DEVELOPMENT AND TESTING OF IN-CONTEXT CONFIDENCE REGIONS FOR GEODETIC SURVEY NETWORKS

E.J. Krakiwsky, D.J. Szabo
Department of Geomatics Engineering
The University of Calgary
2500 University Drive N.W.
Calgary, Alberta T2N 1N4

P. Vaníček
Department of Geodesy and Geomatics Engineering
University of New Brunswick
P.O. Box 4400
Fredericton, New Brunswick E3B 5A3

M.R. Craymer
Geodetic Survey Division,
Geomatics Canada
615 Booth Street
Ottawa, Canada K1A 0E9

Final report prepared for

Geodetic Survey Division
Geomatics Canada
Ottawa, Canada

DSS contract file No. #032SS.23244-3-4264

April 1994
Revised February 1999

PREFACE

This is the final contract report prepared for the Geodetic Survey Division of Geomatics Canada under DSS file No. 032SS.23244-3-4264. The contract was issued on August 1, 1993 to The University of Calgary (U of C) with Dr. E.J. Krakiwsky as the principal investigator, Mr. D.J. Szabo as a research assistant at U of C, and Dr. P. Vaníček as a consultant at the University of New Brunswick under subcontract to U of C. The scientific authority for this contract was Mr. M.R. Craymer.

This report is a joint effort of the Geodetic Survey Division, The University of Calgary, and the University of New Brunswick. We also acknowledge the valuable discussions with Mr. D. Junkins and Mr. M. Pinch on the analysis of geodetic survey networks.

EXECUTIVE SUMMARY

The objectives of the contract were to develop and numerically test in-context absolute and relative confidence regions for geodetic networks. In-context confidence regions are those that relate to many points simultaneously, rather than the conventional notion of speaking about the confidence region about only one point without regard to any others.

An out-of-context test is conducted on some piece of data without regard for the remaining data in the set. An in-context test is conducted on a quantity in the context of being a member of a larger set. Adjustment software, such as GHOST and GeoLab, that use the so-called Tau test of residuals are based on in-context testing. However, we are aware of no software that is capable of performing in-context testing on confidence regions for the estimated coordinate parameters.

Another issue needing clarification is the matter of local versus global testing. Global testing is understood to be a single test involving the entire group of variates under examination. A global test statistic is typically a quadratic form which transforms the variates into a scalar quantity, containing all the information about the group. On the other hand, local testing is the process of testing individual variates in the group, either in-context or out-of-context. Since these tests can be conducted in either parameter or observation space, they should use a consistent approach in both spaces whenever possible.

The development of confidence regions corresponding to one solution is different from the statistical testing of the compatibility (or congruency) of one solution against another. In this report we focus on the development of confidence regions for the analysis of a single network solution, rather than the development of statistical tests for applications such as deformation analyses that require the comparison of two solutions.

The key issue of in-context testing is the formation of a mathematical link between the various statistical tests that may be conducted not only on the estimated parameters but also on the estimated residuals. The consequence of a mathematical link is compatibility of statistical tests throughout observation and parameter space.

Three approaches to the computation of in-context confidence regions were examined during this contract: the Bonferroni, Baarda and projection approaches. The

Bonferroni approach equates the simultaneous probability of the individual in-context confidence regions to a selected global probability level. However, it neglects any correlations between the tested quantities, which can have serious consequences for parameter confidence regions. The Baarda (or Delft) approach uses the relation between Type I and II errors for both global and local testing, but arbitrarily assumes the probability and non-centrality parameters for both local and global Type II errors are the same. Finally, the projection approach simply uses the global confidence region or test and projects it to the individual subspaces for local confidence regions or tests. It uses the global expansion factor for all individual in-context confidence regions and tests, which results in unreasonably large confidence regions that can grow without bound. Strictly speaking this is not an in-context approach as defined above. It is effectively a global test on the individual quantities. That is, the failure of one individual local test also implies the failure of the global test.

To summarize, the projection method tests hypotheses that are different from what we want and its in-context expansion factors are unreasonably large and grow without bound for large networks. Baarda's approach gives relatively large in-context expansion factors which grow without bound (although much more slowly than in the projection approach). The Bonferroni approach yields the smallest and most reasonable expansion factors for in-context confidence regions and tests. The expansion factors are also bounded to reasonable values for even the largest of networks. However, this approach neglects the effects of correlations which can be very large between coordinate parameters in geodetic networks. Primarily because of the smaller expansion factors, we recommend to use the Bonferroni approach for in-context confidence regions and tests, in spite of the neglect of correlations. It is recommended to further investigate the effects of large correlations and possible ways of accounting for them.

The recommended approach for the in-context statistical analysis of the adjustment of a geodetic network is to first chose a global significance level (α) to be used as the basis for all global and local in-context tests and confidence regions. The specific significance levels to use for the various tests and confidence regions are:

- Global test on residuals (variance factor test). Use the global significance level (α).

- Local tests of individual residuals (outlier tests). Use the in-context significance level $\alpha_o = \alpha/\nu$, where ν is the degrees of freedom of the adjustment.

- Global confidence region. For a global confidence region for all points in the network, use the global significance level (α).

- Local absolute (point) confidence regions. For absolute in-context confidence regions at individual points in the network, use the in-context significance level $\alpha_o = \alpha/n$, where n is the number of points being simultaneously assessed.

- Local relative confidence regions. For relative in-context confidence regions between pairs of points in the network, use the in-context significance level $\alpha_o = \alpha/m$, where m is the number of linearly independent pairs of points to be simultaneously assessed.

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Research Objectives

The objectives of the contract were to develop and numerically test in-context absolute and relative confidence regions for geodetic networks. The numerical testing was to be done on medium to large (100 stations to 1000 stations) horizontal and three dimensional geodetic control networks, especially GPS networks. The methodology developed was to be as mathematically rigorous as possible.

1.2 Contract Deliverables

Originally it was expected that some FORTRAN subroutine programs would be written for the Geodetic Survey Division (GSD), to be compatible with program GHOST. It was soon apparent, however, that only modifications to GSD's method of statistical testing was necessary. The parameters for in-context absolute and relative confidence regions can be computed without adjusting any networks, so no real networks were needed for the numerical testing. Instead charts were developed to show the effect of varying numbers of estimated parameters and observables on the expansion factor of the confidence regions.

1.3 Overview of Report

The research results in this report appear in four parts. Chapter 2 is a review of the standard statistical testing methods. Chapter 3 gives an explanation of various methods for in-context testing, including examples, and concludes with a recommended approach for geodetic network analysis. Chapter 4 provides detailed descriptions of the various statistical hypotheses, test statistics, and tests based on the recommended scheme of Chapter 3. Chapter 5 summarizes the conclusions and recommendations for in-context tests and confidence regions.

2. OUT-OF-CONTEXT TESTS AND CONFIDENCE REGIONS

2.1 Out-of-Context Approach

An out-of-context test is conducted on some subset of data without regard for (out of context of) the remaining data in the set [Vaníček and Krakiwsky, 1986]. Similarly an out-of-context confidence region is determined for a point (or subset of points) without regard to the others.

Before proceeding it is necessary to clarify the meaning of the terms global and local testing. Global testing is understood to be a single test involving the entire group of variates under examination. A global test statistic is typically a quadratic form which transforms the variates into a scalar quantity, containing all the information about the group. On the other hand, local testing is the process of testing individual variates in the group, either in-context or out-of-context. Global and local tests can be conducted in either parameter or observation space. In parameter space, tests are usually formed in terms of confidence regions for a single adjustment or congruency (i.e., compatibility) tests between two adjustments.

In the out-of-context approach, exactly the same methods and significance levels are used for both global and local testing or confidence regions. That is, individual residual (outlier) tests on the residuals use the same significance level as for the global (variance factor) test on all the residuals. Similarly individual confidence regions use the same significance levels as for the confidence regions for the entire network. In practice, the 5% significance level is used almost exclusively for all out-of-context testing and confidence regions.

Practically all network adjustment software packages perform out-of-context testing on the residuals and compute out-of-context confidence regions. Although some packages can also perform in-context testing of the residuals (e.g., GHOST and GeoLab), we don't know of any that compute in-context confidence regions. One exception (not really an adjustment program) is the NETVAL network analysis software [Craymer, 1990], which performs in-context congruency (compatibility) tests when comparing two 3D solutions for a network

This section reviews the standard out-of-context tests and confidence regions. These include the global (variance factor) test on all the residuals, local (outlier) tests on individual residuals, absolute confidence regions and congruency tests (global and local), and relative confidence regions.

2.2 Global (Variance Factor) Test on the Residuals

The global or variance factor test on the residuals tests whether all the standardized residuals are normally distributed with a mean of zero and standard deviation of one. It is also commonly interpreted to be a test of whether the estimated variance factor is equal to the a priori variance factor which is usually equal to one. The test is used to determine whether the adjustment as a whole is satisfactory. The statistic used for the global test is the quadratic form of the estimated residuals. It represents the length or norm of the vector of estimated observation residuals.

Mathematically, the quadratic form of the residuals is defined as

$$\|\mathbf{r}\| = \mathbf{r}^T \mathbf{C_r}^{-1} \mathbf{r} \,,$$

where $\mathbf{r}$ is the vector of estimated residuals and $\mathbf{C_r}$ is the covariance matrix of the estimated residuals. However, the $\mathbf{C_r}$ is singular since it has a rank equal to only the degrees of freedom (rank defect equals the number of observations minus the degrees of freedom).

Consequently, the quadratic form of the residuals is approximated by replacing $\mathbf{C_r}$ with the a priori observation covariance matrix $\mathbf{C_l}$; i.e.,

$$\|\mathbf{r}\| \cong \mathbf{r}^T \mathbf{C_l}^{-1} \mathbf{r} = \mathbf{r}^T \mathbf{P_l}\, \mathbf{r} = \nu\, \hat{\sigma}_o^2,$$

where $\mathbf{P_l}$ is the a priori weight matrix of the observations, ν is the degrees of freedom of the adjustment (number of observations minus number of free parameters), and $\hat{\sigma}_o^2$ is the a posteriori variance factor, computed from

$$\hat{\sigma}_o^2 = \frac{\mathbf{r}^T \mathbf{P_l}\, \mathbf{r}}{\nu} \,.$$

The global test thus becomes a test on the estimated variance factor, where the null hypothesis (Ho) is that the estimated variance factor is equal to the a priori variance factor which we here take to have been selected to equal to one. The test is formulated as:

$$\text{Ho:}\ \hat{\sigma}_o^2 = 1$$

$$\text{If}\ \chi^2_{\nu,\alpha/2} \leq \nu\hat{\sigma}_o^2 \leq \chi^2_{\nu,1-\alpha/2}\ \text{accept Ho, otherwise reject Ho.}$$

where α is the global significance level and χ^2 is the value (abscissa) of the Chi-square distribution for ν degrees of freedom. Failure of the test indicates that either (i) there are blunder(s) (outliers) in the data, (ii) the error model (covariance matrix) for the data is incorrect, or (iii) the deterministic model (design matrix) used for the adjustment is incorrect. We note, by the way, that the adjustment of GPS vectors usually fails this test because of (ii).

2.3 Local (Outlier) Tests on the Residuals

The local or outlier tests are conducted on the individual standardized residuals to check if they have a mean of zero and standard deviation of one. The tests are used to identify individual outliers. The standardized residual for the ith observation is given by

$$\frac{r_i}{\sigma_{ri}} ,$$

where r_i is the estimated residual and σ_{ri} is its standard deviation, possibly scaled by the a posteriori variance factor.

The outlier test on the ith observation is then a test of the null hypothesis (Ho) that the tested standardized residual is equal to zero. The test is formulated as

$$\text{Ho:}\ \left|\frac{r_i}{\sigma_{ri}}\right| = 0$$

$$\text{If}\ \left|\frac{r_i}{\sigma_{ri}}\right| \leq N_{1-\alpha/2}\ \text{accept Ho, otherwise reject Ho.}$$

Here, α is the global significance level and N is the value (abscissa) of the normal distribution for the case where the variance factor is considered known. When the variance factor is not known, σ_{ri} is scaled by its estimate and the either the student (t) or Tau (τ) distribution with ν degrees of freedom is used instead of the normal.

It should be noted that there is a symbiosis between the global and local tests on the residuals. A large number of residual outliers can inflate the variance factor, causing it to fail. Using this inflated variance factor in the outlier tests may result in many outliers not being detected in the local tests. To avoid this, the outlier tests should be iterated, rejecting only the very large standardized residuals (preferably) one at a time and repeating the adjustment and testing until no more outliers remain. Of course an investigation of the source of error should be conducted before rejecting an observation as an outlier. The observation should never be rejected without evidence to support the presence of an blunder or systematic error.

Note that the variance factor test could also fail because of an incorrect mathematical model. In such a case the above tests cannot be carried out until a suitable mathematical model is substituted.

2.4 Absolute Confidence Regions and Congruency Tests

Absolute or point confidence regions describe the expected random error in estimated positions, propagated from the random error in the observations. They do not in themselves constitute a statistical test, but they simply portray the hypotheses for the uncertainty in the estimated positions at a specified confidence level (most commonly 95%). Confidence regions are usually expressed in terms of the semi-axes and their orientations, derived from the eigenvalues and eigenvectors of the covariance matrix of the estimated positions.

In the case where the variance factor is considered known, the $(1-\alpha)$ confidence regions is obtained by scaling the covariance matrix by the expansion factor C

$$C = \sqrt{\chi^2_{u,1-\alpha}}$$

where u is the dimension of the point (i.e., 1, 2 or 3) and $\chi^2_{u,1-\alpha}$ is the value (abscissa) of the Chi-square distribution with u degrees of freedom and $1-\alpha$ probability level. When the variance factor is unknown, the covariance matrix is scaled by estimated variance factor and the Fisher distribution used in the following factor

$$C = \sqrt{u\, F_{\nu,u,1-\alpha}}$$

where ν is the degrees of freedom of the adjustment.

Confidence regions are traditionally computed only for individual points, out-of-context of the other points in the network. That is, the global significance level α is used in the expansion factor for each individual out-of-context (local) confidence region.

Two independent solutions for the coordinates can also be tested for congruency (compatibility). The global congruency test for equivalence of solutions is formulated as

$$\text{Ho:} \;\; \Delta\mathbf{x} = \left(\mathbf{x}^{(2)} - \mathbf{x}^{(1)}\right) = \mathbf{0}$$

$$\text{If } \left(\Delta\mathbf{x}^{T}\, C_{\Delta\mathbf{x}}\, \Delta\mathbf{x}\right) \le \chi^2_{u,1-\alpha} \text{ accept Ho, otherwise reject Ho.}$$

where $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ are the independent solutions for the coordinates, u is the total number of parameters (free coordinates) , and $\mathbf{C}_{\Delta\mathbf{x}}$ is the covariance matrix of the coordinate differences which is defined by, assuming solutions (1) and (2) are uncorrelated,

$$\mathbf{C}_{\Delta\mathbf{x}} = \mathbf{C}_{\mathbf{x}^{(1)}} + \mathbf{C}_{\mathbf{x}^{(2)}}$$

When the variance factor is unknown, the covariance matrices are scaled by estimated variance factor and the Fisher distribution used, which gives the following test:

$$\text{Ho:} \;\; \Delta\mathbf{x} = \left(\mathbf{x}^{(2)} - \mathbf{x}^{(1)}\right) = \mathbf{0}$$

$$\text{If } \left(\Delta\mathbf{x}^{T}\, \mathbf{C}_{\Delta\mathbf{x}}\, \Delta\mathbf{x}\right) \le u\, F_{\nu,u,1-\alpha} \text{ accept Ho, otherwise reject Ho.}$$

The local congruency (equivalence) of two solutions for an individual station is obtained similarly. The local congruency test for ith station is defined by

$$\text{Ho:} \;\; \Delta\mathbf{x}_i = \left(\mathbf{x}_i^{(2)} - \mathbf{x}_i^{(1)}\right) = \mathbf{0}$$

$$\text{If } \left(\Delta\mathbf{x}_i^{T}\, C_{\Delta\mathbf{x}_i}\, \Delta\mathbf{x}_i\right) \le \chi^2_{u,1-\alpha} \text{ accept Ho, otherwise reject Ho,}$$

where $\mathbf{x}_i^{(1)}$ and $\mathbf{x}_i^{(2)}$ are the independent solutions for the coordinates of the ith station, u is the number of free coordinates for the station (1, 2 or 3), and $\mathbf{C}_{\Delta\mathbf{x}_i}$ is the covariance matrix of the coordinate differences which is defined by, assuming solutions (1) and (2) are uncorrelated,

$$\mathbf{C}_{\Delta\mathbf{x}_i} = \mathbf{C}_{\mathbf{x}_i}(1) + \mathbf{C}_{\mathbf{x}_i}(2)\ .$$

When the variance factor is unknown, the covariance matrices are scaled by estimated variance factor and the Fisher distribution used, which gives the following test:

$$\text{Ho:}\ \Delta\mathbf{x}_i = \left(\mathbf{x}_i^{(2)} - \mathbf{x}_i^{(1)}\right) = \mathbf{0}\ .$$

$$\text{If}\ \left(\Delta\mathbf{x}_i^T\ \mathbf{C}_{\Delta\mathbf{x}_i}\ \Delta\mathbf{x}_i\right) \leq u\ F_{\nu,u,1-\alpha}\ \text{accept Ho, otherwise reject Ho.}$$

2.5 Relative Confidence Regions

Relative confidence regions represent the uncertainty in the estimated relative position (position difference) between a pair of points. They are derived exactly the same way as absolute confidence regions except that they are based on the covariance matrix $\mathbf{C}_{\Delta\mathbf{x}_{ij}}$ of the position difference between the ith and jth points. This relative covariance matrix is defined by

$$\mathbf{C}_{\Delta\mathbf{x}_{ij}} = \mathbf{C}_{\Delta\mathbf{x}_{ji}} = \mathbf{C}_{\mathbf{x}_i} + \mathbf{C}_{\mathbf{x}_j} - \mathbf{C}_{\mathbf{x}_{ij}} - \mathbf{C}_{\mathbf{x}_{ji}}$$

where $\mathbf{C}_{\mathbf{x}_i}$ and $\mathbf{C}_{\mathbf{x}_j}$ are the covariance matrices for the estimated positions of the ith and jth points, and $\mathbf{C}_{\mathbf{x}_{ij}}$ and $\mathbf{C}_{\mathbf{x}_{ji}}$ are the cross-covariance matrices between the ith and jth points. The same expansion factors are used for the relative confidence as for the absolute confidence regions. Note that there are a maximum of n–1 unique relative confidence regions in a network of n points.

3. IN-CONTEXT APPROACHES

3.1 In-Context Approach

When one intends to test a set of hypotheses in the context of the group to which the tested quantity belongs (e.g., testing individual estimated residuals for outliers), it is important to formulate the hypotheses such that their simultaneous probability is equal to some specified global probability; i.e., the probability that all the individual hypotheses are simultaneously true. Each individual test is then said to be conducted in the context of the other individual tests. This is referred to as “in-context” testing in Vaníček and Krakiwsky [1986] or “max” testing by Pope [1976]. The problem is to determine what significance level α_o should be used for the individual or local tests, given the significance level α used for the global tests.

Recall that this differs from the out-of-context testing only in the significance level used; out-of-context testing uses the global significance level α as well for the individual tests. Most adjustment software performs out-of-context testing of the individual residuals; some exceptions are GHOST and GeoLab. On the other hand, we know of no software that provides in-context confidence regions for the estimated parameters, even as an option (one exception is the network analysis software NETVAL [Craymer, 1990]).

The confidence level for the test on the estimated variance factor is called the global confidence level $1-\alpha$. The confidence level for individual or local residual outlier tests is called the local confidence level $1-\alpha_o$. The same confidence/significance levels can also be used in parameter space for the confidence regions, either global (confidence hyper-ellipsoid for the network as a whole) or local (for individual points). Such in-context confidence regions have already been developed by Vaníček and Krakiwsky [1986, pp. 240-241]. Here, we extend this concept to relative confidence regions and also compare the various methods of deriving the local significance levels.

We note, however, that there is no statistical requirement for the global significance levels to be the same in both observation and parameter space. Intuitively, we suspect the global significance level should take the same value in both parameter and observation space, because the global hypothesis is stated with respect to the same information in each case; the estimated parameters are merely transformed observations. However, the hypotheses are not the same in each space, which leads us also to suspect that the significance level need not be the same. The choice of probability for each of the global significance levels should be justified through previous experience with network analysis.

There are a number of different ways in which in-context testing is done. The Geodetic Survey Division (GSD) currently uses the in-context testing approach of Vaníček and Krakiwsky [1986] for residual outlier testing, which is based on Bonferroni confidence levels [Cook and Prescott, 1981]. The Bonferroni approach is an ancient method in comparison to the others, but has generally not been used by geodesists despite being documented in recent geodetic publications (e.g., Pope [1976], Vaníček and Krakiwsky [1986]). Some people outside of North America, notably Europeans, are using an in-context testing of residuals based on Baarda's reliability theory [Baarda, 1968]. On the other hand, Kubáčková and Kubáček [1993] propose an in-context method of determining confidence regions based on the Scheffe projection approach [Miller, 1969].

All three of these approaches can be used to model the inter-relation of confidence levels to provide a consistent statistical methodology for both absolute and relative confidence regions. Each of the above in-context testing methods were examined and compared from theoretical and numerical perspectives. Specifically, expansion factors for outlier testing and confidence regions are computed for each method and compared.

An important issue in the in-context testing is deciding at the outset what data to include in the global test. Although this is well defined for outlier testing, it may not be so for in-context testing. The simplest interpretation for the global confidence regions is to use all the points in the network. On the other hand, if a more focused analysis is required, (i.e., the re-determination of a certain subset of network points) then the meaning of global could be considered to include only the redetermined points. Reformulation of the global test is one of the issues dealt with in Chapter 4

The number of “free” stations in a network is referred to several times throughout the report. We define a free station as one with at least one unknown parameter (coordinate component) to be estimated. Note that the parameters in the datum constraint equations of the minimally constrained network adjustment are not included in the set of variable parameters.

3.2 Bonferroni Method

The Bonferroni method of in-context testing is based on the hypothesis that all the local tests pass simultaneously. It sets the simultaneous probability of the local tests equal to the probability of the global one. The idea is to make the local tests compatible with the global one; i.e., if a local test fails, the global one should also fail on the same confidence level. The method is based on the joint probability of independent random variables. Given the local probability $1-\alpha_o$ of a series of hypotheses on n independent variables, the simultaneous or global probability $1-\alpha$ that all hypotheses are true is given by the well-known expression (Vaníček and Krakiwsky, 1986, p.230):

$$1-\alpha = \prod_{i=1}^{n}(1-\alpha_o) = (1-\alpha_o)^n ,$$

where n is the number of independent quantities to be simultaneously tested; i.e., number of observations for outlier tests or number of stations considered for confidence regions. Re-arranging, we get the expression for the local confidence level:

$$1-\alpha_o = (1-\alpha)^{1/n} .$$

Note that in observation space, the number of independent observations is actually the number of degrees of freedom ν, not the total number of observations n. This is due to the fact that there are m constraints imposed on the observations by the observation equations. Thus, there may be some justification for using v in place of n when performing in-context outlier testing.

Of course, in practice, neither the residuals nor the estimated parameters are statistically independent random variables. Although it might be possible to transform residuals and parameters into uncorrelated quantities for which the above would apply, such transformed variables would not likely have any simple relation to the quantities we need to test (residuals and parameters).

To circumvent this problem, Vaníček and Krakiwsky (1986, p.230) make use of Bonferroni's inequality for correlated random variables:

$$1-\alpha \leq (1-\alpha_o)^n \quad \text{or} \quad 1-\alpha_o \geq (1-\alpha)^{1/n} .$$

The right hand side of the second expression gives the lower bound for the local confidence level in the case of correlated observations. Although a method has been developed to estimate the upper bound for the one dimensional case [Cook and Prescott, 1981], we are not aware of any method for quantifying the upper bound for correlated multidimensional data.

Using the lower bound of the local confidence level (i.e., assuming the data to be independent) will give smaller confidence levels (larger significance levels), resulting in smaller testing limits. This only results in the detection of false outliers (increased Type I error) and thus errors on the side of caution. Note, however, that this also gives confidence regions that are smaller than they should be.

The in-context expansion factor scales the standard error ellipse (2D) to give the in-context confidence ellipse corresponding to the global confidence level $(1-\alpha)$. It is given by:

$$C_\alpha = \chi^2_{u,1-\alpha_o} \qquad \text{(variance factor known)}$$

$$C_\alpha = u\, F_{u,\nu,1-\alpha_o} \qquad \text{(variance factor unknown)}\ ,$$

where χ^2 is the abscissa of the Chi-square distribution, F is the abscissa of the Fisher distribution, α_o is the local significance level derived from the lower bound of the confidence level ($\alpha_o = 1-(1-\alpha)^{1/m}$), m is the number of local parameters divided by the total number of parameters being tested simultaneously (usually equal to the number of stations), u is the dimension of the individual confidence region (number of local parameters), and ν is the number of degrees of freedom of the adjustment. Figure 1 illustrates the growth of the expansion factor for the 1D in-context local confidence interval as a function of the total number of confidence regions (stations) being tested simultaneously (global significance level $\alpha = 5\%$). This method gives the smallest expansion factors in comparison to the other two methods.

An attempt was also made to formulate a correction to the Bonferroni lower bound that would account for correlations among the parameters. The basic approach was to partition the quadratic form (y) of the estimated parameters into two components (y1 and y2) such that

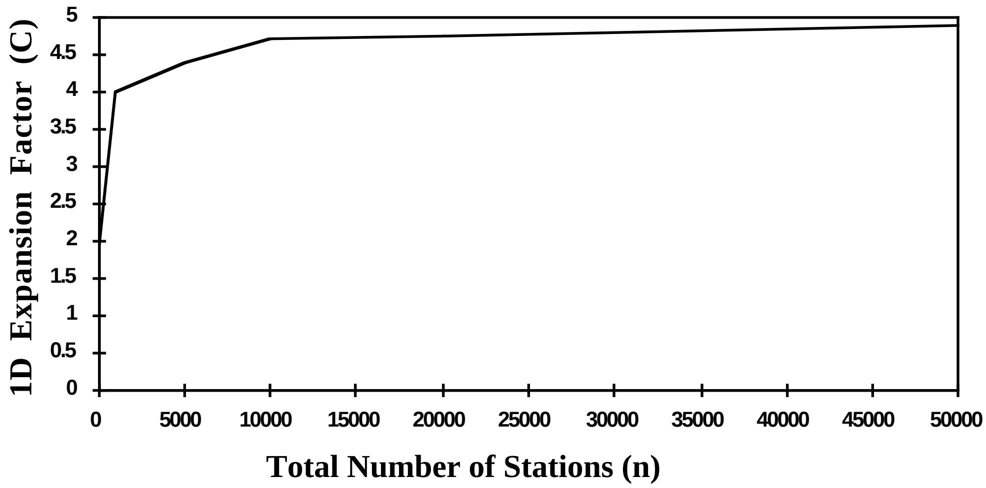

Figure 1. Expansion factor (C) for in-context local confidence intervals (1D) as a function of the total number of stations (n), based on the Bonferroni's method with global confidence level $(1-\alpha) = 95\%$.

$$y = \mathbf{x}^T \mathbf{P}\, \mathbf{x} = y_1 + y_2 ,$$

where

$$y_1 = \mathbf{x}^T \mathbf{P}_1 \mathbf{x} ,$$

$$y_2 = \mathbf{x}^T \mathbf{P}_2 \mathbf{x} .$$

and **P** is the weight matrix computed from original fully populated covariance matrix ($\mathbf{P}=\mathbf{C}^{-1}$). This amounts to partitioning the original weight matrix **P** into $\mathbf{P} = (\mathbf{P}_1+\mathbf{P}_2)$. In particular, we have partitioned **P** so that

$$\mathbf{P}_1 = \text{diag}(\mathbf{C})^{-1} ,$$

$$\mathbf{P}_2 = \Delta\mathbf{P} = \mathbf{P} - \mathbf{P}_1 ,$$

where diag(**C**) is the diagonal part of the original covariance matrix (or block diagonal part for 2D or 3D stations). The first quadratic form (y_1) corresponds to our assumption of independent variables in applying the Bonferroni method and clearly has a Chi-square distribution (due to the use of the diagonal covariance matrix). If the distribution the second term (y_2) can also be shown to have a Chi-square distribution ,then it could be used to correct the first (y_1). But there are some rather basic problems with this approach and further research into this approach is needed.

3.3 Baarda (Reliability) Method

Baarda's method for in-context testing is similar to the Bonferroni approach in that it tries to determine a local significance level α_o that corresponds to the global one α. The difference is that the relation between the two is sought in terms of the probability of both Type I and Type II errors via the assumptions built into Baarda's reliability theory [Baarda, 1968]. Figure 2 illustrates probabilities of both the Type I and Type II errors for an alternative hypothesis for the local tests and the normal distribution. A Type I error (α_o) represents the error in rejecting a good observation as an outlier (a good observation fails the outlier test). A Type II error (β_o) is the error in accepting a outlier as a good observation (i.e., a true outlier passes the outlier test). The difference between the means of the null and alternative hypotheses is called the noncentrality parameter (λ). A similar diagram can be constructed for the Chi-square distribution, where the probabilities of committing Type I and Type II errors are represented by global significance level α and β, respectively.

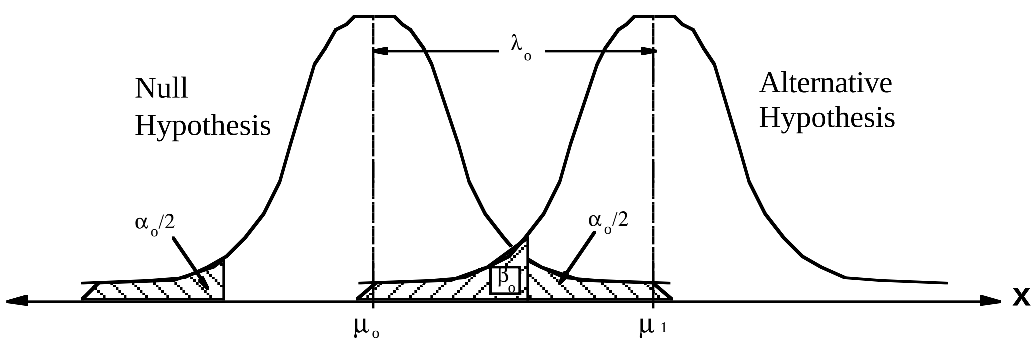

Figure 2. Type I and Type II errors.

In deriving the relation between the local and global probability levels, Baarda assumes that the local and global β and λ values are the same ($\beta=\beta_o$ and $\lambda=\lambda_o$). Given the global probability levels α and β, the global noncentrality parameter λ can be computed. Based on Baarda's assumptions, this then defines the local values $\beta_o=\beta$ and $\lambda_o=\lambda$, which can then be used to compute the local significance level α_o. This is the same procedure as that used in Baarda's reliability analysis and is illustrated in Figure 3. For details of the computations of λ and α_o, see Baarda [1968] or Krakiwsky et al. [1993].

An example of the resulting in-context local expansions factors based on Baarda's local significance levels is given in Figure 4 for the local 1D confidence interval as a function of the total number of local confidence regions (stations) being tested simultaneously. Note that these values are larger than those based on the Bonferroni approach.

$$\begin{array}{ccccc} \alpha & \rightarrow & \beta & \rightarrow & \lambda \\ & & \Downarrow & & \Downarrow \\ \alpha_o & \leftarrow & \beta_o & \leftarrow & \lambda_o \end{array}$$

Figure 3: Graphical illustration of steps in determination of local significance level following Baarda's procedure used for reliability analysis.

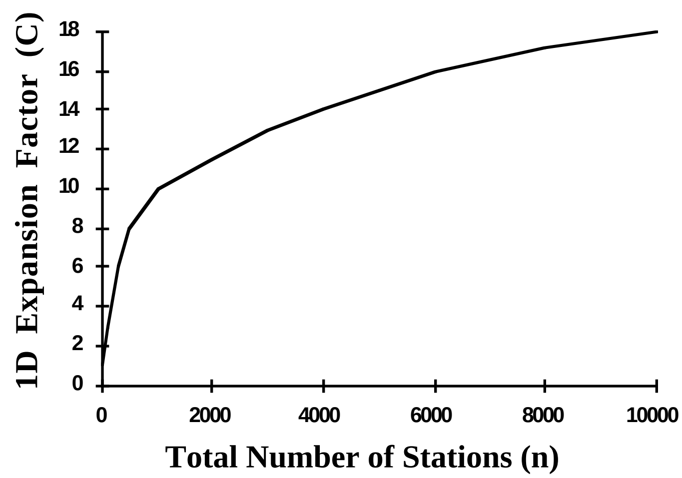

Figure 4. Expansion factor (C) for in-context local confidence intervals (1D) as a function of the total number of stations (n), based on the Baarda's method with global confidence level $(1–\alpha)$ = 95%.

3.4 Scheffe (Projection) Method

The projection method (attributed by us to Scheffe, a well-known statistician) is quite different from the previous two already discussed. The method is based on projecting the global confidence regions onto the local subspaces corresponding to each point in the network. This results in local confidence regions where solutions falling outside the region are also guaranteed to fall outside the global confidence region. However, the converse is not true. There are cases where the global test can fail while all the local tests can pass. This is because the totality of the local tests produces a global one that is larger than the original. This is illustrated in Figure 5 for the case where a 2D confidence region is projected onto the two 1D components.

The Scheffe local confidence regions are obtained by scaling the local confidence regions by the same expansion factor as used for the global one. This obviously produces scale factors that are much larger than either the Bonferroni or Baarda methods, especially for large networks. Moreover, the expansion factor grows without bound as the size of the network increases. Figure 6 shows this behaviour for the expansion factors for the local 2D confidence regions as a function of the total number of stations.

The Scheffe method can also be applied to outlier tests, in which case it would actually be testing the hypothesis that a particular observation does not cause the global (variance factor) test to fail. Thus, the magnitude of a single outlier would need to be very large to cause it to fail in the Scheffe approach. This is different from the usual approach, where the test on the variance factor (the global test) should pass even if a few moderately sized outliers are present.

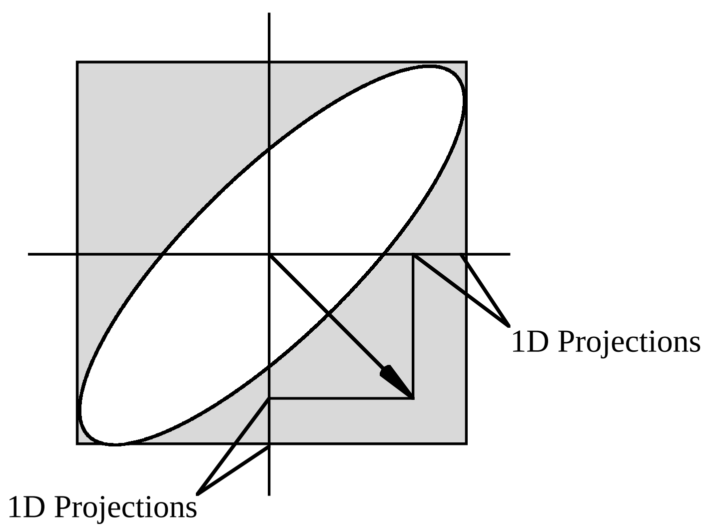

Figure 5. Example of outlier (arrow) in the global test which lies outside the confidence circle, but is not detectable as an outlier in the local Scheffe tests (projections of outlier fall inside projection of confidence region). The shaded areas denote all such regions where outliers in the global would not cause the local tests to fail.

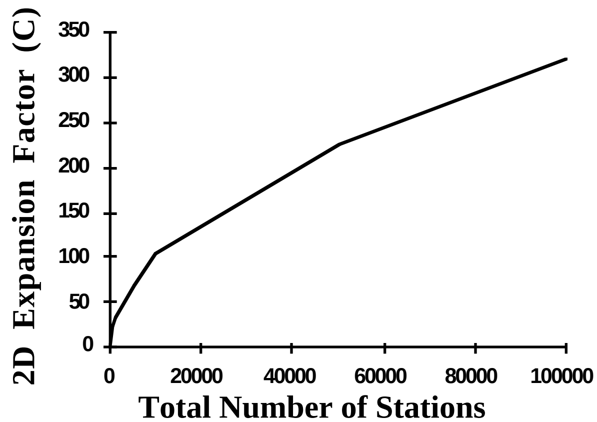

Figure 6. Expansion factor (C) for in-context local confidence intervals (2D) as a function of the total number of stations, based on the Scheffe's projection method with global confidence level (1–α) = 95%.

3.5 Recommended Method

We propose to adopt the Bonferroni in-context confidence level method, because the statistical hypotheses agree with what GSD wishes to test, and the confidence levels err on the side of caution for the case of correlated data. In addition, the expansion factors resulting from Bonferroni confidence levels are the smallest of all the methods – a criteria that Miller [1966] asserts is good for choosing the best testing scheme:

"In any problem it is ethical as well as theoretically sound for the statistician to compute the critical points for both regions and select whichever region is smaller". [Miller, 1966]

The Scheffe projection method is rejected because it is really testing a different hypothesis which should be considered inappropriate for individual in-context testing. Baarda's approach is rejected because of the questionable assumptions made regarding the equivalence of the global and local non-centrality parameters. In addition, there is no clear justification for choosing a particular power of the test (β) in the formula. The usually adopted value of 20% is actually based only on the empirical studies of the Dutch geodetic network and likely doesn't not apply generally to other networks [Teunissen, 1993].

4. IN-CONTEXT CONFIDENCE REGIONS AND STATISTICAL TESTS

4.1 Purpose of Confidence Regions and Congruency Tests

The main purpose for testing the statistical congruency (compatibility) of parameters from two network solutions is to check for network deformations (or incongruencies) between epochs. This may be done for the network as a whole (global) or for individual points or even groups of points (local). The main purpose of confidence regions, on the other hand, is to quantify the precision (by means of the propagation of random errors) of the estimated parameters. Thus the confidence regions do not in themselves constitute a test, rather they describe how precisely the station coordinates are determined.

4.2 Global Confidence Regions and Tests

These confidence regions and tests are the same for both out-of-context and in-context testing (see Chapter 2). The global confidence level (α) is used for these regions and tests. Figure 7 illustrates the growth of the resulting expansion factor as a function of the number of parameters.

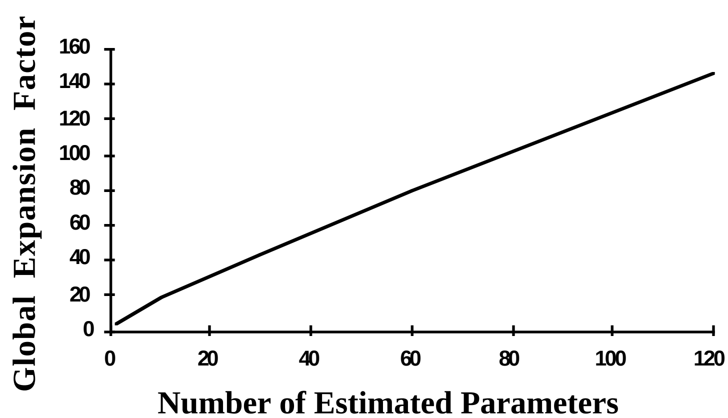

Figure 7. Global expansion factor for a 95% confidence level as a function of the total number of parameters.

4.3 Local (Outlier) Tests on Residuals

The in-context local or outlier tests are conducted on the individual standardized residuals to check if they have a mean of zero and standard deviation of one. The tests are used to identify individual outliers in the context of other observations being present. The test is identical to the out-of-context test (see §2.3) except that the in-context significance level (α_o) is used instead.

Using the Bonferroni in-context approach, the in-context outlier test on the ith observation is then a test of the null hypothesis (Ho) that the standardized residual is equal to zero. The test is formulated as

$$\text{Ho: } \left|\frac{r_i}{\sigma_{ri}}\right| = 0$$

$$\text{If } \left|\frac{r_i}{\sigma_{ri}}\right| \leq N_{1-\alpha/2k} \text{ accept Ho, otherwise reject Ho.}$$

Here, α is the global significance level, k is the number of independent observations being tested simultaneously, and N is the abscissa of the normal distribution for the case where the variance factor is considered known. Either the total number of observations (n) or the degrees of freedom (ν) may be used for k. When the variance factor is not known, σ_{ri} is scaled by its estimate and the either the student (t) or Tau (τ) distribution with ν degrees of freedom is used instead of the normal.

4.4 Local Absolute Confidence Regions and Congruency Tests

The local in-context absolute confidence regions provide confidence regions where the simultaneous probability of all regions equals the global one. The in-context local regions are identical to the local out-of-context ones (see §2.4) except that the expansion factor C for the local in-context confidence regions is defined using the Bonferroni significance level: i.e.,

$$C = \sqrt{\chi^2_{u,1-\alpha/k}} \,,$$

where u is the dimensionality of the point coordinates (i.e., 1, 2 or 3) and $\chi^2_{u,1-\alpha/k}$ is the abscissa from the Chi-square distribution with u degrees of freedom and $1-\alpha/n$ probability level ($k < m$ is the total number of stations to be tested simultaneously in a network of m free stations). When the variance factor is unknown, the covariance matrix is scaled by estimated variance factor and the Fisher distribution used in the following factor

$$C = \sqrt{u\, F_{\nu,u,1-\alpha/k}}$$

where ν is the degrees of freedom of the adjustment.

An example of the magnitude of the in-context expansion factors for 2D local confidence regions is given in Figures 8 and 9 for the 95% confidence level and the case where the variance factor is considered known (i.e., the Chi-square distribution is used).

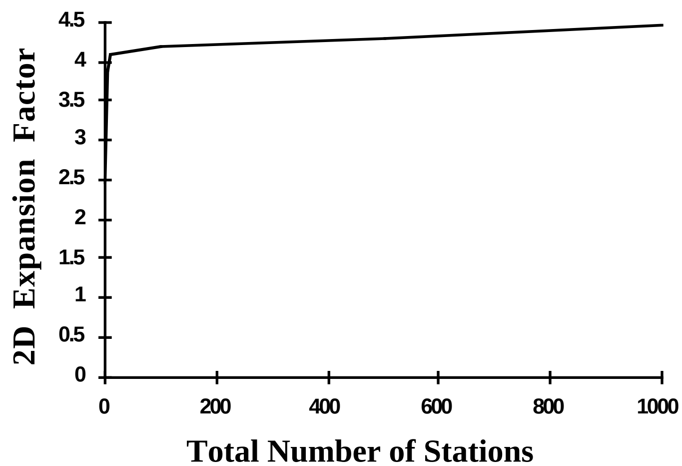

Figure 8. Local 2D expansion factors for a 95% confidence level as a function of the total number of stations.

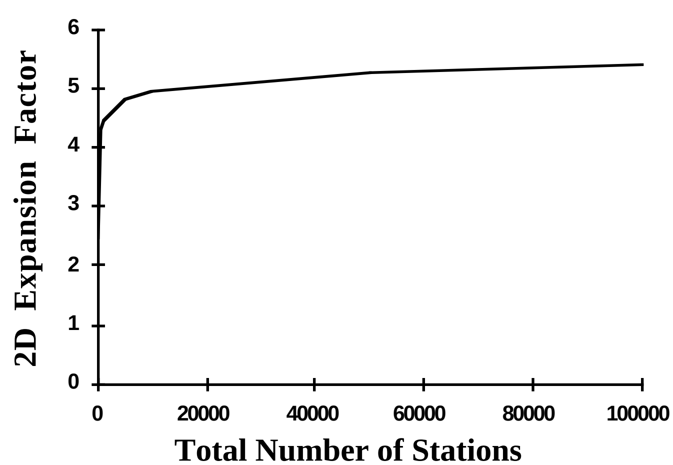

Figure 9. Local 2D expansion factors for a 95% confidence level as a function of the total number of stations in large networks.

Two independent solutions for the coordinates of a station can also be tested for congruency (compatibility). The local congruency test for equivalence is identical to the out-of-context one (see §2.4), except that the Bonferroni in-context confidence level is used. The test is thus formulated as

$$\text{Ho:} \ \Delta \mathbf{x}_i = \left(\mathbf{x}_i^{(2)} - \mathbf{x}_i^{(1)} \right) = \mathbf{0}$$

$$\text{If} \ \left(\Delta \mathbf{x}_i^T \, \mathbf{C}_{\Delta \mathbf{x}_i} \, \Delta \mathbf{x}_i \right) \leq \chi^2_{u,1-\alpha/n} \ \text{accept Ho, otherwise reject Ho.}$$

where $\mathbf{x}_i^{(1)}$ and $\mathbf{x}_i^{(2)}$ are the independent solutions for the coordinates of the ith point, and $\mathbf{C}_{\Delta \mathbf{x}_i}$ is the covariance matrix of the coordinate differences which is defined by, assuming solutions (1) and (2) are uncorrelated,

$$\mathbf{C}_{\Delta \mathbf{x}_i} = \mathbf{C}_{\mathbf{x}_i^{(1)}} + \mathbf{C}_{\mathbf{x}_i^{(2)}} \ .$$

When the variance factor is unknown, the covariance matrices are scaled by estimated variance factor and the Fisher distribution used, which gives the following test:

$$\text{Ho:} \ \Delta \mathbf{x}_i = \left(\mathbf{x}_i^{(2)} - \mathbf{x}_i^{(1)} \right) = \mathbf{0}$$

$$\text{If} \ \left(\Delta \mathbf{x}_i^T \, \mathbf{C}_{\Delta \mathbf{x}_i} \, \Delta \mathbf{x}_i \right) \leq u \, F_{\nu,u,1-\alpha} \ \text{accept Ho, otherwise reject Ho.}$$

4.5 Relative Confidence Regions

In-context relative confidence regions provide confidence regions where the simultaneous probability of all relative regions equals the global absolute confidence region. Again, the in-context relative regions are identical to the out-of-context ones (see §2.5) except that the expansion factor C for the local in-context confidence regions is defined using the Bonferroni significance level: i.e.,

$$C = \sqrt{\chi^2_{u,1-\alpha/k}}$$

where u is the dimension of the point (i.e., 1, 2 or 3), $\chi^2_{u,1-\alpha/n}$ is the abscissa from the Chi-square distribution with u degrees of freedom and $1-\alpha/m$ probability level and k is the total number of unique (linearly independent) relative error ellipses. When the variance factor is unknown, the covariance matrices is scaled by estimated variance factor and the Fisher distribution used in the following factor

$$C = \sqrt{u\ F_{\nu,u,1-\alpha/k}}$$

where ν is the degrees of freedom of the adjustment.

Note that k is the total number of linearly independent relative confidence regions that can be formulated and tested simultaneously. Although k defines the maximum number of independent quantities being simultaneously tested in the expression for the expansion factor, it does not prevent the computation of more or fewer in-context confidence regions. It only defines the scaling of these confidence regions.

Finally, the total number (m) of free stations being tested simultaneously does not necessarily have to equal the total number of stations in the network. An analyst may only be interested in testing a subset of points (perhaps only the new points in a densification network).

5. RECOMMENDATIONS

Although the aim of this report was to develop in-context confidence regions for geodetic networks, additional research into in-context testing of residuals was necessary to maintain a consistent approach throughout the process of network analysis. Consequently, some of the recommendations are with respect to residual testing, and they should be considered along with those for in-context confidence regions for a consistent and compatible methodology for the statistical analysis of both observations and parameters.

The recommendations of this report are summarized as follows:

1. A single, constant global confidence level $(1-\alpha)$ should be chosen and used for all statistical testing and analyses of both observations and parameters.

2. An in-context confidence level of $1-(1-\alpha)^k \cong (1-\alpha/k)$ should be used in the local outlier tests following Vaníček and Krakiwsky [1986], where k represents either the number of observations or the maximum number of independent residuals (i.e., the degrees of freedom of the solution).

3. An in-context confidence level of $1-(1-\alpha)^k \cong (1-\alpha/k)$ should be used in the computation of the expansion factor for the local (individual) in-context absolute confidence regions for each station in a network of m free (non-fixed) stations being tested simultaneously, where k represents the number of stations to be assessed simultaneously $(k \leq m)$.

4. An in-context confidence level of $1-(1-\alpha)^k \cong (1-\alpha/k)$ should be used to compute the expansion factor for the individual in-context relative confidence regions for each pair of stations in a network, where k represents the number of linearly independent pairs of points that can be assessed simultaneously.

5. Investigate further the meaning and usefulness of the global confidence region.

REFERENCES

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

Biacs, F.Z. and W.F. Teskey (1990). Deformation Analysis of Survey Networks with Interactive Hypothesis Testing and Computer Graphics. *CISM Journal*, Vol. 44, No. 4, pp 403-416.

Cook, R.D. and P. Prescott (1981). On the Accuracy of Bonferroni Significance Levels for Detecting Outliers in Linear Models. *Technometrics*, Vol. 23, No.1, pp 59-63.

Craymer, M.R. (1990). NETVAL: Network Validation Software Reference Manual. IBM PC Version 1.0. Geodetic Research Services Ltd., Fredericton, New Brunswick.

Feller, W. (1966). *An Introduction to Probability Theory and Its Applications* (Volume II). John Wiley & Sons, Inc., New York.

Hamilton, W.C. (1964). *Statistics in Physical Science*. Ronald Press Company, New York

Hogg, R.V. and A.T. Craig (1965). *Introduction to Mathematical Statistics* (2nd Ed.). The Macmillan Company, New York.

John, J.A. and P. Prescott (1975). Critical Values of a Test to Detect Outliers in Factorial Experiments. *Applied Statistics*, Vol. 24, No. 1, pp 56-59.

Koch, K.R. (1987). *Parameter Estimation and Hypothesis Testing in Linear Models*, Springer-Verlag, New York

Kok, J.J. (1984). On data snooping and multiple outlier testing. NOAA Technical Report NOS NGS 30, National Geodetic Information Center, Rockville, MD.

Kosters, A.J.M. and J.J. Kok (1984). Statistical testing and quality analysis of observations and transformation parameters in combining 3-dimensional networks. Proceedings of the IAG General Meeting, Symposium 102, The Global Positioning System and Other Radio Tracking Systems, Edinburgh, pp. 215-240.

Krakiwsky, E.J., P. Vaníček and D.J. Szabo (1993). Further Development and Testing of Robustness Analysis. Contract Report, Geodetic Survey Division, Geomatics Canada, Ottawa.

Kubáčková, L. and L. Kubáček (1993). Confidence Regions in Recent Crustal Movements. Contributions of the Geophysical Institute of the Slovak Academy of Sciences.

Miller, K.S. (1966). *Simultaneous Statistical Inference*. McGraw-Hill Book Company, New York.

Pope, A.J. (1976). The Statistics of Residuals and the Detection of Outliers. NOAA Technical Report NOS 65 NGS 1, National Geodetic Information Center, Rockville, MD.

Srivastava, M.S. and E.M. Carter (1983). *An Introduction to Applied Multivariate Statistics*. North-Holland, New York.

Teunissen, P. (1993). Personal communication, Stuttgart, Germany, September 8.

Vaníček, P. and E.J. Krakiwsky (1986). *Geodesy: The Concepts* (2nd Ed.). North-Holland, Amsterdam.

Walpole, R.E. and R.H. Myers (1985). *Probability and Statistics for Engineers and Scientists 3rd Ed*. Macmillan Publishing Company, New York