Enhanced Monitoring of Environmental Processes

M. Mansouri, Marie-France Destain, H. Nounou, and M. Nounou

***Abstract*—The process monitoring systems are often utilized in environmental process operations. Many practical applications used for scheduling, planning or operator training are often complex for direct usage in process monitoring. In this paper, it is proposed to use the generalized likelihood ratio (GLR) based principal components analysis (PCA) for process monitoring and fault detection of environmental processes. The objective is to combine the GLR test with PCA model in order to improve the fault detection performance. GLR-based PCA is a multivariate statistical technique used in multivariate statistical process monitoring and fault detection. PCA reduces the dimensionality of the original data by projecting it onto a space with significantly fewer dimensions. It obtains the principal events of variability in a process. If some of these events change, it can be due to a fault in the process. The data are collected from the crop model in order to calculate the PCA model and the thresholds; Hotelling statistic, T^2, Q statistic and GLR test statistic are used in order to detect the faults. It is demonstrated that the performance of faults detection can be improved by combining GLR test and PCA.**

***Index Terms*—Environmental processes, fault detection, Generalized likelihood ratio test, Principal component analysis.**

I. Introduction

Due to consistent product quality demand and higher requirements in safety, the process monitoring performance has become a key factor in improving productivity and safety.

Process systems are using large amount of data from many variables that are monitored and recorded continuously every day. For these reasons, the problem of fault detection that responses effectively to faults that mislead the process and harm the system reliability represents a key process in such operation of these systems.

Several multivariate statistical techniques for fault detection, analysis of process and diagnosis have been developed and used in practice. These techniques are useful since operation safety and the high quality products are some of the core objectives in the industry applications. Faults detection has been performed manually using data visualization tools [1], but these tools are time consuming for real-time detection in streaming data. Recently, researchers have proposed automated statistical and machine learning methods, such as: nearest neighbor [2], clustering [3], minimum volume ellipsoid [4], convex pealing [5], neural network classifier [6], decision tree [7] and support vector machine classifier [8].

In this paper, generalized likelihood ratio (GLR) - based PCA is proposed to detect the faults in environmental processes representing the crop model. PCA is used to create the model and find linear combinations of variables that describe major trends in data set and the GLR test. Both are utilized to improve faults detection. GLR test has been proposed in order to establish an adaptive system, which reaches three important problems; estimation, fault detection and magnitude compensation of jumps. GLR test is proposed for fault detection of different applications: geophysical signal segmentation [9], signals and dynamic systems [10], incident fault detection on freeways [11], missiles trajectory [12]. Therefore, in the current work it is proposed to exploit the advantages of the GLR test for improved fault detection when a process model is not available.

The rest of the paper is organized as follows. In Section II, a brief introduction to PCA is presented, followed by descriptions of the two main detection indices, T^2 and Q, which are generally used with PCA for fault detection. Then, the GLR test which is used in composite hypothesis testing is described in Section III. Next, the PCA-based GLR used for faults detection integrating PCA modeling and GLR statistical testing, is presented in Section IV. Then, in Section V, the performance of the GLR-based PCA test is illustrated using crop model data. Finally, some concluding remarks are presented in Section VI.

II. Principal Component Analysis (PCA)

Let $X_i \in R^m$ denotes a sample vector of m number of sensors. Also, assume there are *n* samples dedicated to each sensor, a data matrix $X \in R^{nxm}$ is with each row, displaying a sample. Meanwhile, X matrix is scaled to zero mean for covariance-based PCA and at the same time, to unit variance for correlation-based PCA [13]. The *X* matrix can be divided into two matrices: a score matrix S and a loading matrix W through singular value decomposition (SVD):

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

where $S = [s_1 s_2 ... s_m] \in R^{mxm}$ is a transformed variables matrix, $s_i \in R^n$, are the score vectors or principal components, and $W = [w_1 w_2 ... w_3] \in R^{mxm}$ is an orthogonal vectors matrix $w_i \in R^m$ which includes the eigenvectors associated with the covariance matrix of X, i.e., $\sum$, which

Manuscript received April 29, 2015; revised August 31, 2015. This work was supported in part by the Department of Biosystems Engineering, University of Liege.

M. Mansouri and H. Nounou are with the Electrical and Computer Engineering Program, Texas A&M University at Qatar, Qatar (e-mail: majdi.mansouri@qatar.tamu.edu, hazem.nounou@qatar.tamu.edu).

Marie-France Destain is with the Department of Biosystems Engineering, University of Liege, Belgium (e-mail: mfdestain@ulg.ac.be).

M. Nounou is with the Chemical Engineering Program, Texas A&M University at Qatar, Qatar (e-mail: mohamed.nounou@qatar.tamu.edu).

is given by

$$\sum = \frac{1}{n-1} X^T X = W \Lambda W^T \quad (2)$$

with $W\Lambda W^T = W^T \Lambda W = I_n$ and Λ =diag($\lambda_1, \lambda_2, \ldots, \lambda_m$) is a diagonal matrix containing the eigenvalues related to the m PCs, λ_m are simply the eigenvalues of the covariance matrix ($\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m$), and I_n is the identity matrix [14]. It must be noted at this point that the PCA model yields same number of principal components as the number of original variables (m). Nevertheless, for collinear process variables, a smaller number of principal components (l) are required so that most of the variations in the data are captures. Most of the times, a small subset of the principal components (which correspond to the maximum eigenvalues) might carry the most of the crucial information in a data set, which simplifies the analysis.

The effectiveness of the PCA model depends on the number of principal components (PCs) are to be used for PCA. Selecting an appropriate number of PCs introduces a good performance of PCA in terms of processes monitoring. Several methods for determining the number of PCs have been proposed such as; the Scree plot [15], the cumulative percent variance (CPV), the cross validation [16], and the profile likelihood [17]. In this study herein, the cumulative percent variance method is utilized to come up with the optimum number of retained principal components. The cumulative percent variance is computed as follows:

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

When the number of principal components *l* is determined, then, the data matrix X is shown as the following:

$$X = SW = [\hat{S}\ \tilde{S}][\hat{W}\ \tilde{W}]^T \quad (4)$$

where $\hat{S} \in R^{nxl}$ and $\tilde{S} \in R^{nx(m-l)}$ are matrices of l retained principal components and the ($m-l$) ignored principal components, respectively, and the matrices $\hat{W} \in R^{mxl}$ and $\tilde{W} \in R^{mx(m-l)}$ are matrices of l retained eigenvectors and the ($m-l$ ignored eigenvectors, respectively. Using Eq. (4), the following can be written:

$$X = \hat{S}\hat{W}^T + \tilde{S}\tilde{W}^T \quad (5)$$

The matrix $\hat{X}$ represents the modeled variation of X based on first l components.

A. Fault Detection Indices

When using PCA in detecting faults, a PCA model is built utilizing fault-free data. The model is used for fault detection through one of the detection indices (the Hotelling's T^2 and Q statistics), which are presented next.

1) Hotelling's T^2 statistic

The T^2 statistic is a way of measuring the variation captured in the principal components at various time samples, and it is known as ([18]):

$$T^2 = X^T \hat{W} \hat{\Lambda}^{-1} \hat{W}^T X \quad (6)$$

where $\hat{\Lambda}^{-1} = diag(\lambda_1, \lambda_2, ..., \lambda_l)$, is a diagonal matrix containing the eigenvalues related to the l retained PCs. For new real-time data, when the value of T^2 statistic exceeds the threshold, T_α^2 calculated as in ([18]), a fault is detected.

The threshold number used for the T^2 statistic is computed as [18]:

$$T_\alpha^2 = \frac{l(n-1)}{n-1} F_{l,n-l,\alpha}, \quad (7)$$

where α is the level of significance (α usually between 1% and 5%), n is the number of samples in data set, l is the number of retained PCs, and $F_{l,n-l,\alpha}$ is the Fisher F distribution with l and $n-l$ degrees of freedom. These thresholds are computed using faultless data. When the number of observations, n, is high, the T^2 statistic threshold is approximated with a χ^2 distribution with *l* degrees of freedom, i.e., $T_\alpha^2 = \chi_{l,\alpha}^2$.

2) Q statistic or squared prediction error (SPE)

It is possible to detect new events by computing the squared prediction error SPE or Q of the residuals for a new observation. Q statistic [19], [20], is computed as the sum of squares of the residuals. Also, the Q statistic is a measure of the amount of variation not captured by the PCA model, it is defined as [19]:

$$Q = \left\| \tilde{X} \right\|^2 = \left\| X - \hat{X} \right\|^2 = \left\| (I - \hat{W}\hat{W}^T) X \right\|^2. \quad (8)$$

The monitored system, meanwhile, is accepted to be in normal operation if:

$$Q \leq Q_\alpha \quad (9)$$

The threshold Q_α used for the Q statistic can be computed as [14],

$$Q_\alpha = \varphi_1 \left[\frac{h_0 c_\alpha \sqrt{2\varphi_2}}{\varphi_1} + 1 + \frac{\varphi_2 h_0 (h_0 - 1)}{\varphi_1^2} \right] \quad (10)$$

where, $\varphi_i = \sum_{j=l+1}^{m} \lambda_j^i, \{i = 1,2,3\}$, $h_0 = 1 - \frac{2\varphi_1\varphi_3}{\varphi_2^2}$ and c_α is the value of the normal distribution with α is the level of

significance at the instant of an unusual event, when there is a change in the covariance structure of the model, this change is going to be detected by a high value of Q. For new data, the Q statistic is computed and compared to the threshold Q_α [14]. This means a fault is detected when the confidence limit is violated. The threshold value is computed on the assumption that the measurements are independent of time and they are multivariate normally distributed. The Q fault detection index is highly sensitive to errors in modeling and the performance of it is dependent on the number of retained PCs, l, [21].

III. Generalized Likelihood Ratio Test (GLRT)

The faults detection step is done using the residuals computed using PCA. Using the information about the noise distribution of the residuals, a GLR test statistic is formed. To make the decision if a fault is present or not, the test statistic is compared to a threshold from the chi-square distribution.

A. Test Statistic

The GLR test is famous to be a uniformly most powerful test among all invariant tests (shown in Equation (10)). It is basically a hypothesis testing technique which has been utilized successfully in model-based faults detection [9]. Focusing on the following fault detection problem, $Y \in R^n$ is an observation vector formed by one of the two Gaussian distributions: $N(0,\sigma^2 I_n)$ or $N(\theta \neq 0,\sigma^2 I_n)$, where θ is the mean vector (which is the value of the fault) and $\sigma^2 \succ 0$ is the variance (assumed to be known in this problem). The hypothesis test can be shown as:

$$\begin{cases} H_0 = \{Y \sim N(0,\sigma^2 I_n)\} \text{ (null hypothesis);} \\ H_1 = \{Y \sim N(\theta,\sigma^2 I_n)\} \text{(alternative hypothesis).} \end{cases} \quad (11)$$

Here, the GLR method replaces the unknown parameter, θ, by its maximum likelihood estimate. This estimate is computed by maximizing the generalized likelihood ratio $T(Y)$ as shown below:

$$\begin{aligned} T(Y) &= 2\log \frac{\sup_{\theta \in R^n} f_\theta(Y)}{f_{\theta=0}(Y)} \\ &= 2\log \left\{ \frac{\sup_{\theta \in R^n} \exp\left(\frac{\|Y-\theta\|_2^2}{f_{\theta=0}(Y)} \right)}{\exp\left(\frac{\|Y\|_2^2}{2\sigma^2} \right)} \right\} \\ &= \frac{1}{2\sigma^2} \left\{ \min_{\theta \in R^n} \|Y-\theta\|_2^2 + \|Y\|_2^2 \right\} \\ &= \frac{1}{2\sigma^2} \left\{ \|Y\|_2^2 \right\} \end{aligned} \quad (12)$$

where $\hat{\theta} = \arg\min \|Y-\theta\|_2^2 = Y$ is the maximum likelihood estimate of θ, the probability density function of Y is $\frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left(-\frac{\|Y-\theta\|_2^2}{2\sigma^2} \right)$, $\|.\|_2$ represents the Euclidean norm. Because the GLR test utilized the ratio of distributions of the faulty and faultless data; for the case of non-Gaussian variables, non-Gaussian distributions are required to be utilized. It must be noted that, in the derivation mentioned above, maximizing the likelihood function is equivalent to maximizing its natural logarithm since the logarithmic function is a monotonic function. At this stage, the GLR test then decides between the hypotheses H_0 and H_1 as follows:

$$\begin{cases} H_0 & \text{if } T(Y) \prec t_\alpha \\ H_1 & \text{else.} \end{cases} \quad (13)$$

Since distribution of the decision function $T(Y)$ under H_0 allows to design a statistical test with a desired false alarm rate, α, where the threshold t_α is chosen to satisfy the following false alarm probability:

$$P_0(\Lambda(Y) \geq t_\alpha) = \alpha \quad (14)$$

where, $P_0(A)$ represent the probability of an event A when Y is distributed according to the null hypothesis H_0 and α is the desired probability of the false alarm. Since Y is normally distributed, the statistics T is distributed according to the χ^2 law with ($m-l$) degrees of freedom.

B. Statistic

To select an appropriate thresholds for the test statistics shown above, it is crucial to find their distributions. For that purpose, with the Gaussian noise within, the test statistics will be chi-square distributed variables [22]. The normalized residual $\bar{R}$ is distributed as

$$\bar{R} \sim N(\theta,\sigma^2 I_n) \quad (15)$$

where $\theta = 0$ under the null hypothesis (13). Then, the test statistic is distributed as the non-central chi-square distribution as shown below:

$$t_\alpha = \frac{1}{\sigma^2} \left\{ \|Y\|_2^2 \right\} \sim \chi_n^2, \quad (16)$$

and the test statistic is distributed through the central chi-square distribution χ_n^2 with degree of freedom n. The threshold is now chosen from the chi-square distribution therefore the fault-free hypothesis is erroneously rejected with

only a small probability.

IV. Fault Detection Using a GLR-Based PCA Test

In this section, a GLR test to detect faults is derived, and its explicit asymptotic statistics computed using PCA. The objective of the GLR-based PCA fault detection technique is to detect the additive fault, θ, with the maximum detection probability for a given false alarm. Here, the fault detection task can be considered as a hypothesis testing problem with consideration of two possible hypotheses: null hypothesis of no change H_0, where measurements vector X, is fault-free, and the change-point alternative hypothesis H_1, where X contains a fault, and thus X is no longer categorized by the fault-free PCA model (4). For new data, the method needs to pick between H_0 and H_1 for the most efficient detection performance. In the absence of a fault, the residual can be calculated as follows,

$$R = X - \hat{X} \tag{17}$$

while in the presence of an additive fault vector, θ, the residual is computed as,

$$R = X - \hat{X}[+\theta] \tag{18}$$

It is assumed that the residual in Equation (17) is Gaussian. Hence, the fault detection problem consists of detecting the presence of an additive bias vector, θ, in the residual vector, R. The residual vector can be considered as a hypothesis testing problem by focusing on two hypotheses: the null hypothesis H_0, where R is fault-free and the alternate hypothesis H_1, where R contains a fault. The formulation of the hypothesis testing problem can be written as,

$$\begin{cases} H_0 = \{R \sim N(0, \sigma^2 I_n)\} \text{ (null hypothesis);} \\ H_1 = \{R \sim N(\theta, \sigma^2 I_n)\} \text{(alternative hypothesis).} \end{cases} \tag{19}$$

The algorithm which studies the developed GLR-based PCA fault detection technique is presented in Algorithm 1. The GLR- based PCA is proposed to detect the faults in the residual vector obtained from the PCA model, through which the GLR test is used for each residual vector, R.

Algorithm 1: GLR-based PCA fault detection algorithm.

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

Output: GLR statistic T, GLR Threshold t_α

Data preprocessing step:

Standardize: computes data's mean and standard deviation, and standardize it;

PCA running step:

Compute the covariance matrix, Σ;

Calculate the eigenvalues and eigenvectors of Σ and sort the eigenvalues in decreasing order;

Compute the optimal number of principal components to be used using the *CPV* method;

Compute the sum of approximate and residual matrices;

Testing step:

Standardize the new data;

Generate a residual vector, R, using PCA;

Compute the GLR statistic T for the new data;

Compute the GLR statistic threshold t_α: if $T \geq t_\alpha$, then declare a fault.

V. Simulation Results Analysis

Next, the crop model that are used to generate data is described.

A. Crop Model

The original data were issued from experiments carried out on a silty soil in Belgium, with a wheat crop (Triticum aestivum L., cultivar Julius), during the crop seasons 2008-2009 and 2009-2010. The measurements were the results of 4 repetitions by date, each one of them being performed on a small block (2m times 6m) randomly spread over the field to ensure the measurements independence. A wireless monitoring system (eKo pro series system, Crossbow) completed by a micro-meteorological station was used for measuring continuously soil and climate characteristics. Especially, the measurements of soil water content were performed at 20 and 50 cm depth. The plant characteristics (LAI and biomass) were measured at regular intervals (2 weeks) along the crop seasons, since the middle of February (around Julian day 410) till harvest. Each LAI and biomass measurements were the results of four replicates by date of sampling. The LAI is defined as one half of the total leaf area per unit ground surface area [23]. Each LAI sample was collected as a 50 cm linear sample (for a total of 2 meters considering four replicates). The stripped leaves were sticked on a paper sheet and digitalized [24]. The images were segmented using the Meyer and Neto (2008) indices (ExG-ExR) to compute the total green leaf area and the LAI was finally computed as the ratio between this value and the soil reference surface (2 meters times 0.146 meter of inter-row spacing). Each biomass measurement was performed on three adjacent rows of 50 cm (for a total of 6 meters considering the four replicates). The cut samples were dried at laboratory and the total mass was finally weighed. During the season 2008-2009, yields were quite high and close to the optimum of the cultivar. This is mainly explained by the good weather conditions and a sufficient nitrogen nutrition level. The season 2009-2010 was known to induce deep water stresses, and was thus characterized by yield losses. The model for which the methods are tested is Mini-STICS model [25].

The model equations are presented in [26], and the model parameters presented [27]. The dynamic equations indicates the way each state variable changes from one day to another as a function of the current values of the state variables, and of the parameters value. Encoding these equations over time allows for eliminating the intermediate values of the state variables and relate the state variables at any time to the

explanatory variables on each day. The model structure can be derived from the basic conservation laws, namely material and energy balances.

B. Data Generation

Indeed, the findings might depend on the details of the model, on the way/quality the data are generated/measured with and on the specific data which was used. To be independent of these consideration, we are generate dynamic data from the crop model. The model is first used to simulate the responses of the 6 state variables: the leaf-area index LAI; LAI, the biomass growth; MASEC, the grain yield MAFRUIT, the volumetric water content of the soil layer

1; HUR1, the volumetric water content of the soil layer 2; HUR2, the volumetric water content of the soil layer 3; HUR3 as functions of time of the first recorded climatic variable of the crop season 2008-2009. These simulated states are assumed to be noise free. They are then contaminated with zero mean Gaussian errors, i.e., a measurement noise v). The data set consists of 8 random variables, which are generated using the crop model presented in [26]. The generated data were arranged as a matrix X having 297 samples and 6 crop model measurements. The responses of the 6 state variables LAI, MASEC, MAFRUIT, HUR1, HUR2 and HUR3, are shown in Fig. 1.

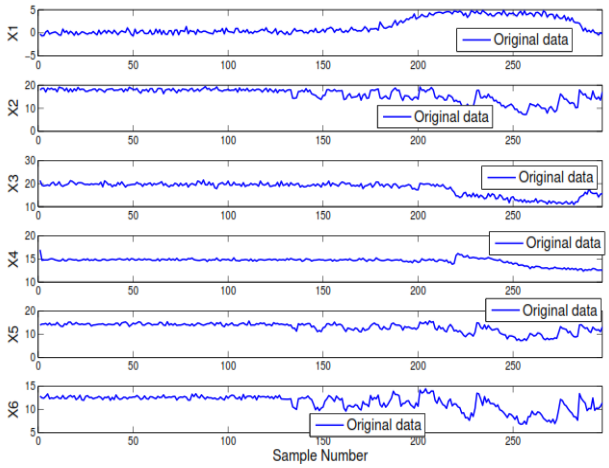

Fig. 1. Original data.

C. Training of PCA Model

As described in Algorithm 1, the PCA-based GLR fault detection method requires constructing a PCA model from fault-free data. Therefore, the fault-free crop model training data described earlier were used to construct a PCA reference model to be used in fault detection. The fault-free crop model data were arranged as a matrix Xtr having 150 rows (samples) and 6 columns (crop model measurements). These data are first scaled (to have zero mean and unit variance), and then are used to construct the PCA model. The responses of the training fault-free data, are shown in Fig. 2. The training fault-free data matrix is used to construct a PCA model. In PCA, most of the crucial variations in the data set are typically captured in the main principal components corresponding to the maximum eigenvalues as shown in Fig. 3. In this study herein, the cumulative percent variance (CPV) method is utilized to find out the optimum number of retained principal components. Utilizing a CPV threshold value of 90%, only the first five principal components of the total variations in the data as displayed in Fig. 3) will be retained. A plot of the decision function of the GLR test T (shown in Fig. 4) confirms that the process operates under normal conditions, where no faults are present.

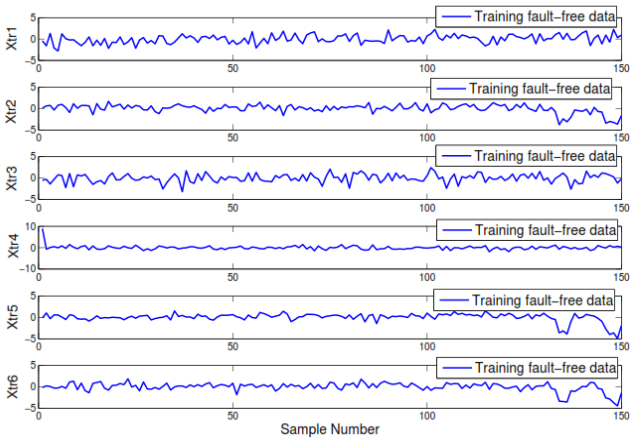

Fig. 2. Training fault-free data.

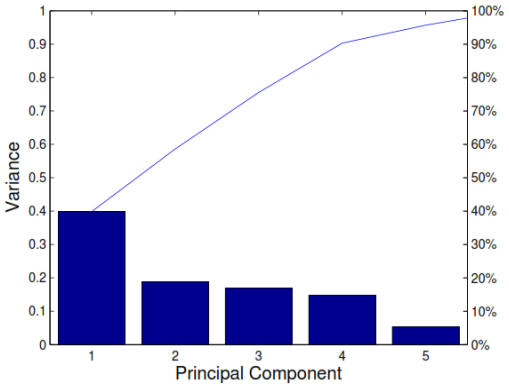

Fig. 3. Variance captured by each principal component.

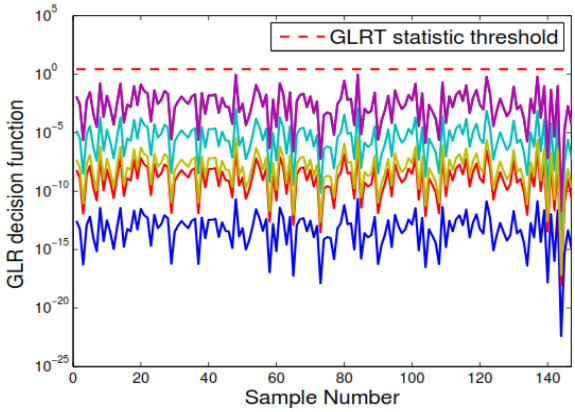

Fig. 4. The time evolution of GLR decision function on a semi-logarithmic scale for the fault-free data.

D. Fault Detection in Crop Model

The PCA model formed utilizing the fault-free data is deployed in this section to detect possible faults with unseen testing data. The data set from tests (which is simulated using the crop model) includes 150 data samples that are free of the training data. Single fault (i.e., in one variable) or multiple faults (i.e., in over two variables) are taken into consideration.

To use the abilities of the various fault detection techniques,

an additive fault (single fault) was introduced in X_1. It consists of a bias of amplitude equal to 10% of the total variation in X_1, (see Fig. 5).

The performance of the various faults detection methods will be compared. Two different cases of faults are going to be taken into consideration. For the first case, it is assumed that the sensor measuring one state variable is damaged by a simple fault. In the second case, it is assumed that the faults occur at the same time at the sensors measuring the state variables.

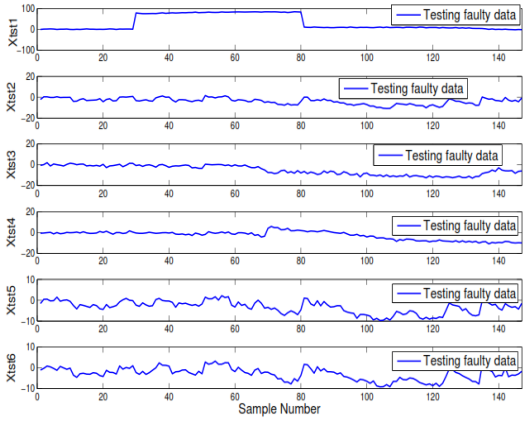

Fig. 5. Testing faulty data X_{test}.

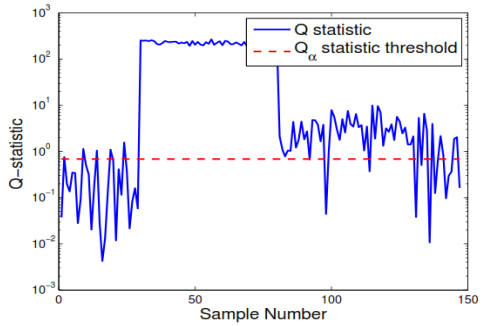

Fig. 6. Fault detection using Q statistic in the presence of simple fault.

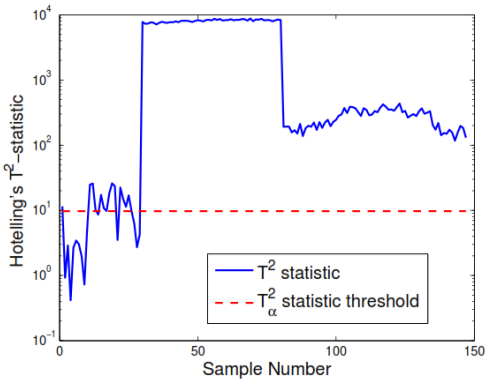

Fig. 7. Fault detection using Hotelling's T^2-statistic in the presence of simple fault.

The conventional PCA based monitoring technique is initially run using the training fault-free data. Based on the first five PCs, T and Q statistics for the conventional PCA algorithm and the GLR-based PCA test algorithm are used for fault detection. Fig. 5 shows the testing faulty data (simple fault in X_1). The results of Q statistic are shown in Fig. 6, where the dotted line represents the detection threshold Q_α, which is found to be 0.6848. Fig. 7 presents the results of the T^2 statistic, where the dotted line represents the detection threshold T^2_α, which is found to be 9.7.

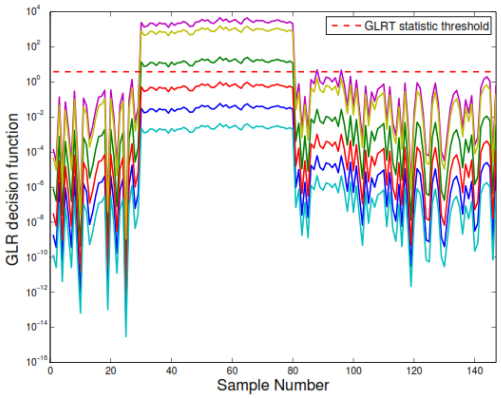

Fig. 8. Fault detection using GLR statistic in the presence of simple fault.

Fig. 6 shows that the Q statistic at the time interval [30 . . . 80] is always above the threshold Q_α, which means that the data fit the PCA model well (since it could capture most of the variations in the data), and verifies that the data belongs to the normal operating region. The process monitoring under fault is presented in Fig. 6 and Fig. 7. Both statistics, T^2 and Q, arise their thresholds when the fault occurs. In this case, the Q statistic detects this fault better than T^2 statistic as these figures show. When the GLR test is applied using the same fault-free data, the GLR threshold value is found to be t_α = 3.841 for a false alarm probability of α = 10%. A plot of the GLR statistic test T (shown in Fig. 8) confirms that the process operates very well compared to the fault detection using T^2 and Q statistics under single fault. We can show that the results of the T^2, Q statistics and the GLR-based PCA test, which are shown in Fig. 6, Fig. 7 and Fig. 8, show the ability of three techniques to detect this additive fault, but with some false alarms for the T^2 and Q tests at the time intervals [1 . . . 30] and [80 . . . 150].

VI. Conclusion

In this paper generalized likelihood ratio (GLR) based principal components analysis (PCA) is used for fault detection in environmental processes. The objective is to combine the GLR test with PCA model in order to improve fault detection performance. The PCA-based GLR was proposed to detect the faults in environmental processes representing the crop model, in which, PCA is used to create the model and find a linear combinations of variables which describe major trends in data set, and the GLR test, is utilized to improve faults detection. It is demonstrated that the performance of faults detection can be improved by combining GLR test and PCA.

Acknowledgment

The authors gratefully acknowledge financial support from the Fonds de la Recherche Scientifique - FNRS.

REFERENCES

[1] M. Mourad and J.-L. Bertrand-Krajewski, "A method for automatic validation of long time series of data in urban hydrology," *Water Science & Technology*, vol. 45, no. 4-5, pp. 263–270, 2002.

[2] J. Tang, Z. Chen, A. W.-C. Fu, and D. Cheung, "A robust outlier detection scheme for large data sets," presented at 6th Pacific-Asia Conf. on Knowledge Discovery and Data Mining, Citeseer, 2001.

[3] R. J. Bolton, D. J. Hand *et al*., "Unsupervised profiling methods for fraud detection," *Credit Scoring and Credit Control VII*, pp. 235–255, 2001.

[4] P. J. Rousseeuw and I. Ruts, "Algorithm as 307: Bivariate location depth," *Applied Statistics*, pp. 516–526, 1996.

[5] I. Ruts and P. J. Rousseeuw, "Computing depth contours of bivariate point clouds," *Computational Statistics & Data Analysis*, vol. 23, no. 1, pp. 153–168, 1996.

[6] F. Gonzalez, D. Dasgupta, and R. Kozma, "Combining negative selection and classification techniques for anomaly detection," in *Proc. the 2002 Congress o Evolutionary Computation*, vol. 1, IEEE, 2002, pp. 705–710.

[7] G. H. John, "Robust decision trees: Removing outliers from databases," *KDD*, 1995, pp. 174–179.

[8] S. Lane, E. Martin, A. Morris, and P. Gower, "Application of exponentially weighted principal component analysis for the monitoring of a polymer film manufacturing process," *Transactions of the Institute of Measurement and Control*, vol. 25, no. 1, pp. 17–35, 2003.

[9] D. Nguyen and B. Widrow, "Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights," in *Proc. International Joint Conference on Neural Networks*, 1990, pp. 21–26.

[10] F. Gustafsson, "The marginalized likelihood ratio test for detecting abrupt changes," *IEEE Transactions on Automatic Control*, vol. 41, no. 1, pp. 66–78, 1996.

[11] A. S. Willsky, E. Chow, S. Gershwin, C. Greene, P. Houpt, and A. Kurkjian, "Dynamic model-based techniques for the detection of incidents on freeways," *IEEE Transactions on Automatic Control*, vol. 25, no. 3, pp. 347–360, 1980.

[12] J. R. Dawdle, A. Willsky, and S. W. Gully, "Nonlinear generalized likelihood ratio algorithms for maneuver detection and estimation," in *Proc. American Control Conference*, 1982, pp. 985–987.

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

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

[15] M. Zhu and A. Ghodsi, "Automatic dimensionality selection from the scree plot via the use of profile likelihood," *Computational Statistics & Data Analysis*, vol. 51, pp. 918–930, 2006.

[16] G. Diana and C. Tommasi, "Cross-validation methods in principal component analysis: A comparison," *Statistical Methods & Applications*, vol. 11, no. 1, pp. 71–82, 2002.

[17] I. Jolliffe, *Principal Component Analysis*, Second edition, Springer, Berlin, 2002.

[18] H. Hotelling, "Analysis of a complex of statistical variables into principal components," *Journal of Educational Psychology*, vol. 24, pp. 417–441, 1933.

[19] J. E. Jackson and G. S. Mudholkar, "Control procedures for residuals associated with principal component analysis," *Technometrics*, vol. 21, no. 3, pp. 341–349, 1979.

[20] S. Qin, "Statistical process monitoring: Basics and beyond," *Journal of Chemometrics*, vol. 17, no. 8/9, pp. 480–502, 2003.

[21] A. Benaicha, M. Guerfel, N. Boughila, and K. Benothman, "New pca-based methodology for sensor fault detection and localization," presented at MOSIM'10, Hammamet, Tunisia, May 10-12 2010.

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

[23] I. Jonckheere *et al*., "Review of methods for in situ leaf area index determination: Part i. theories, sensors and hemispherical photography," *Agricultural and Forest Meteorology*, vol. 121, no. 1, pp. 19–35, 2004.

[24] G. Meyer and J. Neto, "Verification of color vegetation indices for automated crop imaging applications," *Computers and Electronics in Agriculture*, vol. 63, no. 2, pp. 282–293, 2008.

[25] M. Tremblay and D. Wallach, "Comparison of parameter estimation methods for crop models," *Agronomie*, vol. 24, no. 6-7, pp. 351–365, 2004.

[26] M. Mansouri, B. Dumont, and M.-F. Destain, "Modeling and prediction of time-varying environmental data using advanced bayesian methods," *Exploring Innovative and Successful Applications of Soft Computing*, vol. 25, no. 7, pp. 112–137, 2014.

[27] M. Mansouri, B. Dumont, and M.-F. Destain, "Prediction of non-linear time-variant dynamic crop model using bayesian methods," pp. 507–513, 2013.

Majdi Mansouri was born in Kasserine, Tunisia. He received the engineering degree in telecommunications in 2006 from the Higher School of Communication of Tunisia (Tunisia). He received his master degree from the School of Electronic, Informatique and Radiocommunications in Bordeaux (ENSEIRB), France in 2008. He received his PhD degree from the School of Troyes University of Technology in Troyes (UTT), France in 2011. He is currently an assistant research scientist in Texas A&M University at Qatar (TAMUQ).

Majdi Mansouri has over seven years of research and practical experiences in the area of systems engineering and signal processing. His work focuses on the utilization of applied mathematics and statistics concepts to develop algorithms for modeling, estimation and prediction, which can help improve process operations. In academia, He worked extensively on the utilization and development of Bayesian inference to improve filtering, modeling, and state estimation, He worked also on process monitoring and faults detection for chemical and environmental systems. He is the author of over 70 papers (25 international journals, 7 book chapters and 38 international conferences).

Marie-France Destain has a PhD in agricultural engineering. She is teaching precision agriculture in a master of environmental science and technology at University of Liege (Gembloux Agro-Bio Tech, Belgium). She is involved with the development of sensors, models and decision support systems aimed at managing crop inputs in an environmentally sensible manner. She is responsible of several national and international research projects related to engineering in the field of agro-biosystems. She is member of several European networks (ManuFuture-AET, ICT-Agri ERANET, MACSur).

Hazem Nounou received the B.S. degree (*magna cum laude*) from Texas A&M University, College Station, in 1995, and the M.S. and Ph.D. degrees from Ohio State University, Columbus, in 1997 and 2000, respectively, all in electrical engineering. In 2001, he was a development engineer for PDF Solutions, a consulting firm for the semiconductor industry, in San Jose, CA. Then, in 2001, he joined the Department of Electrical Engineering at King Fahd University of Petroleum and Minerals in Dhahran, Saudi Arabia, as an assistant professor. In 2002, he moved to the Department of Electrical Engineering, United Arab Emirates University, Al-Ain, UAE. In 2007, he joined the Electrical and Computer Engineering Program at Texas A&M University at Qatar, Doha, Qatar, where he is currently an Associate Professor. He published more than 80 refereed journal and conference papers. He served as an Associate Editor and in technical committees of several international journals and conferences. His research interests include intelligent and adaptive control, control of time-delay systems, system biology, and system identification and estimation. Dr. Nounou is a senior member of IEEE.

Mohamed Nounou received the BS degree (magna cum laude) from Texas A&M University, College Station, in 1995, and the MS and PhD degrees from the Ohio State University, Columbus, in 1997 and 2000, respectively, all in chemical engineering. From 2000 to 2002, he was with PDF Solutions, a consulting company for the semiconductor industry, in San Jose, CA. In 2002, he joined the Department of Chemical and Petroleum Engineering at the United Arab Emirates University as an assistant professor. In 2006, he joined the Chemical Engineering Program at Texas A&M University at Qatar, Doha, Qatar, where he is currently an associate professor. He has published more than 80 refereed journal and conference papers and book chapters. He also served as an associate editor and in technical committees of several international journals and conferences. His research interests include process modeling and estimation, system biology, and intelligent control. He is a member of the American Institute of Chemical Engineers (AIChE) and a senior member of the IEEE.