A STUDY ON THE SELECTION OF MODEL DIMENSIONS AND SENSITIVITY OF PCA-BASED FAULT DETECTION

Masayuki Tamura * **and Shinsuke Tsujita** *

* *Technical Research Institute, Tokyo Gas Co., Ltd.*
16-25 Shibaura 1-Chome, Minato-ku, Tokyo 105-0023, Japan
m-tamura@tokyo-gas.co.jp

Abstract: A method of determining model dimensions (number of principal components) which maximize the sensitivity of fault detection was studied. In this paper it is shown that the sensitivity of PCA-based fault detection generally depends on the number of principal components. Most of the existing methods give only one value as a recommended number of components, and so sometimes the sensitivity is poor for certain kinds of faults. Among existing methods, although the Variance of Reconstruction Error (VRE) criterion gives a recommended model dimension which depends on the kind of fault, it is not intended to maximize sensitivity. This paper presents a new method of determining the model dimension which maximizes the sensitivity of PCA-based fault detection.

Keywords: Fault detection, Principal component analysis, Model dimension, Number of components

1. INTRODUCTION

Principal component analysis (PCA) is one of the most intensively studied methods for fault detection of plants (Kresta *et al.*, 1991; MacGregor *et al.*, 1994; Kourti, 2002). The selection of model dimensions (number of components) is an ambiguous problem in PCA model building. There have been various comparative studies on methods for determining model dimensions for fault detection. Himes et al. (Himes *et al.*, 1994) compared four methods using the Tennessee Eastman simulation. Valle et al. (Valle *et al.*, 1999) tested 11 methods and concluded that the VRE criterion is preferred. Most of the previous studies use the concept that the best solution is the number of components which is the "true" dimension of the system, which is usually unknown. In this paper, we focus on the sensitivity to fault, which is one of the most important issues in fault detection. Both the Q-statistic (or SPE: Squared prediction error) and T^2-statistic which are used for fault detection, show different behaviors under different model dimensions, and the model dimensions that give the maximum sensitivity of fault detection depend on the kind of fault, as will be shown in this paper. We propose determining the model dimension based on the signal to noise ratio (SNR) of fault detection.

2. BACKGROUND

Data matrix $\mathbf{X}$, which consists of data under the normal operating conditions of a plant, is represented by a PCA model in a lower dimension a as follows.

$$\mathbf{X} = \mathbf{TP}' + \mathbf{E} = \sum_{h=1}^{a} \mathbf{t_h p'_h} + \mathbf{E} \qquad (1)$$

$\mathbf{T}$, $\mathbf{P}$ and $\mathbf{E}$ are score, loading and error matrices with appropriate dimensions respectively. The first term of the right-hand side represents a dimensional model subspace (or representative space), and the second term represents $m-a$ dimensional residual subspace which cannot be expressed by the PCA model. (m is the number of sensors.)

Operation data are analyzed by the constructed model. One of the indices used for fault detection is the Q-statistic:

$$
\begin{aligned}
Q &= \sum_{j=1}^{m}(x_{newj} - \hat{x}_{newj})^2 \qquad (2)\\
&= ||\mathbf{x}'_{new} - \mathbf{x}'_{new}\mathbf{P}_a\mathbf{P}'_a||^2 = ||\mathbf{x}'_{new}\tilde{\mathbf{P}}_a\tilde{\mathbf{P}}'_a||^2
\end{aligned}
$$

$\mathbf{x}_{new}$ is a new measured vector to be diagnosed. $\mathbf{P}_a$ is an $m \times a$ matrix which consists of first a columns of the loading matrix $\mathbf{P}$, and $\tilde{\mathbf{P}}_a$ is an $m \times (m-a)$ matrix which consists of the remaining columns of $\mathbf{P}$. If this statistic shows an unexpectedly large value, it means that the data point went into the residual subspace. This means that correlations between sensors which were observed during normal operation conditions were lost. The other index used for fault detection is Hotelling's T^2-statistic:

$$
T^2 = \mathbf{t}\mathbf{S}^{-1}\mathbf{t}' = \mathbf{x}'_{new}\mathbf{P}_a\mathbf{S}^{-1}\mathbf{P}'_a\mathbf{x}_{new} = \sum_{h=1}^{a}\frac{t_h^2}{s_{t_h}^2} \qquad (3)
$$

$\mathbf{S}$ is a diagonal matrix, which is the estimated covariance matrix of the principal component scores, and $\mathbf{t}$ is the vector which contains the scores at the time of observation. This indicates deviation from normal values inside a dimensional representative space. That is, sensors show unexpected values but the linear relations among them do not necessarily change in this case.

3. DETERMINATION OF MODEL DIMENSION

3.1 Existing methods

Generally, each sensor output includes noise, and so the determination of model dimension a is not straightforward. There are known criteria used in existing methods such as the scree plot, eigenvalue limit, cumulative contribution limit, cross validation, variance of reconstruction error (VRE) criterion, and so on.

The scree plot and eigenvalue limit are based on the concept that components which have small eigenvalues are not important for modeling. In the cumulative contribution limit method, the minimum model dimension which can express, for example, 80% of the variance of the data is selected. Cross validation (Wold, 1978) is a method which uses a part of the training samples for model construction; the remaining samples are compared with the prediction by the model, and when PRESS (prediction residual sum of squares) becomes smaller, the new component is added to the model. All of these methods are popular and detailed descriptions can be found in the literature (Jackson, 1991; Zwick and Velicer, 1986). The VRE criterion was proposed for minimizing the error of fault "reconstruction"; fault reconstruction is a procedure for estimating the true state of the system without a faulty sensor (Dunia and Qin, 1998).

3.2 Proposed method

An observed vector during operation with a fault is written as the vector sum of the normal data and fault vector:

$$
\mathbf{x}_f = \mathbf{x} + \mathbf{f} \qquad (4)
$$

In order to determine the SNR of fault detection by monitoring the Q-statistic, it is reasonable to consider the norm of a projected vector of $\mathbf{x}_f$ onto residual subspace as a signal. It is also natural to regard the statistical control limit as SNR noise. The control limit of the Q-statistic which is shown in a reference (Jackson and Mudholkar, 1979) can be used, or simply a certain percentile of the Q-statistic of sample data (normal condition data) can be used. The SNR for Q-statistic monitoring is defined as:

$$
SNR_Q = \frac{||\mathbf{x}'_f\tilde{\mathbf{P}}_a\tilde{\mathbf{P}}'_a||^2}{Q_a} \qquad (5)
$$

The denominator Q_a is a statistical control limit of the Q-statistic. For monitoring using the T^2-statistic, the SNR can be defined in a similar way:

$$
SNR_{T^2} = \frac{\mathbf{x}'_f\mathbf{P}_a\mathbf{S}^{-1}\mathbf{P}'_a\mathbf{x}_f}{\frac{a(n-1)}{n-a}F_{a,n-a;\alpha}} \qquad (6)
$$

The denominator is an approximation of a control limit of the T^2-statistic. n is the number of samples. $F_{a,n-a;\alpha}$ is $100 \times (1-\alpha)$ percentage point (95% for example) of the F-distribution.

Note that the numerators of both SNR depend on the direction of the fault vector, whereas their denominators are independent of the kind of fault. We propose calculating the model dimension which maximizes these SNR. The recommended dimension will be different for Q and T^2, and it depends on the kind of fault, too.

In order to calculate these SNR, $\mathbf{x}_f$ is needed. In the case of simple sensor faults, fault vector $\mathbf{f}$ is easily determined. If the i-th sensor is faulty, the value of $\mathbf{f}$ is determined as follows:

$$
\mathbf{f} = [0, \cdots, 0, \underset{\uparrow\, i\text{th}}{I}, 0, \cdots, 0] \qquad (7)
$$

This can be added to normal condition data $\mathbf{x}$. In the case of complicated process faults, a reference dataset which includes the fault is needed to obtain $\mathbf{x}_f$.

This means that the proposed method is based on a priori knowledge of the fault if it is a process fault. This is a drawback because usually one of the advantages of PCA in fault detection is that the model can be constructed without information of abnormal situations. However, if a priori information of data of abnormal situations exists, this method is reasonably useful to increase the sensitivity. From another viewpoint, a speculation can be derived from this method. Namely, as far as the model dimension which maximizes the sensitivity of fault detection depends on the kinds of faults generally, models with various (desirably all) dimensions should be monitored if computational load allows.

If there are limitations such as computational load, more than one kind of fault may have to be detected by a single PCA model. In this case the following criterion is proposed. Based on many experiments and simulations, the authors found that there exists a model dimension a_c where SNR_Q takes its maximum and the sensitivity of the Q-statistic for fault detection becomes significantly worse when a model dimension larger than a_c is employed. In contrast, the T^2-statistic shows maximum sensitivity at higher order than $a = a_c$, and selection of fewer components makes the model insensitive to the fault. This is thought to be due to a principal component whose direction is close to $\mathbf{f}$, which is taken into the representative space at a dimension $a_c + 1$ when the dimension of the model is increased. A criterion can be considered based on this characteristic. A model dimension a_c is calculated as:

$$a_c = \underset{a}{\operatorname{argmax}} \left(\sum_i w_i SNR_Q \right) \tag{8}$$

where i is a suffix which denotes the kind of fault and w_i is a positive weight which expresses the significance of the fault. This expression is based on the same concept as the formula which had been proposed for the VRE criterion (Dunia and Qin, 1998). In the case of the SNR criterion proposed here, a pair of PCA models whose model dimension is a_c and $a_c + 1$ are employed and the former is used for Q-statistic monitoring and the latter for T^2-statistic monitoring. This is because the principal component whose direction is close to $\mathbf{f}$ is in the residual subspace at $a = a_c$, and is in the representative space at $a = a_c + 1$. As a result, the sensitivities of the $Q-$statistic and T^2-statistic are high at $a = a_c$ and $a_c + 1$, respectively.

4. RESULT: THE TENNESSEE EASTMAN PLANT SIMULATION

The Tennessee Eastman plant simulation is a plant-wide process control problem developed by Downs and Vogel. The process consists of five major units (a reactor, a product condenser, a vapor-liquid separator, a recycle compressor and a product stripper), as described in the reference (Downs and Vogel, 1993). In this study, the simulation conditions were as follows:

- Sampling interval: 3 minutes
- Data length: 500 points (for each dataset)
- Control: The cascade control system proposed by McAvoy and Ye (McAvoy and Ye, 1994)
- Sensors used for PCA modeling: 16 sensors proposed by Chen and McAvoy (Chen and McAvoy, 1998)
- Fault occurrence time: 1 hour (20 points) after the start of simulations
- Random seed: Changed for every simulation run

4.1 *Definition of sensitivity and fault detection*

In this paper, the sensitivity of fault detection is defined as follows. First, statistical control limits with 95% confidence limit for each PCA model are calculated. Then the Q-statistic and T^2-statistic of test data are calculated and divided by each control limit. Therefore, the value unity means that the statistic is equal to the control limit. In the following analysis we defined that the fault is detected when any of the PCA models has shown six succeeding values larger than one. (The value “six” was determined empirically to suppress false alarms, and has no significant meaning itself.)

4.2 *Sensor faults*

Three kinds of sensor faults are studied here: offset of reactor inlet flow sensor (XMEAS 6), drift of separator pressure sensor (XMEAS 13) and random noise of stripper pressure sensor (XMEAS 16), denoted by SF 1, 2 and 3, respectively. These three sensors are selected because they are not used in the control loop. Note that the use of faulty sensors for feedback control causes complicated dynamics of the system and the faults cannot then be treated as simple sensor faults.

First, a PCA loading matrix was calculated using training data under normal conditions. The scree plot is shown in Fig. 1. It implies that three components should be retained. The numbers of components retained by existing methods are listed in Table 1. The eigenvalue limit method suggests that five components should be retained and cumulative R^2 exceeds 80% with eight components. The cross validation method selected only two components in this case.

Subsequently, SNR were calculated using $\mathbf{f}$ vectors in the form of Eq. 7 for each fault. Model dimensions which maximize SNR_Q and SNR_{T^2} are shown in Table 2, as well as the dimension of the VRE criterion selected.

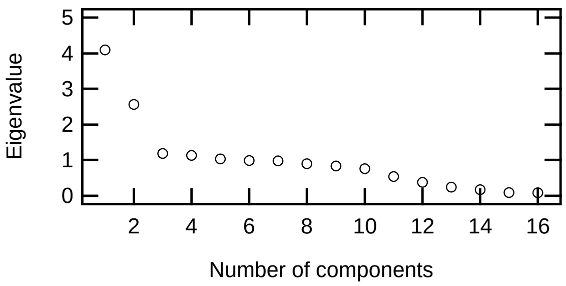

Fig. 1. Scree plot.

Table 1. Model dimensions selected by existing methods

Scree	Eigenvalue limit	Cumulative R^2	Cross validation
3	5	8	2

Table 2. Model dimensions selected by VRE and SNR criteria

Fault	VRE	SNR_Q	SNR_{T^2}
SF1	2	4	8
SF2	11	12	14
SF3	12	13	14

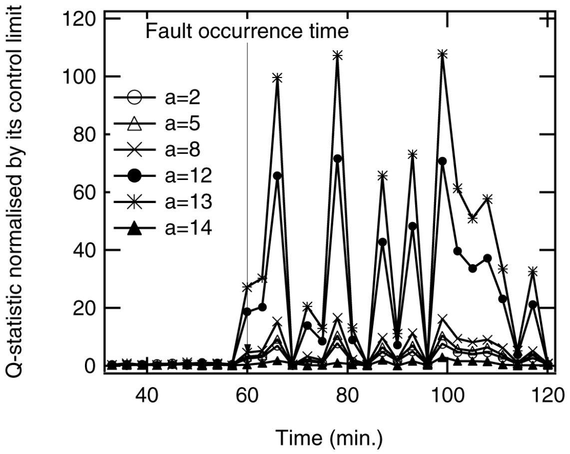

Fig. 2. Q-statistic normalized by its control limits for simulation case SF 3 (random noise of stripper pressure sensor (XMEAS 16)).

An example of the time series plot of sensitivities (Q-statistic and T^2-statistic normalized by their control limit) in the case of SF 3 are shown in Figures 2 and 3, respectively. These values are from a test by data with different random seed (independent data from the samples for modeling and dimension selection procedure). As shown in these figures, models with dimensions $a = 13$ for Q-statistic monitoring and $a = 14$ for T^2-statistic monitoring, which were determined by the SNR criterion, presented the highest sensitivities. The SNR criterion was also the best for the other two sensor faults in terms of the sensitivity.

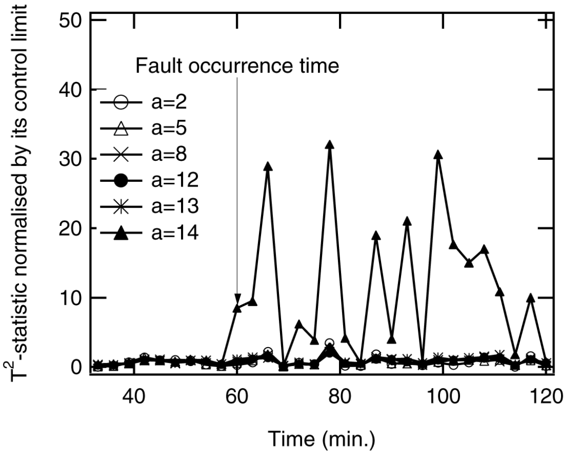

Fig. 3. T^2-statistic normalized by its control limits for simulation case SF 3 (random noise of stripper pressure sensor (XMEAS 16)).

4.3 *Process faults*

Disturbances prepared in the original Tennessee Eastman simulation code (the disturbances are denoted by "IDV" in their program) were used for process fault investigations. First, a PCA model was calculated by training data under normal conditions. The model dimension is tentatively determined by the eigenvalue limit method (here it was $a = 5$). Second, training data which contain each fault were used to obtain a reference for fault directions. Ten points from five points before the point where the fault was first detected were used as $\mathbf{x}_f$. SNR were calculated using these reference data. Note that the proposed method needs only a few samples to know the direction of the fault vector.

Using independent test data (with different random seeds), sensitivities were investigated. Figures 4 and 5 show time series plots of sensitivities for simulation case IDV 10 (random variation of feed temperature of reactant C). Sensitivity to the fault depends on the model dimensions and in this case $a = 14$ was maximum for Q-statistic monitoring and $a = 15$ was maximum for T^2-statistic monitoring, which coincides with the result of dimension selection by the SNR criterion.

Results of other investigated faults are not shown here due to lack of space, however, most of the cases showed that the SNR criterion is the best in terms of the sensitivity.

In case all the faults should be detected by a single PCA model, Eq. 8 is used for model selection. Assuming that all w_is are equal, Eq. 8 gave $a_c = 11$. On the other hand, the VRE criterion gave 2 as a recommended model dimension, which is because of the existence of fault SF 1. Figures 6 and 7 show time series plots of normalized Q-statistic and T^2-statistic. The fault vector of SF 1 is included in the representative space almost perfectly until $a = 10$, and the sensitivities for higher dimension models become poor. This makes u_i, which is an index to be minimized

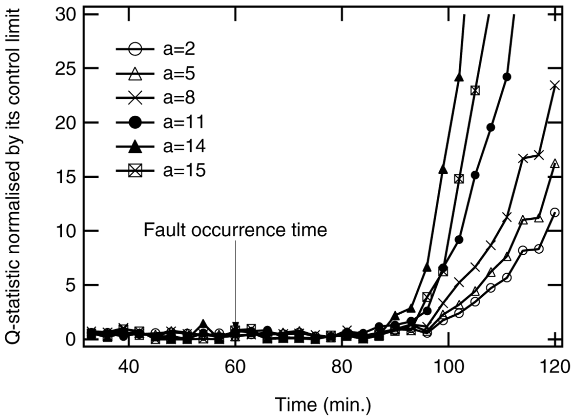

Fig. 4. Q-statistic normalized by its control limits for simulation case IDV 10 (random variation of feed temperature of reactant C).

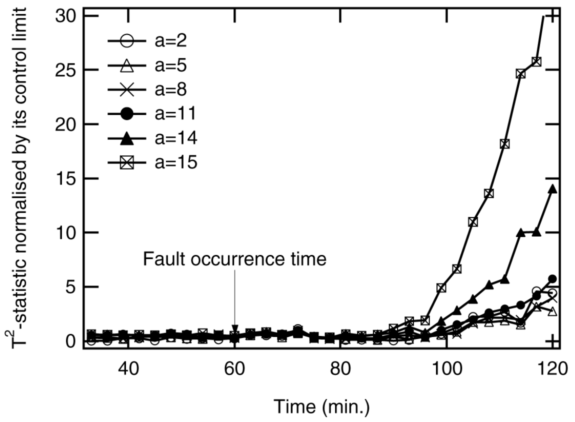

Fig. 5. T^2-statistic normalized by its control limits for simulation case IDV 10 (random variation of feed temperature of reactant C).

for dimension selection (Dunia and Qin, 1998), quite large for large model dimensions. This greatly affects the result and the VRE criterion recommends $a = 2$, which is quite insensitive to other faults.

If only Q-statistic monitoring is performed, the sensitivity for SF 1 is poor when $a = 11$, which is recommended by the method proposed in this paper, is selected. However, if T^2-statistic monitoring with the $a = 12$ PCA model is performed simultaneously as suggested, the sensitivity for SF 1 is also assured (see Figures 6 and 7). T^2-statistic monitoring and Q-statistic monitoring thus work cooperatively with each other.

It is interesting to examine the results of previous studies and compare the behavior of the Q-statistic and T^2-statistic. Himes et al. (Himes *et al.*, 1994) analyzed their data from the Tennessee Eastman plant. They analyzed in a different way and used all the 51 sensors in the simulation. According to their result, Q-control gave the fastest detection of a fault for models with 28–30 components, whereas the T^2-control gave the fastest detection for models with 34 or more components. For another fault, the fastest detection was given

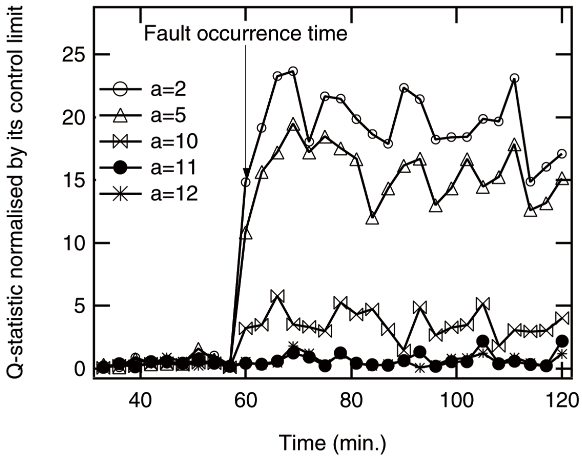

Fig. 6. Q-statistic normalized by its control limits for simulation case SF1 (offset of reactor inlet flow sensor (XMEAS 6)).

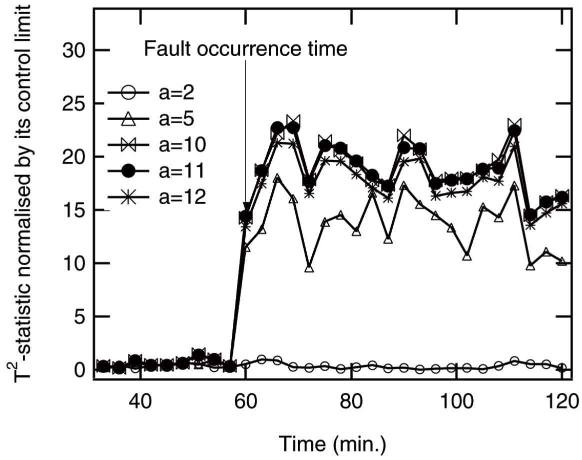

Fig. 7. T^2-statistic normalized by its control limits for simulation case SF1 (offset of reactor inlet flow sensor (XMEAS 6)).

for models with 28–30 components for Q-control and 32–42 components for T^2-control. Their results are consistent with the characteristics identified in this paper.

4.4 *Consideration of false alarm*

False alarms are another important point in fault detection, but were not considered in the analysis given above. Generally, false alarms are expected to increase when many components are retained because the model may include noise which yields large prediction errors. We analyzed the frequency of occurrence and magnitude of false alarms using independent normal data. The data used for this analysis consisted of 10000 points, which corresponds to 500 hours of normal operation of the plant. The result was as expected. The frequency of false alarms (type I error) becomes about 40% for several models with high dimensions although they were designed to show 5% type I error. However, care must be taken when assessing the method presented in this paper with this data. As far as the method is designed to maximize the sensitivity, the important point is the magnitude of false alarms. If the false alarms are not large, it is possible to avoid them by setting the control limit slightly higher. Actually, in our simulation result it was enough to set the control limit at only 2–3 times higher to suppress type I errors below 5%. If the prediction error shows a normal distribution, we can estimate how the set point of the control limit should be changed. If the observed false alarms are 40%, it can be expected from statistical tables of standard normal distributions that setting the control limit at 2–3 times higher makes type I errors about 5 %. In our case, the advantage of selecting the model dimension by the method based on SNR was not lost because the variation of the sensitivity in different model dimensions was much larger than the magnitude of false alarms. This point needs further investigations for generalization.

5. CONCLUSIONS

A new method for model dimension selection of the PCA model based on the sensitivity of fault detection was proposed. The proposed method is based on signal to noise ratios which are defined using the Q-statistic and T^2-statistic, respectively. The method was validated by the Tennessee Eastman plant simulation and the advantage of the new method in terms of sensitivity to faults was shown. On the other hand, the method does not maximize fault reconstruction reliability. It is recommended to use this method and the VRE criterion depending on the purpose of analysis because each method has respective strengths and limitations. An important difference with other existing methods is that existing methods except for the VRE criterion are based on the concept of estimation of the "true" dimension of the system. This is natural in view of the fact that PCA is a tool to explain the hidden mechanism of the system by computation of correlated data. However, our goal of using PCA for fault detection is slightly different and the sensitivity to faults is one of the most important points. Although it does not use a "true" system description, our method is effective in terms of sensitivity. The method uses a priori knowledge of faults to determine SNR. Although this is a drawback, this approach is useful to increase the sensitivity of the method when a priori information of data with faults exists.

REFERENCES

Chen, G. and T. J. McAvoy (1998). Predictive on-line monitoring of continuous processes. *J. Proc. Cont.* **8**, 409–420.

Downs, J. J. and E. F. Vogel (1993). A plant-wide industrial process control problem. *Computers. Chem. Engng.* **17**, 245–255.

Dunia, A. and S. J. Qin (1998). Joint diagnosis of process and sensor faults using principal componenent analysis. *Control Eng. Pract.* **6**, 457–469.

Himes, D. M., R. H. Storer and C. Georgakis (1994). Determination of the number of principal components for disturbance detection and isolation. *Proc. American Control Conference* pp. 1279–1283.

Jackson, J. E. (1991). *A User's Guide to Principal Components*. John Wiley & Sons, Inc.. Hoboken, New Jersey.

Jackson, J. E. and G. S. Mudholkar (1979). Control procedures for residuals associated with principal component analysis.. *Technometrocs* **21**, 341–349.

Kourti, T. (2002). Process analysis and abnormal situation detection: From theory to practice. *IEEE Control Systems Magazine* **22 (5)**, 10–25.

Kresta, J. V., J. F. MacGregor and T. E. Marlin (1991). Multivariative statistical monitoring of process operating performance. *Can. J. Chem. Eng.* **69**, 35–47.

MacGregor, J. F., C. Jaeckle, C. Kiparissides and M. Koutoudi (1994). Process monitoring and diagnosis by multiblock PLS methods. *AIChE J.* **40**, 826–838.

McAvoy, T. J. and N. Ye (1994). Base control for the tennessee eastman problem. *Computers Chem. Engng.* **18**, 383–413.

Valle, S., W. Li and S. J. Qin (1999). Selection of the number of principal components: The variance of the reconstruction eror criterion with a comparison to other methods. *Ind. Eng. Chem. Res.* **38**, 4389–4401.

Wold, S. (1978). Cross-validatory estimation of the number of components in factor and principal components models. *Technometrics* **20**, 397–405.

Zwick, W. R. and W. F. Velicer (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bull.* **99**, 432–442.