Acta Geodyn. Geomater., Vol. 22, No. 4 (220), 531–544, 2025
DOI: 10.13168/AGG.2025.0038
journal homepage: https://www.irsm.cas.cz/acta

ORIGINAL PAPER

SOME NEW INSIGHTS INTO EFFICIENT IMPROVED SINGULAR SPECTRUM ANALYSIS FOR PROCESSING INCOMPLETE GEODETIC TIME SERIES

Kunpu JI and Fengwei WANG *

College of Surveying and Geo-informatics, Tongji University, Shanghai 200092, P.R. China

**Corresponding author's e-mail:* wangfw-foster@tongji.edu.cn

ARTICLE INFO

Article history:
Received 7 August 2025
Accepted 25 November 2025
Available online 9 December2025

Keywords:
Geodetic time series analysis
Singular spectrum analysis
Missing data

ABSTRACT

Improved Singular Spectrum Analysis (ISSA) (Shen et al., 2015) and its computationally optimized variant, efficient ISSA (Ji et al., 2025), are advanced methods for analyzing incomplete time series without interpolation. Previous studies numerically demonstrated that both methods yield equivalent principal component (PC) estimates under a minimum weighted norm criterion in the spectral domain and that efficient ISSA is computationally more efficient, but no formal theoretical proof was provided. This study fills that gap by rigorously proving the equivalence between ISSA and efficient ISSA and presenting a comprehensive analysis of their computational complexity. For each iteration i ($1 \le i \le N - L + 1$, where N is the time series length and L is the window size), efficient ISSA reduces the per-iteration complexity from $O(L^3)$ (ISSA) to $O\left(N_{\bar{S}_i}^3 + LN_{\bar{S}_i}^2 + L^2N_{\bar{S}_i}\right)$, where $N_{\bar{S}_i} < L$ is the number of missing values in the i th interval $[i, i + L - 1]$; under sparse missing data ($N_{\bar{S}_i} \ll L$), the complexity further simplifies to $O(L^2)$. Efficient ISSA is also compared with extended SSA (ESSA) (Ji et al., 2023) in terms of computation time, PC estimation, and signal extraction using both synthetic and real GNSS position time series. Results show that efficient ISSA runs much faster and produces more accurate PC estimates, while ESSA remains slightly superior in signal extraction. When applied iteratively, both methods achieve substantially improved filtering performance, with efficient ISSA delivering comparable reconstruction quality at a lower computational cost. The proposed method is further demonstrated to effectively bridge data gaps in incomplete time series, as illustrated using Gravity Recovery and Climate Experiment (GRACE) and GRACE Follow-On gravity field solutions.

1. INTRODUCTION

Geodetic time series contain abundant signals related to various geophysical processes (Davis et al., 2012; Ming et al., 2016; Kaczmarek and Kontny, 2018; Klos et al., 2021, 2023; Xiang et al., 2022; Huang et al., 2023; Xu et al., 2025). However, these signals are contaminated by different types of noise (Bos et al., 2013; Ming et al., 2017; Li and Shen, 2018; He et al., 2019; Klos et al., 2019; Qiu et al., 2022; Cucci et al., 2023; Huang et al., 2023; Gobron et al., 2024). Singular Spectrum Analysis (SSA) is a widely used, data-driven technique for extracting signals from noisy time series (Vautard and Ghil, 1989; Golyandina and Zhigljavsky, 2013). As a powerful non-parametric method that requires no prior information, SSA decomposes a time series into a sum of interpretable components, including a slowly varying trend, oscillatory patterns, and 'structureless' noise. Owing to these advantages, SSA has been broadly applied to analyze geodetic and geophysical time series (Chen et al., 2013; Xu and Yue, 2015; Xu, 2016; Gruszczynska et al., 2016, 2017; Guo et al., 2018; Klos et al., 2018; Yi and Sneeuw, 2021, 2022; Wang et al., 2018).

However, the presence of missing data limits SSA's applicability in analyzing geodetic time series. To cope with this problem, Schoellhamer (2001) proposed a modified version of SSA, known as SSAM, which directly computes the principal components (PCs) using the available observation data. Although SSAM is efficient, it is empirical in computing PCs. To improve upon this, Shen et al. (2015) developed an improved SSA method called ISSA, which estimates PCs by solving a linear rank-deficient system based on the reconstruction relationship between the missing data and PCs. It has been proved that SSAM is a simplified form of ISSA. While ISSA is more robust than SSAM, its computational cost scales cubically with the window size, which hinders its applications in large-scale data analysis. To improve efficiency, Ji et al. (2023) developed extended SSA (ESSA), which is faster than ISSA but focuses on signal reconstruction through low-rank approximation. Recently, Ji et al. (2025) proposed an efficient ISSA method, which can greatly improve computational efficiency as few parameters are estimated, while maintaining the computed PCs unchanged as the same criterion is used. However, their work relied on numerical experiments rather than formal mathematical proof to assert equivalence between the two methods. Furthermore, a rigorous analysis of time complexity critical for evaluating scalability and practical utility was absent.

This study aims to bridge these theoretical gaps by providing a formal mathematical proof of the

Cite this article as: Ji K, Wang F: Some new insights into efficient improved singular spectrum analysis for processing incomplete geodetic time series. Acta Geodyn. Geomater., 22, No. 4 (220), 531–544, 2025. DOI: 10.13168/AGG.2025.0038

equivalence between ISSA and efficient ISSA, and conducting a comprehensive time complexity analysis of both methods to demonstrate the computational advantage of efficient ISSA theoretically. Additionally, we will conduct a comprehensive comparison between efficient ISSA with ESSA in computational efficiency and computed results including PC estimates (frequency domain) and filtered results (time domain). The remainder of this study is structured as follows: Section 2 briefly reviews the ISSA and efficient ISSA. Section 3 presents the proof of equivalence. Section 4 provides a time complexity analysis showing the computational efficiency of the new method. Section 5 compares efficient ISSA with ESSA in terms of computational cost, PC estimates, and filtered signals using Global Navigation Satellite System (GNSS) position time series. Section 6 applies the efficient ISSA to bridge the data gaps between the Gravity Recovery and Climate Experiment (GRACE) and GRACE Follow-On (GRACE-FO) gravity field solutions. Conclusions are summarized in Section 7.

2. BRIEF REVIEW OF ISSA AND EFFICIENT ISSA

Consider a time series $\boldsymbol{x} = \{x_i \mid 1 \le i \le N\}$ with missing values at certain epochs. Let S and $\overline{S}$ denote the index sets of observed and missing epochs, respectively, satisfying $S \cup \overline{S} = [1, N], S \cap \overline{S} = \emptyset$. The observed data and missing data are represented as $\boldsymbol{x}_1 = \{x_i \mid i \in S\}$ and $\boldsymbol{x}_2 = \{x_i \mid i \in \overline{S}\}$, respectively. Both two methods (i.e., ISSA and efficient ISSA) begin by calculating the Toeplitz lagged covariance matrix $\boldsymbol{C}$ using the observed data $\boldsymbol{x}_1$ (Shen et al., 2015),

$$c(j) = \frac{1}{N_j}\sum_{i \le N-j} x_i x_{i+j} \quad j = 0, 1, 2, \cdots, L-1 \tag{1}$$

where L is the window size, both x_i and x_{i+j} must be observed rather than missed, and N_j represents the number of valid products $x_i x_{i+j}$ within the sample index $i \le N - j$. Then we decompose $\boldsymbol{C} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}}$, where $\boldsymbol{V}$ is an orthogonal matrix satisfying $\boldsymbol{V}\boldsymbol{V}^{\mathrm{T}} = \boldsymbol{V}^{\mathrm{T}}\boldsymbol{V} = \boldsymbol{I}_L$, and $\boldsymbol{\Lambda}$ is a diagonal matrix whose entries are the singular values λ_i $(1 \le i \le L)$ sorting in descending order.

Both two methods require repeatedly solving a specific equation to calculate PCs. In the i-th interval, Shen et al. (2015) establish the following linear system to solve the principal components $\boldsymbol{\xi}_i$ from the non-missing data,

$$\boldsymbol{G}_i\boldsymbol{\xi}_i = \boldsymbol{y}_i \tag{2}$$

where $\boldsymbol{G}_i$, $\boldsymbol{\xi}_i$ and $\boldsymbol{y}_i$ are defined as,

$$\begin{aligned}
\boldsymbol{G}_i(m,n) &= \begin{cases} -\sum_{i+j-1 \in \overline{S}_i} v_{j,m}v_{j,n}, & m \ne n \\ 1 - \sum_{i+j-1 \in \overline{S}_i} v_{j,m}^2 & m = n \end{cases} \\
\boldsymbol{\xi}_i &= [a_{1,i} \quad a_{2,i} \quad \cdots \quad a_{L,i}]^{\mathrm{T}} \\
\boldsymbol{y}_i &= \left[\sum_{i+j-1 \in S_i} x_{i+j-1}v_{j,1} \quad \sum_{i+j-1 \in S_i} x_{i+j-1}v_{j,2} \quad \cdots \quad \sum_{i+j-1 \in S_i} x_{i+j-1}v_{j,L}\right]^{\mathrm{T}}
\end{aligned} \tag{3}$$

where S_i and $\overline{S}_i$ represent the index sets of observed data and missing data, respectively, within the integer interval $[i, i+L-1]$, i.e., $S_i \cap \overline{S}_i = \emptyset$ and $S_i \cup \overline{S}_i = [i, i+L-1]$; $v_{i,j}$ and $a_{i,j}$ are the (i,j)-th element of $\boldsymbol{V}$ and PCs, respectively.

Due to the presence of missing data, $\boldsymbol{G}_i$ is rank-deficient. To ensure a unique solution, Shen et al. (2015) introduced the minimum norm criterion,

$$\min: \boldsymbol{\xi}_i^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{\xi}_i \tag{4}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix whose diagonal elements are the singular values of $\boldsymbol{C}$ sorting in descending order. The corresponding solution is given by,

$$\boldsymbol{\xi}_i = \boldsymbol{\Lambda}\boldsymbol{G}_i^{\mathrm{T}}\left(\boldsymbol{G}_i\boldsymbol{\Lambda}\boldsymbol{G}_i^{\mathrm{T}}\right)^{+}\boldsymbol{y}_i \tag{5}$$

where the subscript '+' denotes the Moore–Penrose inverse. Different from ISSA, the efficient ISSA developed by Ji et al. (2025) start from the equation,

$$\boldsymbol{\xi}_i = \boldsymbol{V}^{\mathrm{T}}\boldsymbol{X}_i = \boldsymbol{B}_{1,i}\boldsymbol{x}_{1,i} + \boldsymbol{B}_{2,i}\boldsymbol{x}_{2,i} \tag{6}$$

where $\boldsymbol{X}_i = (x_i, x_{i+1}, \cdots, x_{i+L-1})^{\mathrm{T}}$, $\boldsymbol{B}_{1,i}$ and $\boldsymbol{B}_{2,i}$ are sub-matrices of $\boldsymbol{V}^{\mathrm{T}}$ comprising its column vectors related to the available epochs and missing epochs with sizes of $L \times N_{S_i}$ and $L \times N_{\overline{S}_i}$, respectively; $\boldsymbol{x}_{1,i} = \{x_j \mid j \in S_i\}$ and $\boldsymbol{x}_{2,i} = \{x_j \mid j \in \overline{S}_i\}$ denote available observed data and missing data in $\boldsymbol{X}_i$, respectively, N_{S_i} and $N_{\overline{S}_i}$ are numbers of elements in S_i and $\overline{S}_i$. The unknowns $\boldsymbol{x}_{2,i}$ in Eq. (6) are estimated using the criterion Eq. (4), yielding,

$$\boldsymbol{\xi}_i = \left[\boldsymbol{I}_L - \boldsymbol{B}_{2,i}\left(\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{B}_{2,i}\right)^{-1}\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\right]\boldsymbol{B}_{1,i}\boldsymbol{x}_{1,i} \tag{7}$$

Once all $\boldsymbol{\xi}_i$ are solved, the signals can be reconstructed using the dominant PCs, just as in standard SSA. This allows us to both filter incomplete time series and fill data gaps using the reconstructed signal values at the missing epochs.

3. PROOF OF THE EQUIVALENCE BETWEEN ISSA AND EFFICIENT ISSA

Since $\boldsymbol{B}_{1,i}$ and $\boldsymbol{B}_{2,i}$ are column-wise submatrices of $\boldsymbol{V}^{\mathrm{T}}$, the orthogonality of $\boldsymbol{V}^{\mathrm{T}}$ implies that,

$$\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{B}_{2,i} = \boldsymbol{0}_{N_{S_i}\times N_{\overline{S}_i}}, \quad \boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{B}_{1,i} = \boldsymbol{0}_{N_{\overline{S}_i}\times N_{S_i}} \tag{8}$$

$$\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{B}_{1,i} = \boldsymbol{I}_{N_{S_i}}, \quad \boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{B}_{2,i} = \boldsymbol{I}_{N_{\overline{S}_i}} \tag{9}$$

$$\boldsymbol{B}_{1,i}\boldsymbol{B}_{1,i}^{\mathrm{T}} + \boldsymbol{B}_{2,i}\boldsymbol{B}_{2,i}^{\mathrm{T}} = \boldsymbol{I}_L \tag{10}$$

where $\boldsymbol{0}_{m\times n}$ denotes an $m \times n$ zero matrix and $\boldsymbol{I}_n$ denotes the $n \times n$ identity matrix. The following derivations will make frequent use of Eqs. (8) - (10).

Observing that $\boldsymbol{G}_i$ and $\boldsymbol{y}_i$ can be expressed as matrix/vector forms,

$$\boldsymbol{G}_i = \boldsymbol{I}_L - \boldsymbol{B}_{2,i}\boldsymbol{B}_{2,i}^{\mathrm{T}} = \boldsymbol{B}_{1,i}\boldsymbol{B}_{1,i}^{\mathrm{T}} \tag{11}$$

$$\boldsymbol{y}_i = \boldsymbol{B}_{1,i}\boldsymbol{x}_{1,i} \tag{12}$$

Given that $\boldsymbol{H}_1 = \boldsymbol{G}_i\boldsymbol{\Lambda}\boldsymbol{G}_i^{\mathrm{T}}$ and $\boldsymbol{H}_2 = \boldsymbol{B}_{1,i}\left(\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\right)^{-1}\boldsymbol{B}_{1,i}^{\mathrm{T}}$, we have,

$$\begin{aligned}\boldsymbol{H}_1\boldsymbol{H}_2\boldsymbol{H}_1 &= \boldsymbol{B}_{1,i}\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\boldsymbol{B}_{1,i}^{\mathrm{T}}\cdot\boldsymbol{B}_{1,i}\left(\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\right)^{-1}\boldsymbol{B}_{1,i}^{\mathrm{T}}\cdot\boldsymbol{B}_{1,i}\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\boldsymbol{B}_{1,i}^{\mathrm{T}}\\ &= \boldsymbol{B}_{1,i}\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\boldsymbol{B}_{1,i}^{\mathrm{T}}\\ &= \boldsymbol{H}_1\end{aligned} \tag{13}$$

$$\begin{aligned}\boldsymbol{H}_2\boldsymbol{H}_1\boldsymbol{H}_2 &= \boldsymbol{B}_{1,i}\left(\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\right)^{-1}\boldsymbol{B}_{1,i}^{\mathrm{T}}\cdot\boldsymbol{B}_{1,i}\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\boldsymbol{B}_{1,i}^{\mathrm{T}}\cdot\boldsymbol{B}_{1,i}\left(\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\right)^{-1}\boldsymbol{B}_{1,i}^{\mathrm{T}}\\ &= \boldsymbol{B}_{1,i}\left(\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\right)^{-1}\boldsymbol{B}_{1,i}^{\mathrm{T}}\\ &= \boldsymbol{H}_2\end{aligned} \tag{14}$$

$$\boldsymbol{H}_1\boldsymbol{H}_2 = \boldsymbol{B}_{1,i}\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\boldsymbol{B}_{1,i}^{\mathrm{T}}\cdot\boldsymbol{B}_{1,i}\left(\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\right)^{-1}\boldsymbol{B}_{1,i}^{\mathrm{T}} = \boldsymbol{B}_{1,i}\boldsymbol{B}_{1,i}^{\mathrm{T}} \tag{15}$$

$$\boldsymbol{H}_2\boldsymbol{H}_1 = \boldsymbol{B}_{1,i}\left(\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\right)^{-1}\boldsymbol{B}_{1,i}^{\mathrm{T}}\cdot\boldsymbol{B}_{1,i}\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\boldsymbol{B}_{1,i}^{\mathrm{T}} = \boldsymbol{B}_{1,i}\boldsymbol{B}_{1,i}^{\mathrm{T}} \tag{16}$$

Since $\boldsymbol{B}_{1,i}\boldsymbol{B}_{1,i}^{\mathrm{T}}$ is a real symmetric matrix (and hence Hermitian), it follows that,

$$(\boldsymbol{H}_1\boldsymbol{H}_2)^* = \boldsymbol{H}_1\boldsymbol{H}_2 \tag{17}$$

$$(\boldsymbol{H}_2\boldsymbol{H}_1)^* = \boldsymbol{H}_2\boldsymbol{H}_1 \tag{18}$$

where the subscript '*' is the Hermitian transpose. Clearly, Eqs. (13), (14), (17), and (18) collectively satisfy the four Moore–Penrose conditions (Penrose, 1955). Consequently, $\boldsymbol{H}_2$ is the unique Moore–Penrose inverse of $\boldsymbol{H}_1$, i.e.,

$$\left(\boldsymbol{G}_i\boldsymbol{\Lambda}\boldsymbol{G}_i^{\mathrm{T}}\right)^{+} = \boldsymbol{B}_{1,i}\left(\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\right)^{-1}\boldsymbol{B}_{1,i}^{\mathrm{T}} \tag{19}$$

Substituting Eq. (19) into Eq. (5) yields,

$$\boldsymbol{\xi}_i = \boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{B}_{1,i}\left(\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\right)^{-1}\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{B}_{1,i}\boldsymbol{x}_{1,i} = \boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\left(\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\right)^{-1}\boldsymbol{x}_{1,i} \tag{20}$$

Furthermore, note that,

$$\begin{aligned}&\left[\boldsymbol{I}_L - \boldsymbol{B}_{2,i}\left(\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{B}_{2,i}\right)^{-1}\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\right]\boldsymbol{B}_{1,i}\cdot\left(\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\right)\\ &= \left[\boldsymbol{I}_L - \boldsymbol{B}_{2,i}\left(\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{B}_{2,i}\right)^{-1}\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\right]\left[\left(\boldsymbol{I}_L - \boldsymbol{B}_{2,i}\boldsymbol{B}_{2,i}^{\mathrm{T}}\right)\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\right]\\ &= \left[\boldsymbol{I}_L - \boldsymbol{B}_{2,i}\left(\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{B}_{2,i}\right)^{-1}\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\right]\left(\boldsymbol{\Lambda}\boldsymbol{B}_{1,i} - \boldsymbol{B}_{2,i}\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\right)\\ &= \boldsymbol{\Lambda}\boldsymbol{B}_{1,i} - \boldsymbol{B}_{2,i}\left(\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{B}_{2,i}\right)^{-1}\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i} - \boldsymbol{B}_{2,i}\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i} + \boldsymbol{B}_{2,i}\left(\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{B}_{2,i}\right)^{-1}\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{B}_{2,i}\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\\ &= \boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\end{aligned} \tag{21}$$

Since $\boldsymbol{B}_{1,i}$ is the column-based submatrix of $\boldsymbol{V}^{\mathrm{T}}$ and has full column rank, we have $\mathrm{Rank}\left(\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\right) = \mathrm{Rank}\left(\boldsymbol{B}_{1,i}\right) = N_S$, where $\mathrm{Rank}(\cdot)$ is the rank operator. It means that $\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}$ is full-rank and invertible. Therefore, we have,

$$\left[\boldsymbol{I}_L - \boldsymbol{B}_{2,i}\left(\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{B}_{2,i}\right)^{-1}\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\right]\boldsymbol{B}_{1,i} = \boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\left(\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{B}_{1,i}\right)^{-1} \tag{22}$$

From Eqs. (7) and (20), we conclude that the filtering matrices of two methods are equivalent, leading to identical $\boldsymbol{\xi}_i$. Notably, the final simplified filtering matrix $\boldsymbol{\Lambda B}_{1,i}\left(\boldsymbol{B}_{1,i}^{\mathrm{T}}\boldsymbol{\Lambda B}_{1,i}\right)^{-1}$ depends solely on observed epochs and is independent of missing data.

4. PROOF OF COMPUTATIONAL SUPERIORITY OF EFFICIENT ISSA OVER ISSA

The computational complexities of ISSA and efficient ISSA comprise shared foundational steps and method-specific iterative processes. Figures 1 and 2 present the flowcharts of the ISSA and efficient ISSA, respectively. It shows that the only difference between the two methods is the calculation of $\boldsymbol{\xi}_i$, highlighted in gray. We present a rigorous comparison of their time complexity, with emphasis on the efficiency improvements enabled by efficient ISSA. Both methods require constructing Toeplitz covariance matrix $\boldsymbol{C}$ ($O(NL)$) and performing eigenvalue decomposition ($O(L^3)$). They then iterative over $K = N - L + 1$ windows. Figure 3 summarizes the time complexity of each step for the ith process, showing that both methods follow three main steps:

Step 1:

Prepares the necessary elements for solving the linear system in *Step 2* and *Step 3*.

Step 2:

Computes the projection matrix that maps $\boldsymbol{y}_i = \boldsymbol{B}_{1,i}\boldsymbol{x}_{1,i}$ to $\boldsymbol{\xi}_i$.

Step 3:

Obtains $\boldsymbol{\xi}_i$ through the multiplication of an $L \times L$ projection matrix and an $L \times 1$ vector.

Due to space limitations, we do not explain the time complexity of each computation step in Figure 3. Instead, we explain a typical step, as the analysis of other steps follows a similar approach. Before the analysis, we introduce the basic principles of time complexity analysis (Cormen et al., 2022): (a) *Matrix multiplication*: For $\boldsymbol{P}_{m\times n}\boldsymbol{Q}_{n\times r}$, the time complexity is $O(mnr)$; (b) *Diagoinal matrix scaling*: For $\boldsymbol{P}_{m\times n}\boldsymbol{R}_{n\times n}$ (diagonal $\boldsymbol{R}$), the time complexity is $O(mn)$; (c)*Dominant term rule*: The highest-order term determines the overall complexity.

Take the computation of $\left(\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{B}_{2,i}\right)^{-1}$ for efficient ISSA as an example. This calculation involves two steps:

***Step 1*:**

Compute term $\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{B}_{2,i}$. (1): $\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}$ scales $LN_{\overline{S}_i}$ elements of $\boldsymbol{B}_{2,i}^{\mathrm{T}}$ with diagonal elements of $\boldsymbol{\Lambda}^{-1}$, requiring $O\left(LN_{\overline{S}_i}\right)$. (2): Multiply $\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}$ (size $N_{\overline{S}_i} \times L$) with $\boldsymbol{B}_{2,i}$ ($L \times N_{\overline{S}_i}$), requiring $O\left(L \times N_{\overline{S}_i}^2\right)$. Therefore, the total complexity of Step 1 is $O\left(L \times N_{\overline{S}_i}^2\right)$ (dominant term).

Step 2:

Compute the inversion of $\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{B}_{2,i}$ with size $N_{\overline{S}_i} \times N_{\overline{S}_i}$, requiring $O\left(N_{\overline{S}_i}^3\right)$.

Therefore, the total complexity for $\left(\boldsymbol{B}_{2,i}^{\mathrm{T}}\boldsymbol{\Lambda}^{-1}\boldsymbol{B}_{2,i}\right)^{-1}$ is $O\left(LN_{\overline{S}_i}^2 + N_{\overline{S}_i}^3\right)$.

Referring back to Figure 3, by focusing on high-order independent terms, we can deduce that for the ith process, the time complexity of ISSA is $O(L^3)$, while the time complexity of efficient ISSA is $O\left(N_{\overline{S}_i}^3 + LN_{\overline{S}_i}^2 + L^2N_{\overline{S}_i}\right)$. It is clear that $O(L^3) > O\left(N_{\overline{S}_i}^3 + LN_{\overline{S}_i}^2 + L^2N_{\overline{S}_i}\right)$. When $N_{\overline{S}_i}$ is sufficiently small (indicating sparse missing data, $N_{\overline{S}_i} \ll L$), $O\left(N_{\overline{S}_i}^3 + LN_{\overline{S}_i}^2 + L^2N_{\overline{S}_i}\right) \approx O\left(L^2N_{\overline{S}_i}\right)$. The improvement of efficient ISSA over ISSA can be quantified as follows:

$$\mathrm{Imp} = \frac{L^3 - \left(N_{\overline{S}_i}^3 + LN_{\overline{S}_i}^2 + L^2N_{\overline{S}_i}\right)}{L^3} = \\ = \left(1 - \sum_{k=1}^{3}\left(\frac{N_{\overline{S}_i}}{L}\right)^k\right) \times 100\ \% \tag{23}$$

It indicates that the improvement depends on the proportion of missing data within the intervals $[i, i + L - 1]$. Specifically, when the total percentage of missing data increases (with fixed N and L), the ratio $N_{\overline{S}_i}/L$ increases, leading to a reduction in the improvement. As L increases (with fixed N and a constant total missing data percentage), the growth rate of $N_{\overline{S}_i}$ is inevitably lower and may even remain constant for certain intervals (e.g., in the case of structured data gaps). This results in a decrease in the ratio $N_{\overline{S}_i}/L$, thereby enhancing the improvement (Imp). These analytical findings align with the numerical experiments presented in Ji et al. (2025).

Noting that Eq. (23) shows how efficient ISSA improves computational efficiency over ISSA when computing $\boldsymbol{\xi}_i$. When $N_{\overline{S}_i} = 0$ for the ith computational process, both Eq. (5) (ISSA) and Eq. (7) (efficient ISSA) reduce to conventional SSA: $\boldsymbol{\xi}_i = \boldsymbol{V}^{\mathrm{T}}\boldsymbol{X}_i$, and no improvement is gained. Nevertheless, as long as missing values exist in the time series $\boldsymbol{x}$, it is unlikely that $N_{\overline{S}_i} = 0$ for all $i = 1$ to K. Therefore, the improvement in efficiency for computing all principal components still holds.

5. COMPARISON OF EFFICIENT ISSA WITH ESSA

In many geophysical and geodetic studies, such as those related to hydrological mass variations (Zhang et al., 2023; Ruan et al., 2024), crustal deformation monitoring (Dong et al., 2006), or spatiotemporal pattern recognition in geophysical fields (Iwamori et al., 2017; Ars et al., 2024), the PCs themselves are of primary scientific interest since they often carry distinct physical meanings. For example, leading PCs may correspond to dominant seasonal cycles, interannual variations, or specific spatial deformation modes. In such cases, the objective is not merely to reconstruct the observed signal with minimal residuals, but rather to extract orthogonal

Algorithm: ISSA

Input:
incomplete time series $\boldsymbol{x}_1$ and indices of missing epochs $\overline{S}$
window size L and reconstruction order r
Output:
PC estimates $\boldsymbol{\xi}_i$ $(i = 1,2,\cdots,K)$
Note:
N is the length of complete time series, and $K = N - L + 1$
Begin
Construct lagged covariance matrix $\boldsymbol{C}$ using available data with Eq. (1)
Perform SVD on $\boldsymbol{C} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^T$
for i from 1 to K
Construct $\boldsymbol{G}_i$ and $\boldsymbol{y}_i$
Estimate $\boldsymbol{\xi}_i$ using Eq. (5)
End
End

Fig. 1 Flowchart of the ISSA method.

Algorithm: Efficient ISSA

Input:
incomplete time series $\boldsymbol{x}_1$ and indices of missing epochs $\overline{S}$
window size L and reconstruction order r
Output:
PC estimates $\boldsymbol{\xi}_i$ $(i = 1,2,\cdots,K)$
Note:
N is the length of complete time series, and $K = N - L + 1$
Begin
Construct lagged covariance matrix $\boldsymbol{C}$ using available data with Eq. (1)
Perform SVD on $\boldsymbol{C} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^T$
for i from 1 to K
Construct $\boldsymbol{B}_{1,i}, \boldsymbol{B}_{2,i}$, and $\boldsymbol{x}_{1,i}$
Estimate $\boldsymbol{\xi}_i$ using Eq. (7)
End
End

Fig. 2 Flowchart of the efficient ISSA method.

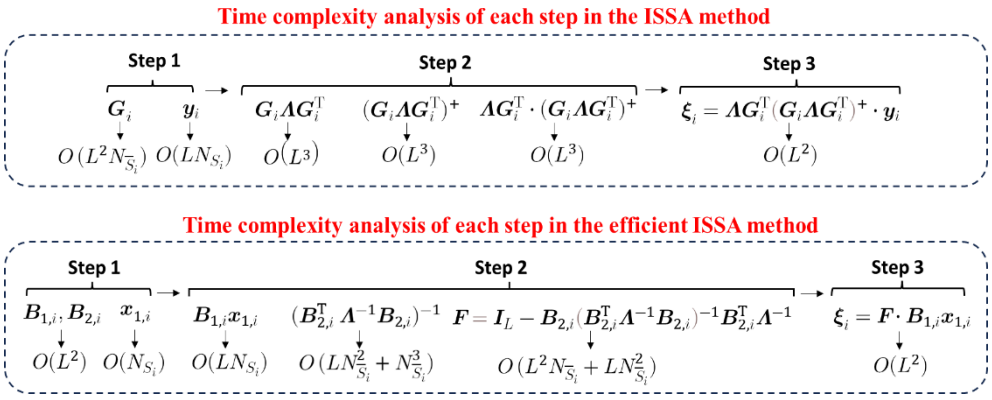

Fig. 3 Time complexity analysis of each step in the ISSA and efficient ISSA methods for the ith proces.

modes that can be physically interpreted and compared across datasets or models. By contrast, when the focus is on improving the fidelity of signal reconstruction, for instance, in noise reduction, data gap filling, or regional signal recovery, methods based on low-rank approximation, such as ESSA, are typically preferred. Therefore, ISSA and ESSA serve different purposes: ISSA emphasizes the stability and interpretability of the PCs under a weighted norm framework, while ESSA prioritizes reconstruction accuracy of the underlying signal.

Since we have theoretically proved that efficient ISSA is equivalent to ISSA in computing PC estimates and is more efficient, there is no need to repeat

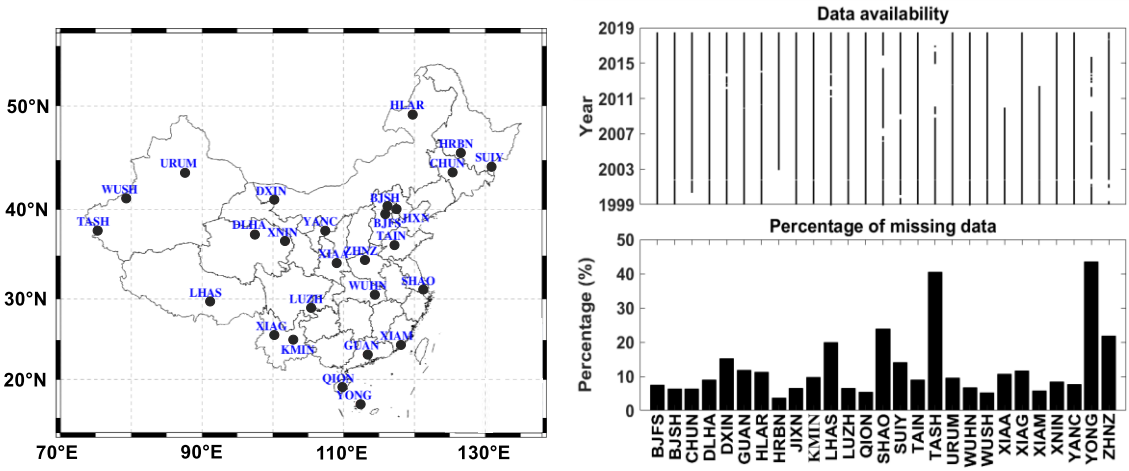

Fig. 4 Geographic locations of 27 stations (left) and data availability and percentage of missing data (right).

numerical experiments already done by Ji et al. (2025). This section focuses on comparing efficient ISSA and ESSA (Ji et al., 2023) in terms of both computational efficiency and results. We follow Ji et al. (2025) by using real position time series of 27 GNSS stations located in Chinese mainland as testing data. Figure 4 (left) shows the geographical distribution of the 27 GNSS stations, while Figure 4 (right) presents the corresponding data availability and percentage of missing data. All computations are performed in MATLAB 2017 on a Windows 10 system equipped with a 12th Gen Intel(R) Core (TM) i7-12700 2.10 GHz processor and 32 GB of memory.

5.1. COMPUTATIONAL EFFICIENCY

In practical applications, the computation time of the two methods depends on four factors: window size (L), reconstruction order (r), time series length (N), and percentage of missing data (MP). The BJFS station position time series is first used as an example. To ensure statistical reliability, each experiment is repeated 20 times.

We start by investigating the impact of varying window sizes and reconstruction orders on computational efficiency. In each experiment, window sizes range from one to four years, and reconstruction orders range from five to eleven, increasing by two. Figure 5 presents the average computation time of the two methods under a fixed window size with varying reconstruction orders (top row), and the relative improvement of efficient ISSA over ISSA (bottom row). Figure 6 shows the average computation time under a fixed reconstruction order with varying window sizes. The results clearly demonstrate that ESSA's time complexity is more senstive to reconstruction order and window size than that of efficient ISSA. This is expected, as ESSA introduces a heavier computational load when building the filtering matrix. This step involves multiple tedious multiplication and summation operations (Ji et al., 2023), and becomes more demanding as the reconstruction order and window size increase. In contrast, for efficient ISSA, the reconstruction order does not affect the time complexity of computing the PCs (see Fig. 3). Even during the signal reconstruction stage, its impact is minimal, which is similar to that in standard SSA.

To further explore how the percentage of missing data and time series length affect computational efficiency, two sets of repeated experiments are performed. The window size is fixed at 365, and the reconstruction order at 5. In the first set, the percentage of missing data is fixed at 30 %, while the time series length varies from 1000 to 6000 in increments of 1000. In the second set, the time series length is fixed at 7022, and the percentage of missing data varies from 10 % to 60 % in increments of 10 %. Figures 7 and 8 show the computation time of efficient ISSA and ESSA under these two scenarios. In all cases, efficient ISSA consistently outperforms ESSA in terms of computational efficiency. Figure 9 shows the relative improvement of efficient ISSA over ESSA based on 50 simulations, and Table 1 summarizes the average computation time and relative improvements. The results show that as the percentage of missing data increases, the efficiency improvement of efficient ISSA becomes more significant. On the other hand, as the time series length increases, the improvement gradually decreases.

5.2. PC ESTIMATES AND FILTERED SIGNALS

The (efficient) ISSA method can theoretically guarantee optimal PC estimates using available observation data, thanks to a minimum norm criterion in the frequency domain. To demonstrate this, we conduct 50 simulation experiments using 27 real datasets, with a window size of one year. For each simulation at every station, a complete time series is generated using the simulation strategy proposed by Ji et al. (2025). Specifically, the signals extracted by efficient ISSA from the 27 real GNSS position time series were regarded as the "true" reference signals. To generate realistic yet controlled observations, Gaussian white noise with zero mean and variance $\hat{\sigma}^2$

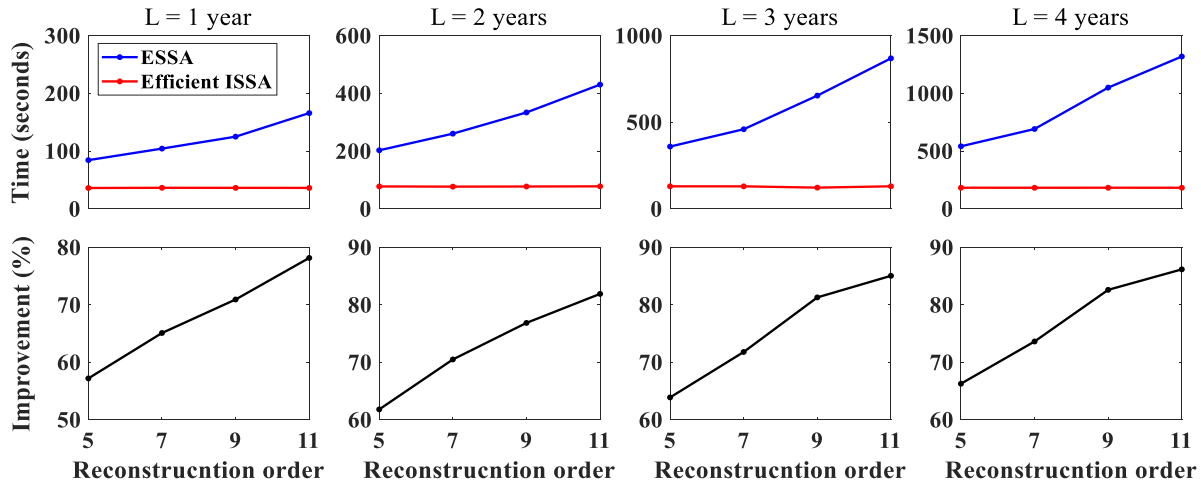

Fig. 5 Computation time of the two methods (first row) and relative improvement (second row) under varying reconstruction orders (fixed window sizes).

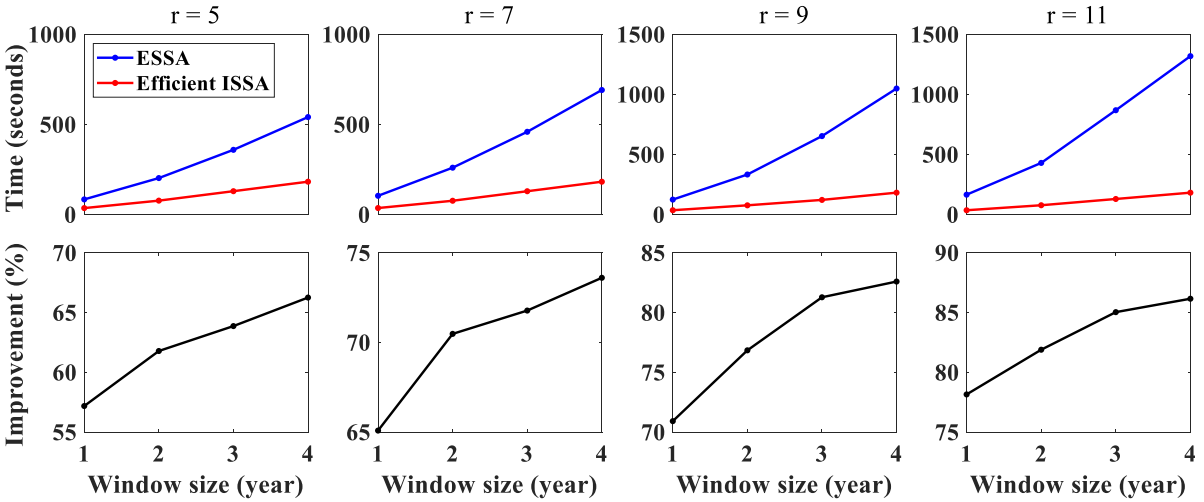

Fig. 6 Computation time of the two methods (first row) and relative improvement (second row) under varying window sizes (fixed reconstruction orders).

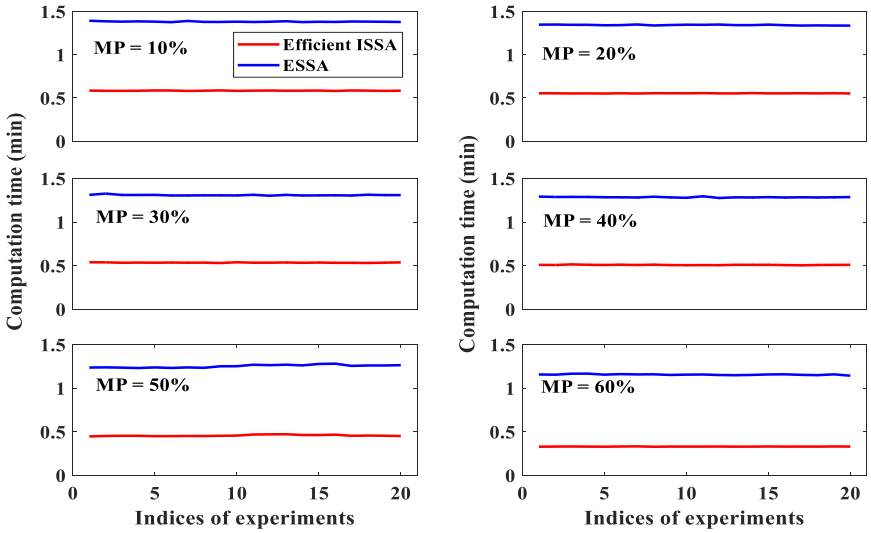

Fig. 7 Computation time of the two methods for 20 simulations under varying percentages of missing data.

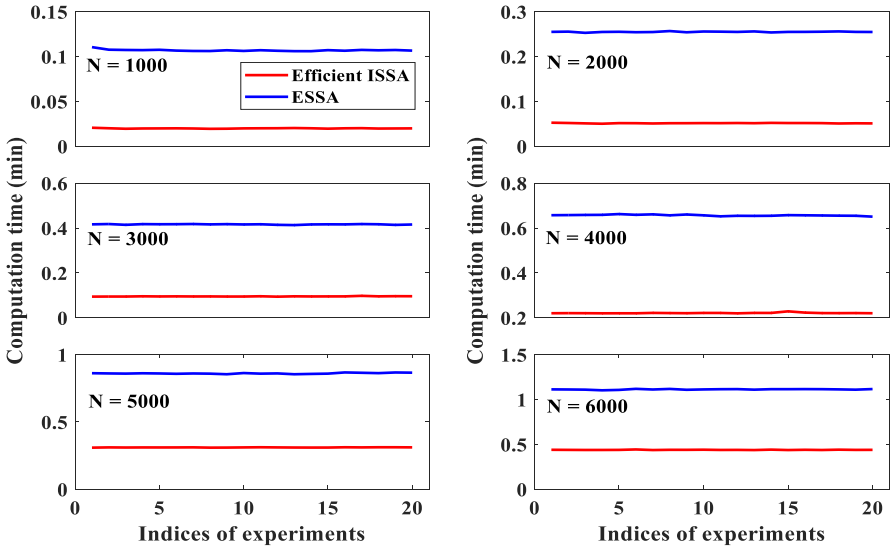

Fig. 8 Computation time of the two methods for 20 simulations under varying time series lengths.

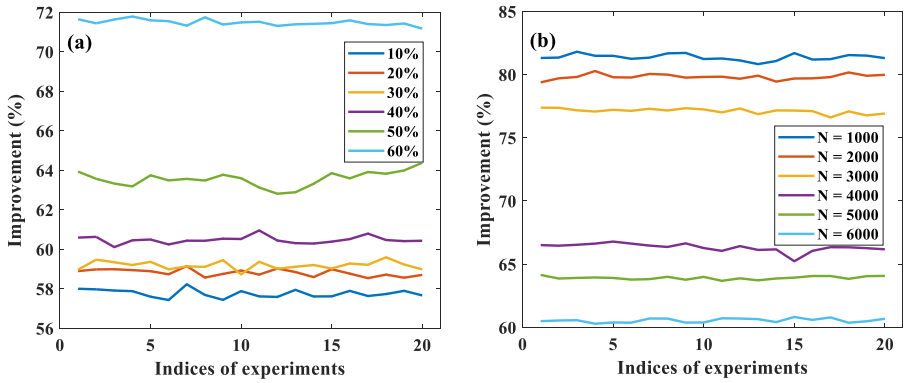

Fig. 9 Relative improvements of efficient ISSA over ESSA under varying percentages of missing data (a) and varying time series lengths (b).

Table 1 Average computation time and relative improvement from 20 simulations T_E and T_{EI} represent computation time of ESSA and efficient ISSA, respectively.

$N = 7022$				MP = 30 %			
MP	T_E (min)	T_{EI} (min)	Imp	N	T_{I} (min)	T_{EI} (min)	Imp
10 %	1.38	**0.58**	57.8 %	1000	0.11	**0.02**	81.4 %
20 %	1.34	**0.55**	58.8 %	2000	0.26	**0.05**	79.8 %
30 %	1.31	**0.53**	59.2 %	3000	0.42	**0.10**	77.1 %
40 %	1.28	**0.51**	60.5 %	4000	0.66	**0.22**	66.3 %
50 %	1.25	**0.46**	63.6 %	5000	0.86	**0.31**	63.9 %
60 %	1.15	**0.33**	71.5 %	6000	1.11	**0.44**	60.6 %

was added, where $\hat{\sigma}$ denotes the fitting error estimated from the real GNSS data. To evaluate the two methods on processing gappy time series, we remove portions of the complete time series based on missing epochs inherited from the real data. The PC estimates derived from the complete dataset using conventional SSA are treated as the true values $\tilde{\boldsymbol{\xi}}_i$. Since these true PCs are known, we compute the root mean squared errors (RMSEs) of the estimates $\boldsymbol{\xi}_i$ using Eq. (24),

$$\mathrm{RMSE}_i = \sqrt{\left\|\left(\tilde{\boldsymbol{\xi}}_i - \boldsymbol{\xi}_i\right)\right\|_2 / K} \tag{24}$$

where $\|\boldsymbol{x}\|_2 = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$ represents the L2-norm operator. Table 2 shows the average RMSE values for the first leading four components. The results clearly show that (efficient) ISSA performs better than ESSA in PCs estimation. It is worth noting that efficient ISSA performs slightly worse than ESSA for some stations and components. This can occur since the two methods

Table 2 RMSE values of the estimated $\boldsymbol{\xi}_i (i = 1, 2, 3, 4)$ (mm).

	ESSA				Efficient ISSA			
	$\boldsymbol{\xi}_1$	$\boldsymbol{\xi}_2$	$\boldsymbol{\xi}_3$	$\boldsymbol{\xi}_4$	$\boldsymbol{\xi}_1$	$\boldsymbol{\xi}_2$	$\boldsymbol{\xi}_3$	$\boldsymbol{\xi}_4$
BJFS	**5.5857**	**5.3809**	**5.5016**	**5.4126**	5.5449	5.3356	5.4551	5.3628
BJSH	**5.0150**	**4.9850**	**5.1238**	**5.1753**	5.0178	4.9606	5.1162	5.1775
CHUN	**8.8681**	**8.7217**	**8.9955**	**9.2750**	8.8725	8.8118	9.0285	9.2754
DLHA	**3.8465**	**3.9534**	**3.7224**	**3.9100**	3.4321	3.4967	3.2617	3.5017
DXIN	**3.9906**	**4.1982**	**4.2418**	**4.0353**	3.9699	4.1508	4.2302	4.0210
GUAN	**8.7878**	**8.9147**	**8.9018**	**8.4430**	8.7814	8.9476	8.8826	8.4673
HLAR	**6.8641**	**7.1583**	**6.8630**	**7.0059**	5.4535	5.8146	5.6372	5.6890
HRBN	**5.7336**	**5.8293**	**5.5401**	**5.6864**	5.6912	5.8201	5.5165	5.6849
JIXN	**5.0702**	**5.0103**	**4.8862**	**4.8184**	5.0518	4.9823	4.8896	4.8184
KMIN	**7.1639**	**7.1578**	**6.8766**	**7.3254**	7.0928	7.1072	6.7609	7.2065
LHAS	**5.2331**	**5.2774**	**5.1875**	**5.3966**	5.0012	5.0754	5.0029	5.1462
LUZH	**5.2017**	**4.8865**	**4.9480**	**5.2786**	5.1846	4.8669	4.9247	5.2564
QION	**9.4496**	**9.6248**	**9.3571**	**9.2583**	9.4055	9.5857	9.3084	9.2280
SHAO	**7.0458**	**7.1802**	**6.9107**	**7.0396**	7.0360	7.1788	6.9165	7.0595
SUIY	**6.5630**	**6.9921**	**6.6789**	**6.8189**	6.5189	6.9360	6.6531	6.7931
TAIN	**5.7931**	**6.0150**	**6.3016**	**5.6865**	5.7639	5.9825	6.2603	5.6777
TASH	**5.8301**	**5.6856**	**5.4942**	**5.7030**	5.8226	5.6841	5.4833	5.6960
URUM	**6.6758**	**6.4589**	**6.7801**	**6.4598**	5.8689	5.6409	5.8003	5.6057
WUHN	**6.8660**	**6.4151**	**6.5248**	**6.2734**	6.8398	6.3811	6.5124	6.2303
WUSH	**5.2816**	**5.4933**	**5.2731**	**5.2700**	5.209	5.4700	5.2331	5.2411
XIAA	**6.0357**	**6.1798**	**6.2934**	**6.5832**	5.4781	5.5102	5.8151	6.0694
XIAG	**7.9406**	**7.9546**	**8.2586**	**8.0256**	7.9168	7.9155	8.2568	7.9667
XIAM	**7.3400**	**7.5037**	**7.4685**	**7.5473**	7.3127	7.4882	7.4283	7.5272
XNIN	**7.3546**	**7.2430**	**7.1986**	**7.2583**	7.2301	7.1674	7.1088	7.1656
YANC	**5.0808**	**5.1136**	**4.9351**	**4.8816**	4.2257	4.3145	4.2065	4.1657
YONG	**11.3792**	**11.4846**	**12.1608**	**11.7952**	11.3628	11.4622	12.1395	11.7759
ZHNZ	**6.1721**	**6.2467**	**6.3619**	**6.3562**	6.1100	6.2161	6.2901	6.2984

estimate PCs in fundamentally different ways. Efficient ISSA directly solves a rank-deficient linear system for each $\boldsymbol{\xi}_i$, yielding a minimum-norm solution, whereas ESSA first estimates the missing values $\boldsymbol{x}_2$ by solving an overdetermined least-squares problem and then substitutes them into the PC equations to obtain $\boldsymbol{\xi}_i$. Consequently, when the system associated with a particular component $\boldsymbol{\xi}_i$ becomes highly rank-deficient, the performance of efficient ISSA may occasionally be slightly inferior to that of ESSA.

On the other hand, since ESSA is based on a low-rank optimal approximation criterion in the time domain, it outperforms ISSA in filtering noisy time series (Ji et al., 2023). However, as shown in Ji et al. (2025), ISSA can be improved by iteratively updating the lagged covariance matrix using previous results. ESSA can also be extended in the same way. To compare the iterative versions of the two methods, we update the lagged covariance matrix at each iteration. The iteration process terminates when the maximum absolute difference between successive extracted signals $\boldsymbol{s}$, denoted as $\max(|\Delta \boldsymbol{s}|)$, falls below 0.001 mm. Table 3 presents the fitting errors of signals extracted by ISSA, ESSA, iterative ISSA, and iterative ESSA, computed using Eq. (25),

$$\hat{\sigma} = \sqrt{\frac{1}{N_S}\sum_{i\in S}(x_i - s_i)^2} \tag{25}$$

A smaller fitting error indicates more effective signal extraction (Ji et al., 2023, 2025). The results show that although original ISSA is less effective than ESSA (Ji et al., 2023), iterative ISSA achieves comparable performance. More importantly, it significantly reduces computational cost, as shown in Table 3.

Finally, to make the comparison clearer, the main properties of ISSA, efficient ISSA, and ESSA are summarized in Table 4.

6. APPLICATION TO GRACE AND GRACE-FO GRAVITY FIELD SOLUTIONS

From March 2002 to October 2017, the monthly time-variable gravity field models released by the GRACE mission have been extensively applied in various fields due to their high spatial resolution and accuracy, including terrestrial water storage (TWS) variations (Syed et al., 2008, 2009), glacier mass balance (Chen et al., 2009; Gao et al., 2015), and global ocean mass change (Tamisiea, 2011). However, following the launch of the GRACE-FO mission in May 2018, a data gap of nearly one year emerged between the two missions, posing new challenges for the continuous monitoring of geophysical signals. In this section, the developed efficient ISSA algorithm is applied to interpolate the missing time-variable gravity field data between the GRACE and GRACE-FO missions.

Table 3 Fitting errors of signals extracted by efficient ISSA and ESSA (mm) and computation time (min).

	Fitting errors				Computation time			
	Non-iterative		Iterative		Non-iterative		Iterative	
	EISSA	ESSA	EISSA	ESSA	ISSA	ESSA	EISSA	ESSA
BJFS	3.8072	3.8060	3.8033	3.8032	1.00	2.80	1.63	6.71
BJSH	3.8919	3.8921	3.8917	3.8918	0.99	2.83	1.20	6.71
CHUN	4.4818	4.4815	4.4814	4.4813	0.90	2.62	1.12	6.24
DLHA	2.5776	2.5763	2.5768	2.5759	1.06	2.80	1.62	6.71
DXIN	3.0380	3.0344	3.0376	3.0367	1.26	2.75	2.51	10.57
GUAN	6.2265	6.2277	6.2251	6.2252	1.33	2.76	2.94	6.64
HLAR	4.5591	4.5574	4.5594	4.5592	1.17	1.99	2.20	5.50
HRBN	4.0711	4.0706	4.0713	4.0709	0.71	2.16	0.86	5.23
JIXN	3.6019	3.6015	3.6024	3.6021	0.98	2.82	1.19	6.68
KMIN	5.0001	4.9991	4.9989	4.9990	1.06	3.51	1.72	11.55
LHAS	3.8024	3.7942	3.7992	3.7963	1.29	2.76	2.76	10.61
LUZH	4.0319	4.0326	4.0317	4.0316	0.98	2.82	1.18	6.68
QION	6.6488	6.6503	6.6489	6.6502	1.02	2.78	1.33	6.65
SHAO	4.2505	4.2480	4.2519	4.2481	1.22	3.71	1.59	7.28
SUIY	4.9318	4.9256	4.9318	4.9269	1.15	2.76	2.15	16.36
TAIN	4.5233	4.5234	4.5226	4.5230	1.01	2.79	1.41	6.65
TASH	3.7928	3.7838	3.7897	3.7870	1.56	2.29	5.80	14.98
URUM	4.7102	4.7003	4.7155	4.7069	1.08	2.04	1.75	5.55
WUHN	4.7160	4.7173	4.7163	4.7174	0.99	2.83	1.21	6.74
WUSH	3.8907	3.8907	3.8911	3.8912	1.04	2.80	1.35	6.64
XIAA	4.3097	4.3116	4.3068	4.3091	0.47	1.38	0.85	3.47
XIAG	6.6760	6.6757	6.6753	6.6751	0.98	2.80	1.21	6.65
XIAM	5.4512	5.4551	5.4517	5.4570	0.57	1.79	0.86	5.67
XNIN	2.9176	2.9172	2.9176	2.9173	1.00	2.81	1.24	6.67
YANC	3.1186	3.1171	3.1172	3.1167	0.99	2.84	1.19	6.71
YONG	7.1110	6.9480	7.0386	6.7279	1.65	1.93	16.25	35.79
ZHNZ	4.4113	4.4027	4.4090	4.4002	1.36	1.91	3.22	9.04

The datasets used in this study include the 96-degree monthly gravity field models ITSG-Grace2018 (April 2002–June 2017) and ITSG-Grace_operational (June 2018–December 2022), both released by the Institute of Geodesy, Graz University of Technology. These models span the period from April 2002 to December 2022, with a total of 34 missing months, accounting for 13.7% of the full dataset. Hydrological models are commonly employed to validate gravity field variations caused by changes in soil moisture, near-surface temperature, snow cover, and other hydrological variables. In this study, the Global Land Data Assimilation System (GLDAS) Noah model is used, with a temporal resolution of one month, a spatial resolution of 1°×1°, and a temporal coverage from April 2002 to May 2023, consistent with the GRACE/GRACE-FO datasets.

Following the recommendations of Yi and Sneeuw (2021), the window length is set to 48 months, and the reconstruction order is determined using the following expression:

$$r' = \begin{cases} \left[4 \times \left(1 - \left(\frac{m}{10}\right)^2\right)^2 + 3\right] \times \left[1 - \frac{1}{2}\left(\frac{l}{96}\right)^2\right], m \le 40 \\ 3 \times \left[1 - \frac{1}{2}\left(\frac{l}{96}\right)^2\right], m > 40 \end{cases} \quad (26)$$

where m and l denote degree and order, and the optimal reconstruction order r is obtained by rounding r'. Figure 10 shows the reconstruction order distribution for degree 96, indicating that the interpolated values closely match the valid observations. Lower-degree coefficients use higher reconstruction orders due to stronger signal components, while higher-degree coefficients use smaller orders because of weaker signals and dominant noise—consistent with theoretical expectations.

To comprehensively evaluate the interpolation performance of the proposed method, analyses are conducted in the frequency, spatial, and temporal domains. Figure 11 presents the results for coefficients C_{22} and S_{22}; Figure 12 compares the global spatial results for April, September, and December across three periods: 2017 (GRACE), 2018 (interpolated), and 2019 (GRACE-FO). Figure 13 shows the mass variation time series over the Tibetan Plateau (TP) and the Yangtze River Basin (YRB), demonstrating excellent consistency between the interpolated and observed data.

To further assess the interpolation performance, the GRACE-derived results were compared with mass variations from the GLDAS Noah hydrological model. Six representative river basins—the Amazon, Congo,

Table 4 Main properties of ISSA, efficient ISSA and ESSA.

Method	Criterion	Advantages	Limitations
ISSA	Minimum-norm criterion in the frequency domain	Produces optimal PC estimates	High computational complexity; lower accuracy in reconstructed signals
Efficient ISSA	Minimum-norm criterion in the frequency domain	Produces optimal PCs with much higher efficiency	Slightly lower accuracy in reconstructed signals
ESSA	Low-rank approximation criterion in the temporal domain	Produces more accurate reconstructed signals	Lower accuracy in PC estimates; time complexity depends on reconstruction order

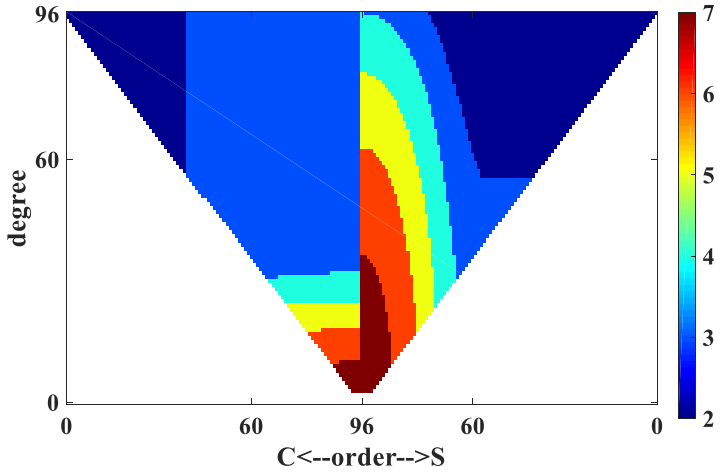

Fig. 10 Determined reconstruction order.

Niger, Orange, Orinoco, and Ob—were analyzed during the missing period as shown in Figure 14. The correlation coefficients between the ESSA-interpolated GRACE data and GLDAS estimates are 0.71, 0.68, 0.53, 0.58, 0.76, and 0.79, respectively. These strong correlations demonstrate that the ISSA-based interpolation provides reliable and robust estimates of mass variations.

7. CONCLUSIONS

This study bridges the theoretical gaps in Ji et al. (2025) by presenting a mathematical proof of equivalence between ISSA and efficient ISSA, along with a detailed time complexity analysis which theoretically demonstrates that efficient ISSA is more computationally efficient than ISSA. Additionally, we compare efficient ISSA with ESSA in terms of computation time, PC estimates, and signal extraction using real and synthetic GNSS position time series. The results show that efficient ISSA runs much faster and produces more accurate PC estimates, while ESSA maintains a slight advantage in signal extraction. When applied iteratively, both methods yield significantly improved filtering performance, with the efficient ISSA delivering comparable reconstruction quality with a lower computational burden. Finally, we demonstrate that the proposed method can also be used to fill data gaps in incomplete time series, as illustrated with GRACE/GRACE-FO gravity field solutions.

ACKNOWLEDGMENTS

This study is sponsored by the National Natural Science Foundation of China (42374017)

DATA AVAILABILITY

The GNSS position time series of 27 stations are provided by the China Earthquake Administration and are available upon reasonable request from the corresponding authors.

REFERENCES

Ars, J.-M. et al.: 2024, Geophysical models integration using principal component analysis: application to unconventional geothermal exploration. Geophys. J. Int., 239, 3, 1789–1798. DOI: 10.1093/gji/ggae357

Bos, M.S., Fernandes, R.M.S., Williams, S.D.P. and Bastos, L.: 2013, Fast error analysis of continuous GNSS observations with missing data. J. Geod., 87, 4, 351–360. DOI: 10.1007/s00190-012-0605-0

Chen, J.L., Wilson, C.R., Blankenship, D. et al.: 2009, Accelerated Antarctic ice loss from satellite gravity measurements. Nat. Geosci., 2, 12, 859–862. DOI: 10.1038/ngeo694

Cormen, T.H., Leiserson, C.E., Rivest, R.L. and Stein, C.: 2022, Introduction to algorithms. MIT Press.

Cucci, D.A., Voirol, L., Kermarrec, G., Montillet, J.P. and Guerrier, S.: 2023, The generalized method of wavelet moments with eXogenous inputs: A fast approach for the analysis of GNSS position time series. J. Geod., 97, 2, 14. DOI: DOI: 10.48550/arXiv.2206.09668

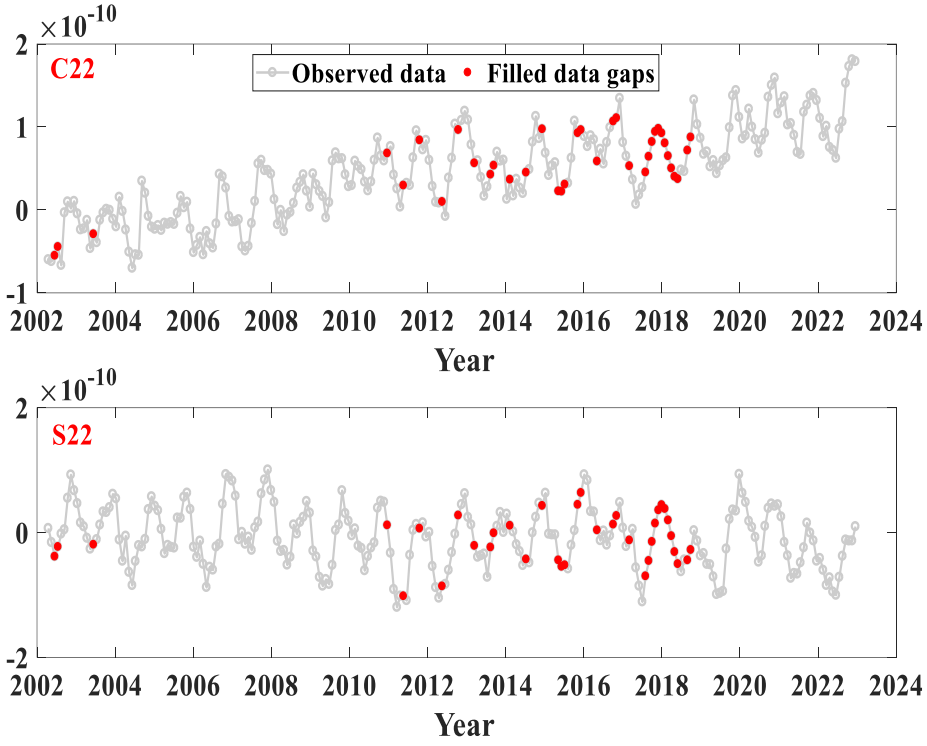

Fig. 11 Interpolation results in the frequency domain.

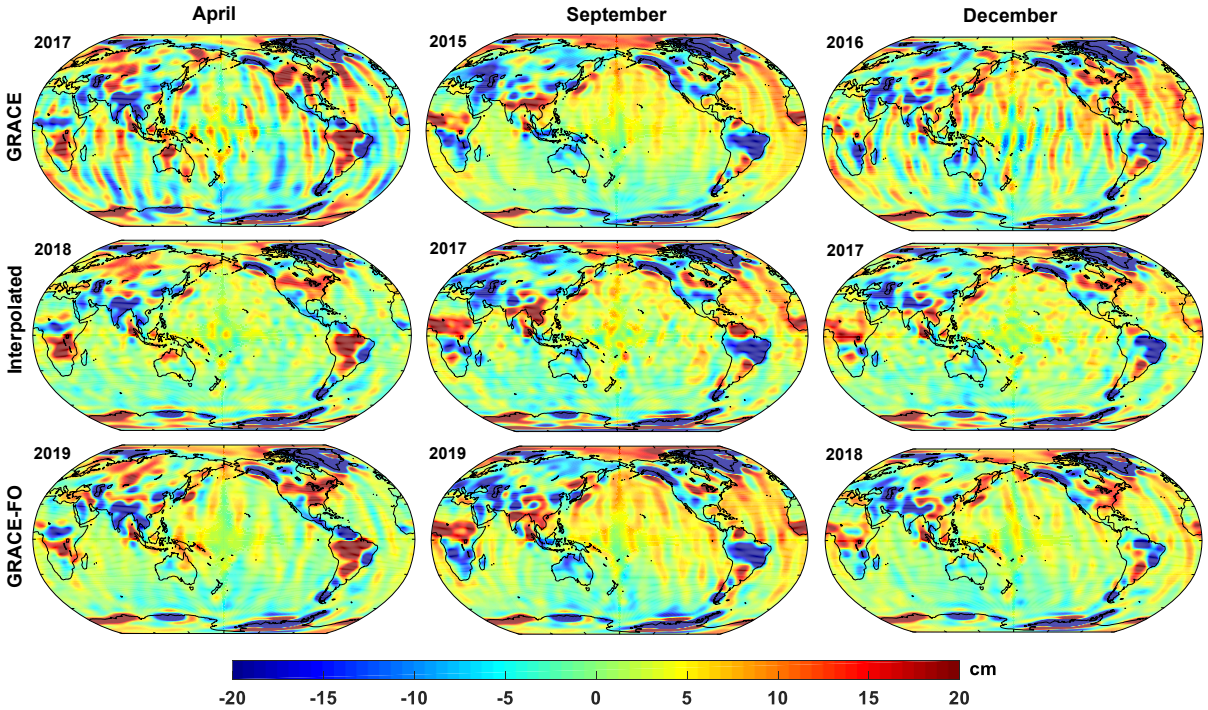

Fig. 12 Interpolation results in the spatial domain.

Davis, J.L., Wernicke, B.P. and Tamisiea, M.E.: 2012, On seasonal signals in geodetic time series. J. Geophys. Res., 117, B1, B01403. DOI: 10.1029/2011JB008690

Dong, D. et al.: 2006, Spatiotemporal filtering using principal component analysis and Karhunen-Loeve expansion approaches for regional GPS network analysis. J. Geophys. Res., Solid Earth, 111, B3. DOI: 10.1029/2005JB003806

Gao, C.C., Lu, Y., Zhang, Z.Z. et al.: 2015, Ice sheet mass balance in Antarctica measured by GRACE and its uncertainty. Chin. J. Geophys., 58, 3, 780–792. DOI: 10.6038/cjg20150308

Gobron, K., Rebischung, P., Chanard, K. and Altamimi, Z.: 2024, Anatomy of the spatiotemporally correlated noise in GNSS station position time series. J. Geod., 98, 5, 34. DOI: 10.1007/s00190-024-01848-z

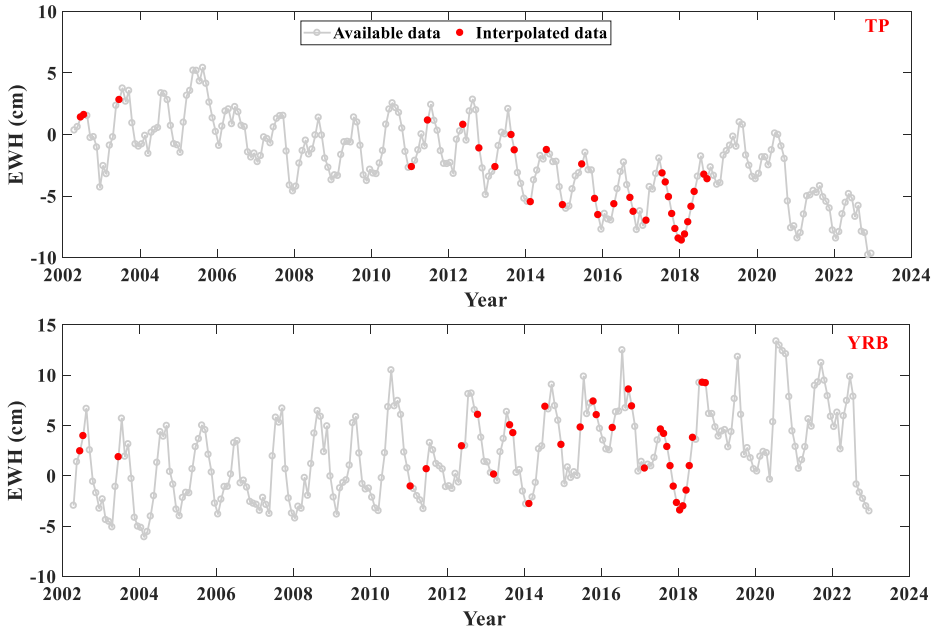

Fig. 13 Interpolation results in the temporal domain.

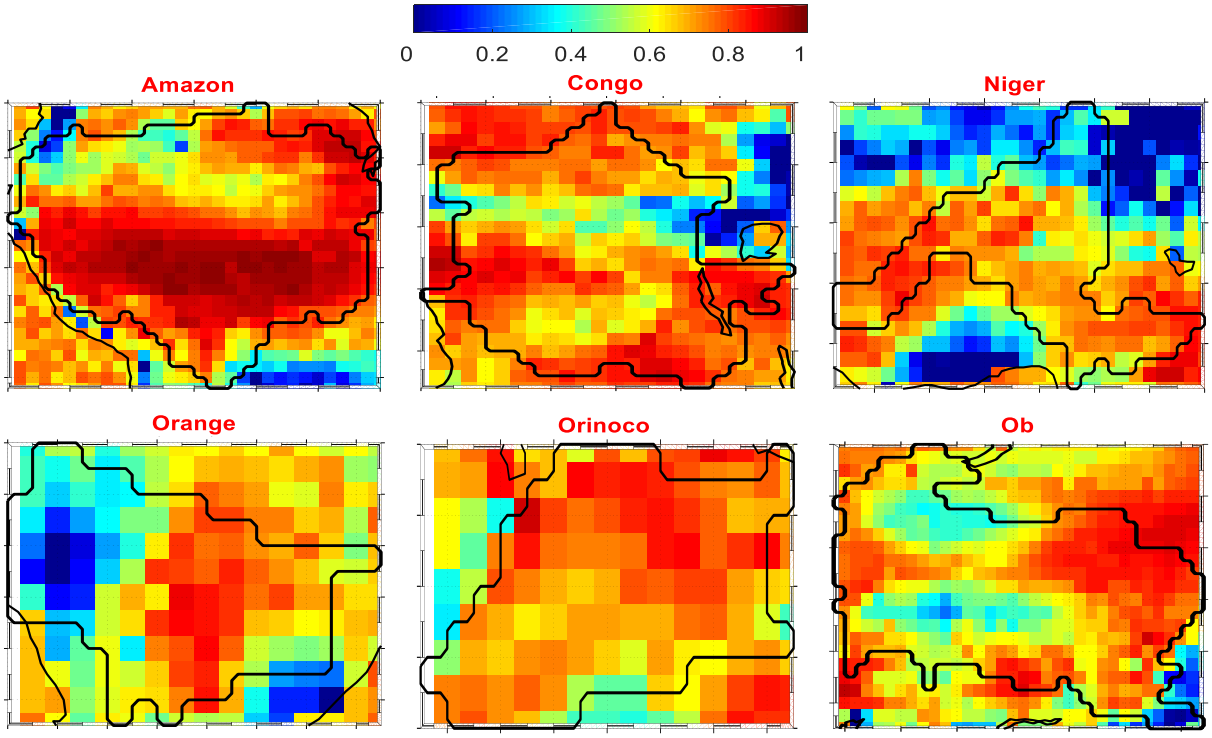

Fig. 14 Correlation coefficients between the interpolation results and the GLDAS Noah results over six river basins.

Golyandina, N. and Zhigljavsky, A.: 2013, Singular spectrum analysis for time series. Springer Briefs in Statistics. Springer. DOI: 10.1007/978-3-642-34913-3

Gruszczynska, M. et al.: 2016, Investigation of time-changeable seasonal components in the GPS height time series: A case study for Central Europe. Acta Geodyn. Geomater., 13, 3, 281–289. DOI: 10.13168/AGG.2016.0010

Gruszczynska, M. et al.: 2017, Deriving common seasonal signals in GPS position time series by using multichannel singular spectrum analysis. Acta Geodyn. Geomater., 14, 3, 267–278. DOI: 10.13168/AGG.2017.0010

Guo, J., Li, W., Chang, X., Zhu, G., Liu, X. and Guo, B.: 2018, Terrestrial water storage changes over Xinjiang extracted by combining Gaussian filter and

multichannel singular spectrum analysis from GRACE. Geophys. J. Int., 213, 1, 397–407. DOI: 10.1093/gji/ggy006

He, X. et al.: 2019, Investigation of the noise properties at low frequencies in long GNSS time series. J. Geod., 93, 9, 1271–1282. DOI: 10.1007/s00190-019-01244-y

Huang, Z., Hou, Z., Huang, J., Sun, X., He, X., Chen, H. and Montillet, J.P.: 2023, A new adaptive WVS based denoising method on GNSS vertical time series. Acta Geodyn. Geomater. 20, 2(210). DOI: 10.13168/AGG.2023.0007

Iwamori, H. et al.: 2017, Classification of geochemical data based on multivariate statistical analyses: Complementary roles of cluster, principal component, and independent component analyses. Geochem. Geophys. Geosyst., 18, 3, 994–1012. DOI: 10.1002/2016GC006663

Ji, K., Shen, Y., Chen, Q. and Wang, F.: 2023, Extended singular spectrum analysis for processing incomplete heterogeneous geodetic time series. J. Geod., 97, 8, 74. DOI: 10.1007/s00190-023-01764-8

Ji, K., Shen, Y., Wang, F. and Chen, Q.: 2025, An efficient improved singular spectrum analysis for processing GNSS position time series with missing data. Geophys. J. Int., 240, 1, 189–200. DOI: 10.1093/gji/ggae381

Kaczmarek, A. and Kontny, B.: 2018, Estimates of seasonal signals in GNSS time series and environmental loading models with iterative Least-Squares Estimation (iLSE) approach. Acta Geodyn. Geomater., 15, 2(190), 131–141. DOI: 10.13168/AGG.2018.0009

Klos, A., Bos, M. S. and Bogusz, J.: 2018, Detecting time-varying seasonal signal in GPS position time series with different noise levels. GPS Solut., 22, 1, 1–11. DOI: 10.1007/s10291-017-0686-6

Klos, A., Bos, M.S., Fernandes, R.M. and Bogusz, J.: 2019, Noise-dependent adaption of the Wiener filter for the GPS position time series. Math. Geosci., 51, 1, 53–73. DOI: 10.1007/s11004-018-9760-z

Klos, A., Dobslaw, H., Dill, R. and Bogusz, J.: 2021, Identifying the sensitivity of GPS to non-tidal loadings at various time resolutions: Examining vertical displacements from continental Eurasia. GPS Solut., 25, 3, 1–17. DOI: 10.1007/s10291-021-01135-w

Klos, A., Kusche, J., Leszczuk, G., Gerdener, H., Schulze, K., Lenczuk, A. and Bogusz, J.: 2023, Introducing the idea of classifying sets of permanent GNSS stations as benchmarks for hydrogeodesy. J. Geophys. Res., 128, 9, e2023JB026988. DOI: 10.1029/2023JB026988

Li, W and Shen, Y.: 2018, The consideration of formal errors in spatiotemporal filtering using principal component analysis for regional GNSS position time series. Remote Sens., 10, 4, 534. DOI: 10.3390/rs1004053

Ming, F., Yang, Y., Zeng, A. and Jing, Y.: 2016, Analysis of seasonal signals and long-term trends in the height time series of IGS sites in China. Sci. China Earth Sci., 59, 6, 1283–1291. DOI: 10.1007/s11430-016-5285-9

Ming, F., Yang, Y., Zeng, A. and Zhao, B.: 2017, Spatiotemporal filtering for regional GPS network in China using independent component analysis. J. Geod., 91, 4, 419–440. DOI: 10.1007/s00190-016-0973-y

Penrose, R.: 1955, A generalized inverse for matrices. Math. Proc. Camb. Philos. Soc., 51, 3, 406–413. DOI: 10.1017/s0305004100030401

Qiu, X. et al.: 2022, Iteration empirical mode decomposition method for filling the missing data of GNSS position time series. Acta Geodyn. Geomater., 19, 4, 271–279. DOI: 10.13168/AGG.2022.0012

Ruan, D. et al.: 2024, Identificaion of groundwater pollution sources and health risk assessment in the Songnen Plain based on PCA-APCS-MLR and trapezoidal fuzzy number-Monte Carlo stochastic simulation model. J. Hydrol., 632, 130897. DOI: 10.1016/j.jhydrol.2024.130897

Schoellhamer, D.H.: 2001, Singular spectrum analysis for time series with missing data. Geophys. Res. Lett., 28, 16, 3187–3190. DOI: 10.1029/2000GL012698

Shen, Y., Peng, F. and Li, B.: 2015, Improved singular spectrum analysis for time series with missing data. Nonlinear Proc. Geophys., 22, 4, 371–376. DOI: 10.5194/npgd-1-1947-2014

Syed, T.H., Famiglietti, J.S. and Chambers, D.P.: 2009, GRACE-based estimates of terrestrial freshwater discharge from basin to continental scales. J. Hydrometeorol., 10, 1, 22–40. DOI: 10.1175/2008JHM993.1

Syed, T.H., Famiglietti, J.S., Rodell, M. et al.: 2008, Analysis of terrestrial water storage changes from GRACE and GLDAS. Water Resour. Res., 44, 2. DOI: 10.1029/2006WR005779

Tamisiea, M.E.: 2011, Ongoing glacial isostatic contributions to observations of sea level change. Geophys. J. Int., 186, 3, 1036–1044. DOI: 10.1111/j.1365-246X.2011.05116.x

Vautard, R. and Ghil, M.: 1989, Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. Physica D, 35, 395–424. DOI: 10.1016/0167-2789(89)90077-8

Wang, F. et al.: 2018, Singular spectrum analysis for heterogeneous time series by taking its formal errors into account. Acta Geodyn. Geomater., 15, 4. DOI: 10.13168/AGG.2018.0029

Xiang, Y. et al.: 2022, Monte Carlo MSSA to modeling GNSS position time series and its implication for crustal deformation monitoring. Acta Geodyn. Geomater., 19, 3, 237–253. DOI: 10.13168/AGG.2018.0029

Xu, C. and Yue, D.: 2015, Monte Carlo SSA to detect time-variable seasonal oscillations from GPS-derived site position time series. Tectonophysics, 665, 118–126. DOI: 10.1016/j.tecto.2015.09.029

Xu, C.: 2016, Reconstruction of gappy GPS coordinate time series using empirical orthogonal functions. J. Geophys. Res., 121, 12, 9020–9033. DOI: 10.1002/2016JB013188

Xu, C., Cui, Y. and Wang, X.: 2025, Velocity estimation of GNSS coordinate time-series accounting for stochastic seasonality revisited. Geophys. J. Int., 243, 2, ggaf350. DOI: 10.1093/gji/ggaf350

Yi, S. and Sneeuw, N.: 2021, Filling the data gaps within GRACE missions using singular spectrum analysis. J. Geophys. Res., 126, 5, e2020JB021227. DOI: 10.1029/2020JB021227

Yi, S. and Sneeuw, N.: 2022, A novel spatial filter to reduce north–south striping noise in GRACE spherical harmonic coefficients. J. Geod., 96, 4, 23. DOI: 10.1007/s00190-022-01614-z

Zhang, L. et al.: 2023, Influence factors and mechanisms of 2015–2016 extreme flood in Pearl River Basin based on the WSDI from GRACE. J. Hydrol. Reg. Stud., 47, 101376. DOI: 10.1016/j.ejrh.2023.101376