ACTA ET COMMENTATIONES UNIVERSITATIS TARTUENSIS DE MATHEMATICA
Volume 16, Number 1, 2012
Available online at www.math.ut.ee/acta/

On the reliability of errors-in-variables models

Burkhard Schaffrin and Sibel Uzun

Abstract. Reliability has been quantified in a simple Gauss–Markov model (GMM) by Baarda (1968) for the application to geodetic networks as the potential to detect outliers – with a specified significance and power – by testing the least-squares residuals for their zero expectation property after an adjustment assuming "no outliers". It was shown that, under homoscedastic conditions, the so-called "redundancy numbers" could very well serve as indicators for the "local reliability" of an (individual) observation. In contrast, the maximum effect of any undetectible outlier on the estimated parameters would indicate "global reliability".

This concept had been extended successfully to the case of correlated observations by Schaffrin (1997) quite a while ago. However, no attempt has been made so far to extend Baarda's results to the (homoscedastic) errors-in-variables (EIV) model for which Golub and van Loan (1980) had found their – now famous – algorithm to generate the total least-squares (TLS) solution, together with all the residuals. More recently, this algorithm has been generalized by Schaffrin and Wieser (2008) to the case where a truly – not just elementwise – weighted TLS solution can be computed when the covariance matrix has the structure of a Kronecker–Zehfuss product.

Here, an attempt will be made to define reliability measures within such an EIV-model, in analogy to Baarda's original approach.

Introduction

In geodetic science, since Baarda's (1968) seminal report, reliability analysis in the frame work of a Gauss–Markov model is firmly based on

- the potential to detect outliers in any given observation ("inner reliability"), and
- the maximum effect of any undetected outlier on the estimated parameter vector ("outer reliability").

Received November 19, 2011.
2010 *Mathematics Subject Classification.* 15A09,15A52, 62F35.
Key words and phrases. Reliability, outlier detection, redundancy numbers, correlated observations, errors-in-variables model, total least-squares.

Originally formulated for uncorrelated observations only, Baarda's concept has been generalized since by Schaffrin (1997) to also cover the case of *correlated observations*.

The case of an EIV-model with outliers, however, has only recently attracted some attention when Schaffrin (2011) analyzed this situation for algorithms used in mobile mapping applications. In particular, hypothesis tests were developed to detect outliers in the elements of the coefficient matrix, should they occur occasionally and not massively there.

After this initial step, Baarda's quest for *reliability measures* in such EIV-models may also find an answer that obviously depends on the willingness to accept errors of the first or second kind, respectively, when making decisions on the basis of such tests. After significance level $(1-\alpha)$ and power of the test $(1-\beta)$ have been fixed, it is mainly the *response answer* of the residuals to any single outlier that matters, a point stressed by Proszynski (2010) already; see also the overview by Chatterjee and Hadi (1988) for models of type Gauss–Markov, with regard to their sensitivity to model changes.

It is this *response answer* that will now be studied for *EIV-models* when the covariance structure for the random errors in the coefficient matrix can be expressed by a certain Kronecker–Zehfuss product in accordance with Schaffrin and Wieser (2008), resulting in reliability measures for both "inner" and "outer reliability."

In Chapter 1, Baarda's approach to reliability measures for Gauss–Markov models will be reviewed both for the cases of the uncorrelated and correlated observations. In Chapter 2, the weighted total-least squares solution (WTLSS) will be presented before, in Chapter 3, an attempt will be made to introduce reliability measures when a single outlier occurs either in the observation vector itself or in the coefficient matrix. Finally, Chapter 4 will provide conclusions and an outlook on further research.

1. A review of reliability analysis for Gauss–Markov models

Let us assume the following Gauss–Markov model (GMM)

$$y = \underset{n\times m}{A}\,\xi + e, \quad e \sim \mathcal{N}(0, \sigma_0^2 Q), \tag{1.1}$$

where

y	denotes the $n \times 1$ observation vector,
ξ	the (unknown) $m \times 1$ parameter vector,
A	the $n \times m$ coefficient matrix with $n > m = \text{rk}A$,
e	the (unknown) normally distributed $n \times 1$ random error vector,
σ_0^2	the (unknown) variance component, and
$Q = P^{-1}$	the $n \times n$ symmetric positive-definite cofactor matrix so that
P	becomes the $n \times n$ symmetric positive-definite weight matrix.

The notation (1.1) indicates that the observations are here assumed to be *unbiased* and, thus, free of outliers due to the *"expectation"* of e being zero $E\{e\} = 0$. Hence, the *"dispersion matrix"* coincides with the mean squared error matrix of e in the sense:

$$\underline{MSE}\{e\} = D\{e\} + E\{e\} \cdot E\{e\}^T = D\{e\} = E\{ee^T\} = \sigma_0^2 Q. \tag{1.2}$$

Consequently, the least-squares solution (LESS) of ξ represents the best linear uniformly unbiased estimate (BLUUE) of ξ, which may be obtained from the *"normal equations"*

$$N\hat{\xi} = c \quad \text{for} \quad [N,\, c] := A^T P[A, y] \tag{1.3}$$

as

$$\hat{\xi} = N^{-1}c \sim \mathcal{N}(\xi, \sigma_0^2 N^{-1}). \tag{1.4}$$

In addition, the best linear prediction (BLIP) of the vector e results in the *"residual vector"*

$$\tilde{e} := y - A\hat{\xi} = (Q_{\tilde{e}}P)y \sim \mathcal{N}(0, \sigma_0^2 Q_{\tilde{e}}), \tag{1.5a}$$

with

$$Q_{\tilde{e}} := P^{-1} - AN^{-1}A^T \tag{1.5b}$$

as its cofactor matrix. Moreover, the (weighted) sum of squared residuals (SSR) can be computed via

$$\Omega := \tilde{e}^T P \tilde{e} = y^T P y - c^T \hat{\xi} \sim \sigma_0^2 \cdot \chi^2(n-m), \tag{1.6a}$$

resulting in the best invariant quadratic uniformly unbiased estimate (BIQUUE) of the variance component

$$\hat{\sigma}_0^2 = \Omega/(n-m) \sim (\sigma_0^2, 2(\sigma_0^2)^2/(n-m)). \tag{1.6b}$$

All the above results can be found in many textbooks on the subject matter such as that by Rao and Toutenburg (1995), for instance.

In case of a single outlier in one of the observations, however, the model (1.1) needs to be modified into

$$y = A\xi + \eta_j \xi_0^{(j)} + e, \quad e \sim \mathcal{N}(0, \sigma_0^2 Q), \tag{1.7}$$

where

η_j is the j-th $n \times 1$ unit vector with $\text{rk}[A, \eta_j] = m + 1$, and
$\xi_0^{(j)}$ is an additional unknown parameter.

The rank condition ensures that any suspected outlier in the j-th observation will not easily be smeared over all or some of the other observations. Obviously, for every $j \in \{1, \ldots, n\}$ there will be a different modified model that needs to be considered, for all of which the model (1.1) can be interpreted as *"constrained model,"* with a vanishing outlier

$$\xi_0^{(j)} = 0. \tag{1.8}$$

Later on, this constraint may serve as the *null hypothesis* which will be tested against its alternative, namely that there is actually one outlier that affects the j-th observation. The smaller the outlier that can still be detected by such hypothesis tests, the better the (local) *reliability* according to Baarda (1968). So, let us derive update formulas for all the above formulas when such an outlier may have occurred.

First, the *"extended normal equations"* will read

$$\begin{bmatrix} N & A^T P\eta_j \\ \eta_j^T PA & \eta_j^T P\eta_j \end{bmatrix} \begin{bmatrix} \hat{\xi}^{(j)} \\ \hat{\xi}_0^{(j)} \end{bmatrix} = \begin{bmatrix} c \\ \eta_j^T Py \end{bmatrix}, \tag{1.9}$$

resulting in the updated estimate of the original parameter vector

$$\hat{\xi}^{(j)} = \hat{\xi} - N^{-1}A^T P\eta_j \cdot \hat{\xi}_0^{(j)}, \tag{1.10}$$

where

$$\hat{\xi}_0^{(j)} = [\eta_j^T (PQ_{\tilde{e}}P)\eta_j]^{-1} \cdot [\eta_j^T (PQ_{\tilde{e}}P)y] \tag{1.11a}$$

$$= \eta_j^T (PQ_{\tilde{e}}P)\eta_j]^{-1} \cdot [\eta_j^T P\tilde{e}] \tag{1.11b}$$

denotes the estimated size of an outlier that may have affected the j-th observation (and only this one). It will be an unbiased estimate with

$$\hat{\xi}_0^{(j)} \sim \mathcal{N}(\xi_0^{(j)}, \sigma_0^2[\eta_j^T (PQ_{\tilde{e}}P)\eta_j]^{-1}) \tag{1.11c}$$

under the model (1.7), leading to the updated residual vector

$$\tilde{e}^{(j)} := y - A\hat{\xi}^{(j)} - \eta_j \cdot \hat{\xi}_0^{(j)} = \tilde{e} - (Q_{\tilde{e}}P)\eta_j \cdot \hat{\xi}_0^{(j)}, \tag{1.12a}$$

with

$$\tilde{e}^{(j)} \sim \mathcal{N}(0, \sigma_0^2[Q_{\tilde{e}} + Q_{\tilde{e}}P\eta_j(\eta_j^T PQ_{\tilde{e}}P\eta_j)^{-1}\eta_j^T PQ_{\tilde{e}}]). \tag{1.12b}$$

Secondly, the (weighted) SSR needs to be updated, too, giving us

$$\Omega_j := (\tilde{e}^{(j)})^T P\tilde{e}^{(j)} = \Omega - R_j \sim \sigma_0^2 \cdot \chi^2(n - m - 1), \tag{1.13a}$$

with

$$R_j := (\hat{\xi}_0^{(j)})^2 \cdot [\eta_j^T (PQ_{\tilde{e}}P)\eta_j] \sim \sigma_0^2 \cdot \chi'^2(1;\, 2\vartheta_j) \tag{1.13b}$$

and the non-centrality parameter

$$2\vartheta_j := (\xi_0^{(j)})^2 \cdot [\eta_j^T (PQ_{\tilde{e}}P)\eta_j]/\sigma_0^2 = E\{R_j/\sigma_0^2\}. \tag{1.13c}$$

The new (modified) estimate of the variance component will thus be obtained from

$$(\hat{\sigma}_0^2)_j := \Omega_j/(n - m - 1) \sim (\sigma_0^2, 2(\sigma_0^2)^2/(n - m - 1)). \tag{1.14}$$

Also note that, under model (1.7), the distribution of Ω will as well become *non-central*, changing (1.6a) into

$$\Omega_{new} \sim \sigma_0^2 \cdot \chi'^2(n - m; 2\vartheta_j). \tag{1.15}$$

Now, following Koch (1999, chapter 44), for instance, the constraint (1.8) may be tested as the j-th *null hypothesis*

$$H_0^{(j)} : \xi_0^{(j)} = 0 \quad \text{vs.} \quad H_a^{(j)} : \xi_0^{(j)} \neq 0 \tag{1.16}$$

as *"alternative"*, using the *test statistic*

$$T_j := R_j/(\hat{\sigma}_0^2)_j = R_j(n-m-1)/\Omega_j \sim F'(1, n-m-1; 2\vartheta_j), \tag{1.17}$$

which would be centrally F-distributed under $H_0^{(j)}$. So, after specifying the error probability α, a fractile T_a can be determined on this basis. Furthermore, after specifying the *power of the test*

$$1 - \beta = \int_{T_\alpha}^{\infty} f'(t; 2\vartheta_j)dt, \tag{1.18}$$

where $f'(t; 2\vartheta_j)$ denotes the probability density function (p.d.f.) of the non-central F-distribution, the non-centrality parameter can be computed numerically and compared with (1.13c), resulting in the formula

$$(\xi_0^{(j)})_{\min} = \sqrt{2\vartheta_j(n-m-1; \alpha, \beta) \cdot \sigma_0^2/[\eta_j^T(PQ_{\tilde{e}}P)\eta_j]} \tag{1.19}$$

for the *"minimum detectible outlier"* that can still be identified by hypothesis testing in the j-th observation with a significance level of $(1-\alpha)$ and a power of $(1-\beta)$. This quantity, beside the redundancy $n-m-1$, mainly depends on the factor

$$\sigma_0^2/[\eta_j^T(PQ_{\tilde{e}}P)\eta_j] = D\{\hat{\xi}_0^{(j)}\} = \overline{\overline{\sigma}}_j^2/\overline{\overline{r}}_j, \tag{1.20a}$$

with

$$\overline{\overline{\sigma}}_j^2 := \sigma_0^2/(\eta_j^T P \eta_j), \quad \text{and} \quad \overline{\overline{r}}_j := [\eta_j^T(PQ_{\tilde{e}}P)\eta_j]/(\eta_j^T P \eta_j). \tag{1.20b}$$

Now, following Schaffrin (1997), the number $\overline{\overline{r}}_j$ can be shown to fall into the interval $[0, 1)$ with a larger number being more desirable than a smaller one. Hence, $\overline{\overline{r}}_j$ may serve as indicator for the *"inner reliability"*, originally defined by Baarda (1968) for *diagonal matrices* $P = \text{Diag}(p_1, \ldots, p_n)$ as

$$r_j := \eta_j^T(Q_{\tilde{e}}P)\eta_j = p_j \cdot [\eta_j^T(Q_{\tilde{e}}P)\eta_j]/(\eta_j^T P \eta_j) = \overline{\overline{r}}_j, \tag{1.21a}$$

with

$$\sigma_j^2 = \sigma_0^2/p_j = \overline{\overline{\sigma}}_j^2 \tag{1.21b}$$

as variance of the j-th observation. Note that, in this case, r_j also defines the response of the estimated outlier $\hat{\xi}_0^{(j)}$ on the j-th residual in accordance with (1.11b) and (1.12a), namely via

$$\begin{aligned}\tilde{e}_j^{(j)} &:= \eta_j^T \tilde{e}^{(j)} = \eta_j^T \tilde{e} - \eta_j^T (Q_{\tilde{e}} P)\eta_j \cdot \hat{\xi}_0^{(j)} \\ &= \tilde{e}_j - r_j \cdot [(p_j \cdot \tilde{e}_j)/(p_j r_j)] = 0,\end{aligned} \tag{1.22a}$$

and thus

$$\tilde{e}_j = r_j \cdot \hat{\xi}_0^{(j)} \quad \text{or} \quad \hat{\xi}_0^{(j)} = \tilde{e}_j / r_j. \tag{1.22b}$$

In the general case, however, the response relationship is more complex as can be found out by translating (1.11b) into

$$\hat{\xi}_0^{(j)} = [(\eta_j^T P \tilde{e})/\eta_j^T P \eta_j)]/\overline{\overline{r}}_j. \tag{1.23}$$

The *"outer reliability"* may now be introduced by the effect that the *"maximum undetectible outlier"* may have on the estimated parameters following (1.10), namely

$$\begin{aligned}&\max_j \left[\left\| N^{-1} A^T P \eta_j \cdot \left(\xi_0^{(j)} \right)_{\max} \right\|_N^2 / \left(\xi_0^{(j)} \right)_{\max}^2 \right] \\ &= \max_j [\eta_j^T (P - P Q_{\tilde{e}} P)\eta_j] = \max_j [(\eta_j^T P \eta_j)(1 - \overline{\overline{r}}_j)].\end{aligned} \tag{1.24}$$

Of course, this number ought to be small in order to ensure that undetectible outliers have no major effects on the outcome in terms of estimated parameters.

2. The weighted total least-squares approach

In the following, let us assume that the elements of the coefficient matrix A are also observed and hence affected by random errors. This provision leads to the so-called *errors-in-variables (EIV) model*

$$y = \underbrace{(A - E_A)}_{n \times m} \xi + e \tag{2.1a}$$

$$\begin{bmatrix} e \\ e_A \end{bmatrix} \sim \mathcal{N}(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \sigma_0^2 \begin{bmatrix} Q & 0 \\ 0 & I_m \otimes Q \end{bmatrix}), \quad Q = P^{-1}, \tag{2.1b}$$

where

E_A	is the $n \times m$ matrix of additional random errors, and
$e_A := \text{vec} E_A$	the same in $nm \times 1$ vector form.

Here, $\otimes$ denotes the "Kronecker–Zehfuss product" of matrices, defined by

$$G \otimes H := [g_{ij} \cdot H] \quad \text{if} \quad G = [g_{ij}]. \tag{2.2}$$

The EIV-model can obviously be transformed into a nonlinear *Gauss–Helmert model (GHM)* by rewriting (2.1a) as

$$y = A\xi + [B_1|B_2]\begin{bmatrix} e \\ e_A \end{bmatrix}, \tag{2.3a}$$

$$B := [B_1|B_2] = \left[I_n| - \left(\xi^T \otimes I_n\right)\right] = B(\xi), \tag{2.3b}$$

which nicely shows how the $n \times n(m+1)$ matrix B depends on the unknown parameter vector ξ.

Now, the total least-squares (TLS) approach is characterized by minimizing the "total (weighted) SSR" subject to the EIV-model (2.1a-b), which can be achieved by making the equivalent Lagrange target function stationary, namely

$$\begin{aligned}\Phi\left(e, e_A, \xi, \lambda\right) := e^T P e + e_A^T \left(I_m \otimes P\right) e_A + \\ +2\lambda^T \left[y - e - \left(\xi^T \otimes I_n\right)\left(\text{vec}A - e_A\right)\right] = \text{stationary}.\end{aligned} \tag{2.4}$$

Then the necessary Euler–Lagrange conditions read:

$$\frac{1}{2}\frac{\partial\Phi}{\partial e} = P\tilde{e} - \hat{\lambda} \doteq 0 \tag{2.5a}$$

$$\frac{1}{2}\frac{\partial\Phi}{\partial e_A} = \left(I_m \otimes P\right)\tilde{e_A} + \left(\hat{\xi} \otimes I_n\right)\hat{\lambda} \doteq 0 \tag{2.5b}$$

$$\frac{1}{2}\frac{\partial\Phi}{\partial \xi} = -A^T\hat{\lambda} + \tilde{E_A^T}\hat{\lambda} \doteq 0 \tag{2.5c}$$

$$\frac{1}{2}\frac{\partial\Phi}{\partial \lambda} = y - \tilde{e} - \left(\hat{\xi}^T \otimes I_n\right)\left(\text{vec}A - \tilde{e_A}\right) = y - \tilde{e} - \left(A - \tilde{E_A}\right)\hat{\xi} \doteq 0 \tag{2.5d}$$

In addition, the sufficient condition for a *multivariate minimum* is also fulfilled, due to the fact that

$$\frac{1}{2}\frac{\partial^2\Phi}{\partial\begin{bmatrix} e \\ e_A \end{bmatrix}\partial\left[e^T, e_A^T\right]} = \begin{bmatrix} P & 0 \\ 0 & I_m \otimes P \end{bmatrix} \quad \text{is positive-definite.} \tag{2.6}$$

From (2.5a-b), it is obvious that both $\tilde{e}$ and $\tilde{e_A}$ can be represented in terms of $\hat{\lambda}$ via

$$\tilde{e} = P^{-1}\hat{\lambda} \quad \text{and} \quad \tilde{e_A} = -(\hat{\xi} \otimes P^{-1})\hat{\lambda}, \tag{2.7a}$$

respectively

$$\tilde{E_A} = -P^{-1}\hat{\lambda}\hat{\xi}^T = -\tilde{e}\hat{\xi}^T, \tag{2.7b}$$

which turns (2.5c-d) into

$$A^T\hat{\lambda} = \tilde{E_A^T}\hat{\lambda} = -\hat{\xi} \cdot \left(\hat{\lambda}^T P^{-1}\hat{\lambda}\right) \tag{2.8a}$$

and

$$P(y - A\hat{\xi}) = P\tilde{e} - P\tilde{E}_A\hat{\xi} = \hat{\lambda}(1 + \hat{\xi}^T\hat{\xi}) \tag{2.8b}$$

such that

$$\hat{\lambda} = P(y - A\hat{\xi})(1 + \hat{\xi}^T\hat{\xi})^{-1} = P\tilde{e}. \tag{2.8c}$$

Summarizing, the *new (nonlinear) normal equations* may be stated as

$$\begin{aligned} c - N\hat{\xi} &= A^T P \left(y - A\hat{\xi}\right) = A^T\hat{\lambda}\left(1 + \hat{\xi}^T\hat{\xi}\right) = -\hat{\xi}\cdot\left(\hat{\lambda}^T P^{-1}\hat{\lambda}\right)\left(1 + \hat{\xi}^T\hat{\xi}\right) \\ &= -\hat{\xi}\cdot\left(y - A\hat{\xi}\right)^T P \left(y - A\hat{\xi}\right)\left(1 + \hat{\xi}^T\hat{\xi}\right)^{-1} =: -\hat{\xi}\cdot\hat{\nu} \end{aligned} \tag{2.9a}$$

with

$$\hat{\nu}\left(1 + \hat{\xi}^T\hat{\xi}\right) = y^T P\left(y - A\hat{\xi}\right) - \hat{\xi}^T\left(c - N\hat{\xi}\right) = y^T P\left(y - A\hat{\xi}\right) + \left(\hat{\xi}^T\hat{\xi}\right)\cdot\hat{\nu} \tag{2.9b}$$

or, in compact form,

$$\hat{\nu} = y^T P y - c^T\hat{\xi} = \tilde{e}^T P\tilde{e} + \tilde{e_A}^T(I_m \otimes P)\tilde{e}_A =: TSSR. \tag{2.10a}$$

Thus, the TLS-solution $\hat{\xi}$ may be obtained by inverting

$$(N - \hat{\nu}I_m)\hat{\xi} = c \tag{2.10b}$$

with a modified matrix in each iteration, or by inverting

$$N\hat{\xi} = c + \hat{\xi}\cdot\hat{\nu} \tag{2.10c}$$

with a modified RHS vector in each iteration. Experience shows that alternating (2.10b) with (2.10a) converges faster than alternating (2.10c) with (2.10a), but requires more operations per iteration step.

Since $\hat{\nu}$ already represents the TSSR (total sum of squared residuals), a convenient estimate of the variance component is provided by

$$\hat{\sigma}_0^2 = \hat{\nu}/(n - m), \tag{2.11}$$

where $(n - m)$ denotes the *redundancy* in the EIV-model (2.1a-b). Note that (2.10a) and (2.10c) can be combined to

$$\begin{bmatrix} N & c \\ c^T & y^T P y \end{bmatrix}\begin{bmatrix} \hat{\xi} \\ -1 \end{bmatrix} = \begin{bmatrix} \hat{\xi} \\ -1 \end{bmatrix}\cdot\hat{\nu}, \tag{2.12}$$

which allows the interpretation of the TLS problem as *"eigenvalue problem"*, in accordance with Golub and van Loan (1980) where $\hat{\nu}$, as TSSR, represents the minimum eigenvalue.

3. Reliability analysis for errors-in-variables models

3.1. Outliers affecting the observation vector. In the following, a single outlier may have occurred in the observation vector y, but not in the observed coefficient matrix A. Thus the family of modified EIV-models would read:

$$y = \underset{n\times m}{A}\,\xi + \eta_j \cdot \xi_0^{(j)} + (e - \underset{n\times m}{E_A}\cdot\xi), \tag{3.1a}$$

$$\begin{bmatrix} e \\ e_A \end{bmatrix} \sim \mathcal{N}(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma_0^2 \begin{bmatrix} Q & 0 \\ 0 & I_m \otimes Q \end{bmatrix}), \tag{3.1b}$$

where, again,

η_j denotes the j-th $n \times 1$ unit vector, and
$\xi_0^{(j)}$ is the unknown size of the potential outlier.

After setting up the Lagrange target function as

$$\begin{aligned} \Phi\left(e, e_A, \xi, \xi_0^{(j)}, \lambda\right) &:= e^T Pe + e_A^T \left(I_m \otimes P\right) e_A \\ +2\lambda^T \left[y - e - \left(\xi^T \otimes I_n\right)(vecA - e_A) - \eta_j \xi_0^{(j)}\right] &= \text{stationary} \end{aligned} \tag{3.2}$$

in analogy to (2.4), a system of *nonlinear normal equations* similar to (2.10c) will be obtained, namely

$$\begin{bmatrix} N & A^T P\eta_j \\ \eta_j^T PA & \eta_j^T P\eta_j \end{bmatrix} \begin{bmatrix} \hat{\xi}^{(j)} \\ \hat{\xi}_0^{(j)} \end{bmatrix} = \begin{bmatrix} c + \hat{\xi}^{(j)} \cdot \hat{\nu}^{(j)} \\ \eta_j^T Py \end{bmatrix}, \tag{3.3a}$$

with

$$\hat{\nu}^{(j)} = y^T P\left(y - A\hat{\xi}^{(j)} - \eta_j \hat{\xi}_0^{(j)}\right) = y^T Py - c^T \hat{\xi}^{(j)} - \left(y^T P\eta_j\right) \cdot \hat{\xi}_0^{(j)} \tag{3.3b}$$

as $(\text{TSSR})_j$, modifying the original TSSR from (2.10a). An estimate of the j-th outlier size is readily derived from (3.3a) as

$$\begin{aligned} \hat{\xi}_0^{(j)} &= \frac{\eta_j^T \left[P - PA\left(N - \hat{\nu}^{(j)} I_m\right)^{-1} A^T P\right] y}{\eta_j^T \left[P - PA\left(N - \hat{\nu}^{(j)} I_m\right)^{-1} A^T P\right] \eta_j} \qquad (3.4a) \\ &= \left[\eta_j^T \left(PQ_{\tilde{e}}P\right)\eta_j\right]^{-1} \left[\eta_j^T \left(PQ_{\tilde{e}}P\right) y - \eta_j^T PAN^{-1} \hat{\xi}^{(j)} \hat{\nu}^{(j)}\right], \qquad (3.4b) \end{aligned}$$

and it may be conjectured that the *test statistic*

$$T_j := \frac{\hat{\nu} - \hat{\nu}^{(j)}}{\hat{\nu}^{(j)}/(n-m-1)} \sim F(1, n-m-1) \tag{3.5}$$

is centrally F-distributed under the j-th *null hypothesis*

$$H_0^{(j)} : \xi_0^{(j)} = 0 \quad \text{vs.} \quad H_a^{(j)} : \xi_0^{(j)} \neq 0 \tag{3.6}$$

as "alternative". On the other hand, using (2.8c)and (2.9a) in conjunction with (3.4b), leads to the expression

$$\hat{\xi}_0^{(j)} = \left[\eta_j^T \left(PQ_{\tilde{e}}P\right)\eta_j\right]^{-1}\eta_j^T P\left[\tilde{e}\left(1+\hat{\xi}^T\hat{\xi}\right)+AN^{-1}\left(\hat{\xi}\cdot\hat{\nu}-\hat{\xi}^{(j)}\hat{\nu}^{(j)}\right)\right] \tag{3.7}$$

for the estimated size of the outlier in the j-th observation y_j. Under the assumption that the difference $(\hat{\xi}\cdot\hat{\nu}-\hat{\xi}^{(j)}\hat{\nu}^{(j)})$ is *negligible* in this context, this means that the response of the outlier on the corresponding residual is only scaled by the factor $(1+\hat{\xi}^T\hat{\xi})$ when compared with the analogous response in the Gauss–Markov model.

We may, therefore, use the scaled quantities

$$\bar{\bar{r}}_j\left(1+\hat{\xi}^T\hat{\xi}\right)^{-1} = \left[\eta_j^T\left(PQ_{\tilde{e}}P\right)\eta_j\right]\left[\left(\eta_j^T P\eta_j\right)\cdot\left(1+\hat{\xi}^T\hat{\xi}\right)\right]^{-1} \tag{3.8a}$$

to indicate *"inner reliability"* for this case in an EIV-model. For similar reasons, the quantity describing the *"outer reliability"* in this case may be approximated by

$$\max_j\left[\left\|N^{-1}A^T P\eta_j\cdot\left(\xi_0^{(j)}\right)_{\max}+N^{-1}\left(\hat{\xi}\cdot\hat{\nu}-\hat{\xi}^{(j)}\cdot\hat{\nu}^{(j)}\right)\right\|_N^2 / \left(\xi_0^{(j)}\right)_{\max}^2\right]$$
$$\approx \max_j\left[\eta_j^T\left(PAN^{-1}A^T P\right)\eta_j\right] = \max_j\left[\left(\eta_j^T P\eta_j\right)\left(1-\bar{\bar{r}}_j\right)\right], \tag{3.8b}$$

which nicely coincides with (1.24) for the Gauss–Markov model.

3.2. Outliers affecting the observed coefficient matrix. Now, in contrast to the Gauss–Markov model, single outliers may have occurred in the observed coefficient matrix, say in its component

$$a_{jk} = \eta_j^T A\eta_k,$$

where

η_j is the j-th $n\times 1$ unit vector, and
η_k is the k-th $m\times 1$ unit vector.

So, the family of modified EIV-models reads:

$$y = \underbrace{(A-\eta_j\xi_0^{(jk)}\eta_k^T)}_{n\times m}\xi + (e-\underbrace{E_A}_{n\times m}\xi), \tag{3.9a}$$

$$\begin{bmatrix} e \\ e_A \end{bmatrix} \sim \mathcal{N}(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma_0^2\begin{bmatrix} Q & 0 \\ 0 & I_m\otimes Q \end{bmatrix}), \tag{3.9b}$$

with $\xi_0^{(jk)}$ as unknown size of the potential outlier. Again, the corresponding Lagrange target function is set up via

$$\Phi\left(e, e_A, \xi, \xi_0^{(jk)}, \lambda\right) := e^T P e + e_A^T \left(I_m \otimes P\right) e_A$$
$$+2\lambda^T \left[y - e - \left(\xi^T \otimes I_n\right)\left(vecA - e_A - (\eta_k \otimes \eta_j)\xi_0^{(jk)}\right)\right] = \text{stationary} \tag{3.10}$$

and, after a few steps, leads to the following system of *nonlinear normal equations*, similar to (3.3a), namely

$$\begin{bmatrix} N & A^T P \eta_j \\ \eta_j^T P A & \eta_j^T P \eta_j \end{bmatrix} \begin{bmatrix} \hat{\xi}^{(jk)} \\ -\hat{\xi}_0^{(jk)} \cdot \hat{\xi}_k^{(jk)} \end{bmatrix} = \begin{bmatrix} c + \hat{\xi}^{(jk)} \cdot \hat{\nu}^{(jk)} \\ \eta_j^T P y \end{bmatrix}, \tag{3.11a}$$

with

$$\hat{\nu}^{(jk)} = y^T P \left(y - A\hat{\xi}^{(jk)} + \eta_j(\hat{\xi}_0^{(jk)} \cdot \hat{\xi}_k^{(jk)})\right)$$
$$= y^T P y - c^T \hat{\xi}^{(jk)} + \left(y^T P \eta_j\right) \cdot (\hat{\xi}_0^{(jk)} \cdot \hat{\xi}_k^{(jk)}) \tag{3.11b}$$

as newly modified $(TSSR)_{jk}$, similar to (3-3b). Consequently, under the assumption

$$\hat{\xi}_k^{(jk)} := \eta_k^T \hat{\xi}^{(jk)} \neq 0, \tag{3.12a}$$

the estimated size of the outlier in this case can be expressed as

$$\hat{\xi}_0^{(jk)} = \frac{-\eta_j^T \left[P - PA\left(N - \hat{\nu}^{(jk)} I_m\right)^{-1} A^T P\right] y}{\eta_j^T \left[P - PA\left(N - \hat{\nu}^{(jk)} I_m\right)^{-1} A^T P\right] \eta_j \cdot (\eta_k^T \hat{\xi}^{(jk)})} \tag{3.12b}$$
$$= -\left[(\eta_k^T \hat{\xi}^{(jk)})\eta_j^T \left(PQ_{\tilde{e}}P\right)\eta_j\right]^{-1} \left[\eta_j^T \left(PQ_{\tilde{e}}P\right) y - \eta_j^T PAN^{-1}\hat{\xi}^{(jk)}\hat{\nu}^{(jk)}\right], \tag{3.12c}$$

leading to the conjecture that the *test statistic*

$$T_{jk} := \frac{\hat{\nu} - \hat{\nu}^{(jk)}}{\hat{\nu}^{(jk)}/(n-m-1)} \sim F(1, n-m-1) \tag{3.13}$$

is centrally F-distributed under the (jk)-th *null hypothesis*

$$H_0^{(jk)} : \xi_0^{(jk)} = 0 \quad \text{vs.} \quad H_a^{(jk)} : \xi_0^{(jk)} \neq 0 \tag{3.14}$$

as "alternative". A simple modification in analogy to (3.7) allows for the representation

$$\hat{\xi}_0^{(jk)} = -\left[(\eta_k^T \hat{\xi}^{(jk)})\eta_j^T \left(PQ_{\tilde{e}}P\right)\eta_j\right]^{-1}$$
$$\cdot \eta_k^T P \left[(y - A\hat{\xi}) + AN^{-1}\left(\hat{\xi} \cdot \hat{\nu} - \hat{\xi}^{(jk)}\hat{\nu}^{(jk)}\right)\right] \tag{3.15}$$

of the estimated outlier size in the j-th row and k-th column of the matrix A. Here, $(y - A\hat{\xi})$ may be rewritten as

$$y - A\hat{\xi} = \tilde{e}(1 + \hat{\xi}^T\hat{\xi}) = -(\hat{\xi}^T \otimes I_n)\tilde{e}_A \cdot (1 + \hat{\xi}^T\hat{\xi})/(\hat{\xi}^T\hat{\xi}), \tag{3.16}$$

following (2.8c) and (2.8b), thus leading to

$$\begin{aligned} \hat{\xi}_0^{(jk)} &= \left[(\eta_k^T \hat{\xi}^{(jk)})\eta_j^T \left(PQ_{\tilde{e}}P\right)\eta_j\right]^{-1} \\ &\quad \cdot \eta_j^T P \left[(\tilde{E}_A\hat{\xi} \cdot (1 + \hat{\xi}^T\hat{\xi})/(\hat{\xi}^T\hat{\xi})) - AN^{-1}\left(\hat{\xi} \cdot \hat{\nu} - \hat{\xi}^{(jk)}\hat{\nu}^{(jk)}\right)\right]. \end{aligned} \tag{3.17}$$

Again, assuming that the difference $(\hat{\xi} \cdot \hat{\nu} - \hat{\xi}^{(jk)}\hat{\nu}^{(jk)})$ is *negligible* in this context, (3.17) shows how the response of the outlier on the corresponding residual is to be scaled, in comparison to the previous case. For the *"inner reliability"*, therefore, the scaled quantities

$$\begin{aligned} &\left[\eta_j^T \left(PQ_{\tilde{e}}P\right)\eta_j\right]\left[\eta_j^T P\left(\hat{\xi}^T \otimes I_n\right)(\eta_k \otimes \eta_j)\right]^{-1} \\ &\quad \cdot \left(1 + \hat{\xi}^T\hat{\xi}\right)^{-1}\left[(\hat{\xi}^T\hat{\xi}) \cdot (\eta_k^T\hat{\xi}^{(jk)})\right] \\ &= \overline{\overline{r}}_j \cdot \left[\left(\hat{\xi}^T\hat{\xi}\right)/\left(1 + \hat{\xi}^T\hat{\xi}\right)\right] \cdot \left[\left(\eta_k^T\hat{\xi}^{(jk)}\right)/\eta_k^T\hat{\xi}\right] \end{aligned} \tag{3.18a}$$

may be applied as indicators for this case in an EIV-model. Similarly, the indicator for the *"outer reliability"* may be approximated by

$$\begin{aligned} &\max_j \left[\left\| N^{-1}A^T P\eta_j \cdot \left(\xi_0^{(jk)}\right)_{\max} \cdot \hat{\xi}_k^{(jk)} \right.\right. \\ &\qquad \left.\left. -\ N^{-1}\left(\hat{\xi} \cdot \hat{\nu} - \hat{\xi}^{(jk)} \cdot \hat{\nu}^{(jk)}\right)\right\|_N^2 / \left(\xi_0^{(jk)}\right)^2_{\max}\right] \\ &\approx \max_j \left[\eta_j^T \left(PAN^{-1}A^T P\right)\eta_j \left(\eta_k^T\hat{\xi}^{(jk)}\right)^2\right] \\ &= \max_j \left[\left(\eta_j^T P\eta_j\right)\left(1 - \overline{\overline{r}}_j\right)\left(\eta_k^T\hat{\xi}^{(jk)}\right)^2\right] \end{aligned} \tag{3.18b}$$

which nicely corresponds to (3.8b).

4. Conclusions and outlook

In this contribution, reliability measures have been studied for the errors-in-variables model, following the ideas of Baarda (1968). It turned out that both the "inner reliability" as well as the "outer reliability" can be measured by the same quantities that served this purpose in the Gauss–Markov model, after proper rescaling. This is a somewhat surprising result insofar as outliers for the coefficient matrix are not allowed in the Gauss–Markov model, at all, but may occur in the EIV-model.

These results have been achieved under the assumption that a certain term in the response relationship between outlier and corresponding residual is small and can be neglected. In a future study, the maximum effect of this neglected term will be investigated in worst case scenarios.

Acknowledgement. The present paper includes investigations that the first author was able to complete in December 2010 when visiting the Hong Kong Polytechnic University upon the invitation of Prof. Xiaoli Ding and Prof. H. Baki Iz. Their support is gratefully acknowledged.

References

Baarda, W. (1968), *A testing procedure for use in geodetic networks*, Neth. Geodetic Comm., Publications on Geodesy **9** (Vol. 2, No. 5), Delft, 97 pages.

Chatterjee, S., and Hadi, A. S. (1988), *Sensitivity Analysis in Linear Regression*, Wiley, New York.

Golub, G. H., and van Loan, C. F. (1980), *An analysis of the total least-squares problem*, SIAM J. Numer. Anal. **17** (1980), 883–893.

Koch, K. R. (1999), *Parameter Estimation and Hypothesis Testing in Linear Models*, 2nd. Edition, Springer, Berlin.

Proszynski, W. (2010), *Another approach to reliability measures for systems with correlated observations*, J. Geodesy **84** (2010), 547–556.

Rao, C. R., and Toutenburg, H. (1995), *Linear Models. Least Squares and Alternatives*, Springer, New York.

Schaffrin, B. (1997), *Reliability measures for correlated observations*, J. Surveying Engrg. **123** (1997), No. 3, 126–137.

Schaffrin, B., and Wieser, A. (2008), *On weighted total least-squares adjustment for linear regression*, J. Geodesy **82** (2008), 415–421.

Schaffrin, B. (2011) *Errors-in-variables for mobile mapping algorithms in the presence of outliers*, Intl. Archives of Photogrammetry, Cartography and Remote Sensing **22** (Kraków 2011), pp. 377–387.

Geodetic Science Program, School of Earth Sciences, The Ohio State University, Columbus/OH, USA

E-mail address: `schaffrin.1@.osu.edu`

E-mail address: `uzun.1@buckeyemail.osu.edu`