Chemical Engineering*, vol. 18, no. 5, pp. 383–413, 1994.