Data-snooping procedure applied to errors-in-variables models

ALI REZA AMIRI-SIMKOOEI[1,2] AND SHAHRAM JAZAERI[3]

1 Section of Geodesy, Department of Surveying Engineering, Faculty of Engineering, University of Isfahan, 81746-73441 Isfahan, Iran (ar_amiri@yahoo.com)

2 Chair Acoustics, Faculty of Aerospace Engineering, Delft University of Technology, Delft, The Netherlands

3 Department of Surveying and Geomatics Engineering, College of Engineering, University of Tehran, Tehran, Iran

Received: October 5, 2012; Revised: January 24, 2013; Accepted: March 22, 2013

ABSTRACT

The theory of Baarda's data snooping - normal and F tests respectively based on the known and unknown posteriori variance - is applied to detect blunders in errors-in-variables (EIV) models, in which gross errors are in the vector of observations and/or in the coefficient matrix. This work is a follow-up to an earlier work in which we presented the formulation of the weighted total least squares (WTLS) based on the standard least squares theory. This method allows one to directly apply the existing body of knowledge of the least squares theory to the errors-in-variables models. Among those applications, data snooping methods in an EIV model are of particular interest, which is the subject of discussion in the present contribution. This paper generalizes the Baarda's data snooping procedure of the standard least squares theory to an EIV model. Two empirical examples, a linear regression model and a 2-D affine transformation, using simulated and real data are presented to show the efficacy of the presented formulation. It is highlighted that the method presented is capable of detecting outlying equations (rather than outlying observations) in a straightforward manner. Further, the WTLS method can be used to handle different TLS problems. For example, the WTLS problem for the conditions and mixed models, the WTLS problem subject to constraints and variance component estimation for an EIV model can easily be established. These issues are in progress for future publications.

Keywords: errors-in-variables model, weighted total least squares, standard least-squares theory, outlier detection, data snooping

1. INTRODUCTION

Gross errors (blunders) are probable to occur in many geodetic measurements. Blunder detection is then an important issue in many engineering applications including engineering surveying and geodesy. So far, many research papers have been published in geodetic literature relating to outlier detection for ordinary least squares problem where the vector of observations is only affected by random/gross errors. We may for example refer to *Marshall and Bethel (1996)*, *Hekimoglu (1997)*, *Amiri-Simkooei (2003)*, *Guo et al. (2007)*, *Hekimoglu and Erenoglu (2009)*, *Koch (1999)*, *Khodabandeh and Amiri-Simkooei (2011)*. The methods have not yet been applied to the errors-in-variables (EIV) models where in contrast to the Gauss-Markov model (GMM), the coefficient matrix is also contaminated by random/gross errors.

In the estimation theory, we usually distinguish between the standard least squares (SLS) and the total least squares (TLS). TLS originates from the work of *Golub and van Loan (1980)* in mathematical literature in which they introduced the errors-in-variables (EIV) models. There are many algorithms in statistical and geodetic literature to solve an EIV model using the weighted total least squares (WTLS), where errors are present in both the observation vector and the coefficient matrix. The reader is referred to *Van Huffel and Vandewalle (1991)*, *Davis (1999)*, *Felus (2004)*, *Schaffrin and Wieser (2008)*, *Schaffrin and Felus (2009)*, *Tong et al. (2011)* and *Shen et al. (2011)*.

All WTLS algorithms presented in the statistical and geodetic literature are based on the assumption that the observations are contaminated only by random noise and hence outliers (gross errors) do not exist in the observation vector and/or the coefficient matrix. In reality, this assumption does not always hold and there may exist outliers in the elements of observation vector and coefficient matrix. *Schaffrin and Uzun (2011)* and *Schaffrin and Uzun (2012)* presented methodology to treat detection of a single outlier in either the observation vector or the coefficient matrix based on the first-order approximation F-distribution test statistic.

Amiri-Simkooei and Jazaeri (2012) have recently formulated the WTLS problem based on the standard least squares theory in which the standard GMM can be considered instead of solving a nonlinear Gauss-Helmert model (GHM) in an iterative manner. The algorithm takes into consideration the complete structure for the covariance matrix of the coefficient matrix. Having this formulation available, one can apply the existing body of the knowledge of the least squares to the WTLS problem. In this contribution, we further explore the high potential capability of the formulation to implement the data snooping procedure (*Baarda, 1968*) in an EIV model. This work is motivated by the fact that although some research is ongoing for outlier detection in an EIV model, a drawback, however, is that the methods proposed can be used when only one outlier appears either in the observation vector or in the coefficient matrix. Further, some of the available methods can provide approximate solutions and hence leave part of the outlier to the estimated parameters.

This paper is organized as follows. In Section 2, a general solution is given to the WTLS problem which is formulated in the standard least squares framework. Section 3 shows the applicability of the algorithm to apply the data snooping procedures for identification of outlying measurements. In Section 4, simulation studies and empirical examples give insight into the efficacy of the proposed algorithm. Finally we make some conclusions in Section 5.

2. WEIGHTED TOTAL LEAST SQUARES (WTLS)

Consider the linear model in which, in addition to the observation vector, the elements of the coefficient matrix are also perturbed by random errors. In this case, the usual GMM is converted to the following EIV model:

$$\boldsymbol{y} = \left(\boldsymbol{A} - \boldsymbol{E}_{\boldsymbol{A}}\right)\boldsymbol{x} + \boldsymbol{e}_{\boldsymbol{y}}, \tag{1}$$

with its stochastic properties characterized by

$$\begin{bmatrix} \boldsymbol{e}_{\boldsymbol{y}} \\ \boldsymbol{e}_{\boldsymbol{A}} \end{bmatrix} := \begin{bmatrix} \boldsymbol{e}_{y} \\ \mathrm{vec}\left(\boldsymbol{E}_{\boldsymbol{A}}\right) \end{bmatrix} \sim \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \sigma_0^2 \begin{bmatrix} \boldsymbol{Q}_{\boldsymbol{y}} & 0 \\ 0 & \boldsymbol{Q}_{A} \end{bmatrix} \right), \tag{2}$$

where $\boldsymbol{e}_{\boldsymbol{y}}$ is the m-vector of the error of observations, $\boldsymbol{A}$ is the $m \times n$ coefficient matrix, $\boldsymbol{E}_{\boldsymbol{A}}$ is the $m \times n$ random error of the coefficient matrix $\boldsymbol{A}$, $\boldsymbol{x}$ is the n-vector of unknown parameters, $\boldsymbol{D}\left(\boldsymbol{e}_{\boldsymbol{y}}\right) = \sigma_0^2 \boldsymbol{Q}_{\boldsymbol{y}}$ and $\boldsymbol{D}\left(\boldsymbol{e}_{\boldsymbol{A}}\right) = \sigma_0^2 \boldsymbol{Q}_{\boldsymbol{A}}$ are the corresponding symmetric and non-negative dispersion matrices of size $m \times m$ and $mn \times mn$ for the vector of observation and coefficient matrix, respectively. The symbol 'vec' denotes the operator that converts a matrix to a column vector by stacking the columns, one column underneath the other. In both expressions, σ_0^2 is the unknown variance factor of the unit weight. Weighted total least squares (WTLS) problem seeks to solve the following optimization problem:

$$\text{minimize } \boldsymbol{e}_{\boldsymbol{y}}^{\mathrm{T}} \boldsymbol{Q}_{\boldsymbol{y}}^{-1} \boldsymbol{e}_{\boldsymbol{y}} + \boldsymbol{e}_{\boldsymbol{A}}^{\mathrm{T}} \boldsymbol{Q}_{\boldsymbol{A}}^{-1} \boldsymbol{e}_{\boldsymbol{A}} \quad \text{subject to } \boldsymbol{y} - \boldsymbol{e}_{\boldsymbol{y}} = \left(\boldsymbol{A} - \boldsymbol{E}_{\boldsymbol{A}}\right)\boldsymbol{x}. \tag{3}$$

Schaffrin and Wieser (2008) introduced the "weighted TLS" to the geodetic community based on the inverse of the covariance matrix. *Amiri-Simkooei and Jazaeri (2012)* developed a novel WTLS algorithm that minimizes the objective function (3) and resembles the standard least squares method. This algorithm is used in the present contribution. Similar to *Schaffrin and Wieser (2008)*, the objective function Φ to be minimized is

$$\Phi := \boldsymbol{e}_{\boldsymbol{y}}^{\mathrm{T}} \boldsymbol{Q}_{\boldsymbol{y}}^{-1} \boldsymbol{e}_{\boldsymbol{y}} + \boldsymbol{e}_{\boldsymbol{A}}^{\mathrm{T}} \boldsymbol{Q}_{\boldsymbol{A}}^{-1} \boldsymbol{e}_{\boldsymbol{A}} + 2\lambda^{\mathrm{T}} \left(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x} - \boldsymbol{e}_{\boldsymbol{y}} + \left(\boldsymbol{x}^{\mathrm{T}} \otimes \boldsymbol{I}_{\boldsymbol{m}} \right) \boldsymbol{e}_{\boldsymbol{A}} \right), \tag{4}$$

with $\boldsymbol{\lambda}$ a m-vector of unknown Lagrange multipliers, ⊗ is the Kronecker product operator of two matrices and $\boldsymbol{I}_{\boldsymbol{m}}$ an identity matrix of size m. Minimizing the objective function Φ given by Eq. (4) results in the following equations (*Amiri-Simkooei and Jazaeri, 2012*)

$$\tilde{\boldsymbol{e}}_{\boldsymbol{y}} = \boldsymbol{Q}_{\boldsymbol{y}} \boldsymbol{Q}_{\tilde{\boldsymbol{y}}}^{-1} \left(\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{x}} \right) \tag{5}$$

and

$$\tilde{\boldsymbol{e}}_{\boldsymbol{A}} = \mathrm{vec}\, \tilde{\boldsymbol{E}}_{\boldsymbol{A}} = -\boldsymbol{Q}_{\boldsymbol{A}} \left(\hat{\boldsymbol{x}} \otimes \boldsymbol{I}_{\boldsymbol{m}} \right) \boldsymbol{Q}_{\tilde{\boldsymbol{y}}}^{-1} \left(\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{x}} \right), \tag{6}$$

where

$$\boldsymbol{Q}_{\tilde{y}} = \boldsymbol{Q}_y + \left(\hat{\boldsymbol{x}}^{\mathrm{T}} \otimes \boldsymbol{I}_m\right)\boldsymbol{Q}_A\left(\hat{\boldsymbol{x}} \otimes \boldsymbol{I}_m\right) \tag{7}$$

is the estimates (due to the presence of $\hat{\boldsymbol{x}}$) for the covariance matrix of the predicted observables $\tilde{\boldsymbol{y}} = \boldsymbol{y} - \tilde{\boldsymbol{E}}_A \boldsymbol{x}$ ("~" and "^" represent the "predicted" and "estimated" quantities, respectively). The least squares estimate $\hat{\boldsymbol{x}}$ is given as

$$\hat{\boldsymbol{x}} = \left(\tilde{\boldsymbol{A}}^{\mathrm{T}} \boldsymbol{Q}_{\tilde{y}}^{-1} \tilde{\boldsymbol{A}}\right)^{-1} \tilde{\boldsymbol{A}}^{\mathrm{T}} \boldsymbol{Q}_{\tilde{y}}^{-1} \tilde{\boldsymbol{y}}\,, \tag{8}$$

where $\tilde{\boldsymbol{A}} = \boldsymbol{A} - \tilde{\boldsymbol{E}}_A$, $\tilde{\boldsymbol{y}} = \boldsymbol{y} - \tilde{\boldsymbol{E}}_A \hat{\boldsymbol{x}}$, and $\boldsymbol{Q}_{\tilde{y}} = \boldsymbol{Q}_y + \left(\hat{\boldsymbol{x}}^{\mathrm{T}} \otimes \boldsymbol{I}_m\right)\boldsymbol{Q}_A\left(\hat{\boldsymbol{x}} \otimes \boldsymbol{I}_m\right)$ play respectively the role of the coefficient matrix ***A***, the observation vector ***y***, and the covariance matrix $\boldsymbol{Q}_y$ of the standard least squares. We thus deal with the *predicted* design matrix $\tilde{\boldsymbol{A}}$, the *predicted* observation vector $\tilde{\boldsymbol{y}}$, and the covariance matrix $\boldsymbol{Q}_{\tilde{y}}$ of the *predicted* observables. Eq. (8) is the estimates for the unknown parameters of the EIV model $\boldsymbol{y} = \left(\boldsymbol{A} - \boldsymbol{E}_A\right)\boldsymbol{x} + \boldsymbol{e}_y$, which looks in fact as if to rewrite it in the form of $\boldsymbol{y} - \boldsymbol{E}_A \boldsymbol{x} = \left(\boldsymbol{A} - \boldsymbol{E}_A\right)\boldsymbol{x} + \boldsymbol{e}_y - \boldsymbol{E}_A \boldsymbol{x}$, or equivalently $\tilde{\boldsymbol{y}} = \tilde{\boldsymbol{A}}\boldsymbol{x} + \tilde{\boldsymbol{e}}$. Eq. (8) has the estimator $\hat{\boldsymbol{x}}$ on both sides. Therefore the least squares estimate $\hat{\boldsymbol{x}}$ is to be obtained in an iterative manner (see Algorithm 1). The estimate for the covariance matrix of the unknown parameters $\hat{x}$ is given by $\boldsymbol{Q}_{\hat{x}} = \left(\tilde{\boldsymbol{A}}^{\mathrm{T}} \boldsymbol{Q}_{\tilde{y}}^{-1} \tilde{\boldsymbol{A}}\right)^{-1}$.

Note 1. The covariance matrix $\boldsymbol{Q}_{\hat{x}}$ was directly obtained based on the assumption that the vector ***x*** is not random. But ***x*** is unknown and therefore one has to use its estimate $\hat{x}$. This will then take the risk to have a random vector in the formulation. This indicates that for application of the error propagation law one has to take into account the randomness of $\hat{x}$, which has been ignored in the formulation. The approximation applied is indeed in conjunction with all nonlinear least squares problems in which the covariance matrix of $\hat{x}$ is considered to be identical to that of $\boldsymbol{\delta}\hat{\boldsymbol{x}}$. Also the above-given covariance matrix $\boldsymbol{Q}_{\hat{x}}$ is identical to that of GHM (see *Amiri-Simkooei and Jazaeri, 2012*). In addition, the unbiasedness property of the linear least squares problem will also be violated in the TLS formulation. This bias is proportional to the observations precision and the geometric properties of the manifold such as curvature (see *Teunissen, 1985, 1990*).

This formulation is similar (but practically different) to that of *Shen et al. (2011)*. We highlight that there is an intrinsic difference between the two algorithms in the sense of computing $\tilde{\boldsymbol{e}}_y$ and $\tilde{\boldsymbol{e}}_A$. Eqs (5) and (6) are used to compute $\tilde{\boldsymbol{e}}_y$ and $\tilde{\boldsymbol{e}}_A$ in our algorithm**,** while *Shen et al. (2011)* use, in addition, the nonlinear Gauss-Newton increment $\boldsymbol{\delta}\hat{\boldsymbol{x}}$ in their formulation.

Having available this formulation, the estimated vectors of the observations and residuals in the model $\tilde{\boldsymbol{y}} = \tilde{\boldsymbol{A}}\boldsymbol{x} + \tilde{\boldsymbol{e}}$ can be easily obtained. One may write

$$\hat{\boldsymbol{y}} = \boldsymbol{P}_{\tilde{A}}\tilde{\boldsymbol{y}} = \tilde{\boldsymbol{A}}\hat{\boldsymbol{x}} = \left(\boldsymbol{A} - \tilde{\boldsymbol{E}}_A\right)\hat{\boldsymbol{x}} \quad \text{and} \quad \hat{\boldsymbol{e}} = \boldsymbol{P}_{\tilde{A}}^{\perp}\tilde{\boldsymbol{y}} = \tilde{\boldsymbol{y}} - \hat{\boldsymbol{y}} = y - \boldsymbol{A}\hat{\boldsymbol{x}} ,$$

where $\boldsymbol{P}_{\tilde{A}} = \tilde{\boldsymbol{A}}\left(\tilde{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{Q}_{\tilde{y}}^{-1}\tilde{\boldsymbol{A}}\right)^{-1}\tilde{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{Q}_{\tilde{y}}^{-1}$ and $\boldsymbol{P}_{\tilde{A}}^{\perp} = \boldsymbol{I} - \boldsymbol{P}_{\tilde{A}}$ are two orthogonal projectors (see *Teunissen, 2000a*). We highlight that $\hat{\boldsymbol{e}}$ is the so-called 'total residuals' of the model, and hence different from $\tilde{\boldsymbol{e}}_y$ in Eq. (5). In the geodetic literature, it is also referred to as the 'misclosure' vector. Similar to the standard least squares theory, the covariance matrices of the least-squares estimates $\hat{\boldsymbol{y}}$ and $\hat{\boldsymbol{e}}$ are

$$\boldsymbol{Q}_{\hat{y}} = \boldsymbol{P}_{\tilde{A}}\boldsymbol{Q}_{\tilde{y}} = \tilde{\boldsymbol{A}}\boldsymbol{Q}_{\hat{x}}\tilde{\boldsymbol{A}}^{\mathrm{T}} \quad \text{and} \quad \boldsymbol{Q}_{\tilde{e}} = \boldsymbol{P}_{\tilde{A}}^{\perp}\boldsymbol{Q}_{\tilde{y}} = \boldsymbol{Q}_{\tilde{y}} - \boldsymbol{Q}_{\hat{y}} ,$$

respectively. Similarly, the variance factor of the unit weight is estimated as

$$\hat{\sigma}_0^2 = \frac{\hat{\boldsymbol{e}}^{\mathrm{T}}\boldsymbol{Q}_{\tilde{y}}^{-1}\hat{\boldsymbol{e}}}{m-n} . \tag{9}$$

We highlight that the method described above is capable of handling different TLS problems. For example, the WTLS problem for the condition and mixed models, the WTLS problem subject to constraints and/or variance component estimation for an EIV model can easily be established. These issues are in progress for future publications. In the present contribution, we apply this algorithm to identify outlying equations in an EIV model (next section).

3. DATA-SNOOPING ON WTLS PROBLEM

In many data analysis tasks in general and geodetic applications in particular a large number of observations are being recorded or sampled. One important step towards obtaining a coherent analysis is the detection of outlaying observations. This guarantees the unbiasedness property of the least-squares estimates. Blunders, if left undetected, lead to unreliable results and make the interpretations difficult. It is therefore important to identify outliers prior to modeling and the final adjustment process. In a linear model, the process of data snooping (removing outlying observations) using the sequential procedure can be summarized into two important steps.

The first step regards the compatibility of the estimated posterior variance factor $\hat{\sigma}_0^2$ with a priori selected variance factor σ_0^2. It is a global test that is a good indicator to assess the correctness results of the least squares because all observations, stochastic and functional model take part into the computation of $\hat{\sigma}_0^2$. The formulation of the hypotheses is as follows (*Teunissen, 2000b*; *Leick, 1998*; *Koch, 1999*)

$$H_0 : \sigma_0^2 = \hat{\sigma}_0^2 \quad \text{versus} \quad H_a : \sigma_0^2 \neq \hat{\sigma}_0^2 . \tag{10}$$

Under the normality of the data, if the numerical value

$$\chi^2 = \frac{\hat{\sigma}_0^2}{\sigma_0^2}\left(m-n\right) \tag{11}$$

is such that $\chi^2 > \chi^2_{m-n,1-(\alpha/2)}$ or $\chi^2 < \chi^2_{m-n,\alpha/2}$, then the zero hypothesis is rejected, otherwise it is not. If the null hypothesis is accepted, it means that everything works properly, otherwise $\hat{\boldsymbol{e}}^T \boldsymbol{Q}_{\tilde{y}}^{-1} \hat{\boldsymbol{e}} = \tilde{\boldsymbol{e}}_y^{\mathrm{T}} \boldsymbol{Q}_y^{-1} \tilde{\boldsymbol{e}}_y + \tilde{\boldsymbol{e}}_A^{\mathrm{T}} \boldsymbol{Q}_A^{-1} \tilde{\boldsymbol{e}}_A$ is too large and hence there is a need to find out what is wrong with the linear system. If the zero hypothesis is rejected it informs that the weight matrix is incorrectly defined, model is incorrect, and/or erroneous measurements are appeared in the observation vector. Therefore, one has to find out the problem using further investigations or statistical tests (see e.g., *Leick 1998, 2004*).

Erroneous measurements can hence lead to the rejection of the above-mentioned hypothesis. The normal and F tests can be used to detect outliers when the a posteriori variance is known and unknown, respectively (*Teunissen, 2000b*; *Koch, 1999*). These tests are applied to detect only one erroneous measurement. In the case of multiple outliers, the testing procedure should be consecutively applied. This procedure is called the ‘data snooping’. The test statistics for data snooping are approximated as follows:

$$w_{i,\mathrm{N}} = \frac{\boldsymbol{c}_i^{\mathrm{T}} \boldsymbol{Q}_{\tilde{y}}^{-1} \hat{\boldsymbol{e}}}{\sigma_0 \sqrt{\boldsymbol{c}_i^{\mathrm{T}} \boldsymbol{Q}_{\tilde{y}}^{-1} \boldsymbol{Q}_{\hat{e}} \boldsymbol{Q}_{\tilde{y}}^{-1} \boldsymbol{c}_i}} \sim \mathrm{N}_{\alpha/2}\left(0,1\right) \tag{12}$$

when the variance factor of the unit weight is known, and

$$w_{i,\mathrm{F}} = \frac{\boldsymbol{c}_i^{\mathrm{T}} \boldsymbol{Q}_{\tilde{y}}^{-1} \hat{\boldsymbol{e}}}{\hat{\sigma}_0 \sqrt{\boldsymbol{c}_i^{\mathrm{T}} \boldsymbol{Q}_{\tilde{y}}^{-1} \boldsymbol{Q}_{\hat{e}} \boldsymbol{Q}_{\tilde{y}}^{-1} \boldsymbol{c}_i}} \sim \sqrt{\mathrm{F}_{1-\alpha,1,df}} \tag{13}$$

when it is unknown, to be estimated under the null hypothesis. In the preceding equations, $\mathrm{N}_{\alpha/2}\left(0,1\right)$ and $\mathrm{F}_{1-\alpha,1,df}$ represent the normal and F tests, respectively, and $\boldsymbol{c}_i$ is the canonical unit vector having one at *i*-th position and zeros elsewhere.

In conjunction with Note 1, it is highlighted that due to the random nature of a few elements in Eqs (11)–(13), the presented distributions are just approximations of real distributions. Some preliminary results indicate that these assumptions are satisfactory for the two applications considered in the present contribution. But, further research is ongoing to see how good these approximations are.

We therefore decide to reject an observation (i.e., mask it as an outlier) if

$$\left|w_{j,\mathrm{N}}\right| \geq \left|w_{i,\mathrm{N}}\right| \quad \forall i = 1, \ldots, m \quad \text{and} \quad \left|w_{j,\mathrm{N}}\right| > \mathrm{N}_{\alpha/2}\left(0,1\right) \tag{14}$$

when the variance factor of the unit weight is known, or

$$\left|w_{j,\mathrm{F}}\right| \geq \left|w_{i,\mathrm{F}}\right| \quad \forall i = 1, \ldots, m \quad \text{and} \quad \left|w_{j,\mathrm{F}}\right| > \sqrt{\mathrm{F}_{1-\alpha,1,df}} \tag{15}$$

when it is unknown. If the covariance matrix $\boldsymbol{Q}_{\tilde{\boldsymbol{y}}}$ is diagonal, the above test statistics can be simplified as $w_{i,\mathrm{N}} = \hat{\boldsymbol{e}}_i / \sigma_0 \sigma_{\hat{\boldsymbol{e}}_i}$ or $w_{i,\mathrm{F}} = \hat{\boldsymbol{e}}_i / \hat{\sigma}_0 \sigma_{\hat{\boldsymbol{e}}_i}$, where $\sigma_{\hat{\boldsymbol{e}}_i} = \sqrt{(\boldsymbol{Q}_{\hat{\boldsymbol{e}}})_{ii}}$ indicates the standard deviation of the ith least squares residual (see, e.g., *Teunissen, 2000b*; *Koch, 1999*). *Koch (1999)* has shown that $\sqrt{\mathrm{F}_{1-\alpha,1,df}}$ and Student's t-test $\mathrm{T}_{m-n,\alpha/2}$ (Pope's Tau-test, *Pope, 1976*) are both equivalent. In this contribution we use the F test.

At the first step, the largest w-test statistic (e.g. standardized residual) is selected and depending on the known or unknown posteriori variance the normal or F test is used to decide whether the measurement is erroneous or not. If the hypothesis is rejected it is indicated that the corresponding 'observation equation' has an outlier. The outlier could be in the observation vector, in the coefficient matrix, and/or in both. The observation equation should be eliminated from the list of the equations. After removing the suspected equation, the new dataset (excluding the blunder) is to be adjusted again and the next possible outlier should be identified. This procedure is repeated until the statistical tests are all accepted.

The weighted total least squares problem along with data snooping procedure can be applied in an iterative procedure (see Algorithm 1). In the following section, two empirical examples using real and simulated data are presented to evaluate the efficiency of the proposed algorithm. Both examples contain gross errors and are used frequently in engineering surveying and geomatic applications.

4. NUMERICAL RESULTS AND DISCUSSIONS

4.1. Linear regression model

Simulated data (one outlier)

As a first example, we consider a simulation study for straight line fit problem in which both variables are observed and thus we deal with an EIV model. The goal is to find the slope *a* and intercept *b* of the regression line

$$y_i - e_{y_i} = a\left(x_i - e_{x_i}\right) + b \,. \tag{16}$$

20 data points are simulated for this model. The x_i coordinates are assumed to be $x_i = i$, $i = 0, \ldots, 19$, and y_i values are calculated based on the line parameters *a* and *b*. The covariance matrices of the observables, i.e. of $\boldsymbol{x}$ and $\boldsymbol{y}$ are set to $\boldsymbol{Q_x} = \sigma_x^2 \boldsymbol{I}_{20}$ and $\boldsymbol{Q_y} = \sigma_y^2 \boldsymbol{I}_{20}$, where σ_x and σ_y indicate standard deviations of the $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively. Both $\boldsymbol{x}$ and $\boldsymbol{y}$ are corrupted by normal random noise.

For the first scenario, one randomly selected elements of $\boldsymbol{x}$ (out of 20), one randomly selected elements of the $\boldsymbol{y}$ or simultaneously both elements of $\boldsymbol{x}$ and $\boldsymbol{y}$ are contaminated by a systematic error (as a blunder) of size $2\sigma_x$, $4\sigma_x$, $6\sigma_x$, $8\sigma_x$ and $2\sigma_y$, $4\sigma_y$, $6\sigma_y$, $8\sigma_y$ for selected elements of $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively. The systematic error (multiples of sigma) is added to the actual line $y = ax + b$, where the line parameters are set to $a = 1$ and $b = 1$. Then the proposed algorithm is applied to detect outliers of the dataset over 10000 independent runs. We can compute the percentage of samples with which the proposed algorithm succeeded to detect outliers of the dataset by

$$\frac{\text{Success}\left(\text{outlier detected}\right)}{10000}\times 100\,.$$

We summarize the results of outlier detection for different standard deviations of σ_x and σ_y in Table 1 (Columns 3–5).

Simulated data (two outliers)

In the second scenario, two randomly selected points are contaminated by gross errors of size $2\sigma_x$, $4\sigma_x$, $6\sigma_x$, $8\sigma_x$ and $2\sigma_y$, $4\sigma_y$, $6\sigma_y$, $8\sigma_y$ individually for selected elements of $\boldsymbol{x}$, $\boldsymbol{y}$ or simultaneously for both $\boldsymbol{x}$ and $\boldsymbol{y}$. The proposed algorithm is applied to detect the two outliers for each data set and repeat it over 10000 independent runs. Numerical results are listed in Table 1 (Columns 6–8) for different standard deviations of σ_x and σ_y.

The results provided in Table 1 show the high performance capability of the proposed approach to detect outlying measurements (or in fact outlying equations). As expected, when the magnitude of the blunder increases the detection power increases. This also happens when blunders exist simultaneously in both x and y. In addition, outlier detection results with one outlying observation have larger success rate compared to those with two outlying observations (Table 1, compare Columns 3–5 with Columns 6–8).

Table 1. Success rate (by means of percentages) of one detected blunder (Columns 4–5) and two detected blunders (6–8) introduced to the line $y = ax + b$ ($a = 1$ and $b = 1$) with different size for blunders and different standard deviations for x_i and y_i observations. Blunders are introduced to x only (Columns 3 and 6), y only (Columns 4 and 7), and both x and y (Columns 5 and 8).

Standard Deviation	Blunder Size	Blunder in					
		x_i	y_i	x_i & y_i	x_i	y_i	x_i & y_i
		One Blunder			Two Blunders		
$\sigma_x = 0.001$ $\sigma_y = 0.001$	2σ	6.85	6.81	82.00	3.30	3.47	69.24
	4σ	82.09	81.65	100.00	69.59	68.93	99.83
	6σ	99.58	99.54	100.00	97.93	98.06	100.00
	8σ	100.00	99.99	100.00	99.89	99.87	100.00
$\sigma_x = 0.002$ $\sigma_y = 0.001$	2σ	21.52	0.57	75.55	12.79	0.18	62.38
	4σ	96.82	22.21	100.00	92.51	12.61	99.82
	6σ	99.99	74.98	100.00	99.76	61.98	100.00
	8σ	100.00	96.85	100.00	100.00	92.30	100.00
$\sigma_x = 0.001$ $\sigma_y = 0.002$	2σ	0.45	21.44	75.68	0.19	12.80	62.18
	4σ	21.68	96.72	99.99	13.55	91.77	99.73
	6σ	74.79	99.99	100.00	62.55	99.86	99.98
	8σ	96.84	100.00	100.00	91.92	100.00	100.00

Real data

In the second example we make use of a real data set presented in *Kelly (1984)*. The data are provided in Table 2 (Columns 2 and 3). "*The data consist of simultaneous pairs of measurements of serum kanamycin levels in blood samples drawn from twenty babies. One of the measurements was obtained by a heelstick method, the other by using an umbilical catheter.*" (*Fekri and Ruiz-Gazen, 2004*).

At first, we assume that there is no blunder in the dataset and both measurements have equal unit variances. We aim to estimate the slope *a* and intercept *b* of the regression line using the presented WTLS algorithm. The results are given in Table 3. The precision of the estimates is also provided. The variance factor of the unit weight is estimated as $\sigma_0^2 = 4.8571$.

Because the a posteriori variance σ_0^2 is unknown in this example, the F test is repeatedly applied to detect blunders. The estimated (total) residuals $\hat{e}$ and the w-test statistics are presented in Table 2 (Columns 4 and 5). The largest value (in the absolute sense) corresponds to the point 2, which is $w_2 = -2.89$. Considering a significance level

Table 2. Twenty observed points (Columns 2 and 3) according to Kelly (1984); estimated least squares residuals $\hat{e}$ along with w-test statistic values before removing outlying equation (Columns 4 and 5), and after removing outlying equation (Columns 6 and 7). Bold numbers indicate outlying equations.

Point No.	Observed Data		Outlier Detection (1)		Outlier Detection (2)	
	x	y	$\hat{e}$	w-test	$\hat{e}$	w-test
1	23	25.2	1.75	0.56	0.68	0.26
2	33.2	26	**–8.33**	**–2.89**	---	---
3	16.3	16.3	–0.01	–0.00	0.42	0.16
4	26.3	27.2	0.23	0.08	–1.57	–0.63
5	20	23.2	2.95	0.94	2.55	0.97
6	20	18.1	–2.15	–0.69	–2.55	–0.97
7	20.6	22.2	1.31	0.42	0.78	0.30
8	18.9	17.2	–1.88	–0.61	–2.03	–0.78
9	17.8	18.8	0.89	0.29	0.99	0.38
10	20	16.4	–3.85	–1.24	–4.25	–1.63
11	26.4	24.8	–2.28	–0.74	–4.10	–1.61
12	21.8	26.8	4.63	1.49	3.83	1.49
13	14.9	15.4	0.58	0.19	1.33	0.52
14	17.4	14.9	–2.58	–0.84	–2.40	–0.94
15	20	18.1	–2.15	–0.69	–2.55	–0.97
16	13.2	16.3	3.30	1.10	4.42	1.75
17	28.4	31.3	2.09	0.73	–0.18	–0.08
18	25.9	31.2	4.65	1.59	2.94	1.23
19	18.9	18	–1.08	–0.35	–1.23	–0.47
20	13.8	15.6	1.96	0.65	2.95	1.17

Table 3. Estimated straight-line parameters along with their standard deviations using data in Table 2.

Parameter	WTLS (This paper)
$\hat{a}$	1.0663138967
$\hat{b}$	−1.0719816068
$\sigma_{\hat{a}}$	0.1517691911
$\sigma_{\hat{b}}$	3.2438766119

Table 4. Estimated straight-line parameters for different methods.

Parameter	*M*-estimates of orthogonal regression	*S*-estimates of orthogonal regression	WTLS (This paper)
$\hat{a}$	1.29	1.29	1.29
$\hat{b}$	−5.26	−5.26	−5.14

of $\alpha = 0.05$ it follows that $\sqrt{F_{1-\alpha,1,df}} = \sqrt{F_{1-0.05,1,18}} = 2.10$. Because $2.89 > 2.10$, it follows that the second equation, i.e., observations at Point 2 (x, y, or both), contains blunder. We now leave out this point and repeat the process. The least squares residuals and the w-test statistics are given in Columns 6 and 7 of Table 2. The largest w-test is 1.75 which, now, is smaller than the critical value of $\sqrt{F_{1-\alpha,1,df}} = \sqrt{F_{1-0.05,1,17}} = 2.11$, indicating that the remaining observations have no more blunder.

Several robust regression estimators instead of the classical estimator, i.e., the least-squares estimator, have been investigated in the literature that are not so strongly affected by outliers. They replaced the squared residuals in the ordinary least-squares estimation by another function of the residuals. *M*-estimators and *S*-estimators are two classes of these robust regression techniques. The *M*- and *S*-estimates of orthogonal regression were reported by *Fekri and Ruiz-Gazen (2004)*. These results along with the results of the proposed data snooping algorithm are presented in Table 4. The results significantly deviate from those provided in Table 3 in which the blunder exists, but they closely follow the results of *M* and *S*-estimates of orthogonal regression. This then highlight the high potential capability of the proposed data snooping algorithm.

4.2. Two-dimensional affine transformation

Figure 1 considers ten data points measured in the start and target coordinate systems. The nominal standard deviation of the measurements in both systems is 0.01. The coordinates of these points are listed in Table 5.

The model for the planar linear affine transformation (six-parameter transformation) is as follows:

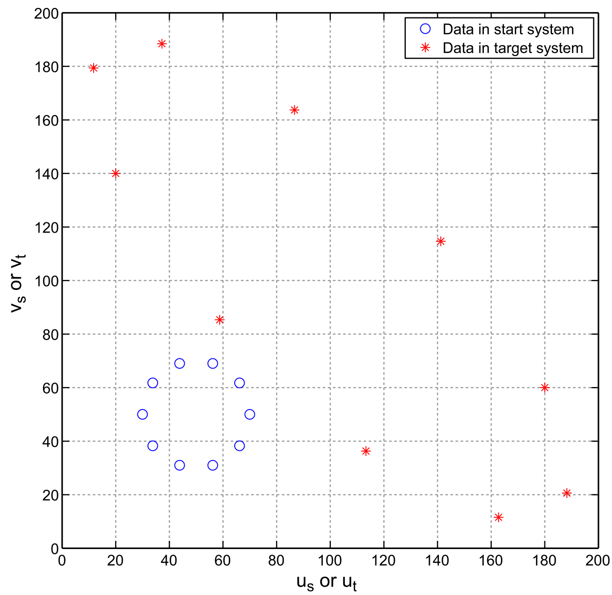

Fig. 1. Coordinates of 10 points in start and target systems.

Table 5. Observed points in start and target coordinate systems.

Point No.	Start System		Target System	
	u_s	v_s	u_t	v_t
1	70.00	49.98	180.00	59.98
2	66.16	61.74	141.21	114.67
3	56.17	69.02	86.70	163.71
4	43.83+0.1	69.01	37.26	188.45
5	33.82	61.77	11.77	179.38
6	30.00	50.00	19.99	140.00
7	33.80	38.25	58.77	85.35
8	43.83	30.97	113.31	36.28
9	56.17	30.98	162.77	11.56
10	66.19	38.24	188.24	20.61

$$\begin{bmatrix} u_t \\ v_t \end{bmatrix} = \begin{bmatrix} u_s & v_s & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & u_s & v_s & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ b_1 \\ c_1 \\ a_2 \\ b_2 \\ c_2 \end{bmatrix}, \tag{17}$$

where the parameters c_1 and c_2 are the shifts along the u and v axes, respectively. The other parameters a_1, a_2, b_1 and b_2 are related to the four physical parameters of a 2-D linear transformation, which include two scales along the u and v axes, one rotation, and one non-perpendicularity (or affinity) parameter. u_t, v_t and u_s, v_s are the components of coordinates in the target and start coordinate systems.

The u_s component of the Point 4 in the source coordinate system is intentionally contaminated by a gross error of size 0.10. We now apply the proposed algorithm to find the possible outlying equations; we expect Eqs (7) and (8) to be masked because the u_s coordinate of the Point 4 appears in these two equations.

The estimated vectors of the (total) residuals $\hat{\boldsymbol{e}}$ along with the w-test values are provided in Table 6. These include results for three cases: those with the erroneous equations (Columns 4 and 5), removing the first outlying equation (Columns 6 and 7), and removing the second outlying equation (Columns 8 and 9). For these three cases, the maximum (in absolute sense) w-test statistic values are 2.79, 3.05, and 1.61, while the corresponding critical values of $\sqrt{\mathrm{F}_{1-\alpha,1,df}}$ are $\sqrt{\mathrm{F}_{1-0.05,1,14}} = 2.14$, $\sqrt{\mathrm{F}_{1-0.05,1,13}} = 2.16$ and $\sqrt{\mathrm{F}_{1-0.05,1,12}} = 2.18$, respectively. The first and second cases contain gross errors in

Table 6. Estimated least squares (total) residuals $\hat{\boldsymbol{e}}$ along with w-test statistic values for three cases: introduced blunder to u_s coordinate of Point 4 (Columns 4 and 5), after removing the first outlier (Columns 6 and 7), and after removing the second outlier (Columns 8 and 9). Bold numbers indicate outlying equations.

Eq. No.	Point No.	Coord.	$\hat{\boldsymbol{e}}$	w-test	$\hat{\boldsymbol{e}}$	w-test	$\hat{\boldsymbol{e}}$	w-test
1	1	u	−0.02	−0.19	−0.03	−0.50	−0.04	−0.89
2		v	0.01	−0.05	0.01	−0.24	0.03	−0.23
3	2	u	0.11	1.25	0.06	0.90	0.03	1.61
4		v	−0.03	0.73	−0.03	0.35	0.02	1.49
5	3	u	0.15	1.12	0.08	0.12	0.02	0.22
6		v	−0.10	0.21	−0.10	−0.88	−0.02	−0.20
7	4	u	**−0.34**	**−2.79**	---	---	---	---
8		v	0.21	−0.76	**0.21**	**3.05**	---	---
9	5	u	0.12	0.63	0.05	−0.64	−0.01	−1.16
10		v	−0.10	−0.15	−0.10	−1.43	−0.02	−1.22
11	6	u	0.04	0.49	−0.00	−0.23	−0.04	−0.40
12		v	−0.01	0.29	−0.01	−0.29	0.04	0.31
13	7	u	0.08	0.74	0.07	0.85	0.06	1.52
14		v	−0.04	0.28	−0.04	0.23	−0.03	0.61
15	8	u	−0.09	−0.56	−0.07	−0.45	−0.06	−0.81
16		v	0.07	0.04	0.07	0.33	0.05	0.27
17	9	u	0.02	0.07	0.04	0.68	0.07	1.23
18		v	−0.02	−0.07	−0.02	0.34	−0.05	0.10
19	10	u	−0.06	−0.77	−0.04	−0.76	−0.03	−1.36
20		v	0.01	−0.52	0.01	−0.49	−0.01	−1.19

Table 7. Estimated planar linear affine transformation parameters for four cases: Case 1) before introducing outlier to the observations, Case 2) after introducing the u_s coordinate of the Point 4 to have a blunder, Case 3) after removing the first outlying equation, and Case 4) after removing the second outlying equation.

Parameter	Case 1	Case 2	Case 3	Case 4	Stand. Dev.
$\hat{a}_1$	3.9997	4.0009	4.0001	3.9994	0.0010
$\hat{b}_1$	−1.9993	−2.0031	−2.0007	−1.9985	0.0012
$\hat{c}_1$	−0.0103	0.0785	0.0217	−0.0296	0.0752
$\hat{a}_2$	−1.9986	−1.9992	−1.9992	−1.9983	0.0010
$\hat{b}_2$	3.9999	4.0018	4.0018	3.9990	0.0012
$\hat{c}_2$	−0.0590	−0.1034	−0.1039	−0.0366	0.0752

the equations. This is what we would expect and is due to the blunder introduced to the u_s coordinate of the Point 4, which appears in these two equations, i.e., u_{t_4} and v_{t_4}.

Table 7 shows the results of the estimated parameters for four cases: 1) before introducing outlier to the observations, 2) after introducing the u_s coordinate of the Point 4 to have a blunder, 3) after removing the first outlier, and 4) after removing the second outlier. The corresponding estimated variance factors of the unit weight are $\hat{\sigma}_0^2 = 1.11$, 6.79, 3.26 and 1.01, respectively. Table 7 shows also the standard deviation of the parameters obtained from the covariance matrix $\boldsymbol{Q}_{\hat{\boldsymbol{x}}} = \hat{\sigma}_0^2 \left(\tilde{\boldsymbol{A}}^{\mathrm{T}} \boldsymbol{Q}_{\tilde{\boldsymbol{y}}}^{-1} \tilde{\boldsymbol{A}} \right)^{-1}$ for Case 4.

5. CONCLUDING REMARKS

The "data snooping" technique was applied to an errors-in-variables (EIV) model in the frame of the standard least squares theory. Our first impression might be that the data snooping is not applicable to the WTLS problem because of intrinsic property of EIV models in which the elements of the coefficient matrix $\boldsymbol{A}$ are also observed and hence affected by random/gross errors. Recently, *Amiri-Simkooei and Jazaeri (2012)* formulated the WTLS problem based on the well-known standard least squares theory. This formulation led us to apply the data snooping procedure to the WTLS problems. The proposed algorithm was shown to be simple in the concept and easy in the implementation. It is also flexible and yields reliable solution in comparison with the available algorithms to detect blunders in EIV models in which gross errors are in the vector of observations and/or in the coefficient matrix. The existing methods are applicable when only one outlier appears, either in the observation vector or in the coefficient matrix. The numerical results showed the high efficiency of the proposed algorithm to apply the usual data snooping techniques for identification of outlying equations. Outliers may then appear in the observation vector, in the coefficient matrix, or in both.

ALGORITHM 1

Iterative algorithm for weighted total least squares problem along with data snooping procedure:

- Choose small value for ε
- Set number of outlying equations $k = 0$
- Do for k
- Initialize $\hat{\boldsymbol{x}} = \hat{\boldsymbol{x}}^{(0)} = \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Q}_{\boldsymbol{y}}^{-1}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Q}_{\boldsymbol{y}}^{-1}\boldsymbol{y}$
- Set iteration counter $i = 0$
  - Do for i
  - Estimate $\hat{\boldsymbol{e}} = \boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{x}}$
  - Calculate $\boldsymbol{Q}_{\tilde{\boldsymbol{y}}} = \boldsymbol{Q}_{\boldsymbol{y}} + \left(\hat{\boldsymbol{x}}^{\mathrm{T}} \otimes \boldsymbol{I}_{\boldsymbol{m}}\right)\boldsymbol{Q}_{\boldsymbol{A}}\left(\hat{\boldsymbol{x}} \otimes \boldsymbol{I}_{\boldsymbol{m}}\right)$
  - Obtain $\tilde{\boldsymbol{E}}_{\boldsymbol{A}} = -\mathrm{vec}^{-1}\left(\boldsymbol{Q}_{\boldsymbol{A}}\left(\hat{\boldsymbol{x}} \otimes \boldsymbol{I}_{\boldsymbol{m}}\right)\boldsymbol{Q}_{\tilde{\boldsymbol{y}}}^{-1}\hat{\boldsymbol{e}}\right)$
  - Predict $\tilde{\boldsymbol{A}} = \boldsymbol{A} - \tilde{\boldsymbol{E}}_{\boldsymbol{A}}$ and $\tilde{\boldsymbol{y}} = \boldsymbol{y} - \tilde{\boldsymbol{E}}_{\boldsymbol{A}}\hat{\boldsymbol{x}}$
  - Increase i: $= i + 1$
  - Estimate $\hat{\boldsymbol{x}}^{(i)} = \left(\tilde{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{Q}_{\tilde{\boldsymbol{y}}}^{-1}\tilde{\boldsymbol{A}}\right)^{-1}\tilde{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{Q}_{\tilde{\boldsymbol{y}}}^{-1}\tilde{\boldsymbol{y}}$
  - Update $\hat{\boldsymbol{x}} := \hat{\boldsymbol{x}}^{(i)}$
  - While $\left\|\hat{\boldsymbol{x}}^{(i)} - \hat{\boldsymbol{x}}^{(i-1)}\right\| > \varepsilon$
  - End do over i
- Estimate $\hat{\sigma}_0^2 = \left(\hat{\boldsymbol{e}}^{\mathrm{T}}\boldsymbol{Q}_{\tilde{\boldsymbol{y}}}^{-1}\hat{\boldsymbol{e}}\right)\Big/(m-n)$
- Obtain $\boldsymbol{Q}_{\hat{\boldsymbol{x}}} = \left(\tilde{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{Q}_{\tilde{\boldsymbol{y}}}^{-1}\tilde{\boldsymbol{A}}\right)^{-1}$
- Estimate $\hat{\boldsymbol{e}} = \boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{x}}$ and $\boldsymbol{Q}_{\hat{\boldsymbol{e}}} = \hat{\sigma}_0^2\left(\boldsymbol{Q}_{\tilde{\boldsymbol{y}}} - \tilde{\boldsymbol{A}}\boldsymbol{Q}_{\hat{\boldsymbol{x}}}\tilde{\boldsymbol{A}}^{\mathrm{T}}\right)$
- Possibly test $\hat{\sigma}_0^2$ (if it is known under H_0)
- Find $\max\left(\left|w_{i,\mathrm{N}}\right|\right)$ or $\max\left(\left|w_{i,\mathrm{F}}\right|\right)$, $i = 1, \ldots, m$
- Reject H_0 if:
  - $\max\left(\left|w_{i,\mathrm{N}}\right|\right) > \mathrm{N}_{\alpha/2}(0,1)$ or
  - $\max\left(\left|w_{i,\mathrm{F}}\right|\right) > \sqrt{\mathrm{F}_{1-\alpha,1,df}}$
- Remove erroneous equation: $m = m - 1$
- Adopt matrices $\boldsymbol{A}$, $\boldsymbol{Q}_{\boldsymbol{y}}$, $\boldsymbol{Q}_{\boldsymbol{A}}$ and $\boldsymbol{y}$
- Increase number of outlying equations: $k = k + 1$
- Repeat while null hypothesis rejected
- End do over k

Acknowledgements: We would like to acknowledge useful comments of the Associate Editor, Ctirad Matyska, and two anonymos reviewers that improved the presentation of this paper. A few useful comments of Burkhard Schaffrin at Ohio State University on an earlier version of this manuscript are also acknowledged.

References

Amiri-Simkooei A.R., 2003. Formulation of L1 norm minimization in Gauss-Markov models. *J. Surv. Eng.*, **129**, 37–43.

Amiri-Simkooei A.R. and Jazaeri S., 2012. Weighted total least squares formulated by standard least squares theory. *J. Geod. Sci.*, **2**, 113−124.

Baarda W., 1968. A testing procedure for use in geodetic networks. *Publications on Geodesy*, **2(5)**, Netherlands Geodetic Commission, Delft, The Netherlands.

Davis T.G., 1999. Total least squares spiral curve fitting. *J. Surv. Eng.*, **125**, 159−176.

Felus Y.A., 2004. Application of total least squares for spatial point process analysis. *J. Surv. Eng.*, **130**, 126−133.

Fekri M. and Ruiz-Gazen A., 2004. Robust weighted orthogonal regression in the errors-in-variables model. *J. Multivariate Anal.*, **88**, 89−108.

Guo J., Ou J. and Wang H., 2007. Quasi-accurate detection of outliers for correlated observations. *J. Surv. Eng.*, **133**, 129−133.

Hekimoglu S., 1997. The finite sample breakdown points of conventional iterative outlier detection procedures. *J. Surv. Eng*., **123**, 15–31.

Hekimoglu S. and Erenoglu C., 2009. New method for outlier diagnostics in linear regression. *J. Surv. Eng.*, **135**, 85–89.

Kelly G., 1984. The influence function in the errors in variables problem. *Ann. Stat.*, **12**, 87−100.

Khodabandeh A. and Amiri-Simkooei A.R., 2011. Recursive algorithm for L1 norm estimation in linear models. *J. Surv. Eng*., **137**, 1–8.

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

Leick A., 1998. *Adjustment Computations*. Department of Surveying Engineering, University of Maine, Orono, ME.

Leick A., 2004. *GPS Satellite Surveying*. 3rd Edition. Wiley, New York.

Marshall J. and Bethel J., 1996. Basic concepts of L1 norm minimization for surveying applications. *J. Surv. Eng*., **122**, 168−179.

Pope A., 1976. *The Statistics of Residuals and Detection of Outliers*. NOAA Technical Report No.66, National Geodetic Survey, Rockville, MD.

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

Schaffrin B. and Felus Y., 2009. An algorithmic approach to the total least-squares problem with linear and quadratic constraints. *Stud. Geophys. Geod*., **53**, 1−16.

Schaffrin B. and Uzun S., 2011. Errors-in-variables for mobile algorithms in the presence of outliers. *Archives of Photogrammetry, Cartography and Remote Sensing*, **22**, 377–387, ISSN 2083-2214 (http://www.sgp.geodezja.org.pl/ptfit/wydawnictwa/krakow2011/APCRS%20vol.%2022%20pp.%20377-387.pdf).

Schaffrin B. and Uzun S., 2012. On the reliability of errors-in-variables models. *Acta et Commentationes Universitatis Tartuensis de Mathematica*, **16(1)**, 69–81.

Shen Y., Li B. and Chen Y., 2011. An iterative solution of weighted total least-squares adjustment. *J. Geodesy*, **85**, 229–238.

Teunissen P.J.G., 1985. The geometry of geodetic inverse linear mapping and nonlinear adjustment. *Publication on Geodesy*, **8(1)**, Netherlands Geodetic Commission, Delft, The Netherlands.

Teunissen P.J.G., 1990. Nonlinear least-squares. *Manuscripta Geodaetica*, **15**, 137–150.

Teunissen P.J.G., 2000a. *Adjustment Theory: an Introduction*. Series on Mathematical Geodesy and Positioning, Delft University Press, Delft University of Technology, Delft, The Netherlands (http://www.vssd.nl).

Teunissen P.J.G., 2000b. *Testing Theory: an Antroduction*. Series on Mathematical Geodesy and Positioning, Delft University Press, Delft University of Technology, Delft, The Netherlands (http://www.vssd.nl).

Tong X.H., Jin Y.M. and Li L.Y., 2011. An improved weighted total least squares method with applications in linear fitting and coordinate transformation. *J. Surv. Eng.*, **137**, 120–128.

Van Huffel S. and Vandewalle J., 1991. *The Total Least-Squares Problem: Computational Aspects and Analysis*. Volume 9 of Frontiers in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 374 pp.