

Statistical Inference in Integrated Geodesy

A. Dermanis & D. Rossikopoulos

Department of Geodesy and Surveying
University of Thessaloniki

IUGG XXth General Assembly
International Association of Geodesy
Vienna, August 11-24, 1991

1. Introduction

Statistical inference is concerned with methods of analysing observations, which are outcomes of random variables, in order to obtain information concerning the parameters of the distribution of these random variables. This broad term covers the estimation of the unknown parameters of the distributions (point estimation), as well as, the estimation of confidence intervals (interval estimation). Here statistical inference is used with respect to its second aspect, which is directly related to the performance of statistical tests of hypotheses (**Lehmann**, 1986). Such tests are an indispensable tool for making inferences about the validity of the models used in data analysis, especially in applications, such as integrated geodesy, where a plethora of models is available.

Integrated geodesy is also a term that has been used with different meanings, varying from a technique for the adjustment of three-dimensional networks, to just another application of the so called mixed linear model (or least squares collocation in geodetic terminology). Here integrated geodesy is viewed as a class of methods for the analysis of discrete geodetic observations which depend not only on a discrete number of unknown parameters (as in classical adjustment), but also on one or more unknown functions.

This concept is a straight-forward generalization of Krarup's original concept (**Eeg and Krarup**, 1975), where not only the gravity potential of the earth but any other type of function is considered. Furthermore, models and solutions of different types are included, in contrast to the traditional standard collocation approach with either its stochastic (minimum mean square error of prediction) or its deterministic (minimum norm) interpretation.

A survey of different modeling and solution alternatives has been presented in **Dermanis and Rossikopoulos** (1988), for cases where the underlying functions are the earth gravity potential considered as a function of time also (in addition to space), and the time-varying shape of the earth masses in the network area.

The present study is concerned with the possibilities of applying statistical tests in order to investigate the "admissibility" of a certain model or to compare competing models whenever possible. Although the presented results will be valid for any underlying functions, the time varying gravity potential and three-dimensional deformation will be primarily investigated.

Another application of statistical tests is the discrimination, not between different models describing the same physical situation, but rather between models describing different states of nature, for example the presence of crustal deformation or not. This latter category includes the presence of observational blunders, so far the more extensive application of hypothesis testing in geodesy.

Hypothesis testing has been introduced to geodesy by the pioneering work of W. Baarda (**Baarda**, 1967, **Baarda**, 1968), for the case of the linear model with only deterministic parameters. Integrated geodesy involves also stochastic parameters (signals) and therefore relates to the mixed linear model or (least squares) collocation in geodetic terminology. The problem of hypothesis testing in relation to collocation has been studied by **Wei** (1987a, 1987b), **Krakiwsky** and **Biacs**(1990) and in relation to the mixed or random effects linear model by **Schaffrin** (1987, 1988) and **Bock and Schaffrin** (1988).

It must be emphasized that statistical tests are more powerful in pointing out inappropriate models, rather than establishing the validity of any model. There is always the possibility that the real world is different than what the model assumes, but the available observations cannot "sense" the difference. As an extreme example consider the most famous geodetic experiment, that of Eratosthenes: the same data could be used as well for the calculation of the distance of the sun from a flat earth!

2. Test of hypotheses

Estimation of fixed (deterministic) parameters and prediction (=estimation of outcomes) of stochastic parameters, is performed in geodesy within the framework of "linear inference". This involves the use of optimality criteria which incorporate only means and covariances of random parameters and thus lead to solutions that are essentially distribution-free. On the contrary, tests of hypotheses depend on the adoption of specific distribution types for the stochastic parameters of the used model and therefore for the observations themselves. It is not necessary to know the specific distribution of the random variables, but only a class of distributions where the actual distribution lies, parametrized through a set of unknown parameters, which are also to be estimated. Such unknown parameters are the deterministic parameters of

the model, parameters in partly known means (trend parameters) and parameters in covariances (variance and covariance components). From the statistical point of view they are simply unknown parameters in the mean vector and covariance matrix of the observations - viewed as random variables whose outcome are the available observational data.

A hypothesis is a relation between the unknown parameters in the mean vector or in the covariance matrix of the observations. Statistical tests are a means for deciding between two competing hypothesis, the null hypothesis H_0 and the alternative hypothesis H_a . The presently available techniques are restricted to linear relations involving either the parameters of the mean vector, or the parameters of the covariance matrix of the observations. Such linear hypotheses are consequently distinguished in the so called "general linear hypothesis" and hypotheses about variance components (usually a single variance component). A further restriction is the assumption that the observations should follow the normal distribution, which may not be valid for the actual observations. The robustness of the various statistical tests with respect to the normality assumption is a subject of statistical research.

In order to test a hypothesis, a statistic is introduced, which is a function of the observations and therefore a random variable itself. On the basis of the distribution of the statistic, as restricted by the null hypothesis, a confidence interval for the statistic is estimated. The null hypothesis is accepted if the outcome of the statistic, computed from the available outcomes of the observations, falls within the confidence interval.

2.1. The general linear hypothesis

The general linear hypothesis is performed in relation to the Gaus-Markov model

$$(1) \quad \mathbf{b} = \mathbf{Ax} + \mathbf{e},$$

$$(2) \quad E\{\mathbf{e}\} = \mathbf{0}, \quad D(\mathbf{e}) = E\{\mathbf{e}\mathbf{e}^T\} = \sigma^2 \mathbf{P}^{-1} \quad \Rightarrow$$

$$(3) \quad E\{\mathbf{b}\} = \mathbf{Ax}, \quad D(\mathbf{b}) = E\{(\mathbf{b} - \mathbf{Ax})(\mathbf{b} - \mathbf{Ax})^T\} = \sigma^2 \mathbf{P}^{-1}$$

The hypotheses to be tested involve k linear relations on the m parameters \mathbf{x}

$$(4) \quad H_0: \mathbf{Hx} = \mathbf{z} \quad \text{vs} \quad H_a: \mathbf{Hx} \neq \mathbf{z}.$$

Let $\hat{\mathbf{x}}, \hat{\mathbf{e}}, \hat{\sigma}^2$, be the solution given by

$$(5) \quad \hat{\mathbf{x}} = (\mathbf{A}^T \mathbf{PA})^{-1} \mathbf{A}^T \mathbf{b},$$

$$(6) \quad \hat{\mathbf{e}} = \mathbf{b} - \mathbf{A}\hat{\mathbf{x}}$$

$$(7) \quad \hat{\sigma}^2 = \frac{\hat{\mathbf{e}}^T \mathbf{P}\hat{\mathbf{e}}}{f}$$

where f are the degrees of freedom.

The test statistic, which follows the F distribution, is

$$(8) \quad F = \frac{(\mathbf{H}\hat{\mathbf{x}} - \mathbf{z})^T (\mathbf{HN}^{-1} \mathbf{H}^T)^{-1} (\mathbf{H}\hat{\mathbf{x}} - \mathbf{z})}{k \hat{\sigma}^2} \sim F_{k,f}$$

The hypothesis $H_0: \mathbf{Hx} = \mathbf{z}$ is accepted over the alternative $H_a: \mathbf{Hx} \neq \mathbf{z}$, when $F \leq F_{k,f}^\alpha$.

A useful form of the general hypothesis involves the comparison of the model $\mathbf{b} = \mathbf{Ax} + \mathbf{e}$ with an "ex-

tended" model $\mathbf{b} = \mathbf{Ax} + \mathbf{Zy} + \mathbf{e}$. The null hypothesis with respect to the extended model is simply $\mathbf{y} = \mathbf{0}$ and the test statistic is (Dermanis, 1987)

$$(9) \quad F = \frac{f-q}{f} \frac{\hat{\mathbf{y}}^T \mathbf{Q}_{\hat{\mathbf{y}}}^{-1} \hat{\mathbf{y}}}{f\hat{\sigma}^2 - \hat{\mathbf{y}}^T \mathbf{Q}_{\hat{\mathbf{y}}}^{-1} \hat{\mathbf{y}}} \sim F_{q, f-q}$$

where q is the number of the additional parameters \mathbf{y} , and

$$(10) \quad \hat{\mathbf{y}} = \mathbf{Q}_{\hat{\mathbf{y}}} \mathbf{Z}^T \mathbf{P} \hat{\mathbf{e}}$$

$$(11) \quad \mathbf{Q}_{\hat{\mathbf{y}}} = [\mathbf{Z}^T \mathbf{P} (\mathbf{P}^{-1} - \mathbf{AN}^{-1} \mathbf{A}^T) \mathbf{PZ}]^{-1}$$

$$(12) \quad \hat{\mathbf{y}}^T \mathbf{Q}_{\hat{\mathbf{y}}}^{-1} \hat{\mathbf{y}} = \hat{\mathbf{e}}^T \mathbf{PZ} [\mathbf{Z}^T \mathbf{P} (\mathbf{I} - \mathbf{AN}^{-1} \mathbf{A}^T) \mathbf{P}]^{-1} \mathbf{Z}^T \mathbf{P} \hat{\mathbf{e}}$$

with a matrix inversion of the order of the additional parameters. In equations (9) to (12) $\hat{\mathbf{e}} = \mathbf{b} - \mathbf{Ax}$ and $\hat{\sigma}^2$ are the residuals and the variance estimate, respectively, from the original model $\mathbf{b} = \mathbf{Ax} + \mathbf{e}$. The hypothesis $H_0: \mathbf{y} = \mathbf{0}$ is accepted over the alternative $H_a: \mathbf{y} \neq \mathbf{0}$, when $F \leq F_{q, f-q}^{\alpha}$.

An alternative form of the same test is

$$(13) \quad F = (f-q) \frac{r^2}{f - qr^2} \sim F_{q, f-q}$$

where

$$(14) \quad r^2 = \frac{\hat{\mathbf{y}}^T \mathbf{Q}_{\hat{\mathbf{y}}}^{-1} \hat{\mathbf{y}}}{q\hat{\sigma}^2}.$$

Data snooping is the special case where $q = 1$, and $\mathbf{Z} = \mathbf{e}_i$, is the i th column of the $n \times n$ identity matrix. In this case

$$(15) \quad F = F_i = (f-1) \frac{r_i^2}{f - r_i^2} \sim F_{1, f-1} \equiv t_{f-1}^2$$

where

$$(16) \quad r_i = \frac{\hat{u}_i}{\hat{\sigma}(\hat{u}_i)}, \quad \hat{\mathbf{u}} = \mathbf{P} \hat{\mathbf{e}}, \quad \hat{\sigma}(\hat{u}_i) = \hat{\sigma}^2 [\mathbf{P} (\mathbf{I} - \mathbf{AN}^{-1} \mathbf{A}^T) \mathbf{P}]_{ii}.$$

Instead of F_i , its root is usually used

$$(17) \quad t_i = \sqrt{F_i} = r_i \sqrt{\frac{f-1}{f - r_i^2}} \sim t_{f-1},$$

which follows Student's t distribution. The hypothesis $H_0: y = 0$ (no blunder), is accepted over the alternative $H_a: y \neq 0$, when $|t_i| \leq t_{f-1}^{\alpha/2}$.

All results related to the general linear hypothesis can easily be extended from the Gauss-Markov linear

model can be easily extended to the mixed linear model

$$(18) \quad \mathbf{b} = \mathbf{Ax} + \mathbf{Gs} + \mathbf{v}$$

$$(19) \quad E\{\mathbf{v}\} = \mathbf{0}, \quad E\{\mathbf{vv}^T\} = \sigma^2 \mathbf{Q}_v,$$

$$(20) \quad E\{\mathbf{s}\} = \mathbf{0}, \quad E\{\mathbf{ss}^T\} = \sigma^2 \mathbf{Q}_s, \quad E\{\mathbf{vs}^T\} = \mathbf{0}$$

by simply setting

$$(21) \quad \mathbf{e} = \mathbf{Gs} + \mathbf{v}$$

$$(22) \quad \mathbf{P}^{-1} = \mathbf{M} \equiv \mathbf{GQ}_s \mathbf{G}^T + \mathbf{Q}_v.$$

Apart from the usual (least squares) collocation solution used in geodesy, some interesting alternative solutions have been introduced by **Schaffrin** (1985). These solutions differ with respect to the prediction principle used for the prediction of the stochastic signals \mathbf{s} , especially when non-zero mean $E\{\mathbf{s}\} \neq \mathbf{0}$ is assumed. While the usual collocation corresponds to inhomogeneous Best Linear Prediction (inhom-BLIP), "robust collocation" results from homogeneous Best Linear Unbiased Prediction (homBLUP). **Schaffrin** (1987) develops tests for the compatibility of two data sets (the second in the form of "stochastic linear restrictions") for both the inhomBLIP and homBLUP case. **Bock** and **Schaffrin** (1988) consider the case where prior information on the parameters leads to a random effects model (collocation without deterministic unknowns, $\mathbf{b} = \mathbf{Gs} + \mathbf{v}$), and present three different solutions, inomBLIP (usual collocation - strong Bayesian solution), homBLIP (robust solution), and homBLUP (weak Bayesian solution). They also develop tests for testing the equivalence of the homBLIP and homBLUP solutions, the compatibility of the prior information, as well as the improvement of the solution (reduction in mean square prediction error) due to the use of prior information, even if the previous test for its compatibility has failed.

The general hypothesis can compare two models, only when the models are "nested", i.e. when one can result from the other through the introduction of a set of constraints. Non-nested models can be compared indirectly through comparison with a third model in which they are both nested. The test of a general hypothesis can be performed in relation with an adjustment solution based on either one of the compared models.

2.2. Variance Component estimation and testing

Variance component estimation, refers to the case where the covariance matrix of the stochastic term \mathbf{e} of the model is not completely known, but rather has the form

$$(23) \quad E\{\mathbf{ee}^T\} = \mathbf{C} = \theta_1 \mathbf{V}_1 + \theta_2 \mathbf{V}_2 + \dots + \theta_q \mathbf{V}_q$$

where $\boldsymbol{\theta} = [\theta_1 \theta_2 \dots \theta_q]^T$, are the unknown variance components, which must be estimated in addition to the other model parameters.

The method of variance component estimation used in geodesy, is essentially, RaoYs Minimum Norm Quadratic Unbiased Estimation - MINQUE (**Rao**, 1970, 1971, **Rao** and **Kleffe**, 1988). In the geodetic literature, the problem has been solved independently by **Helmert** (1907), and in recent times by **Kubik** (1967a, 1967b, 1970). See also **Ebner** (1972), **Grafarend** (1978, 1984, 1985) **Kelm** (1978), **Welsch** (1978), **Förstner** (1979), **Grafarend** et al. (1980), **Sjöberg** (1983a, 1983b, 1984, 1985), **Schwintzer** (1984), **Koch** (1986), **Ou Ziqiang** (1989), **Zhang** (1990), **Kubáčkova** (1990). A review with further references to the literature can be found in **Grafarend** (1985). The relation between the various types of estimation is thoroughly investigated by **Schaffrin** (1983), while **Rao** and **Kleffe** (1988) give conditions under which the various estimation types coincide. A comprehensive presentation of the subject is in-

cluded in **Koch** (1987a).

The estimates

$$(24) \quad \hat{\boldsymbol{\theta}} = [\hat{\theta}_1 \hat{\theta}_2 \dots \hat{\theta}_q]^T$$

of the variance components follow from the solution of the system

$$(25) \quad \mathbf{J}\hat{\boldsymbol{\theta}} = \mathbf{q}$$

where

$$(26) \quad \mathbf{V} = \mathbf{V}_1 + \mathbf{V}_2 + \dots + \mathbf{V}_i + \dots + \mathbf{V}_q .$$

$$(27) \quad \mathbf{W} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{A}(\mathbf{A}^T\mathbf{V}^{-1}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{V}^{-1}$$

$$(28) \quad J_{ik} = \text{trace}\{\mathbf{W}\mathbf{V}_i\mathbf{W}\mathbf{V}_k\}$$

$$(29) \quad q_i = \hat{\mathbf{e}}^T\mathbf{V}^{-1}\mathbf{V}_i\mathbf{V}^{-1}\hat{\mathbf{e}} .$$

$\hat{\mathbf{e}}$ are initial estimates of \mathbf{e} , obtained using \mathbf{V} as the covariance matrix. For models other than those of observation equations and the linear mixed model, \mathbf{VW} is in general the matrix mapping the unknown true residuals \mathbf{e} into their estimates $\hat{\mathbf{e}}$

$$(30) \quad \hat{\mathbf{e}} = \mathbf{VW}\mathbf{e} .$$

When the matrix \mathbf{J} is regular the MINQUE estimates are

$$(31) \quad \hat{\boldsymbol{\theta}} = \mathbf{J}^{-1}\mathbf{q} .$$

and have covariance matrix

$$(32) \quad \mathbf{C}_{\hat{\boldsymbol{\theta}}} = 2\mathbf{J}^{-1} .$$

A disadvantage of the above solution is that it might produce estimates $\hat{\boldsymbol{\theta}}$ which give rise to an estimate $\hat{\mathbf{C}}$ of the covariance matrix, which is not positive-definite.

A general hypothesis for variance components is given by **Rao** and **Kleffe** (1988):

$$(33) \quad H_0: \mathbf{H}\boldsymbol{\theta} = \mathbf{z} \quad \text{vs} \quad H_a: \mathbf{H}\boldsymbol{\theta} \neq \mathbf{z} ,$$

where \mathbf{H} is an $s \times q$ matrix of rank s . The test statistic, which follows the χ^2 distribution, is

$$(34) \quad u = \frac{1}{2}(\mathbf{H}\hat{\boldsymbol{\theta}} - \mathbf{z})^T(\mathbf{H}\mathbf{J}^{-1}\mathbf{H}^T)^{-1}(\mathbf{H}\hat{\boldsymbol{\theta}} - \mathbf{z}) \sim \chi_s^2$$

and the hypothesis $H_0: \mathbf{H}\boldsymbol{\theta} = \mathbf{z}$ is accepted when $u \leq \chi_s^{2(\alpha)}$.

Tests for single components have also appeared in the geodetic literature (**Remmer**, 1971, **Persson**, 1982, **Koch**, 1987b). The test

$$(35) \quad H_0: \theta_i = \theta_0, \text{ vs } H_a: \theta_i \neq \theta_0 ,$$

for $i = 1, 2, \dots, q$, is reduced to the test

$$(36) \quad H_0: \frac{q_i}{f_i \theta_0} = 1 \text{ vs } H_a: \frac{q_i}{f_i \theta_0} \neq 1.$$

where

$$(37) \quad f_i = \text{trace}\{\mathbf{WV}_i\}.$$

$H_0: \theta_i = \theta_0$ is accepted when

$$(38) \quad F_{f_i, \infty}^{1-\alpha/2} \leq q_i \leq F_{f_i, \infty}^{\alpha/2}.$$

An approximate test is developed by **Seifert** (1985) for the hypothesis

$$(39) \quad H_0: \theta_i = 0 \text{ vs } H_0: \theta_i > 0$$

for individual components other than that of the observational noise.

The test is based on the Cholesky decomposition of the inverse covariance matrix

$$(40) \quad \mathbf{C}_{\hat{\theta}}^{-1} = \frac{1}{2} \mathbf{J} = \mathbf{L}^T \mathbf{D}^{-1} \mathbf{L}$$

where \mathbf{L} is an upper triangular matrix with unit diagonal elements. Uncorrelated transformed components follow from

$$(41) \quad \mathbf{z} = \mathbf{L} \hat{\theta}.$$

The hypothesis is accepted when

$$(42) \quad T = \frac{z_i}{z_i - \hat{\theta}_i} \leq F_{f_1, f_2}^{\alpha/2}.$$

The degrees of freedom f_1 and f_2 are approximated by

$$(43) \quad f_1 \approx 2 \frac{z_i(\theta_0)^2}{d_i},$$

$$(44) \quad f_2 \approx 2 \frac{[z_i(\theta_0) - (\theta_0)_i]^2}{(2\mathbf{J}^{-1})_{ii} - d_i},$$

where $\mathbf{z}(\theta_0) = \mathbf{L}\theta_0$ and θ_0 is the vector of the initial values for θ . In particular the values $(\theta_0)_i$ are all divided by the a-priori value σ_0^2 of the noise variance σ^2 , so that the corresponding initial value of σ^2 becomes 1.

A very special case of variance component estimation is that in the Gaus-Markov model, which involves only a single variance component σ^2 . A general check on the overall validity of the model is based on the a-posteriori variance of unit weight $\hat{\sigma}^2$. The hypothesis $H_0: \sigma^2 = \sigma_0^2$ for the variance of unit weight σ^2 , where σ_0^2 is its a-priori value, is tested against $H_a: \sigma^2 \neq \sigma_0^2$, using the statistic

$$(45) \quad u = \frac{f\hat{\sigma}^2}{\sigma_0^2} \sim \chi_f^2$$

and it is accepted when $\chi_f^{2(1-\alpha/2)} \leq u \leq \chi_f^{2(\alpha/2)}$, where α is the adopted level of significance.

2.3. Multivariate models

One of the problems in variance component estimation, is that estimates are meaningful only for a small number of variance components. The reason is that the data do not contain sufficient information for the estimation of both the unknown parameters and too many components. Theoretically, in the extreme case, the whole covariance matrix \mathbf{V} could be treated as unknown, by simply setting

$$(46) \quad \mathbf{V} = \sum_{ij} V_{ij} \mathbf{E}_{ij}, \quad (\mathbf{E}_{ij})_{km} = \delta_{ik} \delta_{jm}.$$

i.e. \mathbf{E}_{ij} is a matrix with 1 as his ij element and 0 elsewhere. In fact more of the applications refer to the so called "design models" (**Graybill**, 1976). These are special models with design matrices having elements 1 or 0 (contribution or not of a fixed or stochastic term), so that repeated observations differing only in the added outcome of a random effect are available. This "repetition" provides the necessary information for a meaningful estimation of the unknown variances in what is also known as "analysis of variance".

The possibility of making inference about a covariance matrix is provided by multivariate statistics, where a random vector is repeatedly observed (**Anderson**, 1958, **Mardia** et al., 1979, **Seber**, 1984). The repetition of the observations provides enough information for testing hypotheses about the mean vector and the covariance matrix, in the same way as in univariate analysis where repeated observations of a single random variable allow inference on its mean and variance. Of more interest to geodesy, is the case where the mean vector \mathbf{m} of the observed random vector \mathbf{b} is known to depend on a set of parameters $\mathbf{m} = \mathbf{A}\mathbf{x}$. In this case the following multivariate linear model is used

$$(47) \quad \mathbf{b}_i = \mathbf{A}\mathbf{x} + \mathbf{e}_i, \quad \mathbf{e}_i \sim (\mathbf{0}, \mathbf{V}), \quad i = 1, 2, \dots, k,$$

where the index i refers to different outcomes of the same random vector, in repeated experiments. Such a model corresponds in repetitions of the geodetic observations under "identical" conditions. The study of crustal motion by means of repeated observations must allow for slightly varying parameters \mathbf{x} , i.e. different parameters \mathbf{x}_i at every repetition (epoch t_i) corresponding to varying coordinates. The model in compact matrix notation becomes

$$(48) \quad [\mathbf{b}_1 \dots \mathbf{b}_k] = \mathbf{A}[\mathbf{x}_1 \dots \mathbf{x}_k] + [\mathbf{e}_1 \dots \mathbf{e}_k] \quad \text{or} \quad \mathbf{B} = \mathbf{A}\mathbf{X} + \mathbf{E}$$

with maximum likelihood estimates

$$(49) \quad \hat{\mathbf{X}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Y}, \quad \hat{\mathbf{E}} = \mathbf{B} - \mathbf{A}\hat{\mathbf{X}}, \quad \hat{\mathbf{V}} = \frac{1}{n} \hat{\mathbf{E}}^T \hat{\mathbf{E}}.$$

A multivariate general hypothesis $H_0: \mathbf{H}\mathbf{X} = \mathbf{Z}$ can be tested (for various choices of test statistics see **Seber**, 1984, **Koch**, 1987a) in order to identify variations in the parameters \mathbf{x}_i corresponding to crustal motions. Also tests related to the covariance matrix \mathbf{V} can be performed (**Koch**, 1987a, p. 426).

Multivariate models have been used in geodesy for the detection of crustal motions, with models where the only stochastic terms are the observational errors. In integrated geodesy, however, the stochastic terms \mathbf{e}_i contain also the influence of random signals \mathbf{s}_i associated with the gravity field

($\mathbf{e}_i = \mathbf{G}\mathbf{s}_i + \mathbf{v}_i$). This poses a problem in relation to the relevance of the multivariate model, since signal variation from epoch to epoch reflects a systematic effect related possibly to changes in position (if not taken into account by proper modeling) and certainly to slow mass redistribution within the earth. It is therefore hardly a variation due to different outcomes from the same population, as is the case with the noise \mathbf{v}_i .

In order to make the argument more clear, assume that repeated identical observations on the same network are performed simultaneously (in fact within a very short time interval) so that the model (47) rather applies, since no crustal motion has occurred. Now at each repetition one would expect different noise \mathbf{v}_i , but the same signals \mathbf{s} from the unaltered gravity field ($\mathbf{e}_i = \mathbf{G}\mathbf{s} + \mathbf{v}_i$). Therefore the multivariate approach is applicable only with models which treat gravity signals as deterministic parameters.

In the stochastic model for the gravity field, randomness lies rather in the random placement of the network within the deterministic gravity field (Moritz, 1978, Sansó, 1978). Multivariate analysis would be relevant only if the network was rigidly shifted by a random rotation to another part of the earth surface!

2.4. Bayesian inference

An interesting alternative to the classical statistical inference is the *Bayesian inference* (Broemeling, 1985, Koch, 1990) which is based on the knowledge of the prior distributions of all parameters, even those which are treated as unknown deterministic parameters in the classical approach. The usual lack of knowledge for the distribution of the unknown parameters, limits the applicability of Bayesian analysis with *informative priors*, but the problem is circumvented by the use of *non-informative priors*.

An essential new feature of Bayesian analysis is the introduction of "less sensitive" tests of hypothesis, where the presence of a function of the parameters within a prescribed set is tested rather than their equality with given values, as in the classical linear hypothesis (Riesmeier, 1984, Koch and Riesmeier, 1985, Koch, 1990, chapter 3.1.7). The degree of arbitrariness associated with the choice of the dimensions of such a set, is somewhat of a restriction to the effective application of such tests.

Bayesian variance component estimation (Koch, 1987b, 1987c, 1990, chapter 3.4) requires the solution of integral equations with numerical techniques, such as Monte-Carlo integration.

3. Inference problems related to integrated geodesy

Integrated geodesy, in the generalized sense used here, deals with the analysis of observations

$$(50) \quad \mathbf{y}^b = \mathbf{F}(\mathbf{x}^a, \mathbf{s}^a) + \mathbf{v}$$

which are contaminated by noise \mathbf{v} and depend on a set of discrete parameters \mathbf{x}^a , as well as a set of signals \mathbf{s}^a , which depend in turn on an underlying function f^a

$$(51) \quad \mathbf{s}^a = \mathbf{s}^a(f^a)$$

in the sense that individual signals are the values of functionals on f^a . Linearization using known approximations \mathbf{x}^0 , f^0 , so that $\mathbf{s}^0 = \mathbf{s}^a(f^0)$, leads to the linearized model

$$(52) \quad \mathbf{y}^b - \mathbf{F}(\mathbf{x}^0, \mathbf{s}^0) = \left. \frac{\partial \mathbf{F}}{\partial \mathbf{x}^a} \right|_0 (\mathbf{x}^a - \mathbf{x}^0) + \left. \frac{\partial \mathbf{F}}{\partial \mathbf{s}^a} \right|_0 (\mathbf{s}^a - \mathbf{s}^0) + \mathbf{v}$$

or, with obvious notation,

$$(53) \quad \mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{G}\mathbf{s} + \mathbf{v}$$

where the individual signals s_i are the values of linear functionals L_i

$$(54) \quad s_i = L_i(f)$$

on the “disturbing function” $f = f^a - f^0$.

Alternative models may arise from different assumptions about the function f . (It is also possible to ignore the dependence of the signals on the function f and to treat \mathbf{s} as discrete deterministic unknown parameters similar to \mathbf{x} .) Any assumption on f has the general form

$$(55) \quad f = m(\mathbf{a}) + \delta f$$

where $m(\mathbf{a})$ is a parametrized trend, depending on a finite set of unknown parameters \mathbf{a} and δf is a stochastic part with known mean and covariance function. It is sufficient to consider zero mean function of δf , since any known non zero part can be included in the trend $m(\mathbf{a})$ or in the approximation f^0 . Special cases to this hybrid (deterministic-stochastic) approach, is the deterministic or analytical approach $f = m(\mathbf{a})$, and the stochastic approach $f = \delta f$. Apart from these three different possibilities, different models may arise within each model class (hybrid, analytical or stochastic) by considering different analytical forms for $m(\mathbf{a})$, or different covariance functions for δf . Restricting $m(\mathbf{a})$ to the linear form

$$(56) \quad m = a_1 f_1 + a_2 f_2 + \dots + a_q f_q$$

where f_1, f_2, \dots, f_q are known base functions, the linear model becomes

$$(57) \quad \mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{G}\mathbf{F}\mathbf{a} + \mathbf{G}\delta\mathbf{s} + \mathbf{v} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{a} + \mathbf{G}\delta\mathbf{s} + \mathbf{v}$$

where

$$(58) \quad F_{ij} = L_i(f_j), \quad \delta s_i = L_i(\delta f), \quad \mathbf{B} = \mathbf{G}\mathbf{F}.$$

From the covariance function of δf

$$(59) \quad C(P, Q) = \sigma(\delta f(P), \delta f(Q))$$

the covariance matrix \mathbf{C}_s of the stochastic signals $\delta\mathbf{s}$, is computed with application of the law of covariance propagation

$$(60) \quad (\mathbf{C}_s)_{ij} = L_{i,P} L_{j,Q} C(P, Q).$$

Dropping δ from $\delta\mathbf{s}$, for the sake of notational simplicity, the model can be separated into a deterministic and a stochastic part

$$(61) \quad \mathbf{b} = [\mathbf{A} \ \mathbf{B}] \begin{bmatrix} \mathbf{x} \\ \mathbf{a} \end{bmatrix} + [\mathbf{G} \ \mathbf{I}] \begin{bmatrix} \mathbf{s} \\ \mathbf{v} \end{bmatrix}, \quad \begin{bmatrix} \mathbf{s} \\ \mathbf{v} \end{bmatrix} \sim \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{C}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_v \end{bmatrix} \right).$$

Any such model suffers from a degree of arbitrariness, associated with the choice of the form and number of the base functions f_i and with the choice of the covariance function $C(P, Q)$. For this reason there is a need for tools from statistical inference which will either help in asserting that the choices made are reasonable, or will be able to compare between such different choices. Assumptions about the trend $m(\mathbf{a})$

can be tested with the help of the usual statistical tests on the deterministic parameters \mathbf{a} . Assumptions about the covariance function C of δf can be made only with the introduction of not one, but a family of covariance functions $C(\boldsymbol{\theta})$ depending on a set of unknown parameters $\boldsymbol{\theta}$. This leads to a model with variance components

$$(62) \quad \mathbf{b} = [\mathbf{A} \ \mathbf{B}] \begin{bmatrix} \mathbf{x} \\ \mathbf{a} \end{bmatrix} + \mathbf{e} \quad \mathbf{e} \sim (\mathbf{0}, \mathbf{V}(\boldsymbol{\theta}))$$

with

$$(63) \quad \mathbf{e} = \mathbf{G}\mathbf{s} + \mathbf{v} \quad \mathbf{V}(\boldsymbol{\theta}) = \mathbf{G}\mathbf{C}_s(\boldsymbol{\theta})\mathbf{G}^T + \mathbf{C}_v.$$

If the covariance function is linear with respect to the parameters $\boldsymbol{\theta}$, then

$$(64) \quad \mathbf{C}_s(\boldsymbol{\theta}) = \sum_{i=1}^p \theta_i \mathbf{Q}_i$$

and the variance factor σ^2 of the noise covariance matrix $\mathbf{C} = \sigma^2 \mathbf{Q}$ is included in the parameters $\boldsymbol{\theta}$ ($\theta_{p+1} = \sigma^2$), the usual linear model with variance components is obtained

$$(65) \quad \mathbf{V}(\boldsymbol{\theta}) = \sum_{i=1}^{p+1} \theta_i \mathbf{V}_i, \quad \mathbf{V}_i = \mathbf{G}\mathbf{Q}_i\mathbf{G}^T, \quad i = 1, \dots, p, \quad \mathbf{V}_{p+1} = \mathbf{Q}.$$

The estimation of the variance components helps in the choice of a particular covariance function $C(\hat{\boldsymbol{\theta}})$ from the family $C(\boldsymbol{\theta})$. Tests on the variance components may be used for the detection of insignificant terms θ_i and the corresponding simplification of the covariance function.

4. Applications of statistical tests to integrated geodesy.

4.1. General testing of the model validity.

Before any particular tests are going to be performed for testing particular features of the model, some general tests must be used. These general tests will indicate if the model is problematic, without identifying in what particular respect. In the most general case where the unknown function is considered to consist of a parametrized trend and a stochastic disturbing part, the model of integrated geodesy has the form

$$(66) \quad \mathbf{b} = [\mathbf{A} \ \mathbf{B}] \begin{bmatrix} \mathbf{x} \\ \mathbf{a} \end{bmatrix} + [\mathbf{G} \ \mathbf{I}] \begin{bmatrix} \mathbf{s} \\ \mathbf{v} \end{bmatrix}, \quad \begin{bmatrix} \mathbf{s} \\ \mathbf{v} \end{bmatrix} \sim \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \sigma^2 \begin{bmatrix} \mathbf{Q}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_v \end{bmatrix} \right)$$

where deterministic and stochastic parameters are separated. The deterministic parameters are the discrete unknown parameters \mathbf{x} and the parameters \mathbf{a} of the trend function. The stochastic parameters are the observational errors \mathbf{v} and the signals \mathbf{s} arising from the stochastic disturbing part of the unknown function. Another useful form of the model is

$$(67) \quad \mathbf{b} = [\mathbf{A} \ \mathbf{B}] \begin{bmatrix} \mathbf{x} \\ \mathbf{a} \end{bmatrix} + \mathbf{e}, \quad \mathbf{e} = \mathbf{G}\mathbf{s} + \mathbf{v} \sim (\mathbf{0}, \sigma^2 [\mathbf{G}\mathbf{Q}_s\mathbf{G}^T + \mathbf{Q}_v]).$$

The well known solution (**Dermanis**, 1987) for the above model is

$$(68) \quad \mathbf{M} = \mathbf{GQ}_s\mathbf{G}^T + \mathbf{Q}_v$$

$$(69) \quad \mathbf{M}^{-1} = (\mathbf{GQ}_s\mathbf{G}^T + \mathbf{Q}_v)^{-1} = \mathbf{Q}_v^{-1} - \mathbf{Q}_v^{-1}\mathbf{G}(\mathbf{G}^T\mathbf{Q}_v^{-1}\mathbf{G} + \mathbf{Q}_s^{-1})^{-1}\mathbf{G}^T\mathbf{Q}_v^{-1}$$

$$(70) \quad \mathbf{N} = \begin{bmatrix} \mathbf{A}^T\mathbf{M}^{-1}\mathbf{A} & \mathbf{A}^T\mathbf{M}^{-1}\mathbf{B} \\ \mathbf{B}^T\mathbf{M}^{-1}\mathbf{A} & \mathbf{B}^T\mathbf{M}^{-1}\mathbf{B} \end{bmatrix}$$

$$(71) \quad \begin{bmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{a}} \end{bmatrix} = \mathbf{N}^{-1} \begin{bmatrix} \mathbf{A}^T\mathbf{M}^{-1}\mathbf{b} \\ \mathbf{B}^T\mathbf{M}^{-1}\mathbf{b} \end{bmatrix}$$

$$(72) \quad \hat{\mathbf{e}} = \mathbf{b} - \mathbf{A}\hat{\mathbf{x}} - \mathbf{B}\hat{\mathbf{a}}$$

$$(73) \quad \begin{bmatrix} \hat{\mathbf{s}} \\ \hat{\mathbf{v}} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_s\mathbf{G}^T \\ \mathbf{Q}_v \end{bmatrix} \mathbf{M}^{-1}\hat{\mathbf{e}}$$

$$(74) \quad \hat{\sigma}^2 = \frac{\hat{\mathbf{e}}^T\mathbf{M}^{-1}\hat{\mathbf{e}}}{f}.$$

The estimated covariance matrices of the above estimates are computed by multiplying with $\hat{\sigma}^2$ the following "cofactor" matrices:

$$(75) \quad \begin{bmatrix} \mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{x}}} & \mathbf{Q}_{\hat{\mathbf{x}}\hat{\mathbf{a}}} \\ \mathbf{Q}_{\hat{\mathbf{a}}\hat{\mathbf{x}}} & \mathbf{Q}_{\hat{\mathbf{a}}\hat{\mathbf{a}}} \end{bmatrix} = \mathbf{N}^{-1}$$

$$(76) \quad \mathbf{Q}_{\hat{\mathbf{e}}\hat{\mathbf{e}}} = \mathbf{M} - [\mathbf{A} \ \mathbf{B}]\mathbf{N}^{-1} \begin{bmatrix} \mathbf{A}^T \\ \mathbf{B}^T \end{bmatrix}$$

$$(77) \quad \begin{bmatrix} \mathbf{Q}_{\hat{\mathbf{s}}\hat{\mathbf{s}}} & \mathbf{Q}_{\hat{\mathbf{s}}\hat{\mathbf{v}}} \\ \mathbf{Q}_{\hat{\mathbf{v}}\hat{\mathbf{s}}} & \mathbf{Q}_{\hat{\mathbf{v}}\hat{\mathbf{v}}} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_s\mathbf{G}^T \\ \mathbf{Q}_v \end{bmatrix} \mathbf{M}^{-1}\mathbf{Q}_{\hat{\mathbf{e}}\hat{\mathbf{e}}}\mathbf{M}^{-1} \begin{bmatrix} \mathbf{GQ}_s & \mathbf{Q}_v \end{bmatrix}.$$

The first test to be applied is the so-called *congruency test* (**Kok**, 1982) $H_0: \sigma^2 = \sigma_0^2$ vs $H_a: \sigma^2 \neq \sigma_0^2$, for the variance of unit weight σ^2 , which is tested using the statistic given by equation (45). The test gives an overall impression on the validity of the model, but further testing is always required, even when the congruency test points to the contrary, since it is possible that effects from different modeling errors cancel each other out, in the computation of $\hat{\sigma}^2$.

The next series of tests to be carried out are those for individual outliers in the observations, using data snooping. Equation (16) becomes in this case ($\mathbf{P} = \mathbf{M}^{-1}$)

$$(78) \quad r_i = \frac{\hat{u}_i}{\hat{\sigma}(\hat{u}_i)}, \quad \hat{\mathbf{u}} = \mathbf{M}^{-1}\hat{\mathbf{e}}, \quad \hat{\sigma}^2(\hat{u}_i) = \hat{\sigma}^2[\mathbf{M}^{-1}\mathbf{Q}_{\hat{\mathbf{e}}\hat{\mathbf{e}}}\mathbf{M}^{-1}]_{ii}$$

and the test statistic is given by eq. (17). Large values of the test statistic t_i for many of the observations, do not necessarily indicate the presence of multiple gross errors (especially for the high quality field work in studies of crustal deformation). They are most probably due to modeling errors, either in the part for the deterministic parameters \mathbf{x} , or in the trend and/or covariance model of the unknown function.

Tests for multiple outliers on observations can be performed, but their interpretation as gross errors is again not possible for the usual high quality of the observations. Such detected outliers signify the existence of model inconsistency between the two sets of observations, those with outliers and those without. It will be shown in section 4.4 that a multiple outlier test is equivalent to tests developed by **Wei** (1987a, 1987b) for consistency in group-wise collocation and related to those of **Schaffrin** (1987) for "stochastic hypotheses".

4.2. Testing the trend function model

In order to test if a set of components in the trend function (56) is significant or not, the hypothesis that the corresponding set of trend parameters a_i are vanishing should be tested. Such a test could be performed in two ways, for adding and removing terms with respect to an available solution. Let

$$(79) \quad \mathbf{b} = [\mathbf{A} \ \mathbf{B}] \begin{bmatrix} \mathbf{x} \\ \mathbf{a} \end{bmatrix} + [\mathbf{G} \ \mathbf{I}] \begin{bmatrix} \mathbf{s} \\ \mathbf{v} \end{bmatrix},$$

be the model of the already available solution with corresponding estimates $\hat{\mathbf{x}}$, $\hat{\mathbf{a}}$, $\hat{\mathbf{s}}$, $\hat{\mathbf{v}}$ and $\hat{\sigma}^2$.

Let $\mathbf{a} = \begin{bmatrix} \mathbf{a}^{(I)} \\ \mathbf{a}_I \end{bmatrix}$, where I stands for a set of indices and (I) for its complement. To test the hypothesis

$H_0: \mathbf{a}_I = \mathbf{0}$ vs $H_a: \mathbf{a}_I \neq \mathbf{0}$, the test statistic according to equation (8) is

$$(80) \quad F = \frac{\hat{\mathbf{a}}_I^T \mathbf{Q}_{\hat{\mathbf{a}}, \hat{\mathbf{a}}_I}^{-1} \hat{\mathbf{a}}_I}{k \hat{\sigma}^2} \sim F_{k, f},$$

where $\mathbf{Q}_{\hat{\mathbf{a}}, \hat{\mathbf{a}}_I}$ is the appropriate submatrix of $\mathbf{Q} = \mathbf{N}^{-1}$, the inverse matrix of the total normal equations. If a single trend parameter a_i is to be tested, the test statistic is

$$(81) \quad t = \frac{\hat{a}_i}{\hat{\sigma}_{\hat{a}_i}} \sim t_f,$$

where $\hat{\sigma}_{\hat{a}_i}^2$ is the corresponding diagonal element of $\mathbf{Q} = \mathbf{N}^{-1}$ multiplied with $\hat{\sigma}^2$.

In the case of a set of additional parameters \mathbf{a}_* not included in the available model, with corresponding coefficient matrix \mathbf{H}_* , the test statistic for $H_0: \mathbf{a}_* = \mathbf{0}$ vs $H_a: \mathbf{a}_* \neq \mathbf{0}$, according to equation (9) is

$$(82) \quad F = \frac{f - q}{q} \frac{\delta\Omega}{f\hat{\sigma}^2 - \delta\Omega} \sim F_{q, f - q}$$

where q is the number of the additional parameters \mathbf{a}_* , and

$$(83) \quad \delta\Omega = \hat{\mathbf{e}}^T \mathbf{M}^{-1} \mathbf{H}_* [\mathbf{H}_*^T \mathbf{M}^{-1} \mathbf{Q}_{\hat{\mathbf{e}}} \mathbf{M}^{-1} \mathbf{H}_*]^{-1} \mathbf{H}_*^T \mathbf{M}^{-1} \hat{\mathbf{e}} = \\ = \hat{\mathbf{v}}^T \mathbf{Q}_v^{-1} \mathbf{H}_* [\mathbf{H}_*^T \mathbf{M}^{-1} \mathbf{Q}_{\hat{\mathbf{e}}} \mathbf{M}^{-1} \mathbf{H}_*]^{-1} \mathbf{H}_*^T \mathbf{Q}_v^{-1} \hat{\mathbf{v}}.$$

When a single additional parameter a_* with coefficient vector \mathbf{h}_* is to be tested, the test statistic becomes

$$(84) \quad t = r_a \sqrt{\frac{f-1}{f-r_a^2}} \sim t_{f-1}$$

where

$$(85) \quad r_a^2 = \frac{\hat{a}_*^2}{\sigma^2(\hat{a}_*)} = \frac{\hat{a}_*^2}{\hat{\sigma}^2 q_{\hat{a}_*}} = \frac{[q_{\hat{a}_*} \mathbf{h}_*^T \mathbf{Q}_v^{-1} \hat{\mathbf{v}}]^2}{\hat{\sigma}^2 q_{\hat{a}_*}} = \frac{[\mathbf{h}_*^T \mathbf{Q}_v^{-1} \hat{\mathbf{v}}]^2}{\hat{\sigma}^2 \mathbf{h}_*^T \mathbf{M}^{-1} \mathbf{Q}_{\hat{\mathbf{e}\hat{\mathbf{e}}}} \mathbf{M}^{-1} \mathbf{h}_*}.$$

The problem of applying statistical tests for the selection of a final set of terms in the trend function has no straightforward solution, since the various choices are not directly comparable. Two choices can be compared only when one is contained (nested) in the other, i.e., when the first follows from the deletion of some of the terms contained in the second. The possibility of adding or subtracting terms one at a time, with the tests developed above, does not lead to a unique solution, because it depends on both the starting solution and the criterion for choosing which term to add or delete. The problem is familiar from the analysis of deformation in classical (non-integrated) horizontal or vertical networks and also from photogrammetry in the case of "additional parameters" describing image deformation. Various strategies have been developed, where the choice criteria is the magnitude of the statistic used in the test, but also correlation with the other parameters. High correlation suggests priority in parameter deletion and low one priority in parameter addition.

In fact eq. (85) cannot be applied when a_* has correlation ± 1 with the other parameters \mathbf{x} and \mathbf{a} , while correlation close to these extremes causes numerical problems. This happens when $\mathbf{h}_* \in R([\mathbf{A} \ \mathbf{B}])$, i.e., in the case where they exist vectors \mathbf{k}_x and \mathbf{k}_a such that $\mathbf{h}_* = \mathbf{A}\mathbf{k}_x + \mathbf{B}\mathbf{k}_a$, with the consequence that the terms \hat{a}_* and $q_{\hat{a}_*}$ in eq. (85) vanish. In this case the angle ϑ of \mathbf{h}_* with (its projection on) the subspace $R([\mathbf{A} \ \mathbf{B}])$ can be a-priori computed from

$$(86) \quad \sin^2 \vartheta = \frac{f \hat{\sigma}_h^2}{\mathbf{h}_*^T \mathbf{M}^{-1} \mathbf{h}_*}$$

where

$$(87) \quad \hat{\sigma}_h^2 = \frac{1}{f} \mathbf{h}_*^T \mathbf{M}^{-1} \mathbf{Q}_{\hat{\mathbf{e}\hat{\mathbf{e}}}} \mathbf{M}^{-1} \mathbf{h}_*$$

is a variance estimate which can also be computed by repeating the adjustment with the vector \mathbf{h}_* in the place of the observations \mathbf{b} .

4.3. Testing the covariance function model

Tests on the validity of the model used for the covariance function require the estimation of variance components θ_i and the implementation of the relevant tests described in section (2.2). The main problem here is the introduction of a model of the form (23) - linear with respect to the parameters θ_i . According to equation (63), and since the functionals in (60) are linear, the total variance $\mathbf{V}(\boldsymbol{\theta}) = \mathbf{G}\mathbf{C}_s(\boldsymbol{\theta})\mathbf{G}^T + \mathbf{C}_v$, is linear in $\boldsymbol{\theta}$, only when the covariance function $C(\boldsymbol{\theta})$ is itself linear in $\boldsymbol{\theta}$. This is hardly the case with the covariance models in use.

When a natural basis $f_i(P)$ exists for the function $f(P)$ (e.g. spherical harmonics for the gravitational potential), the covariance function can take the form

$$(88) \quad C(P, Q) = \sum_k \sigma_k^2 f_k(P) f_k(Q)$$

where *the degree variances* σ_k^2 are either the unknown variance components θ_k , or depend on a smaller number of unknown parameters. If the degree variances are considered as variance components, the desirable linear form for $\mathbf{V}(\boldsymbol{\sigma})$ is obtained, since

$$(89) \quad \{\mathbf{C}(\boldsymbol{\theta})\}_{ij} = \sum_k \sigma_k^2 L_i\{f_k(P)\} L_j\{f_k(Q)\} \equiv \sum_k \sigma^2 \{\mathbf{Q}_k\}_{ij}$$

$$(90) \quad \mathbf{V}(\boldsymbol{\sigma}) = \sum_k \sigma_k^2 \mathbf{G} \mathbf{Q}_k \mathbf{G}^T + \sigma^2 \mathbf{Q}_v \quad (\mathbf{C}_v = \sigma^2 \mathbf{Q}_v)$$

with

$$(91) \quad \boldsymbol{\sigma} = [\dots \sigma_k^2 \dots | \sigma^2]^T.$$

However a large number of terms is required for a sufficient approximation to the covariance function with theoretically infinite number of terms in the series (88), in which case their meaningful estimation is not possible. On the other hand if the degree variances are expressed as functions $\sigma_k^2 = \sigma_k^2(\boldsymbol{\theta})$ of a small number of variance components $\boldsymbol{\theta}$, this dependence is highly non-linear. The problem can be approximately solved by linearization around some initial values, which should be the values that are used anyway in the covariance model, in the usual approach where no variance component estimation takes place.

More general, the covariance function $C(\boldsymbol{\alpha})$ depends on a set of parameters $\boldsymbol{\alpha}$, either through the degree variances, or more often directly when an empirical covariance function is obtained from a statistical analysis of separate sample data (see, e.g., **Meier** (1981) for examples of such parametrized covariance models). Including an unknown factor s^2 , the covariance function has the form

$$(92) \quad C(s^2, \boldsymbol{\alpha}) = s^2 f(\boldsymbol{\alpha})$$

and linearization gives

$$(93) \quad \begin{aligned} C(s^2, \boldsymbol{\alpha}) &= C(s_0^2, \boldsymbol{\alpha}_0) + \frac{\partial C}{\partial s^2} (s^2 - s_0^2) + \sum_i \frac{\partial C}{\partial \alpha_i} (\alpha_i - \alpha_{i0}) = \\ &= s_0^2 f(\boldsymbol{\alpha}_0) + f(\boldsymbol{\alpha}_0) (s^2 - s_0^2) + \sum_i s_0^2 \frac{\partial f}{\partial \alpha_i} \delta \alpha_i = f(\boldsymbol{\alpha}_0) s^2 + \sum_i s_0^2 \frac{\partial f}{\partial \alpha_i} \delta \alpha_i. \end{aligned}$$

The variance components are known the parameters s^2 , $\delta \alpha_i$, $i = 1, 2, \dots, p$, and σ^2 . These components should be estimated and tested for $s^2 = s_0^2$, $\delta \alpha_i = 0$ and $\sigma^2 = \sigma_0^2$. The initial values are those of the model whose validity is tested. The most simple case is the one two variance components which are the factors s^2 and σ^2 , of the covariance function and the observational noise covariance matrix, respectively.

In four-dimensional integrated geodesy the covariance function depends not only on position (usually spatial separation for the isotropic case), but also on time difference. In this case it is advisable to include in the variance component estimation and testing, at least one parameter whose vanishing indicates the lack of temporal correlation (**Rossikopoulos**, 1986).

For truncated covariances with only a small number of "degree variances", tests can be used to separate the cases $\sigma_i^2 = 0$, which means that the $f_i(P)$ component is insignificant. It would be interesting to

have a test of the type $\sigma_i^2 \rightarrow 0$, which means that the $f_i(P)$ component is deterministic and should be included in the trend model. However such a test does not seem to be available.

4.4. Testing the compatibility of two sets of observations.

In order to check the compatibility of two sets of observations, it is sufficient to check for the presence of outliers in one of them, say the second from

$$(94) \quad \begin{aligned} \mathbf{b}_1 &= \mathbf{A}_1 \mathbf{x} + \mathbf{e}_1 \\ \mathbf{b}_2 &= \mathbf{A}_2 \mathbf{x} + \mathbf{y} + \mathbf{e}_2 \end{aligned}$$

or

$$(95) \quad \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix} \mathbf{x} + \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} \mathbf{y} + \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{bmatrix}$$

or

$$(96) \quad \mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{Z}\mathbf{y} + \mathbf{e}$$

Here we assume correlated stochastic terms (residuals) $\mathbf{e}_1, \mathbf{e}_2$, so that the results cover also the case of the mixed model $\mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{G}\mathbf{s} + \mathbf{v}$, by simply setting $\mathbf{e} = \mathbf{G}\mathbf{s} + \mathbf{v}$, with $\mathbf{e} \sim (\mathbf{0}, \mathbf{P}^{-1})$.

The hypothesis $H_0: \mathbf{y} = \mathbf{0}$ vs $H_a: \mathbf{y} \neq \mathbf{0}$ can be tested by direct application of the equations (9) to (14)

$$(97) \quad F = \frac{f - n_2}{n_2} \frac{\hat{\mathbf{y}}^T \mathbf{Q}_{\hat{\mathbf{y}}}^{-1} \hat{\mathbf{y}}}{f \hat{\sigma}^2 - \hat{\mathbf{y}}^T \mathbf{Q}_{\hat{\mathbf{y}}}^{-1} \hat{\mathbf{y}}} = (f - n_2) \frac{r^2}{f - n_2 r^2} \sim F_{n_2, f - n_2}$$

where f are the degrees of freedom, n_2 is the number of observations in the second set and

$$(98) \quad r^2 = \frac{\hat{\mathbf{y}}^T \mathbf{Q}_{\hat{\mathbf{y}}}^{-1} \hat{\mathbf{y}}}{n_2 \hat{\sigma}^2}.$$

It must be emphasized that the above test does not merely check for outliers in the second test, as the case would be if this set was separately analyzed, but the compatibility of the two sets is actually tested, in view of the fact that the outliers are introduced in relation to the simultaneous analysis of both sets.

A different formulation of the test (Appendix A) utilizes the misclosure δ following from the solution of the first set only, which produces an estimate $\hat{\mathbf{x}}^{(1)}$ and a prediction $\tilde{\mathbf{e}}_2^{(1)}$ based on the correlation of \mathbf{e}_2 with \mathbf{e}_1 ,

$$(99) \quad \delta = \mathbf{b}_2 - \mathbf{A}_2 \hat{\mathbf{x}}^{(1)} - \tilde{\mathbf{e}}_2^{(1)}$$

(see **Cook and Weisberg**, 1982, for uncorrelated residuals, and a single outlier).

For the mixed model

$$(100) \quad \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix} \mathbf{x} + \begin{bmatrix} \mathbf{G}_1 & \mathbf{G}_{21} \\ \mathbf{G}_{12} & \mathbf{G}_2 \end{bmatrix} \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}$$

$$\boldsymbol{\delta} = \mathbf{b}_2 - \mathbf{A}_2 \hat{\mathbf{x}}^{(1)} - \mathbf{G}_{21} \tilde{\mathbf{s}}_1^{(1)} - \mathbf{G}_2 \tilde{\mathbf{s}}_2^{(1)}$$

where

$$(101) \quad \mathbf{M}_1 = [\mathbf{G}_1 \quad \mathbf{G}_{12}] \begin{bmatrix} \mathbf{Q}_{s_1} & \mathbf{Q}_{s_1 s_2} \\ \mathbf{Q}_{s_2 s_1} & \mathbf{Q}_{s_2} \end{bmatrix} \begin{bmatrix} \mathbf{G}_1^T \\ \mathbf{G}_{12}^T \end{bmatrix} + \mathbf{Q}_v$$

$$(102) \quad \hat{\mathbf{x}}^{(1)} = (\mathbf{A}_1^T \mathbf{M}_1^{-1} \mathbf{A}_1)^{-1} \mathbf{A}_1^T \mathbf{M}_1^{-1} \mathbf{b}_1$$

$$(103) \quad \begin{bmatrix} \tilde{\mathbf{s}}_1^{(1)} \\ \tilde{\mathbf{s}}_2^{(1)} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_{s_1} & \mathbf{Q}_{s_1 s_2} \\ \mathbf{Q}_{s_2 s_1} & \mathbf{Q}_{s_2} \end{bmatrix} \begin{bmatrix} \mathbf{G}_1^T \\ \mathbf{G}_{12}^T \end{bmatrix} \mathbf{M}_1^{-1} (\mathbf{b}_1 - \mathbf{A}_1 \hat{\mathbf{x}}^{(1)})$$

or in the simpler case with $\mathbf{G}_{12} = \mathbf{0}$, $\mathbf{G}_{21} = \mathbf{0}$

$$(104) \quad \boldsymbol{\delta} = \mathbf{b}_2 - \mathbf{A}_2 \hat{\mathbf{x}}^{(1)} - \mathbf{G}_2 \tilde{\mathbf{s}}_2^{(1)}$$

where

$$(105) \quad \mathbf{M}_1 = \mathbf{G}_1 \mathbf{Q}_{s_1} \mathbf{G}_1^T + \mathbf{Q}_v.$$

$\hat{\mathbf{x}}^{(1)}$ is evaluated from (102) using \mathbf{M}_1 from (105) and

$$(106) \quad \begin{bmatrix} \tilde{\mathbf{s}}_1^{(1)} \\ \tilde{\mathbf{s}}_2^{(1)} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_{s_1} \mathbf{G}_1^T \\ \mathbf{Q}_{s_2 s_1} \end{bmatrix} \mathbf{M}_1^{-1} (\mathbf{b}_1 - \mathbf{A}_1 \hat{\mathbf{x}}^{(1)})$$

In the Appendix we have proved that $\hat{\mathbf{y}} = \boldsymbol{\delta}$, $\mathbf{Q}_{\hat{\mathbf{y}}} = \mathbf{Q}_{\boldsymbol{\delta}}$ and

$$(107) \quad f \hat{\sigma}^2 - \boldsymbol{\delta}^T \mathbf{Q}_{\boldsymbol{\delta}}^{-1} \boldsymbol{\delta} = (f - n_2) \hat{\sigma}_1^2$$

where $\hat{\sigma}_1^2$ is the estimate of σ^2 from only the first set of the observations, so that the test statistic becomes

$$(108) \quad F = \frac{f - n_2}{n_2} = \frac{\boldsymbol{\delta}^T \mathbf{Q}_{\boldsymbol{\delta}}^{-1} \boldsymbol{\delta}}{f \hat{\sigma}^2 - \boldsymbol{\delta}^T \mathbf{Q}_{\boldsymbol{\delta}}^{-1} \boldsymbol{\delta}} = \frac{\boldsymbol{\delta}^T \mathbf{Q}_{\boldsymbol{\delta}}^{-1} \boldsymbol{\delta}}{n_2 \hat{\sigma}_1^2} \sim F_{n_2, f - n_2}.$$

This is exactly one of the statistics derived by **Wei** (1987a, 1987b) for the simpler case of "groupwise collocation", with $\mathbf{G}_1 = \mathbf{I}$, $\mathbf{G}_2 = \mathbf{I}$, $\mathbf{G}_{12} = \mathbf{0}$, $\mathbf{G}_{21} = \mathbf{0}$:

$$(109) \quad \mathbf{b}_1 = \mathbf{A}_1 \mathbf{x} + \mathbf{s}_1 + \mathbf{v}_1$$

$$\mathbf{b}_2 = \mathbf{A}_2 \mathbf{x} + \mathbf{s}_2 + \mathbf{v}_2$$

where the "hypothesis" tested is

$$(110a) \quad H_0: \quad \mathbf{b}_2 = \mathbf{A}_2 \mathbf{x} + \mathbf{s}_2 + \mathbf{v}_2$$

or

$$(110b) \quad H_0: \quad \mathbf{A}_2(\hat{\mathbf{x}}^{(1)} - \mathbf{x}) + (\tilde{\mathbf{s}}_2^{(1)} - \mathbf{s}_2) - \mathbf{v}_2 = \mathbf{A}_2\hat{\mathbf{x}}^{(1)} + \tilde{\mathbf{s}}_2^{(1)} - \mathbf{b}_2 \equiv -\delta .$$

This is a peculiar formulation of a hypothesis involving random variables, lying outside the standard formulation of hypotheses in traditional statistics, which involve only deterministic equations in the non-random parts of the model (means and/or covariances). **Schaffrin** (1987) went even further in formally introducing the, indeed innovating, idea of "stochastic linear hypotheses"! We have shown here that such tests, (which by the way have appeared to be essentially the only new idea in statistical testing for collocation,) are equivalent to a standard classical test, that for outliers in a subset of the observations.

The problem with the above test is how the original data set is going to be divided into two subsets. Therefore the test should be applied only when a natural separation exists. This is the case with prior information from a previous data analysis which is treated through the device of "pseudo-observations". Another such case appears when observations in a network are to be combined with separate gravity related observations in the surrounding area. An example is the incorporation of independent gravity and height information with observations in a GPS network, which is necessary for the determination of orthometric rather than ellipsoidal heights (**Hein**, et. al, 1988).

5. Testing in four-dimensional integrated geodesy

Four-dimensional integrated geodesy models, concerned with the study of crustal deformation with repeated observations in three-dimensional integrated networks, have the general form (**Dermanis** and **Rossikopoulos**, 1988)

$$(111) \quad \begin{bmatrix} \mathbf{b}_0 \\ \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_n \end{bmatrix} = \begin{bmatrix} \mathbf{A}_0 \\ \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_n \end{bmatrix} \mathbf{x}_0 + \begin{bmatrix} \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{A}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{A}_n \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \vdots \\ \mathbf{p}_n \end{bmatrix} + \begin{bmatrix} \mathbf{G}_0 \\ \mathbf{G}_1 \\ \vdots \\ \mathbf{G}_n \end{bmatrix} \mathbf{s}_0 + \begin{bmatrix} \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{G}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{G}_n \end{bmatrix} \begin{bmatrix} \mathbf{q}_1 \\ \mathbf{q}_2 \\ \vdots \\ \mathbf{q}_n \end{bmatrix} + \begin{bmatrix} \mathbf{v}_0 \\ \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{bmatrix}$$

or simply

$$(112) \quad \mathbf{b} = \mathbf{A}^0 \mathbf{x}_0 + \mathbf{G}^p \mathbf{p} + \mathbf{G}^s \mathbf{s}_0 + \mathbf{G}^q \mathbf{q} + \mathbf{v}$$

where \mathbf{x}_0 are the coordinates at a reference epoch t_0 , \mathbf{p} are the coordinate time-variations, \mathbf{s}_0 are the gravity related signals at the reference epoch t_0 , and \mathbf{q} are the time-variations of the gravity signals. In such a model, the only discrete parameters are initially the reference coordinates \mathbf{x}_0 . In addition to the gravity related signals \mathbf{s}_0 and \mathbf{q} , which depend on the gravity potential (now a function of space and time), the coordinate variations are also considered as signals. The reason for this approach is that position variation is in principle defined at every point, and the signals \mathbf{p} depend on a "position variation" function of space and time. Note that the matrices \mathbf{A}_k and \mathbf{G}_k at different epochs are not taken to be identical, in order to include the possibility of missing or additional observations (and even network points) at different epochs t_k .

From the above general model, particular models result by treating each signal group according to one of the following four alternatives:

1. *deterministic approach*: the signals are considered discrete deterministic quantities disregarding their dependence on an underlying function
2. *analytical approach*: the underlying function is modeled as a known deterministic function depending on a new set of discrete parameters
3. *stochastic approach*: the underlying function is modeled as a stochastic process with known mean and covariance function.

4. *hybrid approach*: the underlying function is modeled as the sum of a deterministic function depending on a set of discrete parameters (trend function) and a stochastic part with known mean and covariance function.

The hybrid approach can be considered as the most general one containing the others as special cases: Vanishing trend for the stochastic approach and vanishing (zero mean and covariance) stochastic part for the analytical approach. Even the deterministic approach can be considered as a special case of the analytical approach, where the trend function becomes as a sum of Dirac impulses with the signal values at the particular points and epochs as coefficients. For an account of the different models resulting from combinations of the above alternatives see **Dermanis and Rossikopoulos** (1988), where also a classification of previous work with respect of the above framework is presented.

For the purpose of testing, it is sufficient to consider the most general case where the signals are modeled as the sum of a trend and a stochastic part, as explained in section 3:

$$(113) \quad \mathbf{p} = \mathbf{F}_p \mathbf{a}_p + \delta \mathbf{p}, \quad \delta \mathbf{p} \sim (\mathbf{0}, \mathbf{C}_p),$$

$$(114) \quad \mathbf{s}_0 = \mathbf{F}_s \mathbf{a}_s + \delta \mathbf{s}_0, \quad \delta \mathbf{s}_0 \sim (\mathbf{0}, \mathbf{C}_{s_0}),$$

$$(115) \quad \mathbf{q} = \mathbf{F}_q \mathbf{a}_q + \delta \mathbf{q}, \quad \delta \mathbf{q} \sim (\mathbf{0}, \mathbf{C}_q).$$

Note that the deterministic approach is included in this formulation by setting, e.g., $\delta \mathbf{p} = \mathbf{0}$ and $\mathbf{F}_p = \mathbf{I}$. Replacing the above terms in (112) and introducing the new matrices

$$(116) \quad \mathbf{B}_p = \mathbf{G}^p \mathbf{F}_p, \quad \mathbf{B}_s = \mathbf{G}^0 \mathbf{F}_s, \quad \mathbf{B}_q = \mathbf{G}^q \mathbf{F}_q,$$

we arrive at the following model, which has the same form as eq. (61):

$$(117) \quad \mathbf{b} = [\mathbf{A}^0 \mathbf{B}_p \mathbf{B}_s \mathbf{B}_q] \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{a}_p \\ \mathbf{a}_s \\ \mathbf{a}_q \end{bmatrix} + [\mathbf{G}^p \mathbf{G}^0 \mathbf{G}^q \mathbf{I}] \begin{bmatrix} \delta \mathbf{p} \\ \delta \mathbf{s}_0 \\ \delta \mathbf{q} \\ \mathbf{v} \end{bmatrix},$$

where deterministic and stochastic parameters are separated. The stochastic parameters are considered to have zero means, since any known non-zero mean can be incorporated in the approximate values used for the linearization. They are also considered to be uncorrelated except for the non-zero cross-covariance $\mathbf{C}_{\delta \mathbf{p}, \delta \mathbf{q}}$, which reflects the common origin of crustal motion and gravity field variation, namely the redistribution of masses within the earth. In practice however such a cross-covariance model is very difficult to obtain.

Now the tests described in section 4 can be directly applied to the above general model. Crustal motion can be identified by testing the significance of parameters in \mathbf{a}_p , which may contain individual coordinate variations (deterministic approach) or parameters describing a deformation function, see e.g. **Chen** (1983), **Chrzanowski et al.** (1983), for a purely analytical approach, where strain parameters are used for the description of the deformation function and **Chen et al.** (1990) for a related testing strategy. The existence of related gravity variations can be investigated, as far as its deterministic-trend part is concerned, by testing the significance of parameters in \mathbf{a}_q . Tests for the significance of parameters in \mathbf{a}_s can be used to check the deterministic part of the model used for the gravity potential function at the reference epoch t_0 . All such tests on trend parameters, can be performed using eq. (81) for testing a single parameter and eq. (80) for more than one. Again the problem of deciding between models which include a subset of the original trend parameters, the rest discarded as insignificant, or even additional trend parameters that proved to be significant, has no unique solution as already discusses in section 4.2.

Another problem is the possible masking of (multiple) gross errors as trend parameters and especially

crustal motion parameters. A set of gross errors can be described by a term \mathbf{Zy} , discussed in section 2.1, where the columns of \mathbf{Z} are the columns of the identity matrix \mathbf{I} , corresponding to the observations affected by the gross errors y_i . The possibility of having gross errors absorbed in the trend parameters, say \mathbf{a}_p , depends on the relation between the column spaces $R(\mathbf{Z})$ and $R(\mathbf{B}_p)$. No absorption is possible in the case where $R(\mathbf{Z}) \cap R(\mathbf{B}_p) = \emptyset$ and complete absorption can occur when $R(\mathbf{Z}) \subset R(\mathbf{B}_p)$. The problem is therefore one of "design", since according to eq. (116) the form of \mathbf{B}_p depends on both the network design (matrix \mathbf{G}^p) and the choice of basis functions for the trend (matrix \mathbf{F}_p). The problem of separability of deformations and gross errors has been studied by **Förstner** (1983) and **Zhao Shaorong** (1990).

The stochastic part of the model can be tested in two ways: with respect to the hypothesis of zero means, and with respect of the covariance models used for each one of the stochastic signals $\delta\mathbf{p}$, $\delta\mathbf{s}_0$ and $\delta\mathbf{q}$. To test the mean of any such signal, say $\delta\mathbf{p}$, it is sufficient to test for additive to $\delta\mathbf{p}$ (multiple or single) outliers. This means that the above model (117) must be compared with an "extended model" where a term

$$(118) \quad \mathbf{Zy} = \mathbf{G}^p \mathbf{E} \mathbf{y}$$

has been added, where \mathbf{E} is a matrix with elements 0 or 1, whose columns are those columns of the identity matrix \mathbf{I} which correspond to components of $\delta\mathbf{p}$ suspected for having non-zero means. The test can be performed using equations (9) to (12) with

$$\mathbf{Z} = \mathbf{G}^p \mathbf{E}, \quad \mathbf{A} = [\mathbf{A}^0 \mathbf{B}_p \mathbf{B}_s \mathbf{B}_q], \quad \mathbf{e} = \mathbf{G}^p \delta\mathbf{p} + \mathbf{G}^0 \delta\mathbf{s}_0 + \mathbf{G}^q \delta\mathbf{q} + \mathbf{v},$$

and \mathbf{P}^{-1} being replaced by the covariance matrix of \mathbf{e} .

The covariance function models can be tested with the introduction of additional variance components which must be estimated and tested for significance, as explained in section 4.3.

A remaining problem is the decision on whether a stochastic part should be included at all in the model, i.e. whether its effect is sufficiently described by solely its deterministic trend counterpart, or it is altogether absent. This can be achieved by including a single variance factor for each signal group, i.e. $\mathbf{C}_p = s_p^2 \mathbf{Q}_p$, $\mathbf{C}_{s_0} = s_0^2 \mathbf{Q}_{s_0}$, $\mathbf{C}_q = s_q^2 \mathbf{Q}_q$, and test if a particular one from s_p^2 , s_0^2 or s_q^2 is significant or not.

Multivariate models have been extensively used for the study of crustal deformations, see e.g. **Koch** (1980), **Schaffrin** (1981), **Koch and Fritsch** (1981), **Koch** (1983). The model (111) is essentially a multivariate model, especially when the matrices \mathbf{A}_k and \mathbf{G}_k are identical. As explained in section 2.3, the multivariate model is not applicable with a stochastic treatment of gravity signals (i.e. when the stochastic terms $\delta\mathbf{s}_0$ and $\delta\mathbf{q}$ are included in the model), and similar objections can be raised for the deformation stochastic signal $\delta\mathbf{p}$. In the case that these signals are treated with only deterministic trend parts, the only stochastic part is the observational noise \mathbf{v} . The individual epoch noise components $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_n$, are assumed to be identically distributed, and their common covariance matrix can be estimated within the multivariate model. Since there is no reason for correlation between individual errors, tests can be used on the estimated covariance matrix for the vanishing of its off-diagonal terms, or its equality with a given diagonal covariance matrix when the accuracy of the observations is well known a-priori. Failure of such tests is an indication for the presence of model errors, in the same way that the "congruency test", based on the variance estimate of the univariate case, is used for the same purpose.

It must be mentioned that collocation does not necessarily has to rely on a stochastic interpretation, since a purely deterministic "minimum norm" interpretation is also possible, where the covariance function is simply the reproducing kernel of a Hilbert space where the unknown function is assumed to belong. In this case the "covariance" matrices of the signals should not be taken into account when the covariances of the estimates are evaluated through covariance propagation. This will lead to different distributions for the various statistics used in hypotheses testing. However such an approach suffers from the fact that the arbitrariness associated with the choice of a single function (the one with minimum norm) out of infinite

possible ones, is not reflected at all in the a posteriori distributions.

As a closing remark let it be repeated that statistical tests, when successful, they do not establish by any means the validity of the model used, but merely its appropriateness for the set of observations at hand. One of the objects of network design, or design of experiment in the more general sense, is the detection of additional observations, which will answer questions about shortcomings of the model that cannot be identified from the available data.

Beyond statistical tests, a way to evaluate a particular model, as well as the corresponding experimental design, is the examination of the correlations between the parameter estimates. High correlation points out the inability of separating parameters, in which case the estimates may considerably vary from their true values, as information may flow from one parameter to the other through the channel of correlation.

Appendix A: Consistency test for two sets of observations with correlated residuals.

Consider two sets of observations in the most general case where the stochastic residuals \mathbf{e}_1 , \mathbf{e}_2 are correlated

$$(A1) \quad \begin{aligned} \mathbf{b}_1 &= \mathbf{A}_1 \mathbf{x} + \mathbf{e}_1 \\ \mathbf{b}_2 &= \mathbf{A}_2 \mathbf{x} + \mathbf{e}_2 \end{aligned}$$

with means and covariances

$$(A2) \quad \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{bmatrix} \sim \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{12}^T & \mathbf{Q}_{22} \end{bmatrix} \right)$$

with (collective) weight matrix

$$(A3) \quad \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{12}^T & \mathbf{P}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{12}^T & \mathbf{Q}_{22} \end{bmatrix}^{-1}$$

where

$$(A4) \quad \mathbf{P}_{11} = (\mathbf{Q}_{11} - \mathbf{Q}_{12} \mathbf{Q}_{22}^{-1} \mathbf{Q}_{12}^T)^{-1} = \mathbf{Q}_{11}^{-1} + \mathbf{Q}_{11}^{-1} \mathbf{Q}_{12} \mathbf{P}_{22} \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1}$$

$$(A5) \quad \mathbf{P}_{22} = (\mathbf{Q}_{22} - \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} \mathbf{Q}_{12})^{-1}$$

$$(A6) \quad \mathbf{P}_{11} = -\mathbf{Q}_{11}^{-1} \mathbf{Q}_{12} \mathbf{P}_{22}$$

$$(A7) \quad \mathbf{P}_{12}^T = -\mathbf{P}_{22} \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1}.$$

The solution of only the first set $\mathbf{b}_1 = \mathbf{A}_1 \mathbf{x} + \mathbf{e}_1$, with (individual) weight matrix $\mathbf{P}_1 = \mathbf{Q}_{11}^{-1} \neq \mathbf{P}_{11}$, leads to the estimates

$$(A8) \quad \begin{aligned} \hat{\mathbf{x}}^{(1)} &= (\mathbf{A}_1^T \mathbf{P}_1 \mathbf{A}_1)^{-1} \mathbf{A}_1^T \mathbf{P}_1 \mathbf{b}_1 = \mathbf{N}_1^{-1} \mathbf{A}_1^T \mathbf{P}_1 \mathbf{b}_1 = \\ &= \mathbf{N}_1^{-1} \mathbf{A}_1^T \mathbf{P}_1 (\mathbf{A}_1 \mathbf{x} + \mathbf{e}_1) = \mathbf{x} + \mathbf{N}_1^{-1} \mathbf{A}_1^T \mathbf{P}_1 \mathbf{e}_1 \end{aligned}$$

$$(A9) \quad \hat{\mathbf{e}}_1^{(1)} = \mathbf{b}_1 - \mathbf{A}_1 \hat{\mathbf{x}}^{(1)} = \mathbf{A}_1 \mathbf{x} + \mathbf{e}_1 - \mathbf{A}_1 (\mathbf{x} + \mathbf{N}_1^{-1} \mathbf{A}_1^T \mathbf{P}_1 \mathbf{e}_1) = (\mathbf{I} - \mathbf{A}_1 \mathbf{N}_1^{-1} \mathbf{A}_1^T \mathbf{P}_1) \mathbf{e}_1$$

Due to the correlation between the residuals it is possible to obtain a prediction of \mathbf{e}_2 , which is based on only the first set

$$(A10) \quad \tilde{\mathbf{e}}_2^{(1)} = \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} \hat{\mathbf{e}}_1^{(1)} = \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} (\mathbf{I} - \mathbf{A}_1 \mathbf{N}_1^{-1} \mathbf{A}_1^T \mathbf{P}_1) \mathbf{e}_1 = \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} \mathbf{e}_1 - \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} \mathbf{A}_1 \mathbf{N}_1^{-1} \mathbf{A}_1^T \mathbf{P}_1 \mathbf{e}_1$$

Introducing the estimate $\hat{\mathbf{x}}^{(1)}$ and the prediction $\tilde{\mathbf{e}}_2^{(1)}$ in the second set, the following inconsistency term is obtained

$$(A11) \quad \begin{aligned} \boldsymbol{\delta} &= (\mathbf{b}_2 - \mathbf{A}_2 \hat{\mathbf{x}}^{(1)}) - \tilde{\mathbf{e}}_2^{(1)} = \\ &= \mathbf{A}_2 \mathbf{x} + \mathbf{e}_2 - \mathbf{A}_2 (\mathbf{x} + \mathbf{N}_1^{-1} \mathbf{A}_1^T \mathbf{P}_1 \mathbf{e}_1) - \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} \mathbf{e}_1 + \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} \mathbf{A}_1 \mathbf{N}_1^{-1} \mathbf{A}_1^T \mathbf{P}_1 \mathbf{e}_1 = \\ &= \mathbf{e}_2 - \mathbf{A}_2 \mathbf{N}_1^{-1} \mathbf{A}_1^T \mathbf{P}_1 \mathbf{e}_1 - \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} \mathbf{e}_1 + \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} \mathbf{A}_1 \mathbf{N}_1^{-1} \mathbf{A}_1^T \mathbf{P}_1 \mathbf{e}_1 = \\ &= (\mathbf{e}_2 - \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} \mathbf{e}_1) - (\mathbf{A}_2 - \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} \mathbf{A}_1) \mathbf{N}_1^{-1} \mathbf{A}_1^T \mathbf{P}_1 \mathbf{e}_1 = \\ &= (\mathbf{e}_2 - \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} \mathbf{e}_1) - \mathbf{K} \mathbf{N}_1^{-1} \mathbf{A}_1^T \mathbf{P}_1 \mathbf{e}_1 = \end{aligned}$$

where, using (A7)

$$(A12) \quad \mathbf{K} = \mathbf{A}_2 - \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} \mathbf{A}_1 = \mathbf{A}_2 + \mathbf{P}_{22}^{-1} \mathbf{P}_{12}^T \mathbf{A}_1$$

or

$$(A13) \quad \boldsymbol{\delta} = \begin{bmatrix} -(\mathbf{Q}_{12}^T + \mathbf{K} \mathbf{N}_1^{-1} \mathbf{A}_1^T) \mathbf{Q}_{11}^{-1} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{bmatrix}.$$

The covariance matrix \mathbf{Q}_δ of $\boldsymbol{\delta}$ is obtained by direct application of the propagation law

$$(A14) \quad \begin{aligned} \mathbf{Q}_\delta &= \begin{bmatrix} -(\mathbf{Q}_{12}^T + \mathbf{K} \mathbf{N}_1^{-1} \mathbf{A}_1^T) \mathbf{Q}_{11}^{-1} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{12}^T & \mathbf{Q}_{22} \end{bmatrix} \begin{bmatrix} -\mathbf{Q}_{11}^{-1} (\mathbf{Q}_{12} + \mathbf{A}_1 \mathbf{N}_1^{-1} \mathbf{K}^T) \\ \mathbf{I} \end{bmatrix} = \\ &= \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} \mathbf{Q}_{12} + \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} \mathbf{A}_1 \mathbf{N}_1^{-1} \mathbf{K}^T + \mathbf{K} \mathbf{N}_1^{-1} \mathbf{A}_1^T \mathbf{Q}_{11}^{-1} \mathbf{Q}_{12} + \mathbf{K} \mathbf{N}_1^{-1} \mathbf{K}^T - \\ &\quad - \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} \mathbf{Q}_{12} - \mathbf{K} \mathbf{N}_1^{-1} \mathbf{A}_1^T \mathbf{Q}_{11}^{-1} \mathbf{Q}_{12} - \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} \mathbf{Q}_{12} + \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} \mathbf{A}_1 \mathbf{N}_1^{-1} \mathbf{K}^T + \mathbf{Q}_{12} \end{aligned}$$

or

$$(A15) \quad \mathbf{Q}_\delta = (\mathbf{Q}_{22} - \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} \mathbf{Q}_{12})^{-1} + \mathbf{K} \mathbf{N}_1^{-1} \mathbf{K}^T = \mathbf{P}_{22}^{-1} + \mathbf{K} \mathbf{N}_1^{-1} \mathbf{K}^T.$$

In order to test the compatibility of the two sets, an inconsistency term \mathbf{y} is added in one of the sets (say, the second) and consistency will be checked by testing the hypothesis $H_0: \mathbf{y} = \mathbf{0}$ vs $H_a: \mathbf{y} \neq \mathbf{0}$,

$$(A16) \quad \begin{aligned} \mathbf{b}_1 &= \mathbf{A}_1 \mathbf{x} + \mathbf{e}_1 \\ \mathbf{b}_2 &= \mathbf{A}_2 \mathbf{x} + \mathbf{y} + \mathbf{e}_2 \end{aligned}$$

or

$$(A17) \quad \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix} \mathbf{x} + \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} \mathbf{y} + \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{bmatrix}$$

or, with obvious notation

$$(A18) \quad \mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{Z}\mathbf{y} + \mathbf{e}$$

$$(A19) \quad \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{bmatrix} \sim \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{12}^T & \mathbf{Q}_{22} \end{bmatrix} \right)$$

First it will be proved that $\mathbf{Q}_{\hat{\mathbf{y}}} = \mathbf{Q}_{\delta}$.

Using equation (11)

$$(A20) \quad \begin{aligned} \mathbf{Q}_{\hat{\mathbf{y}}} &= [\mathbf{Z}^T \mathbf{P} (\mathbf{P}^{-1} - \mathbf{A}\mathbf{N}^{-1} \mathbf{A}^T) \mathbf{P}\mathbf{Z}]^{-1} = [\mathbf{W}^T (\mathbf{P}^{-1} - \mathbf{A}\mathbf{N}^{-1} \mathbf{A}^T) \mathbf{W}]^{-1} = \\ &= [\mathbf{W}^T \mathbf{P}^{-1} \mathbf{W} - \mathbf{W}^T \mathbf{A}\mathbf{N}^{-1} \mathbf{A}^T \mathbf{W}]^{-1} \end{aligned}$$

where

$$(A21) \quad \mathbf{W}^T = \mathbf{Z}^T \mathbf{P}, \quad \mathbf{W} = \mathbf{P}\mathbf{Z} = [\mathbf{P}_{12}^T \quad \mathbf{P}_{22}]^T$$

and

$$(A22) \quad \mathbf{W}^T \mathbf{P}^{-1} \mathbf{W} = \mathbf{Z}^T \mathbf{P} \mathbf{P}^{-1} \mathbf{P}\mathbf{Z} = \mathbf{Z}^T \mathbf{P}\mathbf{Z} = \mathbf{P}_{22}.$$

Using the well known matrix identity

$$(A23) \quad (\mathbf{A} - \mathbf{B}\mathbf{D}\mathbf{B}^T)^{-1} = \mathbf{A}^{-1} + \mathbf{A}^{-1} \mathbf{B} (\mathbf{D}^{-1} - \mathbf{B}^T \mathbf{A}^{-1} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{A}^{-1}.$$

with $\mathbf{A} \rightarrow \mathbf{W}^T \mathbf{P}^{-1} \mathbf{W}$, $\mathbf{B} \rightarrow \mathbf{W}^T \mathbf{A}$, $\mathbf{D} \rightarrow \mathbf{N}^{-1}$, it follows, after setting $\mathbf{N} - \mathbf{A}^T \mathbf{W} \mathbf{P}_{22}^{-1} \mathbf{W}^T \mathbf{A} \equiv \mathbf{X}$, that

$$(A24) \quad \begin{aligned} \mathbf{Q}_{\hat{\mathbf{y}}} &= \mathbf{P}_{22}^{-1} + \mathbf{P}_{22}^{-1} \mathbf{W}^T \mathbf{A} (\mathbf{N} - \mathbf{A}^T \mathbf{W} \mathbf{P}_{22}^{-1} \mathbf{W}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W} \mathbf{P}_{22}^{-1} = \\ &= \mathbf{P}_{22}^{-1} + \mathbf{P}_{22}^{-1} [\mathbf{P}_{12}^T \quad \mathbf{P}_{22}] \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix} \mathbf{X}^{-1} [\mathbf{A}_1^T \quad \mathbf{A}_2^T] \begin{bmatrix} \mathbf{P}_{12} \\ \mathbf{P}_{22} \end{bmatrix} \mathbf{P}_{22}^{-1} = \\ &= \mathbf{P}_{22}^{-1} + [\mathbf{P}_{22}^{-1} \mathbf{P}_{12}^T \quad \mathbf{I}] \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix} \mathbf{X}^{-1} [\mathbf{A}_1^T \quad \mathbf{A}_2^T] \begin{bmatrix} \mathbf{P}_{12} \mathbf{P}_{22}^{-1} \\ \mathbf{I} \end{bmatrix} = \\ &= \mathbf{P}_{22}^{-1} + (\mathbf{A}_2 + \mathbf{P}_{22}^{-1} \mathbf{P}_{12}^T \mathbf{A}_1) \mathbf{X}^{-1} (\mathbf{A}_2 + \mathbf{P}_{22}^{-1} \mathbf{P}_{12}^T \mathbf{A}_1)^T \end{aligned}$$

where

$$(A25) \quad \mathbf{X} = \mathbf{N} - \mathbf{A}^T \mathbf{W} \mathbf{P}_{22}^{-1} \mathbf{W}^T \mathbf{A} =$$

$$\begin{aligned}
&= [\mathbf{A}_1^T \mathbf{A}_2^T] \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{12}^T & \mathbf{P}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix} - [\mathbf{A}_1^T \mathbf{A}_2^T] \begin{bmatrix} \mathbf{P}_{12} \\ \mathbf{P}_{22} \end{bmatrix} \mathbf{P}_{22}^{-1} [\mathbf{P}_{12}^T \mathbf{P}_{22}] \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix} = \\
&= \mathbf{A}_1^T (\mathbf{P}_{11} - \mathbf{P}_{12} \mathbf{P}_{22}^{-1} \mathbf{P}_{12}^T) \mathbf{A}_1 = \mathbf{A}_1^T \mathbf{Q}_{11}^{-1} \mathbf{A}_1 = \mathbf{N}_1.
\end{aligned}$$

and therefore

$$(A26) \quad \mathbf{Q}_{\hat{\mathbf{y}}} = [\mathbf{Z}^T \mathbf{P} (\mathbf{P}^{-1} - \mathbf{A} \mathbf{N}^{-1} \mathbf{A}^T) \mathbf{P} \mathbf{Z}]^{-1} = \mathbf{P}_{22}^{-1} + \mathbf{K} \mathbf{N}_1^{-1} \mathbf{K}^T$$

where

$$(A27) \quad \mathbf{K} = \mathbf{A}_2 - \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} \mathbf{A}_1 = \mathbf{A}_2 + \mathbf{P}_{22}^{-1} \mathbf{P}_{12}^T \mathbf{A}_1.$$

Next it will be proved that $\hat{\mathbf{y}} = \delta$.

Using equation (10)

$$(A28) \quad \hat{\mathbf{y}} = \mathbf{Q}_{\hat{\mathbf{y}}} \mathbf{Z}^T \mathbf{P} \hat{\mathbf{e}} = [\mathbf{P}_{22}^{-1} + \mathbf{K} \mathbf{N}_1^{-1} \mathbf{K}^T] [\mathbf{P}_{12}^T \ \mathbf{P}_{22}] \begin{bmatrix} \hat{\mathbf{e}}_1 \\ \hat{\mathbf{e}}_2 \end{bmatrix}.$$

Furthermore

$$\begin{aligned}
(A29) \quad [\mathbf{P}_{12}^T \ \mathbf{P}_{22}] \begin{bmatrix} \hat{\mathbf{e}}_1 \\ \hat{\mathbf{e}}_2 \end{bmatrix} &= \mathbf{P}_{12}^T (\mathbf{b}_1 - \mathbf{A}_1 \hat{\mathbf{x}}) + \mathbf{P}_{22} (\mathbf{b}_2 - \mathbf{A}_2 \hat{\mathbf{x}}) = \mathbf{P}_{12}^T \mathbf{b}_1 - \mathbf{P}_{12}^T \mathbf{A}_1 \hat{\mathbf{x}} + \mathbf{P}_{22} \mathbf{b}_2 - \mathbf{P}_{22} \mathbf{A}_2 \hat{\mathbf{x}} = \\
&= \mathbf{P}_{22} (\mathbf{b}_2 + \mathbf{P}_{22}^{-1} \mathbf{P}_{12}^T \mathbf{b}_1) - \mathbf{P}_{22} (\mathbf{A}_2 + \mathbf{P}_{22}^{-1} \mathbf{P}_{12}^T \mathbf{A}_1) \hat{\mathbf{x}} = \mathbf{P}_{22} (\delta + \mathbf{K} (\hat{\mathbf{x}}^{(1)} - \hat{\mathbf{x}})),
\end{aligned}$$

where it has been taken into account that

$$(A30) \quad \mathbf{b}_2 + \mathbf{P}_{22}^{-1} \mathbf{P}_{12}^T \mathbf{b}_1 = \delta + \mathbf{K} \hat{\mathbf{x}}^{(1)}.$$

Van Mierlo (1988, equations 2.6 and 2.7) has proved that

$$(A31) \quad \hat{\mathbf{x}} = \hat{\mathbf{x}}^{(1)} + \mathbf{G} \delta$$

$$(A32) \quad \mathbf{G} = \mathbf{N}_1^{-1} \mathbf{K}^T (\mathbf{P}_{22}^{-1} + \mathbf{K} \mathbf{N}_1^{-1} \mathbf{K}^T)^{-1}.$$

Taking the above result into account it holds that

$$(A33) \quad [\mathbf{P}_{12}^T \ \mathbf{P}_{22}] \begin{bmatrix} \hat{\mathbf{e}}_1 \\ \hat{\mathbf{e}}_2 \end{bmatrix} = \mathbf{P}_{22} [\mathbf{I} - \mathbf{K} \mathbf{N}_1^{-1} \mathbf{K}^T (\mathbf{P}_{22}^{-1} + \mathbf{K} \mathbf{N}_1^{-1} \mathbf{K}^T)^{-1}] \delta$$

and using the obvious identity

$$(A34) \quad \mathbf{I} = [\mathbf{P}_{22}^{-1} + \mathbf{K} \mathbf{N}_1^{-1} \mathbf{K}^T] [\mathbf{P}_{22}^{-1} + \mathbf{K} \mathbf{N}_1^{-1} \mathbf{K}^T]^{-1}$$

eq. (A32) becomes

$$(A35) \quad [\mathbf{P}_{12}^T \ \mathbf{P}_{22}] \begin{bmatrix} \hat{\mathbf{e}}_1 \\ \hat{\mathbf{e}}_2 \end{bmatrix} = (\mathbf{P}_{22}^{-1} + \mathbf{K} \mathbf{N}_1^{-1} \mathbf{K}^T)^{-1} \delta$$

and finally

$$(A36) \quad \hat{\mathbf{y}} = (\mathbf{P}_{22}^{-1} + \mathbf{KN}_1^{-1}\mathbf{K}^T)(\mathbf{P}_{22}^{-1} + \mathbf{KN}_1^{-1}\mathbf{K}^T)^{-1}\boldsymbol{\delta} = \boldsymbol{\delta}.$$

In order to obtain the test statistic in terms of $\boldsymbol{\delta}$, it will be proved that

$$(A37) \quad f\hat{\boldsymbol{\sigma}}^2 - \boldsymbol{\delta}^T \mathbf{Q}_{\boldsymbol{\delta}}^{-1} \boldsymbol{\delta} = (f - n_2)\hat{\boldsymbol{\sigma}}_1^2.$$

From **van Mierlo** (1988, eq. 2.25):

$$(A38) \quad f\hat{\boldsymbol{\sigma}}^2 = (f - n_2)\hat{\boldsymbol{\sigma}}_1^2 + \mathbf{e}_2^{*T} \mathbf{P}_{22} \mathbf{e}_2^* + (\hat{\mathbf{x}} - \hat{\mathbf{x}}^{(1)})^T \mathbf{N}_1 (\hat{\mathbf{x}} - \hat{\mathbf{x}}^{(1)})$$

where

$$(A39) \quad \mathbf{e}_2^* = \mathbf{e}_2 - \mathbf{Q}_{12}^T \mathbf{Q}_{11}^{-1} \mathbf{e}_1 = \mathbf{e}_2 + \mathbf{P}_{22}^{-1} \mathbf{P}_{12}^T \mathbf{e}_1 = \mathbf{P}_{22}^{-1} [\mathbf{P}_{12}^T \mathbf{P}_{22}] \begin{bmatrix} \hat{\mathbf{e}}_1 \\ \hat{\mathbf{e}}_2 \end{bmatrix} = \mathbf{P}_{22}^{-1} (\mathbf{P}_{22}^{-1} + \mathbf{KN}_1^{-1}\mathbf{K}^T)^{-1} \boldsymbol{\delta}.$$

Using this result it follows that

$$(A40) \quad \mathbf{e}_2^{*T} \mathbf{P}_{22} \mathbf{e}_2^* = \boldsymbol{\delta}^T (\mathbf{P}_{22}^{-1} + \mathbf{KN}_1^{-1}\mathbf{K}^T)^{-1} \mathbf{P}_{22}^{-1} (\mathbf{P}_{22}^{-1} + \mathbf{KN}_1^{-1}\mathbf{K}^T)^{-1} \boldsymbol{\delta}$$

$$(A41) \quad (\hat{\mathbf{x}} - \hat{\mathbf{x}}^{(1)})^T \mathbf{N}_1 (\hat{\mathbf{x}} - \hat{\mathbf{x}}^{(1)}) = \boldsymbol{\delta}^T (\mathbf{P}_{22}^{-1} + \mathbf{KN}_1^{-1}\mathbf{K}^T)^{-1} \mathbf{KN}_1^{-1}\mathbf{K}^T (\mathbf{P}_{22}^{-1} + \mathbf{KN}_1^{-1}\mathbf{K}^T)^{-1} \boldsymbol{\delta}$$

and finally

$$(A42) \quad \begin{aligned} f\hat{\boldsymbol{\sigma}}^2 &= (f - n_2)\hat{\boldsymbol{\sigma}}_1^2 + \mathbf{e}_2^{*T} \mathbf{P}_{22} \mathbf{e}_2^* + (\hat{\mathbf{x}} - \hat{\mathbf{x}}^{(1)})^T \mathbf{N}_1 (\hat{\mathbf{x}} - \hat{\mathbf{x}}^{(1)}) = \\ &= (f - n_2)\hat{\boldsymbol{\sigma}}_1^2 + \boldsymbol{\delta}^T (\mathbf{P}_{22}^{-1} + \mathbf{KN}_1^{-1}\mathbf{K}^T)^{-1} \boldsymbol{\delta} = (f - n_2)\hat{\boldsymbol{\sigma}}_1^2 + \boldsymbol{\delta}^T \mathbf{Q}_{\boldsymbol{\delta}}^{-1} \boldsymbol{\delta}. \end{aligned}$$

References

- Anderson, T.W. (1958): "*An Introduction to Multivariate Analysis*". Wiley, New York.
- 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.
- Bock, Y. and B. Schaffrin (1988): "Deformation Analysis based on Robust Inverse Theory. Part 1: Theoretical Background". Prep. for 17th Conf. on Math. Geophysics, Blanes (Spain), June 1988.
- Broemeling, L.D. (1985): "*Bayesian Analysis of Linear Models*". M. Dekker Inc., New York.
- Chen, Y.Q. (1983): "Analysis of Deformation Surveys - A Generalized Method". University of New Brunswick, Dept. of Surveying Engineering, Technical Report No. 94.
- Chrzanowski, A., Y.Q. Chen and J.M. Secord (1983): "On the Strain Analysis of Tectonic Movements Using Fault Crossing Geodetic Surveys". *Tectonophysics*, 97, 297-315.
- Chen, Y.Q. (1990): "A Strategy for the Analysis of the Stability of Reference Points in Deformation Surveys". *CISM Journal ACSGC*, 44, 2, 141-149.
- Cook, R.D. and S. Weisberg (1982): "*Residuals and Influence in Regression*". Chapman and Hall, Lon-

don.

- Dermanis, A. (1987): "*Adjustment of Observations and Estimation Theory*", vols 1 & 2. Editions Ziti, Thessaloniki (in greek).
- Dermanis, A. and D. Rossikopoulos (1988): "Modeling Alternatives in Four-dimensional Geodesy". In: Proc. of the International Symposium on Instrumentation Theory and Analysis for Integrated Geodesy, Sopron, Mai 16 to 20, Vol. 2, 115-145.
- Ebner, H. (1972): "A posteriori Varianzschätzung für die Koordination unabhängiger Modelle". ZfV, 98, 166-172.
- Eeg, J. and T. Krarup (1975): "Integrated Geodesy", in: Brosowski, B. and E. Martensen (eds.), Methoden und Verfahren der Mathematischen Physik, Band 13, Mathematical Geodesy, Part II, Bibliographisches Inst., Mannheim.
- Förstner, W. (1979): "Ein Verfahren zur Schätzung von Varianz- und Kovarianzkomponenten", AVN, 86, no. 11-12, 446-453.
- Förstner, W. (1983): "Reliability and Dscernability of Extended Gaus-Markov Models". DGK, Reihe A, Heft Nr. 98, 79-102.
- Grafarend, E. (1978): "Shätzung von Varianz und Kovarianz der Beobachtugen in geodätischen Ausgleichungsmodellen". AVN 85, 2, 41-49.
- Grafarend, E. (1984): "Variance-covariance component estimation of Helmert type in the Gaus-Helmert model". ZfV, 109, 1, 34-44.
- Grafarend, E., A. Kleusberg and B. Schaffrin (1980): "An Introduction to the Variance-Covariance Component Estimation of Helmert Type". ZfV, 105, 4, 161-179.
- Grafarend, E.W. (1985): "Variance-covariance Estimation, Theoretical Results and Geodetic Applications". Statistics and Decisions, Supplement Issue 2, 407-441.
- Graybill, F.A. (1976): "*Theory and Application of the Linear Model*". Duxbury Press, North Scituate, Massachusetts.
- Hein, G.W., A. Leick and S. Lambert (1988): "Integrated Processing of GPS and Gravity Data". Univ. of Maine, Dept. of Civil Engineering, Report 87.
- Helmert, F.R. (1907): "*Die Ausgleichung nach der Methode der kleinsten Quadrate*". 2. Auflage, Verlag von B.G. Teubner, Leipzig/Berlin.
- Kelm, R. (1978): "Ist die Varianzshätzung nach Helmert MINQUE?". AVN 85, 2, 49-54.
- Koch, K.-R. (1980): "Modelle für die Parameterschätzung bei der Deformationsanalyse". ZfV, 105, 4, 189-195.
- Koch, K.-R. (1983): "Estimation of Variance and Covariances in the Multivariate and the Incomplete Multivariate Model". DGK, Reihe A, Heft Nr. 98, 53-59.
- Koch, K.-R. (1986): "Maximum likelihood estimate of Variance Components". Bulletin Géodésique, 60, 329-338.
- Koch, K.-R. (1987a): "*Parameter Estimation and Hypothesis Testing in Linear Models*". Springer Verlag.
- Koch, K.-R. (1987b): "Bayesian Inference for Variance Components". Manuscr. Geod., 12, 309-313.
- Koch, K.-R. (1987c): "Bayesian Statistics for Variance Components with Informative and Noninformative priors". Manuscripta Geodaetica, 13, 370-373.
- Koch, K.-R. (1990): "*Bayesian Inference with Geodetic Applications*". Lecture Notes in Earth Sciences No. 31, Springer Verlag.
- Koch, K.-R. and D. Fritsch (1981): "Multivariate Hypothesis Tests for Detecting Recent Crustal Movements". Tectonophysics, 71, 301-313.
- Koch, K.-R. and K. Riesmeier (1985): "Bayesian Inference for the Derivation of Less Sensitive Hypothe-

- sis Tests". *Bulletin Géodésique*, 59, 167-179.
- Kok, J.J. (1982): "Statistical Analysis of Deformation Problems Using Baarda's Testing Procedures". In: 'Forty Years of Thought', Volume in honor of W. Baarda's 65th birthday, Vol. 2, 469-488. Department of Geodesy, Delft University of Technology.
- Krakiwsky, E. and Z. F. Biacs (1990): "Least Squares Collocation and Statistical Testing". *Bulletin Géodésique*, 64, 73-87.
- Kubáčková, L. (1990): "Locally best Estimators of the Second Order Parameters in Fundamental Replicated Structures with Nuisance Parameters". University of Stuttgart, Institute of Geodesy, Technical Report, No. 12.
- Kubik, K. (1967a): "*Schätzfunktionen für Varianzen, Kovarianzen und für andere Parameter in Ausgleichungsaufgaben*". Ph.D. Thesis, Technische Hochschule Wien.
- Kubik, K. (1967b): "Schätzung der Gewichte der Fehlergleichungen beim Ausgleichsproblem nach vermittelnden Beobachtungen". *ZfV*, 92, 173-178.
- Kubik, K. (1970): "The Estimation of Weights of Measured Quantities within the Method of Least Squares". *Bulletin Géodésique*, 95, 21-40.
- Lehmann, E.L. (1986): "*Testing Statistical Hypotheses*", 2nd edition, Wiley.
- Mardia, K.V., J.T. Kent and J.M. Bibby (1979): "*Multivariate Analysis*". Academic Press, London.
- Meier, S. (1981): "Planar Geodetic Covariance Functions". *Reviews of Geophysics & Space Physics*, 19, 4, 673-686.
- Moritz, H. (1978): "The Operational Approach to Physical Geodesy". OSU Dept. of Geodetic Science, Report No.77.
- Ou Ziqiang (1989): "Estimation of Variance and Covariance Components". *Bulletin Géodésique*, 63, 139-148.
- Persson, C.G. (1982): "Adjustment, Weight-Testing and Detection of Outliers in Mixed SFF-SFS Models". *Manuscripta Geodaetica*, 7, 299-323.
- Rao, C.R. (1970): "Estimation of Heteroscedastic Variances in Linear Models". *JASA*, 65, 161-172.
- Rao, C.R. (1971): "Estimation of Variance Components - MINQUE Theory". *Journal of Multivariate Analysis*, 1, 257-275.
- Rao, C.R. and J. Kleffe (1988): "*Estimation of Variance Components and Applications*". North-Holland, Amsterdam.
- Remmer, O. (1971): "A Test of Significance for the Helmert-Kubik-Problem of Weight-Determination". *Bull. Geod.*, 99, 29-36.
- Riesmeir, K. (1984): "*Test von Ungleichungshypothesen in linearen Modellen mit Baye-Verfahren*". DGK, Reihe C, Heft Nr. 292.
- Rossikopoulos, D. (1986): "*Integrated Control Networks*". Ph.D. Dissertation, Department of Geodesy and Surveying, University of Thessaloniki (in Greek).
- Sansó, F. (1978): "The Minimum Mean Square Estimation Error Principle in Physical Geodesy (Stochastic and Non-Stochastic Interpretation)". *Proceedings 7th Symposium on Mathematical Geodesy*, Assisi, 1978.
- Schaffrin, B. (1981): "Best Invariant Covariance Component Estimators and its Application to the Generalized Multivariate Adjustment of Heterogeneous Deformation Observations". *Bull. Géodésique*, 55, 73-85.
- Schaffrin, B. (1983): "*Varianz-Kovarianz-Komponenten Schätzung bei der Ausgleichung heterogener Wiederholungsmessungen*". Dissertation, Universität Bonn. DGK, Reihe C, Heft Nr. 282.
- Schaffrin, B. (1985): "*Das geodätische Datum mit stochastischer Vorinformation*". Habilitationsschrift, Universität Stuttgart. DGK, Reihe C, Heft Nr. 313.

- Schaffrin, B. (1987): "Less Sensitive Tests by Introducing Stochastic Linear Hypotheses". Proc. 2nd Inter. Tampere Conference in Statistics, Tampere, Finland, June 1987.
- Schaffrin, B. (1988): "Tests for Random Effects Based on Homogeneously Linear Predictors". Workshop on "Theory and Practice in Data Analysis", 19-21 August 1988, Berlin.
- Schwintzer, P. (1984): "*Analyse geodätisch gemessener Punktlageänderungen mit gemischten Modellen*". Schriftenreihe des Universitären Studiengangs Vermessungswesen der Universität der Bundeswehr München, Heft 12.
- Seber, G.A.F. (1984): "*Multivariate Observations*". Wiley, New York.
- Seifert, B. (1985): "Estimation and Test of Variance Components Using the MINQUE-Method". *Statistics*, 16, 621-635.
- Sjöberg, L. (1983a): "Unbiased Estimation of Variance-Covariance Components in Condition Adjustment". Univ. of Uppsala, Institute of Geophysics, Dept. of Geodesy, Report No. 19, 223-237.
- Sjöberg, L. (1983b): "Unbiased Estimation of Variance-Covariance Components in Condition Adjustment with Unknowns - A MINQUE Approach". *ZfV*, 108, 9, 382-387.
- Sjöberg, L. (1984): "Non-negative Variance Component Estimation in the Gaus-Helmert Adjustment Model". *Manuscripta Geodaetica* Vol. 9, 247-280.
- Sjöberg, L. (1985): "Adjustment and Variance-Covariance Component Estimation with a Singular Covariance Matrix". *ZfV*, 110, 4, 145-151.
- van Mierlo, J. (1988): "Combined Adjustment versus Sequential and Single Solutions". In: Proceedings Int. Symp. on 'Instrumentation Theory and Analysis for Integrated Geodesy', Sopron, Mai 16-20, Vol. 1, 24-36.
- Wei, M. (1987a): "*Statistische Probleme bei der Kollokation*". Dissertation, Institute für Theoretische Geodäsie, Technische Universität Graz. Mitteilungen der geod. Inst. der Tech. Univ. Graz, Folge 54, Graz, 1987.
- Wei, M. (1987b): "Statistical Problems in Collocation". *Manuscr. Geod.*, 12, 282-289.
- Welsch, W. (1978): "A posteriori Varianzschätzung nach Helmert". *AVN* 85, 2, 55-63.
- Zhang, Wanpeng (1990): "Theoretische Probleme bei Anwendungen von MINQE(I,U) der Varianzkomponenten". *ZfV*, 115, 5, 195-203.
- Zhao Shaorong (1990): "On Separability for Deformation and Gross Errors". *Bull. Géod.*, 64, 383-396.