
An Overview of Adjustment Methods for Mixed Additive and Multiplicative Random Error Models

Yun Shi, Peiliang Xu, and Junhuan Peng

Abstract

Geodetic adjustment theory has been developed on the basis of a linear or nonlinear Gauss-Markov model, in which the random errors of measurements are always assumed to be independent of the true values of measurements themselves and naturally added to the functional model. However, modern geodetic instruments and geodetic imaging systems have clearly shown that the random errors of such measurements consist of two parts: one is of local nature and has nothing to do with the quantity under observation, and the other is proportional to the true value of measurement. From the statistical point of view, these two types of errors are called *additive* and *multiplicative* errors, respectively. Obviously, the conventional geodetic adjustment theory and methods for Gauss-Markov models with additive errors cannot theoretically meet the need of processing measurements contaminated by mixed additive and multiplicative random errors. This paper presents an overview of parameter estimation methods for processing mixed additive and multiplicative random errors. More specifically, we discuss two types of methods to estimate parameters in a mixed additive and multiplicative error model, namely, quasi-likelihood and least-squares-based methods. From this point of view, we extend the conventional adjustment theory and methods and give a solid theoretical foundation to process geodetic measurements contaminated by mixed additive and multiplicative random errors. Finally, we further discuss parameter estimation with prior information.

Keywords

Generalized estimating equation • Least squares • Multiplicative and additive errors • Quasi-likelihood

Y. Shi (✉)

School of Geomatics, Xi'an University of Science and Technology,
Xi'an 710054, PR China

e-mail: shiyun0908@hotmail.com

P. Xu

Disaster Prevention Research Institute, Kyoto University, Uji, Kyoto
611-0011, Japan

e-mail: pxu@rcep.dpri.kyoto-u.ac.jp

J. Peng

School of Land Science and Geomatics, China University of
Geosciences, Beijing, PR China

1 Introduction

Geodetic adjustment theory has been developed by assuming the following model of measurements:

$$\left. \begin{aligned} \mathbf{y} &= \mathbf{f}(\boldsymbol{\beta}) + \boldsymbol{\epsilon} \\ E(\mathbf{y}) &= \mathbf{f}(\boldsymbol{\beta}) \\ D(\mathbf{y}) &= \mathbf{W}^{-1}\sigma^2 \end{aligned} \right\}, \quad (1)$$

where \mathbf{y} is a vector of measurements, $\mathbf{f}(\boldsymbol{\beta})$ is the mathematical or functional model which describes the physical or geometrical relationships between the measurements, $\boldsymbol{\beta}$

is the real-valued vector of unknown parameters to be estimated, ϵ is the random error vector of the measurements. Very often, we also assume that ϵ is of zero mean and variance-covariance matrix $\mathbf{W}^{-1}\sigma^2$, with \mathbf{W} being a given weight matrix of measurements and σ^2 an unknown positive scalar (the variance of unit weight), $E(\cdot)$ and $D(\cdot)$ stand for the expectation and variance-covariance matrix of the measurements, respectively. The most important feature of adjustment model (1) is that the random errors ϵ are added to the functional model $\mathbf{f}(\beta)$. In other words, the sizes or magnitudes of random errors are independent of the true values of measured quantities.

However, in geodetic practice, we know that this assumption is not necessarily always true. For example, we know that the accuracy of an EDM, VLBI and/or GPS baseline is proportional to the length of the baseline itself, namely,

$$\sigma_L^2 = a^2 + b^2 L^2, \quad (2)$$

(see e.g., Ewing and Mitchell 1970; MacDoran 1979; Seeber 2003; Petrov et al. 2010), where both a and b are constants. Physically, the constant a may be more specific to the local environment of stations and b more to the path of propagation of light/electronic waves (see e.g., Xu et al. 2013). From the statistical point of view, the accuracy formula (2) is equivalent to the following representation of random errors:

$$\epsilon_L = \epsilon_a + L \epsilon_b, \quad (3)$$

where ϵ_L is the random error of L , and ϵ_a and ϵ_b stand for the random errors of mean zero and variances a^2 and b^2 , respectively, if ϵ_a and ϵ_b are assumed to be statistically independent. The error representation (3) clearly indicates that the random error ϵ_L is proportional to the measured baseline. In geodetic practice, both ϵ_a and ϵ_b are generally assumed to be normally distributed. For other modern space observation technology such as SLR (see e.g. Pearlman et al. 2002; Seeber 2003) and DORIS (see e.g. Willis et al. 2010), since they essentially utilize electromagnetic waves for observation and go through the same physical media as VLBI and GPS, we conjecture that errors of SLR and DORIS baselines should also show multiplicative error behavior, which will be a topic of research in the future.

Modern geodetic technology also fully utilizes coherent imaging systems such as Synthetic Aperture Radar (SAR) images and Light Detection And Ranging (LiDAR). As is well known, SAR images are contaminated by speckle noise (see e.g. Goodman 1976; Ulaby et al. 1986; Oliver 1991; López-Martínez et al. 2011) and the corresponding observational equation can be represented as follows:

$$y_{ij} = s_{ij}(1 + \epsilon_{ij}), \quad (4)$$

where y_{ij} is the measurement, s_{ij} the true (or noiseless) value of the signal and ϵ_{ij} the random error with zero mean and variance σ^2 . Intensity measurements of SAR type are usually assumed to have a gamma-distribution. Other imaging systems would also produce Gaussian multiplicative random errors (see e.g., Tian et al. 2001). Range measurements of LiDAR are also shown to be contaminated by multiplicative speckle errors (see e.g., Flamant et al. 1984; Wang and Pruitt 1992; Hill et al. 2003).

The paper is organized as follows. Section 2 will first define mixed additive and multiplicative error models. In Sects. 3 and 4, we will discuss two important classes of methods for parameter estimation in mixed additive and multiplicative error models, namely, quasi-likelihood and least squares (LS). Computational algorithms will be briefly given. If the reader is interested in other methods such as cumulant moment and variational methods, he may refer to, Swami (1994) and Aubert and Aujol (2008), for example. Actually, variational methods assume gamma-distributions for intensity measurements and then add an extra smoothness or regularized term to the log-likelihood of the gamma-distributions for de-speckling or de-noising multiplicative random errors, as can be seen in Xu (1999) and Aubert and Aujol (2008). We will then extend the bias-corrected LS method to the case with prior information in Sect. 5. Finally, we will then finish our paper with some concluding remarks in Sect. 6.

2 Mixed Additive and Multiplicative Error Models

We will now extend the conventional Gauss-Markov adjustment model (1) to account for both additive and multiplicative errors. The new starting model of adjustment becomes

$$\left. \begin{aligned} \mathbf{y} &= \mathbf{f}(\beta) \odot (\mathbf{1} + \epsilon_m) + \epsilon_a \\ E(\mathbf{y}) &= \mathbf{f}(\beta) \\ E(\epsilon_m) &= \mathbf{0}, \quad D(\epsilon_m) = \Sigma_m \\ E(\epsilon_a) &= \mathbf{0}, \quad D(\epsilon_a) = \Sigma_a \end{aligned} \right\}, \quad (5)$$

where \mathbf{y} and $\mathbf{f}(\beta)$ have been defined in (1), \odot stands for the Hadamard product of matrices and/or vectors, $\mathbf{1}$ for the vector with all its elements being equal to unity, both of the random errors ϵ_m and ϵ_a are of mean zero and variance-covariance matrices Σ_m and Σ_a , respectively. Since the random vector ϵ_m is multiplied to the true value of measurements $\mathbf{f}(\beta)$, ϵ_m has been naturally called *multiplicative errors*. As in the case of the additive error model (1), ϵ_a in (5) is additive. Accordingly, ϵ_a is called *additive errors*. To illustrate additive and multiplicative random errors, we simulate the random errors of baselines, with the constants a and b set to 0.05 m and 10 ppm, respectively. The generated

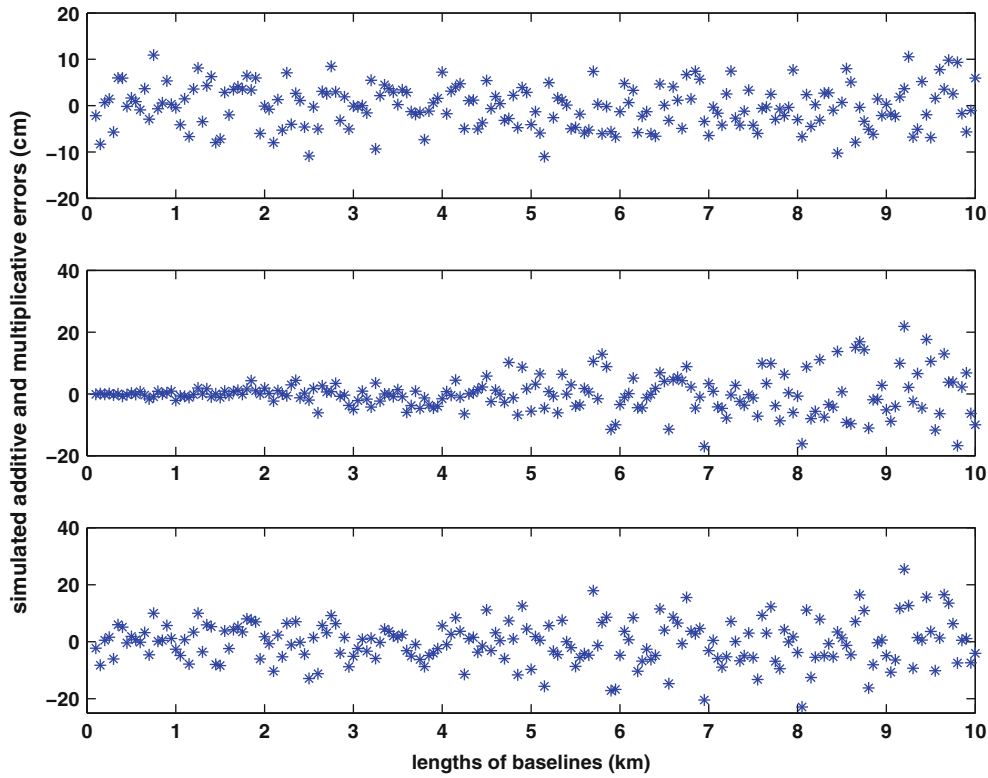


Fig. 1 Illustration of the additive and multiplicative random errors of baselines: the *upper panel* – the additive errors; the *middle panel* – the multiplicative errors; and the *lower panel* – the mixed additive and multiplicative errors

errors with the lengths of baselines are illustrated in Fig. 1. It is obvious from the simulated random errors in the upper panel of Fig. 1 that the additive random errors uniformly scatter over different lengths of baselines. The multiplicative errors in the middle panel of the same figure show a clear trend of fan shape, with the amplitudes of errors increasing with the increase of lengths of baselines.

Assuming that ϵ_m and ϵ_a are statistically independent and applying the error propagation law to each of the measurements \mathbf{y} , we have

$$\sigma_{y_i}^2 = f_i^2(\boldsymbol{\beta})\sigma_{m_i}^2 + \sigma_{a_i}^2, \quad (6)$$

where $\sigma_{y_i}^2$ is the variance of the i th measurement of \mathbf{y} , and $\sigma_{m_i}^2$ and $\sigma_{a_i}^2$ are the i th diagonal elements of $\boldsymbol{\Sigma}_m$ and $\boldsymbol{\Sigma}_a$, respectively. It is obvious from (6) that the larger the true value of measurement $f_i(\boldsymbol{\beta})$, the noisier the corresponding measurement y_i . When applying the same error propagation law to the measurement vector \mathbf{y} , we can obtain the variance-covariance matrix of the measurements \mathbf{y} as follows:

$$\boldsymbol{\Sigma}_y(\boldsymbol{\beta}) = \mathbf{D}_{f\boldsymbol{\beta}}\boldsymbol{\Sigma}_m\mathbf{D}_{f\boldsymbol{\beta}} + \boldsymbol{\Sigma}_a, \quad (7)$$

where $\mathbf{D}_{f\boldsymbol{\beta}}$ is a diagonal matrix with its i th diagonal element being equal to $f_i(\boldsymbol{\beta})$. The elements of $\boldsymbol{\Sigma}_y(\boldsymbol{\beta})$ are obviously the functions of the parameters $\boldsymbol{\beta}$. For simplicity, we will use

$\boldsymbol{\Sigma}_y$ to denote the variance-covariance of \mathbf{y} . If necessary, we can also readily take the correlation between ϵ_m and ϵ_a into account, which will not be discussed in this paper, however.

If $\mathbf{f}(\boldsymbol{\beta})$ is linear, then the error model (5) becomes

$$\left. \begin{aligned} \mathbf{y} &= (\mathbf{A}\boldsymbol{\beta}) \odot (\mathbf{1} + \boldsymbol{\epsilon}_m) + \boldsymbol{\epsilon}_a \\ E(\mathbf{y}) &= \mathbf{A}\boldsymbol{\beta} \\ E(\boldsymbol{\epsilon}_m) &= \mathbf{0}, \quad D(\boldsymbol{\epsilon}_m) = \boldsymbol{\Sigma}_m \\ E(\boldsymbol{\epsilon}_a) &= \mathbf{0}, \quad D(\boldsymbol{\epsilon}_a) = \boldsymbol{\Sigma}_a \end{aligned} \right\}, \quad (8)$$

where \mathbf{A} is a given design matrix, which will be assumed to be of full rank. The model (5) will be called mixed additive and multiplicative error models in the remainder of this paper. In accordance with (8), we can rewrite $\mathbf{D}_{f\boldsymbol{\beta}}$ as $\mathbf{D}_{a\boldsymbol{\beta}}$, whose diagonal elements are equal to $\mathbf{a}_i\boldsymbol{\beta}$, where \mathbf{a}_i is the i th row of the matrix \mathbf{A} .

3 The Quasi-Likelihood Method

The quasi-likelihood method was first proposed by Wedderburn (1974). It has since become a statistical method to estimate the parameters in the model of type (4) and widely applied in many areas of science and engineering. Actually, the multiplicative error model (4) has been better known in statistics as the generalized linear model and

well documented in statistical books (see e.g. McCullagh and Nelder 1989; Heyde 1997, chapter 5.3), if the function of signal s_{ij} can be represented linearly by a number of unknown parameters β .

Wedderburn (1974) started with a set of independent measurements y_i ($i = 1, 2, \dots, n$), with expectations \bar{y}_i and variances $\sigma_i^2(\bar{y}_i)$, and then defined the quasi-likelihood function $\text{QLF}(y_i, \bar{y}_i)$ as follows:

$$\frac{\partial \text{QLF}(y_i, \bar{y}_i)}{\partial \bar{y}_i} = \frac{y_i - \bar{y}_i}{\sigma_i^2(\bar{y}_i)}, \quad (9)$$

where the variance of each y_i is assumed to be the function of \bar{y}_i . By letting the expression (9) equal zero over all the measurements y_i , Wedderburn (1974) was then able to estimate the unknown parameters from the measurements \mathbf{y} . If \bar{y}_i can further be represented linearly by a number of unknown parameters β and if the measurements \mathbf{y} are assumed to be correlated, then (9) can be rewritten as follows:

$$\frac{\partial \text{QLF}(\beta)}{\partial \beta} = \mathbf{A}^T \Sigma_y^{-1}(\beta)(\mathbf{y} - \mathbf{A}\beta), \quad (10)$$

where $\Sigma_y(\beta)$ is the variance-covariance matrix of \mathbf{y} whose elements are all the functions of the unknown parameters β . The quasi-likelihood function is proved to be equal to the maximum likelihood function, if the distribution of y_i is exponential. In general, quasi-likelihood is different from maximum likelihood, however.

Although Wedderburn (1974) defined the quasi-likelihood function $\text{QLF}(y_i, \bar{y}_i)$ through the differential equation (9), $\text{QLF}(y_i, \bar{y}_i)$ is really not required for parameter estimation. Actually, all what we need for parameter estimation is the expression on the right hand side of (9), which is completely defined by y_i , its expectation \bar{y}_i and its variance $\sigma_i^2(\bar{y}_i)$. When the quasi-likelihood method is applied to the mixed additive and multiplicative error model (8), we have the system of normal equations:

$$\mathbf{A}^T \Sigma_y^{-1}(\hat{\beta}_{ql})(\mathbf{y} - \mathbf{A}\hat{\beta}_{ql}) = \mathbf{0}, \quad (11)$$

where $\hat{\beta}_{ql}$ stands for the quasi-likelihood estimate of β . Obviously, the system of normal equations (11) is nonlinear and can generally be solved by using numerical methods. Very often, one can use the Gauss-Newton method to find the solution to (11). The quasi-likelihood estimator $\hat{\beta}_{ql}$ is asymptotically unbiased (see e.g., McCullagh 1983) and its variance-covariance matrix, denoted by $D(\hat{\beta}_{ql})$, is then given approximately by

$$D(\hat{\beta}_{ql}) = (\mathbf{A}^T \Sigma_y^{-1} \mathbf{A})^{-1}. \quad (12)$$

It is seen from (11) that the equation system (11) has completely defined the quasi-likelihood estimator $\hat{\beta}_{ql}$, no matter whether we can or cannot solve for the quasi-likelihood function $\text{QLF}(\mathbf{y}, \beta)$ through the differential equation of type (9). The system of normal equations clearly indicates that an estimator can simply be constructed through a system of equations. As a result of this, the system of equations like (11) has been called *generalized estimating equations* (see e.g., Crowder 1995; Desmond 1997; Heyde 1997; Kukusha et al. 2010; Fitzmauric 1995).

4 Least-Squares-Based Methods

Although quasi-likelihood has become a standard method for parameter estimation in multiplicative error models, its associated quasi-likelihood function may hardly be derived for a general nonlinear function $\mathbf{f}(\beta)$. Even if such a function can indeed be found by solving the corresponding differential equations, it may not be connected with any physically meaningful distribution function. As a result, Xu and Shimada (2000) alternatively proposed LS-based methods to estimate the unknown parameters in the mixed additive and multiplicative error model (8). In this section, we will briefly discuss the ordinary LS, the weighted LS and bias-corrected weighted LS methods. If the reader is interested in the error analysis of adjusted measurements and the corrections of measurements and/or the estimation of the variance of unit weight in multiplicative error models, she/he is referred to Shi et al. (2014).

4.1 The Ordinary LS Method

When applying the ordinary LS method to estimate the unknown parameters β in the model (8), we have the following optimization objective function:

$$\min : F_1(\beta) = (\mathbf{y} - \mathbf{A}\beta)^T (\mathbf{y} - \mathbf{A}\beta). \quad (13)$$

The solution to (13) is the ordinary LS estimate of β , which is denoted by $\hat{\beta}_{LS}$ and given by

$$\hat{\beta}_{LS} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{y}. \quad (14)$$

The variance-covariance matrix of $\hat{\beta}_{LS}$ is then given as follows:

$$D(\hat{\beta}_{LS}) = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \Sigma_y \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-1}, \quad (15)$$

where Σ_y is the variance-covariance matrix of the measurements \mathbf{y} .

If we assume that the signal within a small area of pixels in a coherent image with multiplicative noises is identical, namely, s_{ij} remains unchanged in such a small area, then all the corresponding measurements y_{ij} are of the same variances. In other words, the weights of measurements y_{ij} are all equal to each other. As a result, the estimate of s_{ij} is simply equal to the mean value of all y_{ij} in the area. Actually, it is exactly the local mean filter for de-noising images contaminated by multiplicative noises.

4.2 The Weighted LS Method

When applying the weighted LS method to the model (8), we have the following minimization problem:

$$\min : F_2(\boldsymbol{\beta}) = (\mathbf{y} - \mathbf{A}\boldsymbol{\beta})^T \boldsymbol{\Sigma}_y^{-1} (\mathbf{y} - \mathbf{A}\boldsymbol{\beta}). \quad (16)$$

To derive the weighted LS estimate of $\boldsymbol{\beta}$, we can differentiate $F_2(\boldsymbol{\beta})$ of (16) with respect to $\boldsymbol{\beta}$ and let it be equal to zero. After some lengthy derivations, we can finally obtain the system of normal equations as follows:

$$(\mathbf{A}^T \hat{\boldsymbol{\Sigma}}_y^{-1} \mathbf{A}) \hat{\boldsymbol{\beta}} - \mathbf{A}^T \hat{\boldsymbol{\Sigma}}_y^{-1} \mathbf{y} - \mathbf{G}_1 (\mathbf{A} \hat{\boldsymbol{\beta}} - \mathbf{y}) = \mathbf{0}, \quad (17)$$

where

$$\mathbf{G}_1 = \begin{bmatrix} (\mathbf{A} \hat{\boldsymbol{\beta}} - \mathbf{y})^T \hat{\boldsymbol{\Sigma}}_y^{-1} \mathbf{D}_{ae_1} \boldsymbol{\Sigma}_m \hat{\mathbf{D}}_{a\hat{\boldsymbol{\beta}}} \hat{\boldsymbol{\Sigma}}_y^{-1} \\ (\mathbf{A} \hat{\boldsymbol{\beta}} - \mathbf{y})^T \hat{\boldsymbol{\Sigma}}_y^{-1} \mathbf{D}_{ae_2} \boldsymbol{\Sigma}_m \hat{\mathbf{D}}_{a\hat{\boldsymbol{\beta}}} \hat{\boldsymbol{\Sigma}}_y^{-1} \\ \vdots \\ (\mathbf{A} \hat{\boldsymbol{\beta}} - \mathbf{y})^T \hat{\boldsymbol{\Sigma}}_y^{-1} \mathbf{D}_{ae_t} \boldsymbol{\Sigma}_m \hat{\mathbf{D}}_{a\hat{\boldsymbol{\beta}}} \hat{\boldsymbol{\Sigma}}_y^{-1} \end{bmatrix},$$

\mathbf{D}_{ae_i} is a diagonal matrix with its k th diagonal element being equal to $(\mathbf{a}_k \mathbf{e}_i)$, \mathbf{e}_i is the i th natural basis vector of dimension t , $\hat{\mathbf{D}}_{a\hat{\boldsymbol{\beta}}}$ is the estimate of $\mathbf{D}_{a\boldsymbol{\beta}}$ by replacing $\boldsymbol{\beta}$ with its corresponding weighted LS estimate $\hat{\boldsymbol{\beta}}$. Following Xu et al. (2013), we can solve for the weighted LS estimate of $\boldsymbol{\beta}$ through the following iteration procedures:

$$\hat{\boldsymbol{\beta}}_{k+1} = (\mathbf{A}^T \hat{\boldsymbol{\Sigma}}_{y_k}^{-1} \mathbf{A})^{-1} \{ \mathbf{A}^T \hat{\boldsymbol{\Sigma}}_{y_k}^{-1} \mathbf{y} + \mathbf{G}_{1k} (\mathbf{A} \hat{\boldsymbol{\beta}}_k - \mathbf{y}) \}, \quad k = 0, 1, \dots \quad (18)$$

where $\hat{\boldsymbol{\Sigma}}_{y_k}$ and \mathbf{G}_{1k} stand for computing $\hat{\boldsymbol{\Sigma}}_y$ and \mathbf{G}_1 at the point of $\hat{\boldsymbol{\beta}}_k$.

It is obvious from (17) that the weighted LS estimate $\hat{\boldsymbol{\beta}}$ is nonlinear and is expected to be biased. Xu et al. (2013) derived the bias of $\hat{\boldsymbol{\beta}}$ in the mixed additive and multiplicative error model (8), which is denoted by $\mathbf{b}(\hat{\boldsymbol{\beta}})$ and is simply listed as follows:

$$\mathbf{b}(\hat{\boldsymbol{\beta}}) = E(\mathbf{b}(\boldsymbol{\beta})) = \mathbf{N}^{-1} \mathbf{g}_2, \quad (19)$$

where $\mathbf{N} = \mathbf{A}^T \boldsymbol{\Sigma}_y^{-1} \mathbf{A}$ and \mathbf{g}_2 is given by

$$\mathbf{g}_2 = \begin{bmatrix} \text{tr}\{\mathbf{D}_{ae_1} \boldsymbol{\Sigma}_m \mathbf{D}_{a\boldsymbol{\beta}} \boldsymbol{\Sigma}_y^{-1}\} \\ \text{tr}\{\mathbf{D}_{ae_2} \boldsymbol{\Sigma}_m \mathbf{D}_{a\boldsymbol{\beta}} \boldsymbol{\Sigma}_y^{-1}\} \\ \vdots \\ \text{tr}\{\mathbf{D}_{ae_t} \boldsymbol{\Sigma}_m \mathbf{D}_{a\boldsymbol{\beta}} \boldsymbol{\Sigma}_y^{-1}\} \end{bmatrix}.$$

By limiting themselves to the linear term of $\hat{\boldsymbol{\beta}}$ with respect to the random errors $\boldsymbol{\epsilon}_m$ and $\boldsymbol{\epsilon}_a$, Xu et al. (2013) also derived the first order approximation of the variance-covariance matrix of the weighted LS estimate $\hat{\boldsymbol{\beta}}$, which is denoted by $D_1(\hat{\boldsymbol{\beta}})$ and given as follows:

$$D_1(\hat{\boldsymbol{\beta}}) = (\mathbf{A}^T \boldsymbol{\Sigma}_y^{-1} \mathbf{A})^{-1}. \quad (20)$$

After taking the bias (19) into account, we obtain the approximate mean squared error (MSE) matrix of $\hat{\boldsymbol{\beta}}$ as follows:

$$\begin{aligned} \mathbf{M}(\hat{\boldsymbol{\beta}}) &= D_1(\hat{\boldsymbol{\beta}}) + \mathbf{b}(\hat{\boldsymbol{\beta}}) [\mathbf{b}(\hat{\boldsymbol{\beta}})]^T \\ &= (\mathbf{A}^T \boldsymbol{\Sigma}_y^{-1} \mathbf{A})^{-1} + (\mathbf{A}^T \boldsymbol{\Sigma}_y^{-1} \mathbf{A})^{-1} \mathbf{g}_2 \mathbf{g}_2^T (\mathbf{A}^T \boldsymbol{\Sigma}_y^{-1} \mathbf{A})^{-1}, \end{aligned} \quad (21)$$

where $\mathbf{M}(\hat{\boldsymbol{\beta}})$ stands for the MSE matrix of $\hat{\boldsymbol{\beta}}$.

4.3 The Bias-Corrected Weighted LS Method

Bias analysis in Xu and Shimada (2000) and Xu et al. (2013) clearly indicates that the bias of the weighted LS estimate is solely caused by the non-zero term of derivatives of the variance-covariance matrix $\boldsymbol{\Sigma}_y$ with respect to $\boldsymbol{\beta}$. Thus both works propose deleting the corresponding term in the normal equations, namely the third term on the right hand side of the normal equations (17), to remove the bias from the weighted LS estimate $\hat{\boldsymbol{\beta}}$. As a result, they are able to construct the bias-corrected weighted LS estimate of $\boldsymbol{\beta}$.

When the same idea is applied to the mixed additive and multiplicative error model (8), they derive the bias-corrected weighted LS estimate of $\boldsymbol{\beta}$, which is denoted by $\hat{\boldsymbol{\beta}}_{bc}$ and solved through the following system of normal equations:

$$(\mathbf{A}^T \hat{\boldsymbol{\Sigma}}_y^{-1} \mathbf{A}) \hat{\boldsymbol{\beta}}_{bc} - \mathbf{A}^T \hat{\boldsymbol{\Sigma}}_y^{-1} \mathbf{y} = \mathbf{0}, \quad (22)$$

where $\hat{\boldsymbol{\beta}}_{bc}$ is the bias-corrected WLS estimate of $\boldsymbol{\beta}$. Equivalently, $\hat{\boldsymbol{\beta}}_{bc}$ can be formally rewritten as follows:

$$\hat{\boldsymbol{\beta}}_{bc} = (\mathbf{A}^T \hat{\boldsymbol{\Sigma}}_y^{-1} \mathbf{A})^{-1} \mathbf{A}^T \hat{\boldsymbol{\Sigma}}_y^{-1} \mathbf{y}, \quad (23)$$

which is unbiased up to the second order approximation (Xu et al. 2013). The variance-covariance matrix of $\hat{\beta}_{bc}$ is denoted by $D(\hat{\beta}_{bc})$ and given by

$$D(\hat{\beta}_{bc}) = (\mathbf{A}^T \hat{\Sigma}_y^{-1} \mathbf{A})^{-1}. \quad (24)$$

Because the matrix $\hat{\Sigma}_y$ depends on $\hat{\beta}_{bc}$, (23) is actually a nonlinear system of equations and can, in general, be solved numerically. If the Gauss-Newton method is used to solve for the bias-corrected weighted LS estimate, we have the following iterative formula:

$$\hat{\beta}_{bc}^{k+1} = \hat{\beta}_{bc}^k - (\mathbf{A}^T \hat{\Sigma}_{yk}^{-1} \mathbf{A})^{-1} \mathbf{A}^T \hat{\Sigma}_{yk}^{-1} (\mathbf{A} \hat{\beta}_{bc}^k - \mathbf{y}), \quad k = 0, 1, \dots \quad (25)$$

(see also McCullagh 1983; McCullagh and Nelder 1989; Dennis and Schnabel 1996; Xu et al. 2013).

We should like to note that given some approximate values, say β_0 , Σ_{m0} and Σ_{a0} , we can then replace $\hat{\Sigma}_y$ of (22) with Σ_{y0} (computed at β_0 , Σ_{m0} and Σ_{a0}), which is exactly the conventional practice of adjustment of geodetic networks such as EDM, VLBI and GPS baselines. In other words, the conventional weighted LS adjustment of baseline networks can be interpreted as a special case of the bias-corrected weighted LS method with given approximate values. Nevertheless, the effectiveness of using approximate values would depend on how far away these approximate values deviate from their true values, as also pointed out and demonstrated in Xu et al. (2013).

5 Mixed Additive and Multiplicative Random Error Models with Prior Information

In this section, we will extend the parameter estimation in mixed additive and multiplicative random error models to the case with prior information. Prior information will only be assumed in the form of the first two moments on the unknown parameters β , i.e. its prior mean μ and prior variance-covariance matrix. Bearing the concept of additive and multiplicative random errors in mind, we will accordingly assume two types of prior variance-covariance matrices. As in the case of Gauss-Markov models with additive random errors, the first type of prior variance-covariance matrices is assumed to be independent of β and symbolically denoted by Σ_0 . If prior information on β is obtained from measurements contaminated by multiplicative errors other than the measurements \mathbf{y} in the mixed additive and multiplicative error model (8), then the prior variance-covariance matrix will surely be dependent on β , as can be readily seen in (24), for example. Thus, the second type of prior variance-covariance matrices is assumed to be the functions of β and

denoted by Σ_β . Of course, prior information can also be presented in the form of prior distributions. In this case, one can use Bayesian inference to estimate the unknown parameters. For more information, the reader is referred to Xu (1999).

If the model (8) is combined with the first type of prior variance-covariance matrices, namely, Σ_0 , then the corresponding generalized (weighted) LS objective function will become

$$\min : F_3(\beta) = (\mathbf{y} - \mathbf{A}\beta)^T \Sigma_y^{-1} (\mathbf{y} - \mathbf{A}\beta) + (\beta - \mu)^T \Sigma_0^{-1} (\beta - \mu). \quad (26)$$

According to the bias analysis in Xu and Shimada (2000) and Xu et al. (2013), we know that Σ_y will create a bias in the solution to (26). Since the prior variance-covariance matrix Σ_0 is independent of β , it will not contribute extra terms to the bias of the solution. By following the same rationale as in Sect. 4.3, we can ignore the dependence of Σ_y on β as if it were independent of β and, as a result, readily construct the bias-corrected estimator of β with prior information, denoted by $\hat{\beta}_{bc}^{p0}$, as follows:

$$\hat{\beta}_{bc}^{p0} = (\mathbf{A}^T \hat{\Sigma}_y^{-1} \mathbf{A} + \Sigma_0^{-1})^{-1} (\mathbf{A}^T \hat{\Sigma}_y^{-1} \mathbf{y} + \Sigma_0^{-1} \mu). \quad (27)$$

The first order accuracy of $\hat{\beta}_{bc}^{p0}$ is then given by

$$D(\hat{\beta}_{bc}^{p0}) = (\mathbf{A}^T \Sigma_y^{-1} \mathbf{A} + \Sigma_0^{-1})^{-1}. \quad (28)$$

If the second type of prior information is combined with the measurements from the model (8), we should then have the following optimization problem:

$$\min : F_4(\beta) = (\mathbf{y} - \mathbf{A}\beta)^T \Sigma_y^{-1} (\mathbf{y} - \mathbf{A}\beta) + (\beta - \mu)^T \Sigma_\beta^{-1} (\beta - \mu). \quad (29)$$

Obviously, both of Σ_y and Σ_β will now directly contribute terms to the bias of the optimal solution to the optimization problem (29), according to Xu and Shimada (2000) and Xu et al. (2013). In the similar manner to (27), we can construct the bias-corrected estimator of β with prior mean μ and prior variance-covariance matrix Σ_β , which is denoted by $\hat{\beta}_{bc}^{p1}$ and simply listed as follows:

$$\hat{\beta}_{bc}^{p1} = (\mathbf{A}^T \hat{\Sigma}_y^{-1} \mathbf{A} + \hat{\Sigma}_\beta^{-1})^{-1} (\mathbf{A}^T \hat{\Sigma}_y^{-1} \mathbf{y} + \hat{\Sigma}_\beta^{-1} \mu), \quad (30)$$

which is unbiased up to the second order approximation, and its first order accuracy is given by

$$D(\hat{\beta}_{bc}^{p1}) = (\mathbf{A}^T \Sigma_y^{-1} \mathbf{A} + \Sigma_\beta^{-1})^{-1}. \quad (31)$$

As in the case of the bias-corrected weighted LS estimation, the bias-corrected stochastic inference (or LS

collocation) with prior information in mixed multiplicative and additive error models is essentially of the same form as in the case of purely additive error models, as correctly pointed out by one of the reviewers. However, unlike stochastic inference in additive error models, the bias-corrected stochastic inference with prior information in mixed multiplicative and additive error models requires that $\hat{\beta}_{bc}^{pl}$ of (30) be computed iteratively, since the right hand side of (30) contains the unknowns $\hat{\beta}_{bc}^{pl}$ as well. Nevertheless, if one would simply apply the conventional principle of stochastic inference to mixed multiplicative and additive error models, one would end up with a biased estimator, which would not be of the same form as in the case of purely additive error models.

In the one-dimensional case, namely,

$$y_{ij} = (1 + \varepsilon_{mij})s_{ij} + \varepsilon_{aij},$$

then the bias-corrected LS estimate of s_{ij} with prior information can be rewritten as follows:

$$\hat{s}_{ij} = \mu_{ij} + \frac{\sigma_{\mu}^2}{\sigma_m^2 s_{0ij}^2 + \sigma_a^2 + \sigma_{\mu}^2} (y_{ij} - \mu_{ij}), \quad (32)$$

where μ_{ij} is the prior mean of s_{ij} , σ_{μ}^2 is the prior variance of μ_{ij} , s_{0ij} is some approximate value of s_{ij} , and σ_m^2 and σ_a^2 are the variances of the multiplicative and additive errors ε_{mij} and ε_{aij} , respectively. By properly choosing the values of μ_{ij} , σ_{μ}^2 , s_{0ij} , σ_m^2 and σ_a^2 , one can then construct the filter by Kuan et al. (1985) for image de-noising.

6 Concluding Remarks

Geodetic adjustment has been developed on the basis of Gauss-Markov models with additive random errors. The most important feature of such a Gauss-Markov model with additive random errors is that the accuracy of a measurement has nothing to do with the true value of the measurement. However, geodetic practice has clearly demonstrated that random errors of EDM, VLBI and GPS baselines indeed change with the length of a baseline. In other words, random errors of such types usually consist of two parts: one behaves more or less constant and may reflect only random effects of local nature, while the other is proportional to the length of the baseline and could, very likely, reflect total effect of the propagation path between the two stations. Such error characteristics are part of modern geodetic coherent imaging systems such as SAR and LiDAR. Obviously, the conventional adjustment theory that has been developed on

the assumption of additive random errors cannot theoretically meet the need to process geodetic measurements that are contaminated by multiplicative and/or mixed additive and multiplicative random errors.

In this paper, we have briefly reviewed two types of methods for parameter estimation in mixed additive and multiplicative error models, namely, quasi-likelihood and least-squares-based methods, with or without prior information. Quasi-likelihood, though theoretically connected with distributions, can be used directly for parameter estimation without any assumption on distributions. If there exist multiple solutions to the generalized estimating equations, no criterion is available for quasi-likelihood to pick up the right solution, however. The LS-based methods have a clearly defined objective function. Thus the sense of optimality of LS-based estimates is well defined. For the linear model (8), quasi-likelihood, the ordinary and bias-corrected weighted LS methods can all warrant an unbiased estimate of the unknown parameters, while the weighted LS method will generally lead to a biased estimate. Quasi-likelihood and the bias-corrected weighted LS method are more efficient than the ordinary LS method. We have also extended the bias-corrected LS estimate to the case with prior information, which can either be given in the form of prior mean and a parameter-free or a parameter-dependent prior variance-covariance matrix.

Acknowledgements This work is partially supported by the National Foundation of Natural Science of China (Nos.41204006, 41374016) and the project SKLGED2013-4-8-E, and the Grant-in-Aid for Scientific Research (C25400449). The authors thank the reviewers and the editor very much for their constructive comments, which help clarify some points of the paper.

References

- Aubert G, Aujol J-F (2008) A variational approach to removing multiplicative noise. *SIAM J Appl Math* 68:925–946
- Crowder M (1995) On the use of a working correlation matrix in using generalised linear models for repeated measures. *Biometrika* 82:407–410
- Dennis Jr. JE, Schnabel RB (1996) Numerical methods for unconstrained optimization and nonlinear equations. *SIAM classics in applied mathematics*. SIAM, Philadelphia
- Desmond AF (1997) Optimal estimating functions, quasi-likelihood and statistical modelling. *J Stat Plan Inference* 60:77–123
- Ewing CE, Mitchell MM (1970) Introduction to geodesy. Elsevier, New York
- Fitzmauric GM (1995) A caveat concerning independence estimating equations with multivariate binary data. *Biometrics* 51:309–317
- Flamant PH, Menzies RT, Kavaya MJ (1984) Evidence for speckle effects on pulsed CO₂ lidar signal returns from remote targets. *Appl Optics* 23:1412–1417
- Goodman JW (1976) Some fundamental properties of speckle. *J Opt Soc Am* 66:1145–1150
- Heyde CC (1997) Quasi-likelihood and its applications. Springer, New York

- Hill AC, Harris M, Ridley KC, Jakeman E, Lutzmann P (2003) Lidar frequency modulation vibrometry in the presence of speckle. *Appl Optics* 42:1091–1100
- Kuan DT, Sawchuk AA, Strand TC, Chavel P (1985) Adaptive noise smoothing filter for images with signal-dependent noise. *IEEE Trans Pattern Anal Mach Intell PAMI-7*:165–177
- Kukusha A, Malenkova A, Schneeweiss H (2010) Optimality of quasi-score in the multivariate mean-variance model with an application to the zero-inflated Poisson model with measurement errors. *Statistics* 44:381–396
- López-Martínez C, Fàbregas X, Pipia L (2011) Forest parameter estimation in the Pol-InSAR context employing the multiplicative-additive speckle noise model. *ISPRS J Photogram Remote Sens* 66:597–607
- MacDoran PF (1979) Satellite emission radio interferometric earth surveying series — GPS geodetic system. *Bull Géod* 53:117–138
- McCullagh P (1983) Quasi-likelihood functions. *Ann Stat* 11:59–67
- McCullagh P, Nelder J (1989) *Generalized linear models*, 2nd edn. Chapman and Hall, London
- Oliver CJ (1991) Information from SAR images. *J Phys D Appl Phys* 24:1493–1514
- Pearlman MR, Degnan JJ, Bosworth JM (2002) The international laser ranging service. *Adv Space Res* 30:135–143
- Petrov L, Gordon D, Gipson J, MacMillan D, Ma C, Fomalont E, Walker R, Carabajal C (2010) Precise geodesy with the very long baseline array. *J Geodesy* 83:859–876
- Seeber G (2003) *Satellite geodesy*, 2nd edn. de Gruyter, Berlin
- Shi Y, Xu PL, Peng JH, Shi C, Liu JN (2014) Adjustment of measurements with multiplicative errors: error analysis, estimates of the variance of unit weight, and effect on volume estimation from LiDAR-type digital elevation models. *Sensors* 14:1249–1266. doi:10.3390/s140101249
- Swami A (1994) Multiplicative noise models: parameter estimation using cumulants. *Signal Process* 36:355–373
- Tian H, Fowler B, Gamal El (2001) Analysis of temporal noise in CMOS photodiode active pixel sensor. *IEEE J Solid State Circuits* 36:92–101
- Ulaby F, Kouyate F, Brisco B, Williams T (1986) Textural information in SAR images. *IEEE Trans Geosci Remote Sens* 24:235–245
- Wang JY, Pruitt PA (1992) Effects of speckle on the range precision of a scanning lidar. *Appl Opt* 31:801–808
- Wedderburn R (1974) Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika* 61:439–447
- Willis P, Fagard H, Ferrage P, Lemoine FG, Noll CE, Noomen R, Otten M, Ries JC, Rothacher M, Soudarin L, Tavernier G, Valette JJ (2010) The international DORIS service, toward maturity. In: Willis P (ed) *DORIS: scientific applications in geodesy and geodynamics*. *Adv Space Res* 45(12):1408–1420
- Xu PL (1999) Despeckling SAR-type multiplicative noise. *Int J Remote Sens* 20:2577–2596
- Xu PL, Shimada S (2000) Least squares estimation in multiplicative noise models. *Commun Stat B* 29:83–96
- Xu PL, Shi Y, Peng JH, Liu JN, Shi C (2013) Adjustment of geodetic measurements with mixed multiplicative and additive random errors. *J Geodesy* 87:629–643