Proceedings of the 9th International Symposium on
Dynamics and Control of Process Systems (DYCOPS 2010),
Leuven, Belgium, July 5-7, 2010
Mayuresh Kothare, Moses Tade, Alain Vande Wouwer, Ilse Smets (Eds.)

Multivariate Process Monitoring Using Classical Multidimensional Scaling and Procrustes Analysis

M. Y. M. Yunus and J. Zhang

School of Chemical Engineering and Advanced Material, Newcastle University, Newcastle Upon Tyne, NE1 7RU, UK
(e-mail: m.y.mohd-yunus@ ncl.ac.uk; jie.zhang@ncl.ac.uk)

Abstract: This paper presents a new process monitoring framework using multidimensional scaling. The traditional method of multivariate process monitoring is generally base on principal component analysis (PCA) and is carried out by monitoring the fault detection parameters Hotelling's T^2 and squared prediction errors (SPE). Both indexes are derived directly from multivariate scores in the observation sample configurations. This conventional system was found inappropriately used especially in monitoring highly nonlinear multivariate processes leading to a great number of principal components being selected. Alternatively, classical multidimensional scaling (CMDS) is another technique which can be used in compressing multivariate data by using dissimilarity measures for process monitoring. The proposed process monitoring system is developed based on variable relationships and the dissimilarity measures in terms of variable profiles are used in projecting the multivariate scores. A new monitoring index, which is the resultant vector length different between the new and the normal variable profiles, is introduced. Procrustes analysis (PA) is implemented for on-line process monitoring through a moving-window mechanism. The proposed monitoring method is demonstrated on a simulated continuous stirred tank reactor (CSTR) with recycle system. The results show that the proposed system was efficient as well as effective in detecting various abrupt and incipient faults compared to the linear PCA-based scheme.

Keywords: Multivariate Statistical Process Control, Process Monitoring, Principal Component Analysis, Multidimensional Scaling, Classical Scaling.

1. INTRODUCTION

Multivariate statistical process control (MSPC) or multivariate statistical process monitoring (MSPM) has been shown to be a very effective process monitoring tool. The framework which has been originated from the method of statistical process control (SPC) is aimed to maintain consistent productivity by way of anticipating early warning of possible process malfunctions in the multivariate process (Martin et al., 1996). In this respect, fault detection scheme via monitoring concurrently the Hotelling's T^2 and squared prediction errors (SPE) indexes have been derived directly from the multivariate scores (MacGregor and Kourti, 1995). Qin (2003) emphasizes that both indexes serve different functions, where the first relates to the deviation scales of the current measurement from the targeted mean, whereas the second denotes the consistency of process variables relationships. Bersimis (2007) presented a number of multivariate control charts with a comprehensive explanation on setting the control limits with regard to both indexes.

Usually, principal component analysis (PCA) (Jackson, 1991) is used to obtain the multivariate scores which are linear combinations of the monitored process variables. However, the linear PCA method is not really suitable for reducing the nonlinear multivariate data dimensions as it always ended with a high number of principal components being selected (Zhang et al., 1997). Even though extended algorithms of PCA for dealing with nonlinear processes have been proposed, such as a combination of auto-associative neural network and principal curve (Dong and McAvoy, 1996), the concept is rather complex and its computation is quite demanding.

In addressing the issue, classical multidimensional scaling (CMDS) provides an alternative in compressing multivariate data, where the main reference is defined in terms of dissimilarity measures. However, this dissimilarity measure has to be in the form of variable structure instead of observation correlations, because the fundamental of any process monitoring system should be developed on the basis of variable relationships.

Cox (2001) introduced CMDS, non-metric MDS and also biplots methods as the alternative multivariate techniques for process monitoring which all have been analyzed on a particular gas transportation data. It has been demonstrated that the CMDS algorithms based on the Euclidean dissimilarity scale can construct the same multivariate scores profiles identical to PCA. Other dissimilarity bases such as Mahalanobis and City-block scales were also used where varieties of multivariate configurations have been found accordingly. The main idea of his fault detection technique is by observation on the process samples which are moving away from the main normal cluster. Nevertheless, he did not propose any monitoring statistics. Matheus et al. (2006) developed an on-line CMDS-based process monitoring system by using multiple linear regression (MLR). In their work, MLR is used in relating the original process data with

Copyright held by the International Federation of Automatic Control

the CMDS-based multivariate scores of the normal data. Although the sample trends were used for monitoring, yet, the results were not validated with either T^2 or SPE performance.

This paper presents a new fault detection framework through applying CMDS as the main multivariate data dimensional reduction technique. A new fault detection index (FDI) is proposed accordingly. Procrustes analysis (PA) is used in obtaining the multivariate scores for on-line monitoring. The essential use of PA is to identify the transformation factors which include rotation matrix, compressing or straining scale and translation vector (Borg and Groenen, 1997) between the original normal operating condition (NOC) data and its multivariate scores of variables, in projecting the multivariate profiles for the new measurements.

The paper is organised as follows. Section 2 briefly presents the concept of MDS and the generic outlines of the MDS-based process monitoring procedures. Section 3 discusses on the results demonstrated through application to a simulated continuous stirred tank reactor (CSTR) with recycle system. Section 4 concludes the paper.

2. METHODOLOGIES

2.1 Background of CMDS

Takane (2003) stated that the multivariate points in CMDS are normally arranged in such a way that their distances corresponds to the correlations between the stimuli under study (variables or observations), that is two points are located closely together if their similarity is high, otherwise the distance will be great. Therefore, the main purpose of any MDS algorithms is to measure how well the projected multivariate scores matched as precisely as possible according to the pre-defined dissimilarity scales (Kruskal and Wish, 1978). In this work, the dissimilarity measure have been particularly constructed based on two different scales, Euclidean and city block distances, which are shown respectively by Eq(1) and Eq(2).

Euclidean Distance:

$$\delta_{ij} = \{\sum_a (x_{ia} - x_{ja})^2\}^{\frac{1}{2}} \quad (1)$$

City-block Distance:

$$\delta_{ij} = \sum_a | x_{ia} - x_{ja} | \quad (2)$$

Regarding the proper selection of the reduced number of dimensions, p, Cox and Cox (1994) proposed to use ratio scales on the eigenvalues of the squared dissimilarity matrix (scalar product) in representing the true proportion measure of the corresponded multivariate scores variation in the lower dimensional space as shown in Eq (3).

$$\text{Dimension ratio} = \sum_{i=1}^{p} \lambda_i \Bigg/ \sum_{i=1}^{k-1} \lambda_i \quad (3)$$

Besides, Young and Householder (1938) mentioned that the reduced dimensionality of the multivariate data can be depended either on the rank of **X** or the rank of scalar product of **X**.

2.2 CMDS-based MSPM

In the proposed MSPM framework, the procedures contain two separate phases as shown in Fig. 1.

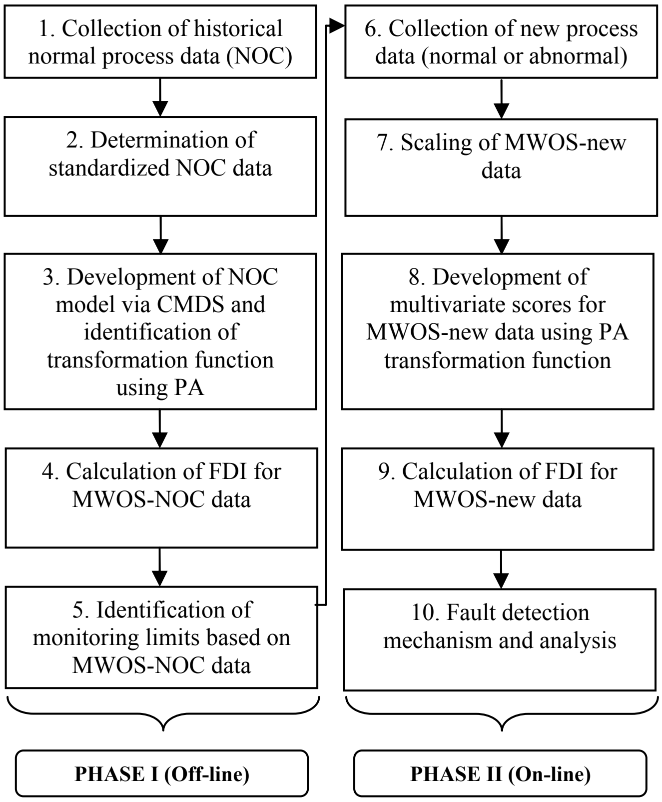

Fig. 1: CMDS-PA-based MSPM frameworks

The first phase concerns with the off-line NOC model building, whereas the second facilitates for on-line process monitoring. A set of NOC data, $\mathbf{X}_{n \times m}$ (n: number of samples and m: number of variables), was determined off-line based on the historical data archive and also scaled to zero mean and unit variance (Steps 1 and 2). In Step 3, CMDS is then applied in order to obtain the variable profiles of NOC following the procedures suggested by Borg and Groenen (1997). Firstly, a squared dissimilarity matrix, Δ^2, with the size of m by m, is determined either by applying Eq(1) or Eq(2) for the Euclidean or city block scale respectively. Then, the double-centering equation is implemented on Δ^2 to obtain $\mathbf{B}_\Delta$ as shown in Eq (4).

$$\mathbf{B}_\Delta = -\frac{1}{2} \mathbf{J}_m \Delta^2 \mathbf{J}_m \quad (4)$$

where $\mathbf{J}_m = (\mathbf{I}_m - \mathbf{1}_m \mathbf{1}_m / m)$, $\mathbf{I}_m$ is an identity matrix, and $\mathbf{1}_m$ is a vector with element of 1 and size m.

Copyright held by the International Federation of Automatic Control

This double-centering operation transforms the dissimilarity matrix into a ratio-scaled matrix, where the origin is relocated onto the centre of the data, thus, a unique configuration of multivariate scores will be obtained accordingly (Torgerson, 1967). Later, $\mathbf{B}_{\Delta}$ is decomposed into $\mathbf{UDU}^{\mathbf{T}}$ where $\mathbf{U}$ is an orthonormal eigenvectors of the double-centred dissimilarity matrix and $\mathbf{D}$ is a diagonal eigenvalues of double-centred dissimilarity matrix such that $d_1 \geq d_2 \geq \ldots \geq d_m$. Finally, the multivariate coordinates of NOC is constructed by using Eq (5).

$$\mathbf{Y}=\mathbf{U}_{+}\mathbf{D}_{+}^{\mathbf{0.5}} \tag{5}$$

where all elements of $\mathbf{D}_{+}^{\mathbf{0.5}}$ and $\mathbf{U}_{+}$ are corresponding to those selected dimensions.

In order to standardize the procedures of projecting the on-line multivariate scores, PA technique is applied which is described as follows (Borg and Groenen, 1997):

i. Computation of the minor product moment between the reconstructed NOC matrix and the modified NOC matrix: $\mathbf{C_{PA}}=\mathbf{Y}^{\mathrm{T}}\mathbf{J_m}\mathbf{X_{mod}}$ where $\mathbf{J_m}$ is from Eq(4) and $\mathbf{X_{mod}}$ is a modified NOC data with size *m* by *p*.
ii. Application of the eigen decomposition on $\mathbf{C_{PA}}$ by way of $\mathbf{C_{PA}}=\mathbf{P_{PA}}\mathbf{V_{PA}}\mathbf{P_{PA}}^{\mathrm{T}}$ where $\mathbf{P_{PA}}$ is a matrix of eigenvectors and $\mathbf{V_{PA}}$ is a matrix of eigenvalues.
iii. Calculation of the optimal rotation matrix, $\mathbf{R} = \mathbf{P_{PA}} \mathbf{P_{PA}}^{\mathrm{T}}$.
iv. Calculation of the optimal dilation scale, $s=(\mathrm{tr}\mathbf{Y}^{\mathrm{T}}\mathbf{JXR})(\mathrm{tr}\mathbf{X}^{\mathrm{T}}\mathbf{JX})$.
v. Calculation of the optimal translation vector, $\mathbf{t}=(\mathbf{Y}-s\mathbf{XR})^{\mathrm{T}}\mathbf{1}/\mathrm{m}$.
vi. The final transformed model of NOC is given by $\mathbf{Y_{PA}}=s\mathbf{X_{mod}}\mathbf{R}+\mathbf{1t}^{\mathrm{T}}$

Transformation factors, $\mathbf{R}$, s and $\mathbf{t}$, emulates the concept of loading factors as in PCA.

In Step 4, the moving windows concept (Kano et al., 2001) is implemented on the NOC data by using the term MWOS (Moving-Window-Observation-Samples)-NOC, $\mathbf{X_{MWOS\text{-}NOC}}$. In particular, this mechanism is operated such that the newly measured sample is added to the data frame by taking the oldest sample out from the data window. In this way, the size of the $\mathbf{X_{MWOS\text{-}NOC}}$ matrix will be maintained at *m* by *p* over the time, especially when a new sample becomes available. This MWOS sample is then applied in Eq (6) to establish the new set of multivariate scores for the NOC data.

$$\mathbf{Y}_{\mathbf{PA(MWOS\ no.}\,k)}=s\mathbf{X}_{\mathbf{(MWOS\ no.}\,k)}\mathbf{R}+\mathbf{1t}^{T} \tag{6}$$

An FDI measuring the change in relationship, C_r, is then defined and it signifies the change in relationships of those monitored variables, as shown in Eq (7).

$$C_r = \sum_{i=1}^{m}\left[v_{new(i)} - v_{NOC(i)}\right]^2 \tag{7}$$

where, v_{new} and v_{NOC} indicate the resultant vector length of the variable scores corresponding to the MWOS-NOC and original NOC data respectively by using Eq(8).

$$v_j = \left[\sum_{i=1}^{p} y_{j(i)}{}^2\right]^{0.5} \tag{8}$$

where y_j is the reconstructed CMDS or PA scores with j=1,2,…,m. The background of Cr is illustrated as in Fig. 2.

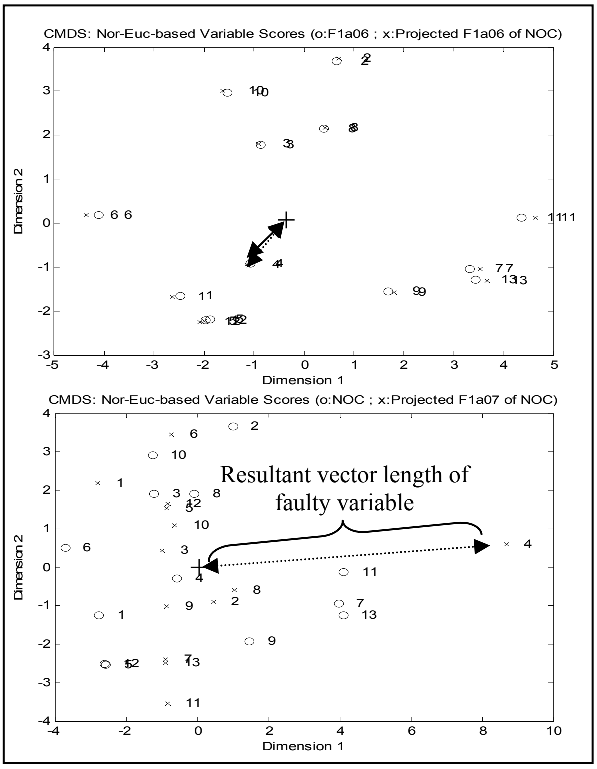

Fig. 2: Dimension plots of CMDS scores of normal vs normal samples (top); dimension plots of CMDS scores of normal vs faulty samples (bottom)

Fig. 2 shows that both resultant vector lengths of variable 4 between the original normal sample (solid line) and the new normal sample (dotted line) from the origin ('+' point) is equivalent to each other (top figure). However, this vector length is substantially increased (dotted line) as a faulty variable occurred in the process as shown in the second plot of Fig. 2. In other words, those faulty variables will move away from the normal cluster whenever a fault condition takes place. Therefore, C_r is actually providing the information of variables correlation consistency in terms of vector length difference between the faulty sample and the original normal model (similar to the SPE concept). Later, the warning (95%) and the control (99%) limits of the monitoring statistics are obtained by assuming that the scores are normally distributed (Step 5).

In on-line monitoring, process data are collected and the newly measured sample is scaled using the means and standard deviations of the NOC data and is then added to the moving window, $\mathbf{X_{MWOS\text{-}new}}$, (Steps 6 and 7). Then, Eq(6) is

Copyright held by the International Federation of Automatic Control

used to obtain the variables profiles in the reduced dimensional space (Step 8), where C_r monitoring statistic is subsequently computed through Eq(7) and Eq(8) (Step 9). Finally, real time monitoring system is carried out by observing the progression of C_r on the monitoring charts and a fault is detected when the index exceeds the 99% monitoring limit for a consecutive number of samples.

3. APPLICATION TO A CSTR WITH RECYCLE

3.1 Case Study

A simulated continuous stirred tank reactor (CSTR) with recycle shown in Fig. 3 was used as the case study. This system operates an irreversible heterogeneous catalytic exothermic reaction in transforming a particular reactant A to product B. Three main control systems have been installed including temperature, level and mixing condition of the vessel in order to sustain the product concentration at a desired setting. In particular, the temperature of the reactor is controlled by manipulating the flow rate of the cold water fed into the heat exchanger through a cascade control system. The flow rate of the product stream is used to control the tank level and the mixing condition is maintained by adjusting the recycle flow rate.

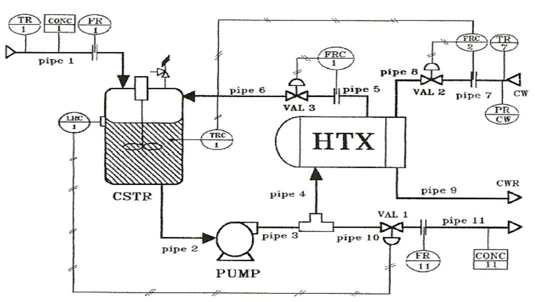

Fig. 3: CSTR System

There are ten on-line measured process variables and three controller outputs have been identified for monitoring as shown in Table 1 (Zhang, 2006).

Table 1. Variables of CSTR system for monitoring

Variables	Variable Names
V1	Tank temperature
V2	Tank level
V3	Flow rate feed
V4	Flow rate inlet
V5	Flow rate cooling
V6	Flow rate outlet
V7	Flow rate recycle
V8	Product concentration
V9	Feed concentration
V10	Tank pressure
V11	Controller 1
V12	Controller 2
V13	Controller 3

3.2 Results

A set of NOC data containing 50 samples was obtained from simulation. The data were standardized to zero mean and unit variance. In the CMDS-PA model, six dimensions were selected to maintain at least 90% of the eigenvalues ratio scale for both Euclidean and city block scale.

A number of faults related to the CSTR system consisting of abrupt fault (magnitude of the faults increased suddenly and maintained over time) and incipient fault (magnitude of the faults increased gradually over time) were then identified as listed in Table 2. In simulating each of those faults, the faulty condition was introduced at sample 2, where the sampling time interval fixed at 4 seconds.

Table 2. Fault list

Fault No.	Descriptions
1	Pipe 1 blockage
2	External feed-reactant flow rate too high
3	Pipe 2 or 3 is blocked or pump fails
4	Pipe 10 or 11 is blocked or control valve 1 fails low
5	External feed-reactant temperature abnormal
6	Control valve 2 fails high
7	Pipe 7, 8, or 9 is blocked or control valve 2 fails low
8	Control valve 1 fails high
9	Pipe 4, 5, or 6 is blocked or control valve 3 fails low
10	Control valve 3 fails high
11	External feed-reactant concentration too low

For the purpose of comparison, a PCA model was also developed for process monitoring based on 6 principal components giving 90% explained data variation. Besides, a CMDS model without PA was also constructed for evaluation. The overall results of Euclidean-scaled CMDS-PA and other models for abrupt fault cases are presented in Table 3, whereas Table 4 summarizes the fault detection performance of city block-scaled CMDS-PA and other models based on incipient fault category.

Table 3. Fault detection time (samples) of CMDS-PA (Euclidean and city block), CMDS and PCA-based MSPM for abrupt and incipient faults based on 99% limit

Abrupt Fault No.	CMDS-PA (Euc and Cit)	CMDS	PCA
1	1	1	1
2	1	1	1
3	1	1	1
4	1	3	1
5	1	1	1
6	1	1	1
7	1	3	1
8	1	1	1
9	1	1	1
10	1	1	1
11	1	1	1

Copyright held by the International Federation of Automatic Control

Table 4. Fault detection time (samples) of CMDS-PA (Euclidean and city block), CMDS and PCA-based MSPM for abrupt and incipient faults based on 99% limit

Incipient Fault No.	CMDS-PA (Euc and Cit)	CMDS	PCA
1	2	3	2
2	4	6	4
3	1	1	1
4	17	23	20
5	8	16	12
6	16	21	20
7	17	19	19
8	15	20	15
9	15	19	17
10	15	19	17
11	4	6	3

The fault detection is defined as the sampling time between a fault being introduced and a monitoring index exceeding its 99% control limit. For instance, if the fault identified at sample 3, then the fault detection time is 1 (3-2). From Table 3, all the process monitoring models including the proposed MDS techniques can detect the abrupt fault cases within 3 sampling time (mostly after 1 sampling time). In the case of incipient fault however (Table 4), all systems took quite longer time in signaling the faults (ranging from 1 to 21).

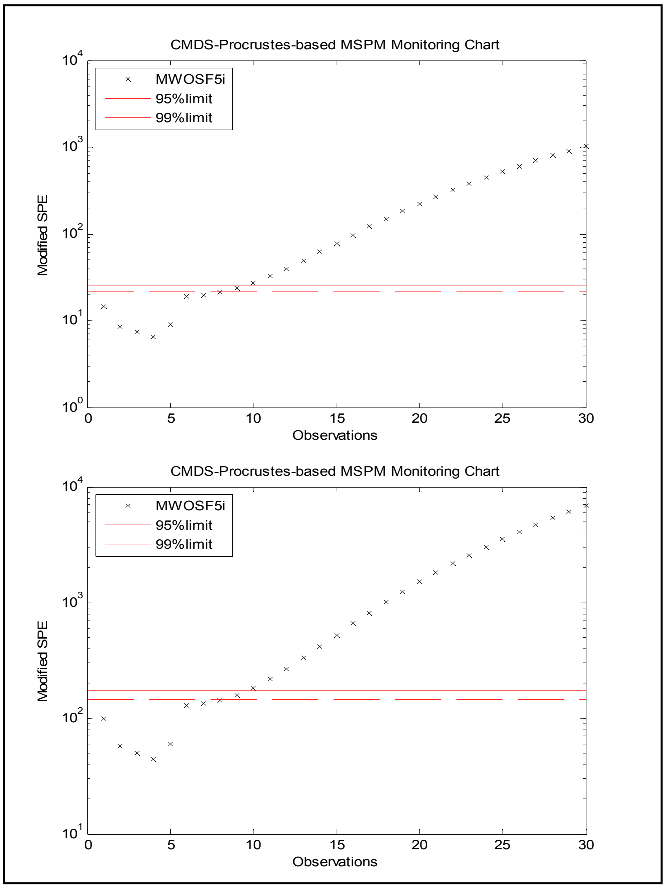

Fig. 4: FDI (C_r) monitoring chart of CMDS-PA for incipient fault number 5 based on Euclidean scale (top) and city block scale (below)

The results in Tables 3 and 4 also indicate that both CMDS-PA-Euclidean and CMDS-PA-City block share the same performance in detecting those specified faults. This is because the FDI in both techniques uniquely having their own control limits respectively as depicted in Fig. 4. Fig. 4 shows the FDI performance of incipient fault number 5, where both CMDS-PA have identified the specified fault at sample 10 (sampling time: 8) but with different 99% monitoring limit values. In compare to the results of standard CMDS, both extended MDS techniques performed significantly better, where they managed to identify the faults much earlier in all of the incipient fault cases and two cases from abrupt fault (fault 4 and 7) . This is due to CMDS-PA used a similar transformation functions (through PA) in projecting the new samples' variable scores as opposed to the CMDS-based system, where different eigenvectors and eigenvalues are used every time for on-line projection. In fact, both of the CMDS-PA systems have shown improved performance with respect to incipient fault number 4, 5, 6, 7, 8, 9 and 10. Other cases of incipient fault showed either both CMDS-PA and PCA have the same performance (fault 1, 2,and 3) or PCA leads by only 1 sampling time detection (fault 11). Nevertheless, these performance variations can be regarded as relatively equivalent as the sampling time interval used is small (4 seconds). Fig. 5 shows the contribution plots of the incipient fault number 5 (abnormal temperature of the reactant from feed stream) based on the CMDS-PA systems.

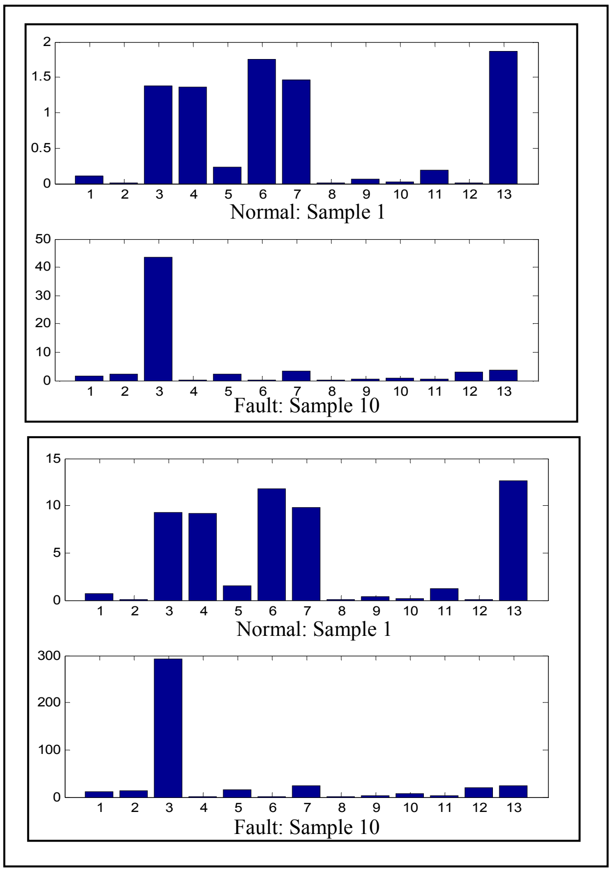

Fig. 5: Contribution plots of CMDS-PA for incipient fault number 5 based on Euclidean scale (top) and city block scale (below)

Copyright held by the International Federation of Automatic Control

In analyzing this particular fault, it was found that the source of the fault is coming from variable number 3. Surprisingly, this phenomenon can be verified by the CMDS-PA, where there was a drastic increased in the variable 3 magnitude on the contribution plots of sample 10 as opposed to the variable 3 value from the normal sample as shown in Fig. 5.

In addition, Fig. 6 illustrates the variable configurations of both CMDS-PA Euclidean and city block scales for incipient fault number 5 based on dimension 1 and 2.

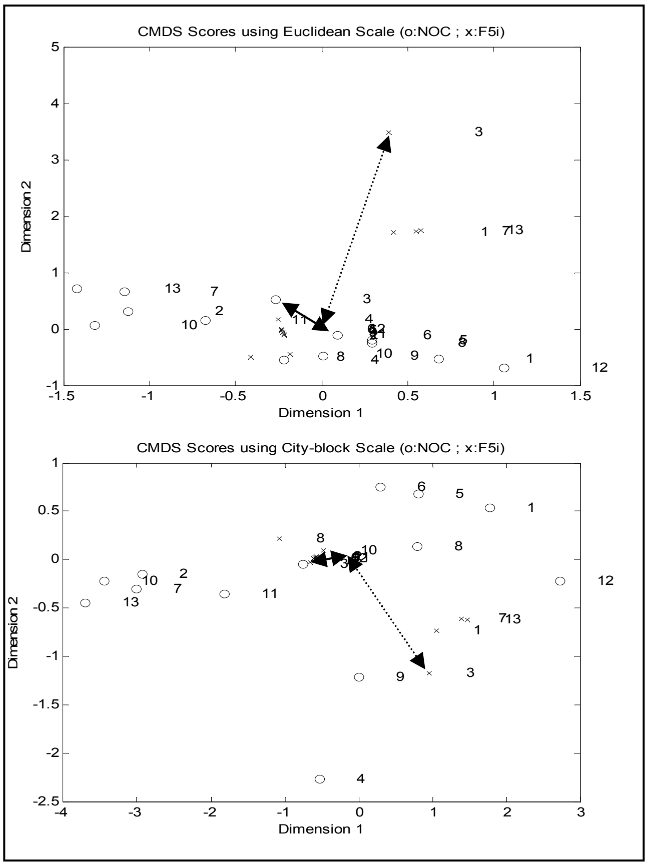

Fig. 6: Configuration plots of CMDS-PA for incipient fault number 5 between sample 1 and sample 12 based on Euclidean scale (top) and city block scale (below)

From Fig. 6, it is obviously shown that the magnitude of the resultant vector length of variable 3 of sample 10 (dotted line) was increased substantially from the origin in compared to the variable vector length of variable 3 of sample 1 (solid line) in both of the CMDS-PA plots. The plots also are able to denote other variables such as variable 1, 7 and 13 of faulty sample seemed to be diverted in great magnitude compared to the normal coordinates. Those were identified as variables which have been affected by the faulty condition of variable 3. Thus, all of these support the fact that CMDS is not only can be utilized alternatively for process monitoring but it also can potentially provide the insight of which variables contribute to the specified fault.

4. CONCLUSIONS

A MSPM framework based on CMDS and PA techniques is proposed. In particular, a new FDI has been introduced, which is derived from the multivariate scores in terms of variable structure. The overall performance of these newly MSPM systems was found to be relatively similar compared to the linear PCA-based MSPM method, which has been demonstrated by a case study on a CSTR system. This proves that the proposed methods can potentially be used as an alternative for process monitoring. Further works is on-going in compressing the multivariate data much lower by using Non-metric MDS technique.

ACKNOWLEDGEMENT

The authors would like to acknowledge Newcastle University UK for providing the funding and facilities for this research. A special appreciation also goes to the Ministry of Higher Education Malaysia and Universiti Malaysia Pahang for sponsoring the study.

REFERENCES

Bersimis, S., Psarakis, S., and Panaretos, J. (2007). Multivariate Statistical Process Control Charts: An Overview. *Quality and Reliability Engineering International,* 23, 517-543.

Borg, I., and Groenen, P. (1997). *Modern Multidimensional Scaling: Theory and Applications*. Springler-Verlag. New York, USA.

Cox, T.F. (2001). *Multidimensional Scaling used in Multivariate Statistical Process Control,* Journal of Applied Statistics, 28, 365-378.

Cox, T.F., and Cox M.A.A. (1994). *Multidimensional Scaling.* Chapman & Hall. London, Great Britain.

Dong, D., and McAvoy, T.J. (1996). Nonlinear Principal Component Analysis-Based On Principal Curves and Neural Networks. *Computer and Chemical Engineering,* 20, 65-78.

Jackson, J.E. (1991). *A User's Guide To Principal Components.* John Wiley and Sons. USA.

Kano, M., Hasebe, S., Hashimoto, I., and Ohno, H. (2001). A New Multivariate Statistical Process Monitoring Method Using Principal Component Analysis. *Computers and Chemical Engineering.* 25, 1103-1113.

Kruskal, J.B., and Wish, M. (1978). *Multidimensional Scaling.* SAGE Publications. California, USA.

MacGregor, J. F., and Kourti, T. (1995). Statistical Process Control of Multivariate Processes. *Control Engineering Practice,* 3, 403 – 414.

Martin., E.B., Morris, A.J., and Zhang, J. (1996). Process Performance Monitoring Using Multivariate Statistical Process Control. *Systems Engineering for Automation,* IEEE Proceedings.

Matheus, J., Dourado, A., Henriques, J., Antonio, M., Nogueira, D. (2006). Iterative Multidimensional Scaling for Industrial Process Monitoring. *Systems, Man, and Cybernetics,* 2006 IEEE International Conference.

Qin, S.J. (2003). Statistical Process Monitoring: Basics and Beyond. *Journal of Chemometrics,* 17, 480-502.

Takane, Y. (2003), Matrices with Special Reference to Applications in Psychometrics. *Linear Algebra and Its Applications,* 388, 341-361.

Torgerson, W.S. (1967). *Theory and Methods of Scaling,* John Wiley & Sons.USA.

Young, G., and Householder, A.S. (1938). Discussion of a Set of Points in terms of Their Mutual Distances. *Psychometrika,* 3 (1), 19-22.

Zhang, J., (2006). Improved On-line Process Fault Diagnosis Through Information Fusion in Multiple Neural Networks. *Computers and Chemical Engineering,,* 30, 558-571.

Zhang, J., Martin, E.B., Morris, A.J. (1997). Process Monitoring Using Non-linear Statistical Techniques. *Chemical Engineering Journal,* 67, 181-189.

Copyright held by the International Federation of Automatic Control