ON THE RELIABILITY OF DATA OBTAINED BY KRIGING

Burkhard SCHAFFRIN
Department of Civil and Environmental Engineering and Geodetic Science
The Ohio State University, Columbus, OH 43210, USA
Schaffrin.1@osu.edu

Working Group IC/14

KEY WORDS: Outlier testing, reliability, Kriging, spatial data

ABSTRACT

Irregularly distributed data may be interpreted as a sample drawn from an underlying spatial process whose properties can be assumed to be known or unknown, depending on the particular situation. In case of a relatively smooth process, one of the various Kriging methods could be employed to derive gridded point data with comparable accuracy *provided that no outliers are present*. Otherwise, they have to be eliminated beforehand or, at least, their influence must be reduced to the level of random uncertainty.

Measures of *reliability* to describe the potential for identifying outliers in suspicious sample points and to quantify the effect of any undetected outliers – well-known for the Gauss-Markov Model – will be introduced for the case of a *spatial process* where the sampled data are supposedly correlated, at least in the spatial sense. In this study, we shall consider Simple as well as Ordinary Kriging which is essentially identical to "least-squares collocation" with (known, resp. unknown) constant trend.

1 INTRODUCTION: A REVIEW OF THE GAUSS-MARKOV MODEL

Originally for the use in geodetic networks, *W. Baarda (1968)* had introduced a testing procedure for outliers that is now known as "data snooping". Shortly afterwards, a theory of *reliability* was developed by *W. Baarda (1976)* which found a number of applications both in geodesy and photogrammetry. We only refer to the work by *H. Pelzer (1980)*, *S. F. El-Hakim (1982)*, and *W. Förstner (1983; 1985)*; for an overview consult, e.g., *K. R. Koch (1988, chapter 44)* or *B. Schaffrin (1988)*.

The original theory was based on a *Gauss-Markov Model*

$$y = \underset{nxm}{A}\,\xi + e\,, \qquad e \sim (0, \sigma_0^2 P^{-1})\,, \tag{1}$$

where y is the n×1 vector of observational increments;

ξ is the m×1 vector of (unknown) parameter increments;

A is the n×m matrix of coefficients with rkA=:q≤min{m,n};

e is the n×1 vector of (unknown) random observation errors;

Q:=P^{-1} is the n×n positive-definite cofactor matrix;

σ_0^2 is the (unknown) variance component.

The weighted LEast-Squares Solution (LESS) can be taken from the normal equations

$$N\hat{\xi} = c \qquad \text{for} \quad [N, c] := A^T P[A, y], \tag{2}$$

but may not be unique unless q=m. The general "solution space" may thus be represented as

$$\hat{\xi} \in \{N^- c \,|\, NN^- N = N, i.e.\ N^- \text{ is g-inverse of N}\}\,, \tag{3}$$

along with the dispersion matrices

$$D\{\hat{\xi}\} \in \{\sigma_0^2 N_{rs}^- \mid NN_{rs}^- N = N, N_{rs}^- NN_{rs}^- = N_{rs}^- = (N_{rs}^-)^T\}. \qquad (4)$$

In any case, the residual vector

$$\tilde{e} = y - A\hat{\xi} = (I_n - AN^- A^T P)\, y = (Q_{\tilde{e}} P)\, y \qquad (5)$$

will be unique with the dispersion matrix

$$D\{\tilde{e}\} = \sigma_0^2 (P^{-1} - AN^- A^T) =: \sigma_0^2 Q_{\tilde{e}}, \qquad (6)$$

and so will be the variance component estimate

$$\hat{\sigma}_0^2 = (n-q)^{-1} \tilde{e}^T P \tilde{e} = (n-q)^{-1} (y^T P y - c^T \hat{\xi}). \qquad (7)$$

Now let us assume that an outlier has occurred in the j-th observation. So, the "true" model should instead have been

$$y = A\xi + \eta_j \delta^{(j)} + e, \qquad e \sim (0, \sigma_0^2 P^{-1}), \qquad (8)$$

with $\eta_j := [0, \ldots, 0, 1, 0, \ldots, 0]^T$ as j-th unit vector, and

$\delta^{(j)}$ as (unknown) size of the outlier.

Consequently, the normal equations should have read

$$\begin{bmatrix} N & A^T P \eta_j \\ \eta_j^T PA & \eta_j^T P \eta_j \end{bmatrix} \begin{bmatrix} \hat{\xi}^{(j)} \\ \hat{\delta}^{(j)} \end{bmatrix} = \begin{bmatrix} c \\ \eta_j^T P y \end{bmatrix} \qquad (9)$$

with the modified "solution space"

$$\hat{\xi}^{(j)} \in \{\hat{\xi} - N^- (A^T P \eta_j) \cdot \hat{\delta}^{(j)} \mid NN^- N = N\} \qquad (10)$$

and the *estimated size* of the outlier

$$\hat{\delta}^{(j)} = (\eta_j^T P \tilde{e})(\eta_j^T P Q_{\tilde{e}} P \eta_j)^{-1} \sim (\delta^{(j)}, \sigma_0^2 (\eta_j^T P Q_{\tilde{e}} P \eta_j)^{-1}) \qquad (11)$$

which turns out to be *uniformly unbiased* as long as

$$\eta_j^T P Q_{\tilde{e}} P \eta_j \neq 0. \qquad (12)$$

Note that $\tilde{e}$ represents the residual vectors (5) from the un-modified Gauss-Markov Model (1) which can be interpreted as a combination of model (8) with the constraint $\delta^{(j)} = 0$.

In most early applications, the observational weight matrix was assumed to be *diagonal*, namely

$$P := \mathrm{Diag}(p_1, \ldots, p_n), \qquad (13)$$

which leads to the particularly appealing formula

$$\hat{\delta}^{(j)} = \tilde{e}_j / r_j \sim (\delta^{(j)}, \sigma_0^2 (p_j r_j)^{-1}) \qquad (14)$$

with the "*redundancy number*"

$$r_j := \eta_j^T Q_{\tilde{e}} P \eta_j = (Q_{\tilde{e}} P)_{jj}. \qquad (15)$$

The name obviously refers to the fact that

$$r_1 + \ldots + r_n = tr(Q_{\tilde{e}} P) = n - q =: r \qquad (16)$$

yields the total redundancy of the model. Moreover, in this special case we have

$$r_j = (Q_{\tilde{e}} P)_{jj} = (P^{1/2} Q_{\tilde{e}} P^{1/2})_{jj} \qquad (17)$$

which now is the j-th diagonal element of a *symmetric idempotent matrix*. For such matrices we know that all diagonal elements lie between the maximum eigenvalue 1 and the minimum eigenvalue 0, thus

$$0 \le r_j \le 1 \quad \text{for all } j \in \{1, \ldots, n\}. \tag{18}$$

Note that this property is lost in the case of correlated observations as was pointed out by *J. Wang/Y. Chen (1994)* and *B. Schaffrin (1997)*. But then we loose the simple relation (14) between the estimated outlier size and the corresponding (un-modified) residual any way, and we have to return to formula (11).

If we look at the change of the corresponding residual, however, that is caused by the presence of an outlier according to (8), we obtain

$$(\delta\tilde{e})_j = \eta_j^T(\delta\tilde{e}) = \eta_j^T(Q_{\tilde{e}}P)\eta_j \cdot \delta^{(j)} = r_j \cdot \delta^{(j)} \tag{19}$$

from (5) after replacing y by $y - \eta_j\delta^{(j)}$. Apparently, this formula holds true *even for non-diagonal P*, thereby telling us that the total effect of the outlier on the corresponding residual is scaled by the factor r_j.

Therefore, when r_j becomes almost 1 (or even bigger) an outlier of the same size becomes more and more visible in the respective residual and, conversely, when r_j is too small (or even negative) the outlier may be hidden. A larger r_j thus should increase the possibility for us to study the effect of an outlier on that particular observation.

Unfortunately, this does *not generally* mean that it would become easier to detect this outlier whenever r_j is larger since, for the detection, we have to rely on the *estimated* outlier size rather than its true size. As a result, we can consider r_j as "measure of reliability" *only in the uncorrelated case* when relation (14) holds. In this case, namely, the non-centrality parameter for the alternative hypothesis, i.e. "outlier in the j-th observation",

$$H_{a_j} : \delta^{(j)} \ne 0 \tag{20}$$

becomes

$$2\vartheta_j = (p_j r_j) \cdot (\delta^{(j)})^2 / \sigma_0^2, \tag{21}$$

and this parameter has to surpass a *certain threshold*, depending on the error probability α_0 and the chosen power of the test $1 - \beta_0$, in order to become *separable* from the null hypothesis, i.e. "no outlier in the j-th observation",

$$H_{0_j} : \delta^{(j)} = 0. \tag{22}$$

We take it from (21) that a larger r_j helps in this respect; thus, the redundancy numbers r_j may well serve as indicators for the "*inner reliability*".

In the general case of *correlated observations*, however, we follow *B. Schaffrin (1997)* and apply the "*normalized reliability numbers*" instead which are defined as the ratio

$$\bar{\bar{r}}_j := (\eta_j^T P Q_{\tilde{e}} P \eta_j)(\eta_j^T P \eta_j)^{-1}, \quad 0 \le \bar{\bar{r}}_j \le 1. \tag{23}$$

These are true generalizations as, for diagonal p, we end up with the original r_j again. According to formula (11), $\bar{\bar{r}}_j$ scales the influence of $\tilde{e}_j$ on the estimated size of the outlier $\hat{\delta}^{(j)}$.

In addition, let us define the "*outer reliability*" by the size of the effect that the maximum non-detectible outlier would have on the estimated parameters, measured in a properly weighted norm. In analogy to formula (10), we readily obtain

$$\left\| \hat{\xi} - \hat{\xi}_{(j)} \right\|_N = (\eta_j^T P A N^- A^T P \eta_j)^{1/2} \cdot \delta_0^{(j)} = \left[\eta_j^T (P - PQ_{\tilde{e}}P)\eta_j\right]^{1/2} \cdot \delta_0^{(j)} = \left[(\eta_j^T P \eta_j)(1 - \bar{\bar{r}}_j)\right]^{1/2} \cdot \delta_0^{(j)}, \tag{24}$$

in general, and in the case of *uncorrelated observations*

$$\left\| \hat{\xi} - \hat{\xi}_{(j)} \right\|_N = \left[p_j(1 - r_j)\right]^{1/2} \cdot \delta_0^{(j)} \tag{25}$$

independent of the chosen g-inverse N^-. Here $\delta_0^{(j)}$ denotes the maximum outlier that cannot be detected at the α_0-level with power $1 - \beta_0$ (where α_0 and β_0 are to be specified in advance).

In the following, we shall generalize the above theory to also cover *spatial processes* which are stochastic in nature - unlike the vector ξ - and (spatially) correlated.

2 RELIABILITY MEASURES FOR SIMPLE KRIGING

Here we begin with the definition of our model of a (stationary) *spatial process* at location $s \in S$

$$X(s) = \mu_0 + e_0(s), \qquad e_0(s) \sim (0, \sigma_X^2), \tag{26}$$

with the process mean μ_0 assumed to be *known*. The process has been sampled at the locations s_i (i=1, …, n) with the result

$$y := [y(s_1), \ldots, y(s_n)]^T = [X(s_1) + e(s_1), \ldots, X(s_n) + e(s_n)]^T =: X + e \tag{27}$$

where y is the n×1 observation vector,

X is the n×1 (unknown) vector of random effects,

e is the n×1 (unknown) vector of observational errors.

Furthermore, $e_0(s)$ denotes the random deviation of the process $X(s)$ from its mean, assumed to be uncorrelated with the vector e everywhere. Thus we have, for $\tau := [1, \ldots, 1]^T$ as n×1 "summation vector",

$$\underset{n\times 1}{e_0} := X - \tau \cdot \mu_0 \sim (0, \Sigma_X), \ \underset{n\times 1}{e} \sim (0, \sigma_e^2 P^{-1}), \ C\{e, e_0\} = 0, \tag{28}$$

where

$$\underset{n\times n}{\Sigma_X} = \begin{bmatrix} \sigma_x^2 & C\{X(s_1), X(s_2)\} & \cdots & C\{X(s_1), X(s_n)\} \\ & \sigma_x^2 & \cdots & C\{X(s_2), X(s_n)\} \\ (symm.) & & \ddots & \vdots \\ & & & \sigma_x^2 \end{bmatrix} \tag{29}$$

Following *N. A. C. Cressie (1993)* or *C. R. Rao/H. Toutenburg (1995)*, e.g., the *Simple Kriging* solution represents the Best inhomogeneously LInear Prediction (inhom BLIP) of X and can be generated by weighted least-squares. The solution is obtained, in the sample points, as

$$\tilde{X} = \tau \cdot \mu_0 + \Sigma_X (\sigma_e^2 P^{-1} + \Sigma_X)^{-1} (y - \tau \cdot \mu_0) \tag{30}$$

and can be shown to be (weakly) unbiased. Thus the *mean square prediction error matrix* can be computed via

$$MSPE\{\tilde{X}\} = D\{\tilde{X} - X\} = \Sigma_X - \Sigma_X (\sigma_e^2 P^{-1} + \Sigma_X)^{-1} \Sigma_X = \sigma_e^2 (P + \sigma_e^2 \Sigma_X^{-1})^{-1}. \tag{31}$$

Furthermore, the *two residual vectors* become

$$\tilde{e} = y - \tilde{X} = \sigma_e^2 P^{-1} (\sigma_e^2 P^{-1} + \Sigma_X)^{-1} (y - \tau \cdot \mu_0) = \sigma_e^2 (P + \sigma_e^2 \Sigma_X^{-1})^{-1} \Sigma_X^{-1} (y - \tau \cdot \mu_0) = \Sigma_{\tilde{e}} (\sigma_e^2 P^{-1})^{-1} (y - \tau \cdot \mu_0) \tag{32}$$

and

$$\tilde{e}_0 = \tau \cdot \mu_0 - \tilde{X} = -\Sigma_X (\sigma_e^2 P^{-1})^{-1} \cdot \tilde{e} = -\Sigma_{\tilde{e}_0} \Sigma_X^{-1} (y - \tau \cdot \mu_0) \tag{33}$$

with their *dispersion matrices*

$$D\{\tilde{e}\} = (\sigma_e^2 P^{-1})(\sigma_e^2 P^{-1} + \Sigma_X)^{-1} (\sigma_e^2 P^{-1}) = \sigma_e^2 P^{-1} - \sigma_e^2 (P + \sigma_e^2 \Sigma_X^{-1})^{-1} =: \Sigma_{\tilde{e}}, \tag{34}$$

$$D\{\tilde{e}_0\} = \Sigma_X (\sigma_e^2 P^{-1} + \Sigma_X)^{-1} \Sigma_X = \Sigma_X - \sigma_e^2 (P + \sigma_e^2 \Sigma_X^{-1})^{-1} =: \Sigma_{\tilde{e}_0}, \tag{35}$$

and the *covariance matrix*

$$C\{\tilde{e}, \tilde{e}_0\} = -(\sigma_e^2 P^{-1})(\sigma_e^2 P^{-1} + \Sigma_X)^{-1} \Sigma_X = -\sigma_e^2 (P + \sigma_e^2 \Sigma_X^{-1})^{-1} =: \Sigma_{\tilde{e}, \tilde{e}_0}. \tag{36}$$

A possible *variance component estimate* can be obtained through

$$\hat{\sigma}_e^2 = (\tilde{e}^T P \tilde{e}) \cdot (n - \tilde{e}_0^T \Sigma_X^{-1} \tilde{e}_0)^{-1}, \tag{37}$$

but may not turn out to be unbiased.

Now, if we assume an outlier in the j-th sample point our modified model will *either* read

$$y = X + \eta_j \delta^{(j)} + e\,, \quad e \sim (0, \sigma_e^2 P^{-1})\,, \quad X \sim (\tau \cdot \mu_0, \Sigma_X)\,, \tag{38}$$

when the outlier is attributed to faulty observations, *or*

$$y = X + e,\ \tau \cdot \mu_0 = X + \eta_k \delta^{(k)} + e_0\,, \quad \begin{bmatrix} e \\ e_0 \end{bmatrix} \sim (\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_e^2 P^{-1} & 0 \\ 0 & \Sigma_X \end{bmatrix}), \tag{39}$$

when the prior information (mean values) is considered faulty.

The *first case* can be handled along similar lines as developed for the Gauss-Markov Model, leading to the *modified solution*

$$\tilde{X}^{(j)} = \tilde{X} - (P + \sigma_e^2 \Sigma_X^{-1})^{-1} (P \eta_j) \cdot \hat{\delta}^{(j)} \tag{40}$$

with the *estimated size* of the outlier

$$\hat{\delta}^{(j)} = (\eta_j^T P \tilde{e}) \sigma_e^2 (\eta_j^T P \Sigma_{\tilde{e}} P \eta_j)^{-1} \sim (\delta^{(\mathrm{j})}, \sigma_e^2 (\eta_j^T P \Sigma_{\tilde{e}} P \eta_j)^{-1}) \tag{41}$$

which, for a *diagonal weight matrix P*, reduces to

$$\hat{\delta}^{(j)} = \tilde{e}_j / \bar{r}_j \sim (\delta^{(\mathrm{j})}, \sigma_e^2 (p_j \bar{r}_j)^{-1}). \tag{42}$$

Here the "*reliability numbers*" $\bar{r}_j$ are defined by

$$\bar{r}_j := \eta_j^T \Sigma_{\tilde{e}} P \eta_j / \sigma_e^2 = (\Sigma_{\tilde{e}} P)_{jj} / \sigma_e^2, \quad 0 < \bar{r}_j \le 1, \tag{43}$$

and may serve to indicate the "inner reliability". In contrast, the "*outer reliability*" would be quantified by

$$\left\| \tilde{X} - \tilde{X}_{(j)} \right\|_{P + \sigma_e^2 \Sigma_X^{-1}} = [\eta_j^T P (P + \sigma_e^2 \Sigma_X^{-1})^{-1} P \eta_j]^{1/2} \cdot \delta_0^{(j)}, \tag{44}$$

in general, and in the case of *uncorrelated observations* by

$$\left\| \tilde{X} - \tilde{X}_{(j)} \right\|_{P + \sigma_e^2 \Sigma_X^{-1}} = [p_j (1 - \bar{r}_j)]^{1/2} \cdot \delta_0^{(j)} \tag{45}$$

where $\delta_0^{(j)}$ again denotes the maximum non-detectible outlier for given α_0 and $1 - \beta_0$ (as explained in section 1).

For *non-diagonal P*, in the above formulas $\bar{r}_j$ would have to be replaced by

$$\bar{\bar{r}}_j := \sigma_e^2 \cdot (\eta_j^T K^{-1} \eta_j)(\eta_j^T P \eta_j)^{-1} \tag{46}$$

for

$$K := \sigma_e^4 (P \Sigma_{\tilde{e}} P)^{-1} = \Sigma_X \Sigma_{\tilde{e}_0}^{-1} \Sigma_X = \sigma_e^2 P^{-1} + \Sigma_X\,. \tag{47}$$

In the *second case*, we obtain the modified solution

$$\tilde{X}^{(k)} = \tilde{X} - (P + \sigma_e^2 \Sigma_X^{-1})^{-1} (\Sigma_X^{-1} \eta_k \sigma_e^2) \cdot \hat{\delta}^{(k)} \tag{48}$$

with the *estimated size* of the outlier

$$\hat{\delta}^{(k)} = (\eta_k^T \Sigma_X^{-1} \tilde{e}_0)(\eta_k^T K^{-1} \eta_k)^{-1} \sim (\delta^{(\mathrm{k})}, (\eta_k^T K^{-1} \eta_k)^{-1})\,. \tag{49}$$

The "*normalized reliability numbers*", in analogy to *B. Schaffrin (1997)*, thus become

$$\bar{\bar{r}}_k := (\eta_k^T K^{-1} \eta_k)(\eta_k^T \Sigma_X^{-1} \eta_k)^{-1}\,, \quad 0 < \bar{\bar{r}}_k \le 1, \tag{50}$$

to indicate "inner reliability" while the "*outer reliability*" is given by

$$\left\| \tilde{X} - \tilde{X}_{(k)} \right\|_{P + \sigma_e^2 \Sigma_X^{-1}} = [\sigma_e^2 (\eta_k^T \Sigma_X^{-1} \eta_k)(1 - \bar{\bar{r}}_k)]^{1/2} \cdot \delta_0^{(k)} \tag{51}$$

with $\delta_0^{(k)}$ denoting the maximum non-detectible outlier, now assumed to affect the mean value of the k-th sample point, i.e. $E\{X(s_k)\} = \mu_0 + \delta_0^{(k)}$, without being noticed. Note that, because of (33), we have the *duality* between $\hat{\delta}^{(k)}$ and $-\hat{\delta}^{(j)}|_{j=k}$ with (46) and (50) indicating where to locate the outlier easier when it occurred.

3 RELIABILITY MEASURES FOR ORDINARY KRIGING

In this section we shall derive similar formulas for a spatial process with *constant, but unknown* mean μ. The model underlying Ordinary Kriging, in compact form, reads

$$y = X + e\,,\ \tau \cdot \mu = X + e_0\,,\ \begin{bmatrix} e \\ e_0 \end{bmatrix} \sim (\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_e^2 P^{-1} & 0 \\ 0 & \Sigma_X \end{bmatrix}), \tag{52}$$

where the (known) value μ_0 has been replaced by the *unknown* μ.

With reference to *N. A. C. Cressie (1993)*, the *Ordinary Kriging* solution represents the Best homogeneously Linear (weakly) Unbiased Prediction (homBLUP) of X and can be obtained from the new model (52) by weighted least-squares, leading to the normal equations in *dual system* form

$$\begin{bmatrix} K & \tau \\ \tau^T & 0 \end{bmatrix} \begin{bmatrix} \overline{\overline{g}} \\ \hat{\mu} \end{bmatrix} = \begin{bmatrix} y \\ 0 \end{bmatrix}, \tag{53}$$

or to the *primal system*

$$\begin{bmatrix} K & -\tau \\ -\tau^T & 0 \end{bmatrix} \begin{bmatrix} \psi \\ \overline{\overline{\nu}}^T \end{bmatrix} = \begin{bmatrix} \Sigma_X \\ -\tau^T \end{bmatrix}, \tag{54}$$

with

$$\tilde{\tilde{X}} = \psi^T \cdot y = \tau \cdot \hat{\mu} + \Sigma_X \cdot \overline{\overline{g}} = \tau \cdot \hat{\mu} + \Sigma_X K^{-1}(y - \tau \cdot \hat{\mu}) \tag{55}$$

and

$$\hat{\mu} = (\tau^T K^{-1} y)(\tau^T K^{-1} \tau)^{-1}. \tag{56}$$

The *mean square prediction error matrix* results in

$$\underline{\mathrm{MSPE}}\{\tilde{\tilde{X}}\} = \Sigma_X - \psi^T \Sigma_X + \overline{\overline{\nu}} \cdot \tau^T = (\Sigma_X - \Sigma_X K^{-1} \Sigma_X) + \overline{\overline{\nu}} \cdot \tau^T (I_n - K^{-1} \Sigma_X) \tag{57}$$

with

$$\overline{\overline{\nu}} = (I_n - \Sigma_X K^{-1}) \tau (\tau^T K^{-1} \tau)^{-1} = (\sigma_e^2 P^{-1}) K^{-1} \tau (\tau^T K^{-1} \tau)^{-1}, \tag{58}$$

$$D\{\hat{\mu}\} = (\tau^T K^{-1} \tau)^{-1}, \quad C\{\tilde{\tilde{X}} - X, \hat{\mu}\} = -\overline{\overline{\nu}}. \tag{59}$$

Furthermore, we obtain the *two residual vectors* as

$$\tilde{\tilde{e}} = y - \tilde{\tilde{X}} = (\sigma_e^2 P^{-1}) K^{-1} (y - \tau \cdot \hat{\mu}) = \Sigma_{\tilde{\tilde{e}}} (\sigma_e^2 P^{-1})^{-1} y \tag{60}$$

and

$$\tilde{\tilde{e}}_0 = \tau\hat{\mu} - \tilde{\tilde{X}} = -\Sigma_X K^{-1} (y - \tau \cdot \hat{\mu}) = -\Sigma_X (\sigma_e^2 P^{-1})^{-1} \tilde{\tilde{e}} = -\Sigma_{\tilde{\tilde{e}}_0} \Sigma_X^{-1} y \tag{61}$$

with their *dispersion matrices*

$$D\{\tilde{\tilde{e}}\} = (\sigma_e^2 P^{-1})[K^{-1} - K^{-1}\tau(\tau^T K^{-1}\tau)^{-1}\tau^T K^{-1}](\sigma_e^2 P^{-1}) =: \Sigma_{\tilde{\tilde{e}}}, \tag{62}$$

$$D\{\tilde{\tilde{e}}_0\} = \Sigma_X [K^{-1} - K^{-1}\tau(\tau^T K^{-1}\tau)^{-1}\tau^T K^{-1}]\Sigma_X =: \Sigma_{\tilde{\tilde{e}}_0}, \tag{63}$$

and the *covariance matrix*

$$C\{\tilde{\tilde{e}},\tilde{\tilde{e}}_0\} = -(\sigma_e^2 P^{-1})[K^{-1} - K^{-1}\tau(\tau^T K^{-1}\tau)^{-1}\tau^T K^{-1}]\Sigma_X =: \Sigma_{\tilde{\tilde{e}},\tilde{\tilde{e}}_0}. \quad (64)$$

A possible, though not necessarily unbiased, *estimate of the variance component* would now be

$$\hat{\sigma}_e^2 = (\tilde{\tilde{e}}^T P\tilde{\tilde{e}})\cdot[(n-1) - \tilde{\tilde{e}}_0^T \Sigma_X^{-1}\tilde{\tilde{e}}_0]^{-1}. \quad (65)$$

In the following let us first assume an *outlier in th j-th observation*, leading to the modified model

$$y = X + \eta_j\delta^{(j)} + e, \;\; \tau\cdot\mu = X + e_0, \;\; \begin{bmatrix} e \\ e_0 \end{bmatrix} \sim \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_e^2 P^{-1} & 0 \\ 0 & \Sigma_X \end{bmatrix}\right). \quad (66)$$

The new reduced normal equations, after eliminating X, would then read

$$\begin{bmatrix} \tau^T K^{-1}\tau & \tau^T K^{-1}\eta_j \\ \eta_j^T K^{-1}\tau & \eta_j^T K^{-1}\eta_j \end{bmatrix}\begin{bmatrix} \hat{\mu}^{(j)} \\ \hat{\tilde{\delta}}^{(j)} \end{bmatrix} = \begin{bmatrix} \tau^T K^{-1} y \\ \eta_j^T K^{-1} y \end{bmatrix} \quad (67)$$

from which we obtain

$$\hat{\tilde{\delta}}^{(j)} = (\eta_j^T P\tilde{\tilde{e}})\{\sigma_e^2\eta_j^T[K^{-1} - K^{-1}\tau(\tau^T K^{-1}\tau)^{-1}\tau^T K^{-1}]\eta_j\}^{-1} \sim (\delta^{(j)}, \sigma_e^2(\eta_j^T P\Sigma_{\tilde{\tilde{e}}} P\eta_j)^{-1}), \quad (68)$$

$$\hat{\mu}^{(j)} = (\tau^T K^{-1}\tau)^{-1}\tau^T K^{-1}(y - \eta_j\hat{\tilde{\delta}}^{(j)}), \quad (69)$$

$$\tilde{\tilde{e}}_0^{(j)} = -\Sigma_X K^{-1}(y - \tau\cdot\hat{\mu}^{(j)} - \eta_j\hat{\tilde{\delta}}^{(j)}) = -\Sigma_{\tilde{\tilde{e}}_0}\Sigma_X^{-1}(y - \eta_j\hat{\tilde{\delta}}^{(j)}), \quad (70)$$

$$\tilde{\tilde{X}} - \tilde{\tilde{X}}^{(j)} = (\tilde{\tilde{e}}_0^{(j)} - \tilde{\tilde{e}}_0) + \tau(\hat{\mu} - \hat{\mu}^{(j)}) = (P^{-1} - \sigma_e^{-2}\Sigma_{\tilde{\tilde{e}}})P\eta_j\cdot\hat{\tilde{\delta}}^{(j)}. \quad (71)$$

In view of (68), an appropriate "*reliability number*" may be defined as

$$\bar{\bar{r}}_j := \sigma_e^2\cdot\eta_j^T[K^{-1} - K^{-1}\tau(\tau^T K^{-1}\tau)^{-1}\tau^T K^{-1}]\eta_j\cdot(\eta_j^T P\eta_j)^{-1} \quad (72)$$

which is the *percentage of the estimated outlier size* found in the corresponding residual $\tilde{\tilde{e}}_j$, thereby measuring the "*inner reliability*".

For the "*outer reliability*" we compute the weighted deviation

$$\left\|\tilde{\tilde{X}} - \tilde{\tilde{X}}_{(j)}\right\|_{(P^{-1}-\sigma_e^{-2}\Sigma_{\tilde{\tilde{e}}})^{-1}} = [\eta_j^T(P - \sigma_e^{-2}P\Sigma_{\tilde{\tilde{e}}}P)\eta_j]^{1/2}\cdot\delta_0^{(j)} \quad (73)$$

following (71), except that the estimated outlier size has been replaced by the maximum non-detectible outlier $\delta_0^{(j)}$ for given α_0 and $1-\beta_0$; see section 1 for more details. Using (62), this formula can be further simplified to

$$\left\|\tilde{\tilde{X}} - \tilde{\tilde{X}}_{(j)}\right\|_{(P^{-1}-\sigma_e^{-2}\Sigma_{\tilde{\tilde{e}}})^{-1}} = [(\eta_j^T P\eta_j)(1-\bar{\bar{r}}_j)]^{1/2}\cdot\delta_0^{(j)} \quad (74)$$

with the obvious reductions in case of a diagonal weight matrix P.

In the dual case of an *outlier in the mean for the j-th point*, i.e. $E\{X(s_j)\} \neq \mu$, we start from the modified model

$$y = X + e, \;\; \tau\cdot\mu = X + \eta_k\delta^{(k)} + e_0, \;\; \begin{bmatrix} e \\ e_0 \end{bmatrix} \sim \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_e^2 P^{-1} & 0 \\ 0 & \Sigma_X \end{bmatrix}\right), \quad (75)$$

but arrive at a similar set of normal equations (after eliminating X first)

$$\begin{bmatrix} \tau^T K^{-1}\tau & -\tau^T K^{-1}\eta_k \\ -\eta_k^T K^{-1}\tau & \eta_k^T K^{-1}\eta_k \end{bmatrix}\begin{bmatrix} \hat{\mu}^{(k)} \\ \hat{\tilde{\delta}}^{(k)} \end{bmatrix} = \begin{bmatrix} \tau^T K^{-1} y \\ -\eta_k^T K^{-1} y \end{bmatrix} \quad (76)$$

from which we obtain the immediate correspondences

$$\hat{\tilde{\delta}}^{(k)} = -\hat{\tilde{\delta}}^{(j)}\,|_{j=k} \quad \text{and} \quad \hat{\mu}^{(k)} = \mu^{(j)}\,|_{j=k}\ . \tag{77}$$

However, since we now inspect $\tilde{\tilde{e}}_0$ rather than $\tilde{\tilde{e}}$ to detect the outlier we have to use relation (61) to represent the estimated outlier size as

$$\hat{\tilde{\delta}}^{(k)} = \frac{-\eta_k^T(\sigma_e^2 P^{-1})^{-1}\tilde{\tilde{e}}}{\eta_k^T[K^{-1} - K^{-1}\tau(\tau^T K^{-1}\tau)^{-1}\tau^T K^{-1}]\eta_k} = (\eta_k^T\Sigma_X^{-1}\tilde{\tilde{e}}_0)(\eta_k^T\Sigma_X^{-1}\Sigma_{\tilde{\tilde{e}}_0}\Sigma_X^{-1}\eta_k)^{-1} \sim (\delta^{(k)}, (\eta_k^T\Sigma_X^{-1}\Sigma_{\tilde{\tilde{e}}_0}\Sigma_X^{-1}\eta_k)^{-1}), \tag{78}$$

thus providing us with the "*normalized reliability numbers*"

$$\bar{\bar{r}}_k := (\eta_k^T\Sigma_X^{-1}\Sigma_{\tilde{\tilde{e}}_0}\Sigma_X^{-1}\eta_k)(\eta_k^T\Sigma_X^{-1}\eta_k)^{-1} \tag{79}$$

which indicate the *percentage of the estimated outlier size* as seen in the corresponding residual $(\tilde{\tilde{e}}_0)_k$.

It is now straight-forward to derive the effect of the maximum non-detectible outlier $\delta_0^{(k)}$ in the weighted norm along similar lines as before, leading to

$$\left\| \tilde{\tilde{X}} - \tilde{\tilde{X}}_{(k)} \right\|_{(P^{-1}-\sigma_e^{-2}\Sigma_{\tilde{\tilde{e}}})^{-1}} = [\sigma_e^2(\eta_k^T\Sigma_X^{-1}\eta_k)(1-\bar{\bar{r}}_k)]^{1/2}\cdot\delta_0^{(k)} \tag{80}$$

4 CONCLUSIONS

We have derived "*normalized reliability numbers*" for spatial processes that indicate the "inner reliability", i.e. the potential to detect outliers *either* by inspecting the observational residuals *or* the residuals for the process mean in the sample points. By a simple relation, measures for the "outer reliability" can also be gained from them. It appears that, generally speaking, outliers can be detected more easily in the context of Simple Kriging than from Ordinary Kriging. This "loss in sensitivity" needs to be studied in the future.

REFERENCES

Baarda, W. (1968): A Testing Procedure for Use in Geodetic Networks, Neth. Geodetic Commission, Publ. on Geodesy, New Series Vol. 2, No. 5, Delft/NL.

Baarda, W. (1976): Reliability and precision of networks, VII. Intl. Course for Surveying Engineering, Darmstadt/Germany.

Cressie, N. A. C. (1993): Statistics for Spatial Data, Wiley: New York etc.

El-Hakim, S. F. (1982): Data snooping with weighted observations, Intl. Archive of Photogrammerty III-24, pp. 126-133.

Förstner, W. (1983): Reliability and discernibility of extended Gauss-Markov models, German Geodetic Comm. A-68, Munich.

Förstner, W. (1985): The reliability of block triangulation, Photogr. Engrg. & Rem. Sens. 51, pp. 1137-1149.

Koch, K. R. (1988): Parameter Estimation and Hypothesis Testing in Linear Models, Springer: New York/Berlin etc.

Pelzer, H. Some criteria for the accuracy and reliability of networks, German Geodetic Comm. B-252, Munich, pp. 55-67.

Rao, C. R. and H. Toutenburg (1995): Linear Models. Least-Squares and Alternatives, Springer: New York/Berlin etc.

Schaffrin, B. (1988): Precision, Reliability and Sensitivity (in Italian); in, F. Crosilla/L. Mussio (eds.), Progrettazione e Ottimizzazione del Rilievo Topografico e Fotogrammetrico de Controllo, Intl. Center of Mechanical Sci, (CISM), Udine/Italy, pp. 29-59.

Schaffrin, B. (1997): Reliability measures for correlated observations, J. of Surveying Engrg. 123, No. 3, pp. 126-133.

Wang, J. and Y. Chen (1994): On the reliability measure of observations, Acta Geodaet. et Cartograph. Sinica, English Edition, pp. 42-51.