https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023

Nonlinear Processes in Geophysics Discussions
Open Access
EGU

Extraction of periodic signals in GNSS vertical coordinate time series using adaptive Ensemble Empirical Modal Decomposition method

Weiwei Li[1], Jing Guo[1]

[1] College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao, 266590, China

Correspondence to: Weiwei Li (whereareou@126.com)

Abstract. Ensemble Empirical Mode Decomposition (EEMD) has been widely used in the data analysis. Adaptive EEMD further improves computational efficiency through the adaptability in the white noise amplitude and set average number. However, its effectiveness of the periodic signal extraction in Global Navigation Satellite System (GNSS) coordinate time series regarding on the inevitable missing data and offsets issues have not been comprehensively validated. It is verified with 5- year time series through 300 simulations for each case. The results show that high accuracy could reach for the overall random missing rate below 15% and avoiding consecutive missing epochs exceeding 30. Meanwhile, offsets should be corrected in advance regardless of their magnitudes. The analysis of vertical component of 13 stations within the Australian Global Sea Level Observing System (GLOSS) monitoring network, demonstrates the advantage in revealing the time-varying characteristics of periodic signals. From the perspectives of correlation coefficients, power spectral density (PSD) indices and signal noise ratio (SNR), the means for adaptive EEMD are 0.36, -0.18 and 0.48, respectively, while for least squares (LS), they are 0.27, -0.50 and 0.23. Meanwhile, significance test of the residuals further substantiate the effectiveness in periodic signal extraction, which shows there is no annual signal remained. Also, the longer the series, the higher accuracy of extracted periodic signal is reasonable concluded via the significance test. Furthermore, the time-varying periodic characteristics is more conducive to analyze the driving factors. Overall, the application of adaptive EEMD could achieve high accuracy in analyzing GNSS time series, but it should be based on the proper dealing with missing data and offsets.

1 Introduction

Global Navigation Satellite System (GNSS) coordinate time series are a crucial type of geophysical data. They can not only be used to obtain accurate positions and velocities of stations (Blewitt and Lavallée, 2002), but also for better studying Earth's dynamic processes such as crustal movement (Manevich et al., 2021), post-ice rebound and sea level change (Md Din et al., 2019), inversion mass change (Willen et al., 2022), and so on (Dong et al., 2002). From these data, we can observe significant periodic variations in addition to long-term trends and noise. These periodic variations are caused by a complex interplay of the Earth's own physical properties and environmental factors, which is particularly obvious in the elevation

https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023


Nonlinear Processes in Geophysics Discussions

direction (Davis et al., 2012; Jiang et al., 2013; Wang et al., 2018). Therefore, accurately extracting periodic signals from GNSS coordinate time series holds significant importance (Van Dam et al., 2001; Sorin et al., 2021). However, in previous studies, the periodic signals extracted by the commonly used LS method did not agree with the actual periodic variations (Bennett, 2008). Therefore, many scholars have conducted studies on periodic signal extraction methods, some of which include Kalman filtering, assuming signals random, which is contrary to the real GNSS time series containing power-law noise (Davis et al., 2012). Also, the high sensitivity of F-test detection to noise may lead to the incorrect determination of the presence of periodic signals (Li and Shen, 2014). The Principal Component Analysis (PCA) method found that most GNSS stations around the world have a nonlinear periodic pattern, with annual and semi-annual terms dominating (Zhang et al., 2019). These studies verified the existence of time-varying-amplitude periodic signals in GNSS time series and separated the signals according to different methods. However, the accurate extraction of tectonic motion information containing various noise, errors and non-tectonic periodic signals from actual observation still necessitates the continuous exploration of novel approaches and methods.

This paper employs an adaptive Ensemble Empirical Mode Decomposition (EEMD) method for signal decomposition capable of decomposing nonlinear and nonstationary signals into multiple physically well-defined Intrinsic Mode Functions (IMFs) with amplitude-varying characteristics over time (Xu, 2022). Adaptive EEMD has significant advantages in terms of noise immunity, stability of IMFs, avoidance of local extremes, adaptivity and consistency of results. It is suitable for signal components of different scales and frequencies, which improves the stability and reliability of decomposition (Wu and Huang, 2009). This method is widespread in data processing. However, there are some limitations, such as high requirements of the completeness of the time series (Agnieszka and Dawid, 2022), and offsets which usually occur in GNSS coordinate time series have not yet been comprehensively assessed. The trend in vertical direction of GNSS stations nearby the tide gauges is essential in obtaining the absolute sea level. However, the trend is affected by the periodic signals contained in GNSS time series (Bos et al., 2010; Klos et al., 2018). Therefore, it is of paramount significance to get the periodic signal of high accuracy for the stations belong to SONEL (Système d'Observation du Niveau des Eaux Littorales), which serves as the GNSS data center for the Global Sea Level Observing System. Therefore, this paper applies adaptive EEMD to the periodicity assessment of Australian stations in GLOSS. Section 2 describes the principles in detail, Section 3 verifies its effectiveness in synthetic time series, and 300 simulations were conducted for each of the experiments in the focused missing data and offset issues. The practical analysis of for the selected 13 stations located in Australia in GLOSS are carried out in Sect. 4, and the conclusions are summarized in Sect. 5.

2 Principle and Method

Due to the uncertain quantity of IMFs, there may be issues of frequency overlap and mode mixing among different IMFs, which can affect the decomposition and analysis of Empirical Modal Decomposition (EMD) signals (Huang et al., 1999). The employed EEMD method effectively mitigates the impact of mode mixing phenomena, enhancing the precision and

https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023


Nonlinear Processes in Geophysics Discussions

robustness of signal decomposition (Zhang et al., 2009). This method is to add white noise to the original signal and use the uniform distribution of Gaussian noise to alter the distribution of extreme points, ensuring signal continuity within each frequency band. Subsequently, the components of the multiple decomposed IMFs are averaged to mitigate the impact of random noise (Huang et al., 1917; Yi et al., 2022). The specific steps are as follows (Peng et al., 1971; Liu et al., 2023):

Step 1: Add the same length Gaussian white noise sequence to the data $x(t)$ multiple times.

$$X_i(t) = x(t) + u \cdot a_i(t)\,, \tag{1}$$

Where u is the amplitude coefficient of white noise added, and $a_i(t)$ is the white noise added for the ith time, where $i = 1,2,\dots,m,\ t = 1,2,\dots,T$.

Step 2: Decompose the original signal $X_i(t)$ with noise to obtain the IMF components of each order by EMD.

$$X_i(t) = \sum_{j=1}^{n} c_{i,j}\,(t) + r_{i,n}(t)\,, \tag{2}$$

Where $c_{i,j}$ is the jth IMF of EMD decomposition with white noise added for the ith time, and $r_{i,n}$ is the remainder after decomposition.

Step 3: Calculate the mean value of the corresponding IMF component obtained from each decomposition, to get the final IMF component of EEMD.

$$IMF_j(t) = \frac{1}{m}\sum_{i=1}^{m} c_{i,j}\,(t)\,, \tag{3}$$

Where $IMF_j(t)$ is the jth IMF component obtained after EEMD decomposition of the decomposed signal.

However, two parameters in Step 1 (added white noise amplitude u and set average number m) in EEMD are crucial, which directly affect the EEMD decomposition. For the too large white noise, it can mask the real signal. For the tiny white noise, its modal effect is not obvious. The larger the set average number, the smaller the decomposition error, but the increase will reduce the computational efficiency. Therefore, it is important to balance both white noise amplitude and set average number, which represents the adaptivity in EEMD (Liu et al., 2023). The adaptive EEMD method can avoid a large number of parameters setting manually and significantly improve decomposition efficiency and accuracy. However, the adaptive EEMD still suffers from boundary effects, which refers the problem of unstable or distorted decomposition due to the lack of extreme points at the end. In this paper, it is mitigated by using the extreme value extension method (Qi et al., 2003).

While, the preprocessing is also essential for EEMD application. This involves selection of enough observations in good continuity, outliers removing, offset detection and missing data filling. Here, the robust IQR method is used for outlier removal (Shu, 2021) and Iterative Empirical Mode Decomposition (IEMD) is utilized for missing data filling (Qiu et al., 2022). The specific IEMD algorithm is seen in Fig. 1. All the missing epochs are initialized with the mean value of the observed epochs to construct a new complete time series $x'(t)$. Subsequently, traditional EMD 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

https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023

Nonlinear Processes in Geophysics
Discussions
Open Access
EGU

correlation coefficient between the IMF component and the time series $x'(t)$ reaches the minimum value for the first time, the corresponding IMF component is the boundary, with k representing the boundary index. Then the signals $\hat{x}'(t)$ are reconstructed by summing the IMF components after the boundary IMF component and residual term. The epochs of missing data are replaced by the corresponding part of reconstructed signals $\hat{x}'(t)$. When the difference of missing data $|\Delta x'(t)_{miss}|$ between two iterations is smaller than ε, the iterative process is terminated, resulting in the final complete time series. Otherwise, the iteration continues. ε represents a threshold (0.005 in this paper) to terminate the iteration process.

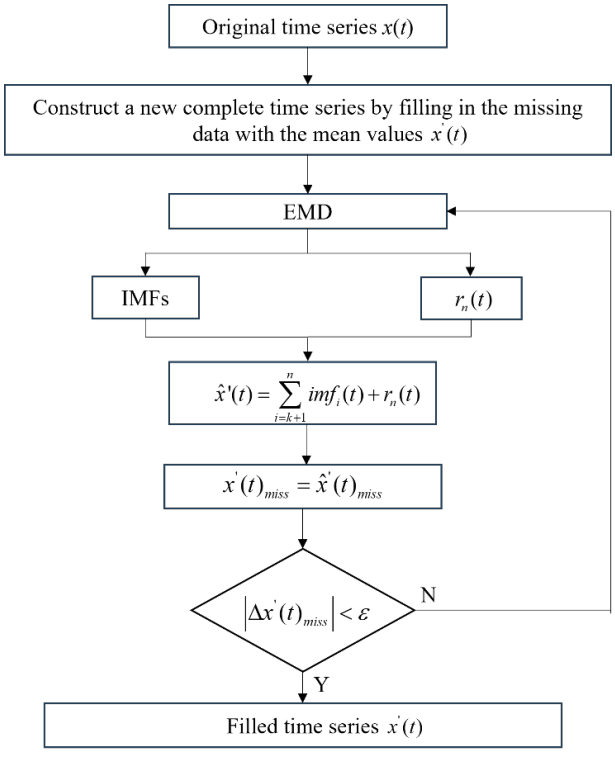


Figure 1: Flow chart of IEMD for filling.

With the complete data, the periodic signals can be extracted using adaptive EEMD method. Firstly, the time series are decomposed into multiple IMFs and corresponding residuals. Secondly, the high-frequency and low-frequency IMFs are determined by the energy value method (Zheng, 2021). Thirdly, the Lomb spectral analysis is applied to the low-frequency IMFs to identify the periodic signals (Ray et al., 2008). Finally, the IMFs containing periodic signals are selected and combined. Meanwhile, the classic least squares (LS) is employed for comparison. Generally, the function model includes annual, semi-annual, and the 1st to 6th draconitic-year periodic signals (Li et al., 2018). Specifically, the LS formulation is as follows:

$$x(t) = a + bt + \sum_{i=1}^{8}(c_i \cdot \sin(2\pi \cdot f_i \cdot t) + d_i \cdot \cos(2\pi \cdot f_i \cdot t) + n(t)), \quad (4)$$

https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023


Nonlinear Processes in Geophysics Discussions

Where t is the observation time, a is the initial position constant, b is the linear rend, c_i, d_i is the coefficient of the periodic signal (c_1 and d_1 represent the annual periodic coefficients, while c_2 and d_2 represent the semi-annual periodic coefficients), f_i is the period frequency, and $n(t)$ is the noise term.

Last but not the least, the assessment of the extracted signal reliability is crucial, which can be measured in different perspectives. Generally, it is measured by correlation coefficients, power spectral density (PSD) index and signal noise ratio (SNR) (Bos et al., 2008; Chen et al., 2020; Ran et al., 2022). Specifically, the higher correlation coefficient of the extracted periodic signal and the original time series, the more accurate extraction. The PSD index of the residuals is used to evaluate the noise characteristics and indirectly reflects the accuracy of the extracted signal, of which the lower absolute values imply more accurate extraction. Similarly, the larger the SNR, the better the periodic model established.

3 Synthetic time series analysis

To test and validate the performance of the adaptive EEMD method, simulation experiments are carried out. In a simplified scenario, only annual and semi-annual periods are simulated as follows,

$$s(t) = a + bt + \sum_{i=1}^{2}(c_i \cdot \sin(2\pi \cdot f_i \cdot t) + d_i \cdot \cos(2\pi \cdot f_i \cdot t)) , \quad (5)$$

Specifically, the 5- year data $x(t)$ is simulated, which is generated with the parameters in Table 1. To be more realistic, not only white noise but also power-law noises are included (Eq. (6)). The magnitudes of white noise $p(t)$ and power-law noise $q(t)$ are 0.6 mm and 2 $\text{mm}^{-0.6}$. That means the power-law spectral index here is -0.60.

Table 1. The Simulated parameters in synthetic time series (mm).

a	b	c_1	d_1	c_2	d_2
5	2	10	10	5	5

$$x(t) = s(t) + p(t) + q(t) , \quad (6)$$

Since missing data and offset are the general cases, these are mainly investigated in this section.

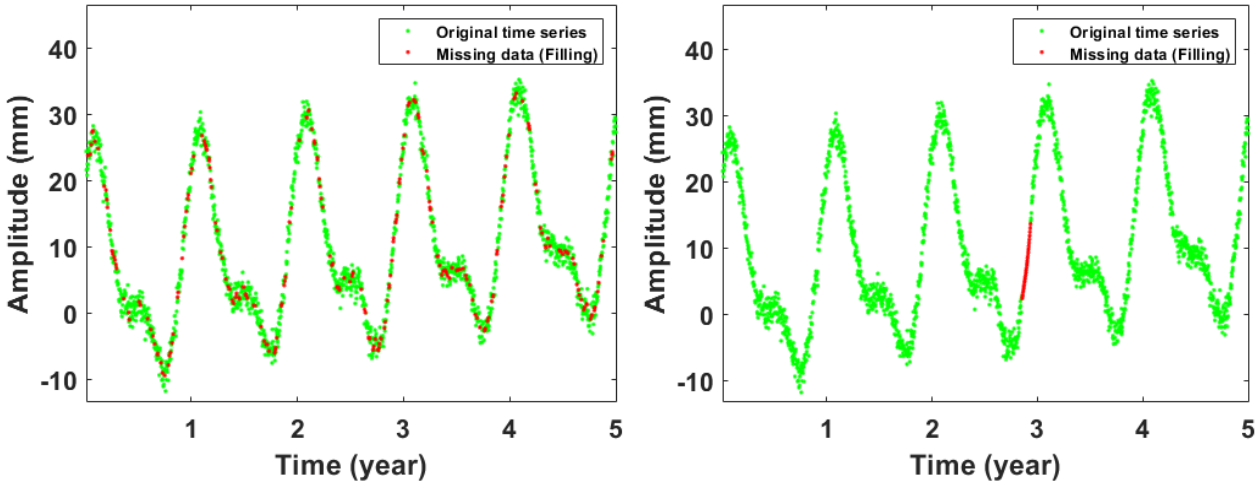
3.1 Impact of missing data on adaptive EEMD

GNSS provides critical daily position time series for geodetic and geophysical studies. However, due to unforeseeable factors such as receiver malfunctions, human errors, or deteriorating environmental conditions, the occurrence of missing data is inevitable (Bao et al., 2021). The presence of missing data can significantly impact time series analysis. Therefore, selecting an appropriate approach is crucial when addressing the issue of missing data in GNSS position time series. To address it, the IEMD method is used in data filling, due to its dual advantage of efficient and accurate. However, this could be influenced by the percentage of missing data. Given the complex actual situations, missing data might appear randomly or consecutively. For randomly missing data, remove 5%, 10%, 15% and 20% of the original data time series; for consecutively

https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023

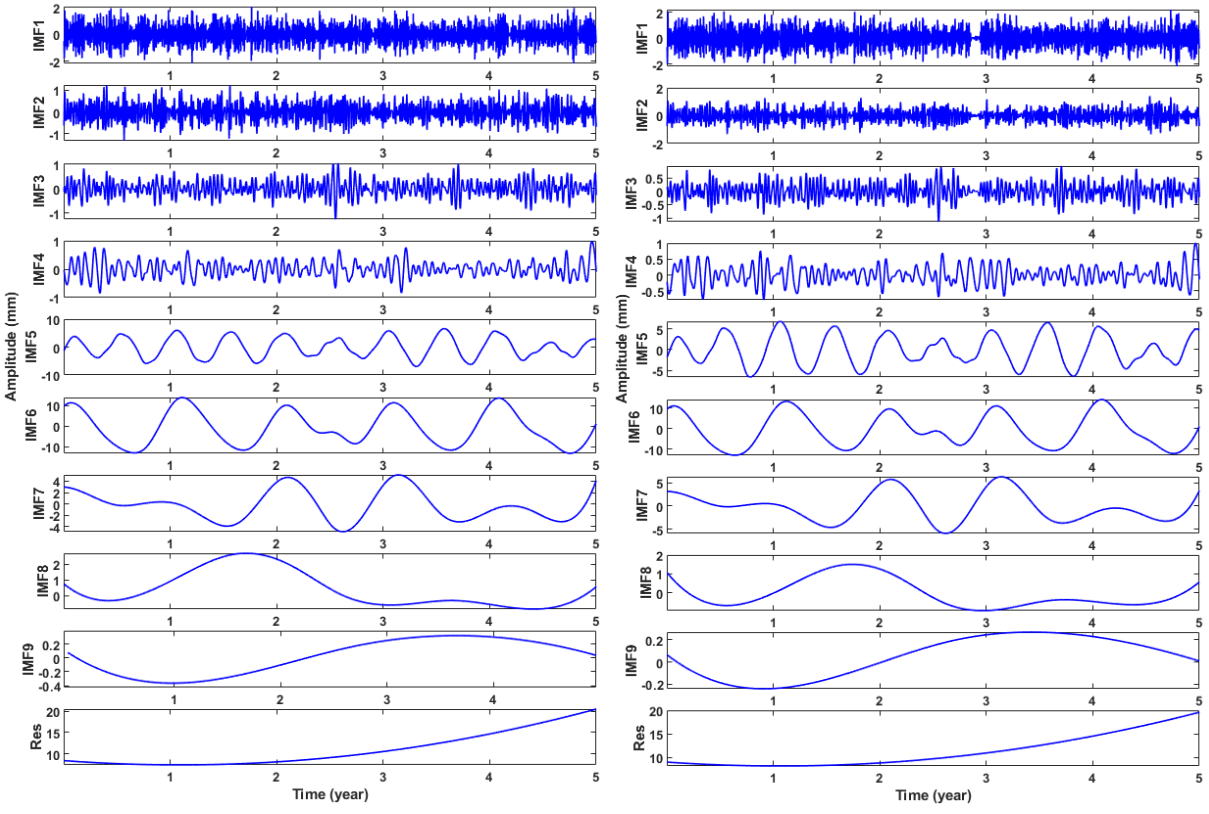
Nonlinear Processes in Geophysics
Discussions
Open Access
EGU

missing data, randomly delete 15, 30, 45 and 60 consecutive days. Here, we present an example with a missing rate of 15% and a consecutive 30 days missing (Fig.2-Fig.5).



140 **(a)** 15%missing. **(b)** 30 consecutive days missing.

Figure 2: Data filling of (a)random missing 15% and (b) consecutive missing 30 days.

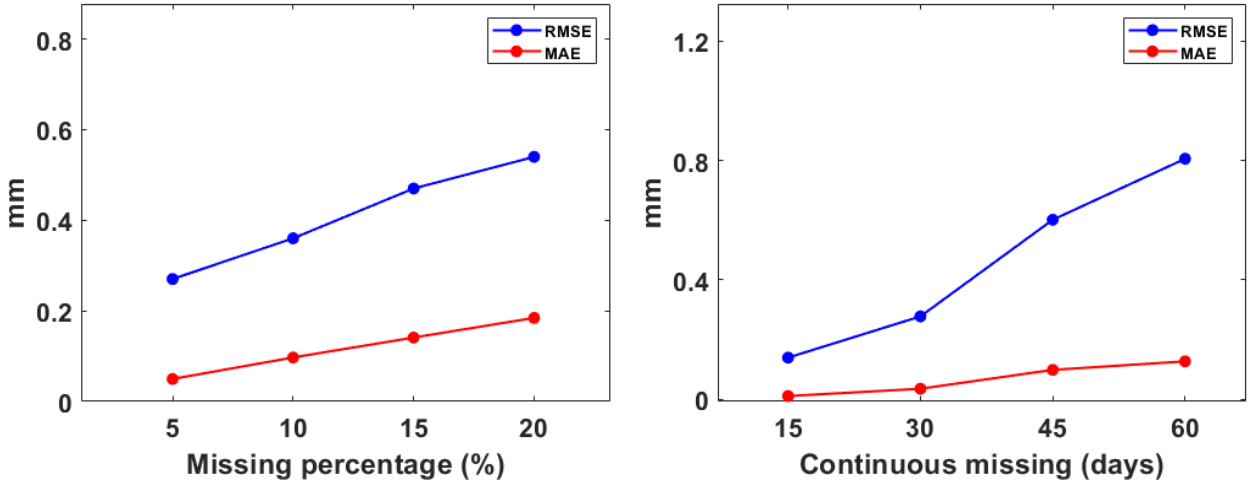


(a) 15%missing. **(b)** 30 consecutive days missing.

https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023

Nonlinear Processes in Geophysics Discussions

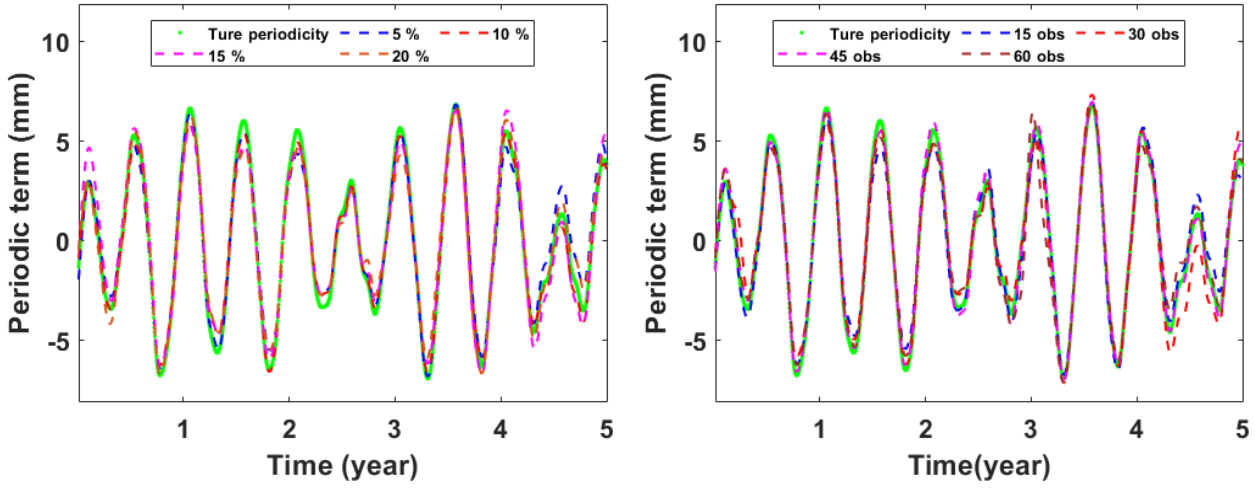
Figure 3: Decomposition of filled series with adaptive EEMD (a) random missing 15% and (b) consecutive missing 30 days.



145

(a) randomly missing. **(b)** consecutively missing.

Figure 4: RMSE and MAE of (a) randomly missing and (b) consecutively missing.



(a) randomly missing. **(b)** consecutively missing.

150 **Figure 5: The obtained periodic signal and the real periodic signal of (a) randomly missing and (b) consecutively missing.**

The results of filling in the missing data are shown in Fig. 2, which keeps the same variation of the original time series. The results of the adaptive EEMD are shown in Fig. 3. The decomposition components indicate that the simulated annual and semi-annual signals are clear. The mean RMSEs and MAEs of 300 simulations in different missing percentages are 155 presented in Fig. 4. As depicted in the figure, it can be observed that with the increase of the missing data percentage, both RMSE and MAE gradually increase, which indicates that the stability of the filling results is decreasing and may lead to the error increased. Especially when consecutively missing data points is between 30 and 45, the increase of RMSE and MAE is

https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023

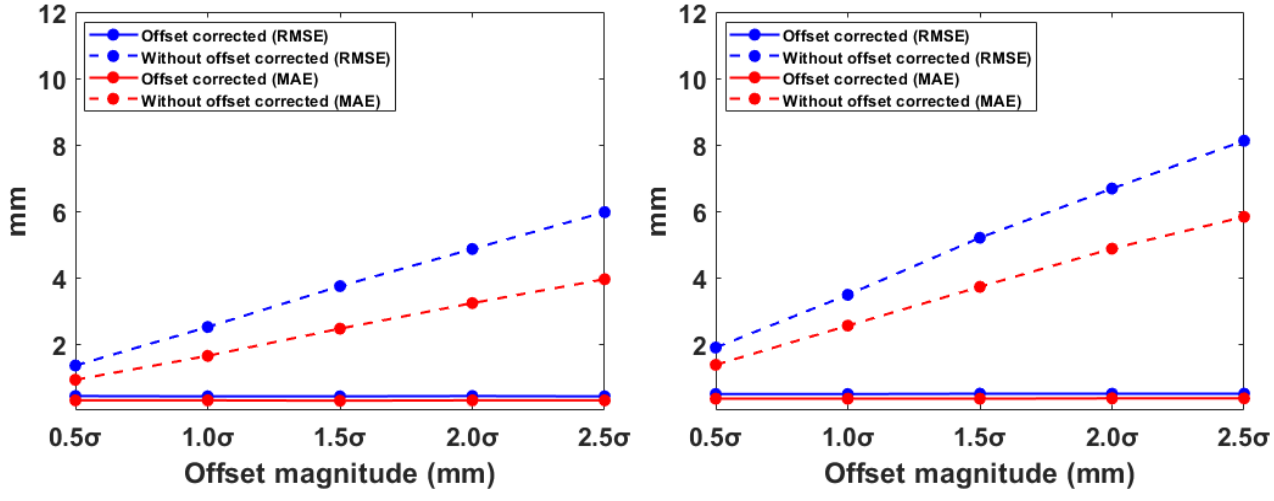
Nonlinear Processes in Geophysics
Discussions
Open Access
EGU

more pronounced. Figure 5 demonstrates the obtained periodic signal with the real periodic signal. As expected, the results indicate that with the increase in the percentage of missing data, the extracted periodic signal gradually deviates from the true values.

Therefore, to ensure the accuracy and reliability of adaptive EEMD, the overall random missing rate should be less than 15%, and the consecutive missing epochs should be less than 30.

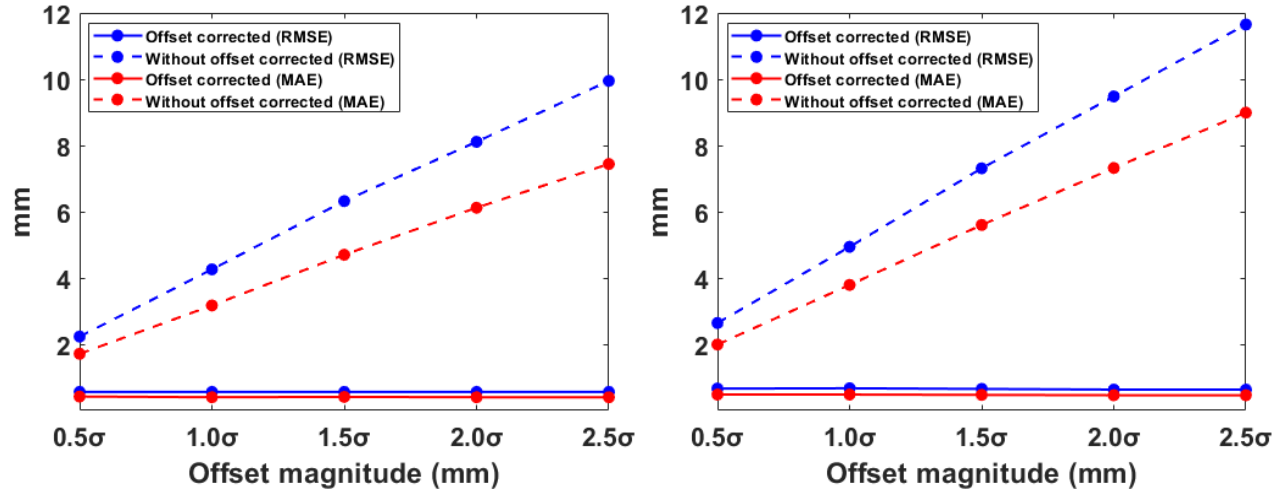
3.2 Impact of offset on adaptive EEMD

The offsets caused by satellite orbit changes, receiver hardware or software problems, data processing algorithm improvements, and other unknown factors are common in GNSS time series (Defraigne et al., 2021). For the known offsets, in data preprocessing, the magnitudes of offsets are estimated according to the epochs of earthquake or instrument replacement, then the known offsets are corrected. However, for the unknown factors, it is difficult to detect the offset time due to the outliers and noises included. While the percentage due to unknown factors can be up to 20%, its influence cannot be ignored (Griffiths et al., 2016). Therefore, it is urgent to identify and correct the offset. Although there are many methods to detect the offset (Feng et al., 2013; Wang et al., 2019), neither is 100% effective. To verify the offsets correction essential or not, tests (with and without offset correction) are carried out. Traditionally, manual visual inspection is combined with the STARS method for offset detection, and the LS method is used for offset correction (Rodionov, 2004; Cai, 2020). Since the number and magnitude are the two factors of offsets, these are variables in the simulations. Specifically, the offset number ranges from 2 to 8, with an interval of 2. The simulated offset magnitudes are 0.5σ, 1.0σ, 1.5σ, 2.0σ and 2.5σ mm (Su et al., 2023), with standard deviation (σ) of the simulated time series (σ=12.70 mm). Also, the simulations are repeated 300 times. Since our focus is the extraction of periodic signals, the effect of simulated offsets on the periodic signals is analyzed. Accordingly, the RMSE and MAE between the periodic signals obtained using the adaptive EEMD method with and without offset correction and the real periodic signal are compared. They are shown in Fig. 6.



(a) 2 offsets. **(b)** 4 offsets.

https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023



(c) 6 offsets. **(d)** 8 offsets.

Figure 6: Periodic signal RMSE and MAE with and without offset correction of different number and magnitude offsets (a) 2 offsets, (b) 4 offsets, (c) 6 offsets and (d) 8 offsets.

It is found that when no correction beforehand, RMSE and MAE rapidly increase as the magnitude and number of offsets increase. With the offset corrected, both RMSE and MAE reveal substantial decrease. This reveals that in adaptive EEMD application, the presence of offsets significantly reduce the accuracy of periodic signal extraction. From the improvement ratio in Table 2, we can see that even when the offset magnitude is 0.5σ and offset number is 2, the ratios for RMSE and MAE are 68.15% and 67.39%, respectively. Meanwhile, as offset number increases, the improvement ratios for RMSE and MAE gradually rise. This indicates the necessity of offset correction even in condition of very small offset. However, these findings are based on the all the offsets detected are accurate, although in reality, it is hard to realize. Thus, it is recommended to perform offset correction for the adaptive EEMD application in the GNSS coordinate time series analysis.

Table2. The improvement ratios of RMSE and MAE with an offset magnitude of 0.5σ.

Number of offsets	Improvement ratio (%)	
	RMSE	MAE
2	68.15	67.39
4	74.07	74.45
6	74.89	75.44
8	75.00	75.88

4 Results and analysis

4.1 Data sources

Thirteen GNSS stations located in the Australian region are selected for the study, and their locations are shown in Fig. 7. The vertical components for these selected stations can be download from the Sonel website

https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023

Nonlinear Processes in Geophysics Discussions

(https://www.sonel.org/spip.php?page=cgps&lang=en) (Łyszkowicz and Bernatowicz, 2019). Table 3 provides essential station information, including longitude and latitude coordinates, the time span, the raw missing rate (%) and the missing rate after gross error elimination (%), among which the longest time span is MAC1 station and the shortest one is BRO1 station. Additionally, the XMIS station exhibits the highest missing rate after outlier removal. The offset epochs and magnitudes available from the Sonel website, as well as offset occurred in 2011.6123 at NORF station determined with STARS method, along with its magnitude estimated by LS, are shown in Table 4.

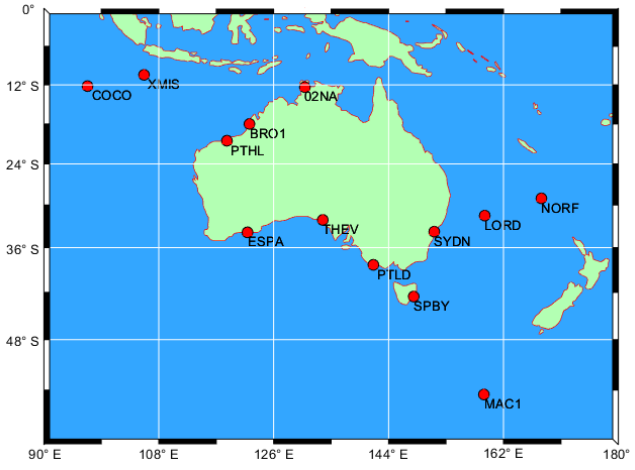


Figure 7: Location of the selected 13 GNSS stations in Australia.

Table 3. GNSS stations information used in this study.

Site	Longitude (°)	Latitude (°)	Span (year)	Length (year)	Raw missing rate (%)	Missing rate after gross error elimination (%)
THEV	133.6968	-32.1286	2014.8726-2020.1189	5.25	3.44	4.27
XMIS	105.6884	-10.4499	2005.4890-2020.9303	15.44	13.15	13.60
SYDN	151.1503	-33.7808	2004.3429-2020.9932	16.65	5.16	5.54
SPBY	147.9308	-42.5464	2008.7336-2013.9959	5.26	4.99	6.03
PTLD	141.6134	-38.3444	2009.7712-2020.9932	11.22	8.07	8.48
PTHL	118.6788	-20.5397	2016.8624-2021.9849	5.12	1.12	1.81
NORF	167.9388	-29.0433	2008.4768-2020.9959	12.52	6.18	6.31
MAC1	158.9358	-54.4995	1996.9850-2013.9959	17.01	8.30	8.98
LORD	159.0611	-31.5198	2009.4836-2020.9686	11.48	8.05	8.69
ESPA	121.8943	-33.8743	2008.4495-2020.9418	12.49	6.16	6.61
COCO	96.8339	-12.1883	1997.7521-2013.9959	16.24	7.02	8.96
BRO1	122.2090	-18.0039	2015.8808-2020.9658	5.09	6.99	8.06
02NA	130.8817	-12.3559	2008.7255-2014.0890	5.36	7.44	7.90

Table 4. Offsets of the investigated 13 GNSS stations.

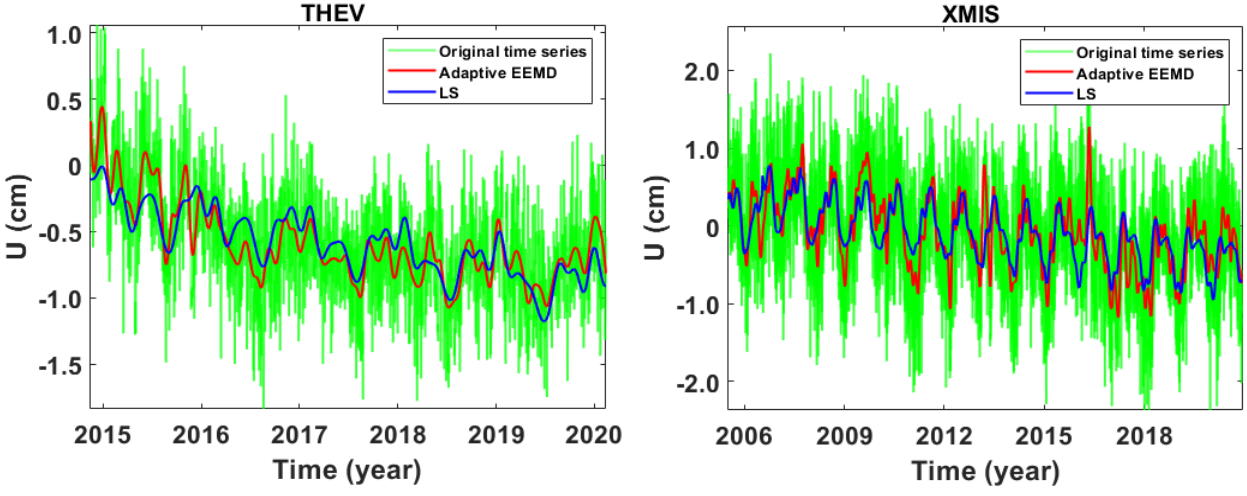
Site	Offset epoch (year)	Offset magnitude (mm)
THEV	\	\
XMIS	2006.5411\2007.6973\2012.2801\2014.4836	1.9\-2.5\0.7\-0.7
SYDN	\	\
SPBY	\	\

https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023

PTLD	2012.1052\2012.3347\2012.9331\2017.4836\2018.3192	0.3\0.9\0.5\8.1\-7.7
PTHL	\	
NORF	2011.6123\2016.7582	3.2\1.8
MAC1	\	\
LORD	\	\
ESPA	2016.1544	-4.8
COCO	\	\
BRO1	\	\
02NA	\	\

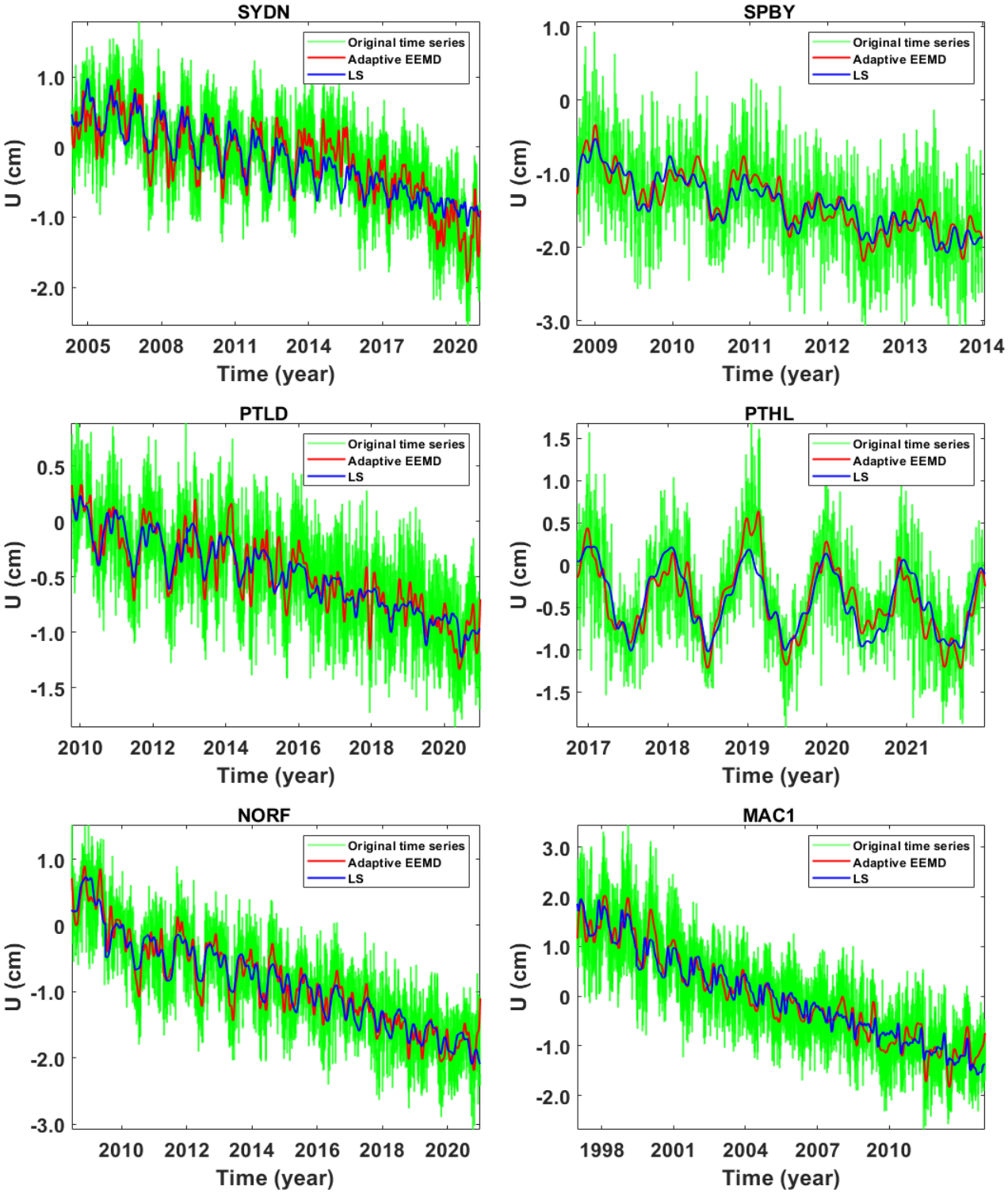
4.2 Data processing and analysis

In this section, we carry out a comparative analysis of the periodic signal in the vertical component of the 13 selected GNSS stations derived by adaptive EEMD method and the LS method. Figure 8 presents the post-processed vertical components (outliers removed, missing data filled and offset corrected) and the extracted signal. It is evident that the adaptive EEMD method provides a superior fit to the original time series due to its time-varying characteristics. For example, the signal of XMIS and ESPA demonstrate apparent peak signal in 2016 and 2015. To further evaluate its performance, Figure 9 displays comprehensive assessment indicators.

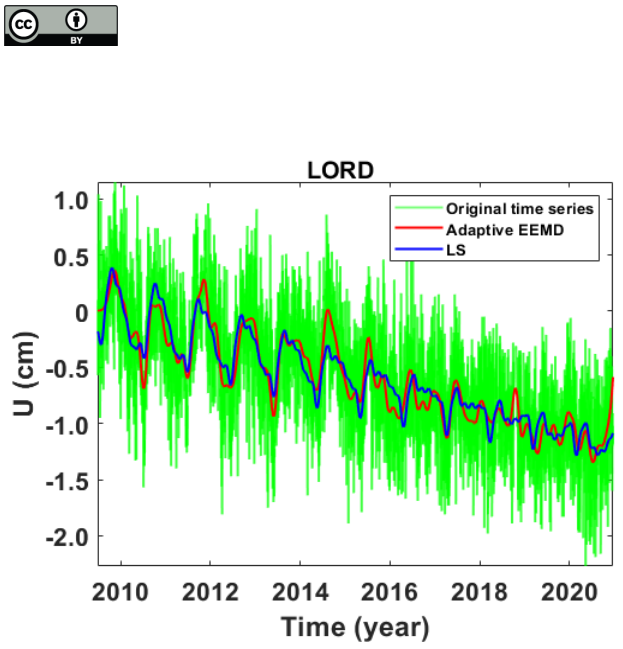


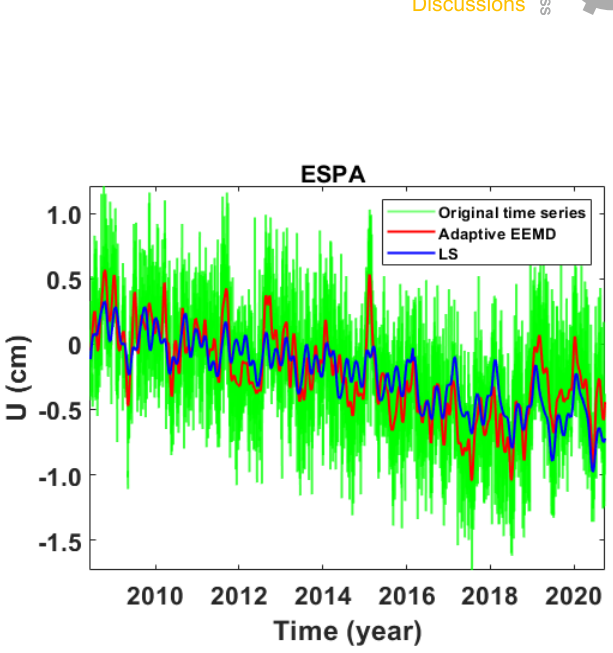
https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023

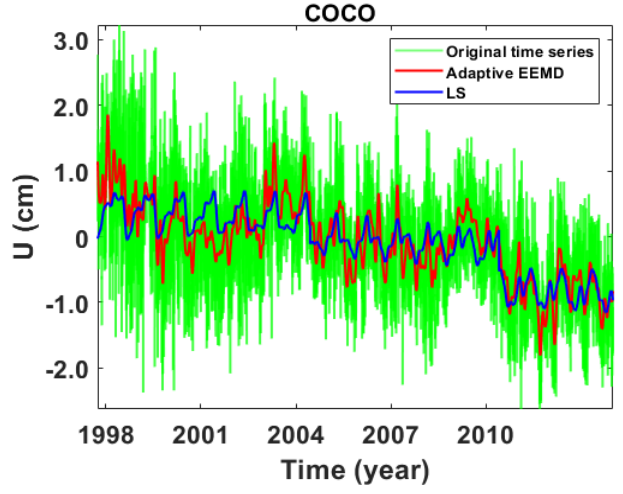
Nonlinear Processes in Geophysics
Discussions
Open Access
EGU

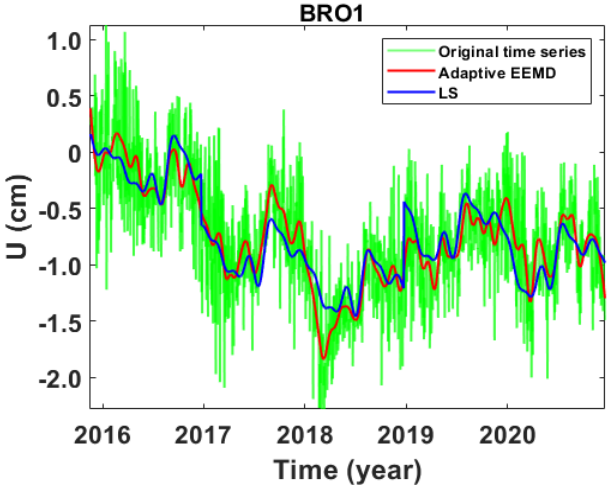


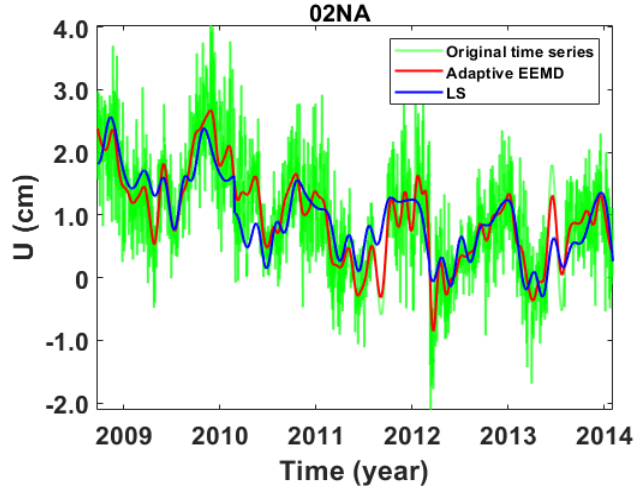
https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023











220

Figure 8: Signals derived by adaptive EEMD and LS of the vertical components at 13 stations.

https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023

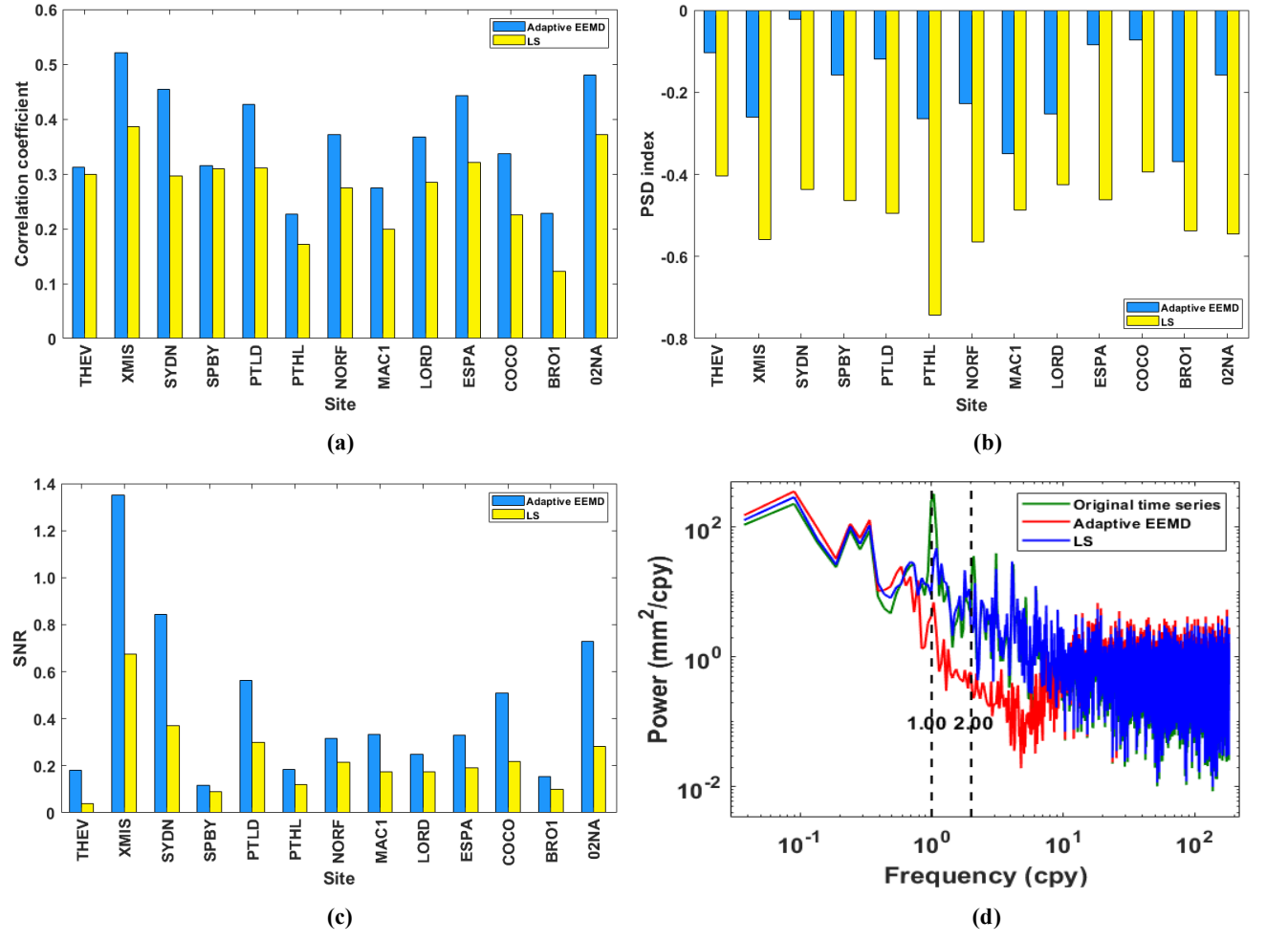


Figure 9: Reliability indicators for periodic model assessing: (a) Correlation coefficient, (b) PSD index, (c)SNR and (d) Stacked comparison of power spectrum analysis.

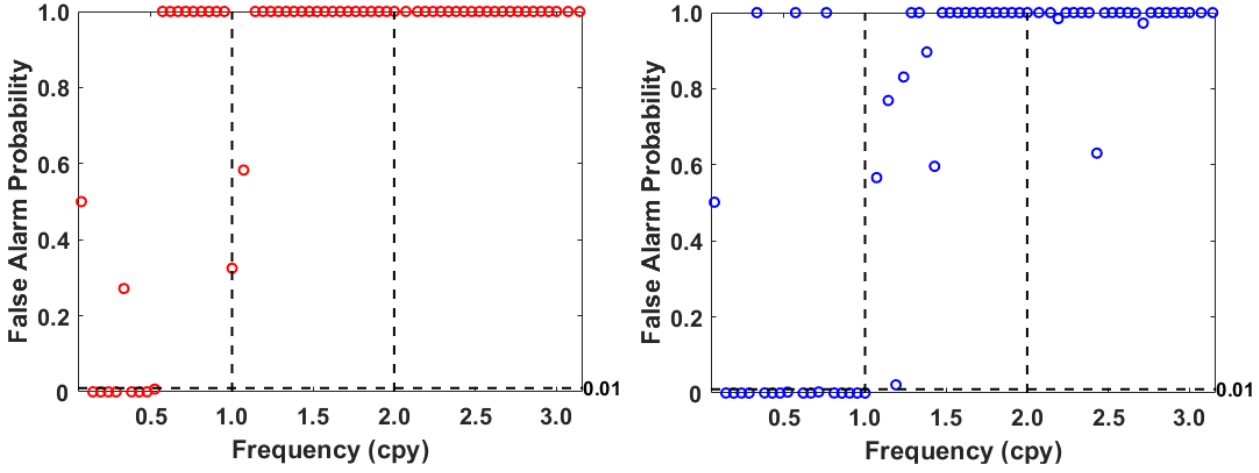
It is observed that the correlation coefficient of the adaptive EEMD method is higher than that of the LS method, while the absolute value of the PSD index of the adaptive EEMD method is lower, which suggests that it can effectively explain the variations of the original time series. And the SNR of the adaptive EEMD methods are all higher than the LS method, which further verifies their excellent performance in capturing signals. In addition, power spectral stacking analysis is performed on residual signals obtained from 13 stations after using two methods for periodic signal extraction, Observations reveal that the periodic signals extracted by the adaptive EEMD method are 'cleaner' and better. However, it is also noted that the residual signal using the adaptive EEMD method still shows an annual peak in Fig. 9(d). To confirm the significance of this signal within the residual signals, power spectral significance test is conducted (Horne et al., 1986; Press et al., 2010), as shown in Fig. 10. The false alarm probability, which is used to estimate the significance of peaks in the power spectrum, indicating the probability of being a result of true signals as opposed to random noise distribution. Generally, lower values (<0.01) indicate

https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023

Nonlinear Processes in Geophysics
Discussions
Open Access
EGU

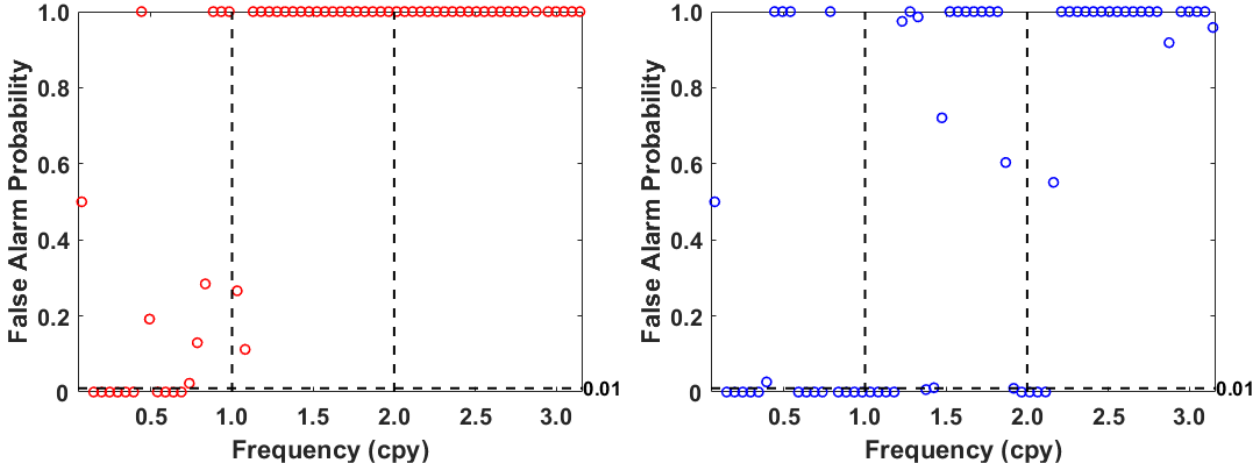
higher probability. The results indicate that 1.00 cpy signal is not significant in the residuals. In contrast, the residual signal using the LS method still shows significant presence of the 1.00 cpy signal. This further highlights the superiority of the adaptive EEMD method in extracting periodic signals.

Moreover, the length of time series in EEMD application is an important factor to consider, which is essential to investigate the impact on periodic signals extracting. Thus, the aforementioned 13 stations are categorized into two groups based on data length: stations with time series exceeding 10 years (mean time series length of 14.13 years) and stations with time series shorter than 10 years (mean time series length of 5.21 years). Also, significance tests are conducted on stacked residual signals for these two groups of stations. As illustrated in Fig. 11-12, the results demonstrate that the longer the time series, the better the performance of both methods in extracting periodic signals. Furthermore, the adaptive EEMD method outperforms in both scenarios, no matter for data longer or shorter than 10 years.



(a) Adaptive EEMD. **(b)** LS.

Figure 10: Significance test of residual signal derived from adaptive (a) EEMD and (b) LS.

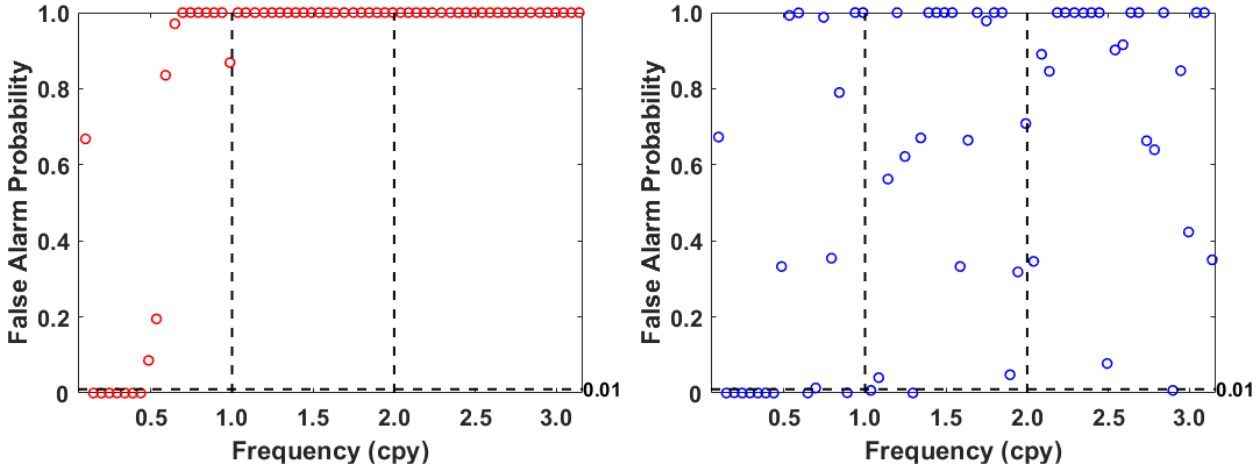


(a) Adaptive EEMD. **(b)** LS.

https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023

Nonlinear Processes in Geophysics
Discussions
Open Access
EGU

Figure 11: Significance test of residual signal derived from (a) adaptive EEMD and (b) LS for stations with data less than 10 years.



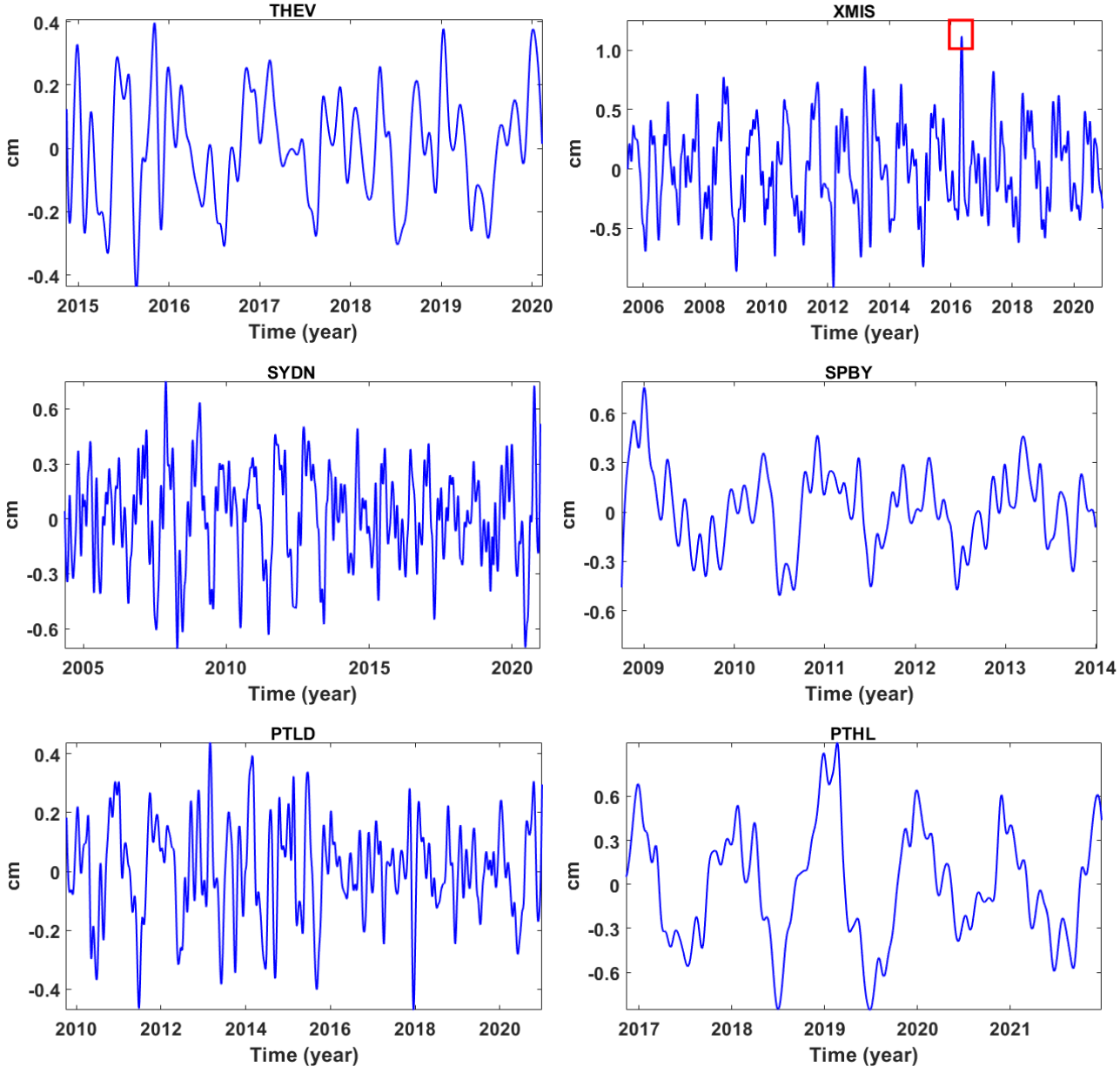
(a) Adaptive EEMD. **(b)** LS.

Figure 12: Significance test of residual signal derived from (a) adaptive EEMD and (b) LS for stations with data more than 10 years.

4.3 Discussion

Figure 13 displays all the periodic signals of the 13 stations in the vertical direction. It is evident that all the 13 stations exhibit significant periodic patterns, but with varying amplitudes. This may be significantly influenced by extreme geophysical phenomenon. It is noteworthy that an anomalous peak is observed in the periodic signal of the XMIS station on 25 May 2016 (approximately 2016.35 years), characterized by an approximately twice larger amplitude of than its mean amplitude (which is approximately 0.5 cm). Similar phenomena are also observed at the ESPA station on 10 February 2015 (approximately 2015.11 years). To find a rational explanation, detailed analysis is conducted and utilized hydrological loading data from the EOST Loading Service (http://loading.u-strasbg.fr/listdata.php?dirn=dicf) (He et al., 2022), which possesses a spatial resolution of $0.5° \times 0.5°$ and a time resolution of 3 hours. Additionally, the daily rainfall time series provided by the Australian Bureau of Meteorology (http://www.bom.gov.au/climate/) is adopted (Pan et al., 2022). Rainfall anomalies are defined as the difference between the daily rainfall and its mean value. The accumulated rainfall anomalies are obtained by summing the daily anomalies with the preceding days' anomalies (Singh et al., 2021). All the possible factors of the two example stations are demonstrated in Fig. 14.

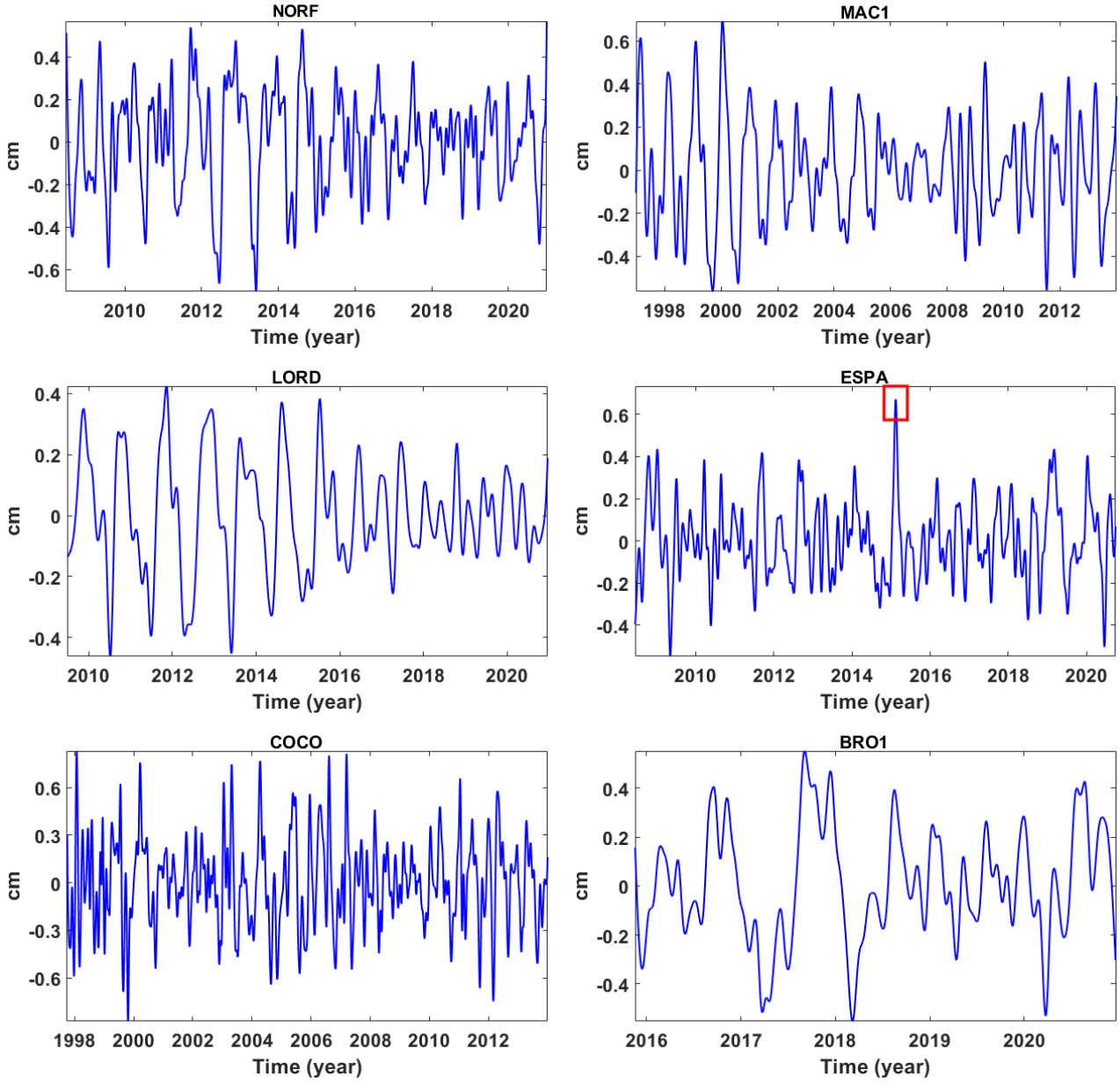
https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023



275

https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023

Nonlinear Processes in Geophysics
Discussions
Open Access
EGU



https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023

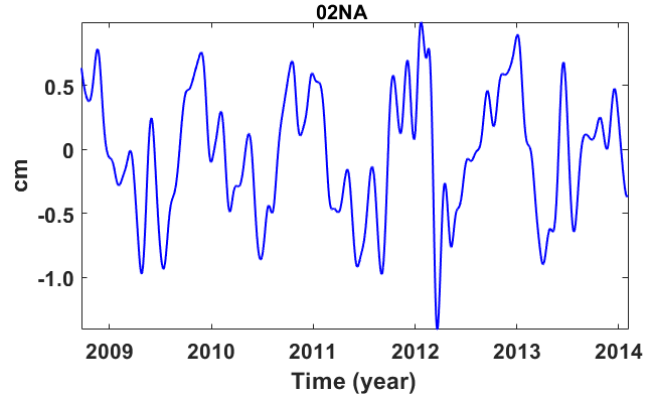
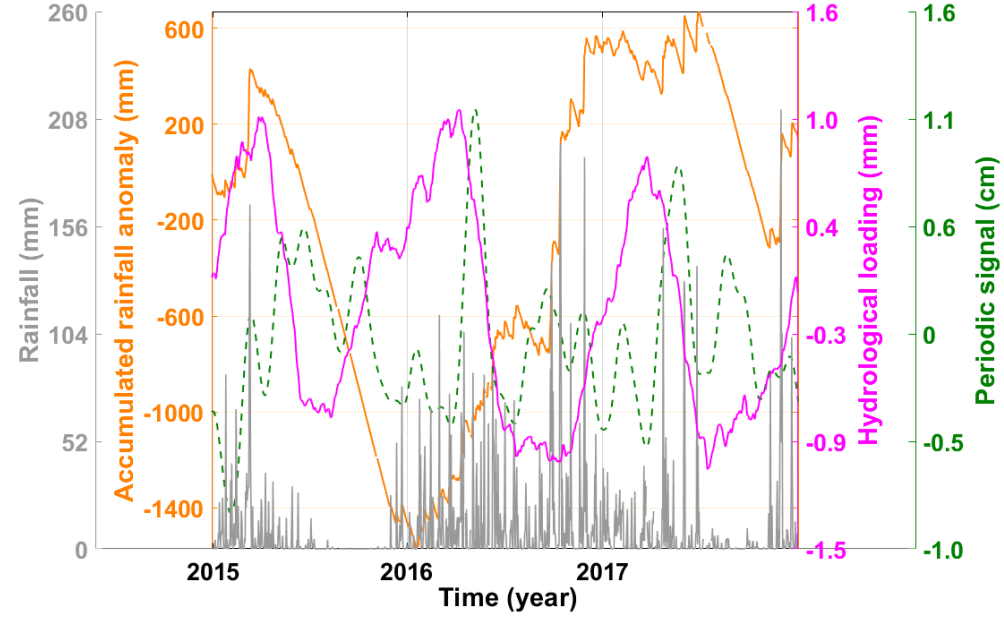


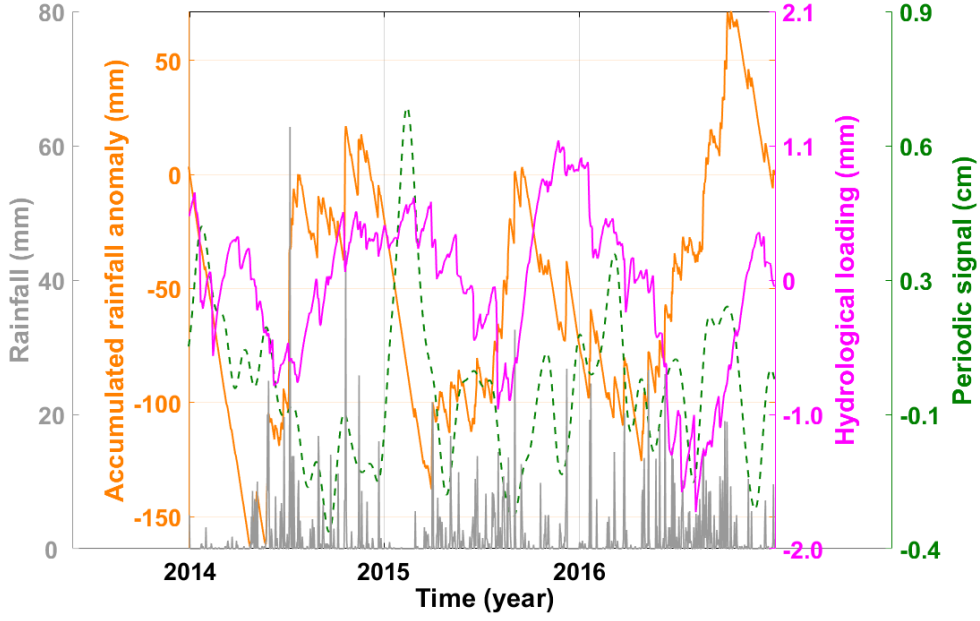
Figure 13: Periodic signals derived by adaptive EEMD of vertical components at 13 stations.

We can see that the extremely low rainfall in the second half of 2015 led to the minimum accumulated rainfall anomaly at XMIS station on 7 May 2016. The increasing hydrological loading responded to the low rainfall but with a time lag reaching its maximum. According to the reports from the Australian Bureau of Meteorology indicated the presence of a strong El Niño phenomenon in 2016 (Kennedy et al., 2017) caused the spread of moist water vapor across the Australian continent. The hydrological loading decreased, which was affected by the rainfall directly. Meanwhile, the hydrological loading shows the visible annual features. The high correlation coefficient of 0.50 between hydrological loading data (detrended) and periodic signals for the year 2016, indicates that hydrological loading is one of the primary factors contributing to periodic changes. In addition, the ESPA station also exhibits a similar phenomenon around February 2015, with extremely low rainfall during the first three months of the year. This leads to a gradual decrease in accumulated rainfall anomalies, reaching its minimum on 29 March 2015. Hydrological loading during the same period reaches its current peak. The correlation coefficient between hydrological loading data (detrended) and the 2015 periodic signal is 0.57. This further emphasizes the significance of hydrological loading variations in relation to periodic changes. In summary, the primary characteristics of periodic variations are induced by hydrological loading, and other factors may also affect the periodic signals to some extent, which provides a direction for future research to explore in depth.

https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023



(a) XMIS station (2015-2017).



(b) ESPA station (2014-2016).

Figure 14: Abnormal periodic signal analysis at (a) XMIS station (2015-2017) and (b) ESPA station (2014-2016).

https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023


Nonlinear Processes in Geophysics Discussions

5 Summary and conclusions

This paper employed the adaptive EEMD method for more efficient and accurate extraction of periodic signals. The simulations are conducted focusing on the inevitable missing data and offsets issues. The results revealed that with the missing data filled with the IEMD method, the adaptive EEMD method exhibits efficiency in signal extraction. As expected, with the percentage of missing data increases, the extracted periodic signals may gradually deviate from the true values. Specifically, for the random data missing, when the missing rate reaches 15%, RMSE and MAE show the rapid increases, at 0.11 and 0.05, respectively.For the consecutive data missing, when the missing lasts for one month, the increases in RMSE and MAE also reach their maximum values, at 0.32 and 0.06, respectively. Therefore, it is recommend that high accuracy could reach for the overall random missing rate below 15% and avoiding consecutive missing epochs exceeding 30 when dealing with GNSS coordinate time series. Regarding the offset, it is apparent that correction beforehand is essential since RMSE and MAE are markedly reduced.That means offset detection and correction is also unignorable for the adaptive EEMD application. For the processing vertical component from 13 GNSS stations in Australia, adaptive EEMD shows its advantage over LS. The mean values for correlation coefficient, spectral index and SNR for the adaptive EEMD are 0.36, -0.18 and 0.48, respectively, while the mean values for LS are 0.27, -0.50 and 0.23. Moreover, significance tests are conducted on the residuals , which further validates that annual signal still exist in the residuals for LS, while not for the EEMD. The periodic signals of the 13 investigated GNSS stations reflect the time-varying characteristics. It is found that the significant variations of the XMIS in 2016 and ESPA in 2015. Their correlation coefficients with hydrological loading are 0.50 and 0.57, respectively, indicating that hydrological loading is one primary factor contributing to the periodic variations. In conclusion, this research highlights the advantages of the adaptive EEMD method for periodic signal extraction no matter in the synthetic or real GNSS coordinate time series analysis. The handling of missing data and offset are firstly comprehensively investigated, which provides helpful solution of preprocessing for EEMD. Furthermore, it reveals the time-varying characteristics of periodic signal, which is favorable in the subsequent trend ananalysis of the GNSS stations in SONEL network. Without doubt, the adaptive EEMD is not limited in the GNSS coordinate time series but also other geophysical data.

Data availability

Any data that support the findings of this study are included within the article.

Author contributions

Weiwei Li: Conceptualization, Methodology, Investigation, Formal Analysis, Writing - Original Draft, Funding acquisition; Jing Guo: Data Curation, Formal Analysis, Writing - Original Draft.

https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023

Competing interests

The authors declare that they have no competing interest.

Acknowledgements

This research was funded by National Natural Science Foundation of China (41904011), Laoshan Laboratory (LSKJ202205104, LSKJ202205104_03) and Shandong Provincial Natural Science Foundation (ZR2019QD003).

References

Agnieszka, W. and Dawid, K.: Modeling seasonal oscillations in GNSS time series with Complementary Ensemble Empirical Mode Decomposition, GPS. Solut., 26, 101, https://doi.org/10.1007/s10291-022-01288-2, 2022.

Bennett, R. A.: Instantaneous deformation from continuous GPS: Contributions from quasi-periodic loads, Geophys. J. Int., 174, 1052–1064, https://doi.org/10.1111/j.1365-246X.2008.03846.x, 2008.

Bos, M. S., Fernandes, R. M. S., Williams, S. D. P. and Bastos, L.: Fast error analysis of continuous GPS observations, J. Geodesy., 82, 157–166, https://doi.org/10.1007/s00190-007-0165-x, 2008.

Bos, M. S., Bastos, L. and Fernandes, R. M. S.: The influence of seasonal signals on the estimation of the tectonic motion in short continuous GPS time-series, J. Geodyn., 49, 205-209, https://doi.org/10.1016/j.jog.2009.10.005, 2010.

Bao, Z., Chang, G., Zhang, L., Chen, G. and Zhang, S.: Filling missing values of multi-station GNSS coordinate time series based on matrix completion, Measurement, 183, 109862, https://doi.org/10.1016/J.MEASUREMENT.2021.109862, 2021.

Blewitt, G. and Lavallée, D.: Effect of annual signals on geodetic velocity, J. Geophys. Res., 107, ETG 9-1-ETG 9-11, https://doi.org/10.1029/2001JB000570, 2002.

Cai, F.: Study on global distribution laws and partial mechanisms of nonlinear variations of GNSS stations' coordinates, M.S. thesis, Strategic Support Force Information Engineering University, https://doi.org/10.27188/d.cnki.gzjxu.2020.000068, 2020.

Chen, B., Bian, J., Ding, K., Wu, H. and Li, H.: Extracting seasonal signals in GNSS coordinate time series via weighted nuclear norm minimization, Remote. Sens-Basel., 12, 2027, https://doi.org/10.3390/rs12122027, 2020.

Dong, D., Fang, P., Bock, Y., Cheng, M. K. and Miyazaki, S.: Anatomy of apparent seasonal variations from GPS-derived site position time series, J. Geophys. Res., 107, ETG 9-1-ETG 9-16, https://doi.org/10.1029/2001JB000573, 2002.

Defraigne, P., Pinat, E. and Bertrand, B.: Impact of Galileo-to-GPS-Time-Offset accuracy on multi-GNSS positioning and timing, GPS. Solut., 25, 45, https://doi.org/10.1007/s10291-021-01090-6, 2021.

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

https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023

Feng, S, and Zheng Z.: Step offset recovery of GNSS continuous station and its effect on GNSS data processing, GNSS World of China, 38, 53-55+82, https://doi.org/10.13442/j.gnss.2013.05.010, 2013.

Griffiths, J. and Ray, J.: Impacts of GNSS position offsets on global frame stability, Geophys. J. Int., 204, 480-487, https://doi.org/10.1093/gji/ggv455, 2016.

Horne, J. H. and Baliunas, S. L.: A prescription for period analysis of unevenly sampled time series, The Astrophysical Journal, 302, 757-763, 1986.

He, Y., Nie, G., Wu, S. and Li, H.: Comparative analysis of the correction effect of different environmental loading products on global GNSS coordinate time series, Adv. Space. Res-Series., 70, 3594-3613, https://doi.org/10.1016/j.asr.2022.08.009, 2022.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N-C. Tung, C. C. and Liu, H. H.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences, 454, 903-995, https://doi.org/10.1098/rspa.1998.0193, 1971.

Huang, N. E., Shen, Z. and Long, S. R.: A new view of nonlinear water waves: the Hilbert spectrum, Annu. Rev. Fluid. Mech., 31, 417-457, https://doi.org/10.1146/annurev.fluid.31.1.417, 1999.

Jiang, W. P., Li, Z., Liu, H. F. and Zhao, Q.: Cause analysis of the non-linear variation of the lGS reference station coordin ate time series inside China, Chinese Journal of Geophysics, 56, 2228-2237, https://doi.org/10.6038/cjg20130710, 2013.

Jia, R. S., Zhao, T. B., Sun, H. M. and Yan, X. H.: Micro-seismic signal denoising method based on empirical mode decomposition and independent component analysis, Chinese Journal of Geophysics, 58, 1013-1023, https://doi.org/10.6038/cjg20150326, 2015.

Kennedy, J., Dunn, R., McCarthy, M., Titchner, H. and Morice, C.: Global and regional climate in 2016, Weather, 72, 219-225, https://doi.org/10.1002/wea.3042, 2017.

Klos, A., Olivares, G., Teferle, F. N., Hunegnaw, A. and Bogusz, J.: On the combined effect of periodic signals and colored noise on velocity uncertainties, GPS. Solut., 22, 1, https://doi.org/10.1007/s10291-017-0674-x, 2018.

Liu, X., Chen, W. and Mao, A.: An adaptive optimization EEMD method and its application in bearing fault detection, ResearchSquare [preprint], https://doi.org/10.21203/rs.3.rs-2615109/v1, 28 February 2023.

Li, M., Guo, S-W., Wei, N., Shi, C. and Gong, H-M.: Study on the draconitic year errors in GPS station coordinates, Progress in Geophysics, 33, 0473-0478, https://doi.org/10.6038/pg2018BB0145, 2018.

Li, W., Shen, Y.: Detection and analysis of velocity and amplitude changes in GNSS coordinate sequences, Journal of Tongji University, 42, 604-10, https://doi.org/10.3969/j.issn.0253-374x.2014.04.017, 2014.

Łyszkowicz, A. and Bernatowicz, A.: Geocentric Baltic Sea level changes along the southern coastline, Adv. Space. Res-Series , 64, 1807-1815, https://doi.org/10.1016/j.asr.2019.07.040, 2019.

https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023

Manevich, A. I., Kaftan, V. I., Losev, I. V. and Shevchuk, R. V.: Improvement of the deformation GNSS monitoring network of the Nizhne-Kansk massif underground research laboratory site, Seismic Instruments., 57, 587–599, https://doi.org/10.3103/S0747923921050042, 2021.

Md Din, A. H., Zulkifli, N. A., Hamden, M. H. and Wan Aris, W. A.: Sea level trend over Malaysian seas from multi-mission satellite altimetry and vertical land motion corrected tidal data, Adv. Space. Res-Series., 63, 3452-3472, https://doi.org/10.1016/j.asr.2019.02.022, 2019.

Peng, W., Dai, W. J., Santerre, R., Cai, C. S. and Kuang, C. L.: GNSS vertical coordinate time series analysis using single-channel independent component analysis method, Physical and Engineering Sciences, 454, 903–995, https://doi.org/10.1007/s00024-016-1309-9, 1971.

Pan, Y., Ding, Hao., Li, J., Shum, C. K., Mallick, R., Jiao, J., Li, M. and Zhang, Y.: Transient hydrology-induced elastic deformation and land subsidence in Australia constrained by contemporary geodetic measurements, Earth. Planet. Sc. Lett., 588, 117556, https://doi.org/10.1016/j.epsl.2022.117556, 2022.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P.: Numerical Recipes: The Art of Scientific Computing, Third Edition, Cambridge University Press, ISBN9780521880688, 2010.

Qiu, X., Wang, F., Zhou, Y. and Zhou, S.: Iteration empirical mode decomposition method for filling the missing data of GNSS position time series, Acta. Geodyn. Geomater., 19, 271–279, https://doi.org/10.13168/AGG.2022.0012, 2022.

Qi, Y., W, L., Yang, Z. and Fu, Z.: Study on methods for improving EMD end effect, Modern Electronics Technique, 36, 50-52, https://doi.org/10.16652/j.issn.1004-373x.2013.22.029, 2013.

Rodionov, N. S.: A sequential algorithm for testing climate regime shifts, Geophys. Res. Lett., 31, 9, https://doi.org/10.1029/2004GL019448, 2004.

Ray, J., Altamimi, Z., Collilieux, X. and Van Dam, T.: Anomalous harmonics in the spectra of GPS position estimates, GPS. Solut., 12, 55-64, https://doi.org/10.1007/s10291-007-0067-7, 2008.

Ran, J., Bian, J., Chen, G., Zhang, Y. and Liu, W.: A truncated nuclear norm regularization model for signal extraction from GNSS coordinate time series, Adv. Spa. Pace. Res., 70, 336-349, https://doi.org/10.1016/j.asr.2022.04.040, 2022.

Shu, Y.: Comparative analysis on outlier elimination methods for GPS coordinate time series, Journal of Navigation and Positioning, 9, 79-85, https://doi.org/10.16547/j.cnki.10-1096.20210412, 2021.

Sorin, N., NorbertSzabolcs, S., Ahmed, E., Michal, A., Zinovy, M., Ilie, N. E., Jacek, K. and Kamil, M.: Implication between geophysical events and the variation of seasonal signal determined in GNSS position time series, Remote. Sens-Basel., 13, 3478-3478, https://doi.org/10.3390/RS13173478, 2021.

Singh, A., Reager, J. T. and Behrangi, A.: Estimation of hydrological drought recovery based on precipitation and Gravity Recovery and Climate Experiment (GRACE) water storage deficit, Hydrol. Earth. Syst. Sc., 25, 511-526, https://doi.org/10.5194/hess-25-511-2021., 2021.

Su, L., Zhai, H., Wang, Q. and Tian, X.: Detecting offsets in GPS coordinate time series based on SSA method, Journal of Geodesy and Geodynamics, 43, 464-466, https://doi.org/10.14075/j.jgg.2023.05.005, 2023.

https://doi.org/10.5194/npg-2023-23
Preprint. Discussion started: 17 October 2023


Nonlinear Processes in Geophysics Discussions

Van Dam, T., Wahr, J., Milly, P. C. D., Shmakin, A. B., Blewitt, G., Lavallée, D. and Larson, K. M.: Crustal displacements due to continental water loading, Geophys. Res. Lett., 28, 651–654, https://doi.org/10.1029/2000GL012120, 2001.

Wang, L. and Herring, T.: Impact of estimating position offsets on the uncertainties of GNSS site velocity estimates, J. Geophys. Res-Sol. Ea., 124, 13452-13467, https://doi.org/10.1029/2019JB017705, 2019.

Wu, Z. and Huang, N. E.: Ensemble empirical mode decomposition: a noise-assisted data analysis method, Advances in Adaptive Data Analysis, 1, 1-41, https://doi.org/10.1142/S1793536909000047, 2009.

Willen, M. O., Horwath, M., Groh, A., Helm, V., Uebbing, B. and Kusche, J.: Feasibility of a global inversion for spatially resolved glacial isostatic adjustment and ice sheet mass changes proven in simulation experiments, J. Geodesy., 96, 75, https://doi.org/10.1007/S00190-022-01651-8, 2022.

Wang, K., Jiang, W., Chen, H., An, X., Zhou, X., Yuan, P. and Chen, Q.: Analysis of seasonal signal in GPS short-baseline time series, Pure. Appl. Geophys., 175, 3485–3509, https://doi.org/10.1007/s00024-018-1871-4, 2018.

Xu, F.: GNSS monitoring data processing based on ensemble empirical mode decomposition method, Geomatics Technology and Equipment, 24, 102-109, https://doi.org/10.20006/j.cnki.61-1363/P.2022.03.020, 2022.

Yi, W. H., Yan, L., Wang, Z. H., Yang, J. H. Tao, T. J. and Liu, L. S.: Geotechnical engineering blasting: a new modal aliasing cancellation methodology of vibration signal de-noising, Earthq. Eng. Eng. Vib., 21, 313-323, https://doi.org/10.1007/s11803-022-2094-3, 2022.

Zheng, J.: Experimental study on the method of IMF component judgment criterion in Empirical Mode Decomposition, Journal of Geomatics, 46 ,33-37, https://doi.org/10.14188/j.2095-6045.2018391, 2021.

Zhang, J., Wang, X. and Hu X.: Analysis of GPS stations' time series based on PCA method, Journal of Geodesy and Geodynamics, 39, 613-619, https://doi.org/10.14075/j.jgg.2019.06.012, 2019.

Zhang, J., Yan, R. Q. and Gao, R. X.: Ensemble Empirical Mode Decomposition for machine health diagnosis, Key. Eng. Mat., 413, 167-174, https://doi.org/10.4028/www.scientific.net/KEM.413-414.167, 2009.