On-Line Fault Detection and Classification for a Compressor Process in the Air Separation Plant

DING-SOU CHEN* and JIALIN LIU**

**New Materials Research and Development Department*
China Steel Corporation
Hsiao Kang, Kaohsiung 81233, Taiwan, R.O.C.
***Department of Information Management*
Fortune Institute of Technology, Kaohsiung 83160, Taiwan, R.O.C.

A data-driven model, distanced-based fuzzy c-means (DFCM), is proposed in this paper for the on-line monitoring of a process with high-dimensional variables, outliers, and time-varying characteristics. In DFCM, principal component analysis (PCA) is used to eliminate collinearity between the process variables. After that, fuzzy rules are generated to reject outliers and cluster the compressed data. When a new event emerges, an adaptive PCA algorithm is utilized to accommodate the new event data and transfer the known event rules to the new PCA subspace without recalculating the trained data. Therefore, only new event data need to be clustered on the new subspace. Owing to the air compressor being the most important and energy consuming units in an air separation plant, the proposed approach has been applied to monitor an air compressor and the results show the challenges of process monitoring can be effectively dealt with.

1. INTRODUCTION

For the purpose of supplying the needs of high purity oxygen, CSC runs some stand-alone air separation plants. The process input is air, coming from the atmosphere, and the products are oxygen, nitrogen and argon. Figure 1 shows a brief flow diagram for an air separation plant, in which the process can be roughly split into two steps. Before entering distillation towers (cold box), air purification has to be carried out in order to remove unwanted ingredients such as, dust, moisture, CO_2, and hydrocarbons, other than the three products, oxygen, nitrogen and argon.

During the purification procedures, the air compressor is one of the most important and energy consuming units. If the output pressure of compressor cannot meet the specification, then the air will become more difficult to liquefy, with the result that the air cannot be separated effectively in the cold box. In addition, the product quantity of the plant depends directly on the quantity of air that the compressor can handle. Therefore, the compressor process monitoring system is the focus of this paper.

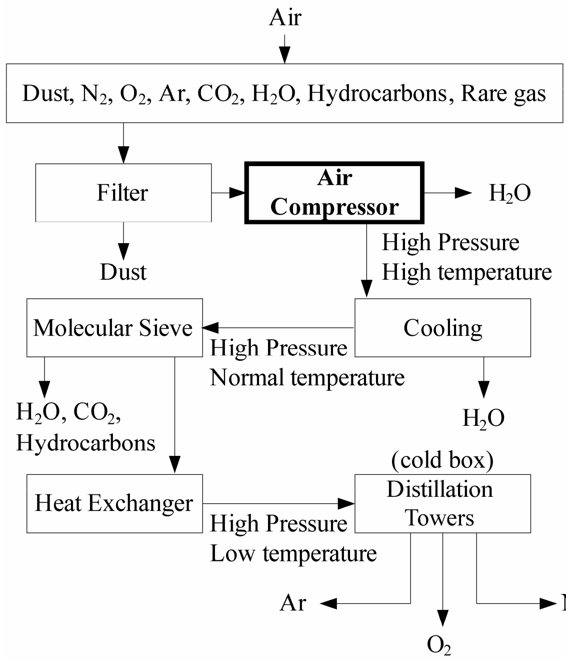

Fig. 1. Brief flow diagram of air separation plant.

2. PROPOSED APPROACH

The proposed approach consists of three phases, comprising (1) the off-line modelling phase, (2) the on-line monitoring and updating phase, and (3) the off-line updating phase, to monitor a process with collinearity, outliers, and time-varying characteristics. In the off-line modelling phase, a distance-based FCM

(DFCM) algorithm has been applied to the score vectors instead of raw data. This approach is not only capable of reducing variable dimensions effectively, but also of gradually discarding outliers to reach a feasible clustering result.

In the on-line monitoring and updating phase, the statistics Q and T^2 of on-line data are examined with their control limits. If any one of the statistics is out of its control limits, alarms should be triggered and the data stored for next off-line model update. On the other hand, the Mahalanobis distances of score vectors to each cluster centre are evaluated with the corresponding cluster boundary distance when the on-line data belongs to the PCA subspace. If any one of the Mahalanobis distances is less than the corresponding cluster boundary, the data point is part of known event groups, and it can then be classified into the known groups according to its membership values whereas the cluster parameters also are updated. Otherwise, the data point is an outlier, even though it is within the PCA subspace, alarms should be triggered and the outlier collected for the next model update.

In the off-line updating phase, the PCA subspace is adapted by using new event data and a recursive PCA algorithm in order to account for known and new events without recalculating the trained data. After that, the cluster parameters on the previous subspace have to be transferred to the new one. A concept of rotating and shifting coordinates of subspace has been applied to transfer the cluster parameters. Therefore, only new event data have to be clustered on the new subspace.

2.1 Off-line Modelling

Consider the data matrix $\mathbf{W} \in R^{m\times n}$ with m rows of observations and n columns of variables. Each column is normalized to zero means and unit variances: $\mathbf{X} = \left(\mathbf{W} - \mathbf{1}\overline{\mathbf{W}}\right)\mathbf{S}^{-1}$ where $\overline{\mathbf{W}}$ is a mean vector, **1** is a column vector in which all elements are one, and S is a diagonal matrix of standard deviations. The eigenvectors (**P**) of the covariance matrix can be obtained from the normalized dataset. The score vectors are the projection of the data matrix **X** to each eigenvector. The data matrix **X** can be decomposed as:

$$\mathbf{X} = \sum_{i=1}^{k} \mathbf{t}_i \mathbf{p}_i^{\mathrm{T}} + \sum_{i=k+1}^{n} \mathbf{t}_i \mathbf{p}_i^{\mathrm{T}} \quad \ldots\ldots(1)$$

The first term of right side of the above equation is the systematic part, which is described by the PCA subspace with the first k eigenvectors, and the second term is remainder of **X** that is orthogonal to the subspace.[1]

The objective of fuzzy clustering is to partition the dataset **T** into c clusters with vague boundaries. The fuzzy c-means (FCM) algorithm is as follows:

(1) Randomly initialize the degrees of membership following the constraints:

$$u_{ij} \in [0,1], \quad 1 \le i \le c, \quad 1 \le j \le m \quad \ldots\ldots(2)$$

$$\sum_{i=1}^{c} u_{ij} = 1, \quad 1 \le j \le m, \qquad 0 < \sum_{j=1}^{m} u_{ij} < m, \quad 1 \le i \le c$$

(2) Compute the cluster centers and covariances:

$$\boldsymbol{\mu}_i^{(k)} = \sum_{j=1}^{m} \left[u_{ij}^{(k-1)}\right]^q \mathbf{t}_j \Big/ \sum_{j=1}^{m} \left[u_{ij}^{(k-1)}\right]^q, \quad i = 1 \ldots c \quad \ldots\ldots(3)$$

$$\boldsymbol{\Sigma}_i^{(k)} = \sum_{j=1}^{m} \left[u_{ij}^{(k-1)}\right]^q \left(\mathbf{t}_j - \boldsymbol{\mu}_i^{(k)}\right)^{\mathrm{T}} \left(\mathbf{t}_j - \boldsymbol{\mu}_i^{(k)}\right) \Big/ \sum_{j=1}^{m} \left[u_{ij}^{(k-1)}\right]^q \quad \ldots(4)$$

(3) Compute the distances and update the degrees of membership:

$$D_{ij,\boldsymbol{\Sigma}_i}^{2(k)} = \left\|\boldsymbol{\Sigma}_i^{(k)}\right\|^{1/n} \left(\mathbf{t}_j - \boldsymbol{\mu}_i^{(k)}\right) \boldsymbol{\Sigma}_i^{-1(k)} \left(\mathbf{t}_j - \boldsymbol{\mu}_i^{(k)}\right)^{\mathrm{T}} \quad \ldots\ldots(5)$$

$$u_{ij}^{(k)} = 1 \Big/ \sum_{l=1}^{c} \left(D_{ij,\boldsymbol{\Sigma}_i}^{2(k)} \Big/ D_{lj,\boldsymbol{\Sigma}_l}^{2(k)}\right)^{2/(q-1)}, \quad i = 1 \ldots c, \quad j = 1 \ldots m \quad \ldots(6)$$

(4) If the norm of membership changes is larger than the predefined tolerance (ε), i.e., $\sum_{j=1}^{m}\sum_{i=1}^{c}\left\|u_{ij}^{(k)} - u_{ij}^{(k-1)}\right\| \ge \varepsilon$, $k = k + 1$ go back to step 2.

The FCM uses the constraints that the memberships of an observation in all clusters must add up to 1, so that the outliers significantly affect the clustering results. In this paper, a boundary distance for the cluster has been used in order to identify outliers of the reference dataset. The boundary distance of the i^{th} cluster is defined as:

$$D_{b,i} \equiv D_{ave,i} + 3D_{std,i} \quad \ldots\ldots(7)$$

$$D_{ave,i} = \sum_{j=1}^{m} u_{ij}^q D_{ij,\boldsymbol{\Sigma}_i}^2 \Big/ \sum_{j=1}^{m} u_{ij}^q$$

$$D_{std,i} = \left(\sum_{j=1}^{m} u_{ij}^q \left(D_{ij,\boldsymbol{\Sigma}_i}^2 - D_{ave,i}\right)^2 \Big/ \left(\sum_{j=1}^{m} u_{ij}^q - 1\right)\right)^{0.5}$$

where $D_{ave,i}$ and $D_{std,i}$, respectively, are the average and standard deviation of the Mahalanobis distances that the data points belong to the i^{th} cluster. When the distances follow a Gaussian distribution, the boundary distance covers 99.87% data points, which belong to the i^{th} cluster. The distances of an outlier to each cluster center should be larger than the respective boundary distances. Therefore, before performing FCM iterations, the outliers should be trimmed according to their distances. The proposed algorithm is as follows:

(1) The initial clustering results of FCM are used to calculate the boundary distance for each cluster

$D_{b,i}^{(k)}$, $i = 1...c$. Set the number of retained observations equaling to the number of the reference data, i.e. $m_{ret}^{(k)} = m$, and $k = 0$.

(2) When the observations are inliers, at least one of the Mahalanobis distances should be less than the corresponding boundary distances. Otherwise, the outliers are discarded.

(3) Perform FCM iterations by using the retained observations, in which the numbers of the retained data points are $m_{ret}^{(k+1)}$. After converging, the boundary distances, $D_{b,i}^{(k+1)}$, $i = 1...c$, can be found.

(4) If the changes of the boundary distances are larger than the predefined tolerance (ε), $\sum_{i=1}^{c}\left\|D_{b,i}^{(k+1)} - D_{b,i}^{(k)}\right\| > \varepsilon$, $k = k + 1$ go back to step 2.

2.2 On-line Model Updating

When the on-line data belong to known event groups, the respective cluster parameters, including centers, covariances, and boundary distances, are updated to accommodate the additional information. Marsili-Libelli and Müller enhanced the FCM with an adaptive capability.[3] The centers and covariance matrices of clusters are updated as follows:

$$\boldsymbol{\mu}_i\big|_{m+1} = \left(\sum_{j=1}^{m} u_{ij}^q \mathbf{t}_j + u_{i,m+1}^q \mathbf{t}_{m+1}\right) \Big/ \left(\sum_{j=1}^{m} u_{ij}^q + u_{i,m+1}^q\right) \text{.......(8)}$$

$$\boldsymbol{\Sigma}_i\big|_{m+1} = \frac{\sum_{j=1}^{m} u_{ij}^q \left(\mathbf{t}_j - \boldsymbol{\mu}_i\right)^{\mathrm{T}}\left(\mathbf{t}_j - \boldsymbol{\mu}_i\right) + u_{i,m+1}^q\left(\mathbf{t}_{m+1} - \boldsymbol{\mu}_i\right)^{\mathrm{T}}\left(\mathbf{t}_{m+1} - \boldsymbol{\mu}_i\right)}{\sum_{j=1}^{m} u_{ij}^q + u_{i,m+1}^q}$$

where $u_{i,m+1}$ is the membership of the new data, which was computed from the unchanged prototypes. Since the $\sum_{j=1}^{m} u_{ij}^q \mathbf{t}_j$, $\sum_{j=1}^{m} u_{ij}^q\left(\mathbf{t}_j - \boldsymbol{\mu}_i\right)^{\mathrm{T}}\left(\mathbf{t}_j - \boldsymbol{\mu}_i\right)$, and $\sum_{j=1}^{m} u_{ij}^q$ can be recursively obtained, the computational requirement of the adaptation is moderate. In this paper, the boundary distances of the clusters also have been adapted:

$$D_{b,i}\big|_{m+1} \equiv D_{ave,i}\big|_{m+1} + 3D_{std,i}\big|_{m+1} \text{ ....(9)}$$

$$D_{ave,i}\big|_{m+1} = \left(\sum_{j=1}^{m} u_{ij}^q D_{ave,i} + u_{i,m+1}^q D_{i,m+1}^2\right) \Big/ \left(\sum_{j=1}^{m} u_{ij}^q + u_{i,m+1}^q\right)$$

$$D_{std,i}\big|_{m+1} = \left\{ \left[\begin{array}{l} \left(\sum_{j=1}^{m} u_{ij}^q - 1\right) D_{std,i}^2 \\ + \sum_{j=1}^{m} u_{ij}^q \left(D_{ave,i} - D_{ave,i}\big|_{m+1}\right)^2 \\ + u_{i,m+1}^q \left(D_{i,m+1}^2 - D_{ave,i}\big|_{m+1}\right)^2 \end{array}\right] \Bigg/ \left(\sum_{j=1}^{m} u_{ij}^q + u_{i,m+1}^q\right) \right\}^{0.5}$$

where $D_{i,m+1}^2$ is the Mahalanobis distance, which is the new data point to the i^{th} prototype; and the $D_{ave,i}\big|_{m+1}$ and $D_{std,i}\big|_{m+1}$, respectively, are the updated average and standard deviation of the distances, which data belong to the i^{th} cluster.

2.3 Off-line Adapting Model

Assuming the data of new events with m' rows of observations, the data matrix is denoted as $\mathbf{W}' \in R^{m' \times n}$. The mean vector ($\overline{\mathbf{W}}'$) and the diagonal matrix of standard deviations ($\mathbf{S}'$) of the new event data have to be prepared for normalizing the data matrix $\mathbf{X}' = \left(\mathbf{W}' - \mathbf{1}\overline{\mathbf{W}}'\right)\mathbf{S}'^{-1}$ with zero means and unit variances. The covariance matrix of the new dataset can be obtained from the normalized data matrix: $\boldsymbol{\Sigma}'_{m'} = \mathbf{X}'^{\mathrm{T}}\mathbf{X}' / (m'-1)$. The mean vector and the standard deviations of combining the reference and the new data can be derived as follows:

$$\overline{\mathbf{W}}^* = \frac{m}{m^*}\overline{\mathbf{W}} + \frac{m'}{m^*}\overline{\mathbf{W}}', \quad \mathbf{S}^* = diag\left[\sigma_1^* \quad \sigma_2^* \quad \cdots \quad \sigma_n^*\right] \text{....(10)}$$

$$\sigma_i^* = \sqrt{\frac{(m-1)\sigma_i^2 + m\overline{w}_i^2 + (m'-1)\sigma_i'^2 + m'\overline{w}_i'^2 - m^*\overline{w}_i^{*2}}{m^*-1}} \text{ ...(11)}$$

where the m^* is the total number of observations in the combined dataset, i.e. $m^* = m + m'$. Based on the updated means and standard deviations, the covariance matrix of the combined dataset is written as:

$$\boldsymbol{\Sigma}_{m^*}^* = \mathbf{X}_{m^*}^{*\mathrm{T}}\mathbf{X}_{m^*}^* / (m^*-1) = \left(\mathbf{X}_m^{*\mathrm{T}}\mathbf{X}_m^* + \mathbf{X}_{m'}^{*\mathrm{T}}\mathbf{X}_{m'}^*\right) / (m^*-1) \text{ ...(12)}$$

where $\mathbf{X}_{m^*}^*$ is the combined data matrix $\mathbf{X}_{m^*}^* = \left[\mathbf{X}_m^{*\mathrm{T}} \quad \mathbf{X}_{m'}^{*\mathrm{T}}\right]^{\mathrm{T}}$. The $\mathbf{X}_m^{*\mathrm{T}}\mathbf{X}_m^*$ of the above equation can be obtained from the covariance matrix of the reference dataset.

$$\mathbf{X}_m^{*\mathrm{T}}\mathbf{X}_m^* = (m-1)\mathbf{A}^{\mathrm{T}}\boldsymbol{\Sigma}_m\mathbf{A} + m\mathbf{Z} \text{ ....(13)}$$

$$\mathbf{A} \equiv \mathbf{S}\mathbf{S}^{*-1}, \quad \mathbf{Z} \equiv \left(\mathbf{S}^{*-1}\right)^{\mathrm{T}}\Delta\overline{\mathbf{W}}^{\mathrm{T}}\Delta\overline{\mathbf{W}}\mathbf{S}^{*-1}, \quad \Delta\overline{\mathbf{W}} \equiv \overline{\mathbf{W}} - \overline{\mathbf{W}}^*$$

where $\boldsymbol{\Sigma}_m$ is the covariance matrix before updating the model. In the same way, the covariance matrix of the new event data can be written based on the mean and standard deviation of the combined dataset. The covariance of the combined dataset can be obtained from the previous covariance matrices.

$$\boldsymbol{\Sigma}_{m^*}^* = (m-1)/(m^*-1)\boldsymbol{\Sigma}_m^* + (m'-1)/(m^*-1)\boldsymbol{\Sigma}_{m'}^* \text{ ...(14)}$$

$$\boldsymbol{\Sigma}_m^* \equiv \mathbf{X}_m^{*\mathrm{T}}\mathbf{X}_m^* / (m-1), \quad \boldsymbol{\Sigma}_{m'}^* \equiv \mathbf{X}_{m'}^{*\mathrm{T}}\mathbf{X}_{m'}^* / (m'-1)$$

The singular value decomposition (SVD) is applied to the updated covariance matrix. The eigenvectors ($\mathbf{P}^*$) can be obtained to span the new PCA subspace.

$$\mathbf{\Sigma}_{m^*}^* = \mathbf{P}^{*\mathrm{T}}\mathbf{\Lambda}^*\mathbf{P}^* \quad \text{.....................................(15)}$$

where $\mathbf{\Lambda}^*$ is the diagonal matrix of the eigenvalues.

When the subspace has been adapted, the known cluster parameters can be transferred from the previous subspace.(2)

$$\boldsymbol{\mu}_i^* = \boldsymbol{\mu}_i\mathbf{C}_{k,k^*} + \mathbf{1}\Delta\overline{\mathbf{W}}\mathbf{S}^{*-1}\mathbf{P}_{k^*}^* \quad \text{.................................(16)}$$

$$\mathbf{\Sigma}_i^* = \mathbf{C}_{k,k^*}^{\mathrm{T}}\mathbf{\Sigma}_i\mathbf{C}_{k,k^*} + \mathbf{C}_{n-k,k^*}^{\mathrm{T}}\tilde{\mathbf{\Sigma}}_i\mathbf{C}_{n-k,k^*} \quad \text{.....................(17)}$$

$$\tilde{\mathbf{\Sigma}}_i^* = \mathbf{C}_{k,n-k^*}^{\mathrm{T}}\mathbf{\Sigma}_i\mathbf{C}_{k,n-k^*} + \mathbf{C}_{n-k,n-k^*}^{\mathrm{T}}\tilde{\mathbf{\Sigma}}_i\mathbf{C}_{n-k,n-k^*} \quad \text{...........(18)}$$

$$\mathbf{C}_{k,k^*} \equiv \mathbf{P}_k^{\mathrm{T}}\mathbf{S}\mathbf{S}^{*-1}\mathbf{P}_{k^*}^*,\ \mathbf{C}_{n-k,k^*} \equiv \mathbf{P}_{n-k}^{\mathrm{T}}\mathbf{S}\mathbf{S}^{*-1}\mathbf{P}_{k^*}^*,\ \mathbf{C}_{n-k,n-k^*} \equiv \mathbf{P}_{n-k}^{\mathrm{T}}\mathbf{S}\mathbf{S}^{*-1}\mathbf{P}_{n-k^*}^*$$

where $\mathbf{P}_{k^*}^*$ are the first k^* terms of eigenvectors, which span the new subspace, and $\Delta\overline{\mathbf{W}} \equiv \overline{\mathbf{W}} - \overline{\mathbf{W}}^*$. The covariance matrix of the $(k+1)^{\text{th}}$ to n^{th} term score vectors for the i^{th} cluster is defined as $\tilde{\mathbf{\Sigma}}_i \equiv \mathbf{T}_{i,n-k}^{\mathrm{T}}\mathbf{T}_{i,n-k}\big/(m_i-1)$. The boundary distance of the i^{th} cluster is given by

$$D_{b,i} = \left\|\mathbf{\Sigma}_i\right\|^{1/k}\left(\mathbf{t}_b - \boldsymbol{\mu}_i\right)\mathbf{\Sigma}_i^{-1}\left(\mathbf{t}_b - \boldsymbol{\mu}_i\right)^{\mathrm{T}} \quad \text{............(19)}$$

where $\mathbf{t}_b$ is an arbitrary data point on the previous subspace. The updated boundary distance of the ith cluster can be derived as follows:

$$D_{b,i}^* = \left\|\mathbf{\Sigma}_i^*\right\|^{1/k^*}\left\|\mathbf{\Sigma}_i\right\|^{-1/k} D_{b,i} \quad \text{.........................(20)}$$

where the $\mathbf{\Sigma}_i^{*-1} \approx \left(\mathbf{C}_{k,k^*}^{\mathrm{T}}\mathbf{\Sigma}_i\mathbf{C}_{k,k^*}\right)^{-1}$ are used.

3. ILLUSTRATIVE EXAMPLE

The compressor process is a 4-stage centrifugal compressor, equipped with one intercooler between stages to cool down the compressed air, as shown in Fig. 2. To reach the minimum shaft work requirement, each stage should have the same compression ratio and the lowest possible inlet temperature. Most of the time, the relations between the motor speed, inlet flowrate, and discharge pressure can be shown in the form of curves, called the performance map. The stable operating region is confined by two lines, the surge line and the choke line. The surge line represents the minimum volumetric flowrate on each performance curve. On the other hand, the choke line represents the maximum volumetric flowrate.

In this paper, twenty-nine measured variables were chosen according to the plant expert's advice. The training data were collected every 5 minutes for 4 days. There were 1152 observations in the training dataset that could be explained by the PCA subspace, in which the captured variances were 88% with 2 PCs, within 99% confidence limits. The clustering results of the projection of the first two score vectors are shown in Fig. 3, in which the grey and the black lines represent the boundary distances of clusters by using FCM and DFCM. It is obvious that the shapes of the FCM clusters were stretched due to the outliers. The situation would be significantly improved by DFCM.

The data after the training dataset were collected every five minutes for 2.5 days. Figure 4 shows the statistic Q and T^2 of the first test dataset, in which the solid and dashed lines represent the 99% and 95% confidence limits, respectively, and the control charts suggested the data belonged to the subspace before the 1235^{th} observation and the sample number between 1400 and 1440. Since a data point may be an outlier,

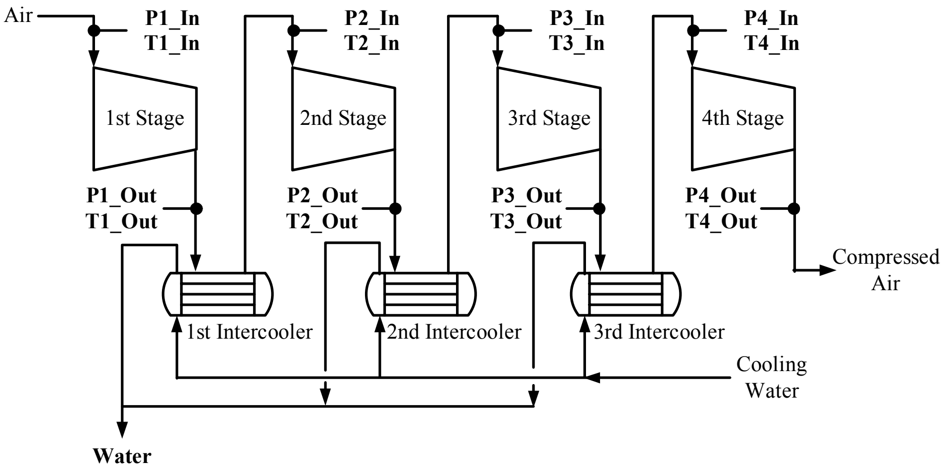

Fig. 2. Air compressor process flow diagram.

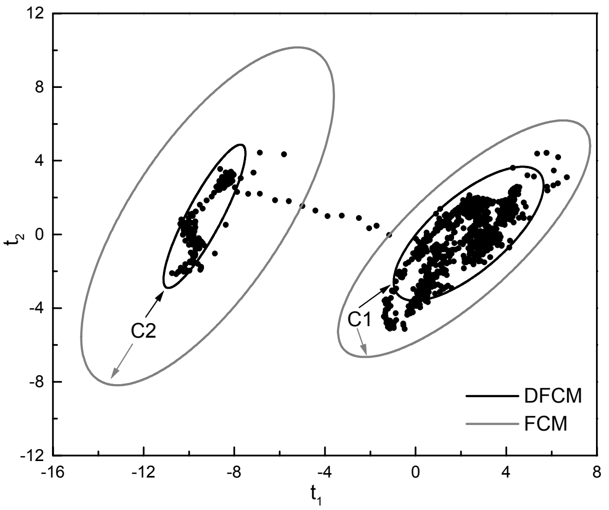

Fig. 3. Cluster results from FCM and DFCM.

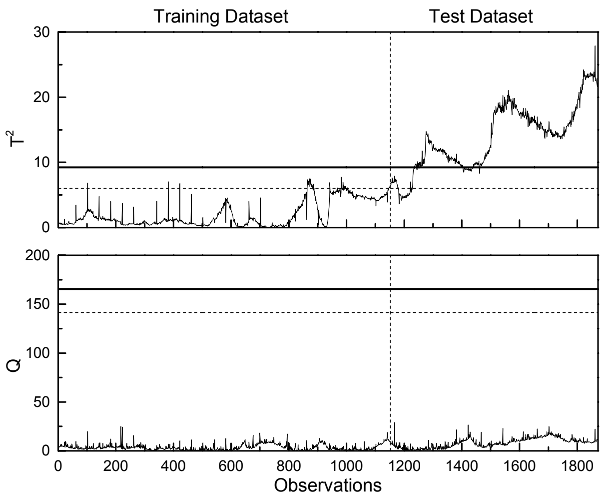

Fig. 4. Process monitoring charts for the first test dataset.

even though it belongs to the PCA subspace, it is necessary to validate whether the data point is part of the known groups, in order to avoid misleading the operator's decisions. So, before classifying the data points which belong to the PCA subspace into the known groups, the Mahalanobis distances of each data point to each cluster centre must be compared with the corresponding cluster boundary distances.

Figure 5 shows the data points that neither belonged to the C1 nor to the C2 cluster after the 1230th observation, and these data points should be stored for the next off-line model update. In Fig. 5, it should be noted that the Mahalanobis distances of the data points which were beyond the subspace were not calculated in the period of the sampling numbers between 1251 and 1400. On the other hand, the data points, in which at least one of the distances was less than the corresponding boundary distance, were classified into the C2 cluster according to their membership values. The centre, covariance, and boundary distance of the cluster C2 were finely adapted by using the additional data, which belonged to the cluster, as shown in Fig. 6. In Fig. 6, the outliers, represented by white points, were collected for the next model update. After that, the PCA subspace was accommodated by using the new event data and the adaptive PCA method. The adaptive PCA, which captured 90% variances with 2 PCs, was applied to the first test dataset.

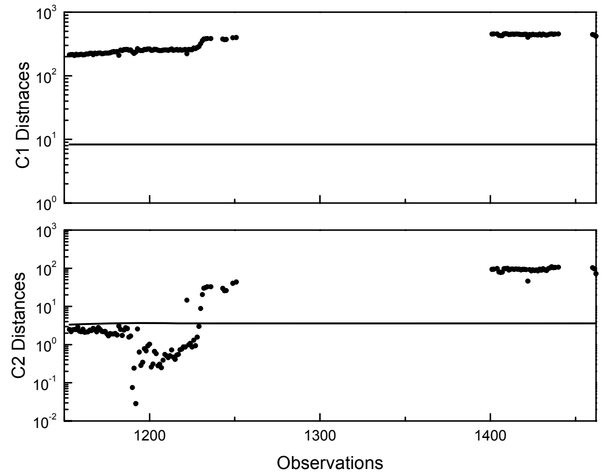

Fig. 5. The Mahalanobis distances of data points.

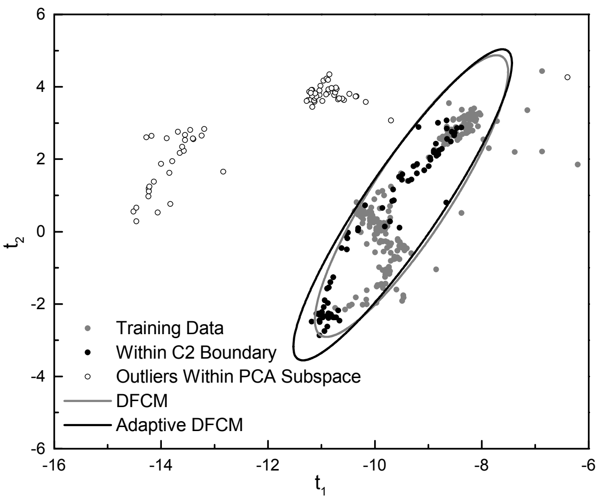

Fig. 6. Adapting C2 cluster with additional information.

Figure 7 shows the Q and T^2 statistics, in which almost all of the statistics of the test data were under their control limits. This shows that the proposed method is capable of adapting the subspace to accommodate new events without recalculating the trained data.

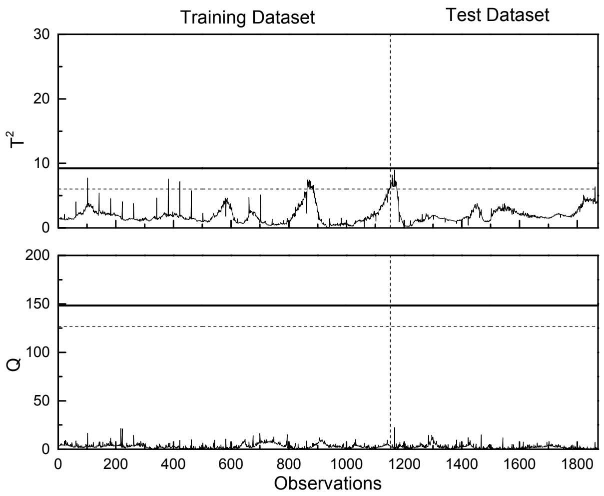

Fig. 7. Statistic Q and T^2 for the training and the first test datasets with the adaptive PCA subspace.

Only the new data need to be clustered on a new PCA subspace, while the known cluster parameters, including cluster centres, covariances, and boundaries, were transferred to the new subspace through rotating and shifting the coordinates of the new subspace. In order to illustrate that the trained data still were covered by the transferred clusters, the first two scores of the training dataset and the first test data, which belonged to C2 cluster, were plotted in Figure 8 with gray points. In Fig. 8, the cluster C1 and C2 were transferred from the previous subspace. The scores of the new event data and the outliers on the previous subspace were clustered by using DFCM. The cluster boundaries are labelled as C3 and C4 in Fig. 8. Eventually, the outliers on the previous subspace, plotted with black points in Fig. 8, were parts of the C3 group

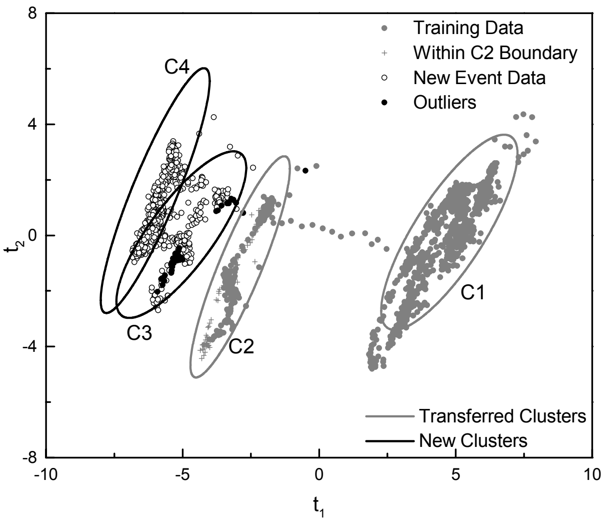

Fig. 8. Clustering results on the new subspace.

that was an unknown event in the training dataset. The operating decisions would be easily misled, if the outliers were not identified in the on-line classification phase.

The second test dataset, which was collected every 5 minutes after the first test dataset for 2.5 days, was examined by utilizing the adaptive PCA subspace. Figure 9 shows most of the statistic Q and T^2 of the second test dataset are within their control limits except for the 2142th observation and the observation numbers between 2407 and 2430, which were stored for the next model update. Figure 10 compares the distances of the data points belonging to the adaptive PCA subspace with each cluster boundary distance. It shows all data points were far from the clusters C1 and C2, and that most of them were within the C4 boundary. The first two scores of the data points also have been plotted in Fig. 11. In Fig. 11, the cluster parameters of C3 and C4 were adapted by the additional information. Figure 11 also shows that the C3 cluster moved slightly toward

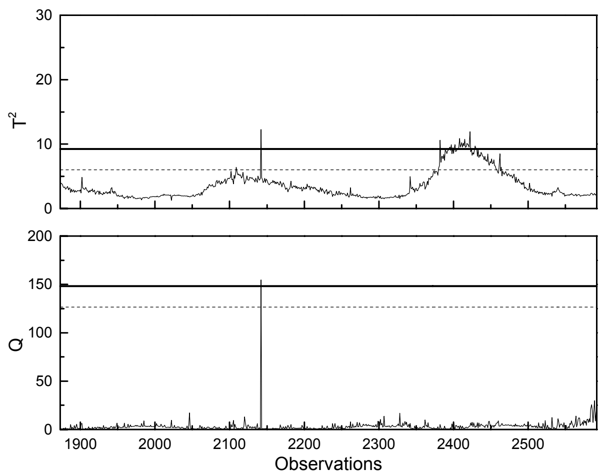

Fig. 9. Process monitoring charts by using the adaptive PCA.

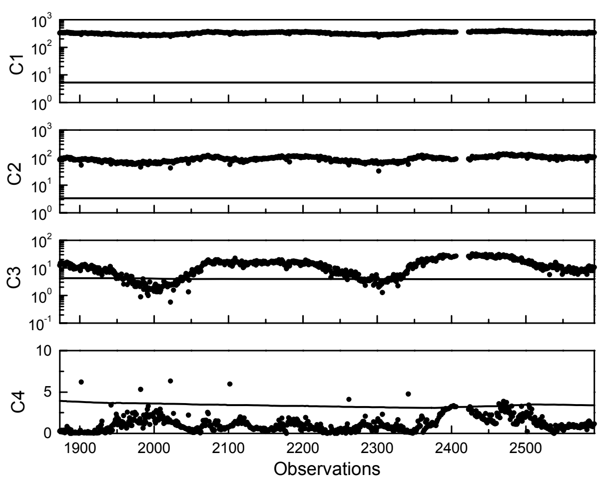

Fig. 10. Data points to each cluster distances.

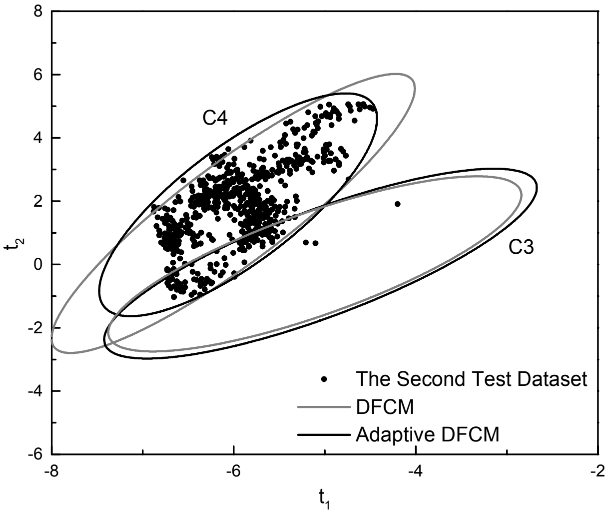

Fig. 11. Adapting clusters by using the second test data.

the C4 group and that the C4 cluster was compactly shrunken in order to accommodate the new information. Therefore, the on-line adaptive method is not only capable of rejecting outliers that mislead the model updating direction, but is also effective in tackling the time-varying process.

To look up the operator logs, the oxygen production rate was decreased after around the 900th observations due to the downstream process reducing its demand. Figure 12 shows the air inlet flowrate was cut to 80% of the normal operating flowrate at that time point. After that, the flowrate further reduced to 70% at the inlet around the 1250th sampling interval. In order to avoid compressor surge, the output pressure had to be decreased simultaneously. Figure 12 shows the output pressure gradually being decreased to 82% of the normal operating pressure. It should be noted that the set point of output pressure was manually adjusted according to the anti-surge control line by field operator. For the purpose of smoothing the transition, the operator carefully adjusted the set point to the low flowrate region. It can be found that the set point changes of output pressure were more frequent than the normal flowrate in Fig. 12. Comparing the clustering results of the training dataset, the clusters C1 and C2 represented the normal operation and 80% represented the normal inlet flowrate condition, as Fig. 12 shows. The first test data were grouped into C3 and C4 clusters, which were partly overlapped as Fig. 8 shows. It is reasonable that the operator finely manipulated the set points of the controlled variables to bring the process to the target. Therefore, the operating regions of the two clusters were close, but not identical. The manipulation resulted in the part of the second test data being shared by C3 and C4 clusters in Fig. 12.

4. CONCLUSIONS

(1) In this paper, an outlier rejection clustering algorithm (DFCM) has been applied to real plant dataset in order to trim the transition data and reach more feasible clustering results.

(2) DFCM can accommodate new event data. When the new PCA subspace has been adapted, the

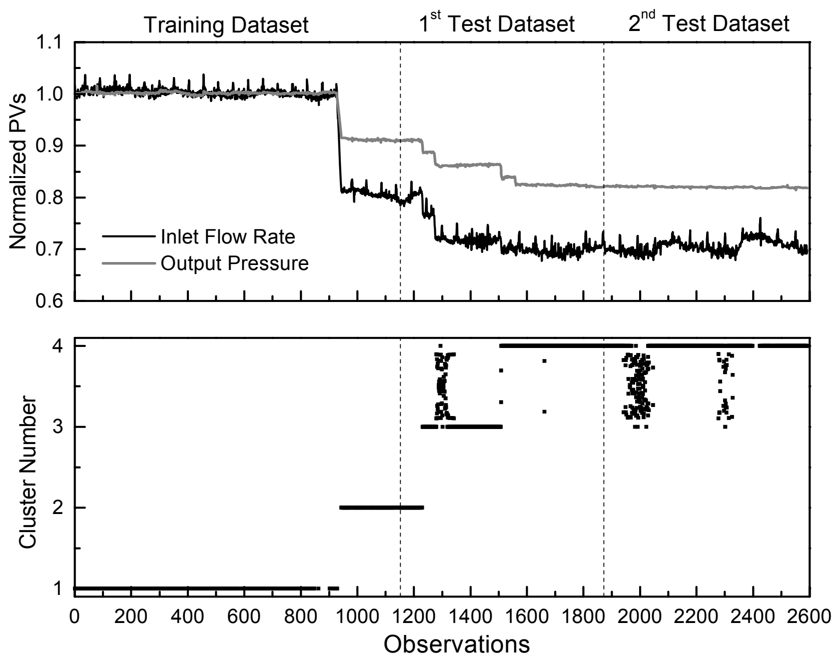

Fig. 12. Compare the operating conditions with the clustering results.

known event clusters can be transferred to the new subspace without recalculating the trained data. This is one of the most desirable characteristics for monitoring a mass production process.

(3) The real plant data from the air compressor process have been demonstrated by using the proposed approach. In this case, new events emerged during on-line monitoring. Before on-line classification, it is necessary to identify the new observations to determine whether they belong to the known event groups; even if they are within the PCA subspace. The results show that the challenges of process monitoring, such as collinearity, outliers, and time-varying characteristics, can be effectively dealt with.

REFERENCES

1. J. E. Jackson: A User's Guide to Principal Components, Wiley, New York, (1991).
2. J. Liu: Process Monitoring Using Bayesian Classification on PCA Subspace. Ind. Eng. Chem. Res. 43, (2004), pp. 7815-7825.
3. S. Marsili-Libelli and A. Műller: Adaptive fuzzy pattern recognition in the anaerobic digestion process. Pattern Recognition Letter, 17, (1996), pp. 651-659.