DIAGNOSIS OF FAULTS WITH VARYING INTENSITIES USING POSSIBILISTIC CLUSTERING AND FAULT LINES[1]

Detroja K. P.[1], Gudi R. D.[2*], Patwardhan S. C.[2]

[1] Interdisciplinary Programme on Systems and Control Engineering,
[2] Department of Chemical Engineering
Indian Institute of Technology Bombay, Powai, Mumbai – 400076, India.
* Email: ravigudi@che.iitb.ac.in

Abstract: In this paper, a new approach for fault detection and isolation that is based on the possibilistic clustering algorithm is proposed. Fault detection and isolation (FDI) is shown here to be a pattern classification problem, which can be solved using clustering and classification techniques. The possibilistic clustering approach was proposed to address some of the short comings of the fuzzy c-means (FCM) algorithm. The probabilistic constraint imposed on the membership value in the FCM algorithm is relaxed in the possibilistic clustering algorithm. Because of this relaxation, the possibilistic approach is shown in this paper to give more consistent results in the context of the FDI tasks. The proposed approach addresses the issue of correctly isolating a fault that may occur with varying intensities. The concept of fault lines is introduced, which in conjunction with possibilistic clustering has been effectively used for FDI. Fault signatures that change as a function of the fault intensities are represented as fault lines, which are shown to be useful to classify faults that can manifest with different intensities. The proposed approach has been validated here through simulations involving a co-polymerization reactor simulation.

Keywords: Possibilistic clustering, Fuzzy c-means clustering, Gustafson-Kessel algorithm, Fault Detection (FDD)

1. INTRODUCTION

Online process monitoring for fault detection and diagnosis (FDD) is very important for ensuring plant safety and product quality. The area of FDD has been very active in recent years. Both, model based and process history based methods have been proposed with a fair amount of success.

In a typical process plant, hundreds of variables are measured every few seconds. These measurements bring in useful signatures about the status of the plant. While model-based methods can be used to detect and isolate signals that indicate abnormal operation, such quantitative cause-effect models may be difficult to develop from the first principles. Methods based on historical data attempt to extract maximum information from the archived data and require minimum physical information of the plant. Multivariate statistical monitoring tools such as PCA (Kresta *et al.*, 1991) are developed to extract information from historical process data so as to carry out the task of FDD easily and more efficiently.

For online process monitoring, it is important to not only be able to say whether the plant operation is aberrant but also to be able to isolate the fault. This is typically done through the use of contribution plots, which assesses the relative contribution of each variable to a suitably formulated error criterion.

While multivariate statistical tools can compress data and perform monitoring in the lower dimension space, they are inherently representation-oriented rather than discrimination oriented. They seek to explain or represent the variance in a data set rather than discriminate between dissimilar subsets in the data (Chiang *et al.*, 2000). Thus, there could be overlaps between the clusters representing fault regions and normal operating regions, leading to higher misclassification rates. Overlaps could also still exist, although to a smaller extent, when tools such as multiple discriminant analysis are used. The latter are discrimination oriented and are generally known to yield directions that enhance discrimination between regions. Thus, it is necessary to look at methods that accommodate such overlaps and analyze these regions so as to provide useful indicators to detect and diagnose faults.

[1] A fuller version of this paper has been communicated for journal review

An alternate approach to the task of fault detection and diagnosis is to mine the archived data and examine patterns in the process variables that indicate the occurrence of a fault. Johannesmayer *et al.* (2002) and Singhal and Seborg (2002) proposed pattern matching methods based on similarity factors which attempted to match patterns in the archived plant database. Typically, parametric faults, sensor and actuator biases and disturbances generate different patterns in the process variables. These patterns or signatures can be classified into different clusters that represent normal or aberrant operation. Subsequently, when deployed online, the plant operation can be classified in terms of the belonging or membership of the new data to the known clusters, based on the similarity of the patterns that the data brings. Thus, the problem is often related towards being able to classify plant operation as normal or belonging to one or more of the faults from the available (measurements and manipulated inputs) plant signatures. This can be effectively done by various pattern recognition and clustering techniques.

Clustering techniques, such as fuzzy c-means clustering (FCM) (Duda *et al.*, 2003) and its variants (Fuzzy Gustafson-Kessel (FGK) algorithm for clustering (Gustafson and Kessel, 1979)) have been very popular in image analysis and pattern classification. Since, fault detection and isolation (FDI) is also a pattern classification problem, these clustering techniques can be effectively used for this task. Attempts have also been made to use different clustering algorithms for the task of FDI. The k-means clustering, which is a hard clustering technique, has been used along with principal components analysis (PCA) and Fisher discriminant analysis (FDA) for the task of FDD in a three step procedure proposed by Peter He et al. (2005). Teppola and Minkkinen (1999) have used adaptive FCM for process monitoring of a waste water plant. They also used possibilistic clustering algorithm for fault detection. Choi *et al.* (2003) used credibilistic clustering algorithm, proposed by Chintalpudi and Kam (1998), based approach for process monitoring.

One of the limitations of the proposed clustering based algorithms is that signatures resulting from the same fault but with differing intensities would confound them and may lead to spurious fault isolation. Another important aspect, relevant for the task of FDD, is related to the issue of identifying and classifying novel faults. It is important to recognize that archived data does not necessarily encompass all possible fault scenarios. Therefore, the FDD algorithm also needs to have learning ability, i.e. when deployed online it should be able to identify the occurrence of new faults and establish relevant signatures or patterns that are representative of the novel fault.

In this paper, we propose to overcome the above difficulties, by using the possibilistic clustering algorithm (Krishnapuram and Keller1, 1993) in conjunction with ‘fault lines’. Possibilistic clustering algorithm is a powerful technique that is similar to probabilistic clustering methods but differs in the nature of the constraint(s) that bind the objective function. Possibilistic clustering algorithm has a number of advantages when compared with conventional FCM algorithm (Krishnapuram and Keller, 1993). In possibilistic clustering, the number of clusters need not be specified accurately and they can be derived during the classification step. In FCM, however, approximate number of clusters/ classes in the data is determined by various cluster validity measures proposed in the literature (Bezdek, 1981). The formulation of the possibilistic clustering algorithm relaxes some constraints on the nature of the membership functions; here we also show that this relaxation gives (i) more consistency in the classification task and (ii) enables the detection of novel (or not-seen) classes. It is also shown here that that the possibilistic clustering algorithm is relatively insensitive to noise and outliers. These features make the possibilistic clustering algorithm more suited for the FDD task. Here, we also introduce the concept of fault lines for handling faults with varying intensities and detecting novel faults. Fault lines are characterized by cluster centre of the normal operation and cluster centre one of the fault clusters.

We demonstrate the suitability of the proposed approach for the FDD task, through simulation case study involving CSTR simulation for solution co-polymerization of methyl-methacrylate (MMA) and vinyl acetate (VA) (Congalidis *et al.*, 1986).

The paper is organized as follows. In the next section we present a brief review of clustering algorithms, namely the FCM/ FGK and PGK algorithm. In the next section the proposed FDI scheme is described in detail. Finally we present a case study to validate the proposed approach.

2. REVIEW OF CLUSTERING ALGORITHMS

The aim of any clustering analysis is to derive a partition of a set of N data points or objects based on some similarity metric, so that the data points/ objects that get clustered into the same group are similar to one another. Clustering algorithms can be broadly classified into hierarchical and non-hierarchical clustering techniques. Here, due to brevity, we briefly review fuzzy c-means clustering and possibilistic clustering algorithms. In fuzzy c-Means (FCM) algorithms, each data point can be a member of more than one cluster, i.e. the membership of a point can take any value between 0 and 1. The following section briefly describes the FCM clustering algorithm.

2.1. *Fuzzy c-means algorithm*

In FCM clustering techniques, a data/ feature point can be a member of more than one cluster with different degrees of membership. If the membership value of j^{th} data point to i^{th} cluster is μ_{ij}, we have the condition that $\mu_{ij} \in [0,1]$. For the data set $\{X \mid x_1, x_2, \ldots, x_N \in X\}$, consisting of c clusters, the constraints imposed on the membership value μ_{ij} are given below.

$$\sum_{i=1}^{c} \mu_{ij} = 1 \qquad (1)$$

where, c is the number of clusters and N is total number of data points in the data set. This is also known as the probabilistic constraint. This constraint requires that total membership of a data/ feature point to all the clusters must be unity. Another important constraint on the clusters is that none of the clusters can be empty, and this constraint can be mathematically represented as shown in Equation (2)

$$0 < \sum_{j=1}^{N} \mu_{ij} < N \qquad (2)$$

The FCM algorithm then minimizes the following objective function subjected to constraints in Equation (1) & (2).

$$J = \sum_{i=1}^{c} \sum_{j=1}^{N} \left(\mu_{ij} \right)^m d_{ij}^2 \qquad (3)$$

where, m is the fuzziness exponent, d_{ij} is the distance between a data point x_j and cluster centre v_i.

The algorithm for FCM starts with some initial guess for either the fuzzy partitioning matrix or the cluster centers and iterates till convergence. The convergence of the FCM algorithm is guaranteed (Bezdek, 1981), but it may converge to a local minima.

The FCM membership of a data point to a cluster depends not only on the distance of that point to the cluster centroid, but also on the distance of that point to other cluster centroids. As will be discussed in detail later, this can cause the algorithm to assign very different memberships to points which are similar as measured by their distances from a cluster center, because their distances from other cluster centers can be different. This problem primarily arises due to the probabilistic constraint described by Equation (1).

The FCM algorithm can be modified in several ways depending on the distance measure chosen. The most commonly used distance measures are Euclidean and Mahalanobis distance. Using these distance measures is equivalent to assuming that the data is oriented in each cluster identically. This may not necessarily be true, for example, the orientation of the data could be spherical in the first cluster and elliptical in the second. From an FDD viewpoint, this could mean higher miss-classification rates and hence poorer diagnosis. To overcome this difficulty, FGK algorithm uses an adaptive distance norm which adapts the similarity measure (norm) according to the shape of the cluster. The algorithm is also quite sensitive to the user specified choice of c, the number of clusters in the data set. The results obtained from FCM/ FGK algorithm largely depends on this number of clusters. If c is not specified correctly, the FCM/ FGK algorithm can arbitrarily split or merge the classes in the data to give exactly c clusters. Different cluster validity measures have therefore been proposed to overcome this difficulty. However, Bezdek (1981) pointed out that the concept of cluster validity is open to interpretation and can be formulated in different ways.

2.2. Possibilistic clustering algorithm

As described earlier, the probabilistic constraint (Equation (1)) imposed on the membership assigned by the FCM/ FGK algorithm brings in problems of classification. These can be broadly enumerated as (i) the points equidistant from the centroid may get very different memberships depending upon the placement of the other clusters, although they are similar as measured by the distance metric, and (ii) the points which are equidistant from all the centroids get the same membership irrespective of their relative positions. To overcome these drawbacks, Krishnapuram and Keller (1993) proposed a new clustering technique called possibilistic clustering, in which the probabilistic constraint on the membership is relaxed. We discuss the possibilistic clustering algorithm in the next section.

In possibilistic clustering, the probabilistic constraint on the objective function in equation (3) is relaxed in possibilistic clustering so as to get membership values, which represent the 'degree of typicality' to a cluster. Simply relaxing the probabilistic constraint produces a trivial solution, i.e. the objective function is minimized by assigning all membership values to 0. Therefore the objective function of Equation (3)(3) is modified as

$$J = \sum_{i=1}^{c} \sum_{j=1}^{N} \left(\mu_{ij} \right)^m d_{ij}^2 + \sum_{i=1}^{c} \eta_i \sum_{j=1}^{N} \left(1 - \mu_{ij} \right)^m \qquad (4)$$

The first term in the equation minimizes the distances of data points from the cluster centers, where as the second term forces the membership values to be as large as possible. In this equation as well, the value of m determines the fuzziness of the final possibilistic partition.

The value of parameter η_i determines the distance at which the membership value of a point in a cluster becomes 0.5. Thus, it needs to be chosen depending on the desired bandwidth of the possibility distribution for each cluster. In practice however, the following definition works well (Krishnapuram and Keller, 1993):

$$\eta_i = \frac{\sum_{j=1}^{N} \left(\mu_{ij} \right)^m d_{ij}^2}{\sum_{j=1}^{N} \left(\mu_{ij} \right)^m} \qquad (5)$$

Updating of the membership values depends on the distance measure chosen. Different distance measures lead to different algorithms. If the distance measure chosen is either Euclidean or Mahalanobis distance, the algorithm gives possibilistic c-means (PCM) membership values. However, if the distance measure is chosen based on scaled Mahalanobis distance and fuzzy covariance matrix, the algorithm gives possibilistic Gustafsson-Kessel (PGK) membership values.

The solution to the objective function in equation (4) leads to the values of memberships as,

$$\mu_{ij} = \frac{1}{1+\left(\frac{d_{ij}^2}{\eta_i}\right)^{1/(m-1)}} \quad (6)$$

The iterative part of the algorithm for possibilistic clustering is very much similar to that of the FCM algorithm, except for the additional parameter η_i which should be estimated from the initial partitioning matrix. However, η_i need not be calculated at every iteration.

Since the parameter η_i is independent of the relative location of the clusters, the membership value μ_{ij} depends only on the distance of a point from the cluster centre (centroid). Hence, unlike in the probabilistic case, the membership of a point in a cluster is determined solely by how far a point is from the centroid and is not coupled with its location with respect to other clusters.

The advantages of PCM/ PGK lie in finding meaningful clusters as defined by dense regions. This happens because each cluster is independent of the other cluster in PCM/ PGK algorithm. Hence, the objective function corresponding to cluster *i* can be formulated as in Equation (7) and the overall objective function is collection of c such objective functions.

$$J_i = \sum_{j=1}^{N}\left(\mu_{ij}\right)^m d_{ij}^2 + \eta_i \sum_{j=1}^{N}\left(1-\mu_{ij}\right)^m \quad (7)$$

It has been shown (Krishnapuram and Keller, 1993) that for a given value of η_i, each of the c sub-objective functions is minimized by choosing the centroid location such that the sum of the memberships is maximized. This makes each cluster centroid to converge to a dense region. Thus, even if the true value of the number of clusters is unknown, the outcome of the algorithm will give c 'good' clusters, i.e. dense regions. Thus, PCM/ PGK have self validating capability which can be very useful when c is not known apriori. When the number of clusters is more than the actual number of clusters in the data set, PCM/ PGK give approximately coinciding clusters, indicating that the actual number of clusters is lesser than specified. This could be interpreted accordingly and the clusters could be collapsed into a single cluster for further analysis.

3. PROPOSED SCHEME FOR FDI

Clustering based approaches are aimed at partitioning the historical data into a number of clusters, e.g. normal operation and different fault operations. Depending on the membership value of the data point to different clusters, the plant operation is declared either normal or otherwise. The shift from normal operating cluster to any fault mode cluster is not instantaneous and the transient response depends on the dynamics of the system. The FCM algorithm would assign different membership values as governed by the probabilistic nature (Equation (1)) to the points even during these transients. This may be useful for example, when the dynamics are to be represented (as shown in Venkat and Gudi, 2002) in a composite modeling methodology, where the memberships essentially weigh the model predictions in each cluster. However, for the FDD task, this may yield erroneous results and misleading interpretations. Possibilistic clustering algorithms (PCM/ PGK) appear to be more suited because these points corresponding to the transition region are not governed by the probabilistic constraint and are assigned low memberships to all the clusters. In the following section, we describe the proposed approach for fault detection and isolation.

3.1. *Data collection and clustering*

The ability of a statistical approach to detect and isolate a fault depends on the availability of rich historical data, containing data corresponding to normal and fault modes of operation. Ideally, the data set used for training the clustering based technique should contain data that represents all possible fault scenarios. In practice however, it may not be possible to have such a data set and the algorithm should have some self-learning abilities.

In general, the historical data consists of measurements of various controlled and manipulated variables at each sampling instant. The clustering approach could either first construct a feature vector from this data or directly work with the measurements. In the former case, the classification is carried out in the space defined by these feature vectors. For example, for incipient fault detection, it may be mandatory to look at the dynamic patterns represented by the feature vector that is constructed from the measurements from the current and past instants. Meel *et al.* (2004) used such an approach to rapidly reject unmeasured disturbances using these pattern recognition techniques, by classifying an appropriately constructed feature vector that was based on apriori knowledge of the dynamics. This approach necessarily requires apriori information of the classification space which is usually difficult to obtain. Alternately, one could directly classify in the space spanned by the measurements (i.e. without constructing feature vectors). This latter approach is taken in this paper.

The clustering algorithm can then be applied on the data. Specifying the exact number of clusters present in the data set is not mandatory for possibilistic clustering approach, as the possibilistic clustering algorithm attempts to search for c good clusters, i.e. dense regions. In the case when the number of clusters specified is more than the actual number of clusters present in the data set, the algorithm will give overlapping clusters indicating that the value of c is over specified. This greatly simplifies the task of clustering of historical data in which the number of clusters present is not known apriori. The outcome of the clustering algorithms would thus yield cluster centroids and fuzzy covariance matrices for each cluster (in case when the GK algorithm is used).

3.2. *Generating Fault lines*

We next discuss the effect of different fault magnitudes and intensities. As mentioned earlier, different fault intensities of the same fault (for example, sensor bias) can manifest in different data vector signatures / paths and would end up into new cluster. In such cases, a methodology, which is still able to classify the fault as a sensor bias (rather than as a novel fault), independent of its magnitude is desirable. Towards this end, we propose the concept of fault lines that characterize the movement of the cluster as a function of the intensity of the fault. When the fault intensity increases, the dynamics and the controller effects result in parallel paths that shift to the fault cluster and eventually end up into new clusters. A fault line could therefore be constructed through the centers of the clusters, beginning from the normal cluster to the fault clusters, and would characterize the behavior of the clusters as a function of increasing intensities of that fault. Assuming that during the training step, data corresponding to a particular fault is available; fault lines can then be constructed to characterize the particular fault.

3.3. Online monitoring and fault detection

For online process monitoring and fault detection, the membership value of the data vector, constructed from the measurements at each instant, to each cluster is calculated from Equation (6) in case of possibilistic clustering approach. High membership values to the normal operating cluster imply that the plant is operating normally. When an abnormal event occurs, these reflect in the signatures of the measured variables, which result in changing memberships of the data vector to the known clusters. An analysis of these memberships would help in the interpretation and classification of the fault scenario. The PCM/ PGK membership value for the normal cluster will assume smaller values close to zero, indicating that a fault may have occurred.

3.4. Fault confirmation and isolation

It is important to recognize that the changing memberships due to the occurrence of a fault are influenced by the inherent system dynamics. The membership profiles can also change due to the occurrence of short-term transients (introduced for example by a control loop), measurement noise or outliers. Thus, it is important to confirm the occurrence of a fault after it is detected. For fault confirmation and isolation, we therefore propose to use a window of M sampling instants over which the membership profiles are analyzed. If the memberships to the normal cluster consistently stay below the user specified threshold for a period exceeding M sampling instants, the occurrence of a fault is confirmed. Similarly, if the memberships to a particular fault cluster assume significant values (above a specified threshold) the fault may be isolated as well.

It should be noted that this will happen only if the fault that has occurred is of the same intensity as in the historical data. In case, if the fault has occurred with a different intensity, the membership value to all known fault clusters will remain close to zero, indicating that a new cluster is formed. As pointed out earlier, the objective function for possibilistic clustering can be seen as a set of c objective functions and the membership value in possibilistic clustering is not influenced by how other clusters are placed. Therefore, it is sufficient to find only the new cluster centre from this newly collected data. Once the new cluster centre is computed, its proximity with the fault lines can be examined. If the new cluster centre is close to one of the fault lines, which are generated from the historical data, the fault may be isolated as the fault associated with that fault line.

The specification of the parameter M has to be carefully done to achieve a compromise between false alarms and sensitivity to the fault occurrence. In general, the choice of M can be made from the closed loop process dynamics or plant operator's experience.

3.5. Novel fault detection

This proposed approach also therefore provides a method to flag novel faults. Low membership value to normal operation cluster for M sampling instants confirms the occurrence of a fault. However, if membership to all known fault clusters and proximity to all known fault lines suggest that the fault that has occurred is indeed novel. Thus, proposed approach enables the classification of the plant operation either as (i) normal operation, (ii) belonging to the known fault scenarios, or (iii) novel faults. The approaches based on other clustering approaches can not provide such crisp division of the plant operation. The new cluster information can be merged with the existing knowledge base and used for future fault diagnosis. Thus an added advantage of the proposed scheme is that it reduces the emphasis on exhaustive historical data. In principle, one can start with just the normal operating data and continue building the monitoring scheme as the new fault events occur.

The role of the fuzziness exponent m in the FDD task also merits some important comments. As mentioned earlier, a higher value of m blurs the distinction between the clusters and makes the cluster boundaries to fade. While monitoring a transition from a normal operating region to a fault mode, with higher values of m, the algorithm would confirm the fault early. However, this high value of m would also increase the incidence of false alarms, which would be indicated when the memberships to the normal cluster decrease. Thus, as in the case of window length M, the value of the fuzziness exponent m should also be chosen as a careful compromise between the requirements of early fault detection and confirmation.

<u>Remark</u>: *The above monitoring strategy is restricted to steady state behavior wherein the points belonging to different operating regions cluster together. For the time varying case, for example in a batch process, the method needs further modifications using manifolds that characterize time varying operation. This aspect is currently under investigation.*

4. CASE STUDY

To validate the proposed approach a simulation case study that is based on co-polymerization reactor is presented here. A 4×5 transfer function matrix model (Congalidis *et al.*, 1986) for a CSTR solution co-polymerization of methyl-methacrylate (MMA) and vinyl acetate (VA) was simulated under closed loop conditions. Based on the RGA analysis, the pairings of controllers were chosen and U_1 was kept constant, effectively resulting in a 4×4 system.

To begin with, the historical data set containing data for (i) normal operation and (ii) for the fault case when sensor Y_1 has developed a bias, was collected. This resulted in two clusters F_0 and F_1 and a fault line corresponding to fault F_1 in the knowledge base. When implemented online, the proposed possibilistic clustering algorithm could easily detect and isolate fault F_1.

In the next step, a positive sensor bias in sensor Y_4 was introduced. As this fault is not part of the archived data that was used for training, it was detected as a novel fault after M (20) samples. The new cluster centre for the newly obtained data was computed and it was found that the new cluster centre (say F_2) was not on the fault line corresponding to fault F_1. Distance of F_2 from F_1 fault line was found to be 2.74 units. Hence, the fault was isolated as novel fault and knowledge base was updated with the new cluster centre. The monitoring scheme now had three fault clusters F_0, F_1 and F_2 along with fault lines for F_1 and F_2. The monitoring scheme could now easily detect and isolate fault F_2 (Figure 1).

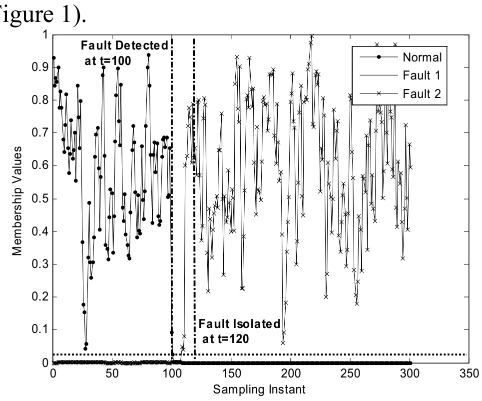

Figure 1: Fault detection and isolation for bias in sensor Y_4

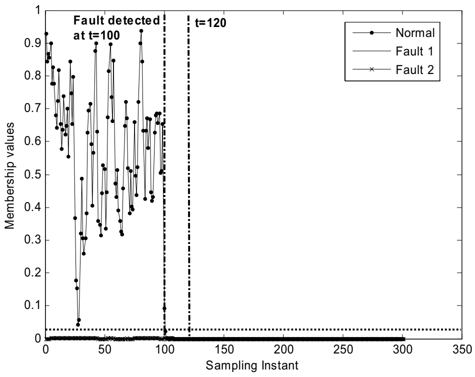

Figure 2: Fault with different intensity than the training set is not be classified as one of the known faults via membership values

As discussed earlier, the same fault can occur with varying intensities during the plant operation. It is therefore important to ascertain that they are not isolated as different faults. To demonstrate the same, a negative bias in sensor Y_4 was introduced. With the proposed approach, it was promptly detected. However, since the intensity of the fault was different than the training set, all the clusters' membership value remained close to zero (Figure 2). Here, fault isolation was performed using the fault lines. The distance of fault lines for F_1 and F_2 were found to be 2.70 and 0.05, indicating that the fault detected is indeed the same fault as F_2 with different intensity.

5. CONCLUSION

A fuzzy clustering and classification based fault detection and diagnosis algorithm was proposed and validated through a simulation case study. The proposed approach is based on the possibilistic clustering methodology of Krishnapuram and Keller (1993) and was found to be vastly superior to other classification methodologies such as the fuzzy c-means and fuzzy credibilistic algorithm. The concept of fault lines was shown to address the difficulty of isolating the same fault with varying intensities. The fault lines were shown to distinguish between scenarios of a novel fault and known fault with different intensities. Thus, the proposed scheme reduced the emphasis on exhaustive historical data and can update the monitoring scheme as the new fault events occur.

REFERENCES

Bezdek, J.C. (1981). *Pattern Recognition with Fuzzy Objective Function*. Plenum Press, New York.

Chiang L. H., Russell E. L. & Braatz R. D. (2000), Fault diagnosis in chemical processes using Fisher discriminant analysis, discriminant partial least squares, and principal component analysis, *Chemometrics & Intelligent Laboratory Systems*, **50**, 243-252.

Chintalapudi K. K. & Kam M. (1998), The Credibilistic fuzzy c-means clustering algorithms, *IEEE Int. Conf. Syst., Man, Cybernetics*, **2**, 2034-2039.

Choi S. W., Yoo C. K., & Lee I. (2003), Overall statistical monitoring of static and dynamic patterns, *Industrial and Engineering chemistry research*, **42**, 108-117.

Congalidis J. P., Richards J. R. & Ray W. H. (1986), Modeling and control of copolymerization reactor, *Proceedings of American Control Conference*, **3**, 1779-1793.

Duda R. O., Hart P. E. and Stork D. G (2003), *Pattern Classification*, Wiley-Interscience publication, New York.

Gustafson, D.E. and Kessel W.C. (1979), Fuzzy clustering with a fuzzy covariance matrix. *Proc. IEEE CDC*, San Diego, CA, USA, 761–766.

Johannesmeyer, M.C., Singhal A. and Seborg D.E. (2002), Pattern Matching in Historical Data, *AIChE J.*, **48**, 2022-2038.

Kresta J. V., MacGregor J. F. & Marlin T. E. (1991), Multivariate statistical monitoring of process operating performance. *Canadian Journal of Chemical Engineering*, **69**, 35-47.

Krishnapuram R. and Keller J. M. (1993), A possibilistic approach to clustering, *IEEE transactions on fuzzy systems*, **1**(2), 98-110.

Meel, A., Venkat, A. and Gudi, R.D. (2003), "Disturbance Classification and Rejection using Pattern Recognition methods", *Ind. and Eng. Chem. Res.*, **42**, 3321-3333.

Peter He Q., Wang J. & Joe Qin S. (2005), A new fault diagnosis method using fault directions in Fisher discriminant analysis, *AIChE J.*, **51**(2), 555-571.

Singhal, A., and Seborg D.E. (2002), Pattern Matching in Multivariate Time Series Databases Using A Moving Window Approach, *Ind. & Eng. Chem. Res.*, **41**, 3822-28.

Teppola P. and Minkkinen P. (1999), Possibilistic and fuzzy c-means clustering for process monitoring in an activated sludge waste-water treatment plant, *Journal of Chemometrics*, **13**, 445-459.

Venkat A. and Gudi, R.D. (2002), "Fuzzy Segregation based Identification and Control of Nonlinear Dynamic Systems", *Ind. and Eng. Chem. Res.*, **41**, 538-552.