STATISTICAL ANALYSIS AND ADJUSTMENT OF THE MASS/ENERGY BALANCES OF PROCESSES.

BY

© YURAIMA ALICIA GARCIA CAMPOS, B.ENG.

A thesis

Submitted to the school of Graduate Studies

in partial fulfilment of the requirements

for the Degree

Master of Engineering

McMaster University

June 1980

STATISTICAL ANALYSIS AND ADJUSTMENT OF THE MASS/ENERGY BALANCES OF PROCESSES.

MASTER OF ENGINEERING — MCMASTER UNIVERSITY

(Chemical Engineering) — Hamilton, Ontario

TITLE: Statistical Analysis and Adjustment of the Mass/Energy Balances of Processes.

AUTHOR: Yuraima Alicia Garcia Campos, B.Eng. (Universidad de Oriente, Venezuela).

SUPERVISOR: Professor C.M. Crowe.

NUMBER OF PAGES: xiii, 140

ABSTRACT

An improved technique for data adjustment to satisfy mass/energy balances for single or multiple processing unit systems has been developed. The method was successfully tested on examples from the literature. With this technique, biased data can be detected and sources of error identified and isolated.

A multivariate normal distribution for the measurement vector $\underline{x}$, was assumed and used to derive a test function vector $\underline{w}$, with univariate normal distribution. A chi-square test on $\underline{w}$ is proposed for detecting biased data. An algorithm for identifying and isolating sources of bias was developed, using standard normal two-sided tests on each element of vector $\underline{w}$ and using the imbalance of the constraints, $\underline{e}$.

In the absence of bias, adjusted values of the measurement vector and/or the estimates of missing variables can be obtained. When chemical reactions take place, the extents of the independent reactions are also estimated.

The main requirements of this technique are the knowledge of the process constraints and measurement statistics.

Implementation of the technique is straightforward and it is suggested that the method be used as a diagnostic aid in process analysis.

ACKNOWLEDGEMENTS

For his guidance and invaluable advice throughout the course of this study, the author is indebted to her supervisor Dr. C.M. Crowe, and, for his assistance in writing this manuscript and his helpful discussion, to her friend, fellow graduate student, Mr. R.C. McFarlane.

Thanks are also due to Fundacion Gran Mariscal de Ayacucho for their financial support; and to Mrs. Amy Stott for her typing.

She also would like to thank her family for their encouragement and moral support throughout the course of this work.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOTATION AND SYMBOLS

All the capital Latin letters (B,C, ...) are reserved for matrices, the lower Latin letters ($\underline{a}$, $\underline{u}$, $\underline{v}$, $\underline{w}$, $\underline{x}$, $\underline{y}$, $\underline{z}$, ...) for column vectors and the letters i, j, k for integer indices. The elements of a matrix or a vector are denoted with the same letter with the appropriate subscripts (b_{ij} or x_j ...). The Greek letters denote a special matrix (ie, eigenvector matrix), a special function (ie., mass balance) or a special number (ie, level of significanoe). Primes denote transposed vectors or matrices and unprimed vectors are column vectors unless specified otherwise. The superscripted letter e has a particular meaning. $\underline{e}^i$ is a vector with its ith element equal to one and with all other elements equal to zero

SYMBOLS

Roman Letters.

A	Partitioned matrix $[B_2 \mid S']$
B_1	n x m coefficient balance matrix for measured variables
B_2	n x ms coefficient balance matrix for missing variables
C	Partitioned matrix defined by 3.5
E	Atom matrix
H	n x n matrix defined by $B_1 \Sigma_x B_1'$
I	Identity matrix

Q	p x n matrix defined by $SH^{-1}(I_n - B_2R)$
0	Matrix or vector of elements equal zero
R	ms x n matrix defined by $(B_2' H^{-1} B_2) B_2' H^{-1}$
S	p x n stoichiometric coefficient matrix
Z	Standard normal variable
$\underline{a}$	Adjustment vector
$\underline{e}$	Imbalance vector defined by 3.30
$\underline{e}^i$	Special vector defined above
$\underline{\ell}$	Vector of Lagrange multipliers
m	Number of measurements
ms	Number of missing variables
n	Number of mass/energy balances
$^9\underline{0}$	Vector of elements equal zero
p	Number of independent chemical reactions.
r	Rank of the covariance matrix of vector $\underline{e}$
$\underline{u}$	Missing variable vector
$\underline{w}$	Test function vector defined by 3.40
$\underline{x}$	Measurement vector
$\underline{y}$	Vector of the adjusted values of $\underline{x}$
$\underline{z}$	Partitioned vector $\begin{bmatrix} \underline{u} \\ \hline \underline{\xi} \end{bmatrix}$

Greek Letters

α	Level of significance, type I error
β	Type II error

$\underline{\gamma}$	Vector of bias
$\underline{\varepsilon}$	Vector of measurement errors
$\underline{\theta}$	Fictitious stream vector for leak, etc ...
λ_i	The ith eigenvalue of Σ
$\underline{\mu}_\bullet$	Vector of true values of $\bullet$
$\underline{\xi}$	Vector of effective extents of individual reactions
σ_i^2	Variance of the ith variable
Λ	Diagonal matrix of eigenvalues of Σ
π_i	Population i
Σ	Covariance matrix of the test function
$\Sigma_\bullet$	Covariance matrix of $\bullet$
Γ	Matrix of eigenvectors of Σ
Γ_1	Matrix of eigenvectors corresponding $\lambda_i \neq 0$
Γ_2	Matrix of eigenvectors corresponding $\lambda_i = 0$

Mathematical Symbols

$E(\bullet)$	Expected value of $\bullet$
$\exp[\bullet]$	Exponential of $\bullet$
$F(\underline{a}, \underline{z}, \ell)$	Lagrange Function
$f(\bullet)$	Probability density function of $\bullet$
$P(\bullet)$	Probability distribution of $\bullet$
$Nr(\underline{\mu}, \Sigma)$	r-variate normally distributed with mean $\underline{\mu}$ and covariance matrix Σ
$\psi(\ )$	Mass/energy balances
$\rho(\bullet)$	Rank of $\bullet$

$\chi^2(1 - \alpha, r)$	Chi-square statistic with r degree of freedom and α significance level
$\| \|$	Determinant
$[\ ]$	Matrix
$'$	Transpose
$\forall$	For all
$\sum_{i=1}^{k}(\cdot)$	Sum of $\cdot$ from 1 to k
$\in$	Belong to
$\notin$	Do not belong to
dim $\{\cdot\}$	Number of independent vector of $\cdot$

CHAPTER I

INTRODUCTION

Process data generally do not satisfy material and/or energy balance equations due to the existence of errors in the data. These errors are referred to as measurement errors and can be classified into two groups: small and gross errors, depending on the source. Those arising from inaccurate reading of instrument responses or normal errors in analytical techniques belong to the group of small errors. Those arising from improper design or installation of instruments, improper calibration of instruments, incomplete information about the process (ie, unsteady state, leaks, etc.), or human errors in reading or reporting data are classified as gross errors.

The presence of either type of error can lead to inaccurate estimation of process parameters and can interfere with the development of the process model. Therefore, it is essential to analyse statistically and compensate for errors in the data to allow for a more precise application of mass and energy models. Data adjustment techniques have evolved in recent years which attempt to solve this problem. Using these techniques small errors can be removed by small adjustment of the data, so that they satisfy the material and/or energy balances. Difficulties arise in applying data adjustment techniques when: (a) insufficient measurements are available and (b) measurements contain gross errors. Insufficient measurement data pose a problem

because these data must be estimated simultaneously to the data adjustment. This problem has been discussed and solved by some workers. Gross errors cause large inconsistencies in the mass/energy balances and require large adjustment. Adjustment techniques compensate for gross errors by distributing the errors over the entire data set. Clearly this is undesirable as it reduces the quality of all measurements. A better approach is to identify and remove the measurements containing gross errors before proceeding to the adjustment stage.

Previous workers (Mah et al (1975), Nogita (1972)) have developed statistical techniques for detecting the presence of gross errors in data. These techniques are based on examining the calculated mass/energy balance or adjustment using the raw data. Nogita identifies the specific measurements in gross error by applying the method of Ripps (1965). Ripps' procedure, however, is cumbersome in its location of the corrupted measurement. The algorithm developed by Mah et al is based on the imbalances around each unit, but it is restrictive in its ability to handle general process data (eg., single unit).

In this study, we have proposed an improved method for the statistical analysis of inconsistency of mass and/or energy balances. The main objective is to detect more efficiently the gross errors in the data. Once identified, the prospected measurements containing gross errors may be removed from the data set. The method proposed in this thesis simultaneously estimates the missing and removed data and adjusts the remaining data to satisfy the mass/energy balances. The main limitation of this method is the fulfilment of the requirements such as,

the number of missing/removed measurements must be less than the number of balance equations and the number of process variables must be greater than the number of balance equations.

This thesis is divided into six chapters which closely follow the development of this study. A review of the past work which has been done on the problem of data adjustment is presented in Chapter II. In addition, preliminary details of the problem are presented to assist the reader in the literature survey. Chapter III describes the mathematical derivation of a general framework to deal with adjustment of data containing gross error and missing variables. A brief resume of statistical fundamentals required for this development is also presented. The computational procedure for computer implementation is described in Chapter IV and is tested on some data from the literature in Chapter V. Finally, in Chapter VI the significance and conclusions of this work are discussed.

CHAPTER II

REVIEW OF PAST WORK

2.1 INTRODUCTION

Data adjustment deals with the problem of process data which are inconsistent due to measurement errors. This chapter presents a review of the literature on this problem of adjusting inconsistent process data. In addition, as a convenience to the reader, we define the problem in general terms in the next section, before the literature review is discussed.

2.2 STATEMENT OF THE GENERAL PROBLEM

In this section the general problem of adjustment of data with missing variables and small errors is considered. The method of solution to this problem will be developed in chapter III.

We shall begin by defining some terms which are used throughout this thesis. A measurement is defined as the numerical value of the process variable obtained either by analytical or process instrumentation methods. A missing variable is one which is required for the mass or energy balances but the measurement is not available. Undeterminable variables arise from situation such as the one shown in Fig. 2.1. If measurements of specific components(or entire streams) are missing in both streams 8 and 9 then these measurements would be undeterminable and for statistical analysis the two streams would have to be considered as one. Then, undeterminable variables can be defined

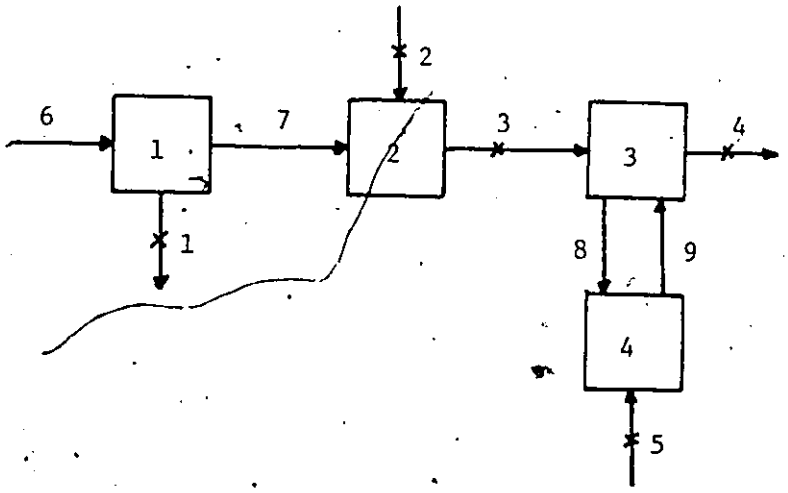


FIGURE 2.1 Undeterminable Variables.

*Measurement

→Missing Variables

as missing variables which can not be estimated from the information available. Consistent data are defined as data which satisfy exactly the mass/energy balances. Biased data are data which contain one or more measurements having gross error.

A process consists of a set of physical and/or chemical processing units which are considered to be systems. The mathematical model of these systems and the interrelationships (mass/energy balance) existing between the process variables of these systems are the usual way of expressing the behaviour of the overall systems. When fitting the model, the parameter estimation should be made using consistent process data. As was asserted before, the process data are usually inconsistent

due to the presence of measurement errors, therefore it is first necessary to adjust the data to satisfy the mass/energy balances. Below, this problem of adjustment of inconsistent process data is defined mathematically.

In general, consider a system of chemical processing units involving n components entering and leaving the system. The steady-state mass/energy balance of the ith component can be represented by:

$$\text{Input-output} + \text{Generation by reaction} = 0 \qquad (2.1)$$

The first two terms of equation 2.1 for all components can be written in matrix form as:

$$\text{Input-output} = B\,\underline{y} \qquad (2.2)$$

where B is the coefficient matrix for mass/energy balance constraints and $\underline{y}$ is the process variable vector (eg., flowrate) whose values are consistent. Any composition change in the system from inlet to outlet due to chemical reactions, may be described by the product of vector $\underline{\xi}$ of effective extents of the independent reactions and the stoichiometric coefficient matrix S, in which the rows represent the chemical reactions and the columns the chemical components taking part in the system. In the energy balances, this matrix represents the heat of reaction. Then, equation 2.1 may be written (Crowe (1979)) as:

$$B\underline{y} + S'\underline{\varepsilon} = 0 \qquad (2.3)$$

Consider the m-vector, $\underline{x}$, of measurement data of flow, material or energy (eg., molar flowrate), the elements of which can contain either gross or small error. The relation between this vector of measured values of process variables and its vector of true values, $\underline{\mu}$, is

$$\underline{x} = \underline{\mu} + \underline{\varepsilon} \qquad (2.4)$$

where $\underline{\varepsilon}$ is the vector of measurement errors.

An adjustment vector, $\underline{a}$, is applied to the measurement vector, $\underline{x}$, in order to satisfy the steady-state mass/energy balance equations expressed by equation 2.3. Then,

$$\underline{y} = \underline{x} - \underline{a} \qquad (2.5)$$

where $\underline{y}$ is referred to as the vector of the values of $\underline{x}$ adjusted by $\underline{a}$.

More commonly not all process variables are measured. Define the missing variables as the ms - vector $\underline{u}$. It is desirable to estimate the vector $\underline{u}$ as well as to adjust the measurement vector $\underline{x}$. The balance coefficient matrix B may be partitioned into two balance coefficient matrices: n x m matrix B_1 corresponding to the measured variables and n x ms matrix B_2 corresponding to the missing variables. Then, equation 2.3 together with 2.5 becomes:

$$\underline{\psi}\,(\underline{a},\underline{u},\underline{\xi}) = B_1(\underline{x} - \underline{a}) + B_2\,\underline{u} + S'\,\underline{\xi} = 0 \qquad (2.6)$$

In general, the problem of data adjustment with missing variables can be described as the optimal estimation of vectors $\underline{a}$, $\underline{u}$ and $\underline{\xi}$ which will satisfy equation 2.6. As stated previously, if the measurement vector $\underline{x}$ contains gross errors, then the adjustment technique will spread these errors over the entire vector $\underline{a}$, reducing the quality of the entire adjusted vector $\underline{y}$. As this is undesirable, it is essential that some method be available for detecting gross errors in the data set before the adjustment. If gross error is detected, the corrupted measurement(s) can be removed from the data and the resulting missing variable(s) included in vector $\underline{u}$ for estimation.

The following section presents a survey of work on the problem of data adjustment which has appeared in the literature. It will be shown that an efficient method for detection of gross error has not been developed to date. Furthermore, most workers have ignored the problem of gross error with missing variables and no one has considered chemical reaction as part of the overall problem.

2.3 LITERATURE SURVEY

A chronological review of the literature pertaining to data adjustment is presented in the following paragraphs. The subject has received little attention in the literature.

Kuehn and Davidson (1961) were among the first who attempted to adjust data containing only small errors, where all the measurements were available. They solved the problem by the Least Squares method

with linear and nonlinear constraints.

Ripps (1965) expanded the above procedure to allow for both types of errors, modifying the Least Squares method to allow a certain number of variables to be discarded as gross errors. Linear constraints only were considered. The idea is to calculate the value of the objective function when successive measurements are removed from the data. The greatest reduction in the objective function corresponds to the measurement(s) which is (are) in gross error. If the analyst has no prior knowledge of which measurements might be in gross error, then the entire data set must be examined. In this case the method is time-consuming.

Vaclavek (1968, 1969) used network characteristic theory to solve the data adjustment problem with small errors. He also considered the case where not all streams are measured; a situation quite common in practice. This technique classifies the data into measured and missing stream variables and through this classification reduces the size of the Least Squares problem. The detection of gross errors is not considered.

Meharg (1972) studied data adjustment with non-linear constraints. His proposed algorithm is basically the same as Kuehn and Davidson's procedure with the constraints linearized by Taylor series. Data containing only small errors were considered and missing variables were not considered.

Nogita (1972) reconsidered data adjustment with gross errors and attempted to detect the bias based on statistical considerations. His contribution was the development of a statistical test criterion of data

inconsistency and statistical data adjustment using a serial elimination algorithm. This algorithm proceeds as follows. (a) Evaluate the test function with the complete set of data. If the value of the test falls within the 90% confidence interval taken from the normal distribution, then the data are considered to be free of gross error. (b) If test (a) fails, remove one element from the measurement vector. (c) Apply Ripps' method to adjust the remaining data and estimate the discarded measurement. (d) Evaluate and record the test function with the adjusted data only. (e) Retrieve the element and remove a new one. (f) Iterate from (a) to (e) over the complete set of data. (g) Eliminate the element which gives the smallest value of the normalized deviation of the test function. (h) Iterate from (a) to (g). The consistent set of data corresponds to the minimum value of the normalized deviation of the test function.

Since this test function takes a linear combination of all adjustments, gross errors can pass unnoticed because of error cancellation. This method is time-consuming because of the need to re-estimate all the parameters after each deletion. The problem of missing variables is not explicitly considered.

Mah et al (1976) further developed the approach proposed by Vaclavek (1969) for mass flow networks. They proposed a technique based on the process flow diagram, in which the data adjustment with missing variables was separated into two distinct problems; (a) data adjustment on a subset of units from the total process, where all streams are measured (Reconciliation) and (b) estimation of missing streams in the

remaining subset of units if they form an open cycle (coaptation). In a closed cycle, missing streams are undeterminable. A gross error detection criterion was developed and used in the identification of the source of gross error. This detection criterion is based on a separate statistic associated with each unit of the reconciliation diagram. Gross error can not pass unnoticed because of error cancellation and the detection of gross error is made before any adjustment calculations are performed. Mah's method, however, is restricted to the adjustment of total mass flow. It can not be applied to any single unit unless a value for a total mass flowrate of each component is available.

The problem of adjustment with chemical reaction has been studied as a separate problem, where stream mass/energy balances are not considered. Vaclavek (1969) took account of the chemical reactions by adding fictitious streams. Murthy (1973, 1974) and Madron et al (1977) described the chemical reaction in terms of the vector of effective extents of reaction and stoichiometric coefficient matrix. In addition, Madron, Veverka and Vanecek (1977) proposed a single statistic chi-square test for bias.

The object of this study was to develop a method for adjusting inconsistent process data, which recognizes all the following needs; (1) efficient detection of gross error by statistical criterion; (2) data adjustment where variables are missing; (3) chemical reaction considered as an integral part of the process and (4) applicable to any single unit or to any number of units. Crowe, (1979) has suggested a data adjustment method which includes chemical reactions and missing

variables.

The following chapter describes the mathematical development of the solution proposed for this complete problem of data adjustment.

CHAPTER III

MATHEMATICAL DEVELOPMENT OF THE TECHNIQUE OF DATA ADJUSTMENT.

3.1 INTRODUCTION;

The main objective of this chapter is to present the mathematical development of the data adjustment technique, considering missing variables and gross errors in the data. Before proceeding, it is convenient to give an outline of the statistical foundations of the mathematical formulation of this study: Multivariate normal distributions and hypothesis testing.

3.2 STATISTICAL FUNDAMENTALS

Knowledge or intuition about a process can lead the process analyst to suspect that a measurement contains a gross error. In such a case the measurement should be discarded or checked. However, in most cases he will be aware of the likely presence of gross error only to the extent that he is faced with inconsistent process calculations. In these cases a judgement about which measurement may be in gross error can only be effected by the use of statistical analysis.

Nonsingular and singular multivariate normal distribution, hypothesis testing, chi-square tests and two-sided normal tests comprise the statistical theory used in the analysis developed below. These topics are discussed in the following subsections.

3.2.1 Multivariate Normal Distribution.

Statistical regression theory requires the assumption of normality of errors in the data. Since the adjustment technique developed in this thesis is based on multivariate regression theory, it is worthwhile to examine the properties of the multivariate normal distribution.

The multivariate normal probability density function of an m-random vector, $\underline{t}$, with mean $\underline{\mu}$ and positive definite covariance matrix, Σ, is defined by

$$f(t) = \frac{1}{(2\pi)^{r/2}|\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(\underline{t}-\underline{\mu})' \Sigma^{-1} (\underline{t}-\underline{\mu})\right] \qquad (3.1)$$

It is developed by analogy with the normal density function.

The notation $f(\underline{t}/\underline{\mu},\Sigma)$ is used to denote equation 3.1 given $\underline{\mu}$ and Σ; and $Nm(\underline{\mu},\Sigma)$ is used to denote the distribution law.

Any linear combination of normal variables is also normally distributed. In other words, if an m-random vector, $\underline{t}$, is distributed according to $Nm(\underline{\mu},\Sigma)$, any linear combination of $\underline{t}$, such as,

$$\underline{t}_* = D\underline{t} \qquad (3.2)$$

is distributed according to $Nn(D\underline{\mu}, D\Sigma D')$, with D being an nxm matrix of rank $n \leq m$.

If $\underline{t}$ has a singular covariance matrix, the distribution of $\underline{t}$ is called a singular or degenerate normal distribution; and the density function 3.1 does not exist. The following subsection is devoted to this type of distribution.

3.2.2 Singular Normal Distribution.

A singular distribution for a vector $\underline{t}$ is a distribution in m-dimensional space which is restricted to a lower dimensional subspace, $\bar{M}$. This means that the probability that $\underline{t}$ has a component in the complementary orthogonal subspace, $\bar{M}^{\perp}$ (excluding the origin), is zero. Therefore, $\underline{t}$ lies in $\bar{M}$. In the case of a singular normal distribution, all of its probability density lies in a linear subspace of m-dimensional space, (for example m=2, then all the probability density lies on the line $t_2 = at_1 + b$ for some $a \neq 0$ and some b). For more details see Anderson (1958).

If an n-random vector $\underline{t}$ is singular normally distributed with mean $\underline{\mu}$ and covariance matrix Σ of rank r, then there is a transformation (except for zero probabilities):

$$\underline{t} = F\underline{w} + \underline{n} \tag{3.3}$$

where F is n x r matrix, (n > r), and the r-vector $\underline{w}$ has a non-singular multivariate normal distribution, $Nr(\nu,T)$, that is

$$f(\underline{w}) \propto \exp[-\frac{1}{2} (\underline{w} - \underline{\nu})' T^{-1} (\underline{w} - \underline{\nu})] \tag{3.4}$$

and it can be said that $\underline{t}$ has a singular normal distribution in n-space, with $E(t) = F(E(\underline{w})) + \underline{n}$ and $\Sigma = F\ T\ F'$

Since Σ is of rank r, there is an $n \times n$ non-singular matrix C such that

$$C \Sigma C' = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] \tag{3.5}$$

In Appendix III it is shown that

$$C = \left[\begin{array}{c} \Lambda_r^{-1/2} \Gamma_1' \\ \hline \Gamma_2' \end{array}\right]$$

where Λ_r is the diagonal matrix of r-nonzero eigenvalues of matrix Σ. The columns of Γ_1 are the normalized eigenvectors corresponding to non-zero eigenvalues and those of Γ_2 are those corresponding to zero eigenvalues.

Then the transformation:

$$C\underline{t} = \underline{g} = \left[\begin{array}{c} \underline{g}_1 \\ \hline \underline{g}_2 \end{array}\right] \tag{3.6}$$

defines a random vector $\underline{g}$ with covariance matrix given by equation 3.5 and mean vector $\underline{\nu}$ given by

$$E(\underline{g}) = C\,\underline{\mu} = \underline{\nu} = \left[\begin{array}{c} \underline{\nu}_1 \\ \hline \underline{\nu}_2 \end{array}\right] \tag{3.7}$$

Since the covariance matrix $\underline{g}_2$ is equal to zero,

$$\underline{g}_2 = \underline{\nu}_2 \tag{3.8}$$

with probability 1. From equation 3.6, $\underline{t}$ can be expressed as

$$\underline{t} = C^{-1}\underline{g} \qquad (3.9)$$

where $C^{-1} = D = [D_1 \vdots D_2]$

By solving the partitioned matrix product $C^{-1}\underline{g}$, equation 3.9 becomes:

$$\underline{t} = D_1\underline{g}_1 + D_2\underline{\nu}_2 \qquad (3.10)$$

and comparing the above equation with equation 3.3, it is found the following set of relations:

$$\begin{aligned} F &= D_1 \\ \underline{w} &= g_1 \\ \underline{n} &= D_2\,\underline{\nu}_2 \end{aligned} \qquad (3.11)$$

In general, an m-random vector $\underline{t}$ with mean $\underline{\mu}$ and covariance matrix Σ of rank r is normally distributed if there is a transformation (3.3), where F is n x r matrix, (n > r), and $\underline{w}$ has a non-singular normal distribution, $Nr(\underline{\nu},T)$.

3.2.3 Hypothesis Testing.

In the usual statistical test, a hypothesis is made about the

population from which the sample is taken. The decision to reject or accept the hypothesis is based on a chosen statistic which is calculated from the sample. The hypothesis is rejected with a known probability of making a false rejection (type I error). The probability is established by the significance level of the test. If the hypothesis is accepted, it is usually on the ground of having insufficient evidence for its rejection, and not with established probability of a false acceptance (type II error). These concepts of type I error, type II error and level of significance (α) are summarized in table 3.1, where the columns are the truth or falsity of the hypothesis and the rows are decision made. The level of significance (α) is the probability of rejecting a true hypothesis and β is the probability of accepting a false hypothesis. Table 3.2 shows the relation between α and β in the decision table.

TABLE 3.1

Types of Error in Hypothesis Testing

DECISION	HYPOTHESIS	
	Actually True	Actually False
To Accept	Correct	β error or type II error
To Reject	α error or Type I error	correct

The choice of level of significance (α) is arbitrary. If α is set so low that there is small chance of type I error, which means using a large critical value of the statistic in question, then there is more chance of type II error. If α is set, β is also set, and an increase in α causes a decrease in β and vice-versa. Therefore the only way to reduce β, when α is fixed, is by increasing the sample size. It is not possible to reduce both errors simultaneously in a single test. Either error is undesirable, of course. It is desirable that both errors be reduced to a minimum. In practice, the 0.05 to 0.01 probability for

TABLE 3.2

Relation Between α and β

DECISION	HYPOTHESIS	
	Actually True	Actually False
To Accept	$1 - \alpha$	β
To Reject	α	$1 - \beta$

type I error usually gives the minimum probabilities of both type of errors. However, there are cases where one type of error is more important to avoid than the other. The value of α is arbitrary depending on the characteristics of the problem in consideration.

One of the most useful tools in statistical analysis is chi-square statistic (χ^2). It can be used to test the goodness of fit of

experimental observations to a hypothesized probability distribution. This statistic can be defined as the sum of the squares of the standarized normal variables (Z_i^2) with s degrees of freedom:

$$\chi^2 = \sum_{i=1}^{s} (Z_i)^2 \qquad (3.12a)$$

with $Z_i = \frac{\text{Independent random variable} - \mu}{\sqrt{\text{variance}}}$.

The distribution of χ^2 depends on s because the z's are standarized. However, a degree of freedom is lost for each constraint placed on the observations or for each parameter estimated from observations. And its probability distribution is

$$P(\chi_*^2) = P(\chi^2 \leq \chi_*^2) = 1 - \alpha \qquad (3.12)$$

where χ_*^2 is available from statistical table with s degrees of freedom.

A two-sided test is a symmetric hypothesis test. It is used to test a value of an estimate against its known or expected value by means of any selected statistic. In this study, the standard normal probability density function is used.

The assumptions required in the mathematical development of this study will be discussed in the following section.

3.3 PRELIMINARY ASSUMPTIONS

In order to simplify the derivation of the technique of data adjustment, some assumptions pertaining to equation 2.3 are made.

Rewrite equation 2.3:

$$B\underline{y} + S' \underline{\xi} = 0 \tag{3.13}$$

where the matrix B is the balance coefficient matrix which characterizes the topology of the system. Its elements are defined in the classical way:

$$b_{i,j} = \begin{cases} 1 & \text{if the ith component is an input variable.} \\ -1 & \text{if the component i is an output variable.} \\ 0 & \text{if the ith component is known to be absent.} \end{cases}$$

The absolute value of each element of S, s_{ij}, is the stoichiometric coefficient of the jth component in the ith reaction in the case where the mass balance is in molar units. It is the product of that coefficient and the molecular weight of the component in the case where mass balance is in mass units. The sign of s_{ij} is negative for reactants, positive for reaction products and zero for chemically inert components.

The set of assumptions is the following:

1. It is assumed that data are available for each component in at least one inlet or outlet stream. Otherwise the method will result in

the missing component being assigned zero in all streams.

2. Each stream vector must have a position for all components and in the same order. A component known to be absent in a particular stream will be assigned a measured value of zero and will not be adjusted by the method. Since the covariance matrix of the measurement vector must be positive definite*, this variable must be assigned a variance greater than zero. The value assigned is arbitrary (in this study a value of 1 is used).

3. Considering the case of missing variables, the matrix B is a partitioned matrix of B_1 and B_2. The rank of B_1 must be equal to the number of mass balance constraints (n) and the rank of B_2 must be equal to the number of missing variables (ms), (see Appendix II). In cases where this assumption is not met, the matrix B_1 and B_2 must be somewhat modified.

4. The chemical nature of the components can be represented by the atom matrix E, where e_{ij} is the number of the jth atom contained in a molecule of the ith component. It is required that

$$S\,E = 0 \qquad (3.14)$$

therefore

$$\rho(S) \leq n - \rho(E) \qquad (3.15)$$

the matrix S is assumed to be the stoichiometric coefficient matrix of the independent set of reactions, and hence, its rank is equal to the

---

* This requirement is discussed below.

number of independent reactions ($\rho(S) = p < n$). In that way the above requirement is satisfied.

5. No reaction exists solely among components which have missing variable in any stream entering or leaving the balance envelope, and the number of missing variables is less than or equal to the difference between the number of components, n, and chemical reactions, p, ($ms < n-p$, see Appendix II).

6. The true value of the vector $\underline{x}$ is unknown, however, the measurement vector, $\underline{x}$, can be assumed to be

$$\underline{x} = \underline{\mu}_x + \underline{\varepsilon} \tag{3.16}$$

where $\underline{\mu}_x$ is the vector of the true value of vector $\underline{x}$ which satisfies the mass/energy balance process. The vector $\underline{\varepsilon}$ is the unknown measurement error vector. It is assumed also that the unknown measurement errors are random variables, multivariate normally distributed with zero means and positive definite covariance matrix, Σ_x. That is:

$$\underline{\varepsilon} \sim Nm\ (\underline{0},\ \Sigma_x) \tag{3.17}$$

Hence, its probability density function, given $\underline{x}$ and Σ_x is

$$f(\underline{\varepsilon}/\underline{x},\ \Sigma_x) = \frac{1}{2\pi^{1/2m}|\Sigma_x|^{1/2}} \exp[-\frac{1}{2}(\underline{\varepsilon}'\ \Sigma_x^{-1}\underline{\varepsilon})] \tag{3.18}$$

and vector $\underline{x}$ is m-variate normally distributed.

In the following section, the ideas discussed above are used to derive an improved method of data adjustment. Particularly, the concepts of singular normal distribution and hypothesis testing are used to develop a method for detecting gross errors.

3.4 STATISTICAL TREATMENT OF MASS/ENERGY BALANCE

This section is divided into three subsections. In the first one, the technique for data adjustment is developed to consider missing variables. This development assumes that gross errors have been detected and removed from the original data. In the second subsection, a test function to detect gross error is derived. And in the third one, the particular cases of the overall problem are discussed.

3.4.1 Derivation of Data Adjustment Technique.

A solution to the problem of adjustment of data with missing variables is presented below.

The problem of data adjustment with missing variables may be considered to be a constrained minimization problem where the total necessary adjustment is to be minimized. As a method of solution, Weighted Least Squares method is used. Hence, the problem is defined as:

$$\underset{\underline{a},\,\underline{z}}{\text{Minimize}}\ \underline{a}'\,\Sigma_x^{-1}\,\underline{a} \qquad (3.19)$$

subject to the mass/energy balance constraints

$$\underline{\psi}(\underline{a}, \underline{z}) = B_1(\underline{x} - \underline{a}) + A\underline{z} = 0 \qquad (3.20)$$

where $A = [B_2 \mid S']$ and $\underline{z} = \begin{bmatrix} \underline{u} \\ \underline{\xi} \end{bmatrix}$. According to the method of Lagrange Multipliers, this is equivalent to minimizing the unconstrained Lagrange Function:

$$F(\underline{a}, \underline{z}, \underline{\ell}) = \underline{a}' \Sigma_x^{-1} \underline{a} + 2\underline{\ell}'(B_1(\underline{x} - \underline{a}) + A\underline{z}) \qquad (3.21)$$

where $\underline{\ell}$ is the vector of Lagrange Multipliers.

A necessary condition for the function $F(\underline{a}, \underline{z}, \underline{\ell})$ to be a minimum at a point is that all first order partial derivatives must be zero at that point:

$$\frac{\partial F}{\partial \underline{a}}(\underline{a}, \underline{z}, \underline{\ell}) = 2\Sigma_x^{-1} \underline{a} - 2 B_1' \underline{\ell} = \underline{0} \qquad (3.22)$$

$$\frac{\partial F}{\partial \underline{z}}(\underline{a}, \underline{z}, \underline{\ell}) = 2A' \underline{\ell} = 0 \qquad (3.23)$$

$$\frac{\partial F}{\partial \underline{\ell}}(\underline{a}, \underline{z}, \underline{\ell}) = 2 B_1(\underline{x} - \underline{a}) + 2A\underline{z} = \underline{0} \qquad (3.24)$$

Equations 3.22 to 3.24 are solved simultaneously to yield the Least Squares estimates:

$$\underline{a} = \Sigma_x B_1' H^{-1} [I_n - AR_1] B_1 \underline{x} \qquad (3.25)$$

$$\underline{z} = -R_1 B_1 \underline{x} \qquad (3.26)$$

where $R_1 = (A' H^{-1} A)^{-1} A' H^{-1}$ and $H = B_1 \Sigma_x B_1'$

substituting $A = [B_2' \,!\, S']$ and $z = \begin{bmatrix} u \\ \xi \end{bmatrix}$

into equations 3.25 and 3.26, respectively, yields after rearrangement:

$$\underline{a} = \Sigma_x B_1' H^{-1} (I_n - B_2 R)(I_n - S'(Q S')^{-1} Q) B_1 \underline{x} \qquad (3.27)$$

$$\underline{u} = R[S'(QS')^{-1} Q' - I_n] B_1 \underline{x} \qquad (3.28)$$

$$\underline{\xi} = -(Q S')^{-1} Q B_1 \underline{x} \qquad (3.29)$$

where $R = (B_2' H^{-1} B_2)^{-1} B_2' H^{-1}$; $Q = SH^{-1}(I_n - B_2 R)$

The derivation of equation 3.27 to 3.29 may be found in Appendix I.

Under assumptions 3 and 4, discussed in the previous section, the matrices $(Q S')$, H and $(B_2' H^{-1} B_2)$ are non-singular (see Appendix II). Under assumption 5, the Least Squares estimates are equivalent to maximum likelihood and minimum variance unbiased estimates.

3.4.2 Detection of Biased Data.

The statistical treatment of mass balance presented in the previous section relied on the assumption that the measured data are subject only to random errors with m-variate normal distribution having zero mean and known positive definite covariance matrix. In practice however, process data may contain types of errors which invalidate this assumption. Malfunction of measurement devices (ie. incorrect calibration etc...), or incomplete information from process equipment (leaks, accumulation, incorrect assumption of steady-state, etc ...) or human error in reading or reporting data result in this kind of error. Errors of this type will be referred to as gross errors and process data containing gross errors, as previously stated, will be called biased data.

Before proceeding with the adjustment of data, therefore, it is essential that the data be tested for bias. This requires testing for multivariate normality of the measurement vector $\underline{x}$. This is a difficult problem and therefore a test function has been developed below which reduces the problem to one of testing for univariate normality,

Consider, in general, the linear mass balance equation

$$\underline{e} = B_1\underline{x} + B_2\underline{u} + S'\underline{\xi} \qquad (3.30)$$

using equation 3.28 and 3.29 in equation 3.30 and rearranging, we get

$$\underline{e} = (I_n - B_2R)(I_n - S'(QS')^{-1}Q)B_1\underline{x} \qquad (3.31)$$

This mass balance equation will be satisfied exactly when the measurement vector $\underline{x}$ is equal to the vector of true value , $\underline{\mu}_x$. This statement is based on the following consideration.

Defining $\underline{\mu}_u$ as the true value of missing variables vector and $\underline{\mu}_\xi$ as the true value of the extent reaction vector, equation 3.30 can be written:

$$B_1\underline{\mu}_x + B_2\underline{\mu}_u + S'\underline{\mu}_\xi = 0 \qquad (3.32a)$$

and according to the property of a linear combination of normal variables, the expected value of vectors $\underline{u}$ and $\underline{\xi}$ are:

$$\begin{bmatrix} E(\underline{u}) \\ \hline E(\underline{\xi}) \end{bmatrix} = \begin{bmatrix} R[S'(QS')^{-1} Q - I_n] \\ \hline -(QS')^{-1}Q \end{bmatrix} B_1 \underline{\mu}_x \qquad (3.32b)$$

Replacing $B_1\underline{\mu}_x$ by its expression obtained from 3.32a:

$$\begin{bmatrix} E(\underline{u}) \\ \hline E(\underline{\xi}) \end{bmatrix} = \begin{bmatrix} \underline{\mu}_u \\ \hline \underline{\mu}_\xi \end{bmatrix} \qquad (3.32c)$$

Since $\underline{e}$ is a linear combination of normal variables $\underline{x}$, $\underline{u}$ and $\underline{\xi}$, then

$$E(\underline{e}) = B_1\underline{\mu}_x + B_2\underline{\mu}_u + S'\underline{\mu}_\xi = 0 \qquad (3.33)$$

using equation 3.32b and 3.32c into equation 3.33, one obtains:

$$E(\underline{e}) = (I_n - B_2R)(I_n - S'(QS')^{-1}Q)B_1\underline{\mu}_x = 0 \qquad (3.34)$$

By comparing Equation 3.34 with 3.31, it can be asserted that the vector of true values of the measurements satisfies the equation 3.31 as said above. But because of the measurement errors, previously defined, the measurement vector is in general inconsistent. Vector $\underline{e}$ is in fact the imbalance or discrepancies of the mass balances of the process. It is n-variate normally distributed with zero means and covariance matrix

$$\Sigma_e = [I_n - B_2R][H - S'(QS')^{-1}QH] \qquad (3.35)$$

which is a singular matrix of rank r equal to the number of constraints minus the size of vector $\underline{z}$ ($\rho(\Sigma_e) = r = n - ms - p$).

To test the validity of normality of vector $\underline{e}$ is equivalent to test the validity of the assumption that the measurement errors are normally distributed random variables. For this purpose the scaled vector $H^{-1/2}\underline{e}$ is constructed. This vector is also normally distributed with mean zero and covariance matrix, Σ

$$\Sigma = H^{-1/2} \Sigma_e H^{-1/2} \tag{3.36}$$

of rank r. It has the following characteristic:

$$(\Sigma)^2 = \Sigma, \tag{3.37}$$

Then, this matrix is idempotent. The eigenvalues of an idempotent matrix are either zero or one; that is

$$A = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] \tag{3.38}$$

and from equation 3.22, it can be proved that the set of eigenvectors corresponding to eigenvalues equal to zero, Γ_2, is the orthonormalized matrix from $H^{-1/2}A$, and the set Γ_1 corresponding to eigenvalues equal to one can be the orthonormalized matrix from the r-linearly independent

columns of Σ or chosen such that the following set of relationships is satisfied:

$$H^{-1/2}\ \Sigma_e\ H^{-1/2}\ \Gamma_1 = \Gamma_1 .$$
$$\Gamma_1'\ \Gamma_1 = I_r \qquad (3.39)$$
$$\Gamma_1'\ \Gamma_2 = \Gamma_2'\ \Gamma_1 = 0$$

Using the theory of singular normal distribution discussed in section 3.2, the vector $\underline{w}$ was determined to be

$$\underline{w} = \Gamma_1'\ H^{-1/2}\ \underline{e} \qquad (3.40)$$

This vector will be distributed according to Nr(0,I), if the assumption of normality of the measurement vector $\underline{x}$ is valid. Since the covariance matrix of $\underline{w}$ is the identity matrix (I), the original problem of testing for multivariate normality has been reduced to testing for univariate normality. The test function vector is the vector $\underline{w}$ and the hypothesis testing is:

$$H_o: x_j \in \pi_1;\ j = 1, \ldots m$$

$$H_1: x_j \in \pi_1;\ j \neq s \text{ and } j = 1, \ldots m$$
$$x_s \in \pi_2$$

where $\pi_1 \sim Nm(\underline{\mu}_x, \Sigma_x)$ and $\pi_2 \sim Nm(\underline{\mu}_x + \underline{\gamma}, \Sigma_x)$ with $\underline{\gamma}$ = vector of bias

If $\underline{x} \in \pi_1$, $\underline{w} \sim Nr\ (0, I)$ and since $\underline{w}'\underline{w}$ has chi-square distribution with r degrees of freedom, at the specified significance level α under H_o:

$$\underline{w}'\underline{w} = \underline{e}'\ H^{-1}\underline{e} < \chi^2\ (1 - \alpha,\ r) \qquad (3.41)$$

The hypothesis H_o is rejected in favour of H_1 if $\underline{w}'\underline{w}$ fails the χ^2 test (3.41). In this event, it is concluded that the data contain one or more gross errors.

It is necessary to provide this test with an allowable error probability, which gives a critical value of $\underline{w}'\underline{w}$. The choice of the value of this error probability depends on process characteristics. However, a type I error probability (α) of 0.10 has been suggested in the literature. And since, it is more important not to accept biased data than to reject unbiased data, the value of 0.10 is also suggested here.

The data have been tested as a whole. Further, it is necessary to identify the measurements having gross error in order to take corrective actions. One obvious approach is to carry through the analysis with and without suspected measurement(s) containing gross error and to compare results (Ripps (1965) and Nogita (1972)). Here it is suggested to use the r-largest values of the imbalance vector $\underline{e}$, through vector $\underline{w}$ for identifying gross errors.

Because the covariance matrix of $\underline{w}$ is the identity matrix, each element of $\underline{w}$ can be examined separately. A standard two-sided test at probability level α on the normal univariate distribution is made for

each element of $\underline{w}$. The critical value of w_i is determined at the significance level of 0.10. The element of $\underline{w}$ which fails the test, corresponds to an element of the r-largest values of the imbalance vector $\underline{e}$ which contains the gross error. The non-zero elements of the row in the matrix B_1 (equation 3.30), corresponding to the elements of the r-largest values of $\underline{e}$ containing gross errors, point to a subset of $\underline{x}$ suspected of having gross error.

The significance test for w_i is given by

$$|w_i| \leq Z(1 - \alpha/2) \tag{3.42}$$

where $Z(1 - \alpha/2)$ is the standard normal probability distribution.

This method is an improvement over the method of Nogita and that of Ripps because it reduces the search for gross errors to a subset of the measurement vector. Additionally, the chi-square test is a convenient and quick indicator of the presence of gross errors.

3.4.3 Formulation of Specific Cases From the General Problem

Three particular cases can be identified in the overall problem of data adjustment with missing variables discussed in the previous subsection. In these cases the mathematical formulation can be reduced to a more simplified form. These cases are discussed below.

(a) No chemical reaction and no missing variables. This is the simplest case. The matrices B_2 and S are absent. Then the Lagrange function is

$$F(\underline{a}, \underline{\ell}) = \underline{a}' \Sigma_x^{-1} \underline{a} + 2\underline{\ell}' B_1 (\underline{x} - \underline{a}) \tag{3.43}$$

and the adjustment vector was determined to be

$$\underline{a} = \Sigma_x B_1' H^{-1} B_1 \underline{x} \tag{3.44}$$

by the same procedure used in subsection 3.4.1. Equations 3.30 and 3.35 reduce to

$$\underline{e} = B_1 \underline{x} \tag{3.45}$$

$$\Sigma_e = H \tag{3.46}$$

From equation 3.46 the scaled imbalance vector's covariance is the identity matrix. Therefore its eigenvector matrix is also the identity matrix and the test function $\underline{w}$ defined by equation 3.40 becomes the product of the imbalance vector associated with each component, $\underline{e}$, and the square root of its covariance matrix, H, with n degrees of freedom

$$\underline{w} = H^{-1/2} \underline{e} \tag{3.47}$$

If any element of $\underline{w}$ falls outside the critical range, it can be said that its corresponding element in vector $\underline{e}$ is too large and either the input or output of that specific component is in gross error. If all elements of $\underline{w}$ fall outside the critical range, then all elements of $\underline{e}$ are too large and that leads us to suspect that there is a leak or accumulation in the system.

Suppose there is a leak in the system (eg., Leaking pump). The fictitious stream $\underline{\theta}$ represents the leak (Fig. 3.1),

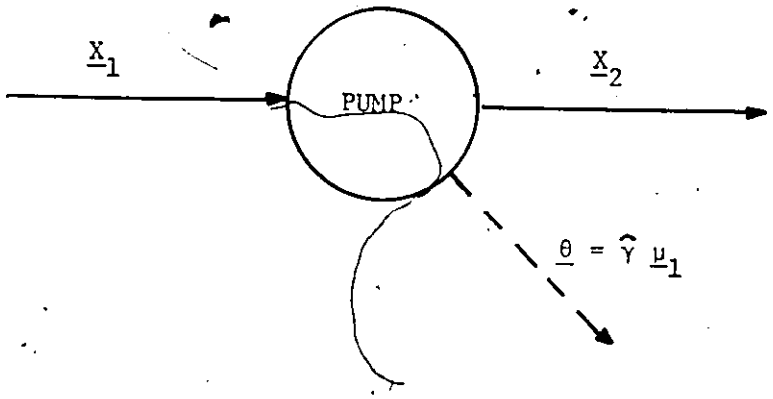


Figure 3.1 Leaks in the system

Then the mass balance is

$$\underline{e} = \underline{x}_1 - \underline{x}_2 = \underline{\theta} \tag{3.48}$$

Since $\underline{\theta}$ is different from zero, all elements of $\underline{w}$ will be out of range for $(\hat{\gamma})$ large enough

(b) No chemical reaction and missing variables. In this case, the Least Squares estimate of $\underline{a}$ and $\underline{u}$ are:

$$\underline{a} = \Sigma_x\ B_1'\ H^{-1}\ [I_n - B_2 R] B_1 \underline{x} \tag{3.49}$$

$$\underline{u} = -\ R B_1 \underline{x} \tag{3.50}$$

and

$$B_2'\ H^{-1}\ \underline{e} = 0 \tag{3.51}$$

the test function $\underline{w}$ is defined by equation 3.40 with

$$\Gamma_1' \Gamma_2 = \Gamma_2' \Gamma_1 = 0 \tag{3.52}$$

and

$$\Gamma_2 = (H^{-1/2} B_2) \text{ Orthonormalized} \tag{3.53}$$

It is still possible to make the same analysis as before on the subset of the elements of vector $\underline{e}$, having excluded the ms-smallest values of $\underline{e}$.

(c) Chemical reaction and no missing variables. This case is basically the same as case (b) with $B_2 = S'$, except that we now effectively test element imbalances, not component imbalances as in the cases (a) and (b). The matrix Γ_1 can be expanded from S.

Changing B_2 to S', equations 3.49 to 3.51 become:

$$\underline{a} = \Sigma_x B_1' H^{-1} [I_n - S' (QS')^{-1} Q] B_1 \underline{x} \tag{3.54}$$

$$\underline{\xi} = - (QS')^{-1} Q B_1 \underline{x} \tag{3.55}$$

$$SH^{-1} \underline{e} = \underline{0} \tag{3.56}$$

where $Q = SH^{-1}$

In Chapter II the existing literature on data adjustment was discussed and in Chapter III the statistical method of solving the problem using weighted Least Square estimation was presented. In the next chapter the computer program designed for this method is described.

CHAPTER IV

COMPUTER IMPLEMENTATION

4.1 INTRODUCTION

For the method of data adjustment developed in this study, a program has been written in Fortran IV and implemented on a Control Data Corporation CDC 6400. The program consists of a main program called ANADATA and a set of subroutines which will be described in the next section. The input to ANADATA consists of: (a) dimensions of the process: the number of measurements, missing variables, constraints, and the maximum number of independent reactions taking place, (b) information on the measured variables: the measurement vector and its covariance matrix, (c) the coefficient matrix of mass/energy balance constraints and, (d) the reaction stoichiometric coefficient matrix.

The program was designed to deal with the most general situations in which there is statistical interaction between measured variables (non-diagonal covariance matrix).

4.2 PROGRAM DESCRIPTION

This section is a general description of the main program and the subroutines to perform the required operations. Figure 4.1 shows the flowchart of ANADATA main program. A complete listing of the program is given in Appendix V.

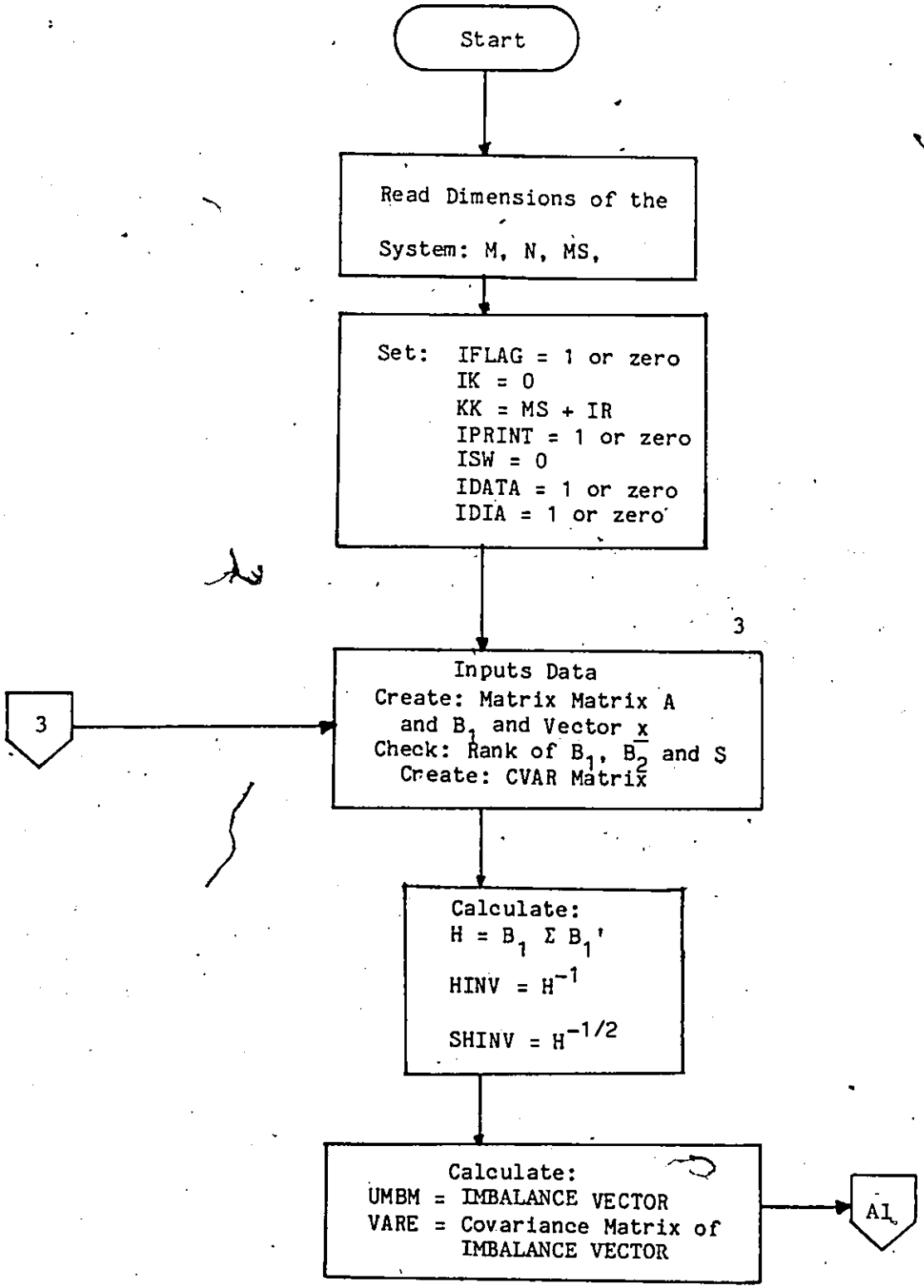


Figure 4.1 ANADATA Main Program Flow Chart

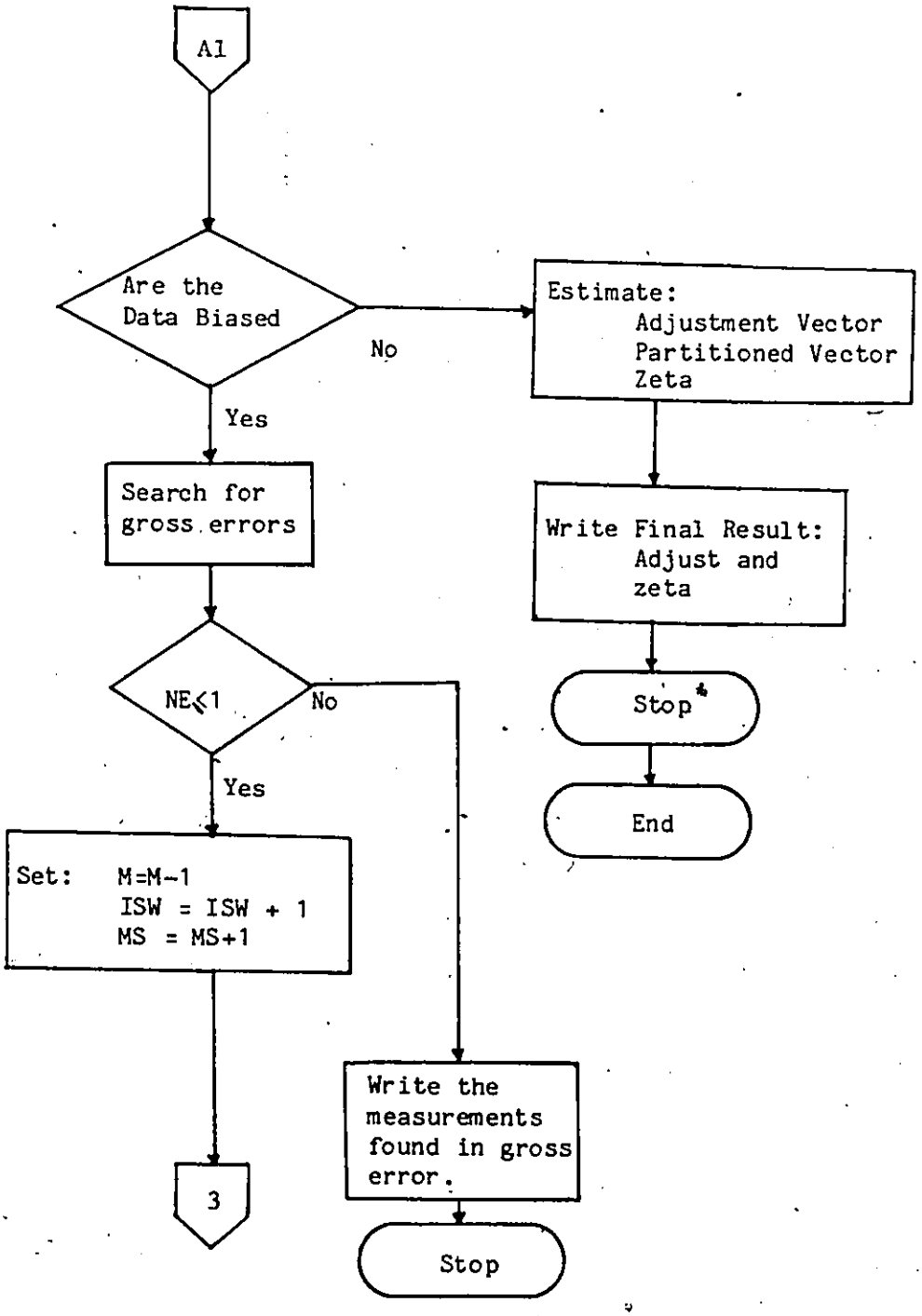


FIGURE 4.1 continued

4.2.1 Main Program

ANADATA program inputs data and coordinates execution of the calculation performed by the subroutines.

After checking the consistency of the input data with assumptions 3 and 4 in section 3.3 (rank of matrices B_1, B_2 and S), execution proceeds in the following order:

(a) The measurement vector $\underline{x}$ is tested for bias by testing the hypothesis stated in chapter III, equation 3.41.

(b) If the hypothesis is accepted no biased data are present. The required adjustment to $\underline{x}$ is calculated and the vector of missing variables and/or the extents of reactions are estimated using equations 3.25 and 3.26 respectively.

(c) If the hypothesis is rejected, the program searches for and outputs the subset of $\underline{x}$ most likely to contain the bias. The program is stopped if more than one measurement is found to be corrupted. At this point, a decision is required by the analyst as to which measurement(s) to exclude from the input data.

(d) If only one corrupted measurement is found, the program removes this particular measurement from the input data and considers it as a missing variable. The program is restarted and execution begins at (a) as before.

The output is the adjusted value of the measurement vector $\underline{x}$ and the estimate of missing variables and/or extenst of reactions.

4.2.2 Subroutines

The program has seven main subroutines to handle all the algebraic

calculations required by the method. Three function subprograms are responsible for estimating the elements of the test function vector, the chi-square statistic associated with the raw data and the rank of a matrix. Six auxiliary subroutines carry out the matrix operations. Three subroutines from the IMSL library and one from the SSP library are also used. These subroutines calculate the eigensystem and inverse of a matrix (IMSL) and matrix factorization and rank determination (SSP).

Here, the seven main subroutines are briefly described:

CALC1H is used to calculate the covariance matrix of $B_1\underline{x}$. It also calculates the inverse of this covariance matrix and its negative square root. These matrices are used in the estimation of the test function elements and adjustment vector.

The method used to calculate the negative square root of the matrix is based on the diagonal factorization of a matrix as follows:

Considering any positive definite matrix D,

$$D^k = \Gamma \Lambda^k \Gamma' ;$$

where k can be any rational number, Γ is the matrix of eigenvectors corresponding to the eigenvalues of D (Λ).

The inverse of the matrix is calculated with the IMSL subroutine LINV2F (using Linear Equation Solution). The flow-chart of CALC1H subroutine is given in Fig. 4.2.

CALC1E estimates the balance vector $\underline{e}$ according to equation 3.31 and the covariance matrix of $\underline{e}$ according to equation 3.35. It is called by the main program before testing the data for bias. Figure 4.3 shows

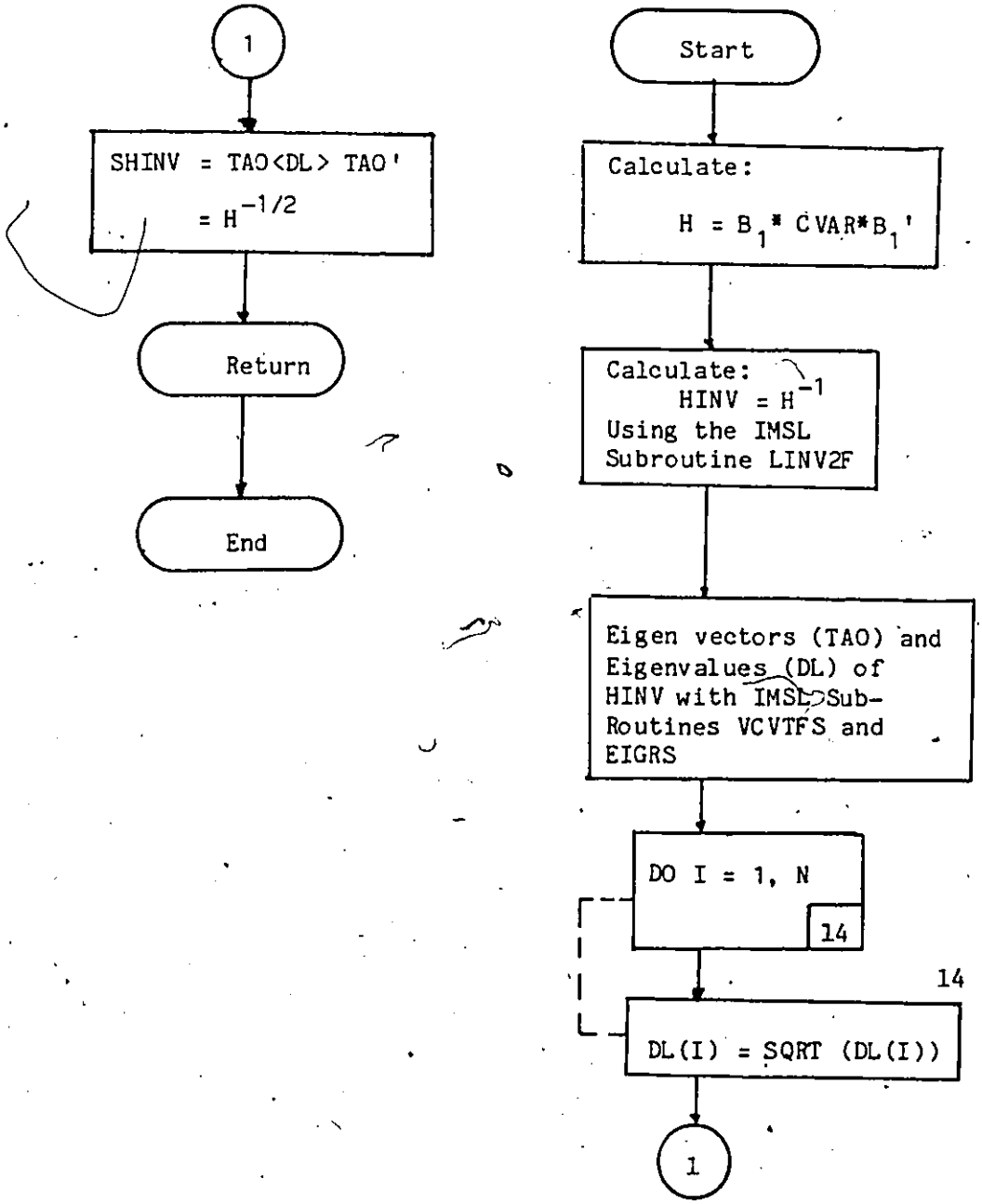


FIGURE 4.2 Subroutine CALC1H Flow Chart.

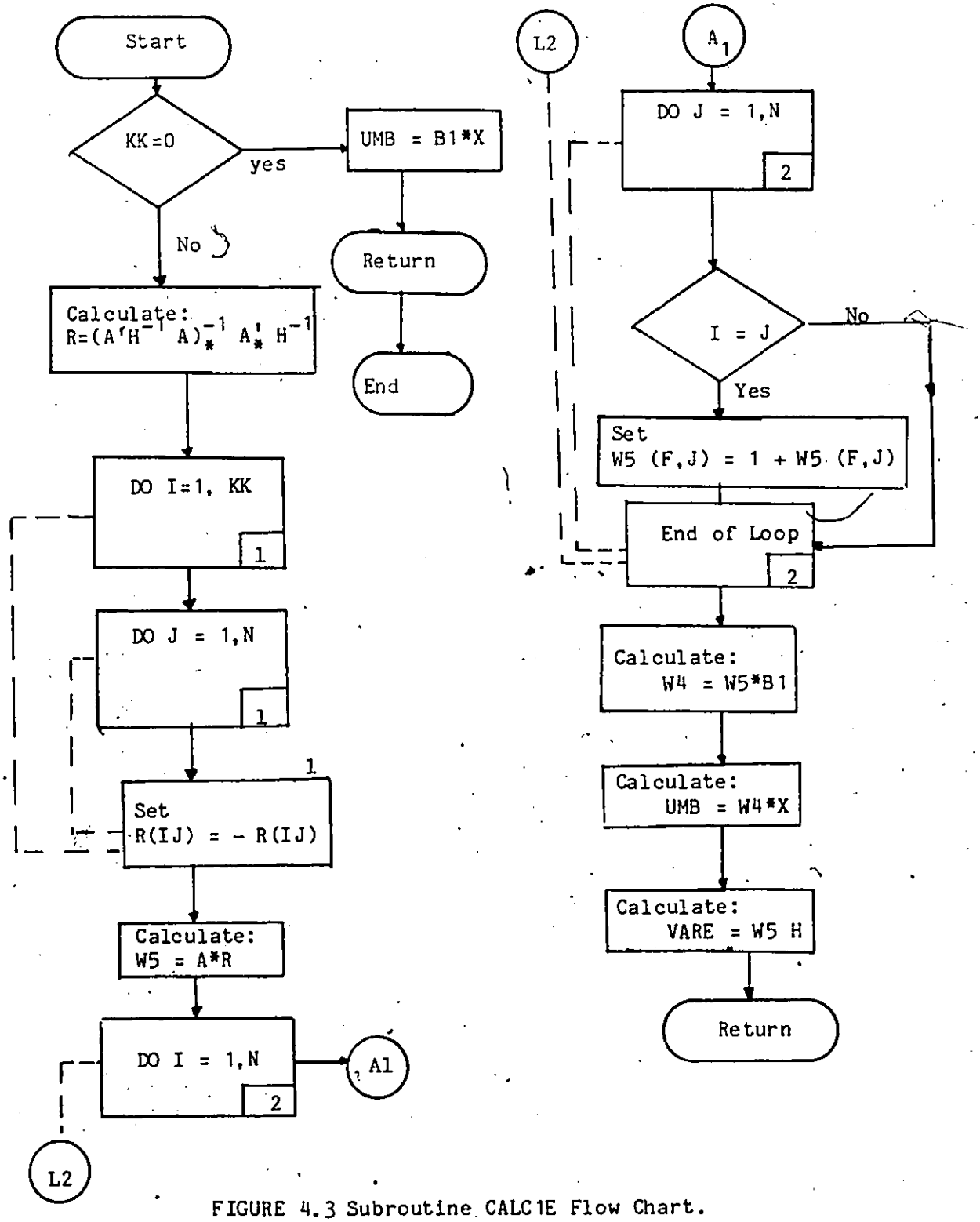


FIGURE 4.3 Subroutine CALC1E Flow Chart.

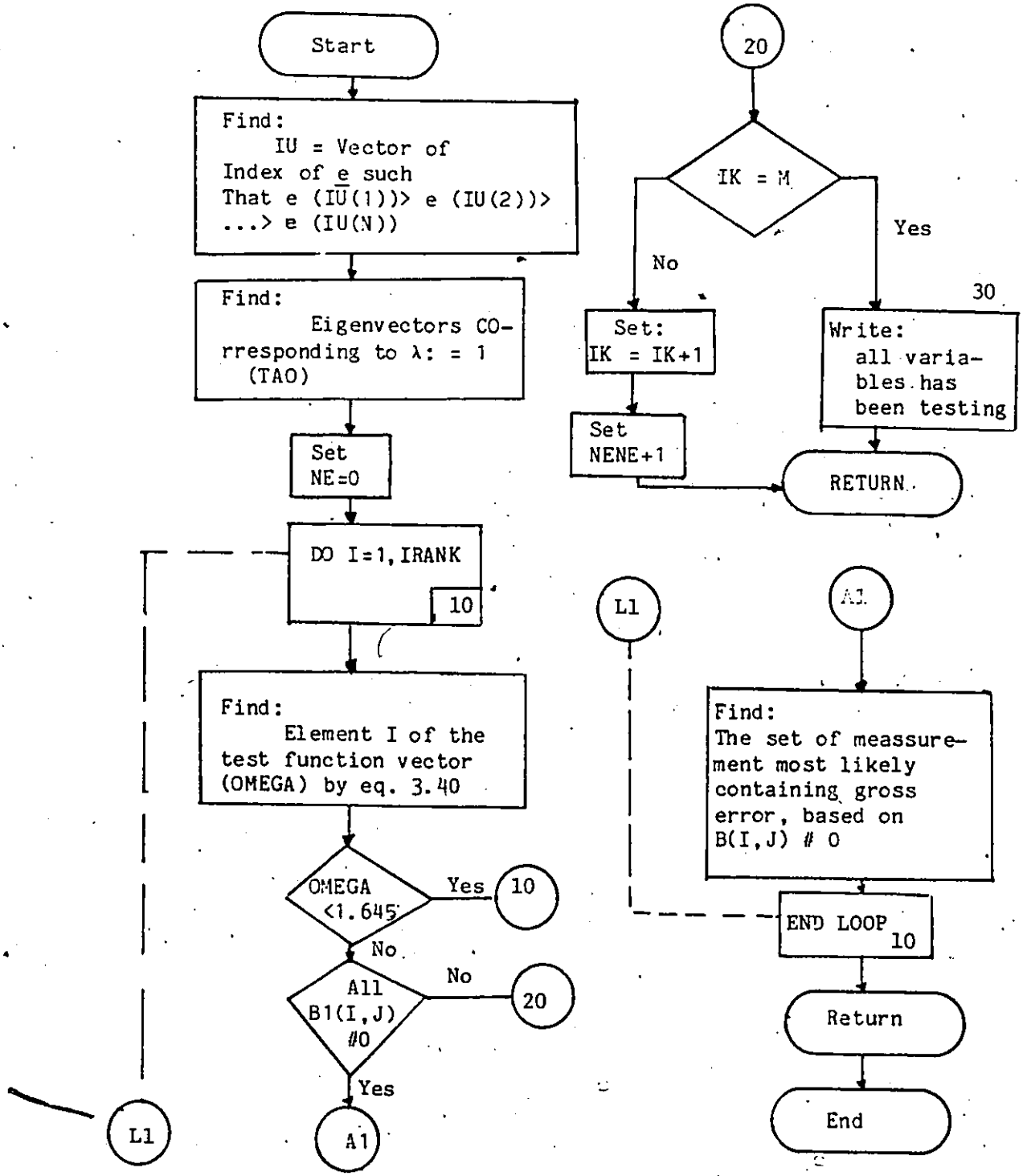


FIGURE 4.4 Subroutine SEARCH Flowchart

the flow-chart of this subroutine.

SEARCH finds the most likely measurement(s) having gross error as described in section 3.4.2. Its flow chart is given in Fig. 4.4.

CALC2E obtains the adjustment vector $\underline{a}$ using equation 3.25 and estimates the missing variables and/or extents of reactions using equation 3.26. It is called by the main program when the hypothesis test has been accepted with a probability of 90% or in the event that one corrupted measurement has been detected and removed from the data. The flow-chart is shown in Fig. 4.5.

RECHAZ is called (a) to check the input data for ranks of matrices B_1, B_2 and S, and to construct matrix A with matrices B_2 and S, and (b) to reject only one measurement having gross error as determined by subroutine SEARCH. The measurement is considered a missing variable by reducing the dimension of $\underline{x}$ by one and increasing the dimension of ZETA by one. The flow-chart is given in Fig. 4.6.

CALC1T is called by subroutine SEARCH to estimate the set of eigenvectors of matrix Σ corresponding to eigenvalues equal to one, according to the procedure described in section 3.4.2. This subroutine must be provided by the user when matrix A exists. The inputs to the calling argument required by this subroutine are: (1) covariance matrix of vector $\underline{e}$ (VARE) calculated by subroutine CALC1E, (2) the partitioned matrix A defined by subroutine RECHAZ, (3) the rank of VARE (IRANK), (4) the number of constraints (N), (5) the size of matrix A (number of columns = KK), and (6) the inverse of $H^{1/2}$ (SHINV). The output is matrix TAO1 of the set of eigenvectors corresponding to eigenvalues

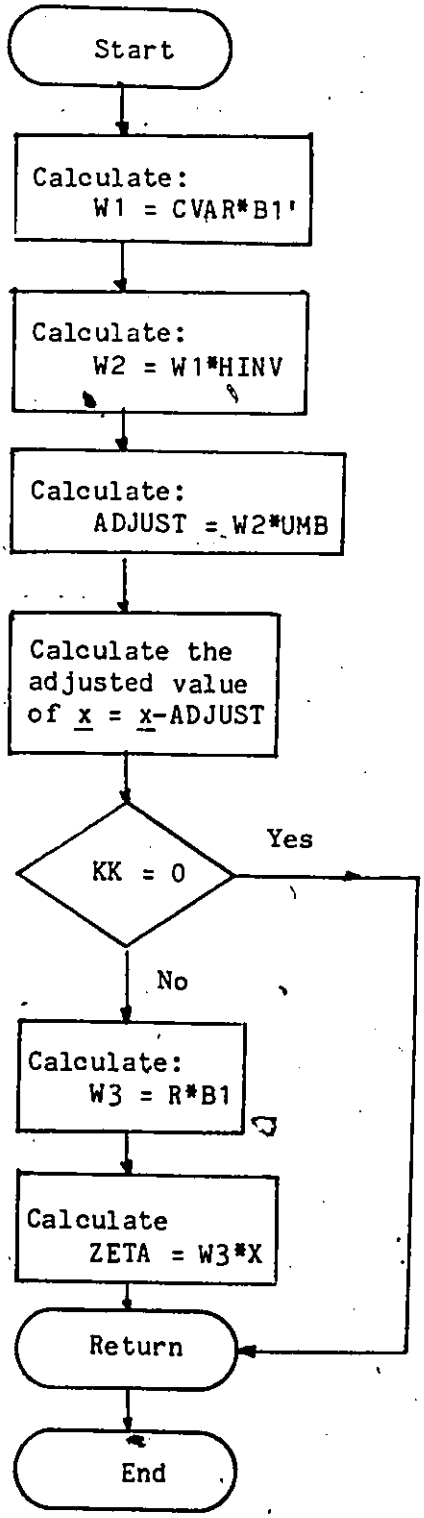


FIGURE 4.5 Subroutine CALC2E Flow Chart.

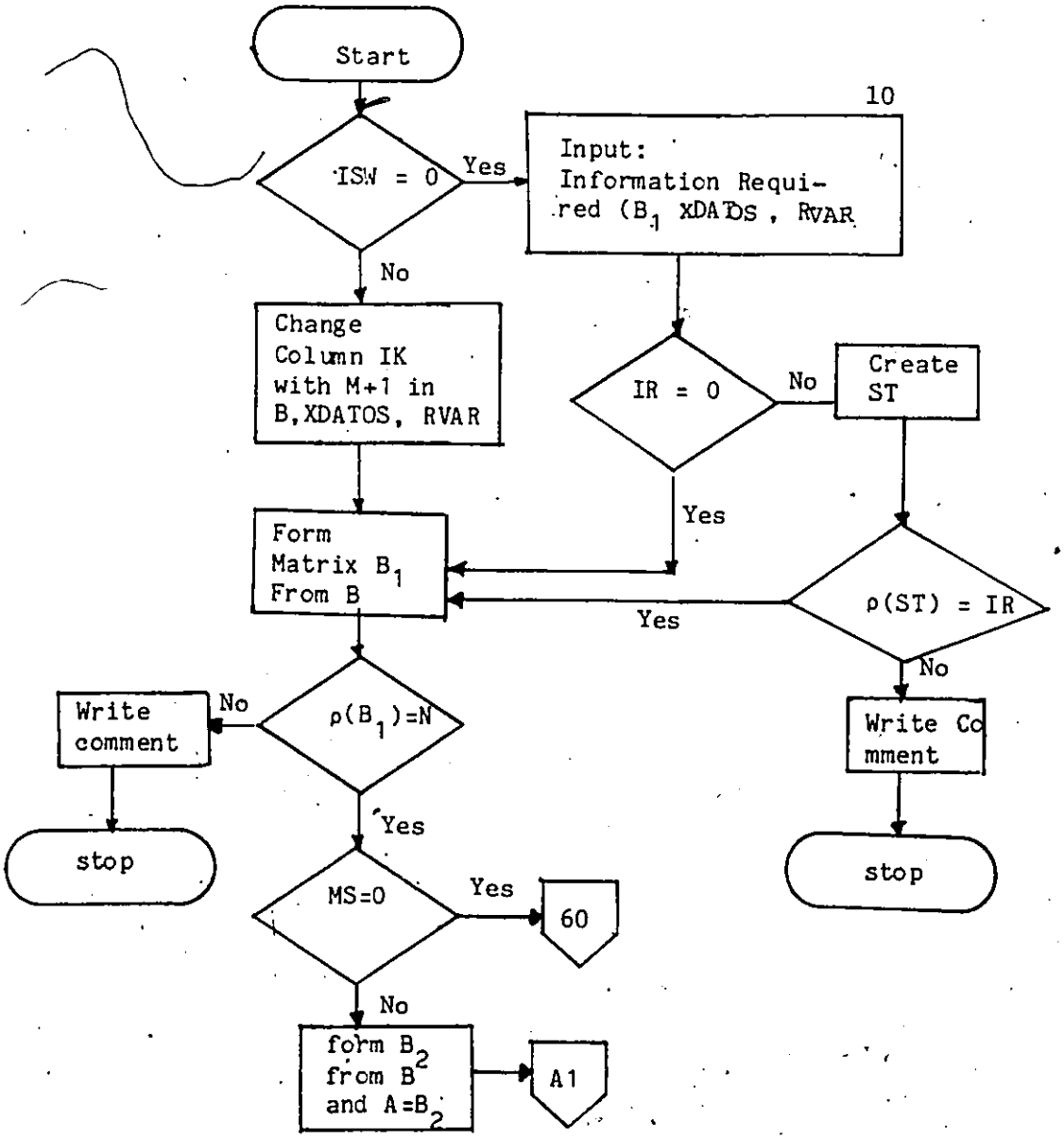


FIGURE 4.6 Subroutine RECHAZ Flowchart.

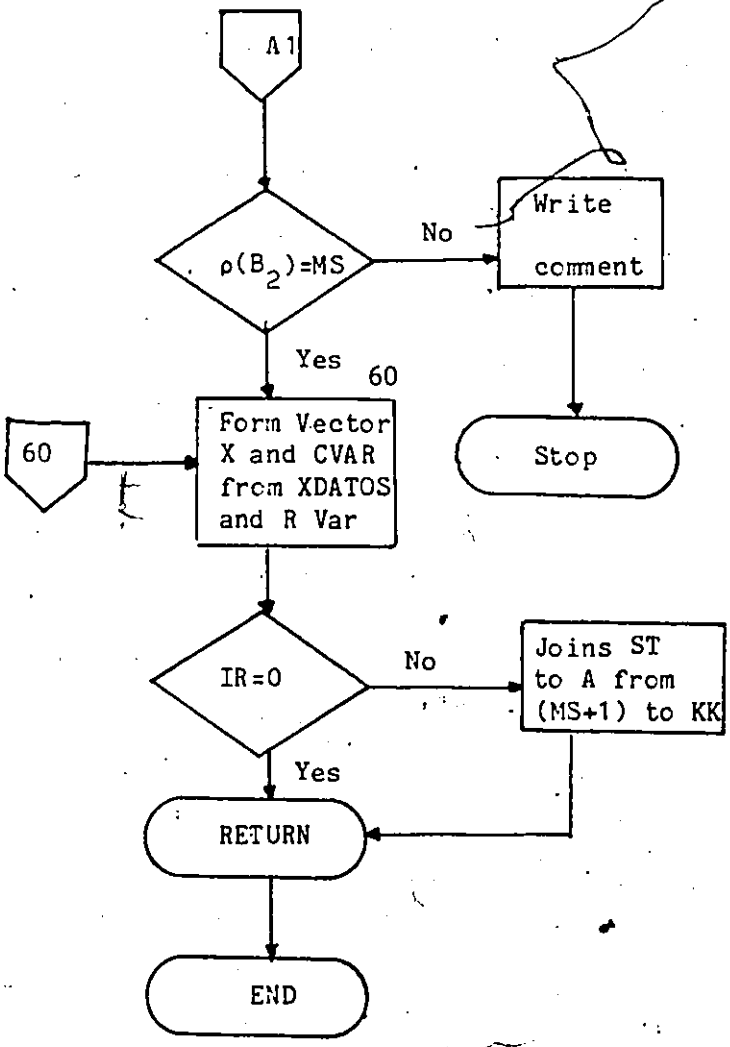


FIGURE 4.6 Continue

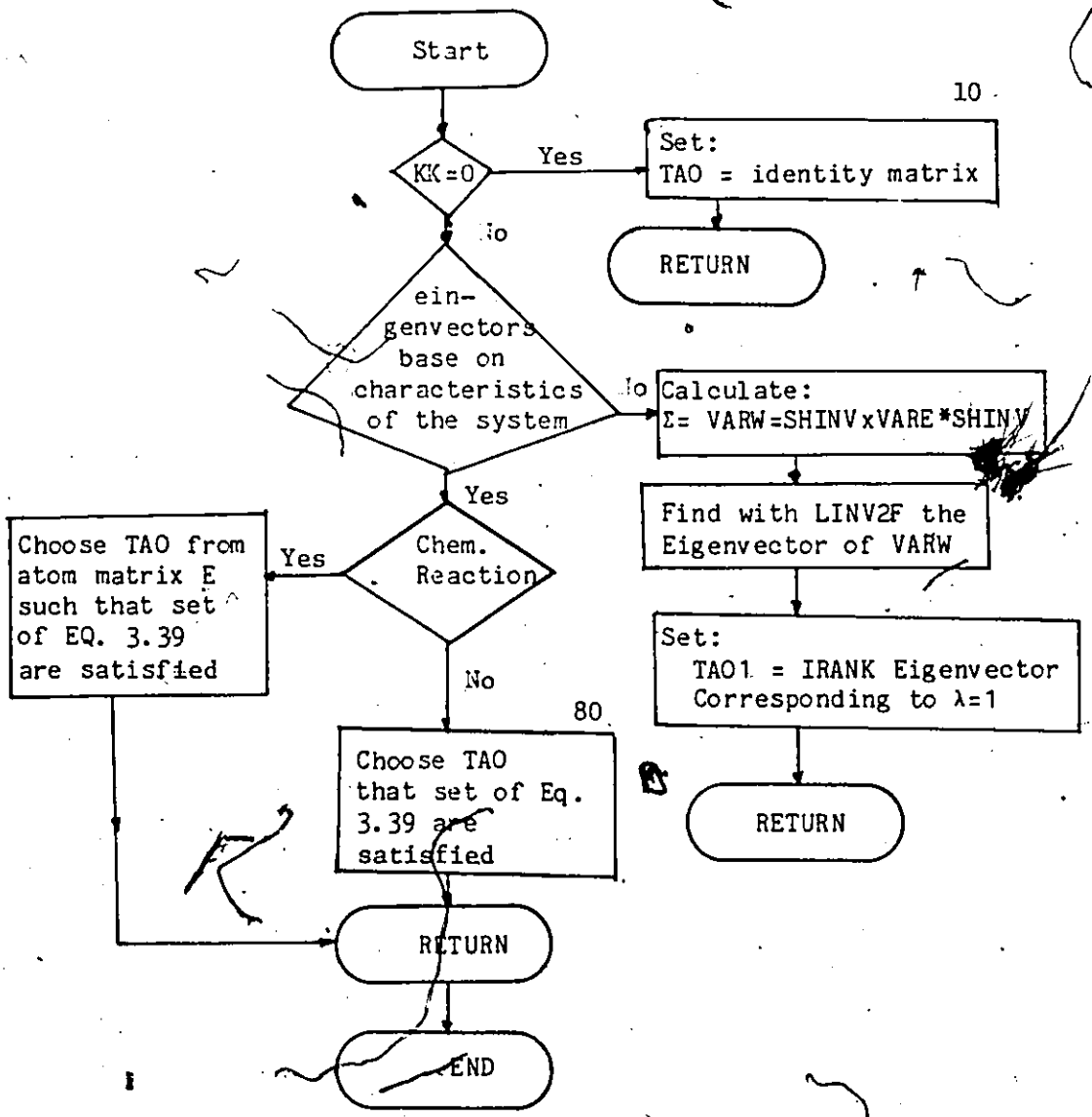


FIGURE 4.7 Subroutine CALCJT Flowchart.

equal one. The dimensions must be: VARE(N,N), A(N,KK), SHINV (N,N) and TAO1(N,IRANK). Any other work arrays needed must be defined as necessary in the subroutine. An example is given in Appendix V and the flow-chart in Fig. 4.7. This flow-chart may be used as a guide to write the subroutine.

ESCRIB writes the final output of the program: the adjusted measurement vector, and the estimates of missing variables and extents of reactions (see Fig. 4.8).

Other details of the program such as parameters definition and dimension requirements are given in the complete listing in Appendix V. It should be noted that the input arrays must be re-dimensioned according to the size of the system. The main program gives the instruction for dimensioning.

In this chapter, the computer program for the method developed in the previous two chapters was described. In the next chapter, the method is tested on four data sets taken from the literature. In chapter VI, the conclusions and significance of this work will be discussed.

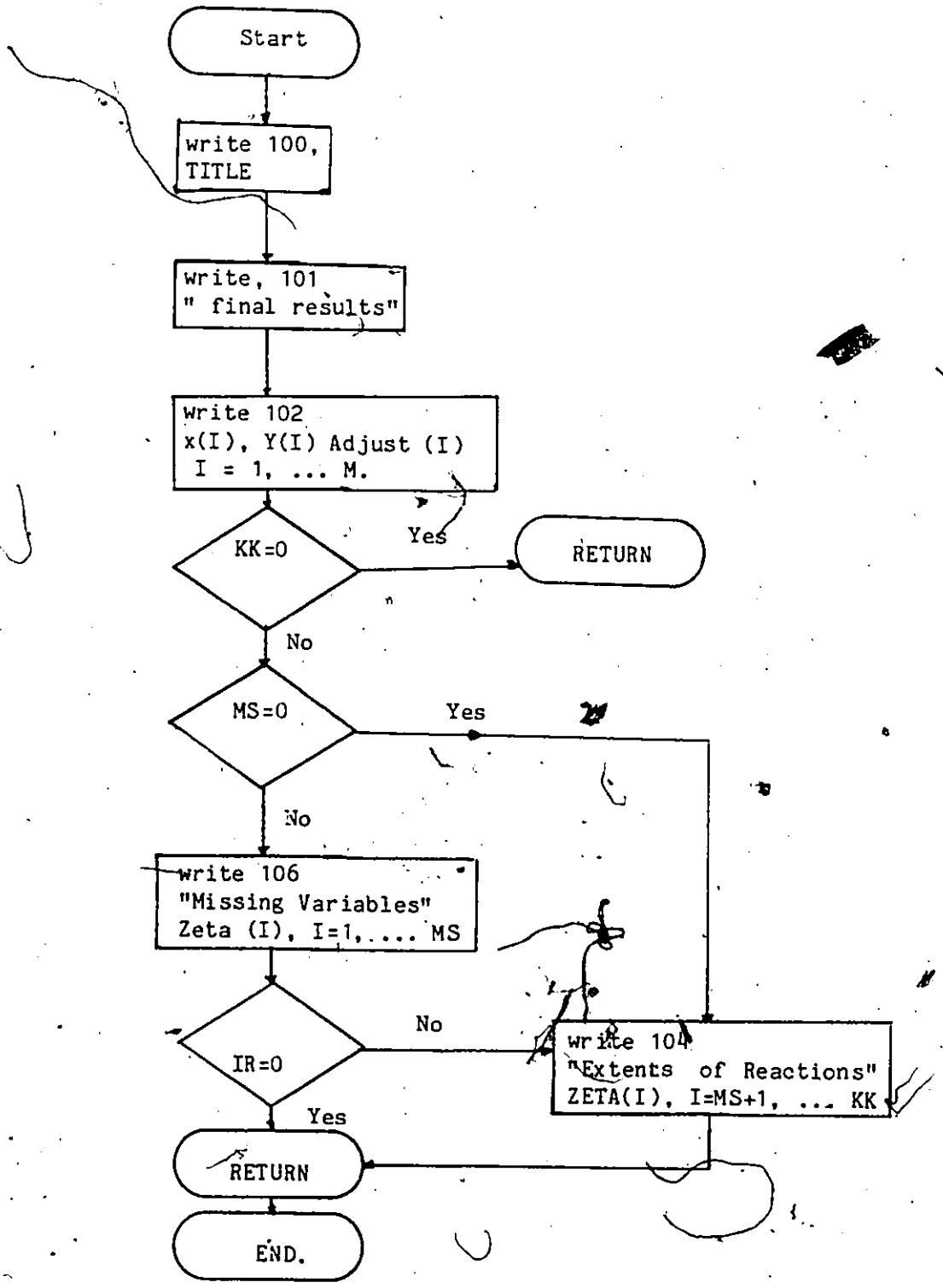


FIGURE 4.8 Subroutine ESCR1B Flowchart.

CHAPTER V

APPLICATION OF THE TECHNIQUE OF DATA ADJUSTMENT

5.1 INTRODUCTION

As discussed in the literature review, previous techniques of data adjustment were unable to adjust process data obtained from both single and multiple unit processes. The examples discussed below were chosen to demonstrate that the technique proposed here can be applied to both cases. Furthermore, the examples demonstrate the applicability of the technique to each of the data adjustment problems discussed in section 3.4.4, viz. biased data, missing variables and chemical reactions.

5.2 CASE STUDIES:

To demonstrate the applicability of the technique presented in this study, four examples from the literature were tested. Flow diagrams for these four processes are shown in Figures 5.1 to 5.3.

In the application of this technique one problem consistently encountered was the lack of availability of the covariance matrix of the data. When the covariance matrix was not available from the literature, a diagonal covariance matrix with elements equal to one was assumed. In almost all of the examples discussed below, the variances of the measurements were known but covariances were not available, therefore a

diagonal covariance matrix with these known elements was used. A basic consideration in the ability of this method to detect biased data is the value of variances of elements. An error would be considered a gross error only if that corresponding measurement falls outside the interval $|x_i| \leq \mu_i \pm 1.65\ \sigma_i$.

Case study I consists of two examples. In example A, the feasibility of the technique to detect and identify gross error caused by leaks in a flow system is evaluated. In example B, Ripps' algorithm is compared with the technique proposed here by using data first tested by Ripps. This example demonstrates the advantages of this technique over Ripps' method in detecting and identifying gross error.

In case study II, the method is applied to data containing missing variables from a multiple unit process.

Case study III is an application of the technique to data obtained from a single processing unit with chemical reaction.

In case study IV, a second test was applied to the data of case study III, after removing one measurement and purposely biasing another.

5.2.1 Case Study I:

Two examples involving multicomponent mass balance at steady-state are discussed. In Figures 5.1a and 5.1b, the flow diagrams of these systems are shown. In Table 5.1 the characteristics of these examples are summarized.

Example A. With this example two features of this technique are discussed: (a) detection of bias, and (2) identification of sources of

TABLE 5.1

Summary of Examples for Case Study I.

	Example A	Example B
System	Pump	Chemical Reactor
Literature Source		Ripps (1965)
Purpose	Detection of bias and identification of sources of bias.	Compare the proposed technique with Ripps' Method
Size of the System:		
Constraints	3	3
Measurements	6	4
Missing Variables	none	none
Chem. reaction	none	reaction unknown
Structure of the system: Coefficient Matrices	$B_1 = [I_3 \vdots -I_3]$	B_1, Given in Table 5.1d. All elements of B_1 are non-zero.

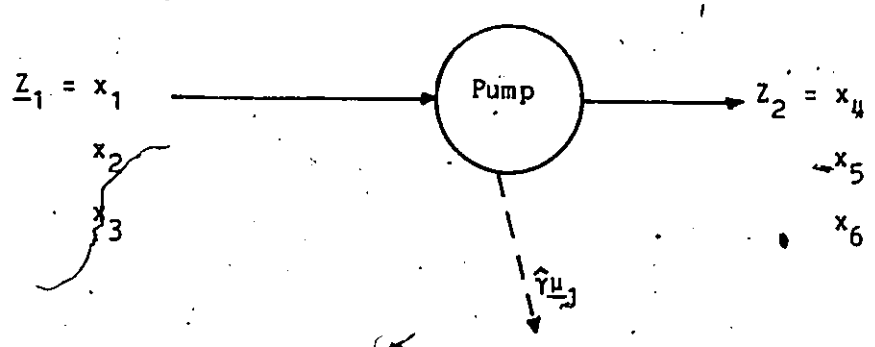


FIGURE 5.1a - Flow Diagram For Example A, Case Study I.

- - - - - - - - fictitious stream

$\hat{\gamma}$ is the ratio of the leak to the true values of $\underline{Z}$.

bias.

Consider the mass balance of a pump (Fig. 5.1a), where $\underline{Z}_i$ represents the input stream (i = 1), and output stream (i = 2). Each stream has the same three components. The true values of elements of $\underline{Z}_i$ ($\underline{\mu}_i$, i = 1,2), were assumed known and consistent. The measurement vector was generated by adding errors of known magnitude to the true values. These errors fell within an interval of ± 30% of the true values and were assigned standard deviations of 5% of the measured flow rate. This information is summarized in Table 5.1a. The covariances of the flow rates (x_i) were assumed zero. The balance coefficient matrix

TABLE 5.1a

Data For Illustrative Example A, Case Study I

COMPONENT (i)	INLET			OUTLET		
	Flow Rate	% Relative error*	s_i	Flow Rate	% Relative error*	s_i**
1	10.00	5.26	5%	7.50	21.05	5%
2	10.00	-4.76	5%	9.50	-9.52	5%
3	99.00	- .70	5%	100.00	.3	5%

* Relative error = $(x_j - \mu_j)/\mu_j$; j = 1, ... 6 and

$\underline{\mu}_1 = \underline{\mu}_2 = (9.50; 10.50, 99.70)^\circ$

** percentage of measured flow rates.

for this case is:

$$B_1 = [I_n \mid -I_n]$$

with n equal to the number of components.

Application of the technique detected the existence of bias in the data (10% significance level), since ($\underline{w}'\underline{w}$ = 8.068) > ($\chi^2(.90,3)$ = 6.251). Source of bias was identified as the existence of gross measurement error. The two-tailed test at α equal to .10 on each

TABLE 5.1b

Results Of Hypothesis Testing For
Example A, Case Study I.

i	without Leak		with simulated Leak*	
	w_i	e_i	w_i	e_i
1	4.	+ 2.5	3.2	+ 2
2	.72	+ .5	1.72	+ 1.2
3	.14	- 1.	2.02	+14.43
$\chi^2(.90,3)$**	16.54		16.70	

* $\underline{Z}_2 = (1-\gamma)\underline{\mu}_1$ = (8., 8.80, 84.57)'; γ = .15

** $\chi^2(.90,3)$ = 6.251

element of $\underline{w}$ resulted in only w_1 lying outside the 90% confidence region. As stated in the theory, section 3.4, this implies that either the inlet or outlet mass flow rate of the first component of the measurement vector could contain gross error. In Table 5.1b, the results of these tests are given. As seen in Table 5.1a, a relative error of 21% of the true value was in fact present in $Z_2(1)$. However, the data adjustment technique informed the analyst, in this case, only that a gross error existed in either or both $Z_1(1)$, $Z_2(1)$. At this point, the analyst should examine the process and the method of measurement, and look for information which will aid him in identifying which measurement(s) contains gross error. Ripps' method was tested and arrived at the same conclusion but only after three iterations of his algorithm. The algorithm of Mah et al is applicable in this case, since the system can be visualized as consisting of multiple units. The method was tested with this data and was successful in detecting the presence of gross error. Like the other methods, the location of the error was determined to be within a subset of two measurements.

In general, when biased data are detected it can be a result of incomplete information on the system (ie, Leaks), and/or the existence of gross measurement errors. As discussed in section 3.4, incomplete information on the system will result in all elements of $\underline{w}$ failing the standard two-tailed test.

By using the same flow processing unit as before, but with simulated "leak" equal to 15% of the true values of the inlet flow-rates (see Fig. 5.1a), the ability of the proposed technique to detect the

possibility of leaks was shown. The results in Table 5.1b verify that all elements of $\underline{w}$ fail the two-tailed test if there is a leak in the system. This result can also mean that all measurements contain gross error. Since this event is unlikely and all elements of $\underline{e}$ are greater than zero, the bias is attributed to a leak.

The leak was created for different values of $\hat{\gamma}$ and it was discovered that for values of $\hat{\gamma}$ less than .15, the method indicated that the biased data was a result of one or two gross measurement errors. However, this result can not be extended to all systems, since the sensitivity of this method to detect gross errors depends on the value of the variances. The larger the variances, the larger the errors which are permissible.

Example B. With this example the proposed technique is compared to Ripps' method in terms of efficiency of detection of gross errors. The same data tested by Ripps (1965) are considered. Four mass [illegible] are measured, two entering and two leaving a chemical reactor [illegible] 5.1b). The data are given in Table 5.1c and the elements of coefficient balance matrix are given in Table 5.1d. The chemical reactions taking place are unknown.

Application of the proposed technique detected bias in the data. The test of bias required only the evaluation of the expression $\underline{e}'H^{-1}\,\underline{e}$, ($\underline{e}'H^{-1}\,\underline{e} \leq \chi^2(.90, 3)$). This test informs the analyst of the presence of bias, whereas Ripps' method requires him to have some knowledge of the system. Since all elements of the coefficient balance matrix are different from zero it is not possible to identify, by the application

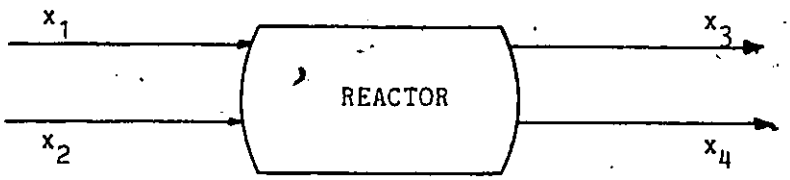


FIGURE 5.1b Flow Diagram for Ripps Example

TABLE 5.1c

Data Tested by Ripps, Case Study I

Variable (i)	Flow-rate (10^3 lb mol/hr) x_i	μ_i	True Relative error*	(Variance)$^{1/2}$ s_i
1	.1850	.1739	.7	.017
2	4.7935	5.0435	-5.0	.05
3	1.2295	1.2175	.5	.024
4	3.8800	4.0000	- .6	.2

*Relative Error = $(x_i - \mu_i)/s_i$

TABLE 5.1d.
Coefficient Balance Matrix For
Ripps Example, Case Study I.

Constraint (i)	Variables 1	2	3	4
1	.1	.3	-.6	-.2
2	.1	.6	-.2	-.7
3	.8	.1	-.2	-.1

TABLE 5.1e

Chi-Square Test . Example B, Case Study I*

Variable rejected (i)	$\underline{w}'\underline{w}$	Suspected measurement in error
0	8.4551	1, 2, 3, 4
1	7.3568	2, 3, 4
2	.9634	2
3	1.5696	3
4	8.4374	1, 2, 3,

* $\chi^2(.90,3) = 6.251$ and $\chi^2(.90,2) = 4.605$

TABLE 5.1f

Adjusted Set of Data For Ripps Example. Case Study I.*

Variable (i)	Adjustment (a_i)	estimated Value y_i
1	.0107	.1751
2		5.0775
3	.0039	1.2256
4	-.1470	4.0270

* x_2 is discarded as measurement contains gross error of standard two-sided test on $\underline{w}$, the subset of $\underline{x}$ suspected of containing the bias. Eventually, this subset of $\underline{x}$ will show up in subsequent chi-square test after sequentially discarding measurement(s). The subset will consist of the discarded measurement(s) when the remaining data pass the chi-square test. For example, if after discarding x_2 the remaining data pass the test, then x_2 is suspected of having gross error. In this example it occurred that other measurements, when removed one at a time, allowed the remaining data to pass the chi-square test. In this case, if the analyst does not have any further information to aid him in exactly locating the gross error, it is suggested to choose as the most likely those measurements which, when removed, give the smallest value of $\underline{e}' H^{-1} \underline{e}$ which passed the test. In

Table 5.1e, it is seen that x_2 is the most probable measurement having gross error. Table 5.1f gives the results of data adjustment when x_2 is discarded. The maximum number of measurements which can be deleted at one time is fixed by the difference between the number of constraints and the measurements, since the rank of B_1 must be less than or equal to the number of constraints. Here, only one measurement can be discarded.

In this example, the proposed technique and Ripps' method both require an iterative search to locate the measurement containing gross error. However, the proposed technique has an advantage as it uses reduced size matrices and fewer matrix operations; each step requires only an evaluation of $\underline{e}' H^{-1} \underline{e}$ whereas Ripps' method requires adjustment of the data each time. As a result the computation time for this technique is considerably less. For this example the evaluation of $\underline{e}' H^{-1} \underline{e}$ required 10 matrix multiplication and 2 inverses at each iteration. The sizes of the matrices involved were 3x3, 3x1 and 3x4 (nxn, nx1 and nxm)#. Ripps' method required in this case, 16 matrix multiplication and 3 inverses at each iteration. The matrix sizes were 7x7, 7x3 and 7x1 (kkxn, kkx1 and kkxkk where kk = m+n-ms).

In summarizing, this example is a case where a complete search of the measurement vector was required to locate the source of bias, since all elements of B_1 were non-zero. Normally, if bias is suggested by the chi-square test and some elements of B_1 are zero, the subset of the measurement vector, containing the bias, will be indicated by subsequent

---

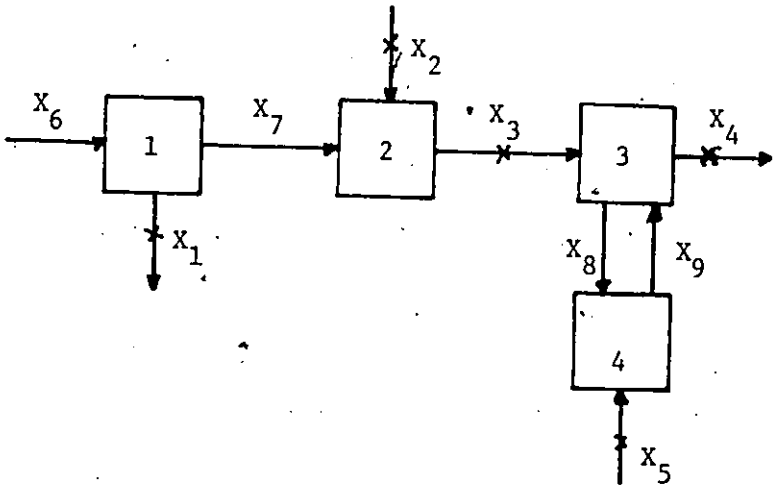
#Calculational details are described in section 3.42

application of the standard two-sided test on each element of the test function vector $\underline{w}$ (see section 3.42). Ripps' method gives no initial indication of the presence or absence of gross error; a search of the measurement vector must be made to reach a conclusion in this regard. For the proposed technique, in the event that bias is not indicated by the chi-square test, the adjustment of the data will be carried out immediately. Clearly the proposed technique allows for a large saving in computational effort over Ripps' method. In Appendix VI, a listing of the program for Ripps' method is given.

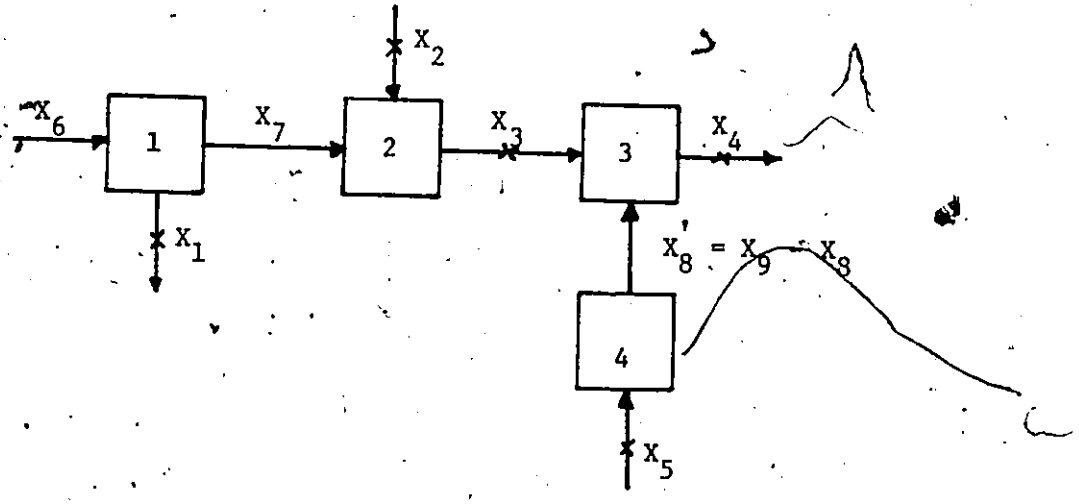
The algorithm of Mah et al can not be used for this system because the structure of B_1 did not allow construction of the coaptation graph required by the method.

5.2.2 Case Study II:

One of the points discussed in this thesis is the adjustment of data when not all process variables are measured. The following example (Vaclavek (1968)), illustrates this aspect of the technique. In Fig. 5.2a the flow diagram of this simple process is shown and Table 5.2 summarizes its characteristics. It consists of four processing units with five measured variables and four missing variables. Two of the four missing variables (x_8 and x_9) form a cycle and are undeterminable. These two variables will be redefined as a new variable x_8' as shown in



a) Original Flow Diagram



b) Modified Flow Diagram

FIGURE 5.2 Flow Diagrams of Example for Case Study II

* Measurements

→ Missing Variables

TABLE 5.2

Summary of Example of Case Study II

System:	Simple process where total flowrate are measured
Literature Source:	Vaclavek (1969)
Purposes	Detection of biased data when some variables are missing and estimation of the missing variables simultaneously with the adjustment of the data
Size of the System: Constraints: Measurements: Missing Variables: Chem. reactions:	 4 5 3 none
Structure of the System: Coefficient matrices:	B_1 and B_2 given in Table 5.2c

Fig. 5.2b, for the purpose of computation. A negative estimate of x_8' implies that x_8 is greater than x_9. It was assumed that all measurements have the same relative error. The information on the measurement vector is summarized in Table 5.2a and the coefficient balance matrix is given in Table 5.2b.

By applying the technique, it was found that there is no sufficient evidence to reject the data as biased; since $\underline{w}'\underline{w} \leq (\chi^2(.90,1) = 2.706)$. The adjusted data are given in Table 5.2c.

Table 5.2a

Data For Case Study II

Variable (i)	Measurement (x_i)	standard deviation ($0.1x_i$)
1	10.	1.
2	20.	2.
3	50.	5.
4	160.	16.
5	100.	10.

5.2b

Incidence Matrix For Case Study II

Unit (i)	Stream j; j = 1, ... 8							
	$B_1(i,j)$					$B_2(i,j)$		
	1	2	3	4	5	6	7	8
1	-1.					1.	-1	
2		1.	-1.				1.	
3			1.	-1				1.
4					1.			-1.

TABLE 5.2c
Adjusted And Estimated Data
Case Study II

Variable (i)	Adjustment (a_i)	Estimated (y_i)
1	0.00	10.00
2	0.00	20.00
3	-.65	50.65
4	6.71	153.28
5	-.262	102.62
6		40.66
7		30.66
8		102.62

Note that in this specific problem, if bias is detected, the technique would reduce the search for error to the subset of (x_3, x_4 and x_5); since the rank of the covariance matrix is one. In other words, the structure of the system and the available information only allow the technique to detect the gross errors in one unit or balance envelope. The balance envelope is chosen according to the information available. The computed value of $\chi^2(.90,1)$ for the chosen envelope will lead to the same conclusion as before.

Again, an aspect which must be considered is that the sensitivity of this method in detecting gross errors increases as the measurement variances decrease. In the example discussed here, the variances of variables x_4 and x_5 are assumed so large that a large error on these measurements will be considered as random. The program was run again assuming reduced values of those variances, and in this case bias was

detected. This observation stresses the importance of not overestimating the variance when it is unknown.

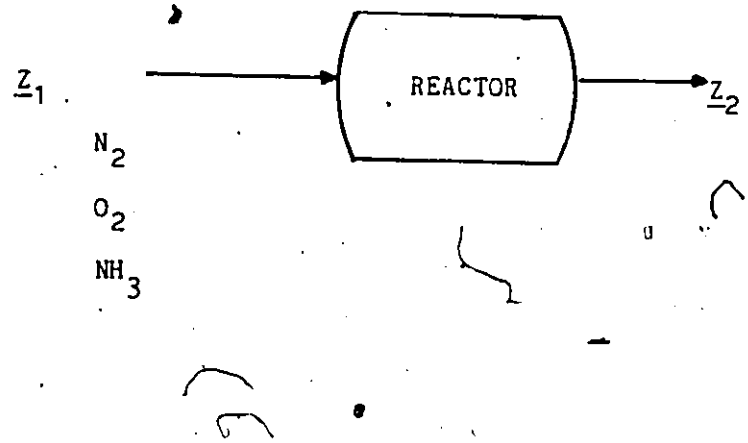


FIGURE 5.3 Flow Diagram of Case Study III

5.2.3 Case Study III:

The purpose of the following example (Murthy (1974)), is to show the applicability of this technique when known chemical reactions are taking place with and without missing variables and gross errors. Table 5.3 summarizes the characteristics of this example.

Consider the mass balance of a reactor (Fig. 5.3), where the catalytic oxidation of ammonia using air occurs. Six chemical substances are taking part in this process: Nitrogen (1), Oxygen (2), Nitrogen monoxide (3), Nitrogen dioxide (4), Water (5) and Ammonia (6). In the absence of the catalyst, the oxidation of ammonia can be described, at least, by the following three linearly independent

stoichiometric equations:

$$4NH_3 + 5O_2 = 4NO + 6H_2O \quad (6.1a)$$

$$4NH_3 + 7O_2 = 4NO_2 + 6H_2O \quad (6.1b)$$

$$4NH_3 + 3O_2 = 2N_2 + 6H_2O \quad (6.1c)$$

TABLE 5.3

Summary of Example for Case Study III

System	Chemical Reactor, Catalytic oxidation of ammonian using air
Literature Source:	Murthy (1974)
Purposes:	Detection of bias in the presence of chemical reactions
Size of the System:	
Constraints:	6
Measurements:	12
Missing Variables:	none
Chem. reactions:	2
Structure of the System: Coefficient matrices:	B_1 and S given in Table 5.2d

Since the rank of the stoichiometric coefficient matrix, S, must be less or equal to the difference between the number of components involved and the rank of the atom matrix, E, ($\rho(S) \leq (n - \rho(E) = 3)$, see Aris (1965))

TABLE 5.3a
Data For Case Study III

Component (i)	Inlet		Outlet	
	Flowrate	s_i	Flowrate	s_i
1	80.00	1	79.00	1
2	20.00	1	6.00	1
3	0.00	1	8.20	1
4	0.00	1	1.7	1
5	0.00	1	15.8	1
6	10.00	1	0.00	1

TABLE 5.3b
Elements of Matrices B_1 and S' For Case Study III

Component (i)	$b_1(i,j)$; j = 1, 2 ... 12												S'(i,j) j = 1,2	
	1	2	3	4	5	6	7	8	9	10	11	12	1	2
1	1.						-1.							
2		1.						-1.					-5.	-7.
3									-1.				4.	
4										-1.				4.
5											-1.		6.	6.
6						1.							-4.	-4.

and the catalyst is known to suppress the reaction represented by equation 6.1c, therefore the catalytic oxidation of ammonia can be described by the first two above equations. The corresponding stoichiometric and balance coefficient matrices are given in Table 5.3b. The data available for all six components taking part in the system are given in Table 5.3a. Covariance is assumed zero. Since, all components require a position in the measurement vector, those components known to be absent in each stream are assigned flowrates of zero. The variances of these positions are assigned any convenient value as they are not significant in the data adjustment calculation, since the elements of the matrix B_1 corresponding to those components are equal to zero. To avoid ill conditioning of the covariance matrix, values close to actual variances of the measurements were chosen. A value of one was found to be satisfactory.

The value of χ^2 statistic associated with the raw data was calculated to be .8024, which is less than the tabulated χ^2 statistic, $\chi^2(.90,4) = 7.779$. Therefore, it can be concluded that the data are not biased. The results of the data adjustment are shown in Table 5.3c. These values are in good agreement with those estimated by Murthy (1974). However, it should be recalled that Murthy's method does not test for gross errors and implicitly accepts that only small random errors are present in the data.

5.2.4 Case Study IV:

Here, the proposed technique is applied to an example known to contain the three data adjustment problems discussed above: missing

variables, chemical reaction and measurement vector with gross error. The characteristics of the example are given in Table 5.4.

The objective is to demonstrate the ability of the method to detect and isolate gross errors while simultaneously treating the problems of missing variables and unknown reaction extents. The data were a modified version of those used in case study III. The measurement

TABLE 5.3c

Adjusted Set of Data, Case Study III

Component (i)	Inlet		Outlet	
	Flowrate	adjust.	Flowrate	adjust.
1	79.500	.500	79.5	-.500
2	19.939	.061	6.061	-.061
3	0.000	0.000	8.400	-.200
4	0.000	0.000	1.931	-.231
5	0.000	0.000	15.496	.305
6	10.331	-.331	0.000	.000

corresponding to the outlet of the second component (x_8) was removed and considered a missing variable. Also, the measurement corresponding to the inlet of the first component (x_1) was purposely biased. These new data are given in Tables 5.4a and 5.4b.

The results from the application of the technique indicate that the

TABLE 5.4

Summary of Example for Case Study IV

System	Chemical reactor. Catalytic oxidation of ammonia using air
Literature Source:	Modification of the data adjusted by Murthy (1974). X_8 is missing and X_1 purposely biased.
Purposes:	Detection of bias considering missing variables and chemical Reactions.
Size of the System:	
Constraints:	6
Measurements:	11
Missing Variables:	1
Chem. Reactions:	2
Structure of the system: Coefficient matrices:	B_1, B_2 and S given in Table 5.2e

data are biased, since ($\underline{w}'\underline{w}$ = 200.296) > ($\chi^2_{(.90,3)}$ = 6.251). From the results of the normal two-sided test performed on the test function $\underline{w}$ (Table 5.4c), the source of bias was identified as one of the measurements contributing to the first element of $\underline{e}$. As in example I.A., the technique was only able to localize the error in the subset (x_1, x_7).

In order to evaluate the method when the measurement containing gross error has been removed, x_1 was removed and the data re-tested. The results are given in Table 5.4d, and were compared with those in Table 5.3c from case study III. The adjusted data and estimates of $\underline{u}$ were found to be in good agreement.

The Case Studies described above have demonstrated the utility of the proposed data adjustment technique in detecting and localizing sources of bias for various system configurations and classes of data adjustment problem. The failure in identifying uniquely the source of gross error results because the source of error is inaccessible solely with balancing of the process.

TABLE 5.4a
Data for Illustrative Example
of Case Study IV*

Component (i)	Flowrate	
	Inlet	Outlet
1	99.00	79.00
2	20.00	missing
3	0.00	8.20
4	0.00	1.70
5	0.00	15.8
6	10.00	0.00

* The values of s_i are the same as in Table 5.1d; $i = 1, \ldots 12$ and $i \neq 6$

TABLE 5.4b

Elements of Matrices B_1, B_2 and S' For Case Study IV

Component (i)	b_1 (i,j); j = 1, ... 11											b_2(i,1)	S'(c,J) j = 1,2	
	1	2	3	4	5	6	7	8	9	10	11	1	1	2
1	1.						-1.							
2		1.										-1.	-5.	-7.
3								-1.					4.	
4									-1.					4.
5										-1.			6.	6.
6						1.							-4.	-4.

TABLE 5.4c

Test Function $\underline{w}$ for Case Study IV

i	$\underline{e}$**	$\underline{w}$*
1	20.00	14.72
2	-.3400	.40
3	-.3066	.145

* $|w_i| \leq$ 1.645 90% confidence region.

** The 3-largest values of the vector (Abs(e_i))

TABLE 5.4d

Results of Data Adjustment Technique

For Case Study IV, after removing x_1

Component (i)	Inlet		Outlet	
	adjust	flowrate	adjust	flowrate
1		79.000	0.000	79.000
2		20.000		6.165
3	0.000	0.000	-.2033	8.4033
4	0.000	0.000	-.2033	1.9033
5	0.000	0.000	.34	15.46
6	-.3067	10.3067	0.000	0.000

A summary of conclusions and significance of this work, as well as recommendations for future work, are presented in the next chapter.

CHAPTER VI

CONCLUSIONS AND RECOMMENDATIONS

6.1 CONCLUSIONS

In the literature, the general problem of data adjustment has been discussed as several separate problems; such as, adjustment of data having small errors only (Kuehn & Davidson (1961)); adjustment of data containing gross errors (Ripps (1965) plus development of a statistical test function for identifying gross error (Nogita (1972)); adjustment of data with missing variables (Vaclavek (1968)) plus derivation of statistical criterion for detecting and isolating the source of bias (Mah et al (1975)); and analysis of data from chemical reactors (Murthy (1977) & Vaclavek (1969)). This study was devoted to developing a technique for statistical analysis and adjustment of inconsistent process data which deals with all the data adjustment problems named above.

The main contributions and conclusions resulting from this study are summarized below:

(a) An improved statistical criterion for detecting biased data has been developed and successfully applied to examples from the literature.

The statistical test for bias is based on the chi-square distribution and has several advantages over the test developed by Nogita. First, it is applicable in cases having missing variables; since this condition was included in its development. Second, it examines the entire measurement vector before any adjustment is made. Nogita's method adjusts the data

before testing for bias and therefore bias might pass unnoticed due to it being spread over the data set. Therefore, the test developed here is more sensitive to gross error detection. And third, Nogita takes for the test function a linear combination of all adjustments while a quadratic form of e is used by this new test. Under Nogita's test the gross error could pass unnoticed because of error cancellation or because its effect is attenuated after adjustment of the data.

(b) A general algorithm has been developed for implementing the bias identification strategy in Chapter III. The effect of bias can be traced and isolated by using a separate statistic for each balance constraint. This algorithm is similar to that developed by Mah et al (1976), but is not limited to one component mass balance systems.

(c) This technique can be applied to multiple processing unit systems as well as single unit systems. Also data from chemical reactors, where the chemical reactions taking place are known, can be analysed and adjusted by this method.

The most relevant weaknesses of this technique are: (a) the difficulty of localizing the measurements containing gross error when all the measured variables contribute to each element of the imbalance vector. However, the serial elimination algorithm suggested here is more efficient than Ripps and Nogita's in terms of computation time and sensitivity to gross error detection; (b) it does not assure that the estimated extents of irreversible reactions and the adjusted and estimated values of the process variables will be posititve. A negative estimated extent where the reaction is reversible implies that the reaction proceeds in the reverse

direction. Therefore, it is recommended that the matrix S be constructed using the maximum number of reversible reactions. The occurrence of negative value of process variables is very rare, since the Least Square weights are the elements of the covariance matrix.

6.2 RECOMMENDATIONS FOR FUTURE WORK

Any further studies done in the analysis of mass/energy balances of chemical processes should consider the following recommendations: (a) Other methods for minimizing the objective function should be examined. Solution by Linear Programming, for example, will ensure positive estimates for extents of reactions and process variables; (b) Alternative formulation of the objective function should be investigated, with the goal of reducing the problem to an unconstrained one; (c) Another extension of the present study is adjustment and estimation of the concentration of a component and/or enthalpy of the streams simultaneously with its total flow rate. This yields a non-linear constraint problem, a situation quite common in practice; (d) The current study has been limited to steady-state operation. The extention of this technique to dynamic systems would be an important contribution. (e) Incorporating a method for selecting the optimal balance envelope into the present technique will provide certain benefits. The size of the system will be reduced, and the amount of computation will be decreased.

REFERENCES

1. Amundson, N.R. (1966), "Mathematical Method in Chemical Engineering", Prentice-Hall, Englewood Cliffs, N.J.

2. Anderson, T.W. (1958), "An Introduction to Multivariate Statistical Analysis", John Wiley and Sons, Inc., New York.

3. Aris, R. (1966), "Introduction to the Analysis of Chemical Reactors", Prentice-Hall, Englewood Cliffs, N.J.

4. Bellman, R. (1970), "Introduction to Matrix Analysis", McGraw Hill Book Co., New York.

5. Crowe, C.M. (1979,1980), Private Communication.

6. Guttman, I. and Wilks, S.S. (1967), "Introductory Engineering Statistics", John Wiley and Sons, Inc., New York.

7. Kuehn, D.R. and Davidson, H. (1961), "Computer Control II. Mathematics of Control", Chem. Eng. Prog., Vol. 57, No. 6, 44.

8. Madron, F. and Vanecek, V. (1977), "Statistical Adjustment of Material Balance of a Chemical Reactor", Collection Czechoslov. Chem. Commun., Vol. 42, 1805.

9. Madron, F., Veverka, V. and Vanecek, V. (1977), "Statistical Analysis of Material Balance of a Chemical Reactor", AIChE J. Vol. 23, No. 4, 482.

10. Mah, R.S., Stanley, G.H. and Downing, D.M. (1976), "Reconciliation and Rectification of Process Flow and Inventory Data", Ind. Eng. Chem., Process Des. Dev., Vol. 15, No. 1, 175.

11. Meharg, E.B. (1972), "Using Data Adjustment", Instruments and Control Systems, Vol. 45, No. 7, 51.

12. Murthy, A.K.S. (1974), "Material Balance Around a Chemical Reactor. II.", Ind. Eng. Chem., Process Des. Dev., Vol. 13, No. 4, 347.

13. Nogita, S. (1972), "Statistical Test and Adjustment of Process Data", Ind. Eng. Chem., Process Des. Dev., Vol. 11, No. 2, 197.

14. Ripps, D.L. (1965), "Adjustment of Experimental Data", Chem. Eng. Prog. Symposium Ser., Vol. 61, No. 55, 8.

15. Stanley, G.M. (1977), "Estimation and Observability in Process Networks", Ph. D. Thesis, Northwestern University, Chem. Eng.

16. Umeda, T., Nishio, M. and Komatsu S. (1974), "A Method for Plant Data Analysis and Parameters Estimation", Ind. Eng. Chem., Process Des. Dev., Vol. 10, No. 2, 236.

17. Vaclavek, V., Kubicek, M., Hlavacek, V. and Marek, M. (1968), "Studies on System Engineering, I. Calculation of Material and Enthalpy Balances on a Computer", Collection Czechoslov. Chem. Commun., Vol. 33, 3653.

18. Vaclaveck, V. (1969), "Studies on System Engineering, II. On the Application of the Calculus of Observations in Calculations of Chemical Engineering Balances", Collection Czechoslov. Chem. Commun., Vol. 34, 364.

19. Vaclavek, V. (1969), "Studies on System Engineering, III. Optimal Choice of the Balance Measurements in Complicated Chemical Engineering Systems", Chem. Eng. Sci., Vol. 24, 947.

20. Vaclavek, V. (1969), "A Note to the Problem of Adjustment of Material Balance of Chemical Reactor", Collection Czechoslov. Chem. Commun., Vol. 34, 2662.

APPENDIX I

CALCULATION INVOLVING PARTITIONED MATRICES

Consider the equation:

$$\underline{z} = -(A' H^{-1} A)^{-1} A' H^{-1} B_1 \underline{x} \qquad (I.1)$$

where $\underline{z} = \begin{bmatrix} \underline{u} \\ \hline \underline{\xi} \end{bmatrix}$ and $A = [\, B_2 \,|\, S' \,]$

In order to solve equation I.1 for $\underline{u}$ and $\underline{\xi}$, substitute $\underline{z}$ and A into I.1:

$$\begin{bmatrix} \underline{u} \\ \hline \underline{\xi} \end{bmatrix} = -\left[\begin{bmatrix} B_2' \\ \hline S \end{bmatrix} H^{-1} [B_2 \,|\, S']\right]^{-1} \begin{bmatrix} B_2' \\ \hline S \end{bmatrix} H^{-1} B_1 \underline{x} \qquad (I.2)$$

Rearranging I.2

$$\begin{bmatrix} \underline{u} \\ \hline \underline{\xi} \end{bmatrix} = -\begin{bmatrix} B_2' H^{-1} B_2 & B_2' H^{-1} S' \\ \hline S H^{-1} B_2 & S H^{-1} S' \end{bmatrix}^{-1} \begin{bmatrix} B_2' H^{-1} \\ \hline S H^{-1} \end{bmatrix} B_1 \underline{x} \qquad (I.3)$$

Let:

$$\begin{bmatrix} B_2' H^{-1} B_2 & B_2' H^{-1} S' \\ \hline S H^{-1} B_2 & S H^{-1} S' \end{bmatrix} = \begin{bmatrix} D_{11} & D_{12} \\ \hline D_{12} & D_{22} \end{bmatrix} \equiv D \quad \text{and}$$

$$D^{-1} = \begin{bmatrix} P_{11} & P_{12} \\ \hline P_{12} & P_{22} \end{bmatrix} = P$$

Then:

$$D P = I \qquad (I.4)$$

Solving equation I.4 gives the following two sets of simultaneous equations:

$$D_{11} P_{11} + D_{12} P_{12}' = I \qquad (I.5a)$$

$$D_{12}' P_{11} + D_{22} P_{12}' = 0 \qquad (I.5b)$$

$$D_{11} P_{12} + D_{12} P_{22} = 0 \qquad (I.6a)$$

$$D_{12}' P_{12} + D_{22} P_{22} = I \qquad (I.6b)$$

Solving the set of simultaneous equations I.5 with $R = (B_2' H^{-1} B_2)^{-1} B_2' H^{-1}$ and $Q = S H^{-1}(I_n - B_2 R)$, we get:

$$P_{21} = - (Q S')^{-1} S R' \qquad (I.7a)$$

$$P_{11} = (B_2' H^{-1} B_2)^{-1} + RS'(Q S')^{-1} S R' \qquad (I.7b)$$

and solving the set 1.6:

$$P_{22} = (Q S')^{-1} \qquad (I.8a)$$

$$P_{12} = - R S'(Q S')^{-1} \qquad (I.8b)$$

Substituting 1.7's and 1.8's into 1.3 and solving the resulting equation:

$$\begin{bmatrix} \underline{u} \\ \underline{\xi} \end{bmatrix} = \begin{bmatrix} R\,(S'\,(QS')^{-1} Q - I_n) \\ -(QS')^{-1} Q \end{bmatrix} B_1 \underline{x} \qquad (I.9)$$

then:

$$\underline{u} = R[S'\,(QS')^{-1} Q - I_n]\, B_1 \underline{x} \qquad (I.10)$$

$$\underline{\xi} = -(QS')^{-1} Q\, B_1 \underline{x} \qquad (I.11)$$

By a similar procedure Equation 3.27 is obtained.

APPENDIX II

PROOF FOR NON-SINGULARITY OF A MATRIX

Let us consider a general vector $\underline{v}$ and a general matrix V which is positive definite. Then defining the quadratic form:

$$q(v) = \underline{v}' V \underline{v} \qquad \text{(II.1)}$$

V is positive definite, if and only if

$$q(\underline{v}) > 0 \; \forall \; \underline{v} \neq \underline{0} \qquad \text{(II.2)}$$

Lemma 1

If there is an m x n matrix D of rank n ($\rho(D) = n \leq m$), then $D' V D$ is positive definite whenever V is positive definite

Proof:

Let

$$\underline{v} = D \hat{\underline{v}} \qquad \text{(II.3)}$$

substituting II.3 into II.1

$$q(\underline{v}) = (D \hat{\underline{v}})' V (D \hat{\underline{v}}) > 0 \; \forall \; \underline{v} \neq \underline{0} \qquad \text{(II.4)}$$

since V is positive definite

Rearranging II.4:

$$q(\underline{v}) = \hat{\underline{v}}' \, (D' \, V \, D) \, \hat{\underline{v}} > 0 \; \forall \; \underline{v} \neq \underline{0} \tag{II.5}$$

From II.5 $\underline{v}$ must be different from zero for all $\hat{\underline{v}}$ different from zero, that is

$$\underline{v} = D \, \hat{\underline{v}} \neq \underline{0} \quad \forall \; \hat{\underline{v}} \neq \underline{0} \tag{II.6}$$

since D has full column rank, that is

$$\rho(D) = n \leqslant m \tag{II.7}$$

Otherwise there exists $\hat{\underline{v}} \neq \underline{0}$ such that $D \, \hat{\underline{v}} = \underline{0}$. This ends the proof of the lemma 1.

It follows from equation II.5 and II.7 that H is positive definite and thus non-singular, if B_1 has a full row rank ($\rho(B_1) = n \leqslant m$). And the matrix $(B_2' \, H^{-1} \, B_2)$ is non-singular only if $\rho(B_2) = ms \leqslant n$.

Theorem 1 (Crowe (1980))

QS' is non-singular matrix provided: (a) no reaction exists solely among components which have missing variables in any stream entering or leaving the balance envelope, and (b) the number of missing variables, ms, is less than or equal to the difference between the number of components n, and chemical reactions, p, (ie., $ms \leqslant n-p$).

Proof

(a) If p x p matrix QS' is non-singular, then the equation

$$QS' \underline{v} = 0 \tag{II:8}$$

has, as unique solution, the trivial solution $\underline{v} = 0$. Otherwise Eq. II.8 has a solution $v \neq 0$.

Let

$$\underline{t} = S' \underline{v} \tag{II.9}$$

Since S' has a full column rank ($\rho(S') = p$), $t \neq 0 \; \forall \; \underline{v} \neq 0$. Substituting Eq. II.9 into Eq. II.8

$$Q\underline{t} = 0 \tag{II.10}$$

The above equation has a solution $t \neq 0$ since the rank of Q is less than or equal to p ($\rho(Q) \leq p$).

$Q = SH^{-1}(I_n - B_2 R)$ was defined in section 3.4, then

$$SH^{-1}(I_n - B_2R)\underline{t} = 0 \tag{II.11}$$

Therefore, a set of solution of Eq. II.11 is

$$\underline{t} = B_2R \, \underline{\alpha} \tag{II.12}$$

for arbitrary $\underline{\alpha}$. Since the rank of $B_2R\underline{\alpha}$ is ms, there is only ms independent vector $\underline{t}$ as defined by Eq. II.12 (dim {t} $\geq$ ms). From Eqs. II.12 and II.9:

$$S'\underline{v} = B_2R\underline{\alpha} \qquad \text{(II.13)}$$

Since B_2 is the incidence matrix corresponding to missing variables, Eq. II.13 represents a chemical reaction existing solely among components which have missing variables in any stream entering or leaving the system. Then

$$QS'\underline{v} = Q\underline{t} = 0 \text{ but } \underline{v} \neq 0 \qquad \text{(II.14)}$$

Therefore, it has been proved that the converse of condition (a) is sufficient for singularity of QS' and therefore condition (a) is necessary for non-singularity of QS'

(b) From properties of rank,

$$\rho(I_n - B_2R) = n-ms \qquad \text{(II.15a)}$$

$$\rho(Q) \leq \text{Min. } (p, n-ms) \qquad \text{(II.15b)}$$

In order that p x p QS' matrix be non-singular

$$\rho(QS') = p, \quad \text{therefore } \rho(Q) = p \tag{II.16}$$

From Eq. 15b, the rank of Q is equal p only if

$$p \leqslant n-ms \text{ or } ms \leqslant n-p \tag{II.17}$$

It is therefore necessary for QS' to be non-singular that condition (b) ($ms \leqslant n-p$) must be hold.

APPENDIX III

PROOF FOR $C = \begin{bmatrix} \Lambda_r^{-1/2} \Gamma_1' \\ \hline \Gamma_2' \end{bmatrix}$

Let C be a non-singular orthogonal matrix such that

$$C \Sigma C' = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] \qquad \text{(III.1)}$$

where r is the rank of the n x n matrix Σ. C exists since Σ is symmetric and idempotent. There exists an (r-n) x n matrix A such that:

$$A \Sigma = 0 \qquad \text{(III.2)}$$

and an r x n matrix $\hat{C}$ such that:

$$C = \begin{bmatrix} \hat{C} \\ \hline A \end{bmatrix} \text{ is non-singular} \qquad \text{(III.3)}$$

Using eq. III.3 in equation III.1 and solving the resulting euqation:

$$\hat{C} \Sigma \hat{C}' = I_r \qquad \text{(III.4)}$$

It is known that

$$\Gamma' \Sigma \Gamma = \Lambda \qquad \text{(III.5)}$$

where Γ is the eigenvector matrix and Λ is the diagonal matrix of eigenvalues for matrix Σ. In this case where Σ is a singular matrix:

$$\Lambda = \left[\begin{array}{c|c} \Lambda_r & 0 \\ \hline 0 & 0 \end{array}\right] \quad \text{and} \quad \Gamma = [\Gamma_1 \,|\, \Gamma_2]$$

with Γ_1 and Γ_2 previously defined in chapter III.

From III.5:

$$\Gamma_1' \Sigma \Gamma_1 = \Lambda_r \qquad \text{(III.6)}$$

and

$$\Gamma_2' \Sigma = 0 \qquad \text{(III.7)}$$

Premultiplying and postmultiplying eq. III.6 by $\Lambda_r^{-1/2}$, we get

$$(\Lambda_r^{-1/2} \Gamma_1') \Sigma (\Gamma_1 \Lambda_r^{-1/2}) = I_r \qquad \text{(III.8)}$$

Rearranging eq. III.8

$$(\Lambda_r^{-1/2} \Gamma_1') \Sigma (\Lambda_r^{-1/2} \Gamma_1')' = I_r \qquad \text{(III.9)}$$

Comparing eq. III.9 with eq. III.4

$$\hat{C} = \Lambda_r^{-1/2} \Gamma_1' \qquad \text{(III.10)}$$

and from eq. III.7

$$A = \Gamma'_2 \qquad \text{(III.11)}$$

Substituting equation III.10 and III.11 into eq. III.3

$$C = \left[\frac{\Lambda_r^{-1/2} \Gamma_1'}{\Gamma_2}\right] \qquad \text{(III.12)}$$

APPENDIX IV

REVIEW OF IDEMPOTENT MATRICES

Idempotent matrices play an important role in the theoretical development of the test function for detecting and identifying gross errors, proposed by this study. Therefore, a brief review of some properties of idempotent matrix are presented here

DEFINITION:

Idempotent matrix is defined as a square symmetric matrix, (D), such that $D^2 = D$. Some of its properties used by the theory of this study are:

Property 1

"If D is n x n symmetric matrix of rank r, then the necessary and sufficient condition that D be idempotent is that r of its eigenvalues are equal to one and the remaining (n - r) are equal to zero".

Proof

Since D = D', there exist an orthogonal transformation, such that:

$$D = \Gamma \Lambda \Gamma' \qquad \text{(IV.1)}$$

where Λ is the diagonal matrix of eigenvalues of D.

Then:

$$D^2 = D * D = (\Gamma \Lambda \Gamma')(\Gamma \Lambda \Gamma') \qquad \text{(IV.2)}$$

$$D^2 = \Gamma \Lambda^2 \Gamma' \qquad \text{(IV.3)}$$

From definition of idempotent matrix:

$$D^2 = D \qquad \text{(IV.4)}$$

Therefore,

$$\Lambda^2 = \Lambda \qquad \text{(IV.5)}$$

Equation IV.5 is true if and only if the elements of Λ are zero or one.

Conversely, if elements of Λ are equal one or zero, then $D * D = D$.

It has been proved the necessary and sufficient condition that the eigenvalues equal one or zero.

Now, it will be proved that the number of the eigenvalues equal one, is the rank of the matrix.

Since D is a matrix or rank r ($r \leq n$), there exists an $n \times (n - r)$ matrix Γ_2 such that:

$$\Gamma_2' D = 0 \qquad \text{(IV.6)}$$

and $n \times r$ matrix Γ_1 such that

$$\Gamma' = [\Gamma_1 \,;\, \Gamma_2]$$

is orthogonal

From equation (IV.1)

$$\begin{bmatrix} \Gamma_1' \\ \hline \Gamma_2' \end{bmatrix} D \begin{bmatrix} \Gamma_1 \vdots \Gamma_2 \end{bmatrix} = \Lambda \qquad \text{(IV.7)}$$

where $\Lambda = \begin{bmatrix} I_p & 0 \\ \hline 0 & 0 \end{bmatrix}$ and p is the number of non-zero eigenvalues.

Solving (IV.7)

$$\Gamma_1' D \Gamma_1 = \begin{bmatrix} I_p & 0 \\ \hline 0 & 0 \end{bmatrix} \qquad \text{(Iv.8)}$$

therefore $p \leq r$. But if $p < r$, D has rank $< r$ since more than $(n - r)$ linear dependencies exist among columns of D. then, p is equal to r.

Then, it can be concluded that r of the eigenvalues of D are equal to one and therefore the remaining $(n - r)$ are equal to zero.

PROPERTY 2

"If D is a symmetric idempotent matrix, then the rank of D is equal to the trace of D."

Proof

The trace of a square matrix is equal to the sum of its eigenvalues:

$$tr\,(D) = \sum_{i=1}^{n} \lambda_i \qquad \text{(IV.9)}$$

From property 1:

$$\lambda_i = \begin{cases} 1 & \text{for } i = 1, \ldots r \text{ with } r = \rho(D) \\ 0 & \text{for } i = r + 1, \ldots n. \end{cases}$$

Then:

$$tr(D) = \sum_{i=1}^{\rho(D)} \lambda_i = \rho(D) \qquad (IV.10)$$

Therefore, the trace of D is equal to the number of eigenvalues equal one, which is the rank of matrix D.

PROPERTY 3

"The only nonsingular symmetric idempotent matrix is the identity matrix"

Proof

From property 1, the n x n matrix D can be expressed as

$$D = \Gamma \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} \Gamma' \qquad (IV.11)$$

Assuming that the rank of matrix D is equal to n, equation IV.11 can be written as:

$$D = \Gamma I_n \Gamma' \qquad (IV.12)$$

Therefore:

$$D = \Gamma \Gamma' = I_n \qquad (IV.13)$$

Then, it can be concluded that the only nonsingular symmetric idempotent matrix is I_n.

PROPERTY 4

"If D is an n x n symmetric idempotent matrix of rank r ($r \leq n$), then D is positive definite if r = n, positive semidefinite if r < n"

Proof

From property 3, if r = n, the matrix D is the identity matrix. I_n is clearly positive definite.

From property 1, the eigenvalues of D are equal to one or zero, ($\lambda_i \geq 0$; i = 1, ... n). Therefore, D is positive semidefinite.

APPENDIX V

PROGRAM LISTING FOR THE DATA ADJUSTMENT TECHNIQUE

V.1 ±ANADATA±, MAIN PROGRAM

```
      PROGRAM ANADATA (INPUT,OUTPUT,TAPE5=INPUT,TAPE6=OUTPUT)
C
C
C*****************************ANADATA*****************************
C
C
C     MAIN PROGRAM FOR ENHANCING  INCONSISTENT PROCESS DATA.
C     WRITTEN BY
C                 Y.A.GARCIA-CAMPOS
C                 DEPART. CHEMICAL ENGINEERING
C                 MCMASTER UNIVERSTTY
C                 HAMILTON-ONTARIO
C                 CANADA
C
C*************************************************************
C
C*GIVEN THE COEFFICIENT  MATRICES FOR MEASUREMENT VECTOR AND MISSING
C*VARIABLE,THE STOICHIOMETRIC COEFFICIENT MATRIX,THE MEASUREMENT VEC
C*TOR,THE COVARIANCE MATRIX AND THE SIZE  OF ALL MATRICES AND VEC-
C*TORS,THIS PROGRAM  DETECTS THE BIAS DATA,OUTPUT THE SUBSET OF MEA-
C*SUREMENT  SUSPECTED OF CONTAINING GROSS ERROR AND ESTIMATE  ALL
C*PARAMETERS STREAMS. THE LIMITS OF THE PROGRAM IS THAT  M.LE.30 AND
C*N.LE.20.THE SIZE OF ALL MATRICES ARE GIVEN BELOW. DIMENSION CAN BE
C*CHANGED  IN THE MAIN  PROGRAM ONLY,BUT IF THE SYSTEM IS OVER THE
C*LIMITS SPECIFIED  BY DIMENSIONS THE COMMON MUST BE CHANGE IN
C*SUBROUTINE RECHAZ TOO.
C
C*FIRST DATA CARD CONTAINS THE  TITLE OF THE SYSTEM.IN 60 CHARACTE-
C*RES,INCLUDED BLANK.
C
C*SECOND CARD CONTAINS  THE VALUES OF N,M,MS,IR IN  FREE FORMAT.
C
C*THIRD DATA SET IS THE ELEMENTS OF MATRICES  B1,B2,ST GIVEN BY CO-
C*LUMNS. 5 VALLES PER CARD. WITH FORMAT F10.0
C
C*THE FORTH SET OF DATA  IS THE VECTOR OF MEASUREMENT. 5 VALUES PER
C*CAR.  WITH FORMAT F10.0
C*THE FIFTH SET IS THE COVARIANCE MATRIX.5 VALUES PER CARD.GIVEN BY
C*COLUMNS. IF THE MATRIX IS DIAGONAL  ONLY THE NON-ZERO ELEMENTS ARE
C*GIVEN.THE PARAMETER IDIA IS SET EQUAL TO ZEROWITH FORMAT F10.0
C
C*************************************************************
C
C*LIST OF PARAMETERS
C*******************
C
C*INPUT
C
C*N=NUMBER OF CONSTRAINS
C*M=NUMBER OF MEASUREMENTS
C*MS=NUMBER OF MISSING VARIABLES
C*IR=NUMBER OF INDEPENDENT CHEM. REACTIONS.
```

```
C     *B=N*(MS+M+IP) VECTOR OF COEFFICIENT MATICES.STORED  BY COLUMNS,
C     * STARTING WITH B1,B2, AND ENDING WITH ST.
C     *XDATOS=M-VECTOR OF MEASUREMENTS
C     *RVAR=M*M-VECTOR OF ELEMENT OF COVARIANCE MATRIX.IF COVARIANCE =0,
C     *      RVAR IS THE M-VECTOR  OF THE SQUARE ROOT OF VARIANCES.
C     *
C     *OUTPUT
C     *
C     *X=M-MEASUREMENT VECTOR EXCLUDING THE ONE HAVING GROSS ERROR
C     *ADJUST=M-ADJUSTMENT VECTOR
C     *ZETA=(MS+IR)- PARTITIONED VECTOR OF MISSING VARIABLES AND EXTENTS
C     *      OF REACTIONS.
C     *Y=M-VECTOR OF ADJUSTED VALUES OF X
C     *UMB=N-IMBALANCE VECTOR
C     *OMEGA=TEST FUNCTION
C     *VARE=N*N COVARIANCE MATRIX OF UMB
C     *IK=INDEX OF MEASUREMENT TO BE REJECTED
C     *MG= NE-VECTOR OF MEASUREMENT TO BE REJECTED
C     *
C     *SPECIAL PARAMETERS
C     *
C     *IFLAG=1,IF B1(I,J)≠0 FOR ALL I AND J.OTHERWISE IFLAG=0
C     *IDIA=1 ,IF COVARIANCE MATRIX OF XDATOS IS NON-DIAGONAL.OTHEWISE
C     *      IDIA=0.
C     *ISW=0,PROGRAM START
C     *   =1, AFTER SEARCH FOR BIAS
C     *KK=IP+MS
C     *IP=N*N+3*N ,ARGUMENT OF IMSL SUBROUTINE,LINV2F
C     *II=N*((N+1)/2), ARGUMENT OF IMSL SUBROUTINE, VCVTFS
C     *IPRINT=1,PRINT  INTERMEDIATE RESULT. OTHERWISE ONLY THE FINAL
C     *       RESULT IS PRINTED.
C     *IDATA=1,PRINT INPUT.OTHERWISE DO NOT DO
C     *IRANK=RANK OF MATRIX VARE=N-KK
C     *
C     ************************************************************************
C     *
C     *WORK ARRAY
C     *
C     *W1....M*N
C     *W2....M*N
C     *W3....KK*M
C     *W4....  N*M
C     *W5....  N*N
C     *W6....KK*N
C     *W7....KK*KK
C     *WKAREA...IP
C     *VARU-----KK*KK
C     *R1.......N + II
C     *SN.......II
C     *DL1......N*N
C     *DL .....N
C     *TAD  ....N*N
C     *TAD1 ....N*IRANK
C     *F1 ......N
C     *VARW ....N*N
C     *
C     *OTHER ARRAYS
C     *
C     *NE=NUMBER OF MEASUREMENT FOUND IN ERROR
C     *CVAR=M*M-COVARIANCE MATRIX OF X
C     *B1=N*M-COEFFICIENT MATRIX  FOR  MEASUREMENT
C     *B2=N*MS COEFFICIENT MATRIX FOR MISSING VARIABLES
C     *ST=N*IR-TRANPOSE OF STOICHIOMETRIC  MATRIX
C     *H=N*N-MATRIX DEFINED BY B1*CVAR*B≠
C     *HINV=N*N- MATRIX OF INVERSE OF H
```

```
C     *SHINV=N*N-SQUARE ROOT OF HINV
C     *R=KK*N-MATRIX DEFINED BY EQ. 3.25 AND 26,
C     *A=N*KK PARTITION MATRIX FORMED BY B2 AND ST
C     *
C     ******************************************************************
C     *
C     *NOTE
C     *****
C     *ALL DIMENSION MUST BE  REVIEWED IF N.GT.20 AND/OR M.GT.30. NOT
C     *ONLY IN THE MAIN PROGRAM,BUT ALSO IN THE SUBROUTINE RECHAZ,SEARCH,
C     *AND SO ON.
C     *
C     *COMMON AND DIMENSIONS
C     *
      COMMON IPRINT,IDIA,ISW
      COMMON /AREA1/IFLAG /AREA2/TITLE(6) /AREA3/B(800),XDATOS(30),RVAR(
     *900),IDATA/AREA4/WW
      DIMENSION  X(30),CVAR(30,30),B1(20,30),B2(20,20),ST(20,10),A(20,30
     *),HINV(20,20),H(20,20),SHINV(20,20),R(30,30),W1(30,20),W2(30,20),W
     23(30,30),W4(20,30),W5(20,20),W6(30,20),W7(30,30),WKAREA(460),VARU(
     330,30),DL1(20,20),R1(220),DL(20),SN(200),TAO(20,20),TAO1(20,20),F1
     4(20),VARW(20,20)
      DIMENSION Y(30),ADJUST(30),ZETA(30)
      DIMENSION MG(20),VARE(20,20)
C     *
C     *DATA... ..AA=CHI-SQUARE (IRANK,ALFA),TABULED VALUE  AND BB FOR
C     *IRANK REDUCED BY ONE
C     *
      DATA AA,BB/6.251,4.605/
C     *
C     *INPUT INFORMATION ON DIMENSION OF THE SYSTEM
C     *
      READ(5,104)TITLE
      READ(5,*)N,M,MS,IR
C     *
C     *DATA SET
C     *
      ISW=0
      IPRINT=1
      IK=0
      IDATA=1
      KK=MS+IR
      IDIA=0
      IFLAG=1
      IP=N*N+3*N
      II=N*((N+1)/2)
      IRANK=N-KK
      CHI2=AA
    3 CALL        RECHAZ (N,M,MS,IR,KK,IK,B1,B2,ST,A,X,CVAR)
      IF(IPRINT.NE.0)WRITE(6,900)
      CALL        CALC1H (N,M,IP,II,B1, CVAR,H,HINV,X, W4,W5,WKAREA,DL,DL
     *1,TAO, R1,SN,SHINV)
      CALL        CALC1E (IP,N,KK,M,B1,A,HINV,R,H,W6,W7,WKAREA,VARU,UMB,X
     *,W4,W5,VARE)
      WW=CHISQ(HINV,UMB,N)
      WRITE(6,999)WW
      IF(WW.LE.CHI2)GO TO 10
      CHI2=BB
      CALL        SEARCH (M,N,KK,IRANK,NF,IK,JMB,VAPE,B1,SHINV,TAO1,F1,MG
     *,MS)
      IF(NE.LE.1)GO TO 40
      WRITE(6,110)
      WRITE(6,1)(MG(I),I=1,NE)
      STOP
   40 CONTINUE
```

```
      M=M-1
      MS=MS+1
      KK=KK+1
      IF(IFLAG.EQ.0)IK=MG(NE)
      ISW=ISW+1
      GO TO 3
   10 CALL        CALC2E (M,N,KK,CVAR,R1,HINV,UMR,X,R,W1,W2,W3,ADJUST,Y,Z
     *ETA)
      CALL        ESCRIB(N,M,MS,IP,KK,ZETA,X,Y,ADJUST)
      IF(IFLAG.EQ.0)STOP
      IF(ISW.EQ.(M+1))STOP
      IK=IK-1
      ISW=ISW+1
      GO TO  3
C     *
C     *FORMATS
C     *
    1 FORMAT(///,10X,10(I3),/)
  110 FORMAT(1H1,////,5X,*MEASUREMENTS FOUND IN GROSS ERROR ARE *,//)
  104 FORMAT(6A10)
  900 FORMAT(///,4X,*TRACING THE PROGRAM *,//)
  999 FORMAT(///,10X,*W=W= *,E12.5)
C     *
      END
      SUBROUTINE CALC1T (KK,N,IRANK,VARE,SHINV,TA01)
C     *
C     ***********************************************************************
C     *
C     *THIS SUBROUTINE CALCULATE THE EINGENVECTORS  FOR EINGENVALUES
C     *EQUAL TO ONE OF MATRIX VARW.
C     *PARAMETERS DEFINED IN MAIN PROGRAM
C     *
C     ***********************************************************************
C     *
C     *DIMENSION PACK
C     *
      DIMENSION VARE(N,N),TA01(N,IRANK), SHINV(N,N)
C     *
C     *DIMENSION OF THE WORK ARRAY
C     *
C     *R1(N),VARW(N,N),SN(N*((N+1)/2),WKAREA(N*N+3*N),AW1(N,N),DL(N), AND
C     *TAO(N,N)  AT LEAST
C     *
      DIMENSION R1(20),VARW(3,3),SN(30),WKAREA(60),AW1(3,3),DL(6),TAO(3,
     *3)
C     *
      IF(KK.EQ.0)GO TO 10
C     *
C     *CALCULATE VARW=SHINV*VARE*SHINV
C     *
      CALL MPROD(N,N,SHINV,N,N,VARE,AW1,0)
      CALL MPROD(N,N,AW1,N,N,SHINV,VARW,0)
C     *
C     *FIND EINGENVECTORS OF VARW
C     *
      CALL VCVTFS(VARW,N,N,SN)
      IJOB=2
      CALL EIGRS(SN,N,IJOB,DL,TAO,N,R1,IER)
C     *
      J1=N-IRANK
      DO 1 I=1,N
      DO 1 J=1,IRANK
      TA01(I,J)=TA(I,(J1+J))
    1 CONTINUE
      WRITE(6,99)
```

```
      DO 2 I=1,N
    2 WRITE(6,*)(TAO(I,J),J=1,N)
      RETURN
   10 CONTINUE
C
C     *SETTING TAO=IDENTITY MATRIX
C
      DO 3 I=1,N
      DO 3 J=1,IRANK
      TAO1(I,J)=0.
      IF(I.NE.J)GO TO 3
      TAO1(I,J)=1.
    3 CONTINUE
C
C     *FORMATS
C
   99 FORMAT(//,1H0,3X,*THE EIGENVECTORS OF VARW ARE *,/)
C
      RETURN
      END
```

V.2 SET OF SUBROUTINES

```
      SUBROUTINE CALC1E (IP,N,KK,M,B1,A,HINV,R,H,W6,W7,WKAREA,VARU,UMB,X
     *,W4,W5,VARE)
C
C     ******************************************************************
C
C     *THIS SUBROUTINE CALCUALATE THE IMBALANCE VECTOR
C
C     *ARGUMENTS ARE DEFINED IN THE MAIN PROGRAM.
C     *W4,W5,W6,W7,WKAREA,VARU ARE WORKS ARRAY.
C
C     ******************************************************************
C
C     *COMMON AND DIMENSIONS
C
      COMMON IPRINT,IDIA,ISW
      DIMENSION R(KK,N),A(N,KK),HINV(N,N),B1(N,M),X(M),UMB(N),VARE(N,N),
     *H(N,N),W4(N,M),W5(N,N),W6(KK,N),W7(KK,KK),VARU(KK,KK),WKAREA(IP)
C
C     *FIND MATRIX R ACCORDING EQ. 3.25 AND 3.26
C
      IF(KK.EQ.0)GO TO 10
      CALL MPROD0(N,KK,A,N,N,HINV,W6)
      CALL MPROD (KK,N,W6,N,KK,A,W7,0)
      IF(IPRINT.EQ.0)GO TO 20
      WRITE(6,8)
      WRITE(6,9)
      DO  4 I=1,KK
    4 WRITE(6,*)( W7(I,J),J=1,KK)
   20 CONTINUE
      IZ=0
      CALL LINV2F (W7,KK,KK,VARU,IZ,WKAREA,IER)
      IF(IPRINT.EQ.0)GO TO 30
      WRITE(6,8)
      WRITE(6,11)
      DO 5 I=1,KK
    5 WRITE(6,*)(VARU(I,J),J=1,KK)
   30 CONTINUE
      CALL MPROD (KK,KK,VARU,KK,N,W6,R,0)
```

```
      DO 1 I=1,KK
      DO 1 J=1,N
    1 R(I,J)=-R(I,J)
      IF(IPRINT.EQ.0)GO TO 40
      WRITE(6,8)
      WRITE(6,12)
      DO 7 I=1,KK
    7 WRITE(6,*)(R(I,J),J=1,N)
   40 CONTINUE
      CALL MPROD (N,KK,A,KK,N,R,W5,0)
C     *
C     *FIND (I-A*R)
C     *
      DO 2 I=1,N
      DO 2 J=1,N
      IF(I.NE.J)GO TO 2
      W5(I,J)=1.+W5(I,J)
    2 CONTINUE
C     *
C     *FIND (I-A*R)*B1*X=IMBALANCE VECTOR
C     *
      CALL MPROD (N,N,W5,N,M,B1,W4,0)
      CALL MPROD (N,M,W4,M,1,X,UMB,C)
C     *
C     *CALCULATE THE COVARIANCE OF  THE IMBALANCE
C     *
      CALL MPROD (N,N,W5,N,N,H,VARE,0)
      IF(IPRINT.EQ.0)RETURN
      WRITE(6,100)
      WRITE(6,*)(UMB(I),I=1,N)
      WRITE(6,8)
      WRITE(6,101)
      DO 3 I=1,N
    3 WRITE(6,*)(VARE(I,J),J=1,N)
      RETURN
   10 CALL MPROD (N,M,B1,M,1,X,UMB,C)
      WRITE(6,100)
      WRITE(6,*)(UMB(I),I=1,N)
C     *
C     *FORMATS
C     *
    8 FORMAT(//)
    9 FORMAT  (1H0,3X,*W7*,/)
   11 FORMAT  (1H0,3X,*VARU*,/)
   12 FORMAT  (1H0,3X,* R*,/)
  100 FORMAT(1H0,3X,*THE IMBALANCE VECTOR*,/)
  101 FORMAT(1H0,3X,*COVARIANCE MATRIX OF UMB*,/)
C
      RETURN
      END
      SUBROUTINE RECHAZ (N,M,MS,IR,KK,IK,B1,B2,ST,A,X,CVAR)
C     *
C     ************************************************************************
C     *
C     *THIS SUBROUTINE CREATE.- B1 AND A FROM DATA INPUT WHE ISW=0
C     *THIS SUBROUTINE CREATE.- B1 AND A AFTER REJECTS ONE MEASUREMENT-
C     *X(IK)-.
C     *PARAMETERS ARE DEFINED IN THE MAIN PROGRAM
C     *
C     ************************************************************************
C     *
C     *COMMON AND DIMENSIONS
C     *
      COMMON IPRINT,IDIA,ISW
      COMMON /AREA2/TITLE(6)/AREA3/B(800),XDATOS(30),RVAR(900),IDATA
```

```
      DIMENSION B1(N,M),B2(N,MS),ST(N,IP),A(N,KK),X(M),CVAR(M,M)
      IF(ISW.EQ.0)GO TO 10
      WRITE(6,100)
      WRITE(6,990)IK
C
C     *CHANGE COLUMN IK WITH M+1 (OLD M=NEW M+1) IN B,XDATOS<AND RVAR
C
      J1=N*(IK-1)
      J2=N*M
      DO 1 I=1,N
      C=B(J2+I)
      B(J2+I)=B(J1+I)
      B(J1+I)=C
    1 CONTINUE
      C=XDATOS(IK)
      XDATOS(IK)=XDATOS(M+1)
      XDATOS(M+1)=C
      IF(IDIA.EQ.0)GO TO 90
      M1=M+1
      J4=M1*M
      J3=M1*(IK-1)
      J5=IK
      J6=M1
      DO 2 I=1,M1
      C=RVAR(J3+I)
      RVAR(J3+I)=RVAR(J4+I)
      RVAR(J4+I)=C
      C=RVAR(J5)
      RVAR(J5)=RVAR(J6)
      RVAR(J6)=C
      J5=IK +I*M1
      J6=M1*(I+1)
    2 CONTINUE
      GO TO 30
   90 C=RVAR(IK)
      RVAR(IK)=RVAR(M+1)
      RVAR(M+1)=C
      GO TO 30
C
C     *INPUT DATA
C
   10 KK1=(KK+M)*N
      READ(5,11)(B(I),I=1,KK1)
      READ(5,11)(XDATOS(I),I=1,M)
        KK1=M
      IF(IDIA.NE.0)KK1=KK1*M
      READ(5,11)(RVAR(I),I=1,KK1)
      WRITE(6,100)
      WRITE(6,104)TTTLE
C
C     *CONSTRUCT MATRIX ST AND CHECK RANK
C
      IF(IP.EQ.0)GO TO 30
      J7=N*(M+MS)
      DO 3 I=1,IP
      DO 3 J=1,N
      J7=J7+1
      ST(J,I)=B(J7)
    3 CONTINUE
      IF(MRANK(ST,N,IP).EQ.IP)GO TO 30
      WRITE(6,999)
      WRITE(6,991)
      STOP
   30 CONTINUE
C
```

```
C     *CONSTRUCT B1,AND  A AND CHECK FOR RANK OF B1 AND B2
C     *
      IKJ=0
      DO 4 J=1,M
      DO 4 I=1,N
      IKJ=IKJ+1
      B1(I,J)=B(IKJ)
    4 CONTINUE
      IF(MRANK(B1,N,M).EQ.N)GO TO 40
      WRITE(6,999)
      WRITE(6,992)
      STOP
   40 IF(IDATA.EQ.0)GO TO 20
C     *
C     *PRINT DATA
C     *
      WRITE(6,101)N,M,KK
      WRITE(6,106)
      DO 16I=1,N
      WRITE(6,*)(B1(I,J),J=1,M)
   16 CONTINUE
   20 CONTINUE
      IF(MS.EQ.0)GO TO 60
      DO 5 J=1,MS
      DO 5 I=1,N
      IKJ=IKJ+1
      B2(I,J)=A(I,J)=B(IKJ)
    5 CONTINUE
      IF(MRANK(B2,N,MS).EQ.MS)GO TO 60
      WRITE(6,999)
      WRITE(6,993)
      STOP
   60 CONTINUE
C     *
C     *FORM CVAR AN X
C     *
      J7=0
      DO 6 I=1,M
      X(I)=XDATOS(I)
      DO 6 J=1,M
      CVAR(J,I)=
      IF(IDATA.EQ.0)GO TO 1000
      CVAR(J,I)=RVAR(J7+J)
      GO TO 6
 1000 IF(I.EQ.J)CVAR(I,J)=RVAR(J)*RVAR(J)
    6 CONTINUE
      IF(IR.EQ.0.AND.IDATA.EQ.0)RETURN
      IF(IDATA.EQ.0)GO TO 70
      WRITE(6,105)
      WRITE(6,*)(X( I ),I=1,M)
      WRITE(6,102)
      DO 15   I=1,M
      WRITE(6,*)(CVAR(I,J),J=1,M)
   15 CONTINUE
   70 CONTINUE
      IF(IR.EQ.0)GO TO 50
      MS1=MS+1
      DO 8 I=1,N
      DO 8 J=1,IR
      A(I,(MS+J))=ST(I,J)
    8 CONTINUE
   50 CONTINUE
      IF(KK.EQ.0)RETURN
      IF(IDATA.EQ.0)RETURN
      WRITE(6,107)
```

```
      DO 17 I=1,N
   17 WRITE(6,*)(A(I,J),J=1,KK)
C
C     *FORMATS
C     *
  100 FORMAT(1H1)
  101 FORMAT(1H0,3X,*NUMBER OF CONSTRAINTS= *,I5,//,1H0,3X,*NUMBER OF M
     1ASUREMENTS= *,I5,//,1H0,3X,*SIZE OF VECTOR ZETA= *,I5,//)
  102 FORMAT(//,1H0,3X,*COVARIANCE MATRIX*,//)
  104 FORMAT(6A10)
  105 FORMAT(//,1H0,3X,*MEASUREMENTS*,//)
  106 FORMAT(1H0,3X,*MATRIX OF COEFFICIENTS*,//)
  107 FORMAT(1H0,3X,*PARTITIONED MATRIX A *,//)
   11 FORMAT( 5E10.0)
  999 FORMAT(///,55(1H*),/,3X,*THIS DATA IS NOT CONSISTENT WITH ASSUMPT
     *ON 3AND 4*,/)
  990 FORMAT(///,3X,*THE MEASUREMENT*,I5,* HAS BEEN REJECTED AS SUSPECT
     *D OF CONTAINING*,/,3X,*GROSS ERROR*,////)
  991 FORMAT(3X,*RANK(ST).NE.IR*)
  992 FORMAT(3X,*RANK(B1).NE.N *)
  993 FORMAT(3X,*RANK(B2).NE.MS*)
C     *
      RETURN
      END
      SUBROUTINE CALC1H (N,M,IP,II,B1, CVAR,H,HINV,X, W4,W5,WKAREA,DL,D
     *L1,TAO, R1,SN,SHINV)
C     *
C     ***************************************************************
C     *
C     *THIS SUBROUTINE  CALCULATE H,ITS INVERSE AND THE SQUARE ROOT OF
C     *ITS INVERSE
C     *
C     *ARGUMENTS DEFINED IN THE MAIN PROGRAM
C     *W4,W5,WKAREA,R1,SN,DL1,DL,AND TAO ARE WORK ARRAYS.
C     *
C     ***************************************************************
C     *
C     *COMMON AND DIMENSIONS
C     *
      COMMON IPRINT,IDIA,ISW
      DIMENSION B1(N,M),X(M),W5(N,N),CVAR(M,M),W4(N,M),H(N,N),HINV(N,N
     *,DL(N),DL1(N,N),R1(IP),SN(II),WKAREA(IP),TAO(N,N)
      DIMENSION SHINV(N,N)
C     *
C     *CALCULATE H=B1*CVAR*B1T
C     *
      IND=0
      IF(IDIA.EQ.0)IND=-1
      CALL MPROD (N,M,B1,M,M,CVAR,W4,IND)
      IF(IPRINT.EQ.0)GO TO 31
      WRITE(6,301)
      DO 22 I=1,N
      WRITE(6,*)(W4(I,J),J=1,M)
   22 CONTINUE
   31 CONTINUE
      CALL MPROD1(N,M,W4,N,M,B1,H)
      IF(IPRINT.EQ.0)GO TO 32
      WRITE(6,304)
      DO 23 I=1,N
      WRITE(6,*)(H (I,J),J=1,N)
   23 CONTINUE
   32 CONTINUE
C     *
C     CALCULATE INVERSE OF H
C     *
```

```
      IZ=0
      CALL LINV2F (H ,N,N,HINV,IZ,WKAREA,IER)
      IF(IPRINT.EQ.0)GO TO 34
      WRITE(6,302)
      DO 25 I=1,N
      WRITE(6,*)(HINV (I,J),J=1,N)
   25 CONTINUE
   34 CONTINUE
C
C*CALCULATE THE SQUARE ROOT OF HINV
C
      CALL VCVTFS (HINV,N,N,SN)
      IJOB=2
      CALL EIGRS(SN,N,IJOB,DL,TAO,N,P1,IER)
      DO 14 I=1,N
      DO 14 J=1,N
      DL1(I,J)=0.
      IF(I.EQ.J)DL1(I,J)=SQRT(DL(I))
   14 CONTINUE
      CALL MPROD(N,N,TAO,N,N,DL1,W5,-1)
      CALL MPROD1 (N,N,W5,N,N,TAO,SHINV)
      IF(IPRINT.EQ.0)RETURN
      WRITE(6,303)
      DO 15 I=1,N
   15 WRITE(6,*)(SHINV(I,J),J=1,N)
C
C*FORMATS
C
  301 FORMAT(/,1H0,3X,*W4*,/)
  302 FORMAT(/,1H0,3X,*HINV*,/)
  303 FORMAT(/,1H0,3X,*SHINV*,/)
  304 FORMAT(/,1H0,3X,*H *,/)
C
      RETURN
      END
      SUBROUTINE CALC2E (M,N,KK,CVAR,B1,HINV,UMB,X,R,W1,W2,W3,ADJUST,Y,
     *ETA)
C
C*********************************************************************
C
C*THIS SUBROUTINE ESTIMATES THE MISSING VARIABLES AND/OR EXTENS OF
C*REACTIONS.
C
C*ARGUMENT DEFINE IN THE MAIN PROGRAM.W1,W2,W3 ARE  WORK ARRAYS
C
C*********************************************************************
C
C*COMMON AND DIMENSIONS
C
      COMMON IPRINT,IDIA,ISW
      DIMENSION W1(M,N),W2(M,N),W3(KK,M),ADJUST(M),CVAR(M,M),HINV(N,N),
     *1(N,M),X(M),Y(M),ZETA(KK),UMB(N),R(KK,N)
C
C*CALCULATE ADJUSMENT
C
      CALL MPROD1 (M,M,CVAR,N,M,B1,W1)
      IF(IPRINT.EQ.0)GO TO 1
      WRITE(6,30)
      DO 22 I=1,M
      WRITE(6,*) (W1(I,J),J=1,N)
   22 CONTINUE
    1 CONTINUE
      CALL MPROD (M,N,W1,N,N,HINV,W2,0)
      CALL MPROD (M,N,W2,N,1,UMB,ADJUST,0)
      IF(IPRINT.EQ.0)GO TO 2
```

```
      WRITE(6,190)
      WRITE(6,*)(ADJUST(I),I=1,M)
    2 CONTINUE
      IND=1
      CALL ADJEM(M,1,X,ADJUST,Y,IND)
      IF(IPRINT.EQ.0)GO TO 5
      WRITE(6,20)
      WRITE(6,*) ( Y( I ),I=1,M)
    5 CONTINUE
      IF(KK.EQ.0)RETURN
      CALL MPROD(KK,N,R,N,M,P1,W3,0)
      CALL MPROD(KK,M,W3,M,1,X,ZETA,0)
      IF(IPRINT.EQ.0)RETURN
      WRITE(6,191)
      WRITE(6,*)(ZETA(I),I=1,KK)
C     *
C     *FORMATS
C     *
   30 FORMAT(/,1H0,3X,*W1*,/)
   20 FORMAT (1H0,3X,*THE ESTIMATED VALUE OF MEASURED STREAMS*,//)
  190 FORMAT(/,1H0,3X,*ADJUSTMENT*,/)
  191 FORMAT(/,3X,*ZETA*,/)
C     *
      RETURN
      END
      SUBROUTINE ORDER (NRX,X,IU)
C     *
C     ***********************************************************************
C     *
C     *FIND VECTOR IU OF INDEX OF ELEMENTS OF X IN DESCENDENT  ORDER
C     *
C     *ARGUMENT
C     *
C     *NRX=SIZE OF VECTOR X
C     *X=VECTOR TO BE ORDERED
C     *IU=VECTOR OF INDEX
C     *
C     ***********************************************************************
C     *
C     *DIMENSIONS
C     *
      DIMENSION X(NRX),IU(NRX)
      DIMENSION X1(100)
      M1=NRX-1
      X1(NRX)=X(NRX)
      IU(NRX)=1
      ZMIN=ABS(X(1))
      DO 3 I=1,M1
      X1(I)=X(I)
      IF  ( ZMIN  .LE.ABS(X(I+1)))GO TO 3
      IU(NRX)=I+1
      ZMIN=ABS(X(I+1))
    3 CONTINUE
      DO 1 I=1,M1
      IMAX=I
      ZMAX=ABS(X1(I))
      DO 2 J= 1,NP)
      IF(ZMAX.GE.ABS(X1(J)))GO TO 2
      IMAX=J
      ZMAX=ABS(X1(J))
    2 CONTINUE
      IU(I)=IMAX
      X1(IMAX)=0.
    1 CONTINUE
      RETURN
```

```
      END
      SUBROUTINE SEARCH (M,N,KK,IRANK,NE,IK,UMB,VARE,B1,SHINV,TA01,F1,MG
     *,MS)
C
C*****************************************************************
C
C*THIS SUBROUTINE SEARCH AND OUTPUT THE MOST LIKELY SUBSET OF VECTOR
C*X CONTAINING GROSS ERRORS.
C*PARAMETERS DEFINED IN THE MAIN PROGRAM
C
C*****************************************************************
C
C*DIMENSIONS AND COMMON
C
      COMMON IPRINT,IDIA,ISW
      COMMON /AREA1/IFLAG/AREA4/WW
      DIMENSION IU(30),B1(N,M),VARE(N,N),SHINV(N,N),TA01(N,IRANK),F1(N),
     *MG(M),MG1(30),UMB(N)
C
C*FIND  IU, VECTOR OF ELEMENT OF U:B,SUCH THAT UMB(IU(1)).GT.UMB(IU(
     *2)).....GT.UMB(IU(N))
C
      I1=I3=NE=0
      IF(IFLAG.NE.0)GO TO 20
      CALL ORDER(N,UMB,IU)
      IF(IRANK.EQ.1)GO TO 40
C
C*FIND EINGENVECTORS OF VARW
C
      CALL CALCIT (KK,N,IRANK,VARE,SHINV,TA01)
C
C*PERFORMS TWO-SIDED TEST ON EACH ELEMENT OF  TEST FUNCTION W
C
      WRITE(6,97)
      DO 10 I=1,IRANK
      DO 11 J=1,N
   11 F1(J)=TA01(J,I)
      W=OMEGA (F1,SHINV,UMB,N)
      WRITE(6,*)W
      IF(W.LE.1.645)GO TO 10
      I1=I1+1
C
C*FIND THE SUBSET OF X CONTAINING GROSS ERRORS
C
      DO 12 K=1,M
      I2=IU(I1)
      IF(B1(I2,K).EQ.0)GO TO 12
      I3=I3+1
      MG1(I3)=K
   12 CONTINUE
   10 CONTINUE
      I4=I3-1
      DO 14 J=1,I3
      IF(MG1(J).EQ.0)GO TO 14
      JJ=J+1
      DO 13 KM=JJ,I4
      IF(MG1(J).EQ.MG1(KM))MG1(KM)=0
      NE=NE+1
      MG(NE)=MG1(J)
   13 CONTINUE
   14 CONTINUE
      GO TO 50
   40 W=SQRT(WW)
      WRITE(6,97)
      WRITE(6,*)W
```

```
      IF(W.LE.1.645)STOP
      I2=IU(1)
      DO 15 K=1,M
      IF(B1(I2,K).EQ.0)GO TO 15
      NE=NE+1
      MG(NE)=K
   15 CONTINUE
   50 CONTINUE
      WRITE(6,98)
      WRITE(6,*)(MG(IJ),IJ=1,NE)
      RETURN
   20 IF(ISW.EQ.0)GO TO 21
      M=M+1
      MS=MS-1
      KK=KK-1
   21 IF(IK.EQ.1)GO TO 30
      IK=M-ISW
      NE=NE+1
      RETURN
   30 WRITE(6,99)
      STOP
C
C     *FORMATS
C
   97 FORMAT(/,1H0,3X,*VALUES OF THE TEST FUNCTION ARE *,/)
   98 FORMAT(///,1H0,3X,*SUBSET OF X PROBABLE IN EPROB*,/)
   99 FORMAT(/////,5X,*NOTE.. ALL THE VARIABLES HAS BEEN TESTED*)
C
      END
      SUBROUTINE ESCRIB(N,M,MS,IR,KK,ZETA,X,Y,ADJUST)
C
C     ******************************************************************
C
C     *THIS SUBROUTINE WRITE THE FINAL RESULTS
C     *PARAMETERS DEFINED IN THE MAIN PROGRAM
C
C     ******************************************************************
C
C     *COMMON AND DIMENSIONS
C
      COMMON              /AREA2/TITLE(6)
      DIMENSION  ZETA(KK),X(M),Y(M),ADJUST(M)
C
      WRITE(6,100)TITLE
      WRITE(6,101)
      WRITE(6,102)
      DO 1 I=1,M
    1 WRITE(6,103)X(I),ADJUST(I),Y(I)
      IF(KK.EQ.0)RETURN
      IF(MS.EQ.0)GO TO 10
      WRITE(6,106)
      DO 2 I=1,MS
    2 WRITE(6,105)ZETA(I)
      IF(IR.EQ.0)RETURN
   10 CONTINUE
      WRITE(6,104)
      MM=MS+1
      DO 3 I=MM,KK
    3 WRITE(6,105)ZETA(I)
C
C     *FORMATS
C
  100 FORMAT(1H1,//,10X,6A10,///)
  101 FORMAT(28X,*FINAL RESULTS *,/,28X,13(1H*),//)
  102 FORMAT(10X,*MEASUREMENTS*,3X,*ADJUSTMENTS*,4X,*ADJUSTED VALUES*,/
```

```
     *10X,45(1H*),/)
  103 FORMAT(10X,3(E12.5,3X),/)
  104 FORMAT(//,10X,*ESTIMATES OF EXTENTS OF REACTION*,/,10X,33(1H*),/
  105 FORMAT(10X,E12.5,/)
  106 FORMAT(//,10X,*ESTIMATES OF MISSING VARIABLES*,/,10X,30(1H*),/)
C
      RETURN
      END
      SUBROUTINE MPROD1(NRA,NCA,A,NRB,NCB,B,C)
C
C*****************************************************************
C
C*THIS SUBROUTINE PERFORMS POSTMULTIPLICATION OF GENERAL MATRIX A
C*BY A TRANPOSE
C
C*ARGUMENT
C
C* NRA AND NRB ARE NUMBER OF ROWS    OF A AND B RESPECTIVELY
C* NCA AND NCB ARE NUMBER OF COLUMNS OF A AND B RESPECTIVELY
C* A= FIRST MATRX (GENERAL MODE)
C* B=SECOND MATRIX (GENERAL MODE)
C* C=MATRIX PRODUCT  (GENERAL MODE)
C
C*****************************************************************
C
      DIMENSION A(NRA,NCA),B(NRB,NCB),C(NRA,NRB)
C
C*PERFORMS A*B#=C
C
      IF (NCA.NE.NCB)GO TO 15
      DO 7 I=1,NRA
      DO 7 J=1,NRB
      C(I,J)=0.
      DO 7 K=1,NCB
      C(I,J)=C(I,J) + A(I,K)*B(J,K)
    7 CONTINUE
      RETURN
   15 WRITE(6,1)
    1 FORMAT(/,1H0,10X,*MATRICES ARE NOT CONFORTABLES*)
      STOP
      END
      SUBROUTINE MPROD0(NRA,NCA,A,NRB,NCB,B,C)
C
C*****************************************************************
C
C*THIS SUBROUTINE PERFORMS PREMULTIPLICATION OF GENERAL  MATRIX  A
C*BY A TRANPOSE
C
C*ARGUMENT
C
C* NRA AND NRB ARE NUMBER OF ROWS    OF A AND B RESPECTIVELY
C* NCA AND NCB ARE NUMBER OF COLUMNS OF A AND B RESPECTIVELY
C* A= FIRST MATRX (GENERAL MODE)
C* B=SECOND MATRIX (GENERAL MODE)
C* C=MATRIX PRODUCT  (GENERAL MODE)
C
C*****************************************************************
C
      DIMENSION A(NRA,NCA),B(NRB,NCB),C(NCA,NCB)
C
C*PERFORMS A#*B=C
C
      IF (NRA.NE.NRB)GO TO 15
      DO 7 I=1,NCA
      DO 7 J=1,NCB
```

```
      C(I,J)=0.
      DO 7 K=1,NRA
      C(I,J)=C(I,J) + A(K,I)*B(K,J)
    7 CONTINUE
      RETURN
   15 WRITE(6,1)
    1 FORMAT(/,1H0,10X,*MATRICES ARE NOT CONFORTABLES*)
      STOP
      END
      SUBROUTINE MPROD (NRA,NCA,A,NRB,NCB,B,C,IJOB)
C     ******************************************************************
C
C     THIS SUBROUTINE PERFORMS MATPIX MULTIPLICATION (A*B)
C
C     ARGUMENT
C
C     NCA AND NCB ARE NUMBER OF COLUMNS OF A AND B RESPECTIVELY
C     NRA AND NRB ARE NUMBER OF ROWS    OF A AND B RESPECTIVELY
C     A= FIRST MATRX (GENERAL MODE)
C     B=SECOND MATRIX (GENERAL MODE)
C     C=MATRIX PRODUCT  (GENERAL MODE)
C     IJOB=0  ,A AND B ARE GENERAL MATRICES
C          1, A IS A DIAGONAL MATRIX
C         -1 , B IS A DIAGONAL MATRIX
C
C     ******************************************************************
C
      DIMENSION A(NRA,NCA),B(NRB,NCB),C(NRA,NCB)
      IF(NCA.EQ.NRB)GO TO 15
      WRITE(6,1)
    1 FORMAT(/,1H0,10X,*MATRICES ARE NOT CONFORTABLES*)
      STOP
   15 IF(IJOB)9,10,11
   10 DO 5 I=1,NRA
      DO 5 J=1,NCB
      C(I,J)=0.
      DO 6 K=1,NCA
      C(I,J)=C(I,J)+A(I,K)*B(K,J)
    6 CONTINUE
      GO TO 16
   11 DO 7 I=1,NRA
      DO 7 J=1,NCB
      C(I,J)=A(I,I)*B(I,J)
    7 CONTINUE
      GO TO 16
    9 DO 8 I=1,NRA
      DO 8 J=1,NCB
      C(I,J)=A(I,J)*B(J,J)
    8 CONTINUE
   16 RETURN
      END
      SUBROUTINE  TRANM(A,T,NR,NC)
C     ******************************************************************
C
C     THIS SUBROUTINE  PERFORMS  THE TRANSPOSE OF  A GENERAL
C     MATRIX A
C     T=TRANSPOSED MATRIX OF A
C     A=MATRIX TO BE TRANSPOSE (GENERAL MODE)
C     NR=NUMBER OF  ROWS OF A
C     NC=NUMBER OF  COLUMNS  OF A
C
C     ******************************************************************
C
```

```
      DIMENSION A(NR,NC),T(NC,NP)
      DO 1 I=1,NR
      DO 1 J=1,NC
      T(J,I)=A(I,J)
    1 CONTINUE
      RETURN
      END
      SUBROUTINE CHANG (A,NRA,NCA,L1,L2,IP)
C     ******************************************************************
C     *
C     * THIS SUBPOUTINE INTERCHANGES POWS OR COLUMNS
C     * A=GENEPAL MATRIX
C     * NRA=NUMBERS OF ROWS
C     * NCA=NUMBERS OF COLUMNS
C     * L1=ROW OR COLUMN TO BE INTERCHANGED WITH L2
C     * IP=PARAMETER TO INDICATE  INTERCHANGE OF ROW (IP.EQ.0)
C     * OR INTERCHANGE  COLUMN (IP.GT.0)
C     *
C     ******************************************************
C     *
      DIMENSION A(NRA,NCA)
      IF(IP)3,3,4
    3 CONTINUE
      DO 1 I=1,NCA
      SA1=A(L1,I)
      A(L1,I)=A(L2,I)
      A(L2,I)=SA1
    1 CONTINUE
      RETURN
    4 CONTINUE
      DO 2 I=1,NRA
      SA2=A(I,L1)
      A(I,L1)=A(I,L2)
      A(I,L2)=SA2
    2 CONTINUE
      RETURN
      END
      SUBROUTINE ADIFM (NRA,NCA,A,B,C,IND)
C     ******************************************************************
C     * THIS SUBPOUTINE  PERFORMS THE ADITION  OR SUBSTRACTION
C     * OF TWO MATRICES OF THE SAME ORDER (A+B OR A-B)
C     * NCA=NUMBER CF COLUMNS  OF A $B$C
C     * NRA=NUMBER CF ROWS     OF A $B$C
C     * IND=PARAMETER TO INDICATE  ADITION (IND=0)
C     * OR SUBSTRACTION (IND.GT.0)
C     * C=MARIX SUM
C     * A*B MUST BE GENEPAL MATRIX
C     *
C     ******************************************************************
C     *
      DIMENSION A(NRA,NCA),B(NRA,NCA),C(NRA,NCA)
      DO 1 I=1,NRA
      DO 1 J=1,NCA
      IF(IND.NE.0)B(I,J)=-B(I,J)
      C(I,J)=A(I,J)+B(I,J)
    1 CONTINUE
      RETURN
      END
      FUNCTION MRANK (A,N,M)
C     ******************************************************
C     *
C     *THIS SUBPROGRAMAN DETERMINE THE RANK OF MATRIX A
```

```
C     *A=MATRIX WHOSE RANK IS GOING TO BE DETERMINED
C     *N=NUMBER OF ROWS
C     *M=NUMBER OF COLUMNS
C     *AUX=WORK ARRAY
C     *
C     ************************************************
C     *
      DIMENSION A(N,M),AUX(400),IROW(50),ICOL(50)
      IK=0
      DO 1 I=1,N
      DO 1 J=1,M
      IK=IK+1
      AUX(IK)=A(I,J)
    1 CONTINUE
C     *
C     *USING MFGR, CALCULATE IRANK
C     *
      EPS=1.0E-07
      CALL MFGR(A,N,M,EPS,IRANK,IROW,ICOL)
      MRANK=IRANK
C     *
C     *SAVE A
C     *
      IK=0
      DO 2 I=1,N
      DO 2 J=1,M
      IK=IK+1
      A(I,J)=AUX(IK)
    2 CONTINUE
      RETURN
      END
      FUNCTION OMEGA (F1,F2,F3,N)
C     *
C     ************************************************************************
C     *
C     *THIS SUBPROGRAM ESTIMATE  THE ELEMENT OF TEST FUNCTION  OMEGA
C     *F1=EIGENVECTOR CORRESPONDING TO EIGENVALUE EQUAL TO ONE
C     *F2=NEGATIVE SQUARE ROOT OF H
C     *F3=IMBALANCE VECTOR E
C     *N=NUMBER OF CONSTRAINTS
C     *AW1=WORK ARRAY
C     *
C     ************************************************************************
C     *
      DIMENSION F1(N),F2(N,N),F3(N),AW1(1,400)
C     *
C     *CALCULATE F1**F2=AW1
C     *
      CALL MPROD0(N,1,F1,N,N,F2,AW1)
C     *
C     *CALCULATE.AW1*F3=OMEGA
C     *
      CALL MPROD(1,N,AW1,N,1,F3,OMEGA,0)
      RETURN
      END
      FUNCTION CHISQ(A,B,N)
C     *
C     ************************************************************************
C     *
C     *THIS SUBPROGRAM DEFINE THE CHI-SQUARE  STATISTIC ASSOCIATED WITH
C     *THE RAW DATA.
C     *A=INVERSE OF H
C     *B=IMBALANCE VECTOR
C     *N=SIZE OF A AND B
C     *AW1=WORK ARRAY
```

```
C     *
C     ****************************************************************************
C     *
      DIMENSION A(N,N),B(N),AW1(1,400)
C     *
C     *CALCULATE B**A=AW1
C     *
      CALL MPRODO(N,1,B,N,N,A,AW1)
C     *
C     *CALCULATE AW1*B=CHISQ
C     *
      CALL MPROD(1,N,AW1,N,1,B,CHISQ,0)
      RETURN
      END
```

APPENDIX VI

PROGRAM LISTING FOR RIPPS METHOD

VI.1 MAIN PROGRAM

```
C
C
C     ADJUSTMENT OF EXPERIMENTAL DATA-SEE RIPPS,C.E.P.SYMP.SER.65,VOL 5
C      PP)8-13(1965)
C
      COMMON II,JJ
      COMMON IPRINT
      DIMENSION M(7),R(25,25),S(7,25),A1(25,25),A2(25,7),A3(7,25),A4(7,
     1),A2T( 7,25),A11(25,25),LL(25),LM(25),ASS(7,7),A0(7,7),A20(25,7),
     221(25,25),B1(25,25),B2(25,7),B3(7,25),WA(25,7),WB(25,7),WC(25,25)
     3BF(25,25),X(25),B(9),A(7,7),BB(25,25)
      DIMENSION TITLE(18)
      DIMENSION AA( 7,7 ),BA(25),SSIG(25)
      DIMENSION SIG(7),AM(9)
      DATA SIG / 3*0.,0.017, 0.050,0.024,0.200/
      WRITE(6,198)
      READ(5,199) TITLE
      WRITE(6,199) TITLE
      WRITE(6,197)
      READ(5,*)II,JJ
      WRITE(6,*)II,JJ
      WRITE(6,197)
C       II=NUMBER OF MEASUREMENTS,JJ=NUMBER OF CONSTRAINTS
      KK=II+JJ
      MM=II
      READ(5,*) N
      WRITE(6,*) N
      WRITE(6,197)
C     N=NUMBER OF MEASUREMENTS WITH GROSS ERRORS
      KA=KK-N
C     M=ROW,COLUMN NUMBERS IN MEASUREMENT MATRIX TO BE INTERCHANGED
      DO 101 IPRINT=1,MM
      IF(IPRINT.GT.1)GO TO 10
      WRITE(6,197)
      DO 100 I=1,KK
      READ(5,*) (A(I,J),J=1,KK)
      WRITE(6,*)(A(I,J),J=1,KK)
      WRITE(6,197)
  100 CONTINUE
      JJ1=JJ+1
      DO 810 MJ=JJ1,KK
      A(MJ,MJ)=A(MJ,MJ)/(SIG(MJ)*SIG(MJ))
  810 CONTINUE
      WRITE(6,811)
  811 FORMAT(1X,*DIAGONAL ELEMENTS OF S MATRIX*,/)
      WRITE(6,*) (A(MJ,MJ),MJ=1,KK)
      WRITE(6,197)
      READ(5,*)(B(I),I=1,KK)
      DO 812 MJ=JJ1,KK
      B(MJ)=B(MJ)/(SIG(MJ)*SIG(MJ))
  812 CONTINUE
      WRITE(6,197)
      WRITE(6,813)
  813 FORMAT (1X,*B MATRIX, CORRECTED FOR VARIANCE*,/)
```

```
      WRITE(6,*)(B(I),I=1,KK)
      DO 700  I=1,KK
      BA(I)=B(I)
      SSIG(I)=SIG(I)
      DO 700 J=1,KK
      AA(I,J) =A(I,J)
  700 CONTINUE
   10 CONTINUE
      IF(IPRINT.EQ.1)GO TO 800
      DO 900 I=1,KK
      B(I)=BA(I)
      SIG(I)=SSIG(I)
      DO 900 J=1,KK
      A(I,J)=AA(I,J)
  900 CONTINUE
  800 CONTINUE
      WRITE(6,198)
      IF(N.NE.1)GO TO 101
      M(N)=IPRINT
      WRITE(6,*) (M(I),I=1,N)
      WRITE(6,197)
      CALL          ADEXP(N,KK,KA,M,R,S,A1,A2,A3,A4,A2T,A11,LL,LM,ASS,A0,A2
     10,A21,B1,B2,B3,WA,WB,WC,BF,X,B,A,BB,SIG,AM)
  197 FORMAT(1H )
  198 FORMAT(1H1)
  199 FORMAT(18A4)
  101 CONTINUE
      CALL EXIT
      END
C
C
C     VI.2 SUBROUTINE ADEXP
C          ****************
C
C
C
      SUBROUTINE ADEXP(N,KK,KA,M,R,S,A1,A2,A3,A4,A2T,A11,LL,LM,ASS,A0,A2
     10,A21,B1,B2,B3,WA,WB,WC,BF,X,B,A,BB,SIG,AM)
      COMMON II,JJ
      COMMON IPRINT
      DIMENSION M(N),R(KA,KK),S(N,KK),A1(KA,KA),A2(KA,N),A3(N,KA),A4(N,
     1),A2T(N,KA),A11(KA,KA),LL(KK),LM(KK),ASS(N,N),A0(N,N),A20(KA,N),A
     21(KA,KA),B1(KA,KA),B2(KA,N),B3(N,KA),WA(KA,N),WB(KK,N),BF(KK,KK),
     3(KK),B(KK),A(KK,KK),BB(KK)
      DIMENSION SIG(KK),AM(KK)
C
C  FORMA((106) TOTAL KK * KK MATRIX
C                KK = NO. OF CONSTRAINTS + TOTAL NO. OF DATA POINTS
C                     (INCLUDING ESTIMATES)
C  (1X,(KK+1)/2(2X,E12.3),/,KK/2(2X,F12.3))
C
C  FORMAT(112)  KA * KA MATRIX
C  FORMAT(132)  KA * N MATRIX
C
      ISW=0
      LB=KK+1
      DO 200 I=1,N
      LA=M(I)+JJ
      LB=LB-1
      WRITE(6,*) LA,LB,KK
      WRITE(6,109)
      CALL CHANG (A,KK,KK,LA,LB,0)
      CALL CHANG (A,KK,KK,LA,LB,1)
      CALL CHANG (B,KK, 1,LA,LB,0)
C     STORE B VECTOR FOR FUTURE USE
```

```
      DO 650 MM=1,KK
      B(MM)=3(MM)
  650 CONTINUE
      CALL CHANG (SIG,KK,1,LA,LB,0)
  200 CONTINUE
      WRITE(6,109)
      DO 201 I=1,KK
      WRITE(6,105) (A(I,J),J=1,KK)
  201 CONTINUE
      WRITE(6,109)
      WRITE(6,106)(B(I),I=1,KK)
      K=KK-N+1
      WRITE(6,109)
      CALL RCUT(A,K,R,S,KK,KK,0)
      KA=K-1
      CALL CCUT(R,K,A1,A2,KA,KK,0)
      LB=N
      CALL CCUT(S,K,A3,A4,LB,KK,0)
      WRITE(6,113)
  113 FORMAT(1X,* A1*,//)
      WRITE(6,112)((A1(I,J),J=1,KA),I=1,KA)
  112 FORMAT (1X,9(2X,E12.3),/,5X,9(2X,E12.3),/,12X,3(2X,E12.3))
      WRITE(6,114)
  114 FORMAT(1X,* A2*,//)
      WRITE(6,132)((A2(I,J),J=1,N ),I=1,KA)
      WRITE(6,115)
  115 FORMAT(1X,* A3*,//)
      WRITE(6,112)((A3(I,J),J=1,KA),I=1,N )
      WRITE(6,116)
  116 FORMAT(1X,* A4*,//)
      WRITE(6,138)((A4(I,J),J=1,N ),I=1,N)
      CALL TRANM (A2,A2T,KA,LB)
      WRITE(6,120)
  120 FORMAT(1X,* A2T*,//)
      WRITE(6,112)((A2T(I,J),J=1,KA),I=1,N)
      CALL MCPY(A1,A11,KA,KA,0)
      CALL MINV(A11,KA,DET,LL,LM)
      WRITE(6,132) DET
  132 FORMAT(1X,2E12.3)
      WRITE(6,117)
  117 FORMAT(1X,* A11*,//)
      WRITE(6,112)((A11(I,J),J=1,KA),I=1,KA)
      CALL MPROD(KA,KA,A11,KA,LB,A2,WA)
      WRITE(6,121)
  121 FORMAT(1X,* WA *,//)
      WRITE(6,132)(( WA(I,J),J=1,N ),I=1,KA)
      CALL MPROD(LB,KA,A2T,KA,LB,WA,ASS)
      WRITE(6,119)
  119 FORMAT(1X,* ASS*,//)
      WRITE(6,*) ASS(N,N)
      CALL ADDFM (LB,LB,A4,ASS,A0,1)
      WRITE(6,118)
  118 FORMAT(1X,* A0 *,//)
      WRITE(6,*) A0(N,N)
  300 CALL MINV(A0,LB,DET,LL,LM)
      WRITE(6,132) DET
      WRITE(6,118)
      WRITE(6,138) ((A0(I,J),J=1,N),I=1,N)
      CALL MPROD (KA,LB,A2,LB,LB,A0,A20)
      WRITE(6,123)
  123 FORMAT(1X,* A20 *,//)
      WRITE(6,132)((A20(I,J),J=1,N ),I=1,KA)
      CALL MPROD (KA,LB,A20,LB,KA,A2T,A21)
      WRITE(6,124)
  124 FORMAT(1X,* A21*,//)
```

```
      WRITE(6,112)((A21(I,J),J=1,KA),I=1,KA)
      CALL MPPOD (KA,KA,A21,KA,KA,A11,A1)
  125 FORMAT(1X,* A1 *,//)
      WRITE(6,125)
      WRITE(6,112)(( A1(I,J),J=1,KA),I=1,KA)
      CALL MPPOD (KA,KA,A11,KA,KA,A1,A21)
      WRITE(6,124)
      WRITE(6,112)((A21(I,J),J=1,KA),I=1,KA)
      CALL MADD(A11,A21,B1,KA,KA,0,C)
      WRITE(6,126)
  126 FORMAT(1X,* B1 *,//)
      WRITE(6,112)(( B1(I,J),J=1,KA),I=1,KA)
      CALL MPPOD (KA,KA,A11,KA,LB,A20,WA)
      WRITE(6,121)
      WRITE(6,132)(( WA(I,J),J=1,N ),I=1,KA)
      C=-1.0
      CALL SMPY(WA,C,B2,KA,LB,0)
      WRITE(6,127)
  127 FORMAT(1X,* B2 *,//)
      WRITE(6,132)(( B2(I,J),J=1,N ),I=1,KA)
      CALL TRANM (B2,B3,KA,LB)
      CALL PTIE(B1,B3,WC,KA,KA,0,0,LP)
      CALL PTIE(B2,A0,WB,KA,LB,0,0,LB)
      CALL CTIE(WC,WB,BF,KK,KA,0,0,LB)
      WRITE(6,108) ISW
      WRITE(6,109)
      WRITE(6,128)
  128 FORMAT(1X,* BF *,//)
      WRITE(6,106) ((BF(I,J),I=1,KK),J=1,KK)
      WRITE(6,139)
      CALL MPPOD (KK,KK,BF,KK,1,B,X)
      WRITE(6,109)
      WRITE(6,107)
      WRITE(6,105)(X(I),I=1,KK)
C      EVALUATE SUM OF SQUARES OF DIFFERENCE BETWEEN MEASURED AND CALC.
      SUM =0.
      IJ=JJ+1
      WRITE(6,133)
  133 FORMAT(1X,*DIFFERENCES BETWEEN MEASURED AND CORRECTED VALUES*)
      DO 500 I=IJ,KK
C     CALCULATE ACTUAL ERROR
      AM(I)=BB(I)*SIG(I)*SIG(I)/2.
      WRITE(6,*) AM(I)
      WRITE(6,*) X(I)
      ADIFF=X(I)-AM(I)
      WRITE(6,137) ADIFF
  137 FORMAT (1X,*ASSUME DIFFERENCE X(I)-M(I) =*,E12.5)
      RDIFF=ADIFF/SIG(I)
      WRITE(6,136) RDIFF
  136 FORMAT(10X,*RELATIVE ERROR X-M/SIG  =*,E12.5,/)
      IF(I.GT.(KK-N))GO TO 500
      SUM=SUM+RDIFF*RDIFF
  500 CONTINUE
      WRITE(6,111) SUM
      IF(ISW.EQ.1) GO TO 400
C     MAKE AA=0
C      ASS IS THE SAME
      ISW=1
C      SET GROSS ERROR MEASUREMENTS=0
      DO 600 I=1,N
      KI=KK-I+1
      B(KI)=0.
  600 CONTINUE
      CALL SMPY(ASS,C,AC,LB,LB,0)
      GO TO 300
```

```
  400 CONTINUE
  101 FORMAT(7(2X,E10.3))
  106 FORMAT(1X,9(2X,E12.3),/,5X,9(2X,E12.3),/,10X,3(2X,E12.3),/)
  107 FORMAT(/,1X,*ADJUSTED DATA,X VECTOR*,/)
  108 FORMAT(/,1X,*FINAL MATRIX,ISW=*,I2,/)
  109 FORMAT(1H )
  111 FORMAT(1X,*SUMOF SQUARES=*,E12.5)
  138 FORMAT(1X,2E12.5)
      RETURN
      END
      SUBROUTINE MPROD (NRA,NCA,A,NRB,NCB,B,C)
C
C
C     THIS SUBROUTINE PERFORMS MATRIX MULTIPLICATION (A*B)
C     NCA AND NCB ARE NUMBER OF COLUMNS OF A AND B RESPECTIVELY
C     NRA AND NRB ARE NUMBER OF ROWS    OF A AND B RESPECTIVELY.
C     A= FIRST MATRX (GENERAL MODE)
C     B=SECOND MATRIX (GENERAL MODE)
C     C=MATRIX PRODUCT  (GENERAL MODE)
C
C
      DIMENSION A(NRA,NCA),B(NRB,NCB),C(NRA,NCB)
      IF(NCA.EQ.NRB)GO TO 15
      WRITE(6,1)
    1 FORMAT(/,1H0,10X,*MATRICES ARE NOT CONFORTABLES*)
      GO TO 16
   15 DO 6 I=1,NRA
      DO 6 J=1,NCB
      C(I,J)=0.
      DO 5 K=1,NCA
      C(I,J)=C(I,J)+A(I,K)*B(K,J)
    6 CONTINUE
   16 RETURN
      END
```

APPENDIX VII

ADDITIONAL RESULTS

STATISTICAL ANALYSIS OF PROCESS DATA.-PIPPS EXAMPLE.

NUMBER OF CONSTRAINTS= 3

NUMBER OF MEASUREMENTS= 4

SIZE OF VECTOR ZETA= 0

MATRIX OF COEFFICIENTS

```
.1 .3 -.5 -.2
.1 .5 -.2 -.7
.3 .1 -.2 -.1
```

MEASUREMENTS

```
.1858 4.7935 1.2295 3.88
```

COVARIANCE MATRIX

```
.000289 0. 0. 0.
0. .0025 0. 0.
0. 0. .000576 0.
0. 0. 0. .04
```

TRACING THE PROGRAM

W4

```
.0000289 .00075 -.0003456 -.008
.0000289 .0015 -.0001152 -.028
.0002312 .00025 -.0001152 -.004
```

H

```
.00203525 .00612201 .00096724
.00612201 .02052593 .00299616
.00096724 .00299616 .000633
```

HINV

```
5506.039823203 -1339.823207916 -2071.622837229
-1339.823207916 483.6504770914 -241.9646347827
-2071.622837229 -241.9646347827 5890.554870825
```

```
   SHINV
70.65545703226 -17.1013970929 -14.87912740173
-17.10139710929 12.61545198732 -5.660659476161
-14.87912740173 -5.660659476161 75.08077898426
   THE IMBALANCE VECTOR
-.05707 -.06722000000001 -.005910000000004

         W#W=    .84547E+01


  THE MEASUREMENT    4 HAS BEEN REJECTED AS SUSPECTED OF CONTAINING
  GROSS ERROR


   NUMBER OF CONSTRAINTS=     3

   NUMBER OF MEASUREMENTS=     3

   SIZE OF VECTOR ZETA=     1


   MATRIX OF COEFFICIENTS

.1 .3 -.5
.1 .5 -.2
.8 .1 -.2


   MEASUREMENTS

.1858 4.7935 1.2295


   COVARIANCE MATRIX

.000289 0. 0.
0. .0025 0.
0. 0. .000576
   PARTITIONED MATRIX A

-.2
-.7
-.1


   TRACING THE PROGRAM


   W4

.0000289 .00075 -.0003456
.0000289 .0025 -.0001152
.0002312 .00025 -.0001152
```

H

.00043525 .00052201 .00016724
.00052201 .00092593 .00019616
.00016724 .00019616 .000233

HINV

8051.966362363 -4034.661835749 -2382.710681232
-4034.661835749 3336.111111111 87.31884057958
-2382.710681292 87.31884057968 5928.566326485

SHINV

82.67524024943 -31.13796270997 -15.72254083578
-31.13796270997 48.53852677224 -3.24804682282
-15.72254083578 -3.24804682282 75.30483867581

W7

803.2696535236

VAPU

.00124491194587

R

-1.807787778428 1.91352588588 .2208943556955

THE IMBALANCE VECTOR

-.05171755056479 -.04848642697671 -.003233775282389

COVARIANCE MATPIX OF UMB

.0003854535221652 .0003477223275782 .0001423417610825
.0003477223275782 .0003159231465237 .0001090161637891
.0001423417610825 .0001090161637891 .0002205508805413

W≠W= .84374E+01

```
 THE MEASUREMENT     3 HAS BEEN REJECTED AS SUSPECTED OF CONTAINING
 GROSS ERROR


    NUMBER OF CONSTRAINTS=      3

    NUMBER OF MEASUREMENTS=     3

    SIZE OF VECTOR ZETA=        1


    MATRIX OF COEFFICIENTS

.1 .3 -.2
.1 .5 -.7
.8 .1 -.1


    MEASUREMENTS

.1858 4.7935 3.88


    COVARIANCE MATRIX

.000289 0. 0.
0. .0025 0.
0. 0. .04
    PARTITIONED MATRIX A

-.6
-.2
-.2


    TRACING THE PROGRAM


    W4

.0000289 .00075 -.008
.0000289 .0015 -.028
.0002312 .00025 -.004
```

H

.00152789 .00605289 .00089812
.00605289 .02050289 .00297312
.00089812 .00297312 .00060996

HINV

25891.53484923 -7215.608007909 -2952.385071675
-7215.608007909 2177.249134948 11.90064260987
-2952.385071675 11.90064260987 5928.608502224

SHINV

154.7626911182 -41.79333323693 -13.90545159411
-41.79333323693 19.87304061262 -5.968977190784
-13.90545159411 -5.968977190784 75.49581597337

W7

7205.82056352

VARU

.0001387767001946

R

1.873668956174 -.5400543650692 -.08095250345138

THE IMBALANCE VECTOR

-.01498035783191 -.0531901192773 .008119880722692

COVARIANCE MATRIX OF UMB

.0017793038793 .006036236795977 .0008814667959766
.006036236795977 .02049733893199 .002967568931992
.0008814667959766 .002967568931992 .0006044089319922

W*W= .15702E+01

W1

.0000289 .0000289 .0002312
.00075 .0015 .00025
-.008 -.029 -.004

ADJUSTMENT

.02017848623049 -.009523899315559 .07070519330146

THE ESTIMATED VALUE OF MEASURED STREAMS

.1655215137695 4.803023899316 3.809294816693

ZETA

1.159350595386

STATISTICAL ANALYSIS OF PROCESS DATA.-RIPPS EXAMPLE.

FINAL RESULTS

MEASUREMENTS	ADJUSTMENTS	ADJUSTED VALUES
.18580E+00	-.20178E-01	.16562E+00
.47935E+01	.95239E-02	.48030E+01
.38800E+01	-.70705E-01	.38093E+01

ESTIMATES OF MISSING VARIABLES

.11594E+01

THE MEASUREMENT 2 HAS BEEN REJECTED AS SUSPECTED OF CONTAINING GROSS ERROR

NUMBER OF CONSTRAINTS= 3

NUMBER OF MEASUREMENTS= 3

SIZE OF VECTOR ZETA= 1

MATRIX OF COEFFICIENTS

.1 -.6 -.2
.1 -.2 -.7
.8 -.2 -.1

MEASUREMENTS

.1858 1.2295 3.88

COVARIANCE MATRIX

.000289 0. 0.
0. .000576 0.
0. 0. .04

PARTITIONED MATRIX A

.3
.6
.1

TRACING THE PROGRAM

W4

.0000289 -.0003456 -.008
.0000289 -.0001152 -.028
.0002312 -.0001152 -.004

H

.00131025 .00567201 .00089224
.00567201 .01962593 .00284615
.00089224 .00284616 .000608

HINV

5842.932232229 -1623.480004719 -2442.194086606
-1623.480004718 543.835708026 -163.3382226081
-2442.194086606 -163.3382226081 5993.272972848

SHINV

78.9010154466 -18.61633043515 -16.46190252835
-18.61633043515 13.00245081333 -5.310764728573
-16.46190252836 -5.310764728573 75.45909167219

W7

120.9924519107

VARU

.008264978386735

R

-6.897721676005 1.463532609384 1.911969371714

THE IMBALANCE VECTOR

.02812237251801 .103164745036 .022487457506

COVARIANCE MATRIX OF UMB

.001066401945194 .004184313890388 .000644290648398
.004184313890388 .01665053778078 .00235026129679
.000644290648398 .00235026129679 .000525350216327

W≠W= .96364E+00

W1

.0000289 .0000289 .0002312
-.0003455 -.0001152 -.0001152
-.008 -.028 -.004

ADJUSTMENT

.01071465982552 .003902688778617 -.1469625940131

THE ESTIMATED VALUE OF MEASURED STREAMS

.1750853301745 1.225597311221 4.026962594013

ZETA

5.07747457506

STATISTICAL ANALYSIS OF PROCESS DATA.-RIPPS EXAMPLE.

FINAL RESULTS

MEASUREMENTS	ADJUSTMENTS	ADJUSTED VALUES
.18580E+00	-.10715E-01	.17509E+00
.12295E+01	-.39027E-02	.12256E+01
.38800E+01	.14696E+00	.40270E+01

ESTIMATES OF MISSING VARIABLES

.50775E+01

THE MEASUREMENT 1 HAS BEEN REJECTED AS SUSPECTED OF CONTAINING GROSS ERROR.

NUMBER OF CONSTRAINTS= 3

NUMBER OF MEASUREMENTS= 3

SIZE OF VECTOR ZETA= 1

MATRIX OF COEFFICIENTS

```
.3 -.6 -.2
.6 -.2 -.7
.1 -.2 -.1
```

MEASUREMENTS

```
4.7935 1.2295 3.88
```

COVARIANCE MATRIX

```
.0025 0. 0.
0. .000576 0.
0. 0. .04
```

PARTITIONED MATRIX A

```
.1
.1
.8
```

TRACING THE PROGRAM

W4

```
.00075 -.0003456 -.008
.0015 -.0001152 -.028
.00025 -.0001152 -.004
```

H

.00203236 .00611912 .00094412
.00611912 .02052304 .00297304
.00094412 .00297304 .00044814

HINV

61835.11111111 11336.11111111 -205525.
11336.11111111 3336.111111112 -46025.00000001
-205525. -46025.00000001 740725.0000001

SHINV

124.4605745415 -4.133055503064 -215.2407822605
-4.133055503064 15.93628833869 -55.36301723449
-215.2407822605 -55.36301723449 831.4633738023

W7

434694.4444445

VARU

.000002300466483481

R

.36140967473 96 .08132788037575 -1.305342194389

THE IMBALANCE VECTOR

-.05890779378874 -.06905779378874 -.02061235030993

COVARIANCE MATRIX OF UMB

.00203233 6995335 .006119096995335 .0009439359626813
.006119095995335 .020523016995 33 .00297285952681
.0009439359626813 .00297285962681 .0004465577014505

W≠W= .72953E+01

NOTE..ALL THE VARIABLES HAS BEEN TESTED

STATISTICAL ANALYSIS OF PROCESS DATA.-CASE II.EXAMPLE.

NUMBER OF CONSTRAINTS= 4

NUMBER OF MEASUREMENTS= 5

SIZE OF VECTOR ZETA= 3

MATRIX OF COEFFICIENTS

```
-1. 0. 0. 0. 0.
0. 1. -1. 0. 0.
0. 0. 1. -1. 0.
0. 0. 0. 0. 1.
```

MEASUREMENTS

```
10. 20. 50. 160. 100.
```

COVARIANCE MATRIX

```
1. 0. 0. 0. 0.
0. 4. 0. 0. 0.
0. 0. 25. 0. 0.
0. 0. 0. 256. 0.
0. 0. 0. 0. 100.
```

PARTITIONED MATRIX A

```
1. -1. 0.
0. 1. 0.
0. 0. 1.
0. 0. -1.
```

TRACING THE PROGRAM

W4

```
-1. 0. 0. 0. 0.
0. 4. -25. 0. 0.
0. 0. 25. -256. 0.
0. 0. 0. 0. 100.
```

H

```
1. 0. 0. 0.
```

```
0. 23. -25. 0.
0. -25. 281. 0.
0. 0. 0. 100.

   HINV

1. 0. 0. 0.
0. .037347195768821 .0033227006911122 0.
0. .0033227006911122 .0038543328017 0.
0. 0. 0. .01

   SHINV

1. 0. 0. 0.
0. .19230890889497 .01310766899916 0.
0. .01310766899916 .0606837854382 0.
0. 0. 0. .1

   W7

1. -1. 0.
-1. 1.037347155768 .0033227006911122
0. .0033227006911122 .013854332801?

   VARU

28.3595800525 27.3595800525 -6.561679790027
27.3595800525 27.3595800525 -6.561679790027
-6.561679790027 -6.561679790027 73.7532808399

   R

-1. -1. -.06561679790027 -.06561679790027
0. -1. -.06561679790027 -.06561679790027
0. 3.996802888651E-14 -.2624671916C1 .7375328083989

   THE IMBALANCE VECTOR

0. .5561579790074 -7.375328083991 -2.624671916009

   COVARIANCE MATRIX OF UMB

0. 0. 0. 0.
0. 1.640419947502 -18.43832020997 -6.561679790027
0. -18.43832020997 207.2467194601 73.75328083989
0. -6.561679790027 73.75328083989 26.24671916001

   W≠W=    .26247E+00
```

W1

```
-1.  0.  0.  0.
 0.  4.  0.  0.
 0. -25. 23.  0.
 0.  0. -256. 0.
 0.  0.  0. 100.
```

ADJUSTMENT

0. 6.945555242055E-13 -.6561679790067 6.719150104985 -2.624671916009

THE ESTIMATED VALUE OF MEASURED STREAMS

10. 20. 50.65616797901 153.280839895 102.624671916

ZETA

40.65616797901 30.65616797901 102.624671916

STATISTICAL ANALYSIS OF PROCESS DATA.-CASE II.EXAMPLE.

FINAL RESULTS

MEASUREMENTS	ADJUSTMENTS	ADJUSTED VALUE
.10000E+02	0.	.10000E+02
.20000E+02	-.69456E-12	.20000E+02
.50000E+02	.65617E+00	.50656E+02
.16000E+03	-.67192E+01	.15328E+03
.10000E+03	.26247E+01	.10262E+03

ESTIMATES OF MISSING VARIABLES

.40656E+02

.30656E+02

.10262E+03

CASE II, With the Values of Variances Changed

```
STATISTICAL ANALYSIS OF PROCESS DATA.-CASE  II.EXAMPLE.
   NUMBER OF CONSTRAINTS=      4

   NUMBER OF MEASUREMENTS=     5

   SIZE OF VECTOR ZETA=        3

   MATRIX OF COEFFICIENTS

-1.  0.  0.  0.  0.
 0.  1. -1.  0.  0.
 0.  0.  1. -1.  0.
 0.  0.  0.  0.  1.

   MEASUREMENTS

10. 20. 50. 160. 100.

   COVARIANCE MATRIX

1.  0.  0.  0.  0.
0.  4.  0.  0.  0.
0.  0.  16.  0.  0.
0.  0.  0.  10.24  0.
0.  0.  0.  0.  9.
   PARTITIONED MATRIX A

1. -1.  0.
0.  1.  0.
0.  0.  1.
0.  0. -1.

   TRACING THE PROGRAM


   W4

-1.  0.  0.  0.  0.
 0.  4. -16.  0.  0.
 0.  0.  16. -10.24  0.
 0.  0.  0.  0.  9.

   H

1.  0.  0.  0.
```

```
0. 20. -16. 0.
0. -16. 25.24 0.
0. 0. 0. 9.

   HINV

1. 0. 0. 0.
0. .09761904761905 .05952380952381 0.
0. .05952380952381 .07440476190476 0.
0. 0. 0. .1111111111111

   SHINV

1. 0. 0. 0.
0. .2925204365146 .1097763264112 0.
0. .1097763264112 .2497076692143 0.
0. 0. 0. .3333333333333


   W7

1. -1. 0.
-1. 1.097619047619 .05952380952381
0. .05952380952381 .1855158730159


   VARU

13.73552780931 12.73552780931 -4.086265607265
12.73552780931 12.73552780931 -4.086265607265
-4.086265607265 -4.086265607265 6.701475595914


   R

-1. -1. -.4540295119183 -.4540295119183
0. -1. -.4540295119183 -.4540295119183
0. 2.66453525916-14 -.255391600454 .744608399546

   THE IMBALANCE VECTOR

0. 4.540295119183 -7.446083995461 -2.55391600454


   COVARIANCE MATRIX OF UMB

0. 0. 0. 0.
0. 7.264472190691 -11.91373439274 -4.086265607265
0. -11.91373439274 19.53852440409 6.701475595914
0. -4.086265607265 6.701475595914 2.298524404086

        W#W=   .28377E+01
```

VALUES OF THE TEST FUNCTION ARE

1.684542801323

SUBSET OF X PROBABLE IN ERROR

3 4

STATISTICAL ANALYSIS OF PROCESS DATA.- CASE III.EXAMPLE.

NUMBER OF CONSTRAINTS= 6

NUMBER OF MEASUREMENTS= 12

SIZE OF VECTOR ZETA= 2

MATRIX OF COEFFICIENTS

```
1. 0. 0. 0. 0. 0. -1. 0. 0. 0. 0. 0.
0. 1. 0. 0. 0. 0. 0. -1. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. -1. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. -1. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. -1. 0.
0. 0. 0. 0. 0. 1. 0. 0. 0. 0. 0. 0.
```

MEASUREMENTS

80. 20. 0. 0. 0. 10. 79. 6. 8.2 1.7 15.8 0.

COVARIANCE MATRIX

```
1. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 1. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 1. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 1. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 1. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 1. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 1. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 1. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 1. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 1. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 1. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 1.
```

PARTITIONED MATRIX A

```
0. 0.
-5. -7.
4. 0.
0. 4.
5. 6.
-4. -4.
```

TRACING THE PROGRAM

H

```
2.  0.  0.  0.  0.  0.
0.  2.  0.  0.  0.  0.
0.  0.  1.  0.  0.  0.
0.  0.  0.  1.  0.  0.
0.  0.  0.  0.  1.  0.
0.  0.  0.  0.  0.  1.
```

HINV

```
.5  0.  0.  0.  0.  0.
0.  .5  0.  0.  0.  0.
0.  0.  1.  0.  0.  0.
0.  0.  0.  1.  0.  0.
0.  0.  0.  0.  1.  0.
0.  0.  0.  0.  0.  1.
```

SHINV

```
.707106781186 5  0.  0.  0.  0.  0.
0.  .707106781186 5  0.  0.  0.  0.
0.  0.  1.  0.  0.  0.
0.  0.  0.  1.  0.  0.
0.  0.  0.  0.  1.  0.
0.  0.  0.  0.  0.  1.
```

W7

```
80.5  69.5
69.5  92.5
```

VARU

```
.035359327217 13  -.026567278287 46
-.026567278287 46  .030772171253 82
```

R

```
0.  -.004587155963303  -.141437308 8685  .106269113 1498  -.052752293 57798  .035168195 71865
0.  .041284403 66972  .106269113 1498  -.123088685 0153  -.025229357 79816  .016819571 86544
```

STATISTICAL ANALYSIS OF PROCESS DATA.- CASE III.EXAMPLE.

FINAL RESULTS

MEASUREMENTS	ADJUSTMENTS	ADJUSTED VALUES
.80000E+02	-.50000E+00	.79500E+02
.20000E+02	-.60550E-01	.19939E+02
0.	0.	0.
0.	0.	0.
0.	0.	0.
.10000E+02	.33089E+00	.10331E+02
.79000E+02	.50000E+00	.79500E+02
.60000E+01	.60550E-01	.60606E+01
.82000E+01	.20031E+00	.84003E+01
.17000E+01	.23058E+00	.19306E+01
.15800E+02	-.30367E+00	.15496E+02
0.	0.	0.

ESTIMATES OF EXTENTS OF REACTION

.21001E+01

.48265E+00

STATISTICAL ANALYSIS OF PROCESS DATA.- CASE IV.EXAMPLE.

NUMBER OF CONSTRAINTS= 6

NUMBER OF MEASUREMENTS= 11

SIZE OF VECTOR ZETA= 3

MATRIX OF COEFFICIENTS

```
1. 0. 0. 0. 0. 0. -1. 0. 0. 0. 0.
0. 1. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. -1. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. -1. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. -1. 0.
0. 0. 0. 0. 0. 1. 0. 0. 0. 0. 0.
```

MEASUREMENTS

```
99. 20. 0. 0. 0. 10. 79. 8.2 1.7 15.8 0.
```

COVARIANCE MATRIX

```
1. 0. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 1. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 1. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 1. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 1. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 1. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 1. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 1. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 1. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 1. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. 0. 1.
```

PARTITIONED MATRIX A

```
0. 0. 0.
-1. -5. -7.
0. 4. 0.
0. 0. 4.
0. 6. 6.
0. -4. -4.
```

TRACING THE PROGRAM.

```
1. 0. 0. 0. 0. 0. -1. 0. 0. 0. 0.
0. 1. 0. 0. 0. 0. 0. 0. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. -1. 0. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. -1. 0. 0.
0. 0. 0. 0. 0. 0. 0. 0. 0. -1. 0.
0. 0. 0. 0. 0. 1. 0. 0. 0. 0. 0.
```

H

```
2. 0. 0. 0. 0. 0.
0. 1. 0. 0. 0. 0.
0. 0. 1. 0. 0. 0.
0. 0. 0. 1. 0. 0.
0. 0. 0. 0. 1. 0.
0. 0. 0. 0. 0. 1.
```

HINV

```
.5 0. 0. 0. 0. 0.
0. 1. 0. 0. 0. 0.
0. 0. 1. 0. 0. 0.
0. 0. 0. 1. 0. 0.
0. 0. 0. 0. 1. 0.
0. 0. 0. 0. 0. 1.
```

SHINV

```
.707106781186 5 0. 0. 0. 0. 0.
0. 1. 0. 0. 0. 0.
0. 0. 1. 0. 0. 0.
0. 0. 0. 1. 0. 0.
0. 0. 0. 0. 1. 0.
0. 0. 0. 0. 0. 1.
```

W7

```
1. 5. 7.
5. 93. 87.
7. 87. 117.
```

VARU

```
1.725 .0125 -.1125
.0125 .03541666666667 -.02708333333333
-.1125 -.02708333333333 .03541666666667
```

R

```
0. 1. -.05 .45 .5 -.4
0. -8.881784197001E-16 -.1416666666667 .1083333333333 -.05 .0333333
0. 0. .1083333333333 -.1416666666667 -.05 .03333333333333
```

THE IMBALANCE VECTOR

20. 1.151079231931E-13 .2033333333333 .2033333333333 -.3400000000001 -.3066666666666

COVARIANCE MATRIX OF UMB

2. 0. 0. 0. 0. 0.
0. 7.105427357601E-15 -3.552713678801E-15 0. 3.552713678801E-15 0.
0. -3.552713678801E-15 .4333333333333 .4333333333333 -.2 .1333333333333
0. 0. .4333333333333 .4333333333333 -.2 .1333333333333
0. -3.329070518201E-15 -.2 -.2 .4 .4
0. 3.552713678801E-15 .1333333333333 .1333333333333 .4 .7333333333333

W≠W= .20029E+03

THE EIGENVECTORS OF VARW ARE

0. 0. 0. 1. 0. 0.
5.329070518201E-15 0. -1. 0. -3.552713678801E-15 0.
.7071057811865 -.2581988897472 1.015364636281E-15 0. -.6396021490668 .1556997888323
-.7071067811865 -.2581988897472 -3.425527462222E-15 0. -.6396021490668 .1556997888323
2.008469021616E-15 -.7745966692415 8.033876086462E-16 0. .4264014327112 .4670993664969
0. .5163977794943 0. 0. 0. .8563488385777

VALUES OF THE TEST FUNCTION ARE

14.14213562373
-.4050813610757
-.3581095143143

SUBSET OF X PROBABLE IN ERROR

1 7