Comparison of Structured and Weighted Total Least-Squares Adjustment Methods for Linearly Structured Errors-in-Variables Models

Yongjun Zhou[1]; Xinjian Kou[2]; Jonathan Li[3]; and Xing Fang[4]

Abstract: The paper focuses on a specific errors-in-variables (EIV) model named the linearly structured EIV (LSEIV) model in which all the random elements of design matrix are in a linear combination of an input vector with random errors. Two existing structured total least-squares (STLS) algorithms named constrained TLS (CTLS) and structured TLS normalization (STLN) are introduced to solve the LSEIV model by treating the input and output vectors as the noisy structure vectors. For comparison purposes, the weighted TLS (WTLS) method is also performed based on the partial EIV model. Approximated accuracy assessment methods are also presented. The plane fitting and Bursa transformation examples are illustrated to demonstrate the accuracy and computational efficiency performance of the proposed algorithms. It shows that the proposed STLS and WTLS algorithms can achieve the same accuracy if the dispersion matrix of the WTLS method is constructed based on the partial EIV model. **DOI: 10.1061/(ASCE)SU.1943-5428.0000190.** *© 2016 American Society of Civil Engineers.*

Author keywords: Errors-in-variables (EIV) model; Structured total least-squares (STLS); Weighted total least-squares (WTLS); Linearly structured.

Introduction

The total least-squares (TLS) method and its extensions have attracted widespread attention since the terminology of TLS was presented by Golub and Van Loan (1980). It is a popular approach for dealing with an overdetermined system of equations $\mathbf{A}\boldsymbol{\xi} \approx \mathbf{y}$ when both $\mathbf{A}$ and $\mathbf{y}$ are subject to errors, which is also known as an errors-in-variables (EIV) model. Essentially, the ordinary TLS (OTLS) does not depend on any prior stochastic assumption of the random variables. Nevertheless, it is proven to be a maximum likelihood estimator (MLE) when the errors of all elements in $[\mathbf{A}\ \mathbf{y}]$ are independent and identically distributed (i.i.d.) (Van Huffel and Vandewalle 1991) from statistic standpoints. However, the i.i.d. assumptions are mostly unrealistic in applications. Therefore, extended methods have been proposed to overcome the disadvantage of the OTLS (Markovsky and Van Huffel 2007). Until recently, a large number of extended TLS methods have been proposed and widely applied in a variety of fields, such as signal processing, system identification, computer vision, geodesy, and so forth. New progress on EIV models with unknown variance components have been published by Xu and Liu (2013, 2014).

[1]Assistant Professor, Schoo1 of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong Univ., Shanghai 200240, People's Republic of China (corresponding author). E-mail: yjzhou@sjtu.edu.cn

[2]Professor, Schoo1 of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong Univ., Shanghai 200240, People's Republic of China.

[3]Professor, Faculty of Environment, Univ. of Waterloo, Waterloo, Ontario, Canada N2L 3G1.

[4]Associate Professor, School of Geodesy and Geomatics, Wuhan Univ., Wuhan 430079, People's Republic of China.

Note. This manuscript was submitted on March 12, 2014; approved on February 18, 2016; published online on May 9, 2016. Discussion period open until October 9, 2016; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Surveying Engineering***, © ASCE, ISSN 0733-9453.

The TLS and its extensions have been widely applied and have become an active topic in the geodetic community. The most frequently discussed extension of TLS in geodesy is weighted TLS (WTLS), in which the full statistical knowledge of the design matrix is taken into account. Quite a number of approaches for solving the WTLS problem have appeared in recent years. Schaffrin and Wieser (2008) developed a WTLS solution to linear regression, in which the cofactor matrix of design matrix (notation $\mathbf{Q_A}$ is used for brevity) is restricted to a certain structure. Xu et al. (2012) extended the ordinary EIV model into a partial one, and presented a nonlinear TLS algorithm. Snow (2012) and Fang (2013) extended the WTLS algorithms in the case of cross-correlation between the design matrix and right-hand side (RHS) vector. Xu et al. (2014) presented a bias-corrected weighted least-squares (LS) method for a partial EIV model. Apart from the unconstrained cases, the constrained TLS (CTLS) or the WTLS solution has also been developed by a number of researchers (Mahboub and Sharifi 2013; Fang 2014a, b, 2015).

In contrast to the WTLS, the structured TLS (STLS) methods take advantage of the special structure of the design matrix when structured matrices are involved. The term STLS was first presented by De Moor (1993) in which $[\mathbf{A}\ \mathbf{y}]$ can be expanded as an affine matrix function of a noisy component vector and a group of fixed basis matrices. The structured matrices widely discussed are the Toeplitz or Hankel matrices (De Moor 1993; Lemmerling and Van Huffel 2001; Markovsky et al. 2005). A number of algorithms have been proposed to solve the STLS problems, in which CTLS (Abatzoglou et al. 1991), the Riemannian singular value decomposition (RiSVD) (De Moor 1993), and the structured total least norm (STLN) algorithms (Rosen et al. 1996; Van Huffel et al. 1996) are most frequently used. The extensions were also made to solve the nonlinear (Rosen et al. 1996; Lemmerling et al. 1996, 2002) or multivariate STLS problems (Kukush et al. 2005). In geodesy, the reformulation of a partial EIV model proposed by Xu et al. (2012) as a nonlinear adjustment problem automatically satisfies the requirement of structure preservation; thus it can be used to solve STLS problems. However, the STLS methods have not been as significantly investigated as the WTLS methods in geodetic literature. Schaffrin and

Felus (2008) developed a STLS procedure for similarity transformation by following a Cadzow algorithm, in which a characteristic matrix mapping a random vector to the design matrix is introduced to preserve the structure. Felus (2006) performed a STLN algorithm for similar and affine coordinate transformation by treating them as STLS problems. However, these studies are restricted to homoscedastic cases and particular applications. Schaffrin et al. (2012) modified the Cadzow step to generate the optimal solution with the desired structure. However, it does not consider the cross-correlation between $\mathbf{A}$ and $\mathbf{y}$. Fang (2014a) modified the WTLS algorithm by only considering the functionally independent errors to preserve the structure. Yet, the solution is based on converting the STLS problem into a WTLS framework.

The relationship between WTLS and STLS methods has also attracted the interests of geodetic researchers. In this paper, the classical STLS methods, including the CTLS algorithm and the STLN algorithm, are introduced but they are not familiar to geodesists yet. Based on these STLS methods, the quality assessment for the parameter estimates has been developed. The derivation presented in this paper on the unknown parameter accuracy is different from that by Xu et al. (2012), which obtained the first-order accuracy formula with the reformulation of EIV models for the first time. Furthermore, the partial EIV model proposed by Xu et al. (2012) is revisited. With their new reformulation, the structure is automatically kept; in addition it is shown that their WTLS method provides the identical results to the conventional STLS method.

There are two motivations for this paper: one is to establish a general STLS adjustment method along with an accuracy assessment method for a specific EIV model, and the other is to compare the proposed WTLS and STLS methods on accuracy and computational efficiency.

Linearly Structured EIV Model Formulation

Because the notations are quite different from different literature, the notations introduced by Schaffrin and Wieser (2008) are closely followed throughout this paper. Minor modifications are made if necessary. The OTLS can be formulated as

$$\min_{\mathbf{E_A},\mathbf{e_y},\boldsymbol{\xi}} \| \mathbf{E_A} \quad \mathbf{e_y} \|_F^2 \tag{1a}$$

$$\text{s.t.}:(\mathbf{A}-\mathbf{E_A})\boldsymbol{\xi}=(\mathbf{y}-\mathbf{e_y}) \tag{1b}$$

where $\|\|_\mathbf{F}$ denotes the Frobenius norm; $\mathbf{A}\in\mathbf{R}^{m\times p}$ $(m>p)$ is the design matrix affected by the random error matrix $\mathbf{E_A}$; $\boldsymbol{\xi}\in\mathbf{R}^p$ is the parameter vector; and $\mathbf{y}\in\mathbf{R}^m$ is the RHS vector corrupted by the random error vector $\mathbf{e_y}$. If one uses the F-norm, the weights of observations are not taken into account. Actually, the probable errors could be sampling errors, human errors, modeling errors, and instrument errors (Golub and Van Loan 1980), which may be heteroscedastic and/or correlated. Here, one only considers the random errors characterized by

$$\begin{pmatrix}\mathbf{e_A}\\ \mathbf{e_y}\end{pmatrix}=\begin{pmatrix}\text{vec}(\mathbf{E_A})\\ \mathbf{e_y}\end{pmatrix}\sim\mathbf{N}\left[\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\end{pmatrix},\ \sigma_0^2\begin{pmatrix}\mathbf{Q_A} & \mathbf{Q_{Ay}}\\ \mathbf{Q_{yA}} & \mathbf{Q_y}\end{pmatrix}\right] \tag{2}$$

where vec() is the operator stacking one column of a matrix underneath the previous one (Schaffrin and Wieser 2008); $\mathbf{N}(\boldsymbol{\mu},\mathbf{C})$ is the normal distribution with expectation vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$; σ_0^2 is an (unknown) variance component; $\mathbf{Q_y}$ is the cofactor matrix of $\mathbf{e_y}$, which is assumed to be a positive definite symmetric matrix; $\mathbf{Q_A}$ is the cofactor matrix of $\mathbf{e_A}$, which is mostly singular because of the fixed or repeated elements; and $\mathbf{Q_{Ay}}$ is the cross cofactor matrix of $\mathbf{e_A}$ and $\mathbf{e_y}$. If $\mathbf{Q_{Ay}}=\mathbf{0}$, it means $\mathbf{A}$ and $\mathbf{y}$ are uncorrelated.

By investigating the structure of design matrix in geodetic applications, such as line fitting, affine transformation and Bursa transformation, one can find that all the random elements in the design matrix consist of coordinates with errors. These error-contaminated coordinates can be equivalently treated as measurements with errors (Xu et al. 2012). In other words, the design matrix is the function of a measurement vector. In such cases, the well-known EIV model can be expressed as (Zhou et al. 2014)

$$\mathbf{y}\approx\mathbf{A}(\mathbf{x})\boldsymbol{\xi} \tag{3}$$

where $\mathbf{x}$ is also called an input vector or variable. The errors of $\boldsymbol{x}$ are the functionally independent errors (Xu et al. 2012; Fang 2014a). Provided that $\mathbf{x}$ is a vector of size q, $\mathbf{A}(\mathbf{x})$ is a function matrix that maps $\mathbf{x}\in\mathbf{R}^q\rightarrow\mathbf{A}\in\mathbf{R}^{m\times p}$. The elements of $\mathbf{A}$ may be constructed in a linear, quadratic, or even higher degree of $\mathbf{x}$ in applications. Here, the linear cases are restricted, such that

$$\mathbf{A}_{ij}(\mathbf{x})=z_{ij}^0+\mathbf{z}_{ij}^\mathrm{T}\mathbf{x};\ \forall i=1\cdots m;\ j=1\cdots p \tag{4}$$

where $\mathbf{A}_{ij}$ is the element in the ith row and jth column of $\mathbf{A}$; $\mathbf{z}_{ij}\in\mathbf{R}^q$ is a constant vector; and z_{ij}^0 is a fixed scalar. $z_{ij}=\mathbf{0}$ makes $\mathbf{A}_{ij}$ a fixed element; otherwise, it is random. The model [Eq. (3)] with the constraint [Eq. (4)] is a specific EIV model if the input and output vectors are corrupted by errors. Here, the concept *linear-structured* introduced by Beck and Ben-Tal (2005) is adopted and called a linear-structured EIV (LSEIV) model. It is a special form of the partial EIV model proposed by Xu et al. (2012). The corresponding stochastic model of LSEIV is given by

$$\mathbf{x}=\mathbf{x_0}+\mathbf{e_x};\ \mathbf{y}=\mathbf{y_0}+\mathbf{e_y} \tag{5}$$

$$\begin{pmatrix}\mathbf{e_x}\\ \mathbf{e_y}\end{pmatrix}\sim\mathbf{N}\left[\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\end{pmatrix},\ \sigma_0^2\begin{pmatrix}\mathbf{Q_x} & \mathbf{Q_{xy}}\\ \mathbf{Q_{yx}} & \mathbf{Q_y}\end{pmatrix}\right] \tag{6}$$

where $\mathbf{x_0}$, $\mathbf{y_0}$ are the true values of $\mathbf{x}$, $\mathbf{y}$; $\mathbf{e_x}$, $\mathbf{e_y}$ are the additive error vectors of $\mathbf{x_0}$, $\mathbf{y_0}$; σ_0^2 is the covariance of unit weight; $\mathbf{Q_x}$, $\mathbf{Q_y}$ are the cofactor matrices of $\mathbf{x}$ and $\mathbf{y}$; and $\mathbf{Q_{xy}}$ is the cross cofactor matrix of $\boldsymbol{x}$ and $\boldsymbol{y}$. It can be proven that the stochastic model [Eq. (6)] can lead to Eq. (2) in a LSEIV model.

Typically, the LSEIV model covers a wide range of geodetic applications. Here, only a few examples are given:

1. For data fitting (or regression), taking line fitting as an example, let there be m ($m \geq 2$) points, and the general LSEIV is given by

$$\mathbf{A}(\mathbf{x})=\begin{pmatrix}1 & x_1\\ \vdots & \vdots\\ 1 & x_m\end{pmatrix},\mathbf{y}=\begin{pmatrix}y_1\\ \vdots\\ y_m\end{pmatrix};\mathbf{x}=\begin{pmatrix}x_1\\ \vdots\\ x_m\end{pmatrix};\boldsymbol{\xi}=\begin{pmatrix}a\\ b\end{pmatrix} \tag{7}$$

2. Geodetic transformation, such as two-dimensional (2D) similarity transformation (Neitzel 2010) or affine transformation (Felus 2004; Tong et al. 2011), and the Bursa datum transformation (Yang 1999;Grafarend and Awange 2003), in which the input vector consists of the coordinates in the source coordinate system and the output vector, are the corresponding coordinates in the target coordinate system. Taking the affine transformation as an example, assuming that there are n ($n \geq 3$) points altogether, it then yields

$$\mathbf{A}(\mathbf{x})=\begin{pmatrix}1 & x_1{}' & y_1{}' & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & x_1{}' & y_1{}'\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 1 & x_n{}' & y_n{}' & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & x_n{}' & y_n{}'\end{pmatrix};\ \mathbf{x}=\begin{pmatrix}x_1{}'\\ y_1{}'\\ \vdots\\ x_n{}'\\ y_n{}'\end{pmatrix}\ \mathbf{y}=\begin{pmatrix}x_1\\ y_1\\ \vdots\\ x_n\\ y_n\end{pmatrix}$$

$$\boldsymbol{\xi}=\begin{pmatrix}a_0\\ a_1\\ a_2\\ b_0\\ b_1\\ b_2\end{pmatrix} \tag{8}$$

3. Dynamic data processing, such as linear time-invariant Kalman filtering (De Moor 1993) or the autoregressive moving average (ARMA) model parameter estimation, are more applications. Consider the moving average time series model (Markovsky et al. 2004) as an example

$$x_i\xi_1+x_{i-1}\xi_2=y_i;\ i=1,\ldots,m \tag{9}$$

where $\boldsymbol{\xi}=(\xi_1\ \xi_2)^T$ is the parameter vector; $\mathbf{x}=(x_1,\cdots,x_m)^T$ is the input time series; $\mathbf{y}=(y_1,\cdots,y_m)^T$ is the output time series; and x_0 is the initial condition. Suppose that the input, the output, and the initial condition are measured with errors, the problem can be formulated as a standard LSEIV model as

$$\mathbf{A}(\mathbf{x})=\begin{pmatrix}x_1 & x_0\\ x_2 & x_1\\ \vdots & \vdots\\ x_m & x_{m-1}\end{pmatrix};\boldsymbol{\xi}=\begin{pmatrix}\xi_1\\ \xi_2\end{pmatrix} \tag{10}$$

From the definition of the LSEIV model with some typical examples, one can find that the input and output vectors in geodesy mostly are 2D or three-dimensional (3D) coordinates. For convenience, measurements (or measurement vector) are used instead of variables. The design matrix is of full column rank. Although the previously mentioned problems belong to the LSEIV model, the structure of each problem is different. Moreover, constraints of parameters may also be imposed if necessary.

General WTLS Solution to LSEIV

As described previously, quite a few WTLS algorithms (Schaffrin and Wieser 2008; Shen et al. 2011; Snow 2012; Xu et al. 2012; Fang 2013) are available for solving the EIV model. In a LSEIV model, the errors of $[\mathbf{A}\ \mathbf{y}]$ are inherited from the measurement errors. If the coefficient matrix is structured, one must extract nonrepetitive random elements (Xu et al. 2012). Xu et al. (2012) have proposed a partial EIV model by selecting functionally independent random elements within the coefficient matrix

$$\mathbf{y}=(\boldsymbol{\beta}^{\mathrm{T}}\otimes\mathbf{I}_n)(\mathbf{h}+\mathbf{B}\bar{\mathbf{a}})+\mathbf{e_y} \tag{11a}$$

$$\mathbf{a}=\bar{\mathbf{a}}+\mathbf{e_a} \tag{11b}$$

where $\mathbf{h}$ is a deterministic constant vector with elements corresponding to the nonrandom elements of $\mathbf{A}$; and $\mathbf{B}$ is a given deterministic matrix that represents the characteristic of matrix $\mathbf{A}$. The m-dimensional vector $\mathbf{a}$ collects all the independent random elements of $\mathbf{A}$, and the true values of $\mathbf{a}$ are denoted by $\bar{\mathbf{a}}$. The weight matrix of $\mathbf{a}$ is usually denoted by $\boldsymbol{\omega}$.

In the partial EIV model, all the variables to be estimated can be written in the form of a compact vector $\boldsymbol{\beta}_{\mathbf{a}}=[\boldsymbol{\beta}^{\mathrm{T}};\bar{\mathbf{a}}^{\mathrm{T}}]$. Compared with the existing EIV models, the partial EIV model proposed by Xu et al. (2012) is advantageous in dealing with the structured coefficient matrix and even generalizing the inequality and equality constraints for both the unknown parameter vector and the independent random elements within the coefficient matrix. For the solution of the partial EIV model, one can either regard the partial EIV model as a standard nonlinear Gauss-Markov model to find the standard nonlinear LS solution or use the TLS solution developed by Xu et al. (2012). Therefore the formulation of the solution will not be presented here again.

STLS Formulation

In this section, the LSEIV model is formulated as a STLS problem and solved by the existing STLS algorithms. The STLS method was first introduced by De Moor (1993), in which $\boldsymbol{A}$ is known to have a linear or affine structure. Ignoring the weight matrix, a STLS problem is given by (Markovsky et al. 2004)

$$\min_{\mathbf{E_A},\mathbf{e_y},\boldsymbol{\xi}}\|\mathbf{E_A}\quad \mathbf{e_y}\|_{\mathbf{F}}^2 \tag{12a}$$

$$\text{s.t.: }(\mathbf{A}-\mathbf{E_A})\boldsymbol{\xi}=(\mathbf{y}-\mathbf{e_y}) \tag{12b}$$

$$[\mathbf{E_A}\ \mathbf{e_y}]\ \text{has the same structure as}\ [\mathbf{A}\ \mathbf{y}] \tag{12c}$$

Examining the definition of Eqs. (1a), (1b), and (12a)–(12c), one can find that it is the constraint [Eq. 12(c)] that distinguishes STLS from OTLS, allowing one to obtain an estimation of $\boldsymbol{\xi}$ when $[\mathbf{A}\ \mathbf{y}]$ are structured. Furthermore, different weights may possibly be assigned to each element. Beck and Eldar (2010) gave a clear description of the STLS problem, in which the design matrix $\mathbf{A}$ can be expressed as

$$\mathbf{A}=\sum_{i=1}^{q}\alpha_i\mathbf{A}_i;\ \alpha_i\in\mathbf{R},i=1,\cdots,q \tag{13}$$

where $\mathbf{A}_1,\ldots,\mathbf{A}_q\in\mathbf{R}^{m\times p}$ are the fixed structure (or basis) matrices; and $\alpha_1,\ldots,\alpha_q$ are the structure components containing errors, and typically q is much smaller than $m\times p$. Assuming $\Delta\alpha_1,\ldots\Delta\alpha_q$ are the unknown perturbations of the structure vector $\alpha_1,\ldots,\alpha_q$, it yields

$$\Delta\mathbf{A}=\sum_{i=1}^{q}\Delta\alpha_i\mathbf{A}_i,(i=1,\cdots,q) \tag{14}$$

where $\Delta\mathbf{A}$ stands for the true errors of $\mathbf{A}$. Eqs. (13) and (14) mean that $\mathbf{A}$ and $\Delta\mathbf{A}$ have the same structures. From the definition of STLS given earlier, an EIV model is a STLS problem only if the design matrix can be expanded as in Eq. (13). From the definition of Eqs. (3) and (4), the design matrix can be expressed as

$$\mathbf{A}=\mathbf{A}_0+x_1\mathbf{A}_1+\cdots+x_q\mathbf{A}_q \tag{15}$$

where

$$\mathbf{A}_0 = \begin{pmatrix} z_{11}^0 & \cdots & z_{1p}^0 \\ \vdots & \ddots & \vdots \\ z_{m1}^0 & \cdots & z_{mp}^0 \end{pmatrix} \quad \mathbf{A}_i = \frac{\partial \mathbf{A}}{\partial x_i} = \begin{pmatrix} z_{11}^i & \cdots & z_{1p}^i \\ \vdots & \ddots & \vdots \\ z_{m1}^i & \cdots & z_{mp}^i \end{pmatrix} (i = 1 \cdots q)$$

It is clear that all the elements in $\mathbf{A}_i$ ($i = 0 \ldots q$) are fixed values that consist of the basis matrices, and the input vector $\mathbf{x}$ is the noisy structured components. Moreover, $\mathbf{y}$ can also be verified to be a structured vector, although whether it is structured or not is not an important issue. Considering the errors in $\mathbf{x}$, it gives

$$\mathbf{A}_{ij}(\mathbf{x}_0 + \mathbf{e}_\mathbf{x}) = z_{ij}^0 + \mathbf{z}_{ij}^\mathrm{T}(\mathbf{x}_0 + \mathbf{e}_\mathbf{x}) = z_{ij}^0 + \mathbf{z}_{ij}^\mathrm{T}\mathbf{x}_0 + \sum_{k=1}^{q} z_{ij}^k \mathbf{e}_x^k \tag{16}$$

where z_{ij}^k and e_x^k are the kth element of the corresponding vectors. As a result, the error counterpart of the design matrix equals

$$\mathbf{E}_\mathbf{A} = e_x^1 \mathbf{A}_1 + \cdots + e_x^q \mathbf{A}_q \tag{17}$$

Eqs. (15) and (17) indicate that $\mathbf{A}$ and $\mathbf{E}_\mathbf{A}$ have the same structures. The CTLS (Abatzoglou et al. 1991), RiSVD (De Moor 1993), and STLN algorithms (Rosen et al. 1996; Van Huffel et al. 1996) are the three most widely used approaches in signal procession literature. Lemmerling (1999) had done an extensive survey and detailed analysis. In geodetic literature, Xu et al. (2012) also directly solved a STLS problem. Fang (2014a) proposed a WTLS-like STLS algorithm, which can be regarded as a new STLS algorithm. To solve a weighted LSEIV model, the CTLS and STLN algorithms are briefly outlined in the following subsections.

CTLS Algorithm Formulation

The CTLS algorithm was presented by Abatzoglou et al. (1991). Actually, the constraint referred by Abatzoglou et al. (1991) is not the so-called constraint in geodetic literatures; it stands for the model functions themselves. The CTLS formulation groups the error components of design matrix and RHS vector into a compact noisy vector $\mathbf{e} \in \mathbf{R}^k$ (k is the order of $\mathbf{e}$).The structure in CTLS is represented by a set of basis matrices $\mathbf{T}_i \in \mathbf{R}^{m\times k}$, in which the corrections of the ith column are defined by $\mathbf{T}_i\mathbf{e}$, such that the CTLS formulation becomes

$$\min_{\mathbf{e},\xi} \mathbf{e}^\mathrm{T}\mathbf{P}_\mathbf{l}\mathbf{e} \tag{18a}$$

$$\text{s.t.: } (\mathbf{A} - [\mathbf{T}_1\mathbf{e}, \cdots, \mathbf{T}_p\mathbf{e}])\xi = \mathbf{y} - \mathbf{T}_{p+1}\mathbf{e} \tag{18b}$$

where $\mathbf{e} = (\mathbf{e}_\mathbf{x}^\mathrm{T} \quad \mathbf{e}_\mathbf{y}^\mathrm{T})^\mathrm{T}$, $\mathbf{P}_\mathbf{l} = \mathbf{Q}_\mathbf{l}^{-1}$ is the weight matrix of vector $\mathbf{e}$. In terms of the matrices $\mathbf{T}_i$, it is the Jacobi matrix of the ith column in A with respect to the measurement vector $\mathbf{l} = (\mathbf{x}^\mathrm{T} \quad \mathbf{y}^\mathrm{T})^\mathrm{T}$, namely

$$\mathbf{T}_i = \frac{\partial \mathbf{A}(:,i)}{\partial \mathbf{l}} = \begin{pmatrix} \frac{\partial \mathbf{A}(1,i)}{\partial l_1} & \cdots & \frac{\partial \mathbf{A}(1,i)}{\partial l_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial \mathbf{A}(m,i)}{\partial l_1} & \cdots & \frac{\partial \mathbf{A}(m,i)}{\partial l_k} \end{pmatrix} (i = 1 \cdots p) \tag{19}$$

$$\mathbf{T}_{p+1} = \frac{\partial \mathbf{y}}{\partial \mathbf{l}} = \begin{pmatrix} \frac{\partial y_1}{\partial l_1} & \cdots & \frac{\partial y_1}{\partial l_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial l_1} & \cdots & \frac{\partial y_m}{\partial l_k} \end{pmatrix} \tag{20}$$

By introducing

$$\mathbf{B} = [\mathbf{A} \ \mathbf{y}];\ \boldsymbol{\eta} = (\xi^\mathrm{T} \quad -1)^\mathrm{T};\ \mathbf{H} = \left[-\sum_{i=1}^{p} \xi_i \mathbf{T}_i \quad \mathbf{T}_{p+1} \right] \tag{21}$$

Eq. (11b) is equal to

$$\mathbf{H}\mathbf{e} + \mathbf{B}\boldsymbol{\eta} = \mathbf{0} \tag{22}$$

Performing the Lagrange multiplier method from Eq. (18a) subject to Eq. (22), the CTLS problem can be changed into an unconstrained minimization problem as seen in Eq. (23) after some simple mathematic transformations (Abatzoglou et al. 1991)

$$\min_{\xi} \boldsymbol{\eta}^\mathrm{T}\mathbf{B}^\mathrm{T}(\mathbf{H}\mathbf{Q}_\mathbf{l}\mathbf{H}^\mathrm{T})^{-1}\mathbf{B}\boldsymbol{\eta} \tag{23}$$

To solve Eq. (23), Abatzoglou et al. (1991) proposed a Newton's method by using analytically derived gradients. Zhou and Pierre (2005) gave a simplified computation scheme of the algorithm. Some minor modifications are made in this paper to make the algorithm suitable for the LSEIV model. The computational procedures are outlined as:

1. Get $\mathbf{C}$ from the Cholesky factorization of the measurement covariance matrix, that is $\mathbf{C}\mathbf{C}^\mathrm{T} = \mathbf{Q}_\mathbf{l}$, let $\mathbf{G}_i = -\mathbf{T}_i\mathbf{C}^\mathrm{T}$ ($i = 1, \ldots, p+1$).
2. Use the OTLS solution as the initial value $\hat{\xi}_{(0)}$, give the convergence tolerance ε, and initialize the iteration counter $j := 0$.
3. Do the following calculation:
   $\mathbf{H} = [\sum_{i=1}^{p} \hat{\xi}_{i(j)}\mathbf{G}_i \quad -\mathbf{G}_{p+1}]$, ($\hat{\xi}_{i(j)}$) stands for the ith element of $\hat{\xi}$ in the jth iteration)

$$\mathbf{u} = (\mathbf{H}\mathbf{H}^\mathrm{T})^{-1}\mathbf{B}\hat{\boldsymbol{\eta}}_{(j)} \quad (\hat{\boldsymbol{\eta}}_{(j)} = (\hat{\xi}_{(j)}^\mathrm{T} \quad -1)^\mathrm{T})$$

$$\bar{\mathbf{A}} = \mathbf{A} - [\mathbf{G}_1\mathbf{H}^\mathrm{T}\mathbf{u}, \ldots, \mathbf{G}_p\mathbf{H}^\mathrm{T}\mathbf{u}];\ \bar{\mathbf{G}} = [\mathbf{G}_1^\mathrm{T}\mathbf{u}, \ldots, \mathbf{G}_p^\mathrm{T}\mathbf{u}];\ \mathbf{v} = \bar{\mathbf{A}}^\mathrm{T}\mathbf{u}$$

$$\mathbf{M} = -\bar{\mathbf{G}}^\mathrm{T}\mathbf{H}^\mathrm{T}(\mathbf{H}\mathbf{H}^\mathrm{T})^{-1}\bar{\mathbf{A}} - (\bar{\mathbf{G}}^\mathrm{T}\mathbf{H}^\mathrm{T}(\mathbf{H}\mathbf{H}^\mathrm{T})^{-1}\bar{\mathbf{A}})^\mathrm{T}$$

$$\mathbf{N} = [\bar{\mathbf{A}}^\mathrm{T}(\mathbf{H}\mathbf{H}^\mathrm{T})^{-1}\bar{\mathbf{A}}]^\mathrm{T} + \bar{\mathbf{G}}^\mathrm{T}[\mathbf{H}^\mathrm{T}(\mathbf{H}\mathbf{H}^\mathrm{T})^{-1}\mathbf{H} - \mathbf{I}_k]\bar{\mathbf{G}}$$

4. Update the parameter as

$$\hat{\xi}_{(j+1)} = \hat{\xi}_{(j)} + (\mathbf{M}\mathbf{N}^{-1}\mathbf{M} - \mathbf{N})^{-1}(\mathbf{v} - \mathbf{M}\mathbf{N}^{-1}\mathbf{v})$$

5. Set $j := j + 1$, repeat Steps 3–4 until $\hat{\xi}$ converges.

It should be noted that the Newton algorithm cannot guarantee the convergent to the global optimal solution due to the nonconvex nature of the STLS problem. However, one still assumes it can correctly converge by assuming that the starting point is close enough. The interested reader can refer to Abatzoglou et al. (1991) for a detailed description.

STLN Algorithm Formulation

The STLN is another widely used STLS algorithm. It is formulated to minimize the error in either the L_1, L_2 or the L-norm (Rosen et al. 1998). Only the L_2 normal STLN problem is considered to keep the cost function consistent with the LS criteria. The STLN uses a vector $\boldsymbol{\alpha}$ to represent the functionally independent errors of $\mathbf{E_A}$ and $\boldsymbol{\beta}$ to represent the functionally independent errors of $\boldsymbol{e_y}$ that are not already contained in $\boldsymbol{\alpha}$. The STLN is used to find the solution by minimizing the cost function of $\boldsymbol{\alpha}^T\mathbf{P}_\alpha\boldsymbol{\alpha} + \boldsymbol{\beta}^T\mathbf{P}_\beta\boldsymbol{\beta}$ (where $\mathbf{P}_\alpha$ and $\mathbf{P}_\beta$ are the weight matrices of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, respectively). It is obvious that $\mathbf{e_x}$ and $\mathbf{e_y}$ in a LSEIV model are $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. Because the STLN algorithm is only taking that $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are uncorrelated into account, improvements are made for general consideration. The improved STLN formulation is given by

$$\min_{\mathbf{e},\xi} \frac{1}{2}(\mathbf{eP_le}) \tag{24a}$$

$$\text{s.t.: } \mathbf{A}\xi - \mathbf{Ge_x} = \mathbf{y} - \mathbf{e_y} \tag{24b}$$

with $\mathbf{G} \in \mathbf{R}^{m\times q}$, $\mathbf{Ge_x}=\mathbf{E_A}\xi$; and the other notations are the same as those defined in Eqs. (18a) and (18b). The total residual vector $\mathbf{r}$ is defined as

$$\mathbf{r}(\tilde{\mathbf{e}},\hat{\xi}) = -\mathbf{y} + \tilde{\mathbf{e}}_\mathbf{y} + \mathbf{A}\xi - \mathbf{G}\tilde{\mathbf{e}}_\mathbf{x} + (\mathbf{A}-\mathbf{E_A})\delta\hat{\xi} \tag{25}$$

where $\tilde{\mathbf{e}}$ is the predicted value of the $\mathbf{e}$; $\hat{\xi}=\xi+\delta\hat{\xi}$, where ξ is an approximated value close to $\hat{\xi}$; and $\delta\hat{\xi}$ are the corresponding corrections. Considering $\mathbf{Ge_x}=\mathbf{E_A}\xi$, one gets

$$\mathbf{Ge_x}=\mathbf{E_A}\xi=(\xi^T\otimes\mathbf{I}_m)\mathbf{e_A} = (\xi^T\otimes\mathbf{I}_m)\mathbf{Je_x} \tag{26}$$

It leads to

$$\mathbf{G} = (\xi^T\otimes\mathbf{I}_m)\mathbf{J} \tag{27}$$

A couple of algorithms have been proposed (Van Huffel et al. 1996; Rosen et al. 1996; 1998) to solve the STLN problem. They can be classified into two strategies: transforming the constrained optimization problem into an unconstrained one (Xu et al. 2012; Fang 2014c) and approximating the constraint [Eq. (24b)] in each step iteratively by a linear approximation while leaving the cost function unchanged. Here is a brief introduction to the maximal weight algorithm presented by Van Huffel et al. (1996), in which the constraint was unified into the cost function by assigning a maximal weight, which makes the constrained problem into an unconstrained one. Eqs. (24a) and (24b) are equal to

$$\min_{\mathbf{e},\delta\xi} : \left\| \begin{matrix}\omega\mathbf{r}(\tilde{\mathbf{e}},\xi)\\ \mathbf{P}_\mathbf{l}^{\frac{1}{2}}\tilde{\mathbf{e}}\end{matrix}\right\|_\mathbf{F}^2 = \|\mathbf{U}\boldsymbol{\delta}+\bar{\mathbf{r}}\|_\mathbf{F}^2 \tag{28}$$

with

$$\mathbf{U}=\begin{pmatrix}\omega[\mathbf{G} \quad -\mathbf{I}_m] & -\omega(\mathbf{A}-\mathbf{E_A})\\ \mathbf{P}_\mathbf{l}^{\frac{1}{2}} & \mathbf{0}\end{pmatrix}$$

where ω is the maximal weight of Eq. (24b); and $\|\|_F$ is the Frobenius norm. Eq. (27) is a traditional LS adjustment problem. It can be solved by an iterative numerical method. The algorithm is summarized as

1. Choose a large weight ω and the convergence tolerance ε and set $\mathbf{E_A}=\mathbf{0}$, $\mathbf{e_x}=\mathbf{e_y}=\mathbf{0}$. Compute the initial value of ξ, and construct matrix $\mathbf{G}$ through the equation $\mathbf{Ge_x}=\mathbf{E_A}\xi$, $\hat{\mathbf{r}} = y - \mathbf{A}\xi$
2. Repeat
   a. Compute the corrections $\boldsymbol{\delta} = -(\mathbf{U}^T\mathbf{U})^{-1}\mathbf{U}^T\bar{\mathbf{r}}$;
   b. Set $\hat{\xi}:=\hat{\xi}+\delta\hat{\xi}$, $\tilde{\mathbf{e}}:=\tilde{\mathbf{e}}+\delta\tilde{\mathbf{e}}$; and
   c. Construct $\mathbf{E_A}$ from $\tilde{\mathbf{e}}_\mathbf{x}$ and $\mathbf{G}$ from $\hat{\xi}$, compute $\hat{\mathbf{r}}=\mathbf{y}-\tilde{\mathbf{e}}_\mathbf{y} - (\mathbf{A}-\mathbf{E_A})\hat{\xi}$ until ($\|\delta\hat{\xi}\| \le \varepsilon$).

From the two STLS algorithms given earlier, one can find that their difference lies in how to expand the LSEIV model. The CTLS algorithm formulated the STLS problem by expanding [$\mathbf{Ay}$] with multiplication of a noisy vector and the basis matrices from column to column and making STLS an unconstrained optimizing problem to be solved by a standard Gauss-Newton method. The STLN formulates the STLS problem by assigning a maximal weight to the model functions and treats the corrections of measurements and parameters as unknowns to be solved like an ordinary LS adjustment problem. The problem is that such a maximal weight may easily cause the overflow or truncated error in computation. Therefore, CTLS is preferable in solving a STLS problem.

Accuracy Assessment of STLS

The accuracy formula in Xu et al. (2012) is of the first-order approximation. The formula of the estimated accuracy and others also are of the same type, although these formulas might have different formulations. Because the STLS method actually is a nonlinear LS adjustment by treating the random vectors $\mathbf{x}$ and $\mathbf{y}$ as measurements, the variance component of unit weight gives

$$\hat{\sigma}_0^2 = \frac{\tilde{\mathbf{e}}^T\mathbf{P_l}\tilde{\mathbf{e}}}{m-p} \tag{29}$$

Once the estimates are achieved, Eq. (22) can be rewritten as

$$\hat{\mathbf{H}}\mathbf{e}+\tilde{\mathbf{A}}(\hat{\xi}+\delta\hat{\xi})-\tilde{\mathbf{y}}=\mathbf{0} \tag{30}$$

where $\hat{\mathbf{H}}=\left[\sum_{i=1}^p \hat{\xi}_i\mathbf{T}_i \quad -\mathbf{T}_{p+1}\right]$; $\tilde{\mathbf{A}}=\mathbf{A}+[\mathbf{T}_1\tilde{\mathbf{e}}_\mathbf{x},\cdots,\mathbf{T}_p\tilde{\mathbf{e}}_\mathbf{x}]$; $\tilde{\mathbf{y}}=\mathbf{y}+\tilde{\mathbf{e}}_y$; and $\delta\hat{\xi}$ is the perturbation vector of $\hat{\xi}$. Eq. (30) under the cost function of Eq. (18a) is a conditional adjustment with parameters. It gives

$$\delta\hat{\xi} = [\tilde{\mathbf{A}}^T(\hat{\mathbf{H}}\mathbf{Q_l}\hat{\mathbf{H}}^T)^{-1}\tilde{\mathbf{A}}]^{-1}\tilde{\mathbf{A}}^T(\hat{\mathbf{H}}\mathbf{Q_l}\hat{\mathbf{H}}^T)^{-1}(\tilde{\mathbf{y}}-\tilde{\mathbf{A}}\hat{\xi}) \tag{31}$$

With respect to $\tilde{\mathbf{y}}-\tilde{\mathbf{A}}\hat{\xi} = \hat{\mathbf{H}}\mathbf{e}$, through error propagation law, one can get

$$\mathbf{Q}_{\hat{\xi}}=[\tilde{\mathbf{A}}^T(\hat{\mathbf{H}}\mathbf{Q_l}\hat{\mathbf{H}}^T)^{-1}\tilde{\mathbf{A}}]^{-1} \tag{32}$$

As a result, Eqs. (29) and (32) provide a method for accuracy assessment of the CTLS method, and that of the STLN method can be established in a similar way.

To conclude this section, a simple comparison between the previously proposed methods is given. The STLS algorithm attempts to minimize the sum of the weighted squared errors of the noisy component vector. It is equal to the well-known LS criteria aimed at minimizing the sum of weighted measurement errors in geodesy. The WTLS seeks to minimize the sum of weighted squared errors of $\mathbf{A}$ and $\mathbf{y}$. It can be proven that the proposed WTLS and STLS algorithms in a LSEIV model can be converted into the same unconstrained optimization problem (Abatzoglou et al. 1991; Zhou et al. 2014). The differences only arise from numerical computation.

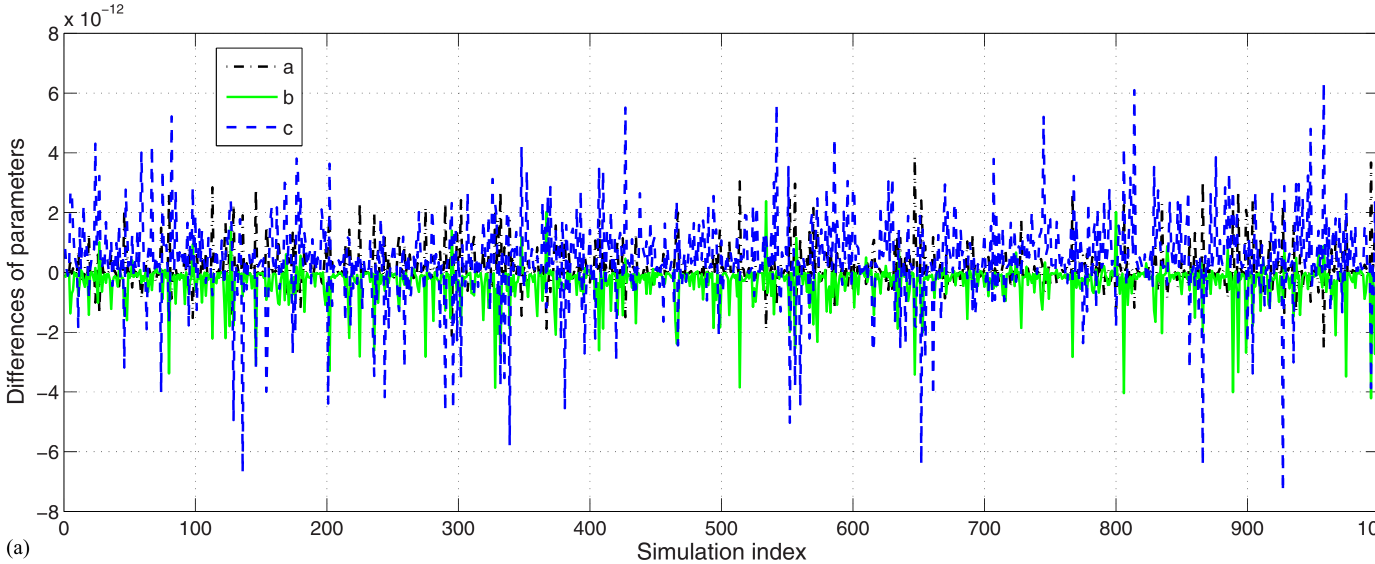
Table 1. RMSE of Different Estimators

Method	*a*	*B*	*c*
TLS	0.0055292410	0.0109221503	0.0526734212
WTLS	0.0053276355	0.0065547310	0.0376027566
CTLS-STLS	0.0053276355	0.0065547310	0.0376027566
STLN-STLS	0.0053276355	0.0065547310	0.0376027565

It should be stressed that both WTLS and STLS are nonconvex optimization problems; consequently, the devised algorithms are not guaranteed to converge to a global optimum. Appropriate initial values or extra improvements are needed to guarantee the validity.

Case Studies

Data experiments are performed in this section to examine the accuracy and efficiency performance of the proposed STLS algorithms for a LSEIV model, and comparisons are made with that of the WTLS methods as well.



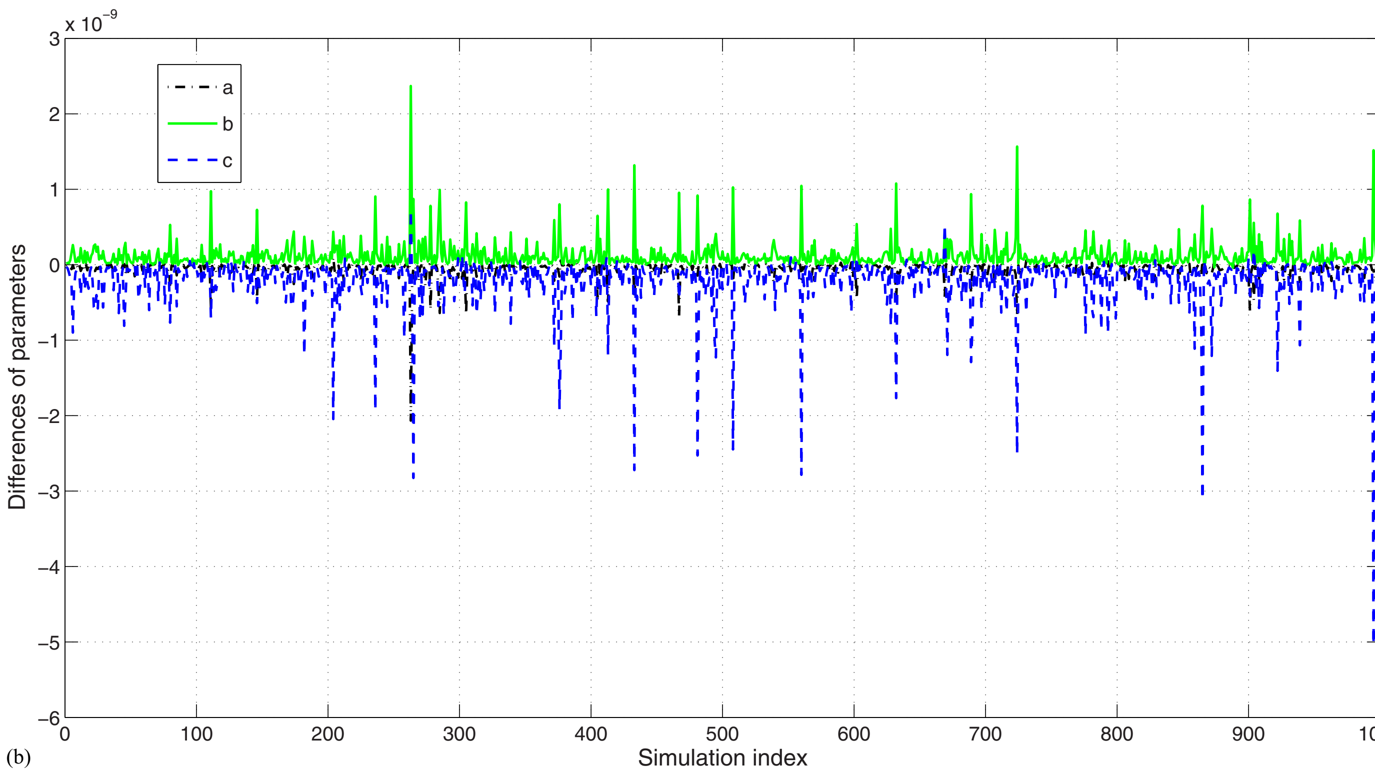


Fig. 1. Differences of parameters estimated by WTLS and STLS methods: (a) differences between the WTLS and CTLS algorithms; (b) differences between the WTLS and STLN algorithms

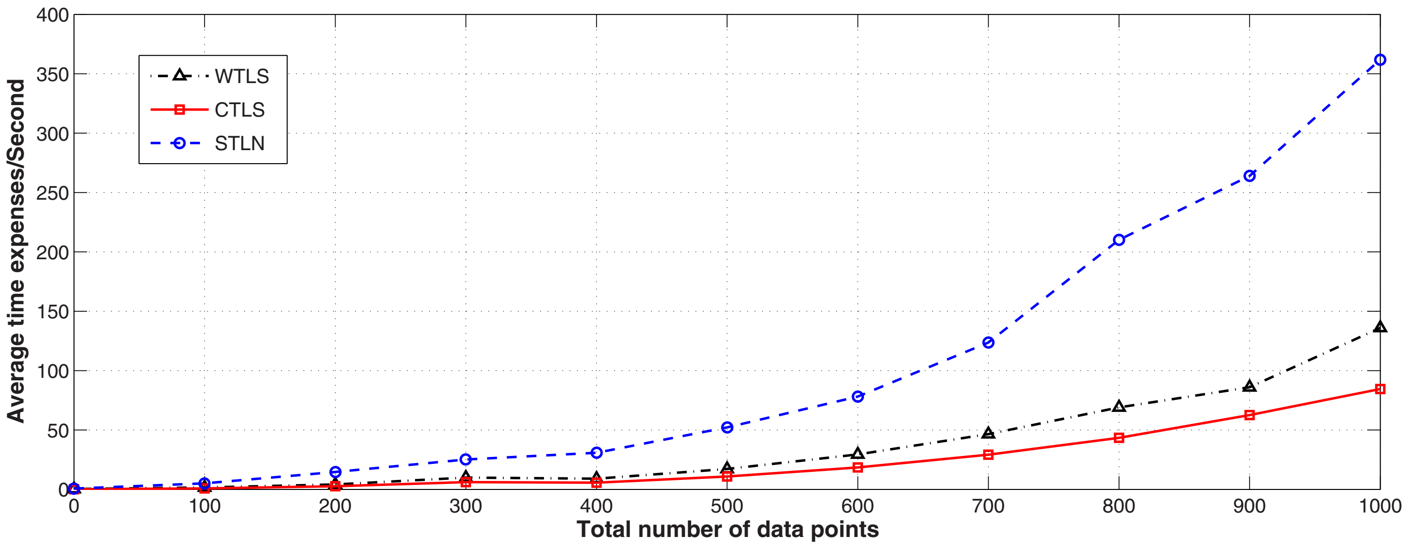


Fig. 2. Computational efficiency of different algorithms

Plane Fitting

As a first example, it is demonstrated that the implementation of the STLS algorithms for plane fitting are from simulated light detection and ranging (LiDAR) data. Assuming an area of 10×8 m^2 are scanned by an airborne LiDAR with a spatial resolution of 1 m, 80 ideal points at a regular grid with coordinates $\{(x_i^0, y_i^0) \in N;\ x_i^0 = (0 \ldots 9);\ y_i^0 = (0 \ldots 7)\}$ are generated. The noise-free height z_i^0 is given by

$$z_i^0 = ax_i^0 + by_i^0 + c \tag{33}$$

where a, b, and c are the parameters of plane with true values of a = 0.05, b = 0.2, and c = 100. The errors of different points are independent, but the errors within a point are correlated to simulate the correlations between **A** and **y**. The errors of an arbitrary point i are normally distributed with zero mean and covariance matrix given by

$$\mathbf{V}_i = \begin{pmatrix} \sigma_{x_i}^2 & \rho_1 \sigma_{x_i} \sigma_{y_i} & \rho_2 \sigma_{x_i} \sigma_{z_i} \\ \rho_1 \sigma_{x_i} \sigma_{y_i} & \sigma_{y_i}^2 & \rho_3 \sigma_{y_i} \sigma_{z_i} \\ \rho_2 \sigma_{x_i} \sigma_{z_i} & \rho_3 \sigma_{y_i} \sigma_{z_i} & \sigma_{z_i}^2 \end{pmatrix} \tag{34}$$

where $\mathbf{V}_i$ is a symmetric positive definite matrix with $\sigma_{x_i}^2 = \sigma_{y_i}^2 = 0.5^2$ and $\sigma_{z_i}^2 = 0.1^2$; and ρ_1, ρ_2, and ρ_3 are the correlation coefficients randomly generated within [−0.5 +0.5] at each simulation. OTLS and proposed WTLS, CTLS, and STLN algorithms are implemented for comparison purposes. The RMS error (RMSE) of each parameter defined by Eq. (35) from 1,000 trials is chosen as the criterion for accuracy comparison

$$\text{RMSE} = \pm \sqrt{\frac{\sum_{k=1}^{m} (\hat{\xi}_i^k - \xi_i^0)^2}{m}} \tag{35}$$

where $\hat{\xi}_i^k$ is the estimated value of the ith parameter in the kth trial; ξ_i^0 is the true value of the ith parameter; and m is the total amount of simulations, which is m = 1,000.

Table 2. Coordinates of the Local System

Site name	X (m)	Y (m)	Z (m)	σ^2 (m^2)
Solitude	4157222.543	664789.307	4774952.099	0.14330
Buoch Zeil	4149043.336	688836.443	4778632.188	0.15510
Hohenneuffen	4172803.511	690340.078	4758129.701	0.15030
Kuhlenberg	4177148.376	642997.635	4760764.800	0.14000
Ex Mergelaec	4137012.190	671808.029	4791128.215	0.14590
Ex Hof Asperg	4146292.729	666952.887	4783859.856	0.14690
Ex Kaisersbac	4138759.902	702670.738	4785552.196	0.12200

Table 3. Coordinates of WGS 84

Site name	X (m)	Y (m)	Z (m)	σ^2 (m^2)
Solitude	4157870.237	664818.678	4775416.524	0.01030
Buoch Zeil	4149691.049	688865.785	4779096.588	0.00380
Hohenneuffen	4173451.354	690369.375	4758594.075	0.00060
Kuhlenberg	4177796.064	643026.700	4761228.899	0.01140
Ex Mergelaec	4137659.549	671837.337	4791592.531	0.00680
Ex Hof Asperg	4146940.228	666982.151	4784324.099	0.00002
Ex Kaisersbac	4139407.506	702700.227	4786016.645	0.00410

Because the WTLS and STLS algorithms need iterative computations, all the iterations start from the solutions acquired by the OTLS method, and $\varepsilon = 10^{-10}$ is adopted by Schaffrin and Wieser (2008) and chosen as the convergence tolerance. In terms of the STLN algorithm, the maximum weight is set as $\omega = 10^6$. As the weight will be squared in the normal function formulation, it means the equivalent weight of the conditional function is at least 10^9 times than that of the measurements. All the algorithms are performed with the software *MATLAB* 6.5.

The RMSE of previously mentioned algorithms are listed in Table 1, and all the results are kept 10 digits after decimals for comparison purposes. It shows that the RMSE of the WTLS and two STLS algorithms are almost equal except for the OTLS method. The differences of parameters estimated by different methods in each trial are illustrated in Fig. 1. It shows the largest difference between WTLS and CTLS is less than 10^{-11}, which is within the given convergence tolerance. It can be concluded that the differences are mainly caused by numerical computation. The differences between STLN and WTLS are about a 10^{-8} order of magnitude, which is lager that that of CTLS. This is because the maximum weight assigned in computation is not the theoretical maximum, whereas increasing the maximum weight may deteriorate the ill-condition of the normal matrix. Therefore, it is of vital importance in choosing an appropriate maximum weight to balance the accuracy and numerical stability in a STLN algorithm.

To compare the computational efficiency of each algorithm, the total amount of points varying from 100 to 1,000 with a step of 100 is simulated. The average time cost of 100 trials at each quantity level is introduced for comparison. All the computations are performed by a Thinkpad X230 computer (Lenovo, Beijing), with Intel Core i5-3210M CPU and 2.5G RAM. The average time is shown in Fig. 2. The results indicate that the CTLS algorithm costs less time than those of the other two algorithms, and the STLN has the worst computational efficiency. The tendency implies the differences will increase significantly with the increasing amount of points because the corrections of measurements and parameters are solved simultaneously (which increase the order of normal matrix significantly) in a STLN algorithm, whereas the other two algorithms only solve the parameters in normal function and update the measurements afterward. The differences between CTLS and WTLS are not so significant because that they are mainly caused by the iterative formulas.

Bursa Transformation

Bursa model for 3D datum transformation is given by (Yang 1999):

$$\begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix} = \begin{pmatrix} X_i' \\ Y_i' \\ Z_i' \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} + \mu \begin{pmatrix} X_i' \\ Y_i' \\ Z_i' \end{pmatrix} + \begin{pmatrix} 0 & \omega_z & -\omega_y \\ -\omega_z & 0 & \omega_x \\ \omega_y & -\omega_x & 0 \end{pmatrix} \begin{pmatrix} X_i' \\ Y_i' \\ Z_i' \end{pmatrix} \quad (36)$$

where t_x, t_y, and t_z are the three translations; ω_x, ω_y, and ω_z are the three rotations and μ is the scale parameter; and $(X_i, Y_i, Z_i)^T$ and $(X_i', Y_i', Z_i')^T$ are the 3D coordinates of the source and target coordinates of the ith point, respectively. Let there be n ($n \geq 3$) point correspondences all together; thus Eq. (36) can be rewritten in a general LSEIV model as

$$\mathbf{y} = \begin{pmatrix} X_1 \\ Y_1 \\ Z_1 \\ \vdots \\ X_n \\ Y_n \\ Z_n \end{pmatrix}; \mathbf{A}(\mathbf{x}) = \begin{pmatrix} 1 & 0 & 0 & X_1' & 0 & -Z_1' & Y_1' \\ 0 & 1 & 0 & Y_1' & Z_1' & 0 & -X_1' \\ 0 & 0 & 1 & Z_1' & -Y_1' & X_1' & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & X_n' & 0 & -Z_n' & Y_n' \\ 0 & 1 & 0 & Y_n' & Z_n' & 0 & -X_n' \\ 0 & 0 & 1 & Z_n' & -Y_n' & X_n' & 0 \end{pmatrix}$$

$$\mathbf{x} = \begin{pmatrix} X_1' \\ Y_1' \\ Z_1' \\ \vdots \\ X_n' \\ Y_n' \\ Z_n' \end{pmatrix}; \xi = \begin{pmatrix} t_x \\ t_y \\ t_z \\ \mu \\ \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} \quad (37)$$

The coordinates of seven stations of the local and global reference systems (WGS 84) adopted from Grafarend and Awange (2003) are illustrated in Tables 2 and 3, respectively. Assuming that the three components of an arbitrary point are isotropic and independent, then $\sigma_{X_i}^2 = \sigma_{Y_i}^2 = \sigma_{Z_i}^2 = \sigma_i^2$, $\sigma_{X_iY_i} = \sigma_{X_iZ_i} = \sigma_{Y_iZ_i} = 0$. The corresponding errors of each point adopted from Grafarend and Awange (2003) are also shown in Tables 2 and 3.

From Eq. (37), $\mathbf{A}$ is a matrix of size $3n \times 7$, and the noisy component is $\mathbf{l} = (X_1', Y_1', Z_1' \cdots, X_n', Y_n', Z_n'; X_1, Y_1, Z_1, \ldots, X_n, Y_n, Z_n)^T \in \mathbf{R}^k$ ($k = 6n$). It is expanded in the Appendix to be solved by the proposed STLS algorithms. Because the translate parameters have a large quantity compared with the first example, the convergence tolerance $\varepsilon = 10^{-6}$ is chosen. In terms of the STLN algorithm, the maximum weight is still $\omega = 10^6$. The results are listed in Table 4.

The three translates are kept in four digits and the other parameters are kept in 10 digits after the decimal. The results in Table 4 indicate that the difference between the three translates estimated by the mentioned algorithms are within ±0.2

Table 4. Parameters and Their Errors of Different Methods

Methods	CTLS	STLN	WTLS	Nonlinear
tx/m	641.8393	641.8393	641.8392	641.8377
	±9.0327	±9.0327	±9.0327	
ty/m	68.4728	68.4728	68.4729	68.4743
	±10.5317	±10.5317	±10.5317	
tz/m	416.2155	416.2155	416.2154	416.2159
	±9.0495	±9.0495	±9.0495	
μ	1.0000056111	1.0000056111	1.0000056111	1.0000056112
	±0.0000010829	±0.0000010829	±0.0000010829	
ω_x/rad	-0.0000048371	-0.0000048371	-0.0000048371	-0.0000048373
	±0.0000014865	±0.0000014865	±0.0000014865	
ω_y/rad	0.0000043444	0.0000043444	0.0000043444	0.0000043441
	±0.0000016806	±0.0000016806	±0.0000016806	
ω_z/rad	0.0000047797	0.0000047797	0.0000047797	0.0000047798
	±0.0000013181	±0.0000013181	±0.0000013181	

mm, and the other parameters are completely the same. Comparing with the nonlinear solutions of Grafarend and Awange (2003), it is found that the differences of three translates are within ±2 mm, and the differences of scale and rotation parameters can reach the 10^{-9} level. In terms of the adjustment results, one finds that the TLS is basically of no use for geodetic coordinate transformation unless one pursues high precise results at the millimeter level. This agrees with the viewpoint given in Xu et al. (2014), especially when the signal-to-noise ratio is extremely large. It shows that the solution of WTLS and STLS methods can achieve a high accuracy close to the nonlinear ones. The error of each parameter estimated by the proposed accuracy assessment method is also equal. As a result, the proposed WTLS and STLS methods can achieve the same accuracy. In terms of computational efficiency, the proposed WTLS algorithm needs more than 500 iteration times compared with three times for CTLS and five times for STLN. Such situations are from the severe ill-condition $[\mathrm{Cond}(\mathbf{A}) = 2.77 \times 10^9]$ of the problem, except for the methodology differences described previously. In such cases, Tikhonov regularization (Tikhonov and Arsenin 1977), biased estimation (Hoerl and Kennard 1970) and/or truncated SVD (Xu 1998) methods can be implemented to overcome the ill-posedness.

Conclusions and Recommendations

The studies are restricted to a specific EIV model called the LSEIV model. The WTLS, CTLS, and STLN algorithms along with approximated accuracy assessment methods are introduced. Comparisons of accuracy and computational efficiency are made between the proposed STLS and WTLS methods. Through theoretic analysis and data experiments the conclusions maybe summarized as follows:

1. For a LSEIV model, the STLS can achieve the same accuracy as that of the WTLS method proposed in Xu et al. (2012). Furthermore, the formula of the variance factor is derived based on the STLS strategy.
2. To more fully investigate the potential of STLS adjustment methods for a range of geodetic applications, a number of areas need further studying, such as nonlinear structured EIV models, the ill-condition of STLS problems, the variance component estimation of STLS problems, and so forth.

Appendix. STLS Formulation of the Bursa Model

In the CTLS solution to the Bursa transformation parameters estimation, the design matrix $\mathbf{A}$ can be formulated as

$$\underset{m\times p}{\mathbf{A_0}} = \begin{pmatrix} 1&0&0&0&0&0&0\\ 0&1&0&0&0&0&0\\ 0&0&1&0&0&0&0\\ \vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots\\ 1&0&0&0&0&0&0\\ 0&1&0&0&0&0&0\\ 0&0&1&0&0&0&0 \end{pmatrix};\ \underset{m\times k}{\mathbf{T}_{p+1}} = \begin{pmatrix} \overbrace{0\ \cdots\ 0}^{m} & \overbrace{1\ \cdots\ 0}^{m}\\ \vdots\ \ddots\ \vdots & \vdots\ \cdots\ \vdots\\ 0\ \cdots\ 0 & 0\ \cdots\ 1 \end{pmatrix}$$

$$\underset{m\times k}{\mathbf{T}_1} = \underset{m\times k}{\mathbf{T}_2} = \underset{m\times k}{\mathbf{T}_3} = \mathbf{0};\ \underset{m\times k}{\mathbf{T}_i} = \begin{pmatrix} \underset{m\times m}{\mathbf{U}_i} & \underset{m\times m}{\mathbf{0}} \end{pmatrix} (i=4,5,6,7) \tag{38}$$

where

$$\underset{m\times m}{\mathbf{U}_4} = \begin{pmatrix} 1&0&0&0&0&0&0\\ 0&1&0&0&0&0&0\\ 0&0&1&0&0&0&0\\ \vdots&\vdots&\vdots&\ddots&\vdots&\vdots&\vdots\\ 0&0&0&0&1&0&0\\ 0&0&0&0&0&1&0\\ 0&0&0&0&0&0&1 \end{pmatrix};$$

$$\underset{m\times m}{\mathbf{U}_5} = \begin{pmatrix} 0&0&0&\cdots&0&0&0\\ 0&0&1&\cdots&0&0&0\\ 0&-1&0&\cdots&0&0&0\\ \vdots&\vdots&\vdots&\ddots&\vdots&\vdots&\vdots\\ 0&0&0&\cdots&0&0&0\\ 0&0&0&\cdots&0&0&1\\ 0&0&0&\cdots&0&-1&0 \end{pmatrix}$$

$$\underset{m\times m}{\mathbf{U}_6} = \begin{pmatrix} 0&0&-1&\cdots&0&0&0\\ 0&0&0&\cdots&0&0&0\\ 1&0&0&\cdots&0&0&0\\ \vdots&\vdots&\vdots&\ddots&\vdots&\vdots&\vdots\\ 0&0&0&\cdots&0&0&-1\\ 0&0&0&\cdots&0&0&0\\ 0&0&0&\cdots&1&0&0 \end{pmatrix};$$

$$\underset{m\times m}{\mathbf{U}_7} = \begin{pmatrix} 0&1&0&\cdots&0&0&0\\ -1&0&0&\cdots&0&0&0\\ 0&0&0&\cdots&0&0&0\\ \vdots&\vdots&\vdots&\ddots&\vdots&\vdots&\vdots\\ 0&0&0&\cdots&0&1&0\\ 0&0&0&\cdots&-1&0&0\\ 0&0&0&\cdots&0&0&0 \end{pmatrix}$$

In the STLN solution, the matrix $\mathbf{G}$ in Eq. (24*b*) is

$$\underset{m\times m}{\mathrm{G}} = \begin{pmatrix} \mu&\omega_z&-\omega_y&\cdots&0&0&0\\ -\omega_z&\mu&\omega_x&\cdots&0&0&0\\ \omega_y&-\omega_x&\mu&\cdots&0&0&0\\ \vdots&\vdots&\vdots&\ddots&\vdots&\vdots&\vdots\\ 0&0&0&\cdots&\mu&\omega_z&-\omega_y\\ 0&0&0&\cdots&-\omega_z&\mu&\omega_x\\ 0&0&0&\cdots&\omega_y&-\omega_x&\mu \end{pmatrix} \tag{39}$$

Acknowledgments

The authors thank the anonymous reviewers for their valuable comments and constructive suggestions on this manuscript. The work was partially supported by the National Natural Science Foundation of China (41274012, 41404005, and 41474006).

References

Abatzoglou, T. J., Mendel, J. M., and Harada, G. A. (1991). "The constrained total least squares technique and its applications to harmonic superresolution." *IEEE Trans. Signal Process.*, 39, 1070–1087.

Beck, A., and Ben-Tal, A. (2005). "A global solution for the structured total least squares problem with block circulant matrices." *SIAM J. Matrix Anal. Appl.*, 27(1), 238–255.

Beck, A., and Eldar, Y. C. (2010). "Structured total maximum likelihood: An alternative to structured total least squares." *SIAM J. Matrix Anal. Appl.*, 31(5), 2623–2649.
De Moor, B. (1993). "Structured total least squares and L_2 approximation problems." *Linear Algebra Appl.*, 188–189, 163–205.
Fang, X. (2013). "Weighted total least squares: necessary and sufficient conditions, fixed and random parameters." *J. Geod.*, 87(8), 733–749.
Fang, X. (2014a). "A structured and constrained total least-squares solution with cross-covariances." *Stud. Geophys. Geod.*, 58(1), 1–16.
Fang, X. (2014b). "On non-combinatorial weighted total least squares with inequality constraints." *J. Geod.*, 88(8), 805–816.
Fang, X. (2014c). "A total least squares solution for geodetic datum transformations." *Acta Geod. Geophys.*, 49(2),189–207.
Fang, X. (2015). "Weighted total least-squares with constraints: a universal formula for geodetic symmetrical transformations." *J. Geod.*, 89(5), 459–469.
Felus, Y. A. (2004). "Application of total least squares for spatial point process analysis." *J. Surv. Eng.*, 10.1061/(ASCE)0733-9453(2004)130:3(126), 126–133.
Felus, Y. A. (2006). "On linear transformations of spatial data using the structured total least norm principle." *Cart. Geogr. Inf. Sci.*, 33(3), 195–205.
Golub, G. H., and Van Loan, C. F. (1980). "An analysis of the total least squares problem." *SIAM J. Numer. Anal.*, 17(6), 883–893.
Grafarend, E. W., and Awange, L .J. (2003). "Nonlinear analysis of the three-dimensional datum transformation [conformal group C7(3)]." *J. Geod.*,77(1–2), 66–76.
Hoerl, A. E., and Kennard, R. W. (1970). "Ridge regression: Biased estimation for nonorthogonal problems." *Technometrics*, 12(1), 55–67.
Kukush, A., Markovsky, I., and Van Huffel, S. (2005). "Consistency of the structured total least squares estimator in a multivariate errors-in-variables model." *J. Stat. Plann. Infer.*, 133(2), 315–358.
Lemmerling, P. (1999). "Structured total least squares: Analysis, algorithms and applications." Ph.D. thesis, ESAT/SISTA, Katholieke Univ., Leuven, Netherlands.
Lemmerling, P., and Van Huffel, S. (2001). "Analysis of the structured total least squares problem for Hankel/Toeplitz matrices." *Numer. Algorithm*, 27(1), 89–114.
Lemmerling, P., Van Huffel, S., and De Moor, B. (1996). "Structured total least squares methods in signal processing." *Proc., 8th European Signal Processing Conf. (EUSIPCO'96)*, Trieste, Italy, 819–822.
Lemmerling, P., Van Huffel, S., and De Moor, B. (2002). "The structured total least-squares approach for non-linearly structured matrices." *Numer. Linear Algebra Appl.*, 9(4), 321–332.
Mahboub, V., and Sharifi, M. A. (2013). "On weighted total least-squares with linear and quadratic constraints." *J. Geod.*, 87(3), 279–286.
Markovsky, I., and Van Huffel, S. (2007). "Overview of total least-squares methods." *Signal Process.*, 87(10), 2283–2302.
Markovsky, I., Van Huffel, S., and Kukush, A. (2004). "On the computation of the multivariate structured total least squares estimator." *Numer. Linear Algebra Appl.*, 11(5–6), 591–608.
Markovsky, I., Van Huffel, S., and Pintelon, R. (2005). "Block-Toeplitz/Hankel structured total least squares." *SIAM J. Matrix Anal. Appl.*, 26(4), 1083–1099.
MATLAB 6.5 [Computer software]. MathWorks, Natick, MA.
Neitzel, F. (2010). "Generalization of total least-squares on example of unweighted and weighted 2D similarity transformation." *J. Geod.*, 84(12), 751–762.
Rosen, J. B., Park, H., and Glick, J. (1996). "Total least norm formulation and solution for structured problems." *SIAM J. Matrix Anal. Appl.*, 17(1), 110–126.
Rosen, J. B., Park, H., and Glick, J. (1998). "Structured total least norm for nonlinear problems." *SIAM J. Matrix Anal. Appl.*, 20(1), 14–30.
Schaffrin, B., and Wieser, A. (2008). "On weighted total least-squares adjustment for linear regression." *J. Geod.*, 82(7), 415–421.
Schaffrin, B., and Felus, Y. A. (2008). "On the multivariate total least squares approach to empirical coordinate transformation. Three algorithms." *J. Geod.*, 82(6), 373–383.
Schaffrin, B., Neitzel, F., Uzun, S., and Mahboub, V. (2012). "Modifying Cadzow's algorithm to generate the optimal TLS-solution for the structured EIV-model of a similarity transformation." *J. Geod. Sci.*, 2(2), 98–106.
Shen, Y. Z., Li, B. F., and Chen, Y. (2011). "An iterative solution of weighted total least-squares adjustment." *J. Geod.*, 85(4), 229–238.
Snow, K. (2012). "Topics in total least-squares adjustment within the errors-in-variables model: Singular cofactor matrices and prior information." Ph.D. thesis, Ohio State Univ., Columbus, OH.
Tikhonov, A. N., and Arsenin, V. I. (1977). *Solutions of ill-posed problems*, V. H. Winston, Columbia, MD.
Tong, X., Jin, Y., and Li, L. (2011). "An improved weighted total least squares method with applications in linear fitting and coordinate transformation." *J. Surv. Eng.*, 10.1061/(ASCE)SU.1943-5428.0000055, 120–128.
Van Huffel, S., and Vandewalle, J. (1991). *The total least squares problem: Computational aspects and analysis*, SIAM, Philadelphia.
Van Huffel, S., Park, H., and Rosen, J. B. (1996). "Formulation and solution of structured total least norm problems for parameter estimation." *IEEE Trans. Signal Process.*, 44(10), 2464–2474.
Xu, P. L. (1998). "Truncated SVD methods for discrete linear ill-posed problems." *Geophys. J. Int.*, 135(2), 505–514.
Xu, P. L., and Liu, J. N. (2013). "Variance components in errors-in-variables models: Estimability, stability and bias analysis." *Invited talk, VIII Hotine-Marussi Symposium on Mathematical Geodesy*, Rome.
Xu, P. L., and Liu, J. N. (2014). "Variance components in errors-in-variables models: Estimability, stability and bias analysis." *J. Geod.*, 88(8), 719–734.
Xu, P. L., Liu, J. N., and Shi, C. (2012). "Total least squares adjustment in partial errors-in-variables models: algorithm and statistical analysis." *J. Geod.*, 86(8), 661–675.
Xu, P. L., Liu, J. N., Zeng, W. X., and Shen, Y. Z. (2014). "Effects of errors-in-variables on weighted least squares estimation." *J. Geod.*, 88(7), 705–716.
Yang, Y. X. (1999). "Robust estimation of geodetic datum transformation." *J. Geod.*,73(5), 268–274.
Zhou, N., and Pierre, J. W. (2005). "Estimation of autoregressive parameters by the constrained total least square algorithm using a bootstrap method." *Proc., 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing*, 417–420.
Zhou, Y. J., Kou, X. J., Zhu, J. J., and Li, J. (2014). "A Newton algorithm for weighted total least-squares solution to a specific errors-in-variables model with correlated measurements." *Stud. Geophys. Geod.*, 58(3), 349–375.