

Research Article

Anomaly Detection and Localization for Process Security Based on the Multivariate Statistical Method

Imen Hamrouni ¹, Hajer Lahdhiri,² Khaoula Ben Abdellafou,^{3,4} Ahamed Aljuhani,⁵ Okba Taouali ^{2,5} and Kais Bouzrara²

¹National Engineering School of Sousse, Electronics and Micro Electronic Laboratory, Faculty of Sciences of Monastir, Monastir, Tunisia

²National Engineering School of Monastir, Monastir, Tunisia

³Department of Computer Science, Faculty of Computers and Information Technology, University of Tabuk, Tabuk, Saudi Arabia

⁴University of Sousse, ISITCom, MARS Research Laboratory, LR17ES05, 4011, Hammam Sousse, Tunisia

⁵Department of Computer Engineering, Faculty of Computers and Information Technology, University of Tabuk, Tabuk, Saudi Arabia

Correspondence should be addressed to Imen Hamrouni; imenhamrouni.eniso@gmail.com

Received 21 February 2021; Revised 16 December 2021; Accepted 21 December 2021; Published 3 February 2022

Academic Editor: Abdesselem Boulkroune

Copyright © 2022 Imen Hamrouni et al. 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.

Anomaly detection is very important for system monitoring and security since successful execution of these engineering tasks depends on access to validated data. The localization of the variable causing the fault is very essential. Indeed, the localization of the fault is defined as the ability to determine the source of the fault on a system. Generally, the identification of faults is linked to the detection procedure implemented. Therefore, it is very important to choose the adequate fault detection model to locate fault. For nonlinear uncertain systems, the most performed fault detection method is reduced rank interval kernel principal component analysis (RRIKPCA), which enhances the computational skill by downgrading the kernel matrix dimension. We have proposed in this article a new fault localization technique for uncertain systems, named partial RRIKPCA, which combines the benefits of the RRIKPCA technique and the principle of partial localization. The principal of this method involves selecting partially reduced rank data subsets and then building more accurate models with fewer PCs and isolating faults with higher precision. The proposed fault isolation method is applied for monitoring air quality monitoring network (AIRLOR) data.

1. Introduction

Industrial processes are uncertain, caused by the imprecision of measurement; for more precision, data with incertitude became an interval-valued form [1]. For this reason, many researchers deal with this problem of uncertainties, and several linear PCA models for interval-valued data are presented in literature [2–4]. The center range PCA method is the supreme current interval multivariate technique.

Though in the nonlinear situation, author in [5] proposed the nonlinear interval-valued data-based KPCA approach (IKPCA). On the other hand, in the case of interval-valued data, there are limited available recent

methods for dimensionality diminution with IKPCA; the reduced interval kernel principal component analysis (RIKPCA) [6], and the interval reduced rank KPCA based on kernel generalized likelihood ratio test (IRR-KGLRT) [7].

The next important step after fault detection is the identification or localization, the goal of the identification is to determine the source and the cause of the fault, and this stage is related to the detection way employed. Some techniques based on PCA have been proposed in the literature, like localization of defects by PCA which is the approach of contribution diagrams [8]. Unlike the reconstruction approach [9], the contribution approach does not require any information on the defect to generate the plots.

As an alternative to contribution diagrams, the authors [10] proposed the use of reconstruction-based contributions (BCRs) that address the aforementioned gap. We can cite other approaches based on PCA such as the technique of localization of defects using the structuring of the residues [11]; this technique consists to form a set of residues, so that every residue is not vulnerable to all defects. So, we get for every defect a sign that simplifies the fault localization. An extension of this approach based on maximizing the compassion of residues to defects, known as partial PCA, was suggested by [12]-[13].

These majors fault localization treat only to certain systems; to deal with this limitation, the key objective of this study is to develop a new fault detection and localization approach applied to the nonlinear uncertain process that is capable to treat data learning characterized by nonlinear uncertain large datasets, which is called partial reduced rank IKPCA.

This technique consists of generating RRIKPCA models with reduced sets of variables. Thus, the aim of this technique of localization is to generate fault detection indices sensitive to certain faults and insensitive to others.

This study contains 5 sections: in Section 2, we present briefly review to the interval reduced rank (RRIKPCA) method. The proposed partial interval-reduced rank KPCA for interval-valued data is introduced in Section 3. Section 4 illustrates the utility of the proposed fault detection and identification using an air quality (AIRLOR) application.

2. RRIKPCA Review

2.1. Interval-Valued Data, Interval Midpoints-Radii. The interval form of data sample $x_j(k)$ is composed by the lower bound (LB) of the interval $\underline{x}_j(k)$, and the upper bound (UB) $\bar{x}_j(k)$ is a result of the incertitude of measurement:

$$[x_i(\kappa)] = [\underline{x}_i(\kappa), \bar{x}_i(\kappa)],$$

$$[X] = \begin{pmatrix} [\underline{x}_1(1), \bar{x}_1(1)] & \dots & [\underline{x}_i(1), \bar{x}_i(1)] & \dots & [\underline{x}_M(1), \bar{x}_M(1)] \\ \vdots & & \vdots & & \vdots \\ [\underline{x}_1(\kappa), \bar{x}_1(\kappa)] & & [\underline{x}_i(\kappa), \bar{x}_i(\kappa)] & & [\underline{x}_M(\kappa), \bar{x}_M(\kappa)] \\ \vdots & & \vdots & & \vdots \\ [\underline{x}_1(N), \bar{x}_1(N)] & \dots & [\underline{x}_i(N), \bar{x}_i(N)] & \dots & [\underline{x}_M(N), \bar{x}_M(N)] \end{pmatrix} \quad \kappa = 1, \dots, N; i = 1, \dots, M. \quad (1)$$

The center of an interval $x_i^c(\kappa)$ and its radius $x_i^r(\kappa)$ are given by

$$x_i^c(\kappa) = \frac{1}{2} (\bar{x}_i(\kappa) + \underline{x}_i(\kappa)),$$

$$x_i^r(\kappa) = \frac{1}{2} (\bar{x}_i(\kappa) - \underline{x}_i(\kappa)). \quad (2)$$

A new numerical data matrix can be reformed based on radius and centers:

$$[x_i(\kappa)] = [x_i^c(\kappa) - x_i^r(\kappa), x_i^c(\kappa) + x_i^r(\kappa)],$$

$$X_{CR} = [X^C \quad X^R], \quad (3)$$

with $X^C = [x_1^c, x_2^c, \dots, x_r^c] \in R^{r \times M}$ and $X^R = [x_1^r, x_2^r, \dots, x_N^r] \in R^{r \times M}$, and the new data are $x_{CR} = [x^c \quad x^r] \in R^{2M}$.

The average and the variance of $[X_i]$ are given by

$$M([X_i]) = \frac{1}{N} \sum_{\kappa=1}^N \frac{\bar{x}_i(\kappa) + \underline{x}_i(\kappa)}{2},$$

$$\text{VAR}([X_i]) = \langle [X_i], [X_i] \rangle = \sum_{\kappa=1}^N \frac{1}{3} (\underline{x}_i^2(\kappa) + \underline{x}_i(\kappa)\bar{x}_i(\kappa) + \bar{x}_i^2(\kappa)), \quad (4)$$

with the vector $[X_i]$ given by

$$[X_i] = [[x_i(1)] \quad \dots \quad [x_i(\kappa)] \quad \dots \quad [x_i(N)]]^T$$

$$= ([\underline{x}_i(1), \bar{x}_i(1)], [\underline{x}_i(2), \bar{x}_i(2)], \dots, [\underline{x}_i(N), \bar{x}_i(N)])^T. \quad (5)$$

The standardization of $[X_i(\kappa)]$ is given by

$$\frac{[x_i(\kappa)] - M([X_i(\kappa)])}{\sqrt{\text{VAR}([X_i(\kappa)])}} = \left[\frac{[\underline{x}_i(\kappa)] - M([X_i(\kappa)])}{\sqrt{\text{VAR}([X_i(\kappa)])}}, \frac{[\bar{x}_i(\kappa)] - M([X_i(\kappa)])}{\sqrt{\text{VAR}([X_i(\kappa)])}} \right]. \quad (6)$$

2.2. Fault Detection Using IKPCA Based on Interval Midpoints-Radii. The authors in [5, 14] suggested to involve uncertainties to the KPCA approach, and two models are

developed: model based on LB and UB and model based on midpoints and radii.

The kernel matrix K is presented by

$$K = \begin{bmatrix} \phi_1^T \phi_1 & \dots & \phi_1^T \phi_N \\ \vdots & \dots & \vdots \\ \phi_N^T \phi_1 & \dots & \phi_N^T \phi_N \end{bmatrix}. \quad (7)$$

Remember that ϕ is a nonlinear projection function in the characteristic space H (feature space): $H, \phi: x_i \in R^m \rightarrow \phi_i = \phi(x_i) \in R^h$.

$$\phi(x_i)^T \phi(x_{i'}) = k(x, x_{i'}), \quad (8)$$

with k as the core function defined as [15]

$$k(x_i, x_{i'}) = \exp\left[\frac{-\|x_i - x_{i'}\|^2}{2\sigma^2}\right]. \quad (9)$$

So, the kernel matrix K can be defined with $k(x, x_{i'})$ elements as follows:

$$K = \begin{bmatrix} k(x_1, x_1) & \dots & k(x_1, x_N) \\ \vdots & \dots & \vdots \\ k(x_N, x_1) & \dots & k(x_N, x_N) \end{bmatrix}. \quad (10)$$

To compute the selected PC (ℓ), we apply the cumulative percent variance (CPV) [16]. The cumulative percent variance (CPV) can be expressed as

$$\text{CPV}(\ell) = \frac{\sum_{j=1}^{\ell} \lambda_j}{\sum_{j=1}^m \lambda_j} 100\%, \quad (11)$$

where λ_j is the j^{th} eigenvalue.

The number ℓ of selected principal components is chosen if the CPV is higher than 95%.

$P = [\vartheta_1, \dots, \vartheta_{\ell}] \in R^{N \times \ell} = [1/B_1 \rho^T A_1, \dots, 1/B_{\ell} \rho^T A_{\ell}]$ is the matrix of ℓ principal loading eigenvectors in the feature space, and $\tilde{P} = [\vartheta_{\ell+1}, \dots, \vartheta_N] = [1/B_{\ell+1} \rho^T A_{\ell+1}, \dots, 1/B_N \rho^T A_N] \in R^{N \times (N-\ell)}$, the $N - \ell$ last principal loading eigenvectors [10].

The mapped data are arranged as $\rho = [\phi(x_1) \phi(x_2) \dots \phi(x_N)]^T$.

The new kernel matrix K_{CR} -based center and range in the new form is given by

$$K_{\text{CR}} = \begin{bmatrix} k(x_{\text{CR},1}, x_{\text{CR},1}) & \dots & k(x_{\text{CR},1}, x_{\text{CR},r}) \\ \vdots & \dots & \vdots \\ k(x_{\text{CR},r}, x_{\text{CR},1}) & \dots & k(x_{\text{CR},r}, x_{\text{CR},r}) \end{bmatrix}, \quad (12)$$

where

$$\begin{aligned} k(x_{\text{CR},k}) &= (k(x_{\text{CR},1}, x_{\text{CR},k}), \dots, k(x_{\text{CR},N}, x_{\text{CR},k}))^T, \\ P_{\text{CR}} &= [B_{\text{CR},1}^{-1} \rho_{\text{CR},1}^T A_{\text{CR},1}, \dots, B_{\text{CR},\ell}^{-1} \rho_{\text{CR},\ell}^T A_{\text{CR},\ell}], \\ \tilde{P}_{\text{CR}} &= [B_{\text{CR},\ell+1}^{-1} \rho_{\text{CR},\ell+1}^T A_{\text{CR},\ell+1}, \dots, B_{\text{CR},N}^{-1} \rho_{\text{CR},N}^T A_{\text{CR},N}], \end{aligned} \quad (13)$$

where A_{CR} and B_{CR} are the eigenvector and the eigenvalue of the matrix K_{CR} , and ρ_{CR}^T is the mapped data.

Now, we present the squared prediction error (Q) statistics for fault detection [17, 18].

Fault detection index $Q_{r\text{CR}}$ is given by

$$Q_{\text{CR}}(k) = k_{\text{CR}}(x_{\text{CR},k}, x_{\text{CR},k}) - k^T(x_{\text{CR},k}) V_{\text{CR}} k(x_{\text{CR},k}), \quad (14)$$

where $V_{\text{CR}} = P_{\text{CR}} P_{\text{CR}}^T$.

If $Q_{\text{CR}}(k) > \theta_{\beta}^{Q_{\text{CR}}}$, a fault is detected, where $\theta_{\beta}^{Q_{\text{CR}}}$ represent the control limits.

2.3. RRIKPCA Method. In the case of reduced rank IKPCA constructed on the centers and ranges of intervals RRIKPCA_CR, a new reduced data matrix is given by $X_{\text{RCR}} = [X_{\text{RCR}}^C \ X_{\text{RCR}}^R] \in R^{R \times 2m}$, where $X_{\text{RCR}}^C = [x_1^C, \dots, x_i^C, \dots, x_R^C]^T \in R^{R \times M}$ and $X_{\text{RCR}}^R = [x_1^R, \dots, x_i^R, \dots, x_R^R]^T \in R^{R \times M}$.

This reduction of data is a result of a selection by keeping the most useful observations that designate properly the system procedure this selection detailed in [18].

For each novel observation, the kernel matrix [19] is reorganized:

The general form of the kernel matrix is

$$K_R^t = \begin{bmatrix} K_R^{t-1} & k_{x_t} \\ k_{x_t}^T & k(x_1, x_t) \end{bmatrix} \in R^{R \times R}, \quad (15)$$

with $k_{x_t} = [k(x_1, x_t) \ \dots \ k(x_i, x_t) \ \dots \ k(x_s, x_t)]$.

The procedure of the RRIKPCA_CR model is illustrated in the flowchart (Figure 1).

3. Localization by Partial RRIKPCA

The principle of the partial RRIKPCA fault localization method consists subsets distinguishing data where some variables are excluded compared to the original data. So, some variables are missing in the partial RRIKPCA; in celebration of the residual structuring approach, an RR- IKPC is applied to a reduced vector.

The partial RRIKPCA technique has four steps: first, we create an extremely localizable incidence matrix; second, we apply the RRIKPCA model; third, we build a set of partial RRIKPCA models, each corresponding to a row of the theoretical signature matrix; and finally, we define the control thresholds to decide the fault detection. This procedure is shown in Figure 2.

After identifying the partial RRIKPCA models, they can be used for the isolation of defaults.

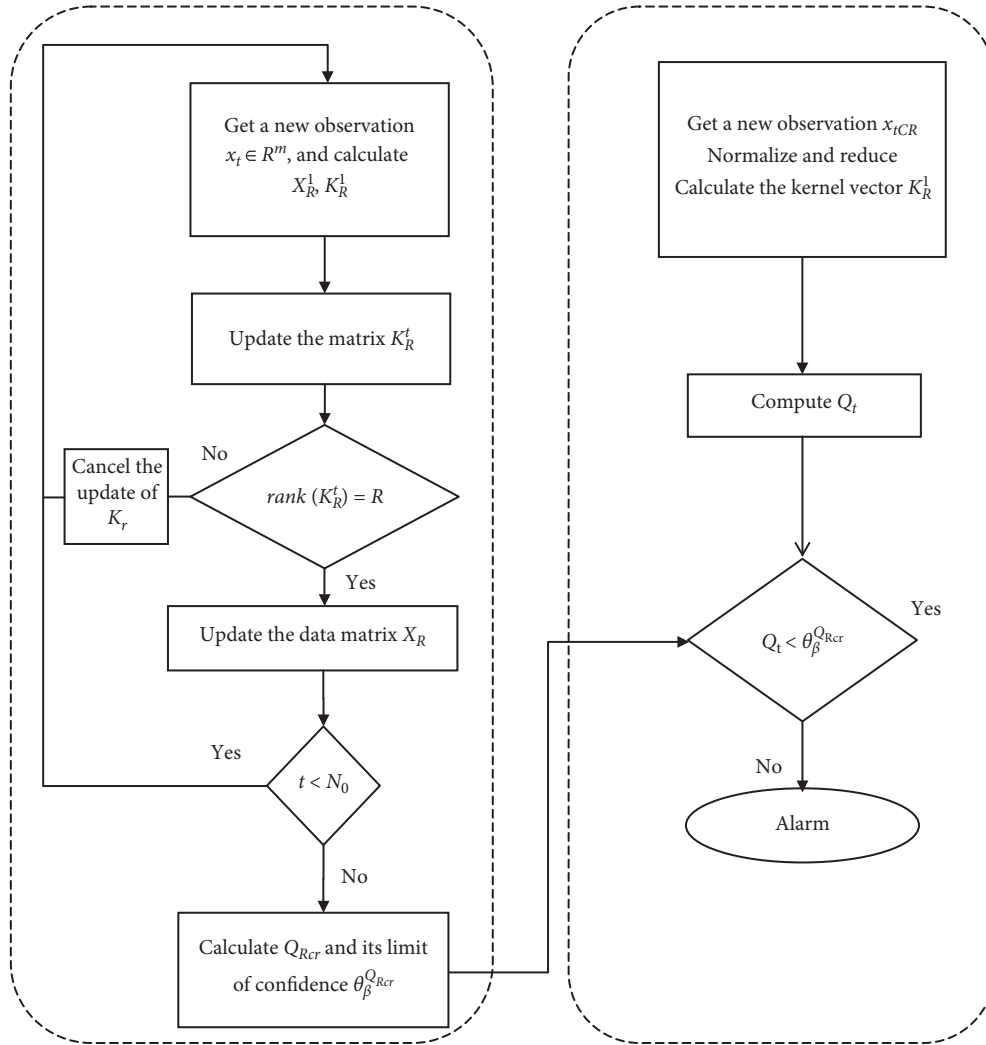


FIGURE 1: RRIKPCA_CR method flowchart.

The procedure of the isolation is shown in Figure 3, and summarize in three steps: first, we must determine the index for every partial RR-KPCA models at each time t . Then, we evaluate the index to their proper confidence limits and form the experimental sign of the fault.

$$SE_j = \begin{cases} 0, & \text{if } Q_j(t) \leq \text{Seuil}_j(t), \\ 1, & \text{if } Q_j(t) > \text{Seuil}_j(t). \end{cases} \quad (16)$$

The final step is to balance the experimental signature of the defect with the columns of the incidence matrix to arrive at a location decision.

The advantage of this approach is that it makes the fault localization easy; the principle of this approach is to build a set of residues, so that each residue is sensitive to certain

faults and not to others. Thus, for each fault, a theoretical signature is obtained which makes it possible to easily identify the defective variable.

The major disadvantage of this approach is that there is no systematic method for choosing partial interval reduced PCAs. In addition, we often face with the problem of the insensitivity of the residues to certain faults, which results in most cases of false locations.

4. Simulation Results

The air quality monitoring network AIRLOR (Figure 4), operating in Lorraine, France, consists in measuring and controlling the percentage of appearance of some reactive gases such as the ozone concentration O_3 and the nitrogen

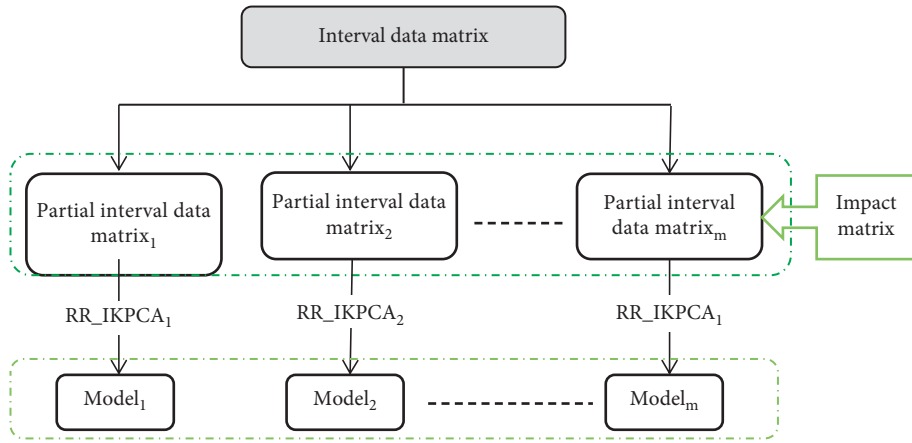


FIGURE 2: Procedure for structuring indices using the partial RRIKPCA method.

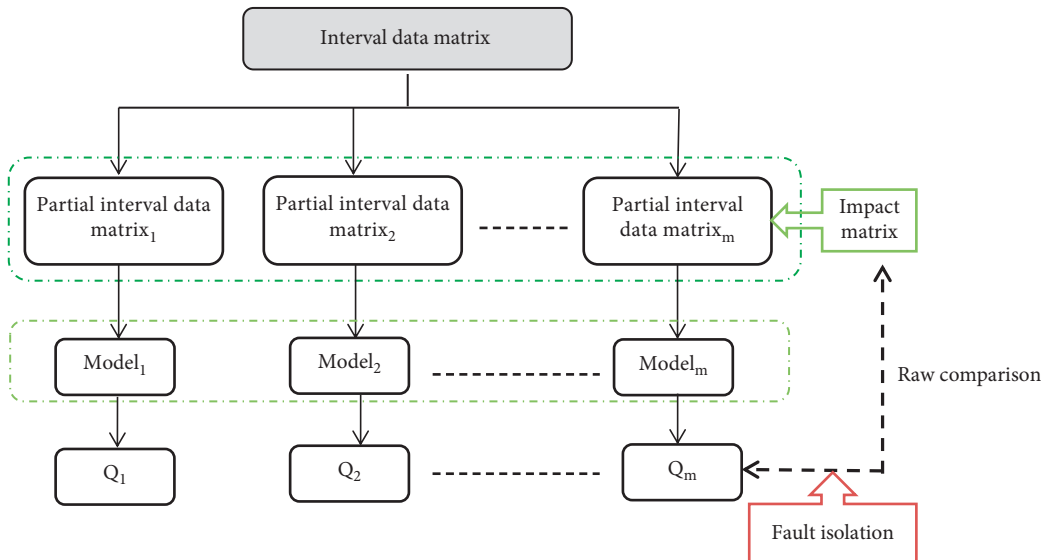


FIGURE 3: The localization procedure by the partial RR-KPCA method.



FIGURE 4: Air quality monitoring station. A variation δx (7% of the variation range of each variable) has been added to the data generated to simulate the presence of measurement uncertainties.

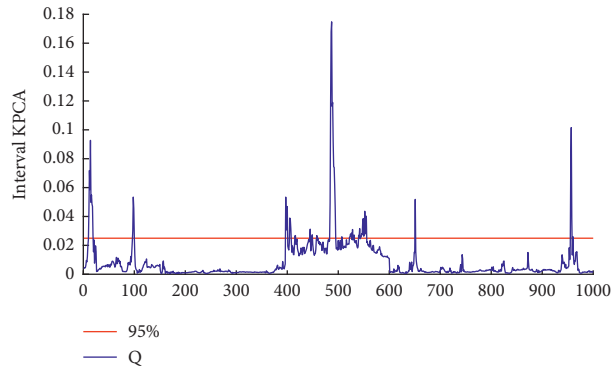


FIGURE 5: Q using the IKPCA_{CR} method with fault 18.

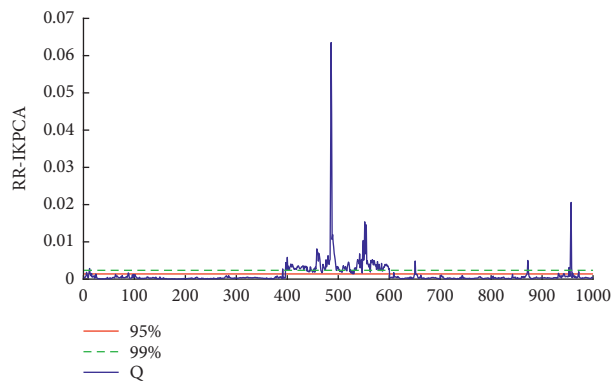


FIGURE 6: Q using the RR_IKPCA_{CR} method with fault 18.

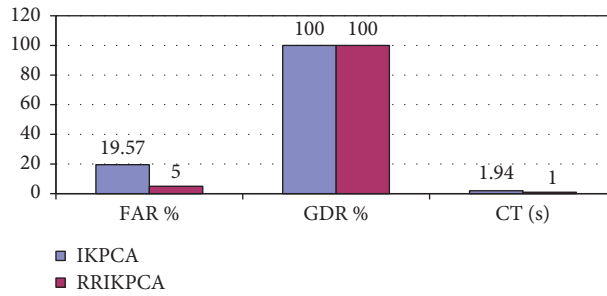


FIGURE 7: Performances and comparison between methods in different cases (AIRLOR). Locating faults using the partial RRIKPCA method.

TABLE 1: Theoretical signature matrix.

	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}	d_{11}	d_{12}	d_{13}	d_{14}	d_{15}	d_{16}	d_{17}	d_{18}
r_1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
r_2	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
r_3	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
r_4	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
r_5	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1
r_6	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1
r_7	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
r_8	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1
r_9	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1
r_{10}	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1
r_{11}	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1
r_{12}	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1

TABLE 1: Continued.

	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}	d_{11}	d_{12}	d_{13}	d_{14}	d_{15}	d_{16}	d_{17}	d_{18}
r_{13}	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1
r_{14}	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1
r_{15}	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1
r_{16}	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1
r_{17}	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1
r_{18}	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0

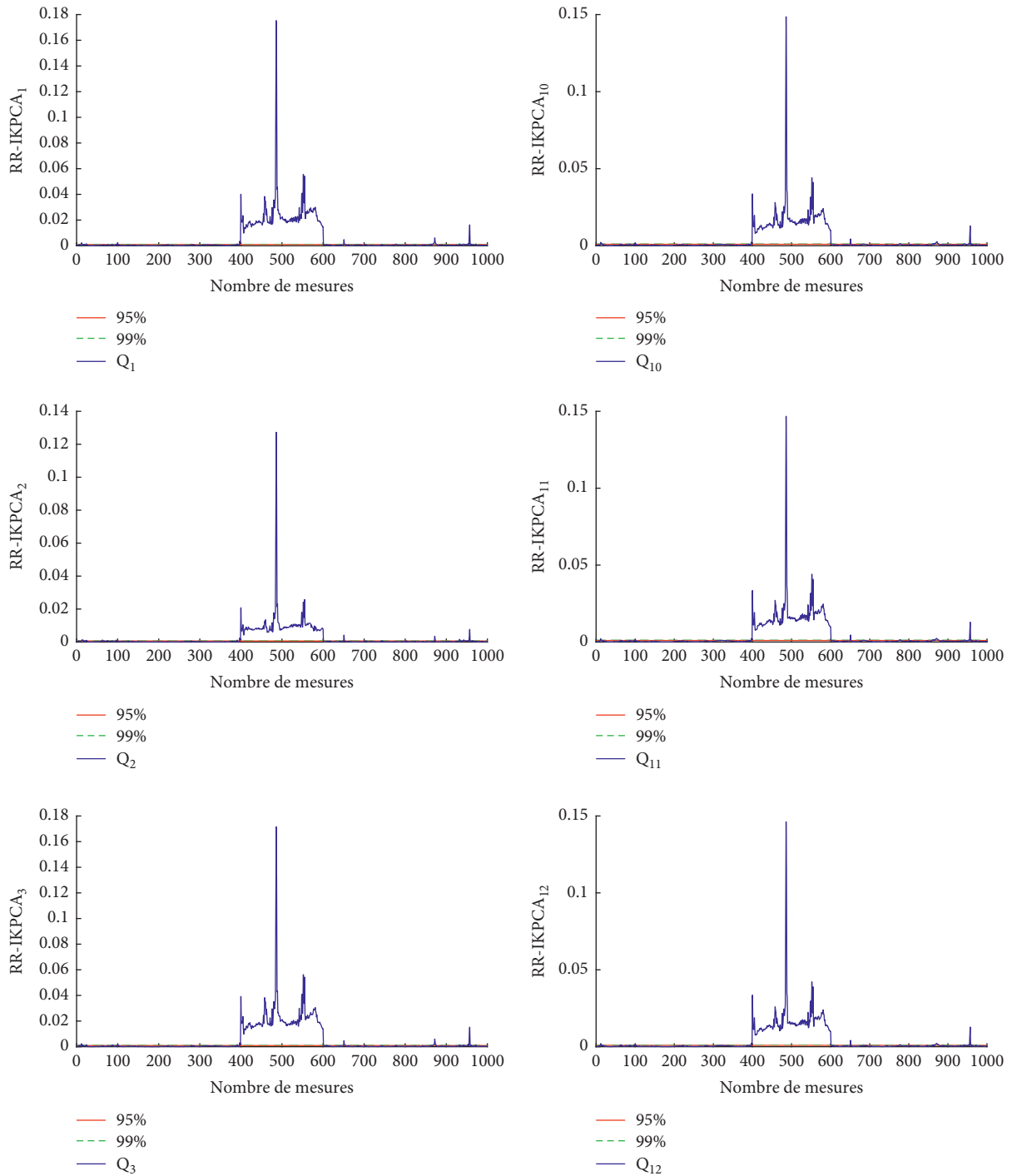


FIGURE 8: Continued.

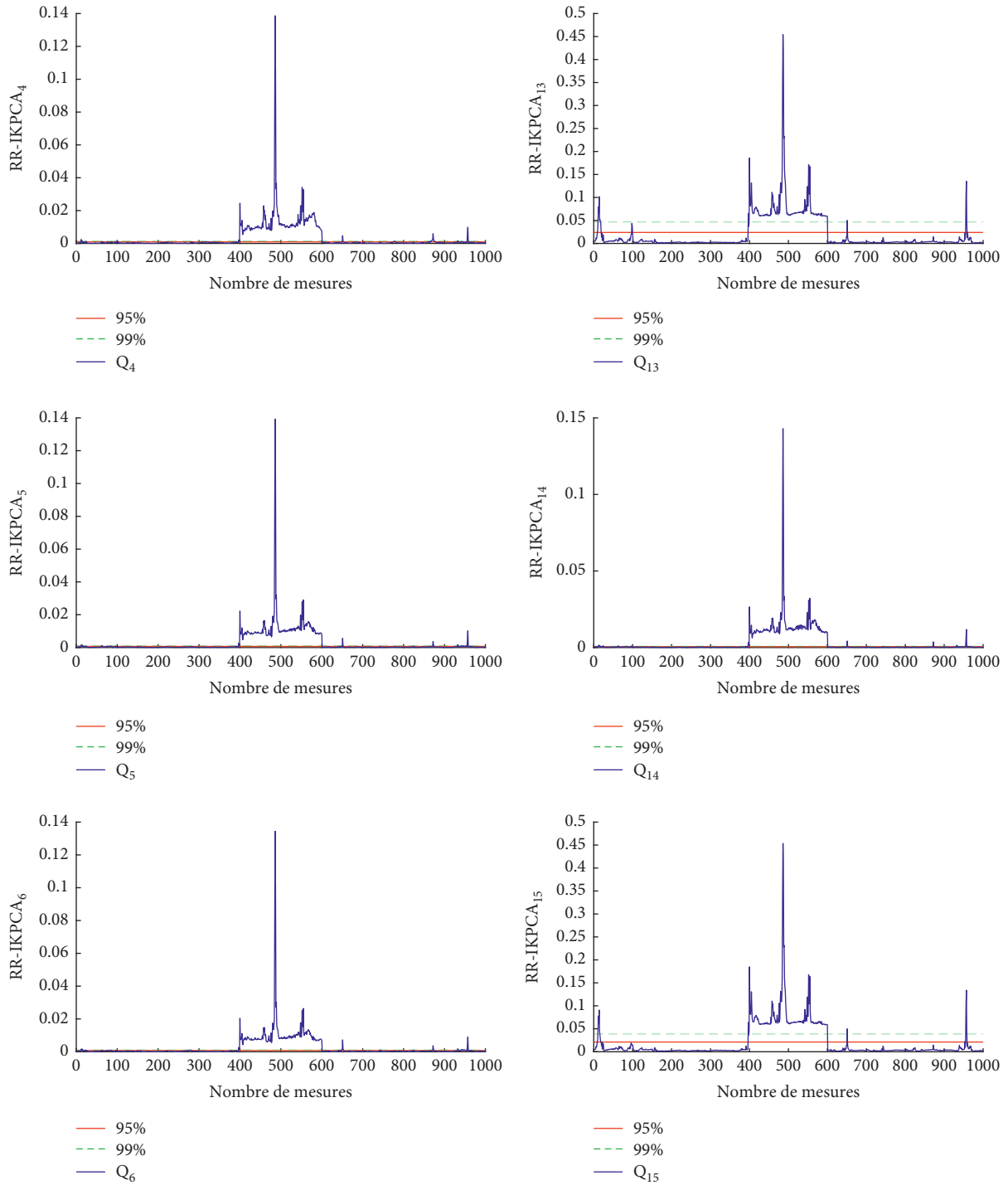


FIGURE 8: Continued.

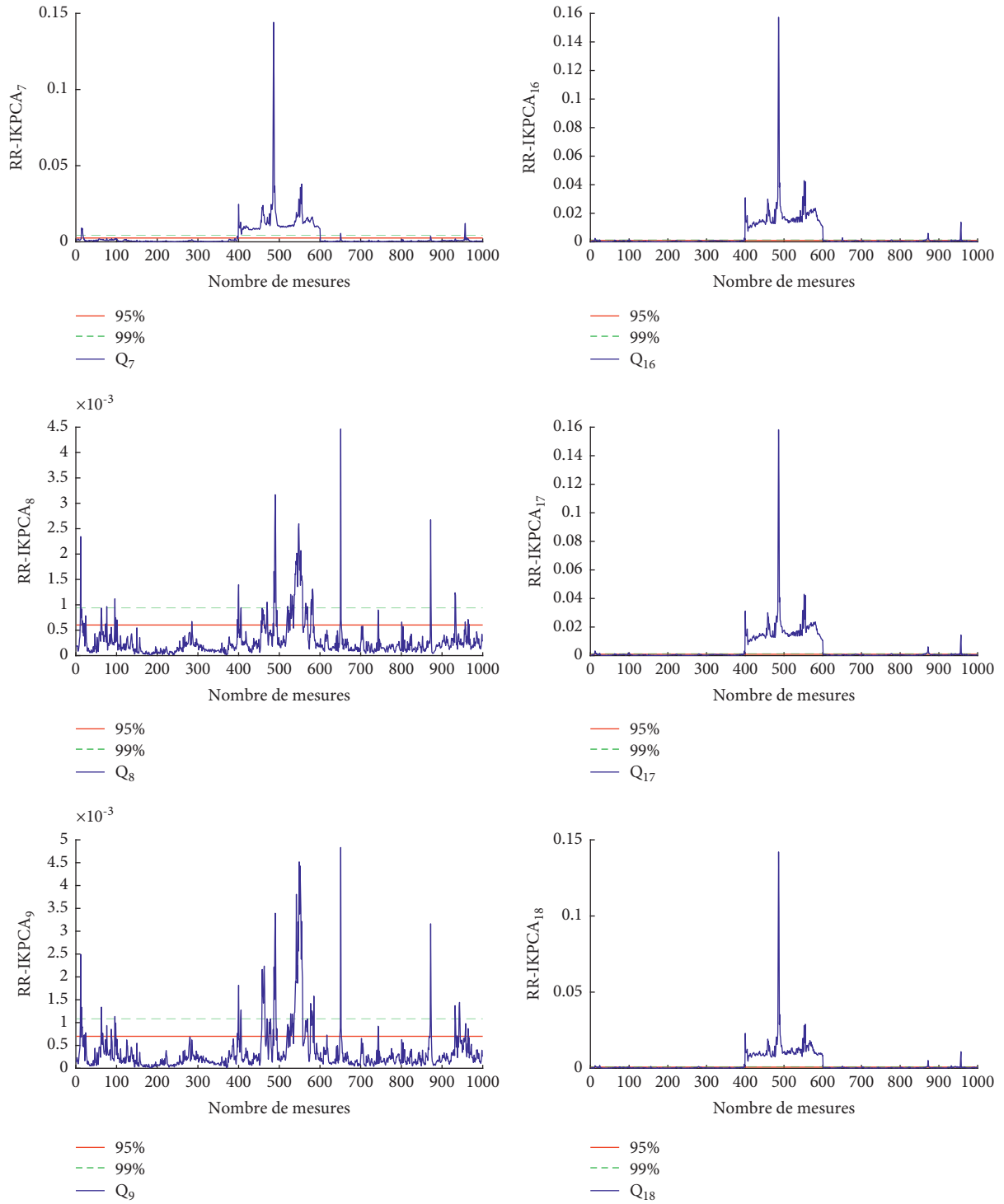


FIGURE 8: Time evolution of Q index corresponding to the 18 partial RRIKPCA models in the case of fault 9.

oxides NO and NO₂ [20]. In this study, our work focuses on the analysis of data from six stations located on different sites in order to detect and locate faults.

$$x(k) = \frac{v_1(k)v_2(k)v_3(k)}{O_3(k)NO(k)NO_2(k)station1} \cdots \frac{v_{10}(k)v_{11}(k)v_{12}(k)}{O_3(k)NO(k)NO_2(k)station4} \cdots \frac{v_{16}(k)v_{17}(k)v_{18}(k)}{O_3(k)NO(k)NO_2(k)station5} . \quad (17)$$

For these systems, three faults have been experienced, fault 8, 9, and 18, adding, respectively, 40% of the standard variation of (NO₂) from station 4 and 50% of the standard variation of (O₃) from station 6 from observations 400–600.

To have good fault localization, we need a good fault detection approach. Comparing Figures 5 and 6, it is very remarkable that the best performance of RRIKPCA_CR is in accordance to IKPCA_CR.

The best one named RRIKPCA_CR is chosen in the following for the localization step.

A comparative examination between the used method and the conventional method is illustrated to estimate the efficiency of the developed monitoring techniques.

The fault detection performances considered for the comparative study are the false alarm rate (FAR), the good detection rate (GDR), and the cost time (CT).

Figure 7 shows the monitoring success of IKPCA and RRIKPCA. Based in this figure, it is clear that the methods based on reduced rank kernel principal component analysis for interval-valued data process provide a good result, especially based on the mean value of FAR and GDR and cost time values compared to other fault detection methods. Therefore, we can observe that the developed method based on center and range approach with reduced rank ensures a good result for the fault detection performances. So, we will choose to use this efficient method in the fault localization part.

To illustrate these localization approaches by the partial RRIKPCA method, we applied the localization of faults on the AIRLOR system. The matrix incidence developed for this application is given in Table 1.

Thus, the procedure for structuring indices by partial RRIKPCA, a set of 18 RRIKPCA partials models, has been generated. Each partial RRIKPCA is insensitive to two variables.

An additive fault (d_o) is introduced between the times 400 and 600 with an amplitude equal to 25% of the variation range of this variable.

Figure 8 shows the time evolution of the 18 indices Q ; from this figure, it is shown that the signature is structured as follows: (1 1 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 1).

We see that this signature experimental is identical to the ninth column of the theoretical signature matrix (Table 1). Therefore, the variable nine is the offending variable. Thus, we can conclude that the fault localization procedure with the partial RRIKPCA method is validated with sucked for the AIRLOR process.

5. Conclusion

Majors fault localization treats only certain systems; to deal with this limitation, we have anticipated in this study a novel fault localization approach applied to the nonlinear uncertain process that is capable to treat data learning characterized by nonlinear uncertain large datasets. The proposed method is called partial reduced rank IKPCA; this method combines the benefits of downsizing of the kernel matrix in the characteristic space assured by RRIKPCA and the principle of partial localization. The obtained results exposed that the application of the partial RRIKPCA method guarantees the good detection and localization of the defects. The proposed method is incapable to locate faults correctly in the case of fast uncertain dynamic systems because the RRIKPCA model is static; to surmount this limitation and as future plant, we propose to extend these methods in the case of dynamic uncertain systems.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] G. E. P. Box, "Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification," *The Annals of Mathematical Statistics*, vol. 25, no. 2, pp. 290–302, 1954.
- [2] T. Ait-Izem, M.-F. Harket, M. Djeghaba, and F. Kratz, "Sensor fault detection based on principal component analysis for interval-valued data," *Quality Engineering*, vol. 30, no. 4, pp. 635–647, 2018.
- [3] A. Chouakria, *Extension des Methodes d'analyse Factorielle des Donnes de Type Intervalle*, Ph.D. dissertation, Universit Paris-Dauphine, Paris, France, 1998.
- [4] J. Le-Rademacher and L. Billard, "Symbolic covariance principal component analysis and visualization for interval-valued data," *Journal of Computational & Graphical Statistics*, vol. 21, no. 2, pp. 413–432, 2012.
- [5] M. F. Harkat, M. Mansouri, M. Nounoua, and H. Nounou, "Fault detection of uncertain nonlinear process using interval-valued data-driven approach," *Chemical Engineering Science*, vol. 205, pp. 36–45, 2019.
- [6] I. Hamrouni, H. Lahdhiri, K. B. Abdellafou, and O. Taouali, "Fault detection of uncertain nonlinear process using reduced interval kernel principal component analysis (RIKPCA),"

- International Journal of Advanced Manufacturing Technology*, vol. 106, pp. 4567–4576, 2020.
- [7] H. Lahdhiri and O. Taouali, “Interval valued data driven approach for sensor fault detection of nonlinear uncertain process,” *Measurement*, vol. 171, Article ID 108776, 2020.
- [8] T. Kourti and J. F. Mac Gregor, “Process analysis, monitoring and diagnosis, using multivariate projection methods,” *Chemometrics and Intelligent Laboratory Systems*, vol. 28, pp. 3–21, 1995.
- [9] S. Qin, “Statistical process monitoring: basics and beyond,” vol. 17, no. 3, pp. 480–502, 2003.
- [10] C. F. Alcalá and S. J. Qin, “Reconstruction based contribution for process monitoring with kernel principal component analysis,” in *Proceedings of the 2010 American Control Conference Industrial & Engineering Chemistry Research*, pp. 7849–7857, Baltimore, MD, USA, July 2010.
- [11] J. Gertler and T. McAvoy, “Principal component analysis and parity relations – a strong duality,” in *Proceedings of the IFAC Conference SAFEPROCESS*, pp. 837–842, Hull, UK, May 1997.
- [12] Y. Huang, J. Gertler, and T. McAvoy, “Fault isolation by partial PCA and partial NLPCA,” in *Proceedings of the IFAC’99, 14th Triennial World congress*, pp. 545–550, Beijing, China, July 1999.
- [13] M. F. Harakat, “Détection et Localisation de Défauts par Analyse en Composantes Principales,” de l’Institut National Polytechnique de Lorraine, France, Europe, Thèse de doctorat, 2003.
- [14] C. Chakour, A. Benyounes, and M. Boudiaf, “Diagnosis of uncertain nonlinear systems using interval kernel principal components analysis: application to a weather station,” *ISA Transactions*, vol. 83, pp. 126–141, 2018.
- [15] B. Schölkopf, A. Smola, and K. R. Müller, “Nonlinear component analysis as a kernel eigenvalue problem,” *Revue/Neural Computation*, vol. 10, pp. 1299–1319, 1998.
- [16] S. Valle, W. Li, and S. Qin, “Selection of the number of principal components: the variance of the reconstruction error criterion with a comparison to other methods,” *Industrial & Engineering Chemistry Research*, vol. 38, pp. 4389–4401, 1999.
- [17] M.-F. Harkat, G. Mourot, and J. Ragot, “An improved pca scheme for sensor fdi: application to an air quality monitoring network,” *Journal of Process Control*, vol. 16, no. 6, pp. 625–634, 2006.
- [18] H. Lahdhiri, I. Elaissi, O. Taouali, and M. F. Harakat, H. Messaoud, “Nonlinear process monitoring based on new reduced Rank-KPCA method,” *Stochastic Environmental Research and Risk Assessment*, vol. 32, pp. 1833–1848, 2017.
- [19] H. Lahdhiri, M. Said, K. B. Abdellafou, O. Taouali, and M. F. Harakat, “Supervised process monitoring and fault diagnosis based on machine learning methods,” *International Journal of Advanced Manufacturing Technology*, vol. 102, pp. 2321–2337, 2019.
- [20] M. F. Harkat, M. Mansouri, M. Nounoua, and H. Nounou, “Enhanced data validation strategy of air quality monitoring network,” vol. 160, pp. 183–194.