



Total M_{split} estimation

Zbigniew Wiśniewski¹

Received: 10 July 2021 / Accepted: 9 September 2022 / Published online: 18 October 2022
© The Author(s) 2022

Abstract

M_{split} estimation is a method that enables the estimation of mutually competing versions of parameters in functional observation models. In the presented study, the classical functional models found in it are replaced by errors-in-variables (EIV) models. Similar to the weighted total least-squares (WTLS) method, the random components of these models were assigned covariance matrix models. Thus, the proposed method, named Total M_{split} (TM_{split}) estimation, corresponds to the basic rules of WTLS. TM_{split} estimation objective function is constructed using the components of squared M_{split} and WTLS estimation objective functions. The TM_{split} estimation algorithm is based on the Gauss–Newton method that is applied using a linear approximation of EIV models. The basic properties of the method are presented using examples of the estimation of regression line parameters and the estimation of parameters in a two-dimensional affine transformation.

Keywords Geodetic errors-in-variables · Weighted total least-squares · M_{split} estimation

1 Introduction

Wiśniewski (2009) proposed a method for estimating parameters in split functional models of geodetic observations. Such a split occurs when two functional models that differ from each other in terms of mutually competing versions of the same parameter can correspond to a single observation. For example, in a network deformation analysis carried out based on an aggregate set of observations obtained during two measurement epochs, any given observation from this set corresponds to one of the following models: a functional model of observations from the first measurement epoch or a functional model of observations obtained during the second measurement epoch. Another example concerns the sets containing outliers. In that case, any given observation from this set can be a “good” observation or a wrong observation with a functional model appropriate for it.

The proposed method called M_{split} estimation assumes that if a particular observation already occurs, it brings two mutually competing pieces of f -information (Jones and Jones 2000) determined in relation to two versions of the same parameter (Wiśniewski 2009, 2010). M_{split} estimators of

these versions are the quantities that minimise the aggregate information being the product of competing information. Similar assumptions are also adopted in the maximum likelihood method (ML-method) (e.g. Rao 1973; Wiśniewski 2017). However, this method does not allow the existence of several versions of the parameters in a functional model relating to the same observation. From this perspective, M_{split} estimation can be regarded as a particular kind of development of the ML method. In the absence of competing versions of the parameter, M_{split} estimators become ML-estimators. Since the study conducted by Huber (1964), a generalisation of the ML-method has been very popular, known as M-estimation, in which f -information is replaced by certain arbitrary functions. A similar substitution can also be observed in M_{split} estimation. M_{split} estimation based on L1 norm condition was developed to take advantage of this possibility (Wyszkowska and Duchnowski 2019).

The general theory of M_{split} estimation was developed without detailed assumptions about probabilistic observation models, which enables the creation of M_{split} estimation varieties corresponding to specific models of this nature. The most commonly accepted probabilistic model of geodetic observations is the normal distribution. The family of normal distributions corresponds to the basic variant of M_{split} estimation called “squared M_{split} estimation”. This variant of M_{split} estimation can be regarded as a particular type of expansion of the least-squares (LS) method (Wiśniewski 2009). In

✉ Zbigniew Wiśniewski
zbyszewk@uwm.edu.pl

¹ Department of Geodesy, Faculty of Geoengineering,
University of Warmia and Mazury, 10-719 Olsztyn, Poland

the absence of competing parameter versions, squared M_{split} estimators become LS-estimators. Whenever M_{split} estimation is mentioned further on, this will indicate this particular variant.

M_{split} estimation was applied *inter alia* in the analysis of geodetic network deformation (Duchnowski and Wiśniewski 2012, 2014; Zienkiewicz 2014; Zienkiewicz et al. 2017; Wiśniewski and Zienkiewicz 2016). In these problems, M_{split} estimation is particularly effective in identifying stable potential reference points (PRPs) (Nowel 2019). Janowski and Rapiński (2013) applied M_{split} estimation in 3D modelling, primarily for the detection of surface structures (e.g. roof planes) of engineering structures. The modelling was carried out based on laser scanning data. A similar problem is also analysed in a study by Janicka et al. (2020), where M_{split} estimation was proposed as a means of detecting and determining the displacements of adjacent planes. Laser scanning data also provided the basis for determining the terrain profiles using M_{split} estimation (Błaszczak-Bąk et al. 2015; Wyszowska et al. 2021).

M_{split} estimation can provide an alternative to M -estimation which is robust to gross errors. The possibility of such applications was indicated in the following studies (Wiśniewski 2009, 2010; Yang et al. 2010; Ge et al. 2013; Janicka and Rapinski 2013; Amiri-Simkooei et al. 2017). Wiśniewski and Zienkiewicz (2021a, b) demonstrated that with properly established, competitive functional models, the robustness of M_{split} estimators to gross errors is their inherent feature. The robustness of these estimators in a wider context (e.g. to poorly chosen models) was analysed in detail by Duchnowski and Wiśniewski (2019, 2020).

In both the traditional models and split functional models constructed on their basis (in M_{split} estimation), it is assumed that only observations are affected by random errors. Currently, an errors-in-variables (EIV) model, in which design matrix elements are also affected by random errors, is applied in many geodetic problems. For example, this model was applied in geodetic datum transformation (Teunissen 1988; Davis 1999; Acar et al. 2006; Akyilmaz 2007; Schaffrin and Felus 2008; Mahboub 2012; Fang 2015; Aydin et al. 2018; Mercan et al. 2018)) as well as in remote sensing (Felus and Schaffrin 2005), in a function approximation (Wang and Zhao 2019), in linear regression (Schaffrin and Wieser 2008; Amiri-Simkooei and Jazaeri 2012; Zeng et al. 2018; Lv and Sui 2020) and in the least-squares collocation (Schaffrin 2020; Wiśniewski and Kamiński 2020). The effect of the random design matrix on the weighted LS estimate is presented in Xu et al. (2014). This paper also proposed a bias-corrected weighted LS estimate for the EIV model. A developed EIV stochastic model and an estimation of the model's components (using, *inter alia*, the MINQUE method) are presented in Xu and Liu (2014).

The estimation of parameters in functional models extended to the EIV form is most commonly carried out using the total least-squares (TLS) method. The optimisation problem of this method as well as its solution based on the singular value decomposition (SVD) was presented by Golub and Loan (1980). TLS using the SVD procedure was developed and adapted to geodetic purposes as well (e.g. Felus 2004; Akyilmaz 2007; Schaffrin and Felus 2008). Another way to solve the TLS optimisation problem, based on a nonlinear Lagrange function, is proposed in Schaffrin et al. (2006).

In the practical applications of the TLS method, besides having effective algorithms at one's disposal, the possibility for taking into account random weights of EIV model components is also very important. The basic solutions in this regard were presented by van Huffel and Vandewalle (1991), who established the generalised total least-squares (GTLS) method. On the other hand, Schaffrin and Wieser (2008) proposed an expansion of the TLS method, in which weights were derived from the adopted covariance matrix models (stochastic models). In the method proposed in the cited study, called "the weighted total least-squares (WTLS) method", stochastic models can apply to both observation vectors and the vectors created from random errors affecting the design matrix.

WTLS is still developed and analysed. For example, Fang (2013) analysed necessary and sufficient conditions for WTLS optimality. Amiri-Simkooei (2017, 2018) presented the theory behind the constrained weighted total least-squares (CWTLS) method. The WTLS optimisation problem was also formulated and solved using a second-order approximation function (Wang and Zhao 2019). Due to their nonlinear nature, the WTLS estimators are biased. Bias-corrected versions of these estimators are presented in studies by (Xu et al. 2012; Tong et al. 2015). Moreover, an important problem in WTLS theory and practice is the assessment of the accuracy of the determined estimators. What might be helpful in this regard are the strategies for determining the covariance matrix of WTLS estimates (Amiri-Simkooei et al. 2016) and methods for estimating the variance components in EIV models (Xu and Liu 2014).

In the optimisation problem of the WTLS method based on the Lagrange approach, the objective function is minimised with the conditions defined by the nonlinear EIV model. An iterative algorithm to solve this problem was proposed in (Schaffrin and Wieser 2008). Shen et al. (2011), based on Newton–Gauss algorithm of nonlinear LS adjustment (Pope 1974), proposed another iterative method for solving WTLS problems, which is easier in practical applications. In this model, a nonlinear EIV model is replaced with a linear approximation, which significantly facilitates the organisation of a corresponding computational algorithm.

The origin of TLS or WTLS estimation is the LS-method which is neutral for all observations. The WTLS estimators'

lack of robustness to gross errors is, therefore, an inherent feature, which may restrict the scope of the practical application of these estimators. Therefore, a robust estimation of EIV model parameters is of interest to many authors. For example, Wang et al. (2016) proposed a robust total least-squares (RTLS) method in which the robustness of WTLS estimators was obtained by means of the application of weight functions adopted in robust M-estimation. Another proposal, based on the least trimmed squares (LTS) method, was presented by Lv and Sui (2020). In that method, the authors used the inherent robustness of estimators minimising the sum of the squared orthogonal errors. They called the LTS version adjusted to EIV models “total least trimmed squares” (TLTS).

The current study will apply the EIV model in a basic M_{split} estimation variant. When referring to the WTLS method, models of covariance matrices of random components of this model will also be taken into account. According to the basic principles of M_{split} estimation, it will be assumed that the parameters in EIV model will have two mutually competing versions (which consequently leads to the split of this model). The objective function of the proposed method, called the “Total M_{split} (TM_{split}) estimation”, will be created through the application of the Lagrange approach (Schaffrin and Wieser 2008) using the approach adopted in (Shen et al. 2011), i.e. the split EIV models will be replaced with their linear approximations.

The paper is organised as follows. As the proposed method is an expansion of M_{split} estimation, which takes into account the basic assumptions used in the WTLS method, it appears necessary to review the theoretical foundations of both these methods. These foundations, set in the context relevant to this study, are provided in Sect. 2. The theory behind TM_{split} estimation and its algorithm are provided in Sect. 3. In Sect. 4, examples of the method application will be provided. TM_{split} estimation will be applied to estimate parameters in competing bias models (Sect. 4.1). The obtained results will be compared with classical M_{split} estimators calculated in Wiśniewski (2010). In Sect. 4.2, the data provided in Neri et al. (1989) and also used, *inter alia*, in studies by Schaffrin and Wieser (2008), Shen et al. (2011) and Mahboub (2012), will be used to determine TM_{split} estimators of linear regression parameters. It will be assumed that the basic set of “good” observations is disturbed by “strange” observations for which a corresponding regression line also exists. Moreover, in this Chapter, the behaviour of TM_{split} estimators will be checked in the event that the set contains one observation affected by gross error with different values. The determined TM_{split} estimators will be compared with the WTLS estimators published in the cited studies. TM_{split} estimators’ robustness to gross errors is additionally analysed using an example of a two-dimensional affine transformation (Sect. 4.3). The data for this example are derived from (Lv and Sui 2020). The RTLS and TLTS estimators showed in

the cited paper will be compared with TM_{split} estimators. The paper concludes with a summary.

2 Review of M_{split} and WTLS estimation

2.1 M_{split} estimation

Let $\mathbf{y} = \mathbf{A}\mathbf{X} + \mathbf{v}$ be a functional model of $\mathbf{y} = [y_1, \dots, y_n]^T$ observation vector, where \mathbf{A} is the $n \times m$ coefficient matrix ($\text{rank}(\mathbf{A}) = m$), \mathbf{X} is the m -vector of unknown parameters to be estimated, and \mathbf{v} is n -vector of random observation errors. In M_{split} estimation, this model is split into two models:

$$\mathbf{y} = \mathbf{A}\mathbf{X}_\alpha + \mathbf{v}_\alpha \quad \text{and} \quad \mathbf{y} = \mathbf{A}\mathbf{X}_\beta + \mathbf{v}_\beta \quad (1)$$

where \mathbf{X}_α and \mathbf{X}_β are mutually competing versions of the same vector of parameters \mathbf{X} . The vectors $\mathbf{v}_\alpha, \mathbf{v}_\beta$ are respective versions of the vector \mathbf{v} , which result from the observation errors and the errors of the functional models.

M_{split} estimators of parameters \mathbf{X}_α and \mathbf{X}_β are quantities that minimise the following general objective function (Wiśniewski 2009).

$$\varphi(\mathbf{X}_\alpha, \mathbf{X}_\beta) = \sum_{i=1}^n \rho_\alpha(v_{i\alpha}) \rho_\beta(v_{i\beta}) \quad (2)$$

where ρ_α and ρ_β are arbitrary functions. In the context of cross-weighting that is natural in M_{split} estimation, function (2) can also be expressed in the following form:

$$\varphi(\mathbf{X}_\alpha, \mathbf{X}_\beta) = \sum_{i=1}^n \rho_\alpha(v_{i\alpha}) w_\alpha(v_{i\beta}) = \sum_{i=1}^n \rho_\beta(v_{i\beta}) w_\beta(v_{i\alpha}) \quad (3)$$

where $w_\alpha(v_{i\beta}) = \rho_\beta(v_{i\beta})$ and $w_\beta(v_{i\alpha}) = \rho_\alpha(v_{i\alpha})$ are now regarded as a special type of weight function. The specific character of weighting is that the contribution of function $\rho_\alpha(v_{i\alpha})$ to the optimisation problem is enhanced (or weakened) by the weight function whose argument is quantity $v_{i\beta}$ competing in relation to $v_{i\alpha}$ (and vice versa). The weight functions are not like those in M-estimation, which are modified to make the estimator robust. Mutual “cross weighting” functions $w_\alpha(v_{i\beta})$ and $w_\beta(v_{i\alpha})$ are applied to determine mutually competitive estimates related to the same observation set (Wiśniewski 2009). In the case of M_{split} estimation, one supposes that the observation set might be a mixture of realisations of two different random variables that differ from each other in the parameters of the functional models. One of those variables might be regarded as a “strange” one and its realisations as outliers in a particular case. Then results of M_{split} estimation are estimates of the parameters of the

“good” variable (like in robust M-estimation) but also estimates of the parameters of the “strange” variable.

This study will use the basic variant of M_{split} estimation, in which $\rho(v_\alpha) = v_{i\alpha}^2 q_i^{-1}$ and $\rho(v_\beta) = v_{i\beta}^2 q_i^{-1}$. The quantities q_i are diagonal elements of the \mathbf{Q}_y cofactor matrix occurring in the $\mathbf{C}_y = \sigma_0^2 \mathbf{Q}_y$ covariance matrix model (σ_0^2 —unknown variance component). The adopted functions can be associated (although this is not necessary) with normal distributions as probabilistic observation models. Taking these functions into account, based on Eqs. (2) and (3), the following will be recorded (Wiśniewski 2009; Zienkiewicz 2018a, 2018b)

$$\begin{aligned} \varphi(\mathbf{X}_\alpha, \mathbf{X}_\beta) &= \sum_{i=1}^n \rho_\alpha(v_{i\alpha}) \rho_\beta(v_{i\beta}) = \sum_{i=1}^n v_{i\alpha}^2 v_{i\beta}^2 q_i^{-2} \\ &= \sum_{i=1}^n v_{i\alpha}^2 w_\alpha(v_{i\beta}) = \sum_{i=1}^n v_{i\beta}^2 w_\beta(v_{i\alpha}) \\ &= \mathbf{v}_\alpha^T \mathbf{W}_\alpha(\mathbf{v}_\beta) \mathbf{v}_\alpha = \mathbf{v}_\beta^T \mathbf{W}_\beta(\mathbf{v}_\alpha) \mathbf{v}_\beta \end{aligned} \tag{4}$$

where

$$\begin{aligned} w_\alpha(v_{i\beta}) &= v_{i\beta}^2 q_i^{-2}, \quad w_\beta(v_{i\alpha}) = v_{i\alpha}^2 q_i^{-2} \\ \mathbf{W}_\alpha(\mathbf{v}_\beta) &= \text{Diag}(w_\alpha(v_{1\beta}), \dots, w_\alpha(v_{n\beta})), \\ \mathbf{W}_\beta(\mathbf{v}_\alpha) &= \text{Diag}(w_\beta(v_{1\alpha}), \dots, w_\beta(v_{n\alpha})) \end{aligned} \tag{5}$$

The solution to the optimisation problem $\varphi(\mathbf{X}_\alpha, \mathbf{X}_\beta) \rightarrow \min$ includes such quantities $\hat{\mathbf{X}}_\alpha$ and $\hat{\mathbf{X}}_\beta$ (M_{split} estimators) for which the following is true:

$$\begin{aligned} \frac{1}{2} \frac{\partial \varphi(\mathbf{X}_\alpha, \mathbf{X}_\beta)}{\partial \mathbf{X}_\alpha} \Big|_{\mathbf{X}_\alpha = \hat{\mathbf{X}}_\alpha, \mathbf{X}_\beta = \hat{\mathbf{X}}_\beta} &= \mathbf{A}^T \mathbf{W}_\alpha(\tilde{\mathbf{v}}_\beta) \tilde{\mathbf{v}}_\alpha = \mathbf{A}^T \mathbf{W}_\alpha(\tilde{\mathbf{v}}_\beta) (\mathbf{y} - \mathbf{A} \hat{\mathbf{X}}_\alpha) = \mathbf{0} \\ \frac{1}{2} \frac{\partial \varphi(\mathbf{X}_\alpha, \mathbf{X}_\beta)}{\partial \mathbf{X}_\beta} \Big|_{\mathbf{X}_\alpha = \hat{\mathbf{X}}_\alpha, \mathbf{X}_\beta = \hat{\mathbf{X}}_\beta} &= \mathbf{A}^T \mathbf{W}_\beta(\tilde{\mathbf{v}}_\alpha) \tilde{\mathbf{v}}_\beta = \mathbf{A}^T \mathbf{W}_\beta(\tilde{\mathbf{v}}_\alpha) (\mathbf{y} - \mathbf{A} \hat{\mathbf{X}}_\beta) = \mathbf{0} \end{aligned} \tag{6}$$

where $\tilde{\mathbf{v}}_\alpha = \mathbf{y} - \mathbf{A} \hat{\mathbf{X}}_\alpha$ and $\tilde{\mathbf{v}}_\beta = \mathbf{y} - \mathbf{A} \hat{\mathbf{X}}_\beta$ are residual vectors. The above equations are solved by means of iteration. The iterative procedure can be organised in such a manner that in the steps $l = 1, \dots, s$, the following quantities are determined (Wiśniewski and Zienkiewicz 2021a, 2021b):

$$\begin{aligned} \mathbf{X}_{\alpha(l+1)} &= (\mathbf{A}^T \mathbf{W}_\alpha(\mathbf{v}_{\beta(l)}) \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W}_\alpha(\mathbf{v}_{\beta(l)}) \mathbf{y}, \quad \mathbf{v}_{\alpha(l+1)} = \mathbf{y} - \mathbf{A} \hat{\mathbf{X}}_{\alpha(l+1)} \\ \mathbf{X}_{\beta(l+1)} &= (\mathbf{A}^T \mathbf{W}_\beta(\mathbf{v}_{\alpha(l)}) \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W}_\beta(\mathbf{v}_{\alpha(l)}) \mathbf{y}, \quad \mathbf{v}_{\beta(l+1)} = \mathbf{y} - \mathbf{A} \hat{\mathbf{X}}_{\beta(l+1)} \end{aligned} \tag{7}$$

(the iterative procedure using gradients and Hessians of the function $\varphi(\mathbf{X}_\alpha, \mathbf{X}_\beta)$ is presented in Wiśniewski 2009, 2010). The iterative process defined by Eq. (7) is convergent and ends for such $l = s$ that $\mathbf{X}_{\alpha(s)} = \mathbf{X}_{\alpha(s-1)}$ and $\mathbf{X}_{\beta(s)} = \mathbf{X}_{\beta(s-1)}$. Then, $\hat{\mathbf{X}}_\alpha = \mathbf{X}_{\alpha(s)}$ and $\hat{\mathbf{X}}_\beta = \mathbf{X}_{\beta(s)}$.

In M_{split} estimation, the stochastic model $\mathbf{C}_v = \mathbf{C}_y = \sigma_0^2 \mathbf{Q}_y$, similar to the functional model, is split. The split

results in covariance matrices $\mathbf{C}_{v_\alpha} = \sigma_{0\alpha}^2 \mathbf{Q}_y$ and $\mathbf{C}_{v_\beta} = \sigma_{0\beta}^2 \mathbf{Q}_y$, which are two versions of the covariance matrix \mathbf{C}_v (Wiśniewski and Zienkiewicz 2021b). The invariant and unbiased estimators of variance coefficients $\sigma_{0\alpha}^2$ and $\sigma_{0\beta}^2$ are the following quantities (Wiśniewski and Zienkiewicz, 2021a, b):

$$\begin{aligned} \hat{\sigma}_{0\alpha}^2 &= \frac{\tilde{\mathbf{v}}_\alpha^T \overline{\mathbf{Q}}_{v_\alpha}^{-1} \mathbf{Q}_y \overline{\mathbf{Q}}_{v_\alpha}^{-1} \tilde{\mathbf{v}}_\alpha}{\text{Tr}(\mathbf{N}_\alpha^T \mathbf{N}_\alpha)} \quad \text{and} \\ \hat{\sigma}_{0\beta}^2 &= \frac{\tilde{\mathbf{v}}_\beta^T \overline{\mathbf{Q}}_{v_\beta}^{-1} \mathbf{Q}_y \overline{\mathbf{Q}}_{v_\beta}^{-1} \tilde{\mathbf{v}}_\beta}{\text{Tr}(\mathbf{N}_\beta^T \mathbf{N}_\beta)} \end{aligned} \tag{8}$$

where

$$\begin{aligned} \mathbf{N}_\alpha &= \mathbf{Q}_y \overline{\mathbf{Q}}_{v_\alpha}^{-1} \mathbf{M}_\alpha, \quad \mathbf{N}_\beta = \mathbf{Q}_y \overline{\mathbf{Q}}_{v_\beta}^{-1} \mathbf{M}_\beta \\ \mathbf{M}_\alpha &= \mathbf{I}_n - \mathbf{A} (\mathbf{A}^T \overline{\mathbf{Q}}_{v_\alpha}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \overline{\mathbf{Q}}_{v_\alpha}^{-1}, \\ \mathbf{M}_\beta &= \mathbf{I}_n - \mathbf{A} (\mathbf{A}^T \overline{\mathbf{Q}}_{v_\beta}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \overline{\mathbf{Q}}_{v_\beta}^{-1} \end{aligned} \tag{9}$$

and $\overline{\mathbf{Q}}_{v_\alpha} = [\mathbf{W}_\alpha(\mathbf{v}_\beta)]^{-1} \mathbf{Q}_y$, $\overline{\mathbf{Q}}_{v_\beta} = [\mathbf{W}_\beta(\mathbf{v}_\alpha)]^{-1} \mathbf{Q}_y$, \mathbf{I}_n denotes an $n \times n$ identity matrix (Tr—matrix trace).

2.2 Weighted TLS method

The total least-squares method is applied where the classical model $\mathbf{y} = \mathbf{A} \mathbf{X} + \mathbf{v}$ is replaced by the EIV model of the following form:

$$\mathbf{y} = (\mathbf{A} - \mathbf{E}) \mathbf{X} + \mathbf{v} = \mathbf{A} \mathbf{X} + \mathbf{v} - \mathbf{E} \mathbf{X} \tag{10}$$

where \mathbf{E} is an $n \times m$ random matrix corresponding to the matrix \mathbf{A} being observed. The random vector corresponding to observation vector \mathbf{y} in the EIV model was denoted as \mathbf{v} . If we assume that $\mathbf{v} = \mathbf{y} - \mathbf{A} \mathbf{X}$ is the error vector in the classical functional model, then $\mathbf{v} = \mathbf{v} + \mathbf{E} \mathbf{X}$. It should be considered that $\mathbf{E} \mathbf{X} = (\mathbf{X}^T \otimes \mathbf{I}_n) \mathbf{e}$, where $\mathbf{e} = \text{vec}(\mathbf{E})$ is a vector formed from successive columns of matrix \mathbf{E} (\otimes —the Kronecker product, \mathbf{I}_n —an identity matrix of dimensions $n \times n$). Moreover, in order to simplify further notation, additional designations $\mathbf{X}_\otimes = \mathbf{X} \otimes \mathbf{I}_n = (\mathbf{X}^T \otimes \mathbf{I}_n)^T$, $\mathbf{X}_\otimes^T = \mathbf{X}^T \otimes \mathbf{I}_n = (\mathbf{X} \otimes \mathbf{I}_n)^T$ are introduced. Model (10) can then be expressed in the following form:

$$\mathbf{y} = \mathbf{A} \mathbf{X} + \mathbf{v} - \mathbf{E} \mathbf{X} = \mathbf{A} \mathbf{X} + \mathbf{v} - (\mathbf{X}_\otimes^T \otimes \mathbf{I}_n) \mathbf{e} = \mathbf{A} \mathbf{X} + \mathbf{v} - \mathbf{X}_\otimes^T \mathbf{e} \tag{11}$$

In the WTLS method, in addition to the stochastic model of the observation vector $\mathbf{C}_y = \sigma_0^2 \mathbf{Q}_y$, a stochastic model of \mathbf{e} vector is also adopted. In the simplest case, it can be assumed that such a model is the expression $\mathbf{C}_e = \sigma_0^2 \mathbf{Q}_e$, where \mathbf{Q}_e is the known cofactor matrix. However, there are examples in which not all columns of matrix \mathbf{A} are affected by random

disturbances (e.g. in the linear regression analysis). In that case, matrix \mathbf{Q}_e can be subject to appropriate decomposition $\mathbf{Q}_e = \mathbf{Q}_0 \otimes \mathbf{Q}_x$, where \mathbf{Q}_x denotes a nonnegative definite diagonal matrix of size $n \times n$ (Schaffrin and Wieser 2008; Shen et al. 2011).

In the Lagrange approach applied in Schaffrin and Wieser (2008) and Lv and Sui (2020), the WTLS method is based on the following objective function:

$$\varphi(\mathbf{v}, \mathbf{e}, \boldsymbol{\lambda}, \mathbf{X}) = \mathbf{v}^T \mathbf{Q}_y^{-1} \mathbf{v} + \mathbf{e}^T \mathbf{Q}_e^+ \mathbf{e} - 2\boldsymbol{\lambda}^T (\mathbf{y} - \mathbf{A}\mathbf{X} - \mathbf{v} + \mathbf{X}_{\otimes}^T \mathbf{e}) \tag{12}$$

where $\boldsymbol{\lambda}$ denotes an $n \times 1$ vector of ‘‘Lagrange multipliers’’. Matrix \mathbf{Q}_e^+ is the Moore–Penrose inverse of \mathbf{Q}_e . For example, when $\mathbf{Q}_0 = \text{Diag}(0, 1)$, hence $\mathbf{Q}_e = \text{Diag}(\mathbf{0}, \mathbf{Q}_x)$, and \mathbf{Q}_x is regular, then $\mathbf{Q}_e^+ = \text{Diag}(\mathbf{0}, \mathbf{Q}_x^{-1})$ (e.g. Rao 1973; Felus 2004). The iteration procedures that solve the optimisation problem $\varphi(\mathbf{v}, \mathbf{e}, \boldsymbol{\lambda}, \mathbf{X}) \rightarrow \min$ with the nonlinear conditions $\mathbf{y} - \mathbf{A}\mathbf{X} - \mathbf{v} + \mathbf{X}_{\otimes}^T \mathbf{e} = \mathbf{0}$ are, in general, complicated. The application of a similar objective function in M_{split} estimation will generate numerical procedures and computational algorithms of even greater complexity. From the perspective of computational process optimisation, however, the approach adopted in Shen et al. (2011), also applied in Wang et al. (2016), is particularly interesting for the purposes of the study. The iterative method proposed in the study is based on the Newton–Gauss algorithm of nonlinear LS adjustment, proposed by Pope (1974). In this method, the EIV model (11) is replaced, in the j -th iteration, by a linear approximation of the following form:

$$\begin{aligned} \mathbf{y} &= \mathbf{A}\mathbf{X} + \mathbf{v} - \mathbf{E}\mathbf{X} \\ &= \mathbf{A}\mathbf{X}^j + \mathbf{A}^j \delta \mathbf{X} + \mathbf{v} - \mathbf{E}\mathbf{X}^j \\ &= \mathbf{A}\mathbf{X}^j + \mathbf{A}^j \delta \mathbf{X} + \mathbf{v} - (\mathbf{X}_{\otimes}^j)^T \mathbf{e} \end{aligned} \tag{13}$$

where $\mathbf{A}^j = \mathbf{A} - \tilde{\mathbf{E}}^j$, $\mathbf{X} = \mathbf{X}^j + \delta \mathbf{X}$ and $\mathbf{X}_{\otimes}^j = \mathbf{X}^j \otimes \mathbf{I}_n$. $\tilde{\mathbf{E}}^j$ are residual matrices built on the basis of the residual vector $\tilde{\mathbf{e}}^j$ determined in the j -th iteration. Vector $\delta \mathbf{X}$ is a small quantity to be solved in the iteration. After taking into account the condition $\mathbf{y} - \mathbf{A}\mathbf{X}^j - \mathbf{A}^j \delta \mathbf{X} - \mathbf{v} + (\mathbf{X}_{\otimes}^j)^T \mathbf{e} = \mathbf{0}$ resulting from model (13), the objective function (12) can be rewritten in the following form (Shen et al. 2011).

$$\varphi(\mathbf{v}, \mathbf{e}, \boldsymbol{\lambda}, \delta \mathbf{X}) = \mathbf{v}^T \mathbf{Q}_y^{-1} \mathbf{v} + \mathbf{e}^T \mathbf{Q}_e^+ \mathbf{e} - 2\boldsymbol{\lambda}^T (\mathbf{y} - \mathbf{A}\mathbf{X}^j - \mathbf{A}^j \delta \mathbf{X} - \mathbf{v} + (\mathbf{X}_{\otimes}^j)^T \mathbf{e}) \tag{14}$$

The minimum of this function is obtained through satisfying the following Euler–Lagrange necessary conditions (Shen et al. 2011)

$$\left. \frac{1}{2} \frac{\partial \varphi}{\partial \mathbf{v}} \right|_{\mathbf{v}=\tilde{\mathbf{v}}, \mathbf{e}=\tilde{\mathbf{e}}, \delta \mathbf{X}=\delta \hat{\mathbf{X}}, \boldsymbol{\lambda}=\hat{\boldsymbol{\lambda}}} = \mathbf{Q}_y^{-1} \tilde{\mathbf{v}} + \hat{\boldsymbol{\lambda}} = \mathbf{0}$$

$$\begin{aligned} \left. \frac{1}{2} \frac{\partial \varphi}{\partial \mathbf{e}} \right|_{\mathbf{v}=\tilde{\mathbf{v}}, \mathbf{e}=\tilde{\mathbf{e}}, \delta \mathbf{X}=\delta \hat{\mathbf{X}}, \boldsymbol{\lambda}=\hat{\boldsymbol{\lambda}}} &= \mathbf{Q}_e^+ \tilde{\mathbf{e}} - \mathbf{X}_{\otimes}^j \hat{\boldsymbol{\lambda}} = \mathbf{0} \\ \left. \frac{1}{2} \frac{\partial \varphi}{\partial \delta \mathbf{X}} \right|_{\mathbf{v}=\tilde{\mathbf{v}}, \mathbf{e}=\tilde{\mathbf{e}}, \delta \mathbf{X}=\delta \hat{\mathbf{X}}, \boldsymbol{\lambda}=\hat{\boldsymbol{\lambda}}} &= (\mathbf{A}^j)^T \hat{\boldsymbol{\lambda}} = \mathbf{0} \\ \left. \frac{1}{2} \frac{\partial \varphi}{\partial \boldsymbol{\lambda}} \right|_{\mathbf{v}=\tilde{\mathbf{v}}, \mathbf{e}=\tilde{\mathbf{e}}, \delta \mathbf{X}=\delta \hat{\mathbf{X}}, \boldsymbol{\lambda}=\hat{\boldsymbol{\lambda}}} &= \mathbf{y} - \mathbf{A}\mathbf{X}^j - \mathbf{A}^j \delta \hat{\mathbf{X}} - \tilde{\mathbf{v}} + (\mathbf{X}_{\otimes}^j)^T \mathbf{e} = \mathbf{0} \end{aligned} \tag{15}$$

The solution to the equations contained in Eq. (15) are the following quantities:

$$\begin{aligned} \hat{\boldsymbol{\lambda}} &= -(\mathbf{Q}_l^j)^{-1} (\mathbf{y} - \mathbf{A}\mathbf{X}^j - \mathbf{A}^j \delta \hat{\mathbf{X}}) \\ \delta \hat{\mathbf{X}}^{j+1} &= \left[(\mathbf{A}^j)^T (\mathbf{Q}_l^j)^{-1} \mathbf{A}^j \right]^{-1} (\mathbf{A}^j)^T (\mathbf{Q}_l^j)^{-1} (\mathbf{y} - \mathbf{A}\mathbf{X}^j) \\ \mathbf{X}^{j+1} &= \hat{\mathbf{X}}^j + \delta \hat{\mathbf{X}}^{j+1} \\ &= \left[(\mathbf{A}^j)^T (\mathbf{Q}_l^j)^{-1} \mathbf{A}^j \right]^{-1} (\mathbf{A}^j)^T (\mathbf{Q}_l^j)^{-1} (\mathbf{y} - \mathbf{E}^j \mathbf{X}^j) \end{aligned} \tag{16}$$

and

$$\begin{aligned} \tilde{\mathbf{v}}^{j+1} &= -\mathbf{Q}_y (\mathbf{Q}_l^j)^{-1} (\mathbf{y} - \mathbf{A}\hat{\mathbf{X}}^j - \mathbf{A}^j \delta \hat{\mathbf{X}}^{j+1}) \\ \tilde{\mathbf{e}} &= \mathbf{Q}_e \mathbf{X}_{\otimes}^j (\mathbf{Q}_l^j)^{-1} (\mathbf{y} - \mathbf{A}\hat{\mathbf{X}}^j - \mathbf{A}^j \delta \hat{\mathbf{X}}^{j+1}) \end{aligned} \tag{17}$$

where

$$\mathbf{Q}_l^j = \mathbf{Q}_y + (\mathbf{X}_{\otimes}^j)^T \mathbf{Q}_e \mathbf{X}_{\otimes}^j \tag{18}$$

($\tilde{\mathbf{v}}$ —residual vector corresponding to the observation vector \mathbf{y}). Shen et al. (2011) also apply the iterative process on the assumption that $\mathbf{E}^j \delta \mathbf{X}$ is a negligible quantity. Then, $\mathbf{A}^j \delta \mathbf{X} = (\mathbf{A} - \mathbf{E}^j) \delta \mathbf{X} = \mathbf{A} \delta \mathbf{X}$ and $\mathbf{y} = \mathbf{A}\mathbf{X}^j + \mathbf{A} \delta \mathbf{X} + \mathbf{v} - \mathbf{E}^j \mathbf{X}^j$.

The estimation of variance coefficient σ_0^2 , common to stochastic models $\mathbf{C}_y = \sigma_0^2 \mathbf{Q}_y$ and $\mathbf{C}_e = \sigma_0^2 \mathbf{Q}_e$, is also of interest in WTLS. Schaffrin and Wieser (2008) proposed a biased estimator in the following form:

$$\hat{\sigma}_0^2 = (\mathbf{y} - \mathbf{A}\hat{\mathbf{X}})^T \mathbf{Q}_l^{-1} (\mathbf{y} - \mathbf{A}\hat{\mathbf{X}}) / (n - m) \tag{19}$$

A correction of the estimator (19) by means of introducing a bias term δb to it was presented by Shen et al. (2011). More complex EIV stochastic models are also being under consideration. For example, Xu and Liu (2014) introduced a model containing variance components and proposed a way to estimate these components.

3 Total M_{split} estimation

3.1 Theoretical foundations

Let it be assumed that according to the M_{split} estimation rules, the EIV model (10) is split into two mutually competing

models

$$\begin{aligned} \mathbf{y} &= (\mathbf{A} - \mathbf{E})\mathbf{X}_\alpha + \mathbf{v}_\alpha = \mathbf{A}\mathbf{X}_\alpha + \mathbf{v}_\alpha - \mathbf{E}\mathbf{X}_\alpha \\ \mathbf{y} &= (\mathbf{A} - \mathbf{E})\mathbf{X}_\beta + \mathbf{v}_\beta = \mathbf{A}\mathbf{X}_\beta + \mathbf{v}_\beta - \mathbf{E}\mathbf{X}_\beta \end{aligned} \tag{20}$$

By applying the approach used in Shen et al. (2011) and Wang et al. (2016), see Eq. (13), the above models in the j -th iteration will be replaced with the following linear approximations:

$$\begin{aligned} \mathbf{y} &= \mathbf{A}\mathbf{X}_\alpha^j + \mathbf{A}^j\delta\mathbf{X}_\alpha + \mathbf{v}_\alpha - \mathbf{E}\mathbf{X}_\alpha^j \\ &= \mathbf{A}\mathbf{X}_\alpha^j + \mathbf{A}^j\delta\mathbf{X}_\alpha + \mathbf{v}_\alpha - (\mathbf{X}_{\otimes\alpha}^j)^T \mathbf{e} \\ \mathbf{y} &= \mathbf{A}\mathbf{X}_\beta^j + \mathbf{A}^j\delta\mathbf{X}_\beta + \mathbf{v}_\beta - \mathbf{E}\mathbf{X}_\beta^j \\ &= \mathbf{A}\mathbf{X}_\beta^j + \mathbf{A}^j\delta\mathbf{X}_\beta + \mathbf{v}_\beta - (\mathbf{X}_{\otimes\beta}^j)^T \mathbf{e} \end{aligned} \tag{21}$$

where $\mathbf{X}_\alpha = \mathbf{X}_\alpha^j + \delta\mathbf{X}_\alpha$, $\mathbf{X}_\beta = \mathbf{X}_\beta^j + \delta\mathbf{X}_\beta$ and $\mathbf{X}_{\otimes\alpha}^j = \mathbf{X}_\alpha^j \otimes \mathbf{I}_n$, $\mathbf{X}_{\otimes\beta}^j = \mathbf{X}_\beta^j \otimes \mathbf{I}_n$. After determining the quantities \mathbf{E}^j , \mathbf{X}_α^j and \mathbf{X}_β^j , the models that are valid in the j -th iteration, contained in Eq. (20), take the following forms:

$$\begin{aligned} \mathbf{y} &= (\mathbf{A} - \mathbf{E}^j)\mathbf{X}_\alpha^j + \mathbf{v}_\alpha^j = \mathbf{A}^j\mathbf{X}_\alpha^j + \mathbf{v}_\alpha^j \\ \mathbf{y} &= (\mathbf{A} - \mathbf{E}^j)\mathbf{X}_\beta^j + \mathbf{v}_\beta^j = \mathbf{A}^j\mathbf{X}_\beta^j + \mathbf{v}_\beta^j \end{aligned} \tag{22}$$

Total M_{split} estimators of parameters \mathbf{X}_α and \mathbf{X}_β are quantities $\hat{\mathbf{X}}_\alpha$ and $\hat{\mathbf{X}}_\beta$ which minimise the following objective function:

$$\varphi(\mathbf{X}_\alpha, \mathbf{X}_\beta) = \mathbf{v}_\alpha^T \mathbf{W}_\alpha(\mathbf{v}_\beta)\mathbf{v}_\alpha + \mathbf{e}^T \mathbf{Q}_e^+ \mathbf{e} = \mathbf{v}_\beta^T \mathbf{W}_\beta(\mathbf{v}_\alpha)\mathbf{v}_\beta + \mathbf{e}^T \mathbf{Q}_e^+ \mathbf{e} \tag{23}$$

This function [like functions Eqs. (3) and (4)] is not designed to make the method robust and also does not “predict” outliers among observations of elements of matrix \mathbf{A} . In view of the following conditions resulting from Eq. (21)

$$\begin{aligned} \mathbf{y} - \mathbf{A}\mathbf{X}_\alpha^j - \mathbf{A}^j\delta\mathbf{X}_\alpha - \mathbf{v}_\alpha + (\mathbf{X}_{\otimes\alpha}^j)^T \mathbf{e} &= \mathbf{0} \\ \mathbf{y} - \mathbf{A}\mathbf{X}_\beta^j - \mathbf{A}^j\delta\mathbf{X}_\beta - \mathbf{v}_\beta + (\mathbf{X}_{\otimes\beta}^j)^T \mathbf{e} &= \mathbf{0} \end{aligned} \tag{24}$$

the original objective function (23) will be supplemented to the following form:

$$\begin{aligned} \varphi(\mathbf{v}_\alpha, \mathbf{v}_\beta, \mathbf{e}, \mathbf{X}_\alpha, \mathbf{X}_\beta, \boldsymbol{\lambda}_\alpha, \boldsymbol{\lambda}_\beta) &= \mathbf{v}_\alpha^T \mathbf{W}_\alpha(\mathbf{v}_\beta)\mathbf{v}_\alpha + \mathbf{e}^T \mathbf{Q}_e^+ \mathbf{e} \\ &\quad - 2\boldsymbol{\lambda}_\alpha^T (\mathbf{y} - \mathbf{A}\mathbf{X}_\alpha^j - \mathbf{A}^j\delta\mathbf{X}_\alpha - \mathbf{v}_\alpha + (\mathbf{X}_{\otimes\alpha}^j)^T \mathbf{e}) \\ &\quad - 2\boldsymbol{\lambda}_\beta^T (\mathbf{y} - \mathbf{A}\mathbf{X}_\beta^j - \mathbf{A}^j\delta\mathbf{X}_\beta - \mathbf{v}_\beta + (\mathbf{X}_{\otimes\beta}^j)^T \mathbf{e}) \\ &= \mathbf{v}_\beta^T \mathbf{W}_\beta(\mathbf{v}_\alpha)\mathbf{v}_\beta + \mathbf{e}^T \mathbf{Q}_e^+ \mathbf{e} \\ &\quad - 2\boldsymbol{\lambda}_\alpha^T (\mathbf{y} - \mathbf{A}\mathbf{X}_\alpha^j - \mathbf{A}^j\delta\mathbf{X}_\alpha - \mathbf{v}_\alpha + (\mathbf{X}_{\otimes\alpha}^j)^T \mathbf{e}) \end{aligned}$$

$$- 2\boldsymbol{\lambda}_\beta^T (\mathbf{y} - \mathbf{A}\mathbf{X}_\beta^j - \mathbf{A}^j\delta\mathbf{X}_\beta - \mathbf{v}_\beta + (\mathbf{X}_{\otimes\beta}^j)^T \mathbf{e}) \tag{25}$$

where $\boldsymbol{\lambda}_\alpha$ and $\boldsymbol{\lambda}_\beta$ are Lagrange multiplier vectors corresponding to conditions (24). It is established that the Euler–Lagrange necessary conditions have the following forms in the optimisation problem $\varphi(\mathbf{v}_\alpha, \mathbf{v}_\beta, \mathbf{e}, \mathbf{X}_\alpha, \mathbf{X}_\beta, \boldsymbol{\lambda}_\alpha, \boldsymbol{\lambda}_\beta) \rightarrow \min$:

$$\left. \frac{1}{2} \frac{\partial \varphi}{\partial \mathbf{v}_\alpha} \right|_\Omega = \mathbf{W}_\alpha(\tilde{\mathbf{v}}_\beta)\tilde{\mathbf{v}}_\alpha + \hat{\boldsymbol{\lambda}}_\alpha = \mathbf{0} \tag{26a}$$

$$\left. \frac{1}{2} \frac{\partial \varphi}{\partial \mathbf{v}_\beta} \right|_\Omega = \mathbf{W}_\beta(\tilde{\mathbf{v}}_\alpha)\tilde{\mathbf{v}}_\beta + \hat{\boldsymbol{\lambda}}_\beta = \mathbf{0} \tag{26b}$$

$$\left. \frac{1}{2} \frac{\partial \varphi}{\partial \mathbf{e}} \right|_\Omega = \mathbf{Q}_e^+ \tilde{\mathbf{e}} - \mathbf{X}_{\otimes\alpha}^j \hat{\boldsymbol{\lambda}}_\alpha - \mathbf{X}_{\otimes\beta}^j \hat{\boldsymbol{\lambda}}_\beta = \mathbf{0} \tag{26c}$$

$$\left. \frac{1}{2} \frac{\partial \varphi}{\partial \delta\mathbf{X}_\alpha} \right|_\Omega = (\mathbf{A}^j)^T \hat{\boldsymbol{\lambda}}_\alpha = \mathbf{0} \tag{26d}$$

$$\left. \frac{1}{2} \frac{\partial \varphi}{\partial \delta\mathbf{X}_\beta} \right|_\Omega = (\mathbf{A}^j)^T \hat{\boldsymbol{\lambda}}_\beta = \mathbf{0} \tag{26e}$$

$$\left. \frac{1}{2} \frac{\partial \varphi}{\partial \hat{\mathbf{X}}_\alpha} \right|_\Omega = \mathbf{y} - \mathbf{A}\mathbf{X}_\alpha^j - \mathbf{A}^j\delta\hat{\mathbf{X}}_\alpha - \tilde{\mathbf{v}}_\alpha + (\mathbf{X}_{\otimes\alpha}^j)^T \tilde{\mathbf{e}} = \mathbf{0} \tag{26f}$$

$$\left. \frac{1}{2} \frac{\partial \varphi}{\partial \hat{\mathbf{X}}_\beta} \right|_\Omega = \mathbf{y} - \mathbf{A}\mathbf{X}_\beta^j - \mathbf{A}^j\delta\hat{\mathbf{X}}_\beta - \tilde{\mathbf{v}}_\beta + (\mathbf{X}_{\otimes\beta}^j)^T \tilde{\mathbf{e}} = \mathbf{0} \tag{26g}$$

The set of simultaneous substitutions: $\mathbf{v}_\alpha = \tilde{\mathbf{v}}_\alpha$, $\mathbf{v}_\beta = \tilde{\mathbf{v}}_\beta$, $\mathbf{e} = \tilde{\mathbf{e}}$, $\delta\mathbf{X}_\alpha = \delta\hat{\mathbf{X}}_\alpha$, $\delta\mathbf{X}_\beta = \delta\hat{\mathbf{X}}_\beta$, $\boldsymbol{\lambda}_\alpha = \hat{\boldsymbol{\lambda}}_\alpha$, $\boldsymbol{\lambda}_\beta = \hat{\boldsymbol{\lambda}}_\beta$, introduced to simplify the notation, was denoted as Ω . Based on Eqs. (26a)–(26c), the following residual vectors are determined:

$$\begin{aligned} \tilde{\mathbf{v}}_\alpha &= -\mathbf{W}_\alpha^{-1}(\tilde{\mathbf{v}}_\beta)\hat{\boldsymbol{\lambda}}_\alpha \\ \tilde{\mathbf{v}}_\beta &= -\mathbf{W}_\beta^{-1}(\tilde{\mathbf{v}}_\alpha)\hat{\boldsymbol{\lambda}}_\beta \\ \tilde{\mathbf{e}} &= \mathbf{Q}_e(\mathbf{X}_{\otimes\alpha}^j \hat{\boldsymbol{\lambda}}_\alpha + \mathbf{X}_{\otimes\beta}^j \hat{\boldsymbol{\lambda}}_\beta) \end{aligned} \tag{27}$$

By substituting the quantities obtained above to Eqs. (26f) and (26g), a system of normal equations relating to vectors $\hat{\boldsymbol{\lambda}}_\alpha$ and $\hat{\boldsymbol{\lambda}}_\beta$ is obtained. The system has the following form:

$$\begin{aligned} \Theta_\alpha^j \hat{\boldsymbol{\lambda}}_\alpha + \Theta_{\alpha\beta}^j \hat{\boldsymbol{\lambda}}_\beta &= -(\mathbf{y} - \mathbf{A}\mathbf{X}_\alpha^j - \mathbf{A}^j\delta\hat{\mathbf{X}}_\alpha) \\ \Theta_{\beta\alpha}^j \hat{\boldsymbol{\lambda}}_\alpha + \Theta_\beta^j \hat{\boldsymbol{\lambda}}_\beta &= -(\mathbf{y} - \mathbf{A}\mathbf{X}_\beta^j - \mathbf{A}^j\delta\hat{\mathbf{X}}_\beta) \end{aligned} \tag{28}$$

where

$$\begin{aligned} \Theta_\alpha^j &= \mathbf{W}_\alpha^{-1}(\tilde{\mathbf{v}}_\beta) + (\mathbf{X}_{\otimes\alpha}^j)^T \mathbf{Q}_e \mathbf{X}_{\otimes\alpha}^j, \quad \Theta_{\alpha,\beta}^j = (\mathbf{X}_{\otimes\alpha}^j)^T \mathbf{Q}_e \mathbf{X}_{\otimes\beta}^j \\ \Theta_{\beta\alpha}^j &= (\Theta_{\alpha\beta}^j)^T = (\mathbf{X}_{\otimes\beta}^j)^T \mathbf{Q}_e \mathbf{X}_{\otimes\alpha}^j, \quad \Theta_\beta^j = \mathbf{W}_\beta^{-1}(\tilde{\mathbf{v}}_\alpha) + (\mathbf{X}_{\otimes\beta}^j)^T \mathbf{Q}_e \mathbf{X}_{\otimes\beta}^j \end{aligned} \tag{29}$$

After introducing block matrices

$$\begin{aligned} \Theta^j &= \begin{bmatrix} \Theta_{\alpha}^j & \Theta_{\alpha\beta}^j \\ \Theta_{\beta\alpha}^j & \Theta_{\beta}^j \end{bmatrix}, \\ \widehat{\mathbf{A}} &= \mathbf{I}_2 \otimes \mathbf{A} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{A} \end{bmatrix}, \\ \widehat{\mathbf{y}} &= \mathbf{1}_2 \otimes \mathbf{y} = \begin{bmatrix} \mathbf{y} \\ \mathbf{y} \end{bmatrix} \end{aligned} \tag{30}$$

($\mathbf{1}_2 = [1, 1]^T$) and after including the mutually competing quantities being determined in combined vectors, i.e. having introduced vectors

$$\begin{aligned} \boldsymbol{\lambda} &= [\boldsymbol{\lambda}_{\alpha}^T, \boldsymbol{\lambda}_{\beta}^T]^T, \delta \mathbf{X} = [\delta \mathbf{X}_{\alpha}^T, \delta \mathbf{X}_{\beta}^T]^T, \\ \mathbf{X} &= [\mathbf{X}_{\alpha}^T, \mathbf{X}_{\beta}^T]^T \end{aligned} \tag{31}$$

system of Eqs. (28) can also be expressed as follows:

$$\Theta^j \hat{\boldsymbol{\lambda}} = -(\widehat{\mathbf{y}} - \widehat{\mathbf{A}} \mathbf{X}^j - \widehat{\mathbf{A}}^j \delta \widehat{\mathbf{X}}) \tag{32}$$

The solution to Eq. (32) is the combined Lagrange multiplier vector of the following form:

$$\hat{\boldsymbol{\lambda}} = \begin{bmatrix} \hat{\boldsymbol{\lambda}}_{\alpha} \\ \hat{\boldsymbol{\lambda}}_{\beta} \end{bmatrix} = -(\Theta^j)^{-1} (\widehat{\mathbf{y}} - \widehat{\mathbf{A}} \mathbf{X}^j - \widehat{\mathbf{A}}^j \delta \widehat{\mathbf{X}}) \tag{33}$$

After substituting Eq. (33) to conditions (26d) and (26e), jointly recorded as

$$\left. \begin{aligned} (\mathbf{A}^j)^T \hat{\boldsymbol{\lambda}}_{\alpha} &= \mathbf{0} \\ (\mathbf{A}^j)^T \hat{\boldsymbol{\lambda}}_{\beta} &= \mathbf{0} \end{aligned} \right\} \Leftrightarrow (\widehat{\mathbf{A}}^j)^T \hat{\boldsymbol{\lambda}} = \mathbf{0} \tag{34}$$

the following normal equation is obtained:

$$(\widehat{\mathbf{A}}^j)^T (\Theta^j)^{-1} (\widehat{\mathbf{y}} - \widehat{\mathbf{y}} \mathbf{X}^j - \widehat{\mathbf{A}}^j \delta \widehat{\mathbf{X}}) = \mathbf{0} \tag{35}$$

The solution to this equation is vector

$$\delta \widehat{\mathbf{X}}^{j+1} = \begin{bmatrix} \delta \widehat{\mathbf{X}}_{\alpha}^{j+1} \\ \delta \widehat{\mathbf{X}}_{\beta}^{j+1} \end{bmatrix} = \left((\widehat{\mathbf{A}}^j)^T (\Theta^j)^{-1} \widehat{\mathbf{A}}^j \right)^{-1} (\widehat{\mathbf{y}} - \widehat{\mathbf{A}} \mathbf{X}^j) \tag{36}$$

which represents an evaluation of the combined vector of $\delta \mathbf{X}$ increments in the $(j + 1)$ iteration. In order to determine the combined vector $\mathbf{X} = \mathbf{X}^j + \delta \mathbf{X}$ that is valid in this iteration, it will be taken into account that $\mathbf{A} = \mathbf{A}^j + \mathbf{E}^j$, and thus, $\widehat{\mathbf{A}} = \widehat{\mathbf{A}}^j + \widehat{\mathbf{e}}^j$, where $\widehat{\mathbf{E}}^j = \mathbf{I}_2 \otimes \mathbf{E}^j$. Equation (35) then takes the following form:

$$(\widehat{\mathbf{A}}^j)^T (\Theta^j)^{-1} (\widehat{\mathbf{y}} - \widehat{\mathbf{A}}^j \mathbf{X}^j - \widehat{\mathbf{E}}^j \mathbf{X}^j - \widehat{\mathbf{A}}^j \delta \widehat{\mathbf{X}}) = \mathbf{0} \tag{37}$$

which yields the following:

$$\begin{aligned} \widehat{\mathbf{X}}^{j+1} &= \begin{bmatrix} \widehat{\mathbf{X}}_{\alpha}^{j+1} \\ \widehat{\mathbf{X}}_{\beta}^{j+1} \end{bmatrix} = \widehat{\mathbf{X}}^j + \delta \widehat{\mathbf{X}}^{j+1} \\ &= \left((\widehat{\mathbf{A}}^j)^T (\Theta^j)^{-1} \widehat{\mathbf{A}}^j \right)^{-1} (\widehat{\mathbf{A}}^j)^T (\Theta^j)^{-1} (\widehat{\mathbf{y}} - \widehat{\mathbf{E}}^j \mathbf{X}^j) \end{aligned} \tag{38}$$

Based on Eqs. (26a) and (26b), jointly recorded as

$$\left. \begin{aligned} \tilde{\mathbf{v}}_{\alpha} &= -\mathbf{W}_{\alpha}^{-1} (\tilde{\mathbf{v}}_{\beta}) \hat{\boldsymbol{\lambda}}_{\alpha} \\ \tilde{\mathbf{v}}_{\beta} &= -\mathbf{W}_{\beta}^{-1} (\tilde{\mathbf{v}}_{\alpha}) \hat{\boldsymbol{\lambda}}_{\beta} \end{aligned} \right\} \Leftrightarrow \tilde{\mathbf{v}} = -\widehat{\mathbf{W}}^{-1} \hat{\boldsymbol{\lambda}} \tag{39}$$

the combined residual vector that is valid in the $(j + 1)$ iteration can be determined:

$$\tilde{\mathbf{v}}^{j+1} = \begin{bmatrix} \tilde{\mathbf{v}}_{\alpha}^{j+1} \\ \tilde{\mathbf{v}}_{\beta}^{j+1} \end{bmatrix} = \widehat{\mathbf{W}}^{-1} (\tilde{\mathbf{v}}_{\alpha}^j, \tilde{\mathbf{v}}_{\beta}^j) (\Theta^j)^{-1} (\widehat{\mathbf{y}} - \widehat{\mathbf{A}} \widehat{\mathbf{X}}^j - \widehat{\mathbf{A}}^j \delta \widehat{\mathbf{X}}^{j+1}) \tag{40}$$

where

$$\widehat{\mathbf{W}}^{-1} (\tilde{\mathbf{v}}_{\alpha}^j, \tilde{\mathbf{v}}_{\beta}^j) = \text{Diag}(\mathbf{W}_{\alpha}^{-1} (\tilde{\mathbf{v}}_{\beta}^j), \mathbf{W}_{\beta}^{-1} (\tilde{\mathbf{v}}_{\alpha}^j)) \tag{41}$$

On the other hand, based on Eq. (26c), expressed in the following form:

$$\tilde{\mathbf{e}} = \mathbf{Q}_{\mathbf{e}} (\mathbf{X}_{\otimes \alpha}^j \hat{\boldsymbol{\lambda}}_{\alpha} + \mathbf{X}_{\otimes \beta}^j \hat{\boldsymbol{\lambda}}_{\beta}) = \mathbf{Q}_{\mathbf{e}} \mathbf{X}_{\otimes}^j \hat{\boldsymbol{\lambda}} \tag{42}$$

the following residual vector is determined:

$$\tilde{\mathbf{e}}^{j+1} = -\mathbf{Q}_{\mathbf{e}} \mathbf{X}_{\otimes}^j (\Theta^j)^{-1} (\widehat{\mathbf{y}} - \widehat{\mathbf{A}} \widehat{\mathbf{X}}^j - \widehat{\mathbf{A}}^j \delta \widehat{\mathbf{X}}^{j+1}) \tag{43}$$

where $\mathbf{X}_{\otimes}^j = [(\mathbf{X}_{\otimes \alpha}^j)^T, (\mathbf{X}_{\otimes \beta}^j)^T]^T$. Taking that vector, we can obtain $\tilde{\mathbf{E}}^{j+1}$, and hence, the matrix $\mathbf{A}^{j+1} = \mathbf{A} - \tilde{\mathbf{E}}^{j+1}$.

Variance coefficient estimators in TM_{split} estimation should be derived by referring to the split EIV models (e.g. by applying the theory presented in Wiśniewski and Zienkiewicz (2021b)). However, this problem that requires additional, detailed theoretical and empirical analyses is beyond the scope of this paper. With minor random disturbances of matrix \mathbf{A} , variance coefficient estimators appropriate for M_{split} estimation can also be recommended here. Then, in Eq. (8), the vectors $\tilde{\mathbf{v}}_{\alpha}$ and $\tilde{\mathbf{v}}_{\beta}$ should be replaced by the vectors $\tilde{\mathbf{v}}_{\alpha} = \mathbf{y} - (\mathbf{A} - \tilde{\mathbf{E}}) \widehat{\mathbf{X}}_{\alpha}$ and $\tilde{\mathbf{v}}_{\beta} = \mathbf{y} - (\mathbf{A} - \tilde{\mathbf{E}}) \widehat{\mathbf{X}}_{\beta}$. These estimators of variance coefficients stay unbiased; however, they lose their invariance according to growing values of $\tilde{\mathbf{E}} \widehat{\mathbf{X}}_{\alpha}$ and $\tilde{\mathbf{E}} \widehat{\mathbf{X}}_{\beta}$.

3.2 Algorithm

Total M_{split} estimation algorithm contains the following basic elements: *Step 0*—a starting step, *Step 1*—iterative calculation of M_{split} estimators for valid split EIV models (with internal iterations $l = 0, \dots, s$), *Step 2*—updating the EIV models’ parameters, and return to *Step 1* (until the adopted criterion for stopping the iterative process, $j = 0, \dots, k$), *Step 3*—adopting the final values of Total M_{split} estimators. Each of these steps is described in more detail below.

Step 0: Similar to M_{split} estimation, the iterative process can also be initiated here using the following classical least-squares (LS) estimators

$$\hat{\mathbf{X}}_{LS} = (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W} \mathbf{y} \quad \text{and} \quad \tilde{\mathbf{v}}_{LS} = \mathbf{y} - \mathbf{A} \hat{\mathbf{X}}_{LS} \quad (44)$$

where $\mathbf{W} = \mathbf{Q}_y^{-1}$ is the weight matrix. Therefore, the following are adopted: $\mathbf{X}_{\alpha(0)}^0 = \hat{\mathbf{X}}_{LS}$, $\mathbf{X}_{\beta(0)}^0 = \hat{\mathbf{X}}_{LS}$, $\mathbf{v}_{\alpha(0)}^0 = \tilde{\mathbf{v}}_{LS}$, $\mathbf{v}_{\beta(0)}^0 = \tilde{\mathbf{v}}_{LS}$. Moreover, $\mathbf{E}^0 = \mathbf{0}$, $\mathbf{A}^0 = \mathbf{A}$ and $\delta \hat{\mathbf{X}}^0 = \mathbf{0}$.

Step 1: Calculate M_{split} estimators $\hat{\mathbf{X}}_{\alpha}^j$ and $\hat{\mathbf{X}}_{\beta}^j$:

step 1(1): Based on the valid vector $\mathbf{v}_{\beta(l)}^j$, the following weight matrix is constructed

$$\mathbf{W}_{\alpha}(\mathbf{v}_{\beta(l)}^j) = \text{Diag}\left(\left(\mathbf{v}_{1\beta(l)}^j\right)^2 q_1^{-2}, \dots, \left(\mathbf{v}_{n\beta(l)}^j\right)^2 q_n^{-2}\right) \quad (45)$$

and the following is calculated:

$$\begin{aligned} \mathbf{X}_{\alpha(l+1)}^j &= \left(\mathbf{A}^T \mathbf{W}_{\alpha}(\mathbf{v}_{\beta(l)}^j) \mathbf{A}\right)^{-1} \mathbf{A}^T \mathbf{W}_{\alpha}(\mathbf{v}_{\beta(l)}^j) \mathbf{y} \\ \mathbf{v}_{\alpha(l+1)}^j &= \mathbf{y} - \mathbf{A}^j \mathbf{X}_{\alpha(l+1)}^j \end{aligned} \quad (46)$$

step 1(2): Based on the vector $\mathbf{v}_{\alpha(l+1)}^j$, the following weight matrix is constructed

$$\mathbf{W}_{\beta}(\mathbf{v}_{\alpha(l+1)}^j) = \text{Diag}\left(\left(\mathbf{v}_{1\alpha(l+1)}^j\right)^2 q_1^{-2}, \dots, \left(\mathbf{v}_{n\alpha(l+1)}^j\right)^2 q_n^{-2}\right) \quad (47)$$

and the following is calculated:

$$\begin{aligned} \mathbf{X}_{\beta(l+1)}^j &= \left(\mathbf{A}^T \mathbf{W}_{\beta}(\mathbf{v}_{\alpha(l+1)}^j) \mathbf{A}\right)^{-1} \mathbf{A}^T \mathbf{W}_{\beta}(\mathbf{v}_{\alpha(l+1)}^j) \mathbf{y} \\ \mathbf{v}_{\beta(l+1)}^j &= \mathbf{y} - \mathbf{A}^j \mathbf{X}_{\beta(l+1)}^j \end{aligned} \quad (48)$$

step 1(3): Repeat steps 1(1) and 1(2) until

$$\begin{aligned} \left\| \mathbf{X}_{\alpha(l+1)}^j - \mathbf{X}_{\alpha(l)}^j \right\| &< \varepsilon_0 \quad \text{and} \\ \left\| \mathbf{X}_{\beta(l+1)}^j - \mathbf{X}_{\beta(l)}^j \right\| &< \varepsilon_0 \quad (\text{for a given } \varepsilon_0) \end{aligned} \quad (49)$$

Once criterion (49) has been satisfied, the following M_{split} estimators in the j -th iteration:

$$\hat{\mathbf{X}}_{\alpha}^j = \mathbf{X}_{\alpha(l+1)}^j, \quad \hat{\mathbf{X}}_{\beta}^j = \mathbf{X}_{\beta(l+1)}^j \quad (50)$$

and residual vectors

$$\tilde{\mathbf{v}}_{\alpha}^j = \mathbf{y} - \mathbf{A}^j \hat{\mathbf{X}}_{\alpha}^j, \quad \tilde{\mathbf{v}}_{\beta}^j = \mathbf{y} - \mathbf{A}^j \hat{\mathbf{X}}_{\beta}^j \quad (51)$$

are adopted.

Step 2: For the determined iterative M_{split} estimators $\hat{\mathbf{X}}_{\alpha}^j$, $\hat{\mathbf{X}}_{\beta}^j$, and residual vectors $\tilde{\mathbf{v}}_{\alpha}^j$, $\tilde{\mathbf{v}}_{\beta}^j$, weight matrices that are valid in the j -th iteration are determined:

$$\begin{aligned} \mathbf{W}_{\alpha}(\tilde{\mathbf{v}}_{\beta}^j) &= \text{Diag}\left(\left(\tilde{\mathbf{v}}_{1\beta}^j\right)^2 q_1^{-2}, \dots, \left(\tilde{\mathbf{v}}_{n\beta}^j\right)^2 q_n^{-2}\right) \\ \mathbf{W}_{\beta}(\tilde{\mathbf{v}}_{\alpha}^j) &= \text{Diag}\left(\left(\tilde{\mathbf{v}}_{1\alpha}^j\right)^2 q_1^{-2}, \dots, \left(\tilde{\mathbf{v}}_{n\alpha}^j\right)^2 q_n^{-2}\right) \end{aligned} \quad (52)$$

Based on these matrices and taking $\hat{\mathbf{X}}_{\otimes\alpha}^j = \hat{\mathbf{X}}_{\alpha}^j \otimes \mathbf{I}_n$, $\hat{\mathbf{X}}_{\otimes\beta}^j = \hat{\mathbf{X}}_{\beta}^j \otimes \mathbf{I}_n$, we can determine matrix Θ^j , Eq. (30); then, the combined vector of increments is calculated:

$$\delta \hat{\mathbf{X}}^{j+1} = \left((\hat{\mathbf{A}} - \hat{\mathbf{E}}^j)^T (\Theta^j)^{-1} (\hat{\mathbf{A}} - \hat{\mathbf{E}}^j) \right)^{-1} (\hat{\mathbf{y}} - \hat{\mathbf{A}} \hat{\mathbf{X}}^j) \quad (53)$$

This vector enables the calculation of Lagrange multiplier vector that is valid in the $(j + 1)$ iteration

$$\hat{\lambda}^{j+1} = -(\Theta^j)^{-1} \left(\hat{\mathbf{y}} - \hat{\mathbf{A}} \hat{\mathbf{X}}^j - (\hat{\mathbf{A}} - \hat{\mathbf{E}}^j) \delta \hat{\mathbf{X}}^{j+1} \right) \quad (54)$$

and then the calculation of combined residual vectors

$$\tilde{\mathbf{v}}^{j+1} = \hat{\mathbf{W}}^{-1} (\tilde{\mathbf{v}}_{\alpha}^j, \tilde{\mathbf{v}}_{\beta}^j) (\Theta^j)^{-1} \left(\hat{\mathbf{y}} - \hat{\mathbf{A}} \hat{\mathbf{X}}^j - (\hat{\mathbf{A}} - \mathbf{E}^j) \delta \hat{\mathbf{X}}^{j+1} \right) \quad (55)$$

and

$$\tilde{\mathbf{e}}^{j+1} = -\mathbf{Q}_e \hat{\mathbf{X}}_{\otimes}^j (\Theta^j)^{-1} \left(\hat{\mathbf{y}} - \hat{\mathbf{A}} \hat{\mathbf{X}}^j - (\hat{\mathbf{A}} - \mathbf{E}^j) \delta \hat{\mathbf{X}}^{j+1} \right) \quad (56)$$

Based on vector $\tilde{\mathbf{e}}^{j+1}$, it is necessary to build a matrix of disturbances $\tilde{\mathbf{E}}^{j+1}$ and to calculate the vector of the following parameters which are valid in the $(j + 1)$ iteration:

$$\hat{\mathbf{X}}^{j+1} = \left((\hat{\mathbf{A}} - \hat{\mathbf{E}}^j)^T (\Theta^j)^{-1} (\hat{\mathbf{A}} - \mathbf{E}^j) \right)^{-1} (\hat{\mathbf{A}} - \hat{\mathbf{E}}^j)^T (\Theta^j)^{-1} (\hat{\mathbf{y}} - \hat{\mathbf{E}}^j \hat{\mathbf{X}}^j) \quad (57)$$

Step 3: Repeat Step 1 and Step 2 until

$$\left\| \hat{\mathbf{X}}^{j+1} - \hat{\mathbf{X}}^j \right\| < \varepsilon_0 \quad (58)$$

Once criterion (58) has been satisfied, it is assumed that TM_{split} estimator of the combined vector of $\mathbf{X} = [\mathbf{X}_\alpha^T, \mathbf{X}_\beta^T]^T$ parameters is vector $\hat{\mathbf{X}} = \hat{\mathbf{X}}^{j+1}$. The blocks of this vector are TM_{split} estimators $\hat{\mathbf{X}}_\alpha$ and $\hat{\mathbf{X}}_\beta$, the ‘‘removal’’ of which from vector $\hat{\mathbf{X}}$ may be facilitated by relationships

$$\hat{\mathbf{X}}_\alpha = \mathbf{D}_{X_\alpha} \hat{\mathbf{X}} \quad \text{and} \quad \hat{\mathbf{X}}_\beta = \mathbf{D}_{X_\beta} \hat{\mathbf{X}} \tag{59}$$

where $\mathbf{D}_{X_\alpha} = [\mathbf{I}_m, \mathbf{0}_{m,m}]$ and $\mathbf{D}_{X_\beta} = [\mathbf{0}_{m,m}, \mathbf{I}_m]$ ($\mathbf{0}_{m,m}$ —zero matrix with dimensions of $m \times m$). Once the iterative process is complete, the final residual vectors $\tilde{\mathbf{v}} = \tilde{\mathbf{v}}^k$ and $\tilde{\mathbf{e}} = \tilde{\mathbf{e}}^k$ are also determined. Based on vector $\tilde{\mathbf{v}} = [\tilde{\mathbf{v}}_\alpha^T, \tilde{\mathbf{v}}_\beta^T]^T$, two versions of the residual vector corresponding to the observation vector \mathbf{y} , i.e.

$$\tilde{\mathbf{v}}_\alpha = \mathbf{D}_{v_\alpha} \tilde{\mathbf{v}} \quad \text{and} \quad \tilde{\mathbf{v}}_\beta = \mathbf{D}_{v_\beta} \tilde{\mathbf{v}} \tag{60}$$

where $\mathbf{D}_{v_\alpha} = [\mathbf{I}_n, \mathbf{0}_{n,n}]$ and $\mathbf{D}_{v_\beta} = [\mathbf{0}_{n,n}, \mathbf{I}_n]$, are obtained. On the other hand, vector $\tilde{\mathbf{e}}$ provides the basis for the construction of residual matrix $\tilde{\mathbf{E}}$ corresponding to matrix \mathbf{A} , hence also $\hat{\mathbf{A}} = \mathbf{A} - \tilde{\mathbf{E}}$.

The given above process of iterative determination of TM_{split} estimates is also described by the flowchart presented in Fig. 1

4 Examples

4.1 Example 1: competitive models of systematic errors

In one of the examples provided in a study by Wiśniewski (2010), it was assumed that $y_i, i = 1, \dots, n$ were observations of a certain value of Y disturbed not only with random errors v_i but also with systematic errors $s_i = s(t_i) = a + bt_i$ (e.g. Wiśniewski 1985; Kubáčková and Kubáček 1991; Yang and Zhang 2005). The problem, however, is that two versions of this model can be used: $s_\alpha(t_i) = a_\alpha + b_\alpha t_i$ and $s_\beta(t_i) = a_\beta + b_\beta t_i$, whereas it is not known which of them concerns specific observation y_i . For this reason, in M_{split} estimation, the classical observation model

$$y_i = Y + s(t_i) + v_i = (Y + a) + bt_i + v_i = X + bt_i + v_i \tag{61}$$

is split into the following models:

$$\begin{aligned} y_i &= (Y + a_\alpha) + b_\alpha t_i + v_{i\alpha} = X_\alpha + b_\alpha t_i + v_{i\alpha} \\ y_i &= (Y + a_\beta) + b_\beta t_i + v_{i\beta} = X_\beta + b_\beta t_i + v_{i\beta} \end{aligned} \tag{62}$$

where $X = Y + a$. In these models, two mutually competing versions of the parameters occur, namely $X_\alpha = Y + a_\alpha$,

$X_\beta = Y + a_\beta$. Observations were simulated with the assumption of theoretical values of the parameters X_α, b_α and X_β, b_β . Theoretical observations $\bar{y}_i, i = 1, \dots, 10$, were affected by Gaussian errors with the expected value of 0, and standard deviation of σ_y . For theoretical values $X_\alpha = 6.0, b_\alpha = 0.5, X_\beta = 3.0, b_\beta = 1.0$, and standard deviation $\sigma_y = 0.14$, the set of observations presented in Table 1 was obtained. The table also presents M_{split} estimates, namely $\hat{X}_\alpha, \hat{b}_\alpha, \hat{X}_\beta, \hat{b}_\beta$, the mutual competitive residuals $\hat{v}_{i\alpha}, \hat{v}_{i\beta}$ and weights related to such residuals $w_\alpha(\hat{v}_{i\beta}) = \hat{v}_{i\beta}^2 q_i^{-2}, w_\beta(\hat{v}_{i\alpha}) = \hat{v}_{i\alpha}^2 q_i^{-2}$ (for $q_i^{-2} = \sigma_y^{-4}$). A graphical illustration of the set of observations, and a graphical interpretation of the obtained results (as compared to LS-estimators determined using model (61)), are shown in Fig. 2. For the sake of clarity, the figure shows the competitive residuals only of the observation y_{10} . In this figure, the mutually competing results of M_{split} estimation are conventionally denoted as $M_{\text{split}(\alpha)}$ and $M_{\text{split}(\beta)}$ (when it is convenient, these notations will also be used further on in this paper).

The models contained in Eq. (62) will now be replaced with EIV models of the following form:

$$\begin{aligned} y_i &= X_\alpha + b_\alpha(t_i - e_{t_i}) + v_{i\alpha} \\ y_i &= X_\beta + b_\beta(t_i - e_{t_i}) + v_{i\beta} \end{aligned} \tag{63}$$

where e_{t_i} is a random error affected to the variable t_i . For n observations, based on Eq. (63), models $\mathbf{y} = (\mathbf{A} - \mathbf{E})\mathbf{X}_\alpha + \mathbf{v}_\alpha$ and $\mathbf{y} = (\mathbf{A} - \mathbf{E})\mathbf{X}_\beta + \mathbf{v}_\beta$ are constructed, where

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} 1_1 & t_1 \\ \vdots & \vdots \\ 1_n & t_n \end{bmatrix}, \\ \mathbf{E} &= \begin{bmatrix} 0_1 & e_{t_1} \\ \vdots & \vdots \\ 0_n & e_{t_n} \end{bmatrix} = [\mathbf{0}_n, \mathbf{e}_t], \\ \mathbf{X}_\alpha &= \begin{bmatrix} X_\alpha \\ b_\alpha \end{bmatrix}, \quad \mathbf{X}_\beta = \begin{bmatrix} X_\beta \\ b_\beta \end{bmatrix} \end{aligned} \tag{64}$$

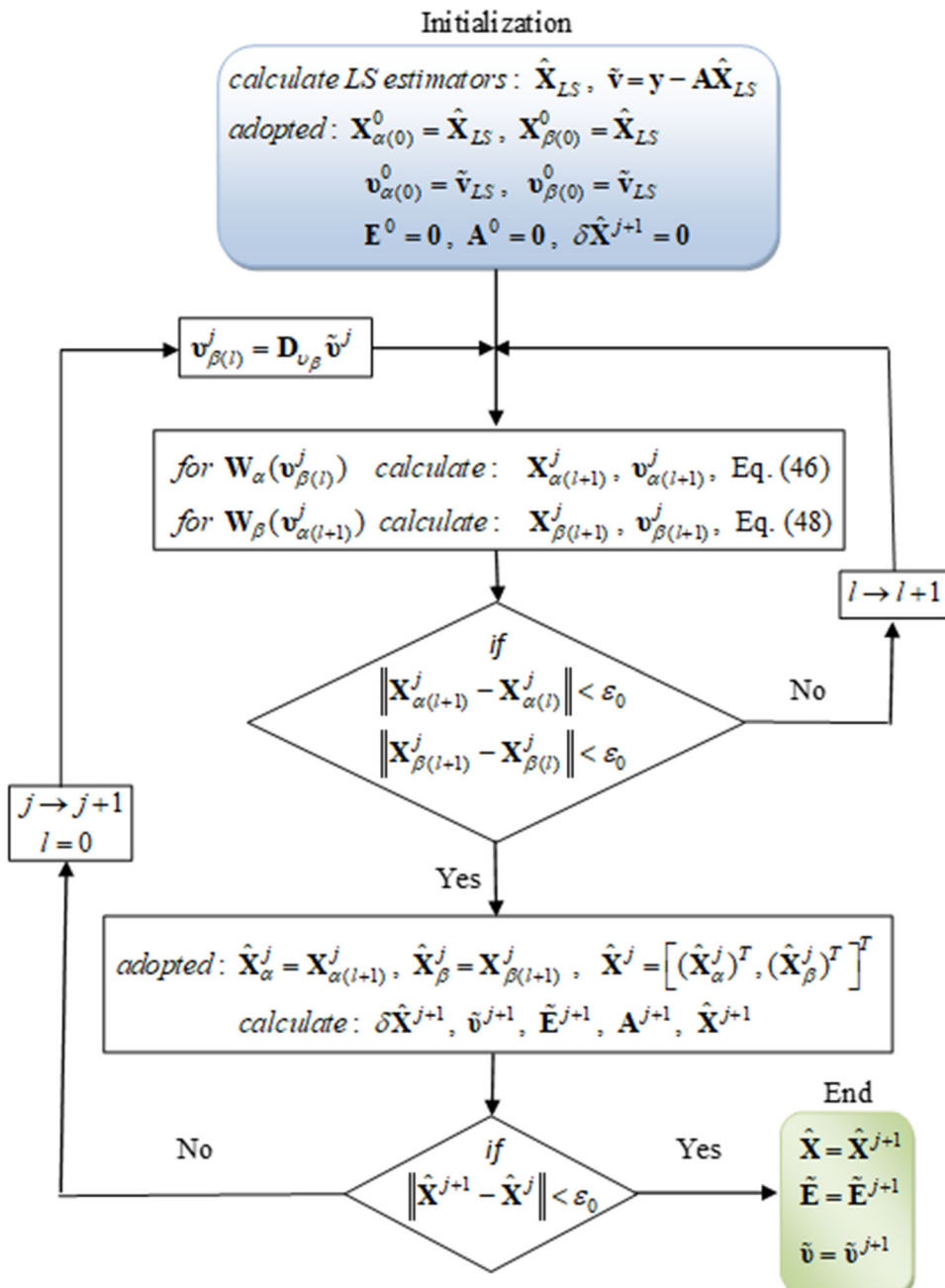
(the first column of matrix \mathbf{A} is not random). Vector $\mathbf{e} = \text{vec}(\mathbf{E})$, built from matrix \mathbf{E} columns, has the following form:

$$\mathbf{e} = [0_1, \dots, 0_n, e_{t_1}, \dots, e_{t_n}]^T = [\mathbf{0}_n^T, \mathbf{e}_t^T]^T \tag{65}$$

In view of the structure of this vector, \mathbf{Q}_e cofactor matrix, similarly as in Schaffrin and Wieser (2008) and Shen et al. (2011), will be expressed in the following form:

$$\mathbf{Q}_e = \mathbf{Q}_0 \otimes \mathbf{Q}_x = \mathbf{Q}_0 \otimes \mathbf{Q}_{e_t} \quad \text{with} \quad \mathbf{Q}_0 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{66}$$

Fig. 1 Flochart of the Total M_{split} estimation



where \mathbf{Q}_{e_i} is the cofactor matrix of vector \mathbf{e}_i . Note that here \mathbf{Q}_{e_i} is regular, but \mathbf{Q}_0 is not; thus, \mathbf{Q}_e is not regular too (the similar situation is in the example given by Schaffrin and Wieser 2008). Random errors e_{t_i} are simulated as Gaussian quantities with the expected value of 0, and standard deviation of σ_e . M_{split} and TM_{split} estimators of model (63) parameters will be determined for four values of standard deviation: $\sigma_e = 0$ (variant I), $\sigma_e = 0.13$ (variant II), $\sigma_e = 0.28$ (variant III), $\sigma_e = 0.37$ (variant IV). In each of these variants, observations $y_i, i = 1, \dots, 10$, are as in the example cited above (Table 1). The data adopted for the calculations are listed in Table 2, while the M_{split} and TM_{split}

estimators obtained for these data are presented in Table 3. Additionally, the residuals \tilde{e}_{t_i} are presented in Table 4. Table 5 shows the competitive residuals and the respective weights for variant III.

In variant I, since matrix \mathbf{A} is not disturbed by random disturbances, TM_{split} estimators are equal to M_{split} estimators. With an increase in the σ_e standard deviation value, the differences between these estimators increase, with TM_{split} estimators remaining close to the theoretical values (as do M_{split} estimators for the variant without random errors of matrix \mathbf{A} , $\sigma_e = 0$). This is well illustrated by the norm values of the vector of differences between the vector of theoretical

Table 1 Observed data and results of M_{split} estimation (Wiśniewski 2010)

True value and observations				Results of M_{split} estimate			
i	\bar{y}_i	v_i	$y_i = \bar{y}_i + v_i$	$\hat{v}_{i\alpha}$	$w_\alpha(\hat{v}_{i\beta})$	$\hat{v}_{i\beta}$	$w_\beta(\hat{v}_{i\alpha})$
1	4.0	0.1	4.1	- 2.55	9.37	0.06	16,989.77
2	7.0	0.2	7.2	0.05	12,484.64	2.19	5.94
3	6.0	- 0.2	5.8	- 1.85	89.74	- 0.18	8906.32
4	8.0	0.0	8.0	- 0.15	2813.93	1.04	56.40
5	8.0	0.1	8.1	- 0.54	70.95	0.16	772.26
6	9.0	- 0.1	8.9	- 0.24	0.23	0.00	152.64
7	10.0	0.2	10.2	0.56	259.73	0.32	817.40
8	10.0	- 0.1	9.9	- 0.24	2392.68	- 0.96	146.35
9	12.0	- 0.1	11.9	1.26	11.56	0.07	4168.21
10	11.0	0.2	11.2	0.07	6730.35	- 1.61	12.01
	$X_\alpha = 6.0,$ $X_\beta = 3.0$		$b_\alpha = 0.5$ $b_\beta = 1.0$		$\hat{X}_\alpha = 6.16,$ $\hat{X}_\beta = 3.06$		$\hat{b}_\alpha = 0.50$ $\hat{b}_\beta = 0.97$

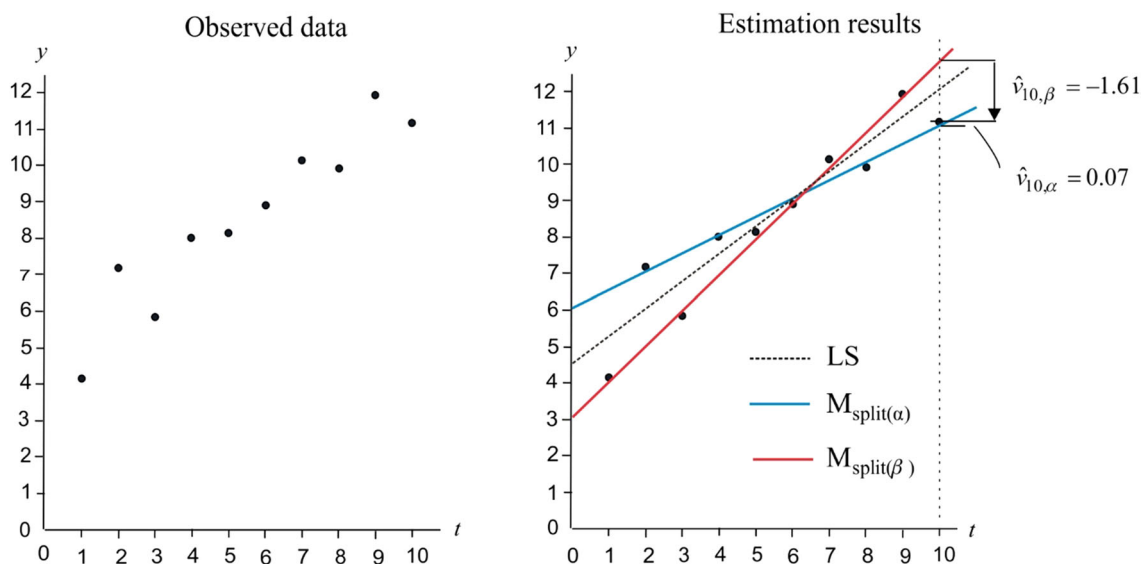


Fig. 2 Set of observations and results of M_{split} estimation (as compared to the results of LS-estimation). The example competitive residuals $\hat{v}_{10,\alpha}$ and $\hat{v}_{10,\beta}$ for the observation $y_{10} = 11.2$

parameters $\mathbf{X} = [\mathbf{X}_\alpha^T, \mathbf{X}_\beta^T]^T$ and the vector of obtained estimators $\hat{\mathbf{X}} = [\hat{\mathbf{X}}_\alpha^T, \hat{\mathbf{X}}_\beta^T]^T$, provided in the last row of Table 3.

In general, the iterative process involved in the determination of TM_{split} estimators ended after 4 to 6 steps of “external” iteration. In each of these steps, 6 to 7 “internal” iterations were carried out resulting in M_{split} estimators that are valid for this step. The course of the iterative process in Total M_{split} estimation, based on the example of variant II, is presented in Fig. 3.

The given examples apply one observation set, respectively. The additional analyses will be based on Monte Carlo (MC) simulations. The main objective is to determine the

empirical accuracy of Total M_{split} estimation and the measures of its efficacy.

The accuracy of M_{split} estimation can be determined by applying asymptotical covariance matrices proposed in Wiśniewski and Zienkiewicz (2021b). The diagonal elements allow us to compute the estimated standard deviation $\sigma_{\hat{X}_\alpha}, \sigma_{\hat{b}_\alpha}, \sigma_{\hat{X}_\beta}, \sigma_{\hat{b}_\beta}$ of the respective M_{split} estimates. To apply such an approach to Total M_{split} estimation, one should develop the theory presented in the paper mentioned, which is beyond the scope of the present paper. Based on simulated observation sets, an empirical way could be an alternative for the analytical assessment in question. Total M_{split} estimates $\hat{X}_\alpha^k, \hat{b}_\alpha^k, \hat{X}_\beta^k, \hat{b}_\beta^k$ are computed for each $k = 1, \dots, N$ simulation,

Table 2 Observed data for Total M_{split} estimation (variant I, II, III, IV)

i	Observations y_i (Wiśniewski 2010)	True value \bar{t}_i	Errors e_{t_i}				Observations t_i			
			I	II	III	IV	I	II	III	IV
1	4.1	1.0	0	0.1	0.2	-0.4	0	1.1	1.2	0.6
2	7.2	2.0	0	-0.1	-0.3	0.3	0	1.9	1.7	2.3
3	5.8	3.0	0	-0.2	-0.2	0.4	0	2.8	2.8	3.4
4	8.0	4.0	0	0.1	0.4	-0.5	0	4.1	4.4	3.5
5	8.1	5.0	0	0.0	-0.1	-0.3	0	5.0	4.9	4.7
6	8.9	6.0	0	0.2	0.4	-0.4	0	6.2	6.4	5.6
7	10.2	7.0	0	-0.1	-0.2	0.3	0	6.9	6.8	7.3
8	9.9	8.0	0	-0.2	0.3	0.3	0	7.8	8.3	8.3
9	11.9	9.0	0	0.0	-0.4	0.5	0	9.0	8.6	9.5
10	11.2	10.0	0	0.2	-0.1	-0.2	0	10.2	9.9	9.8

Table 3 Results of M_{split} and Total M_{split} estimation (variant I, II, III, IV)

Parameter	True value	Estimate							
		Variant							
		I ($\sigma_e = 0$)		II ($\sigma_e = 0.13$)		III ($\sigma_e = 0.28$)		IV ($\sigma_e = 0.37$)	
		M_{split}	TM_{split}	M_{split}	TM_{split}	M_{split}	TM_{split}	M_{split}	TM_{split}
X_α	6.0	6.16	6.16	6.25	6.13	6.36	6.07	5.66	6.07
b_α	0.5	0.50	0.50	0.48	0.49	0.47	0.49	0.63	0.50
X_β	3.0	3.06	3.06	3.02	3.01	2.87	2.89	3.48	3.13
b_β	1.0	0.97	0.97	0.99	1.00	1.04	1.06	0.77	0.92
$\sigma_{0\alpha}$	-	0.731	0.731	0.510	0.360	2.023	0.549	2.372	0.608
$\sigma_{0\beta}$	-	1.403	1.403	0.334	0.155	0.619	0.229	1.655	0.745
$\ \widehat{\mathbf{X}} - \mathbf{X}\ $		0.173	0.173	0.251	0.131	0.386	0.144	0.649	0.168

Table 4 Residuals \tilde{e}_t column of A

Variant	\tilde{e}_t									
I	0	0	0	0	0	0	0	0	0	0
II	0.02	-0.27	0.02	0.27	-0.07	0.20	-0.25	0.10	0.13	-0.14
III	0.02	-0.59	0.03	0.48	-0.04	0.05	-0.11	0.54	0.10	-0.48
IV	-0.27	0.06	0.69	-0.32	-0.49	-0.01	-0.16	0.67	0.19	-0.36

which is a base for determining the following MC-estimates of the parameters $X_\alpha, b_\alpha, X_\beta, b_\beta$:

$$\hat{X}_\alpha^{MC} = \frac{1}{N} \sum_{k=1}^N \hat{X}_\alpha^k, \quad \hat{b}_\alpha^{MC} = \frac{1}{N} \sum_{k=1}^N \hat{b}_\alpha^k, \quad (67)$$

$$\hat{X}_\beta^{MC} = \frac{1}{N} \sum_{k=1}^N \hat{X}_\beta^k, \quad \hat{b}_\beta^{MC} = \frac{1}{N} \sum_{k=1}^N \hat{b}_\beta^k$$

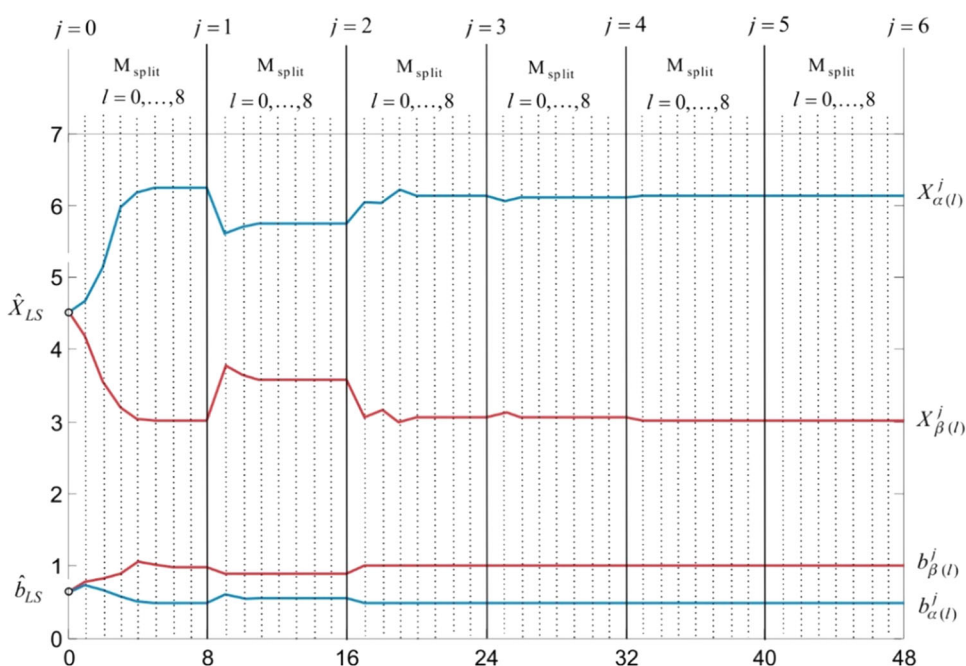
The Monte Carlo estimators of the parameter standard deviations can be computed in the following way (e.g. Koch 2013; Nowel 2016; Lv and Sui 2020).

$$\hat{\sigma}_{\hat{X}_\alpha}^{MC} = \sqrt{\frac{1}{N} \sum_{k=1}^N (\hat{X}_\alpha^k - \hat{X}_\alpha^{MC})^2},$$

Table 5 Residuals and weights in M_{split} and Total M_{split} estimations (variant III, $\sigma_e = 0.28$)

i	M_{split}				TM_{split}			
	$\hat{v}_{i\alpha}$	$w_\alpha(\hat{v}_{i\beta})$	$\hat{v}_{i\beta}$	$w_\beta(\hat{v}_{i\alpha})$	$\hat{v}_{i\alpha}$	$w_\alpha(\hat{v}_{i\beta})$	$\hat{v}_{i\beta}$	$w_\beta(\hat{v}_{i\alpha})$
1	-2.66	$1.5 \cdot 10^{-6}$	0.00	$1.5 \cdot 10^6$	-2.82	0.72	-0.02	$2.1 \cdot 10^4$
2	0.00	$8.0 \cdot 10^3$	2.11	$4.6 \cdot 10^{-2}$	0.05	$1.7 \cdot 10^4$	2.56	6.16
3	-1.77	$3.4 \cdot 10^{-5}$	0.00	$6.4 \cdot 10^3$	-1.87	1.01	0.02	$9.1 \cdot 10^3$
4	-0.01	$2.1 \cdot 10^3$	1.04	$4.6 \cdot 10^{-2}$	-0.42	$8.0 \cdot 10^2$	0.56	$4.5 \cdot 10^2$
5	-0.43	$1.7 \cdot 10^{-2}$	0.00	$4.0 \cdot 10^2$	-0.55	48.34	0.14	$7.8 \cdot 10^2$
6	-0.34	3.66	-0.81	$5.9 \cdot 10^{-3}$	-0.45	$1.0 \cdot 10^3$	-0.62	$5.3 \cdot 10^2$
7	0.75	$6.6 \cdot 10^{-3}$	0.00	$1.2 \cdot 10^3$	0.66	$1.8 \cdot 10^2$	0.26	$1.1 \cdot 10^3$
8	0.00	$3.1 \cdot 10^3$	-1.34	$3.3 \cdot 10^{-2}$	-0.34	$6.6 \cdot 10^3$	-1.60	$3.0 \cdot 10^2$
9	1.70	$2.3 \cdot 10^{-4}$	0.00	$6.3 \cdot 10^3$	1.52	20.72	0.09	$6.0 \cdot 10^3$
10	0.00	$1.6 \cdot 10^4$	-2.91	$9.2 \cdot 10^{-4}$	0.21	$1.0 \cdot 10^4$	-1.96	$1.1 \cdot 10^2$

Fig. 3 Iterative process resulting in TM_{split} estimators of parameters X_α , b_α and parameters X_β , b_β competing in relation to them (j —the number of “external” iterations, l —the number of “internal” iterations)



$$\begin{aligned}
 \hat{\sigma}_{\hat{b}_\alpha}^{MC} &= \sqrt{\frac{1}{N} \sum_{k=1}^N (\hat{b}_\alpha^k - \hat{b}_\alpha^{MC})^2} \\
 \hat{\sigma}_{\hat{X}_\beta}^{MC} &= \sqrt{\frac{1}{N} \sum_{k=1}^N (\hat{X}_\beta^k - \hat{X}_\beta^{MC})^2}, \\
 \hat{\sigma}_{\hat{b}_\beta}^{MC} &= \sqrt{\frac{1}{N} \sum_{k=1}^N (\hat{b}_\beta^k - \hat{b}_\beta^{MC})^2}
 \end{aligned} \tag{68}$$

M_{split} estimation and its several developments (including Total M_{split} estimation) focus on optimal fitting the competitive functional models in the observation set. The efficacy of M_{split} estimation, like the efficacy of other methods, can be described by the differences between the parameter estimates obtained and the actual parameter values. Considering N simulations, the efficacy is usually measured by the root mean squared error (RMSE) (e.g. Kargoll et al. 2018; Lv and Sui 2020). Here, the efficacy of M_{split} or Total M_{split} estimates is determined by following RMSEs

Such quantities determine the accuracy of the estimates obtained. They can also be used to compare the stability of M_{split} and Total M_{split} estimators.

$$\text{RMSE}_{\hat{X}_\alpha} = \sqrt{\frac{1}{N} \sum_{k=1}^N (\hat{X}_\alpha^k - X_\alpha)^2},$$

$$\begin{aligned}
 \text{RMSE}_{\hat{b}_\alpha} &= \sqrt{\frac{1}{N} \sum_{k=1}^N (\hat{b}_\alpha^k - b_\alpha)^2} \\
 \text{RMSE}_{\hat{X}_\beta} &= \sqrt{\frac{1}{N} \sum_{k=1}^N (\hat{X}_\beta^k - X_\beta)^2}, \\
 \text{RMSE}_{\hat{b}_\beta} &= \sqrt{\frac{1}{N} \sum_{k=1}^N (\hat{b}_\beta^k - b_\beta)^2} \tag{69}
 \end{aligned}$$

Additionally, the global root mean squared error, concerning the whole parameter vector, is determined (e.g. Wiśniewski 2014)

$$\text{RMSE}_{\hat{\mathbf{X}}} = \sqrt{\frac{1}{N} \sum_{k=1}^N (\hat{\mathbf{X}}^k - \mathbf{X})^T (\hat{\mathbf{X}}^k - \mathbf{X}) / r} \tag{70}$$

where r is the number of estimated parameters (here $r = 4$).

The simulated errors $v_i^k, e_{i,i}^k, i = 1, \dots, 10$ (for each $k = 1, \dots, N$) affect the theoretical observations \bar{y}_i or the elements $a_{i,2}$ of the second column of the matrix \mathbf{A} (the theoretical observations and the theoretical matrix \mathbf{A} stay the same). The simulations are performed by applying the Gaussian random generators of the MatLab system, $\sigma_y \text{randn}(n, 1)$ or $\sigma_e \text{randn}(n, 1)$, respectively.

First, let us examine the efficacy of M_{split} estimates, which apply the models of (62). Since matrix \mathbf{A} is constant, only observation errors are simulated. The computations are determined in several variants of the observation standard deviation, i.e. $\sigma_y = 0.05, 0.1, 0.2, 0.3$. Table 6 presents results obtained for $N = 3000$ and the theoretical parameter values $\mathbf{X} = [X_\alpha, b_\alpha, X_\beta, b_\beta]^T = [6.0, 0.5, 3.0, 1.0]^T$. Figure 4 shows M_{split} estimates obtained in each simulation and MC estimates (for $\sigma_y = 0.1$).

The accuracy and efficacy of the estimates \hat{b}_α and \hat{b}_β are the most satisfying. The values of $\hat{\sigma}_{b_\alpha}^{MC}, \hat{\sigma}_{b_\beta}^{MC}$ and $\text{RMSE}_{\hat{b}_\alpha}, \text{RMSE}_{\hat{b}_\beta}$ are relatively small for all values of the observation standard deviations. The values obtained for the estimates $\hat{X}_\alpha, \hat{X}_\beta$ are higher; however, they are still acceptable.

Let us now use the models (63) to examine how random disturbances of matrix \mathbf{A} might influence the accuracy and efficacy of the estimates. The observation errors are simulated assuming the constant standard deviation $\sigma_y = 0.1$, whereas the errors $e_{t,i}$ in several variants, in which $\sigma_e = 0, 0.05, 0.1, 0.2, 0.3$. Table 7 presents the results for $N = 3000$.

Both accuracy and efficacy of M_{split} estimates decrease when the coefficient matrix is disturbed with random errors. That effect can be reduced by using Total M_{split} estimation, for which the measures in questions are smaller. As in the previous case, TM_{split} estimates of the parameters b_α and b_β have the most satisfying accuracy and efficacy. The efficacy

of Total M_{split} estimation is confirmed by the parameter estimates obtained in each simulation. The example estimates and the MC estimates are presented in Fig. 5 (for $\sigma_e = 0.2$).

4.2 Example 2: linear regression

Schaffrin and Wieser (2008) as well as Shen et al. (2011) and Mahboub (2012) applied WTLS for the estimation of the intercept ξ_1 and slope ξ_2 of the regression line

$$y_i = \xi_1 + (x_i - e_i)\xi_2 - v_i \tag{71}$$

Let it now be assumed that the set of observations not only contains observations concerning model (71) but also observations for which the regression line differs in parameters ξ_1 and ξ_2 (in Total M_{split} estimation, these will be parameters $\xi_{1\beta}$ and $\xi_{2\beta}$). These observations will hereinafter be referred to conventionally as outliers. In contrast to the classical approach, the deviation here is of a different nature, and is not necessarily related to the existence of gross errors. If the assignment of observation y_i to its respective regression line is not known, then this observation may correspond to both the following model

$$y_i = \xi_1 + (x_i - e_i)\xi_2 - v_i = \xi_{1\alpha} + (x_i - e_i)\xi_{2\alpha} - v_{i\alpha} \tag{72}$$

and to the model that is competing in relation to it, namely:

$$y_i = \xi_{1\beta} + (x_i - e_i)\xi_{2\beta} - v_{i\beta} \tag{73}$$

For the estimation of parameters in models (72) and (73), Total M_{split} estimation will be applied. The calculations will be carried out using the data provided in (Neri et al. 1989) and also used in (Schaffrin and Wieser 2008; Shen et al. 2011; Mahboub 2012). These data will be supplemented with two variants of bias. In the first of these variants, regression line (73) with theoretical parameters $\xi_{1\beta} = 2.0, \xi_{2\beta} = 0.75$ is adopted, while in the second variant, $\xi_{1\beta} = 4.5, \xi_{2\beta} = -0.70$ is adopted. For the outliers, weights W_{x_i} were determined based on the information on the weights of coordinates x_i concerning the group of original observations (where necessary, also using interpolation). Moreover, the outliers are assigned equal weights $W_{y_i} = 50$, which corresponds to the standard deviation $\sigma_y = 0.14$. The values of these weights are very important to the success of Total M_{split} estimation. Too high accuracy (high weights) of the added observations may cause the original observations to be “ignored”, while too low accuracy (low weights) may cause the added observations to be “ignored”. In the presented example, satisfying results that were not much different from each other were obtained for $40 < W_{y_i} < 80$. Original observations ($i = 1, \dots, 10$) (Neri et al. 1989), outliers ($i = 11, 12, 13$), and the

Table 6 M_{split} estimates and their accuracy and efficacy (for sets without random disturbances in coefficients in the functional models)

	\hat{X}_α^{MC}	\hat{b}_α^{MC}	\hat{X}_β^{MC}	\hat{b}_β^{MC}	RMSE $_{\hat{X}}$
	$\hat{\sigma}_{\hat{X}_\alpha}^{MC}$	$\hat{\sigma}_{\hat{b}_\alpha}^{MC}$	$\hat{\sigma}_{\hat{X}_\beta}^{MC}$	$\hat{\sigma}_{\hat{b}_\beta}^{MC}$	
	RMSE $_{\hat{X}_\alpha}$	RMSE $_{\hat{b}_\alpha}$	RMSE $_{\hat{X}_\beta}$	RMSE $_{\hat{b}_\beta}$	
$\sigma_y = 0.05$	5.93	0.52	3.05	0.97	0.08
	0.20	0.06	0.14	0.07	
$\sigma_y = 0.1$	0.21	0.07	0.15	0.07	0.09
	5.95	0.52	3.04	0.99	
$\sigma_y = 0.2$	0.25	0.06	0.17	0.05	0.25
	0.26	0.06	0.17	0.06	
$\sigma_y = 0.3$	5.79	0.57	3.14	0.93	0.32
	0.42	0.11	0.31	0.11	
$\sigma_y = 0.3$	0.47	0.13	0.34	0.13	0.32
	5.77	0.58	3.15	0.91	
	0.52	0.12	0.39	0.12	
	0.56	0.15	0.42	0.15	

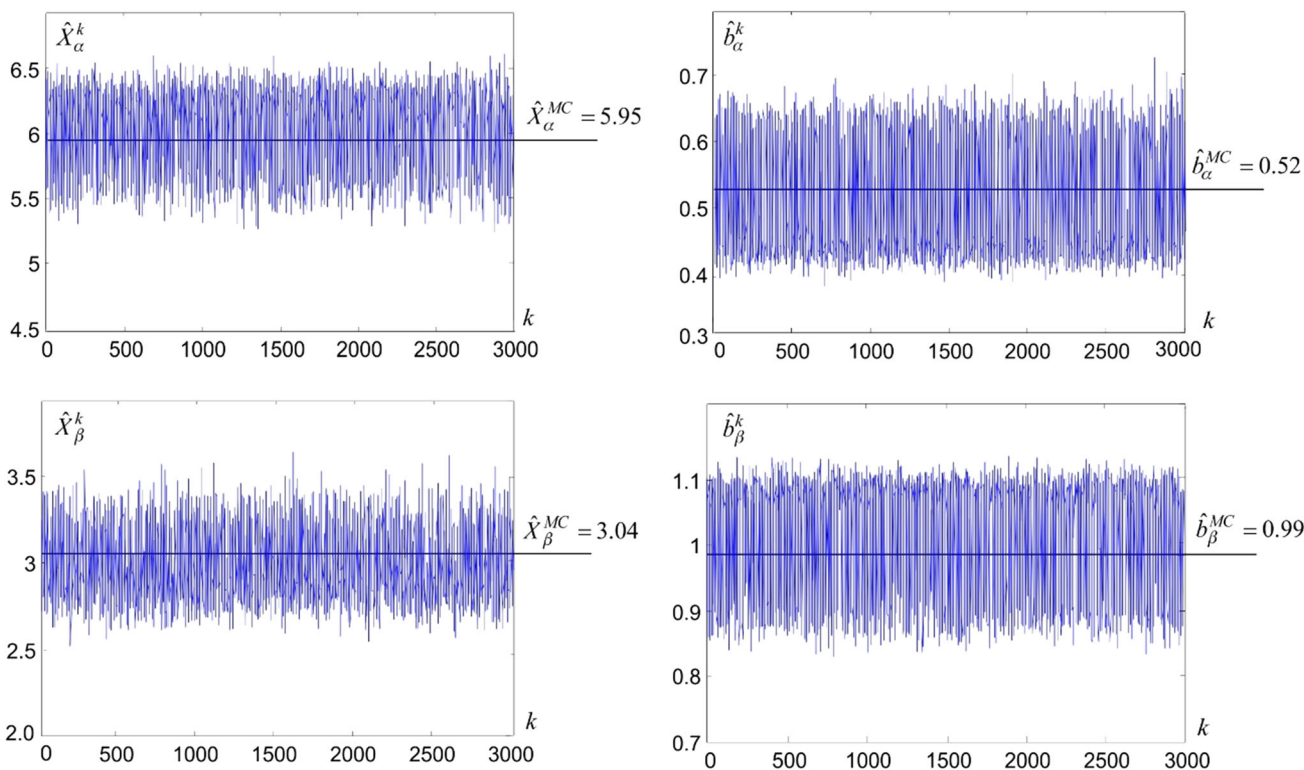


Fig. 4 M_{split} estimates obtained in each simulation $k = 1, \dots, 3000$ (for $\sigma_y = 0.1$ and the theoretical values of parameters $X_\alpha = 6.0, b_\alpha = 0.5, X_\beta = 3.0, b_\beta = 1.0$)

weights W_{x_i} and W_{y_i} , corresponding to these observations, are listed in Table 8. The results of model (71) parameter estimation using WTLS for original observations, transcribed from Shen et al. (2011), are provided in columns 2 and 3 of Table 9 [the estimate $\hat{\sigma}_0$ is computed by applying Eq. (19)]. Based on original observations and using models (72) and

(73), TM_{split} estimators of the parameters occurring there were calculated (column 4, Table 9). A graphical interpretation of the original set of observations and the location of regression lines (determined based on the WTLS and TM_{split} estimators) in this set are shown in Fig. 6. On the other hand, the WTLS and TM_{split} estimators determined for the sets

Table 7 M_{split} estimates and their accuracy and efficacy (for sets with random disturbances in coefficients in the functional models)

	M_{split}					TM_{split}				
	\hat{X}_α^{MC}	\hat{b}_α^{MC}	\hat{X}_β^{MC}	\hat{b}_β^{MC}	RMS $_{\hat{X}}$	\hat{X}_α^{MC}	\hat{b}_α^{MC}	\hat{X}_β^{MC}	\hat{b}_β^{MC}	RMS $_{\hat{X}}$
	$\hat{\sigma}_{\hat{X}_\alpha}^{MC}$	$\hat{\sigma}_{\hat{b}_\alpha}^{MC}$	$\hat{\sigma}_{\hat{X}_\beta}^{MC}$	$\hat{\sigma}_{\hat{b}_\beta}^{MC}$		$\hat{\sigma}_{\hat{X}_\alpha}^{MC}$	$\hat{\sigma}_{\hat{b}_\alpha}^{MC}$	$\hat{\sigma}_{\hat{X}_\beta}^{MC}$	$\hat{\sigma}_{\hat{b}_\beta}^{MC}$	
	RMSE $_{\hat{X}_\alpha}$	RMSE $_{\hat{b}_\alpha}$	RMSE $_{\hat{X}_\beta}$	RMSE $_{\hat{b}_\beta}$		RMSE $_{\hat{X}_\alpha}$	RMSE $_{\hat{b}_\alpha}$	RMSE $_{\hat{X}_\beta}$	RMSE $_{\hat{b}_\beta}$	
$\sigma_e = 0$	5.95	0.52	3.04	0.99	0.09	5.95	0.52	3.04	0.99	0.09
	0.25	0.06	0.17	0.05		0.25	0.06	0.17	0.05	
	0.26	0.06	0.17	0.06		0.26	0.06	0.17	0.06	
$\sigma_e = 0.05$	5.83	0.56	3.12	0.94	0.17	5.97	0.51	3.01	0.99	0.08
	0.31	0.09	0.22	0.09		0.20	0.04	0.14	0.04	
	0.36	0.11	0.25	0.11		0.20	0.04	0.14	0.04	
$\sigma_e = 0.1$	5.79	0.57	3.15	0.92	0.20	5.97	0.51	3.01	0.99	0.09
	0.33	0.09	0.24	0.10		0.21	0.04	0.15	0.04	
	0.39	0.12	0.28	0.12		0.21	0.04	0.15	0.04	
$\sigma_e = 0.2$	5.76	0.58	3.16	0.92	0.24	5.93	0.51	3.01	0.99	0.15
	0.38	0.10	0.31	0.23		0.32	0.06	0.24	0.06	
	0.45	0.13	0.35	0.13		0.33	0.07	0.24	0.06	
$\sigma_e = 0.3$	5.75	0.58	3.14	0.92	0.29	5.88	0.53	3.03	0.99	0.21
	0.44	0.11	0.39	0.12		0.41	0.08	0.35	0.09	
	0.50	0.14	0.42	0.14		0.43	0.09	0.35	0.09	

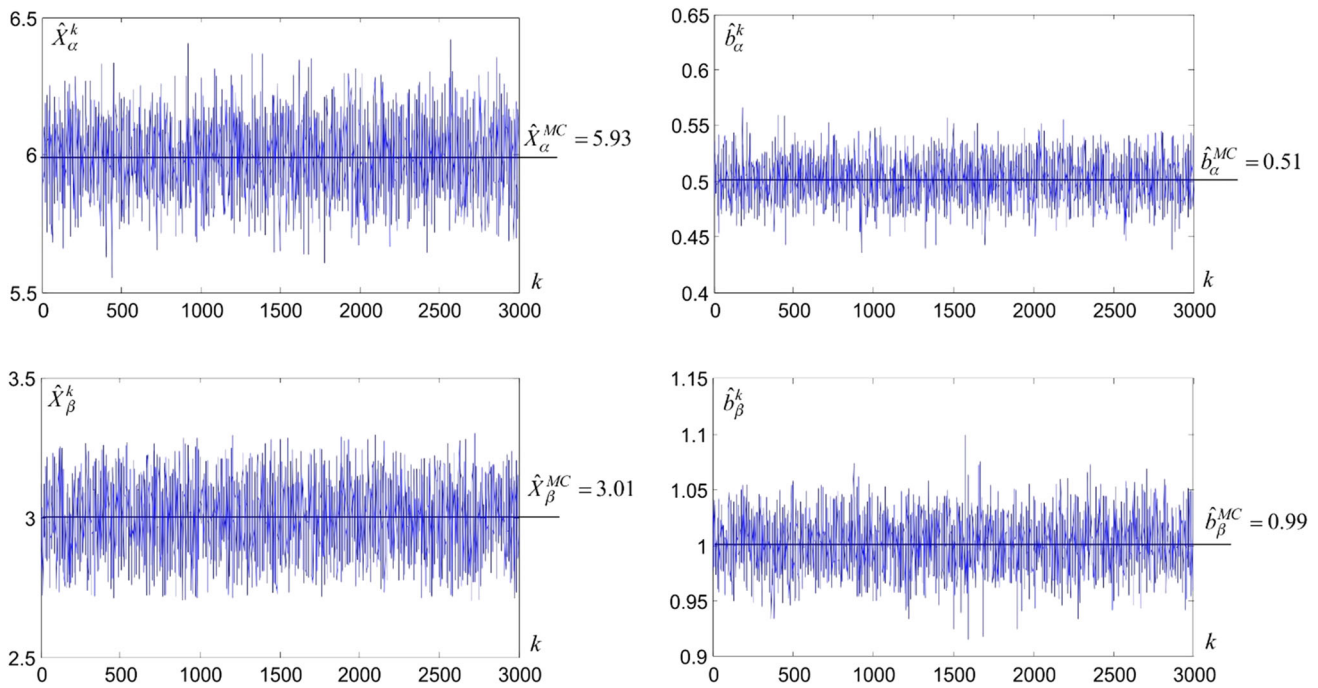


Fig. 5 TM_{split} estimates obtained in each simulation $k = 1, \dots, 3000$ (for $\sigma_y = 0.1, \sigma_e = 0.2$ and the theoretical values of parameters $X_\alpha = 6.0, b_\alpha = 0.5, X_\beta = 3.0, b_\beta = 1.0$)

Table 8 Observed data and corresponding weights

i	x_i	W_{x_i}	y_i	W_{y_i}
<i>Original data</i> (Neri et al. 1989)				
1	0.0	1000.0	5.9	1.0
2	0.9	1000.0	5.4	1.8
3	1.8	500.0	4.4	4.0
4	2.6	800.0	4.6	8.0
5	3.3	200.0	3.5	20.0
6	4.4	80.0	3.7	20.0
7	5.2	60.0	2.8	70.0
8	6.1	20.0	2.8	70.0
9	6.5	1.8	2.4	100.0
10	7.4	1.0	1.5	500.0
<i>Outliers variant I</i>				
11	1.0	550.0	2.8	60.0
12	4.0	120.0	5.3	60.0
13	5.5	50.0	6.3	60.0
<i>Outliers variant II</i>				
11	1.8	500.0	3.2	60.0
12	3.3	200.0	2.2	60.0
13	5.2	80.0	1.4	60.0

Table 9 WTLS and Total M_{split} estimation results for the sets containing biases (additional regression line) (for TM_{split} : $\xi_1 := \xi_{1\alpha}$, $\xi_2 := \xi_{2\alpha}$)

Parameter	Estimate						
	Original data			With outliers			
	Exact solution (Neri et al.)	WTLS (Shen et al.)	TM_{split}	Variant I		Variant II	
WTLS				TM_{split}	WTLS	TM_{split}	
ξ_1	5.4799	5.4799	5.9436	4.9393	5.4604	3.7366	5.8531
ξ_2	- 0.4805	- 0.4805	- 0.5213	- 0.2870	- 0.5009	- 0.2412	- 0.5696
$\hat{\xi}_{1\beta}$	-	-	4.8702	-	2.0222	-	4.1870
$\hat{\xi}_{2\beta}$	-	-	- 0.4012	-	0.7753	-	- 0.5566
σ_0 , Eq. (19)	-	0.2039	-	0.3987	-	0.8180	-
$\sigma_{0\alpha}$	-	-	0.1662	-	0.2744	-	0. 5041
$\sigma_{0\beta}$	-	-	0.4758	-	0.7663	-	0.8274

extended to include outliers are provided in other columns of Table 9. These sets and the corresponding regression lines are shown in Fig. 7.

The TM_{split} estimators determined for the original set of observations may appear not wholly satisfactory (column 4, Table 9). Due to the lack of outliers, Total M_{split} estimation, predicting the existence of two mutually competing regression lines, forces two regression lines to fit into the set of observations (see Fig. 6). It should be noted that the regression line established by WTLS estimators lies

between these lines. The average values of TM_{split} estimators $\hat{\xi}_{1M_{\text{split}}} = (\hat{\xi}_{1\alpha} + \hat{\xi}_{1\beta})/2 = 5.4069$ and $\hat{\xi}_{2M_{\text{split}}} = (\hat{\xi}_{2\alpha} + \hat{\xi}_{2\beta})/2 = -0.4612$, as compared to WTLS estimators $\hat{\xi}_1 = 5.4799$ and $\hat{\xi}_2 = -0.4805$, can already be considered satisfactory. For both sets containing outliers, TM_{split} estimators $\hat{\xi}_{1\alpha}$ and $\hat{\xi}_{2\alpha}$ are close to the corresponding WTLS estimators obtained for the original set of observations. On the other hand, estimators $\hat{\xi}_{1\beta}$ and $\hat{\xi}_{2\beta}$ are close to the true values of parameters $\xi_{1\beta} = 2.0$, $\xi_{2\beta} = 0.75$ (Variant I) and $\xi_{1\beta} = 4.5$, $\xi_{2\beta} = -0.70$ (Variant II). In such cases, WTLS estimators yielded no good answers. This is particularly true

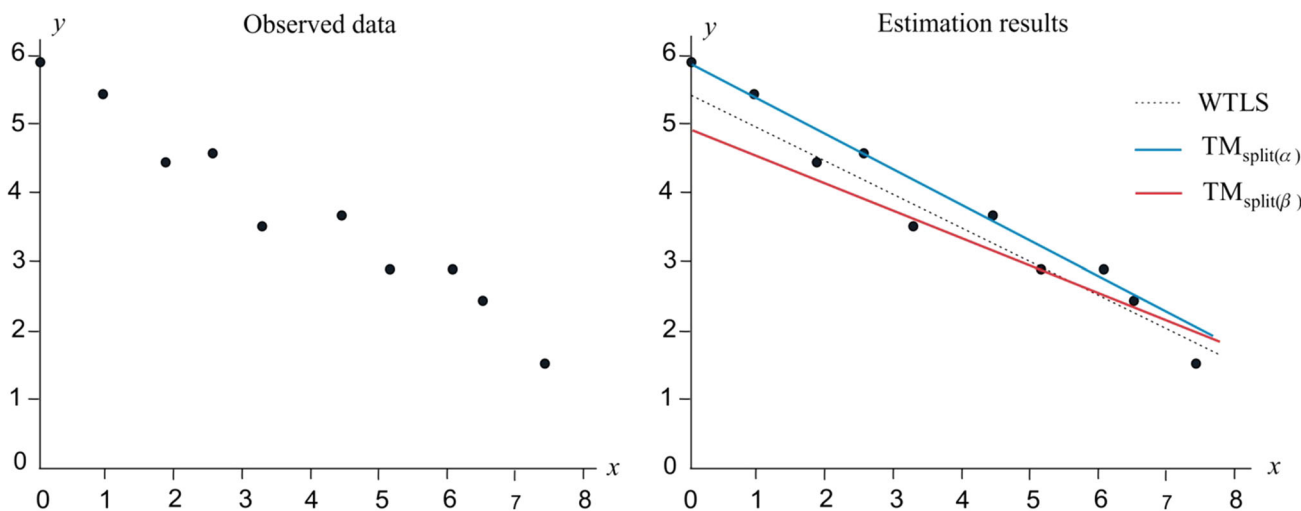


Fig. 6 A set of original observations (Shen et al. 2011) and the regression line position were determined based on WTLS and TM_{split} estimators

for Variant II for which the relevant comparisons are particularly unfavourable. The results obtained using WTLS are not surprising, as the lack of WTLS estimators’ robustness to observations is their inherent feature.

The example presented above concerned a situation where outliers can be assigned a regression line that is appropriate for them. In practice, however, the outlying of observations may relate to single observations and result from, for example, the effect of gross errors. In order to check the response of TM_{split} and WTLS estimators to such errors, one of the observations from the original set will be affected by gross error with a few value versions. For example, let it be assumed that such an observation is $y_5 = 3.5$ ($x_5 = 3.3$) with weights $W_{x_5} = 200$ and $W_{y_5} = 20$ (Table 8). This observation will be affected by gross error with values of $g = 1, g = 2, g = 5, g = 10$, respectively. The data adopted for the calculations are provided in Table 10. On the other hand, the TM_{split} and WTLS estimators of parameters ξ_1 and ξ_2 are provided in Table 11.

TM_{split} estimators $\hat{\xi}_{1\alpha}$ and $\hat{\xi}_{2\alpha}$ for each accepted gross error value are satisfactory, especially in comparison with WTLS estimators being obtained, which are unacceptable even for small gross errors. The quantities $\hat{\xi}_{1\beta}$ and $\hat{\xi}_{2\beta}$ are competing in relation to $\hat{\xi}_{1\alpha}$ and $\hat{\xi}_{2\alpha}$. These are estimators of the parameters of the $y_5 = \xi_{1\beta} + \xi_{2\beta}x_5$ regression line on which the observation affected by gross error should lie. By using the equation $\tilde{y}_5 = \hat{\xi}_{1\beta} + \hat{\xi}_{2\beta}x_5$, it is possible to calculate the prediction of observation y_5 affected by gross error. The prediction of this observation for the adopted gross error values, as compared to its simulated values, is presented in Table 12. A graphical interpretation of the obtained results is provided in Fig. 8.

4.3 Example 3: Two-dimensional affine transformation

The estimators of parameters in the EIV model, which are robust to gross errors, include *inter alia* robust total least-squares (RTLS) and total least trimmed squares (TLTS) estimators (Wang et al. 2016; Lv and Sui 2020). One of the examples provided in Lv and Sui (2020) concerns a two-dimensional affine transformation carried out based on observations affected by gross errors. In that case, the authors applied the following transformation model:

$$\begin{bmatrix} u_t \\ v_t \end{bmatrix} = \begin{bmatrix} u_s & v_s & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & u_s & v_s & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ b_1 \\ c_1 \\ a_2 \\ b_2 \\ c_2 \end{bmatrix} \tag{74}$$

where u_s, v_s and u_t, v_t are the coordinates of the common points in the start and target coordinate systems, while $a_1, b_1, c_1, a_2, b_2, c_2$ are the parameters being determined. Table 13 presents 15 observation points simulated in the start and target coordinate systems (Lv and Sui 2020). Some of these observations (highlighted in bold) are affected by gross errors. For these data, Lv and Sui (2020) applied TLTS estimation (using two different algorithms yielding the same results) as well as TLTS with RTLS as a starting point (hereafter denoted as TLTS/RTLS). The authors compared the obtained estimators with WTLS and RTLS estimators. These results may now be supplemented with TM_{split} estimators (the seventh column of Table 14). TM_{split} estimators will also be calculated for the data without gross errors. All the cited and calculated estimators are provided in Table 14.

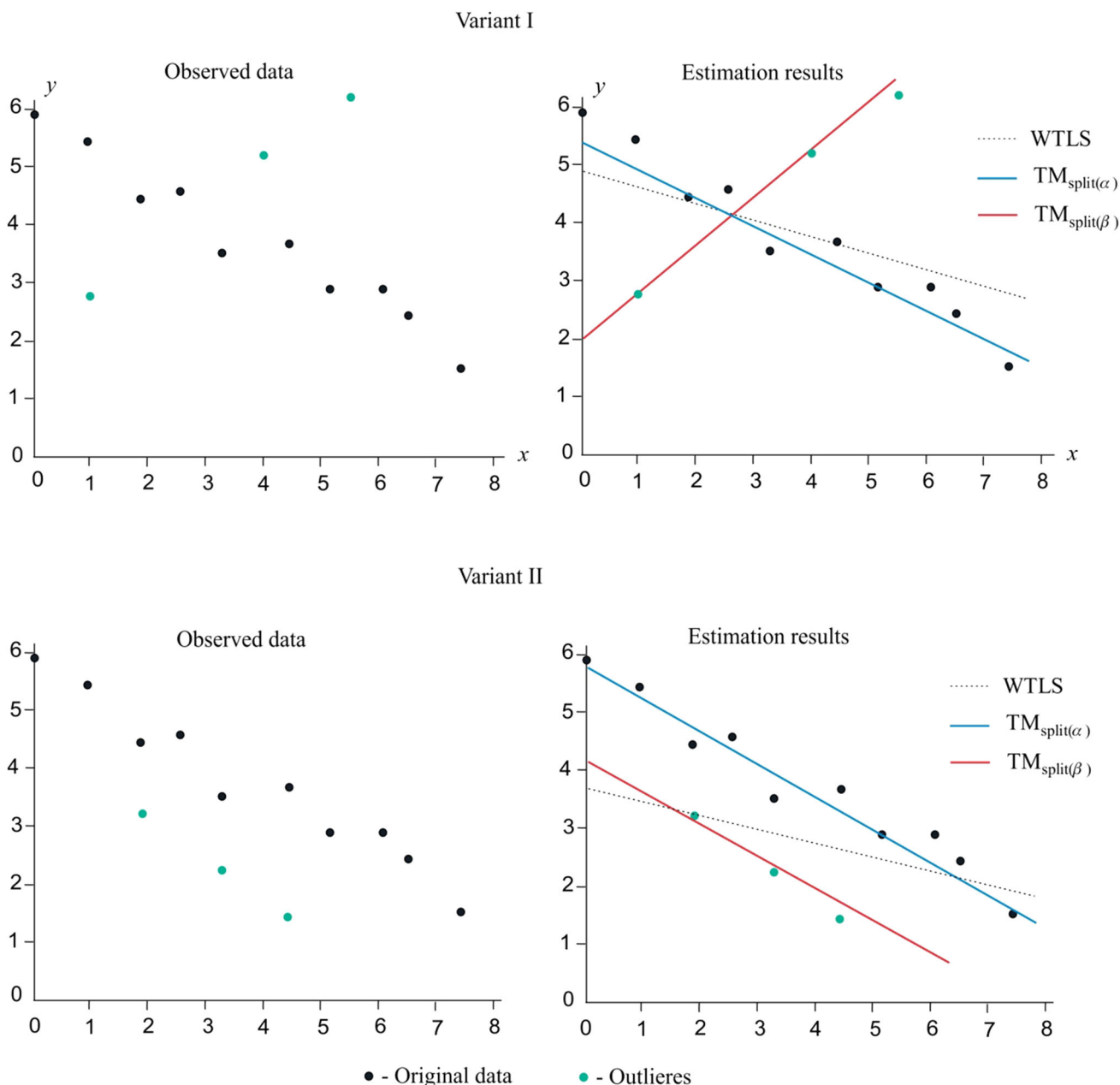


Fig. 7 Sets containing outliers (green) and the regression line position determined based on WTLS and TM_{split} estimators

Table 14 shows that $TM_{\text{split}(\alpha)}$ estimates are generally the closest to TLTS estimates. However, in the case of parameters a_1 , b_1 , c_1 and b_2 , $TM_{\text{split}(\alpha)}$ estimators are closest to TLTS/RTLS estimators and thus to the true parameter values. The interpretation of $TM_{\text{split}(\beta)}$ estimators is similar to that in the previous example, i.e. these are estimators of the parameters of the model for observations affected by gross errors. It is worth noting here that in the absence of gross errors, both versions of TM_{split} estimators differ slightly from each other.

5 Summary

By using M_{split} estimation, it is possible to determine the estimators of mutually competing parameters in classical functional models. However, there are cases in geodetic practice in which classical models need to be replaced with EIV models. The method proposed in this paper, called “Total M_{split} estimation”, is an development of M_{split} estimation that accepts such models. The Total M_{split} estimation objective function was determined by applying the Lagrange approach, as in the case of WTLS method. The mutually competing

Table 10 Observations and weights corresponding to them (observation y_5 , highlighted in bold, is affected by gross error g with different values)

i	x_i	W_{x_i}	y_i					W_{y_i}
			$g = 0$	$g = 1$	$g = 2$	$g = 5$	$g = 10$	
1	0.0	1000.0	5.9	5.9	5.9	5.9	5.9	1.0
2	0.9	1000.0	5.4	5.4	5.4	5.4	5.4	1.8
3	1.8	500.0	4.4	4.4	4.4	4.4	4.4	4.0
4	2.6	800.0	4.6	4.6	4.6	4.6	4.6	8.0
5	3.3	200.0	3.5	4.5	5.5	8.5	13.5	20.0
6	4.4	80.0	3.7	3.7	3.7	3.7	3.7	20.0
7	5.2	60.0	2.8	2.8	2.8	2.8	2.8	70.0
8	6.1	20.0	2.8	2.8	2.8	2.8	2.8	70.0
9	6.5	1.8	2.4	2.4	2.4	2.4	2.4	100.0
10	7.4	1.0	1.5	1.5	1.5	1.5	1.5	500.0

Table 11 TM_{split} and WTLS estimators determined for the set containing an observation affected by gross error (for $TM_{split}:\xi_1 := \xi_{1\alpha}$, $\xi_2 := \xi_{2\alpha}$)

Parameter	Estimate									
	$g = 0$		$g = 1$		$g = 2$		$g = 5$		$g = 10$	
	TM_{split}	WTLS	TM_{split}	WTLS	TM_{split}	WTLS	TM_{split}	WTLS	TM_{split}	WTLS
ξ_1	5.94	5.48	5.53	6.13	5.57	6.84	5.80	9.21	5.83	13.76
ξ_2	-0.52	-0.48	-0.53	-0.59	-0.45	-0.71	-0.52	1.11	-0.54	-1.97
$\xi_{1\beta}$	4.87	-	6.42	-	9.44	-	11.85	-	19.25	-
$\xi_{2\beta}$	-0.40	-	-0.60	-	-1.24	-	-1.04	-	-1.80	-

Table 12 Prediction of an observation affected by gross error

Gross error	Prediction of an observation	Simulated observation
$g = 1$	$\tilde{y}_5 = \hat{\xi}_{1\beta} + \hat{\xi}_{2\beta}x_5 = 6.42 - 0.60 \cdot 3.3 = 4.44$	$y_5 = 4.5$
$g = 2$	$\tilde{y}_5 = \hat{\xi}_{1\beta} + \hat{\xi}_{2\beta}x_5 = 9.44 - 1.24 \cdot 3.3 = 5.34$	$y_5 = 5.5$
$g = 5$	$\tilde{y}_5 = \hat{\xi}_{1\beta} + \hat{\xi}_{2\beta}x_5 = 11.85 - 1.04 \cdot 3.3 = 8.41$	$y_5 = 8.5$
$g = 10$	$\tilde{y}_5 = \hat{\xi}_{1\beta} + \hat{\xi}_{2\beta}x_5 = 19.25 - 1.80 \cdot 3.3 = 13.31$	$y_5 = 13.5$

EIV models occurring in this function were replaced with their linear approximations. This enabled the construction of a relatively simple yet efficient algorithm for determining TM_{split} estimators. The basis of this algorithm is the iterative updating of EIV models (external iterations) based on M_{split} estimators (internal iterations) obtained in the previous iterative step. The proposed algorithm is efficient in all cases presented in the paper. It concerns both results and the flow of the iterative process. The problems with the convergence of the external iterations might occur because of the linear

approximation of EIV models applied. It is especially evident when the errors disturbing the matrix A are too big.

The examples presented in the paper showed that the properties of TM_{split} estimators are, in general, similar to the properties of M_{split} estimators. If the elements of matrix A are not affected by random errors, then TM_{split} and M_{split} estimators are equal to each other. The possibility of determining estimators of competing parameters is of particular importance when the sets of observations are a mixture of the realisations of two random variables with mutually competing positional parameters. Such a situation occurs in Examples 1 and 2, where each observation group can be assigned a corresponding regression line. The problem, however, is that it is not known which of these lines is the best for a particular observation. TM_{split} estimators, similarly like M_{split} estimators for classical models, yield satisfactory results here. Due to their theoretical origin (neutral LS-method), WTLS estimators are not robust to gross errors. Where the sets are realisations of only a single random variable, Total M_{split} estimation further offers two mutually competing solutions. These are forced solutions, yet so close to each other that even in such a situation, it is possible to evaluate the functional model parameters (e.g. after the calculation of average values of respective estimators).

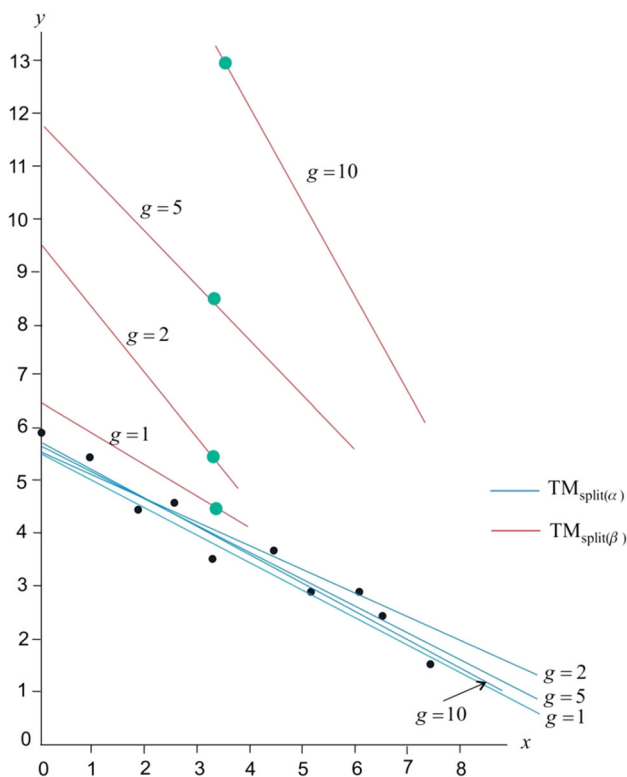


Fig. 8 Regression lines determined using Total M_{split} estimation based on the set containing a single observation affected by gross error g with different values

Total M_{split} estimation can also be applied for the estimation of EIV model parameters in the case where the outlying of observations results from their being affected by gross

errors. From the perspective of the M_{split} and TM_{split} estimation theory, such a case is not significantly different from that discussed earlier. In the second part of Example 2, it was shown that the determined TM_{split} estimators enabled the determination of not only the regression line appropriate for “good” observations ($TM_{split(\alpha)}$ solution), but also of the regression line on which the observation affected by gross error lies ($TM_{split(\beta)}$ solution).

In EIV models, it is assumed that matrix \mathbf{A} is observed as well. Therefore, its elements can also be affected by gross errors arising from various reasons. Such a situation is the case in Example 3, in which certain coordinates are affected by gross errors, both in the start and target systems in two-dimensional affine transformation. In this example, TM_{split} estimators are close to robust TLTS estimators and, for certain transformation parameters, they are also close to the results of TLTS estimation with RTLS estimation as a starting step.

The estimates’ accuracy is an important issue (especially in comparing the estimation methods). The accuracy of M_{split} estimation can be determined by applying asymptotical covariance matrices. To apply such an approach to Total M_{split} estimation, one should develop the theory, which is beyond the scope of the present paper. Section 4.1 presents the assessments of the M_{split} and Total M_{split} estimates accuracy (empirical standard deviations) obtained from the Monte Carlo simulations. Generally, the accuracy of M_{split} estimates decreases with the growing standard deviation of errors disturbing the matrix \mathbf{A} . Total M_{split} estimates have smaller standard deviations than respective M_{split} estimates in such a context. Similar relations concern the measures of efficacy,

Table 13 Observed points with outliers in the start and target coordinate systems (Lv and Sui 2020). Boldface numbers indicate outliers

Point No	Start system		Target system	
	u_s	v_s	u_t	v_t
1	99.893	− 0.106	319.940	− 29.856
2	99.899	125.022	− 55.206	94.918
3	99.951	249.936	− 430.084	219.994
4	200.134	− 0.015	519.953	− 129.895
5	200.019	125.103	144.887	− 4.903
6	200.104	250.013	− 230.107	120.036+1
7	299.999	0.084	719.965	− 230.037
8	299.944	125.050+7	345.088	− 104.950
9	299.973	249.996	− 30.266	19.973−0.6
10	150.111	62.520	232.425	− 17.450
11	150.090	187.393	− 142.441	107.521
12	249.999	62.435	432.673	− 117.527
13	249.915	187.558	57.729	7.494
14	300.127−9	187.435	157.599	− 42.344
15	150.170+2	249.887	− 330.273	170.002

Table 14 A comparison of the estimated parameters from different methods (for TM_{split} : $a_1 := a_{1\alpha}$, $b_1 := b_{1\alpha}$, $c_1 := c_{1\alpha}$, $a_2 := a_{2\alpha}$, $b_2 := b_{2\alpha}$, $c_{2\alpha} := c_2$)

Parameter	True value	Estimate					
		WTLS (Lv and Sui)	RTLS (Lv and Sui)	TLTS (Lv and Sui)	TLTS/RTLS (Lv and Sui)	TM_{split}	TM_{split} Without gross errors
a_1	2	2.05035	2.04090	2.00333	2.00097	2.00160	2.00184
b_1	-3	-2.99938	-3.00236	-2.99890	-3.00062	-3.00001	-3.00062
c_1	120	111.90862	113.66743	119.21405	119.78826	119.60660	119.5534
a_2	-1	-1.02106	-1.02431	-1.00141	-1.00015	-1.00166	-0.99783
b_2	1	0.99935	1.00091	0.99935	0.99998	1.00006	1.00108
c_2	70	73.52109	73.81469	70.39.033	70.08370	70.38717	69.47631
$a_{1\beta}$	-	-	-	-	-	2.14284	2.00340
$b_{1\beta}$	-	-	-	-	-	-3.08044	-3.00194
$c_{1\beta}$	-	-	-	-	-	104.32632	119.49778
$a_{2\beta}$	-	-	-	-	-	-1.09872	-1.00161
$b_{2\beta}$	-	-	-	-	-	0.95548	1.00004
$c_{2\beta}$	-	-	-	-	-	98.43220	70.38461

namely values of RMSEs. Thus, when the competitive functional models are supplemented with EIV models, then M_{split} estimation should be replaced by Total M_{split} estimation. It is especially advisable when disturbances of matrix \mathbf{A} have large standard deviations (here, the application of Total M_{split} estimation is justified for $\sigma_e = \sigma_y$).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Acar M, Özlüdemir MT, Akyılmaz O, Çelik RN, Ayan T (2006) Deformation analysis with total least squares. *Nat Hazards Earth Syst Sci* 6:663–669
- Akyılmaz O (2007) Total least squares solution of coordinate transformation. *Surv Rev* 39(303):68–80
- Amiri-Simkooei AR (2017) Weighted total least squares with singular covariance matrices subject to weighted and hard constraints. *J Surv Eng* 143(4):04017018
- Amiri-Simkooei AR (2018) Weighted total least squares with constraints: alternative derivation without using Lagrange multipliers. *J Surv Eng* 144(2):06017005
- Amiri-Simkooei A, Jazaeri S (2012) Weighted total least squares formulated by standard least squares theory. *J Geod Sci* 2(2):113–124
- Amiri-Simkooei AR, Asce M, Zangeneh-Nejad F, Asgari J (2016) On the covariance matrix of weighted total least-squares estimates. *J Surv Eng* 142(3):04015014
- Amiri-Simkooei AR, Alaei-Tabatabaei SM, Zangeneh-Nejad F, Voosoghi B (2017) Stability analysis of deformation-monitoring network points using simultaneous observation adjustment of two epochs. *J Surv Eng* 143(1):04016020
- Aydin C, Mercan H, Uygur SÖ (2018) Increasing numerical efficiency of iterative solution for total least-squares in datum transformations. *Stud Geophys Geod* 62:223–242
- Błaszczak-Bąk W, Janowski A, Kamiński W, Rapiński J (2015) Application of the M_{split} method for filtering airborne laser scanning data-sets to estimate digital terrain models. *Int J Remote Sens* 36(9):2421–2437
- Davis TG (1999) Total least squares spiral curve fitting. *J Surv Eng* 125(4):159–176
- Duchnowski R, Wiśniewski Z (2012) Estimation of the shift between parameters of functional models of geodetic observations by applying M_{split} estimation. *J Surv Eng* 138:1–8
- Duchnowski R, Wiśniewski Z (2014) Comparison of two unconventional methods of estimation applied to determine network point displacement. *Surv Rev* 46(339):401–405
- Duchnowski R, Wiśniewski Z (2019) Robustness of $M_{split(q)}$ estimation: a theoretical approach. *Stud Geophys Geod* 63:390–417
- Duchnowski R, Wiśniewski Z (2020) Robustness of squared $M_{split(q)}$ estimation: empirical analyses. *Stud Geophys Geod* 64:153–171
- Fang X (2013) Weighted total least squares: necessary and sufficient conditions, fixed and random parameters. *J Geod* 87:733–749
- Fang X (2015) Weighted total least-squares with constraints: a universal formula or geodetic symmetrical transformations. *J Geod* 89:459–469
- Felus F (2004) Application of total least squares for spatial point process analysis. *J Surv Eng* 130(3):126–133
- Felus Y, Schaffrin B (2005) Performing similarity transformations using the errors-in-variables-model. In: *Proceedings of the ASPRS meeting, Washington, DC, May 2005*
- Ge Y, Yuan Y, Jia N (2013) More efficient methods among commonly used robust estimation methods for GPS coordinate transformation. *Surv Rev* 45:229–234

- Golub GH, van Loan CF (1980) An analysis of the total least squares problem. *SIAM J Numer Anal* 17:883–893
- Huber PJ (1964) Robust estimation of location parameter. *Ann Math Stat* 43(4):1041–1067
- Janicka J, Rapiński J (2013) M_{split} transformation of coordinates. *Surv Rev* 45:269–274
- Janicka J, Rapiński J, Błaszczak-Bąk W, Suchocki C (2020) Application of the M_{split} estimation method in the detection and dimensioning of the displacement of adjacent planes. *Remote Sens* 12:3203
- Janowski A, Rapiński J (2013) M_{split} Estimation in laser scanning data modeling. *J Indian Soc Remote Sens* 41:15–19
- Jones GA, Jones JM (2000) Information and coding theory. Undergraduate Mathematics Series. Springer, Heidelberg
- Kargoll B, Omidalizarandi M, Loth I, Paffenholz JA, Alkhatib H (2018) An iteratively reweighted least-squares approach to adaptive robust adjustment of parameters in linear regression models with autoregressive and t-distributed deviations. *J Geod* 92:271–297
- Koch RK (2013) Robust estimation by expectation maximization algorithm. *J Geod* 87:107–116
- Kubáčková L, Kubáček L (1991) Optimum processing of measurements from a group of instruments affected by drift. *Manuscr Geod* 16:148–154
- Lv Z, Sui L (2020) The BAB algorithm for computing the total least trimmed squares estimator. *J Geod* 94(12):110
- Mahboub V (2012) On weighted total least-squares for geodetic transformations. *J Geod* 86:359–367
- Mercan H, Akyilmaz O, Aydin C (2018) Solution of the weighted symmetric similarity transformations based on quaternions. *J Geod* 92:1113–1130
- Neri F, Saitta G, Chiofalo S (1989) An accurate and straightforward approach to line regression analysis of error-affected experimental data. *J Phys Ser E: Sci Instr* 22:215–217
- Nowel K (2016) Application of Monte Carlo method to statistical testing in deformation analysis based on robust M-estimation. *Surv Rev* 48(348):212–223
- Nowel K (2019) Squared $M_{\text{split}(q)}$ S-transformation of control network deformations. *J Geod* 93:1025–1044
- Pope A (1974) Two approaches to nonlinear least squares adjustments. *Can Surv* 28(5):663–669
- Rao CR (1973) Linear statistical inference and its applications. Wiley, New York
- Schaffrin B (2020) Total least-squares collocation: an optimal estimation technique for the EIV-model with prior information. *Mathematics* 8(6):971
- Schaffrin B, Felus Y (2008) On the multivariate total least-squares approach to empirical coordinate transformations. Three algorithms. *J Geod* 82:373–383
- Schaffrin B, Wieser A (2008) On weighted total least-squares adjustment for linear regression. *J Geod* 82:415–421
- Schaffrin B, Lee IP, Felus Y, Choi YS (2006) Total least-squares for geodetic straight-line and plane adjustment. *Boll Geod Sci Aff* 65:141–168
- Shen Y, Li B, Chen Y (2011) An iterative solution of weighted total least-squares adjustment. *J Geod* 85:229–238
- Teunissen PJG (1988) The nonlinear 2D symmetric Helmert transformation: an exact nonlinear least squares solution. *Bull Geod* 62:1–16
- Tong X, JinY ZS, Li L, Liu S (2015) Bias-corrected weighted total least-squares adjustment of condition equations. *J Surv Eng* 141(2):04014013
- van Huffel S, Vandewalle J (1991) The total least squares problem: computational aspects and analysis. SIAM, Philadelphia
- Wang L, Zhao Y (2019) Second-order approximation function method for precision estimation of total least squares. *J Surv Eng* 145(1):04018011
- Wang B, Liu J, Liu C (2016) A robust weighted total least squares algorithm and its geodetic applications. *Stud Geophys Geod* 60:177–194
- Wiśniewski Z (1985) The effect of the asymmetry of geodetic observation error distribution on the results of adjustment by the least squares method. *Geod Cartogr* 34(11):11–21
- Wiśniewski Z (2009) Estimation of parameters in a split functional model of geodetic observations (M_{split} estimation). *J Geod* 83:105–120
- Wiśniewski Z (2010) $M_{\text{split}(q)}$ estimation: estimation of parameters in a multi split functional model of geodetic observations. *J Geod* 84:355–372
- Wiśniewski Z (2017) M-P estimation applied to platykurtic sets of geodetic observations. *Geod Cartogr* 66(1):117–135
- Wiśniewski Z, Kamiński W (2020) Estimation and prediction of vertical deformations of random surfaces, applying the total least squares collocation method. *Sensors* 20(14):3913
- Wiśniewski Z, Zienkiewicz MH (2016) Shift- M_{split} estimation in deformation analyses. *J Surv Eng* 142:1–13
- Wiśniewski Z, Zienkiewicz MH (2021a) Empirical analyses of robustness of the square M_{split} estimation. *J Appl Geod* 15(2):87–104
- Wiśniewski Z, Zienkiewicz MH (2021b) Estimators of covariance matrices in $M_{\text{split}(q)}$ estimation. *Surv Rev* 53(378):263–279
- Wyszowska P, Duchnowski R (2019) M_{split} estimation based on L1 norm condition. *J Surv Eng* 145(3):04019006
- Wyszowska P, Duchnowski R, Dumalski A (2021) Determination of terrain profile from TLS data by applying M_{split} estimation. *Remote Sens* 13:31
- Xu P, Liu J (2014) Variance components in errors-in-variables models: estimability, stability and bias analysis. *J Geod* 88:719–734
- Xu P, Liu J, Shi C (2012) Total least squares adjustment in partial errors-in-variables models: algorithm and statistical analysis. *J Geod* 86:661–675
- Xu P, Liu J, Zeng W, Shen Y (2014) Effects of errors-in-variables on weighted least squares estimation. *J Geod* 88:705–716
- Yang Y, Zhang S (2005) Adaptive fitting of systematic errors in navigation. *J Geod* 79:43–49
- Yang Y, Gao W, Zhang X (2010) Robust Kalman filtering with constraints: a case study for integrated navigation. *J Geod* 84:373–381
- Zeng W, Fang X, Lin Y, Huang X, Zhou Y (2018) On the total least-squares estimation for autoregressive model. *Surv Rev* 50:186–190
- Zienkiewicz MH (2014) Application of M_{split} estimation to determine control points displacements in networks with unstable reference system. *Surv Rev* 47:174–180
- Zienkiewicz MH (2018a) Deformation analysis of geodetic networks by applying M_{split} estimation with conditions binding the competitive parameters. *J Surv Eng* 145(2):04019001
- Zienkiewicz MH (2018b) Determination of an appropriate number of competitive functional models in a square $M_{\text{split}(q)}$ estimation by applying the modified Baarda approach. *Surv Rev* 52(370):13–23
- Zienkiewicz MH, Hejbudzka K, Dumalski A (2017) Multi split functional model of geodetic observations in deformation analyses of the Olsztyn castle. *Acta Geodyn Geomater* 14:195–204