THE CANADIAN INSTITUTE OF SURVEYING

ASSOCIATION CANADIENNE DES SCIENCES GEODESIQUES

PAPERS FOR THE CIS ADJUSTMENT AND ANALYSIS SEMINARS

JULY 1983

STATISTICAL POST-ANALYSIS OF LEAST SQUARES ADJUSTMENT RESULTS

Robin R. Steeves
Surveys and Mapping Branch
Department of Energy, Mines & Resources
Ottawa, Ontario, K1A 0E9

and

Clive S. Fraser
Division of Surveying Engineering
The University of Calgary
Calgary, Alberta, T2N 1N4

ABSTRACT

A review of statistical procedures commonly used for the evaluation of least squares adjustment results is given. We begin with a summary of the underlying theoretical concepts after which various statistical evaluation procedures are developed. Rather than giving an exhaustive coverage of the multitude of statistical testing procedures, we attempt to give a clear presentation of those which we feel are most common. A numerical example of a geodetic application of these procedures is given.

TABLE OF CONTENTS

1. INTRODUCTION

We examine the subject of statistical post analysis of least squares adjustment results with emphasis on basic concepts rather than mathematical rigor. Most of the mathematical details are left to the many textbooks, where they are treated in detail. By "statistical post analysis" we mean simply a statistical evaluation of results of a least-squares adjustment.

We feel that a good understanding of practical applications of post analysis can be gained by (i) being aware of a few fundamental underlying concepts and (ii) going through numerical examples of practical applications. Of course, a certain amount of mathematical detail is also necessary. It must be pointed out that only the most common statistical methods used in least-squares post analysis will be discussed here; there is a wealth of other methods, which is continually growing.

Our first endeavour will be to discuss, in this chapter, a few very important underlying concepts on which both the method of least squares adjustment and its statistical analysis are based. The importance of these basic concepts should not be underestimated; their thorough understanding is a prerequisite for the understanding necessary to properly apply statistical evaluation methods in practice. Although we illustrate these concepts using the example of repeated measurements of the length of a desk, they are fully applicable to the more complex applications of least squares adjustments.

In Chapter 2, we will describe the post analysis procedures and their use. Finally, in Chapter 3, a numerical example is given, which demonstrates the use of the methods in the case of an actual small horizontal geodetic network.

1.1 Conceptual Foundations of Least Squares and its Post Analysis

The foundations of the method of least squares and its statistical post analysis are very closely related, as we shall now demonstrate.

Let us consider that we have made n measurements of the length of a desk. On the basis of these n observations we wish to determine (or "estimate") one value which "best" represents the length x of the desk.

Consider the case of n=2 in which our observations are ℓ_1 = 984 mm and ℓ_2=986 mm. Suppose we arbitrarily choose 950 mm as the estimate for x. Is this a "good" choice? Most of us agree that a better estimate would be "closer" to the observations. It is natural to define the "best" estimate as that value which is "closest" to the observations in some way.

We can look at this in another way. Since our observations are subject to "random errors" (assuming absence of "blunders" or "systematic" errors), i.e. each observation contains a "small distortion", we may consider the observations to be composed of two parts: the portion ℓ' which we would observe in the absence of random (or other) errors and small "residuals" v_i, i=1, 2, ..., n. That is, we may represent the observations $\underline{\ell}$ by

$$\underline{\ell} + \underline{v} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \ell' \qquad (1.1)$$

or

$$\underline{v} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \underline{\ell}' - \underline{\ell} \qquad (1.2)$$

For the case of n=2, we have explicitly

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} \ell' \\ \ell' \end{pmatrix} - \begin{pmatrix} \ell_1 \\ \ell_2 \end{pmatrix} \qquad (1.3)$$

In our previous discussion, we decided that the "best" estimate for x is that ℓ' such that $[\ell' \ \ell']^T$ is closest to $[\ell_1 \ \ell_2]^T$. Considering $[\ell' \ \ell']^T$

and $[\ell_1 \; \ell_2]^T$ as points in a two dimensional Cartesian coordinate system, the "distance" d between these two points may be defined by

$$d^2 = (\ell'-\ell_1)^2 + (\ell'-\ell_2)^2$$

$$= v_1^2 + v_2^2 \qquad (1.4)$$

(We do not consider other possible definitions of distance here). That particular value for ℓ' which makes d^2(or d) the smallest can thus be considered as the "best" estimate for x. That is, the best estimate for x is that ℓ' which minimizes the sum of squares of the residuals, or simply, which gives the "least squares".

By setting the first derivatives (with respect to ℓ') of equation (1.4) to zero we get the least-squares estimate for x as

$$\hat{x} = \frac{1}{2}\,(\ell_1+\ell_2) \qquad (1.5)$$

The corresponding least squares estimate of the residuals are then

$$\hat{v}_1 = \hat{x}-\ell_1$$

and (1.6)

$$\hat{v}_2 = \hat{x}-\ell_2$$

The major point of the previous discussion is the following: The method of least squares is based on selecting an estimate $\hat{x}$ of x which is, in a particular way, "best". In order to even define the method of least squares, we thus had to first define a method of measuring "goodness". That is, even before the method of least squares is defined, we are concerned about how good our estimates will be, which is precisely the essence of post analysis! We will see presently that the same basic measure of "goodness" is used to evaluate the results of least-squares adjustment.

1.2 Further Basic Concepts

We continue using the example of n measurements of a desk's length to illustrate some additional basic concepts and definitions. We denote our n observations by $\ell_1, \ell_2, \ldots, \ell_n$, or, in vector notation, simply as $\underline{\ell}$. In practice, we will find that our n observations take on values of $d_1, d_2, \ldots, d_m$ where $m \leq n$. For n sufficiently large, we usually have m much less than n.

There are certain numbers c_i, i=1, 2, ..., m of observations $\underline{\ell}$ having values d_i, i=1, 2, ..., m respectively. We observe in practice that some values of $\underline{\ell}$ occur more frequently than others i.e., in retrospect, we can say that some values of $\underline{\ell}$ are more "likely" to occur than others. This gives rise to the definition of "experimental probability" as

$$p_i = \frac{c_i}{n} . \tag{1.7}$$

Generalization of equation (1.5) to the case of n observations gives

$$\hat{x} = \frac{1}{n} \sum_{i=1}^{n} \ell_i \tag{1.8}$$

as the least squares estimate of the "mean" or "average" of our sample. This is equivalent to writing

$$\hat{x} = \frac{1}{n} \sum_{i=1}^{n} c_i d_i$$

$$= \sum_{i=1}^{n} \frac{c_i}{n} d_i$$

or,

$$\hat{x} = \sum_{i=1}^{n} p_i d_i \tag{1.9}$$

Up to this point, we have been working entirely with "discrete" numbers. It is much more convenient, in mathematics, to work with "continuous" functions. For this reason we now make a very useful "mathematical abstraction": We assume that the discrete function $p_i = c_i/n$ (c_i is a function of d_i) can be well represented by a smooth, continuous function over the entire interval $(-\infty, \infty)$. We call this function the "probability density function" (pdf) which is assumed (or "postulated") for ℓ.

Denoting the pdf of the continuous variable ℓ by $\phi(\ell)$, we can define the mean M of ℓ, analogous to equation (1.9), as

$$M = \int_{-\infty}^{\infty} \ell \cdot \phi(\ell) d\ell \qquad (1.10)$$

This result is a very useful concept in mathematical statistics. It is referred to as the "mathematical expectation" of ℓ and is denoted by $E(\ell)$, the "expected value of ℓ".

We can take this concept further and define the "variance" of ℓ as the average squared deviation

$$\sigma_\ell^2 = \int_{-\infty}^{\infty} (\ell-M)^2 \cdot \phi(\ell) d\ell \qquad (1.11)$$

Note the similarity of this definition with the "measure of goodness" we have defined for the least-squares method (equation 1.4). Let us use σ_ℓ^2 as a new measure of goodness for determining M. Writing the variance in terms of the expectation operator, we get:

$$\begin{aligned} \sigma_\ell^2 &= E[(\ell-M)^2] \\ &= E[\ell^2 - 2\ell M + M^2] \\ &= E(\ell^2) - 2M \cdot E(\ell) + M^2 \end{aligned}$$

Differentiating this result with respect to M (and setting the derivative to zero) results in

$$-2E(\ell) + 2M = 0$$

or

$$M = E(\ell) .$$

That is $E(\ell)$ is a "least-squares" or "minimum-variance" estimator.

We can generalize our discussion to a vector of continuous variables (or a "multivariate")

$$\underline{\ell} = [\ell_1 \ \ell_2 \ \ell_3 \ \ldots \ \ell_n]^T$$

whose postulated pdf is n-dimensional, which simply results in n-dimensional expectation operators (and therefore n-dimensional integration). We get, analogous to equations (1.10) and (1.11) respectively:

$$\underline{M} = E(\underline{\ell}) \tag{1.12}$$

and

$$\underline{C} = E[(\underline{\ell}-E(\underline{\ell}))\ (\underline{\ell}-E(\underline{\ell}))^T] \tag{1.13}$$

Explicity for n=2, we have

$$\begin{bmatrix} M_1 \\ M_2 \end{bmatrix} = E \begin{bmatrix} \ell_1 \\ \ell_2 \end{bmatrix}$$

and

$$\underline{C} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}$$

$$= \begin{bmatrix} E[(\ell_1-E(\ell_1))^2] & E[(\ell_1-E(\ell_1))(\ell_2-E(\ell_2))] \\ E[(\ell_2-E(\ell_2))(\ell_1-E(\ell_1))] & E[(\ell_2-E(\ell_2))^2] \end{bmatrix}$$

where $\underline{C}$ is called the "covariance matrix" of $\underline{\ell}$ and σ_{ij} the covariance of elements ℓ_i and ℓ_j. It can easily be shown that $\sigma_{ij} = \sigma_{ji}$, i.e. the covariance matrix is always symmetric.

At this point in our discussion, we can note two important things. First, we see that the covariance matrix can be considered as nothing but a generalization of the "goodness" criterion that we adopted when considering a "best" estimate for the representative value of a sample of observations (equation 1.4). Second, since we have adopted the covariance matrix as a measure of "goodness" for our estimates, it is not surprising that it is the basis for all our statistical post analysis. Indeed, the conceptual beginning of the notion of the covariance matrix was based on the definition of the goodness of our estimates.

We may also note here that we have not yet specified a particular probability density function (such as, for example, the "normal distribution") for our observations. We see, therefore, that the method of least squares is independent of this choice but is closely related to the concept of variance, variance being a "parameter" of such a pdf (as we shall discuss next).

1.3 Fundamentals of Statistical Testing

Let us consider again our sample of n observations of the length of a desk. If we draw the "histogram" of these observations we see, as we obtain more and more observations, that the histogram tends to take the general shape of the "normal distribution". On this basis, we may postulate that our observations have a normal pdf with a certain mean (or average) and certain variance. We call the mean and variance "parameters" of the pdf; the normal distribution is completely specified by these two parameters.

The mathematical abstraction we are making here can be interpreted as follows: we are representing the process of measuring the length of the desk by a "random variable". A random variable is simply a variable whose probability of taking on a specific value is given by its pdf. The basic assumption that this is a valid thing to do is known as "the basic postulate of mathematical statistics" and is the basis for all our statistical testing procedures. We can illustrate the basic ideas in statistical testing with a simple example.

Let us assume that we have made n measurements of the length of a desk. On the basis of the basic postulate of mathematical statistics, we represent the process of measuring the desk's length by the random variable ℓ whose pdf is $\phi(\ell)$. We can now make different assumptions or "hypothesis" about $\phi(\ell)$ and its parameters. For example, let us assume (or hypothesize) that $\phi(\ell)$ is the normal distribution (perhaps based on previous experience).

This basic assumption, which we wish to test, is called the "null hypothesis" and is denoted by H_o. We now want to devise a procedure to test (statistically) our null hypothesis on the basis of our collected measurements. That is, we seek to confirm or reject the assumption that our measurements are normally distributed. The testing procedure we use in this case is the Chi-square (χ^2) goodness of fit test (see section 2.2). Using this test, we compute a "statistic" y on the basis of our measurements and compare it with a value c which we obtain from tables of the χ^2 distribution. If $y < c$, we have (statistical) evidence that our null hypothesis should be accepted; if $y > c$, we have evidence that H_o should be rejected.

In the above example, we applied a "one-sided" statistical test. More common is the "two-sided" statistical test which results in computing two "critical values" c_1 and c_2 between which our statistic should lie if our null hypothesis is to be accepted. For example, let us make the null hypothesis that the mean of $\phi(\ell)$ is a specific value, say 5. We will assume that we have convinced ourselves previously that $\phi(\ell)$ is the normal distribution whose variance is 4 (say). Figure 1.1 depicts a normal distribution with mean 5 and variance 4. From the figure we can easily make the following observations. If our null hypothesis is true, the value of the statistic

$$M = \frac{1}{n} \sum_{i=1}^{n} \ell_i$$

(i.e. the sample mean) should lie within the interval from 0 to 10 with a probability equal to the area under the normal curve in this interval, i.e. within the shaded portion of Figure 1.1.

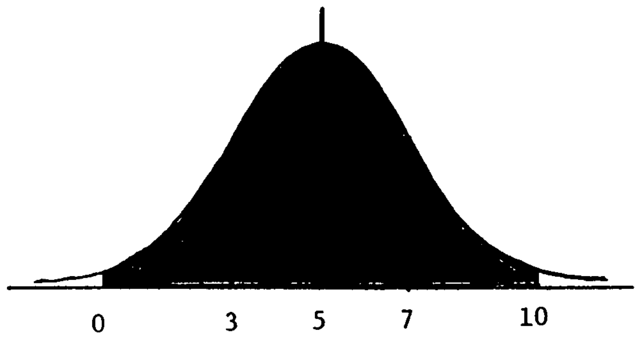


Figure 1.1

There is a chance, however, that M will fall outside this interval even if H_0 is true. The probability of this latter occurrence is equal to the area under the unshaded "tail" regions of Figure 1.1. We can say this another way. Let α equal the unshaded area of Figure 1.1, then the shaded area is $1-\alpha$. If H_0 is actually true, there is a $100(1-\alpha)\%$ probability that M will lie in the interval 0 to 10 and a $100\alpha\%$ probability that it will lie outside this interval. That is, we have a $100\alpha\%$ probability of rejecting H_0 when it is actually true. We see that the significance of this risk decreases as α decreases. Also note that by setting $\alpha=0$, the limits of our interval extend to $-\infty$ and $+\infty$ and we would always accept H_0. In practice, we set the "significance level" α to a small value (e.g. 0.05) which results in a "confidence level" $(1-\alpha)$ of nearly unity. In this way, we will have a probability of $(1-\alpha)\cdot 100 = 95\%$ (if $\alpha=0.05$) of accepting H_0 if it is true. The interval between the critical values is referred to as the "confidence interval" or, more generally, the "confidence region".

We will elaborate on the above mentioned testing procedures in the following chapter of this paper. Note that our tests will follow the basic procedure, as outlined above, of: (1) making a hypothesis about the shape of a pdf (e.g. that it is the normal distribution), and/or about one or more of its parameters; (2) choosing a testing procedure based on a statistic computed from our measurements; (3) comparing the computed statistic with bounds of a confidence region.

2. STATISTICAL POST ANALYSIS PROCEDURES

We will now work our way through a hypothetical least squares adjustment and discuss various statistical testing procedures along the way. By considering various situations which may occur in practice, we will see the repeated application of the basic principles discussed in Chapter 1. In order that our discussion can be related to the real world of geodetic practice, we will discuss our hypothetical least squares adjustment mainly in the context of two and three-dimensional networks. We assume the reader is familiar with setting up the observation equations and performing the basic adjustment procedure (see the seminar paper by Krakiwsky & Gagnon [this volume]).

2.1 χ^2 Test on the Variance Factor

One of the first considerations we must make in performing a least squares adjustment, which has a direct influence on all of our post analysis statistics, is whether we know the scale of the covariance matrices of the observations. Before considering this in the context of our hypothetical example, we will consider it in general.

We denote the covariance matrix of our observations by $\underline{C}_\ell$ and its inverse by $\underline{C}_\ell^{-1}$. When we say the scale of $\underline{C}_\ell$ is not known, this is equivalent to saying that an unknown factor has been extracted from $\underline{C}_\ell$. That is, we do not know the covariance matrix $\underline{C}_\ell$ in this case, we know only the relative values of its elements. This can be expressed in terms of the "weight matrix"

$$\underline{P} = \sigma_o^2 \, \underline{C}_\ell^{-1} \tag{2.1}$$

We call σ_o^2 the "variance factor". Note that we must always know or assume the relative values of the elements of $\underline{P}$ but for the least squares estimation of the parameters, knowledge of the scale of $\underline{C}_\ell^{-1}$ is not necessary since

$$
\begin{aligned}
\hat{\underline{x}} &= (\underline{A}^t\underline{P}\underline{A})^{-1}\underline{A}^t\underline{P}\underline{w} \\
&= (\underline{A}^t\ \sigma_o^2\ C_\ell^{-1}\ \underline{A})^{-1}\underline{A}^t\ \sigma_o^2\ \underline{C}_\ell^{-1}\underline{w} \\
&= \frac{1}{\sigma_o^2}\ (\underline{A}^t\underline{C}_\ell\ \underline{A})^{-1}\ \sigma_o^2\underline{A}^t\underline{C}_\ell^{-1}\ \underline{w} \\
&= (\underline{A}^t\underline{C}_\ell^{-1}\underline{A})^{-1}\underline{A}^t\underline{C}_\ell^{-1}\underline{w}
\end{aligned}
\tag{2.2}
$$

However, we must eventually know the scale of $\underline{C}_\ell$ in order to perform any statistical post analysis. In the case of unknown σ_o^2, the only source of information about σ_o^2 is given by its least-squares estimate

$$\hat{\sigma}_o^2 = \frac{\underline{r}^t\underline{P}\ \underline{r}}{\nu} \tag{2.3}$$

where $\hat{\underline{r}}$ is the vector of least-squares estimates of the residuals and ν is the number of degrees of freedom of the adjustment (ν = number of observation equations minus number of unknown parameters). $\hat{\sigma}_o^2$ is called the "estimated variance factor" or sometimes the "a posteriori variance factor". In practice, it is rare that σ_o^2 is unknown. Let us consider now, the case in which σ_o^2 is known. For the adjustment, we have shown (equation 2.2)) that only relative values of $\underline{C}_\ell$ need be used in the computations of $\hat{\underline{x}}$. This means that we can set σ_o^2 equal to any value we like and use $\underline{P} = \sigma_o^2\underline{C}_\ell^{-1}$ in the computations. In this case, after computing $\hat{\sigma}_o^2$, we can test the null hypothesis $H_o:\sigma_o^2=\hat{\sigma}_o^2$. This test will indicate whether $\hat{\sigma}_o^2$ is in agreement with our assumed value σ_o^2. Assuming that $\underline{r}$ has a normal distribution (which should also be tested), it can be shown that the statistic $\nu\hat{\sigma}_o^2/\sigma_o^2$ has a χ^2 distribution with ν degrees of freedom. The two tailed $100(1-\alpha)\%$ confidence region for $\hat{\sigma}_o^2$ is thus given by

$$\frac{\nu\hat{\sigma}_o^2}{\chi^2_{\nu,\ \alpha/2}} < \sigma_o^2 < \frac{\nu\hat{\sigma}_o^2}{\chi^2_{\nu,\ 1-\alpha/2}} \tag{2.4}$$

If σ_o^2 falls outside this interval, we have statistical evidence for suspecting one of the following:

1. σ_o^2 was incorrectly chosen or assumed; or

2. the model chosen to relate the observations to the unknown parameters was incomplete or not correct, or the observations contain systematic components which are not modelled properly.

In many cases, the second of the above reasons for failure of this test can be confirmed and appropriate action taken. In the event, however, that it cannot, we must assume that σ_o^2 was incorrectly chosen and proceed with further statistical procedures using $\hat{\sigma}_o^2$ since it is now the only information we have about the scale of $\underline{C}_\ell$ (since we reject the incorrectly chosen σ_o^2). Of course, we are assuming here that the normality of the estimated residuals has been confirmed, a subject which we now discuss.

2.2 χ^2 Goodness of Fit Test on the Estimated Residuals

In the previous section, we saw that the χ^2 test on the variance factor is based on the assumption that the estimated residuals are normally distributed. This assumption can be tested using the χ^2 goodness of fit test. We give here both a brief summary of situations which may arise for the general adjustment case, and an outline of how to perform this test.

There are two basic cases which may arise depending on whether the variance factor is known or unknown. The fundamental assumption we make here, as in all our statistical post analysis, is that the observations have a normal distribution. (Note that other assumptions may be made about the pdf of the observations, which we do not consider here.)

Under the assumption that the observations have a normal pdf, it can be shown that the least-squares estimates of the residuals will also have a normal distribution. In general, however, each residual has a different variance and thus a different normal distribution. In order to transform the residuals into variables which all have the same distribution, we divide them by their standard deviations. This process is called "standardization" of the residuals. Note that we assume the mean of each pdf is zero. In the case of the variance factor known, the standardized residuals

$$\hat{r}_i^* = \frac{\hat{r}_i}{\sigma_{\hat{r}_i}} \qquad (2.5)$$

have a normal distribution with mean zero and standard deviation equal to unity. We have tacitly assumed in our above discussion that either $\sigma_{\hat{r}_i}$ or $\hat{\sigma}_{\hat{r}_i}$ have been computed. However, because of the large amount of computations involved, it is usually considered too expensive to actually compute the covariance matrix of the estimated residuals. For this reason Pope [1976] has suggested that, for the parametric adjustment case, we can compute approximate standard deviations of the residuals using;

$$\hat{\sigma}_{\hat{r}_i} \simeq \left(\frac{n-u}{n}\right)^{\frac{1}{2}} \frac{\hat{\sigma}_o}{\sigma_o} \sigma_{\ell_i} \qquad (2.6)$$

where u is the number of parameters comprising the linear functional model.

Having determined the values of $\hat{r}_i^*$, we form a histogram of standardized residuals and examine its "goodness of fit" to a normal distribution of zero mean and unit variance. The (1-α) confidence interval for the tested statistic y is given as

$$0 < y < \chi^2_{m-1,\alpha} \qquad (2.7)$$

where

$$y = \sum_{i=1}^{m} \frac{(a_i - e_i)^2}{e_i} \qquad (2.8)$$

in which a_i are actual historgram counts for each of the m selected classes, and e_i are the theoretical counts from the normal pdf. In applying equation (2.7) we assume that σ_o^2 is known. For the case of σ_o^2 unknown the degrees of freedom value associated with the test is m-2 rather than m-1.

2.3 Tests for Outlying Estimated Residuals

In this section we describe the statistical testing procedure employed to detect outliers in the vector of estimated residuals $\hat{\underline{r}}$. These outliers, which are most often referred to as "gross errors", are usually due to observational blunders or unmodelled systematic errors.

The test for outlying estimated residuals involves an examination of the null hypothesis that the individual residuals $\hat{r}_i$ belong to a sample having a pdf of $n(0,\sigma^2_{\hat{r}_i})$ or $n(0,\hat{\sigma}^2_{\hat{r}_i})$, depending on whether σ^2_o is known or unknown. Once again, the tests essentially involve an examination of the standardized residuals. In the more commonly encountered case of σ^2_o known we make the following test:

$$\hat{r}^*_i = \left| \frac{\hat{r}_i}{\sigma_{\hat{r}_i}} \right| < n_{1-\alpha/2} \tag{2.9}$$

Alternatively in the case where σ^2_o is unknown we employ the test

$$\hat{r}^*_i = \left| \frac{\hat{r}_i}{\hat{\sigma}_{\hat{r}_i}} \right| < \tau_{\nu,1-\alpha/2} \tag{2.10}$$

where $\hat{\sigma}_{\hat{r}}$ can be obtained from equation (2.6) and τ_ν is the Tau pdf with ν degrees of freedom. In adjustments with large observational redundancy (e.g. large photogrammetric bundle adjustments), τ_ν can be replaced for practical purposes by the student pdf t_ν or the normal pdf n.

In the above tests we are examining each residual by itself, out of the context of the other residuals. Since we are really concerned about the characteristics of the entire set of residuals we should base our confidence intervals on simultaneous (or "in context") probability statements, that all of the standardized residuals are simultaneously within a specified interval with a certain probability. The resulting tests, at a (1-α) confidence level, are equivalent to equations (2.9) and (2.10) except that $\alpha/2$ is replaced by $\alpha/2N$ where N is the number of residuals. Also, since we have ignored the correlations between residuals, it turns out that the probability associated with our test is not $1-\alpha$, but somewhat higher. The best we can do in this case is to replace the equality in our probability statement with the inequality $\geq$ (greater than or equal to). This result is called "Bonferroni's inequality" (see Vaníček and Krakiwsky [1981] for more details).

One further test for gross errors needs to be outlined, this being the concept of "data-snooping" which was developed by Baarda [1967; 1968]. Data-snooping is a testing procedure for outliers which is very much tied to the quality of the adjustment model, i.e. to its reliability. Of importance in this regard is the internal reliability, which gives the magnitude of a gross error that will just be detectable at a certain probability level. The test again involves the standardized residual

$$\hat{r}_i^* = \frac{\hat{r}_i}{\hat{\sigma}_{\hat{r}_i}} \tag{2.11}$$

where $\hat{\sigma}_{\hat{r}_i}$ is obtained from the estimated covariance matrix of the residuals, and not via equation (2.6). The null hypothesis that $\hat{r}_i$ is free from gross error is rejected when

$$\left| \hat{r}_i^* \right| > c \tag{2.12}$$

where c is a critical value for a chosen confidence level. The value of c is dependent on the probability α of making a type I error, and on the probability β of making a type II error. For given values of α and β, c may be computed by a procedure outlined in Baarda [1968]. Two commonly adopted values for c are 3.3 and 4.1, the latter corresponding to $\alpha = 0.1\%$ and $\beta = 20\%$.

2.4 Confidence Regions for the Estimated Parameters

Once the parameters $\underline{\hat{x}}$ are determined (via equation 2.2), we can use the covariance matrix $\underline{C}_{\hat{x}} = \sigma_0^2(\underline{A}^t\underline{C}_\ell^{-1}\underline{A})^{-1}$ to establish confidence regions around these estimated values. Within the u-dimensional parameter space it is possible to formulate a multivariate pdf associated with the parameters, thus giving us a direct measure of the degree of trust we may place in the results. This pdf is postulated as

$$\phi_{\underline{x}} = n(\underline{\hat{x}},\underline{C}_{\hat{x}}) = \frac{1}{(2\pi)^{n/2}(\det \underline{C}_{\hat{x}})^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\underline{x}-\underline{\hat{x}})^t\underline{C}_{\hat{x}}^{-1}(\underline{x}-\underline{\hat{x}})\right] \qquad (2.13)$$

From equation (2.13) we can show that a region of constant probability is bounded by a u-dimensional hyperellipsoid, centred at $\underline{\hat{x}}$. The function which describes the hyperellipsoid for any chosen confidence level is given as

$$y = (\underline{x}-\underline{\hat{x}})^t \; \underline{C}_{\hat{x}}^{-1} \; (\underline{x}-\underline{\hat{x}}) \qquad (2.14)$$

where, for known σ_0^2, the quadratic form has a χ^2 distribution with u degrees of freedom. The probability that any point $\underline{x}$ lies within the hyperellipsoid is given by the expression

$$\mathrm{pr}(\chi^2{}_u < y) = 1 - \alpha \qquad (2.15)$$

We determine the parameters of the hyperellipsoid (i.e. semiaxes and their orientations) through a singular value decomposition of $\underline{C}_{\hat{x}}$. For the two-dimensional case the region of constant probability for y=1 is simply the well-known standard point error ellipse, formulae for the computation of which are given in the seminar paper by Mepham and Nickerson [this volume]. The following table, Table 2.2, lists the probabilities of a point $\underline{x}$ lying within the hyperellipsoid for varying values of $y^{\frac{1}{2}}$, for both u=2 and u=3.

Table 2.2

u = 2 error ellipse		u = 3 error ellipsoid	
pr	$y^{\frac{1}{2}}$	pr	$y^{\frac{1}{2}}$
0.394	1	0.199	1
0.500	1.177	0.500	1.538
0.900	2.146	0.900	2.500
0.950	2.447	0.950	2.700
0.990	3.035	0.990	3.368

For example, the 95% confidence region (significance level α=0.05) for a point with adjusted two-dimensional coordinates is obtained by multiplying the semiaxes of the standard ellipse (y=1) by the factor $y^{\frac{1}{2}}$=2.447. Further, we may say that there is only a 20% probability that an adjusted point in three-dimensional space will lie within the standard point error ellipsoid.

2.5 Tests for Compatibility of Estimated Parameters with Existing Independent Estimates

There are a number of practical adjustment situations in which we may have independently determined values $\underline{x}$ for at least some of the parameter estimates $\hat{\underline{x}}$. For example, in a photogrammetric self-calibration adjustment we may recover estimates of the lens distortion coefficients. These can then be compared to the corresponding values obtained in a laboratory calibration. In the simplest case, we usually wish to examine the null hypothesis that $\underline{x}$ and $\hat{\underline{x}}$ are compatible on a given level of probability 1-α.

Two cases can be distinguished for the test of the null hypothesis H_o: $\hat{\underline{x}} = \underline{x}$. The first is that σ_o^2 is known, and the second is that it is unknown. Since the former is the most commonly encountered situation in practical geodetic adjustments we first present this case.

In testing whether $\underline{x}$ lies within a given confidence region about $\hat{\underline{x}}$, we employ the statistic

$$y = (\underline{x}-\hat{\underline{x}})^t \, \underline{C}_{\hat{\underline{x}}}^{-1} (\underline{x}-\hat{\underline{x}}) \tag{2.16}$$

which was introduced in the previous section. The pdf of y is χ^2 with u degrees of freedom. Thus, for a given significance level α, $\underline{x}$ and $\hat{\underline{x}}$ may be assumed compatible if $y < \chi^2_{u,\alpha}$.

For the case where the variance factor is unknown we make use of the estimated variance factor $\hat{\sigma}_0^2$. We then test the value of the statistic

$$y = \frac{1}{u} (\underline{x}-\hat{\underline{x}})^t \, \hat{\underline{C}}_{\hat{\underline{x}}}^{-1} \, (\underline{x}-\hat{\underline{x}}) \tag{2.17}$$

against the tabulated value $F_{u,\nu,\alpha}$ from the F (Fisher) pdf.

In a number of adjustment problems (e.g. photogrammetric self-calibration) it is often practicable, especially in cases where there is very limited correlation between the parameters, to employ a simple one-dimensional test of

$$y = \frac{(\underline{x}_i-\hat{\underline{x}}_i)^2}{\sigma^2_{\hat{x}_i}} \quad \text{or} \quad y = \frac{(\underline{x}_i-\hat{\underline{x}}_i)^2}{\hat{\sigma}^2_{\hat{x}_i}} \tag{2.18}$$

for the cases of σ_0^2 known and unknown, respectively. Note here that

$$\chi^2_{1,\alpha} = F_{1,\infty,\alpha} \tag{2.19}$$

3. NUMERICAL EXAMPLE

The practical example we consider is the horizontal network NET 299, Geodetic Survey of Canada, in Nova Scotia. A diagram of the network, which comprises eight stations (Camperdown III is treated as "fixed"), is shown in Figure 3.1. The observations, which included one azimuth, 19 distances and 34 directions were adjusted by the adjustment program GEOPAN (Steeves, 1978).

In working through the numerical example we will follow the order in which the post-analysis tests were outlined in the previous chapter.

3.1 The χ^2 Test on the Variance Factor

The estimated variance factor (equation 2.3) was computed to be

$$\hat{\sigma}_o^2 = \frac{\underline{r}^t \underline{P}\ \underline{r}}{\nu} = 3.864$$

From this value we compute the 95% (α=0.05) confidence region for (equation 2.4):

$$\frac{\nu\hat{\sigma}_o^2}{\chi^2_{\nu;\alpha/2}} < \sigma_o^2 < \frac{\nu\hat{\sigma}_o^2}{\chi^2_{\nu,1-\alpha/2}}$$

For ν=32 and α=0.05 we find that the computed region is

$$2.38 < \sigma_o^2 < 6.45$$

But, σ_o^2 has been assumed a priori to have a value of 1.00. Thus, the test fails. We shall see that the reason for this failure is the presence of an observation blunder. Once this outlier is removed, the variance factor test passes.

3.2 The χ^2 Goodness of Fit Test on the Estimates Residuals

Initially, all residuals are standardized using equation (2.6). The goodness of fit test can then be carried out with all standardized residuals included or with just standardized residuals from different types of observations (distances or directions and angles). Here we will consider only the first case.

The 2-class histogram used for our χ^2 goodness of fit test is shown in Figure 3.2, along with the bell-shaped curve of the normal distribution. A summary of the test is given below, this listing having been provided by the GEOPAN program.

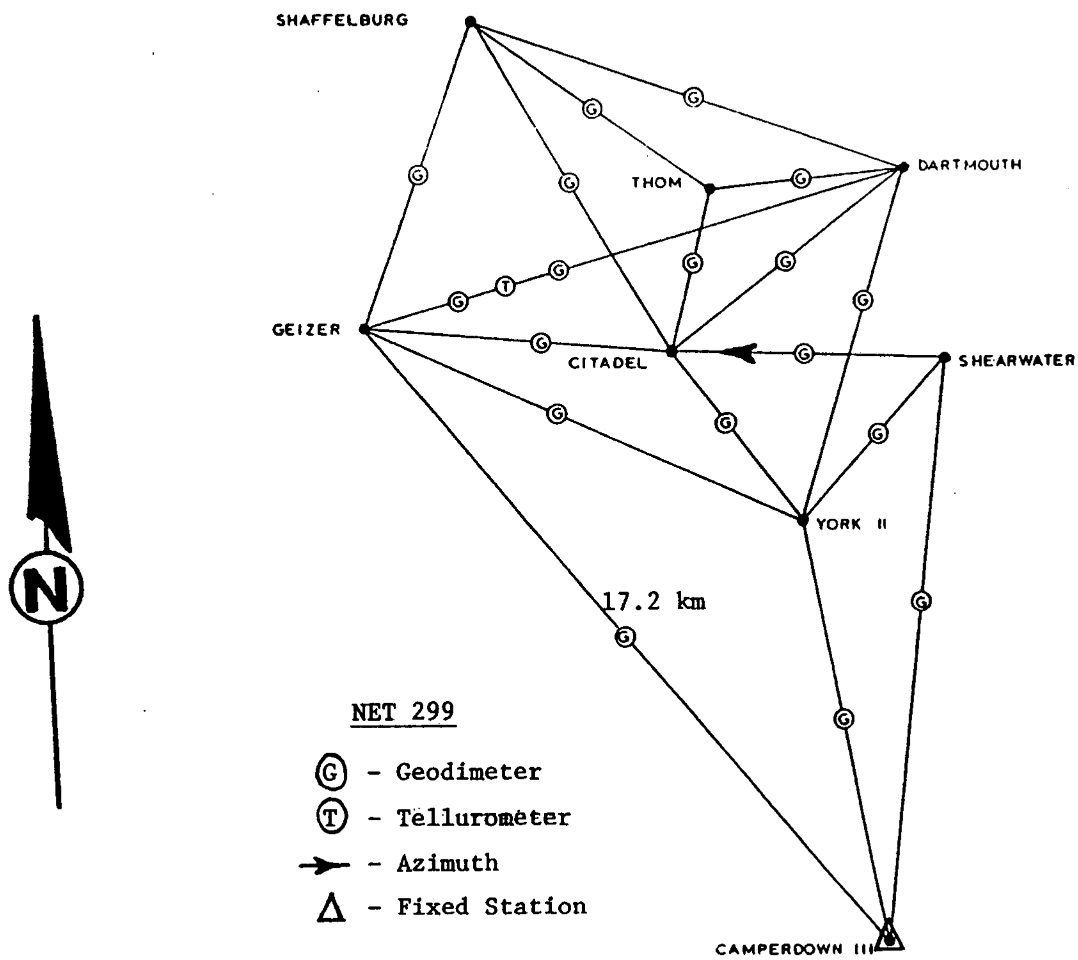

SHAFFELBURG
DARTMOUTH
THOM
GEIZER
CITADEL
SHEARWATER
YORK II
CAMPERDOWN III
17.2 km
N
NET 299
G - Geodimeter
T - Tellurometer
- Azimuth
Δ - Fixed Station

Figure 3.1

CHI-SQUARE GOODNESS OF FIT TEST

ON THE STANDARDIZED RESIDUALS (ALL RESIDUALS INCLUDED)

THE NUMBER OF CLASSES IS 2
THE NUMBER OF DEGREES OF FREEDOM FOR THE TEST IS 1

SUMMARY OF THE COMPUTATION OF THE CHI-SQUARE STATISTIC

CLASS INTERVAL	OBSERVED FREQ.(O)	EXPECTED FREQ.(E)	(O-E)	(O-E)**2	(O-E)**2/E
(-5.0 , 0.0)	26	26	0	0	0.00
(0.0 , 5.0)	28	26	2	4	.15

TOTAL (CHI-SQUARE STATISTIC) --> .15

THE CHI-SQUARE CRITICAL VALUE AT THE 95.000 % CONFIDENCE LEVEL IS --> 3.84

.15 IS LESS THAN 3.84

THE TEST PASSES

(SEE HISTOGRAM ON NEXT PAGE)

NOTE: THE HISTOGRAM IS FIRST PLOTTED WITH 2 CLASSES (THAT USED IN THE GOODNESS OF FIT TEST); THEN WITH 20 CLASSES SO THAT A MORE DETAILED REPRESENTATION OF THE ACTUAL RESIDUAL DISTRIBUTION IS GIVEN.

The test passes and thus we accept the hypothesis that the residuals are normally distributed. Now, to obtain a more detailed picture of the distribution of the standardized residuals we plot the histogram with 20 classes. This plot is shown in Figure 3.3. From the figure it can be seen that one standardized residual is removed from the zero mean by greater than 4.5 times the unit standard error. We now naturally suspect that this observation is a blunder or gross error. In order to test this hypothesis we carry out the test for outlying estimated residuals.

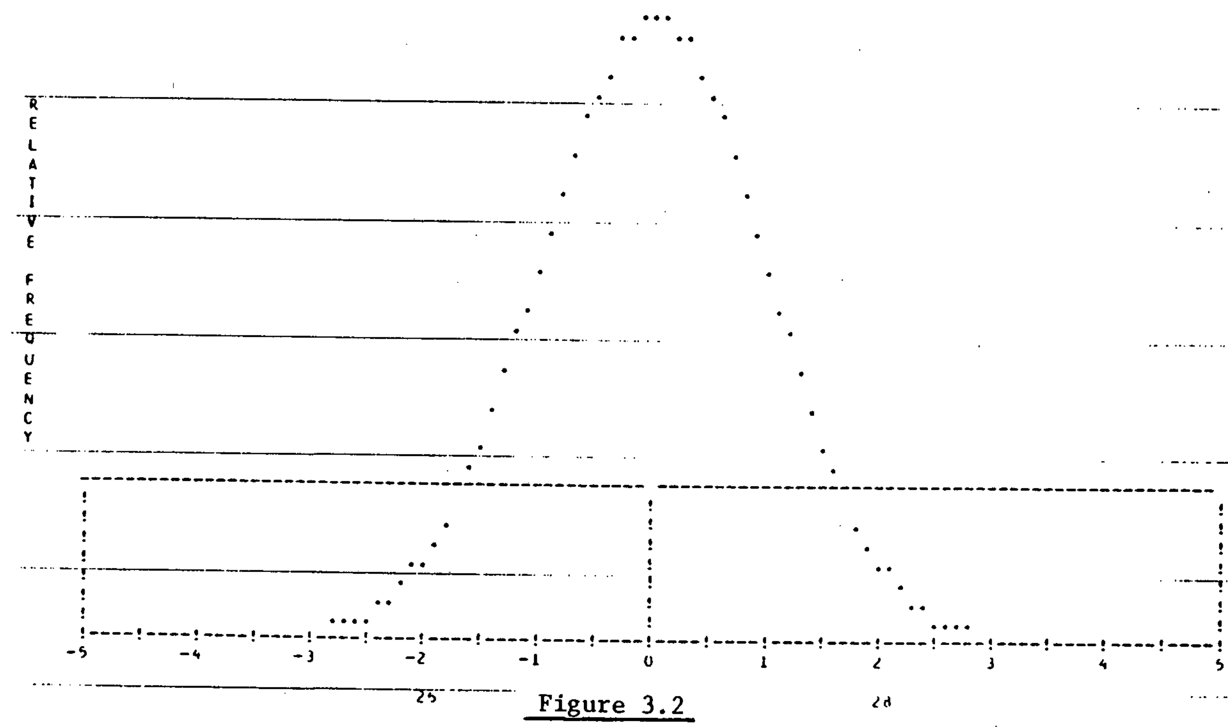
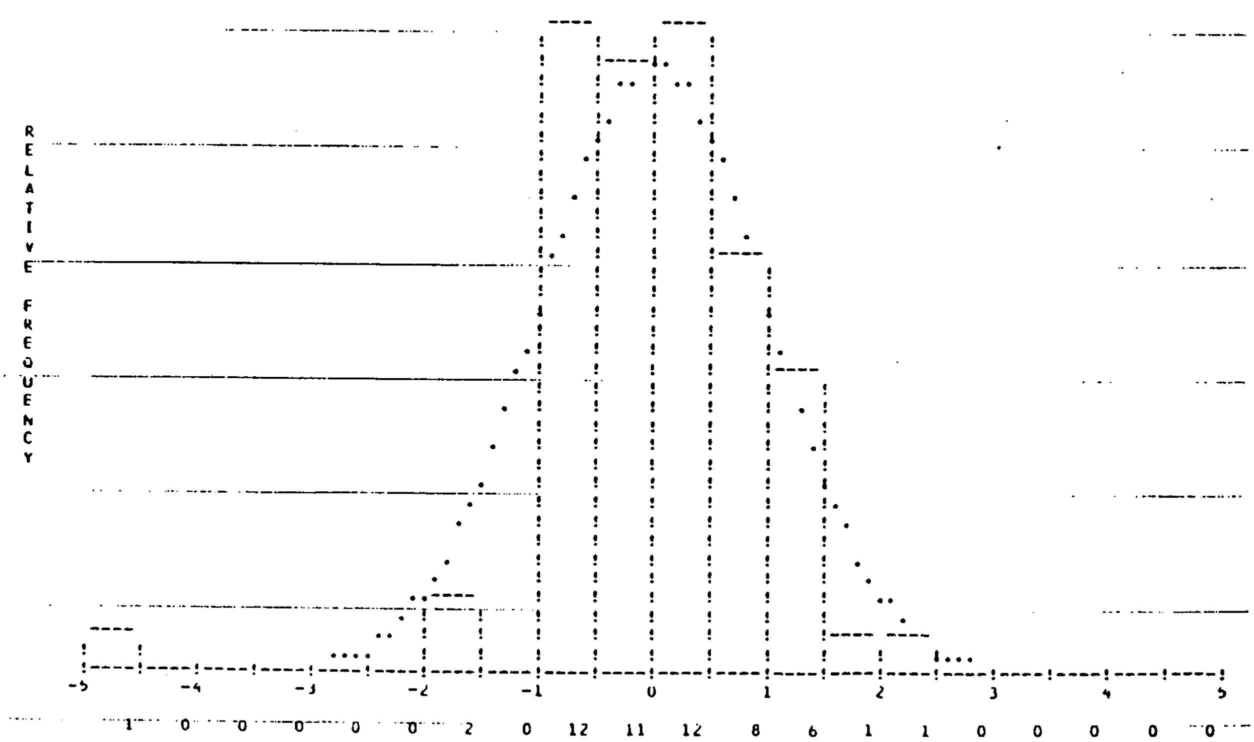


Figure 3.2

HISTOGRAM OF THE STANDARDIZED RESIDUALS (ALL RESIDUALS INCLUDED)

(WITH CLASSES AS USED IN THE GOODNESS OF FIT TEST; A MORE DETAILED REPRESENTATION IS PLOTTED ON THE NEXT PAGE)



HISTOGRAM OF THE STANDARDIZED RESIDUALS (ALL RESIDUALS INCLUDED)

Figure 3.3

3.3 Test for Outlying Estimated Residuals

Here we employ the test (equation 2.10):

$$\hat{r}_i^* < \tau_{32,0.975}$$

where equation (2.6) has been used to compute the approximate standard deviation $\hat{\sigma}_{\hat{r}_i}$, and α=0.05 (95% confidence level). The GEOPAN summary shown below indicates that one distance observation should be rejected. If in fact the distance observation is rejected and the network is readjusted the estimated variance factor $\hat{\sigma}_o^2$ is found to be in agreement with σ_o^2 at the 95% confidence level.

SUMMARY OF REJECTION OF RESIDUALS AT THE 95.000 % CONFIDENCE LEVEL

(TAU MAX CRITERION USED)

COMPUTED FACTOR FOR STANDARD DEVIATION OF RESIDUAL = 3.1084

REJECTED RESIDUALS:

	OBSERVATION	AT	FROM	TO	RESIDUAL	STD.DEV RESIDUAL	CRITICAL POINT	
41	DISTANCE	SCHAFF	SCHAFF	CITADEL	-.1049	.0223	.0694	REJECT

1 RESIDUALS (1 % OF THE OBSERVATIONS) WERE FLAGGED FOR REJECTION

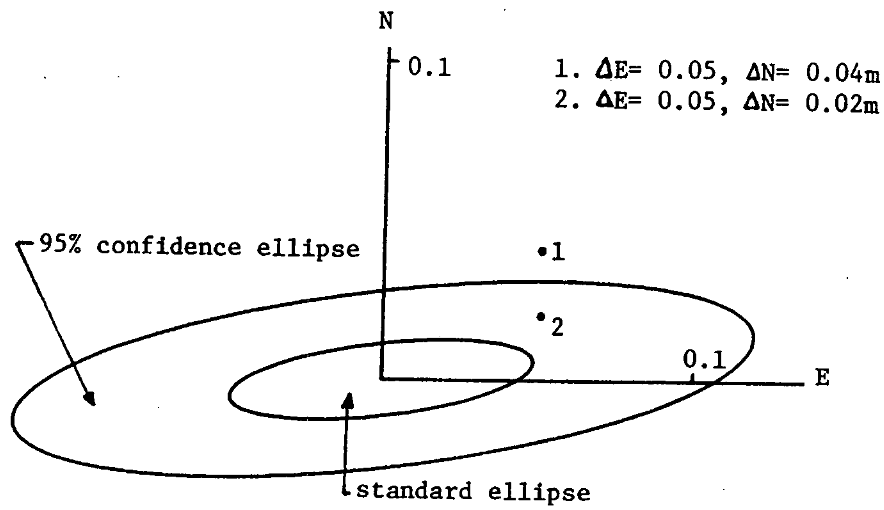
**** WARNING **** OBSERVATIONS CORRESPONDING TO REJECTED RESIDUALS HAVE BEEN USED IN THIS ADJUSTMENT

3.4 Confidence Regions for the Parameters

The standard point error ellipse, or confidence region, is shown for the station York II in Figure 3.4. Also shown in the figure is the 95% confidence ellipse, this having been obtained by multiplying the semiaxes of the standard ellipse by the factor $y^{\frac{1}{2}}$ = 2.447 (see table 2.2). The covariance matrix (in units of metres squared) for the point York II was computed to be

$$\underline{C_{\hat{x}}} = \begin{bmatrix} \sigma_E^2 & \sigma_{EN} \\ \sigma_{NE} & \sigma_N^2 \end{bmatrix} = \begin{bmatrix} 0.00237 & 0.000302 \\ 0.000302 & 0.000165 \end{bmatrix}$$

where E relates to the Easting and N to the Northing.



station : YORK II

Figure 3.4

3.5 Test for Compatibility of Estimated Parameters With Existing Independent Estimates

We now consider the situation where we have two independent estimates for the coordinates of York II, and we wish to examine the hypothesis that these values are compatible with the coordinates computed in the network adjustment. One of the independently estimated points is removed from the computed position of York II by an amount $\Delta E = 0.05$ m and $\Delta N = 0.04$ m, whereas the other has coordinates which differ from the adjusted values by $\Delta E = 0.05$ m and $\Delta N = 0.02$ m. We can plot these two points on Figure 3.4 and note that one lies within the 95% confidence region, while the other is outside the ellipse. Intuitively, we then assume that at a 95% confidence level one independently estimated point is compatible with the adjusted position, whereas the other is not.

The verification of this assumption is obtained by computing the statistic (equation 2.16)

$$y = \begin{bmatrix} \Delta E & \Delta N \end{bmatrix} \begin{bmatrix} \sigma_E^2 & \sigma_{EN} \\ \sigma_{NE} & \sigma_N^2 \end{bmatrix} \begin{bmatrix} \Delta E \\ \Delta N \end{bmatrix}$$

and testing it against $\chi^2_{2,\alpha}$. For each of the above mentioned cases, the following is found:

1. $\Delta E = 0.05$ m, $\Delta N = 0.04$ m

   $y = 10.0 > \chi^2_{2,0.05} = 5.99$

   REJECT COMPATIBILITY

2. $\Delta E = 0.05$ m, $\Delta N = 0.02$ m

   $y = 2.5 < \chi^2_{2,0.05} = 5.99$

   ACCEPT COMPATIBILITY

REFERENCES

Baarda, W. (1967): Statistical Concepts in Geodesy. Netherlands Geodetic Commission, Publications on Geodesy, New Series, Vol. 2, No. 4.

Baarda, W. (1968): A Testing Procedure for Use in Geodetic Networks. Netherlands Geodetic Commission, Publications on Geodesy, New Series, Vol. 2, No. 5.

Krakiwsky, E.J. & P. Gagnon (this volume): Least-Squares Adjustment. Papers for the CIS Adjustment and Analysis Seminars.

Mepham, M.P. & B. Nickerson (this volume): Preanalysis. Papers for the CIS Adjustment and Analysis Seminars.

Pope, A.J. (1976): The Statistics of Residuals and the Detection of Outliers. NOAA Technical Report NOS 65 NGS 1, U.S. Dept. of Commerce, Rockville, U.S.A.

Steeves, R.R. (1978): A Users' Manual for Program GEOPAN. Dept. of Surveying, University of New Brunswick, Technical Report No. 54, Fredericton.

Vaníček, P. & E.J. Krakiwsky (1982): Geodesy: The Concepts. North-Holland, Amsterdam, The Netherlands, 656 pp.