Performance Studies of the Measurement Test for Detection of Gross Errors in Process Data

The measurement test proposed by Mah and Tamhane (1982) allows the gross error associated with a measurement to be directly identified without a separate procedure. In this paper a comprehensive evaluation of this test was carried out based on two different definitions of its power. The influence of constraints, network configuration, position of measurement, magnitudes of gross error and standard deviations, number of measurements, and other factors were summarized as rules and guidelines for the application of this test. The simulation procedure developed in this investigation may be used to design a gross error detection scheme for any specific application.

C. IORDACHE, R. S. H. MAH and A. C. TAMHANE

Northwestern University
Evanston, IL 60201

SCOPE

Process measurements are subject to two types of errors: random errors which are commonly assumed to be independently and normally distributed with zero mean, and gross errors which are caused by nonrandom events such as leaks, depositions, and inadequate accounting of departures from steady state operations as well as by measurement biases and malfunctioning instruments. In comparison with the random errors there should normally be a very small number of gross errors present in any given set of data. Nonetheless their presence invalidates the statistical basis of reconciliation procedures which are used to enhance the accuracy of process data burdened with random errors. It is therefore important to detect, identify, and remove gross errors before final data reconciliation.

Three types of statistical tests have been proposed for gross error detection. Test statistics based on the residuals or imbalances of the constraints either individually (normal distribution test) or collectively (chi-square test) have been proposed by Reilly and Carpani (1963), Almasy and Sztano (1975) and Mah et al. (1976). However, in order to identify the sources or locations of the gross errors these tests must be supplemented by an identification algorithm or procedure. This requirement is obviated in the measurement test proposed by Mah and Tamhane (1982).

In this paper we derive certain useful results concerning the partition of information for data reconciliation and gross error detection, and the uniqueness of the test statistics for the measurement test. We study the performance of this test as measured by its power. The power of the test is the probability of correctly detecting and identifying gross errors when they are present in the process data. Mah and Tamhane (1982) gave an upper bound on the power of the measurement test when exactly one gross error is present in the measurements. In this paper we continue to restrict our consideration to the presence of only a single gross error. We explore the influence of different parameters on the power of the test. Our aim is to evaluate the actual performance of the test and to provide guidelines on its applicability. Since the power of the measurement test is difficult to evaluate analytically, it is estimated under different conditions using Monte Carlo simulations.

CONCLUSIONS AND SIGNIFICANCE

The results on the information partition between data reconciliation and gross error detection were discussed without proofs in an earlier paper (Mah and Tamhane, 1982). In this paper we provide a derivation of these results and broaden the treatment to generalized linear reconciliation with missing measurements. In general, V and W, the covariance matrices of residuals, will be rank deficient, and also the test statistics associated with different measurements may be numerically equal, leading to possible ambiguity in the identification of a gross error. The necessary and sufficient conditions for nonidentifiability are established (Theorem 3 and corollary) and a procedure is given for prescribing the modified level of significance used in the test.

A comprehensive evaluation of the measurement test was carried out based on two different definitions of the power of this test. It was found that the effect of constraints, network configuration, and position of measurement could be adequately accounted for by the constraint matrix B. Factors which make the columns of B more proportional tend to reduce the

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

powers associated with these columns (streams) and to make them equal to one another. Factors which make them less proportional tend to improve the powers.

Probably the single most important factor affecting the powers of the test is the numerical ratio of gross error to standard deviation, δ/σ. For problems with unequal measurement standard deviations the powers are also dependent on the range and distribution of σ's. The influence of these factors and of the degrees of adjacent vertices and the number of measured variables on the powers as defined by Eqs. 42 and 43 are summarized by rules which are based on simulation results.

From the viewpoint of enhancing the power of the measurement test in gross error detection our findings suggest that one:

1. Avoid making the columns of the constraint matrix proportional, for instance by incorporating appropriate component material or energy balances.
2. Equalize the standard deviations of measurement errors as much as possible by selectively improving measurements with large standard deviations.
3. Consider that if it is important to detect a gross error associated with a stream connected to nodes with large degrees—i.e., $(d_{ij} + d_{oj})$ is large—then duplicate instrumentation may be necessary.

Finally, through our investigation we have developed a comprehensive simulation procedure for estimating the powers of the test. This procedure may be used to design the gross error detection scheme for any specific application.

GENERALIZED LINEAR RECONCILIATION

Three classes of state variables and parameters are involved in the generalized linear reconciliation. First, there is a class of state variables and unknown parameters x which are related to the measured variables $\tilde{y}$ through a measurement matrix D. In the absence of gross errors this relationship may be expressed as

$$\tilde{y} = y + \epsilon = Dx + \epsilon \tag{1}$$

where ϵ denotes an $(s \times 1)$ vector of measurement errors, which is commonly assumed to be normally distributed with a zero mean vector and a known covariance matrix Q. The measurement matrix D is an $(s \times p)$ matrix of known constants with full column rank $p \leq s$.

In general, there may also be two other classes of state variables and parameters in addition to x. Let c' be a vector of parameters whose values are known precisely, and let u be a vector of parameters which are not directly related to the measurements through Eq. 1. Then the generalized constraints are

$$A_1 x + A_2 u = c' \tag{2}$$

where A_1 is an $(n \times p)$ matrix of known constants and A_2 is an $(n \times m)$ matrix of known constants. The general linear reconciliation problem is the least-squares estimation of x and u,

$$\min_{x,u} (y - \tilde{y})^T Q^{-1} (y - \tilde{y}) \tag{3}$$

subject to the affine constraints, Eq. 2.

Let $(\hat{x}, \hat{u})$ be the solution to this least-squares estimation problem and let P be a $(t \times n)$ matrix such that

$$PA_2 = 0. \tag{4}$$

In other words, the rows of P are orthogonal to the column space of A_2. We shall refer to P as a projection matrix. A reconciliation problem is *unconstrained* if x can be estimated without reference to the constraints, Eq. 2.

Theorem 1. The reconciliation problem defined by Eqs. 1, 2, and 3 is constrained, if and only if A_2 is of less than full row rank.

Proof: Suppose the reconciliation problem is unconstrained. Then for any x whatsoever, which is estimated by Eq. 3, the constraints of Eq. 2 can always be satisfied by a suitable choice of u. That is to say, Eq. 2 can always be rearranged to yield

$$A_2 u = c' - A_1 x = v. \tag{5}$$

Clearly, v lies in the column space of A_2, and yet v can be arbitrarily chosen so long as it remains in R^n. The necessary and sufficient condition for this to be true is that dim (column space of A_2) = rank$(A_2) = n$. Hence, the theorem.

Note that when A_2 is of full row rank, P is a null matrix.

In the subsequent treatment we shall assume that

$$\text{rank}(A_2) = q \tag{6}$$

and

$$t = n - q. \tag{7}$$

Theorem 2. The reconciliation problem defined by Eqs. 1, 2, and 3 may be reduced to a reconciliation problem involving no unmeasured variables. The unmeasured variables may be computed from the estimates of the measured variables in a subsequent step.

Proof: Premultiplying constraint Eq. 2 by the projection maxtrix P, we obtain

$$PA_1 x = Bx = Pc' = c \text{ (say)}. \tag{8}$$

Let x^* be the solution to the reconciliation problem defined by Eqs. 1, 3, and 8, and let u^* be computed by solving the equation,

$$A_2 u = c' - A_1 x^*. \tag{9}$$

Then clearly, (x^*, u^*) is also the solution $(\hat{x}, \hat{u})$ to the reconciliation problem defined by Eqs. 1, 2, and 3.

Theorem 2 is essentially a restatement of the decomposition result obtained by Crowe et al. (1983). In the special case of no unmeasured variables $P = I$, as expected. For the flow reconciliation considered by Mah et al. (1976) P is the first $(N - q)$ rows of the inverse of $[A_{13} A_2]$ in their Eq. A2.

In view of Theorem 2, we need only consider reconciliation with all associated variables measured. Without any loss of generality but with the benefit of much simplification we shall henceforth consider the reconciliation problem defined by Eqs. 1, 3, and the constraints.

$$Bx = c. \tag{10}$$

Analytical solutions for the constrained and unconstrained problems were given as Eqs. 3 and 4 in an earlier paper by Mah and Tamhane (1982), where b, b_o, and A should be replaced by x, x_o, and B in our present notation.

GROSS ERROR DETECTION BY MEASUREMENT TEST

In addition to the random measurement errors the raw process data may also contain gross errors which are caused by nonrandom events. Since gross errors are only defined with respect to the measured variables, we cannot possibly make any statement about gross errors associated with unmeasured variables. We shall therefore assume that a decomposition of the type discussed in the previous section has already been carried out to transform the constraints from Eqs. 2 to 10.

The different statistical tests for gross error detection have been reviewed by Mah (1982) and Crowe et al. (1983). In this paper we shall be concerned only with the measurement test proposed by Mah and Tamhane (1982). The advantage of this test is that the gross error is directly identified with the measurement itself. The assumption here is that there is no gross error in the process model. Examples of model gross errors are leaks from nodes and departures from steady state operations, both of which may be viewed as missing arcs in the process digraph. If we can make this assumption, then the source of a gross error may be identified directly by this test.

Let r be the vector of measurement adjustments or residuals,

$$r = \tilde{y} - \hat{y} = \tilde{y} - D\hat{x}. \tag{11}$$

It can be shown that

$$r = (I - DM)\tilde{y} - DNc, \tag{12}$$

$$E(r) = 0, \tag{13}$$

and

$$\mathrm{cov}(r) = V = (I - DM)Q(I - DM)^T \tag{14}$$

where

$$M = (I - NB)(D^TQ^{-1}D)^{-1}D^TQ^{-1}. \tag{15}$$

and

$$N = (D^TQ^{-1}D)^{-1}B^T[B(D^TQ^{-1}D)^{-1}B^T]^{-1}. \tag{16}$$

Since r is a linear transformation of y which obeys multivariate normal distribution, r is also normally distributed. We can standardize the test statistics by dividing each element of r by its standard deviation.

Tamhane (1982) has shown that for a nondiagonal covariance matrix Q, a vector of test statistics with the maximal power for detecting a single gross error is obtained by premultiplying r by Q^{-1}. Let

$$d = Q^{-1}r \tag{17}$$

Then d is normally distributed with zero mean and covariance matrix W, where

$$W = Q^{-1}VQ^{-1}. \tag{18}$$

The statistics

$$z_i = d_i/\sqrt{w_{ii}} \tag{19}$$

may be used to test for gross error in the ith measurement. It will be concluded that a gross error is present in the ith measurement, if

$$|z_i| > k \tag{20}$$

where k is some critical constant. The choice of k will be discussed in a later section of this paper. This test will be referred to as the measurement test. In the case of a diagonal covariance matrix, the test is unaltered by this premultiplication.

RANK RELATIONSHIP IN ESTIMATION

We shall now return to the general linear reconciliation in order to examine the relationship between data reconciliation and gross error detection. For this purpose we shall refer to the model, Eq. 1, the constraints, Eq. 10, and the objective function, Eq. 3.

For the unconstrained case, since $\hat{y} = D\hat{x}_o$ the vector of adjusted measurements $\hat{y}$ clearly lies in a p-dimensional subspace of R^s. This subspace is spanned by the column vectors of D. The transformed residual vector d, given by Eq. 17, must lie in the $(s - p)$ dimensional subspace orthogonal to D.

When the constraints of Eq. 10 are introduced, further restrictions are imposed. Let us consider first the case of linear (homogeneous) constraints ($c = 0$). Recall that

$$B = PA_1 \tag{21}$$

where A_1 is an $(n \times p)$ matrix of rank at least $(n - q)$ and P is an $(n - q) \times n$ matrix of full row rank, and $n \leq p$. Therefore,

$$\mathrm{rank}(B) = n - q = t. \tag{22}$$

Let us write the constraints, Eq. 10, as

$$B_1x_1 + B_2x_2 = 0 \tag{23}$$

where B_2 is a nonsingular $(t \times t)$ matrix. It follows that

$$x_2 = -B_2^{-1}B_1x_1, \tag{24}$$

$$\hat{y} = D_1\hat{x}_1 + D_2\hat{x}_2 = [D_1 - D_2B_2^{-1}B_1]\hat{x}_1, \tag{25}$$

and

$$\mathrm{rank}(D_1 - D_2B_2^{-1}B_1) = p - t. \tag{26}$$

In other words, $\hat{y}$ now lies in a $(p - t)$ dimensional subspace and the transformed residual vector d now lies in the $(s - p + t)$ dimensional subspace orthogonal to it.

The above discussion pertains to homogeneous constraints. For affine constraints ($c \neq 0$), we can always make the linear coordinate transformation, $x' = x - x_a$ and $y' = y - Dx_a$ where $Bx_a = c$. In terms of the transformed variables, we have $y' = Dx'$ subject to $Bx' = 0$, which is of the same form as before. Therefore, for affine constraints we arrive at the same conclusion that the dimension of the estimation space is $(p - t)$ and the dimension of the residual space is $(s - p + t)$.

The trade-off between the estimation space and the residual space was previously discussed without proof by Mah and Tamhane (1982) and will not be repeated here for the sake of brevity.

The rank of V can be established in the following way. Since we can always transform a linearly constrained estimation problem into an unconstrained problem, we may write for this purpose without loss of generality,

$$M = (D^TQ^{-1}D)^{-1}D^TQ^{-1} \tag{27}$$

and

$$V = [I - D(D^TQ^{-1}D)^{-1}D^TQ^{-1}] \times Q[I - D(D^TQ^{-1}D)^{-1}D^TQ^{-1}]^T \tag{28}$$

with the understanding that D is now an $s \times (p - t)$ matrix. It follows that

$$\mathrm{rank}(V) = \mathrm{rank}[I - D(D^TQ^{-1}D)^{-1}D^TQ^{-1}] \tag{29}$$

Now $[I - D(D^TQ^{-1}D)^{-1}D^TQ^{-1}]$ is an idempotent matrix whose eigenvalues can only be 0's and 1's (Lapidus, 1962, p. 219).

Therefore, its rank must be equal to the sum of its eigenvalues, which is in turn equal to its trace (Lapidus, 1962, p. 212). That is,

$$\begin{aligned}\text{rank}[I - D(D^TQ^{-1}D)^{-1}D^TQ^{-1}] &= tr[I - D(D^TQ^{-1}D)^{-1}D^TQ^{-1}] \\ &= tr[I_s] - tr[D(D^TQ^{-1}D)^{-1}D^TQ^{-1}] \\ &= tr[I_s] - tr[(D^TQ^{-1}D)^{-1}D^TQ^{-1}D] \\ &= tr[I_s] - tr[I_{p-t}] = s - p + t. \quad (30)\end{aligned}$$

Finally, since rank(Q) = s, clearly rank(W) = rank(V).

UNIQUENESS OF TEST STATISTICS

With reference to Eq. 19 it is possible to have the same $|z_i|$ for different measurements, making the measurements indistinguishable for gross error identification using the measurement test. The specific conditions for this occurrence are established in the following theorem for the case $D = I$.

Theorem 3. Let b_i and b_j be two columns of B. Then $|z_i| = |z_j|$ for all $\tilde{y}$ if and only if there exists a nonzero constant h such that $b_i = h\,b_j$.

Proof: Substituting $D = I$ in Eqs. 12, 15, and 16 we have

$$\begin{aligned} r &= QB^T(BQB^T)^{-1}(B\tilde{y} - c) \\ &= QB^T(BQB^T)^{-1}B(\tilde{y} - \hat{y}). \end{aligned} \quad (31)$$

To simplify the notation let us write

$$J = (BQB^T)^{-1}. \quad (32)$$

Then from Eqs. 17 and 19

$$\begin{aligned} z_i &= d_i/\sqrt{w_{ii}} \\ &= \frac{b_i^TJB(\tilde{y} - \hat{y})}{(b_i^TJb_i)^{1/2}}, \quad 1 \le i \le s. \end{aligned} \quad (33)$$

Therefore, the theorem is true if and only if

$$\frac{B^TJb_i}{(b_i^TJb_i)^{1/2}} = \pm \frac{B^TJb_j}{(b_j^TJb_j)^{1/2}} \quad (34)$$

which is true if and only if for some nonzero h

$$B^TJb_i = h\,B^TJb_j$$

or

$$B^TJ(b_i - h\,b_j) = 0. \quad (35)$$

Clearly,

$$b_i = h\,b_j \quad (36)$$

is a sufficient condition.

To show that it is also a necessary condition let us suppose that Eq. 36 is not satisfied. Then Eq. 35 implies that the columns of B^TJ must be linearly dependent. But since rank(B) = t and rank(J) = t, rank(B^TJ) = t. Therefore, B^TJ must be an ($n \times t$) matrix of full column rank. Hence Eq. 36 is also a necessary condition for $|z_i| = |z_j|$.

Corollary. Let a_i and a_j be two columns of A_1. Then if $a_i = h\,a_j$ for some nonzero h, $|z_i| = |z_j|$ for all $\tilde{y}$.

The corollary follows directly from Theorem 3 and definition Eq. 21. The necessary and sufficient condition for $|z_i| = |z_j|$ may be restated as

$$P(a_i - h\,a_j) = 0. \quad (37)$$

Note, however, that although $a_i = h\,a_j$ is a sufficient condition, it is not a necessary condition. Equation 37 will also be satisfied if ($a_i - h\,a_j$) lies in the nullspace of P, i.e., the column space of A_2.

If Q is diagonal, then

$$d_i/\sqrt{w_{ii}} = r_i/\sqrt{v_{ii}}. \quad (38)$$

In that case Theorem 3 holds also for the test statistics obtained from the untransformed residuals r.

It may be interesting to note that one physical situation fulfilling the condition of Theorem 3 arises if we apply the measurement test to two streams linking the same pair of nodes in the process digraph. This could happen as a result of applying the measurement test to a subgraph of the original process digraph. For instance, if we apply the test to node b in Figure 3, then all the other physical nodes are in effect lumped with the environment node. (See Appendix for all the figures and tables.) So far as the test is concerned both streams 2 and 4 link node b to the environment node. In fact, if the subgraph consists of only one node, then all its stream flow rates have the same $|z_i|$, and the measurement test becomes identical with the nodal test.

SIGNIFICANCE LEVEL USED IN MULTIPLE TESTS

It should be clear from Eq. 30 and the rank of W that W will be a full rank matrix if and only if $t = n = p$ and $q = 0$. Except for this case W will be a singular matrix, but we may still have a set of distinct values for $|z_i|$. In our earlier note (Mah and Tamhane, 1982) we pointed out that even in the situation in which a single gross error is suspected but the source of the gross error is unknown, it is appropriate to choose k as the critical point $Z_{\beta/2}$ instead of $Z_{\alpha/2}$ in Eq. 20, where

$$\beta = 1 - (1 - \alpha)^{1/s}. \quad (39)$$

and $Z_{\beta/2}$ and $Z_{\alpha/2}$ denote the upper $\beta/2$ and $\alpha/2$ points of the standard normal distribution, respectively. If the number of distinct values of $|z_i|$ is s' ($\le s$), then a modified level of significance

$$\beta' = 1 - (1 - \alpha)^{1/s'} \quad (40)$$

should be used to give a less conservative test. In general, it is not possible to replace s' in Eq. 40 by a smaller (less conservative) number s'' = rank (W) = $s - p + t$.

The choice $\beta = \{1 - (1 - \alpha)^{1/s}\}$ given by Eq. 39 is based on the following inequality of Sidak (1967): Let $z_i, z_2, \ldots, z_s$ have a joint normal distribution with zero means, unit variances and arbitrary, possibly singular, correlation matrix. Then

$$P\{|z_i| \le k, \ldots, |z_s| \le k\} \ge \prod_{i=1}^{s} P\{|z_i| \le k\}. \quad (41)$$

The assumptions for Sidak's result hold true under H_o: There are not gross errors present. Now we would like to set the probability of type I error less than or equal to α. The lefthand side (LHS) of Eq. 41 is the probability of not committing a type I error. It should, therefore, be set greater than or equal to $1 - \alpha$. By choosing each probability in the righthand side (RHS) equal to $(1 - \alpha)^{1/s}$ we obtain a product which is equal to $1 - \alpha$ and which is also a lower bound on the LHS. This choice gives rise to $k = Z_{\beta/2}$ where $\beta = 1 - (1 - \alpha)^{1/s}$. It is a conservative bound in the sense that P {at least one $|z_i| > k$} $\le \alpha$ under H_o.

If there are linear dependencies among the z_i's, then it is not known, in general, how to sharpen the above inequality except for the special case of $z_i = \pm z_j$. In this case, $|z_i| \le k$ implies $|z_j| \le k$ and therefore we can take the LHS probability in Eq. 41 over only s' ($\le s$) distinct $|z_i|$'s and thus the product on the RHS over only those distinct $|z_i|$'s. This fact is used in our Eq. 40.

Notice that this simplification is not obtained even in the special cases of proportional z_i's and of z_i's related by a linear equation.

For instance, if $z_i = 2z_j$, $|z_i| \le k$ does not imply $|z_j| \le k$. Crowe (1984) has speculated that s' in Eq. 40 should be replaced by rank (W). However, the validity of this procedure is questionable in view of the discussion above.

POWER OF THE MEASUREMENT TEST

The above results turn out to be very helpful in interpreting the computer simulation results in evaluating the power of the measurement test. Depending on the assumption made, the power of the measurement test may be defined in two different ways:

Case A. At most one gross error is known to be present but its location is not known beforehand. If a gross error is present in the ith measurement, then the power of the test for the ith measurement,

$$P_{ai} = P[|z_i| \ge |z_j|, \text{for all } j \ne i \text{ and } |z_i| > k]. \tag{42}$$

Case B. No prior knowledge is available on the number and locations of gross errors. If, in fact, only one gross error is present in the ith measurement, an alternative definition is

$$P_{bi} = P[|z_j| \le k, \text{for all } |z_j| \ne |z|_{\max} \text{ and } |z_i| = |z|_{\max} > k] \tag{43}$$

where

$$|z|_{\max} = \max_{1 \le j \le s} |z_j|. \tag{44}$$

This definition takes into account the cases in which two or more columns of matrix B are proportional, giving rise to two or more z_i's having equal value $|z|_{\max}$ (Theorem 3 and corollary). However, it is tacitly assumed that among the $|z_j|$'s tied for $|z|_{\max}$ and exceeding the critical value k, $|z_i|$ will be correctly identified as the one associated with the gross error. The definition also ignores the fact that if some $|z_j|$ is tied with $|z_i|$ for $|z|_{\max}$, then measurement j may be incorrectly identified as containing a gross error.

SIMULATION EXPERIMENTS

The power of the test in either form P_{ai} or P_{bi} is difficult to evaluate analytically because of the correlations among the z_i's. But estimates of these values may be obtained by computer simulation. In each simulation a measurement vector $\tilde{y}$ is generated according to the following equation:

$$\tilde{y} = y + \epsilon + \delta \tag{45}$$

where y is the vector of true values, ϵ is the vector of random errors, and δ is the vector of gross errors. We only consider δ vectors with one nonzero element $\delta_i > 0$ corresponding to the measurement in which a gross error is present. In our simulations the measured variables are mass flow rates in flow networks. Given a set of "true" values, normally distributed univariate random errors ϵ_i with standard deviations as prescribed for the measurements are generated using subroutine GGNML (FORTRAN 77 IMSL Library). The random errors are generated in a univariate manner, because Q is assumed to be diagonal for the sake of simplicity and clarity.

A simulation run consists of generating N_T realizations of the measurement vector $\tilde{y}$ according to Eq. 45 with a gross error δ placed in a particular measurement and repeating this process for all s measurements. In a given realization the s vectors $\tilde{y}$ corresponding to gross errors placed in s measurements are generated from the same ϵ vector. This is done as a variance reduction technique to yield more precise comparisons between the powers associated with different measurements. For a given $\tilde{y}$ vector corresponding to a gross error in the ith measurement the vector of test statistics $z = (z_1, z_2, \ldots z_s)^T$ is computed according to Eq. 33, where

$$z_j = \frac{b_j^T JB(\tilde{y} - \hat{y})}{(b_j^T J b_j)^{1/2}} = \frac{b_j^T JB(\epsilon + \delta)}{(b_j^T J b_j)^{1/2}}, \quad 1 \le j \le s. \tag{46}$$

Note that we assume further $D = I$ (or $y = x$), which allows us to illustrate the results predicted by Theorem 3 and its corollary.

If N_{ai} is the number of realizations satisfying the condition,

$$|z_i| \ge |z_j| \text{ for all } j \ne i \text{ and } |z_i| > k \tag{47}$$

and N_{bi} is the number of realizations satisfying the condition,

$$|z_j| \le k, \text{for all } |z_j| \ne |z|_{\max} \text{ and } |z_i| = |z|_{\max} > k \tag{48}$$

then the powers P_{ai} and P_{bi} may be estimated as follows:

$$\hat{P}_{ai} = N_{ai}/N_T \tag{49}$$

and

$$\hat{P}_{bi} = N_{bi}/N_T. \tag{50}$$

Since Eq. 48 is a more restrictive condition than Eq. 47,

$$\hat{P}_{ai} \ge \hat{P}_{bi}. \tag{51}$$

Since the outcome of each simulation realization is 0 or 1 and N_{ai} and N_{bi} are sums of these 0–1 variables, $\hat{P}_{ai}$ and $\hat{P}_{bi}$ are sample means of Bernoulli distributed variables. We may construct $(1 - \alpha)$ level two-sided confidence intervals on P_{ai} and P_{bi}, using the following formula (Hines and Montgomery, 1980):

$$\hat{P}_i - E \le P_i \le \hat{P}_i + E \tag{52}$$

where E is the error in the estimate P_i

$$E = Z_{\alpha/2} \sqrt{\hat{P}_i(1 - \hat{P}_i)/N_T}, \tag{53}$$

and P_i may be P_{ai} or P_{bi}. If a confidence level $(1 - \alpha)$ and an upper limit on the estimation error E are specified, Eq. 53 may be used to determine the minimum value of N_T. In our simulation we chose $\alpha = 0.05$ ($Z_{\alpha/2} = 1.96$) and $E \le 0.01$. Then using the fact that E is maximum when $\hat{P}_i = 0.5$, we obtain

$$N_T \ge 9{,}600.$$

On this basis we set the total number of realizations at 10,000. In practice the accuracy obtained is better than the second decimal place indicated by this choice. Three decimal places have been chosen to report all simulation results.

In evaluating the simulation results we need a criterion to determine statistically significant differences. Suppose that for the same run estimated powers $\hat{P}_{ai}$ and $\hat{P}_{ai'}$ are to be compared. For a particular realization let z and z' be the corresponding vectors of test statistics. Since z and z' are based on the same ϵ, they are correlated, and hence $\hat{P}_{ai}$ and $\hat{P}_{ai'}$ are correlated. Therefore we apply McNemar's test (Conover, 1980), which is a technique for comparing correlated proportions. This test is illustrated in the tableau below for comparing the estimated powers of two flow rates in a flow network:

	Stream i'	
Stream i	Reject $H_{oi'}$	Not Reject $H_{oi'}$
Reject H_{oi}	N_A	N_B
Not Reject H_{oi}	N_C	N_D

Estimates of the powers associated with the two streams are given by

$$\hat{P}_i = (N_A + N_B)/N_T; \quad \hat{P}_{i'} = (N_A + N_C)/N_T \tag{54}$$

We reject H_o: $P_i = P_{i'}$ and conclude that P_i and $P_{i'}$ are different if

$$(N_B - N_C)^2/(N_B + N_C) > \chi^2_{\alpha,1} \tag{55}$$

where $\chi^2_{\alpha,1}$ is the upper α point of the χ^2 distribution with 1 degree of freedom. In our computation α is set at 0.05. McNemar's test is applied to all estimates associated wtih the same network. Where appropriate, the results are listed in the last columns of Tables 3–8.

For runs involving different networks, different seeds are used to generate random error vectors ϵ. The test vectors z for different runs are independent. Hence the corresponding power estimates are also independent. The test of the equality of two independent proportions (Hines and Montgomery, 1980) is used in comparing the estimates of powers associated with streams corresponding to different networks. The results of this test are used to compare runs 7.1 and 8.

Since P_{ai} and P_{bi} depend on the distributions of the z_i's, we must examine Eq. 46 to determine factors influencing the power of the test. These factors may be classified as follows:

1. Constraints, network configuration, and stream position. These factors are summarized by the information contained in the constraint matrix B.
2. Magnitude of the gross error.
3. Magnitudes and distribution of standard deviations of measurements, which information is contained in the covariance matrix Q.
4. Number of measurements or size of the network.

These factors are further discussed in the next section.

Many simulation experiments were carried out to discern the influence of these factors. Twelve of these computer runs covering eight different configurations are listed in the Appendix. In these experiments all stream flow rates were "measured," i.e., B is the same as A_1 (A for short). Standard deviations were taken to be the same for all measurements except in Run 7.2 which was designed to study the effect of standard deviations. The critical value k in Eq. 20 was given by $Z_{\beta/2}$ with β computed using Eq. 40 and α = 0.1.

To facilitate cross reference, the first digit of the run number is the same as the number of the figure showing the flow network. The data and results are shown in tables with the same number and labels a and b. For instance, run 1.1 is based on the flow network shown in Figure 1. The run data are shown in Table 1.1a and the simulation results in Table 1.1b.

To expedite the discussion and interpretation, we shall first present an analysis and summary of the influence of the different factors and then refer to the simulation results to illustrate these points.

FACTORS INFLUENCING THE POWER OF THE TEST

For a single gross error $\delta_i = \delta$ in the ith measurement,

$$E(z_j) = w_{ji}\delta/\sqrt{w_{jj}}, \quad 1 \le j \le s \tag{56}$$

where w_{ji} is the (j,i)th element of the matrix

$$W = A^T J A. \tag{57}$$

Equation 57 is a simplification of Eq. 18 for the case $D = I$.

Now suppose that variances are equal for all measurements, then

$$E(z_j) = \frac{a_j^T(AA^T)^{-1}a_i}{[a_j^T(AA^T)^{-1}a_j]^{1/2}}\frac{\delta}{\sigma} \tag{58}$$

In this case the influence of the common standard deviation is decoupled from that related to the constraint matrix A and the gross error δ. The impact of these factors on the power of the test can be summarized by the following rules.

1. *Constraints and network configurations.* The most notable effect of these factors is in the way they affect the proportionality in the columns of the constraint matrix A. If two or more columns of A are proportional, the $|z_i|$'s and the P_{ai}'s and P_{bi}'s are the same for the corresponding streams (Theorem 3 and corollary). Furthermore the powers associated with these streams will be lower than those associated with the nonproportional streams. Constraints and configurations which tend to make the columns of A more proportional tend to reduce the powers associated with the corresponding streams and to make them equal to one another.

2. *Position of a gross error in the network.*

(a) P_{ai} is generally inversely proportional to the total number of streams incident with the nodes adjacent to the stream containing a gross error. In other words,

$$P_{aj} \propto 1/(d_{ij} + d_{oj}) \tag{59}$$

where d_{ij} is the number of streams (regardless of signs) incident with the node for which stream j is an inflow and d_{oj} is the corresponding number for the node for which stream j is an outflow. External streams are considered to be incident with the environment node.

For a stream adjacent to a node of degree 2 (i.e., d_{ij} or d_{oj} is 2), the above rule (a) is replaced by the following rules:

(b) The power associated with such a stream is always greater than that associated with any other stream with the same (d_{ij} + d_{oj}) for which $d_{ij} \neq 2$ and $d_{oj} \neq 2$.

(c) Two such streams adjacent to a common node have the same power regardless of the degrees of their other adjacent nodes.

(d) The power is enhanced if the nodes of degree 2 form a sequence. The longer the sequence, the higher the power P_{ai} associated with the incident streams.

P_{bi} obeys Rules 2c and 2d, but shows considerable deviation from Rules 2a and 2b, especially at high values of δ/σ.

3. *Magnitude of the gross error.* This parameter has a profound effect on the powers. As the ratio δ/σ increases, P_{ai} increases monotonically, but P_{bi} goes through a maximum.

4. *Number of measurements.* At the same level of significance α, the power of the test is affected by the number of distinct $|z_i|$'s, which corresponds approximately to the number of measured variables or the size of the process network. According to Eq. 40, as the number of distinct $|z_i|$'s increases, β decreases, and hence, $Z_{\beta/2}$ increases, making it more difficult to reject H_{oi}. In general, for a stream with a particular value of (d_{ij} + d_{oj}) the larger the network size, the smaller the power associated with that stream.

SIMULATION RESULTS AND DISCUSSION

The flow networks for the various runs are shown in Figures 1–8; data and results for each run are given in the tables accompanying the figures. These are grouped with the Appendix, which lists the key features and purpose of each run.

Same Variance for All Measurements

Runs 1.1, 1.2, 2.1, 2.2 and 2.3 were made to verify the influence of the constraints. To start with, in run 1.1 the columns of the constraint matrix A are not proportional (constant multiples of each other). In run 1.2 columns 2 and 3 of A are made nearly proportional by changing flow rates and the coefficients of the second constraint—a component material balance. As a result, the powers associated with streams 2 and 3 decrease and become almost equal in accordance with rule 1. In run 2.1 A is an incidence matrix. Columns 6 and 7 of A corresponding to two parallel streams connecting the same pair of nodes are identical, and the corresponding powers are also the same in accordance with Theorem 3 and its corollary. In run 2.2 the column proportionality is eliminated by the addition of one more constraint—a component material balance. The powers for streams 6 and 7 are substantially improved as predicted by rule 1. On the other hand, if we keep the coefficients approximately the same as in run 2.2 but change the true flow rates drastically as in run 2.3, the powers remain unchanged.

Runs 3–6, 7.1 and 8 are made to evaluate the influence of the position of a gross error in a process network. In all these runs only mass conservation constraints are used, i.e., A is an incidence matrix. Streams with the same powers as determined by McNemar's test are bracketed together. For instance, streams (1,2,4,5) in run 3 and streams (1,2,3,7) and (4,6) in run 4 have the same powers. They illustrate rules 2a and 2c, whereas rule 2b is illustrated by a comparison of streams 1 and 2 with stream 5 in run 4 and of streams 4 and 5 with stream 7 in run 8. Rules 2c and 2d are illustrated by streams (1,2), (6,7) and (14,15,16) in run 5, and also by streams (21,22) and (8,9,10) in run 6.

Although the rules on the position of a gross error are deduced from empirical observation, they are consistent with the following explanation: It is clear from Eq. 46 that a gross error in the ith position would affect the test statistic z_j in other positions. The number of streams adjacent to stream j as given by $(d_{ij} + d_{oj})$ is a measure of the potential for the gross error to spread. The greater the smearing of the gross error, the less the probability of detecting it. Hence, rule 2a. By contrast, streams incident to a common node of degree 2 must share the same measurement information. If the standard deviations are the same for each measurement, then the power must also be the same (rule 2c). In a sequence of nodes of degree 2, each additional stream contributes another measurement to the common pool of information without any accompanying loss or dissipation, which is the basis for rules 2b and 2d.

The rules on position of a gross error also account for the effect of a network configuration. Runs 7.1 and 8 are based on networks with the same number of nodes and streams but differently connected. The test of equality of independent proportions shows that the powers are significantly different in streams 2, 4, 5, 6 and 8 which have adjacent nodes of different degrees in the two runs. The orientation of the stream has no effect on the power (see, for example, the power associated with stream 3 in the two runs).

We should point out that although the above rules by and large account for the major effects, deviations from these rules do occur in some complex networks, possibly as a result of correlations among the z_i's which are neglected in the above analysis. Notably, in run 5 streams 3, 4 and 12 have the same $(d_{ij} + d_{oj})$ but not the same powers. Again in the same run the power associated with stream 8 is higher than that associated with stream 4 even though $(d_{ij} + d_{oj})$ is lower for the latter stream.

Rule 4 is illustrated by comparing the power associated with the following streams: Streams (4,6) in run 4, streams (1,3,6,8) in run 8 and streams (4,15,16,17) in run 6. In all these cases $\delta/\sigma = 3.5$ and $d_{ij} + d_{oj} = 7$. But the powers are ranked in accordance with rule 4.

The dependency of the power of the test on the ratio δ/σ is given by Eq. 58. This dependency is illustrated in Figure 9 using data taken from Run 4. The actual data points lie exactly on the lines. They are omitted from the plot for clarity. At low values of δ/σ both P_{ai} and P_{bi} increase with increase in δ/σ. But for large δ/σ the smearing effect of the gross error causes more than one set of $|z_i|$'s to exceed the critical value k, which disqualifies the correct gross error detection according to the condition set in Eq. 43. For this reason, while P_{ai} continues to increase P_{bi} goes through a maximum with further increase of δ/σ. The definition (Eq. 43) can of course be modified to allow for such instances to be counted, but at the expense of increasing type I error. One such modified definition is the upper bound on the power of the test previously given by Mah and Tamhane (1982, Eq. 21).

Different Variances for Different Measurements

In the general case in which the variances are different for different measurements, Eq. 46 shows that W depends on the standard deviations as well as on the constraints, network configuration, and stream positions. We would therefore expect the power of the test to depend on all these factors as well as on the ratio δ/σ. In order to highlight the dependency on the standard deviations a configuration is chosen to make the effect of position as uniform as possible. Such a network is shown in Figure 7 for which only two categories of positions are allowed. Only mass conservation constraints were used, i.e., A is an incidence matrix. Runs 7.2.1 and 7.2.2 share the same set of σ_j's arranged in two different orders. Similarly, runs 7.2.3 and 7.2.4 share another set of σ_j's also arranged in two different orders. The ratio $\sigma_{max}/\sigma_{min}$ is larger for the second set than for the first set. The effects of the distribution of σ_j's are highlighted in Table 7.2c.

In general, the range of powers is greater, the larger the ratio $\sigma_{max}/\sigma_{min}$. For instance, in run 7.2.1 $\delta/\sigma = 3.5$, $\sigma_{max}/\sigma_{min} = 4$, and $P_{a,max}/P_{a,min} = 6.86$, while in run 7.2.3 for the same δ/σ, but $\sigma_{max}/\sigma_{min} = 500$, we have $P_{a,max}/P_{a,min} = 167$.

Equation 58 shows that if the σ's are the same for all measurements, $E(z_j)$ is a linear function of δ/σ, but is not otherwise a function of δ. But if σ's are different, then the $E(z_j)$ is a function not only of δ/σ, but also of δ. However, its dependency on δ is not easily predicted analytically. Our results show that in that case for the same δ/σ, the lower the value of δ, the lower the power.

The distribution of σ_j's among the streams also influences the powers. Table 7.2c shows that the powers for streams with the same δ/σ and σ can still be quite different depending on the distribution of other σ_j's. The effect is illustrated by the comparison of the following pairs of streams: Stream 3 for runs 7.2.1 and 7.2.2, stream 4 for runs 7.2.3 and 7.2.4, stream 5 for runs 7.2.1 and 7.2.2, stream 5 for runs 7.2.3 and 7.2.4, and stream 6 for runs 7.2.1 and 7.2.2.

As a comparison, the upper bound on the power of the test was also computed for all the runs, although we did not choose to report the values in order to economize on the space. It was found that the upper bound did not, in general, give a close estimate of the power of the test. It gave a good estimate of P_{ai} only when σ's were the same or almost the same for all measurements.

CLOSING REMARKS

The problem of how to detect and identify gross errors in process data has been tackled by numerous investigators for at least two decades, beginning with Ripps in 1965. The use of different statistical tests has been proposed in published literature since 1972 (Nogita, 1972; Almasy and Sztano, 1975; Mah et al., 1976; Madron et al., 1977; Romagnoli and Stephanopoulos, 1980; Mah and Tamhane, 1982; Crowe et al., 1983). However, without exception no evaluation of these tests has been given. Typically, the authors would apply their proposed test to one or more benchmark problems, one of which is usually Ripps' problem, and claim, implicitly or explicitly, that the method "works." Since the underlying problem is stochastic in nature, a single instance has little significance. A "success" does not mean that the test is good any more than a "failure" implies that it is bad for detecting gross errors.

In this paper the problem of how to evaluate these tests is directly addressed for the first time. This was done in the context of the measurement test. The results of our investigation provide some insight into the influences of various problem parameters on the power of the test for gross errors. Two definitions of power were used.

The power of the test is profoundly affected by the ratio δ/σ. P_{ai} increases with increasing ratio δ/σ, but P_{bi} goes through a maximum. If the standard deviations of measurements are unequal, the powers will also be dependent on the range and distribution of σ_j's. The ratio $\sigma_{max}/\sigma_{min}$ is a measure of the spread of powers. The larger this ratio, the greater the range of powers. For the same δ/σ ratio, the smaller the δ, the lower the power associated with that measurement.

In this paper we have confined our attention to the detection of only one gross error. For scenarios involving the possible presence of more than one gross error, a plausible measure of the performance of the test is the ratio of expected number of gross errors correctly identified in all runs to the total number of gross errors introduced in the same runs.

ACKNOWLEDGMENT

This work was supported by the National Science Foundation Grant CPE 8115161.

NOTATION

A = an $(n \times p)$ matrix of known constants in general and an $(N \times S)$ incidence matrix in particular
A_1 = an $(n \times p)$ coefficient matrix of known constants
A_2 = an $(n \times m)$ coefficient matrix of known constants
a_i, a_j = ith or jth column of A
B = PA_1, a $(t \times p)$ transformed constraint matrix, Eq. 21
b_i, b_j = ith or jth column of B
c, c' = vectors of known constants
d = an $(s \times 1)$ transformed residual vector defined by Eq. 17
d_{ij} = number of streams incident with the node for which stream j is an inflow
d_{oj} = number of streams incident with the node for which stream j is an outflow
D = measurement matrix $(s \times p)$ with full column rank p, defined by Eq. 1
E = error in estimation of the power by an average value
H_{oi} = null hypothesis for the stream i
h = a scalar constant
I = identity matrix
i, i' = subscript for the stream containing a gross error
J = a nonsingular matrix defined by $(BQB^T)^{-1}$
j = general subscript for the streams
k = decision criterion (critical value of the test statistic)
M = matrix defined by Eq. 15
m = number of unmeasured streams or variables
N = number of process nodes
N = matrix defined by Eq. 16
N_A, N_B, N_C, N_D = number of favorable cases explained in McNemar's tableau
N_{ai} = number of realizations satisfying the condition in Eq. 47
N_{bi} = number of realizations satisfying the condition in Eq. 48
N_T = total number of realizations
n = number of rows of matrix A
P = a $(t \times n)$ projection matrix, defined by Eq. 4
P_{ai} = power of the test in the case (a), defined by Eq. 42
P_{bi} = power of the test in the case (b), defined by Eq. 43
$\hat{P}_{ai}$ = estimated value of P_{ai}
$\hat{P}_{bi}$ = estimated value of P_{bi}
p = number of columns of matrix A
Q = an $(s \times s)$ covariance matrix of measurement errors
q = rank of the constraint matrix A_2
r = an $(s \times 1)$ vector of residuals defined by Eq. 11
S = number of arcs or streams
s = number of measured variables
s' = number of distinct values of the test statistic z
s'' = rank (W)
t = $(n - q)$, number of rows of P
u = an $(m \times 1)$ vector of unmeasured variables
V = an $(s \times s)$ covariance matrix of residuals r, defined by Eq. 14
v = a vector defined by Eq. 5
W = an $(s \times s)$ covariance matrix of transformed residuals, defined by Eqs. 18 and 57
w_{ii} = ith diagonal element of W
x = vector of true flow rates or state variables
$\hat{x}$ = a $(p \times 1)$ vector of estimated flow rates or state variables
$\hat{x}_o$ = the unconstrained least-squares estimate of x
y = an $(s \times 1)$ vector of true values of measured variables
$\hat{y}$ = an $(s \times 1)$ vector of estimated values of measured variables
$\tilde{y}$ = an $(s \times 1)$ vector of measured or simulated variables
$Z_{\alpha/2}$ = the upper $\alpha/2$ point of the standard normal distribution function
z_i = test statistic for the ith measurement, based on the transformed residuals, defined in Eq. 19
z_{max} = maximum value among all test statistics z_j

Greek Letters

α = level of significance
β, β' = modified level of significance, defined by Eqs. 39 and 40
δ = an $(s \times 1)$ vector of gross errors
δ_i = gross error in the ith measurement
ϵ = an $(s \times 1)$ vector of generated random measurement errors
σ = an $(s \times 1)$ vector of standard deviations of measurements
σ_i = standard deviation of the ith measurement

Other Symbols

$E(.)$ = expected value of
$P(.)$ = probability of
$tr(.)$ = trace of the matrix
$cov(.)$ = covariance of

APPENDIX: List of Runs

Run	Network shown in Figure	Key Features of Data	Used to Demonstrate Effects of	Remarks
1.1	1	$\sigma_j = 0.25$ for all j	Constraints	
1.2	1	$\sigma_j = 0.25$ for all j; Two almost proportional columns in A	Constraints	Flow rates ratios different from run 1.1
2.1	2	$\sigma_j = 0.25$ for all j; A = incidence matrix; Two identical columns in A	Constraints	
2.2	2	$\sigma_j = 0.25$ for all j	Constraints	One more constraint (component balance) added to the matrix A of run 2.1.
2.3	2	$\sigma_j = 0.25$ for all j	Constraints	Flow rates ratios different from run 2.2
3	3	$\sigma_j = 0.25$ for all j; A = incidence matrix	Position	
4	4	$\sigma_j = 0.25$ for all j; A = incidence matrix	Position, δ/σ	
5	5	$\sigma_j = 0.25$ for all j; A = incidence matrix	Position	
6	6	$\sigma_j = 0.25$ for all j; A = incidence matrix; Two identical columns in A	Position	
7.1	7	$\sigma_j = 0.25$ for all j; A = incidence matrix	Configuration, Position	
7.2.1	7	σ_j's not all same; different values and distribution of σ_j; A = incidence matrix	Values and distribution of σ_j among the streams	Data for σ given in Table 7.2a; Computer results presented in Tables 7.2b-7.2c
7.2.2	7			
7.2.3	7			
7.2.4	7			
8	8	$\sigma_j = 0.25$ for all j; A = incidence matrix	Configuration, Position	Same number of nodes and streams as in Fig. 7 but differently arranged

LITERATURE CITED

Almasy, G.A., and T. Sztano, "Checking and Correction of Measurements on the Basis of Linear System Model," *Problems of Control and Information Theory*, **4**, 57 (1975).

Conover, W. J., *Practical Nonparametric Statistics*, 133, Wiley & Son, New York (1980).

Crowe, C. M., Y. Garcia Campos, and A. Hrymak, "Reconciliation of Process Flow Rates by Matrix Projection. I: The Linear Case," *AIChE J.*, **29**, 881 (1983).

Crowe, C. M., "Reconciliation of Process Flow Rates by Matrix Projection. II: The Nonlinear Case." AIChE Natl. Meet., Anaheim, CA (May, 1984).

Hines, W. W., and D. C. Montgomey, *Probability and Statistics in Engineering and Management Science*, 252, 297, Wiley & Sons, New York (1980).

Lapidus, L., *Digital Computation for Chemical Engineers*, McGraw-Hill, New York (1962).

Madron, F., V. Veverka, and V. Vanecek, "Statistical Analysis of Material Balance of a Chemical Reactor," *AIChE J.*, **23**, 482 (1977).

Mah, R. S. H., "Design and Analysis of Performance Monitoring Systems," *Chemical Process Control II*, D. E. Seborg and T. F. Edgar, Eds., 525, Engineering Foundation, New York (1982).

Mah, R. S. H., G. M. Stanley, and D. M. Downing, "Reconciliation and Rectification of Process Flow and Inventory Data," *I&EC Proc. Des. Dev.*, **15**, 175 (1976).

Mah, R. S. H., and A. C. Tamhane, "Detection of Gross Errors in Process Data," *AIChE J.*, **28**, 828 (1982).

Nogita, S., "Statistical Test and Adjustment of Process Data," *I&EC Proc. Des. Dev.*, **11**, 197 (1972).

Reilly, P. M., and R. E. Carpani, "Application of Statistical Theory of Adjustment to Material Balances," 13th Can. Chem. Eng. Cong., Montreal, Que. (Oct., 1963).

Romagnoli, J. A., and G. Stephanopoulos, "Rectification of Process Measurement Data in the Presence of Gross Errors," *Chem. Eng. Sci.*, **36**, 1,849 (1980).

Sidak, Z., "Rectangular Confidence Regions for the Means of Multivariate Normal Distributions," *J. Amer. Statist. Assoc.*, **62**, 626 (1967).

Tamhane, A. C., "A Note on the Use of Residuals for Detecting an Outlier in Linear Regression," *Biometrika*, **69**(2), 488 (1982).

Manuscript received Jan. 10, 1984; revision received July 10, 1984 and accepted July 23.

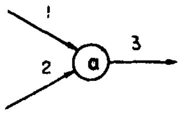

Figure 1. Network for runs 1.1–1.2.

Table 1.1b. Computer Results for Run 1.1

Stream		1	2	3
$\delta/\sigma = 3.5$	$\hat{P}_a$	0.900	0.559	0.361
	$\hat{P}_b$	0.760	0.080	0.078
$\delta/\sigma = 5.0$	$\hat{P}_a$	0.996	0.768	0.615
	$\hat{P}_b$	0.724	0.023	0.064

Table 1.1a. Data for Run Run 1.1

Stream	1	2	3
Matrix	1	1	−1
A	0.7	0.3	−0.4
y	1	3	4

Table 1.2a. Data for Run 1.2

Stream	1	2	3
Matrix	1	1	−1
A	0.7	0.3	−0.316
y	1	24	25

TABLE 1.2b. COMPUTER RESULTS FOR RUN 1.2

Stream		1	2	3
$\delta/\sigma = 3.5$	$\hat{P}_a$	0.918	0.348	0.326
	$\hat{P}_b$	0.880	0.006	0.008
$\delta/\sigma = 5.0$	$\hat{P}_a$	0.998	0.504	0.490
	$\hat{P}_b$	0.947	0.004	0.004

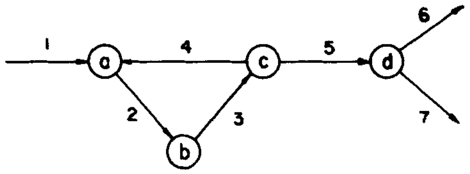

Figure 2. Network for runs 2.1–2.3.

TABLE 2.1a. DATA FOR RUN 2.1

Stream		1	2	3	4	5	6	7
Matrix *A*	*a*	1	−1		1			
(= incidence	*b*		1	−1				
matrix)	*c*			1	−1	−1		
	d					1	−1	−1
y		2	3	3	1	2	1	1

TABLE 2.1b. COMPUTER RESULTS FOR RUN 2.1

Stream		1	2	3	4	5	6	7
$\delta/\sigma = 3.5$	$\hat{P}_a$	0.648	0.596	0.597	0.508	0.656	0.431	0.431
	$\hat{P}_b$	0.420	0.310	0.313	0.292	0.419	0.360	0.360
$\delta/\sigma = 5.0$	$\hat{P}_a$	0.936	0.907	0.903	0.844	0.939	0.788	0.788
	$\hat{P}_b$	0.381	0.242	0.245	0.326	0.380	0.599	0.599

TABLE 2.2a. DATA FOR RUN 2.2

Stream		1	2	3	4	5	6	7
Matrix	*a*	1	−1		1			
A	*b*		1	−1				
	c			1	−1	−1		
	d					1	−1	−1
	d					0.5	−0.05	−0.95
y		2	3	3	1	2	1	1

TABLE 2.2b. COMPUTER RESULTS FOR RUN 2.2

Stream		1	2	3	4	5	6	7
$\delta/\sigma = 3.5$	$\hat{P}_a$	0.632	0.579	0.580	0.492	0.638	0.811	0.807
	$\hat{P}_b$	0.424	0.313	0.317	0.298	0.423	0.694	0.691
$\delta/\sigma = 5.0$	$\hat{P}_a$	0.932	0.900	0.897	0.835	0.936	0.989	0.990
	$\hat{P}_b$	0.419	0.262	0.261	0.346	0.421	0.763	0.765

TABLE 2.3a. DATA FOR RUN 2.3

Stream		1	2	3	4	5	6	7
Matrix	*a*	1	−1		1			
A	*b*		1	−1				
	c			1	−1	−1		
	d					1	−1	−1
	d					0.5	−0.1	−0.9
y		10	100	100	90	10	5	5

TABLE 2.3b. COMPUTER RESULTS FOR RUN 2.3

Stream		1	2	3	4	5	6	7
$\delta/\sigma = 3.5$	$\hat{P}_a$	.632	.579	.580	.492	.638	.811	.807
	$\hat{P}_b$	.424	.313	.317	.298	.423	.694	.691

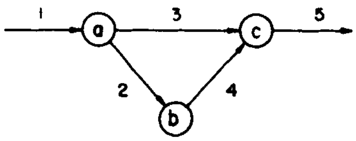

Figure 3. Network for run 3.

Table 3a. Data for Run 3

Stream		1	2	3	4	5
Matrix A	a	1	−1	−1		
(= incidence	b		1		−1	
matrix)	c			1	1	−1
y		7	2	5	2	7

Table 3b. Computer Results for Run 3

Stream		1	2	3	4	5	Power is statistically same for:
$d_{oj} + d_{ij}$		2 + 3	3 + 2	3 + 3	2 + 3	3 + 2	
$\delta/\sigma = 3.5$	$\hat{P}_a$	0.606	0.605	0.496	0.605	0.606	(1,2,4,5)
	$\hat{P}_b$	0.294	0.295	0.255	0.292	0.296	(1,2,4,5)
$\delta/\sigma = 5.0$	$\hat{P}_a$	0.905	0.907	0.824	0.903	0.903	(1,2,4,5)
	$\hat{P}_b$	0.209	0.202	0.267	0.202	0.214	(1,2,4,5)

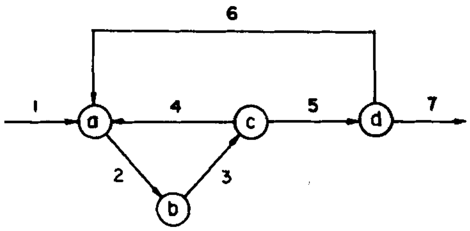

Figure 4. Network for run 4.

Table 4a. Data for Run 4

Stream		1	2	3	4	5	6	7
Matrix A	a	1	−1		1		1	
(= incidence	b		1	−1				
matrix)	c			1	−1	−1		
	d					1	−1	−1
y		1	3	3	1	2	1	1

Table 4b. Computer Results for Run 4

Stream		1	2	3	4	5	6	7	Power is statistically same for:
$d_{oj} + d_{ij}$		2 + 4	4 + 2	2 + 3	3 + 4	3 + 3	3 + 4	3 + 2	
$\delta/\sigma = 2$	$\hat{P}_a$	0.168	0.166	0.172	0.121	0.156	0.117	0.167	(1,2,3,7) and (4,6)
	$\hat{P}_b$	0.119	0.120	0.118	0.081	0.106	0.077	0.115	(1,2,3,7) and (4,6)
$\delta/\sigma = 3.5$	$\hat{P}_a$	0.556	0.563	0.564	0.429	0.524	0.425	0.563	(1,2,3,7) and (4,6)
	$\hat{P}_b$	0.301	0.297	0.299	0.239	0.269	0.230	0.292	(1,2,3,7) and (4,6)
$\delta/\sigma = 5.0$	$\hat{P}_a$	0.884	0.887	0.889	0.777	0.856	0.777	0.887	(1,2,3,7) and (4,6)
	$\hat{P}_b$	0.245	0.245	0.247	0.285	0.238	0.286	0.241	(1,2,3,7) and (4,6)

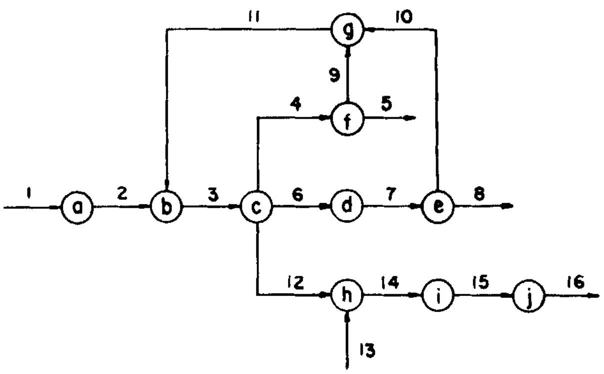

Figure 5. Network for run 5.

TABLE 5a. DATA FOR RUN 5

Stream	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
y	3	3	5	2	1	2	2	1	1	1	2	1	1	2	2	2

TABLE 5b. COMPUTER RESULTS FOR RUN 5, $\delta/\sigma = 3.5$

Stream	1	2	3	4	5	6	7	8	Power is statistically same for:
$d_{oj} + d_{ij}$	5 + 2	2 + 3	3 + 4	4 + 3	3 + 5	4 + 2	2 + 3	3 + 5	
$\hat{P}_a$	0.528	0.532	0.438	0.395	0.407	0.526	0.533	0.442	P_a: (1,2,6,7),(3,8,9),(4,5,13),(9,10,11),
$\hat{P}_b$	0.389	0.394	0.324	0.277	0.286	0.388	0.397	0.320	(10,11,12),(14,15),(15,16)
Stream	9	10	11	12	13	14	15	16	
$d_{oj} + d_{ij}$	3 + 3	3 + 3	3 + 3	4 + 3	5 + 3	3 + 2	2 + 2	2 + 5	
$\hat{P}_a$	0.450	0.461	0.460	0.470	0.401	0.568	0.573	0.585	P_b: (1,2,6,7),(3,8,9),(4,5,13),(9,11),
$\hat{P}_b$	0.325	0.341	0.334	0.334	0.281	0.439	0.446	0.451	(10,11,12),(14,15,16)

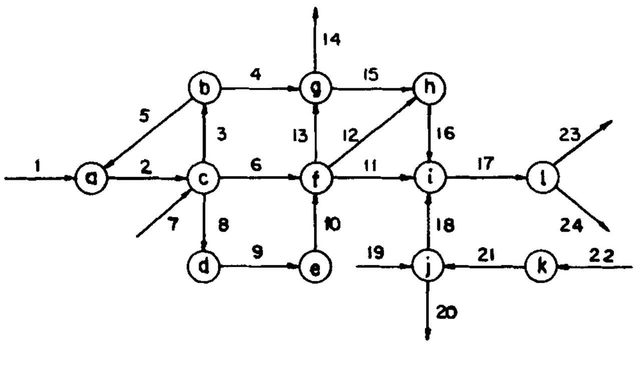

Figure 6. Network for run 6.

TABLE 6a. DATA FOR RUN 6

Stream	1	2	3	4	5	6	7	8	9	10	11	12
y	2	3	2	1	1	1	2	2	2	2	1	1
Stream	13	14	15	16	17	18	19	20	21	22	23	24
y	1	1	1	2	4	1	1	2	2	2	1	3

TABLE 6b. COMPUTER RESULTS FOR RUN 6, $\delta/\sigma = 3.5$

Stream	1	2	3	4	5	6	7	8	Power is statistically same for:
$d_{oj} + d_{ij}$	8 + 3	3 + 5	5 + 3	3 + 4	3 + 3	5 + 5	8 + 5	5 + 2	
$\hat{P}_a$	0.323	0.293	0.291	0.365	0.337	0.269	0.249	0.513	P_a: (1,15,16),(2,3),(4,17),(5,16),(6,11,12,14),(7,13),
$\hat{P}_b$	0.212	0.197	0.204	0.267	0.223	0.200	0.177	0.413	(8,9,10),(17,18),(19,20),(21,22),(23,24)
Stream	9	10	11	12	13	14	15	16	
$d_{oj} + d_{ij}$	2 + 2	2 + 5	5 + 4	5 + 3	5 + 4	4 + 8	4 + 3	3 + 4	
$\hat{P}_a$	0.513	0.523	0.280	0.278	0.253	0.269	0.319	0.333	P_b: (1,3,5,15,16),(2,3,6,11,14,23,24),(4,17,18),
$\hat{P}_b$	0.419	0.420	0.194	0.181	0.175	0.194	0.216	0.223	(7,12,13),(8,9,10),(19,20),(21,22),(23,24)
Stream	17	18	19	20	21	22	23	24	
$d_{oj} + d_{ij}$	4 + 3	4 + 4	8 + 4	4 + 8	2 + 4	8 + 2	3 + 8	3 + 8	
$\hat{P}_a$	0.353	0.350	0.180	0.180	0.378	0.387	0.231	0.231	
$\hat{P}_b$	0.263	0.262	0.148	0.148	0.225	0.230	0.193	0.193	

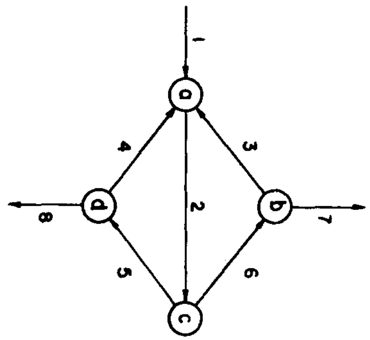

Figure 7. Network for runs 7.1, and 7.2.1–7.2.4.

TABLE 7.1a. DATA FOR RUNS 7.1–7.2

Stream		1	2	3	4	5	6	7	8
Matrix A	a	1	−1	1	1				
(= incidence	b			−1			1	−1	
matrix)	c		1			−1	−1		
	d				−1	1			−1
y		2	4	1	1	2	2	1	1

TABLE 7.1b. COMPUTER RESULTS FOR RUN 7.1

Stream		1	2	3	4	5	6	7	8	Power is statistically same for:
$d_{oj} + d_{ij}$		3 + 4	4 + 3	3 + 4	3 + 4	3 + 3	3 + 3	3 + 3	3 + 3	
$\delta/\sigma = 3.5$	$\hat{P}_a$	0.393	0.399	0.397	0.395	0.466	0.467	0.465	0.471	(1,2,3,4)
	$\hat{P}_b$	0.211	0.218	0.211	0.211	0.247	0.249	0.248	0.251	(5,6,7,8)
$\delta/\sigma = 5.0$	$\hat{P}_a$	0.746	0.748	0.745	0.749	0.818	0.819	0.818	0.819	(1,2,3,4)
	$\hat{P}_b$	0.264	0.269	0.261	0.263	0.265	0.265	0.264	0.259	(5,6,7,8)

TABLE 7.2a. STANDARD DEVIATIONS FOR RUN 7.2

Stream σ-vector	1	2	3	4	5	6	7	8
7.2.1	0.5	0.2857	0.5	0.125	0.2	0.125	0.5	0.2857
7.2.2	0.125	0.5	0.5	0.2857	0.2	0.125	0.2857	0.5
7.2.3	0.01	0.125	0.5	1.0	1.0	2.0	5.0	0.5
7.2.4	0.5	0.01	5.0	1.0	1.0	0.125	0.5	2.0

TABLE 7.2b. COMPUTER RESULTS FOR RUN 7.2

Stream		1	2	3	4	5	6	7	8
Run 7.2.1	σ-vector	0.5	0.2857	0.5	0.125	0.2	0.125	0.5	0.2857
$\delta/\sigma = 2.0$	$\hat{P}_a$	0.158	0.144	0.165	0.030	0.105	0.029	0.165	0.147
	$\hat{P}_b$	0.117	0.068	0.116	0.016	0.062	0.014	0.121	0.071
$\delta/\sigma = 3.5$	$\hat{P}_a$	0.551	0.504	0.560	0.084	0.389	0.087	0.576	0.519
	$\hat{P}_b$	0.325	0.143	0.316	0.039	0.189	0.039	0.329	0.147
$\delta/\sigma = 5.0$	$\hat{P}_a$	0.882	0.796	0.893	0.192	0.741	0.207	0.902	0.810
	$\hat{P}_b$	0.315	0.063	0.283	0.070	0.217	0.080	0.301	0.068
$\delta/\sigma = 8$	$\hat{P}_a$	0.999	0.958	0.998	0.536	0.983	0.574	0.999	0.960
	$\hat{P}_b$	0.018	0.0002	0.012	0.104	0.018	0.107	0.011	0.0001
Run 7.2.2	σ-vector	0.125	0.5	0.5	0.2857	0.2	0.125	0.2857	0.5
$\delta/\sigma = 3.5$	$\hat{P}_a$	0.059	0.694	0.683	0.347	0.220	0.058	0.443	0.736
	$\hat{P}_b$	0.031	0.301	0.321	0.152	0.099	0.035	0.183	0.514
Run 7.2.3	σ-vector	0.01	0.125	0.5	1.0	1.0	2.0	5.0	0.5
$\delta/\sigma = 3.5$	$\hat{P}_a$	0.005	0.014	0.121	0.684	0.545	0.799	0.836	0.134
	$\hat{P}_b$	0.001	0.006	0.014	0.314	0.154	0.640	0.771	0.063
Run 7.2.4	σ-vector	0.5	0.01	5.0	1.0	1.0	0.125	0.5	2.0
$\delta/\sigma = 3.5$	$\hat{P}_a$	0.083	0.007	0.836	0.497	0.746	0.013	0.091	0.792
	$\hat{P}_b$	0.015	0.002	0.772	0.129	0.249	0.002	0.018	0.639

TABLE 7.2c. COMPARATIVE RESULTS FOR RUN 7.2

Stream		3			4		5				6			7	
Run		7.2.1	7.2.2	7.2.3	7.2.3	7.2.4	7.2.1	7.2.2	7.2.3	7.2.4	7.2.1	7.2.2	7.2.4	7.2.1	7.2.4
σ_i		0.5	0.5	0.5	1.0	1.0	0.2	0.2	1.0	1.0	0.125	0.125	0.125	0.5	0.5
$\delta/\sigma = 3.5$	$\hat{P}_a$	0.560	0.683	0.121	0.684	0.497	0.389	0.220	0.545	0.746	0.087	0.058	0.013	0.576	0.091
	$\hat{P}_b$	0.316	0.321	0.014	0.314	0.129	0.189	0.099	0.154	0.249	0.039	0.035	0.002	0.329	0.018

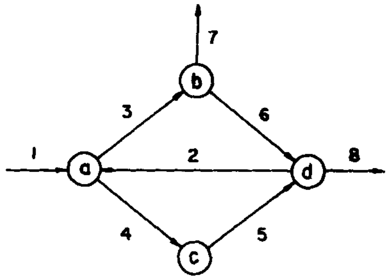

Figure 8. Network for run 8.

TABLE 8a. DATA FOR RUN 8

Stream		1	2	3	4	5	6	7	8
Matrix A	a	1	1	−1	−1				
(= incidence	b			1			−1	−1	
matrix)	c				1	−1			
	d		−1			1	1		−1
y		2	1	2	1	1	1	1	1

TABLE 8b. COMPUTER RESULTS FOR RUN 8

Stream $d_{oj} + d_{ij}$		1 3 + 4	2 4 + 4	3 4 + 3	4 4 + 2	5 2 + 4	6 3 + 4	7 3 + 3	8 4 + 3	Power is statistically same for:
$\delta/\sigma = 3.5$	$\hat{P}_a$	0.401	0.332	0.404	0.520	0.527	0.400	0.422	0.396	(1,3,6,8)
	$\hat{P}_b$	0.202	0.178	0.205	0.271	0.273	0.203	0.212	0.197	and (4,5)

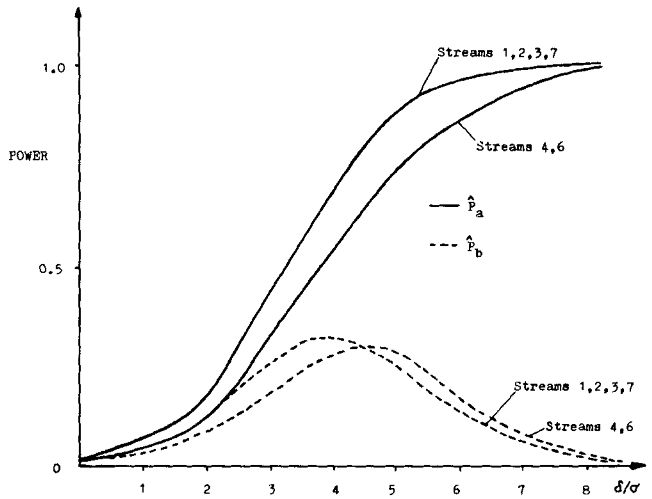

Figure 9. Dependence of the power of the test on the ratio δ/σ for run 4.