**Reply to the comments on preprint npg-2023-23 entitled 'Extraction of periodic signals in GNSS vertical coordinate time series using adaptive Ensemble Empirical Modal Decomposition method'**

We would like to express our appreciation to the editor and the reviewers for their comments and valuable suggestions, which are helpful for further improving this paper. We have revised the manuscript very carefully and responded point by point to the comments as below. (**C** and **R** indicate comment and response, respectively)

**Reply to Reviewer#1:**
Thank you very much for your suggestions.

**C1:** The abstract and introduction are not clear enough: too limited introduction to the concept of EMD, GNSS data etc. Moreover, some sentences are obscure and in few cases there are grammar mistakes. Can the authors add background for the reader, be more inclusive in the references they include and amend errors? Follow my point-by-point notes in the attached pdf file.
**R1:** Thanks for your comments. We have extensively revised the abstract and introduction sections to provide a clearer overview of the concepts related to EMD and GNSS data. Additionally, we have carefully reviewed and corrected grammar mistakes, and added more relevant content to the reference list. Specifically, we followed the detailed notes in the attached pdf and made modifications point by point, highlighting these modifications in red in the revision.

**C2:** Some transient trends in GNSS time series are not connected with seasonal trends, but they also may include other contributions such as
- tectonic effects such as afterslip. I know authors investigate signals of GNSS stations in Australia and surrounding regions, but I think that at least a short discussion about this topic should be added; moreover, I suggest to check whether some remote signal can be identified or not. In Australia large earthquakes do not occur, but along the surrounding subduction zones even M9+ earthquakes take place, producing long-term rearrangement of stress at even large distances (I know this because of my direct expertise).
Check, for instance, Blewitt, G., Kreemer, C., Hammond, W. C., Plag, H.-P., Stein, S., & Okal, E. (2006). Rapid determination of earthquake magnitude using GPS for tsunami warning systems *Geophysical Research Letters*, 33, L11309. https://doi.org/10.1029/2006GL026145
- both solid and ocean tides can impact on GNSS time series, check, for instance, Zaccagnino, D., Vespe, F., & Doglioni, C. (2020). Tidal modulation of plate motions. *Earth-science reviews*, 205, 103179. and Ide, S., & Tanaka, Y. (2014). Controls on plate motion by oscillating tidal stress: Evidence from deep tremors in western Japan. *Geophysical Research Letters*, 41(11), 3842-3850.
Can the author add a short discussion about these topics?
**R2:** The trends derived from GNSS time series are affected by multiple factors. As you said, some transient trends are not connected with seasonal signal but also the earthquakes and tides. As you suggested, we checked whether remote signal can be identified at the 13 stations we analyzed here. We find that XMIS station located in the northwest indeed affected by remote earthquakes during the investigated period. Specifically, one is about 2000 km far in $M_w$ 8.6 and the other is about 1000 km in $M_w$ 8.5. Thanks for your suggestion. We added one short discussion on the two

contributors in the introduction.

'The trend terms are composed with long-term trend and transient trend. For the transient trend, it mainly includes the effect of earthquake and tide. For the 26 December 2004 Sumatra earthquake ($M_w$ 9.2–9.3) caused the afterslip of larger than 10 mm within minutes as far away as India (Blewitt et al., 2006). It is usually modelled with logarithmic or exponential functions (Hetland and Hager, 2006; Nishimura, 2014; Tobita, 2016). Moreover, modulations in plate motion compatible with the Solar year, the period of the Lunar perigee and Lunar nodes, which clearly the influence of lunar and solar tidal forces on the plate motion (Zaccagnino et al., 2020). Also, this motion by oscillating tide can be evidenced by the deep tremors (Ide and Tanaka., 2014). Since our focus is the long-term trend, we mainly investigate the impact of periodic signals on it.'

**C3:** Subfigures should be highlighted using letters (such as A, B; C).
**R3:** All the subfigures are marked with letters in the revision.

**C4:** To prove that EMD works better than other techniques, a test should be done in addition to the already performed one. For instance, at least RMSE or Chi Squared of fits in Figure 8.
**R4:** Yes, to show the advantage of adaptive EEMD, the evaluation metrics including correlation coefficients (CC), power spectral density index ($\kappa$) and signal-to-noise (SNR) are given in Figure 9 in the manuscript. As you suggested, we added one more root mean square error (RMSE) indictor, which is displayed in Fig. 9(B). We can see that the RMSE values of the adaptive EEMD method are consistently lower than those of the LS method. This further emphasizes the superior performance of the adaptive EEMD method in extracting periodic signals.

**C5:** Check the grammar mistakes pointed in the pdf.
**R5:** Thanks. We checked through the manuscript and corrected the grammar mistakes.

**C6:** p1-line 7: Provide a short introduction to empirical mode decomposition.
**R6:** Thanks for your suggestion. We have added brief introduction of EMD in the abstract. 'Empirical Modal Decomposition (EMD) is an efficient tool of extracting signal from stationary or non-stationary time series, which is enhanced in the stability and robustness by Ensemble Empirical Mode Decomposition (EEMD).'

**C7:** p1-line 10: '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.' This sentence is very confused.
**R7:** We are sorry for the confused sentence. It is revised as 'In order to thoroughly investigate their impacts, we simulated 5 years of daily time series data with different missing data percentage or different number of offsets and conducted 300 times for each simulation. The results show that high accuracy could reach for the overall random missing rate below 15% and avoiding consecutive missing exceeding 30 days.'

**C8:** p1-line 14: why ','?
**R8:** We are sorry for this mistake. We have removed ','.

**C9:** p1-line 16: change 'signal noise ratio' to 'signal-to-noise'
**R9:** Accepted and modified.

**C10:** p1-line 19: I cannot understand the meaning of this sentence: Furthermore, the time-varying periodic characteristics is more conducive to analyze the driving factors.
**R10:** We are sorry for the misleading. This sentence is revised as 'Moreover, driving factors is more effectively facilitated by the time-varying periodic characteristics compared with the constant periodic signal derived by LS.'

**C11:** p1-line 26: check this spelling: Md Din.
**R11:** Checked and corrected.

**C12:** p1-line 27: change 'so on' to 'other applications'
**R12:** Accepted and modified.

**C13:** p1-line 28: change 'periodic variations' to 'fluctuations/periodic signals'
**R13:** Thanks for your comment. We have changed 'periodic variations' to 'periodic signals'.

**C14:** p2-line 49: change 'assessed' to 'included'
**R14:** Accepted and modified.

**C15:** p2-line 49: A new paragraph??
**R15:** Thanks for your suggestion. We have started a new paragraph as recommended.

**C16:** p2-line 57: change' Sect. 5'to 'in the last one'. Avoid repetitions, please.
**R16:** Thank you for your comment. We have made revisions to avoid repetitions.

**C17:** p3-line 62: change 'is to' to 'consists in adding'
**R17:** Accepted and modified.

**C18:** p3-line 80: rephrase this sentence: but the increase will reduce the computational efficiency.
**R18:** The sentence is rewritten as 'however, it leads to a reduction in computational efficiency.' in the revision.

**C19:** p3-line 86: While, ... :while what?
**R19:** Thanks. This sentence is revised as 'While preprocessing is also essential for adaptive EEMD application, which involves the selection of good continuity and sufficient data, outliers removing, offset detection and missing data filling.'

**C20:** p3-line 89: 'of the observed epochs' Which ones?
**R20:** Thanks for your comment. The phrase 'of the observed epochs' refers to the epochs with valid observations. This sentence is revised as 'All the missing data are firstly filled with the mean of the observed data to construct an initial complete time series $x'(t)$.'

**C21:** p4-line 100: in this figure, in the sum from i = k+1 to n, imf should be IMF.
**R21:** We have made the necessary modification in the flowchart, including changing 'imf' to 'IMF' in the view of consistency.

**C22:** p4-line 109: eventual seismic events should be taken into account both in terms of offsets and afterslip. I think that a simple short discussion is enough since authors are considering the surrounding of Australia, where large earthquakes are rare.
**R22:** We agree with you. From the holistic perspective, the effects of seismic events should be considered. We have incorporated the terms of offsets and afterslip into the functional model (Equation 4).

$$x(t) = a + bt + \sum_{i=1}^{8}(c_i \cdot \sin(2\pi \cdot f_i \cdot t) + d_i \cdot \cos(2\pi \cdot f_i \cdot t)) + \sum_j e_i \cdot H(t - t_j) + \sum_z g_z \cdot (1 - e^{-\frac{(t-t_z)}{\tau_z}}) + n(t)\,, \tag{4}$$

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, others are draconitic-year periodic coefficients), $f_i$ is the frequency, $e_i$ is the offset magnitude, $t_j$ is the moment of offset occurrence, $H$ is the Heaviside function, $g_z$ is the afterslip magnitude, $t_z$ is the seismic event occurrence time, $\tau_z$ is the relaxation time and $n(t)$ is the noise term.

**C23:** p5-line 136: Add a little bit of background. What are authors doing?
**R23:** We are sorry for the left-out background. Due to the lack of the true data at the missing epochs, it is difficult to assess the effectiveness of the method. Therefore, we removed the existed data and regarded them as the missing. In this case, we have the true data in the missing epochs. In other word, the missing data with different percentages are generated artificially in the simulations. The way we do in the simulations is added in the revision. 'Since there are no true data at the missing epochs, it is impossible to evaluate the performance. Therefore, the missing data in the simulations are artificial deletion from the original data. In practice, data missing may occur randomly or consecutively (Shen et al., 2014). To be more realistic, randomly missing and consecutively missing cases are simulated through deleting data from original data.'

**C24:** p7-line 157: change 'error increased' to 'larger errors'
**R24:** Accepted and modified.

**C25:** p10-line 199: Cite this data and reference also in the data availability section. Additionally, p21-line 327: Nevertheless, please, report in this section both reference and link to data, please.
**R25:** Thank you for your suggestion. We have incorporated the citations of all the data used in the data availability section.

**C26:** p11-line 212: Add a test to prove this. A simple comparison between RMSE and chi Squared fit may be useful.
**R26:** Thanks for your comment. We have incorporated the comparison of RMSE and presented the relevant results in Fig. 9 (B). This sentence is revised as 'It is apparent that the signal extracted is significantly different, from which adaptive EEMD method shows its advantage in time-varying

signal extraction.'

**C27:** p14-line 232: Once again, a test is needed to prove this.
**R27:** Thanks. As you suggested, we added RMSE test to substantiate this result. Additionally, we have included an explanation in the relevant section of the manuscript. 'To further evaluate its performance, comprehensive assessment indicators are displayed in Fig. 9. It is observed that the higher CC (Fig.9A), the lower RMSE (Fig.9B), the lower absolute of $\kappa$ (Fig.9C), the higher SNR (Fig.9D) of the adaptive EEMD, which shows its outstanding advantage over LS.'

**C28:** p19-line 288: 0.5 is not a very high correlation. Can authors explain if they apply a filter? The correlation between the rainfall signal and vertical GNSS series without any filter?
**R28:** We are sorry for the misleading. The correlation coefficient of 0.5 does not represent the correlation between rainfall data and GNSS vertical time series. Rainfall would not directly affect the periodic signal but cause hydrological loading variation. Therefore, when we are trying to look for the drive factors of the peak annual signal, we calculated the correlation between the periodic signal extracted using adaptive EEMD and hydrological loading (detrended). The adaptive EEMD can be considered a form of filtering of the original data since the periodic signals composed of the specific IMFs. Additionally, hydrological loadings contain power at long periods, which would be mistaken for secular tectonic trends (Jiang et al., 2013), therefore the trend is removed.
**Reference**
Jiang, W., Li, Z., Van Dam, T. and Ding, W.: Comparative analysis of different environmental loading methods and their impacts on the GPS height time series, J. Geodesy., 87, 687-703, https://doi.org/10.1007/s00190-013-0642-3, 2013.

**C29:** p21-line 308: Add with space.
**R29:** Accepted and modified.

**C30:** p21-line 310: 30 what? 30%?
**R30:** We are sorry for this mistake. The text has been revised to specify '30 days'.

**C31:** p21-line 324: change 'is not limited' to 'can also be applied not only to'
**R31:** Accepted and modified.

**C32:** p24-line 395: Ami Hassan Md Din
**R32:** Mistake corrected. Thanks.

**Reply to Reviewer#2:**
Thank you very much for your suggestions.

**C1:** As minor comments, we would like the authors to present their thoughts regarding the processing of not only vertical GPS data, but also horizontal components. As is known, horizontal components almost always contain strong trends reflecting slow movements of tectonic plates. It would be interesting to read about how the authors would highlight harmonic components against the background of strong trends. With the exception of simple preliminary elimination of the general trend with subsequent analysis of the remainder.

**R1:** The absolute sea level is related with the vertical movement of the tide gauges. That is why we focus on the vertical component of GNSS stations which distributed nearby these tide gauges. Yes, the strong trends exist in horizontal components of GNSS stations. If we don't eliminate them beforehand, the decomposition of EEMD seems failure. Especially the last IMF with long term period cannot be distinguished with the reminder. We have tried data processing such as standardization. It cannot still weaken the trend effect in the data. Only by weakening the trend effect, can the accuracy of the harmonic components be guaranteed. Because the crucial step in EEMD is the formation of upper and lower envelope lines which is based on data. However, when the trend is strong, the spacing between the signal's maxima and minima becomes small or unclear, potentially causing EMD to overly focus on the trend component and overlook other inherent harmonic features. Given an extreme example, it cannot be decomposed for the data only linear signal included. Up to now, we think although the elimination of the general trend is simple, it is regarded as the effective way in weakening its effect. We are sorry that we have not any idea of the weakening trend effect except this. Besides, we conducted an extensive literature survey, which studies on the horizontal components. Some directly utilize detrended time series (Klos et al., 2020; Dong et al., 2022), while others incorporate a step to remove trend components before extracting periodic signals (Xu and Yue, 2015; Klos et al., 2018). Thanks for your comment. We will consider more on this topic.

**References**

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

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

Klos, A., Bogusz, J., Bos, M. S. and Gruszczynska, M.: Modelling the GNSS time series: different approaches to extract seasonal signals, Geodetic time series analysis in earth sciences, 211-237, https://doi.org/10.1007/978-3-030-21718-1_7, 2020.

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