



# Monte-Carlo-based uncertainty propagation in the context of Gauss–Markov model: a case study in coordinate transformation

Propagação de incertezas baseada no método Monte Carlo no contexto do modelo de Gauss Markov: um estudo de caso em transformação de coordenadas

V. F. Rofatto<sup>1,2\*</sup>; M. T. Matsuoka<sup>1,2,3</sup>; I. Klein<sup>4,5</sup>; M. R. Veronez<sup>6</sup>

<sup>1</sup>Geography Institute/Quality Control Lab, Federal University of Uberlândia, 38500-000, Monte Carmelo, MG, Brazil

<sup>2</sup>Graduate Program in Remote Sensing, Federal University of Rio Grande do Sul, 91501-970, Porto Alegre, RS, Brazil

<sup>3</sup>Graduate Program in Agriculture and Geospatial Information, Federal University of Uberlândia, 38500-000, Monte Carmelo, MG, Brazil

<sup>4</sup>Land Surveying Program, Federal Institute of Santa Catarina, 88020-300, Florianópolis, SC, Brazil

<sup>5</sup>Graduate Program in Geodetic Sciences, Federal University of Paraná, 80060-000, Curitiba, PR, Brazil

<sup>6</sup>Graduate Program in Applied Computing/Advanced Visualization Laboratory (Vizlab), 93022-750, Unisinos University, São Leopoldo, RS, Brazil

\*vprofatto@gmail.com

(Recebido em 14 de agosto de 2019; aceito em 12 de setembro de 2019)

One of the main tasks of professionals in the Earth sciences is to convert coordinates from one reference frame into another. The coordinates transformation is widely needed and applied in all branches of modern geospatial activities. To convert from one reference frame to another, it is necessary initially to determine the parameters of a coordinate transformation model. A common way to estimate the transformation parameters is using the least-squares theory within a linearized Gauss-Markov Model (GMM). Another approach arises from a numerical method. Here, a Monte Carlo Method (MCM) is used to infer the uncertainty of estimators. This method is based on: (i) the assignment of probability distributions to the coordinates in the two reference frames, (ii) the determination of a discrete representation of the probability distribution for the transformation parameters, and (iii) the determination of the associated uncertainties from this discrete representation of the estimates of the transformation parameters. In this contribution, we compare the weighted least-squares within the GMM (WLSE-GM) and the proposed method regarding the transformation problem of computing 2D similarity transformation parameters. The results show that, the transformation parameters uncertainties are higher for the LS-MC than the WLSE-GM. This is due to the fact that the WLSE-GM solution does not take into account the uncertainties associated with the system matrix. In future studies the Monte Carlo method should be applied to the nonlinear least-squares solution.

Keywords: Uncertainty Propagation, Monte Carlo, Coordinate Transformation.

Uma das principais tarefas dos profissionais envolvidos com as Ciências da Terra é a transformação de pontos cujas coordenadas estão vinculadas em referenciais distintos. A transformação é necessária e amplamente aplicada em diversas atividades que lidam com informações espaciais. Para realizar tal conversão, inicialmente, deve-se dispor dos parâmetros de transformação. Geralmente, os parâmetros de transformação são estimados por meio da aplicação do método dos mínimos quadrados (MMQ) sob a condição do modelo de Gauss-Markov (GMM). Outra abordagem surge de um método numérico. Aqui, o método Monte Carlo (MCM) é utilizado para inferir as incertezas das estimativas do MMQ quando as coordenadas em ambos os referenciais são tratadas como observações e, portanto, sujeitas a erros aleatórios. Três aspectos principais do método proposto são: (i) especificação da distribuição de probabilidade das coordenadas em ambos os referenciais; (ii) determinação empírica da distribuição acumulada dos parâmetros de transformação, e (iii) estimativa das incertezas dos parâmetros de transformação. Nesta contribuição, comparamos os mínimos quadrados ponderados (WLSE-GM) e o método proposto dentro do contexto de GMM. A comparação foi realizada por meio de um exemplo numérico de transformação de similaridade no espaço 2D. Os resultados mostram que as incertezas dos parâmetros de transformação são maiores para o método proposto que o WLSE-GM. Isso se deve ao fato de que a solução WLSE-GM não leva em conta as incertezas associadas à matriz *design*. Em estudos futuros, o método Monte Carlo deve ser aplicado à solução não linear de mínimos quadrados.

Palavras-chaves: Propagação de Incertezas, Monte Carlo, Transformação de Coordenadas.

## 1. INTRODUCTION

One of the main tasks of professionals in the Earth sciences is to convert coordinates from one reference frame into another. The coordinates transformation is widely needed and applied in all branches of modern geospatial activities. Those fields of application range from satellite positioning [1] to photogrammetry [2].

In general, the process of conversion is performed in two steps. First, the transformation parameters are estimated from a redundant set of points of coordinates given in two different reference frames known as control points (also referenced as identical or homologous points). Second, the transformation parameters estimated in the first step are used for the conversion of the points from one frame to another. Moreover, the law of propagation of uncertainty may also be applied in order to estimate the uncertainties of transformed points [3]. Typically, the standard uncertainty component is quantified by variance or standard deviation.

A variety of transformation models are available, such as: Body rigid, Similarity, Orthogonal, Affine, Polynomial, Projective, among others. An overview of different transformation models may be found in Greenfeld (1997) [4]. In the majority of cases, the transformation problem is formulated as a mathematical problem of unknown transformation parameters in a standard Gauss-Markov model (GMM). In that case, the transformation parameters can easily be computed by the application of the least-squares adjustment method (e.g., Ghilani, 2010, pp. 345-368) [3].

However, in the last decades many new computation methods have been developed to deal with the problem of coordinate transformation. Here, we list some recent contributions:

- Robust method: Janicka & Rapinski (2013) [5] used a robust method, showing that coordinates of control points are less influential to outliers.

- Multiple hypothesis testing and Akaike information criterion for transformation model selection: Lehmann (2014) [6] showed how to select a proper transformation model by hypothesis testing, and also extended the approach so that not only parameters themselves can be tested but also constraints on those parameters.

- Computational-Intelligence algorithms: the use of algorithms to search for a solution for a problem by natural or artificial agents through the use of collective methods has been investigated in the literature. For instance, Konakoglu et al. (2016) [7] investigated the performances of three different artificial neural network models (Feed Forward Back Propagation, Cascade Forward Back Propagation, and Radial Basis Function Neural Network) concerning 2D coordinate transformation. Civicioglu (2012) [8] compared different computational-intelligence algorithms, such as Differential Search Algorithm, Artificial Bee Colony, Self-Adaptive Differential Evolution Algorithm, Adaptive Differential Evolution Algorithm, Particle Swarm Optimization Algorithm, and many others.

- A more recent development is associated with adjusting the ‘error-in-variables’ (EIV) model by a total least-squares approach. In that case, the coordinates of the control points in both reference frames are observations and, therefore, are subjected to random errors. However, based on the example of a planar similarity transformation, Neitzel (2010) [9] showed that the total least-squares solution can be easily obtained from a rigorous evaluation of the Gauss-Helmert model. Discussions regarding total least-squares algorithms can be seen in [10 - 18].

Despite the countless contributions made over the years, there is continuing research on the subject, mainly because there has been an increase in computational power. In this sense, apart from the purely analytical approaches, we proposed an alternative method that incorporates the least-squares estimation (LS) into the Gauss-Markov model (GMM) together with a Monte Carlo method (MCM). We denominated it the Least-Squares-based Monte Carlo method (LS-MC). Here, we restrict ourselves to the Gauss-Markov model (GMM) for the problem of coordinate transformation, because it is the most typical situation in geodesy and geophysics. This contribution is also confined to regular models in which no singular matrices occur. However, singular cases can be treated in an analogous manner.

In the sense of standard GMM, the random errors are only assigned to the coordinates of control points in a unique frame, whereas the elements of the system matrix are treated as error-free quantities. This poses the problem, into GMM, of assessing the uncertainties associated with transformation parameters in the case where the coordinates in the two reference frames are

observations, and therefore subjected to random errors. This problem may be formalized as the integration of parameters estimation and uncertainty propagation in the context where the random errors contaminate both the observations vector and system matrix within GMM.

Moreover, replacing physical parameters with synthetic ones in order to simplify the transformation model to a linear equation system has been a common practice in geodesy and geophysics. However, the linearization of the model that relates the synthetic parameters to the physical ones is still necessary in order to have the uncertainties of the estimated physical parameters. In that case, the first order of the Taylor series expansion can be applied to solve it. Nevertheless, the linearized model can often provide an inadequate representation of the uncertainties of parameters when the model itself is highly non-linear and/or the uncertainty of measurements is very low [19]. Analytically expressible solutions are ideal only in cases where they do not introduce any approximation. Therefore, a more appropriate method would be based on the propagation of probability density functions (PDFs) using a numerical method, specifically a Monte Carlo method (MCM).

Another problem also encountered in geophysical and geodetic applications concerns the construction of the confidence interval (CI) associated with the transformation parameters. The CI (also referred to as the coverage interval or in some cases the coverage region) is closely related to statistical significance testing. For example, if for some estimated parameter  $\theta$  one wants to test the null hypothesis that  $\theta = 0$  against the alternative that  $\theta \neq 0$ , then this test can be performed by determining whether the confidence interval for  $\theta$  contains 0. More generally, given the availability of a hypothesis testing procedure that can test the null hypothesis  $\theta = \theta_0$  against the alternative that  $\theta \neq \theta_0$  for any value of  $\theta_0$ , then a confidence interval with confidence level  $CI = 1 - \alpha$  can be defined as containing any number  $\theta_0$  for which the corresponding null hypothesis is not rejected at significance level  $\alpha$  [20]. The CI of a least-squares parameter is often constructed by multiplying the standard deviation of a parameter by the appropriate critical value, i.e., by a quantile value of the probability distribution associated with that parameter. This quantile is computed for a given probability. For example, the probability that a Gaussian random variable is within 2 standard deviations from its mean is approximately 95%. This means that 95% of the time the most probable value will be contained in that interval. Generally, standard normal and t-student distributions are used to predict the CI in which the unknown true value of a parameter can exist. The critical values of these distributions are found in widespread statistical lookup tables or computed by popular scientific software. In this contribution, we will show how LS-MC can be used to construct the CI without the need to resort to tabulated critical values.

The outline of the paper is as follows: First, we introduce the mathematical problem of determining the transformation parameters using weighted least-squares in the case where uncertainties exist in both the observations vector and coefficient matrix of the Gauss-Markov model (GMM), possibly of full rank. Second, we show the main aspects to implement the Least-Squares-based Monte Carlo method. Finally, we present a numerical example using the 2D Similarity Transformation to illustrate the proposed LS-MC method. It is important to mention that, although we restrict ourselves to a coordinate transformation problem, the proposed LS-MC is a generally applicable method.

## 2. THE GAUSS-MARKOV MODEL IN THE CONTEXT OF COORDINATE TRANSFORMATION

We start with the following transformation model:

$$\begin{bmatrix} X_i \\ Y_i \end{bmatrix} = s \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \end{bmatrix} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} c \\ d \end{bmatrix} \quad (1)$$

with  $i, \dots, n$  the number of homologous points of the coordinate systems XY and xy. It is important to mention that the number of observations is  $2n$ , where  $n$  is the number of control points in both reference frames. In that case, the scale factor  $s$  and the rotation angle  $\alpha$  are given in function of the synthetic parameters as follow:

$$s = \sqrt{a^2 + b^2} \quad (2)$$

$$\alpha = \tan^{-1}\left(\frac{b}{a}\right) \quad (3)$$

The transformation given by Eq. (1) employs four physical parameters: a single scale factor "s" along the X- and Y-axes, one rotation angle  $\alpha$ , and two translations (*shifting*) of the coordinate origin,  $T_x$  and  $T_y$ . This transformation model is known as the 2D Similarity Planar Transformation Model and it is often used in applications of geodesy, geophysics, engineering, computer vision, geographical information science, and related branches of science and engineering. Figure 1 shows a schematic diagram of the 2D similarity transformation model. In geodesy and geophysics, we most often replace the physical parameters by synthetic ones in order to simplify the transformation model to a linear equation system, i.e., we take  $a = s \times \cos(\alpha)$ ,  $b = s \times \sin(\alpha)$ ,  $c = T_x$  and  $d = T_y$ .

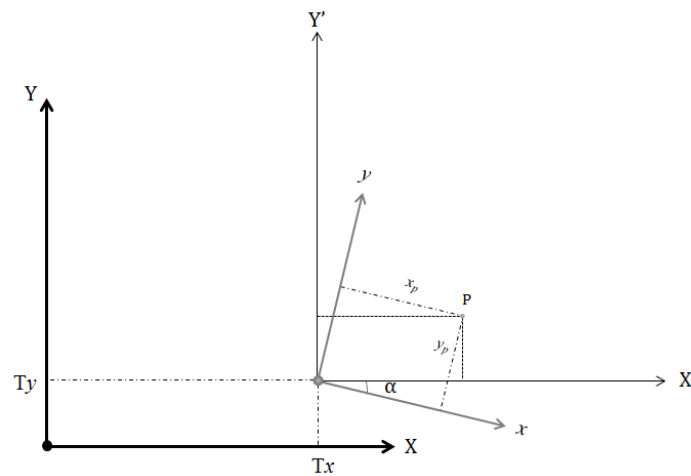


Figure 1: Schematic diagram of the Similarity Transformation Model.

The transformation Eq. (1) can be easily formulated as a standard Gauss-Markov model (GMM) as follow:

$$\mathbf{y}^* = \mathbf{A}\mathbf{X} \quad (4)$$

with

$$\mathbf{y}^* = \mathbf{y} - \mathbf{e}_y, \text{ with } E\{\mathbf{e}_y\} = \mathbf{0} \quad (5)$$

where  $\mathbf{y} \in \mathbb{R}^n$  the vector of observations,  $\mathbf{A} \in \mathbb{R}^{n \times u}$  the Jacobian matrix (also called design matrix or coefficient matrix) of full rank  $u$ ,  $\mathbf{X} \in \mathbb{R}^u$  the unknown parameter vector, and  $\mathbf{e}_y \in \mathbb{R}^n$  is an unknown random vector of normally distributed observation errors and  $E$  denote the operators of expectation. In this case, the overall redundancy (also referred to as degrees of freedom)  $df$  of the model in Eq. (4) is  $df = n - u$ , i.e., the Eq. (4) represents an over-determined system of equations.

In that case, the mathematical problem of determining the unknown transformation parameters is formulated under assumption that the coordinates  $[X_i, Y_i]$  are the observations and therefore subjected to random errors, whereas the coordinates  $[x_i, y_i]$  are error-free quantities that are treated as constants in the GMM, i.e.:

$$\mathbf{y} = \begin{bmatrix} X_1 \\ Y_1 \\ X_2 \\ Y_2 \\ \vdots \\ X_n \\ Y_n \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} x_1 & y_1 & 1 & 0 \\ y_1 & -x_1 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & 1 & 0 \\ y_n & -x_n & 0 & 1 \end{bmatrix} \quad (6)$$

It is also possible to formulate  $[x_i, y_i]$  in function of  $[X_i, Y_i]$  so that in this case  $[x_i, y_i]$  are random variables and  $[X_i, Y_i]$  are error-free quantities. In any case, GMM in Eq. (4) corresponds to a supposedly valid model describing the physical reality of the observations without bias and that only the measurement vector  $\mathbf{y}$  are affected by random errors. In that case, the associated stochastic model is assumed to be:

$$\mathbf{e}_y \sim N(\boldsymbol{\mu}_y = \mathbf{0}, \boldsymbol{\Sigma}_y = \sigma^2 \mathbf{W}^{-1}) \quad (7)$$

where  $N(\cdot)$  represents the multivariate normal distribution function,  $\boldsymbol{\mu}_y \in \mathbb{R}^n$  is a vector of zeros means,  $\sigma^2$  is the a priori variance factor, which may be either known or unknown,  $\mathbf{W} \in \mathbb{R}^{n \times n}$  is a known weight matrix, and  $\boldsymbol{\Sigma}_y \in \mathbb{R}^{n \times n}$  is the covariance matrix of the observations.

When it is assumed to be ‘true’, the model in (1) is used to estimate the unknown parameters, typically in a least-squares approach. In that case, the objective function that should be minimized is given by:

$$\mathbf{e}_y^T \mathbf{W} \mathbf{e}_y \rightarrow \min \quad (8)$$

The least-squares solution for the estimated parameters reads:

$$\hat{\mathbf{X}} = (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W} \mathbf{y} \quad (9)$$

The uncertainties of the least-squares solution for the estimated parameters are given by:

$$D\{\hat{\mathbf{X}}\} = \sigma^2 (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \quad (10)$$

with  $D$  denote the operators of dispersion.

The least-squares residuals (estimated observation errors) are always unique and we obtain them by:

$$\hat{\mathbf{e}}_y = \mathbf{y} - \mathbf{A} \hat{\mathbf{X}} \quad (11)$$

The unbiased estimate of the variance factor  $\sigma^2$  is given by:

$$\hat{\sigma}^2 = \frac{\mathbf{e}_y^T \mathbf{W} \mathbf{e}_y}{n - u} \quad (12)$$

The symbol “ $\hat{\cdot}$ ” represents the weighted least-squares estimator (WLSE) for the parameters  $\mathbf{X}$ ,  $\mathbf{e}_y$ , and  $\sigma^2$ . The element  $\hat{\sigma}^2$  is referred to as posteriori variance component. If there are only random errors in the observations, the WLSE is the best linear unbiased estimator (BLUE) for the unknown parameters while if the observational errors follow the multivariate normal distribution with zeros means and a symmetric positive-definite variance-covariance matrix, the WLSE coincides with the maximum likelihood estimator [21]. Since the WLSE is applied to a GMM model, we referred to it as WLSE-GM.

In the WLSE-GM, the uncertainties associated with the physical parameters of rotation  $\alpha$  and scale  $s$  are computed using the classical law of uncertainty propagation. However, in that case, it is

not possible to obtain these uncertainties directly because the physical quantities are not linear with respect to the synthetic parameters, as can be seen in Eq. (2) and (3). In that case, the Taylor series expansion is often applied to linearize these equations, where the higher order terms are neglected. Therefore, the uncertainties of the physical parameters are approximated by linearization.

Until now, we assume the coordinates from one reference frame as observations. However, often in practice, both coordinates of control points,  $[X_i, Y_i]$  and  $[x_i, y_i]$ , are observations and therefore are subjected to random errors. In that case, the model is given by:

$$\mathbf{y} - \mathbf{e}_y = \mathbf{A}^* \mathbf{X} \quad (13)$$

with

$$\mathbf{A}^* = \mathbf{A} - \mathbf{E}_A \quad (14)$$

and

$$\mathbf{E}_A \sim N(\boldsymbol{\mu}_A = \mathbf{0}, \boldsymbol{\Sigma}_A = \sigma_A^2 \mathbf{W}_A^{-1}) \quad (15)$$

where  $\mathbf{E}_A \in \mathbb{R}^{n \times u}$  is the corresponding random noise matrix of the functional matrix  $\mathbf{A}$ ,  $\boldsymbol{\mu}_A \in \mathbb{R}^n$  is a vector of zeros means,  $\sigma_A^2$  is the a priori variance factor, which may be either known or unknown,  $\mathbf{W}_A \in \mathbb{R}^{n \times n}$  is the weight matrix, and  $\boldsymbol{\Sigma}_A \in \mathbb{R}^{n \times n}$  is the covariance matrix (note: the subscript “A” refers to elements of matrix  $\mathbf{A}$ ).

In the case in which the coordinates of the control points in both reference frames are subjected to random errors, the matrix  $\mathbf{A}^*$  for the similarity model is given as follows:

$$\mathbf{A} = \begin{bmatrix} x_1 - e_{x_1} & y_1 - e_{y_1} & 1 & 0 \\ y_1 - e_{y_1} & -x_1 - e_{x_1} & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_n - e_{x_n} & y_n - e_{y_n} & 1 & 0 \\ y_n - e_{y_n} & -x_n - e_{x_n} & 0 & 1 \end{bmatrix} \quad (16)$$

The solution of Eq. (13) can be treated as a total least-squares approach in the sense of Error-in-Variables model. In that case, [9] has shown that the total least-squares solution can be obtained easily from a rigorous evaluation of the Gauss–Helmert model. A good discussion about total least-squares algorithms can be seen in [9, 10, 18]. Different from the approaches mentioned so far, in the following sections, we present an alternative method based on the Monte Carlo method.

### 3. IMPLEMENTATION OF THE LEAST-SQUARES METHOD BASED ON THE MONTE CARLO METHOD FOR COORDINATE TRANSFORMATION

Previously, the major drawbacks of Monte Carlo methods (MCM) were the low speed and high computation requirements. These limitations have been surmounted over the years, mainly by the rapid development of computers, which now allow this technique to be used efficiently on personal computers or even on handheld computers. One of the advantages of the MCM is that one can reduce or even dispense analytical formulae [22 - 25]. Since powerful computers are available everywhere, here we use the MCM to calculate the statistical properties of the transformation parameters.

The basic idea of this contribution is to provide an alternative method to include in the estimation process of a Gauss-Markov Model (GMM) the random errors of the coordinates in two reference frames using pseudo-random number generators based on a probability density function (PDF) of observations. The result of thousands of parameter estimations makes it possible to describe the frequency distributions of random computer experiments performed using pseudo-random numbers. We called this method LS-MC.

A random number generator is an algorithm that generates a deterministic sequence of numbers, which simulates a sequence of independent and identically distributed numbers chosen uniformly

between 0 and 1. It is ‘random’ in the sense that the generated sequence of numbers passes the statistical tests for randomness. For this reason, random number generators are typically referred to as pseudo-random number generators (PRN). The PRN is part of many machine learning and data mining techniques [26]. A good generator produces numbers that are not distinguishable from truly random numbers in limited computation time. This is, in particular, true for Mersenne Twister, a popular generator with a long period length of  $2^{19937} - 1$  [27].

Typically, in geodesy, the random errors of good measurements are normally distributed with expectation zero. In order to have normal random errors, uniformly distributed random number sequences (produced by the Mersenne Twister algorithm, for example) are transformed into a normal distribution using the Box–Muller transformation [28]. Here, we use Box–Muller, which has been used in geodesy for the MCM [29, 30].

In essence, the MCM replaces random variables by PRN, probabilities by relative frequencies, and expectations by arithmetic means over large sets of such numbers. A computation with one set of PRN is a Monte Carlo trial [31], also referred to as the number of Monte Carlo experiments [32].

In other words, the key of LS-MC is making repeated draws from PDFs of control points coordinates in both frames, so that for each draw the transformation parameters are calculated by the least-square method. Therefore, the distribution function of transformation parameters can be derived by LS-MC, which makes the method attractive because we no longer need to resort the well-known tabulated statistical distributions, such as normal distribution, and so on. The LS-MC procedure is summarised as a diagram in Figure 2.

The LS-MC procedure is as follow:

- (1) Define the mathematical relation among all quantities known to be involved in a model. Here, it would be defined by the transformation model, for example, Similarity given by the expression (1).
- (2) Set up the number of MCM trials, here denoted by  $M$ .
- (3) Specify the PDFs to the input quantities, i.e., it should be provided the PDFs of the control points coordinates in both reference frames, e.g., a joint normal distribution given by the expression (7) and another one by (15).
- (4) Generate  $M$  vectors, by drawing randomly from the PDFs assigned to the input quantities according to step (3). In the case of 2D transformation, we would have  $\mathbf{v}_1, \dots, \mathbf{v}_M$  belonging to the  $xy$  coordinate system and  $\mathbf{w}_1, \dots, \mathbf{w}_M$  to the  $XY$  system so that for each vector  $\mathbf{v}$  there would be coordinates  $x_1, y_1, \dots, x_n, y_n$  and for each vector  $\mathbf{w}$  there would be coordinates  $X_1, Y_1, \dots, X_n, Y_n$ .
- (5) For each  $\mathbf{v}$  and  $\mathbf{w}$  vectors compute the least-square solution according to Eq. (9) and (11), yielding  $M$  draws of transformation parameters  $\mathbf{X}_k^*$  and random vector observation errors  $\mathbf{e}_{y_k}^*$ , respectively, i.e.:

$$\mathbf{X}_k^* = \left( \mathbf{A}_k^{*T} \mathbf{W} \mathbf{A}_k^* \right)^{-1} \mathbf{A}_k^{*T} \mathbf{W} \mathbf{y}_k^*, (\forall k = 1, \dots, M) \quad (17)$$

$$\mathbf{e}_{y_k}^* = \mathbf{y}_k^* - \mathbf{A}_k^* \mathbf{X}_k^* (\forall k = 1, \dots, M) \quad (18)$$

where the symbol ‘\*’ represents continuous random variables. The matrix  $\mathbf{A}^*$  and the vector  $\mathbf{y}^*$  may be defined according to the expressions (14), and (5), respectively. The output quantities  $\mathbf{X}_k^*$  and  $\mathbf{e}_{y_k}^*$  are computed for each of the  $M$  draws from the PDFs of the coordinates of the control points.

- (6) Use these  $M$  values to calculate an estimate  $\hat{\mathbf{X}}$  of  $\mathbf{X}$  and the associated covariance matrix  $\Sigma_{\hat{\mathbf{X}}}$ . Similarly, compute an estimate  $\hat{\mathbf{e}}_y$  of  $\mathbf{e}_y$  and the covariance matrix  $\Sigma_{\hat{\mathbf{e}}_y}$ , i.e.:

$$\hat{\mathbf{X}} = \frac{1}{M} \sum_{k=1}^M \mathbf{X}_k^* \quad (19)$$

$$\Sigma_{\hat{\mathbf{X}}} = \frac{1}{M-1} \left[ (\mathbf{X}_1^* - \hat{\mathbf{X}})(\mathbf{X}_1^* - \hat{\mathbf{X}})^T + \dots + (\mathbf{X}_M^* - \hat{\mathbf{X}})(\mathbf{X}_M^* - \hat{\mathbf{X}})^T \right] \quad (20)$$

$$\hat{\mathbf{e}}_y = \frac{1}{M} \sum_{k=1}^M \mathbf{e}_{y_k}^* \quad (21)$$

$$\Sigma_{\hat{\mathbf{e}}} = \frac{1}{M-1} \left[ (\mathbf{e}_{y_1}^* - \hat{\mathbf{e}}_y)(\mathbf{e}_{y_1}^* - \hat{\mathbf{e}}_y)^T + \dots + (\mathbf{e}_{y_M}^* - \hat{\mathbf{e}}_y)(\mathbf{e}_{y_M}^* - \hat{\mathbf{e}}_y)^T \right] \quad (22)$$

- (7) Sort  $M$  values from (19) and (21) into strictly increasing order for each parameter of  $\mathbf{X}$  and  $\mathbf{e}_y$ , respectively. The sorted values are denoted by  $\mathbf{X}_s^*$  and  $\mathbf{e}_{y_s}^*$ , for  $s = 1, \dots, M$ , respectively. These  $\mathbf{X}_s^*$  and  $\mathbf{e}_{y_s}^*$  values provide a discrete representation of the cumulative density functions (CDFs), say  $G_{\mathbf{X}_s^*}$  and  $G_{\mathbf{e}_{y_s}^*}$  for  $\mathbf{X}$  and  $\mathbf{e}_y$ , respectively. In other words, we get the discrete representations of CDFs of each transformation parameter as well as for each unknown random observation errors in the sense of the GMM.
- (8) From the discrete representation  $G_{\mathbf{X}_s^*}$  by the step above, form a left-hand  $X_{low}$  and a right hand  $X_{high}$  endpoints of a coverage interval for  $\mathbf{X}$  to a given probability  $p = 1 - \alpha$ , where  $\alpha$  is the significance level. The probability  $p$  provides a region where the true parameter is most likely to lie. In that case, the left-hand endpoint is formed by taking  $X_{low} = \mathbf{X}_{(r)}^*$  and the right-hand endpoint by  $X_{high} = \mathbf{X}_{(r+q)}^*$ , where  $q = (p) \times M$  and  $r = (M - q)/2$ . Different from the well-known confidence interval, the coverage interval contains the value of a quantity with a given probability  $p$  based on the information available. For example, for  $M = 10^4$  trials and  $\alpha = 0.05$ , we obtain  $p = 0.95$ ,  $q = 9500$ , and  $r = 250$ . In that case, the coverage interval was defined by the 0.025- and 0.975-quantiles. This is the case for a probabilistically symmetric  $100 \times p\%$  coverage interval and for the case where the distribution is single-peaked. A good discussion about coverage intervals (also referred to as expanded uncertainties) can be seen in [33].



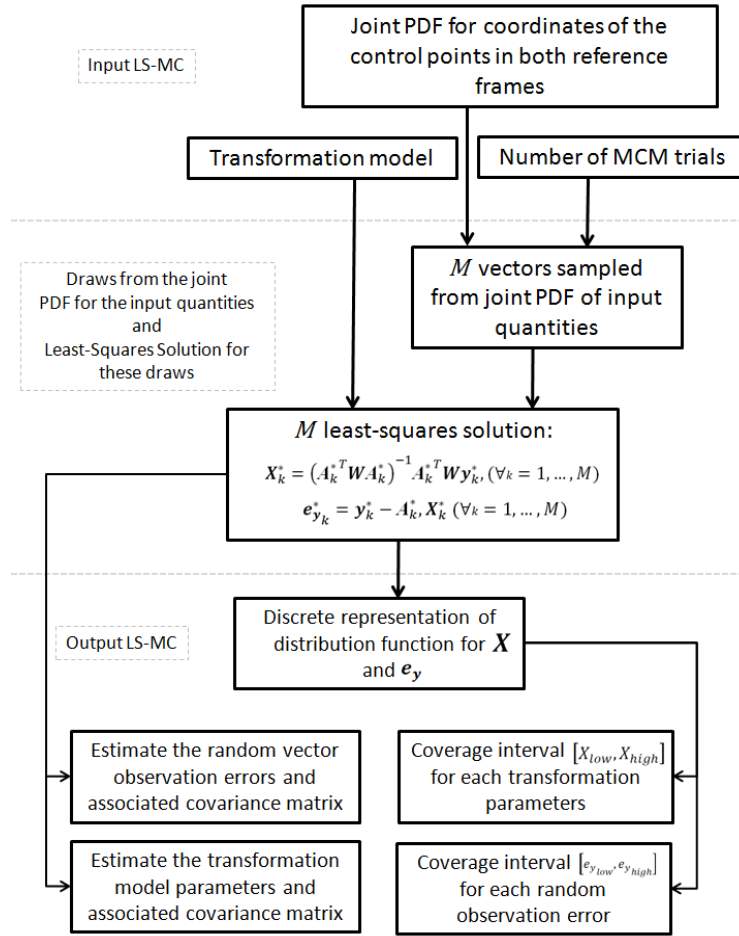


Figure 2: Main stages of LS-MC for coordinate transformation problem in the sense of GMM.

#### 4. GENERATION OF OBSERVATION ERRORS AND NUMBER OF MONTE CARLO TRIALS

It is necessary to assume the stochastically properties of the observation errors. These properties are completely derived from the probability distribution of these errors. Here, the random error vectors  $\mathbf{e}_y$  and the random noise matrix  $\mathbf{E}_A$  of the functional matrix  $\mathbf{A}$  have been synthetically generated based on a multivariate normal distribution according to Eq. (7) and (15), respectively. The design matrix  $\mathbf{A}^*$  for the similarity transformation is given by (16). This assumption is very reasonable because, typically, the random errors of the good measurements are normally distributed with expectation zero. Here, we use the Mersenne Twister algorithm to generate a sequence of random numbers and the Box–Muller method to transform it into a normal distribution. It is important to emphasize that unlike standard practice, which considers only the coordinates  $XY$  as those that contain random errors, here the coordinates of the control points in both reference frames are subjected to random errors. For a synopsis of transformation model selection and observation error model selection, see Lehmann (2014) [6] and Lehmann (2015) [29], respectively.

For this purpose, we *arbitrary* choose  $M=10^6$  trials for generating both random error vectors  $\mathbf{e}_y$  and random noise matrix  $\mathbf{E}_A$ . It has been investigated that the results presented here do not change significantly when the computations are repeated with different PRN (see Figure 3). For the multivariate normal PRN, we can use directly MATLAB R2018b PRN generator “*mvnrnd*”. In our analyses, Intel Core i5-4200M CPU @2.5GHz and MATLAB R2018b are used for the experiments.

Figure 3 shows an example of the standard deviation of the LS-MC as a function of the number of MCM trials for the computation of the synthetic parameter  $a$  from expression (1). In this example, each set of MCM trials was run 1000 times to calculate the standard deviation ( $\sigma$ ) of the LS-MC. This standard deviation represents the variability of the PRN for each set of MCM trials.

It is important to mention that the PRN generator is reseeded each time. Note that the standard deviations tend to be reduced proportionally to  $M^{-1/2}$ . The standard deviation for  $M=10^6$  trials was better than  $10^{-5}$  unit. As pointed out [34], the quality of MCM does not depend on the dimension of the data space. The quality of these calculated results generally improves with the number of draws made. The question of how to find the optimal number of MCM trials to evaluate the quality of experimental results has already been addressed by Rofatto et al. (2018) [30].

The computation time for  $M=10^6$  trials took a few seconds on a personal computer operating at 2.5 GHz. Hence, we are not overwhelmed with the computational workload.

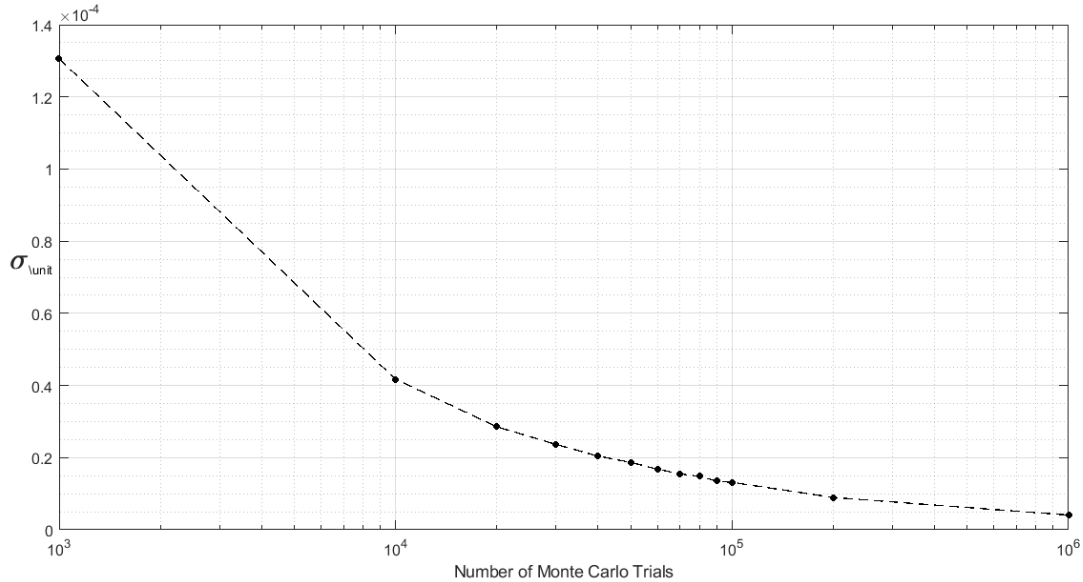


Figure 3: Empirical standard-deviation ( $\sigma$ ) of LS-MC for each set of MCM trials run 1000 times.

## 5. EXAMPLE OF LS-MC APPLIED TO A SIMILARITY TRANSFORMATION OF PLANAR COORDINATES

Here, we intended to demonstrate how the LS-MC can be applied in terms of coordinate transformation. The coordinates in the frame XY and xy are presented in Table 1, which are taken from Ghilani (2010, pp. 370, Table 18.1) [3]. This numerical case refers to the 2D similarity transformation given by Eq. (1). Here, the transformation parameters are obtained by both the LS-MC and WLSE-GM methods.

Table 1: Numerical example of coordinates in the frames XY and xy from Ghilani (2010) [3].

Points	Coordinates in XY frame		Coordinates in xy frame	
	X [m]	Y [m]	x [m]	y [m]
A	1049422.40	51089.20	121.622	-128.066
B	1049413.95	49659.30	141.228	187.718
C	1049244.95	49884.95	175.802	135.728

Four different scenarios are considered here (and summarized in Table 2):

Table 2: Different scenarios for variances of the coordinates in the frames XY and xy.

Number of scenarios	Error Model								
First scenario	$\mathbf{e}_y \sim N(\boldsymbol{\mu}_y = \mathbf{0}, \boldsymbol{\Sigma}_y)$ $\mathbf{E}_A \sim N(\boldsymbol{\mu}_A = \mathbf{0}, \boldsymbol{\Sigma}_A = \boldsymbol{\Sigma}_y)$ $\sigma^2 = \sigma_A^2 = 1.00\text{e-}6 \text{ m}^2$								
Second scenario	$\mathbf{e}_y \sim N(\boldsymbol{\mu}_y = \mathbf{0}, \boldsymbol{\Sigma}_y = 100 \times \boldsymbol{\Sigma}_A)$ $\mathbf{E}_A \sim N(\boldsymbol{\mu}_A = \mathbf{0}, \boldsymbol{\Sigma}_A)$ $\sigma^2 = 100 \times \sigma_A^2 = 1.00\text{e-}4 \text{ m}^2$								
Third scenario	$\mathbf{e}_y \sim N(\boldsymbol{\mu}_y = \mathbf{0}, \boldsymbol{\Sigma}_y)$ $\sigma_{x_A}^2 = 1.00\text{e-}4 \text{ m}^2, \sigma_{y_A}^2 = 4.41\text{e-}4 \text{ m}^2,$ $\sigma_{x_B}^2 = 2.56\text{e-}4 \text{ m}^2, \sigma_{y_B}^2 = 1.00\text{e-}4 \text{ m}^2,$ $\sigma_{x_C}^2 = 3.24\text{e-}4 \text{ m}^2, \sigma_{y_C}^2 = 2.25\text{e-}4 \text{ m}^2.$								
Fourth scenario	$\mathbf{e}_y \sim N(\boldsymbol{\mu}_y = \mathbf{0}, \boldsymbol{\Sigma}_y)$ $\mathbf{E}_A \sim N(\boldsymbol{\mu}_A = \mathbf{0}, \boldsymbol{\Sigma}_A \neq \boldsymbol{\Sigma}_y)$ <table style="width: 100%; border: none;"> <tr> <td style="text-align: center;">Frame XY</td> <td style="text-align: center;">Frame xy</td> </tr> <tr> <td><math>\sigma_{x_A}^2 = 1.00\text{e-}4 \text{ m}^2, \sigma_{y_A}^2 = 4.41\text{e-}4 \text{ m}^2</math></td> <td><math>\sigma_{x_A}^2 = 8.84\text{e-}5 \text{ m}^2, \sigma_{y_A}^2 = 3.36\text{e-}5 \text{ m}^2</math></td> </tr> <tr> <td><math>\sigma_{x_B}^2 = 2.56\text{e-}4 \text{ m}^2, \sigma_{y_B}^2 = 1.00\text{e-}4 \text{ m}^2</math></td> <td><math>\sigma_{x_B}^2 = 8.84\text{e-}5 \text{ m}^2, \sigma_{y_B}^2 = 1.16\text{e-}5 \text{ m}^2</math></td> </tr> <tr> <td><math>\sigma_{x_C}^2 = 3.24\text{e-}4 \text{ m}^2, \sigma_{y_C}^2 = 2.25\text{e-}4 \text{ m}^2</math></td> <td><math>\sigma_{x_C}^2 = 9.60\text{e-}5 \text{ m}^2, \sigma_{y_C}^2 = 4.62\text{e-}5 \text{ m}^2</math></td> </tr> </table>	Frame XY	Frame xy	$\sigma_{x_A}^2 = 1.00\text{e-}4 \text{ m}^2, \sigma_{y_A}^2 = 4.41\text{e-}4 \text{ m}^2$	$\sigma_{x_A}^2 = 8.84\text{e-}5 \text{ m}^2, \sigma_{y_A}^2 = 3.36\text{e-}5 \text{ m}^2$	$\sigma_{x_B}^2 = 2.56\text{e-}4 \text{ m}^2, \sigma_{y_B}^2 = 1.00\text{e-}4 \text{ m}^2$	$\sigma_{x_B}^2 = 8.84\text{e-}5 \text{ m}^2, \sigma_{y_B}^2 = 1.16\text{e-}5 \text{ m}^2$	$\sigma_{x_C}^2 = 3.24\text{e-}4 \text{ m}^2, \sigma_{y_C}^2 = 2.25\text{e-}4 \text{ m}^2$	$\sigma_{x_C}^2 = 9.60\text{e-}5 \text{ m}^2, \sigma_{y_C}^2 = 4.62\text{e-}5 \text{ m}^2$
Frame XY	Frame xy								
$\sigma_{x_A}^2 = 1.00\text{e-}4 \text{ m}^2, \sigma_{y_A}^2 = 4.41\text{e-}4 \text{ m}^2$	$\sigma_{x_A}^2 = 8.84\text{e-}5 \text{ m}^2, \sigma_{y_A}^2 = 3.36\text{e-}5 \text{ m}^2$								
$\sigma_{x_B}^2 = 2.56\text{e-}4 \text{ m}^2, \sigma_{y_B}^2 = 1.00\text{e-}4 \text{ m}^2$	$\sigma_{x_B}^2 = 8.84\text{e-}5 \text{ m}^2, \sigma_{y_B}^2 = 1.16\text{e-}5 \text{ m}^2$								
$\sigma_{x_C}^2 = 3.24\text{e-}4 \text{ m}^2, \sigma_{y_C}^2 = 2.25\text{e-}4 \text{ m}^2$	$\sigma_{x_C}^2 = 9.60\text{e-}5 \text{ m}^2, \sigma_{y_C}^2 = 4.62\text{e-}5 \text{ m}^2$								

1. First scenario: In the first situation, as observation error models we choose are multivariate normal distribution  $\mathbf{e}_y \sim N(\boldsymbol{\mu}_y = \mathbf{0}, \boldsymbol{\Sigma}_y)$  and  $\mathbf{E}_A \sim N(\boldsymbol{\mu}_A = \mathbf{0}, \boldsymbol{\Sigma}_A)$ , so that  $\boldsymbol{\Sigma}_A = \boldsymbol{\Sigma}_y$ . In this example, the coordinates in frames XY and xy are regarded as equally weighted uncorrelated observations with  $\sigma^2 = \sigma_A^2 = 1 \text{ mm}^2$ . This is a case that may rarely occur in practice since it implies that the coordinates in the frames XY and xy are computed with the same instrument precision, measurement techniques, and working conditions.
2. Second scenario: In the second one, we also choose the multivariate normal distribution, so that  $\mathbf{e}_y \sim N(\boldsymbol{\mu}_y = \mathbf{0}, \boldsymbol{\Sigma}_y)$  and  $\mathbf{E}_A \sim N(\boldsymbol{\mu}_A = \mathbf{0}, \boldsymbol{\Sigma}_A)$ . The coordinates in frames XY and xy are regarded as equally weighted uncorrelated observations, but here  $\sigma^2 = 100 \times \sigma_A^2 = 100 \text{ mm}^2$ . This means that the coordinates in the frame XY are less precise than the coordinates in frame xy. In practical terms, for example, this would mean that the coordinates in the frame XY and xy are computed with a different instrument precision, but with the same measurement technique and work conditions.
3. Third scenario: the variances between the coordinates for a given frame are not constant. In that case, the coordinates in the frame XY are regarded as uncorrelated observations, but with different variances, whereas the coordinates in the frame xy are considered error-free values. This means that only the coordinates in the frame XY are subjected to random error, so that  $\mathbf{e}_y \sim N(\boldsymbol{\mu}_y = \mathbf{0}, \boldsymbol{\Sigma}_y)$ . This is a typical case of a free-total station problem, where total-station must be georeferenced, i.e., the position and orientation of the total station in the required reference system must be established before its use for detail surveying or staking out. The coordinates established prior serve as control points for the establishment of the total station. Generally, GNSS (*Global Navigation Satellite Systems*) observations, such as carrier phase and code pseudorange, are processed to estimate the coordinates of the control points, which normally lead to different variances of the coordinates of control points.
4. Fourth scenario: This case is similar to the third situation described above, i.e.,  $\mathbf{e}_y \sim N(\boldsymbol{\mu}_y = \mathbf{0}, \boldsymbol{\Sigma}_y)$  and  $\mathbf{E}_A \sim N(\boldsymbol{\mu}_A = \mathbf{0}, \boldsymbol{\Sigma}_A)$ . However, now not only are the variances of coordinates in the frame XY different, but also the variances of coordinates in xy frame. On the same example of free-station in the third situation mentioned above, now not only the coordinates of control points are subjected to random errors, but also the coordinates of the total-station. The positional uncertainties of the coordinates computed by total station are different for each surveyed point because they depend on the distance and the bearing. Therefore, the

variances of the coordinates in the frame of the total station will also be different from each other, i.e., the block-diagonal elements (variances) of the covariance matrix will not be equal.

The results confirm that there are no significant differences between the LS-MC and the conventional WLSE-GM in the estimation of the transformation parameters for all scenarios studied. This result is expected because both the LS-MC and WLSE-GM are based on the same objective function given by the expression (8). Moreover, the random noise matrix  $E_A$ , associated with the coordinates in the frame  $xy$ , follows a multivariate normal distribution with zeros means  $\mu_A = \mathbf{0}$ . In that case, the expectation of the matrix  $A^*$  given in Eq. (16) is equal to the matrix  $A$  given in Eq. (6), i.e.:

$$E\{A^*\} = E\{A\} - E\{\Sigma_A\} = A \tag{23}$$

where  $E\{\cdot\}$  is the expectation operator.

On the other hand, the uncertainties of the transformation parameters computed by the LS-MC are different from those obtained by the WLSE-GM. According to the results displayed in Figure 4 it is possible to note that the standard uncertainties associated with the transformation parameters are higher for the LS-MC than the WLSE-GM solution. This is due to the fact that the WLSE-GM does not take into account the uncertainties associated with the coordinates in the frame  $xy$ . In other words, the elements of the matrix  $A^*$ , which are composed by the coordinates in the  $xy$  reference frame, are treated as error-free quantities in the WLSE-GM solution, whereas in the LS-MC the uncertainties associated with the elements of the  $A^*$  matrix are propagated through model to the uncertainties of the transformation parameters.

In terms of uncertainty, the highest difference between the LS-MC and the WLSE-GM occurred in the fourth scenario: approximately 21 mm (~63%) for shifting  $T_x$ , 18 mm (~55%) for shifting  $T_y$ , 6 arc-seconds (~71%) for rotation  $\alpha$ , 48 ppm (~39%) for scale  $s$ , 49 ppm (~40%) for synthetic parameter  $a$ , and 130 ppm (~71%) for synthetic parameter  $b$ .

Moreover, there are no significant differences between the standard deviations estimated by the WLSE-GM and LS-MC in the third scenario, as can be seen in Figure 4, since the errors in the  $xy$  system are not considered. In that case, the first-order terms of Taylor series expansion used in the linearization of the model for the propagation of the uncertainties of the synthetic parameters to the physical parameters were consistent with the LS-MC. Although we have not applied it here, a least-squares solution in the sense of nonlinear Gauss-Helmert models can also be an alternative for the case where the coordinates in both frames contain random errors [9].

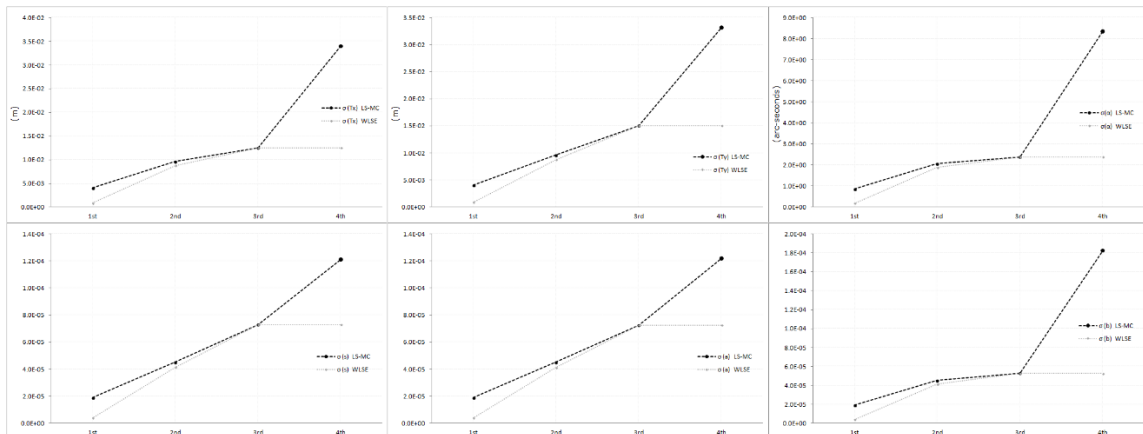


Figure 4: Standard-deviation of the transformation parameters computed by the WLSE-GM for the four scenarios considered.

A coverage interval for the transformation parameters was also determined according to the step (8). From the resulting cumulative distribution of the transformation parameters displayed in Fig. 5, we extracted the 0.025 and 0.975 quantiles as a good approximation to the coverage interval

$[X_{low}, X_{high}]$  for  $\alpha = 0.05$ . The markers represent the 0.025, 0.5, and 0.975 percentiles for each scenario.

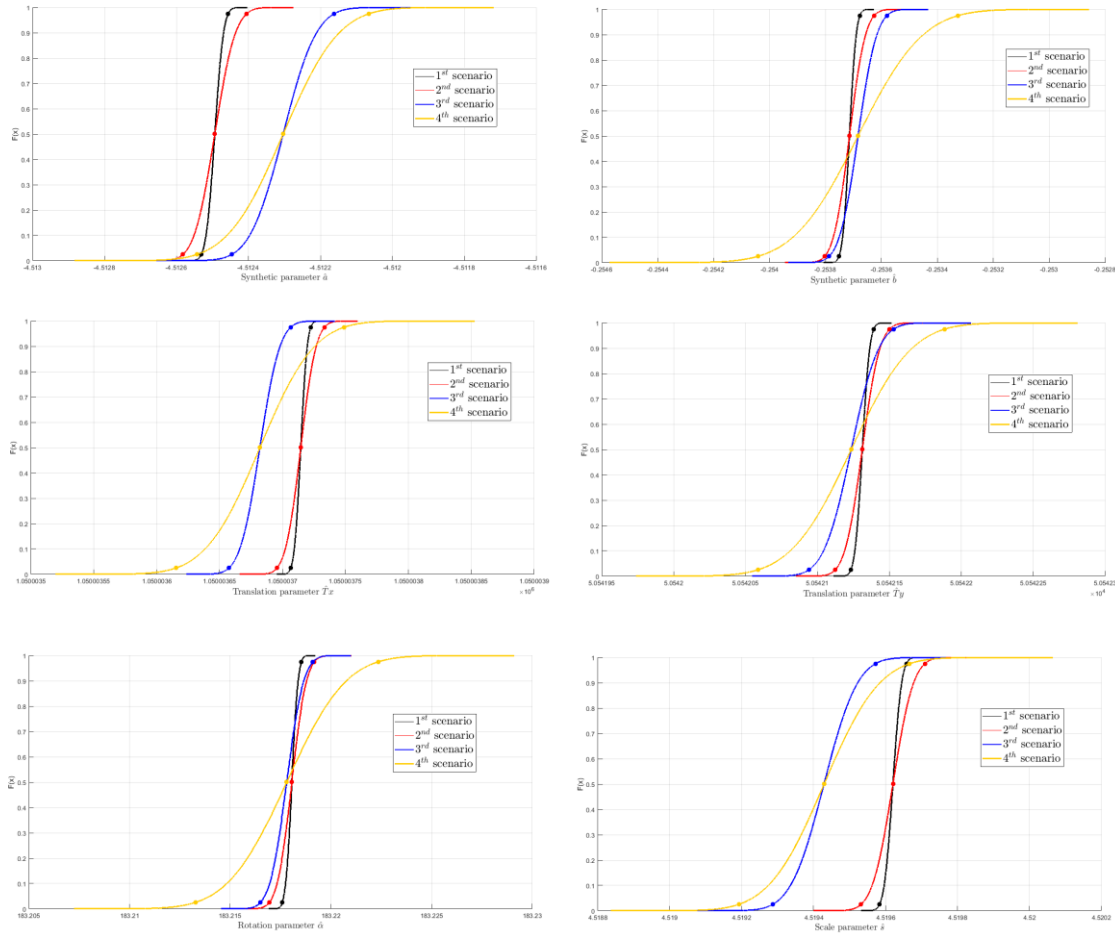


Figure 5: Cumulative distribution of the transformation parameters and 0.025, 0.5, and 0.975 percentiles for each scenario

To construct the confidence intervals for the solutions derived from the WLSE-GM, it was necessary to find the theoretical critical value from a known PDF. As the observations set has less than 30 values, the critical value was calculated based on a *t-student distribution* as follows (Ghilani, 2010, pp.78) [3]:

$$\hat{\mathbf{X}} \pm |t_{\alpha/2,df}| \left( \sqrt{D\{\hat{\mathbf{X}}\}} \right) \tag{24}$$

where  $\hat{\mathbf{X}}$  and  $D\{\hat{\mathbf{X}}\}$  are the estimated transformation parameters and its associated covariance matrix, respectively. They are given by Eq.(9) and Eq.(10), respectively. The  $t_{\alpha/2}$  is the theoretical critical value from the *t*-distribution based on *df* degrees of freedom of the adjustment and  $\alpha/2$  percentage points. The modulus operator  $|\cdot|$  is to avoid negative values. The *t* value that locates an  $\alpha/2$  area in both the upper and lower tails of the distribution can be found in the statistical tables or computed by popular scientific software. In our case, we use the function “*inv*” from the MATLAB. Since  $df = n - u = 2$  and  $\alpha = 0.05$ , we obtain  $|t_{0.025,2}| = 4.3$ . This leaves 0.025 in each of the upper- and lower-tail areas, i.e., 0.975 and 0.025 quantiles, respectively.

The most typical error distributions of real observations seem to be leptokurtic as is the case of the Laplace distribution [29, 35, 36, 37]. Because of that, Laplace observation error distributions were also generated here for both  $\mathbf{e}_y \sim L(\boldsymbol{\mu}_y = \mathbf{0}, \boldsymbol{\Sigma}_y)$  and  $\mathbf{E}_A \sim L(\boldsymbol{\mu}_A = \mathbf{0}, \boldsymbol{\Sigma}_A)$ . Laplace distribution is used whenever normality is rejected (Lehmann, 2015). In the standard Laplace case,

we use uniformly distributed PRN generated by MATLAB PRN generator “unidrnd” and apply a transformation by the inverse CDF (Tanizaki, 2004, pp. 122) [34].

Table 3 displays the width of confidence intervals for the WLSE-GM and LS-MC based on normal and Laplace distributions. These widths were computed by taking the largest value  $X_{high}$  minus the smallest value  $X_{low}$  in an interval for  $\alpha = 0.05$ . It is important to mention that the construction of the coverage interval by the LS-MC is totally empirical, based on the distribution of the observations, as can be seen in the step (8) from the previous section about the LS-MC algorithm, whereas the confidence interval is based on the theoretical inference of the parameters defined in Eq. (24).

Tabela 3: Width of coverage intervals of the transformation parameters for the WLSE-GM and LS-MC.

2D Transformation parameters							
Scenarios	Method	$a$ [ppm]	$b$ [ppm]	$T_x$ [cm]	$T_y$ [cm]	$\alpha$ ["]	$s$ [ppm]
First scenario	$WLSE_{X_{high}-X_{low}}$	35.475	35.475	0.755	0.755	1.619	35.475
	$LS-MC_{X_{high}-X_{low}}^{normal}$	74.797	74.711	1.59	1.59	3.409	74.821
	$LS-MC_{X_{high}-X_{low}}^{Laplace}$	80.188	80.000	1.710	1.706	3.652	80.143
Second scenario	$WLSE_{X_{high}-X_{low}}$	354.751	354.751	7.545	7.545	16.19	354.751
	$LS-MC_{X_{high}-X_{low}}^{normal}$	177.626	177.258	3.773	3.773	8.09	177.623
	$LS-MC_{X_{high}-X_{low}}^{Laplace}$	187.783	188.178	3.988	3.997	8.589	187.816
Third scenario	$WLSE_{X_{high}-X_{low}}$	625.018	453.981	10.794	12.943	20.587	627.124
	$LS-MC_{X_{high}-X_{low}}^{normal}$	285.013	206.895	4.912	5.895	9.380	285.829
	$LS-MC_{X_{high}-X_{low}}^{Laplace}$	307.898	223.707	5.318	6.387	10.153	309.024
Fourth scenario	$WLSE_{X_{high}-X_{low}}$	625.018	453.981	10.794	12.943	20.587	627.124
	$LS-MC_{X_{high}-X_{low}}^{normal}$	476.861	714.544	13.322	12.957	32.712	474.226
	$LS-MC_{X_{high}-X_{low}}^{Laplace}$	502.018	763.301	14.091	13.730	34.915	498.606

It was expected that the confidence intervals for the WLSE-GM would be larger than the coverage interval for the LS-MC for all scenarios because the  $t$ -student distribution used to construct the confidence intervals for the WLSE-GM is less conservative than normal and Laplace distribution. In this case, conservative means that a distribution provides a larger range than others. However, this did not happen in the first scenario, where the intervals for the WLSE-GM were smaller than those for the LS-MC. This is due to the fact that the effect of the uncertainties associated with the coordinates in the  $xy$  frame was higher than the selection of the distribution of the observation errors on the width of the intervals. For this case, in general, the intervals were nearly two times larger for the LS-MC than the WLSE-GM. Regarding the other scenarios (second, third, and fourth), the confidence intervals for the WLSE-GM were larger than the coverage intervals for the LS-MC.

Note that there is a slight difference between the normal and Laplace distributions in the widths of the intervals. In general, the lengths of the intervals derived from the LS-MC based on the Laplace distribution is larger than the one based on normal distribution. The reason is that we generated Laplacian observation errors and the solution for transformation parameters were based on the L2 norm minimization (given by Eq. 8). In this case, the solution given in the L2 norm minimization as a “best linear unbiased estimation” (BLUE) is expected to provide the least root mean squares values for the estimated parameters, independent of the error distribution. However, compared with the “linear” estimation, a non-linear estimation, like the L1 norm minimization, could perform better [29].

Moreover, the coverage interval in the case of the LS-MC was higher in the fourth scenario for both the normal and Laplace observation error models with approximately 14 cm for  $T_x$ , 13 cm for  $T_y$ , 34 arc-seconds for rotation  $\alpha$ , 486 ppm for scale  $s$ , 489 ppm for synthetic parameter  $a$ , and 739 ppm for synthetic parameter  $b$ . As the WLSE-GM considers the coordinates from  $xy$  frame error-free quantities, coverage intervals in the third and fourth scenarios were the same for this method. Note that both third and fourth scenarios are based on the same uncertainties for the observation vectors.

## 6. CONCLUSIONS AND OUTLOOK

It has been shown that a Monte Carlo Method (MCM) can be used in combination with least-squares to solve a coordinate transformation problem in the context of the Gauss-Markov Model (GMM). The method was called LS-MC. The main highlights of the LS-MC in the context of the transformation parameters are:

1. Contrary to the purely standard GMM, the uncertainties associated with the coordinates in both frames are propagated through the model to the uncertainties of the transformation parameters. This can also be accomplished by the least-squares of a rigorous Gauss-Helmert model (GH) and/or also by Total Least Squares (TLS).
2. In contrast to solutions based on the GMM and GH models, the uncertainties of the estimated physical parameters are obtained by means of the discrete representation of their distributions.
3. Since powerful computers are available everywhere, it is no longer necessary to resort to statistical tables in order to construct the coverage interval. Therefore, the coverage interval can be computed numerically.
4. Although we restricted ourselves to the Laplace and normal distributions for the observation errors, in some applications it is also possible to assign a non-Gaussian distribution to the input data (e.g., asymmetric, uniform, triangle, among others).

However, there are also some aspects of the method that deserve to be studied in the future, for example:

1. The number of Monte Carlo trials: as the computations are stochastic, being based on random draws, there is no guarantee that any specific pre-assigned MCM trial number suffices. Therefore, a procedure that selects the MCM trials adaptively, i.e., as the trials progress, will be studied. However, this problem should not be important. Although we have set the number of MCM trials to  $10^6$ , the results presented here did not change significantly when the computations were repeated with different pseudo-random numbers.
2. The coverage interval is formed by considering each parameter individually, without taking into account the correlation between the variables. In some geodetic and geophysical applications, there are multivariate models, namely models with more than one output quantity. Such quantities are generally mutually correlated because they depend on common input quantities. In this case, it is more complicated to form the discrete representation of the probability distribution because it would be necessary to deal with multidimensional data ordering [38]. This will be investigated in the next study.
3. The proposed method is based on Least-Squares Estimation within the Gauss-Markov Model. Therefore, it lacks the robustness or insensitivity to outliers in the observations. In the future, quality control procedures for detection and identification of outliers, such as data snooping, should be integrated with the LS-MC.
4. Although the idea of the Monte Carlo simulation to infer the uncertainty of estimators is correct, in the future works the Monte Carlo method should be applied to the nonlinear least-squares solution.

## 7. ACKNOWLEDGMENTS

We appreciate the comments and suggestions of the anonymous reviewers. We also appreciate

the support of all members of the research group “Quality Control in Geodesy” (<http://dgp.cnpq.br/dgp/espelhogruppo/3674873915161650>).

## 8. REFERENCES

1. Zhang P, Xu C, Hu C, Chen Y. Coordinate transformations in satellite navigation systems. In: *Lecture Notes in Electrical Engineering*. doi:10.1007/978-3-642-27296-7\_40. 2012.
2. Zhao H, Zhang B, Wu C, Zuo Z, Chen Z. Development of a Coordinate Transformation method for direct georeferencing in map projection frames. *ISPRS J Photogramm Remote Sens.* 2013;77:94-103, doi:10.1016/j.isprsjprs.2012.12.004.
3. Ghilani CD. *Adjustment Computations: Spatial Data Analysis*. New Jersey: John Wiley & Sons; 2010. 699 p.
4. Greenfeld JS. Least Squares Weighted Coordinate Transformation Formulas and Their Applications. *J Surv Eng.* 1997; 123:147-161, doi:10.1061/(asce)0733-9453(1997)123:4(147).
5. Janicka J, Rapinski J. Msplit transformation of coordinates. *Surv Rev.* 2013;45:269-274, doi:10.1179/003962613X13726661625708.
6. Lehmann R. Transformation model selection by multiple hypotheses testing. *J Geod.* 2014;88:1117-1130, doi:10.1007/s00190-014-0747-3.
7. Konakoglu B, Cakır L, Gökalp E. 2D coordinate transformation using Artificial Neural Networks. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 3rd International GeoAdvances Workshop; 2016; Istanbul, Turkey: p. 183-186, doi:10.5194/isprs-archives-XLII-2-W1-183-2016.*
8. Civicioglu P. Transforming geocentric cartesian coordinates to geodetic coordinates by using differential search algorithm. *Comput. Geosci-UK.* 2012;46:229-247, doi:10.1016/j.cageo.2011.12.011.
9. Neitzel F. Generalization of total least-squares on example of unweighted and weighted 2D similarity transformation. *J Geod.* 2010;84:751-762, doi:10.1007/s00190-010-0408-0.
10. Schaffrin B, Felus Y.A. On the multivariate total least-squares approach to empirical coordinate transformations. Three algorithms. *J Geod.* 2008;82:373-383, doi:10.1007/s00190-007-0186-5.
11. Schaffrin B, Wieser A. On weighted total least-squares adjustment for linear regression. *J Geod.* 2008; 82:415-421, doi:10.1007/s00190-007-0190-9.
12. Amiri-Simkooei AR, Jazaeri S. Weighted total least squares formulated by standard least squares theory. *J Geod Sci.* 2012;2:113-124, doi:10.2478/v10156-011-0036-5.
13. Amiri-Simkooei AR, Jazaeri, S. Data-snooping procedure applied to errors-in-variables models. *Stud Geophys Geod.* 2013;57:426-441, doi:10.1007/s11200-012-0474-2.
14. Fang X. Weighted total least-squares with constraints: a universal formula for geodetic symmetrical transformations. *J Geod.* 2015;59:459-469, doi:10.1007/s00190-015-0790-8.
15. Hu C, Chen Y, Zhu WD. Generalised total least squares solution based on pseudo-observation method. *Surv Rev.* 2016;123:157-167, doi:10.1179/1752270614y.0000000155.
16. Mahboub V. On weighted total least-squares for geodetic transformations. *J Geod.* 2012;86:359-367, doi:10.1007/s00190-011-0524-5.
17. Mahboub V, Sharifi MA. On weighted total least-squares with linear and quadratic constraints. *J Geod.* 2013;87:279-286, doi:10.1007/s00190-012-0598-8.
18. Mihajlović D, Cvijetinić Ž. Weighted coordinate transformation formulated by standard least-squares theory. *Surv Rev.* 2017;49:328-345, doi:10.1080/00396265.2016.1173329.
19. Wyszowski P. Propagation of uncertainty by Monte Carlo simulations in case of basic geodetic computations. *Geocart.* 2017;66(2):333-346, doi: 10.1515/geocart-2017-0022.
20. Cox DR, Hinkley DV. *Theoretical Statistics*. New York: Chapman & Hall/CRC; 1974.
21. Teunissen PJG. *Testing Theory: An Introduction*. Series on Mathematical Geodesy and Positioning. Delft: VSSD Press; 2006.
22. Heslop D. A Monte Carlo investigation of the representation of thermally activated single-domain particles within the Day plot. *Stud Geophys Geod.* 2005;49:163-176.
23. Koch, KR. Determining the maximum degree of harmonic coefficients in geopotential models by Monte Carlo methods. *Stud Geophys Geod.* 2005;49:259-275.
24. Koch, KR. Expectation Maximization algorithm and its minimal detectable outliers. *Stud Geophys Geod.* 2017;61:1-18.
25. Erdogan B, Hekimoglu S, Durdag UM, Ocalan T. Empirical estimation of the power of test in outlier detection problem. *Stud Geophys Geod.* 2019;63:55-70.
26. Altioik T, Melamed B. *Simulation Modelling and Analysis with ARENA*. Burlington: Academic Press; ISBN: 978-0-12-370523-5. 2007.



27. Matsumoto M, Nishimura T. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Trans Model Comput Simul.* 1998;8:3-30.
28. Box GEP, Muller ME. A Note on the Generation of Random Normal Deviates. *Ann Math Stat.* 1958; doi:10.1214/aoms/1177706645.
29. Lehmann R. Observation error model selection by information criteria vs. normality testing. *Stud Geophys Geod.* 2015;59:489-504, doi:10.1007/s11200-015-0725-0.
30. Rofatto VF, Matsuoka MT, Klein I, Veronez MR, Bonimani ML, Lehmann R. A half-century of Baarda's concept of reliability: a review, new perspectives, and applications. *Surv Rev.* 2018; doi:10.1080/00396265.2018.1548118
31. JCGM 102:2011. Evaluation of measurement data – Supplement 2 to the “Guide to the expression of uncertainty in measurement” – Extension to any number of output quantities, JCGM, (online at <https://www.bipm.org/en/publications/guides/gum.html>); 2011.
32. Lehmann R, Scheffler T. Monte Carlo-based data snooping with application to a geodetic network. *J Appl Geod.* 2011;5:123-134, doi:10.1515/JAG.2011.014.
33. Cox MG, Siebert BRL. The use of a Monte Carlo method for evaluating uncertainty and expanded uncertainty. *Metrologia.* 2006;43:178-188.
34. Tanizaki H. Computational methods in statistics and econometrics. New York: Marcel Dekker; 2004.
35. Hampel FR. Robust statistics: A brief introduction and overview. In: Carosio A, Kutterer H, editors. *First International Symposium on Robust Statistics and Fuzzy Techniques in Geodesy and GIS*; 2001; Institute of Geodesy and Photogrametry, ETH Zurich, Zurich, Switzerland.
36. Huber PJ. *Robust Statistics*. New York: John Wiley & Sons Inc; 2009. ISBN 978-0-470-12990-6.
37. Wisniewski Z. M-estimation with probabilistic models of geodetic observations. *J Geod.* 2014;88:941-957, doi: 10.1007/s00190-014-0735-7.
38. Barnett V. The Ordering of Multivariate Data. *J Royal Stat Soc.* 1976;139:318–355, doi:10.2307/2344839.