The effect of an unknown data bias in least-squares adjustment: some concerns for the estimation of geodetic parameters

C. Kotsakis
Department of Geodesy and Surveying, Aristotle University of Thessaloniki
University Box 440, Thessaloniki, GR-54124, Greece, Email: kotsaki@topo.auth.gr

Abstract

Least-squares (LS) estimation is a standard tool for the optimal processing of geodetic data. In the framework of global gravity field modelling, for example, such methods are extensively applied for the determination of geoid solutions from CHAMP and GRACE data via the estimation of a large set of spherical harmonic coefficients. Frequently, in geodetic applications additional nuisance parameters need to be included in the least-squares adjustment procedure to account for external biases and disturbances that have affected the available measurements. The objective of this paper is to expose a trade-off which exists in the LS inversion of linear models that are augmented by additional parameters in the presence of unknown systematic errors in the input data. Specifically, it is shown that if a linear model of full rank is extended by a scalar parameter to account for a common bias in the data, the LS estimation accuracy of *all* the other model parameters automatically worsens. Some simple numerical examples are also given to demonstrate the significance of this accuracy degradation in the geodetic practice.

1. Introduction

Least-squares (LS) estimation is a standard tool for the optimal processing of geodetic data. Its use is commonly associated with a linear(-ized) model of observation equations

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v} \tag{1}$$

$$E\{\mathbf{v}\} = \mathbf{0} \;,\; E\{\mathbf{v}\mathbf{v}^{\mathrm{T}}\} = \mathbf{C} \tag{2}$$

where $\mathbf{y}$ is a vector of observations, $\mathbf{v}$ is a vector of zero-mean random errors with a covariance (CV) matrix $\mathbf{C}$, $\mathbf{A}$ is a matrix of known coefficients, and $\mathbf{x}$ is a parameter vector that needs to be estimated from the given data. Despite its simplistic linear character and its inherent restriction for additive noise, the above model is frequently employed in every field of geodetic research (e.g., *Dermanis and Rummel* 2000).

A problem that is often encountered in geodetic data analysis with the aforementioned general model is the presence of external disturbances (biases) in the available measurements. In cases where the effect of such disturbances can be determined *a-priori* with sufficiently high accuracy (i.e. with an uncertainty that is significantly lower than the data noise level), the original measurements $\mathbf{y}$ should be replaced with a 'corrected' set that is compatible with the theoretical model of Eqs. (1) and (2). If, on the other hand, the magnitude of the external biases is completely unknown (or, at least, poorly known), then the model of Eq. (1) should be augmented by additional parametric terms that describe the systematic effects in the input data. An integrated LS estimation procedure can then lead to optimal estimates for the original model parameters **x** and the additional bias-related parameters that will be included in the data adjustment process.

Bias modeling and elimination has been a topic of continuous research interest in geodesy, by both theoreticians and practitioners. An extensive discussion, related mostly to the mathematical details of bias treatment in geodetic data analysis is given in *Kukuča* (1987). Some interesting aspects from the theory of nuisance parameter elimination in LS adjustment models can be found in *Schaffrin and Grafarend* (1986). A comparison of different approaches for dealing with systematic effects that arise from the integration of heterogeneous geodetic data sets is given in *Schaffrin and Baki-Iz* (2001). The problem of bias modeling and elimination has also been treated in the framework of various particular fields of geodetic research, including gravimetric data processing (e.g. *Kubáčková and Kubáček* 1993, *Harnisch* 1993), altimetric data processing (e.g. *Tscherning and Knudsen* 1986, *Gaspar et al.* 1994), GPS data processing (e.g. *Satirapod et al.* 2003, *Jia et al.* 2000) and global gravity field modelling (e.g. *Lerch* 1991).

Motivated by the fact that most geodetic observations are carried under complex physical conditions which do not correspond exactly to the standard mathematical models that we often employ for their analysis, the purpose of this paper is to expose an important trade-off which exists in the LS inversion of linear models that are augmented by additional (nuisance) parameters in the presence of unknown systematic errors in the input data. In particular, it is shown that if a linear model of full rank is extended by a single scalar parameter to account for a common bias in the data, then the LS estimation accuracy of *all* the other model parameters automatically worsens. Two different cases are considered in this study, depending whether *some* or *all* of the observations are affected by a common unknown bias. Some numerical examples are also given to demonstrate the significance of this accuracy degradation in the geodetic practice.

2. LS parameter estimation with different types of models

In this section, we examine the results obtained from the LS inversion of three different models for the linear(-ized) system of observation equations. Specifically, we study the cases where *none*, *all*, or *some* of the input measurements are affected by a common unknown bias.

2.1 Linear model with non-biased data

This is the archetypical case whose formulation and description have already been given in the previous section; see Eqs. (1) and (2). For the sake of economy in our discussion, we shall assume that the design matrix **A** has full column rank and that the CV matrix **C** is fully known. In such a case, the LS inversion of the system of observation equations in Eq. (1) leads to the well-known optimal solution

$$\hat{\mathbf{x}} = (\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{y} \tag{3}$$

The quality of the above unbiased estimate is generally described by its associated CV matrix

$$\mathbf{C}_{\hat{\mathbf{x}}} = E\left\{(\hat{\mathbf{x}} - \mathbf{x})(\hat{\mathbf{x}} - \mathbf{x})^{\mathrm{T}}\right\} = (\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})^{-1} \tag{4}$$

whose diagonal elements give a measure of the statistical accuracy for the estimated model parameters $\hat{\mathbf{x}}$.

2.2 Linear model with uniformly biased data – Case I

In the presence of a common unknown bias in the measurements **y**, the functional model of Eq. (1) should be modified as follows

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \beta\mathbf{s} + \mathbf{v} \tag{5}$$

where β is a scalar parameter that describes the systematic offset in the input data, and **s** corresponds to a vector with all its values equal to one, $\mathbf{s} = \left[1 \ \ \ldots \ \ 1\right]^{\mathrm{T}}$. The data noise **v** follows the same stochastic behaviour as in the case of the simple model with non-biased data.

Assuming that the partitioned matrix $[\mathbf{A} \ \ \mathbf{s}]$ has full column rank, the LS inversion of the augmented system in Eq. (5) leads to the following optimal estimates for the original model parameters **x** and the bias parameter β

$$\hat{\beta} = k\,\mathbf{s}^{\mathrm{T}}\mathbf{C}^{-1}(\mathbf{y} - \mathbf{A}\,\mathbf{x}^{b}) \tag{6}$$

$$\hat{\mathbf{x}} = \mathbf{x}^{b} - \hat{\beta}\,\mathbf{q} \tag{7}$$

where the auxiliary quantities $\mathbf{x}^{b}$, $\mathbf{q}$ and k are defined by the expressions

$$\mathbf{x}^{b} = (\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{y} \tag{8}$$

$$\mathbf{q} = (\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{s} \tag{9}$$

$$k = \left(\mathbf{s}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{s} - \mathbf{q}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})\mathbf{q}\right)^{-1} \tag{10}$$

Note that the term $\mathbf{x}^{b}$ corresponds to the LS estimate that we would obtain if we ignored the bias presence in the input data. Applying the (co-)variance propagation law to the unbiased estimators of Eqs. (6) and (7), we obtain the CV matrix of the estimated parameter vector $\hat{\mathbf{x}}$

$$\mathbf{C}_{\hat{\mathbf{x}}} = E\left\{(\hat{\mathbf{x}} - \mathbf{x})(\hat{\mathbf{x}} - \mathbf{x})^{\mathrm{T}}\right\} = (\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})^{-1} + k\,\mathbf{q}\,\mathbf{q}^{\mathrm{T}} \tag{11}$$

and also the variance of the estimated bias parameter $\hat{\beta}$

$$\sigma_{\hat{\beta}}^2 = E\left\{(\hat{\beta}-\beta)^2\right\} = k \tag{12}$$

The last two formulae can easily be derived by using standard rules of matrix calculus.

2.3 Linear model with uniformly biased data – Case II

A useful generalization of the results given in the previous section is obtained if we assume that only *a part* of the data is affected by a common unknown bias. In this case, the standard model of Eq. (1) should be modified in the following way

$$\underbrace{\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix}}_{\mathbf{A}} \mathbf{x} + \beta \begin{bmatrix} \mathbf{s} \\ \mathbf{0} \end{bmatrix} + \underbrace{\begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}}_{\mathbf{v}} \tag{13}$$

The total data vector $\mathbf{y}$ is now partitioned into two subsets, $\mathbf{y}_1$ and $\mathbf{y}_2$, whereas the submatrices $\mathbf{A}_1$ and $\mathbf{A}_2$ provide the corresponding partition for the design matrix $\mathbf{A}$ of the initial model in Eq. (1). Only the first group of measurements is affected by a common unknown bias, which is denoted again by the scalar parameter β. The symbol $\mathbf{s}$ denotes a vector of ones, whereas $\mathbf{0}$ is the zero vector. Finally, the subvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ contain the random errors of the two corresponding data subsets, with the following stochastic characteristics

$$E\{\mathbf{v}_1\} = E\{\mathbf{v}_2\} = \mathbf{0} \tag{14}$$

$$E\{\mathbf{v}\mathbf{v}^{\mathrm{T}}\} = \mathbf{C} = E\left\{\begin{bmatrix} \mathbf{v}_1\mathbf{v}_1^{\mathrm{T}} & \mathbf{v}_1\mathbf{v}_2^{\mathrm{T}} \\ \mathbf{v}_2\mathbf{v}_1^{\mathrm{T}} & \mathbf{v}_2\mathbf{v}_2^{\mathrm{T}} \end{bmatrix}\right\} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_2 \end{bmatrix} \tag{15}$$

For simplicity, we assume zero cross-correlation in the data noise between the two groups of measurements. Assuming also that the augmented model of Eq. (13) has full rank, its LS inversion yields the following optimal solution

$$\hat{\beta} = \lambda \mathbf{s}^{\mathrm{T}} \mathbf{C}_1^{-1} (\mathbf{y}_1 - \mathbf{A}_1 \mathbf{x}^b) \tag{16}$$

$$\hat{\mathbf{x}} = \mathbf{x}^b - \hat{\beta}\,\mathbf{p} \tag{17}$$

where the auxiliary quantities $\mathbf{p}$ and λ are defined by the equations

$$\mathbf{p} = (\mathbf{A}^{\mathrm{T}} \mathbf{C}^{-1} \mathbf{A})^{-1} \mathbf{A}_1^{\mathrm{T}} \mathbf{C}_1^{-1} \mathbf{s} \tag{18}$$

$$\lambda = \left(\mathbf{s}^{\mathrm{T}}\mathbf{C}_1^{-1}\mathbf{s} - \mathbf{p}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})\mathbf{p}\right)^{-1} \tag{19}$$

The auxiliary term $\mathbf{x}^b$ has already been defined in Eq. (8). Applying the (co-)variance propagation law to the unbiased estimators of Eqs. (16) and (17), we get the CV matrix for the estimated parameter vector $\hat{\mathbf{x}}$

$$\mathbf{C}_{\hat{\mathbf{x}}} = E\left\{(\hat{\mathbf{x}}-\mathbf{x})(\hat{\mathbf{x}}-\mathbf{x})^{\mathrm{T}}\right\} = (\mathbf{A}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{A})^{-1} + \lambda\,\mathbf{p}\mathbf{p}^{\mathrm{T}} \tag{20}$$

and also the variance of the estimated bias parameter $\hat{\beta}$

$$\sigma_{\hat{\beta}}^2 = E\left\{(\hat{\beta}-\beta)^2\right\} = \lambda \tag{21}$$

All the previous formulae can easily be verified by using standard rules of matrix calculus. A brief discussion on the estimation performance that is achieved by employing each of the three previous models follows in the next section.

3. Remarks

By comparing the results from Eqs. (4), (11) and (20), it is seen that the accuracy of the estimated parameters $\hat{\mathbf{x}}$ becomes worse when an overparameterized model is used for the LS adjustment of the input data. Indeed, in the cases where biased measurements are processed with an augmented linear model, the CV matrix of $\hat{\mathbf{x}}$ "increases" by the amount $k\mathbf{q}\mathbf{q}^{\mathrm{T}}$ (or $\lambda\mathbf{p}\mathbf{p}^{\mathrm{T}}$), with respect to the result obtained from the LS adjustment with non-biased data. Since the scalar quantities k and λ are positive, the diagonal elements of $\mathbf{C}_{\hat{\mathbf{x}}}$ in Eqs. (11) and (20) will always be larger than the corresponding elements of $\mathbf{C}_{\hat{\mathbf{x}}}$ in Eq. (4).

It can thus be concluded that the presence of an additional bias parameter in a linear model causes an overall degradation in the LS estimation accuracy for the rest of the model parameters. Note that the noise level is assumed to be the same in the cases of non-biased and biased data.

Furthermore, it is interesting to point out that the extent of the accuracy reduction is completely independent of the magnitude of the bias that has affected the input data. That is evident from the fact that the terms $k\mathbf{q}\mathbf{q}^{\mathrm{T}}$ and $\lambda\mathbf{p}\mathbf{p}^{\mathrm{T}}$ do not depend on β. Hence, any constant data bias (regardless of how small or large its true value really is) that is taken into account by adding an extra parameter within the LS adjustment framework, will cause the <u>same</u> degradation on the estimation accuracy for the rest of the model parameters.

4. Numerical examples

Two simple examples are given in this section to demonstrate the practical significance of the theoretical results that were previously discussed.

4.1 Gravity network adjustment

The first example refers to the LS adjustment of a simulated gravity network, as shown in Figure 1. The observables consist of eight gravity differences that are specified explicitly in Table 1, along with their corresponding accuracy. Note that all observations are considered uncorrelated in this example.

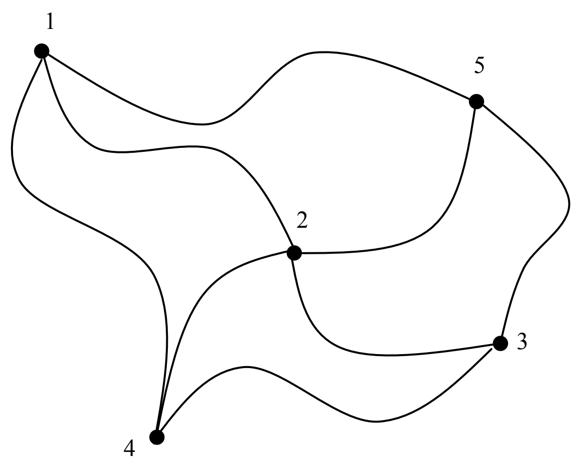

Figure 1. A simulated gravity network.

Table 1. The simulated observables, along with their corresponding accuracy, for the gravity network of Figure 1.

i	**Observables**	**Measurement accuracy** σ_i (in μgal)
1	$\Delta g_{12} = g_2 - g_1$	11
2	$\Delta g_{23} = g_3 - g_2$	10
3	$\Delta g_{34} = g_4 - g_3$	15
4	$\Delta g_{14} = g_4 - g_1$	16
5	$\Delta g_{51} = g_1 - g_5$	17
6	$\Delta g_{52} = g_2 - g_5$	16
7	$\Delta g_{53} = g_3 - g_5$	15
8	$\Delta g_{24} = g_4 - g_2$	10

Assuming a known gravity value for *point 1*, which is held fixed for the purpose of datum definition, the unknown parameter vector **x** in the network adjustment includes the absolute gravity values g_i for the four other network points. Using the information provided in Table 1, the CV matrix $\mathbf{C}_{\hat{\mathbf{x}}}$ of the estimated parameters, for the case of non-biased data, can be computed from Eq. (4). The result is

$$\mathbf{C}_{\hat{\mathbf{x}}} = \begin{bmatrix} 74.225 & 63.862 & 55.921 & 48.589 \\ 63.862 & 116.945 & 63.138 & 65.136 \\ 55.921 & 63.138 & 100.259 & 42.254 \\ 48.589 & 65.136 & 42.254 & 125.248 \end{bmatrix} [\mu gal^2] \tag{22}$$

If a common unknown bias is assumed to exist in all observations, then the extended model of Eq. (5) must be used for the LS network adjustment. In this case, the CV matrix $\mathbf{C}_{\hat{\mathbf{x}}}$ for the estimated gravity values is obtained from Eq. (11), which yields the result

$$\mathbf{C}_{\hat{\mathbf{x}}} = \begin{bmatrix} 104.602 & 126.932 & 147.301 & 22.803 \\ 126.932 & 247.795 & 252.794 & 11.617 \\ 147.301 & 252.794 & 375.151 & -35.318 \\ 22.803 & 11.617 & -35.318 & 147.138 \end{bmatrix} [\mu gal^2] \tag{23}$$

whereas the standard deviation of the estimated bias parameter $\hat{\beta}$ according to Eq. (12) is

$$\sigma_{\hat{\beta}} = 10.81 \ \mu gal \tag{24}$$

Table 2. Standard deviation for the estimated gravity values in the network of Figure 1.

Model parameters (x)	**Estimation accuracy obtained from the *simple* LS adjustment model**	**Estimation accuracy obtained from the *extended* LS adjustment model**
g_2	8.62 μgal	10.23 μgal
g_3	10.81 μgal	15.74 μgal
g_4	10.01 μgal	19.37 μgal
g_5	11.19 μgal	12.13 μgal

By comparing the diagonal elements from the CV matrices in Eqs. (22) and (23), it can be deduced that an average degradation of about 4.2 μgal in the accuracy of the estimated gravity values occurs, when the augmented model is used. The actual decrease in the estimation accuracy for each individual gravity value is analytically shown in Table 2.

4.2 Gravity determination from repeated free-fall experiments

For our second example, we consider the determination of absolute gravity through repeated observations of a free-falling mass *m* at a certain point *P* (see Figure 2).

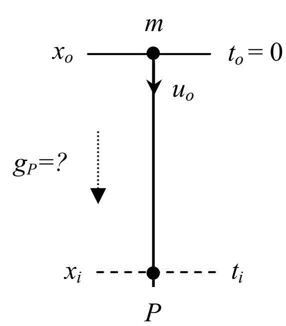

Figure 2. Free-fall experiment for absolute gravity determination.

Assuming a homogeneous gravity field within the region of the experiment, we have the following well-known equation

$$x_i = x_o + u_o t_i + \frac{1}{2} g t_i^2 \tag{25}$$

which describes the Newtonian motion of the test mass. The quantities x_o and u_o denote the known initial position and velocity, x_i is the observed vertical position of the test mass at the time instant t_i, and g corresponds to the gravity acceleration at the observation point.

The model of Eq. (25) provides the basic observation equation that we can use to determine the gravity value g via the LS adjustment of multiple repeated measurements $\{x_i\}$ of the vertical position of the test mass. For simplicity, we assume that (i) the time values t_i are known without any error, and (ii) the initial conditions are such that $x_o = 0$ and $u_o = 0$.

Our simulated experiment consists of twenty-five different drops of the test mass. The recorded times t_i are shown in Table 3, along with the corresponding accuracy for each of the observed positions x_i. Using the information from Table 3, we can easily construct the design matrix **A** and the data noise CV matrix **C** for the underlying adjustment problem. The standard deviation of the estimated gravity value is then computed by the general formula of Eq. (4), which yields the result

$$\sigma_{\hat{g}} = 0.281\, mgal \tag{26}$$

Table 3. Accuracy of the simulated observables x_i and the corresponding time intervals t_i for the free-falling mass experiment.

i	t_i (in sec)	σ_{x_i} (in mm)
1	0.516	0.001
2	0.474	0.002
3	0.487	0.002
4	0.502	0.002
5	0.467	0.002
6	0.491	0.002
7	0.487	0.001
8	0.501	0.001
9	0.512	0.001
10	0.493	0.003
11	0.479	0.003
12	0.492	0.003
13	0.514	0.002
14	0.505	0.002
15	0.496	0.002
16	0.493	0.003
17	0.495	0.003
18	0.510	0.003
19	0.501	0.002
20	0.486	0.002
21	0.483	0.003
22	0.495	0.003
23	0.510	0.002
24	0.511	0.002
25	0.495	0.002

In the case where a common unknown bias exists in all available observations x_i, the following extended model must be employed for the optimal determination of the gravity acceleration g

$$x_i = x_o + u_o t_i + \frac{1}{2} g t_i^2 + \beta \tag{27}$$

where β corresponds to the common bias in the measurements. The standard deviation of the LS estimated gravity acceleration is now computed from the general formula of Eq. (11), which gives the result

$$\sigma_{\hat{g}} = 5.473\, mgal \tag{28}$$

whereas the standard deviation of the bias estimate, according to Eq. (12), is

$$\sigma_{\hat{\beta}} = 6.83 \times 10^{-3}\ mm \tag{29}$$

From the numerical results given in Eqs. (26) and (28), it is evident that the standard deviation of the estimated gravity acceleration $\hat{g}$ increases by almost a factor of 20, when a single bias parameter is included in the LS adjustment model for the repeated observations $\{ x_i \}$!

5. Conclusions

The elimination of a common unknown bias from a set of geodetic measurements can be achieved by extending the classic linear(-ized) model $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v}$ with an additional scalar parameter that describes the systematic offset in the data values. In this way, optimal unbiased estimates for the model parameters $\mathbf{x}$ can be obtained when employing the standard LS estimation criterion with a biased data set.

Nevertheless, as it was explained in this paper, the statistical accuracy of the estimated model parameters $\hat{\mathbf{x}}$ will always be worse when an additional bias parameter is included in the LS adjustment process. Geodesists should be aware of this important trade-off, since in many applications the theoretical models often need to be augmented with additional nuisance parameters in order to account for external disturbances in the observations.

References

Dermanis A, Rummel R (2000) Data analysis methods in geodesy. In A. Dermanis, A. Gruen and F. Sanso (eds) *Geomatic Methods for the Analysis of Data in Earth Sciences*. Lecture Notes in Earth Sciences Series, 95: 17-92, Springer Verlag, Berlin Heidelberg.

Gaspar P, Ogor F, Le Traon P-Y, Zanife O-Z (1994) Estimating the sea state bias of the TOPEX and POSEIDON altimeters from crossover differences. *J Geoph Res*, 99(C12): 24,981-24,994.

Harnisch G (1993) Systematic errors affecting the accuracy of high precision gravity measurements. *IAG Symposia*, vol. 112, Springer-Verlag, Berlin Heidelberg, pp. 200-204.

Jia M, Tsakiri M, Stewart M (2000) Mitigating multipath errors using semi-parametric models for high precision static positioning. *IAG Symposia*, vol. 121, Springer-Verlag, Berlin Heidelberg, pp. 393-398.

Kubáčková L, Kubáček L (1993) A group of gravimeters, stochastical problems and their solution. *IAG Symposia*, vol. 112, Springer-Verlag, Berlin Heidelberg, pp. 275-278.

Kukuča J (1987) Systematic influences and their elimination. In: Kubáčková L, Kubáček L and Kukuča J: *Probability and statistics in geodesy and geophysics*. Elsevier, Amsterdam, pp. 257-289.

Lerch FJ (1991) Optimum data weighting and error calibration for estimation of gravitational parameters. *Bull Geod*, 65: 44-52.

Satirapod C, Wang J, Rizos C (2003) Comparing different global positioning system data processing techniques for modeling residual systematic errors. *J Surv Eng*, 129(4): 129-135.

Schaffrin B, Grafarend E (1986) Generating classes of equivalent linear models by nuisance parameter elimination. *Manuscr Geod*, 11: 262-271.

Schaffrin B, Baki-Iz H (2001) Integrating heterogeneous data sets with partial inconsistencies. *IAG Symposia*, vol. 123, Springer-Verlag, Berlin Heidelberg, pp. 49-54.

Tscherning CC, Knudsen P (1986) Determination of bias parameters for satellite altimeter by least-squares collocation. *Proceedings of the 1st Hotine-Marussi Symposium on Mathematical Geodesy*, Rome, June 3-6, 1985, pp. 833-851.