A Composite Statistical Test for Detecting Changes of Steady States

A statistical procedure for determining whether process variables have undergone a change of steady states is developed by using a composite statistical test. Factors influencing the power of the test and the probability of Type I errors were studied through theoretical analysis, computer simulation, and application to plant data. They include the number of variables to be tested together, the number of measurements used in time averaging, the levels of significance of the tests, and the appropriate formulation of the hypotheses. The procedure developed here is useful in process data reconciliation and in applications that require an analysis of process trends.

S. Narasimhan, R. S. H. Mah, and A. C. Tamhane
Northwestern University
Evanston, IL 60201

J. W. Woodward, J. C. Hale
E. I. du Pont de Nemours and Company
Engineering Department
Wilmington, DE 19898

Introduction

Steady state is one of the most important and common assumptions made about a process. Depending on whether this assumption is made, an entirely different treatment of the process may follow. Much of the pilot plant data for process modeling and process design is collected when steady state conditions are attained. Preliminary data are taken only for the purpose of ascertaining the state of the operating conditions. Likewise, process data for plant material and energy balances and yield accounting are normally taken under presumed conditions of steady state.

On the other hand, in some applications changes of process conditions are the very essence of the analysis. Process dynamics is obviously central to process control applications. But the detection of steady states can also be useful in historical data recording. Hale and Sellars (1981) have proposed and implemented various heuristic algorithms for data compression to reduce the storage required for process data without losing trend information about the process behavior. In data compression schemes, data are recorded only when their numerical values undergo significant changes. In their schemes the magnitude of significant changes are prespecified. But detection of a change of steady states could provide an alternative procedure that takes into account the variability in the data in determining when the changes are significant.

In process data reconciliation and estimation the treatment is different depending on whether steady state conditions are assumed. For steady state processes the estimation procedure (Mah, 1982; Tamhane and Mah, 1985) take no account of past data. But for quasi-steady state processes the estimation technique utilizes past data as well as steady state material and energy conservation constraints (Stanley and Mah, 1977).

In all of the foregoing applications it is important to be able to determine whether a process is in a steady state. In a strict sense steady state conditions almost never prevail in practice. A judgment is made by the process engineer or operator as to when a steady state is supposedly attained. Previous work in this area has been limited. In this paper we propose a statistical procedure for determining if a change of steady states has occurred.

In our investigation a change in steady state is considered to take place if one or more of the true values of the measured process variables undergo a change. Notice that in this model the variability in the measurements as given by the measurement error covariance matrix may change without affecting a change of steady states. A combination of multivariate statistical tests is used to detect changes in the state of the variables using measurements on these variables. The variables are grouped together based on the knowledge that they are physically related and on the expectation that they are likely to change their states together. The measurements are assumed to be sampled at regular time intervals. A set of N successive measurement vectors constitutes a period, and N is referred to as the period size. It is assumed that a change of steady states may occur from one period to the next, but that within each period the process is in a steady state. The tests are based on sample statistics computed for each time period.

Computer simulation, theoretical analysis, and application to plant data are used to arrive at the following main conclusions of this study: The method proposed appears to provide a practica-

Correspondence concerning this paper should be addressed to R. S. H. Mah.

ble approach for detecting changes in steady states. Application to plant data indicates that the test should be formulated such that only when the combined change of the true values of the variables exceeds a prespecified threshold amount, should a change of steady states be assumed to take place. Selection of the variables to be tested together is an important consideration. The power of the composite test to detect a change in the state of the variables increases as we increase the number of variables to be tested together, provided these variables change their states simultaneously. The strategy of application of the test depends on how the process is expected to change states. The strategy proposed in this study is intended for quasi-steady state processes (Stanley and Mah, 1977).

Model and Assumptions

In this study process variables are considered to be in a steady state if their true values do not change with time. Our aim is to detect changes in the state of the variables given their measurements. The following assumptions are made in this regard:

(a) All process variables are measured directly.

(b) Measurements contain only random errors, which are normally distributed with mean $\underline{0}$.

(c) A time period is defined to consist of a set of N successive measurement vectors. Within each period the process variables are assumed to be steady.

(d) The covariance matrix of measurement errors is assumed to be unknown and is allowed to change from one period to the next.

(e) Successive measurement vectors are assumed to be mutually independent.

Using these assumptions, the measurement model can be described by

$$\underline{\tilde{x}}_{ki} = \underline{x}_k + \underline{v}_{ki} \quad i = 1, 2, \ldots, N \tag{1}$$

$$\underline{v}_{ki} \sim \mathcal{N}(\underline{0}, \underline{Q}_k) \quad i = 1, 2, \ldots, N \tag{2}$$

$$E[\underline{v}_{ki}\underline{v}'_{lj}] = \underline{0} \quad \text{for all } i, j, k, l \text{ with either } i \neq j \text{ or } k \neq l \tag{3}$$

Here the subscript k refers to the kth period, $\underline{x}_k$ is the vector of true values, $\underline{\tilde{x}}_{ki}$ is the ith vector of measurements in the kth period, and $\underline{v}_{ki}$ is the corresponding normally distributed vector of random errors with mean $\underline{0}$ and covariance matrix $\underline{Q}_k$. Thus the $\underline{\tilde{x}}_{ki}$ are independently distributed as

$$\underline{\tilde{x}}_{ki} \sim \mathcal{N}(\underline{x}_k, \underline{Q}_k), \quad i = 1, 2, \ldots, N \tag{4}$$

Assumptions (c) and (e) allow us to use the sample covariance matrix as an estimate of $\underline{Q}_k$. The estimate of $\underline{Q}_k$ is more complicated if steady state conditions are not assumed (Almasy and Mah, 1984).

Test Procedure

Let us consider consecutive time periods k and $k + 1$. Let us also consider a group of p variables that are chosen to be tested simultaneously. The test for a change in the true values of the variables has two forms, depending on whether the covariance matrices in the two periods are equal. The proposed test procedure proceeds in two stages. In the first stage we apply a test to determine whether the covariance matrices $\underline{Q}_k$ and $\underline{Q}_{k+1}$ are equal. Depending on the outcome of this test we apply the appropriate test in the second stage to determine whether the true vectors $\underline{x}_k$ and $\underline{x}_{k+1}$ are equal. We refer to this two-stage test procedure as the composite test. Figure 1 is a schematic representation of the composite test. The details of the composite test are given in the Appendix. We will refer to the different tests as test 1, test 2A, and test 2B, as defined in Figure 1. In our application we choose the period size to be equal for all periods. We also choose the same level of significance α_2, for test 2A and test 2B, while test 1 is applied using a level of significance α_1.

It is assumed that in practice a subset of the measured variables is chosen based on the specific application. The composite test is applied successively to pairs of consecutive time periods for this subset of variables. For example, we apply the composite test to this subset for periods 1 and 2, then for periods 2 and 3, and so on. This strategy of applying the test is appropriate for a quasi-steady state process (Stanley and Mah, 1977), which remains essentially at a steady state for a long interval of time and changes quickly from one steady state to another. To detect slow drifts in the variables, other strategies may be used. Such alternative strategies are discussed in the final section, Closing Remarks.

Evaluation of the Test Procedure

The performance of the composite test is characterized by two quantities:

1. The probability of its Type I error
2. Its power.

These performance characteristics are defined for each subset of variables as follows:

$Pr\{\text{Type I error}\} = Pr\{H_0$ of test 2A or test 2B is rejected when the variables in the subset are in a steady state$\}$

Power $= Pr\{H_0$ of test 2A or test 2B is rejected when one or more variables in the subset have changed their states$\}$

In the above definitions H_0 of test 2A or 2B is the hypothesis that the variables are in a steady state. Based on these performance characteristics we study the effect of the following factors on the performance of the composite test:

- The level of significance, α_1 of test 1.
- The dimension of the subset p.
- The period size, N.

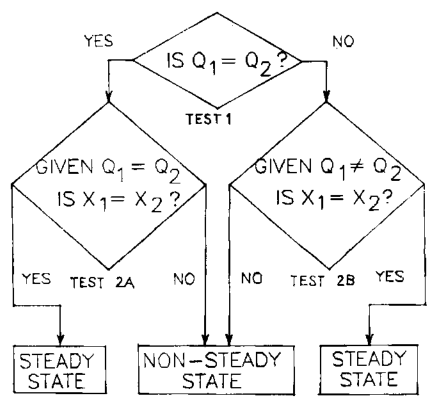

Figure 1. Composite statistical test.

It is difficult to obtain values of the probability of Type I error or power of the composite test analytically, although for test 2A this is possible. We shall therefore use a combination of simulation and analytical methods to study the effect of the factors listed above. First we shall study the effect of α_1 through computer simulation. Henceforth, we refer to the performance characteristics of the composite test simply as the probability of Type I error or power, respectively.

Effect of α_1

The choice of α_1 affects the behavior of test 1, which may affect the performance of the composite test. This effect is explained as follows.

Let us consider consecutive periods, say, periods 1 and 2. Ideally, if we know that $\underline{Q}_1 = \underline{Q}_2$, we can apply test 2A. Otherwise we can apply test 2B. Since test 1 is used to determine whether the covariance matrices are equal, there is a probability α_1 that test 2B may be fed with cases for which $\underline{Q}_1 = \underline{Q}_2$. Similarly there is a probability β_1, the probability of type II error of test 1, that test 2A may be fed with cases for which $\underline{Q}_1 \neq \underline{Q}_2$. For specified values of $\underline{Q}_1$, $\underline{Q}_2 (\underline{Q}_1 \neq \underline{Q}_2)$, p, and N, β_1 is uniquely determined by choosing α_1. Thus, even when $\underline{Q}_1 \neq \underline{Q}_2$, we interpret the effect of test 1 as the effect of α_1 on the performance of the composite test. A brief description of the simulation procedure used to evaluate the effect of α_1 follows. A more detailed description is given in Narasimhan (1984).

For a particular simulation, we choose values for p, N, α_1, α_2, $\underline{Q}_1$, $\underline{Q}_2$, and the vector of differences, $\underline{\delta}$ in the true values of the variables in the two periods. Given these values, the vector $\underline{d}$ defined in Eq. A15 and the sample covariance matrices $\underline{S}_1$ and $\underline{S}_2$ defined in Eq. A21 are generated on the computer. Efficient methods for generating these quantities are described in Narasimhan (1984). The test statistics defined in Eqs. A2 and A14 are calculated and the composite test is applied. Each such application is a simulation trial, and a simulation run consists of N_T trials. If N_R is the total number of simulation trials rejected by test 2A and test 2B, then we calculate the proportion $\hat{s}$ given by

$$\hat{s} = \frac{N_R}{N_T} \tag{5}$$

If in a simulation run $\underline{\delta} = \underline{0}$, then this implies that there is no change in the true values of variables and therefore we are simulating an idealized steady state situation. The proportion $\hat{s}$ calculated for this simulation run is an estimate of the probability of Type I error of the composite test. On the other hand if $\underline{\delta} \neq \underline{0}$, then we are simulating a nonsteady state situation for a magnitude of change $\underline{\delta}$ in the true values of the variables. The proportion $\hat{s}$ calculated for this case is an estimate of the power of the composite test.

The following four simulation runs were used to evaluate the effect of α_1.

1. $\underline{Q}_1 = \underline{Q}_2 = \underline{I}, \underline{\delta} = \underline{0}$
2. $\underline{Q}_1 = \underline{I}, \underline{Q}_2 = 2.0\underline{I}, \underline{\delta} = \underline{0}$
3. $\underline{Q}_1 = \underline{Q}_2 = \underline{I}, \underline{\delta} \neq \underline{0}$
4. $\underline{Q}_1 = \underline{I}, \underline{Q}_2 = 2.5\underline{I}, \underline{\delta} \neq \underline{0}$

Simulation runs 1 and 2 were used to evaluate the effect of α_1 on the probability of Type I error, and simulation runs 3 and 4 were used to evaluate the effect of α_1 on the power of the composite test. In all the simulation runs we used the values of $p = 3$ and $N = 30$. For each simulation run the composite test was applied using two different values of α_2 and three different values of α_1 in the range 0.01 to 0.1. The results of the simulation showed that for fixed values of p, N, $\underline{Q}_1$, $\underline{Q}_2$, α_2, and $\underline{\delta}$, the value of α_1 in the range 0.01 to 0.1 does not significantly affect the probability of Type I error or the power of the composite test. Therefore any value of α_1 in this range would be acceptable; we adopt the choice of $\alpha_1 = 0.05$ from now on.

Instead of performing extensive simulations for different combinations of $\underline{Q}_1$, $\underline{Q}_2$, p, and N, the effect of other factors p and N can be analyzed theoretically by examining their effect on test 2A and test 2B. Depending on the values of $\underline{Q}_1$ and $\underline{Q}_2$, the following three cases can occur.

Case 1: $\underline{Q}_1 = \underline{Q}_2$. For this case, the probability that test 2B will be applied is small, since α_1 is small. Therefore, the overall performance of the composite test is determined by test 2A.

Case 2: $\underline{Q}_1$ moderately different from $\underline{Q}_2$. Both test 2A and test 2B influence the performance of the composite test but their relative role depends on the power of test 1.

Case 3: $\underline{Q}_1$ highly different from $\underline{Q}_2$. The power of test 1 in this case is high and therefore the performance of the composite test is determined by test 2B.

Effect of p and N on the probability of Type I error

Let all the variables be at steady state. For cases 1 and 3 described above, there is a high probability that the correct test (test 2A or test 2B) is applied in the second stage and therefore the probability of Type I error of the composite test is approximately equal to α_2. For case 2 it has been shown (Ito and Schull, 1964) that when the period sizes are equal, as in our case, test 2A is robust to small differences in $\underline{Q}_1$ and $\underline{Q}_2$ and gives a probability of Type I error approximately equal to α_2. Moreover, the robustness of test 2A is enhanced for larger period sizes. We also note that the probability of Type I error of test 2B is always controlled at α_2. Therefore even for case 2 and hence for all cases, the probability of Type I error of the composite test is approximately equal to α_2, and independent of p, N, $\underline{Q}_1$, or $\underline{Q}_2$.

Effect of p on power

Here we attempt to answer the following question: From a given set of variables, how many variables should be chosen and tested together so that the power of the test for detecting changes in steady states is maximized? We are therefore interested in the effect of the dimension of the subset p on the power of the test. It is assumed that once a subset of p variables is chosen, any remaining variables are dismissed from further consideration.

The dimension of the subset, p, affects the power of the test in two counteracting ways. As p increases, the test criteria $T^2_{p,f_A}(\alpha_2)$ for test 2A and $T^2_{p,f_B}(\alpha_2)$ for test 2B increase, which causes a reduction in the power of the composite test. On the other hand, the noncentrality parameter defined in Eq. A24 may increase with p, which causes an increase in the power of the test. The net effect on the power is difficult to ascertain if the variables change by different extents. However, for the particular case of the same relative change in all variables, we can easily show as follows that the power of test 2A increases with p.

Let $\underline{D}$ be a matrix such that

$$\underline{D}\underline{D}' = \underline{Q}_{av} \tag{6}$$

where

$$\underline{Q}_{av} = (\underline{Q}_1 + \underline{Q}_2)/2 \tag{7}$$

Consider the case where the vector of changes in the true values of the variables $\underline{\delta}$ is such that

$$\underline{D}^{-1}\underline{\delta} = \Delta \underline{1} \tag{8}$$

where $\underline{1}$ is a vector with elements ± 1 and Δ is a positive number. We refer to Δ as the relative extent by which each variable changes. If the covariance matrices of the measurement errors are diagonal, then Eq. 8 implies that $|\delta_i|/\sigma_{i,av}$ is the same for all variables, where $\sigma^2_{i,av}$ are the diagonal elements of $\underline{Q}_{av}$. So long as the relative extents are the same, i.e., Eq. 8 is satisfied, the noncentrality parameter may be obtained from Eq. A24 to be

$$\tau = Np(\Delta)^2/2 \tag{9}$$

which increases with p.

We first note that the power of the composite test is determined by test 2A (resp., test 2B) when $\underline{Q}_1 = \underline{Q}_2$ (resp., $\underline{Q}_1$ highly different from $\underline{Q}_2$), whereas both test 2A and test 2B determine the power of the composite test when $\underline{Q}_1$ is moderately different from $\underline{Q}_2$. When $\underline{Q}_1 = \underline{Q}_2$, the power of test 2A can be obtained analytically as described in the Appendix for given values of α_2, p, N, and τ. Power curves for test 2A plotted against Δ for different values of p are shown in Figure 2. For any given value of Δ it is observed that the power increases with p. Due to the robustness of test 2A when the period sizes are equal (Ito and Schull, 1964), these power curves are valid for test 2A even when $\underline{Q}_1$ is moderately different from $\underline{Q}_2$. The power of test 2B when $\underline{Q}_1 \neq \underline{Q}_2$ can be estimated only through simulation. Power curves for test 2B obtained through simulation are also shown in Figure 2 for particular choices of $\underline{Q}_1$ and $\underline{Q}_2$. We again observe that the power of test 2B increases with p for a given value of Δ. Thus for the case when all variables change by the same relative extent Δ, the power of the composite test increases with p, whatever the values of $\underline{Q}_1$ and $\underline{Q}_2$. This indicates that by grouping variables that change their states simultaneously, we may be able to increase the power of the test.

A second question that may be important with regard to the effect of p is the following: Given a set of M variables, how should the variables be grouped together and tested without discarding any variable? The two extreme possibilities are:

1. Testing each variable independently
2. Testing all the variables together.

It should be noted that in this case, the dimension of the group, p, and the number of groups formed, g (such that $pg = M$), both affect the performance of the composite test. This question is not addressed in our investigation, but is a possible direction for future research.

Effect of N on power

The value of N should be chosen greater than p, so that an estimate of the covariance matrix $\underline{Q}_k$ can be obtained from the data in every period k. It can be observed from Eq. A24 that for any nonzero magnitude of change $\underline{\delta}$ in the true values of the variables, the noncentrality parameter increases with N, which increases the power of test 2A or 2B. It can also be observed that the denominator degrees of freedom (defined by f_A and f_B in Eqs. A17 and A20, respectively) increase with N. This decreases the test criteria of test 2A or test 2B, as may be verified by reference to their tabulated values, with resultant increase in their powers. Therefore the power of the composite test also increases with N.

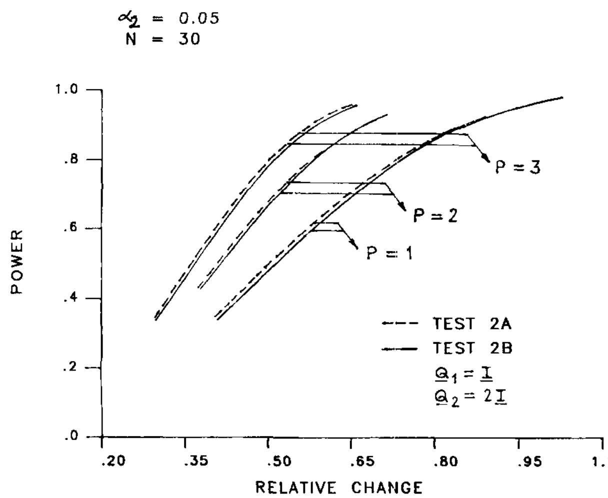

Figure 2. Power as a function of Δ for different *p*.

Application to Plant Data

The steady state detection test was applied to plant data to examine the practical utility of the composite test. Plant data were obtained from an industrial plant at Beaumont, Texas, for the process represented by Figure 3. Measurements were made at 1 min intervals for 20 variables over a period of 35 days. Fifteen of these variables are flow variables, which are shown by double slashes in Figure 3; the remaining five variables are temperature variables, which are shown by single slashes. The recorded data, which contain more than one steady state, were used to evaluate the performance of the composite test.

A preliminary analysis of the data was first conducted to test for the assumption of multivariate normality. We applied the modified Raleigh test recommended by Koziol (1983) to our data. We tested this assumption for a measurement vector of dimension $p = 3$, which is the maximum dimension subsequently used in our application. Twenty-four random samples, each con-

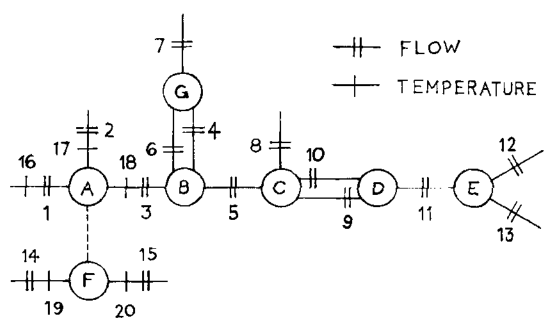

Figure 3. Process graph.

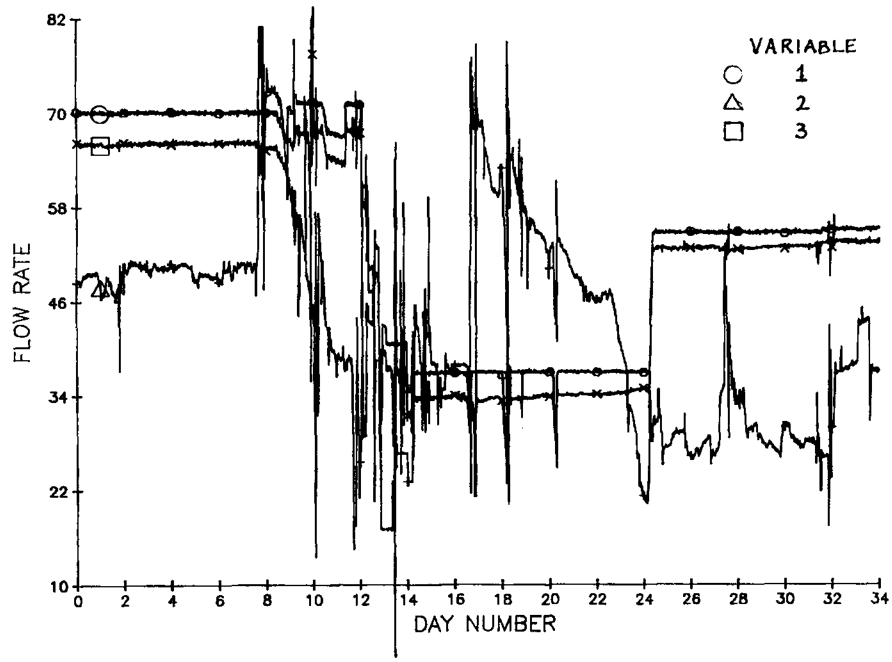

Figure 4. Plot of 15 min averages.

sisting of 100 measurement vectors were tested. The results showed that in 14 out of the 24 cases the null hypothesis of multivariate normality is not rejected for a level of significance less than or equal to 0.3. In the remaining 10 cases the null hypothesis is rejected even for a level of significance equal to 0.01. It is possible, however, that in these cases the hypothesis of multivariate normality is rejected because the data used are not obtained under steady state conditions, which is an assumption basic to the Raleigh test. Evidence is also available (Seber, 1984) that tests 2A and 2B are quite robust to nonnormality, particularly when the period sizes are equal. Based on the above results and considerations, we may reasonably assume multivariate normality for our data.

In order to examine the actual performance of the test a basis for comparison has to be chosen. A perception of steady state was obtained by plotting measurements of variables averaged over different time periods. The plots for variables 1, 2, and 3 are shown in Figures 4 to 6, where the time period for averaging is 15 min, 4 h, and 1 day, respectively. For the convenience of display a different scale was used for variable 2 in these figures.

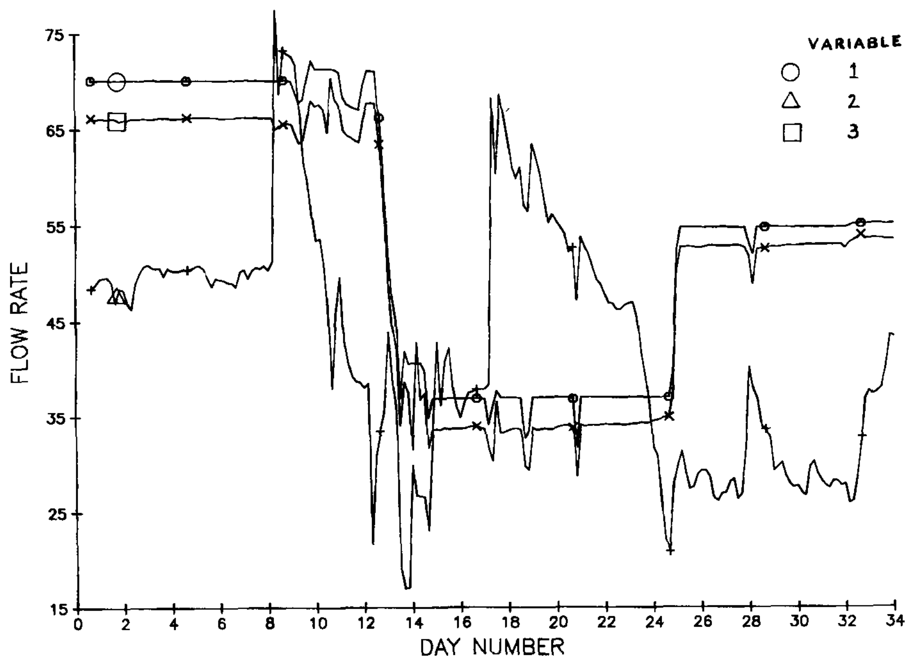

Figure 5. Plot of four-hour averages.

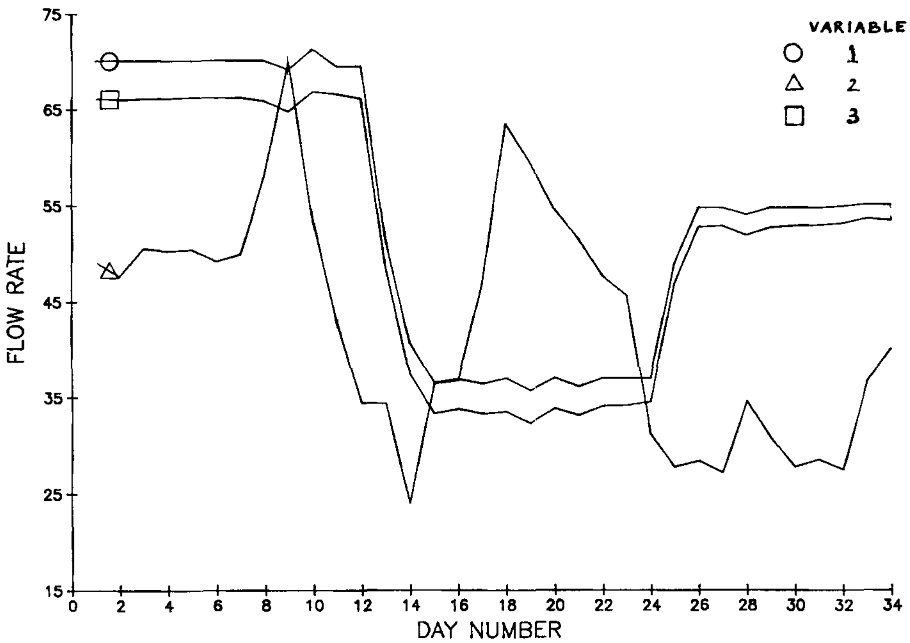

Figure 6. Plot of one-day averages.

The value of variable 2 shown in these figures is 18 times the actual flow rate. If a short time period is used in data averaging, the pattern of behavior of the variables is obscured by high-frequency noises. In the context of our application this situation is in evidence with time periods of 15 min or less. On the other hand, if a very long time period is used, all changes are obliterated by the averaging and no useful conclusions can be reached. In our case, time periods of eight days or more seem to fall into this category. In between these extremes and over a wide range of time periods, however, there is an unambiguous pattern of steady states recognizable by any reasonable observer. In our application we can conclude that variables 1, 2, and 3 are steady for the first eight days, after which there is a transient period of six days. We can also conclude that variables 1 and 3 exhibit almost identical behavior.

The performance of the test as given by the probability of Type I error and power was examined for different values of p, N, and α_2. The following input specifications are necessary for this purpose:

- The time intervals over which the variables are steady and nonsteady, respectively.
- A set of variables that behave similarly (change their states together).

Table 1. Parameters Used in Applying Composite Test to Plant Data

No. variables chosen	10
Variables chosen	1, 2, 3, 4, 14, 15, 16, 17, 19, 20
Dimension of measurement vector p	1, 2, and 3
Level of significance, test 1	0.05
Levels of significance, test 2A or 2B	0.10, 0.05, 0.01
Period sizes	15, 30, and 60

Ten variables that behave similarly were chosen by examining their plots to form a pool from which subsets of variables are later selected. The first 184 h was specified as the interval of steady state. The specified length of the nonsteady state interval was 136 h, beginning 208 h from the starting day. The composite test was applied to the selected variables over the specified time intervals, for different values of p, N, and α_2. Table 1 gives the values of the different parameters used for this application.

In order to compare readily the results of applying the test for different values of p, N, and α_2, the results obtained for different subsets are pooled together and reported in terms of a proportion $\hat{r}$, which is given by

$$\hat{r} = \frac{\text{Total number of subsets rejected by the composite test}}{\text{Total number of subsets}} \tag{10}$$

While the proportion $\hat{s}$ defined by Eq. 5 is an estimate of the performance characteristics of the composite test for one subset of variables, the proportion $\hat{r}$ used here is an average estimate over all subsets of the performance characteristics for chosen values of p, N, and α_2. It should be noted that when the variables are steady $\hat{r}$ gives an unbiased estimate of the probability of Type I error of the composite test. If the variables are not steady, then $\hat{r}$ is the average estimate of the power of the test for a given choice of p, N, and α_2. The value of $\hat{r}$ was calculated separately for the two specified time intervals. On the basis of the specified time intervals these proportions represent estimates of the probability of Type I error and the average power of the test, respectively.

Discussion of results

The results are presented in Figures 7 and 8. In these figures the proportion of subsets rejected is plotted against the period size N, with p as a parameter. Figure 7 corresponds to the time

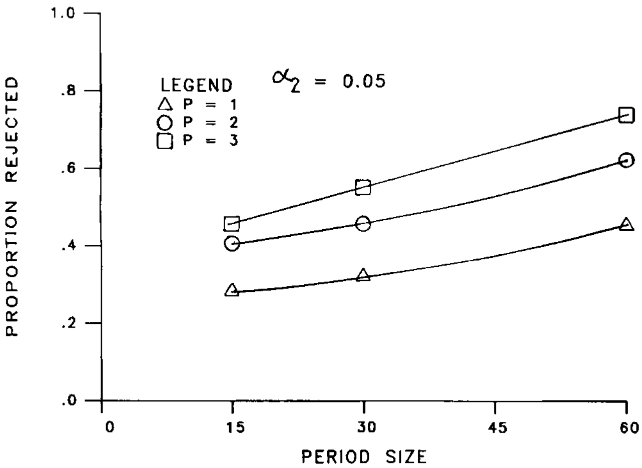

Figure 7. Proportion of trials rejected when variables are steady.

interval of steady state and therefore gives the variation of the estimate of probability of Type I error with N and p. Figure 8 gives the variation of the estimate of power with N and p. The results are presented for $\alpha_2 = 0.05$. Similar results were obtained for $\alpha_2 = 0.01$ and $\alpha_2 = 0.10$ (Narasimhan, 1984), with the estimated probability of Type I error and power decreasing with α_2 as expected.

From Figure 8 we observe that the power increases as p or N increases, which agrees with the theoretical conclusions reached for the effects of p and N on the power of the test in the earlier subsections "Effect of p on power" and "Effect of N on power."

From Figure 7 we observe the following:

- The estimated probability of Type I error is much higher than the chosen value of α_2.
- The estimated probability of Type I error increases with p and N.

Due to our subjectivity in specifying the interval of steady state, the estimated probability of Type I error may be higher than α_2. However, the second of these observations contradicts the earlier theoretical conclusion that the probability of Type I

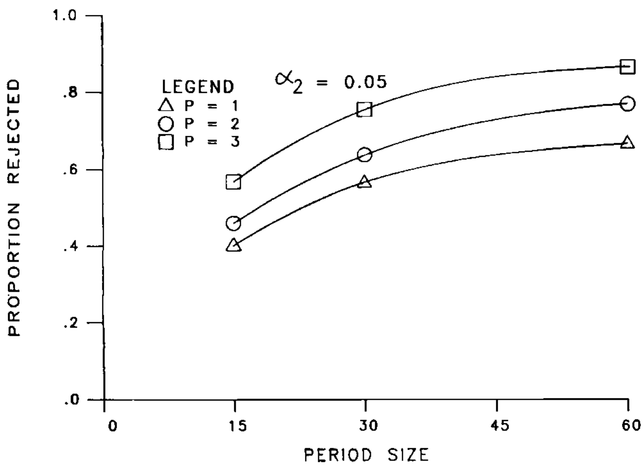

Figure 8. Proportion of trials rejected when variables are unsteady.

error is independent of p and N. This apparent contradiction may be explained if over the time interval of steady state specified, the variables are not truly steady but change by small extents that are detected by the test. Therefore, the proportion $\hat{r}$ estimated for this time interval also represents an estimate of the power of the test, which increases with p and N.

The hypotheses of test 2A and test 2B can be reformulated to allow small changes in the variables to go undetected. One approach for doing this is described and tested in the following section.

Modified formulation of hypotheses

Let $\underline{\delta}$ be the vector of changes in the true values from period k to $k + 1$. The hypotheses of test 2A and test 2B are reformulated as

$$H_0 : \underline{\delta}'(\underline{Q}_{av})^{-1}\underline{\delta} \le p(\Delta^*)^2$$
$$H_1 : \underline{\delta}'(\underline{Q}_{av})^{-1}\underline{\delta} > p(\Delta^*)^2 \qquad (11)$$

where $\underline{Q}_{av}$ is defined in Eq. 7. The original formulation of the hypotheses described by Eq. A13 in the appendix is a special case of the above formulation with Δ^* equal to zero. The composite test is applied as before with the modification that the null hypothesis of test 2A or test 2B is rejected if the corresponding test statistic exceeds the α_2 point of the corresponding noncentral T^2 distribution (with degrees of freedom p and f_A or p and f_B) and noncentrality parameter $\tau = Np(\Delta^*)^2/2$. The percentage point of the noncentral T^2 distribution is obtained using Eq. A26.

The reformulated hypotheses, Eq. 11, imply that if the weighted sum of squares of the changes in the true values of the variables represented by $\underline{\delta}'(\underline{Q}_{av})^{-1}\underline{\delta}$ is greater than $p(\Delta^*)^2$, then the state of the variables is deemed to have changed. In practical applications Δ^* may be specified by the user based on the changes in the true values of the variables that can be tolerated, although this may require knowledge about the behavior of the process variables.

Note that the threshold value on $\underline{\delta}'(\underline{Q}_{av})^{-1}\underline{\delta}$ is chosen directly proportional to p since it represents the total weighted sum of changes in p variables that can be tolerated and is therefore expected to increase with p. The composite test was applied with the reformulated hypotheses for different values of Δ^*.

Figures 9 and 10 show the variation of the estimated probability of Type I error with N and p, respectively, for different values of Δ^*. Figure 9 shows that the estimated probability of Type I error becomes largely independent of N for Δ^* greater than $\sqrt{0.2}$. Moreover, as Δ^* becomes greater than 1.0 the estimated probability of Type I error is nearly equal to α_2. Figure 10 shows that the estimated probability of Type I error still depends slightly on p, although the extent to which it depends is reduced when Δ^* is greater than $\sqrt{0.2}$. Figures 11 and 12 show the variation in the estimated power with N and p, respectively, for different values of Δ^*. The estimated power decreases as Δ^* increases. Thus there exists a trade-off between the probability of Type I error and power as Δ^* increases.

Ideally we would like to choose a value of Δ^* such that the following conditions are met:

- The actual probability of Type I error is independent of N and p, and is close to the chosen value of α_2.
- The actual power of the test is high.

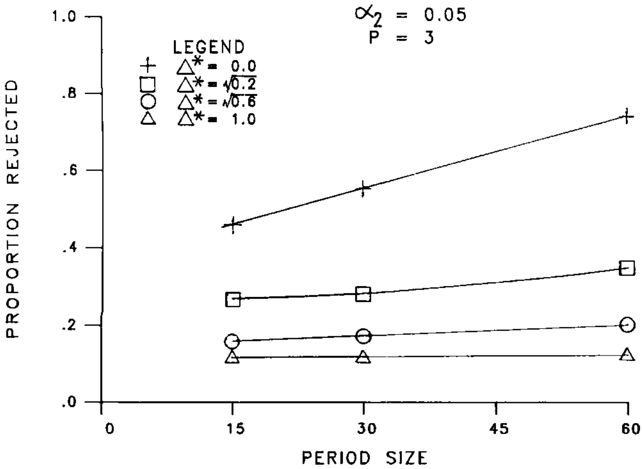

Figure 9. Proportion of subsets rejected when variables are steady, using reformulated hypothesis.

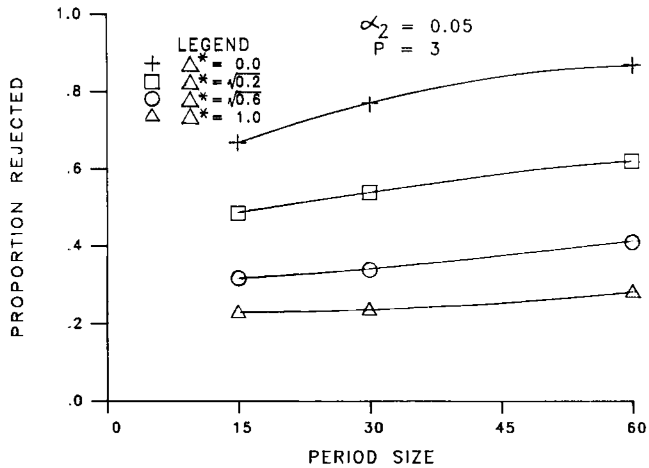

Figure 11. Proportion of subsets rejected when variables are unsteady, using reformulated hypothesis.

A high value of Δ^* meets the first condition but gives an unacceptably low power. A low value of Δ^*, on the other hand, satisfies the second condition at the expense of a high probability of Type I error. The exact choice of Δ^* depends on the application. If it is more important not to miss a change of steady state than to mispredict a change when no change has actually taken place, then we would demand a higher power at the expense of a higher probability of Type I error. In this case $\Delta^* = \sqrt{0.2}$ seems to be a good choice.

Closing Remarks

To summarize, for the original formulation of the hypotheses ($H_0 : \underline{\delta} = \underline{0}$ vs. $H_1 : \underline{\delta} \neq \underline{0}$) the power of the test may be increased by increasing p, provided all the variables change their states simultaneously, but the nature of the process may limit the number of variables which can be grouped together. Another way to increase the power of the test is to increase N by either increasing the frequency of sampling or by choosing a longer time period. However, the time period should not be chosen so long as to violate the assumption of steady state within each time period. The choice of α_1 in the range 0.01–0.10, does not affect the probability of Type I error or the power of the composite test. By increasing α_2 the power can be increased but only at the expense of higher probability of Type I error. In practical applications, it may be necessary to adopt the reformulated hypotheses, Eq. 11. The choice of Δ^* can be made with a view to make the probability of type I error almost independent of N and p.

The strategy for applying the composite test depends on how we expect the process variables to change their states. For quasi-steady state (QSS) process variables, which remain steady for long intervals of time and which change relatively quickly from one steady state to another, the composite test can be applied to pairs of consecutive time periods as in this study. A continuous process is normally operated under a steady state, and changes of steady state are deliberately imposed by changes of feedstock, desired product slate, and/or operating conditions. It is usually desirable to accomplish the changes in the shortest time possible. For these reasons QSS processes are a very important, if not the dominant, class of processes. However, if we are interested in detecting slow drifts in variables, then we may apply the composite test to periods 1 and 2, 1 and 3, and so on until we detect for the first time a change in steady state, say for example, between periods 1 and k. We may then continue applying the

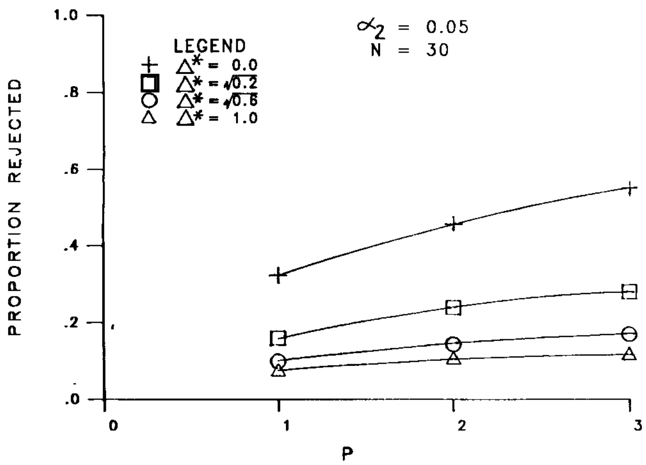

Figure 10. Proportion of subsets rejected when variables are steady, using reformulated hypothesis.

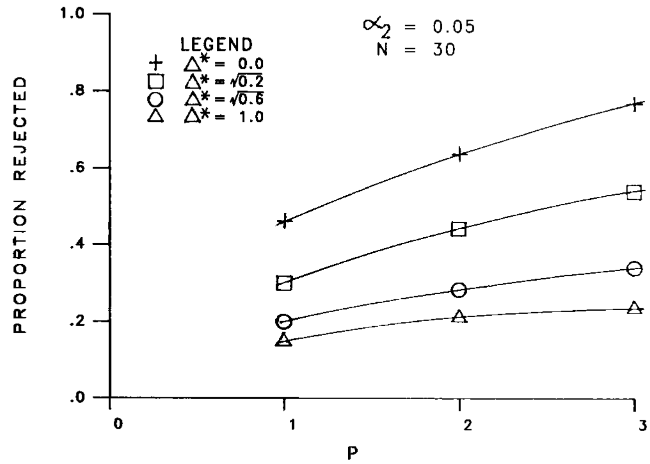

Figure 12. Proportion of subsets rejected when variables are unsteady, using reformulated hypothesis.

test to periods k and $k+1$, k and $k+2$, and so on. Other strategies may also be considered.

The method proposed in this study for detecting changes in steady state represents a preliminary effort in this area. It may be possible to achieve these objectives through alternative methods. The development of such methods is open to future research.

Acknowledgment

This work was supported by the National Science Foundation Grant No. CPE 8115161 and by the Du Pont Educational Foundation.

Notation

$\underline{A}_k$ = $N-1$ times the sample covariance matrix, Eq. A4
$\underline{d}$ = difference of average values of measurements, Eq. A15
$\underline{D}$ = square root of matrix $(\underline{Q}_1 + \underline{Q}_2)$, Eq. 6
f = parameter, Eq. A7
f_A = denominator degrees of freedom for test 2A
f_B = denominator degrees of freedom for test 2B
g = number of groups such that $pg = M$
$\underline{I}$ = identity matrix
k = subscript for the kth period
M = total number of process variables measured
n = $N-1$
N = number of measurements in a period (period size)
N_R = number of simulation trials rejected
N_T = number of simulation trials performed
p = number of variables in a subset
$\underline{Q}(\underline{Q}_k)$ = true measurement error covariance matrix
$\underline{Q}_{av}$ = covariance matrix, Eq. 7
$\hat{r}$ = proportion, Eq. 10
$\hat{s}$ = proportion, Eq. 5
$\underline{S}_k$ = sample covariance matrix
W = test statistic for test 1, Eq. A2
w_2 = parameter, Eq. A9
$\underline{x}_k$ = vector of true values
$\bar{\underline{x}}_k$ = average value of measurements
$\tilde{\underline{x}}_{ki}$ = ith measurement vector in period k
Z = quantity, Eq. A12
$\underline{0}$ = vector with all elements having value 0
$\underline{1}$ = vector with elements ± 1

Greek letters

α_1 = level of significance of test 1
α_2 = level of significance of test 2A or test 2B
β_1 = probability of Type II error of test 1
τ = noncentrality parameter, Eq. A24
σ_i^2 = diagonal elements of $\underline{Q}$
$\sigma_{1i}^2(\sigma_{2i}^2)$ = diagonal elements of $\underline{Q}_1\,(\underline{Q}_2)$
$\sigma_{i,av}^2$ = diagonal elements of $\underline{Q}_{av}$
$\underline{\delta}$ = vector of changes in the true values
δ_i = change in true value for the ith variable
Δ = relative change in each variable, Eq. 8
Δ^* = parameter, Eq. 11
ρ = parameter, Eq. A8

Other symbols

$F_{p,f}$ = F distribution with degrees of freedom p and f
$F_{p,f,\tau}$ = noncentral F distribution with degrees of freedom p, f, and noncentrality parameter τ
$\mathcal{N}(.,..)$ = normal distribution with mean . and covariance matrix . .
$Pr\{.\}$ = probability of .
$T_{p,f}^2$ = T^2 distribution with degrees of freedom p and f
$T_{p,f,\tau}^2$ = noncentral T^2 distribution with degrees of freedom p, f, and noncentrality parameter τ
χ_f^2 = chi-squared distribution with f degrees of freedom
$|.|$ = determinant of matrix .
$\sim$ = "is distributed as"

Appendix: The Composite Test

The composite test consists of three separate tests applied to sample statistics derived from two consecutive periods of time. The two periods are denoted by subscripts 1 and 2. The formulae are given for the case of equal period sizes.

Test 1: Test of sample covariance matrices

This test (Anderson, 1957, pp. 247–256) is used to test whether the covariance matrices of two normal distributions are equal. In our case the hypotheses are formulated as

$$H_0 : \underline{Q}_1 = \underline{Q}_2$$
$$H_1 : \underline{Q}_1 \neq \underline{Q}_2 \qquad (A1)$$

The test statistic is given by

$$W = \frac{2^{pn}|\underline{A}_1|^{0.5n}|\underline{A}_2|^{0.5n}}{|\underline{A}_1 + \underline{A}_2|^n} \qquad (A2)$$

where

$|\underline{A}_k|$ denotes the determinant of $\underline{A}_k$, $k = 1, 2$

$$n = N - 1 \qquad (A3)$$

$$\underline{A}_k = \sum_{i=1}^{N} (\tilde{\underline{x}}_{ki} - \bar{\underline{x}}_k)(\tilde{\underline{x}}_{ki} - \bar{\underline{x}}_k)', \quad k = 1, 2 \qquad (A4)$$

$$\bar{\underline{x}}_k = \frac{1}{N}\sum_{i=1}^{N} \tilde{\underline{x}}_{ki} \qquad (A5)$$

The null hypothesis is rejected if $W \leq W(\alpha_1)$ where $W(\alpha_1)$ is the test criterion at a level of significance α_1. The value of $W(\alpha_1)$ cannot be obtained exactly but may be approximated using the following equations.

Let W_r be the random variable that has the same distribution as W. Then

$$Pr\{-2\rho \log W_r \leq Z\} = Pr\{\chi_f^2 \leq Z\} + w_2[Pr\{\chi_{f+4}^2 \leq Z\} - Pr\{\chi_f^2 \leq Z\}] \qquad (A6)$$

where

$$f = p(p+1)/2 \qquad (A7)$$

$$\rho = 1 - (2p^2 + 3p - 1)/[4n(p+1)] \qquad (A8)$$

$$w_2 = p(p+1)[(p-1)(p+2)(1.75/n^2) - 6(1-\rho)^2]/48\rho^2 \qquad (A9)$$

We want the level of significance to be α_1 for this test. This means that

$$Pr\{W_r \leq W(\alpha_1)\} = \alpha_1 \qquad (A10)$$

Since ρ is positive

$$Pr\{-2\rho \log W_r \leq Z\} = 1 - \alpha_1 \qquad (A11)$$

where

$$Z = -2\rho \log W(\alpha_1) \qquad (A12)$$

We can iteratively obtain Z for a given value of α_1 using Eqs. A6 and A11.

Test 2A: Test of sample means for equal but unknown covariance matrices

The Hotelling T^2 test (Anderson, 1957, pp. 101–115) is used to test whether the means of two normal distributions are equal given that their covariance matrices are equal but unknown. The hypotheses in our case are

$$H_0 : \underline{x}_1 = \underline{x}_2$$
$$H_1 : \underline{x}_1 \neq \underline{x}_2 \qquad (A13)$$

The test statistic is given by

$$T^2_{p,f_A} = 0.5N\underline{d}'\underline{S}^{-1}\underline{d} \qquad (A14)$$

where

$$\underline{d} = \overline{\underline{x}}_1 - \overline{\underline{x}}_2 \qquad (A15)$$

$$\underline{S} = (\underline{A}_1 + \underline{A}_2)/2n \qquad (A16)$$

$$f_A = 2n \qquad (A17)$$

and n is defined in Eq. A3.

The null hypothesis is rejected if

$$T^2_{p,f_A} \geq T^2_{p,f_A}(\alpha_2) \qquad (A18)$$

T^2_{p,f_A} is distributed under the null hypothesis as $pf_A/(f_A - p + 1)$ times an F random variable with p and $f_A - p + 1$ degrees of freedom. Therefore $T^2_{p,f_A}(\alpha_2)$, the upper α_2 point of T^2_{p,f_A} under the null hypothesis, is given by

$$\frac{(f_A - p + 1)}{pf_A} T^2_{p,f_A}(\alpha_2) = F_{p,f_A - p + 1}(\alpha_2) \qquad (A19)$$

This is used to obtain $T^2_{p,f_A}(\alpha_2)$ for a chosen level of significance α_2.

Test 2B: Test of sample means for unequal and unknown covariance matrices

This test, given by Yao (1965), is similar to the T^2 test and is used when the covariance matrices of the two samples are unequal. The alternative hypotheses are still given by Eq. A13. The test statistic is computed according to Eq. A14 with f_A replaced by f_B, which is the approximate degrees of freedom given by

$$\frac{1}{f_B} = \frac{1}{n}\sum_{k=1}^{2}\left(\frac{\underline{d}'\underline{S}^{-1}\underline{S}_k\underline{S}^{-1}\underline{d}}{2\underline{d}'\underline{S}^{-1}\underline{d}}\right)^2 \qquad (A20)$$

where

$$\underline{S}_k = \underline{A}_k/(N - 1) \quad k = 1, 2 \qquad (A21)$$

$$\underline{S} = (\underline{S}_1 + \underline{S}_2)/2 \qquad (A22)$$

The test is now applied as in the case of test 2A. This test is approximate in that its probability of Type I error is approximately equal to α_2.

Power of test 2A

The power of test 2A to detect a given change in the true values of the variables is obtained as follows.

Let $\underline{Q}_1 = \underline{Q}_2 = \underline{Q}$ be the covariance matrix in each period. Let the vector of differences $\underline{\delta}$ in the true values in the two periods be given by

$$\underline{\delta} = \underline{x}_1 - \underline{x}_2 \qquad (A23)$$

The noncentrality parameter is defined as

$$\tau = N\underline{\delta}'(\underline{Q}_1 + \underline{Q}_2)^{-1}\underline{\delta} = 0.5N\underline{\delta}'(\underline{Q})^{-1}\underline{\delta} \qquad (A24)$$

Under H_1, T^2 is distributed as $pf_A/(f_A - p + 1)$ times a noncentral F variable with noncentrality parameter (Anderson, 1954, pp. 114). Let this be denoted by $T^2_{p,f_A,\tau}$. If the percentage points of the noncentral F distribution are available (Tiku, 1966), then the power of test 2A for any given value of α_2, p, N, and τ is given by

$$Pr\{T^2_{p,f_A,\tau} \geq T^2_{p,f_A}(\alpha_2)\} \qquad (A25)$$

where

$$\frac{(f_A - p + 1)}{pf_A} T^2_{p,f_A,\tau} \sim F_{p,f_A - p + 1,\tau} \qquad (A26)$$

The righthand side being the noncentral F distribution. This procedure was used to obtain the power for test 2A in Figure 2.

Literature Cited

Almasy, G. A., and R. S. H. Mah, "Estimation of Measurement Error Variances from Process Data," *I & EC Proc. Des. Dev.*, **23**, 779 (1984).

Anderson, T. W., *An Introduction to Multivariate Statistical Analysis*, Wiley, New York (1957).

Hale, J. C., and H. L. Sellars, "Historical Data Recording for Process Computers," *Chem. Eng. Prog.*, **77**(1), 38 (1981).

Ito, K., and W. J. Schull, "On the Robustness of the T_0^2 Test in Multivariate Analysis of Variance when Variance-Covariance Matrices are Not Equal," *Biometrika*, **51**, 71 (1964).

Koziol, J. A., "On Assessing Multivariate Normality," *J. Roy. Statist. Soc.*, **B45**(3), 358 (1983).

Mah, R. S. H., "Design and Analysis of Performance Monitoring Systems," *Chemical Process Control II*, D. E. Seborg and T. F. Edgar, eds., Engineering Foundation, New York (1982).

Narasimhan, S., M. S. Thesis, "Detection of Steady States and its Application to Process Data Reconciliation," Northwestern Univ., Evanston, IL (1984).

Seber, G. A. F., *Multivariate Analysis*, Wiley, New York (1984).

Stanley, G. M., and R. S. H. Mah, "Estimation of Flows and Temperatures in Process Networks," *AIChE J.*, **23**(9), 642 (1977).

Tamhane, A. C., and R. S. H. Mah, "Data Reconciliation and Gross Error Detection in Chemical Process Networks," *Technometrics*, **27**(4), 409 (1985).

Tiku, M. L., "A Note on Approximating to the Noncentral F Distribution," *Biometrika*, **53**, 606 (1966).

Yao Ying, "An Approximate Degrees-of-Freedom Solution to the Multivariate Behrens-Fisher Problem," *Biometrika*, **52**, 139 (1965).

Manuscript received May 21, 1985, and revision received Feb. 2, 1986.