Statistical Process Monitoring and Disturbance Diagnosis in Multivariable Continuous Processes

Anne Raich and Ali Çinar
Dept. of Chemical and Environmental Engineering, Illinois Institute of Technology, Chicago, IL 60616

Detecting out-of-control status and diagnosing disturbances leading to the abnormal process operation early are crucial in minimizing product quality variations. Multivariate statistical techniques are used to develop detection methodology for abnormal process behavior and diagnosis of disturbances causing poor process performance. Principal components and discriminant analysis are applied to quantitatively describe and interpret step, ramp and random-variation disturbances. All disturbances require high-dimensional models for accurate description and cannot be discriminated by biplots. Diagnosis of simultaneous multiple faults is addressed by building quantitative measures of overlap between models of single faults and their combinations. These measures are used to identify the existence of secondary disturbances and distinguish their components. The methodology is illustrated by monitoring the Tennessee Eastman plant simulation benchmark problem subjected to different disturbances. Most of the disturbances can be diagnosed correctly, the success rate being higher for step and ramp disturbances than random-variation disturbances.

Introduction

Maintaining high product quality in chemical process manufacturing has become a widespread concern. Efforts to manufacture more on-target product and to reduce variation in product properties have lead to the increased use of statistical process monitoring and quality control techniques in a variety of industries. The traditional statistical process monitoring/control (SPC) approach adopted in many industrial environments relies on the use of Shewhart charts (Montgomery, 1991; Wetherhill and Brown, 1991). Shewhart charts are developed for each variable monitored and it is assumed that the process is in control as long as all variables are in control. If any one of the variables moves out of control, then the process is declared to be out of control. This approach is appealing to most plant personnel and parallels an old process control philosophy that concentrates on keeping each variable at its setpoint in order to keep the process at the desired operating point. However, most chemical process operations are multivariable continuous processes with collinearities among process variables. The collinearities generate strong departures from the assumptions utilized in developing this approach and limits its usefulness in multivariable processes.

Correspondence concerning this article should be addressed to A. Çinar.

An example where a two-variable process (x_1 and x_2) is monitored by using univariate Shewhart charts (Figure 1a) or the control ellipse based on Hotelling T^2 (Figure 1b) illustrates the pitfalls of using univariate charts. Based on the Shewhart charts, the process is out of control at samples 16, 17, and 19. In reality, based on multivariate analysis, sample 16 is inside the control ellipse and the process is in control. Furthermore, sample 18, which indicates that the process is in control based on the Shewhart charts, is outside the control ellipse and the process is out of control. A probability level of 0.95 is used for both the univariate and the multivariate charts. The erroneous decisions about samples 16 and 18 do not change when the univariate chart limits are adjusted for joint (simultaneous) confidence by using techniques such as Bonferroni inequality. The use of other types of univariate charts such as cumulative sum (CUSUM) or exponentially weighted moving average (EWMA) charts would result in similar decisions. The issue is the use of univariate SPC charts for monitoring multivariable processes that have collinearity among variables.

Statistical process monitoring and control involve three activities: *detection* of the out-of-control status, *identification* of the variable(s) responsible for the process to go out of control, and *diagnosis* of the source cause for the abnormal be-

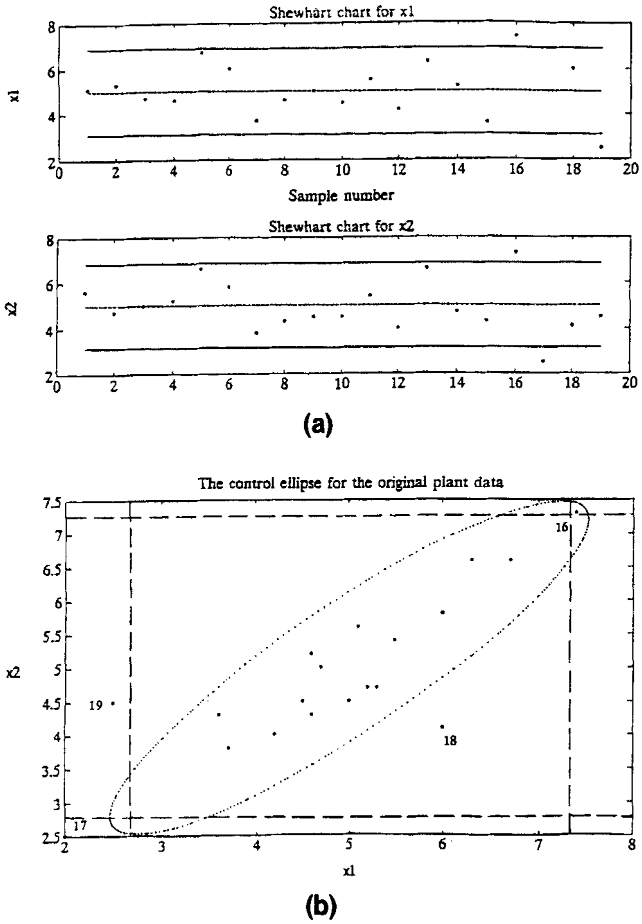

Figure 1. (a) Univariate Shewhart charts for variables x_1 and x_2; (b) multivariate control ellipse based on Hotelling T^2 for x_1 and x_2.

havior. Monitoring focuses on the detection and identification activities, while diagnosis provides the information for the intervention or control stage. The terminology varies across different disciplines, for example, in aerospace and system theory communities the terms fault detection and isolation are favored (Frank, 1990). In univariate processes, the SPC chart accomplishes the first two activities simultaneously. In multivariate processes, a univariate chart may identify a variable responsible for indicating the process to be out of control, but it is not a reliable tool for detection of the out-of-control status as shown in the illustrative example. Hotelling T^2 can detect the out-of-control status reliably, but offers no assistance in identifying the variables responsible for this status.

Advances in SPC techniques for multivariable processes have evolved in two different directions. One group of methods involves the development of efficient techniques for detection of the out-of-control status. CUSUM techniques for T^2 (Alwan, 1986; Crosier, 1988), multivariate CUSUM techniques (Woodall and Ncube, 1985; Healy, 1987; Pignatiello and Runger, 1990), and multivariate exponentially weighted moving average EWMA methods (Lowry et al., 1992) have been proposed for faster detection of smaller deviations. Once an out-of-control status is detected, other techniques such as univariate Shewhart charts have been utilized in order to identify the specific variables that cause the out-of-control status (Alt, 1985; Doğanaksoy et al., 1991). The second group of techniques focuses on the integration of detection and identification. These techniques utilize residual based methods. Principal components analysis (Jackson, 1992; Jackson and Mudholkar, 1979; Johnson and Wichern, 1992; Kresta et al., 1991; Wold et al., 1987), multivariate linear regression (Hawkins, 1991) and partial least squares regression (Hoskuldsson, 1988; Skagerberg et al., 1992; Wise et al., 1990; Negiz and Çinar, 1992) have been utilized for generating the residuals. This study presents an integral statistical methodology for automated detection of abnormal process operation and discrimination between several source causes by utilizing principal components analysis and discriminant-analysis techniques. It extends the methodology reported in the second group of techniques. Like most of the methods discussed in the references cited, the proposed methods are developed for monitoring processes from their steady-state operation. Consequently, lack of significant autocorrelation, stationarity, and ergodicity must be established before utilizing these methods. The methodology is illustrated by monitoring the Tennessee Eastman industrial challenge problem (Downs and Vogel, 1990). The method does not rely on visual inspection of plots; consequently, it is suitable for processes described by large sets of variables.

An important concern in industry is the detection and diagnosis of *multiple simultaneous disturbances*. In a real process, combinations of disturbances may occur. The best intervention policy may need to take into account each of the contributing disturbances. Diagnosis should be able to identify major contributors and correctly indicate which, if any, secondary disturbances are occurring (Raich and Çinar, 1995). Most monitoring techniques are developed based on the assumption of a single fault occurrence at a time. Hierarchical artificial neural networks were proposed recently to diagnose multiple simultaneous faults (Watanabe et al., 1994). In this study, several statistical measures are utilized to assess the overlap between models describing single-disturbance effects. The similarity between models indicates the potential for confusion and masking of multiple-disturbance effects. Quantitative measures to compare multivariable models permit decisions about usefulness and discrimination capability of the models. They also provide *a priori* information about disturbances that is likely to be masked by other disturbances. The Tennessee Eastman process is used to illustrate the methodology proposed.

Statistical Monitoring of Multivariable Processes

In multivariate processes, Hotelling T^2 statistic reliably detects the out-of-control status, but offers no assistance in identifying the variables responsible for indicating that the process is out of control. The identification of these variables is usually necessary in order to facilitate the diagnosis stage. Consequently, if diagnosis can be facilitated through some other means, the identification activity may be less crucial. The approach presented in this article utilizes the detection and the diagnosis activities in order to determine the disturbances that affect a multivariable process and cause the process to go out of control. The method proposed belongs to the class of *residuals-based techniques* such as Kalman filter techniques that have been successful in fault diagnosis (Frank, 1990).

Standard SPC methods use statistical hypothesis testing repeatedly in order to decide if a significant change from normal operation has occurred. The simplest tests such as the Shewhart charts are for a single variable at a time, and can detect changes in the mean value and/or range of a variable. However, the simultaneous use of univariate tests on several process variables can fail to recognize out-of-control conditions in multivariable processes. Hotelling suggested the T^2 test to monitor deviations from a target in multivariable processes, considering a process to be on target whenever the quadratically weighted distance from the desired (average) measurement vector is less than a probability threshold value. Also by using the covariance of the measurements in normal operation for scaling, quadratically weighted distances can be used in discrimination analysis to distinguish between different modes of process operation. Hotelling's statistics and probability descriptions of the possible modes can be compared to decide on a likely class of operation.

A considerable mathematical difficulty is posed by the quadratic weighing, which requires the inversion of the covariance matrix. When measurements are correlated, inversion tends to place unwarranted importance on the least important changes in operation, and can be mathematically unstable. One modification to overcome the inversion problem is to work in the reduced-dimension space described by principal components. Such an approach has the appealing characteristic of considering the largest linear trends in variation while filtering out insignificant measurement noise. Multivariate process monitoring based on principal components has been proposed by Kresta et al. (1991) based on visual inspection of 2- or 3-dimensional principal components plots. This approach has limitations when the process description necessitates several principal components, since projections to 2-dimensional plots do not provide a clear picture.

This study focuses on the development of methodology that yields quantitative information summarizing process behavior and eliminates the need for visual inspection. It will be illustrated by using the Tennessee Eastman problem.

Geometry of principal components models

Principal components analysis (PCA) techniques are used to develop a model describing variation under normal operating conditions (NOC). This model is utilized to detect outliers from NOC, as excessive variation from normal target and as unusual patterns of variation. Operation under various known types of upsets can also be modeled if sufficient historical data are available, and these fault models can then be used to isolate source causes of faulty operation based on similarity to previous upset behavior.

New Coordinate System. Principal components (PCs) are a new set of axes in which the data points are randomly scattered, usually with fewer dimensions than the original measurements (Figure 2). For a data set that is well described by two PCs, the data can be displayed in a plane. The scatter is enclosed in an ellipse whose axes are the PC *loadings*. However, for higher than two or three dimensions, graphical displays do not provide a clear picture.

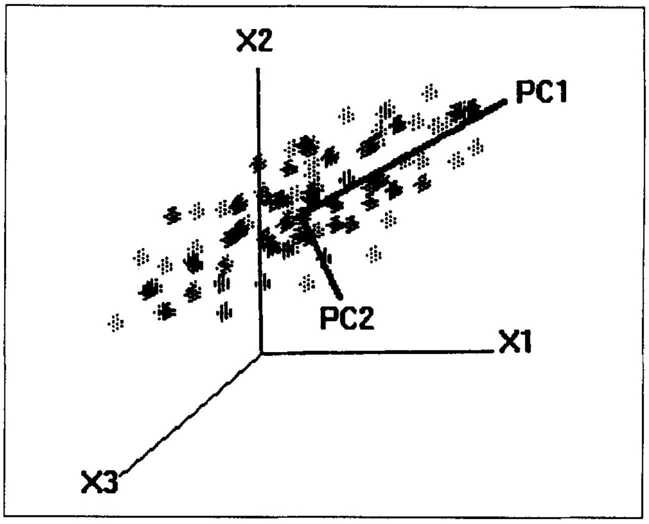

Figure 2. Principal components of multivariate data.

Monitoring Space. The envelope of NOC can be described by using the principal components (Figure 3). For data using two PCs, observations can be displayed with a control ellipse to monitor process performance: points located inside the ellipse are considered in control with respect to the PCs. The boundary can be placed statistically using Hotelling's T^2 statistic for location along PCs, called *scores*. A third dimension is incorporated to account for the magnitude of the model residuals. Large residuals indicate that the model is not capable of describing the process, and process behavior is affected by characteristics not included in the model. For data requiring more than two dimensions, the same statistical measures can be used for quantitative testing of points in a hyperdimensional elliptical envelope. In addition, the possible loss of information by using fewer PCs than original variables should also be considered. Data points that have larger than typical error after use of PCs, called *residuals*, can also be considered as outliers from NOC.

Disturbance Isolation. Using PCs for several sets of data under different operating conditions (NOC and with various upsets), statistics can be computed to describe distances of

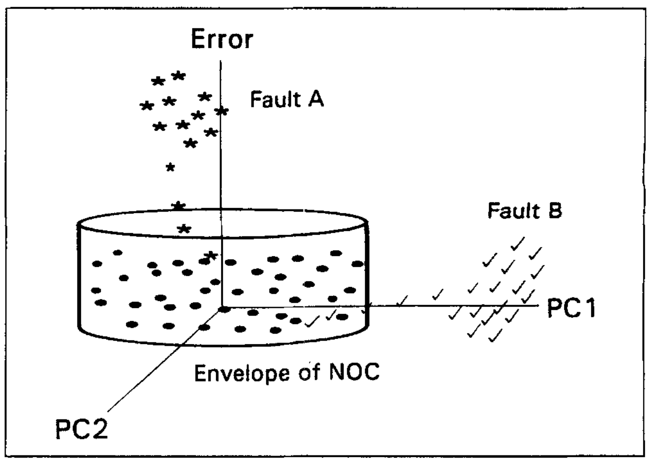

Figure 3. Envelope of normal operation region based on PCA.

the current operating point to regions representing other areas of operation. Both score distances and model residuals are used to measure such distance-based statistics as defined by Eqs. 3 and 4.

Statistical tools

Principal Components. A coordinate transformation of data matrix X, composed of n correlated observations of p variables, to fewer independent score variables T, can be performed using singular-value decomposition:

$$X'X = P'\Sigma P \qquad T = XP. \tag{1}$$

X is usually generated by mean centering and variance scaling raw process data. The decomposition acts on the covariance matrix of X and produces the vectors (PCs) for a change of coordinates. Each additional coordinate is chosen to be orthogonal to the previously selected PCs, and its direction provides the best explanation of the remaining unexplained variation in X. The covariance matrix Σ of the independent variables in the new coordinate system is diagonal. The elements of Σ indicate how much of the overall variation in X occurs in the corresponding direction in P.

Dimensionality Reduction. The number of PCs used to describe the data is often less than the original number of variables. This is justified by considering the amount of variation explained by each dimension. Several methods have been proposed to select the number of PC dimensions (Jackson, 1991; Himes et al., 1994). The guideline used in this study is that a PC should describe more than the average variation of all remaining PCs with less variation. Using the corresponding elements of Σ to test each coordinate vector in P, the dimensions can be discarded that describe less variation than the smallest PC for which

$$\frac{\lambda_k}{\lambda_{\text{pool},k}} > F_{1,p-k,\alpha}, \tag{2}$$

where λ_k denotes the variation along the PC under consideration, $\lambda_{\text{pool},k}$ is the average of all smaller variations, and $F_{1,p-k,\alpha}$ is a confidence threshold using the F statistic at probability level α. Such a test looks for an elliptical rather than spherical shape for the confidence envelope; if the boundary is spherical, PCs offer no advantage over the original variables.

The first PCs describe the common features in the data, while the higher PCs highlight the subtle differences. Consequently, for discrimination and diagnosis, the inclusion of the higher PCs may be very useful. Unfortunately, this increases the number of PCs to a value that makes graphical presentations impossible. The graphical interpretation of the F-test is shown in Figure 4. Starting with 22 PCs calculated by using observations collected from an example process (the Tennessee Eastman process), the test value goes above the threshold when the number of PCs is 18, indicating that 18 PCs have a significant contribution to the description of process behavior. However, PCs 17 and 16 are below the threshold. Consequently, reducing the number of significant PCs to 15 may also be considered.

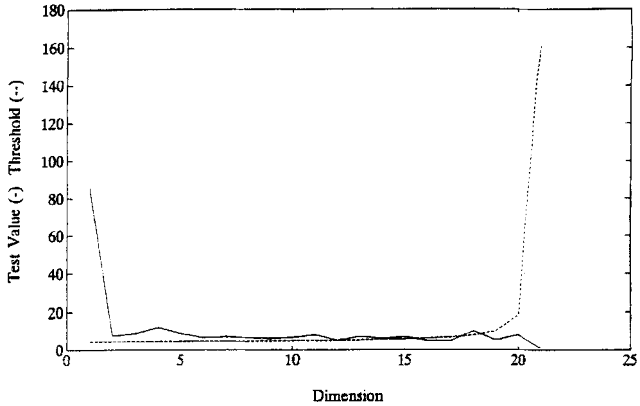

Figure 4. *F*-test for number of dimensions to retain.

Monitoring

Monitoring the status of the process is based on the score distance and on the residuals.

Score Distance. Testing if an observation is within the control envelope is done by using PC scores and Hotelling's T^2 statistics s:

$$s < s_\alpha \quad \text{with} \quad s = t'\Sigma^{-1}t$$
$$\text{and} \quad s_\alpha = \frac{p(n-1)}{(n-p)}F_{p,n-p,\alpha}, \tag{3}$$

where $t = xP$ is the location of original observation x in the PC space and $F_{p,n-p,\alpha}$ is the F statistic at confidence level α.

Model Residual. Testing that an observation does not have large variation unexplained by the PC model is accomplished by using a threshold r_α derived from the standard normal confidence limit:

$$r < r_\alpha \quad \text{with} \quad r = t'(I - PP')t$$
$$\text{and} \quad r_\alpha = a(b + cz)^d, \tag{4}$$

where r denotes the residual, z the standard normal threshold, I is the identity matrix, and a, b, c, d are parameters for correcting sampling effects. Denoting the eigenvalues of the covariance matrix by λ_i, and the subscript of the first (largest) omitted eigenvalue by $k+1$ (model dimension is k):

$$a = \sum_{i=k+1}^{p} \lambda_i \qquad b = 1 + [\theta_2 h_o(h_o - 1)]/a^2$$
$$c = \left(\sqrt{2\theta_2} h_o\right)/a \qquad d = 1/h_o \tag{5}$$

with (Jackson and Mudholkar, 1979)

$$\theta_2 = \sum_{i=k+1}^{p} \lambda_i^2 \qquad \theta_3 = \sum_{i=k+1}^{p} \lambda_i^3 \qquad h_o = (1 - 2a\theta_3)/(3\theta_2^2).$$

The method is suitable for high-dimensional data since these are quantitative tests and do not rely on visual presentation

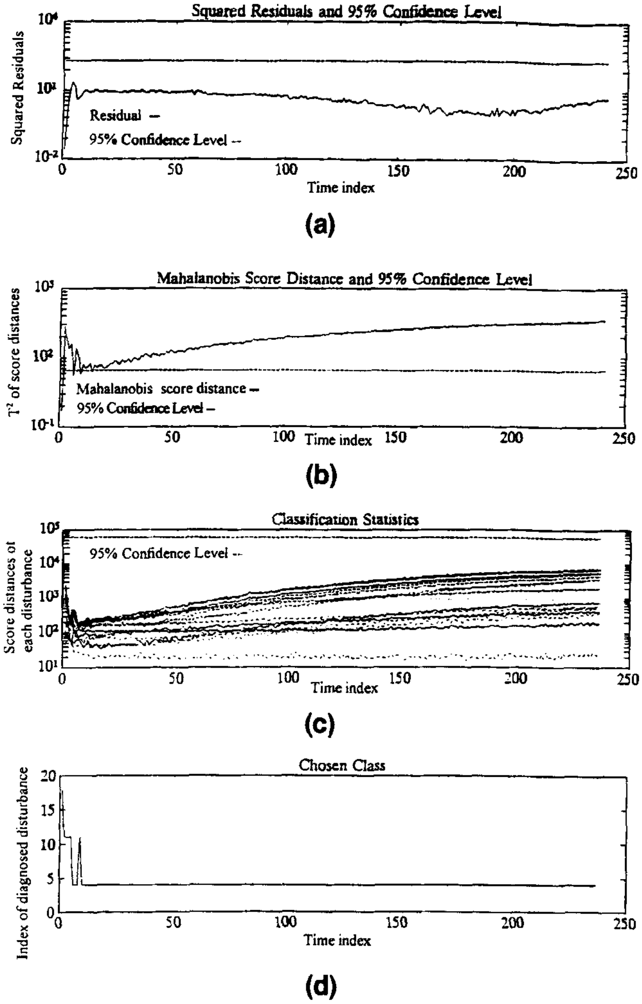

Figure 5. Detection and diagnosis of process upsets.

(a) Detection of outliers based on residuals. (b) Detection based on T^2 test of scores. (c) Diagnosis statistics considering each possible disturbance. (d) Index of chosen disturbance for each observation.

and interpretation. However, the calculated statistics and thresholds can be presented in Shewhart-style control charts to follow the testing over a period of operation (Figure 5). In Figure 5a, data collected over a period of upset are tested against the NOC model residual threshold, with no faults shown. The same data are also tested against the NOC score distance threshold (Figure 5b). Here faulty behavior is indicated within 20 samples.

Disturbance isolation

PC models for specific disturbances can be developed by utilizing historical data sets collected when that disturbance was active. When current measurements exhibit out-of-control behavior, a likely cause for this behavior can be assigned by pattern matching by using scores, residuals, or a combination of the two.

Score Discriminant. Assuming that the PCA models retain the variation important to discrimination between possible causes in scores that have independent normal distributions, the maximum likelihood is that x is from the fault model i based on the minimum distance given by the maximum of d_i:

$$d_i(t) = -0.5t'\Sigma_i^{-1}t + 0.5\ln[\det \Sigma_i] + \ln(p_i), \qquad (6)$$

where $t = xP_i$ is the location of original observation x in PCA space for fault model i, Σ_i is the covariance along PCs for fault model i, and p_i is the adjustment for overall occurrence likelihood of fault i (Johnson and Wichern, 1992).

Figure 5 illustrates the disturbance isolation process. Score discriminants are calculated using PCA models for the various known faults (Figure 5c); this semilog plot shows the negative of the discriminant. As with residuals and scores from NOC, these statistics are smooth over time, which is consistent with the step nature of the underlying cause. The most likely disturbance is chosen over time by selecting the disturbance corresponding to the maximum discriminant (curve with the lowest magnitude). Figure 5d reports the disturbance selected at each sampling time. Disturbance number 3, which was the correct disturbance, has been reported consistently after the first ten sampling times.

Residual Discriminant. Assuming that observations will not be well described by PCA models for other disturbances but will be within the residual threshold of their own class, it is most likely that x is from the fault model i with minimum

$$r_i/r_{i,\alpha} \quad \text{where} \quad r_i = t_i'(I - PP')t_i, \qquad (7)$$

where r_i is the residual computed by using the PCA model for fault i and $r_{i,\alpha}$ is the residual threshold at level 100α based on the PCA model for fault i.

Combined Discriminant. Combining the information available in both the scores and the residuals usually improves the diagnosis accuracy (Naes and Isaksson, 1991). Comparing the combined information to the confidence limits of each fault model, x is most likely to be from the fault model i with minimum

$$c_i\left(\frac{r_i}{r_{i,\alpha}}\right) + (1-c_i) * \left(\frac{s_i}{s_{i,\alpha}}\right), \qquad (8)$$

where s_i and r_i are the score distance and the residual based on the PCA model for fault i, $s_{i,\alpha}$ and $r_{i,\alpha}$ are the score distance and residual thresholds using PCA model for fault i, and c_i is a weighting factor between 0 and 1. In order to weigh scores and residuals according to the amount of variation in the data explained by each, c_i equal to the fraction of total variance explained by scores is used in this study. The combined discriminant value calculated gives an indication of the degree of certainty for the diagnosis; statistics less than 1 indicate a good fit to the chosen model. If no model results in a statistic less than 1, none of the models provide an adequate match to the observation.

Summary of the procedure

Design. Historical data sets collected during normal plant operation and operation under specific disturbances are utilized to generate the principal component models. The model describing process behavior under normal operating condi-

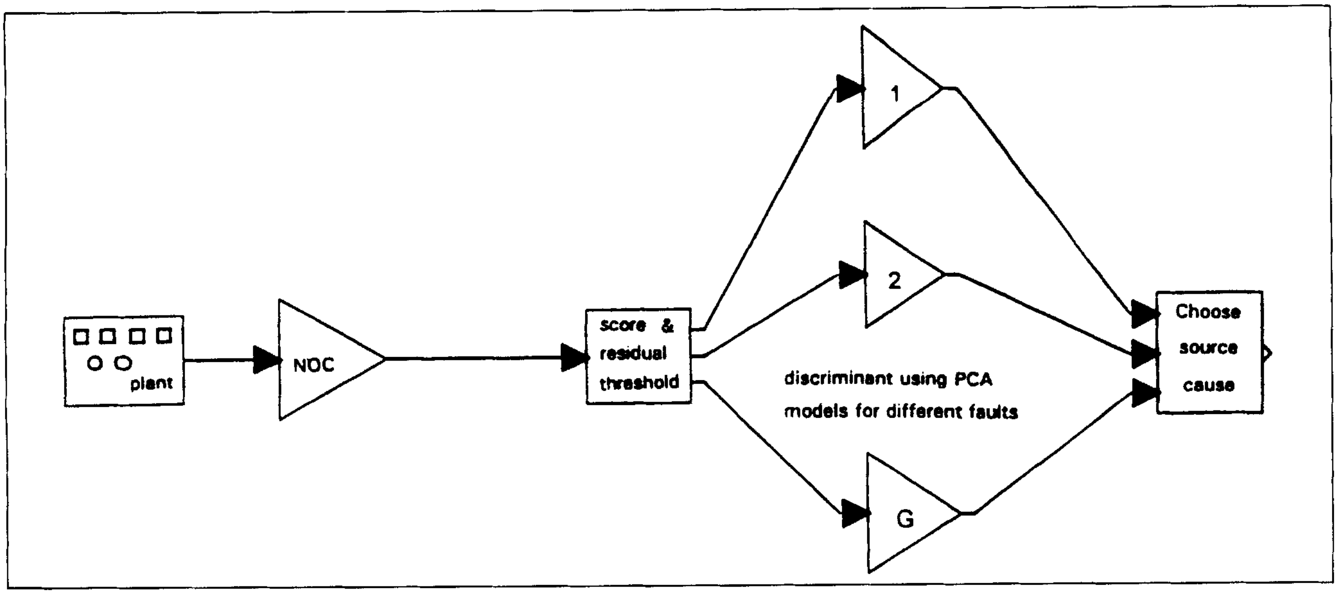

Figure 6. Process information monitoring and disturbance diagnosis.

tions is used with new data to decide if the current operation is in control (Figure 6). Measurements from the plant are used with the PCA model to test if an observation indicates upset from NOC. If score or residual tests exceed their statistical limits, further PCA models for different types of faulty behavior are used to calculate discriminant statistics. The model giving the maximum or minimum discriminant is the most likely of the known source causes for the observed upset.

Implementation. When new process data become available, the monitoring and disturbance diagnosis test is initiated. If there is no significant evidence that the process is out of control, further analysis is not necessary and the procedure is concluded. If there is significant evidence that the process is out of control, then the PC models for each disturbance are used to carry out the score and residuals tests, and discriminant analysis is performed by using various disturbance PC models to diagnose the source cause of abnormal behavior (Figure 6).

Discrimination and Diagnosis of Multiple Disturbances

Multivariate statistical methods of model development have been explored for many years, but comparison of models for different data sets is difficult. A statistical basis for comparison of continuous process models has many areas of application. For monitoring of acceptable performance, many models can be built. For example, data can be chosen from different runs or at different sampling rates. A method is needed to decide if the resulting models are significantly different on an appropriate data set. For a batch process, it may be useful to have a statistical basis to identify significant changes in the process model. In disturbance discrimination, where process behavior due to different disturbances is described by different models, it is useful to have a quantitative measure of similarity or overlap between models, and to identify the overall likelihood of successful diagnosis.

Comparison of multivariate models

In comparing multivariate models, much work has been done for testing significant differences between means when covariance is constant. Testing for differences in covariance is much more difficult yet more crucial; diagnosis can be successfully done, whether or not means are different, as long as there is a difference in covariance (Fukunaga, 1990). For example, ellipses A and B in Figure 7 have similar orientation (covariance), while ellipses B and C overlap with each other's center (mean). Although means of B and C overlap, most points in their ellipses can still be distinguished.

Testing for eigenvalue models of covariance adds new complications, since the statistical characteristics are not well known, even for the most common distributions. Simplifying

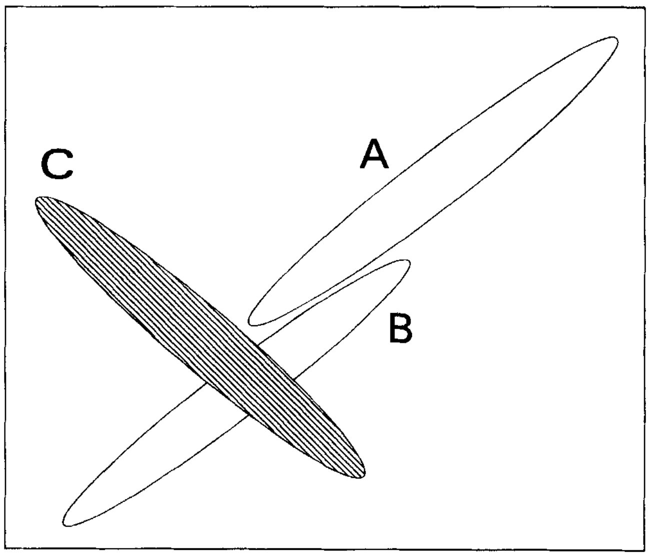

Figure 7. Overlap of groups with similar mean or covariance.

assumptions for narrowly restricted special cases can be made, with significant loss of generality (McLachlan, 1992).

Angles between Different Coordinate Systems. Consider the definition of angles between vectors in different coordinate systems. The minimum angle between an arbitrary vector in one principal component space P_1 and the most nearly parallel vector to it in a second PC space P_2 is given by the largest eigenvalue s_1 of S where

$$L'SL = P_1'P_2P_2'P_1 \tag{9}$$

and L gives a set of consensus coordinates. The angles between the most parallel coordinate directions (loadings) in the two coordinate systems are defined the same way. Using eigenvalue decomposition, the angles α_k between coordinate directions in coordinate sets P_1 and P_2 correspond to eigenvalues s_k where

$$\alpha_k = \cos^{-1}\left(\sqrt{s_k}\right) \tag{10}$$

and s_k is the kth largest eigenvalue in S (Krzanowski, 1979).

Angles between closest points in covariance coordinate systems are visually interpretable in two or three dimensions, and computational procedure can be extended to higher dimensional systems. A statistic could be developed to describe conical confidence thresholds for angles between coordinate axes of different models. However, such a statistical analysis is currently prohibitively complex beyond two or three dimensions (Watson, 1983). Instead, possible simple geometric benchmarks include minimum angle between models and sum of squares of the cosines of angles between model axes. For high-dimensional models, the minimum angle between models can quickly become trivial. However, successful discrimination generally requires consideration of higher dimensions, since considerable similarity between models is often found within the first few dimensions. The inclusion of similar lower dimension components can make the minimum angle as small as computational inaccuracy allows.

Similarity Index. An alternative model discrimination method can be developed by using the sum of cosines to define a fraction of similarity between models. Exact similarity can be easily described. Models of data with identical covariance structures for the same number of independent variables would correspond to coordinate axes P_1 and P_2 that are identical and orthonormal. Their products $P_1'P_2$ and $P_2'P_1$ are both equal to the identity matrix. The eigenvector decomposition described earlier would result in a set of eigenvalues s_k all equal to one for corresponding angles all equal to zero (Krzanowski, 1979). The sum of squares of cosines would then be equal to the number of dimensions considered. This suggests a general measure of similarity f as the sum of squares of cosines divided by the number of dimensions, p:

$$f = \sum_{k=1}^{p} s_k/p \tag{11}$$

The similarity index, f, ranges from zero to one, increasing as models become more similar. The similarity factor f provides a quantitative measure of difference in covariance directions between models and a description of overall geometric similarity in spread.

The similarity index can be used to evaluate discrimination models, a threshold value can be selected to indicate where mistakes in classification of data from the two models involved may occur. The similarity index can also be used to compare models built from different operating runs of the same process for monitoring systematic changes in process variation during normal operation. Another possible application is in batch processes, where use of the similarity index could provide a way to check if PCA model orientation around a moving mean varies in time.

Overlap of Means. The other important statistical test in comparing multivariate models is for differences in means. This corresponds to comparison of origin of coordinates rather than the coordinate directions. Many statistical tests have been developed for testing means, but most of them can become numerically unstable when significant correlation exists between variables. To work around the instability, overlap between eigenvalue-based models can be evaluated.

Target factor analysis methodology can place a likelihood on whether or not a candidate vector is a contributor to the model of a multivariate data set. A numerically viable approach (Malinowski, 1989) considers the situation from the distribution of quadratic forms, such as Mahalanobis weighting. A statistic is defined to test if a specific vector is significantly inside the confidence region containing the modeled data or if it could be responsible for an eigenvalue-based model. With respect to overlap of means, the test can be applied to determine whether the mean from one model, μ_1, significantly overlaps the region of data from another (second) model. In this case, the test statistic examines the estimate using the second model:

$$m = (rr')/[(1-\rho)t\Sigma^{-1}t']$$

where

$$t = (\mu_1 - \mu_2)P \qquad \text{and} \qquad r = tP' - \mu_1. \tag{12}$$

Here, t is the approximation of μ_1 by the second model, r is the residual error in μ_1 unexplained by the second model, P is the second model, transforming p measurements to k dimensions, $\boldsymbol{\Sigma}$ is the covariance in k model directions for the second model, μ_2 is the mean of n modeled observations in the second model, and ρ is the fraction of explained variation of data used to build the second model. Significant overlap is indicated when the statistic m is smaller than a threshold from the F distribution, $m_\alpha = F_{\alpha;(p-k),(n-k)}$. In chemometric applications, a confidence level of 95% has been satisfactorily used (Malinowski, 1989). Mean overlap analysis can be used to test if an existing PCA model fits a new set of observations or if two PCA models are analogous.

Diagnosis of multiple disturbances

Similarities between models for disturbances can be quantitatively measured by the methods presented. Comparison of models for individual disturbances and their combinations can provide information for extending the diagnosis methods to multiple simultaneous disturbances.

If there were no overlap between the regions spanned by two different disturbances, two alternative schemes might handle multiple disturbances modeled by PCA. In one method, the combination disturbance is idealized as being located between the regions of the underlying component disturbances: allocations of membership to the different independent disturbances contributing to the combination may provide diagnosis of underlying disturbances. The second is based on a more general extension of the discrimination scheme by introducing new models for each multiple-disturbance combination of interest. This approach could describe combinations of disturbances that produce more complicated models than the consensus of component models.

The measures of similarity in model center and direction of spread described earlier (Eqs. 11 and 12) can be useful to determine the independence of the models used in diagnosis. By comparing models for multiple disturbances to models of their components, similarity measures can give insight into the diagnosis of multiple disturbances and masking of contributing disturbances.

Masking of Multiple Disturbances. For realistic process data, it is reasonable to expect that models of single disturbances produce some overlap with each other. Inherent physical relations, such as mass and energy balances, are in effect in all process operation regimes. Further correlation is due to control schemes, which maintain certain levels by imposing relationships between variables. Similarity measures may point out that the process models exclude a few combinations of variables that describe these inherent physical relationships. Since such overlap is likely to exist for most processes under closed-loop control, the multiple-disturbance scenario is further complicated for such processes.

When the region spanned by the model for one (*outer*) disturbance contains the model for another (*inner*) disturbance, their combination will not be perfectly diagnosed. Idealizing the two disturbance regions as concentric spheres, the *inner* model region is enveloped by the *outer* model. As a result, only the *outer* disturbance will be diagnosed and the *inner* disturbance will be masked.

Random Variation or Noise Disturbances. Random-variation disturbances, which move a process less drastically off target than step or ramp disturbances, may be less successfully diagnosed in a single-disturbance scenario. Consequently, similarity measures should indicate that the random variation disturbances have more overlap with other models, particularly with each other. Similarity measures may also systematically indicate which random variation disturbance would be masked when multiple disturbances occur. Figure 8 shows an example of two random variation disturbances (+ and ×) and their combination (∗) plotted against the first two PCs for NOC of the process. The quantitative information is generated from the Tennessee Eastman process described in detail in the next section. Samples of operation with the combination of these two are enclosed within the data points of each of the single disturbances. In general, the *inner* disturbance is masked; its mean overlaps the region for the combination, which is most similar in spread to the *outer* disturbance model. Combinations of random variation disturbances with ramp or step disturbances manifest the same pattern. The ramp or step disturbances tend to be the *outer* models, this is consistent with moving the process off its control target or nominal operating point. As the *outer* model, the ramp or step disturbance masks the secondary random variation disturbances.

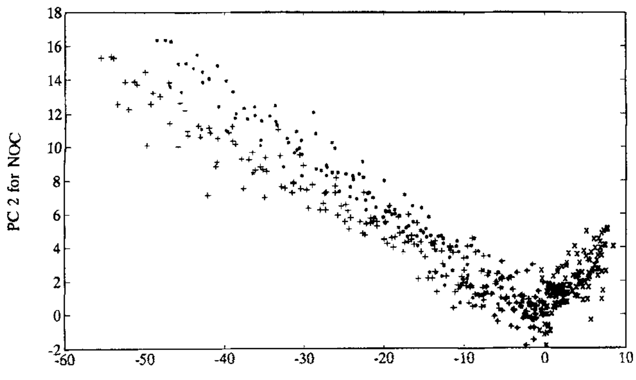

Figure 8. Trajectories of two individual ramp disturbances D (×), H (+), and their simultaneous occurrence (∗) plotted in the PC1 and PC2 plane for normal operation.

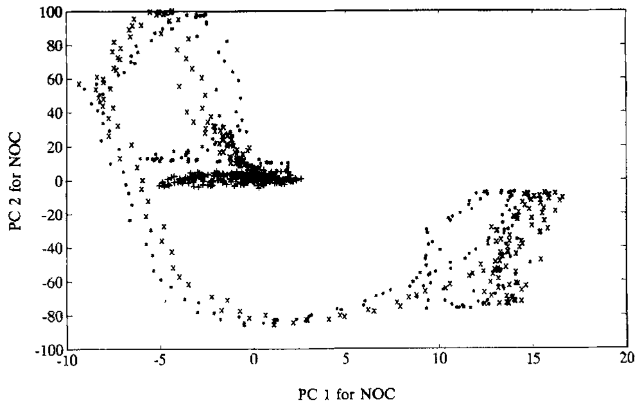

Figure 9. Trajectories of two individual random variation disturbances C (×), G (+), and their simultaneous occurrence (∗) plotted in the PC1 and PC2 plane for normal operation.

Using these patterns, a successful diagnosis method for multiple random-variation disturbances would be to incorporate only the single-disturbance models in the analysis framework. The individual disturbances will continue to be successfully diagnosed. When their combination occurs, the dominating *outer* disturbance would be diagnosed.

Step or Ramp Disturbances. In combination with each other, step and ramp disturbances show less predictable masking patterns. Figure 9 shows an example of a step and a ramp disturbance (+ and ×) and their combination (∗). Although some cases follow the simple *inner* and *outer* pattern, some models of single disturbances have mutually overlapping means and highly similar spreads, but the model of the combination disturbance does not overlap with either of the single-disturbance means.

A quantitative indicator of when this type of effect might occur is the angle between the means of disturbances, a and b, using a nominal in-control operating point as the vertex:

$$\cos(\theta) = (a'b)/(\|a\|\|b\|) \qquad \text{where} \qquad \|a\| = \sqrt{a'a}\,. \quad (13)$$

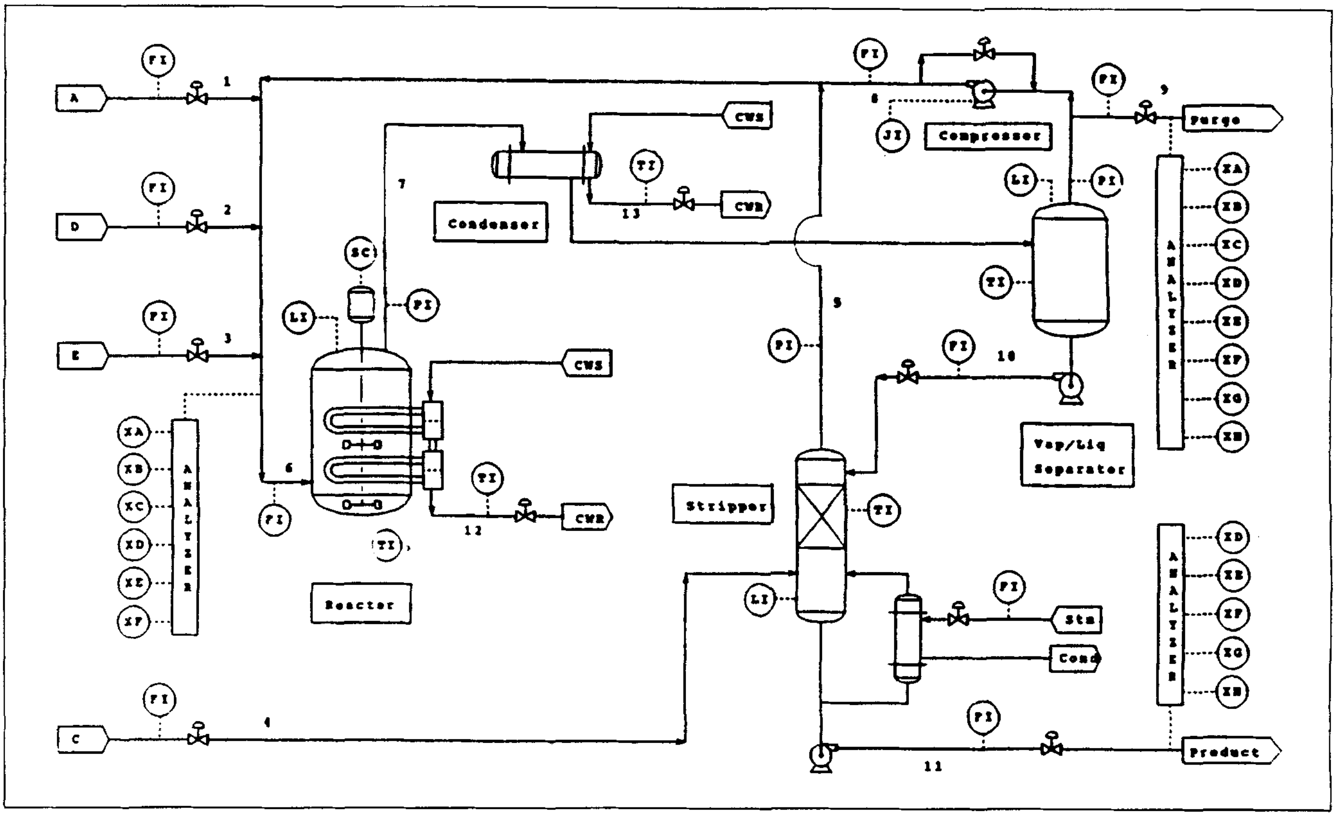

Figure 10. Tennessee Eastman test problem plant flowsheet.

As the cosine of the angle approaches zero, the effect becomes more pronounced and overlap of the combination with the component single-disturbance models is reduced. This is demonstrated in Figure 9, where samples of operation with the combination of two-step disturbances lie between the data points of the single disturbances, which are within 90 degrees of each other.

Since combinations of step or ramp disturbances are less likely to overlap with single-disturbance models, models of combinations may be included in the discrimination framework to correctly diagnose the occurrence of multiple disturbances.

Misdiagnosis of Additional Disturbances. In the analysis of random variation disturbances, multiple disturbance models are generally *outer* to single-disturbance models. As a result, secondary disturbances may be diagnosed when not actually present if multiple-disturbance models are included in the discrimination framework. Alternatively, if a single-disturbance model is *outer* to the multiple-disturbance model, the existence of more than one disturbance may not be detectable.

The angle between means of single- and multiple-disturbance models can also serve as a quantitative measure of the likelihood of such misdiagnosis. Overlap of the combination with the component single-disturbance models and misdiagnosis increase as the cosine of the angle approaches one. While large angular separation lessens the success of diagnosing the contributors to a combination disturbance by a scheme based on single-disturbance discrimination, a scheme including combination disturbances can be successfully implemented to diagnose these cases with higher success.

The similarity measures defined can serve as indicators of the success in diagnosing combinations of disturbances. The measures can identify which combinations of disturbances may be masked and which combinations may be falsely diagnosed. Similarity measures can provide information about the success rates of different diagnosis schemes incorporating single and combinations of disturbances. Using these guidelines, multiple disturbances occurring in a process can be analyzed *a priori* with respect to their components, and accommodated within the diagnosis framework described earlier.

Description of the Simulated Process and Data

The process described in the Tennessee Eastman Industrial Challenge Problem was formulated as a realistic model of an industrial chemical operation. Based on an actual proprietary process, the detailed Fortran model includes five major unit operations (reactor, condenser, separator, stripper, and compressor) with recycle, multiple inputs, and outputs (Figure 10). There are 22 continuous-process measurements (Table 1), 12 manipulated variables (Table 2), and 19 composition measurements sampled less frequently. Since the process is open-loop unstable, proportional-integral (PI) control loops were used to stabilize the system by pairing reactor pressure (Measurement 7) with flow rates of reactants (manipulated variables 1–5) and reactor temperature (Measurement 9) with coolant flow rate (manipulated variable 10). The proportional gain parameters (K_c) for the pressure loop were -0.02, -0.02, -0.03, -0.01, and 0.02, respectively, for manipulated variables 1–5, and $\tau_i = 2.4 \times 10^7$ s for all controllers. For the temperature loop, the PI controller parame-

Table 1. Tennessee Eastman Continuous Process Measurements

No.	Description	Stream
1	Feed flow rate	1
2	Feed flow rate	2
3	Feed flow rate	3
4	Feed flow rate	4
5	Recycle flow rate	8
6	Reactor feed rate	6
7	Reactor pressure	—
8	Reactor level	—
9	Reactor temperature	—
10	Purge rate	9
11	Product separator temperature	—
12	Product separator level	—
13	Product separator pressure	—
14	Product separator underflow	10
15	Stripper level	—
16	Stripper pressure	—
17	Stripper underflow	11
18	Stripper temperature	—
19	Stripper steam flow rate	14
20	Compressor work	—
21	Reactor coolant outlet temperature	12
22	Separator coolant outlet temperature	14

ters were $K_c = 20$, $\tau_i = 2.4 \times 10^7$ s. A good plantwide control would necessitate more sophisticated multivariable control strategies. Data generated during simulation depends on the control system used. Consequently, the control strategies utilized may affect the successful diagnosis rate of specific disturbances. The focus of the present study is disturbance diagnosis, and the Tennessee Eastman process is used for illustrating the detection and diagnosis methodologies. Consequently, an elaborate control system design was not attempted. In an actual implementation, obviously one should develop a satisfactory control system and then generate data for detection and diagnosis methods. For example, the same disturbances with a different controller scheme may still be marginally detectable (De Veaux et al., 1994).

For this study, data were recorded at 10-minute intervals in 40-hour runs. Normal variation during operation was introduced as pseudo random binary sequence (PRBS) inputs of less than 2% of nominal on each of the separate manipulated variables and control setpoints. Figure 11 shows samples of typical input variations and measurements: steam flow rate to stripper (input 9), reactor coolant flow rate (input 10), stripper temperature (measurement 18), and reactor outlet temperature (measurement 21). Some of the measurements are

Table 2. Tennessee Eastman Manipulated Variables

No.	Description	Stream
1	Feed flow	2
2	Feed flow	3
3	Feed flow	1
4	Feed flow	4
5	Compressor recycle valve	—
6	Purge valve	9
7	Separator pot liquid flow	10
8	Stripper liquid product flow	11
9	Stripper steam valve	14
10	Reactor cooling water flow	12
11	Condenser cooling water flow	13
12	Agitator speed	—

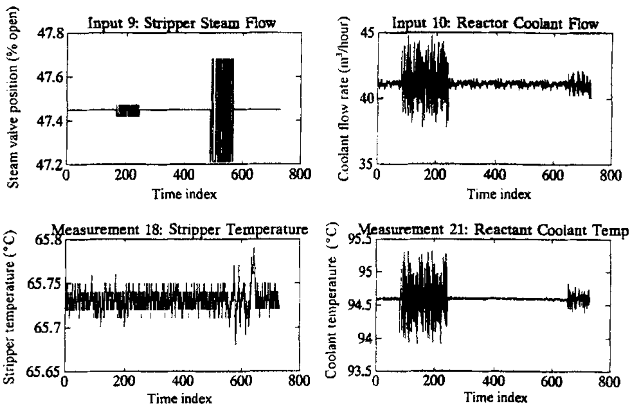

Figure 11. Sample manipulated inputs and process outputs for modeling normal operation.

Input variations: (a) stripper steam flow and (b) reactor coolant flows, output responses: (c) stripper temperature and (d) reactor coolant outlet temperature.

significantly correlated, as shown by the pattern of their scatter when plotted against each other. Process measurements show significant correlation, as demonstrated between separator temperature and separator coolant outlet temperature (Figure 12). While the process remains in a region of normal operation, most of the measurements show only weak autocorrelation, with much of the autocorrelation appearing in the earlier dimensions of PCs (Figure 13a). After the first five lags, autocorrelation under normal operation is almost small enough to be insignificant. Considering the input signals with persistent variation (PRBS), some autocorrelation is expected in the data for operation in NOC. Since the contributions from these autocorrelated scores to the overall distance are small, Hotelling's statistic is still an acceptable approximation.

Additional data were generated with 20 separate process upsets and one combined upset, including step, ramp, and random disturbances (Table 3). Figure 13b shows an example where autocorrelation of measurements under process upsets (demonstrated with upset 7) is considerably larger than under normal operation; significant autocorrelation along at least the first PC from the NOC model was seen with all process upsets.

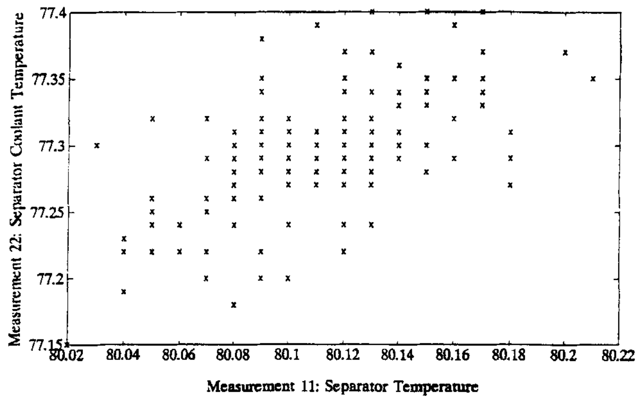

Figure 12. Correlation between measurements under normal operation.

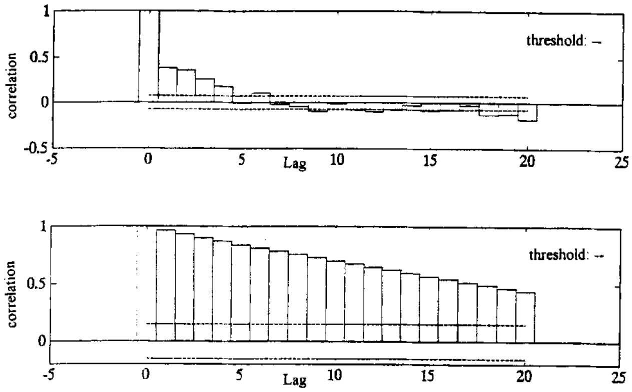

Figure 13. Autocorrelation on first PC for normal operating conditions: (a) under normal operating conditions; (b) under upset 7.

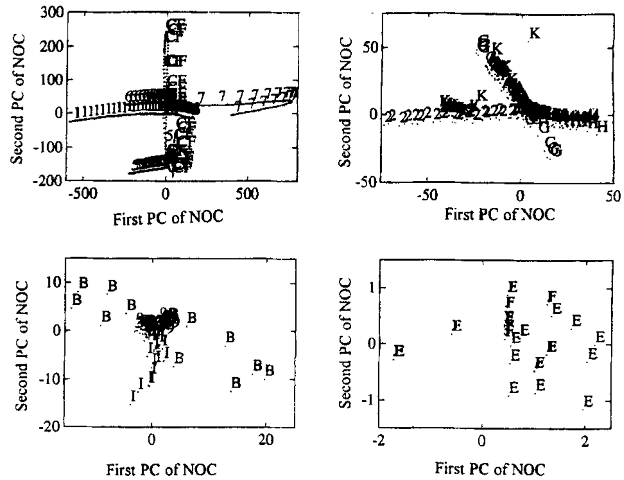

Figure 14. 21 Types of upsets as biplots in the plane of PC1 and PC2 of NOC.

Coordinate scales are reduced to zoom on the origin of the plots.

The need to include dynamic effects into the PC models depends on the severity of autocorrelation. A pseudo-steady-state approach may be adequate particularly with multiple measurements (Wise et al., 1990). Autocorrelation can be explicitly incorporated into time series models estimated by PC methods on combinations of lagged variables (Ku et al., 1994). Although the step, ramp, and noisy process upsets studied have definite dynamic components during departure from the NOC region, they are detectable using a simple PC model. Detection against the NOC model is not sufficient to pinpoint diagnosis; since all upsets studied showed similar autocorrelation on at least the first PC, even a model incorporating NOC dynamics would not necessarily provide an adequate basis for diagnosis.

PC models for each of the individual upsets provide the potential for comparisons useful for diagnosis. Observations from step and ramp upsets tend to form clusters away from the NOC region, such as disturbances 1, 2, and 7 (Figure 14a,b). Observations from random variation upset tend to overlap each other and the NOC region such as disturbances 9, B, E, and F (Figure 14c,d). In this scenario, PC models for the step and ramp upsets may be well separated even without taking dynamics into account. Random-variation upsets, which are likely to have less autocorrelation, may also be adequately described by simple PC models, although their visual discrimination from biplots is challenging (Figure 14d).

Discussion of Results

Detection and diagnosis of single disturbances

The method proposed was able to detect all 21 types of process upsets with different rates of success. Detection rates using residuals and scores from the NOC model are listed in Table 4. For all types of upsets, score detection was superior to residual detection. This is due to the fact that the particular upsets studied lie in the directions described by scores,

Table 3. Tennessee Eastman Process Disturbances

No.	Disturbance Description	Stream	Type
1	Feed ratio	4	Step
2	Feed composition	4	Step
3	Feed temperature	2	Step
4	Reactor coolant inlet temperature	12	Step
5	Condenser coolant inlet temperature	13	Step
6	Feed loss	1	Step
7	Feed header pressure loss	4	Step
8	Feed composition	4	Random
9	Feed temperature	2	Random
A	Feed temperature	4	Random
B	Reactor coolant inlet temperature	12	Random
C	Condenser coolant inlet temperature	13	Random
C and F	Random condenser coolant inlet temperature and sticking condenser coolant valve	13	Random
D	Drift in reaction kinetics	—	Ramp
E	Sticking reactor coolant valve	12	Random
F	Sticking condenser coolant valve	13	Random
G	Unknown fault	—	Random
H	Unknown fault	—	Ramp
I	Unknown fault	—	Ramp
J	Unknown fault	—	Random
K	Unknown fault	—	Random

Table 4. % Diagnosis of Disturbances

Fault	Type	Residuals	Scores
1	Step	58	98
2	Step	0	93
3	Step	0	6
4	Step	0	99
5	Step	20	99
6	Step	100	100
7	Step	91	100
8	Random	0	83
9	Random	0	17
A	Random	0	89
B	Random	0	55
C	Random	0	90
C and F	Random	0	90
D	Ramp	0	58
E	Random	0	5
F	Random	0	3
G	Random	0	93
H	Ramp	0	68
I	Ramp	0	71
J	Random	0	3
K	Random	0	4

Table 5. % Success in Diagnosis of Disturbances with New Data

Disturbance	Type	Score	Residual	Combination
1	Step	94	0	89
2	Step	41	0	0
3	Step	0	100	92
4	Step	0	0	0
5	Step	37	0	48
6	Step	70	0	73
7	Step	93	0	51
8	Random	1	0	0
9	Random	14	0	0
A	Random	57	0	25
B	Random	0	0	0
C	Random	7	0	9
C&F	Random	33	0	41
D	Ramp	0.7	0	0
E	Random	0	100	62
F	Random	0	14	0
G	Random	33	0	13
H	Ramp	76	0	66
I	Ramp	0	0	0
J	Random	0	14	0
K	Random	73	0	0.9

and indicates that the NOC model developed includes most types of phenomena that are likely to occur. The types of upsets that are least detectable are random variations, and only their extreme ranges are likely to be outside the normal operating region. In general, step disturbances affect the mean value of a product variable, random variations affect its spread, and ramp changes influence both its mean and spread. Prevention of changes in the mean is frequently the main concern. Consequently, the diagnosis scheme is relatively successful in providing information about the more crucial class of disturbances.

Observations under various upset conditions are shown in Figure 14 by plotting data from several types of upsets on the plane of the first two PCs. Various degrees of zooming in on the origin ($PC_1 = 0$, $PC_2 = 0$) shows the extent of data cluttering for various disturbances. Visual classification is not practical. Diagnosis of source cause using individual disturbance models achieved different levels of success for different discriminants and types of disturbances.

The success rates in diagnosing the correct source cause of observations under different upsets are listed in Table 5 for data not used in model building. Only a few upsets were well diagnosed using the residual discriminant. Comparison of Tables 4 and 5 shows that these upsets had been poorly detected. In fact, the small sample size caused the PCA models for these disturbances to be statistically unreliable.

The criteria used for choosing dimensionality (Eq. 2) may also be suspect for these cases, especially since the resultant residual thresholds were unusually large. Since the score discriminant was less sensitive to this type of effect, using different numbers of dimensions can improve diagnosis. A more comprehensive guard against such bias would be to ensure that all disturbance models are built using data records of comparable size.

Table 6 compares use of the different discriminants with data used in training and separate test sets for step or ramp upsets and random disturbances. Trends common to the two sets are visible. The score discriminant outperformed the other discriminants for both step/ramp and noise disturbances, both with training and testing data sets. A combined discriminant was less successful than the score discriminant, mostly due to worse diagnosis of source causes of random disturbances. Average run length (ARL) studies with many sets of data collected from Monte-Carlo simulations would offer more detailed information, but the deterministic nature of the simulation program made generation of many testing sets infeasible. Typical run lengths for detection of the out-of-control condition and for diagnosis of the correct disturbance are given in Table 7. In general, step- and ramp-type disturbances were more easily diagnosed, with a success rate of around 50%; some upsets were over 90% correctly diagnosed. Noise or random-type disturbances were less easily diagnosed, with only one in three correctly diagnosed.

Table 6. Comparison of % Success of Diagnosis of 21 Process Upsets with Different Methods

	Overall		Step/Ramp		Random	
Method	Training	Testing	Training	Testing	Training	Testing
Score	47	40	51	47	43	30
Combination	35	26	45	36	23	13
Residual	0.8	0.7	0.8	0.8	0.9	0.7

Diagnosis of multiple simultaneous disturbances

Comparison of Models. A low similarity index was found to correctly indicate few misclassifications between groups. The actual diagnosis between samples from 21 disturbances is shown in Table 8. The actual disturbance introduced to the process is listed in the first column. The other columns of Table 8 indicate the disturbance diagnosed by the proposed methodology. For example, the first row shows that disturbance 1 was diagnosed 230 times as disturbance 1, 7 times as disturbance 15, and once as disturbance 21.

Comparing Table 8 to Table 9, which shows the similarity indexes between the groups, 80% of diagnosed points fall into a bin with similarity index greater than 0.6. For example, row 5 in Table 8 lists 68 points correctly diagnosed as disturbance 5, 63 incorrectly diagnosed as disturbance 12, 108 misdiagnosed as disturbance 13, and 1 point misclassified as disturbance 21. The corresponding similarity indexes are 1, 0.94, 0.94, and 0.91, which are all higher than the 0.6 cutoff and the corresponding models are among the models most similar to the model for disturbance 5.

The remaining misclassifications are generally into groups 15, 16 and 20. Notable misclassifications in the absence of high-similarity indexes include misdiagnosis of observations as belonging to groups 15, 16 and 20, which had low-similarity indexes to most other models. These upsets were poorly

Table 7. Sample Run Lengths for Detection and Diagnosis of a Disturbance

	Detection		Diagnosis	
Disturbance	Training	Testing	Training	Testing
3	10	5	11	8
6	1	1	33	54
E	27	27	28	28
G	18	18	28	28

Table 8. Matrix of Diagnosis of Disturbances to Tennessee Eastman Control Problem

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	230	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	1
2	0	185	0	0	0	0	0	0	0	0	1	1	0	0	6	16	0	0	0	16	16
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	22	0	0	0	22	0
4	0	0	0	57	0	0	0	0	0	0	167	0	0	0	1	0	1	8	0	0	0
5	0	0	0	0	68	0	0	0	0	0	0	63	108	0	0	0	0	0	0	0	1
6	0	0	0	0	0	241	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	194	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	71	0	0	0	0	0	0	113	0	0	0	0	0	0	0	16	0	0	0	16	0
9	0	0	0	0	0	0	0	0	1	1	0	0	0	0	22	27	0	0	0	27	0
10	0	0	0	0	0	0	0	0	0	126	3	2	2	0	3	2	24	0	1	2	0
11	0	0	0	0	0	0	0	0	0	0	77	0	2	0	10	6	1	0	0	6	2
12	0	0	0	0	0	0	0	0	0	0	0	45	169	0	0	3	0	0	0	3	2
13	0	0	0	0	0	0	0	0	0	0	0	45	169	0	0	3	0	0	0	3	2
14	0	0	0	0	0	0	0	1	0	0	5	13	25	53	4	14	0	0	0	14	7
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	6	0	0	0	6	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	6	0	0	0	6	0
17	0	0	0	0	0	0	0	0	0	39	3	0	0	0	15	3	151	0	0	3	0
18	0	0	0	0	0	0	0	0	0	0	4	2	0	0	5	11	1	184	0	11	1
19	0	0	0	0	0	0	0	0	0	0	0	42	17	0	17	1	0	0	91	1	3
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	6	0	0	0	6	0
21	0	0	0	0	0	0	0	1	0	0	0	5	5	0	2	7	0	0	0	7	192

Entries indicate the number of observations from disturbance indicated in the first column that were diagnosed as belonging to the disturbance indicated by the column number. Entries off the diagonal are misdiagnosed.

detected, and consequently the groups did not have an appropriate amount of data available. Insufficient sample size causes these models to be statistically unreliable and affects interpretation of the similarity index. An illustration of this fact is observed by comparing the ninth row in Table 8 and Table 9, where 1 point is misdiagnosed as disturbance 10, 22 points as disturbance 5, 27 points as disturbance 16, and 27 points as disturbance 20. Of these misclassifications, Table 9 shows only model 10 to have a high-similarity index. Adding up the number of observations classified along rows in Table 8, it can be seen that groups 15, 16 and 20 had less than 20, only one-tenth of the average number of samples per group. To avoid misinterpretation of similarity index, models should be screened to ascertain that similar numbers of observations are used for all groups.

Comparing Tables 9 and 10, calculations of mean overlap were not as strong an indicator of pairwise misclassification as the similarity index. As Figure 7 has demonstrated, similarity of covariance is more crucial for distinguishing between points from two groups. In analyzing many groups, however, the test of mean overlap was more useful as a measure of overall similarity. By counting the number of groups with

Table 9. Similarity Factor Between Disturbance Models for Tennessee Eastman Control Problem

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	1	0.93	0.71	0.92	0.91	0.97	0.93	0.98	0.69	0.66	0.89	0.89	0.90	0.89	0.06	0.00	0.81	0.91	0.85	0.00	0.88
2	0.93	1	0.74	0.89	0.78	0.93	0.92	0.93	0.73	0.77	0.91	0.89	0.89	0.94	0.06	0.00	0.85	0.88	0.93	0.00	0.84
3	0.71	0.74	1	0.73	0.73	0.69	0.70	0.71	0.71	0.68	0.75	0.71	0.71	0.75	0.32	0.25	0.73	0.73	0.74	0.25	0.69
4	0.92	0.89	0.73	1	0.92	0.92	0.93	0.92	0.68	0.74	0.92	0.94	0.94	0.93	0.07	0.00	0.81	0.93	0.88	0.00	0.88
5	0.91	0.78	0.73	0.92	1	0.91	0.93	0.90	0.65	0.68	0.88	0.94	0.94	0.88	0.00	0.00	0.76	0.90	0.82	0.00	0.91
6	0.97	0.93	0.69	0.92	0.91	1	0.96	0.98	0.69	0.67	0.89	0.90	0.90	0.89	0.00	0.00	0.77	0.90	0.87	0.00	0.88
7	0.93	0.92	0.70	0.93	0.93	0.96	1	0.94	0.66	0.55	0.89	0.93	0.93	0.89	0.00	0.00	0.76	0.91	0.89	0.00	0.89
8	0.98	0.93	0.71	0.92	0.90	0.98	0.94	1	0.68	0.67	0.91	0.89	0.89	0.92	0.00	0.00	0.77	0.92	0.88	0.00	0.87
9	0.69	0.73	0.71	0.68	0.65	0.69	0.66	0.68	1	0.68	0.71	0.66	0.66	0.71	0.35	0.22	0.76	0.76	0.74	0.22	0.65
10	0.66	0.77	0.68	0.74	0.68	0.67	0.55	0.67	0.68	1	0.78	0.72	0.72	0.79	0.00	0.00	0.84	0.80	0.78	0.00	0.67
11	0.89	0.91	0.75	0.92	0.88	0.89	0.89	0.91	0.71	0.78	1	0.92	0.92	0.95	0.00	0.00	0.85	0.95	0.90	0.00	0.88
12	0.89	0.89	0.71	0.94	0.94	0.90	0.93	0.89	0.66	0.72	0.92	1	0.99	0.88	0.09	0.07	0.80	0.94	0.85	0.07	0.91
13	0.90	0.89	0.71	0.94	0.94	0.90	0.93	0.89	0.66	0.72	0.92	0.99	1	0.88	0.11	0.07	0.80	0.94	0.84	0.07	0.91
14	0.89	0.94	0.75	0.93	0.88	0.89	0.89	0.92	0.71	0.79	0.95	0.88	0.88	1	0.09	0.10	0.87	0.94	0.93	0.10	0.85
15	0.06	0.06	0.32	0.07	0.00	0.00	0.00	0.00	0.35	0.00	0.00	0.09	0.11	0.09	1	0.88	0.00	0.00	0.06	0.88	0.00
16	0.00	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.22	0.00	0.00	0.07	0.07	0.10	0.88	1	0.07	0.00	0.00	1.00	0.00
17	0.81	0.85	0.73	0.81	0.76	0.77	0.76	0.77	0.76	0.84	0.85	0.80	0.80	0.87	0.00	0.07	1	0.88	0.91	0.07	0.77
18	0.91	0.88	0.73	0.93	0.90	0.90	0.91	0.92	0.76	0.80	0.95	0.94	0.94	0.94	0.00	0.00	0.88	1	0.89	0.00	0.90
19	0.85	0.93	0.74	0.88	0.82	0.87	0.89	0.88	0.74	0.78	0.90	0.85	0.84	0.93	0.06	0.00	0.91	0.89	1	0.00	0.84
20	0.00	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.22	0.00	0.00	0.07	0.07	0.10	0.88	1.00	0.07	0.00	0.00	1	0.00
21	0.88	0.84	0.69	0.88	0.91	0.88	0.89	0.87	0.65	0.67	0.88	0.91	0.91	0.85	0.00	0.00	0.77	0.90	0.84	0.00	1

Factors close to 1 indicate a high similarity between the directions of the models. The matrix is symmetric since comparison of the models designated in rows and columns is insignificant.

Table 10. Matrix of Overlap of Means for Disturbances to Tennesse Eastman Control Problem

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	1	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
2	0	1	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	1	1	1	1	1	1	0	0	0	0	1	1	0	0	0	0	1	0	0	1
4	1	0	0	1	0	1	1	1	0	0	1	0	1	1	0	0	0	0	0	0	1
5	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
6	1	0	0	0	1	1	1	1	1	0	0	1	1	1	0	0	0	1	1	0	1
7	1	0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0
8	0	0	0	0	0	1	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0
9	1	0	0	1	0	1	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0
10	0	1	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
11	1	1	0	1	0	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	1
12	0	0	1	0	1	1	1	0	0	0	0	1	0	1	0	0	0	0	0	0	1
13	0	0	1	0	1	1	1	0	0	0	0	0	1	1	0	0	0	1	0	0	1
14	1	0	1	0	1	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0
15	1	1	0	1	1	1	1	1	0	1	1	0	0	1	1	0	1	1	1	0	1
16	1	1	0	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	0	1	1
17	0	1	0	0	0	1	1	1	0	0	0	1	1	0	0	0	1	0	0	0	0
18	0	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0
19	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0
20	1	1	0	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	0	1	1
21	0	0	1	0	1	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1

Zero indicates no overlap, 1 indicates that the mean of the disturbance indicated down the column is significantly overlapping the model of disturbance indicated across the corresponding row.

overlapping means, an estimate of overall success in diagnosis was calculated as 80%. Comparing the estimate to discrimination rules of several formats with realistic data, this estimate was found to approach the upper limit on overall success in classification, which was around 70%.

The data sets used with these similarity measures included over 20 groups of more than 20 correlated variables, generally with over 200 observations in each group. A set of such large dimensions is beyond the scope of most simplifying assumptions, enabling a theoretically rigorous statistical estimate. By completely random diagnosis, a success rate of less than 5% would be expected between over 20 groups.

Diagnosis of Multiple Disturbances. Success rates in classification were maintained when including four combinations of two disturbances in the discrimination framework. When compared to 20 single-disturbance models, these disturbances were identified as one of their contributing upsets. To identify the second contributing disturbance, additional PCA models of these combinations were added to the discriminant pool.

Using similarity indexes and mean overlap, models of step or ramp combinations overlapped less with other models than combinations of random upsets. Consequently, step/ramp combinations were better diagnosed than random combinations. Table 10 shows success rates using the single- and multiple-disturbance models to diagnose combinations of step and ramp disturbances. Combinations of step upsets 6, 7, and H and ramp upset D were correctly identified for over 60% of samples, more than the average of all single disturbances. The combination of random disturbances C and F, which had been included in the base case reported in Tables 4–6 and 11, was successfully classified 41% of the time.

Table 11. % Successful Diagnosis for Multiple Disturbances

Disturbance Combinations	Scores	Combined
6 and 7 (step and step)	96	92
7 and D (step and ramp)	96	85
D and H (ramp and step)	72	62
C and F (random and random)	33	41

Conclusions

The multivariate statistical tools presented are useful in monitoring performance of a large-scale continuous process. They utilize charts similar to Shewhart charts as opposed to biplots that can be used only when a few PCs are enough to describe the process behavior. Step, ramp, and random upsets in operation are detectable using a simple linear model based on historical records of normal operating conditions. Additional simple statistical models of operation under various upsets can be constructed and used to diagnose possible source causes of unusual process behavior. The method is consistent when tested with new data sets. It is suitable for simultaneously monitoring many variables. Diagnosis between a large group of disturbances, with nearly as many possibilities as measured variables, can be performed. Step and ramp upsets that would move the process off target are more easily diagnosed than random-variation disturbances, which widen the process quality spread without shifting the mean.

Methods have been proposed to assess the capability to discriminate multiple disturbances that occur simultaneously. This necessitates the development of measures for evaluating the discrimination capability between different disturbances. Such measures are also useful for evaluating the confidence in discriminating *a priori* between any two disturbances when single disturbances are assumed to occur.

Besides providing quantitative comparisons to the outcomes of different diagnosis schemes, the proposed methods proved useful in evaluating models for a realistic process. By quantifying overlap between models, limits on the usefulness of the models could be found. The tools could prove to be valuable to evaluate models of other processes with changing statistical characteristics or different operating strategies. Providing general quantitative methods for comparison, the techniques presented here are able to accommodate large-sized data sets with correlated variables.

Multivariate statistical techniques offered valuable insight in the operation of the plantwide Tennessee Eastman process. Summarizing data in PCA models, a variety of single and combination process upsets could be handled. Success in diagnosis depended on the nature of the upset. Similarity measures or sample-size arguments are useful in predicting when diagnosis tools may not be successful.

Acknowledgments

The authors acknowledge the Illinois Institute of Technology Research Institute and the American Association of University Women who provided funding in support of this work. The authors are grateful to Professor Larry Ricker for supplying the MEX files linking the Fortran simulation routines and MATLAB.

Notation

h_o = intermediate term used to calculate residual threshold
P_1, P_2 = PCA loading matrix for models 1 and 2
P = matrix of PCA model loadings
P_i = PCA loading matrix for ith disturbance
s_α = $100\alpha\%$ confidence threshold for Hotelling's T^2 statistic on scores
x = single measurement of all variables as a column vector
x_1, x_2 = observations of variables 1 and 2
θ_2 = intermediate term used to calculate residual threshold
θ_3 = intermediate term used to calculate residual threshold
Σ = eigenvalues of scores as diagonal matrix
τ_i = controller integral time constant

Subscripts

i = index of disturbance under consideration as diagnosis
k = index of PC direction under consideration for inclusion the model

Literature Cited

Alt, F. B., "Multivariate Quality Control," *Encyclopedia of Statistical Sciences*, S. Kotz and N. L. Johnson, eds., Wiley, New York (1985).
Alwan, L. C., "Cusum Quality Control-Multivariate Approach," *Commun. Stat.—Theory Methods*, **15**, 3531 (1986).
Crosier, R. B., "Multivariate Generalizations of Cumulative Sum Quality-Control Schemes," *Technometrics*, **30**, 291 (1988).
De Veaux, R. D., J. Vinson, and L. H. Ungar, "Fault Diagnosis Using PCA and PLS," AIChE Meeting, Paper 231d, San Francisco (1994).
Downs, J. J., and E. F. Vogel, "A Plant-Wide Industrial Control Problem," AIChE Meeting, Paper 24a, Chicago (1990).
Doğanaksoy, N., F. W. Faltin, and W. T. Tucker, "Identification of Out of Control Quality Characteristics in a Multivariate Manufacturing Environment," *Commun. Stat.—Theory Methods*, **20**, 2775 (1991).
Frank, P. M., "Disturbance Diagnosis in Dynamic Systems Using Analytical and Knowledge-Based Redundancy," *Automatica*, **26**, 459 (1990).
Fukunaga, K., *Statistical Pattern Recognition*, 2nd ed., Academic Press, San Diego, CA (1990).
Hawkins, D. M., "Multivariate Quality Control Based on Regression Adjusted Variables," *Technometrics*, **33**, 61 (1991).
Healy, J. D., "A Note on Multivariate CUSUM Procedures," *Technometrics*, **29**, 409 (1987).
Himes, D. M., R. H. Storer, and C. Georgakis, "Determination of the Number of Principal Components for Disturbance Detection and Isolation," *Proc. Amer. Control Conf.*, 1279 (1994).
Hoskuldsson, A., "PLS Regression Models," *J. Chemometrics*, **2**, 211 (1988).
Jackson, J. E., *A User's Guide to Principal Components*, Wiley, New York (1992).
Jackson, J. E., and G. S. Mudholkar, "Control Procedures for Residuals Associated with Principal Components Analysis," *Technometrics*, **21**, 341 (1979).
Johnson, R. A., and D. W. Wichern, *Applied Multivariate Statistical Analysis*, 3rd ed., Prentice-Hall, Englewood Cliffs, NJ (1992).
Kresta, J., J. F. MacGregor, and T. Marlin, "Multivariate Statistical Monitoring of Process Operating Performance," *Can. J. Chem. Eng.*, **69**, 35 (1991).
Krzanowski, W. J., "Between-Groups Comparison of Principal Components," *J. Amer. Stat. Assn.*, **74**, 703 (1979).
Ku, W., R. H. Storer, and C. Georgakis, "Disturbance Detection and Isolation by Dynamic Principal Components Analysis," Int. Chemometrics Internet Conf., Session No. 8 (1994).
Lowry, C. A., W. H. Woodall, C. W. Champ, and S. E. Rigdon, "A Multivariate Exponentially Weighted Moving Average Control Chart," *Technometrics*, **34**, 46 (1992).
Malinowski, E., "Statistical F-Tests for Abstract Factor Analysis and Target Testing," *J. Chemometrics*, **3**, 49 (1989).
McLachlan, G. J., *Discriminant Analysis and Statistical Pattern Recognition*, Wiley, New York (1992).
Montgomery, D. C., *Introduction to Statistical Quality Control*, 2nd ed., Wiley, New York (1991).
Naes, T., and T. Isaksson, "Splitting of Calibration Data by Cluster Analysis," *J. Chemometrics*, **5**, 49 (1991).
Negiz, A., and A. Çinar, "On the Detection of Multiple Sensor Abnormalities in Multivariate Processes," *Proc. Amer. Control Conf.*, p. 2364, Chicago (1992).
Pignatiello, J. J., and G. C. Runger, "Comparisons of Multivariate CUSUM Charts," *J. Qual. Technol.*, **22**, 173 (1990).
Raich, A., and A. Çinar, "Multivariate Statistical Methods for Monitoring Continuous Processes: Assessment of Discrimination Power of Disturbance Models and Diagnosis of Multiple Disturbances," *Chemometrics Intell. Lab. Syst.*, **30**, 37 (1995).
Skagerberg, B., J. F. MacGregor, and C. Kipparisides, "Multivariate Data Analysis Applied to Low Density Polyethylene Reactors," *Chemometrics Intell. Lab. Syst.*, **14**, 341 (1992).
Watanabe, K., S. Hirota, L. Hou, and D. M. Himmelblau, "Diagnosis of Multiple Simultaneous Faults via Hierarchical Artificial Neural Networks," *AIChE J.*, **40**, 839 (1994).
Watson, G. S., *Statistics on Spheres*, Wiley, New York (1983).
Wetherill, G. B., and D. W. Brown, *Statistical Process Control Theory and Practice*, Chapman & Hall, New York (1991).
Wise, B., N. L. Ricker, D. F. Veltkamp, and B. R. Kowalski, "A Theoretical Basis for the Use of Principal Component Models for Monitoring Multivariate Processes," *Process Contr. Quality*, **1**, 41 (1990).
Wold, S., K. Esbensen, and P. Geladi, "Principal Components Analysis," *Chemometrics Intell. Lab. Syst.*, **2**, 37 (1987).
Woodall, W. H., and M. M. Ncube, "Multivariate CUSUM Quality-Control Procedures," *Technometrics*, **27**, 285 (1985).

Manuscript received Dec. 19, 1994, and revision received May 30, 1995.