Fast implementation of least squares variance component estimation for diagonal matrices
applications to GNSS time series

Mashhadizadeh-Maleki, Saeed; Amiri-Simkooei, Alireza

DOI
10.1007/s10291-024-01778-5

Publication date
2025

Document Version
Final published version

Published in
GPS Solutions

GPS Solutions (2025) 29:24
https://doi.org/10.1007/s10291-024-01778-5

ORIGINAL ARTICLE

Fast implementation of least squares variance component estimation for diagonal matrices: applications to GNSS time series

Saeed Mashhadizadeh-Maleki[1] · Alireza Amiri-Simkooei[1,2]

Received: 12 July 2024 / Accepted: 7 November 2024

Abstract

The study of long-term GNSS time series provides valuable insights for researchers in the field of earth sciences. Understanding the trends in these time series is particularly important for geodynamic researchers focused on earth crust movements. Functional and stochastic models play a crucial role in estimating trend values within time series data. Various methods are available to estimate variance components in GNSS time series. The least squares variance component estimation (LS-VCE) method stands out as one of the most effective approaches for this purpose. We introduce an innovative method, which streamlines calculations and simplifies equations, and therefore significantly boosting the processing speed for diagonal(ized) cofactor matrices. The method can be applied to the GNSS time series of linear stochastic models consisting of white noise, flicker noise and random walk noise. Moreover, unlike the conventional approaches, our method experiences high computational efficiency even with an increase in the number of colored noise components in time series data. For GNSS time series, this variable transformation has been applied to both univariate and multivariate modes, preserving the optimal properties of LS-VCE. We conducted simulations on daily time series spanning 5, 10, 15, and 20 years, employing two general and fast modes with one and two colored noise components plus white noise. The computation time for estimating variance components was compared between the two modes, revealing a notable decrease in processing time with the fast mode.

Keywords Least squares variance component estimation (LS-VCE) · Multivariate LS-VCE · Fast LS-VCE · Variable transformation

Introduction

Proper analysis and subsequent interpretation of GNSS time series have proven to be valuable in the field of earth sciences and geophysics. Also, the time series analysis methods have been widely used by researchers across various disciplines including meteorology and ionosphere modeling (Amiri-Simkooei and Asgari 2012; Brack et al. 2021) and troposphere studies (Ma et al. 2021; Mirmohammadian et al. 2022; Xu et al. 2023). Daily GNSS position time series analysis, in particular, is employed in geodynamic and geophysical problems to estimate crustal deformation patterns (Montillet et al. 2015; He et al. 2017) and in seismology to aid in analyzing pre- and post-earthquake deformation dynamics (Moghtased-Azar and Grafarend 2009; Gualandi et al. 2017).

For accurate modeling of time series, their stochastic characteristic can be determined using variance component estimation (VCE) methods. Teunissen and Amiri-Simkooei (2008) introduced the least squares variance component estimation (LS-VCE) to estimate different noise components in a stochastic model in general and in GNSS position time series in particular. Since then, significant progress on the implementation and applications of LS-VCE to GNSS coordinate time series have been made (Khazraei et al. 2017; Moghtased-Azar et al. 2014; Amiri-Simkooei et al.

✉ Alireza Amiri-Simkooei
a.amirisimkooei@tudelft.nl

Saeed Mashhadizadeh-Maleki
Saeed.Mashhadizadeh@gmail.com

1 Department of Geomatics Engineering, Faculty of Civil Engineering and Transportation, University of Isfahan, Isfahan, Iran

2 Department of Control and Operations, Faculty of Aerospace Engineering, Delft University of Technology, Delft 2629 HS, The Netherlands

Published online: 18 November 2024

2007). LS-VCE, applied in its univariate and multivariate modes, facilitates realistic estimation of noise components (Amiri-Simkooei et al. 2019; Farzaneh et al. 2021). However, a persisting challenge concerning GNSS position time series analysis is the high computational burden for long time spans. This contribution is a step forward in tackling this challenge in GNSS time series analysis.

Various research studies have developed different approaches to estimate variance components from observations (1971), Sjoberg (1983), Koch (1986), Rao and Kleffe (1988), and Xu et al. (2007). In multiple studies focused on permanent GPS stations, researchers discovered that the time series of coordinates exhibit not only white noise but also flicker noise (FL). This finding was highlighted in studies conducted by Zhang et al. (1997), Mao et al. (1999), Nikolaidis et al. (2001), Williams et al. (2004), Amiri-Simkooei et al. (2007), and Teferle et al. (2008). Recent research suggests the presence of power-law noise, excluding only white noise, as highlighted by Amiri-Simkooei et al. (2007), Langbein (2008), King and Williams (2009), and Santamaria-Gómez et al. (2011).

The LS-VCE method is renowned for its flexibility in estimating variance components, offering various constraints such as the non-negativity of these components (Moghtased-Azar et al. 2014; Amiri-Simkooei 2016). Also, statistical tests like the w-test can be employed to validate the variance components estimated with LS-VCE (Amiri-Simkooei 2007). However, the calculation of variance components with LS-VCE has a high volume of calculations and complexity. For the first time, our innovative approach introduces a fast implementation that mitigates the above challenges. The method can be used for the stochastic models with diagonal(ized) cofactor matrices, applied to GNSS time series using a variable transformation technique. This will significantly enhance the method's efficiency and effectiveness in univariate and multivariate modes. With the proposed variable transformation, there is no hindrance in utilizing statistical tests nor non-negative LS-VCE to validate the outcomes. The key distinction lies in the quicker and simpler estimation process in both modes.

In the study of GPS coordinates using a stochastic model, it has been found that it is not simply a combination of white noise with colored noise. Recent research studies by He et al. (2019), Dmitrieva et al. (2015), Langbein (2012), Langbein and Svarc (2019) and Tehranchi et al. (2021) have suggested the presence of additional noise such as FL and RW. This research presents a variable transformation technique to sub-optimal estimation of these variance components. When comparing the sub-optimal estimated variance components with the unbiased estimated variance components, there is a minor difference. However, this difference can be acceptable considering the advantages it brings, such as reducing computational load, faster estimation of variance components, and the small discrepancy between the sub-optimal and unbiased estimated values.

The above development can also be linked to other techniques applied to the analysis of GNSS position time series. The maximum likelihood estimation (MLE) method is also a commonly used approach to estimate the variance components in GNSS time series. Many researchers, including Williams (2003) and Bos et al. (2008, 2013), have employed this method in their studies. Bos et al. (2013), by means of the Hector software, increased its calculation speed by a factor of 10 to 100, compared to the general MLE mode. The latest version of Hector, called 'hectorp', is based on the REstricted Maximum Likelihood (REML) principle, which was recommended by Gobron et al. (2022). Furthermore, Tehranchi et al. (2021) also introduced a method based on the REML principle that increased the implementation speed of variance components by 2 to 14 times compared to the Hector software.

The objectives of this article are as follows: (**1**) Enhancing the speed of calculations by reducing the computational load of the LS-VCE method in both univariate and multivariate modes. This advancement aims to facilitate the faster estimation of variance components for time series data with a large number of observations. (**2**) Streamlining the equations of the LS-VCE method and simplifying the inversion operation through the utilization of a variable transformation technique. This approach aims to simplify the overall methodology and improve its computational efficiency. (**3**) The LS-VCE method is known for its simplicity and flexibility. To maintain these properties, we offer a variable transformation approach that ensures certain conditions are preserved, including non-negative estimation of variance components. (**4**) In cases where the time series consists of multiple noise components, we use also use the above variable transformation solution. It is worth noting that in time series with a large amount of data, the difference between the sub-optimal estimation of variance components and their unbiased estimated values in univariate and multivariate modes can be disregarded. By accomplishing these objectives, this research intends to contribute to the optimization of the LS-VCE method, enabling more efficient and accurate analysis of time series data.

In the second section of this article, we will delve into the theory of how to improve the speed of calculation, and its impact on the estimation of univariate variance components and multivariate modes. Moving on to the third section, we will explore the application of variable transformation to the estimation equations of multivariate variance components. We will also present a comparison between the computational time and the changes in estimated components when the variable transformation is not applied. Finally, in the fourth section, we will conclude our discussion, emphasizing the benefits of this approach and exploring potential ideas to further enhance the methodology.

Fast implementation of LS-VCE

LS-VCE is a robust technique used to estimate the variance components of GNSS time series. It offers numerous advantages in this specific domain (Amiri-Simkooei 2009). We employ the following linear model of observation Eq.

$$\mathrm{E}(y) = Ax,\ \mathrm{D}(y) = Q_y = \sum\nolimits_{k=1}^{p} \sigma_k Q_k \tag{1}$$

where the E and D are the expectation and distribution operators, y is a vector of m observations, x is a vector of n unknown parameters, A the design matrix, is assumed to be full-column rank, and Q_y is the $m \times m$ covariance matrix of the observations y. In time series analysis, it is usually assumed that Q_y is not given and it is expressed as an unknown linear combination of some given co-factor matrices Q_k. The variance components are denoted by σ_k, $k = 1, 2, \ldots, p$, which can be estimated using LS-VCE as (Amiri-Simkooei 2007):

$$\hat{\sigma} = N^{-1}\ell \tag{2}$$

where N is a $p \times p$ normal matrix and ℓ is a p-vector, with the following entries

$$n_{ij} = \frac{1}{2}\mathrm{tr}\left(Q_y^{-1}P_A^{\perp} Q_i Q_y^{-1} P_A^{\perp} Q_j\right),\ i, j = 1, \ldots, p \tag{3}$$

and

$$\ell_i = \frac{1}{2}\hat{e}^T Q_y^{-1} Q_i Q_y^{-1}\hat{e} \tag{4}$$

with the least squares residuals

$$\hat{e} = P_A^{\perp} y \tag{5}$$

and the orthogonal projector

$$P_A^{\perp} = I - P_A = I - A\left(A^T Q_y^{-1} A\right)^{-1} A^T Q_y^{-1} \tag{6}$$

The matrix N is said to be 'normal' if it commutes with its conjugate transpose, i.e. if $NN^* = N^*N$. The conjugate transpose is simply the transpose for real matrices, so N should satisfy $NN^T = N^T N$, which is simply verified for the symmetric matrix N.

To compute the variance components, it is necessary to start with initial values and proceed with an iterative procedure until the convergence is achieved. This approach involves multiple matrix inversions and iterative calculations, which can result in significant computational requirements in terms of time and computational resources. Moreover, when dealing with a large number of data points, the dimensions of the matrix Q_y become excessively large, which escalates the computational burden.

Fast implementation of univariate LS-VCE

The main factor that significantly affects the calculation time of variance components is the inverse computation of the covariance matrix Q_y. There are various methods available to accelerate the inversion operation. One approach that improves the speed of inversion is the diagonalization of Q_y. Assuming that the cofactor matrices $(Q_k,\ k = 1, \ldots, p)$ are diagonal, the m-vector q_y (consisting of diagonal entries of Q_y) will have the following form:

$$Q_y = \sum\nolimits_{k=1}^{p} \sigma_k Q_k,\ \ q_y = \sum\nolimits_{k=1}^{p} \sigma_k q_k,\ \ Q_y = \mathrm{diag}(q_y) \tag{7}$$

which is inverted to

$$q_y^{-1} = \left(\frac{1}{q_y}\right),\ \ Q_y^{-1} = \mathrm{diag}\left(q_y^{-1}\right) \tag{8}$$

The simplification above will greatly facilitate the subsequent estimation of Variance Component Estimation (VCE). For example, for a diagonal covariance matrix, the estimated unknowns along with their covariance matrix can be obtained as: $\hat{x} = Q_{\hat{x}} T_1^T y$ and $Q_{\hat{x}} = \left(T_1^T A\right)^{-1}$, respectively, where $T_1 = A \odot \left(u_n^T \otimes q_y^{-1}\right)$, with $\odot$ and $\otimes$ are the Hadamard (element-wise) and Kronecker products, respectively, and $u_n = [1, \ldots, 1]^T$ the summation vector. To simplify the computation of Eq. (3), we denote $q_i^{-1} = q_i \odot q_y^{-1} = \mathrm{diag}(Q_i Q_y^{-1})$, $q_j^{-1} = q_j \odot q_y^{-1} = \mathrm{diag}(Q_j Q_y^{-1})$, and $q_{ij}^{-1} = q_i^{-1} \odot q_j^{-1} = \mathrm{diag}(Q_i Q_y^{-1} Q_j Q_y^{-1})$ (see Appendix)

$$n_{ij} = H_1 + 2H_2 + H_3 \tag{9}$$

where

$$\begin{cases} H_1 = \frac{1}{2}\mathrm{sum}\left(q_{ij}^{-1}\right) \\ H_2 = \frac{-1}{2}\mathrm{tr}\left(Q_{\hat{x}} T_1^T T_2\right) \\ H_3 = \frac{1}{2}\mathrm{tr}\left(Q_{\hat{x}} T_1^T T_4 Q_{\hat{x}} T_1^T T_3\right) \end{cases} \tag{10}$$

where $T_2 = A \odot \left(u_n^T \otimes q_{ij}^{-1}\right)$, $T_3 = A \odot \left(u_n^T \otimes q_j^{-1}\right)$ and $T_4 = A \odot \left(u_n^T \otimes q_i^{-1}\right)$.

The dimensions of matrices H_2 and H_3 before the trace operation are $n \times n$. Assuming that the matrix Q_y

is diagonal, the trace operation also occurs faster. This is because instead of performing the trace operation on an $m \times m$ matrix, it is performed on an $n \times n$ matrix. If we denote $\hat{e} = y - A\hat{x}$ and rewrite ℓ after applying the mentioned variable transformations, we can write it as follows:

$$\ell_i = \frac{1}{2}\mathrm{sum}(\hat{e} \odot T_5) \tag{11}$$

where q_{iy}^{-1} and T_5 are $q_y^{-1} \odot q_i \odot q_y^{-1}$, $\left(u_n^T \otimes q_{iy}^{-1}\right) \odot \hat{e}$, respectively.

Fast implementation of multivariate LS-VCE

Many studies have illustrated that GNSS time series exhibit temporal and spatial correlation. Taking into account these correlations enhances the functional model, which includes parameters such as offsets and harmonics, and subsequently improves the stochastic model (Gobron et al. 2022, 2024; Amiri-Simkooei and Asgari 2012; Amiri-Simkooei et al. 2019). Amiri-Simkooei (2009) presented the formulation of the multivariate variance components using the LS-VCE method. This method allows for the consideration of spatiotemporal correlation in time series when estimating the variance components using the least squares approach. If we express the model of observation equations and the stochastic noise model of the time series in the multivariate case, it can be written as follows:

$$\mathrm{E}\left(\mathrm{vec}\left(Y\right)\right) = \left(I_r \otimes A\right)\mathrm{vec}\left(X\right), \quad \mathrm{D}\left(\mathrm{vec}\left(Y\right)\right) = Q_{\mathrm{vec}(Y)} = \Sigma \otimes \mathrm{Q} \tag{12}$$

where

$$X = [x_1, \ldots, x_r]; \quad Y = [y_1, \ldots, y_r] \tag{13}$$

the Y and X represent the observation and unknown matrices, respectively, which are formed by aligning the observations and unknowns of r time series together, vec is the vector operator. The matrix A denotes the design matrix, which is considered to be identical for all time series, and the covariance matrix of all time series has a Kronecker structure as $\Sigma \otimes Q$. The $r \times r$ matrix Σ represents the spatial correlation among the time series. To express the aforementioned Eq. (12) in a practical form, $\mathrm{D}\left(\mathrm{vec}\left(Y\right)\right) = \Sigma \otimes \sum_{k=1}^{p} s_k Q_k$ will be employed, and the variance components will be estimated for p noise terms (Amiri-Simkooei 2009). The matrix Σ can be estimated as

$$\hat{\Sigma} = \frac{\hat{E}^T Q^{-1}\hat{E}}{m-n}, \quad \hat{E} = P_A^{\perp} Y \tag{14}$$

where $\hat{E}$ denotes the residual matrix, and m and n respectively indicate the number of observations and unknowns. Using initial values and iterative cycles, we can iteratively estimate the variance components as $\hat{s} = N^{-1}\ell$, where

$$n_{ij} = \frac{r}{2}\mathrm{tr}\left(Q^{-1}P_A^{\perp}Q_i Q^{-1}P_A^{\perp}Q_j\right) \tag{15}$$

and

$$\ell_i = \frac{1}{2}\mathrm{tr}(\hat{E}^T Q^{-1}Q_i Q^{-1}\hat{E}\Sigma^{-1}) \tag{16}$$

When calculating the normal matrix N in the multivariate case, we may use Eq. (15). If we assume a diagonal variance-covariance matrix Q, the equations can be simplified similarly to the univariate case. The only difference is that you will use $\frac{r}{2}$ instead of $\frac{1}{2}$ in the equations. Furthermore, when calculating the vector ℓ, if $T_6 = Q_y^{-1}\hat{E} = \left(u_n^T \otimes q_y^{-1}\right) \odot \hat{E}$, it will follow that

$$\ell_i = \frac{(m-n)}{2}\mathrm{tr}\left(\hat{E} \odot T_5(T_6 \odot \hat{E})^{-1}\right) \tag{17}$$

Similar to the univariate case, in estimating the variance components using the least squares method in the multivariate case, an iterative cycle is also used. Additionally, to calculate it according to Eq. (6), the matrix inversion Q_y^{-1} (in fact Q^{-1} for multivariate analysis) is performed multiple times. However, using the proposed diagonalization algorithm, the cofactor matrices (and hence Q) become diagonal. Therefore, multivariate GNSS analysis using LS-VCE can benefit from two sorts of simplifications: (i) The proposed Kronecker formulation of the covariance matrix, called 'more practical formulation' by Amiri-Simkooei (2016). In this case the covariance matrix structure $\mathrm{D}\left(\mathrm{vec}\left(Y\right)\right) = \Sigma \otimes Q = \Sigma \otimes \sum_{k=1}^{p} s_k Q_k$ can significantly reduce the computational load of the algorithm. (ii) The diagonalization algorithm to make the $m \times m$ matrix Q diagonal, can also substantially decrease the computational time (similar to univariate analysis).

Computational complexity

The speed of algorithm execution is determined by the complexity of calculations, which are typically based on mathematical equations related to the problem at hand. One effective strategy for improving algorithm efficiency is to simplify mathematical operations. To compute the computational complexity, we assume $m \gg n$, which is usually the case for time series analysis. One important operation in the implementation of VCE is the inversion of Q_y. This article

Table 1 Specifications of functional model components

	East (mm)	North(mm)
Intercept (a)	1.356	1.356
Linear rate $(b)\,(mm/year)$	1	1
Fourier component (annual) (c_1)	2	2
Fourier component (semi-annual)(c_2)	1	1
Fourier component (annual) (d_1)	3	3
Fourier component (semi-annual) (d_2)	2	2

Table 2 Specifications of stochastic model (noise) used in simulated time series

	East	North
WN amplitude (mm)	5	6
FL amplitude $\left(mm/year^{1/4}\right)$	15	18
RW amplitude $\left(mm/year^{1/2}\right)$	8	9.6

introduces a technique where instead of performing the inversion operation on a matrix with dimensions of $\times m$, the inversion of vector domains with the same dimensions is used, effectively translating the operation to a division. As a result, the complexity of computing the inversion of the Q_y reduces from $\mathcal{O}\left(m^3\right)$ to $\mathcal{O}\left(m\right)$. Another factor that contributes to the complexity of calculations is the multiplication of matrices with large dimensions. In case of the original formulation with full covariance matrices, the multiplications of $m\times\,m$ and $m\times\,n$ are needed, which introduce the complexity of $\mathcal{O}=\max\left(\mathcal{O}\left(m^2\right),\mathcal{O}\left(mn\right)\right)=\mathcal{O}(m^2)$. For the diagonal covariance matrices, the complexity is $\mathcal{O}=\max\left(\mathcal{O}\left(n^3\right),\mathcal{O}\left(mn\right)\right)=\mathcal{O}\left(mn\right)$.

Considering the above complexities, it follows that the complexity of original formulation is $\mathcal{O}\left(\mathrm{m}^3\right)$, whereas that with the diagonal covariance matrices is $\mathcal{O}\left(mn\right)$. As $m\gg\,n$, it follows that $\mathcal{O}\left(m^3\right)\gg\,\mathcal{O}\left(mn\right)$, indicating that the diagonal structure has a significantly lower computational burden compared to the original full covariance matrices. This will be validated using the numerical results.

Applications and results

In this article, we describe the implementation of a fast method for estimating variance components using least squares. As LS-VCE can provide negative variances, we used the non-negative variant of LS-VCE to guarantee non-negative variance component (Amiri-Simkooei 2016). We applied this method to both univariate and multivariate cases, and performed the analysis on 500 simulated time series with varying numbers of data points (5, 10, 15 and 20 years of daily data points). The computations and implementation were carried out using the MATLAB software. We compared the computation speed with and without variable transformation. To simulate the time series, we considered different stochastic models based on previous studies that examined the various noise components present in GNSS time series. The functional model used in our simulations consists of an intercept (a), a trend component (b), and the coefficients c_i and d_i of the periodic signals (Fourier components), as expressed in the following equation:

$$y\left(t\right)=\;a+bt+\sum\,{}_{i=1}^{2}c_i\cos\omega\,_it+d_i\sin\omega\,_it \tag{18}$$

where $\omega\,_1=2\pi$ and $\omega\,_2=4\pi$ are the angular frequencies of the annual and semi-annual signals, respectively and the time t is expressed in decimal years. To generate simulated time series, we utilized the specifications outlined in Table 1 for the functional model. These simulated time series were enhanced by incorporating both white and colored noise, the details of which can be found in Table 2. Our approach involved generating the simulated time series by combining the chosen stochastic model with one or more colored noise, in addition to the white noise, while adhering to the functional model's specifications. For illustration purposes, Fig. 1 presents a sample of daily simulated time series for the East and North regions, spanning 20 years of daily data points. This particular example was created by employing the specifications mentioned in Table 1 for the functional model, and incorporating colored noise components such as FL and RW, specified in Table 2.

Fast univariate LS-VCE with two noise components

Numerous studies have provided evidence that daily GNSS time series display both white noise, characterized by a spectral index of zero, and colored noise, known as FL, which has a spectral index of 1 (Williams et al. 2004; Amiri-Simkooei 2007; Davis et al. 2012). Additionally, researchers have noted the presence of power-law noise with a spectral index of 0.5, alongside white noise (Williams et al. 2004; He et al. 2019; Tehranchi et al. 2021). Notably, all of these studies introduce only one additional noise component alongside white noise and emphasize the importance of estimating the variances of these two components. In this section, we address this issue by considering a stochastic model that incorporates both white noise and FL noise.

$$Q_y=\sigma\,_w^2Q_w+\sigma\,_f^2Q_f \tag{19}$$

where $Q_w=I$ and Q_f are the cofactor matrices of white noise and FL noise, respectively, and $\sigma\,_w^2$ and $\sigma\,_f^2$ represents the variance components of those noise components. It is known that Q_y is a full matrix, as the cofactor matrix Q_f is full. This indicates that the original implementation of VCE on this problem is computationally expensive. One way to

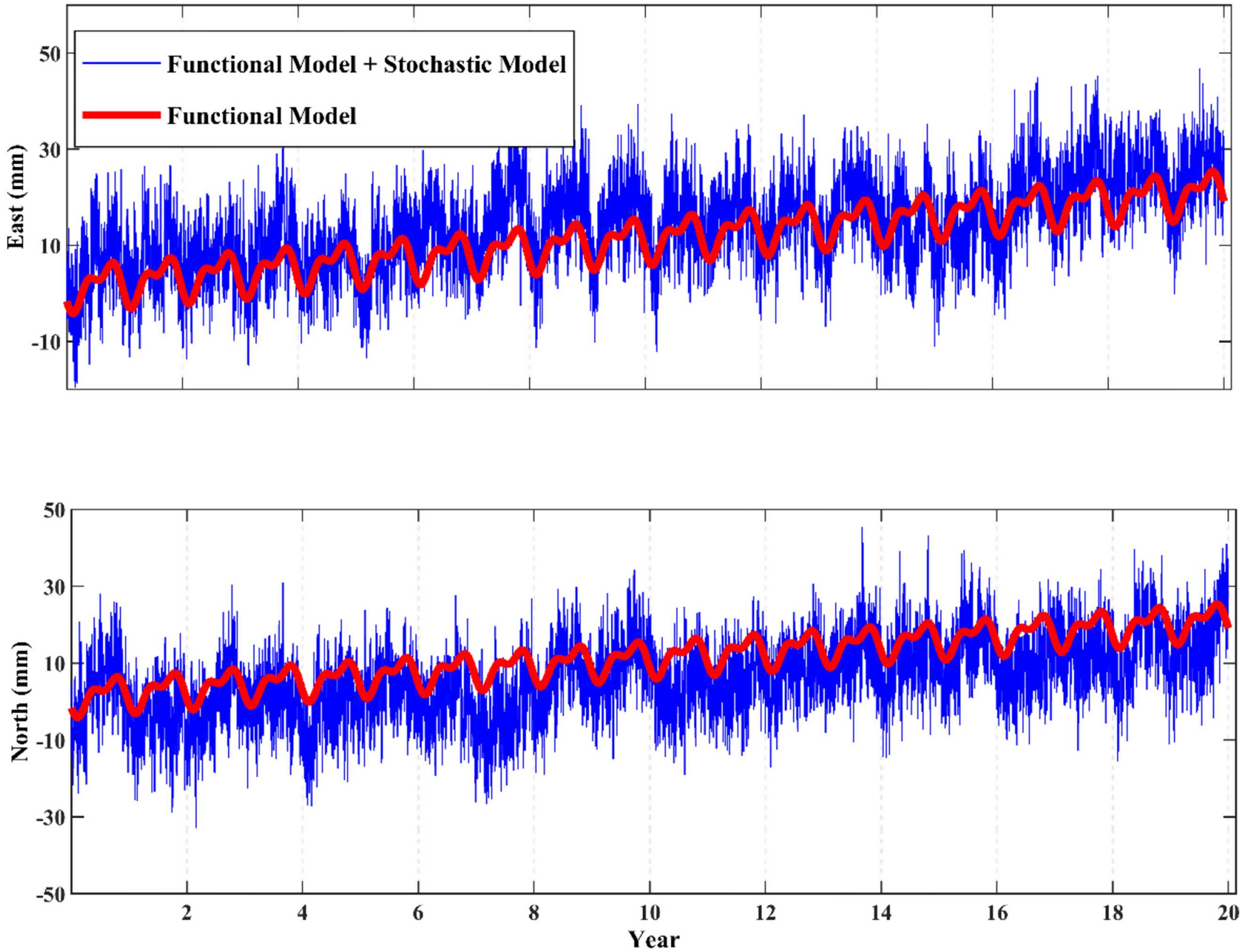

Fig. 1 Simulated daily time series of East (up) and North (down) with 20 year of data points, including colored noise FL and RW (blue line), along with the corresponding functional model time series (red line)

accelerate the estimation of variance components is to make the covariance matrix of observations diagonal. This is achieved by diagonalization of Q_f. In this regard, using the Singular Value Decomposition (SVD) operator, the observation covariance matrices can be diagonalized, which will be further explained as the variable transformation method.

One of the properties of SVD is that it decomposes a matrix into three parts. For instance, if we consider the symmetric matrix M, its SVD representation is $M = U\Lambda U^T$, where Λ is a diagonal matrix containing the eigenvalues of matrix M arranged in descending order along the diagonal. The matrices U represent the eigenvectors of matrix M. An important property of the matrix U is that it is a orthogonal matrix, indicating that $UU^T = U\,U^T = I$, indicating that $U^T = U^{-1}$. Therefore the SVD of the matrix Q_f is

$$Q_f = U_f \Lambda_f U_f^T \tag{20}$$

This with Eq. (19) gives the following linear model of observation equations:

$$\mathrm{E}(y) = Ax, \quad Q_y = \sigma_w^2 I + \sigma_f^2 U_f \Lambda_f U_f^T \tag{21}$$

We now apply an admissible non-singular transformation of the observations y as

$$y_f = U_f^T y \Leftrightarrow y = U_f y_f \tag{22}$$

with the covariance matrix $Q_{y_f} = U_f^T Q_y U_f$. This admissible transformation changes the basis of y using SVD, which allows a one-to-one mapping, so it is possible to revert to the original observations y if required. After a few simple mathematical operations, Eq. (21) can be represented as the following equivalent form:

$$\mathrm{E}(y_f) = A_f x, \quad Q_{y_f} = \sigma_w^2 I + \sigma_f^2 \Lambda_f \tag{23}$$

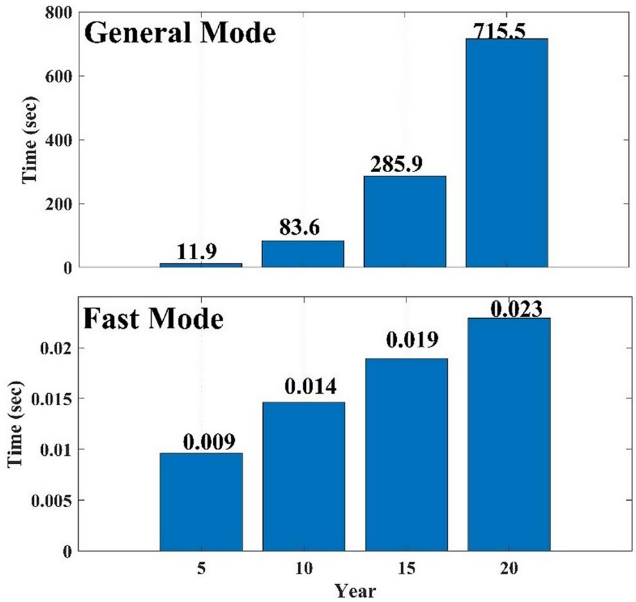

Fig. 2 Mean computational time to estimate variance components using the general non-negative univariate LS-VCE method compared to its equivalent fast mode

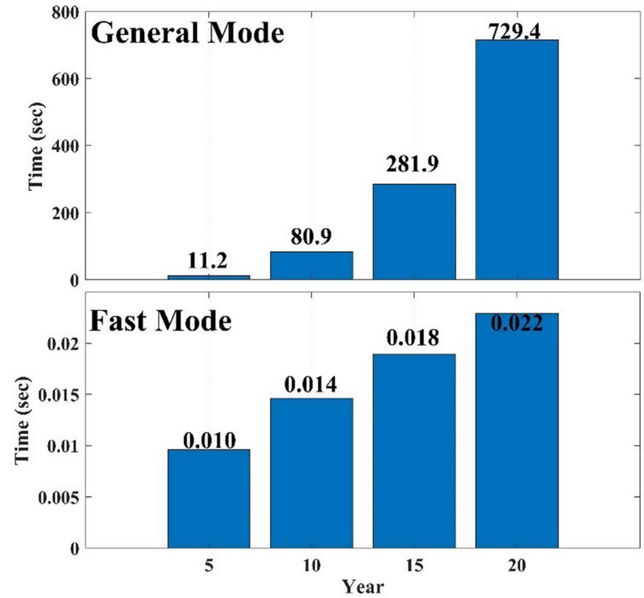

Fig. 3 Computational time to estimate two variance components using the general non-negative multivariate LS-VCE method compared to the fast mode

which has a diagonal covariance matrix as both I and Λ_f are diagonal. We also note that the estimate $\widehat{x}$ and its covariance matrix $Q_{\widehat{x}}$ are invariant of under non-singular transformation. This in particular has no effect on the site velocity and its uncertainty in the GNSS time series analysis. The formulation of the fast implementation of LS-VCE, presented in Sect. 2, can be applied to estimate the two variance components σ_w^2 and σ_f^2.

In the context of observation equations, if we assume that there are only two sources of noise in the stochastic model (like white plus FL noise), the calculation of variance components does not require any bias. Additionally, applying the above transformation does not introduce any changes or alterations to the estimated values of the variance components. Therefore, the variance components remain equal to their respective values in the general case, regardless of the specific variable transformation applied.

We conducted VCE using the least squares method in two scenarios: fast and general. This analysis involved 500 simulated time series, with the corresponding values listed in Tables 1 and 2 for the functional and stochastic models. These models solely consisted of white noise and FL noise components. The estimation of variance components was performed on time series of varying lengths of 5, 10, 15 and 20 years. The calculations were carried out on a unified system with the Intel Core i3-2100 specifications.

Figure 2 demonstrates that as the length of time series increases, the computation time increases in both cases. However, the rate of increase in the computation time of the variance components in the general mode is significantly higher than in the fast mode. Additionally, in the time series with lengths of 5, 10, 15, and 20 years of daily data, the fast method is approximately 1200, 6000, 15,000, and 31,000 times faster than the general mode.

Fast multivariate LS-VCE with two noise components

The multivariate least squares method is recommended for estimating variance components due to its simultaneous consideration of spatial and temporal correlation in computations. This method can identify the types of variance components with greater power and provide more accurate estimates of their values (Amiri-Simkooei 2009).

Using the values provided in Tables 1, 2, 500 daily time series for GNSS coordinates in East and North components were simulated, ranging from 5, 10, 15, to 20 years. These simulations incorporated a white noise plus FL noise stochastic model. The matrix of observations, for which variance components will be calculated, is equivalent to $Y = [East, \; North]$. The variance components were estimated using the multivariate least squares method in the two modes, general and fast. The time taken to compute the variance components was also accounted for. In Fig. 3, it is evident that the calculation time increases as the number of data points rises in both modes. Moreover, as the number of time series data points increases, the disparity in quick and overall calculation times becomes more pronounced. This discrepancy highlights the effectiveness of the diagonalization method introduced in this article, which allows for efficient and rapid calculation of variance components in time series data with a high volume of data points. By implementing this suggested fast mode, the issue of high calculation burden in the general mode is effectively addressed.

Fast univariate LS-VCE with three noise components

Numerous researchers investigating GNSS time series have found that the underlying stochastic model should consist of a linear combination of white noise and two-colored noise components. Some studies have also noted the presence of RW noise alongside white noise and FL noise in the time series of permanent GNSS stations (Tehranchi et al. 2021). Moreover, additional research has highlighted the existence of RW in conjunction with FL noise (He et al. 2019). The consistent theme across these studies is the consideration of a stochastic model comprising three distinct types of noise. To further explore this issue and assess the efficacy of this sub-optimal estimation, we conducted simulations using time series that combined white noise, FL and RW noise. These time series were simulated and computed 500 times. Our analysis included comparisons of calculation speed and the sub-optimal estimation of variance components against their unbiased estimated values.

If we define the number of noise components in Eq. (1) as three, we will obtain:

$$Q_y = \sigma_w^2 Q_w + \sigma_f^2 Q_f + \sigma_R^2 Q_R \tag{24}$$

where Q_R represents cofactor matrix of RW, while σ_R represents its variance component. If we apply the transformation SVD to Eq. (22), we will have $Q_y = \sigma_w^2 I + \sigma_f^2 \Lambda_f + \sigma_R^2 U_f^T U_R \Lambda_R U_R^T U_f$. Since U_R and U_f are distinct, $U_R^T U_f$ is not expected to be exactly $I_{m\times m}$. However, our numerical results show that $U_R^T U_f$ can be well approximated by an identity matrix $I_{m\times m}$. Consequently, when the number of noise components is 3, their LS estimates can be approximated. By employing this assumption, Eqs. (10) and (11) can be utilized for the univariate mode to expedite the estimate of the variance components.

As with the two-noise mode, by applying the above transformation method, we observed an increase in the speed of calculations. However, the variance component values were bias estimated, leading to discrepancies between the estimated and actual values, as depicted in Fig. 4. This discrepancy was generated by the noise type across 500 time series. As illustrated in Fig. 4, the difference between the bias estimate values obtained through variable transformation diminishes as the length of the time series increases.

Figure 4 shows the dispersion of the difference between the values of the variance components and the real value simulated using the univariate non-negative-LSVCE method, which decreases with the increase in the number of data, and the Root Mean Square (RMS) of the difference has also been calculated for each noise component. As the number of data increases, the RMS values of the bias estimated components also decrease.

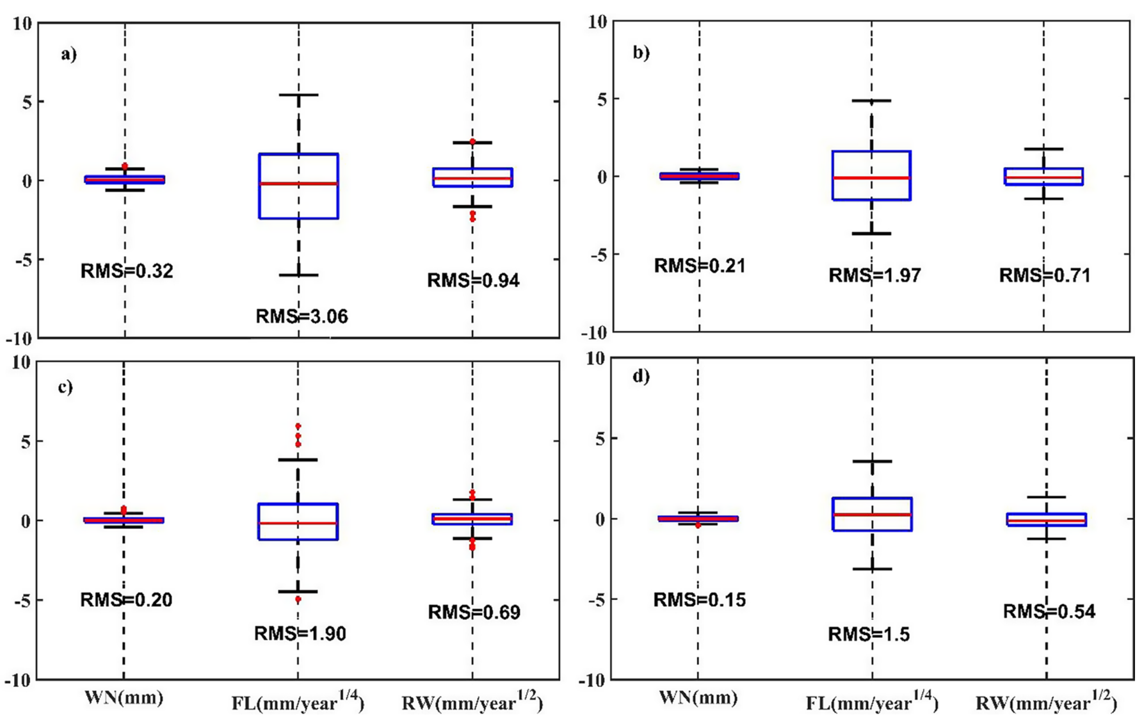

Fig. 4 Boxplot of the discrepancy between the bias estimated variance components with fast non-negative univariate LS-VCE and the actual values for four scenarios, with time series lengths of 5 years (**a**), 10 years (**b**), 15 years (**c**), and 20 years (**d**)

Fast multivariate LS-VCE with three noise components

In Section "Fast univariate LS-VCE with three noise components", we discussed the three-noise component model in GNSS time series. Section "Fast multivariate LS-VCE with two noise components" introduced multivariate mode variable transformation equations, which are utilized in the explanations provided in Sect. "ast multivariate LS-VCE with three noise components". In this section, we analyze the variance components of simulated time series featuring the three- noise model through the above diagonalization method. We then assess the variance component differences from the estimated values and compare the computational efficiency.

Figure 5 illustrates the comparison of calculation speeds for variance components in both univariate and multivariate modes. The figure also allows for a comparison between the fast and general modes in terms of the duration required for calculating variance components. The data presented indicates that as the number of noise components increases, there is a notable increase in the calculation time. However, in the fast mode, the calculation time remains relatively stable regardless of the number of noise components. This suggests that by employing variable transformation, equations can be simplified, and the time needed to calculate variance components does not experience a significant increase even if the number of noise components increases.

Figure 5 also illustrates that as the length of the time series increases in general mode, the amount of the calculation time is higher compared to the fast mode. This trend is evident in both univariate and multivariate modes. If we want to determine the computational burden to estimate variance components of 500 time series, we must sum-up the calculation duration in univariate mode 500 times. In the multivariate mode, the variance components are estimated in one run. As shown in Fig. 5, the calculation time for the variance components in the multivariate mode is much faster than in the univariate mode.

According to Fig. 6 (similarly in Fig. 4), as the dataset size grows, the variance between the estimated and genuine values diminishes. Moreover, it is noticeable that the RMS values of the variance components between the estimated and actual values decrease as the dataset size escalates. When comparing the RMS of univariate and multivariate modes, it is evident that the multivariate model provides a more accurate estimation than the univariate model across all data sets.

Discussion

Since its introduction by Teunissen and Amiri-Simkooei in 2008, the LS-VCE method has undergone numerous applications and enhancements by various researchers. Some shortcomings of the original method include the lack of incorporation of temporal and spatial correlations in time series analysis and the challenge of accurately estimating

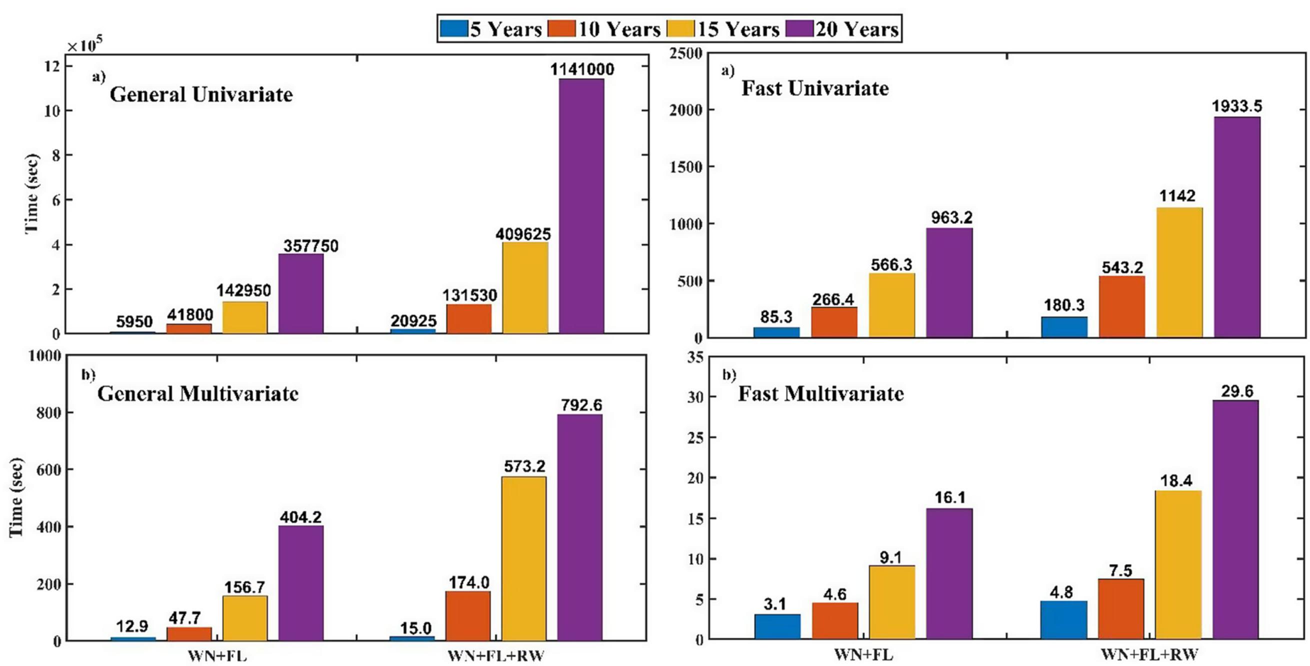

Fig. 5 Comparison of sum of the computational time on calculating variance components of 500 time series in the presence of two or three noise components using univariate LS-VCE and fast non-negative univariate LS-VCE modes (**a**), multivariate LS-VCE and fast multivariate LS-VCE modes (**b**)

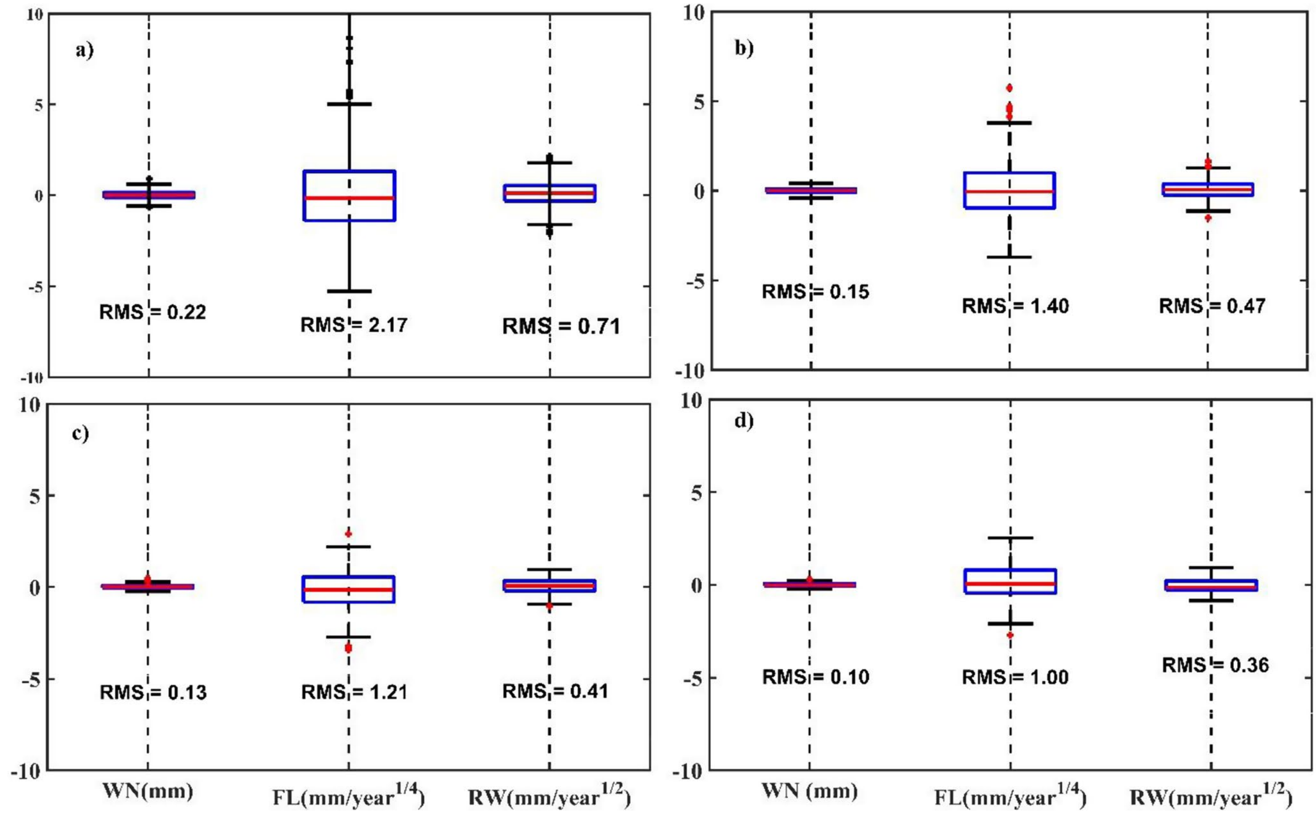

Fig. 6 Boxplot of the discrepancy between the bias estimated variance components with fast non-negative multivariate LS-VCE and the actual values for four scenarios, with time series lengths of 5 years (**a**), 10 years (**b**), 15 years (**c**), and 20 years (**d**)

variance components. Additionally, the method is criticized for its computational intensity and time-consuming nature. Over the years, significant progress has been made in addressing the deficiencies related to variance component estimation and incorporating temporal and spatial correlations in time series analysis, as detailed in the literature. However, the computational burden and time requirements of the method remain areas of concern that warrant further attention.

With the innovative approach described in this article, the issue of time-consuming calculation of variance components has been substantially addressed, although not entirely resolved. One key assumption in this study is that the white noise cofactor matrix is identity across different scenarios, leading to the estimation of variance components for white noise and FL that align with the general mode. If the white noise cofactor matrix is variable, the values of the variance components in the two-noise scenario are bias estimated (sub-optimal). This bias has been shown not to be significant for GNSS time series analysis. Solving this problem is one of the problems that dear readers can think about and solve. Also, another problem that can be solved in the future is that in the second scenario (three-noise scenario) the variance components are unbiased estimate instead of bias estimation.

Conclusions

Trend estimation in GPS time series plays a crucial role in geodynamic and geological studies. It is well-known that GPS time series are often affected by colored noise, which significantly impact the trend estimation and its uncertainty. To address this, VCE methods can be employed. The non-negative-LSVCE method stands out as a top choice for accurately identifying, diagnosing, and estimating the variance components of these stochastic models. Additionally, this method allows for simultaneous estimation of both the functional model and the random model.

In this study, we introduced an innovative approach to speed up the LS-VCE method for estimating variance components. Our proposed diagonalization method streamlines calculations, simplifies equations, and decreases the computational load, thereby promoting wider adoption of this method for VCE. The utilization of fast LS-VCE enables expedited processing of long-term time series data, facilitating the estimation of variance components in GNSS time series. This advancement contributes to a deeper comprehension of the noise characteristics within such time series and enhances the precision of component estimation over time.

We conducted a comparison of simulated GNSS time series under two scenarios: (a) with the presence of colored noise alongside white noise and (b) with the presence of two-colored noise alongside white noise, focusing on the computation speed of variance components in general and fast modes. Notably, we observed a significant acceleration in speed when utilizing the fast mode. Additionally, our analysis of the impact of increased number of noise led us to the conclusion that implementing the suggested variable transformation greatly streamlines and expedites the calculation of variance components through the LS-VCE method. It is therefore highly recommended to employ the LS-VCE method with the proposed variable transformation for estimating or approximating the variance components due to its superior detection capabilities compared to alternative methods. Furthermore, statistical tests on the estimated variance components validate the effectiveness of this approach, with no adverse effects on the statistical analysis process introduced by the proposed variable transformation.

In GNSS time series analysis, to estimate variance components, we suggest to use multivariate mode with non-negative LS-VCE. This fast implementation of LS-VCE can benefit from the possible diagonal(ized) structures of the cofactor matrices in the stochastic model. By comparing Figs. 4 and 6, we observe that this mode has a lower RMS and dispersion compared to the univariate mode. In scenario b, the bias induced due to increasing the number of noise components (from 2 to 3) becomes negligible for long time series.

Appendix A

Rewriting the equations.

In order to reach the equations resulting from variable transformation, we have to Substituting Eq. (6) into Eq. (3), for $i, j = 1, \ldots, p$ we will have:

$$n_{ij} = \frac{1}{2}\mathrm{tr}(Q_y^{-1}\left(I - A\left(A^TQ_y^{-1}A\right)^{-1}A^TQ_y^{-1}\right) Q_iQ_y^{-1}\left(I - A\left(A^TQ_y^{-1}A\right)^{-1}A^TQ_y^{-1}\right)Q_j) \tag{25}$$

when Eq. (22) is expanded:

$$\begin{aligned} n_{ij} = \frac{1}{2}\mathrm{tr}(&Q_y^{-1}Q_iQ_y^{-1}Q_j - Q_y^{-1}A\left(A^TQ_y^{-1}A\right)^{-1}A^TQ_y^{-1}Q_iQ_y^{-1}Q_j \\ &- Q_y^{-1}Q_iQ_y^{-1}A\left(A^TQ_y^{-1}A\right)^{-1}A^TQ_y^{-1}Q_j \\ &+ Q_y^{-1}A\left(A^TQ_y^{-1}A\right)^{-1}A^TQ_y^{-1}Q_iQ_y^{-1}A\left(A^TQ_y^{-1}A\right)^{-1}A^TQ_y^{-1}Q_j) \end{aligned} \tag{26}$$

and using the properties of the trace operator $\left(\mathrm{tr}\left(AB\right) = \mathrm{tr}\left(B^TA\right)\right)$, we will have:

$$n_{ij} = \frac{1}{2}\begin{bmatrix} \mathrm{tr}\left(Q_y^{-1}Q_iQ_y^{-1}Q_j\right) \\ -2\mathrm{tr}\left(Q_y^{-1}Q_iQ_y^{-1}A\left(A^TQ_y^{-1}A\right)^{-1}A^TQ_y^{-1}Q_j\right) + \\ + \\ \mathrm{tr}\left(Q_y^{-1}A\left(A^TQ_y^{-1}A\right)^{-1}A^TQ_y^{-1}Q_iQ_y^{-1}A\left(A^TQ_y^{-1}A\right)^{-1}A^TQ_y^{-1}Q_j\right) \end{bmatrix} \tag{27}$$

due to the fact that the diagonal Q_y is assumed and to simplify Eq. (24), we use $T_1, T_2, T_3$and T_4 variables and Hadamard multiplication, which are obtained as a result of Eqs. (9) and (10).

To rewrite and simplify Eq. (4), We used the feature of diagonally of the Q_y and Hadamard product and took advantage of the change of the variable $T_5 = \left(u_n^T \otimes q_{iy}^{-1}\right) \odot \widehat{e}$ and $q_{iy}^{-1} = q_y^{-1} \odot q_i \odot q_y^{-1}$to simplify the Eq. (4):

$$\ell_i = \frac{1}{2}\widehat{e}^TQ_y^{-1}Q_iQ_y^{-1}\widehat{e} = \frac{1}{2}\mathrm{sum}(\widehat{e} \odot \ T_5) \tag{28}$$

Author contributions S.M-M performed experiments, analyzed data and carried out the implementation and prepared figures; AR. A-S. is the supervisor and wrote the manuscript;

Data availability No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

References

Amiri-Simkooei A (2007) Least-squares variance component estimation: theory and GPS applications

Amiri-Simkooei AR (2009) Noise in multivariate GPS position time-series. J Geodesy 83:175–187

Amiri-Simkooei AR (2016) Non-negative least-squares variance component estimation with application to GPS time series. J Geodesy 90:451–466

Amiri-Simkooei AR, Asgari J (2012) Harmonic analysis of total electron contents time series: methodology and results. GPS Solutions 16:77–88

Amiri-Simkooei AR, Tiberius CC, Teunissen PJ (2007) Assessment of noise in GPS coordinate time series: methodology and results. J Geophys Res Solid Earth 112:B7

Amiri-Simkooei AR, Hosseini-Asl M, Asgari J, Zangeneh-Nejad F (2019) Offset detection in GPS position time series using multivariate analysis. GPS Solutions 23:1–12

Bos MS, Fernandes RMS, Williams SDP, Bastos L (2008) Fast error analysis of continuous GPS observations. J Geodesy 82(3):157–166

Bos MS, Fernandes RMS, Williams SDP, Bastos L (2013) Fast error analysis of continuous GNSS observations with missing data. J Geodesy 87(4):351–360

Brack A, Männel B, Wickert J, Schuh H (2021) Operational multi-GNSS global ionosphere maps at GFZ derived from uncombined code and phase observations. Radio Sci 56(10):1–14

Davis JL, Wernicke BP, Tamisiea ME (2012) On seasonal signals in geodetic time series. J Geophys Res Solid Earth 117:B1
Dmitrieva K, Segall P, DeMets C (2015) Network-based estimation of time-dependent noise in GPS position time series. J Geodesy 89(6):591–606
Farzaneh S, Safari A, Parvazi K (2021) Improving dam deformation analysis using least-squares variance component estimation and Tikhonov regularization. J Surv Eng 147(1):04020024
Gobron K, Rebischung P, de Viron O, Demoulin A, Van Camp M (2022) Impact of offsets on assessing the low-frequency stochastic properties of geodetic time series. J Geodesy 96(7):46
Gobron K, Rebischung P, Chanard K, Altamimi Z (2024) Anatomy of the spatiotemporally correlated noise in GNSS station position time series. J Geodesy 98(5):34
Gualandi A, Avouac JP, Galetzka J, Genrich JF, Blewitt G, Adhikari LB, Liu-Zeng J (2017) Pre-and post-seismic deformation related to the 2015, Mw7. 8 Gorkha earthquake. Nepal Tectonophysics 714:90–106
He X, Montillet JP, Fernandes R, Bos M, Yu K, Hua X, Jiang W (2017) Review of current GPS methodologies for producing accurate time series and their error sources. J Geodyn 106:12–29
He X, Bos MS, Montillet JP, Fernandes RMS (2019) Investigation of the noise properties at low frequencies in long GNSS time series. J Geodesy 93(9):1271–1282
Khazraei SM, Nafisi V, Amiri-Simkooei AR, Asgari J (2017) Combination of GPS and leveling observations and geoid models using least-squares variance component estimation. J Surv Eng 143(2):04016023
King MA, Williams SD (2009) Apparent stability of GPS monumentation from short-baseline time series. J Geophys Research: Solid Earth 114:B10
Koch KR (1986) Maximum likelihood estimate of variance components: ideas by AJ Pope (in memory of Allen J. Pope, 11.10. Bull Gæodésique 60:329–338 1939–29.08. 1985
Langbein J (2008) Noise in GPS displacement measurements from Southern California and Southern Nevada. J Geophys Research: Solid Earth 113:B5
Langbein J (2012) Estimating rate uncertainty with maximum likelihood: differences between power-law and flicker–random-walk models. J Geodesy 86(9):775–783
Langbein J, Svarc JL (2019) Evaluation of temporally correlated noise in global navigation satellite system time series: geodetic monument performance. J Geophys Res Solid Earth 124(1):925–942
Mao A, Harrison CG, Dixon TH (1999) Noise in GPS coordinate time series. J Geophys Re Solid Earth 104(B2):2797–2816
Mirmohammadian F, Asgari J, Verhagen S, Amiri-Simkooei A (2022) Multi-GNSS-weighted interpolated tropospheric delay to improve long-baseline RTK positioning. Sensors 22(15):5570
Moghtased-Azar K, Grafarend EW (2009) Surface deformation analysis of dense GPS networks based on intrinsic geometry: deterministic and stochastic aspects. J Geodesy 83(5):431–454
Moghtased-Azar K, Tehranchi R, Amiri-Simkooei AR (2014) An alternative method for non-negative estimation of variance components. J Geodesy 88:427–439
Montillet JP, Williams SDP, Koulali A, McClusky SC (2015) Estimation of offsets in GPS time-series and application to the detection of earthquake deformation in the far-field. Geophys J Int 200(2):1207–1221
Nikolaidis RM, Bock Y, de Jonge PJ, Shearer P, Agnew DC, Van Domselaar M (2001) Seismic wave observations with the global positioning system. J Geophys Res Solid Earth 106(B10):21897–21916
Rao CR, Kleffe J (1988) Estimation of variance components and applications
Santamaria-Gómez A, Bouin MN, Collilieux X, Wöppelmann G (2011) Correlated errors in GPS position time series: implications for velocity estimates. J Geophys Research: Solid Earth 116:B1
Sjoberg LE (1983) Unbiased estimation of variance-components in condition adjustment with unknowns-a MINQUE approach. ZfV 108:9
Teferle FN, Williams SD, Kierulf HP, Bingley RM, Plag HP (2008) A continuous GPS coordinate time series analysis strategy for high-accuracy vertical land movements. Phys Chem Earth Parts A/B/C 33(3–4):205–216
Tehranchi R, Moghtased-Azar K, Safari A (2021) Fast approximation algorithm to noise components estimation in long-term GPS coordinate time series. J Geodesy 95(2):18
Teunissen PJ, Amiri-Simkooei AR (2008) Least-squares variance component estimation. J Geodesy 82:65–82
Williams SDP (2003) The effect of coloured noise on the uncertainties of rates estimated from geodetic time series. J Geodesy 76:483–494
Williams SD, Bock Y, Fang P, Jamason P, Nikolaidis RM, Prawirodirdjo L, Johnson DJ (2004) Error analysis of continuous GPS position time series. J Geophys Res Solid Earth 109:B3
Xu P, Liu Y, Shen Y, Fukuda Y (2007) Estimability analysis of variance and covariance components. J Geodesy 81:593–602
Xu C, Zhu Y, Xu X, Kong J, Yao Y, Shi J, Li X (2023) Regional tropospheric correction model from GNSS–Saastamoinen–GPT2w data for Zhejiang Province. Atmosphere 14(5):815
Zhang J, Bock Y, Johnson H, Fang P, Williams S, Genrich J, Behr J (1997) Southern California Permanent GPS Geodetic Array: error analysis of daily position estimates and site velocities. J Geophys Res Solid Earth 102(B8):18035–18055