Uncertain Dynamic Process Monitoring Using Moving Window PCA for Interval-Valued Data

M-Faouzi HARKAT * Tarek AIT-IZEM **,***
Frédéric KRATZ *** Majdi Mansouri ****
Mohamed Nounou **** Hazem Nounou *

* *Chemical Engineering Program, Texas A& M University at Qatar, Doha, Qatar*
** *Automatics and signals laboratory Annaba (LASA), Department of electronics, Badji Mokhtar University, P.O.Box 12, 23000, Algeria*
*** *Laboratoire PRISME, INSA Centre Val-de-Loire, 88 boulevard Lahitolle 18020 Bourges Cedex, France*
**** *Electrical and Computer Engineering Program, Texas A&M University at Qatar, Doha, Qatar*

Abstract: In this paper, we present a new process monitoring approach for uncertain, or highly noisy systems, which is based on the well known Moving Window Principal Component Analysis (MWPCA) extended to the interval case. We propose to use The Midpoints-Radii PCA (MRPCA) for modelling, which independently exploits two PCAs on the center and radius matrices of the system's sensor interval-valued data. Furthermore, by changing the size and the shift of the window, Both center and radius model parameters are updated on-line; thus deriving a new Moving Window Midpoints-Radii PCA (MWMRPCA) approach. Based on the updated MWMRPCA, an interval SPE statistic and its control limit are calculated and updated through time, and are used for monitoring the state of the process. The performances of the proposed approach is illustrated by an application to the detection of faults on the Tennesse Eastman Process (TEP).

Keywords: Dynamic Systems, Moving Window PCA, Interval Data, Midpoints-Radii PCA, Fault detection, SPE statistic.

1. INTRODUCTION

Multivariate statistical approaches based on principal component analysis (PCA) have been widely applied for fault diagnosis to improve process quality and productivity. The key idea of PCA is to extract linear structure from high-dimensional data by finding new principal axes. In other words, PCA provides a statistical model of the system while reducing the dimensionality of the used dataset. For process monitoring, PCA uses several statistics to detect changes in the system behaviour. An abnormal situation will cause the corresponding statistic to exceed the control limits (Qin, 2003), (Harkat et al., 2006).

A major limitation of PCA-based monitoring is that the PCA model, once built from the data, is time-invariant, while most real industrial processes are time-varying. The time-varying characteristics of industrial processes include changes in the mean, changes in the variance; and changes in the correlation structure among variables, and even changes in the number of significant principal components (PCs). When a static PCA model is used to monitor time-varying processes, false alarms often occurs, which significantly compromise the reliability of the monitoring system. Thus, a dynamic variant of PCA, such as the moving window PCA (MWPCA) (Wang et al., 2005) or recursive PCA (RPCA) (Li et al., 2000) has to be used to handle these type of processes.

PCA based diagnosis strategies have been developed for the analysis of single-valued variables. But in real life, there are many situations in which the use of these single-valued variables may cause severe loss of information, and thus compromise the reliability of the approach. In this case, more robust strategy can be achieved by describing the process measurements by interval-valued data. Therefore, An interval-valued PCA model is to be used to handle the new nature of data.

In the last two decades or so, many researchers investigated the possibility to extend PCA to interval-valued data. (Cazes et al., 1997), and (Chouakria, 1998), proposed the first interval-valued PCA approaches, known as the centers PCA (CPCA) and the vertices PCA (VPCA) methods. The centers method relies on the interval centers to computes the principal components (PCs), while the vertices method computes the PCs using the vertices of the observed hyper-rectangles. Another approach is the midpoints-radii PCA (MRPCA) introduced by (Palumbo and Lauro, 2003), which treats midpoints and interval ranges as two separate variables, and performs based on two separate models of these variables. (D'Urso and Giordani, 2004), introduced another approach using least squares for MRPCA, and (Gioia and Lauro, 2006), put forward an analytical interval PCA based on an interval-valued covariance matrix. (Le-Rademacher and Billard, 2012), employed symbolic covariance to extend the classical PCA, and (Wang et al., 2012) proposed the complete-information principal component analysis (CIPCA), which is a variant of CPCA with a new covariance calculation method. Based on the different interval PCA methods found in the literature, (Ait-Izem et al., 2015) and (Ait-Izem et al., 2017), proposed to apply these approaches for monitoring of uncertain systems by modelling the sensors uncertainties in the form of interval-valued data.

In this paper, we propose a new moving window interval PCA approach for on-line monitoring of time-varying systems subject to uncertainties. The interval PCA model used is the MRPCA model, which is the most adapted to the task, as it decomposes the interval-valued data matrix into center and radius matrix in order to handle each separately. Based on the moving window principle, the MRPCA model is updated on-line with shifting of time window, where different parameters are calculated, i.e. normalization parameters, MRPCA model parameters, and detection statistics with their limits. The new Moving Window Midpoints-Radii PCA (MWMRPCA) is then more suited for monitoring of time-varying systems as it follows the changes in such processes. The rest of the paper is organized as follows: Static PCA, in its conventional and interval case, is presented in section 2 along with their application to fault detection. Section 3 is dedicated to the MWPCA and to the introduction of the new MWMRPCA algorithm. Then, an application on the Tennessee Eastman Process (TEP) benchmark data is given in section 5, and finally conclusions are presented in the last section.

2. STATIC PCA

In this section, we present time-invariant PCA models for single-valued, and interval-valued data. The interval MRPCA model is chosen, among other interval PCA methods, for its high suitability with the on-line application.

2.1 Classical PCA

In classical PCA, the raw data matrix $X \in \Re^{n\times m}$, after standardization, i.e. reducing the data to zero mean and unit variance, is decomposed as follows:

$$X = TP^T + E \tag{1}$$

$T \in n \times \ell$, $P \in m \times \ell$ and $E \in n \times m$ represent, respectively, principal components matrix, loading matrix and residuals matrix. Where ℓ is the number of principal components ($\ell < m$). The Euclidean norm of the residual matrix E must be minimized for a given number of components. This criterion is satisfied when the columns of P are the eigenvectors corresponding to the ℓ largest eigenvalues of the covariance matrix Σ of X. Thus, PCA can be viewed as a linear mapping from $\Re^m$ to a lower dimensional space $\Re^\ell$. The mapping has the form:

$$T = P^T X \tag{2}$$

Where P is the eigenvectors matrix, or the coefficients of the linear transformation. The projection can be reversed back to $\Re^m$ with:

$$\hat{X} = TP^T \tag{3}$$

Where $\hat{X}$ is the estimate of the initial data matrix X. Accordingly, the residual matrix E is the difference between X and $\hat{X}$:

$$E = X - \hat{X} \tag{4}$$

2.2 Midpoints-Radii PCA for interval-valued data

In the case of interval PCA, the data matrix $X \in \mathbb{R}^{n\times m}$ is represented by interval-valued variables, denoted $[X_j]$,$j = 1, \ldots, m$. According to the uncertainties or approximation error of the process sensors, and for each time sample k, the measurement takes the form $[\underline{x}_j(k), \overline{x}_j(k)]$, where the bounds of the interval are given by:

$$\underline{x}_j(k) = x_j^c(k) - x_j^r(k) \tag{5}$$

$$\overline{x}_j(k) = x_j^c(k) + x_j^r(k) \tag{6}$$

the global interval data matrix is then constructed as :

$$[X] = [[X_1][X_2], \ldots, [X_j]] \tag{7}$$

In a real application, $x_j^c(k)$ is the center of the interval, which is given by the the estimate single-valued measurement provided by the sensor, and $x_j^r(k)$, also denoted $\delta_x(k)$ is the uncertainty of measurement or the range of the interval, which usually presents the sensor precision given by the manufacturer.

The midpoints-radii PCA (MRPCA) on interval-valued data, introduced in (Palumbo and Lauro, 2003), is based on a midpoints (X^c) and radius (X^r) representation, rather than the standard interval representation $[\underline{X}, \overline{X}]$ of the interval-valued data matrix $[X]$. MRPCA performs two PCA's on these two matrices, and thus depends on two models, centers and radius.

$$X^c\Sigma^{-1}P^c = \lambda^c P^c \tag{8}$$

$$X^r\Sigma^{-1}P^r = \lambda^r P^r \tag{9}$$

Where λ^c, P^c and λ^r, P^r are, respectively, the eigenvalues and eigenvectors of the two eigen-decompositions of midpoints and range matrices, and Σ is the covariance matrix given by:

$$\Sigma = \left(X'^c X^c\right) + \left(X'^r X^r\right) + \left(\left|X'^c X^r\right| + \left|X'^r X^c\right|\right) \tag{10}$$

Thus, based on classical PCA models of centers and radius matrices, principal components for centers T^c and for radius T^r are given by:

$$\begin{cases} T^c = X^c P^c_{\ell_c} \\ T^r = X^r P^r_{\ell_r} \end{cases} \tag{11}$$

and interval estimates of centers $\hat{X}^c$ and radius $\hat{X}^r$ are given by:

$$\begin{cases} \hat{X}^c = X^c P^c_{\ell_c} P^{cT}_{\ell_c} \\ \hat{X}^r = X^r P^r_{\ell_r} P^{rT}_{\ell_r} \end{cases} \tag{12}$$

Where ℓ_c and ℓ_c are the number of principal components for the midpoints model and the radius model respectively.

The interval form for the components $[T]$ and the estimates $[\hat{X}]$ can be obtained from the midpoins-radii representation through a rotation of coordinates. This rotation is performed based on matrix $A = QP^T$ (Palumbo and Lauro, 2003), given the following singular value decomposition:

$$X^{cT} X^r = P \Lambda^{cr} Q^T \tag{13}$$

Hence, Interval components can be calculated as:

$$\begin{cases} \underline{T} = T^c - AT^r \\ \overline{T} = T^c + AT^r \end{cases} \tag{14}$$

and interval estimates given by:

$$\begin{cases} \underline{\hat{X}} = \hat{X}^c - A\hat{X}^r \\ \overline{\hat{X}} = \hat{X}^c + A\hat{X}^r \end{cases} \tag{15}$$

The interval residual matrix $[E]$ is, as in classical PCA, given by the difference between $[X]$ and $[\hat{X}]$:

$$[E] = [X] - [\hat{X}] \tag{16}$$

Thus, according to interval arithmetic, the bounds of the interval residual matrix are given by:

$$\begin{cases} \underline{E} = \underline{X} - \overline{\hat{X}} \\ \overline{E} = \overline{X} - \underline{\hat{X}} \end{cases} \tag{17}$$

2.3 Fault detection using MRPCA for interval-valued data

In general, the PCA based fault detection scheme uses the squared prediction error (SPE) statistic, which is defined as:

$$SPE(k) = \|\mathbf{x}(k) - \hat{\mathbf{x}}(k)\|^2 = \|\mathbf{e}(k)\|^2 \tag{18}$$

where $\mathbf{e}(k)$ is the residual vector given by the difference between the measurement vector and its estimate from the PCA model. A faulty condition is declared if the SPE index exceeds its control limit δ^2_α determined statistically (Jackson and Mudholkar, 1979),(Nomikos and MacGregor, 1995) for a significance level α.

For the interval-valued PCA case, several indices are proposed as extentions of classical SPE to interval valued data (Ait-Izem et al., 2017),(Benaicha et al., 2013). A first method of calculation treats separately both bounds of the residuals in computing interval SPE, which we denote $\overline{\underline{SPE}}$, and is given by:

$$\begin{cases} \underline{SPE}(k) = \|\underline{\mathbf{e}}(k)\|^2 = \underline{\mathbf{e}}(k)^T \underline{\mathbf{e}}(k) \\ \overline{SPE}(k) = \|\overline{\mathbf{e}}(k)\|^2 = \overline{\mathbf{e}}(k)^T \overline{\mathbf{e}}(k) \end{cases} \tag{19}$$

Given that $\underline{\mathbf{e}}(k)$ is the interval valued residual computed from the MRPCA model. In the presence of faults, decisions are made when both bounds in Eq. 19 exceed their detection threshold.

Another index introduced in (Ait-Izem et al., 2017) is denoted $ISPE$ index, and is given by:

$$ISPE(k) = \|[\mathbf{e}(k)]\|^2 = \sum_{j=1}^{m} \|[e_j(k)]\|^2 \tag{20}$$

given that $[\mathbf{x}(k)] = [[\underline{x}_1(k), \overline{x}_1(k)], \ldots, [\underline{x}_m(k), \overline{x}_m(k)]]$ is the interval measurement vector; and that

$$\|[e_j(k)]\|^2 = \frac{1}{3}\left(\underline{e}^2_j(k) + \underline{e}_j(k)\overline{e}_j(k) + \overline{e}^2_j(k)\right) \tag{21}$$

The control limits for both statistics, i.e. $\underline{SPE}_{lim}$ and $\overline{SPE}_{lim}$ for $\overline{\underline{SPE}}$ index, and $ISPE_{lim}$ for $ISPE$ index, can be computed from their approximate distribution as detailed in (Nomikos and MacGregor, 1995), based on Box's approximation for quadratic forms (Box, 1954), the example of the $ISPE$ limit is given by the following:

$$ISPE_{lim} = g\chi^2_{h,\alpha} \tag{22}$$

Where g is a weighting parameter included to account for the magnitude of the $ISPE$ and h accounts for the degrees of freedom with a significance level of $1-\alpha$, typically selected to be 95% to 99%. The parameters g and h can be estimate as:

$$g = \frac{b}{2a}, \; h = \frac{2a^2}{b} \tag{23}$$

Given that a is the estimate mean of $ISPE$, and b is its estimated variance.

In comparison to classical PCA, a PCA model for interval-valued data offers more robustness toward uncertainties of sensor measurements, due to the interval nature of data. The interval-valued model considers that every variation inside the interval is a normal process variation. In other words, the radius δX of data is considered as a safe zone, were small magnitude offsets, i.e. uncertainties $\zeta < \delta X$ are not detected and are considered as normal process variation, while high magnitude offsets $f > \delta X$ are considered as faults.

3. DYNAMIC PCA

In dynamic PCA algorithms, new measurements are used to update the PCA model on-line. The updating scheme should include: mean, covariance, principal components and the confidence limit for the monitoring statistic. Several algorithms for dynamic PCA can be found in the literature. In this paper, we investigate the possibility to extend the Moving Window PCA (MWPCA) algorithm to the interval-valued data case.

3.1 Moving window PCA

In MWPCA, a data window of fixed length is moved in real time to update the PCA model once a new normal sample is available. Thus, for a window of size w, the treated data matrix at time sample k is $X_k = [x(k-w+1), x(k-w+2), \ldots, x(k)]$, and at time $k+1$ becomes $X_{k+1} = [x(k-w+2), x(k-w+3), \ldots, x(k+1)]$. MWPCA algorithm then updates various parameters, including : normalization parameters, PCA model parameters, and monitoring statistics control thresholds. The overall strategy for time-varying monitoring using MWPCA is given as follows:

(1) Collect training data from the process under NOC. Normalize the training data using its mean and standard deviation. Then compute covariance matrix and carry out an eigen-decomposition to obtain the PCA

model for the process. Choose the number of components. Finally, Determine the control limits for the used monitoring statistic.
(2) Obtain a new testing sample and normalize it using scaling parameters from the moving data window.
(3) Evaluate the monitoring statistic for the normalized testing sample using the PCA model obtained in Step 1, and check with the corresponding control limit from step 1. If not exceeded, the measurement is considered normal.
(4) If the measurement is normal, update the moving window by shifting it by one time sample (add new measurement and delete the oldest measurement). Repeat from Step 2. Else, the measurement is faulty, stop update after three consecutive faulty measurements.

3.2 Moving Window MRPCA for interval-valued data

Dynamic variants of PCA, such as MWPCA, were developed in order to take into account the dynamics or variations of the system to be monitored, but which remain less robust due to the effect of measurements uncertainties. To overcome this problem, we propose in this section a dynamic approach to PCA for interval-valued data for diagnosis purpose. The idea is to exploit the MRPCA model for interval-valued data combined with the well known Mowing window PCA algorithm.

First, let us recall the normalization procedure for interval-valued data, each interval-valued measurement is normalized as follows:

$$\frac{[x_j(k)] - \left(\bar{X^c}_j\right)}{\sqrt{VAR([X_j])}} = \left[\frac{\underline{x}_j(k) - \left(\bar{X^c}_j\right)}{\sqrt{VAR([X_j])}}, \frac{\overline{x}_j(k) - \left(\bar{X^c}_j\right)}{\sqrt{VAR([X_j])}}\right] \quad (24)$$

Where $\bar{X}_j^c$ is the centers mean of the j-th interval-valued variable, and $VAR([X_j])$ is the interval variance defined as follows (Wang et al., 2012):

$$VAR([X_j]) = \sum_{k=1}^{n} \frac{1}{3}\left(\underline{x}^2{}_j(k) + \underline{x}_j(k)\overline{x}_j(k) + \overline{x}^2{}_j(k)\right) \quad (25)$$

Since the MRPCA model is constituted of two sub-models of center and radius, we propose to apply a moving window scheme to both models in parallel. The proposed MWMRPCA algorithm is detailed in the following algorithm:

Offline mode

(1) Construct the center (X^c) and radius (X^r) matrices from the initial interval-valued data matrix $[X]$
(2) Obtain the initial values for normalization parameters, i.e. centers mean m_0^c and interval variance D_0
(3) Compute the initial models parameters (for centers and raidus): covariance Σ, eigenvector matrices P_0^c and P_0^r, the eigenvalues matrices Λ_0^c and Λ_0^r, the rotation matrix A_0, and the number of principal components for the two sub-models ℓ_0^c and ℓ_0^r; based on the training data block;
(4) Compute interval residuals, and $ISPE_0$ index then obtain the control threshold $ISPE_{lim,0}$.
(5) Determine the initial size for moving window w.

Online mode
At sample time k, use the previous values of time sample $(k-1)$ of m_{k-1}^c, D_{k-1}, Λ_{k-1}^c, Λ_{k-1}^r, P_{k-1}^c, P_{k-1}^r, A_0, $SPE_{lim,k-1}$, ℓ_0^c and ℓ_0^r;

(1) Compute interval residuals and $ISPE_k$ index, for a new sample $[x(k)]$ after standardizing the interval using m_{k-1}^c, D_{k-1};
(2) If $ISPE_k > ISPE_{lim,k-1}$ go to step 3, else go to step 5;
(3) Check if the new sample is an outlier. For instance, if $ISPE_{lim,t-s} > ISPE_{lim,k-t-1}$, (t=1,2) (i.e. if three consecutive out-of control signals have been generated), the new sample is not an outlier. Otherwise, it is an outlying sample;
(4) If it is an outlier, go to step 5. Otherwise, consider the current condition to be abnormal. Typically, the process condition cannot be determined by examining one sample. In this case, the model is retained without updating and the current sample is stored. Then, if the process condition is proven to be normal subsequently, the model can be updated in a blockwise manner;
(5) recompute residuals and $ISPE_k$;
(6) Shift the window and calculate new normalization and model parameters for the new window (update)
(7) update the number of principal components to retain for the two sub-models, and $SPE_{lim,k}$.

4. APPLICATION

The performance of the proposed monitoring scheme, using a MWMRPCA model, is illustrated through an application on a classical example: the Tennessee Eastman Process (TEP), whose scheme is shown in Figure 1. This process is developed by Eastman Chemical Company to provide a simulation of a real industrial process for the testing of control and/or monitoring methods. There are five major units in TEP simulation (as shown in fig 1 a reactor, a separator, a stripper, a condenser, and a compressor. The process has 12 manipulated variables, 22 continuous process measurements, and 19 composition measurements sampled less frequently. which all have Gaussian noise. Corresponding to different production rates, there are six modes of process operation. The

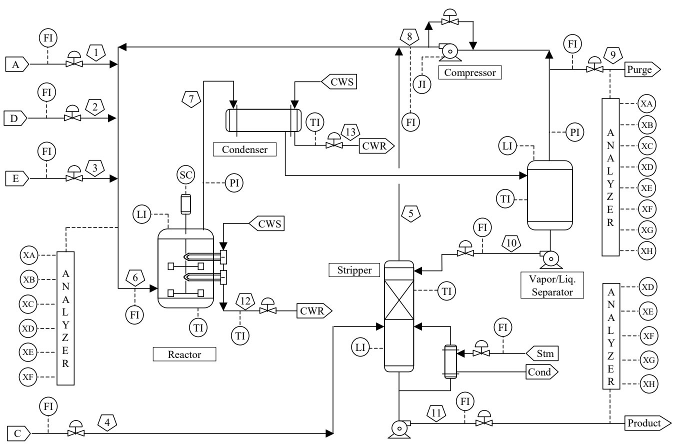

Fig. 1. Tennessee Eastman Process

TEP process was run for 1 hours, and we collected 1000 samples from 22 variables. Considering an uncertainty δ_{x_i} of measurements, of the order of 5% of the measurements for each variable, we construct the new interval-valued data matrix of the process, with δ_{x_i} being the radius of the data intervals. The first 150 samples were used to construct the initial MWMRPCA model, while the rest of the samples will be processed one by one, by sliding the window each time, in testing phase (online monitoring). The model parameters are updated for each normal sample, where the the number of components is selected according to the Cumulative Percentage Variance criterion (PCV), so that the explained variance represents approximately 95% of the total variance, given as follows:

$$CPV(\ell) = 100 \left(\frac{\sum_{j=1}^{\ell} \lambda_j}{\sum_{j=1}^{m} \lambda_j} \right) \tag{26}$$

In order to compare the performance of the new MWMR-PCA monitoring strategy, a classical dynamic PCA model using the sliding window PCA (MWPCA) (Wang et al., 2005) approach, as well as a static classical PCA model were developed based on the same set of data. Subsequently, two types of outliers, in two different variables, and in two different moments were simulated, that is:

(1) An uncertainty ζ_{x_9} in reactor temperature, explained by the variable x_9, simulated as bias of amplitude smaller than the radius ($\zeta_{x_9} < \delta_{x_9}$), from time sample 200 until moment 300.
(2) A fault f_{x_3} in the E feed rate, which is given by variable x_3, simulated as a relatively large bias compared to the uncertainty/radius ($f_{x_3} > \delta_{x_3}$, from time sample 700 to the end.

An interesting property of the PCA for interval-valued data, applied to the diagnosis, is that the defined radius of the data acts as a safe-zone. In other words, the PCA model is insensitive to outliers with amplitude smaller than the radius (uncertainty), because considered as a normal variation of the process. As for the outliers with greater amplitude than the radius, they are normally detected as faults.

Indeed, by inspecting the figures 2 and 3, which represent the detection indices $\overline{SPE}$ and $ISPE$ for the MWMRPCA model, we can clearly notice the absence of the bias ζ_{x_9} representing the uncertainty (present between time 200 and 300), as well as the presence/detection of the f_{x_3} faults simulated from moment 700. This is due to the nature of the interval model, which considers the bias ζ_{x_9} as a normal variation of the process. The MWMRPCA model thus manages to follow the dynamics of the system, to detect faults, while providing more robustness with respect to measurement uncertainties. Note that the $\overline{SPE}$ index depends on two thresholds, one for each of its bounds, while the $ISPE$ depends only on a single threshold. The latter has better performance with regard to the number of false alarms, and in distinguishing the fault.

According to the SPE criterion calculated in the case of the MWPCA model, represented in the figure 4, we can notice that the uncertainty ζ_{x_9} present in the variable x_9 has mislead the dynamic MWPCA model. More precisely, the simulated bias ζ_{x_5} is considered as a fault by the MWPCA model thus resulting in a stopping of model update. The impact is a huge rate of false detections from the moment of the presence of uncertainties (from time sample 200). A conventional dynamic PCA monitoring strategy, MWPCA in this case, can be thus insufficient to handle highly noisy systems. Figure 5, represents the SPE index combined with static PCA, which demonstrates its total incapacity in detecting faults in a variant system.

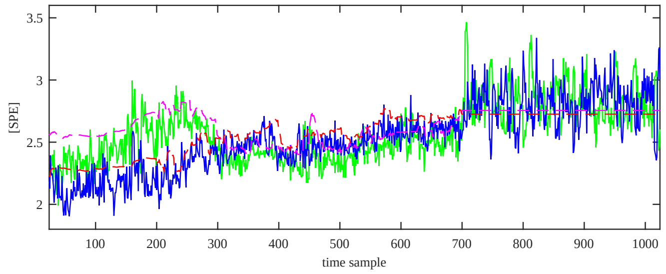

Fig. 2. Fault detection using $\overline{SPE}$ indicator with MWM-RPCA model for TEP data

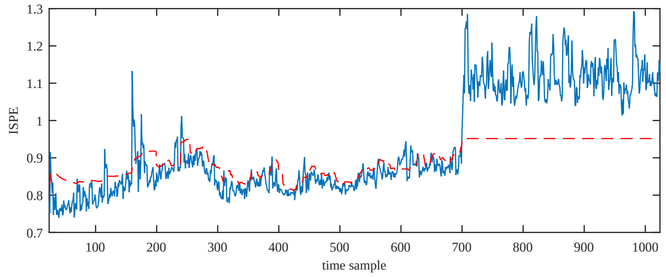

Fig. 3. Fault detection using $ISPE$ indicator with MWM-RPCA model for TEP data

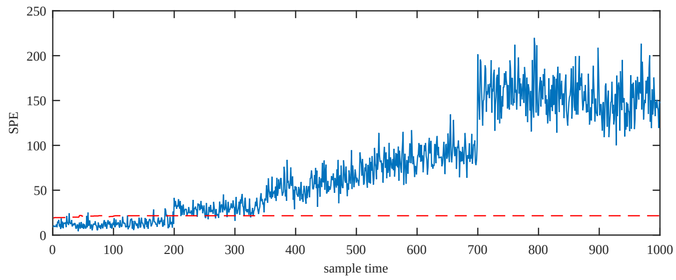

Fig. 4. Fault detection using SPE indicator with MWPCA model for TEP data

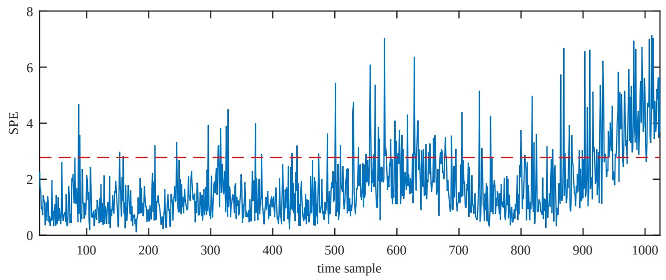

Fig. 5. Fault detection using SPE indicator with Classical PCA model for TEP data

5. CONCLUSION

In this work, a new algorithm based on moving window PCA is presented, for modelling and monitoring of dynamic processes. The model used is an interval variant of PCA, called MRPCA, which treats separately the matrices of the centers and the radii extracted from the interval-valued data matrix. Applied on-line, the proposed algorithm verifies for each sample the presence or absence of faults, using interval statistical indices ($\underline{SPE}$ and $ISPE$). After that, the update of the model and the normalization parameters is performed. The presented approach improves not only the robustness toward measurement uncertainties, but also allows to have a confidence zone defined by the radius of the interval, making it possible to consider any additional amount of information within this zone as a part of the normal operation of the system. Applied to the Tenessee Eastman Process, the proposed algorithm demonstrates good performance over the static PCA model, and the MWPCA model with sliding window.

REFERENCES

Ait-Izem, T., Harkat, M.F., Djeghaba, M., and Kratz, F. (2015). Fault detection and isolation using interval principal component analysis methods. IFAC-PapersOnLine, 48(21), 1402 – 1407. doi: https://doi.org/10.1016/j.ifacol.2015.09.721. 9th IFAC Symposium on Fault Detection, Supervision andSafety for Technical Processes SAFEPROCESS 2015.

Ait-Izem, T., Harkat, M.F., Djeghaba, M., and Kratz, F. (2017). Sensor fault detection based on principal component analysis for interval-valued data. Quality Engineering. doi:10.1080/08982112.2017.1391288.

Benaicha, A., Mourot, G., Benothman, K., and Ragot, J. (2013). Fault detection and isolation with interval principal component analysis. 162–167. Proceedings Engineering and Technology PET. International Conference on Control, Engineering and Information Technology.

Box, G. (1954). 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. 25.

Cazes, P., Chouakria, A., Diday, E., and Schektman, Y. (1997). Extension de l'analyse en composantes principales à des données de type intervalle. Revue de Statistique Appliquée, 45(3), 5–24.

Chouakria, A. (1998). Extension des Methodes d'analyse Factorielle à des Données de Type Intervalle. Ph.D. thesis, Université Paris-Dauphine.

D'Urso, P. and Giordani, P. (2004). A least squares approach to principal component analysis for interval valued data. Chemometr Intell Lab Syst, 70(2), 179,192.

Gioia, F. and Lauro, C. (2006). Principal component analysis on interval data. Computational Statistics, 21, 343–363.

Harkat, M.F., Mourot, G., and Ragot, J. (2006). An improved pca scheme for sensor fdi: Application to an air quality monitoring network. Journal of Process Control, 16(6), 625 – 634. doi: https://doi.org/10.1016/j.jprocont.2005.09.007.

Jackson, J.E. and Mudholkar, G.S. (1979). Control procedures for residuals associated with principal component analysis. Technometrics, 21(3), 341–349. doi: 10.1080/00401706.1979.10489779.

Le-Rademacher, J. and Billard, L. (2012). Symbolic covariance principal component analysis and visualization for interval-valued data. Journal of Computational and Graphical Statistics, 21(2), 413–432. doi:10.1080/10618600.2012.679895.

Li, W., Yue, H., Valle-Cervantes, S., and Qin, S. (2000). Recursive pca for adaptive process monitoring. Journal of Process Control, 10(5), 471 – 486. doi: https://doi.org/10.1016/S0959-1524(00)00022-6.

Nomikos, P. and MacGregor, J.F. (1995). Multivariate spc charts for monitoring batch process. Technometrics, 37, 414.

Palumbo, F. and Lauro, N. (2003). A pca for interval-valued data based on midpoints and radii. in New Developments in Psychometrics, eds. H. Yanai, A. Okada, K. Shigemasu, Y. Kano, and J. Meulman, Tokyo, 641–648.

Qin, S.J. (2003). Statistical process monitoring: basics and beyond. Journal of Chemometrics, 17(8-9), 480–502. doi:10.1002/cem.800.

Wang, H., Guan, R., and Wu, J. (2012). Cipca: Complete-information-based principal component analysis for interval-valued data. Neurocomput., 86, 158–169. doi: 10.1016/j.neucom.2012.01.018.

Wang, X., Kruger, U., and Irwin, G.W. (2005). Process monitoring approach using fast moving window pca. Industrial & Engineering Chemistry Research, 44(15), 5691–5702. doi:10.1021/ie048873f.