International Journal of Control, Energy and Electrical Engineering (CEEE)
Vol.11 pp.8-11

Accurate Quality Control Charts via Sparsity Reconstruction For Multimode Process Monitoring

Wafa Bougheloum
Laboratory of Automatics and Signals of Annaba (LASA)
Faculty of Engineering
University Badji-Mokhtar of Annaba
P.O. Box. 12, Annaba, 23000, Algeria
Email: bougheloumwafa@gmail.com

Messaoud Ramdani
Department of Electronics
Faculty of Engineering
University Badji-Mokhtar of Annaba
P.O. Box. 12, Annaba, 23000, Algeria
Email: messaoud.ramdani@univ-annaba.org

***Abstract*—Conventional process monitoring often assumes that process data follow a Gaussian distribution with linear correlation. However, this type of constraint can not be satisfied in practice because many industrial processes are nonlinear in nature. This paper proposes an improved multivariate statistical process monitoring scheme based on the Stacked Sparse Autoencoder (SSAE) to detect and reconstruct the potential invalid/missing data. The Squared Prediction Error (SPE) index is used for novelty detection with an adaptive non-parametric confidence limit derived from the kernel density estimate (KDE) to reduce false alarms. Based on the reconstruction principle, an improved sensor validity index (SVI) is proposed to identify the faulty/missing data. The experimental results using both synthetic and actual data from a drinking-water treatment plant, reveal the effectiveness of the proposed scheme and show its capacity to detect and identify sensor failures.**

***Index Terms*—Process monitoring, Multivariate Statistical Process Control, Data Reconstruction, Sparsity reconstruction.**

I. Introduction

Data mining can extract hidden and useful information from large databases, where potential relationships can be used for automated anomaly detection and related problem root cause (RC) localization. In fact, statistical process control (SPC) charts allow the visualization of process evolution and the detection of the abnormal changes. However, most traditional SPC such as PCA works optimally only in the situation where the correlations are linear, which is most of the time an approximation. Recently, deep learning has been very successful in many applications. Which is an unsupervised algorithm able to find a better representation with a deep learning architecture [1]. As one of deep learning methods, we propose the stacked sparse autoencoder (SSAE) [2] to reconstruct the input data. After reconstruction, the novelty detection is carried out by using the Q-statistics [3] in conjunction with an adaptive non-parametric confidence limit derived from the kernel density estimate (KDE) in order to reduce the false alarm detection rates. After the detection of an abnormal event, we are in the obligation to know which sensor is defective using basically the contribution plots or the reconstruction principle.

This paper focuses on the use of Stacked Sparse Autoencoder (SSAE), which is trained to reconstruct the input data collected from the normal operating mode.

The rest of this paper is organized as follows: Section II outlines the Sparsity Reconstruction based Process Monitoring (SRPM) strategy. In section III, experimental results are carried out using both synthetic and actual data from a drinking-water treatment plant to show the effectiveness of the proposed schemes.

II. Sparsity Reconstruction based Process Monitoring (SRPM)

Deep Learning recently achieved outstanding performance for a variety of tasks. It has been applied successively in the field of image processing and visual analysis, but its application in the field of process control is still rare. The deep neural network models are used as feature extraction tools which contain a hidden layer called bottleneck layer. First, the input vector $X_i = \{1, 2, 3, ..., N\}$ is transformed into a hidden part represented by the function h_i, as follows:

$$h_i = f(x_i) = sigm(W_1 x + b_1) \tag{1}$$

Where W_1 and b_1 are respectively the weight and the bias between the input layer and the hidden part and $sigm(x)$ is a sigmoid function.

In the decoding layer, h_i is mapped to the output denoted by $\hat{x}$. Where we use the activation function shown as follows:

$$\hat{x}_i = g(h_i) = sigm(W_2 h + b_2) \tag{2}$$

Where W_2 and b_2 are respectively the weight and the bias between the hidden part and the output layer ($\hat{x}$). The bottleneck network whose training criterion involves a sparsity penalty in the botleneck layer is called Stacked Sparse Autoencoder (SSAE). The power of this network is to predict its output (estimation of the input) as close as possible to its input, this by optimizing the cost function defined by:

$$J = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{2}\left\|\hat{x}_i - x_i\right\|^2\right) + \frac{\lambda}{2}\sum_{i=1}^{N}\left\|W_i\right\|^2 + \beta\sum_{j=1}^{m} KL(\rho\| \hat{\rho}_j) \tag{3}$$

Where m is the number of the hidden node. λ and β are the coefficient that determine the weight decay and the sparsity penalty terms, respectively.

In the equation (3), the first term represents the reconstruction

ISSN 2356-5608

International Journal of Control, Energy and Electrical Engineering (CEEE)
Vol.11 pp.8-11

error, the second is the regularization term and the last is Sparsity Penalty term, where $KL(\rho\| \hat{\rho}_i)$ is the Kullback-Leibler divergence, it is used to calculate the difference between ρ and $\hat{\rho}_i$, where ρ and $\hat{\rho}_i$ are the constraint used during learning.
The back propagation algorithm is used to minimize the cost function and to find the appropriate parameters W_1, W_2, b_1, b_2.

A. Anomaly Detection

Anomaly detection is the process of identifying unexpected items or events in data sets, which differ from the norm. The common measure is the squared prediction error (SPE):

$$Q = SPE = \sum_{i=1}^{N} (x_i - \hat{x}_i)^2 \tag{4}$$

Where N is the number of samples. In this study, we propose an adaptive confidence limit using a fuzzy clustering algorithm, namely the FCM [4] to divide the normal data of the whole operating regime into smaller number of possibly simpler local operating regimes.

1) Parametric χ^2 Distribution (δ^2_α): The system is considered in its normal operating conditions if $SPE \leq \delta^2_\alpha$. On the other hand, if $SPE > \delta^2_\alpha$, the system is considered defective, where δ^2_α is specified for the SPE control limit [5], which can be calculated using a weighted χ^2 distribution:

$$\delta = \mathrm{g}\chi^2_{\mathrm{h},\alpha} \quad g = \frac{v}{2m} \quad h = \frac{2m^2}{v} \tag{5}$$

Where m and v are the estimated mean and variance of SPE, respectively.

2) Adaptive confidence limit based on Kernel Density Estimation (KDE) ($AUCL_{KDE}$): Kernel density estimation (KDE) is a very powerful tool for the nonparametric estimation of the probability density function of a random variable at any point of the support. Given a sample matrix with n variables and m samples, the KDE of the density function $f(x)$ at any point x is defined as follows:

$$f(x) = \frac{1}{mh} \sum_{j=1}^{n} K\left(\frac{x - x_j}{h}\right) \tag{6}$$

Where h is the bandwidth parameter and K is a Kernel function that integrates to one and has zero mean.
Many classification algorithms for different problems have been proposed. Fuzzy c-means (FCM) is a grouping method that allows a data item to belong to two or more clusters. It is based on the minimization of the following objective function:

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^m \left\| x_i - c_j \right\|^2 \tag{7}$$

Where m is is the fuzziness parameter, u_{ij} is the degree of membership of x_i in the cluster j and x_i is the ith of the measured data and c_j is the center of the cluster.
The clustering algorithm that has been used in this article is the Gustafson-Kessel (GK) [6].

B. Fault Identification

1) Contribution Plots: There are several methods of faults identification. For this purpose, contribution plots can be used. The contribution of variable j to the Q statistic is calculated as follows:

$$C_{ijk}^{Q} = e_{ijk}^2 \tag{8}$$

Where $e = (x_i - \hat{x}_i)$.

2) Sensor Validity Index (SVI): To be able to reconstruct the faulty data, it is necessary to determine the fault in a unique way (figure 1). The approach consists in predicting the

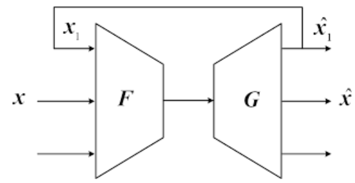

Fig. 1: Reconstruction Principle

measurement $\hat{x}_j$ of the process, by replacing the j^{th} process variable by the predicted one and repeating the operation until convergence of the algorithm as follow:

$$\tilde{x}_i = \xi_j^T G(F(x_j)) \tag{9}$$

Where $\tilde{x}_i = (x_1, x_2, ..., \hat{x}_j, ..., x_m)$, ξ_j^T is the j^{th} column of the identity matrix. The Sensor Validity Index (SVI) is the measure of sensor performance where standard range should exist regardless of the number of principal components of the disturbances or faults [7], it is defined as follows: $\eta_j^2(k) = \frac{SPE_j(k)}{SPE(k)}$ Where SPE is the quadratic global prediction error computed before reconstruction and SPE_j is the jth quadratic prediction error computed after reconstruction [8]. The validity index of a faulty sensor must converge towards zero.

III. Experimental results

A. Synthetic data

We use dataset [9] containing three variables where t is uniformly distributed in the interval $[-1, 1]$; ε_i denotes the Gaussian white noises with zero means and standard deviation of 0.01 and 1000 samples collected to build SSAE model.

$$\begin{aligned} x_1 &= t^2 + 0.3\sin(2\pi t) + \varepsilon_1 \\ x_2 &= t + \varepsilon_2 \\ x_3 &= t^3 + t + 1 + \varepsilon_3 \end{aligned} \tag{10}$$

After creating the model, we check the evolution of SPE under normal conditions, where statistical and adaptive threshold is calculated. The result is shown in figures 2. For the clarity of the results we will use a window from 0 to 250 samples. We simulate a fault in one of the sensors and we notice the evolution of SPE with the two thresholds as it is illustrated in figure 3. By examining the figures in normal and faulty

ISSN 2356-5608

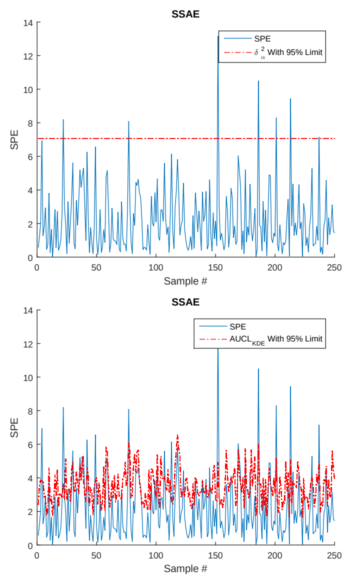

Fig. 2: SPE: data in normal state (statistical and adaptive threshold)

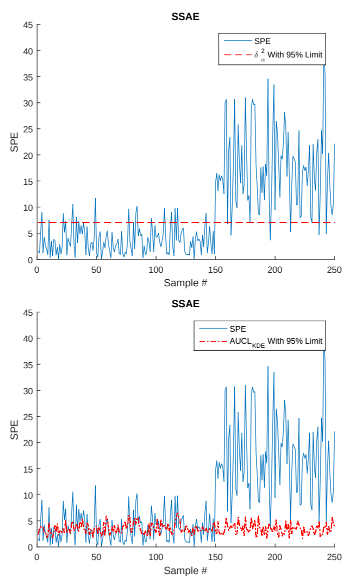

Fig. 3: SPE: data in faulty state (statistical and adaptive threshold)

state, we can see a false alarm in our data caused generally by outlier measures. The SSAE model is able to detect the fault, the detection was at 150 sample. Also, we can identify the faulty sensor via the reconstruction principle Fig 4, where the corresponding SVI tends to zero, as shown in 5.

B. Case study: Drinking Water Treatment Plant

Traditional drinking water treatment treating surface water usually include four important processes: flocculation, sedimentation, filtration and disinfection, as shown in figure 6.

Adding chemicals to water may be the most important process in the surface treatment plant. The main function of the unit is

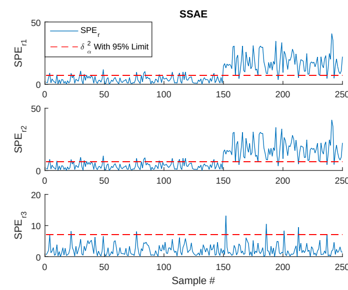

Fig. 4: Fault isolation using reconstruction principle (fault in the 3rd sensor)

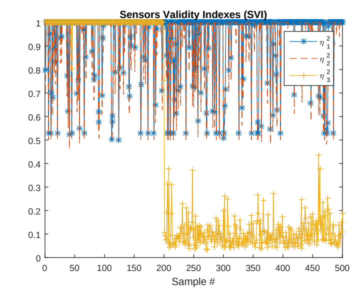

Fig. 5: Fault identification using Sensor Validity Index (SVI) (Fault in the 3rd sensor)

chemical coagulation, in which chemicals, usually aluminum or iron salts, are added into the water for the purpose of producing flocs of colloidal particles and deposition of other contaminants. In our study, the model inputs consist of raw water parameters, while the model output is the best dose of coagulant to achieve the required quality of treated water. This section focuses on the development of self-coagulation control based on the total water parameters to calculate the required dose. The plant we studied is the drinking water treatment plant of Oued Athmania, located at the northwest of the Constantine Province, in the Mila City in East of Algeria. It is responsible for the distribution of drinking water to many citizens at and around Constantine (Algeria) [10].

The data consists of raw water and treated water parameters: Turbidity, Temperature, PH and O_2, so in total we have eight parameters implying that we have 8 sensors to monitor. The measurements are collected from a data acquisition unit

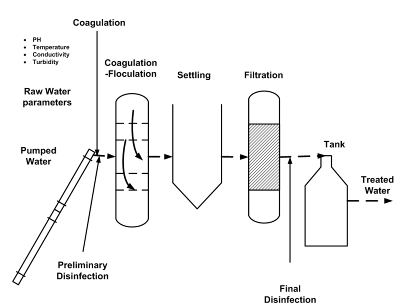

Fig. 6: The typical drinking water processing units

ISSN 2356-5608

(SCADA system) covering a period of 356 days, including different periods. This dataset is used to build the SSAE model. The result obtained for SPE in normal conditions with the two types of thresholds is illustrated in figure 7. For the clarity of the results we will use a window from 0 to 200 samples. We

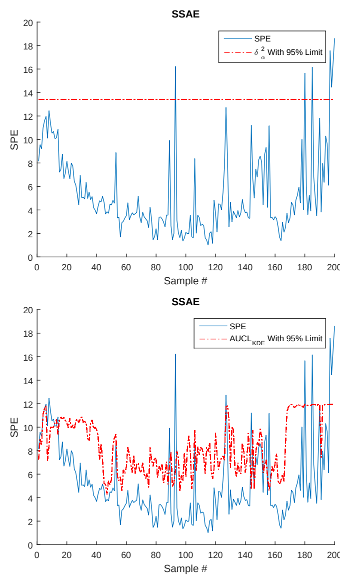

Fig. 7: SPE evolution for the the healthy state.

inject a fault in the dissolved oxygen O_2 of the treated water and we observe the evolution of the SPE, which is shown in figure 8. It is clear that the SSAE model is able to detect and isolate the fault, the detection starts from the $120th$ sample and the faulty sensor was indicated as the $8th$ sensor which is the O_2 of the treated water (TW), 9.

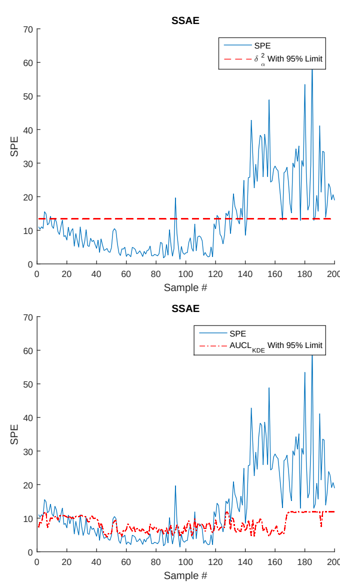

Fig. 8: SPE evolution in faulty state

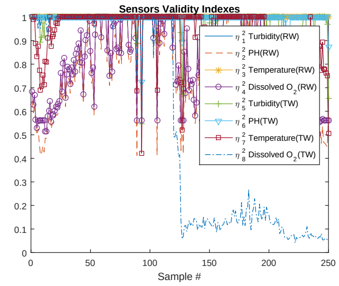

Fig. 9: SVI's Raw Water (RW) and Treated Water (TW) sensors

IV. Conclusions

In this paper, an accurate quality control charts strategy based on Sparsity Reconstruction is proposed for multimode process monitoring. A Stacked Sparse Autoencoder (SSAE) was build and used to reconstruct the input data. The novelty detection of upsets or abnormal events is carried out by an adaptive upper control limit, which is well suited for non Gaussian processes. To identify the defective sensor, an improved Sensor Validy Index (SVI) based on the reconstruction principle is proposed. Experimental results obtained on synthetic and actual data from a drinking water treatment plant reveal the effectiveness of the proposed system and demonstrate its ability to detect and identify sensor failures.

References

[1] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *science*, vol. 313, no. 5786, pp. 504–507, 2006.

[2] J. Xu, L. Xiang, Q. Liu, H. Gilmore, J. Wu, J. Tang, and A. Madabhushi, "Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 1, pp. 119–130, jan 2016.

[3] D. White, B. Goodlin, A. Gower, D. Boning, H. Chen, H. Sawin, and T. Dalton, "Low open-area endpoint detection using a PCA-based t/sup 2/ statistic and q statistic on optical emission spectroscopy measurements," *IEEE Transactions on Semiconductor Manufacturing*, vol. 13, no. 2, pp. 193–207, may 2000.

[4] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Computers & Geosciences*, vol. 10, no. 2-3, pp. 191–203, jan 1984.

[5] P. Nomikos and J. F. MacGregor, "Multivariate spc charts for monitoring batch processes," *Technometrics*, vol. 37, no. 1, pp. 41–59, 1995.

[6] J. V. De Oliveira and W. Pedrycz, *Advances in fuzzy clustering and its applications*. John Wiley & Sons, 2007.

[7] R. Dunia, S. J. Qin, T. F. Edgar, and T. J. McAvoy, "Use of principal component analysis for sensor fault identification," *Computers & Chemical Engineering*, vol. 20, pp. S713–S718, jan 1996.

[8] K. Bouzenad, M. Ramdani, N. Zermi, and K. Mendaci, "Use of NLPCA for sensors fault detection and localization applied at WTP," in *2013 World Congress on Computer and Information Technology (WCCIT)*. IEEE, jun 2013.

[9] X. Zhang and Y. Li, "Multiway principal polynomial analysis for semiconductor manufacturing process fault detection," *Chemometrics and Intelligent Laboratory Systems*, vol. 181, pp. 29–35, oct 2018.

[10] K. Mendaci, M. Ramdani, and T. Benzaraa, "Nonlinear multivariate statistical process monitoring of a water treatment plant," in *2013 5th International Conference on Modeling, Simulation and Applied Optimization (ICMSAO)*. IEEE, apr 2013.

ISSN 2356-5608