Multivariate SPC Methods for Process and Product Monitoring

THEODORA KOURTI and JOHN F. MacGREGOR

McMaster University, Hamilton, ON L8S 4L7, Canada

Statistical process control methods for monitoring processes with multivariate measurements in both the product quality variable space and the process variable space are considered. Traditional multivariate control charts based on χ^2 and T^2 statistics are shown to be very effective for detecting events when the multivariate space is not too large or ill-conditioned. Methods for detecting the variable(s) contributing to the out-of-control signal of the multivariate chart are suggested. Newer approaches based on principal component analysis and partial least squares are able to handle large ill-conditioned measurement spaces; they also provide diagnostics which can point to possible assignable causes for the event. The methods are illustrated on a simulated process of a high pressure low density polyethylene reactor, and examples of their application to a variety of industrial processes are referenced.

Introduction

The objective of statistical process control (SPC) is to monitor the performance of a process over time in order to detect any unusual events that may occur. By finding assignable causes for them, improvements in the process and in the product quality can be achieved by eliminating the causes or improving the process or its operating procedures. Most SPC procedures currently in practice are based on charting only a small number of variables, usually the final product quality variables, $\mathbf{y}$. These approaches are often inadequate for modern process industries where massive amounts of data are being collected continually on perhaps hundreds of process variables, $\mathbf{x}$. Measurements on temperatures, pressure, flowrates etc., are available every few seconds. However, final product quality variables, $\mathbf{y}$, such as polymer molecular weights or melt index, cut points in distillations, etc., are usually available on a much less frequent basis. In any effective scheme for monitoring and diagnosing operating performance, both process and product data should be used to extract information.

Dr. Kourti is a Senior Research Engineer in the McMaster Advanced Control Consortium, in the Chemical Engineering Department.

Dr. MacGregor is a Professor in the Chemical Engineering Department and in the McMaster Advanced Control Consortium. He is a Member of ASQC.

Neither the process variables nor the product quality variables are independent of one another. Only a few underlying events are driving a process at any time, and measurements on all these variables are simply different reflections of the same underlying events. Multivariate methods that treat all the data simultaneously can extract information on the *directionality* of the process variations, that is on how all the variables are behaving relative to one another. When special events occur in processes they affect not only the magnitude of the variation in variables but also their relationship to each other (i.e., the direction of variation). These events are often difficult to detect by looking only at the magnitude of each process variable independently because the signal-to-noise ratio is very low in each variable. Multivariate methods can extract confirming information from observations on many variables and can reduce the noise levels through averaging.

In this paper we examine the application of some traditional multivariate statistical process control charts to monitoring process variables as well as final product quality variables. We present some new developments, based on contribution plots, for detecting the variable(s) that caused an out-of-control signal in multivariate charts. We then relate the traditional approaches for monitoring and diagnosis to approaches based on multivariate statistical projection methods such as principal component analysis (PCA) and partial least squares (PLS). Both of the above approaches are applied for monitoring and diagnosis of a simulated case of a high pressure low density polyethylene (LDPE) reactor. The particular simulation has been chosen because it can serve as a common example to all the methods discussed in this paper. A large number of applications of the projection methods to monitoring a variety of industrial processes (from the petrochemical to polymer and steel industry) are referenced in the text, so that readers may find examples closer to their interests.

Traditional Multivariate SPC Approaches

The SPC approach for process monitoring, currently in practice in several industries, is to chart a small number of variables, usually the final product quality variables ($\mathbf{y}$), and examine them one at a time (Shewhart, cumulative sum (CUSUM) or exponentially-weighted moving average (EWMA) charts). However, when the quality of a product is defined by more than one property, all the properties should be studied collectively. Multivariate SPC charts developed for this purpose have been based on the chi-squared statistic or on Hotelling's T^2 statistic (Hotelling (1947), Alt (1985), Jackson (1980, 1991), Montgomery (1996)).

Given a vector of measurements $\mathbf{z}$ on n normally distributed variables, with an in-control covariance matrix $\mathbf{\Sigma}$, one can test whether the current mean of the multivariate process is at its population mean $\boldsymbol{\mu}$ by computing the statistic

$$\chi^2 = (\mathbf{z} - \boldsymbol{\mu})'\mathbf{\Sigma}^{-1}(\mathbf{z} - \boldsymbol{\mu}). \quad (1)$$

This statistic will be distributed as a central chi-squared distribution with n degrees of freedom if the mean is equal to $\boldsymbol{\mu}$. A multivariate chi-squared chart can be constructed by plotting χ^2 versus time with an upper control limit (UCL) given by $\chi^2_{\alpha,n}$, where α is an appropriate significance level for performing the test (e.g., $\alpha = 0.01$ or 0.05). The χ^2 statistic in (1) represents the directed or weighted distance (Mahalanobis distance) of any point from $\boldsymbol{\mu}$.

For two variables z_1, z_2, if we set (1) equal to $\chi^2_{\alpha,n}$ the solution on z_1, z_2-space is an ellipse centered at $\boldsymbol{\mu}$. All the points on the ellipse have the same Mahalanobis distance, $\chi^2_{\alpha,n}$. In this case, instead of using the chi-squared chart, one can plot the variables against each other and check if the point falls within the elliptical control region, with the same degree of confidence—the perimeter of the ellipse corresponds to the UCL of the chart (Shewhart (1931), Jackson (1956)).

When the in-control covariance matrix $\mathbf{\Sigma}$ is not known and must be estimated from a limited amount of data, it is more appropriate to plot Hotelling's T^2 statistic given by

$$T^2 = (\mathbf{z} - \boldsymbol{\mu})'\mathbf{S}(\mathbf{z} - \boldsymbol{\mu}) \quad (2)$$

where $\mathbf{S}$ is an estimate of $\mathbf{\Sigma}$. An upper control limit T^2_{UCL} is then obtained based on the F distribution and will depend upon the degrees of freedom available for the estimate $\mathbf{S}$ (Alt (1985), Tracy, Young, and Mason (1992), Wierda (1994a)). For a single new $n \times 1$ multivariate observation vector and an estimate $\mathbf{S}$ based on m past multivariate observations

$$T^2_{\mathrm{UCL}} = \frac{(m^2-1)n}{m(m-n)} F_\alpha(n, m-n) \quad (3)$$

where $F_\alpha(n, m-n)$ is the upper $100\alpha\%$ critical point of the F distribution with $(n, m-n)$ degrees of freedom.

Alternatively, other types of multivariate charts, such as Multivariate CUSUM and Multivariate EWMA charts may be used (Lowry, Woodall, Champ, and Rigdon (1992), Sparks (1992), Wierda (1994a)).

The T^2 in (2) can also be expressed in terms of the principal components of the multinormal variables (Mardia, Kent, and Bibby (1979), Jackson (1991))

$$T^2 = \sum_{a=1}^{n} \frac{t_a^2}{\lambda_a} = \sum_{a=1}^{n} \frac{t_a^2}{s_a^2} \quad (4)$$

where λ_a, $a = 1, 2, \ldots, n$, are the eigenvalues of $\mathbf{S}$, and t_a are the scores from the principal component transformation. s_a^2 is the variance of t_a (the variances of the principal components are the eigenvalues of $\mathbf{S}$). Each score t_a can be expressed as

$$t_a = \mathbf{p}_a'(\mathbf{z} - \boldsymbol{\mu}) = \sum_{j=1}^{n} p_{a,j}(z_j - \mu_j) \quad (5)$$

where $\mathbf{p}_a$ is the eigenvector of $\mathbf{S}$ corresponding to λ_a, and $p_{a,j}$, z_j, μ_j are elements of the corresponding vectors, associated with the j^{th} variable. Therefore a T^2 statistic on the initial set of variables is equivalent to a T^2 statistic on the scores when all the principal components are utilized.

A Chemical Process

We consider here the first two zones of a multizone tubular reactor for the production of LDPE. A review on LDPE processes can be found in Kiparissides, Veros, and MacGregor (1993). Details on the simulation can be found in Veros, Papadakis, and

Kiparissides (1993). The schematic of the reactor is shown in Figure 1. Ethylene and a solvent are fed together with the initiator to the first section. Additional injections of ethylene, solvent, and initiators are made between the sections. Each section is jacketed and cooled by a cooling medium flowing countercurrently. The operating conditions in the reactor influence the molecular properties of the polymer produced (weight and number average molecular weights and long and short chain branching), and these in turn affect the behavior of the polymer in its final application. The productivity variable of interest is the conversion per pass. None of the product properties are measured on-line, and in practice many of them are either not measured at all or are measured infrequently. However, many on-line process measurements $\mathbf{x}$, such as the temperature profile down the reactor, the coolant temperatures and the solvent and initiator flowrates, are available on a frequent basis. The five product variables $\mathbf{y}$ of interest are listed in Table 1, together with the process measurements $\mathbf{x}$ assumed to be available for this study. Although the entire temperature profile is usually available for each reactor section, we use a common industrial practice of summarizing the profile in each section by its inlet, maximum, and outlet temperatures together with the position of the maximum. In this paper we assume that the process and quality

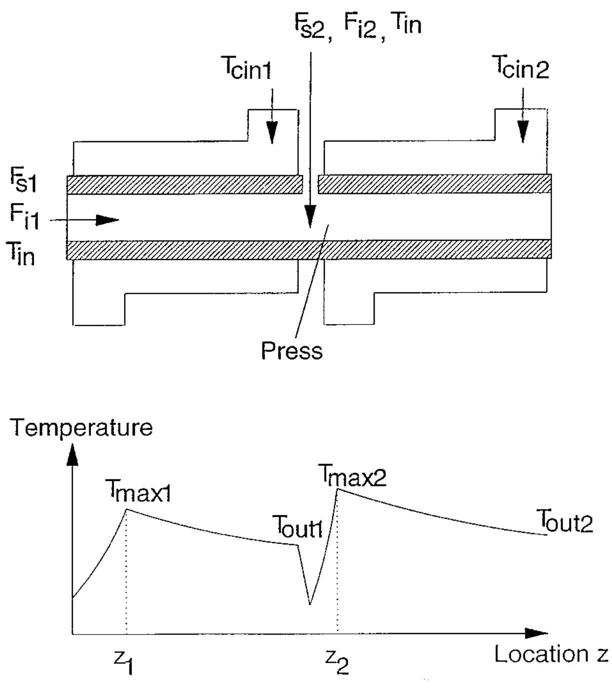


FIGURE 1. Schematic of a Two Zone LDPE Reactor with a Typical Temperature Profile.

TABLE 1. Variables in the LDPE Reactor

Process Variables		
$T_{\max 1}$	$T_{\max 2}$	Maximum Temperature of the Reactor Mixture (K)
$T_{\text{out}1}$	$T_{\text{out}2}$	Outlet Temperature of the Reactor Mixture (K)
$T_{\text{c\_in}1}$	$T_{\text{c\_in}2}$	Inlet Temperature of the Coolant (K)
z_1	z_2	Position of $T_{\max 1}$ and $T_{\max 2}$ in the Reactor (% of Reactor Length)
F_{i1}	F_{i2}	Flowrate of the Initiators (g/s)
F_{s1}	F_{s2}	Flow of the Solvent (% of Ethylene)
T_{in}		Inlet Temperature of the Feed Mixture (K)
Press		Pressure of the Reactor (atm)
Product Quality Variables		
CONV		Cumulative Conversion of Monomer
MW_n		Number Average Molecular Weight
MW_w		Weight Average Molecular Weight
LCB		Long Chain Branching/1000 atom C
SCB		Short Chain Branching/1000 atom C

Indices 1 and 2 in the variables refer to zones 1 and 2.

data are available in steady-state conditions. More details on the choice of the variables for monitoring this process can be found in MacGregor, Jaeckle, Kiparissides, and Koutoudi (1994).

We will examine the behavior of various monitoring tools in several cases, where problems developed in different parts of the reactor:

Case A: The problem appears in zone 1, and zone 2 may or may not be affected.

Case B: The problem appears in zone 2, and zone 1 was not affected.

Case C: The problem appears simultaneously in both zones.

Two very common industrial problems are considered:

1. An increase of the impurity content of the feed. Impurities in ethylene (concentrations at the parts per million level, ppm) inhibit the polymerization process.
2. Fouling of the reactor walls by sticky polymer which impedes heat transfer and cooling of the reactor.

Both impurities and fouling cannot be measured directly. When these problems occur they affect several process variables and eventually the product quality.

Their existence can therefore be detected by the effect they have on the process and quality variables. This effect is not a simple shift of mean of one or more variables; both the magnitudes and the relationships of the variables to each other (i.e., direction of process variability) will change.

In the following examples the type of problem that occurred (i.e., assignable cause) will be determined from the group of process variables that are affected.

Monitoring and Diagnosis in the Product Space

Traditional multivariate charts and new developments on detecting the variables contributing to the out-of-control signal, are first discussed for monitoring product properties only. Figure 2 gives the T^2 chart for the five measured polymer properties (listed in Table 1) that define the quality of the LDPE product. The unnumbered observations in this chart were obtained by simulating normal operating conditions; the numbered observations correspond to simulated problematic operation caused by increasing levels of impurities in the ethylene feed in the reactor. The dashed line (- - -) in the T^2 charts corresponds to the 99% limit and the dotted line ($\cdots$) to the 95% limit. Observations 22–25 show the T^2 response when the impurity levels increase in the feed of zone 1 only (case A), 41–44 when impurities enter zone 2 only (case B), and 60–63 when impurities enter both zones (case C). Notice that the presence of impurities in the ethylene feed had an effect on the product, and that this effect could be detected in all cases, by following the T^2 calculated from the product quality properties. Had we used univariate charts, only points 24, 25, 44, and 61–63 would have been detected as out of control; points 23, 42, 43, and 60 would have been missed. (The Bonferroni type of limits, discussed below, have been used for the univariate charts.)

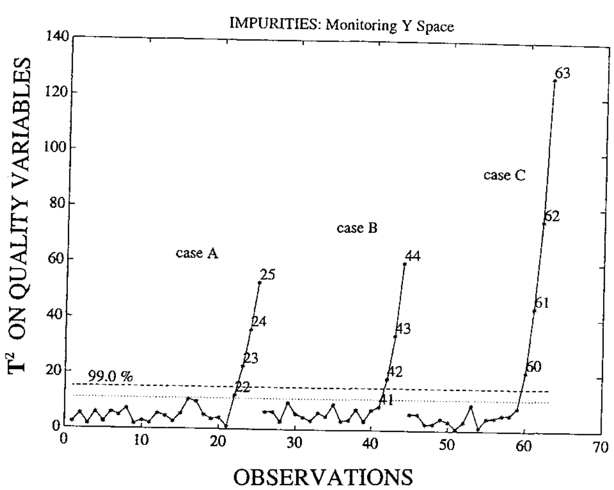


FIGURE 2. T^2 Chart on the Quality Variables. Disturbances are due to Impurities.

Once a problem (or an unusual event) has been detected by the multivariate chart it is important to diagnose an assignable cause for it. This involves two steps: *first* find which variable(s) contributed to the out-of-control signal and *second* determine what happened in the process to upset the behavior of these variables.

Isolation of Variables Responsible for the Out-of-Control Signal

Multivariate monitoring charts can detect an unusual event but do not provide a reason for it. It is difficult to determine which one of the n variables or which subset of them is responsible for an out-of-control signal. Wierda (1994a) lists several attempts reported in the literature to solve this issue. We comment on a few of these approaches, since most others are variations of them.

Murphy (1987) used a procedure similar to discriminant analysis, in which the original set of variables n is divided into subsets, and he tried to detect the subset n_1 of the variables that is responsible for the out-of-control signal. Usually there is no a priori knowledge of how to select subsets, but several suggestions are summarized by Wierda (1994a). This procedure is lengthy, especially if several variables are involved.

Wierda (1994b) discusses using the "step-down" procedure instead of the Hotelling's T^2 chart. The procedure requires ordering the variables in subsets. It checks for a shift in the mean of a subset given that the previous subset was acceptable. The procedure requires prior knowledge on the distribution of the means of the variables. Variables whose mean values rarely shift are put first in the order. A review of the step-down procedure can also be found in Marden and Perlman (1990). Other approaches to detection and diagnosis of multivariate quality control data based on regression adjustments have been proposed (Hawkins (1993), Wade and Woodall (1993), Zhang (1985)). They rely upon plotting various residuals of quality variables after trying to remove the effects of other out-of-control variables by a regression. Recently Mason, Tracy, and Young (1995) suggested a method for decomposing T^2 and showed how the procedure suggested by Murphy and the step-down approach are subsets of their method. They demonstrate its applications for cases where three to five quality properties are involved. All of these procedures may not be practical when a large number of variables are involved, or when there is no clear hierarchical structure (ordering).

In 1980, a diagnostic approach was introduced by Jackson, in which if an unusual event was detected by the T^2 chart, univariate charts for both the normalized PCA scores (t_a/s_a) and the variables were consulted to diagnose the cause of the alarm. If there are only a few variables involved, then by looking at their univariate charts one could detect a gross error. When the scores have some physical significance then a group of variables out of control may affect only a specific principal component and from physical interpretation of the component one could diagnose the problem. The idea of monitoring scores was criticized as being of little value if there is no physical significance to the scores; since each component is a linear combination of the original variables it is not always easy to interpret the results. However, this is no longer the case if one goes one step further and utilizes the contribution plots, discussed below.

Alt (1985) suggested using univariate "Shewhart-like" charts for the individual variables but with Bonferroni-type limits (i.e., replace $\alpha/2$ with $\alpha/(2n)$ in the control limit calculation), to check if it is possible to isolate the variables that have caused the problems. This is a method of opening up the limits to get the desired type I error. They are conservative bounds yielding a type I error of at most α. These bounds would handle the type I error problem for independent variables but not for correlated variables. Alt (1985) also emphasized that all the diagnostic approaches should *follow* the T^2 chart and not replace it. Jackson (1991) modified his 1980 procedure and utilized univariate *score* plots with Bonferroni bounds for diagnosis; since scores are independent these bounds solve the type I error problem.

Many of the above diagnostic procedures have been concerned with providing correct type I error limits. However, once a multivariate chart has detected an event, there is no longer a need to be concerned with precise control limits which control the type I error (α) on the univariate charts. A deviation at the chosen level of significance has already been detected. Diagnostic univariate plots are only used to help decide on the variables which are causing the deviation. Therefore limits at 3σ are just rough guidelines to help determine large errors. Any variables close to these limits should be investigated. In what follows we give several suggestions for detecting the variables that contributed to the out-of-control signal.

Bar Plots of Normalized Errors of the Variables

With this approach n univariate plots are being replaced with one graph. When an observation $\mathbf{x}_k$ falls out of limits in the multivariate chart, the normalized error for each variable j is calculated as

$$(x_{k,j} - \mu_j)/s_j,$$

and these normalized errors are plotted on a common graph. The variables with the highest normalized deviation(s) for a given observation can be detected at a glance. Fuchs and Benjamini (1994) have recently proposed multivariate profile charts where both the T^2 statistic and the bar plots of the normalized variables are displayed in a very efficient way for visual inspection. Figure 3a shows the normalized errors for the five quality variables for point 25 of Figure 2. When the value of the normalized variable exceeds ± 3, the variable plots out of the $\pm 3\sigma$ in a univariate chart. For this example, MW_w plots out of $\pm 3\sigma$. Figure 3b shows the normalized errors plot for point 23 which was detected out of control in the T^2 chart of Figure 2. Notice that no normalized error is above 3. The reason that normalized error plots cannot detect the problem for point 23 is that the error in MW_w is still small, and the variables are highly correlated. Using score plots, described below, will overcome this and generally detect the problem much more readily.

Bar Plots of Normalized Scores

This approach is preferable for correlated variables. The normalized scores (t_a/s_a) of $\mathbf{x}_k$ are plotted on a common graph, and normalized scores with high values are detected. Since the scores are independent, Bonferroni-type limits could also be used as guides for "high values". Once the score(s) that contribute to the out of limit T^2 signal are detected, contribution plots (discussed below) can be used to find the quality variables, numerically responsible for the signal.

For point 23 of Figure 2 the normalized scores are shown in Figure 4a. Notice that scores 2 and 5 have relatively high values, in spite of the fact that the individual normalized quality variables (Figure 3b) are not exceptionally large. Again, the limits for the normalized scores are roughly used as guides here,

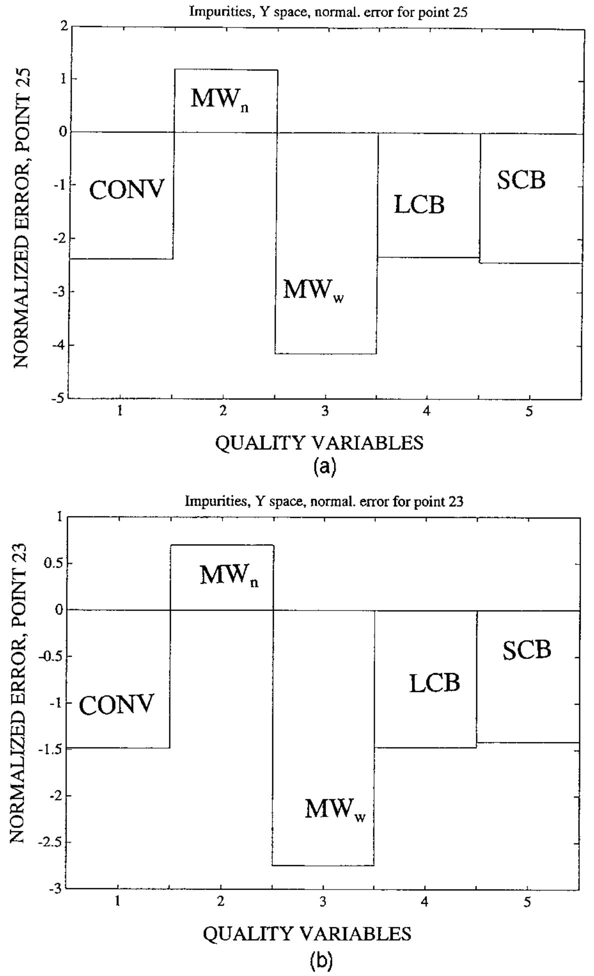


FIGURE 3. Normalized Errors for the Quality Variables, for Observations of Figure 2, (a) for Point 25 and (b) for Point 23.

but not for precise control of type I error; 2.7σ would have been the value equivalent to Bonferroni limits for 95% confidence on type I error for five variables (Alt (1985)).

Jackson (1991) reports a somewhat similar procedure suggested to him by R. Thomas (Burroughs Corp.) during a short course: "represent the value of Hotelling's T^2 of an observation in a histogram form by plotting the individual contributions of the scores squared $(t_a/s_a)^2$—stack bar graph—to indicate the nature of the out-of-control situations". Here, by instead using the normalized scores t_a/s_a, Bonferroni limits can be used on the chart as rough guidelines for detecting large t_a/s_a values.

Variable Contributions to Individual Scores

One set of diagnostic tools, recently suggested by

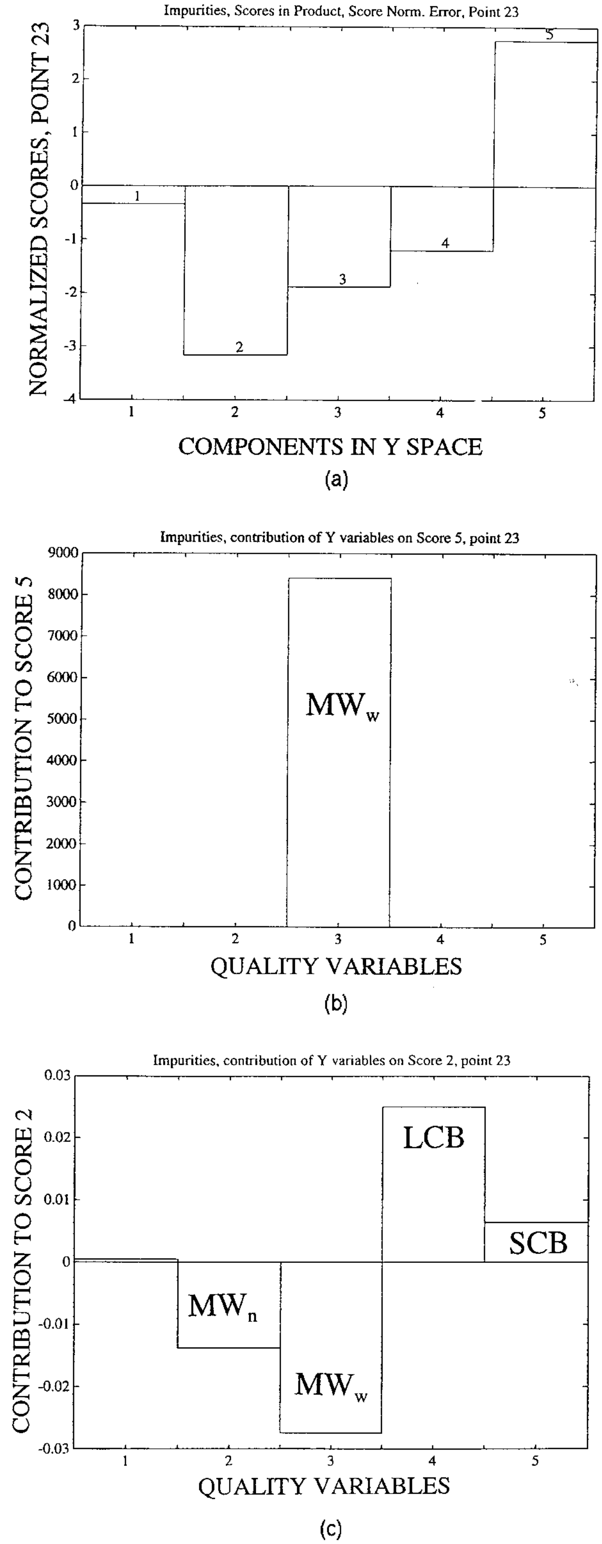


FIGURE 4. Fault Diagnosis using Scores in Quality Space. (a) Normalized Scores for Point 23 of Figure 2, (b) Variable Contributions to Score 5, and (c) Variable Contributions to Score 2.

Miller, Swanson, and Heckler (1993) and MacGregor et al. (1994) are contribution plots on scores. A contribution plot indicates how each variable involved in the calculation of that score contributes to it. Contribution plots can be constructed for normalized scores with high values; the variables on this plot that appear to have the largest contributions to it, should be investigated.

From (5), the contribution of each variable z_j to the score of the principal component a is

$$p_{a,j}(z_j - \mu_j). \tag{6}$$

The variable contribution plots for scores 5 and 2 of Figure 4a are shown in Figures 4b and 4c, respectively. Variable MW_w is mainly responsible for score 5; variables MW_w and LCB have high values for score 2. However score 2 has a negative value (see Figure 4a); it is this value (squared) that contributes to the calculation of the T^2. Therefore variables with negative contributions are more significant in driving this value high; positive contributions will only make it smaller. Here we suggest that variables with high contributions but *with the same sign as the score* should be investigated. For the case of score 2, high negative contributions are important and that leads to variable MW_w. (From simulations of this particular process, indeed, we confirm that once impurities keep building up it is variable MW_w that will be mostly affected.)

The variable contribution plots together with the score plots is a powerful tool of detecting variables contributing to the out-of-control signal. Notice that although T^2 signaled for point 23, the univariate charts on the quality variables could not detect the problem with MW_w. However the combination of the score plots and the variable contribution plots did.

Overall Average Variable Contributions

Very often more than one score may have high values. Therefore it is suggested here that an "overall average contribution" per variable is calculated, over all the scores with high values.

For each one of the normalized scores with high values (say above 2.5) calculate the variable contribution, but keep only the contributions with the same sign as the score. Sum the values over all the K high scores, $K \leq n$ as shown below:

1. Repeat for all the K high scores $K \leq n$.
   i. Calculate contribution of a variable z_j in the normalized score $(t_a/s_a)^2$

$$\text{cont}_{a,j} = \frac{t_a}{s_a^2} p_{a,j}(z_j - \mu_j) = \frac{t_a}{\lambda_a} p_{a,j}(z_j - \mu_j) \tag{7a}$$

   ii. $\text{cont}_{a,j}$ is set equal to zero if it is negative (i.e., its sign is opposite to the value of the score t_a).
2. Calculate the total contribution of variable z_j

$$\text{CONT}_j = \sum_{a=1}^{K} (\text{cont}_{a,j}). \tag{7b}$$

Using this procedure on the scores 2 and 5 of Figure 4a, one would have obtained MW_w as the responsible variable for the out of limits signal of point 23 in Figure 3. Another example where this procedure is applied is shown later in Figures 7a and 7b.

Notice from (4) and (5) that T^2 can be written as a sum of the contributions of all variables over all the scores

$$\begin{aligned} T^2 &= \sum_{a=1}^{n} \frac{t_a}{s_a^2} \sum_{j=1}^{n} p_{a,j}(z_j - \mu_j) \\ &= \sum_{j=1}^{n} \left\{ \sum_{a=1}^{n} \frac{t_a}{\lambda_a} \left[p_{a,j}(z_j - \mu_j) \right] \right\}. \end{aligned} \tag{8}$$

It has been suggested (Nomikos (1995)) that a graph with the contributions of each variable to T^2 could be constructed; these contributions could be positive or negative. However, from our experience, determining the variables responsible for the out-of-control signal from such a graph is ambiguous. The discriminating ability of the approach described by (7a) and (7b) is more powerful; it unravels the variable(s) responsible for an increase in *both* the value of T^2 *and* the values of the high normalized scores. (The approach in (7) is *not* equivalent to plotting *all* the contributions from (8) and keeping the ones with positive values; rather to plotting contributions in (8) with the summation *over the high scores only* and keeping the positive values.)

General Diagnostic Procedure

The approach for diagnosing the variable(s) responsible for an out-of-control signal in a multivariate T^2 chart, initially suggested by Jackson (1980) could be extended and utilized as follows:

1. Out-of-limit signal detected: T^2 value for observation $\mathbf{x}_k$, above T^2_{UCL} (see (3)).
2. Check normalized scores of observation $\mathbf{x}_k$: find scores with high values. (Bonferroni limits could be used on the score charts as rough guides; if α is

the significance level for T^2, use α/n for the score charts).

3. Calculate the variable contributions for these high scores (see (6) and (7)); investigate variables with high contributions.

With this approach one does not necessarily have to plot the normalized scores and the variable contributions and consult graphs. The whole procedure can be easily programmed in a computer. In step 2, the individual scores can be checked against some prespecified limits, and in step 3 the variable contributions can then be calculated for the high scores.

With this procedure there is no need to interpret the scores for physical significance; their role is to facilitate the calculation of variables contributing to the large value of T^2. In general this approach will point to a variable or a group of variables that contribute *numerically* to the out-of-control signal; these variables and any variables highly correlated with them should be investigated in order to assign causes.

Monitoring and Diagnosing in the Process Space

Traditionally the product quality properties are charted to determine if the process is in a state of "statistical control". However, product quality data may not be available frequently, but only every few hours. If a problem develops in the process, out-of-control product is being produced and will go unnoticed at least until the results from the laboratory become available. In the LDPE example none of the product properties are available on-line. However, many on-line process measurements such as the temperature profile down the reactor, the coolant temperatures, and the solvent and initiator flowrates are available on a frequent basis. The signature of any special events or faults occurring in the process that will eventually affect the product should also appear in these process data ($\mathbf{x}$). Therefore monitoring the process may be preferable.

There are several other reasons why monitoring the process is advantageous. Sometimes, only a few properties of the product are measured, but these are not sufficient to define entirely the product quality. For example, only the relative viscosity (RV) may be charted in nylon production, although there are other properties (amine end groups) that affect the dye properties of the product. If process problems that affect the amine groups occur, they will not be detected by following RV only. In these cases the process data may contain information about these events.

Finally, even if product quality measurements are frequently available, monitoring the process may facilitate assigning causes for an event. When monitoring product quality only, even if we determine which quality variable was "unusual" and contributed to the out-of-control signal, it may still be difficult to determine what went wrong in the process. For example, in the LDPE process by following the product variables it was determined that for point 23, MW_w is the major contributor to the out-of-control signal. However, there may be several reasons (combinations of process conditions) that might have caused molecular weight to change. Therefore although we have answered the question "Which variable is responsible for the out-of-control signal?", it is not always easy to answer the question "What process conditions have caused this?" Monitoring the process would bring us closer to the answer.

In what follows we show how to utilize traditional approaches and extend their application to monitoring process variables. The values of the process variables collected from historical data and corresponding to conditions that produced acceptable product are used as a training set to calculate the mean and covariance matrix. Then each new set of process variables is checked against the "in-control" conditions. By looking at the process as well as the quality variables we are truly doing statistical *process* control (SPC), as opposed to statistical *quality* control (SQC).

Figure 5 shows the response of the T^2 calculated for the fourteen process variables in the LDPE reactor (see Table 1) for the same simulated operating conditions as those of Figure 2. Observations 22–25 show the T^2 values when increasing levels of impurities enter zone 1 (case A); 41–44 when impurities enter zone 2 only (case B), and observations 60–63 when impurities enter simultaneously in both zones (case C). Notice that, for all the cases, the upset in the process that eventually led to the upset in the product quality was detected by monitoring the T^2 statistic of the process variables only. In fact the upset is detected *earlier* (i.e., at a lower impurity concentration) by monitoring the process; notice that points 22 and 41 with small amounts of impurities are detected as problematic in the process, although in the product they are around or even below the 95% limit. Furthermore the T^2 statistic on the process variables detects the problem on-line and one

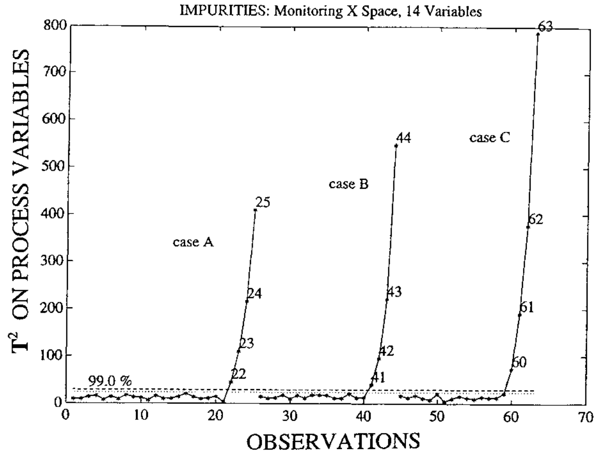


FIGURE 5. T^2 Chart on the Process Variables. Disturbances are due to Impurities.

does not have to wait for the infrequent laboratory tests on the quality properties. Once a problem (or an unusual event) has been detected by the multivariate chart on the process variables one can diagnose an assignable cause for it. The procedures described earlier for identifying the variables contributing to an out-of-control signal can be applied for multivariate charts on process variables. Once we determine the *process* variables responsible for an out-of-control signal it is much easier to diagnose sources for the process deviation, as illustrated below.

The normalized errors for all fourteen process variables for point 23 of Figure 5 are plotted in Figure 6. Two variables seem to have large errors: the location of the hot spot in zone 1, z_1, has moved further up the reactor and the maximum temperature, $T_{\max 1}$ has decreased. Because there are strong correlations among the process variables, plots of normalized errors may not be sufficient to detect all the variables with "unusual behavior". Figure 7a shows the normalized scores for point 23. The variable contributions for scores 1, 2, 3, and 5 which lie beyond the 3σ limits are plotted in Figure 7b. The maximum temperature in the reactor in the first zone $T_{\max 1}$, and the location of the hot spot in that zone, z_1 appear to be variables with important contributions, together with the temperature of the coolant in the reactor, $T_{\text{c\_in}1}$, and the outlet temperature of the reactor mixture, $T_{\text{out}1}$. The combination of the normalized scores and the variable contribution plots uncovered two more variables with "unusual" behavior than those detected by the normalized errors of the variables. From knowledge of the process and polymerization histories the explanation when this combination of variables appears to have problems, is that something has affected the extent of the reaction. The problem (assignable cause) could be either in a decrease in the efficiency of the initiator or in an increase in the level of impurities in the reactor. The contribution plots may not explicitly reveal the cause of the event, but *they point to the group of process variables that are no longer consistent with normal operating conditions*, and thereby focus the attention of the operators or the engineers and allow them to use process knowledge to deduce possible causes.

Blocking the Process Variables

When dealing with several sections of a unit or even several units, one may have to deal with many process variables. In these cases it may make sense to break the process variables into logical sections within which the variables are highly correlated but between which there is much less correlation, and examine the T^2 response for the variables of each section separately. These sections could correspond to physical units, or even sections of a unit.

In the LDPE example, since the process consists of two zones, the process variables can be separated into two groups and one may monitor the T^2 response for each zone separately. The inlet temperature of the feed mixtures in each zone and the pressure are common in both of the sections, hence these two variables were included in both groups, such that each zone is described by eight process variables (see Table 1). The objective of this variable separation is to isolate and more easily diagnose a problem when it occurs. Each chart gives the picture of the operation

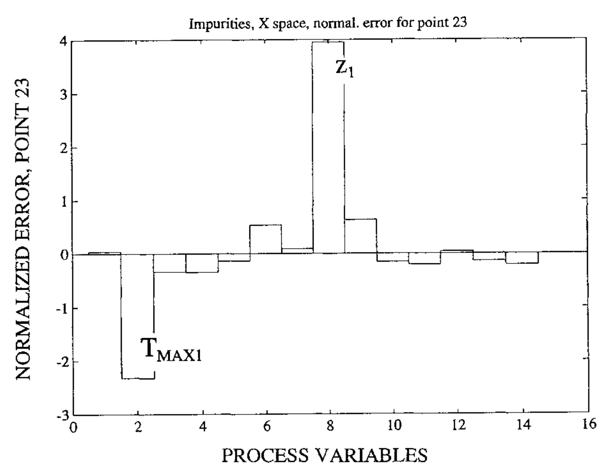
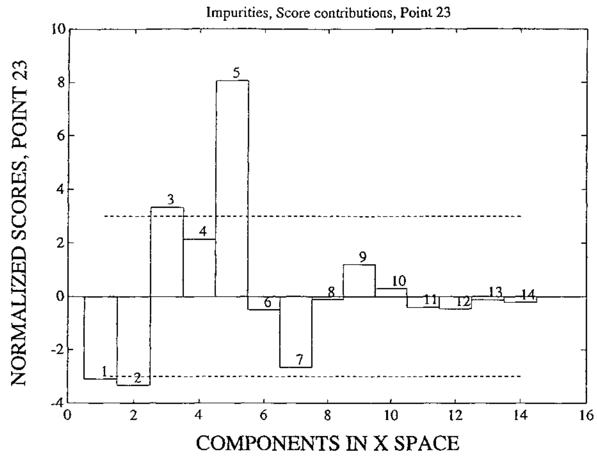
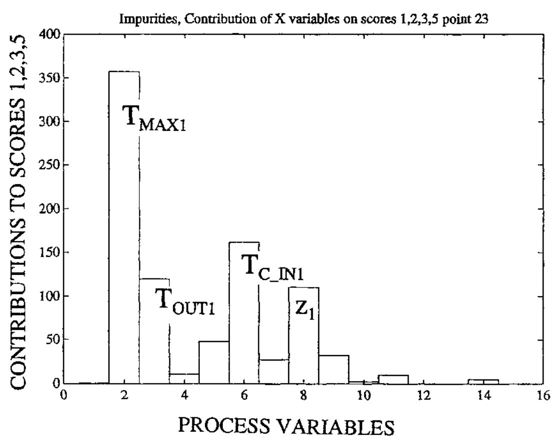


FIGURE 6. Normalized Errors for the Process Variables for Point 23 of Figure 5.



(a)



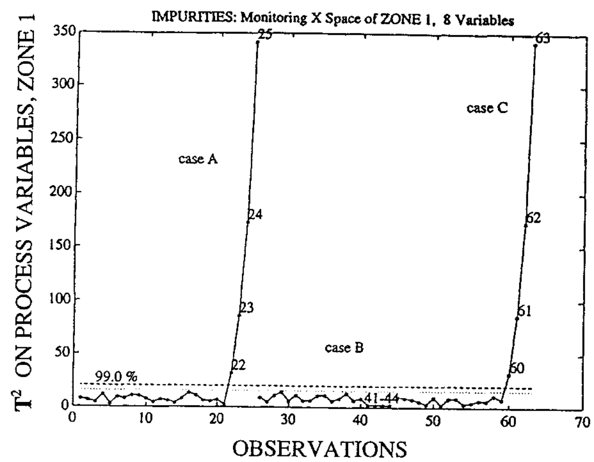
(b)

FIGURE 7. Fault Diagnosis using Scores in Process Space. (a) Normalized Scores for Point 23 of Figure 5, and (b) Variable Contributions to all High Scores.

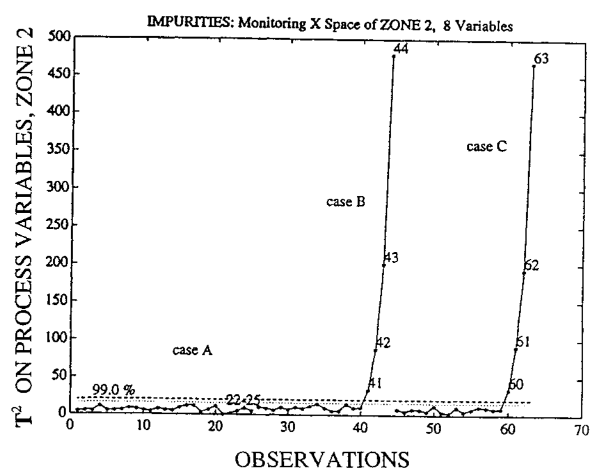
history for each individual zone. Figure 8a shows the T^2 chart for zone 1 only, for the same simulated conditions of Figures 2 and 5. Notice that the chart indicates that zone 1 appears to have problems for points 22–25 and 60–63. That is, for cases A and C, indeed a disturbance (impurities in the ethylene feed) went through zone 1 in both cases. For points 41–44 (case B) no abnormality is detected for zone 1. During this time impurities entered zone 2 but there was no effect on the process variables of zone 1 (their values should remain within normal operating range), and therefore the T^2 calculated from the process variables of zone 1 should remain within limits. Figure 8b shows the T^2 chart for zone 2. Notice that the chart indicates that there is nothing unusual for points 22–25 in zone 2. Impurities had entered in zone 1, for these observations. However, the effect of the impurities is not expected to propagate from zone to zone since they are consumed in the section which they enter (zone 1) and have no effect on zone 2 as it is correctly shown on the graph.

Once an out-of-control signal is detected in the chart of the blocked variables one can proceed with diagnosis as discussed earlier. For zone 1 and observation 23, the variables contributions (not shown here) for the high normalized scores (scores calculated for variables of zone 1 only) show that variables $T_{\max1}$, z_1, $T_{\text{c\_in1}}$, and T_{out1} are mainly responsible for the out-of-control signal.

By separating the variables into groups we were able to clearly isolate the section with the problem.



(a)



(b)

FIGURE 8. T^2 Charts on Blocks of Process Variables in the LDPE Reactor. Disturbances are due to Impurities, (a) Zone 1 and (b) Zone 2.

Of course there are processes in which a problem in one section may cause operational problems in subsequent sections. In such cases the T^2 chart approach should be utilized to detect the problem in the earliest (first) section in which it appears. This is illustrated with the following example, where the disturbance is an increasing level of fouling of the reactor walls.

Figures 9a–9d give the T^2 charts on the product variables $\mathbf{y}$ and on the process variables $\mathbf{x}$ for both zones, as well as the T^2 charts for zone 1 and zone 2, for a different type of disturbance. In this example, observations 22–26 simulate process conditions when fouling of the wall occurs in zone 1 (case A); for observations 42–46 fouling occurs in zone 2 (case B); and for observations 62–66 fouling occurs simultaneously in both zones (case C). The unnumbered observations correspond to normal operating conditions. Notice from Figure 9a, that for this particular case, fouling in zone 1 had a significant effect on the product properties; fouling in zone 2 had a moderate effect (brings T^2 above the 95% limits but below the 99%). The T^2 chart on all the process variables (Figure 9b) detected the problems for all cases (including case B, where it shows that observations 42–46 are out of the normal operating region, although they had a moderate effect on the product). Monitoring zone 1 only (Figure 9c) indicates correctly that there were problems in observations 22–26 and 62–66. Figure 9d indicates that zone 2 experienced disturbances in all three cases. Even though fouling of the walls was not observed in zone 2 for case A, the correlations of the variables in zone 2 are slightly affected by the disturbance that occurred in zone 1 (i.e., fouling effects are carried through). However, by monitoring

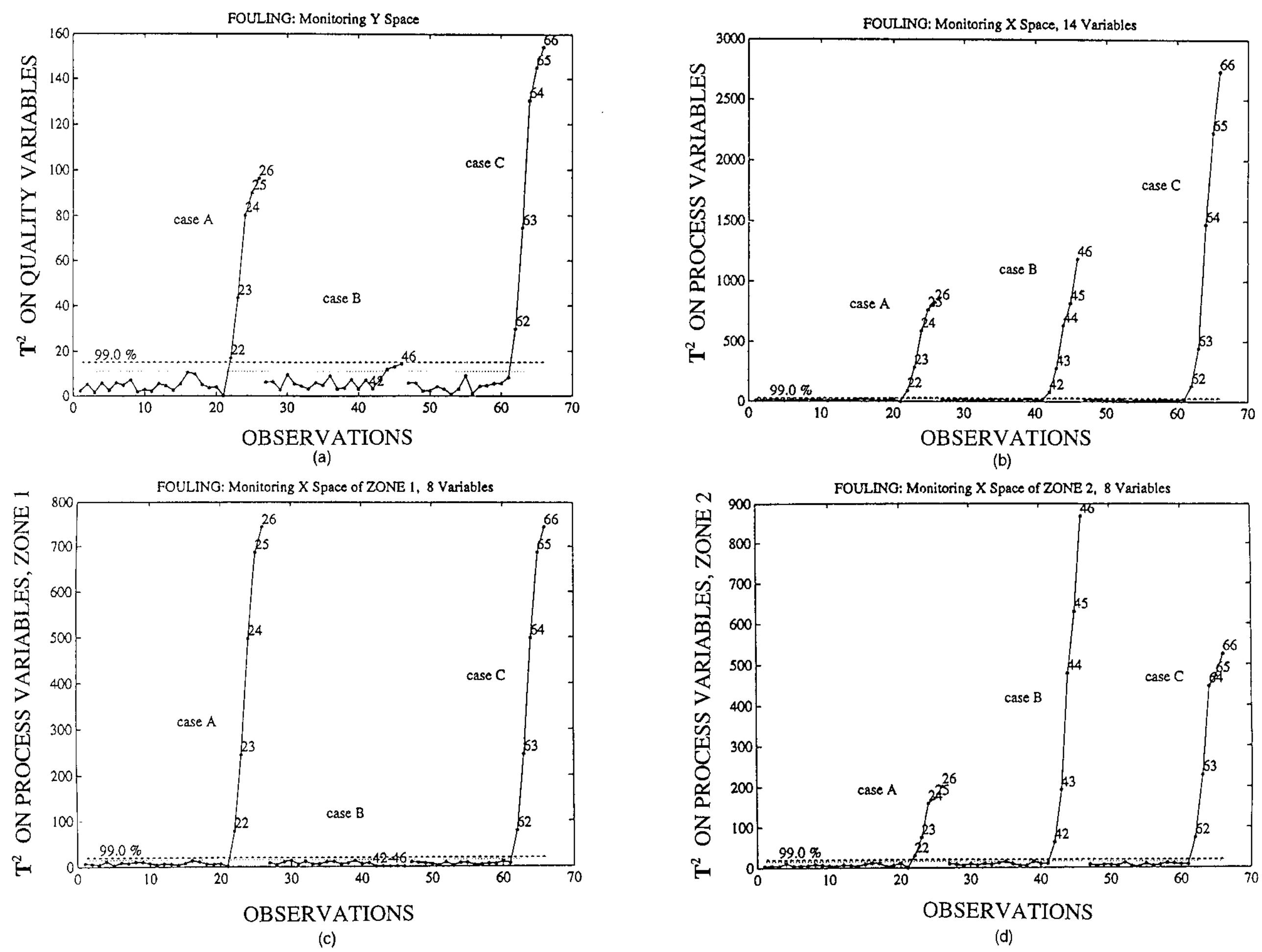


FIGURE 9. Process Monitoring in the LDPE Reactor, using T^2 Charts. The Disturbance is Fouling. (a) T^2 Chart on Quality Variables, (b) T^2 Chart on Process Variables, (c) T^2 Chart on Blocked Process Variables, Zone 1, and (d) T^2 Chart on Blocked Process Variables, Zone 2.

the sections separately we were able to detect the earliest section of the unit (zone 1) that had problems.

Blocking the variables and monitoring a separate T^2 per block (unit) is a good approach when there are not significant correlations between the variables of the different blocks. If there are correlations between two or three blocks (i.e., recycles), an independent T^2 per block will still give a good indication of the performance of the block; but an overall T^2 (global T^2) that includes the variables from all the blocks is needed to monitor the performance of the plant (i.e., to detect inconsistent behavior among blocks).

Projection Methods

A problem with utilizing traditional χ^2 and T^2 approaches is that they may not be practical for very high-dimensional systems with collinearities. Many highly collinear variables lead to ill-conditioned data which implies that χ^2 or T^2 becomes very sensitive to deviations of the components corresponding to small eigenvalues. The inversion of the covariance matrix (needed for the calculation of χ^2 or T^2) may be a problem. Breaking up the process into sections may resolve the problem of dealing with large systems; but this is not always feasible as discussed earlier in systems with correlated units. A common procedure for reducing the dimensionality of the variable space is the use of projection methods like principal component analysis (PCA) or partial least squares (PLS). Another problem with using the T^2 approach (either in the x- or y-space) is that it cannot easily handle missing data. If a sensor fails, or an analysis on one or more of the variables is messed up, then the whole set of good measurements on all the other variables cannot be used. However the available good measurements might still be a sufficient indication of the performance for that observation. Multivariate projection methods handle missing data more easily and utilize the remaining information to test for an out-of-control situation.

The basic concepts and algorithms of PCA and PLS (Geladi and Kowalski (1986)) and their use in multivariate monitoring of process operating performance (Wise and Ricker (1989), Wise (1991), Kresta, MacGregor, and Marlin (1991), Skagerberg, MacGregor, and Kiparissides (1992), MacGregor et al. (1994), Nomikos and MacGregor (1994, 1995), MacGregor and Kourti (1995)) have been extensively presented in the literature. A brief description of the methods is given below.

In the PCA approach a set of data $\mathbf{Z}$ is collected when the process was known to operate "in control". This set can be either process ($\mathbf{Z} = \mathbf{X}$) or product quality ($\mathbf{Z} = \mathbf{Y}$) variables. The method is equivalent to performing principal component transformation to the original set of data $\mathbf{Z}$ ($m \times n$) and keeping only the A (out of n) principal components with the highest variation. This way the original data set is projected to an orthogonal space, which, in general, is of a lower dimension (i.e., $A << n$)

$$\mathbf{Z} = \sum_{a=1}^{A} \mathbf{t}_a \mathbf{p}_a' + \sum_{a=A+1}^{n} \mathbf{t}_a \mathbf{p}_a' = \widehat{\mathbf{Z}} + \mathbf{E}. \tag{9}$$

The location of this A-dimensional space with respect to the original coordinates is given by the loadings $\mathbf{p}_a$, which are the eigenvectors, corresponding to the A largest eigenvalues of the covariance matrix of the original data set $\mathbf{Z}$. $\widehat{\mathbf{Z}}$ indicates the estimate of $\mathbf{Z}$ based on the first A principal components. The value of A can be determined by cross-validation (Wold (1978)). The location of the projection of an observation onto the A-dimensional space is given by its score t_a, which are linear combinations of the original variables. The squared perpendicular distance of an observation $\mathbf{z}_i$ from the projection space, called squared prediction error (SPE), gives a measure of how close the observation is to this A-dimensional space:

$$\mathrm{SPE}_{\mathbf{z}_i} = \sum_{j=1}^{n} e_{ij}^2 = \sum_{j=1}^{n} (z_{ij} - \widehat{z}_{ij})^2 \tag{10}$$

where e_{ij} is an element of the i^{th} row of matrix $\mathbf{E}$ and $\widehat{z}_{ij}$ is calculated from the PCA model with A components. SPE measures the contribution of the components that have not been utilized.

From data collected when the process is "in control", the loadings corresponding to the normal operating region are determined. For each new observation, $\mathbf{z}_{\text{new}}$, the scores and the SPE are calculated and checked against "in-control" operation. For monitoring the scores, T^2 charts based on the first A principal components are constructed:

$$T_A^2 = \sum_{a=1}^{A} \frac{t_{\text{new},a}^2}{s_a^2} = \sum_{a=1}^{A} \left[\frac{\mathbf{p}_a'(\mathbf{z}_{\text{new}} - \boldsymbol{\mu})}{s_a} \right]^2. \tag{11}$$

An upper control limit can be calculated from (3) with $(A, m - A)$ degrees of freedom.

For the SPE a chart is also constructed,

$$\mathrm{SPE}_{\mathbf{z}_{\text{new}}} = \sum_{j=1}^{n} (z_{\text{new},j} - \widehat{z}_{\text{new},j})^2 \tag{12}$$

where

$$\widehat{\mathbf{z}}_{\text{new}} = \sum_{a=1}^{a=A} t_{\text{new},a}\ \mathbf{p}_a$$

is computed from the reference PCA model. Upper control limits for the SPE statistic can be computed using approximate results for the distribution of quadratic forms (Jackson (1991), Nomikos and MacGregor (1995)). This is also often referred to as a Q statistic (Jackson (1991)).

The above procedure can be used for monitoring either process variables ($\mathbf{Z} = \mathbf{X}$) or product quality variables ($\mathbf{Z} = \mathbf{Y}$). Both SPE and T_A^2 are necessary for monitoring. The statistic on the scores checks if a new observation remains within the normal operating region in the projection space, and SPE checks if the distance of the new observation from the projection space is within acceptable limits. If a totally new type of special event occurs which was not present in the reference data used to develop the in-control PCA model, then new components will appear and the new observations will move away from the plane. Such new events can be detected by plotting the SPE.

One of the advantages of using T_A^2 rather the traditional T^2 in cases when the variables are highly correlated and $\mathbf{\Sigma}$ is ill-conditioned, is illustrated here. Note that the traditional Hotelling T^2 in (4) can be written as:

$$T^2 = \sum_{a=1}^{A} \frac{t_a^2}{s_a^2} + \sum_{a=A+1}^{n} \frac{t_a^2}{s_a^2} = T_A^2 + \sum_{a=A+1}^{n} \frac{t_a^2}{s_a^2}. \quad (13)$$

By scaling each t_a by the reciprocal of its variance each term plays an equal role in the computation of T^2 irrespective of the amount of variance it explains in the $\mathbf{Z}$ matrix. When the number of variables n is large, $\mathbf{\Sigma}$ is often singular and cannot be inverted, nor can all the eigenvectors be obtained. Even if $\mathbf{\Sigma}$ can be inverted, the last components $a = A+1, \ldots, n$ in (13) explain very little of the variance of $\mathbf{Z}$ and generally represent random noise. By dividing these t_a's by their very small variances, slight deviations in these t_a's (sometimes even due to round-off error) which have almost no effect on $\mathbf{Z}$ will be magnified and lead to an out-of-control signal in T^2. Therefore T_A^2 based on the first A cross-validated components provides a test for deviations in the variables that are of greatest importance to the variance of $\mathbf{Z}$. The NIPALS algorithm can calculate sequentially the first A principal components without calculating all the eigenvalues (Geladi and Kowalski (1986)).

The use of the first A principal components with T^2 is discussed by Fuchs and Benjamini (1994). However they use a full-rank T^2 chart to detect out-of-control observations and then examine normalized scores of the first A principal components to diagnose the problem; they rely on the physical interpretation of the principal components for an out-of-control signal and not on the contribution plots discussed below.

The PCA approach for monitoring *process variables* is used when product quality data are not available in the historical data set. In the PLS approach for monitoring process variables the model is developed from historical data sets with measurements from both the process, $\mathbf{X}$, and the quality, $\mathbf{Y}$, variables obtained during in-control operation. PLS simultaneously reduces the dimensions of $\mathbf{X}$ and $\mathbf{Y}$, to find latent vectors for $\mathbf{X}$ that not only explain the variation in the process data, but that variation in $\mathbf{X}$ which is most predictive of the product quality data, $\mathbf{Y}$. This is accomplished by working on the sample covariance $\mathbf{X}'\mathbf{Y}\mathbf{Y}'\mathbf{X}$ matrix. In the most common version of PLS (Höskuldsson (1988)), the first PLS latent variable $t_1 = \mathbf{w}_1'\mathbf{x}$ is that linear combination of the x variables that maximizes the covariance between it and $\mathbf{Y}$. The first PLS loading vector $\mathbf{w}_1$ is the first eigenvector of the sample covariance matrix $\mathbf{X}'\mathbf{Y}\mathbf{Y}'\mathbf{X}$. Once the scores $\mathbf{t}_1 = \mathbf{X}\mathbf{w}_1$ for the first component have been computed (i.e., the values of that latent variable for the n observations) the columns of $\mathbf{X}$ are regressed on $\mathbf{t}_1$ to give a regression or loading vector

$$\mathbf{p}_1 = \frac{\mathbf{X}'\mathbf{t}_1}{\mathbf{t}_1'\mathbf{t}_1}$$

and the $\mathbf{X}$ matrix is deflated to give residuals

$$\mathbf{X}_2 = \mathbf{X} - \mathbf{t}_1\mathbf{p}_1'.$$

$\mathbf{Y}$ is also deflated as

$$\mathbf{Y}_2 = \mathbf{Y} - \mathbf{t}_1\mathbf{q}_1'$$

where $\mathbf{q}_1$ is obtained by regressing the columns of $\mathbf{Y}$ on $\mathbf{t}_1$,

$$\mathbf{q}_1' = \frac{\mathbf{t}_1'\mathbf{Y}}{\mathbf{t}_1'\mathbf{t}_1}.$$

The second latent variable is then computed as $t_2 = \mathbf{w}_2'\mathbf{x}$ where $\mathbf{w}_2$ is the first eigenvector of $\mathbf{X}_2'\mathbf{Y}_2\mathbf{Y}_2'\mathbf{X}_2$ and so on. As in PCA, both the score vectors $(\mathbf{t}_1, \mathbf{t}_2, \ldots)$ and the loading vectors $(\mathbf{w}_1, \mathbf{w}_2, \ldots)$ are orthogonal. For large ill-conditioned data sets, it is usually convenient to calculate the PLS latent variables sequentially via the NIPALS algorithm (Geladi and Kowalski (1986)) and to stop based on cross-validation criteria. An advantage of the NIPALS algorithm is that it handles missing data (i.e., observation vectors from which some variable measurements are missing) in a very simple and straightforward way (Kresta, MacGregor, and Marlin (1991), Nelson, MacGregor, and Taylor (1996)).

In the modelling stage, $\mathbf{p}_a$'s, $\mathbf{w}_a$'s, and $\mathbf{q}_a$'s are calculated sequentially as discussed. In the monitoring stage, for each new observation, $\mathbf{x}_{\text{new}}$, the t_a's are calculated (using $\mathbf{p}_a$'s, $\mathbf{w}_a$'s, and $\mathbf{q}_a$'s) sequentially by deflating $\mathbf{x}_{\text{new}}$ each time the score for a component has been calculated. The multivariate control chart is again a T_A^2 chart on the first A latent variables, t_a, $a = 1, \ldots, A$; the difference with PCA is that these latent variables have been computed utilizing information from both the $\mathbf{X}$ and $\mathbf{Y}$. Therefore when using PLS, T_A^2 monitors the variation in the process variables which is more relevant to the product quality variables (Kresta, MacGregor, and Marlin (1991)). A chart on the squared prediction error in the x-space, $\text{SPE}_{\mathbf{x}_{\text{new}}}$ (12), where $\widehat{\mathbf{x}}_{\text{new}}$ is now computed from the reference PLS model, is used to detect if the covariance structure of the process has changed. A change in the pattern needs investigation because it means that the process conditions are now different than the time that the model was developed. This change may or may not have effects on the product. Control limits for the T_A^2 and $\text{SPE}_{\mathbf{x}}$ charts are chosen in the same manner as previously discussed. In both PCA and PLS the scores t_a are linear combinations of the process variables. If the process variables are normally distributed then the scores are also normally distributed. When the process variables are not normally distributed, the scores (being linear combinations of these variables) are expected to be closer to a multinormal distribution by the central limit theorem.

Monitoring and Diagnosis utilizing Scores and SPE

An essential part of SPC is to establish control charts to detect special events as they occur, and to diagnose possible causes for them while the information is fresh. The philosophy applied in developing multivariate SPC procedures based on projection methods, is the same as that used for the traditional univariate or multivariate charts. An appropriate reference set is chosen which defines the normal or "in-control" operating conditions for a particular process. In other words, a PCA or PLS model is built based on data collected from various periods of plant operation when performance was good. Any periods containing variations arising from special events that one would like to detect in the future are omitted at this stage. The choice of this reference set is critical to the successful application of the procedure, as discussed by Kresta, MacGregor, and Marlin (1991).

A PLS model was built to relate the fourteen process variables, $\mathbf{X}$, with the five quality variables, $\mathbf{Y}$, under normal operating conditions of the LDPE reactor. Three latent variables could explain 90% in the $\mathbf{Y}$. The 99% limits for the scores and the SPE were defined; then the monitoring scheme was tested for simulated cases of fouling. Figure 10 shows the behavior of the process on a projection of the process variables on the plane defined by the two latent variables t_2 and t_3. The unnumbered asterisks denote normal operation. The numbered ones show the process behavior under fouling. The ellipses show confidence limits (99% solid line, 95% dotted line). Fouling occurred for observations 22–26, 42–46, and 62–66, as described earlier for Figures 9a–9d. During normal operation the projected points move randomly within the limits. Notice, however, that during an "unusual" event the direction of the process (i.e., the relationship among the variables) is changing; different events (i.e., increasing fouling at different sections of a unit) have different direction. The T_3^2 chart for the three latent variables is shown in Figure 11a and the SPE chart in Figure 11b. Notice that by monitoring the T_3^2 chart on the scores and the SPE simultaneously, one could detect the problematic observations in all cases, utilizing information collected from only the x-space (i.e., process variables).

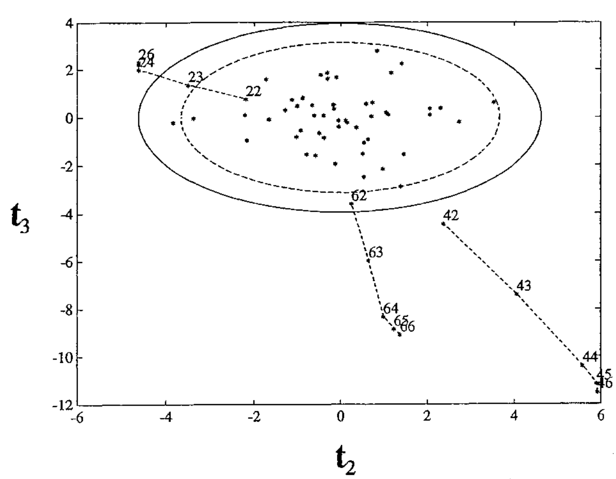
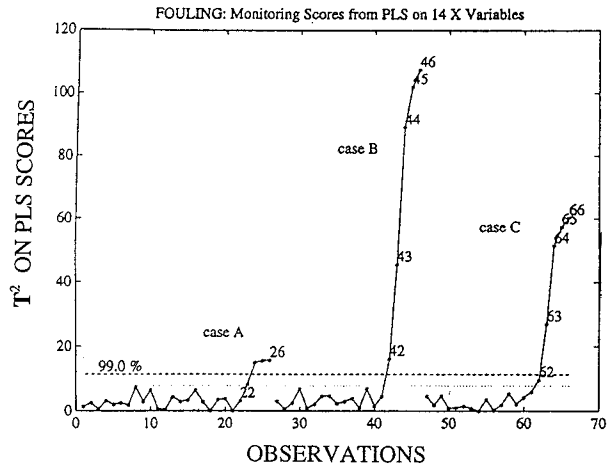
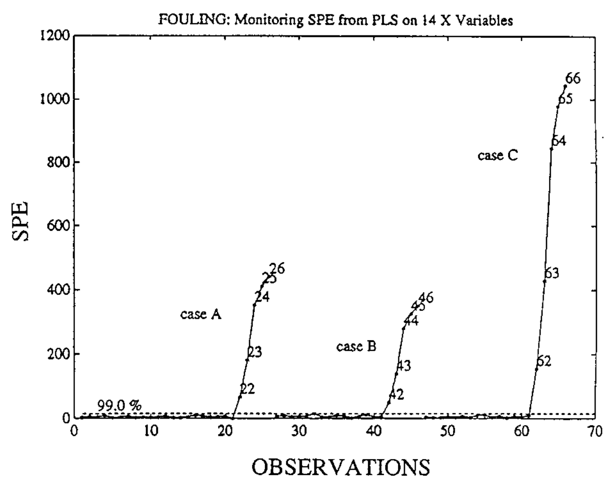


FIGURE 10. Process Monitoring in the LDPE Reactor, using PLS. The Disturbance is Fouling. Projection on the $t_2 - t_3$ Plane.



(a)



(b)

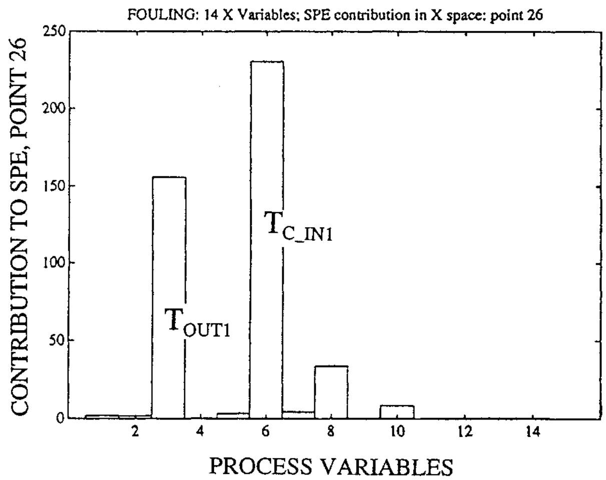
FIGURE 11. Process Monitoring in the LDPE Reactor, using PLS. The Disturbance is Fouling. (a) T_3^2 Chart on 3 PLS Scores, and (b) SPE Chart.

Once a deviation from the normal operating region has been detected on-line, an assignable cause for this deviation must be found and corrected. One set of diagnostic tools are the contribution plots (on the SPE or on the scores) discussed by Miller, Swanson, and Heckler (1993) and MacGregor et al. (1994). Contributions to PCA scores have been discussed earlier (equations 4 or 5); the contribution of a variable x_j to the PLS scores is given as $w_{a,j}(x_{\text{new},(a-1),j})$, where $\mathbf{w}_a$ is the weight at component a and $\mathbf{x}_{\text{new},(a-1)}$ is the value of $\mathbf{x}_{\text{new}}$ after deflation from the previous $a-1$ components, with $\mathbf{x}_{\text{new},0} = \mathbf{x}_{\text{new}} - \boldsymbol{\mu}$. The total contribution to all scores can again be computed as in (7). (The value of s_a, $a = 1, 2, \ldots, A$, can be calculated from the corresponding t_a's of the reference set; notice that in PLS the s_a's are related to the eigenvalues of $\mathbf{X}'\mathbf{Y}\mathbf{Y}'\mathbf{X}$.) Contribution plots on the SPE (a bar plot of each term in the summation of (12)) indicate how each variable contributes to the SPE for a specific observation. When an observation has a value of the $\text{SPE}_{\mathbf{x}_{\text{new}}}$ beyond the control limit, the variables on this plot that appear to have larger contributions than usual (and variables highly correlated to them) could be investigated as responsible for the deviation. An example is shown in Figure 12a which gives the contribution plot for observation 26 of Figure 11b. Recall that this observation corresponds to fouling in zone 1. The contribution plot shows that the major process variables contributing to the $\text{SPE}_{\mathbf{x}_{26}}$ are the coolant inlet and the reacting mixture outlet temperatures of zone 1 (as one might expect in a fouled reactor). The contribution plot for the SPE for point 66 (corresponding to fouling in both zones) indicates (Figure 12b) that the inlet coolant and reacting mixture outlet temperatures for both zones were responsible, consistent with what is expected.

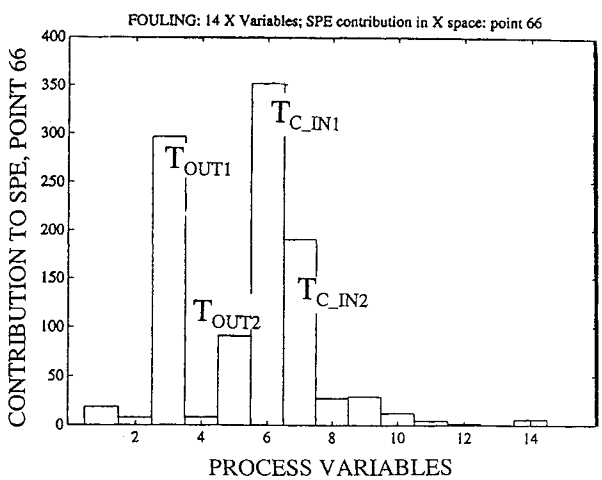
Multi-Block PLS

When a large number of variables is included in the x-space, the monitoring and diagnosing charts discussed in the previous sections may be difficult to interpret. The combined use of multi-block PLS (MB-PLS) and contribution plots may facilitate this task. In the MB-PLS approach, sets of process variables $\mathbf{X}$ are broken into meaningful blocks $\mathbf{X}_1, \mathbf{X}_2, \ldots$; usually each block corresponds to a process unit or a section of a unit. The principles behind multi-block data analysis methods and their algorithms can be found in Wold (1982) and Wangen and Kowalski (1988). Multi-block PLS is not equivalent to performing PLS on each block separately; all the blocks are considered together. MacGregor et al. (1994) discuss an application of MB-PLS to process monitoring and diagnosing for the LDPE reactor. In this case each block corresponds to one reactor zone. Plots of t_1 vs t_2 and SPE, obtained for each block of the process, were utilized to detect an abnormal event in the zone in which it occurred; then contribution plots were successfully used to assign causes for it. The following example demonstrates a monitoring procedure utilizing a chart of T_A^2 calculated from the scores and the SPE obtained from the MB-PLS model.

Figures 13a and 13b give respectively the T_3^2 chart (calculated from scores t_1, t_2 and t_3 of block 1) and



(a)



(b)

FIGURE 12. Process Monitoring in the LDPE Reactor, using PLS. The Disturbance is Fouling. (a) Contribution of Process Variables to SPE, Point 26, and (b) Contribution of Process Variables to SPE, Point 66.

the SPE chart for block 1 (corresponding to zone 1) for the same simulated process conditions of Figures 9a and 9b (i.e., fouling in different sections). Notice that by monitoring block 1 (zone 1, only) problems are only detected for observations 22–26 and 62–66 when zone 1 was affected by fouling. Observations 42–46 show a normal operation for this zone. Figures 13c and 13d give the T_3^2 and the SPE charts for block 2 (zone 2), for the same simulated fouling cases. Notice that problems were detected in zone 2 for observations 42–46 and 62–66. Notice also that the diagnostics of zone 2, in this case, do not show a problem for zone 2 for observations 22–26; recall that when a traditional multivariate chart on the process variables was used for this zone (Figure 9d) it showed a problem for these observations; MB-PLS successfully isolated the effect from zone 1, and showed clearly that in 22–26 the problem occurs only in zone 1. Utilizing the contribution plots for fault diagnosis (not shown here) has revealed that important variables for the fouling cases were the coolant inlet temperature and the reacting mixture outlet temperature, in the corresponding zones.

Although the monitoring and diagnosis procedures based on MB-PLS and PLS gave comparable results for this system with only fourteen variables, MB-PLS offers an advantage when bigger systems with tens or hundreds of variables are involved.

Discussion on the Projection Approaches

Control charts developed from the multivariate statistical projection methods (i.e., score charts $(t_1, t_2, t_3, \ldots)$ and SPE charts) have usually been based on a minimum of assumptions. The only essential one is that the PCA or PLS model has been built on a representative set of data that captures all common cause variations. The statistical variation about this model that captures "in-control" behavior, provides a reference distribution from which one can obtain control limits. Neither the assumptions of normality, nor independence are needed if one has enough data. (See for example the historical reference distribution concepts discussed in Chapters 2 and 3 of the text by Box, Hunter, and Hunter (1978).) The construction of control limits for the multivariate charts, based on reference distributions obtained from historical data, are discussed in Nomikos and MacGregor (1995).

If one wished to detect a very specific type of deviation, then more powerful charts (e.g., the "cuscore charts" of Box and Ramirez (1991)) could be devised, that are based on the specific nature of this deviation. However, to construct a whole set of charts based on statistical tests for specific types of special events would assume that one has a priori knowledge of all such events, and how they will affect the distribution of the variables. Then one would have to assess not only the ability of each chart to detect the event for which it is specifically designed, but also assess its robustness to false signals when any of the other events occur.

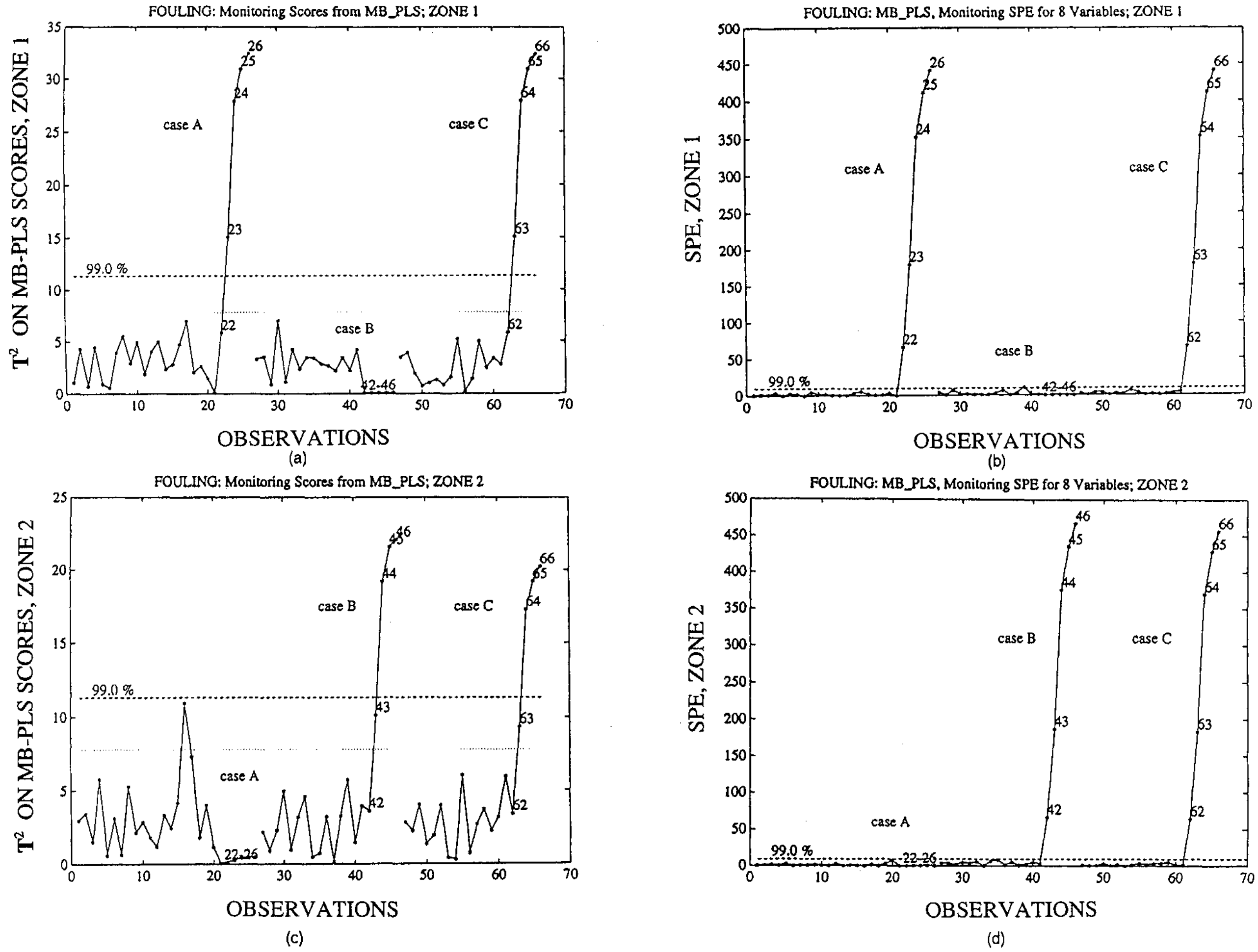


FIGURE 13. Process Monitoring in the LDPE Reactor, using MB-PLS. The Disturbance is Fouling. (a) T_3^2 Chart on 3 MB-PLS Scores, Zone 1, (b) SPE Chart, Zone 1, (c) T_3^2 Chart on 3 MB-PLS Scores, Zone 2, and (d) SPE Chart, Zone 2.

A typical multivariable process having hundreds of variables, may have dozens of special events which might occur, most of them affecting the mean and covariance structure of the variables in an unknown way. For example, special events arising from changes in impurity concentrations and fouling in the LDPE reactor did not simply affect a few process variable means. Rather, they simultaneously affected the mean and the covariance structure of variables in ways that were not known a priori.

For this reason we believe that using multivariate control charts which simply test whether the behavior of future data is consistent with past "in-control" behavior (captured by a PCA or PLS model) is a preferable approach when dealing with large multivariate processes. However, when a deviation is detected, since the charts involve no specific assumptions about the nature of the specific event, they must be supplemented with diagnostic procedures such as contribution plots. The contribution plots may not explicitly reveal the cause of the event, but *they point to the group of process variables that are no longer consistent with normal operating conditions*, and thereby focus the attention of the operators or the engineers and allow them to use process knowledge to deduce possible causes.

Concluding Remarks

Multivariate SPC monitoring methods can be applied to process variables as well as final quality variables. When the number of process variables is not too great, traditional control charts based on the χ^2 and T^2 statistics can be very effective for the early detection of special events. When the dimension of the problem becomes large, charts based on multivariate statistical projection methods (PCA and PLS) and their multi-block versions are generally much more effective. These latter methods can easily handle systems with many collinear and redundant variables and can accommodate missing observations in a straightforward fashion (Nelson, MacGregor, and Taylor (1996)).

Furthermore by monitoring the process variables as well as the quality variables, the diagnosis of assignable causes is much easier. By interrogating the underlying multivariable model for the in-control process, contribution plots for the process variables are easily constructed. Although these plots will not unequivocally reveal the assignable cause(s) they will at least point to that group of process variables that contributed most to the detected out-of-control situation, thereby aiding the engineer or operator in their search for assignable causes. The inherent condition in all the methods discussed here is that multivariate SPC charts (either traditional or based on projection methods) of process variables or product quality variables will detect only those unusual process events that will affect the variables that have been included in the chart.

There has been a significant amount of work done on monitoring continuous processes using projection methods (Kourti, Lee, and MacGregor (1996)). Industrial applications include Catalytic Cracking in Petroleum Refining (Slama (1991)), Applications in Mineral Processing (Hodouin, MacGregor, Hou, and Franklin (1993); Tano, Samskog, Gärde, and Skagerberg (1993); and Tano, Samskog, and Andreasson (1995)), a Ceramic Melter (Wise and Ricker (1989)), a Pulp Digester (Dayal, MacGregor, Taylor, Kildaw, and Marcikic (1994)), Photographic Paper Manufacturing (Miller, Swanson, and Heckler (1993)), and Cold Rolling and Annealing in the Steel Industry (Vaculik (1994, 1995)). Simulation studies have been performed on a Distillation Column and a Fluidized Bed Rector for Hydrocracking (Kresta, MacGregor, and Marlin (1991)), and on a Low Density Polyethylene Reactor (Skagerberg, MacGregor, and Kiparissides (1992), MacGregor et al. (1994)).

The above ideas have also been extended to monitoring time varying batch processes by Nomikos and MacGregor (1994, 1995); Kourti, Nomikos, and MacGregor (1995); and Kourti, Lee, and MacGregor (1996), with several applications on polymerization reactors and other processes. Finally, unpublished work (where the authors have been involved) includes applications of the projection methods to the monitoring and analysis of processes from the pharmaceutical industry, the semiconductor industry, and the tire and rubber and other chemical industries.

Acknowledgments

The authors wish to thank the nineteen companies of the McMaster Advanced Control Consortium for their financial support and their active involvement in joint research projects, which provide the opportunity to test new research ideas in industry.

References

ALT, F. B. (1985). "Multivariate Quality Control" in *Encyclopedia of Statistical Sciences, Vol. 6* edited by N. L. Johnson and S. Kotz. John Wiley & Sons, New York, NY.

BOX, G. E. P.; HUNTER, W. G.; and HUNTER, J. S. (1978). *Statistics for Experimenters.* John Wiley & Sons, New York, NY.

BOX, G. E. P. and RAMIREZ, J. (1991). "Cumulative Score Charts". Report No. 58, Center for Quality and Productivity Improvement, University of Wisconsin, Madison, WI.

DAYAL, B.; MACGREGOR, J. F. ; TAYLOR, P. A.; KILDAW, R.; and MARCIKIC, S. (1994). "Application of Feedforward Neural Networks and Partial Least Squares Regression for Modelling Kappa Number in a Continuous Kamyr Digester". *Pulp and Paper Canada* 95, pp. 26–32.

FUCHS, C. and BENJAMINI, Y. (1994). "Multivariate Profile Charts for Statistical Process Control". *Technometrics* 36, pp. 182–195.

GELADI, P. and KOWALSKI, B. R. (1986). "Partial Least Squares Regression: A Tutorial". *Analytica Chimica Acta* 185, pp. 1–17.

HAWKINS, D. M. (1993). "Regression Adjustment for Variables in Multivariate Quality Control". *Journal of Quality Technology* 25, pp. 170–182.

HODOUIN, D.; MACGREGOR, J. F.; HOU, M.; and FRANKLIN, M. (1993). "Multivariate Statistical Analysis of Mineral Processing Plant Data". *CIM Bulletin, Mineral Processing* 86, pp. 23–34.

HÖSKULDSSON, A. (1988). "PLS regression Methods". *Journal of Chemometrics* 2, pp. 211–228.

HOTELLING, H. (1947). "Multivariate Quality Control" in *Techniques of Statistical Analysis* edited by Eisenhart, Hastay, and Wallis. McGraw Hill, New York, NY.

JACKSON, J. E. (1956). "Quality Control Methods for Two Related Variables". *Industrial Quality Control* 12(7), pp. 4–8.

JACKSON, J. E. (1980). "Principal Components and Factor Analysis: Part I - Principal Components". *Journal of Quality Technology* 12, pp. 201–213.

JACKSON, J. E. (1991). *A User's Guide to Principal Components.* John Wiley & Sons, New York, NY.

KIPARISSIDES, C.; VEROS, G.; and MACGREGOR, J. F. (1993). "Mathematical Modelling, Optimization and Quality Control of High Pressure Ethylene Polymerization Reactors". *Journal of Macromolecular Science, Reviews in Macromolecular Chemistry and Physics* C33 (4), pp. 437–527.

KOURTI, T.; LEE, J.; and MACGREGOR, J. F. (1996). "Experiences with Industrial Applications on Projection Methods for Multivariate Statistical Process Control". *Computers in Chemical Engineering* 20 Supplementary, pp. S745–S750.

KOURTI, T.; NOMIKOS, P.; and MACGREGOR, J. F. (1995). "Analysis, Monitoring and Fault Diagnosis of Batch Processes Using Multi-block and Multiway PLS". *Journal of Process Control* 5, pp. 277–284.

KRESTA, J.; MACGREGOR, J. F.; and MARLIN, T. E. (1991). "Multivariate Statistical Monitoring of Process Operating Performance". *Canadian Journal of Chemical Engineering* 69, pp. 35–47.

LOWRY, C. A.; WOODALL, W. H.; CHAMP, C. W.; and RIGDON, S. E. (1992). "A Multivariate Exponentially Weighted Moving Average Control Chart". *Technometrics* 34, pp. 46–53.

MACGREGOR, J. F.; JAECKLE, C.; KIPARISSIDES, C.; and KOUTOUDI, M. (1994). "Process Monitoring and Diagnosis by Multi-Block PLS Methods". *Journal of the American Institute of Chemical Engineers* 40, pp. 826–838.

MACGREGOR, J. F. and KOURTI, T. (1995). "Statistical Process Control of Multivariate Processes". *Control Engineering Practice* 3, pp. 403–414.

MARDEN, J. I. and PERLMAN, M. D. (1990). "On the Inadmissibility of Step-Down Procedures for the Hotelling T^2 Problem". *Annals of Statistics* 18, pp. 172–190.

MARDIA, K. V.; KENT, J. T.; and BIBBY, J. M. (1979). *Multivariate Analysis.* Academic Press Ltd., San Diego, CA.

MASON, R. L.; TRACY, N. D.; and YOUNG, J. C. (1995). "Decomposition of T^2 for Multivariate Control Chart Interpretation". *Journal of Quality Technology* 27, pp. 99–108.

MILLER, P.; SWANSON, R. E.; and HECKLER, C. F. (1993). "Contribution Plots: The Missing Link in Multivariate Quality Control". Presented at Annual Fall Technical Conference of the American Society for Quality Control (Milwaukee, WI) and the American Statistical Association (Alexandria, VA).

MONTGOMERY, D. C. (1996). *Introduction to Statistical Quality Control,* 3rd ed. John Wiley & Sons, New York, NY.

MURPHY, B. J. (1987). "Selecting Out Of Control Variables with the T^2 Multivariate Control Procedure". *The Statistician* 36, pp. 571–583.

NELSON, P.; MACGREGOR, J. F.; and TAYLOR, P. A. (1996). "Missing Data Methods in PCA and PLS: Score Calculations with Incomplete Observations". *Chemometrics and Intelligent Laboratory Systems* 35 (in press).

NOMIKOS, P. (1995). Private Communication.

NOMIKOS, P. and MACGREGOR, J. F. (1994). "Monitoring of Batch Processes using Multi-way Principal Component Analysis". *Journal of the American Institute of Chemical Engineers* 40, pp. 1361–1375.

NOMIKOS, P. and MACGREGOR, J. F. (1995). "Multivariate SPC charts for Monitoring Batch Processes". *Technometrics* 37, pp. 41–59.

SHEWHART, W. A. (1931). *Economic Control Quality of Manufactured Product.* D. Van Nostrand Co., Inc., New York, NY.

SKAGERBERG, B.; MACGREGOR, J. F.; and KIPARISSIDES, C. (1992). "Multivariate Data Analysis Applied to Low Density Polyethylene Reactors". *Chemometrics and Intelligent Laboratory Systems* 14, pp. 341–356.

SLAMA, C. F. (1991). "Analysis of Industrial FCCU data using PCA and PLS". M.Eng. Thesis, McMaster University, Hamilton, ON, Canada.

SPARKS, R. S. (1992). "Quality Control with Multivariate Data". *Australian Journal of Statistics* 34, pp. 375–390.

TANO, K.; SAMSKOG, P.-O.; and ANDREASSON, B. (1995). "Multivariate Modelling and On-Line Data Presentation for Process Monitoring at LKAB". Presented at the International Federation of Automatic Control Symposium on Automation in Mining Mineral and Metal Processing. Sun City, South Africa.

TANO, K.; SAMSKOG, P.-O.; GÄRDE J.-C.; and SKAGERBERG, B. (1993). "Partial Least Squares Modelling of Process Data at LKAB: Predicting Chemical Assays in Iron Ore for Process Control". Presented at the International Symposium on the Application of Computers and Operations Research in the Mineral Industries. Canadian Institute of Mining, Metallurgy and Petroleum, Montreal, PQ, Canada.

TRACY, N. D.; YOUNG, J. C.; and MASON, R. L. (1992). "Multivariate Control Charts for Individual Observations". *Journal of Quality Technology* 24, pp. 88–95.

VACULIK, V. (1994). "Multivariate Data Analysis Using PCA/PLS with Extensions to Performance Monitoring at Dofasco". Presented at the Advanced Modelling and Control Seminar by the Association of Iron and Steel Engineers, Cleveland, OH.

VACULIK, V. (1995). "Applications of Multivariate Projection Methods in the Steel Industry". M.S. Thesis, Chemical Engineering Department, McMaster University, Hamilton, ON, Canada.

VEROS, G.; PAPADAKIS, M.; and KIPARISSIDES, C. (1993). "Mathematical Modeling of High Pressure Tubular LDPE Copolymerization Reactors". *Polymer Reaction Engineering* 3, pp. 427–460.

WADE, M. R. and WOODALL, W. H. (1993). "A Review and Analysis of Cause Selecting Charts". *Journal of Quality Technology* 25, pp. 161–169.

WANGEN, L. E. and KOWALSKI, B. (1988). "A Multi-block Partial Least Squares Algorithm for Investigating Complex Chemical Systems". *Journal of Chemometrics* 3, pp. 3–20.

WIERDA, S. J. (1994a). "Multivariate Statistical Process Control—Recent Results and Directions for Future Research". *Statistica Neerlandica* 48, pp. 147–168.

WIERDA, S. J. (1994b). *Multivariate Statistical Process Control.* Groningen Theses in Economics, Management and Organization. Wolters-Noordhoff, Groningen, The Netherlands, pp. 95–124.

WISE, B. M. (1991). "Adapting Multivariate Analysis for Monitoring and Modelling of Dynamic Systems". Ph.D. Dissertation, University of Washington, Seattle, WA.

WISE, B. M. and RICKER, N. L. (1989). "Feedback Strategies in Multiple Sensor Systems". *American Institute of Chemical Engineers Symposium Series on Process Sensing* 85, pp. 19–23.

WOLD, H. (1982). "Soft Modelling. The Basic Design and Some Extensions" in *Systems under Indirect Observation Vol. 2* edited by K. G. Joreskog and H. Wold. North Holland, Amsterdam, Chapter 1.

WOLD, S. (1978). "Cross-Validatory Estimation of the Number of Components in Factor and Principal Component Models". *Technometrics* 20, pp. 397–405.

ZHANG, G. X. (1985). "Cause Selecting Control Charts—A New Type of Quality Control Charts". *The QR Journal* 12, pp. 221–225.

Key Words: *Multivariate Analysis, Multivariate Control Charts, Principal Components, Statistical Process Control.*