Weighted total least squares applied to mixed observation model

A. R. Amiri-Simkooei*, S. Mortazavi and J. Asgari

This contribution presents the weighted total least squares (WTLS) formulation for a mixed errors-in-variables (EIV) model, generally consisting of two erroneous coefficient matrices and two erroneous observation vectors. The formulation is conceptually simple because it is formulated based on the standard least squares theory. It is also flexible because the existing body of knowledge of the least squares theory can directly be generalised to the mixed EIV model. For example, without any derivation, estimate for the variance factor of unit weight and a first approximation for the covariance matrix of the unknown parameters can directly be provided. Further, the constrained WTLS, variance component estimation and the theory of reliability and data snooping can easily be established to the mixed EIV model. The mixed WTLS formulation is also attractive because it can simply handle the two special cases of EIV models: the conditioned EIV model and the parametric EIV model. The WTLS formulation has been applied to three examples. The first two examples are simulated ones, the results of which are shown to be identical to those obtained by the non-linear Gauss-Helmert method. Further, the covariance matrix of the WTLS estimates is shown to closely approximate that obtained through a large number of simulations. The third is a real example of which two object points are photographed by three terrestrial cameras. Three scenarios are employed to show the efficiency of the proposed formulation on this last example.

Keywords: Weighted total least squares, Errors-in-variables model, Mixed observation model, Non-linear Gauss-Helmert model

Introduction

Total least squares (TLS), developed to handle errors-in-variables (EIV) models, have attained much attention recently. The TLS approach has been applied to linear system theory, audio processing, remote sensing and geodetic datum transformation (Teunissen, 1988; Markovsky and Van Huffel, 2007; Shen et al., 2011). The EIV model is widely used in geodetic data applications, such as straight line fitting and geodetic 2D and 3D coordinate transformations (Krystek and Anton, 2007; Fang, 2013; Fang, 2014c; Fang, 2015; Amiri-Simkooei et al., 2014).

The terminology ‘TLS’ appears only recently in the statistical and geodetic literature. This estimation principle, however, is not new and has a long history in the literature where it is also known as ‘orthogonal regression’, ‘Deming regression’ and ‘EIV’ (Markovsky and Van Huffel, 2007). The EIV model is widely used in geodetic data applications, such as straight line fitting and geodetic 2D and 3D coordinate transformations (Krystek and Anton, 2007; Fang, 2013; Fang, 2014c; Fang, 2015; Amiri-Simkooei et al., 2014).

The terminology ‘TLS’ appears only recently in the statistical and geodetic literature. This estimation principle, however, is not new and has a long history in the literature where it is also known as ‘orthogonal regression’, ‘Deming regression’ and ‘EIV’ (Markovsky and Van Huffel, 2007). The TLS method was originally introduced by Adcock (1878) for univariate problems in which errors in both dependent and independent variables were considered independent and the ratio of their variances was assumed to be one. Kummell (1879) made the method more general with the assumption of an arbitrary variance ratio. Their ideas remained unnoticed until 1937, when their work flourished again under Koopmans (1937) and Deming (1934) for linear regression analysis of economic time series. In statistics, the terminology of Deming regression is also used, which is named after W. Edwards Deming, who tried to find the line of best fit for a 2D dataset having errors in both x and y axes (Deming, 1931; Deming, 1934; Deming, 1964). Deming regression is a simple case of the orthogonal regression. Serious work on TLS originates from the work of Golub and Van Loan (1980) in mathematical literature in which they introduced the EIV models. For more information, we may refer to a series of text books by Kendall and Stuart (1973) and Kendall and Stuart (1979).

In geodetic literature, an EIV model treated as the 2D non-linear symmetric Helmert transformation was introduced by Teunissen (1988) in which the exact solution was given using a rotational invariant covariance structure. Later, many other researchers contributed to the solution of the EIV models. We may, for instance, refer to Felus (2004), Acar et al. (2006), Akylmaz (2007), Schaffrin and Wieser (2008), Schaffrin and Felus (2008, 2009), Fang (2011, 2013, 2014a, b, c, 2015), Tong et al. (2011), Shen et al. (2011), Xu et al. (2012, 2014), Xu and Liu (2014), Zhang et al. (2013) and Shi et al. (2015). These references mainly
solve a so-called parametric observation equation of EIV model. In many Geomatics applications, one may also deal with other models – conditioned model and/or mixed models, for instance, Teunissen (2000), Mikhail and Ackermann (1976). The WTLS methods are thus to be generalised to these models. Attempts have been made in the past to develop theories to the conditioned WTLS problem. Schaffrin and Wieser (2011) proposed a TLS method for the condition equations, where both the design matrix and observation vector are affected by homocedastic errors. Tong et al. (2014) proposed a method that can handle the conditioned EIV model in which arbitrary covariance matrices can be used.

Amiri-Simkooei and Jazaeri (2012) formulated the weighted total least squares (WTLS) problem based on the standard least squares theory. An alternative derivation was provided by Jazaeri et al. (2014). This formulation allows one to apply the existing body of knowledge of the least squares theory to the EIV models. Based on this formulation, a series of publications have appeared in the geodetic community. Amiri-Simkooei and Jazaeri (2013) applied the data snooping procedure to the EIV models. The theory of least squares variance component estimation was also applied to the EIV models by Amiri-Simkooei (2013). Amiri-Simkooei et al. (2014) applied the WTLS formulation to the linear regression problem with fully correlated coordinates. Amiri-Simkooei (2015) formulated the WTLS problem subjected to weighted and hard constraints. This contribution presents another application, namely, the WTLS formulation on a general mixed model for which an iterative algorithm is presented.

A few characteristics of the mixed WTLS formulation are highlighted. (1) The problem is formulated such that the covariance matrices $Q_\Lambda$ and $Q_\rho$ of the coefficient matrices have a general structure. (2) The mixed WTLS problem is formulated using the standard least squares theory. The existing body of knowledge of the least squares theory can directly be applied to the mixed EIV model. (3) The general formulation can handle two special cases, namely, the parametric EIV model and the conditioned EIV model. (4) The formulation is shown to be conceptually simple and practically efficient and have algorithmically low complexity. Numerical examples with different scenarios are employed to demonstrate the efficacy of the proposed algorithm in geodetic applications.

This paper is organised as follows. In the next section, we briefly review the WTLS method presented by Amiri-Simkooei and Jazaeri (2012), which is formulated in the standard least squares framework. We then generalise this algorithm to the mixed EIV model in a later section. Finally, we use our algorithm to solve three examples and show that the results are identical to those obtained using the non-linear Gauss-Helmert (NLGH) model. The conclusions are presented in the last section.

**WTLS**

Before we continue with the WTLS on the combined (or mixed) observation model, we briefly explain the theory of the WTLS on the parametric observation model. We explain the WTLS solution based on the standard least squares theory, which was introduced by Amiri-Simkooei and Jazaeri (2012). We will then generalise its formulation to a general mixed model.

In the WTLS problem, the Gauss-Markov model is replaced by an EIV model expressed as

$$y = (A - E_\Lambda)x + e_y$$  \hspace{1cm} (1)

with its stochastic properties characterised by

$$
\begin{bmatrix}
    e_y \\
    e_\Lambda
\end{bmatrix} := 
\begin{bmatrix}
    e_y \\
    \text{vec}(E_\Lambda)
\end{bmatrix} \sim \begin{bmatrix}
    0 \\
    0
\end{bmatrix} \sigma_0^2 
\begin{bmatrix}
    Q_y \\
    0
\end{bmatrix} \sigma_\Lambda^2 
\begin{bmatrix}
    0 \\
    Q_\Lambda
\end{bmatrix}
$$  \hspace{1cm} (2)

where $y$ is the $m$ vector of observations, $e_y$ is the $m$ vector of observational errors, $A$ is the $m \times n$ design matrix, $E_\Lambda$ is the corresponding $m \times n$ matrix of random errors, $x$ is the $n$-vector of unknown parameters, $D(e_y) = \sigma_0^2 Q_y$ and $D(e_\Lambda) = \sigma_\Lambda^2 Q_\Lambda$ are the corresponding symmetric positive-definite and non-negative dispersion matrices of size $m \times m$ and $mn \times mn$, respectively. In both expressions, $\sigma_0^2$ is the unknown variance factor of the unit weight, which, for the sake of simplicity is assumed to be $\sigma_0^2 = 1$.

The following target (Lagrange) function is used (Amiri-Simkooei and Jazaeri, 2012)

$$
\Phi := e_y^T Q_y^{-1} e_y + e_\Lambda^T Q_\Lambda e_\Lambda + 2x^T (y - Ax - e_y) + (x^T \otimes I_m) e_\Lambda
$$  \hspace{1cm} (3)

with $\lambda$ as an $m$-vector of unknown Lagrange multipliers and $I_m$ as an identity matrix of size $m$. $\otimes$ is the Kronecker product of two matrices. Owing to the possible singularity of $Q_\Lambda$, an arbitrary-generalised inverse $Q_\Lambda$ is used in the target function. One then has $Q_\Lambda Q_\Lambda^T Q_\Lambda = Q_\Lambda$ (see Bjørnhammar, 1973; Teunissen, 1985b). Then, the first partial derivatives of equation (3), with respect to the vectors $e_y$, $e_\Lambda$, $\lambda$, and $x$, are set to zero from which the predicted residual vector $\hat{e}_y$ and $\hat{e}_\Lambda$ are obtained. After using these equations and the first partial derivative of equation (3) with respect to $\lambda$, the Lagrange multipliers vector $\hat{\lambda}$ is estimated. Finally, the least squares estimate $\hat{x}$ is obtained as

$$
\hat{x} = (A - E_\Lambda)^T Q_\Lambda^{-1} (A - E_\Lambda)^{-1} (A - E_\Lambda)\hat{y}
$$  \hspace{1cm} (4)

where $Q_\Lambda = Q_y + (\hat{\lambda}^T \otimes I_m) Q_\Lambda (\hat{\lambda}^T \otimes I_m)$ approximates the actual unknown covariance matrix $Q_\Lambda = Q_y + (x^T \otimes I_m) Q_\Lambda (x^T \otimes I_m)$. Considering $A = A - E_\Lambda$ and $\hat{y} = y - E_\Lambda \hat{x}$, equation (4) is then rewritten as

$$
\hat{x} = (\hat{A}^T Q_\Lambda^{-1} \hat{A})^{-1} \hat{A}^T Q_\Lambda^{-1} \hat{y}
$$  \hspace{1cm} (5)

This is the WTLS formulation, which is similar to the standard least squares method. Further information is provided by Amiri-Simkooei and Jazaeri (2012). Equation (5) was also proved by Jazaeri et al. (2014) without the use of Lagrange multipliers. Considering the cross-covariances between $y$ and $A$, similar formulae to equation (5) have also been presented in a series of publications by Fang (2011, 2013, 2014a, 2014b, 2014c, 2015).
**WTLS on mixed observation model**

We now apply the WTLS formulation to the mixed observation model expressed as

\[
(B - E_B)(y - e_y) + (A - E_A)x = w - e_w
\]

or, for later use, as

\[
B y - E_B(y - e_y) - (B - E_B)e_y - E_A e_y + (A - E_A)x = w - e_w
\]

with its stochastic properties characterised by

\[
\begin{bmatrix}
 e_y \\
e_A \\
e_B \\
e_w \\
\end{bmatrix} =
\begin{bmatrix}
 e_y \\
\text{vec}(E_A) \\
\text{vec}(E_B) \\
 e_w \\
\end{bmatrix} =
\begin{bmatrix}
 0 & 0 & 0 & 0 \\
 0 & Q_A & 0 & 0 \\
 0 & 0 & Q_B & 0 \\
 0 & 0 & 0 & Q_w \\
\end{bmatrix}
\]

where \(y\) is the \(m\) vector of observations, \(e_y\) is the \(m\) vector of observational errors, \(A\) is the \(k \times n\) design matrix, \(E_A\) is the corresponding \(k \times n\) matrix of random errors, \(B\) is the \(k \times m\) coefficient matrix, \(E_B\) is the corresponding \(k \times m\) matrix of random errors, \(x\) is the \(n\) vector of unknown parameters, \(w\) is the \(k\) vector of possible additional observables, \(e_w\) is its corresponding random errors, \(D(e_y) = \sigma_y^2 Q_y\), \(D(e_A) = \sigma_A^2 Q_A\), \(D(e_B) = \sigma_B^2 Q_B\) and \(D(e_w) = \sigma_w^2 Q_w\) are the corresponding symmetric dispersion matrices of size \(m \times m\), \(kn \times kn\), \(km \times km\) and \(k \times k\), respectively. In all expressions, \(\sigma_i^2\) is the (unknown) variance factor of the unit weight (for now, we assume \(\sigma_i^2 = 1\)). Without the loss of the generality, for the sake of convenience, the covariance matrix in equation (8) is assumed to have a block diagonal structure; cross-covariance matrices among the above-mentioned four vectors are assumed to be absent.

Similar to equation (3), the target (Lagrange) function to be minimised is of the form

\[
\Phi := e_y^T Q_y^{-1} e_y + e_A^T Q_A^{-1} e_A + e_B^T Q_B^{-1} e_B + 2 \lambda^T (B y - E_B(y - e_y) - (B - E_B)e_y - E_A e_y + (A - E_A)x - w + e_w)
\]

with \(\lambda\) a \(k\) vector of unknown Lagrange multipliers. Owing to the possible singularity of \(Q_A\) and \(Q_B\), arbitrary-generalised inverses \(Q_A^L\) and \(Q_B^L\) are used in the target function. One then has \(Q_A Q_A^L Q_A = Q_A\) and \(Q_B Q_B^L Q_B = Q_B\). To solve the unknown parameters in equation (9), one has to obtain its first partial derivatives with respect to the vectors \(e_y, e_A, e_B, e_w, \lambda\) and \(x\) (see Petersen and Pedersen, 2012). For the sake of convenience, we obtain the partial derivatives with respect to their transpose vectors as

\[
\frac{1}{2} \frac{\partial \Phi}{\partial e_y^T} = Q_y^{-1} e_y - (B - E_B) \lambda = 0
\]

and [because \(\hat{E}_A \lambda = (\lambda^T \otimes I_k) \lambda\), with \(I_k\) an identity matrix of size \(k\)]

\[
\frac{1}{2} \frac{\partial \Phi}{\partial e_A^T} = Q_A^{-1} e_A - (\lambda^T \otimes I_k) \lambda = 0
\]

and [because \(\hat{E}_B \lambda = (\lambda^T \otimes I_k) \lambda\), with \(I_k\) an identity matrix of size \(k\)]

\[
\frac{1}{2} \frac{\partial \Phi}{\partial e_B^T} = Q_B^{-1} e_B - (\lambda^T \otimes I_k) \lambda = 0
\]

and [because \(\hat{E}_B (y - \hat{e}_y) = \hat{E}_B y = (\lambda^T \otimes I_k) \lambda\)],

\[
\frac{1}{2} \frac{\partial \Phi}{\partial e_w^T} = Q_w^{-1} e_w - (\lambda^T \otimes I_k) \lambda = 0
\]

and

\[
\frac{1}{2} \frac{\partial \Phi}{\partial \lambda^T} = Q_y^{-1} e_y - (B - E_B) \lambda = 0
\]

and

\[
\frac{1}{2} \frac{\partial \Phi}{\partial x^T} = 2 \lambda^T (B y - E_B(y - e_y) - (B - E_B)e_y - E_A e_y + (A - E_A)x - w + e_w) = 0
\]

where

\[
\hat{d} = (\hat{A}^T \otimes I_k) \lambda = (\lambda^T \otimes I_k) \lambda\]

and

\[
\hat{e}_w = -Q_w \hat{d}
\]

with \(\hat{d}\) being obtained as

\[
\hat{d} = (\hat{A}^T \otimes I_k) \lambda = (\lambda^T \otimes I_k) \lambda
\]

and

\[
\hat{e}_w = -Q_w \hat{d}
\]

with \(\hat{d}\) being obtained as

\[
\hat{d} = (\hat{A}^T \otimes I_k) \lambda = (\lambda^T \otimes I_k) \lambda
\]

and

\[
\hat{e}_w = -Q_w \hat{d}
\]

with \(\hat{d}\) being obtained as

\[
\hat{d} = (\hat{A}^T \otimes I_k) \lambda = (\lambda^T \otimes I_k) \lambda
\]

and

\[
\hat{e}_w = -Q_w \hat{d}
\]
is the TLS residuals of the mixed EIV model. Substituting equation (20) into equations (16)–(19) gives the predicted residual vectors as

\[
\begin{align*}
\tilde{e}_y &= -Q_eB^TQ_e^{-1}\hat{\varepsilon} \\
\tilde{e}_A &= -Q_A(\hat{\gamma}\otimes I_f)Q_e^{-1}\hat{\varepsilon} \\
\tilde{e}_B &= -Q_B(\hat{\gamma}\otimes I_f)Q_e^{-1}\hat{\varepsilon} \\
\tilde{e}_w &= Q_wQ_e^{-1}\hat{\varepsilon}
\end{align*}
\]

Equation (15), with equations (20)–(22) follows

\[
\tilde{A}^TQ_e^{-1}\tilde{A}\hat{x} = \tilde{A}^TQ_e^{-1}(w - B\hat{y} + E\tilde{e}_w)
\]

where \( \tilde{A} = A - \tilde{E}_A \). One solution to the preceding equation is: \( \hat{x} = (\tilde{A}^TQ_e^{-1}\tilde{A})^{-1}\tilde{A}^TQ_e^{-1}(w - B\hat{y} + E\tilde{e}_w) \). This formulation looks similar to the standard least squares formulation. But, the so-called normal matrix is not in general symmetric and positive definite. To make a symmetric positive-definite normal matrix, after a few simple mathematical operations, with \( A = A + \tilde{E}_A \), one obtains the estimated unknown parameters vector as

\[
\hat{x} = (\tilde{A}^TQ_e^{-1}\tilde{A})^{-1}\tilde{A}^TQ_e^{-1}v = N^{-1}u
\]

where \( N = \tilde{A}^TQ_e^{-1}\tilde{A} \) is the \( n \times n \) normal matrix, and \( u = \tilde{A}^TQ_e^{-1}v \). Further, we have

\[
\tilde{v} = w - \tilde{E}_A\hat{x} - B\hat{y} + E\tilde{e}_w = w - E\tilde{x} - B\hat{y} - E\tilde{e}_w
\]

This is a formulation in the framework of the standard least squares problem \( y = Ax + e \) in which \( A = A + \tilde{E}_A \) plays the role of the design matrix. \( Q_e \) plays the role of the covariance matrix \( Q_e \), and \( v = w - E\tilde{x} - B\hat{y} + E\tilde{e}_w = w - E\tilde{x} - B\hat{y} + E\tilde{e}_w \) plays the role of the vector \( y \). Therefore, equation (6) can be rewritten in the form of \( \tilde{v} = Ax + \tilde{e} \) in which \( \tilde{e} \) is the total residuals of the model defined as \( \tilde{e} = e_w - E\tilde{x} - B\hat{y} - E\tilde{e}_w \) and the remaining terms are compared with the standard linear model \( y = Ax + e \) in Table 1. We note that while \( Q_e = Q_{\tilde{e}} \) is the (unknown) covariance matrix of \( \tilde{v} = w - E\tilde{x} - B\hat{y} - E\tilde{e}_w \), \( Q_e \) is not the covariance matrix of \( \tilde{v} = w - E\tilde{x} - B\hat{y} - E\tilde{e}_w \); \( Q_e \) just approximates \( Q_{\tilde{e}} \).

Because this formulation is based on the standard least squares theory, the covariance matrix of the estimated parameters \( \hat{x} \) is obtained as

\[
Q_{\hat{x}} = (\tilde{A}^TQ_e^{-1}\tilde{A})^{-1}
\]

A first approximation of the preceding equation is

\[
Q_{\hat{x}} = (\hat{A}^T\hat{Q}_e^{-1}\hat{A})^{-1}
\]

Further, in the analogy with the standard least squares theory, the variance factor of the unit weight is estimated as

\[
\sigma_0^2 = \frac{\tilde{e}^TQ_e^{-1}\tilde{e}}{k - n}
\]

Using the above WTLS formulation, the existing body of knowledge of the least squares theory can be applied to the mixed EIV model \( \tilde{v} = Ax + \tilde{e} \). There is only one complication. We note that, in the formulation, \( x, \tilde{E}_A, \tilde{v} \) and \( \tilde{e} \) are unknown. The best one can do is to use their estimates/predictions instead of the actual unknown parameters. Therefore, from the implementation point of view, the above procedure should be solved through iterations in which all the terms involved, i.e. the design matrix, the covariance matrix and the observable vector are unknown random variables, also to be determined through the iterative procedure. Solving such non-linear problem will bias the least squares estimates along with their precision description (see Teunissen, 1985a; Teunissen, 1990). It is relevant to explain the difference between \( x \) and \( \hat{x} \) by the terminologies of actuality (what is ‘as it is’) and reality (what is ‘as we view it’). We have to live in the real world! The unknown actual \( x \) can best be replaced by its counterpart estimate \( \hat{x} \) in the real world (Table 1). The mixed EIV model can be implemented by the algorithm presented in Figure 1.

The general WTLS formulation of the mixed model allows one to handle different special cases. For example, this formulation considers solving the adjustment problems with the parametric and the conditioned EIV models. These two models are the special cases of the more general mixed model (Teunissen, 2000; Teunissen et al., 2004). We now consider these two cases for an EIV model.

**Special cases**

- **If** \( B = I \) and \( w = 0 \), \( E_w \) and \( e_w \) will be equal to zero because the matrix \( B \) and the vector \( w \) are constant and hence they are error-free following \( Q_{\tilde{e}} = 0 \) and \( Q_w = 0 \).
Under these assumptions, equation (6) is converted to the parametric or standard EIV model. We then have

\[ y - e_y = (A - E_A) \bar{x} \]  \hspace{1cm} (30)

In this case, one has \( Q_y = Q_y + (\bar{y}^T \otimes I_m) Q_A (\bar{y} \otimes I_m) \) and \( \bar{y} = y - E_A \bar{x} \), which are identical to \( Q_y \) and \( \bar{y} \) in Amiri-Simkooei and Jazaeri (2012), respectively. Equation (5) can then directly be followed.

- If \( A = 0 \), \( E_A = 0 \) and \( Q_A = 0 \), equation (6) reduces to the conditioned model as

\[ (B - E_B)(y - e_y) = w - e_w \]  \hspace{1cm} (31)

We thus have \( Q_y = Q_y + (\bar{y}^T \otimes I_k) Q_A (\bar{y} \otimes I_k) + BQ_e B^T \) and \( \bar{y} = y - B \bar{x} + E_B \bar{e}_y \). Finally, the presented formulation can solve the TLS adjustment of the conditioned equations.

**Numerical results and discussions**

To verify the efficacy of the above formulation, three case studies are considered. In the first and second example, we use simulated data set. The WTLS problem is solved on the mixed observation model in this case. In the third example, we solve a photogrammetric example presented by Mikhail and Ackermann (1976) using three scenarios and show the capabilities of the presented formulation.

**Example 1**

A simulated example is considered to compare our results with those of the classical mixed adjustment model of NLGH method. For this purpose, we define the actual unknown parameters vector as \( x = [2 \ 7]^T \). We then define the matrices \( A \) and \( B \) and the vectors \( y \) and \( w \).
assuming that all their elements contain random errors. These matrices/vectors are provided in Table 2. We consider the covariance matrices of these elements to be diagonal. Table 3 lists the diagonal elements (variances) of the covariance matrices of the simulated data.

We then perform the least squares adjustment with both methods (namely, our WTLS algorithm and the NLGH method) with a threshold of $\delta = 10^{-12}$. The estimated unknown parameter vector $\hat{x}$ along with its covariance matrix $Q_\hat{x}$ are then compared (Table 4). Both methods provide the same solution vector and covariance matrix, and only the convergence rates are different. Starting from the same initial value, to obtain the above-specified threshold $\delta$ for both methods, the NLGH method converges after six iterations while the proposed WTLS formulation solves the problem within four iterations. This highlights the superiority of our linearly structured WTLS algorithm over the NLGH method.

**Example 2**

We now use simulated data over 100,000 runs. For this purpose, the actual unknown parameter vector is defined as $x = [x_1, x_2]^T$. We then define $A$, $B$, $f$, and $w$ as the actual and error-free elements of the matrices and vectors $A$, $B$, $y$, and $w$. In each simulation run, for all of the matrices and vectors involved, we produce an error vector using ‘randn’ in MATLAB based on variances in Table 3. We then add them up to the error-free matrices and vectors and perform the adjustment using the proposed WTLS formulation, with the threshold of $\delta = 10^{-12}$. Table 5 shows the average of the estimated unknown parameters.

Two strategies are used to obtain the covariance matrix of the parameters. The first strategy is to directly use equation (28). The second strategy is used to estimate the covariance matrix of the parameters via the simulation process. For this purpose, we use the simulation over 100,000 independent runs. For each run, the line parameters $x_1$ and $x_2$ are estimated by the WTLS algorithm.

### Table 3 Variances for elements of matrices and vectors of weighted total least squares (WTLS) mixed observation model as presented in Table 2

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\sigma_{x_1}^2 \times 10^{-5}$</th>
<th>$\sigma_{x_2}^2 \times 10^{-5}$</th>
<th>$\sigma_{x_1}^2 \times 10^{-5}$</th>
<th>$\sigma_{x_2}^2 \times 10^{-5}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>0.05</td>
<td>0.06</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>$x_2$</td>
<td>0.07</td>
<td>0.08</td>
<td>0.07</td>
<td>0.08</td>
</tr>
</tbody>
</table>

### Table 4 Estimated unknown parameters $\hat{x}$ along with their standard deviations and correlation coefficient using data in Tables 2 and 3; non-linear Gauss-Helmert adjustment (middle column), weighted total least squares (WTLS) formulation presented in this paper (right column)

<table>
<thead>
<tr>
<th>Parameter/standard deviation</th>
<th>Non-linear Gauss-Helmert model (NLGH)</th>
<th>WTLS on mixed model (This paper)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>1.98391963</td>
<td>1.98391963</td>
</tr>
<tr>
<td>$x_2$</td>
<td>1.98391963</td>
<td>1.98391963</td>
</tr>
<tr>
<td>$\sigma_{x_1}$</td>
<td>0.03145580</td>
<td>0.03145580</td>
</tr>
<tr>
<td>$\sigma_{x_2}$</td>
<td>0.0346796</td>
<td>0.0346796</td>
</tr>
<tr>
<td>$\sigma_{x_1,x_2}$</td>
<td>0.00290695</td>
<td>0.00290695</td>
</tr>
</tbody>
</table>

### Table 5 Average of weighted total least squares (WTLS) estimate $\hat{x}$ along with its covariance matrix $Q_\hat{x}$ based on 100,000 simulated runs; two strategies are employed to compute $Q_\hat{x}$

<table>
<thead>
<tr>
<th>$\hat{x}_{\text{avg}}$</th>
<th>$Q_\hat{x}$ based on residual matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{x}$</td>
<td>$\left( A^TQ_\hat{x}^{-1}A \right)^{-1}$</td>
</tr>
<tr>
<td>$Q_\hat{x}$</td>
<td>$Q_\hat{x}$</td>
</tr>
<tr>
<td>$2000050374$</td>
<td>0.00343 $-0.00288$</td>
</tr>
<tr>
<td>$699918736$</td>
<td>0.00288 $0.00264$</td>
</tr>
</tbody>
</table>

The difference between the estimates and the actual parameters makes a $100,000 \times 2$ residual matrix as $E_\hat{x} = (\hat{x}_j - x_j) \cdot (\hat{x}_j - x_j)^T$, $j = 1:100,000$, where $\hat{x}_j$ and $x_j$ denote the estimated parameters of the $j$th run. The covariance matrix of the estimated parameters $\hat{x}$ is then estimated by $Q_\hat{x} = E_\hat{x}^T E_\hat{x} / (m - n)$, where $m = 100,000$, $n = 0$ (Teunissen and Amiri-Simkooei, 2008; Amiri-Simkooei, 2009). The results of these two strategies are also presented in Table 5. As observed, the results closely follow each other. This indicates that the first approximation of the covariance matrix of the parameters based on equation (28) is indeed a good approximation that may fulfill the requirements for many practical applications.

### Example 3

The third is a real example provided by Mikhail and Ackermann (1976). A similar example has also been solved using an EIV model of conditioned equations by Tong et al. (2014). Figure 2 shows a simplified problem of two object points A and B, which are photographed by three terrestrial cameras $S_1$, $S_2$, and $S_3$ with the principle distance of $f = 100$ mm. All the distances $l_1$, $l_2$, $l_3$, $l_4$, $l_5$, $l_6$, $y_1$, and $y_3$ are observed (see Table 6).

Based on the geometry, one can write six observation equations as follows:

\[
\begin{align*}
I_1 & - x_2 - f x_1 = 0 \\
I_2 & - f x_3 = 0 \\
I_3 & - f y_1 + f x_1 = 0 \\
I_4 & - f y_1 + f x_3 = 0 \\
I_5 & - f y_1 - f y_2 + f x_1 = 0 \\
I_6 & - f x_4 - f y_1 - f y_2 + f x_3 = 0
\end{align*}
\]
We now solve this problem using three scenarios (three cases). In the first case, we estimate the coordinates of both object points A and B. In the second case, we eliminate the coordinate of point A from the observation equations and estimate the coordinate of point B. In the last case, we eliminate the coordinates of both points A and B from the equations and solve the conditioned adjustment problem.

**First case**

To estimate the coordinates of the points A and B with the presented WTLS formulation, the mixed EIV model is \((B - E_B)(y - e_y) + (A - E_A)x = w - e_w\), with

\[
A = \begin{bmatrix}
-f & l_1 & 0 & 0 \\
0 & 0 & -f & l_2 \\
f & l_3 & 0 & 0 \\
0 & 0 & f & l_4 \\
f & l_5 & 0 & 0 \\
0 & 0 & f & l_6 \\
\end{bmatrix}, \quad B = \begin{bmatrix}
0 & 0 \\
0 & 0 \\
-f & 0 \\
-f & 0 \\
-f & 0 \\
-f & 0 \\
\end{bmatrix}
\]

and

\[
y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \quad w = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
\]

In this case, the matrix \(B\) and the vector \(w\) are error-free. We now estimate the coordinates of both points with the NLGH and WTLS methods with the threshold of \(\delta = 10^{-12}\). Table 7 provides the results. As observed, both methods yield the same values for the unknown parameters and their standard deviations. We, however, note that starting from the same initial values, the problem is converged to the final solution in nine and seven iterations for the NLGH and WTLS algorithms, respectively. Faster convergence rate of the WTLS compared to the NLGH has already been reported by Amiri-Simkooei and Jazaeri (2012). This is also verified for the application considered in this contribution. The estimated variance factor of the unit weight is \(\hat{\sigma}_0^2 = 0.822842\).

**Second case**

We may now reformulate the above equations. One can solve for the coordinates of point A using the first and third observation equations and substitute them to the other equations. Therefore, we can write \(x_1\) and \(x_2\) as follows:

\[
x_1 = l_2 y_2/l_1 + l_3 \quad \text{and} \quad x_2 = f y_1/l_4 + l_5\]

Substituting the above equations in equation (32) follows

\[
\begin{align*}
l_2 x_4 - f x_3 &= 0 \\
l_4 x_4 - f y_1 + f x_3 &= 0 \\
(l_3 - l_5) y_1 + (l_1 + l_3) y_2 &= 0 \\
l_6 x_4 - f y_1 - f x_2 + f x_3 &= 0
\end{align*}
\]

The coordinates of point B are the only unknown parameters of this problem. To solve this problem, the above equations can be formulated in the matrix form of a mixed EIV model as \((B - E_B)(y - e_y) + (A - E_A)x = w - e_w\), with

\[
A = \begin{bmatrix}
-f & l_2 \\
f & l_4 \\
0 & 0 \\
f & l_6 \\
\end{bmatrix}, \quad B = \begin{bmatrix}
0 & 0 \\
0 & 0 \\
-l_3 & l_1 + l_3 \\
-f & 0
\end{bmatrix}
\]

and

\[
y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad x = \begin{bmatrix} x_3 \\ x_4 \end{bmatrix}, \quad w = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
\]

In this case, both matrices \(A\) and \(B\) contain random errors. We solved this problem with the WTLS method and used the same threshold \(\delta = 10^{-12}\) (see Table 8). As can be seen, in both cases, the estimated coordinates of the point B are the same (Tables 7 and 8). The estimated variance factor of the unit weight is again \(\hat{\sigma}_0^2 = 0.822842\).

**Third case**

One can obtain alternative values for the coordinates of points A and B using the first four equations. Substituting these parameters into the last two equations, one obtains two condition equations as follow

\[
\begin{align*}
(l_3 - l_5) y_1 + (l_1 + l_3) y_2 &= 0 \\
(l_4 - l_6) y_1 + (l_2 + l_4) y_2 &= 0
\end{align*}
\]
The proposed WTLS method can also solve this adjustment problem. For this purpose, we have $A = 0$, because there are no unknown parameters left in these equations. Therefore, one can only solve this problem to obtain the least squares estimates of the observations. In this case, we have $(B - E_0)(y - e_0) = w - c_0$, with

$$B = \begin{bmatrix} l_1 - l_5 & l_1 + l_3 & 0 \\ l_4 - l_6 & l_2 + l_4 & 0 \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad w = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Again, this problem was solved by using two methods of the WTLS and NLGH methods with the threshold of $\delta = 10^{-12}$. Table 9 shows the estimated observations. The results show that both methods have the same values for the estimated observations. Therefore, the WTLS method proposed in this contribution can solve the conditioned EIV model as a special case of the mixed EIV model. We also note that the same results presented in Table 9 for the estimated observations can accordingly be obtained from the mixed model in Cases 1 and 2.

### Conclusions and outlook

We showed that WTLS problem on mixed observation model can be formulated in the standard least squares framework. Three examples were solved using the WTLS and the NLGH method. It was shown that the results of these two methods were identical. We, however, showed that the WTLS algorithm converges to the optimal solution faster than the NLGH method. Therefore, the WTLS algorithm has a faster convergence rate compared to the classical NLGH method. Another advantage of the WTLS formulation is that it has a linearly structured iterative formulation, which helps one to avoid the linearisation of the observation equations.

The mixed EIV problem was formulated in the standard least squares framework. Therefore, one can estimate the covariance matrix of the estimated unknown parameters by inverting the normal equations matrix. In addition, one can generalise the orthogonal projectors of the standard least squares from which estimates for the (total) residuals and observations (along with their covariance matrix) and the variance of the unit weight can directly be derived. Also, the constrained WTLS on mixed model, variance component estimation for a mixed EIV model and the theory of reliability and data snooping can easily be established to the mixed WTLS problem.

### Acknowledgement

Authors would like to acknowledge useful comments of two anonymous reviewers, which improved the presentation and quality of the paper.

### References


