Multi-Scale PLS Modeling for Industrial Process Monitoring

Mohammad Sadegh Emami Roodbali
Petroleum University of Technology
Ahwaz, Iran

Mehdi Shahbazian
Petroleum University of Technology
Ahwaz, Iran

ABSTRACT
In the process monitoring procedure, Data-driven (statistical) methods usually rely on the process measurements. In most industrial process this measurements has a multi-scale substance in time and frequency. Therefore the statistical methods which are proper for one scale may not be able to detect events at several scales. A Multi-Scale Partial Least Squares (MSPLS) algorithm consists of Wavelet Transforms for extracting multi-scale nature of measurements and Partial Least Squares (PLS) as a popular technique of statistical monitoring methods. In this paper the MSPLS algorithm is applied for monitoring of the Tennessee Eastman Process as a benchmark. To show the advantages of MSPLS, its process monitoring performance is compared with the standard PLS and is proved that MSPLS can be a more efficient technique than standard PLS for fault detection in industrial processes.

Keywords
Process, monitoring, fault, wavelet, PLS, multi-scale.

1. INTRODUCTION
The primary aim of process monitoring is to ensure the success of planned operation by detecting the abnormal events during the operation. In process control subject these abnormal events usually called fault. Identification and diagnosis of these faults and recovery of the process are the further steps of process monitoring loop [1].

There are a lot of methods for process monitoring such as analytical methods, artificial intelligence and statistical approaches. Among these various methods, there are methods that do not assume the form of model a-priori and rely only on historical process data. These methods are called data-driven or statistical methods. Although the model-based methods have been usually used in process monitoring, modeling and simulation are very difficult and sometimes impossible tasks in many real industrial processes. Therefore, the data-driven techniques can be a useful and practical tool for process monitoring.

Partial Least Squares (PLS) is one of the most common and widely used methods for statistical process monitoring. PLS usually used where the process data can be grouped into two sub-blocks as predictor and predicted (or response) blocks which contain process variables and quality variables respectively [1, 2]. For monitoring task, when the PLS model is constructed, the abnormal process conditions can be easily detected and diagnosed based on the statistic monitoring indices.

PLS similar to the other advanced statistical methods, such as PCA, FDA and CVA have been successfully applied in many process control applications. However, the problem is that the data blocks described above contain the process measurements which have a multi-scale nature in time and frequency, and in the other side, most statistical methods are proper for detecting events at only one scale [3]. To overcome this problem, a multi-scale process monitoring method called MSPCA (multi-scale principal component analysis) was suggested for extracting relationship between the measurements by wavelet analysis [4]. After that, many researchers studied on multi-scale techniques for process monitoring and suggested new further multi-scale techniques. Multi-scale PLS (MSPLS) is one of these process monitoring techniques that can adapt to the scale of relevant signal using Wavelet analysis [3].

In this paper, a multi-scale PLS model is constructed using the proper Wavelet Transform and Standard PLS model. The monitoring ability of the MSPLS is tested by monitoring of Tennessee Eastman Process (TEP) as a benchmark and the results are compared with the standard PLS method.

2. PLS MODEL
Partial Least Squares (PLS) Analysis works on the assumption that the focus of analysis is on which aspects of the signal in one matrix are related directly to signals in another matrix. The construction of a PLS model involves the decomposition of both the input or predictor X (n samples of m variables) and the output or predicted Y (n samples of p variables) data blocks in such a way as to maximize the covariance between them. A linear relationship can then be developed between the two sets of latent variables [5].

Mathematically, the outer model, describing the projection stage of the algorithm, is given by:

$$X = TP^T + E = \sum_{h=1}^{a} t_h p_h^T + E \qquad (1)$$

$$Y = UQ^T + F = \sum_{h=1}^{a} u_h q_h^T + F \qquad (2)$$

Where $T \in R^{n \times a}$ and $U \in R^{n \times a}$ use $t_1 ... t_a$ and $u_1 ... u_a$ (the score vectors) as columns, while P^T ($a \times m$) and Q^T ($a \times p$) use $p_1^T ... p_a^T$ and $q_1^T ... q_a^T$ (the loading vectors) as rows. E and F represent the residuals of the X and Y blocks, respectively [5].

The inner relationship, for estimating the Y block scores, for the linear case can then be described by:

$$\hat{u}_h = b_h t_h \qquad (3)$$

So the regression parameter b_h can be derived from training data as:

$$b_h = \frac{u_h^T t_h}{t_h^T t_h} \qquad (4)$$

The goal of partial least squares is to predict Y from X and to describe the common structure underlying the two variables [6]. For monitoring aim, PLS use the training data sets to construct the relevant model which can be able to predict the new depended variables from new independent variables. Any abnormal behavior of system then could be detected by comparing the new variables with the corresponding variable predicted in the PLS model or using the monitoring indices such as Hotelling's T^2 and squared prediction error (SPE).

3. MSPLS MODEL DESCRIPTION
3.1 Wavelet Transform
Multi-scale process monitoring algorithms usually use wavelet Transform to decompose the original process measurements into their multi-scale components according to time and frequency characteristics [2]. A wavelet Transform involves the decomposition of a signal function or vector into simpler, fixed building blocks at different scales and positions [7].

A family of wavelet basis functions may be represented as

$$\psi_{su}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right) \qquad (5)$$

Where *s* and *u* represent the dilation and translation parameters, respectively, and $\psi(t)$ is the mother wavelet. For most practical applications to measured data, the wavelet dilation and translation parameters are discretized dyadically, and the family of wavelets is represented as

$$\psi_{mk} = 2^{\frac{-m}{2}} \psi\left(2^{-m} t - k\right) \qquad (6)$$

Where $\psi(t)$ is the mother wavelet, *m* and *k* are the dilation and translation parameters, respectively.

Any signal can be decomposed into its contributions from multiple scales as a weighted sum of dyadically discretized orthonormal basis functions:

$$y(t) = \sum_{m=1}^{L} \sum_{k=1}^{N} d_{mk} \psi_{mk} + \sum_{k=1}^{N} a_{Lk} \varphi_{Lk}(t) \qquad (7)$$

Where *y* is a time signal, d_{mk} represents the wavelet or detail signal coefficient at scale *m* and location *k*, and a_{Lk} represents the scaled signal or scaling function coefficient of φ_{Lk} (t) at

the coarsest scale *L* and location *k* [3]. The scaling function φ (t) is responsible for capturing the coarse (low frequency) contents of the signal. mong the efficient methods for decomposing a signal on a family of wavelet basis functions based on convolution with the corresponding filters, Mallat developed an efficient way to implement the DWT (Discretized Wavelet Transform) iteratively using the scaling (low-pass) filter H and the wavelet (high-pass)filter G as follows [8]:

$$d_s = G d_{s-1} \qquad (8)$$

$$a_s = H a_{s-1} \qquad (9)$$

Where a_s is the approximation coefficient vector at scale s representing the lower frequency information contained in the upper level approximation coefficient vector a_{s-1} while d_s is the detail coefficient vector representing the higher frequency content. Once the wavelet type and the decomposition level S are specified, the original signal f can be efficiently decomposed into its multi-scale components which consist of one approximation coefficient vector at the coarsest scale S, and S detail coefficient vectors at all the scales from the finest one to the coarsest one [2].

3.2 MSPLS Algorithm
Please For on-line process monitoring, the wavelet decomposition described previously, should be performed concerning each process variable in on-line mode. To implement the on-line DWT the scheme of moving window is used [3, 4, 7] where a dyadic length of the window corresponding to the selected decomposition level S (length=2^S) is always maintained by including the most recent measurements while discarding the oldest ones. It must be considered that the translation invariance characteristic of the on-line DWT, which is not usually guaranteed in the off-line DWT, is necessary in multi-scale statistical analysis [9].

To actualize an on-line multi-scale version of PLS that could fully use the functional properties of the wavelet analysis mentioned above, the predictor data matrix $X \in R^{n \times m}$ and the response data matrix $Y \in R^{n \times p}$ are first decomposed into the following multi-scale matrix $X \in R^{n \times m(s+1)}$ and $Y \in R^{n \times p(s+1)}$ using the on-line DWT [2]:

$$X_W = \left[X_1 X_2 ... X_s X_{s+1}\right] \qquad (10)$$

$$Y_W = \left[Y_1 Y_2 ... Y_s Y_{s+1}\right] \qquad (11)$$

Where each sub-block $X_S \in R^{n \times m}$ or $Y_S \in R^{n \times p}$ (s=1, 2… S) Contains the detail coefficients at scale s and the last sub-blocks X_{S+1} and Y_{S+1} contain the approximation coefficients at the coarsest scale S. In this study the same decomposition level is employed for both X and Y.

Figure 1 shows the schemes of the MSPLS applied in this study. As shown in this figure, the PLS model is formed once, separately on each corresponding pair of sub blocks X_S and Y_S to describe the localized features in the given scale, and

another time on X_W and Y_W to describe the global relationship between X and Y which is common in the entire scale blocks.
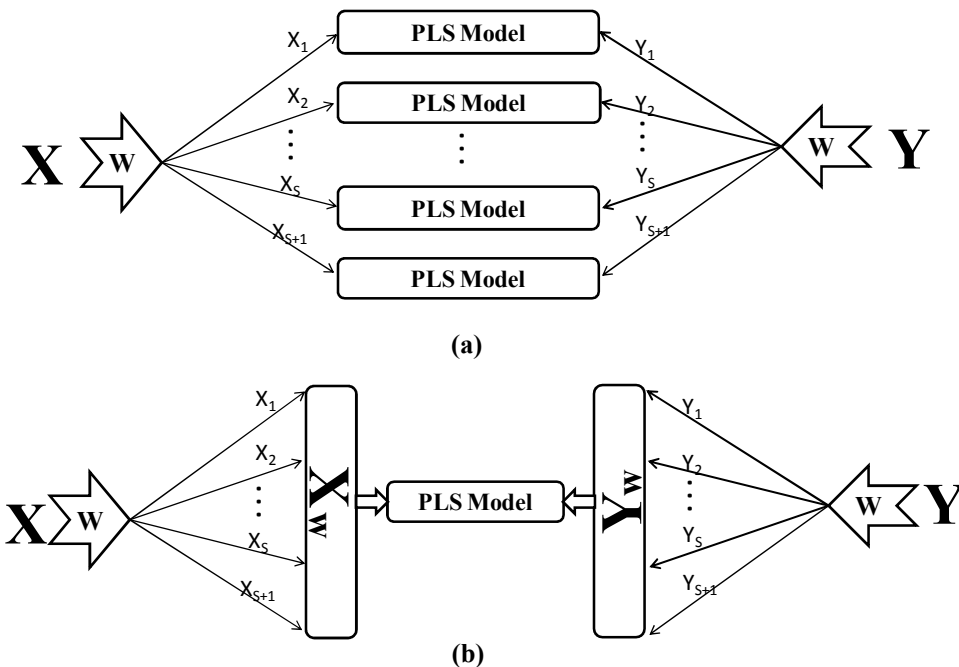

Figure 1. a) The scheme of MSPLS at scale-level. b) The scheme of MSPLS at global level.

So, for online process monitoring via MSPLS, the predictor and predicted data blocks (X and Y) should be decomposed using online DWT into their relevant scales and then, the monitoring procedure using PLS should be implemented for overall data blocks X_W and Y_W containing all scales and also for each scale separately.

4. TENNESSEE EASTMAN PROCESS

The Tennessee Eastman Process (TEP) was created by the Eastman Chemical Company to provide a realistic industrial process for evaluating process control and monitoring methods [10]. The test process is based on a simulation of an actual industrial process where the components, kinetics, and operating conditions have been adjusted for specific aims. The process consists of five major units: a reactor, condenser, compressor, separator, and stripper; and, it contains eight components: A, B, C, D, E, F, G, and H [1].
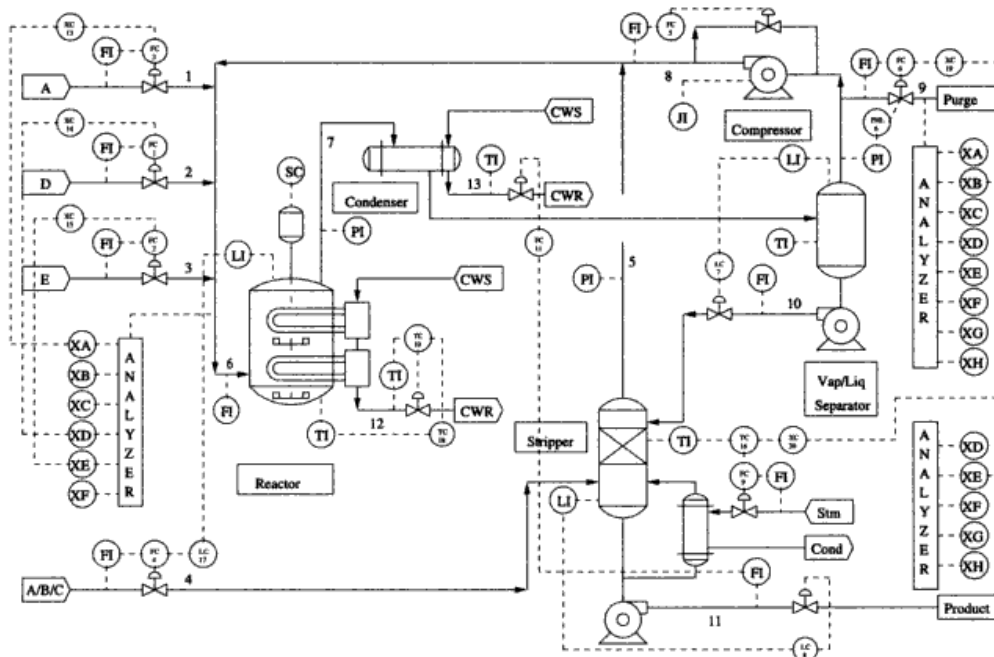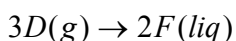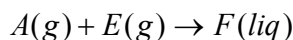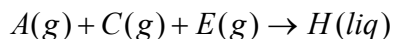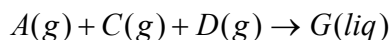
Figure 2 the processes flow sheet for the Tennessee Eastman Process [11].

As shown in Figure 2, the gaseous reactants A, C, D, and E and the inert B are fed to the reactor where the liquid products G and H are formed and the species F is a by-product of the reactions. The reactions in the reactor are:

$$A(g) + C(g) + D(g) \rightarrow G(liq)$$

$$A(g) + C(g) + E(g) \rightarrow H(liq)$$

$$A(g) + E(g) \rightarrow F(liq)$$

$$3D(g) \rightarrow 2F(liq)$$

The labels in Figure 2 represent flow meters (FI), thermometers (TI), pressure gauges (PI), level detectors (LI), agitator speed control (SC), steam supply (Stm), and cooling water supply/recycle (CWS/CWR) (Wilson et al., 2000). The reactions are exothermic, irreversible, and nearly first-order with respect to the reactant concentrations. The reaction rates are Arrhenius functions of temperature where the reaction for G has higher activation energy than the reaction for H, resulting in a higher sensitivity to temperature [1]. Here is the sufficient description was need to this study and the further description of TEP exists in the other literatures [3, 12, 13].

The process contains 53 variables containing 41 measured and 12 manipulated variables. In this study, the measurements of components F, G and H have been selected as product variables to forming a matrix as response block, Y. The other variables formed the predictor block, X. It should be noted that the process measurements has not identical sampling interval and time delay and all of them include Gaussian noise which should be considered in data preparing process.

Among various preprogrammed faults in the Tennessee Eastman Process simulation we selected three known faults to compare the monitoring ability of MSPLS algorithm with the standard PLS algorithm. The training set applied in this study is consisting of 500 observations for each variable which was generated with no fault. Three sets of data have been used for testing, each containing a separable known fault. Fault 1 is connected to the step change in the cooling water. Fault 2 is a low drift in the reaction kinetics, and Fault 3 is associated with one of the sticking valves.

5. MONITORING OF TENNEESSE EASTMAN PROCESS USING MSPLS

Although the MSPLS algorithm described earlier [2] has been used in this study because of its ability to use in online application, simplicity and realizable structure, but there are some other algorithms for multi-scale extension of statistical methods proposed by other researchers. In order to build a MSPLS algorithm for online monitoring application, some prior conditions should be first specified. The type of mother wavelet for DWT and the optimum decomposition level are the most important of these prior conditions. In this study the 'Haar' wavelet was used because of its simplicity and usual use in multi scale statistical process control. Maulud et al. [14] obtained the optimum decomposition level of wavelet for multi-scale PCA based on the variance explained by the first principle component of PCA model which was constructed using the approximation block generated at every decomposition level. Based on this systematic approach, but using PLS model instead PCA, the optimum decomposition level for this study was obtained (three). Therefore each data block was decomposed into one approximation block and three detail blocks.

The number of latent variables in PLS model is an important factor that should be selected correctly. Many researchers use cross validation or other efficient statistical approaches to estimate the optimum number of latent variables [15]. In this study, using the cross validation technique, the number of latent variables for standard PLS and MSPLS were obtained three and nine, respectively.

Process monitoring abilities of the standard PLS model and the MSPLS model has been compared using the simulated Tennessee Eastman Process data sets containing three different faults which described earlier in this study. The same sampling interval, three minutes, was used to collect the all simulated data to allow fast monitoring and good comparison between techniques that either consider or ignore serial correlations.

As mentioned earlier, three faulty data blocks have been used each containing 980 observations as testing data. In order to execute the MSPLS model, each process measurement vector were first decomposed using the on-line DWT, and then the monitoring indices, Hotelling's T^2 and squared prediction

error SPE^X and SPE^Y were calculated and the chart for each of them was drawn versus time. The control limits associated with all the monitoring indices mentioned above was calculated from their characteristic distributions. The detailed equations for the control limits are given elsewhere [16]. When calculating the overall fault detection rate, it was assumed that the given fault was successfully detected whenever the violation of a control limit occurred in any of the global level monitoring diagrams (T^2, SPE^X and SPE^Y) [2]. It was assumed that these calculations were done without any time delay.

6. RESULTS AND DISCUSSION

All fault detection rates that were calculated for each test are presented in table 1. As it is shown in this table, the fault detection precision of the MSPLS is clearly better than the standard PLS. Also, to compare the monitoring abilities of MSPLS with the standard PLS, the monitoring charts of standard PLS, MSPLS at global level and MSPLS at scale level are presented together for each above mentioned fault, separately.
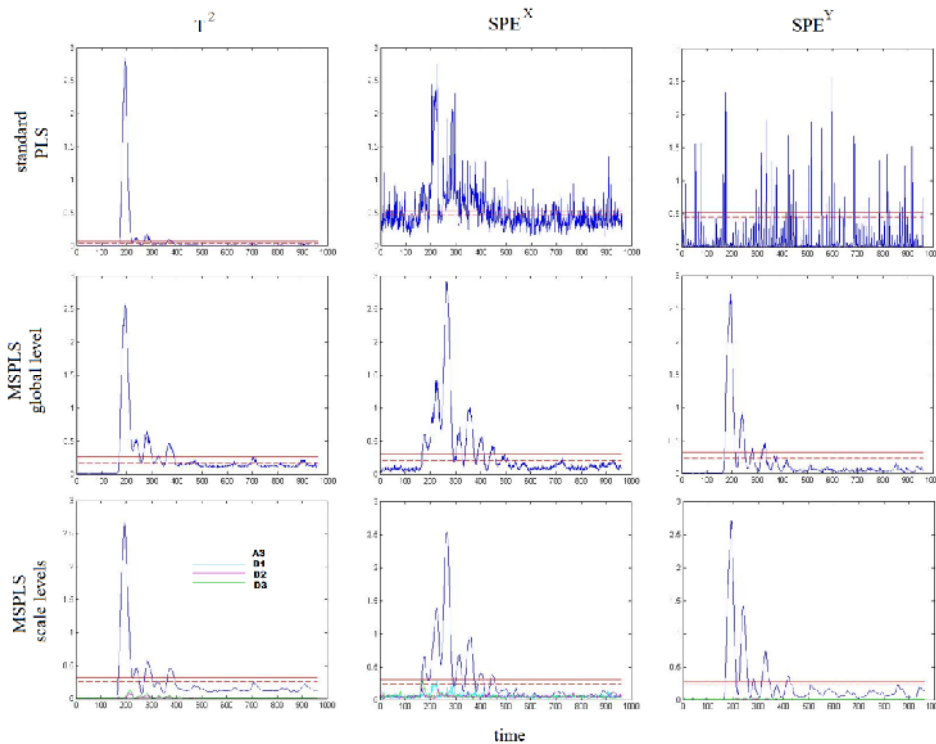

Figure 3. Monitoring abilities of the standard PLS and MSPLS for fault 1. Red solid and dashed lines represent the 99% and 95% confidence limits, respectively

The monitoring charts for Fault 1 (connected to the step change in the cooling water) are presented in Figure 3. It is obvious form this Figure that the MSPLS at global level could clearly detect the fault while the standard PLS was not able to do it as well as the MSPLS. This prominence of MSPLS can also be seen in Table 1 where the fault detection rate of MSPLS at global level (0.92) is more than this corresponding value for standard PLS (0.79). From the monitoring charts of MSPLS at scale level it is obvious that the fault is mostly detected in approximation block A3. The sources of detected fault can be

further identified by exploring the contributions profiles of this approximation block.

Table 1: Fault detection rates of MSPLS and standard PLS for the simulated faults (calculated using 95% control limits). A3: approximation block at level 3, D3: detail block at level 3, D2: detail block at level 2, D1: detail block at level 1.

		MSPLS				
	PLS	Global level	A3	D3	D2	D1
Fault1	0.79	0.92	0.88	0.04	0.06	0.09
Fault2	0.72	0.89	0.93	0.08	0.20	0.20
Fault3	0.45	0.73	0.52	0.56	0.41	0.70

Figure 4 represents the monitoring charts for Fault 2 (a low drift in the reaction kinetics). Similar to the previous charts, it can be seen that the MSPLS at global level could detect the fault better than standard PLS method. It is confirmed at table 1 where the fault detection rate of the global level MSPLS (0.89) is more than standard PLS (0.72). Again, the monitoring charts of the MSPLS at scale level give this point that the fault is mostly detected at approximation block A5, which can be helpful for further monitoring steps.
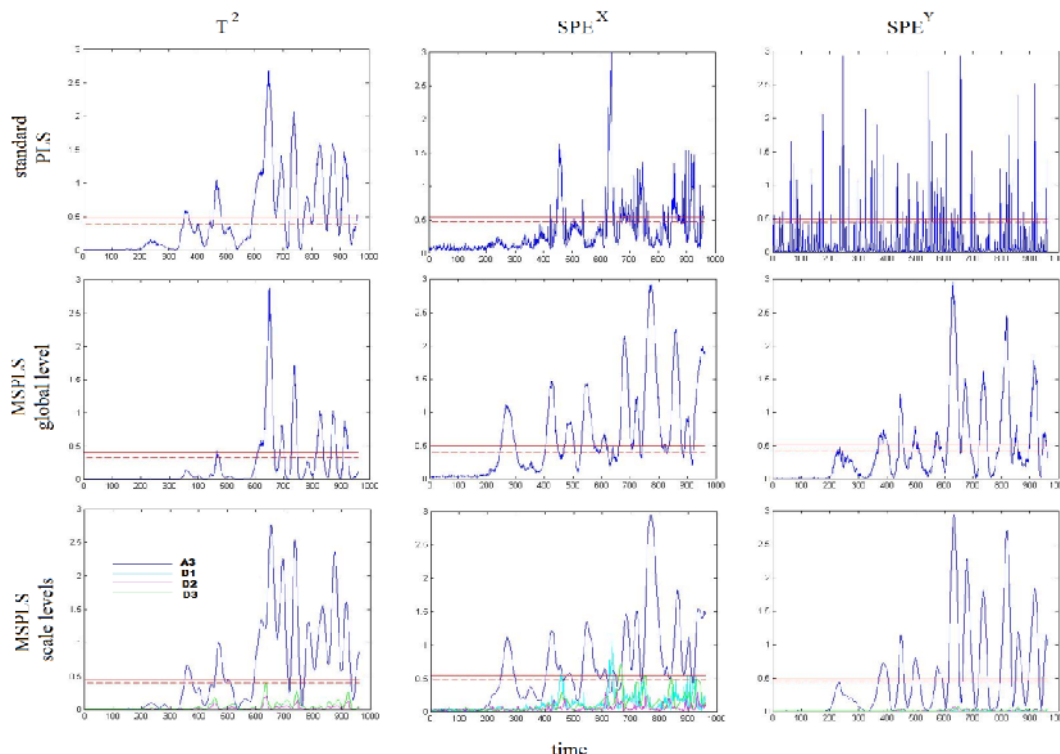

Figure 4. Monitoring abilities of the standard PLS and MSPLS for fault 2. Red solid and dashed lines represent the 99% and 95% confidence limits, respectively.

The most superiority of MSPLS technique can be seen in Figure 5 where the standard PLS is not able to detect Fault 3(associated with one of the sticking valves) adequately, but the MSPLS at global level can finely do it. As it can be seen in table 1, the fault detection rate of the standard PLS (0.45) is not acceptable for a reliable monitoring while the MSPLS detects this fault with a satisfactory detection rate (0.73). the monitoring charts of MSPLS at scale level in Figureure.5 also specifies that for the further identifying of the fault, the monitoring index SPE^X of detail block D1 should be explored.
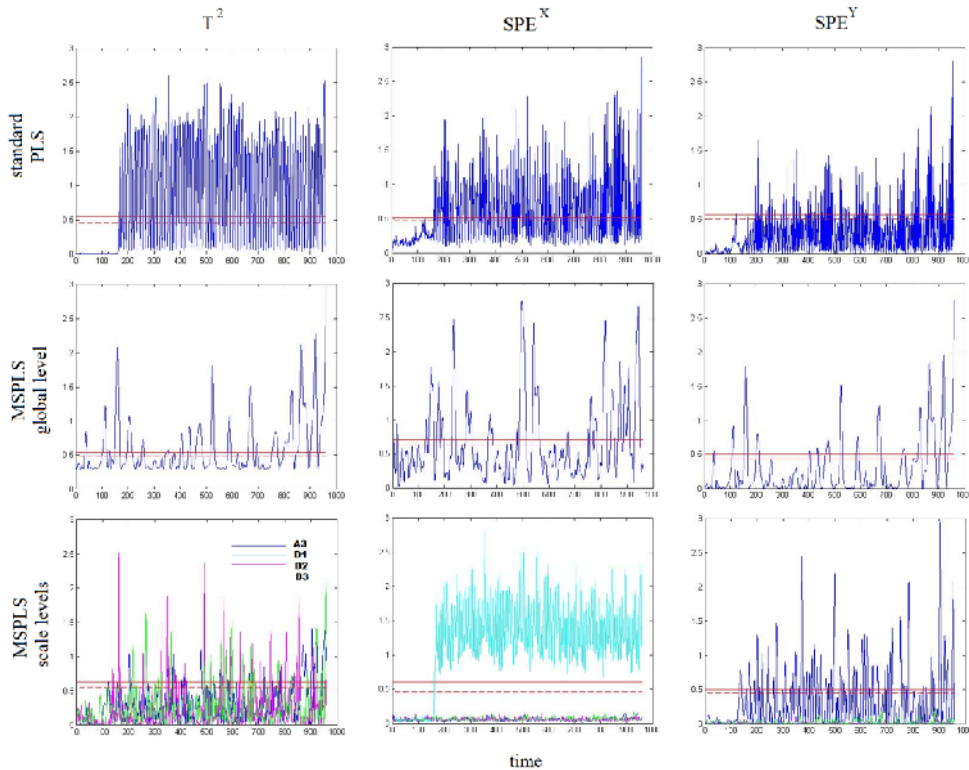
Figure 5. Monitoring abilities of the standard PLS and MSPLS methods for fault 3. Red solid and dashed lines represent the 99% and 95% confidence limits, respectively.

7. CONCLUSIONS

In most industrial processes, the process measurements have a multi-scale nature in time and frequency. Therefore, it is required to apply an optimum way for monitoring of these processes to take into consideration this multi-scale and multivariate nature of process dynamics. Unlike the most of MSPLS algorithms proposed heretofore, which had not very superiority than the standard PLS, the MSPLS algorithm used in this study proved that can be a more efficient technique than the standard PLS for monitoring of industrial processes. A suitable MSPLS algorithm which takes advantage of useful properties of wavelet analysis, in addition to improving the performance of PLS technique to fault detection, can be an advantageous method to further identifying the fault sources.

8. REFERENCES

[1] L.H Chaing,., E.L Russel, and R.D Braatz,.. Fault Detection and Diagnosis in Industrial Systems, Springer, London, 2001.

[2] H.W Lee, M.W Lee, and J.M. Park., Multi-scale extension of PLS for advanced on-line process monitoring, Chemometrics and Intelligent Laboratory Systems, 98:201-212, 2009.

[3] H.B. Aradye, B.R. Bakshi, R.A Strauss,.and J. F. Davis, Multi-scale SPC using wavelet, AIChEJ, 44(7):1596-1610, 2003.

[4] B. R. Bakshi, Multiscale PCA with application to multivariate statistical process monitoring. AIChEJ, 44: 1596–1610, 1998.

[5] D. J. H. Wilson and G. W. Irwin. PLS modelling and fault detection on the Tennessee Eastman benchmark, International Journal of Systems Science, 31(11): 1449 – 1457, 2000.

[6] H. Abdi, Partial Least Squares Regression. In Lewis-Beck, M., A. Bryman, & T. Futing, editors, Encyclopedia of Social Sciences Research Methods. Thousand Oaks: Sage, 2003.

[7] S. Yoon and J.F. MacGregor, principal component analysis of multi scale data for process monitoring. AIChEJ, 50(11):2892-2903, 2004

[8] S.G. Mallat, A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans. Pattn Anal. Mach. Intell., 11: 674-693, 1989.

[9] G. P Nason and *T.* Sapatinas, Wavelet packet transfer function modelling of nonstationary time series. Statist and Computing, 12: 45–56, 2002.

[10] J.J., Downs and E.F Vogel,. A plant-wide industrial process control problem. Computer and chemical engineering 17: 245-255, 1993.

[11] P.R. Lyman, Plant-wide control structure for Teenesse Eastman problem. Computers and chemical engineering,19:321-331, 1995.

[12] I., Hashimoto, M., Kano, and K. Nagao, A new method for process monitoring using principal component analysis, AIChE, Annual meeting, paper 224a, 1999.

[13] W.E Larimore and D.E. Seborg, Process monitoring and identification of dynamic systems using statistical techniques. Los Angeles, CA, USA, 1997.

[14] A. Maulud, D. Wang, and J.A. Romagnoli, Process Control 16: 671–683, 2006.

[15] A. Simoglou, E.B. Martin, and A.J. Morris, Comput. Chem. Eng. 26 909–920, 2002.

[16] X. Wang, U.K. Kruger and B. Lennox, Control Eng. Pract. 11 613–632, 2003.

[17] Y Zhang and C. Ma, Fault diagnosis of nonlinear processes using multi scale KPCA and multi scale KPLS, Chemical engineering science, 66:64-72, 2011

[18] Y.Zhang, S.J.Qin,. improved nonlinear fault detection technology and statistical analysis. AIChEJ, 45(12): 3270-3220, 2008.