Acta Geodyn. Geomater., Vol. 19, No. 4 (208), 271–279, 2022
DOI: 10.13168/AGG.2022.0012

journal homepage: https://www.irsm.cas.cz/acta

ORIGINAL PAPER

ITERATION EMPIRICAL MODE DECOMPOSITION METHOD FOR FILLING THE MISSING DATA OF GNSS POSITION TIME SERIES

Xiaomeng QIU [1], Fengwei WANG [2] *, Yunqi ZHOU [3] and Shijian ZHOU [4]

[1] *Gandong College, Fuzhou, PR, China*
[2] *State Key Laboratory of Marine Geology, Tongji University, Shanghai, PR, China*
[3] *School of Geographical Sciences, University of Bristol, Bristol, United Kingdom*
[4] *Nanchang Hangkong University, Nanchang, PR, China*

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

ARTICLE INFO

Article history:
Received 24 April 2022
Accepted 9 October 2022
Available online 25 October 2022

Keywords:
GNSS
Iteration EMD
Missing data
Interpolation

ABSTRACT

The Global Navigation Satellite System (GNSS) can provide the daily position time series for the geodesy and geophysical studies. However, due to various unpredictable factors, such as receiver failure or bad observation conditions, missing data inevitably exist in GNSS position time series. Most traditional time series analysis methods require the time series should be completed. Therefore, filling the missing data is a valuable step before analyzing the GNSS time series. In this study, a new method named Iteration Empirical Mode Decomposition (Iteration EMD) is proposed to fill the missing data in GNSS position time series. The simulation experiments are performed by randomly removing different missing percentages of the synthetic time series, with the added different types noise. The results show that Iteration EMD approach performs well regardless of high or low missing percentage. When the missing percentage increases from 5 % to 30 % with a step of 5 %, all the Root Mean Square Errors (RMSE) and Mean Absolute Errors (MAE) of Iteration EMD are smaller than Interpolation EMD. The relative improvements at different percentages of Iteration EMD relative to Interpolation EMD are significant, especially for the high missing percentage. The real GNSS position time series of eight stations were selected to further evaluate the performance of Iteration EMD with an average missing percentage 8.15 %. Principal Component Analysis (PCA) was performed on the filled time series, which is used to assess the interpolation performance of Iteration EMD and Interpolation EMD. The results show that Iteration EMD can preserve variance 75.9 % with the first three Principal Components (PC), more than 66.5% of interpolation EMD. Therefore, we can conclude that Iteration EMD is an efficient interpolation method for GNSS position time series, which can make full use of available information in existing time series to fill the missing data.

1. INTRODUCTION

With the continuous development of GNSS, high-precision GNSS position time series can provide valuable basic data for the geodesy and geodynamics studies. GNSS observations have been widely used in different fields, such as region crustal deformation, global plate tectonic movement, post-glacial rebounded and other geophysical phenomena (Wang and Gu, 2013; Uzel et al., 2013; Chen et al., 2013; Jiang et al., 2018). In all these applications, certain geophysical signal components can be extracted from GNSS position time series using the parametric or non-parametric model. However, most of these modeling methods such as Wavelet Transform (Miller and Shirzaei, 2015), Independent Component Analysis (ICA) (Feng et al., 2021) and PCA (He et al., 2015), require the GNSS position time series should be completed without missing data. Unfortunately, due to a range of reasons, such as equipment changes or failures, bad observation conditions and gross errors detection, missing data are very common in GNSS position time series (Xu, 2016). Therefore, it is important to find a suitable approach to process the GNSS position time series with missing data.

There exist some studies for processing the incomplete GNSS position time series, which can be divided into three categories: 1) Filling the missing data before analyzing the incomplete time series; 2) Directly using an improved time series analysis method for processing the incomplete time series without interpolating the missing data (Wang et al., 2016); 3) Filling the data gaps iteratively with zero or mean value as the initial values (Kondrashov and Ghil, 2006). Utilizing some interpolation approaches to fill the missing data is normally more considered (Zhang and Long, 2021), which is an important pre-processing step before analyzing the time series. There exist some conventional interpolation methods, such as nearest neighborhood interpolation, linear interpolation and cubic spline interpolation (Goudarzi et al., 2013). However, the performance of filling data gaps using these interpolation approaches is basically depend on the lengths of time series and gaps and availability of neighboring data (Wang et al., 2021). Recently, some

Cite this article as: Qiu X, Wang F, Zhou Y, Zhou S: Iteration empirical mode decomposition method for filling the missing data of GNSS position time series. Acta Geodyn. Geomater., 19, No. 4 (208), 271–279, 2022. 0. DOI: 10.13168/AGG.2022.0012

new interpolation methods are adopted to fill the missing data, which include Expectation Maximization (EM) (Nguyen, 2021), regularized EM (RegEM) (Schneider, 2001), Kriged Kalman Filter (Li et al., 2018) and missForest (Emmanuel et al., 2021; Zhang et al., 2021). The performance of EM, RegEM and Kriged Kalman Filter usually depends on the tuning parameters or specification of a parametric model and makes assumptions about the data distribution, such as uniform or normal distribution (Zhang et al., 2021). MissForest is computationally expensive to implement in high dimensions (Tang and Ishwarm, 2017).

Many researchers focus on processing the incomplete time series without interpolation, probabilistic PCA (pPCA) (Gruszczynski et al., 2018) and Variational Bayesian PCA (VBPCA) (Li et al., 2020) are used to estimate and extract Common Mode Error (CME) from the incomplete multiple-station GNSS time series. Shen et al. (2015) developed an improved Singular Spectrum Analysis (SSA) for processing the incomplete time series and obtained a satisfactory performance. Choi et al. (2018) considered an improvement of Variational Mode Decomposition (VMD) in the presence of missing data. Though the above-mentioned methods perform well for analyzing the incomplete time series, there still exist the problem that some significant information may be missed (Wang et al., 2016). PCA is applicable to large networks, not for a single station (Shen et al., 2014). SSA has a phase shift phenomenon and is easily affected by the problem of window length and reconstruction order parameter (Vautard and Yiou, 1992). VMD presets the number of mode functions and quadratic penalty factors based on experience, so it is difficult to achieve better decomposition effect for the measured GNSS time series (Chen et al., 2021).

Empirical Mode Decomposition (EMD) is suitable to process the nonlinear and non-stationary time series for extracting signals (Huang et al., 1998). The traditional EMD cannot directly process the incomplete time series. Self-consistency is an effective method for filling the missing data, which combined with EMD to analyze the incomplete time series produces stable decomposition results (Kim and Oh, 2016; Tarpey and Flury, 1996). As a non-parameter decomposition method, EMD can decompose the original time series into a range of Intrinsic Mode Functions (IMF) based on the time domain characteristics of original time series (Ma et al., 2022). Each IMF represents an independent component (Nelsen et al., 2018), which includes various signals, such as seasonal signals and trend term (Qiu et al., 2021). EMD, a spectrum analysis method, is more in line with the characteristics of signal for low frequency and noise for high frequency. Considering that missing data can be interpolated with iteration methods, we proposed a new approach to fill the missing data, which we have named Iteration EMD in this contribution. In addition, the performance of this method for interpolating the missing data is assessed with the comparison to Interpolation EMD, which performs the traditional EMD method after filling the missing data using the cubic spline (Dyer, 2001).

The remaining sections of this paper are organized as follows, the details of Iteration EMD are presented in Section 2. Simulation experiments of missing data interpolation efficiency are presented in Section 3. The Iteration EMD for filling the missing data in real position time series are presented in Section 4. The conclusions are summarized in Section 5.

2. METHODOLOGY

2.1. TRADITIONAL EMPIRICAL MODE DECOMPOSITION

EMD is an adaptive time-frequency analysis method, which is mainly used for nonlinear and non-stationary time series. The basic principle of EMD is to decompose the time series into several IMFs and a residual sequence. An IMF should satisfy two conditions: (1) For the whole data set, the number of extrema and the number of zero crossings must either equal or differ at most by one; (2) At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero. Essentially, the solution process of EMD is a "sieve" process, from which the IMFs with frequencies from high to low are obtained, and finally a monotonic residual sequence that can't be decomposed remains. The detailed procedures of EMD are presented as follows:

1. Find out all the local maxima points and local minima points of the original time series $x(t)$ and use the cubic spline interpolation to obtain the upper envelopes $\mu_{\max}(t)$ and lower envelopes $v_{\min}(t)$. The mean series $m_1(t)$ is generated by averaging the upper envelopes and lower envelopes.

$$m_1(t)=\frac{\mu_{\max}(t)+v_{\min}(t)}{2} \tag{1}$$

2. The difference between the original time series and $m_1(t)$ can be expressed:

$$h_1(t)=x(t)-m_1(t) \tag{2}$$

3. Judging whether $h_1(t)$ satisfies the two basic conditions of IMF. If it is satisfied, set $h_1(t)$ as the first IMF component of the original time series. If not, repeat steps 1 and 2, until it satisfies the two conditions of IMF. The first IMF component is denoted as $c_1(t)$. However, in actual time series decomposition, it is difficult to strictly satisfy the condition that the average of upper envelope and lower envelope equals zero for the IMF components. The threshold

expression for IMF component to stop filtering is as follows:

$$SD = \sum_{t=1}^{N} \left[\frac{|h_{k-1}(t) - h_k(t)|^2}{h_{k-1}^2(t)} \right] \tag{3}$$

where $h_{k-1}(t)$ and $h_k(t)$ are the two adjacent data sequences in the screening process of each IMF component, N denotes the length of time series, SD is the threshold for stopping filtering of each IMF component, which usually set 0.2 ~ 0.3 (Huang et al., 1998).

4. $c_1(t)$ is separated from the original time series to generate a new series.

$$r_1(t) = x(t) - c_1(t) \tag{4}$$

5. Treat $r_1(t)$ as a new time series and repeat the above steps to obtain n IMF components and a residual sequence only when the residual sequence satisfies the monotonic condition. Hence, the original time series can be expressed as:

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t) \tag{5}$$

where n denotes the number of IMF components, i is the index of IMF component, $r_n(t)$ represents the residual sequence.

2.2. ITERATION EMPIRICAL MODE DECOMPOSITION FOR FILLING THE MISSING DATA

Normally, there exist some missing data in GNSS position time series due to some reasons such as bad receiver, power failures and other aspects. For the EMD method, how to fill the missing data accurately is valuable for analyzing the GNSS position time series. In this study, we propose a new method named Iteration EMD to analyze the GNSS position time series. For the GNSS position time series, we assume that the complete time series $\boldsymbol{x}(t), t = 1, 2, \cdots, N$ consists of observed data and missing data. The specific steps of Iteration EMD algorithm are as follows,

1. All the missing epochs are first filled with the mean value of the observed epochs to construct a new complete time series $x'(t)$.
2. According to Equations (1)-(5), traditional EMD can decompose $x'(t)$ into a series of IMF components with frequency from high to low and one residual term. Generally, the boundary IMF component is determined by using correlation coefficient criterion during the process of noise reduction (Jia et al., 2015). When the correlation coefficient between the IMF component and the time series $x'(t)$ gets the minimum value for the first time, the corresponding IMF component is the boundary IMF component. Then the signals $\hat{x}'(t)$ are reconstructed by summing the IMF components after the boundary IMF component and residual term.

$$\hat{x}'(t) = \sum_{i=k+1}^{n} c_i(t) + r_n(t) \tag{6}$$

where k denotes the index of boundary IMF component.

3. The epochs of missing data are replaced by the corresponding part of reconstructed signals $\hat{x}'(t)$. The difference of the filled missing data $\Delta\hat{x}'(t)_{\text{miss}}$ between two following iterations should satisfy the following condition.

$$|\Delta\hat{x}'(t)_{\text{miss}}| \le \varepsilon \tag{7}$$

where ε denotes a small value, which is used to terminate the iterative procedure. Through experimental comparison, it is suggested to be chosen as 0.2 in this study.
If the condition of Equation (7) is not met, then update the missing data of the time series $x'(t)$.

4. The steps 2-3 are iteratively processed using traditional EMD until meet the condition of Equation (7). Considering that with the number of iterations increases, the interpolation accuracy is not improved much, but the iteration procedure needs more time. Therefore, if the condition of Equation (7) does not satisfy, we will terminate the iteration procedure when the iteration number reaches 10.
5. Appropriate evaluation indexes are used to test the performance of Iteration EMD and Interpolation EMD.

3. SIMULATION EXPERIMENTAL ANALYSIS

To test and verify the performance of the Iteration EMD approach, simulation experiments were carried out. The observation model of synthetic position time series without nonlinear changes, such as offsets, co-seismic or post-seismic, can be expressed as follows,

$$\begin{aligned} s(t) = a + bt + c\sin(2\pi t) + d\cos(2\pi t) + \\ + e\sin(4\pi t) + f\cos(4\pi t) \end{aligned} \tag{8}$$

where, $s(t)$ represents the position at epoch t (unit of year), a is the initial position constant, b is the linear trend, c and d are the coefficients of annual periodic motion, e and f are coefficients of semi-annual periodic motion (Zhou et al., 2022). Then the time series are generated as follows,

Table 1 **S**imulated time series parameters (mm).

a	b	c	d	e	f
5	2	10	10	5	5

$$x(t) = s(t) + e(t) \tag{9}$$

where, $e(t)$ is the simulated different types noise.

The parameters presented in Table 1 are used to generate the synthetic GNSS position time series based on the Equations (8)-(9). The missing data are generated by randomly removing some epochs from the synthetic at different missing percentages. By this way, we evaluate the performance of Iteration EMD for filling the simulated incomplete time series. Since the true values of missing data are known, the RMSE and MAE of the difference between the true values and filled data gaps are computed to evaluate the performance of Iteration EMD and Interpolation EMD. The smaller RMSE and MAE are, the better the method filled the missing data.

To understand the performances of two approaches for filling the gaps at different missing percentages, we randomly remove (1) 5 %, (2) 10 %, (3) 15 %, (4) 20 %, (5) 25 %, (6) 30 % of the original time series as missing data to perform the simulation experiments. To avoid the contingency of the experimental results, we repeated the simulation experiment 300 times. The missing data are filled by Iteration IEMD and Interpolation EMD for each simulation. Besides, when adopting different missing percentages, the same processing procedure was adopted.

3.1. WHITE NOISE

In this section, we only add the white noise to the simulated signals to generate the synthetic time series, and then use Iteration EMD and Interpolation EMD to fill the missing data, respectively. Here we take a missing percentage of 15 % as an example for the detailed description, the results of one simulation experiment are shown in Figure 1. From Figure 1, it is found that the filled missing data of Iteration EMD were closer to the true signal than Interpolation EMD. The mean RMSEs and MAEs of 300 simulations at different missing percentages are presented in Figure 2. As can be seen from Figure 2, all mean RMSEs and MAEs of Iteration EMD are smaller than Interpolation EMD regardless of what the missing percentages is, indicating that Iteration EMD outperforms than Interpolation EMD for filling the missing data. Besides, the RMSE and MAE will be larger with the increasing missing percentage. When the missing percentage is up to 30 %, the mean RMSE and MAE of Interpolation EMD are 2.4 mm and 1.7 mm, obviously larger than 1.7 mm and 1.3 mm of Iteration EMD, with the relative improvements of 30.3 % and 20.1 %, respectively. Figure 3 presents the relative improvements of all RMSEs and MAEs at the different missing percentages. The mean relative improvements of RMSE and MAE for Iteration EMD with respect to Interpolation EMD are 19.5 % and 12.7 %, respectively. Therefore, it is reasonable to conclude that Iteration EMD can better filling the missing data than Interpolation EMD, mainly due to that cubic spline is basically dependent on the lengths of time series and gaps, availability of neighboring data, and so on (Semiromi and Koch, 2019).

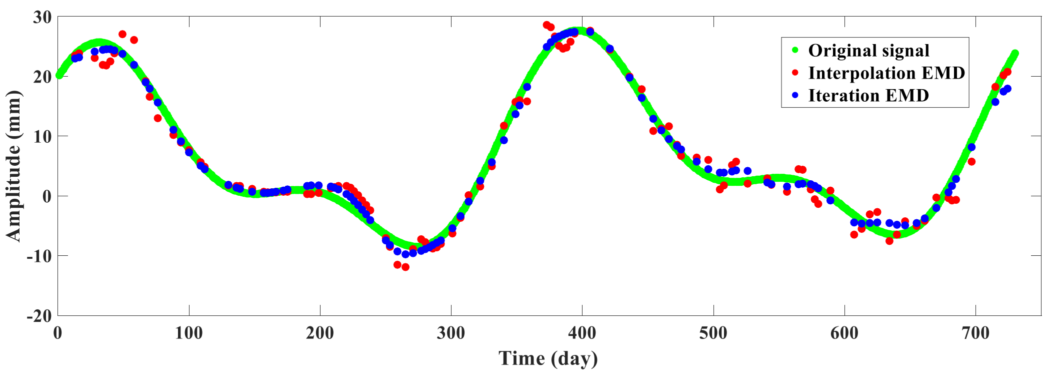

Fig. 1 The reconstructed signals from all available data and filled the missing data by Iteration EMD and Interpolation EMD at the missing percentage of 15 % (White noise).

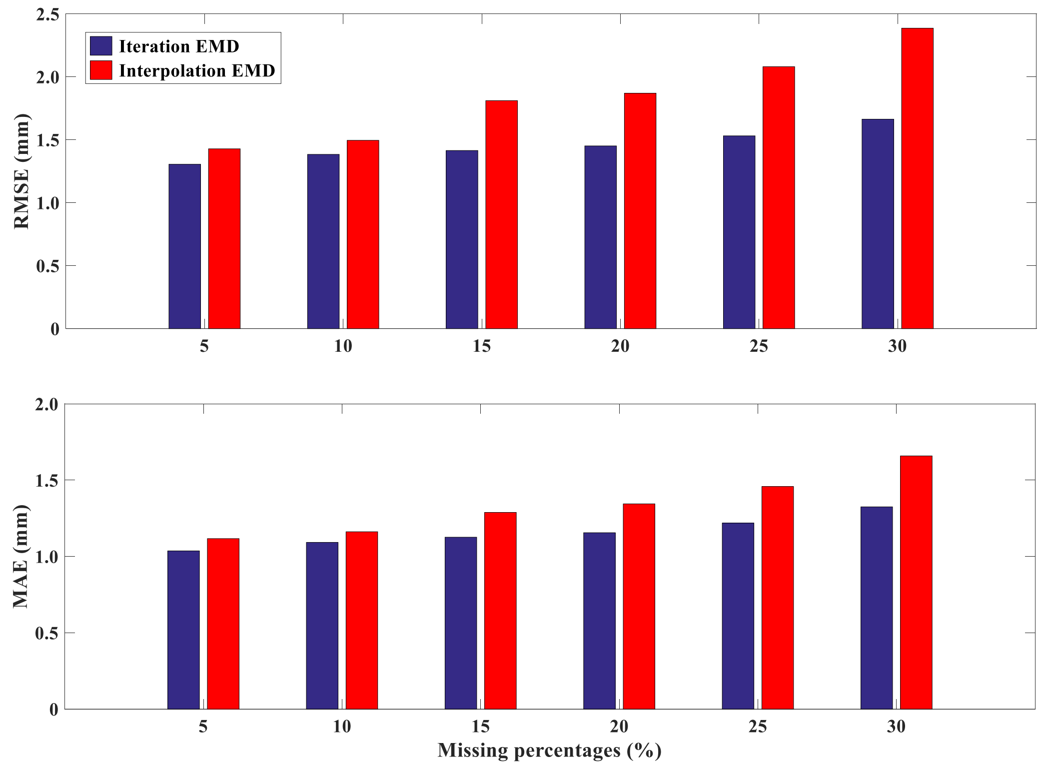

Fig. 2 Mean RMSEs and MAEs of 300 simulations at different missing percentages.

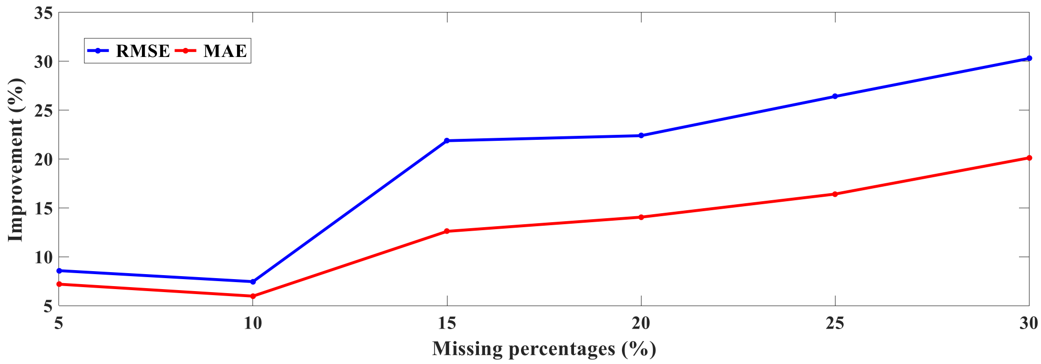

Fig. 3 The relative improvements of RMSE and MAE for Iteration EMD with respect to Interpolation EMD.

3.2. WHITE NOISE AND POWER LAW NOISE COMBINATION

Many studies found that there exists not only white noise, but also colored noise in GNSS position time series (Bogusz and Klos, 2016), and that white noise plus power law noise model is common (Bos et al., 2020). Therefore, the combination of white noise and power law noise was employed to generate the original time series. Here we randomly remove 15 % of the original time series as missing data, and fill them using the Iteration EMD and Interpolation EMD, shown in Figure 4. It can be found that the filled missing data of Iteration EMD are closer to the true signals than Interpolation EMD. The RMSE of Interpolation EMD and Iteration EMD are 2.7 mm, 2.0 mm, respectively. Figure 5 shows the mean RMSEs and MAEs of 300 simulations at different missing percentages when adding the mixed noise including white noise and power law noise, which is similarly to those of Figure 2. Figure 6 presents the relative improvements of RMSE and MAE for Iteration EMD relative to Interpolation EMD. Consequently, all results suggest similar conclusion that Iteration EMD filling the missing data more accurately than Interpolation EMD.

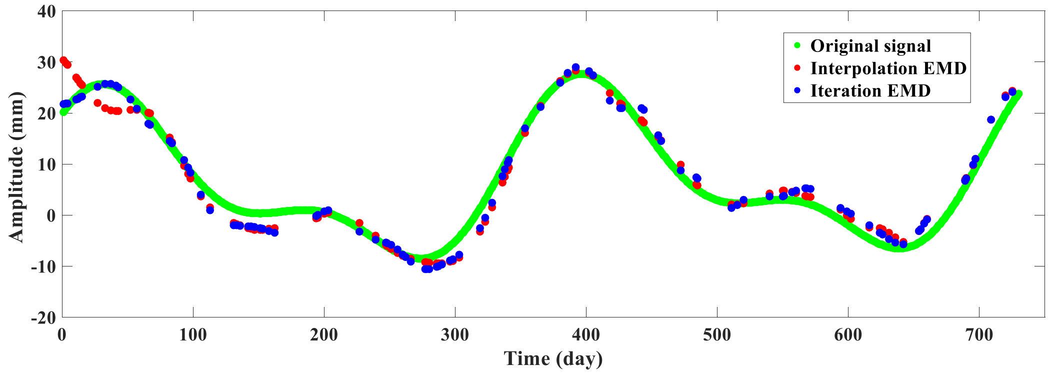

Fig. 4 The reconstructed signals from all available data and filled the missing data by Iteration EMD and Interpolation EMD at the missing percentage of 15 % (Mixed noise).

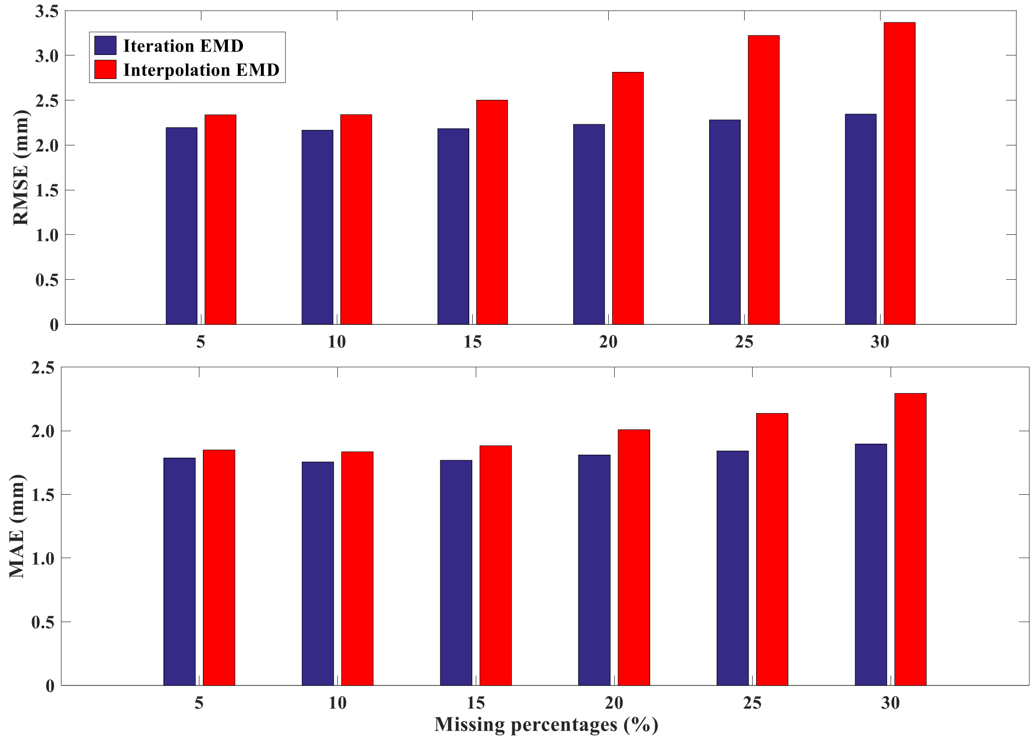

Fig. 5 Mean RMSEs and MAEs of 300 simulations at different missing percentages.

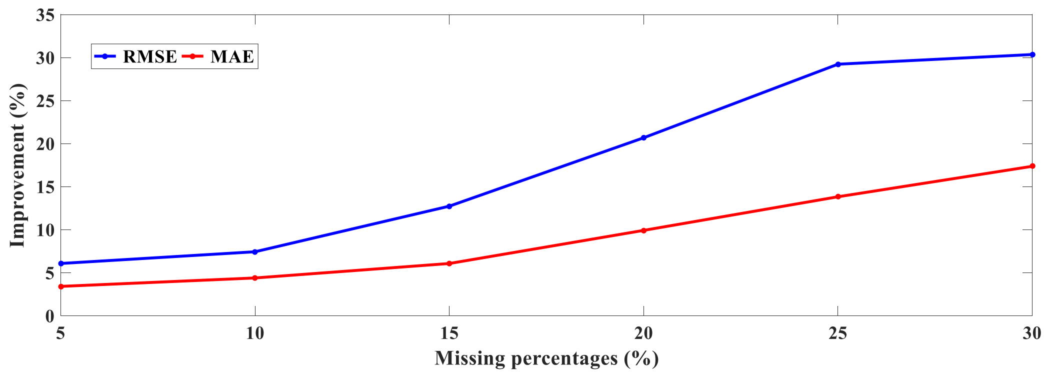

Fig. 6 The relative improvements of RMSE and MAE for Iteration EMD with respect to Interpolation EMD.

4. REAL POSITION TIME SERIES ANALYSIS

To further evaluate the performance of Iteration EMD with comparison of Interpolation EMD, the up component of GNSS position time series of eight stations over the period from 2017 to 2019 were selected to analyze, which contained missing data in each station. The position time series used in this study were downloaded from data products of National Earthquake Data Center (https://www.eqdsc.com). Outliers were eliminated by Interquartile Range (IQR) criterion (Langbein and Bock, 2004) and the missing percentages of eight stations were shown in Figure 7. From Figure 7, we can see that SXTY station has the highest missing percentage, and the missing percentages of most stations are less than 10 %. The average missing percentage of eight stations in the up component is 8.15 %.

Then the Iteration EMD and Interpolation EMD are used to fill the missing data of the position time series. Since the true values of missing data in real position time series are unknown, it is impossible to use the two above used evaluation indexes to verify the interpolation effect when analyzing the real GNSS position time series of eight stations. Similar to that the time series after interpolation should keep the original variance maximization direction as much as possible to extract the CME of the observation network accurately. We will evaluate the performance of Iteration EMD and Interpolation EMD by comparing the variance derived from the filled time series of eight stations using the PCA method (Ming et al., 2016). The equation for calculating the variance in each principal direction is as follows (Tripathi and Govindaraju, 2008).

$$v_j = w_j^{\mathrm{T}} \boldsymbol{S} w_j \tag{10}$$

where w_j denotes the j-th principal direction, $\boldsymbol{S} = \boldsymbol{X}\boldsymbol{X}^{\mathrm{T}} / (\mathrm{Nrows} - 1)$ is the covariance matrix, $\boldsymbol{X} = [x_1'(t), x_2'(t), \cdots, x_8'(t)]$ represents the complete time series after interpolation for eight stations, Nrows is the number of rows of the matrix $\boldsymbol{X}$.

To determine the interpolation efficiency of the two methods, we perform the PCA approach on the filled time series without missing data of eight stations. Table 2 shows the proportion of variance of the first three PCs in the filled time series for Iteration EMD and Interpolation EMD approaches. The first three PCs of Iteration EMD can preserve 75.9 % of total variance, more than 66.5 % of Interpolation EMD, indicating that Iteration EMD has better interpolation ability. The improvement of Iteration EMD relative to Interpolation EMD is 14.1 % for the sum of the variances with the first three PCs.

5. CONCLUSIONS

Filling the missing data is a crucial pre- processing step in GNSS time series analysis. We proposed a new approach named Iteration EMD to fill the missing data of position time series, with the comparison of Interpolation EMD. The simulation results show that all the RMSEs and MAEs of Iteration EMD are smaller than those of Interpolation EMD, indicating that Iteration EMD can fill the missing data closer to the true signal than Interpolation EMD regardless of the missing percentage. Besides, the relative improvements will increase especially for the high missing percentage. For the real GNSS position time series of eight stations, PCA approach was used to evaluate the filled position time series by Iteration EMD and Interpolation EMD. The results show that the first three leading PCs of Iteration EMD can obtain

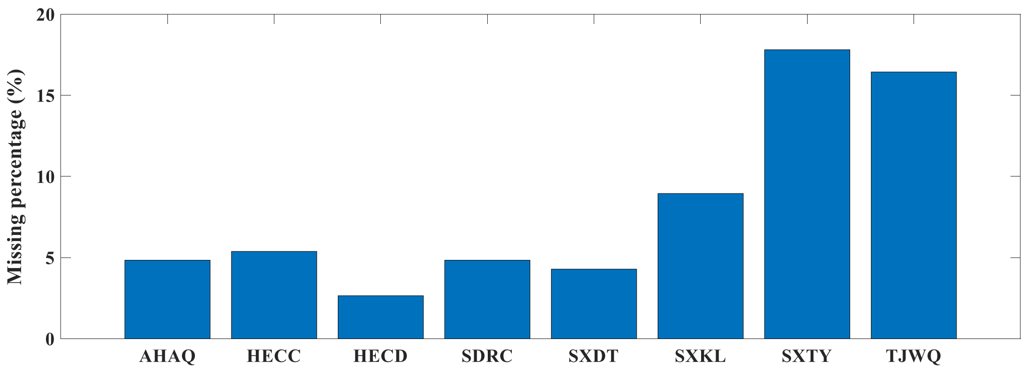

Fig. 7 The statistics missing percentages of eight stations.

Table 2 The proportion of the variance of the first three PCs of two interpolation methods (%).

Method	PC1	PC2	PC3	Sum
Iteration EMD	46.8	16.6	12.5	75.9
Interpolation EMD	37.5	15.0	14.0	66.5

75.9 % of total variance, 9.4 % larger than 66.5 % of Interpolation EMD. Therefore, it is reasonable to believe that Iteration EMD can fill the missing data more accurately and obtain more signal information than Interpolation EMD, which is beneficial to the subsequent analysis of GNSS position time series.

ACKNOWLEDGMENTS

This work is mainly funded by the Natural Science Foundation of China (42064001) and Science and Technology Project of the Education Department of Jiangxi Province (209507). We acknowledge data products of National Earthquake Data Center for providing raw position time series (https://www.eqdsc.com).

REFERENCES

Bogusz, J. and Klos, A.: 2016, On the significance of periodic signals in noise analysis of GPS station coordinates time series. GPS Solut., 20, 4, 655–664. DOI: 10.1007/s10291-015-0478-9

Bos, M.S., Montillet, J.-P., Williams, S.D.P. and Fernandes, R.M.S: 2020, Introduction to geodetic time series analysis. In: Montillet, J.-P. and Bos, M.S. (eds.), Geodetic time series analysis in Earth sciences. Springer Geophysics. DOI: 10.1007/978-3-030-21718-1_2

Chen, Q., van Dam, T., Sneeuw, N., Collilieux, X., Weigelt, M. and Rebischung, P.: 2013, Singular spectrum analysis for modeling seasonal signals from GPS time series. J. Geodyn., 72, 25–35. DOI: 10.1016/j.jog.2013.05.005

Chen, X., Yang, Z., Tian, Z., Yang, B. and Liang, P.: 2021, A denoising method for GNSS time series based on GA-VMD and multi-scale permutation entropy. Geomat. Inf. Sci. Wuhan Univ. DOI:10.13203/j.whugis20210215

Choi, G., Oh, H. and Kim, D.: 2018, Enhancement of variational mode decomposition with missing values. Signal Process., 142, 75–86. DOI: 10.1016/j.sigpro.2017.07.007

Dyer, S.A. and Dyer, J.S.: 2001, Cubic-spline interpolation.1. IEEE Instrum. Meas. Mag., 4, 1, 44–46. DOI: 10.1109/5289.911175

Emmanuel, T., Maupong, T., Mpoeleng, D., Semong, T., Mphago, B. and Tabona, O.: 2021, A survey on missing data in machine learning. J. Big Data, 8, 1, 140. DOI: 10.1186/s40537-021-00516-9

Feng, T., Shen, Y. and Wang, F.: 2021, Independent component extraction from the incomplete coordinate time series of regional GNSS networks. Sensors, 21, 5, 1569. DOI: 10.3390/s21051569

Goudarzi, M.A., Cocard, M., Santerre, R. and Woldai, T.: 2013, GPS interactive time series analysis software. GPS Solut., 17, 4, 595–603. DOI: 10.1007/s10291-012-0296-2

Gruszczynski, M., Klos, A. and Bogusz, J.: 2018, A filtering of incomplete GNSS position time series with probabilistic principal component analysis. Pure Appl. Geophys., 175, 5, 1841–1867. DOI: 10.1007/s00024-018-1856-3

He, X., Hua, X., Yu, K., Xuan, W., Lu, T., Zhang. W. and Chen, X.: 2015, Accuracy enhancement of GPS time series using principal component analysis and block spatial filtering. Adv. Space Res., 55, 5, 1316–1327. DOI: 10.1016/j.asr.2014.12.016

Huang, N.E., Long, S.R., Wu, M.L.C., Shih, H.H., Zheng, Q.N., Yen, N.C., Tung, C.C., Liu, H.H. and Shen, Z.: 1998, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. A: Math., Phys. Eng. Sci., 454, 1971, 903–995. DOI: 10.1098/rspa.1998.0193

Jia, R., Zhao, T., Sun, H. and Yan, X.: 2015, Microseismic signal denoising method based on empirical mode decomposition and independent component analysis. Chin. J. Geophys., 58, 3, 1013–1023. DOI: 110.6038/cjg20150326

Jiang, W., Wang, K., Li, Z., Zhou, X., Ma, Y. and Ma, J.: 2018, Prospect and theory of GNSS coordinate time series analysis. Geomat. Inf. Sci. Wuhan Univ., 43, 12, 2112–2123. DOI: 10.13203/j.whugis20180333

Kim, D. and Oh, H.S.: 2016, Empirical mode decomposition with missing values. Springerplus, 5, 1, 2016. DOI: 10.1186/s40064-016-3692-1

Kondrashov, D. and Ghil, M.: 2006, Spatio-temporal filling of missing points in geophysical data sets. Nonlinear Process. Geophys., 13, 2, 151–159. DOI: 10.5194/npg-13-151-2006

Langbein, J. and Bock, Y.: 2004, High-rate real-time GPS network at Parkfield: Utility for detecting fault slip and seismic displacements. Geophys. Res. Lett., 31, 15. DOI: 10.1029/2003GL019408

Li, N., Dai, W., Santerre, R. and Kuang, C.: 2018, A MATLAB-based Kriged Kalman Filter software for interpolating missing data in GNSS coordinate time series. GPS Solut., 22, 1, 1–8. DOI: 10.1007/s10291-017-0689-3

Li, W., Jiang, W., Li, Z., Chen, H., Chen, Q., Wang, J. and Zhu, G.: 2020, Extracting common mode errors of regional GNSS position time series in the presence of missing data by variational Bayesian principal component analysis. Sensors, 20, 8, 2298. DOI: 10.3390/s20082298

Ma, Y., Cheng, J., Wang, P., Wang, J. and Yang, Y.: 2022, A novel Lanczos quaternion singular spectrum analysis method and its application to bear fault diagnosis with multi-channel signals. Mech. Syst. Signal Process., 168, 108679. DOI: 10.1016/j.ymssp.2021.108679

Miller, M.M. and Shirzaei, M.: 2015, Spatiotemporal characterization of land subsidence and uplift in Phoenix using InSAR time series and wavelet transform. J. Geophys. Res., Solid Earth, 120, 8, 5822–5842. DOI: 10.1002/2015JB012017

Ming, F., Zeng, A., Gou, W. and Xu, K.: 2016, Analysis of the sensitivity of data-driven interpolation algorithm to GNSS coordinate time series. Geomatics Science and Engineering, 36, 5, 4–9, (in Chinese).

Nelsen, B., Williams, D.A., Williams, G.P. and Berrett, C.: 2018, An empirical mode-spatial model for environmental data imputation. Hydrology, 5, 4, 63. DOI: 10.3390/hyrology5040063

Nguyen, L.: 2021, Handling missing data with expectation maximization algorithm. GRDJE, 6, 11, 9–32.

Qiu, X., Wang, F., Zhou, Y., Zhou, S.: 2021, Weighted empirical mode decomposition for processing GNSS position time series with the consideration of formal errors. Acta Geodyn. Geomater., 18, 3(203), 399–408. DOI: 10.13168/AGG.2021.0028

Schneider, T.: 2001, Analysis of incomplete climate data: estimate of mean values and covariance matrices and imputation of missing values. J. Clim., 14, 5, 853–871. DOI: 10.1175/1520-0442(2001)014

Semiromi, M.T. and Koch, M.: 2019, Reconstruction of groundwater levels to impute missing values using singular and multichannel spectrum analysis: application to the Ardabil Plain, Iran. Hydrol. Sci. J., 64, 14, 1711–1726. DOI: 10.1080/02626667.2019.1669793

Shen, Y., Li, W., Xu, G. and Li, B.: 2014, Spatiotemporal filtering of regional GNSS network's position time series with missing data using principal component analysis. J. Geod., 88, 1, 1–12. DOI: 10.1007/s00190-013-0663-y

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

Tang, F. and Ishwarm, H.: 2017, Random forest missing data algorithms. Stat. Anal. Data Min., 10, 6, 363–377. DOI: 10.1002/sam.11348

Tarpey, T. and Flury, B.: 1996, Self-consistency: A fundamental concept in statistics. Stat. Sci., 11, 3, 229–243. DOI: 10.1214/ss/1032280215

Tripathi, S. and Govindaraju, R.S.: 2008, Engaging uncertainty in hydrologic data sets using principal component analysis: BaNPCA algorithm. Water Resour. Res., 44, 10. DOI: 10.1029/2007WR006692

Uzel, T., Eren, K., Guial, E., Tiryakioglu, I., Dindar, A.A. and Yilmaz, H.: 2013, Monitoring the tectonic plate movement in Turkey based on the national continuous GNSS network. Arab. J. Geosci., 6, 9, 3573–3580. DOI: 10.1007/s12517-012-0631-5

Vautard, R., Yiou, P. and Ghil, M.: 1992, Singular spectrum analysis: A toolkit for short, noisy chaotic signals. Phys. D: Nonlinear Phenom., 58, 1–4, 95–126. DOI: 10.1016/0167-2789(92)90103-T

Wang, X., Cheng, Y., Wu, S. and Zhang, K.: 2016, An effective toolkit for the interpolation and gross error detection of GPS time series. Surv. Rev., 48, 348, 202–211. DOI: 10.1179/1752270615Y.0000000023

Wang, W. and Gu, G.: 2013, Advantages of GNSS in monitoring crustal deformation for detection of precursors to strong earthquakes. Positioning, 4, 1, 11–19. DOI: 10.4236/pos.2013.41003

Wang, F., Shen, Y., Chen, Q. and Wang, W.: 2021, Bridging the gap between GRACE and GRACE follow-on monthly gravity field solutions using improved multichannel singular spectrum analysis. J. Hydrol., 594, 125972. DOI: 10.1016/j.jhydrol.2021.125972

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

Zhang, S., Gong, L., Zeng, Q., Li, W., Xiao, F. and Lei, J.: 2021, Imputation of GPS coordinate time series using missforest. Remote Sens., 13, 12, 2312. DOI: 10.3390/rs13122312

Zhang, Y. and Long, Q.: 2021, Fairness in missing data imputation. Proc. 38th International Conference on Machine Learning, PMLR 139.

Zhou, W., Ding, K., Liu, P., Lan, G. and Ming, Z: 2022, Spatiotemporal filtering for continuous GPS coordinate time series in mainland China by using independent component analysis. Remote Sens., 14, 2904. DOI: 10.3390/rs14122904