Acta Geod Geophys (2015) 50:353–363
DOI 10.1007/s40328-014-0096-y

Comparing the estimates of the variance of unit weight in multiplicative error models

Yun Shi · **Peiliang Xu**

Received: 10 November 2014 / Accepted: 18 December 2014 / Published online: 7 January 2015

Abstract Multiplicative error models should become more and more important in geodesy, since modern measurement technology on the basis of electromagnetic wave has clearly demonstrated that measurements of this type contain two types of random errors: fixed random errors and baseline-length dependent random errors. Although a number of the estimators of the variance of unit weight are derived from the least-squares-based adjustment methods for multiplicative error models recently, we know very little about their statistical performances. We first derive the variances of the estimates of the variance of unit weight in multiplicative error models. We find that the second order term of random errors will not affect the unbiasedness of an estimate of the variance of unit weight, if such a term is generated from the nonlinearity of models and/or least-squares-based nonlinear objective functions. The result is surprising, since the second order term of random errors has been well known to create the biases in both the estimate of parameters and the measurement corrections in the literature of nonlinear adjustment and nonlinear regression. Simulations are carried out to confirm the statistical analysis and to numerically compare the performances of different estimates of the variance of unit weight in multiplicative error models. From the simulation results, we recommend the estimate of the variance of unit weight with the bias-corrected weighted LS solution, followed by the two estimates with the ordinary LS solutions and the first estimate with the weighted LS solution.

Keywords Multiplicative error models · Bias-corrected weighted LS method · Variance of unit weight

Y. Shi
School of Geomatics, Xi'an University of Science and Technology, Xi'an 710054, 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

1 Introduction

Conventional geodetic adjustment theory and methods have almost always been established on the basis of Gauss–Markoff model, which can be represented by the addition of random errors and functional models. The starting observation equations of such types can be written as follows:

$$\boldsymbol{y} = \mathbf{f}(\boldsymbol{\beta}) + \boldsymbol{\varepsilon} \tag{1}$$

where $\boldsymbol{y}$ is a vector of measurements, $\mathbf{f}(\boldsymbol{\beta})$ the vector of linear and/or nonlinear functional models, $\boldsymbol{\beta}$ the vector of unknown parameters to be estimated, $\boldsymbol{\varepsilon}$ the random error vector of the measurements $\boldsymbol{y}$. An important feature of model (1) is that the random errors $\boldsymbol{\varepsilon}$ do not change with the magnitude of the true values of the signals $\mathbf{f}(\boldsymbol{\beta})$. In other words, $\boldsymbol{\varepsilon}$ is independent of the model parameters $\boldsymbol{\beta}$. Since the random errors $\boldsymbol{\varepsilon}$ disturb the measurements $\boldsymbol{y}$ in an additive manner, model (1) is also called additive error models.

Measurement technology on the basis of electromagnetic wave has fundamentally advanced modern earth-space observation. A huge amount of data can now be readily collected with a higher accuracy. As a result, adjustment theory and methods have to be further developed to meet the need of modern earth-space observation in order to better understand the nature of data and to extract information from data more efficiently and more accurately. In fact, electromagnetic-wave-based geodetic measurements have unambiguously demonstrated by themselves that measurements of these types are contaminated by both additive and multiplicative random errors (see, e.g. Xu et al. 2013; Shi et al. 2014). The vector form of a multiplicative error model can be symbolically represented as follows:

$$\boldsymbol{y} = \mathbf{f}(\boldsymbol{\beta}) \odot (1 + \boldsymbol{\varepsilon}) \tag{2}$$

where the symbols $\boldsymbol{y}$, $\boldsymbol{\beta}$, $\mathbf{f}(\boldsymbol{\beta})$ and $\boldsymbol{\varepsilon}$ have been defined as in model (1), $\odot$ is the Hadamard product of matrices and/or vectors of the same size. In other words, given two matrices $\boldsymbol{A} = (a_{ij})$ and $\boldsymbol{B} = (b_{ij})$, the Hadamard product of $\boldsymbol{A}$ and $\boldsymbol{b}$ is defined as $\boldsymbol{A} \odot \boldsymbol{B} = (a_{ij}b_{ij})$. In model (2), $\mathbf{1}$ stands for an n-dimensional vector with all its elements being equal to unity. In this paper, for simplicity of discussion but without loss of generality, we will assume that all the random errors of $\boldsymbol{\varepsilon}$ are of mean zero and stochastically independent with the same variance, i.e. the variance-covariance matrix of $\boldsymbol{\varepsilon}$ is equal to $\boldsymbol{I}\sigma^2$, where $\boldsymbol{I}$ is an identity matrix. If each of the functional models $f_i(\cdot)$ is linear with respect to $\boldsymbol{\beta}$, i.e. $f_i(\boldsymbol{\beta}) = \boldsymbol{x}_i^T\boldsymbol{\beta}$, $\boldsymbol{x}_i^T$ is a t-dimensional vector and $\boldsymbol{x}_i^T = (\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \ldots, \boldsymbol{x}_{it})$, then model (2) can be rewritten as follows:

$$\boldsymbol{y} = (\mathbf{X}\boldsymbol{\beta}) \odot (\mathbf{1} + \boldsymbol{\varepsilon}) \tag{3}$$

where $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)^T$. It is easy to see from model (3) that the random errors of measurements in a multiplicative error model are proportional to the true values of measurements. The larger the absolute values of measurements, the noisier the measurements. On the contrary, if the true values of measurements are small, then the random errors of the measurements are also proportionally small. The variance-covariance matrix of the measurements $\boldsymbol{y}$ is denoted by $\boldsymbol{\Sigma}_y$, which can also be written as $\boldsymbol{D}_1^2\sigma^2$, or $\boldsymbol{D}_y\sigma^2$, $\boldsymbol{D}_1 = diag(\boldsymbol{x}_i^T\boldsymbol{\beta})$ are the functions of the unknown parameters $\boldsymbol{\beta}$, $diag(\cdot)$ stands for a diagonal matrix.

Multiplicative errors are also known as speckles in electronic and optical literature (see, e.g. Flamant et al. 1984; Goodman 1976; Ulaby et al. 1986; Wang and Pruitt 1992; Xu 1999). On the other hand, multiplicative error model (3) has been called generalized linear models in statistics (see, e.g. Wedderburn 1974; McCullagh and Nelder 1989; Heyde 1997). A standard method to handle multiplicative errors in statistics is the quasi-likelihood. Although multiplicative errors should dominate almost all geodetic measurements of modern types, they have been investigated only recently in geodesy (see, e.g. Xu et al. 2013; Shi et al. 2014). Unlike the standard method of quasi-likelihood in statistics, geodesists approach to handle multiplicative error models from the point of view of the conventional least squares (LS) principle (see, e.g. Xu 1999; Xu and Shimada 2000; Xu et al. 2013; Shi et al. 2014). The advantage of using the LS principle is twofold: (i) we do not assume probability distributions for measurements; and (ii) the LS objective function is well defined. The LS-based estimation procedures and their error evaluations up to the first order approximation can also be found in the geodetic publications mentioned in the above. Shi et al. (2014) systematically investigated the estimation of the variance of unit weight in multiplicative error models and constructed five different estimators for it, starting with the three LS-based estimators of parameters. Nevertheless, they did not go further to study the statistical aspects of these estimators. As an important supplement to Shi et al. (2014), we will carry out a statistical analysis of their estimators of the variance of unit weight and carry out numerical simulations to compare their performances.

This paper is organized as follows. Section 2 will briefly discuss three LS-based estimation methods for parameter estimation and then directly follow Shi et al. (2014) to present five different estimators of the variance of unit weight in multiplicative models. By assuming that the multiplicative random errors are normally distributed, we will derive the variances of these five estimators and briefly compare their statistical aspects in Sect. 3. Finally, in Sect. 4, we will conduct numerical simulations to confirm our theoretical analysis and to compare the performances of the estimators of the variance of unit weight.

2 Estimation of the variance of unit weight in multiplicative error models

2.1 LS-based parameter estimation

We will now briefly outline three LS-based parameter estimation methods. For more details, the reader is referred to Xu and Shimada (2000), Xu et al. (2013) and Shi et al. (2014).

If we apply the ordinary LS method to the multiplicative error model (3), we can readily obtain the ordinary LS estimator of $\boldsymbol{\beta}$, which is denoted by $\hat{\boldsymbol{\beta}}_{LS}$ and given as follows:

$$\hat{\boldsymbol{\beta}}_{LS} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}. \tag{4a}$$

If we apply the weighted LS method to (3), then we obtain the corresponding weighted LS estimator of $\boldsymbol{\beta}$ as follows:

$$\hat{\boldsymbol{\beta}}_{WLS}=(\boldsymbol{X}^T\hat{\boldsymbol{D}}_{\boldsymbol{y}}^{-1}\boldsymbol{X})^{-1}\mathbf{X}^T\hat{\boldsymbol{D}}_{\boldsymbol{y}}^{-1}\boldsymbol{y}+(\boldsymbol{X}^T\hat{\boldsymbol{D}}_{\boldsymbol{y}}^{-1}\boldsymbol{X})^{-1}\begin{bmatrix}(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}_{WLS})^T\hat{\boldsymbol{P}}_1(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}_{WLS})\\ \vdots \\ (\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}_{WLS})^T\hat{\boldsymbol{P}}_t(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}_{WLS})\end{bmatrix} \tag{4b}$$

where $\hat{\boldsymbol{\beta}}_{WLS}$ is the weighted LS estimator of $\boldsymbol{\beta}$. Since $\hat{\boldsymbol{\beta}}_{WLS}$ is known to be biased (see also Xu and Shimada 2000; Xu et al. 2013), it does not possess the optimal statistical properties of unbiasedness and minimum variance, as in the case of linear models with additive random errors. It is also well known from Xu and Shimada (2000) and Xu et al. (2013) that the bias of the weighted LS estimate $\hat{\boldsymbol{\beta}}_{\mathrm{WLS}}$ is completely attributed to the dependence of the variance-covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{y}}$ of the measurements on the unknown parameters $\boldsymbol{\beta}$. As a result, one can construct an almost unbiased estimate of parameters by directly removing the term of derivatives arising from the dependence of $\boldsymbol{\Sigma}_{\boldsymbol{y}}$ on $\boldsymbol{\beta}$ from the nonlinear normal equations.

In other words, by directly deleting the bias term from the nonlinear normal equations of $\hat{\boldsymbol{\beta}}_{\mathrm{WLS}}$, we can derive a new almost unbiased estimate of $\boldsymbol{\beta}$. Such an estimator, called the bias-corrected weighted LS estimator, is denoted by $\hat{\boldsymbol{\beta}}_{bc}$ and given as follows:

$$\hat{\boldsymbol{\beta}}_{bc} = (\boldsymbol{X}^{\boldsymbol{T}} \hat{\boldsymbol{D}}_{\boldsymbol{y}}^{-1} \boldsymbol{X})^{-1} \boldsymbol{X}^{\boldsymbol{T}} \hat{\boldsymbol{D}}_{\boldsymbol{y}}^{-1} \boldsymbol{y} \tag{4c}$$

2.2 Estimation of the variance of unit weight

In order to estimate the variance of unit weight in the multiplicative error model (3), our basic idea is to first use a quadratic form of the residuals or equivalently the measurement corrections, then apply the expectation operator to it, and finally construct an estimate of the variance of unit weight according to the relationship between the expectation of the quadratic form and σ^2. Thus, Shi et al. (2014) were able to construct five different estimators of σ^2. In what follows, we will directly give these estimators in association with the three LS-based parameter estimation methods.

In the case of the ordinary LS method, Shi et al. (2014) gave two estimators for the variance of unit weight, which are given, respectively, as follows:

$$\hat{\sigma}^2_{LS1} = \boldsymbol{V}^{\boldsymbol{T}}_{LS} \boldsymbol{V}_{LS} / r_{LS1} \tag{5a}$$

$$\hat{\sigma}^2_{LS2} = \boldsymbol{V}^{\boldsymbol{T}}_{LS} \hat{\boldsymbol{D}}_{\boldsymbol{y}}^{-1} \boldsymbol{V}_{LS} / r_{LS2} \tag{5b}$$

where

$$\boldsymbol{V}_{LS} = \boldsymbol{X}\hat{\boldsymbol{\beta}}_{LS} - \boldsymbol{y}$$

$$r_{\boldsymbol{LS}}\mathbf{1} = tr\left\{\hat{\boldsymbol{D}}_{\boldsymbol{y}} - \hat{\boldsymbol{D}}_{\boldsymbol{y}}\boldsymbol{X}(\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\boldsymbol{T}}\right\},$$

$$r_{LS2} = n - 2t + tr\left\{\hat{\boldsymbol{D}}_{\boldsymbol{y}}^{-1}\boldsymbol{X}(\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\boldsymbol{T}}\hat{\boldsymbol{D}}_{\boldsymbol{y}}\boldsymbol{X}(\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\boldsymbol{T}}\right\}$$

In the cases of the weighted LS and bias-corrected weighted LS estimates, if we simply ignore the biases of the weighted LS estimate and its corresponding residuals, then the estimates of the variance of unit weight for both the weighted LS and bias-corrected weighted LS estimates should be identical. As a result, we can directly write these two estimators for the variance of unit weight as follows:

$$\hat{\sigma}^2_{WLS1} = \boldsymbol{V}^{\boldsymbol{T}}_{WLS} \hat{\boldsymbol{D}}_{\boldsymbol{y}}^{-1} \boldsymbol{V}_{WLS} / (n - t) \tag{6a}$$

$$\hat{\sigma}^2_{\boldsymbol{bc}} = \boldsymbol{V}^{\boldsymbol{T}}_{\boldsymbol{bc}} \hat{\boldsymbol{D}}_{\boldsymbol{y}}^{-1} \boldsymbol{V}_{\boldsymbol{bc}} / (n - t) \tag{7}$$

where

$$\boldsymbol{V}_{WLS} = \boldsymbol{X}\hat{\boldsymbol{\beta}}_{WLS} - \boldsymbol{y}$$

$$\boldsymbol{V}_{bc} = \boldsymbol{X}\hat{\boldsymbol{\beta}}_{bc} - \boldsymbol{y}.$$

If we take the bias of the weighted LS estimate $\hat{\boldsymbol{\beta}}_{WLS}$ into account, we may like to first remove the biases of the corrections of measurements and then use the bias-corrected corrections of measurements to estimate the variance of unit weight. As a result, we can construct the second estimate of σ^2 for the weighted LS estimate, which is simply given as follows:

$$\hat{\sigma}^2_{WLS2} = (\boldsymbol{V}_{WLS} - \boldsymbol{X}\hat{\boldsymbol{b}}_{\beta})^{\boldsymbol{T}} \hat{\boldsymbol{D}}_{\boldsymbol{y}}^{-1} (\boldsymbol{V}_{WLS} - \boldsymbol{X}\hat{\boldsymbol{b}}_{\beta}) / (n - t) \tag{6b}$$

where $\hat{\boldsymbol{b}}_\beta$ is the estimate of the bias of the weighted LS estimate. According to Xu and Shimada (2000), we have the theoretical relationship:

$$\mathrm{E}(\boldsymbol{b}_\beta) = \left(\boldsymbol{X}^T \boldsymbol{D}_y^{-1} \boldsymbol{X}\right)^{-1} \boldsymbol{c}, \tag{8a}$$

where the elements of the vector $\boldsymbol{c}$ are given by:

$$c_i = tr\left\{\boldsymbol{P}_i\left[\boldsymbol{D}_y - \boldsymbol{X}\left(\boldsymbol{X}^T \boldsymbol{D}_y^{-1} \boldsymbol{X}\right)\boldsymbol{X}^T\right]\right\}\sigma^2, \tag{8b}$$

$$\boldsymbol{P}_i = \boldsymbol{diag}\left[x_{ji} \Big/ \left(\boldsymbol{x}_j^T \boldsymbol{\beta}\right)^3\right], \tag{8c}$$

which can be computed by replacing the unknown parameters with their corresponding estimates. Since the elements of c contain the unknown variance of unit weight σ^2, we can either iteratively solve for the estimate of σ^2 or substitute (8) into (6b) and then solve for σ^2. In this latter case, if we neglect the term with σ^4, we can obtain a directly computable formula for σ^2as follows:

$$\begin{aligned}\hat{\sigma}^2_{WLS2} &= \boldsymbol{V}^T_{WLS}\hat{\boldsymbol{D}}_y^{-1}\boldsymbol{V}_{WLS} \Big/ \left[n - t + 2\boldsymbol{V}^T_{WLS}\hat{\boldsymbol{D}}_y^{-1}\boldsymbol{X}(\boldsymbol{X}^T\hat{\boldsymbol{D}}_y^{-1}\boldsymbol{X})^{-1}\boldsymbol{h}\right] \\ &= \hat{\sigma}^2_{WLS1}/(1 + c_m)\end{aligned} \tag{9a}$$

where $\boldsymbol{h} = \boldsymbol{c}/\sigma^2$,and the coefficient c_m is given by

$$c_m = 2\boldsymbol{V}^T_{WLS}\hat{\boldsymbol{D}}_y^{-1}\boldsymbol{X}(\boldsymbol{X}^T\hat{\boldsymbol{D}}_y^{-1}\boldsymbol{X})^{-1}\boldsymbol{h} \Big/ (n - t). \tag{9b}$$

Practically, since c_m is computed with the corrections of measurements $\boldsymbol{V}_{WLS}$, it can be negative. Nevertheless, the expectation of $\boldsymbol{V}_{WLS}$ is known to be proportional to $\boldsymbol{X}(\boldsymbol{X}^T\hat{\boldsymbol{D}}_y^{-1}\boldsymbol{X})^{-1}\boldsymbol{h}$. Thus, the expectation value of c_m is theoretically positive.

3 Variances of the estimates of the variance of unit weight in multiplicative error models

Although the variance of the estimate of variance of unit weight with the weighted LS method has been well documented in the literature of linear models with additive random errors, nothing has ever been done about statistical aspects of this quantity in the case of multiplicative random errors. To better understand the statistical properties and performances of the five estimators of the variance of unit weight in the previous section, we will now derive the variances of these five estimators. Our starting point is the following theorem on the quadratic form of a normally distributed random vector.

Theorem 1 (Searle 1971) *Given a normally distributed vector* $\boldsymbol{x}$*, i.e.* $\boldsymbol{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$*,then the covariance of the quadratic forms* $\boldsymbol{x}^T\boldsymbol{P}\boldsymbol{x}$ *and* $\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x}$ *is given by*

$$cov(\boldsymbol{x}^T\boldsymbol{P}\boldsymbol{x}, \boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x}) = 2tr(\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{Q}\boldsymbol{\Sigma}) + 4\boldsymbol{\mu}^T\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{Q}\boldsymbol{\mu}. \tag{10a}$$

If $\boldsymbol{P} = \boldsymbol{Q}$ *and* $\boldsymbol{\mu} = \boldsymbol{0}$*, then the variance of the quadratic form* $\boldsymbol{x}^T\boldsymbol{P}\boldsymbol{x}$ *is given by*

$$\sigma^2(\boldsymbol{x}^T\boldsymbol{P}\boldsymbol{x}) = 2tr(\boldsymbol{P}\boldsymbol{\Sigma})^2. \tag{10b}$$

To start with, we assume that the random errors are normally distributed with zero mean. In order to compute the variance of an estimate of the variance of unit weight in the multiplicative error model (3), the basic idea is to represent the estimate of the variance of unit weight as a quadratic form of a normally distributed random vector and then use Theorem 1 to find the required variance.

In what follows, we will derive the variances of the five estimators of the variance of unit weight in the previous section in association with the three LS-based parameter estimation methods.

First of all, we denote $\boldsymbol{\varepsilon}_y = (\boldsymbol{X\beta}) \odot \boldsymbol{\varepsilon}$ and then rewrite (3) as follows:

$$\boldsymbol{y} = \boldsymbol{X\beta} + (\boldsymbol{X\beta}) \odot \boldsymbol{\varepsilon} = \bar{\boldsymbol{y}} + \boldsymbol{\varepsilon_y} \tag{11}$$

To derive the variances of the two estimators of the variance of unit weight with the ordinary LS method, we rewrite the LS corrections of measurements as:

$$\boldsymbol{V}_{\mathrm{LS}} = \left[\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T - \boldsymbol{I}\right]\boldsymbol{\varepsilon_y}.$$

Because any quadratic form of $\boldsymbol{V}_{\mathrm{LS}}$ is a second order function of the random errors $\boldsymbol{\varepsilon_y}$, if we confine ourselves only to the second order approximation of $\boldsymbol{\varepsilon_y}$, then all the other terms in the two estimators (5) of the variance of unit weight can be directly substituted with their true values. The two quadratic forms with the corrections $\boldsymbol{V}_{\mathrm{LS}}$ of measurements from the ordinary LS method can be respectively represented as follows:

$$\boldsymbol{V}_{LS}^{\mathrm{T}}\boldsymbol{V}_{\mathrm{LS}} = \boldsymbol{\varepsilon}_y^T\left[\boldsymbol{X}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}} - \boldsymbol{I}\right]\boldsymbol{\varepsilon}_y^T,$$

$$\boldsymbol{V}_{\mathrm{LS}}^{\mathrm{T}}\hat{\boldsymbol{D}}_{\boldsymbol{y}}^{-1}\boldsymbol{V}_{\mathrm{LS}} \approx \boldsymbol{\varepsilon}_y^T\left[\boldsymbol{X}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}} - \boldsymbol{I}\right]\boldsymbol{D}_y^{-1}\left[\boldsymbol{X}\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}} - \boldsymbol{I}\right]\boldsymbol{\varepsilon}_y^T$$

Bearing in mind that $\mathrm{E}(\boldsymbol{\varepsilon_y}) = \boldsymbol{0}$, $\mathrm{D}(\boldsymbol{\varepsilon_y}) = \boldsymbol{D_y}\,\sigma^2$ and (12b) of Theorem 1,we can readily write the variances for the two estimators (5a) and (5b) with the ordinary LS estimate of parameters as follows:

$$\boldsymbol{\sigma}^2(\hat{\sigma}_{\mathrm{LS1}}^2) = 2\mathrm{tr}\left\{\left[\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T - \boldsymbol{I}\right]\boldsymbol{D_y}\right\}^2\sigma^4\Big/r_{\mathrm{LS1}}^2, \tag{12a}$$

$$\sigma^2(\hat{\sigma}_{\mathrm{LS2}}^2) = 2\mathrm{tr}\left\{\left[\boldsymbol{X}\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}} - \boldsymbol{I}\right]\boldsymbol{D}_y^{-1}\left[\boldsymbol{X}\left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}} - \mathbf{I}\right]D_y\right\}^2\sigma^4\Big/r_{\mathrm{LS2}}^2, \tag{12b}$$

where $\sigma^2(\hat{\sigma}_{\mathrm{LS1}}^2)$ and $\sigma^2(\hat{\sigma}_{\mathrm{LS2}}^2)$ stand for the variances of $\hat{\sigma}_{\mathrm{LS1}}^2$ and $\hat{\sigma}_{\mathrm{LS2}}^2$, respectively.

To find the variances of $\hat{\sigma}_{\mathrm{WLS1}}^2$in association with the weighted LS method, we expand $\hat{\boldsymbol{\beta}}_{\mathrm{WLS}}$with respect to the random errors $\boldsymbol{\varepsilon}_y$ up to the first order approximation as follows:

$$\hat{\boldsymbol{\beta}}_{\mathrm{WLS}} = \boldsymbol{\beta} + (\boldsymbol{X}^T\boldsymbol{D}_y^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{D}_y^{-1}\boldsymbol{\varepsilon}_y. \tag{13}$$

The corresponding corrections of measurements up to the first order approximation of the errors can be directly written as:

$$\boldsymbol{V}_{\mathrm{WLS}} = \left[\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{D}_y^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{D}_y^{-1} - \boldsymbol{I}\right]\boldsymbol{\varepsilon_y}. \tag{14}$$

Since

$$\left[\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{D}_y^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{D}_y^{-1} - \boldsymbol{I}\right]^T\boldsymbol{D}_y\left[\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{D}_y^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{D}_y^{-1} - \boldsymbol{I}\right]$$
$$= \boldsymbol{D}_y^{-1} - \boldsymbol{D}_y^{-1}\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{D}_y^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{D}_y^{-1}$$

and bearing in mind that $\mathrm{E}(\boldsymbol{\varepsilon_y}) = \boldsymbol{0}$, $\mathrm{D}(\boldsymbol{\varepsilon_y}) = \boldsymbol{D}_y\,\sigma^2$ and (12b) of Theorem 1,we have

$$\left\{\left[\boldsymbol{D}_y^{-1} - \boldsymbol{D}_y^{-1}\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{D}_y^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{D}_y^{-1}\right]\boldsymbol{D}_y\right\}^2 = \boldsymbol{I} - \boldsymbol{D}_y^{-1}\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{D}_y^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T.$$

As a result, the variances of the two estimates of the variance of unit weight from the weighted LS estimate of parameters, i.e., $\hat{\sigma}^2_{WLS1}$ and $\hat{\sigma}^2_{WLS2}$, are respectively equal to:

$$\sigma^2(\hat{\sigma}^2_{\mathrm{WLS1}}) = 2\sigma^4 \Big/ (n-t), \tag{15a}$$

$$\sigma^2(\hat{\sigma}^2_{\mathrm{WLS2}}) = 2\sigma^4 \Big/ (n-t) \Big/ (1+c_{\mathrm{m}})^2. \tag{15b}$$

It is very interesting to note that although the weighted LS estimate $\hat{\boldsymbol{\beta}}_{\mathrm{WLS}}$ and its corresponding corrections of measurements in the multiplicative error model (3) are well known to be biased (see, e.g. Xu and Shimada 2000; Xu et al. 2013), both the estimates of the variance of unit weight $\hat{\sigma}^2_{WLS1}$ and $\hat{\sigma}^2_{WLS2}$ are clearly unbiased up to the second order of approximation. This result is theoretically significant and clearly cannot be found in any literature, either geodetically or statistically, at least, to our best knowledge. Actually, in the case of ill-posed inverse problems, Xu et al. (2006) and Xu (2009) show that a positive regularization parameter makes the estimate of parameters and its corresponding corrections/residuals of measurements biased. They also show that correcting the biases of the residuals is necessary and can be very significant to estimate the variance components in an ill-posed inverse problem. If we compare our new results here with those in Xu et al. (2006) and Xu (2009), we can conclude that the necessity and significance of correcting the biases of the estimated parameters and the residuals depend on the source of bias. If the biases of the residuals are generated from the second order term of the random errors due to the nonlinearity of models and/or objective functions, then such bias-generated second order terms of random errors will be turned into the terms higher than the third order approximation in the quadratic form of the residuals; as a result, these terms higher than the third order will not affect the estimation of the variance of unit weight up to the second order approximation. When comparing (15a) with (15b), we can see that if $-2 < c_{\mathrm{m}} < 0$, the variance of $\hat{\sigma}^2_{WLS1}$is smaller than that of $\hat{\sigma}^2_{WLS2}$; otherwise, the variance of $\hat{\sigma}^2_{WLS1}$ is larger than that of $\hat{\sigma}^2_{WLS2}$.

Since the bias-corrected weighted LS estimate $\hat{\boldsymbol{\beta}}_{bc}$ and the weighted LS estimate $\hat{\boldsymbol{\beta}}_{WLS}$ are different in that $\hat{\boldsymbol{\beta}}_{bc}$ does not contain the second order terms of the random errors. In other words, both $\hat{\boldsymbol{\beta}}_{bc}$ and $\hat{\boldsymbol{\beta}}_{\mathrm{WLS}}$ are essentially identical up to the first order approximation of the random errors. Thus, without any further derivation, we can directly write the variance of the estimate of unit weight with the bias-corrected weighted LS estimate $\hat{\boldsymbol{\beta}}_{bc}$as follows:

$$\boldsymbol{\sigma}^2(\hat{\sigma}^2_{bc}) = 2\boldsymbol{\sigma}^4 \Big/ (n-t). \tag{16}$$

4 Numerical simulations and result analysis

To demonstrate our theoretical analysis and to further understand numerically the statistical aspects of the estimates of the variance of unit weight in the multiplicative error model (3), we will use the linear regression example of Xu and Shimada (2000) to conduct a large scale of simulations in this section. The regression line is given as follows:

$$y_i = (\beta_1 + x_i\beta_2)(1+\varepsilon_i), \tag{17}$$

where both the parameters β_1 and β_2 take the true values of 5, the variables x_i take values in the interval [−1, 3], ε_i are the random errors, which are assumed to be independent and normally distributed with mean zero and a standard deviation of 0.4, i.e. $\sigma = 0.4$. Based on the starting model (17), we simulate 100 measurements, with the corresponding x_i uniformly

Table 1 The true values of parameters and their means from the experiments for the three LS-based methods

Parameters	β_1	β_2
True values	5.0	5.0
ordinary LS method	4.9904	4.9997
weighted LS method	5.7850	5.8015
bias-corrected WLS method	4.9963	4.9978

sampled in the interval $[-1, 3]$. To make sure that the mean values of the parameter estimates of both β_1 and β_2 from the independent experiments are accurate up to the order of 0.01, we finally set the number of experiments to 2000, each with a different set of simulated random errors.

From these 2,000 independent experiments, we obtain 2000 sets of $\hat{\boldsymbol{\beta}}_{\mathrm{LS}}$, $\hat{\boldsymbol{\beta}}_{\mathrm{WLS}}$, $\hat{\boldsymbol{\beta}}_{\mathrm{bc}}$ and five estimates of the variance σ^2 of unit weight, i.e.$\hat{\sigma}^2_{LS1}$, $\hat{\sigma}^2_{LS2}$, $\hat{\sigma}^2_{WLS1}$, $\hat{\sigma}^2_{WLS2}$ and $\hat{\sigma}^2_{bc}$. Thus, we can respectively compute the mean values of the estimated parameters as follows:

$$\bar{\boldsymbol{\beta}}_{\mathbf{M}} = \frac{1}{2{,}000}\sum_{i=1}^{2{,}000}\hat{\boldsymbol{\beta}}^{i}_{\mathbf{M}} \tag{18}$$

for the two parameters $\boldsymbol{\beta}$, where the subscript $\mathbf{M}$ stands for each of the three LS-based methods. In the case of the estimates of the variance of unit weight, we have the mean values:

$$\bar{\sigma}^2_K = \frac{1}{2{,}000}\sum_{i=1}^{2{,}000}\hat{\sigma}^2_{iK}, \tag{19a}$$

and the variances of $\hat{\sigma}^2_K$:

$$\hat{\sigma}^2(\hat{\sigma}^2_K) = \frac{1}{1{,}999}\sum_{i=1}^{2{,}000}(\hat{\sigma}^2_{iK} - \bar{\sigma}^2_K)^2, \tag{19b}$$

where the subscript K stands for each of *LS1*, *LS2*, *WLS1*, *WLS2* and *bc*.

Listed in Table 1 are the true values of the two parameters and their mean values of the estimates from the 2,000 experiments for the three LS-based methods, as computed by using (18). Obviously, the ordinary LS and bias-corrected weighted LS methods produce the almost unbiased estimates of the two parameters, but the weighted LS method results in the significant biases for both the estimated parameters. These simulations further confirm the results reported in Xu and Shimada (2000). With the true values and the mean values of the estimated parameters, we can also compute the true and mean values of measurements, which are shown in Fig.1, together with a sample of randomly generated measurements. Figure 1 clearly indicates that the ordinary LS and bias-corrected weighted LS methods can unbiasedly compute the adjusted measurements but the weighted LS method will bias the adjusted measurements.

We now summarize the statistics from the numerical simulations of the estimates of the variance of unit weight in Table 2. The values in the rows of $\bar{\sigma}^2_K$ and $\hat{\sigma}(\hat{\sigma}^2_K)$ are computed by using formulae (19a) and (19b), respectively. The values in the last row of Table 2 are computed by using the theoretical formulae (12a), (12b), (15a), (15b) and (16), respectively. Comparing the numbers in the row of $\bar{\sigma}^2_K$ with the true value of σ^2, we can see that $\hat{\sigma}^2_{LS1}$, $\hat{\sigma}^2_{LS2}$, $\hat{\sigma}^2_{WLS1}$ and $\hat{\sigma}^2_{bc}$ perform very well to reconstruct the variance of unit weight. They work much better than $\hat{\sigma}^2_{WLS2}$, which may imply that this estimate may be biased due to the neglect of

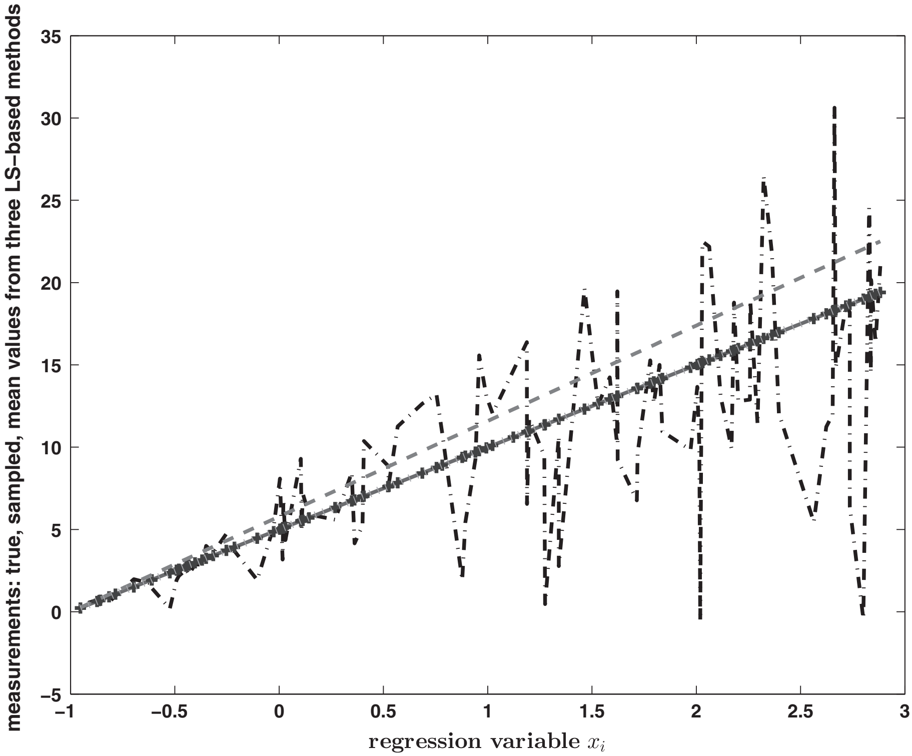

Fig. 1 The true and mean values of measurements from the simulation experiments, together with a set of randomly simulated measurements: *(green) dotted line*—true values; *(blue) mark* '+'—mean values with the ordinary LS method; *(red) solid line*—mean values with the bias-corrected weighted LS method; *(pink) dashed line*—mean values with the weighted LS method; and *(black) dash-dotted line*—a set of randomly simulated measurements. We note that *the three lines* for the true values and the mean values with both the ordinary LS and bias-corrected weighted LS methods are almost overlapped in this figure. (Color figure online)

the term with σ^4. $\hat{\sigma}^2_{LS1}$ has a better performance than $\hat{\sigma}^2_{LS2}$, if we compare their accuracy of estimates in the last two rows of Table 2. If we further take the accuracy of these estimates (row of $\hat{\sigma}(\hat{\sigma}^2_K)$ in Table 2) into account, we can see that $\hat{\sigma}^2_{bc}$ performs the best, followed by $\hat{\sigma}^2_{WLS1}$, $\hat{\sigma}^2_{LS1}$ and $\hat{\sigma}^2_{LS2}$. The results of $\hat{\sigma}^2_{WLS1}$ also clearly indicate that it can be used to almost unbiasedly estimate the variance of unit weight, even though the weighted LS method is known to produce significant biases in both the estimated parameters, the corrections of measurements and the adjusted measurements; its difference from the true value of σ^2 is within two times of the standard deviation ($\sigma(\hat{\sigma}^2_{WLS1}) = 0.0229$). If we compare the values in the last two rows of Table 2, we can conclude that the theoretical formulae (12a), (15a) and (16) can correctly predict the accuracy of the estimates of σ^2, while (12b) and (15b) are a bit too optimistic.

5 Conclusions

The variance of unit weight is one of the most important quantities in adjustment and statistical testing in linear models (see, e.g. Koch 1999; Gui 1995). Shi et al. (2014) derived a number of

Table 2 The true values and the statistics of the estimates of the variance of unit weight

σ^2	$\hat{\sigma}^2_{LS1}$	$\hat{\sigma}^2_{LS2}$	$\hat{\sigma}^2_{WLS1}$	$\hat{\sigma}^2_{WLS2}$	$\hat{\sigma}^2_{bc}$
True values	0.1600	0.1600	0.1600	0.1600	0.1600
$\bar{\sigma}^2_K$	0.1614	0.1719	0.1386	0.1048	0.1613
$\hat{\sigma}(\hat{\sigma}^2_K)$	0.0348	0.1769	0.0200	0.0148	0.0271
$\sigma(\hat{\sigma}^2_K)$	0.0299	0.0816	0.0229	0.0014	0.0229

the estimators for the variance of unit weight in multiplicative error models recently. However, very little is known about their statistical performances. To understand the statistical aspects of these estimators, we have approximately derived their variances in multiplicative error models. In the case of the weighted LS method, we have found for the first time that the second order term of random errors does not affect the unbiasedness of an estimate of the variance of unit weight up to the second order approximation, if such a term is generated from the nonlinearity of models and/or least-squares-based nonlinear objective functions. The result is not reported in any literature, either statistically or geodetically, at least, to our best knowledge. This is a very interesting result, since the second order term of random errors has been well known to create the biases in both the estimate of parameters and the measurement corrections in nonlinear adjustment and nonlinear regression.

Numerical simulations have been carried out to confirm our statistical analysis and to compare the performances of different estimates of the variance of unit weight in multiplicative error models. The simulation results have clearly demonstrated that $\hat{\sigma}^2_{LS1}$, $\hat{\sigma}^2_{LS2}$,$\hat{\sigma}^2_{WLS1}$ and $\hat{\sigma}^2_{bc}$ can be safely used to estimate the variance of unit weight. In particular, the results of $\hat{\sigma}^2_{WLS1}$ confirm our theoretical conclusion of almost unbiasedness. The simulation results of $\hat{\sigma}^2_{WLS2}$ may indicate that this estimator could be biased. The theoretical formulae (12a), (15a) and (16) can properly predict the accuracy for their corresponding estimates of σ^2, while (12b) and (15b) are too optimistic. The simulations have also clearly shown that the variances of the estimates of the variance of unit weight can be estimated through Monte Carlo simulation. Finally, from the simulation results, we recommend estimating the variance of unit weight by using the bias-corrected weighted LS method, followed by the two estimates of the same quantity by using the ordinary LS method and the first estimate (6a) by using the weighted LS method. We should like to note that if additive and multiplicative random errors are assumed to have the same variance of unit weight, then the results in this paper are also valid for mixed additive and multiplicative error models.

Acknowledgments This work is partially supported by the Chinese National Science Foundation (Projects Nos. 41204006 and 41020144004) and a Grant-in-Aid for Scientific Research C25400449.

References

Flamant PH, Menzies RT, Kavaya MJ (1984) Evidence for speckle effects on pulsed CO_2 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

Gui QM (1995) The ridge estimate and the Stein estimate of the variance of unit weight. J PLA Inst Surveing Mapp 3:227–230 in Chinese with English abstract

Heyde CC (1997) Quasi-likelihood and its applications. Springer, New York

Koch KR (1999) Parameter estimation and hypothesis testing in linear models. Springer-Verlag, Berlin

McCullagh P, Nelder J (1989) Generalized linear models, 2nd edn. Chapman and Hall, London

Searle SR (1971) Linear models. Wiley, New York

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

Ulaby F, Kouyate F, Brisco B, Williams T (1986) Textural information in SAR images. IEEE Trans Geosci Rem Sens 24:235–245

Wang JY, Pruitt PA (1992) Effects of speckle on the range precision of a scanning lidar. Appl Optics 31:801–808

Wedderburn R (1974) Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. Biometrika 61:439–447

Xu PL (1999) Despeckling SAR-type multiplicative noise. Int J Rem Sens 20:2577–2596

Xu PL (2009) Iterative generalized cross-validation for fusing heteroscedastic data of inverse ill-posed problems. Geophys J Int 179:182–200

Xu PL, Shimada S (2000) Least squares estimation in multiplicative noise models. Commun Stat B29:83–96

Xu PL, Shen YZ, Fukuda Y, Liu YM (2006) Variance component estimation in inverse ill-posed linear models. J Geod 80:69–81

Xu PL, Shi Y, Peng JH, Liu JN, Shi C (2013) Adjustment of geodetic measurements with mixed multiplicative and additive random errors. J Geod 87:629–643