GPS TOOLBOX

GTS_CME: an open-source MATLAB-based software for the analysis of common mode errors in GNSS coordinate time series

Zien Xiao[1,2] · Bin Liu[1,2] · Wujiao Dai[3] · Xiaojun Ma[1,2] · Xuemin Xing[1,2] · Yabo Luo[1,2]

Abstract

Regional global navigation satellite system (GNSS) network analyses are often affected by spatially correlated errors, also known as common mode errors (CMEs). To better understand and process the CMEs in GNSS observations and to enhance the efficiency and accuracy of data processing, we developed a GNSS coordinate time series analysis software named GTS_CME based on mature principal component analysis (PCA) and independent component analysis (ICA) spatiotemporal filtering methods using MATLAB language. In addition, GTS_CME integrates numerous functions to assist users in the comprehensive analysis of GNSS coordinate time series. It includes standard preprocessing capabilities such as outlier removal, offset detection and removal, and continuous missing value interpolation. Moreover, the software features tools for retrieving surface mass loading data and calling Hector, which can be used for surface mass loading correction, noise characteristic analysis, and accuracy evaluation of GNSS coordinate time series.

Zien Xiao and Bin Liu have contributed equally to this work.

✉ Bin Liu
    binliu@csust.edu.cn

1   Engineering Laboratory of Spatial Information Technology of Highway Geological Disaster Early Warning in Hunan Province, Changsha University of Science and Technology, Changsha 410114, Hunan, China

2   School of Traffic and Transportation Engineering, Changsha University of Science and Technology, Changsha, China

3   School of Geosciences and Info-Physics, Central South University, Changsha, China

Introduction

Global navigation satellite system (GNSS) observations have been used in the field of geodesy and geophysics to investigate a range of geophysical effects (Bock and Melgar 2016; White et al. 2022). However, errors with spatial correlations, also referred to as "common-mode errors (CMEs)," usually affect the GNSS coordinate time series derived from GNSS observations (Wdowinski et al. 1997). CMEs are thought to be generated by an incompletely estimated mass loading (isostatic water and atmosphere), an imprecise atmospheric model, GNSS satellite orbit inaccuracy, and other causes (Dong et al. 2006). Several alternative regional GNSS network filtering approaches have been deployed over the last 20 years to mitigate the influence of CME and increase the signal-to-noise ratio of GNSS coordinate time series. Examples include stacking (Wdowinski et al. 1997), weighted-stacking (Nikolaidis 2002), PCA (principal component analysis) and Karhunen–Loeve expansion (KLE) (Dong et al. 2006), and ICA (independent component analysis) (Ming et al. 2016).

Nowadays, many specialized and practical software tools are available for processing and analyzing GNSS coordinate time series data, which have been widely applied in various

fields. For instance, iGPS (Tian 2011) and TSanalyzer (Wu et al. 2017) are visualization and preprocessing toolboxes developed based on IDL and Python, respectively. GMIS (Liu et al. 2017b) adopts spatio-temporal Kalman interpolation to handle continuous missing values, whereas JUST (Ghaderpour 2021) are used for discontinuity detection. Hector and CATS, on the other hand, are utilized for noise characteristic analysis and trend estimation (Bos et al. 2012; Williams 2008). GNSS2TWS (Jiang et al. 2022) is a dedicated toolbox for land water storage inversion. Moreover, He et al. (2020) developed GNSS-TS-NRS, employing the EMD decomposition algorithm for GNSS coordinate time series denoising. It can also utilize stacking and PCA for spatiotemporal filtering in GNSS datasets. However, it lacks a specific focus on CMEs analysis and lacks post-processing features such as noise characteristic analysis, which are important for describing the stochastic properties of GNSS coordinate time series (He et al. 2019, 2021). We developed a new software named GTS_CME, which is dedicated to extracting, analyzing, and interpreting CMEs in the GNSS coordinate time series. The software incorporates the widely adopted ICA spatiotemporal filtering method and integrates various pre- and post-processing functions to help users comprehensively analyze GNSS coordinate time series.

Software platform and installation

GTS_CME is an open-source software developed using MATLAB and available for download at https://figshare.com/projects/GTS_CME/165658. The downloadable package includes installation files, code files, sample data, a user's manual, and an example video. Following the download, there are two ways to install GTS_CME: (1) Open Matlab, locate and double-click the "GTS_CME.mlappinstall" file in the directory for program installation. Select GTS_CME from the MATLAB app drop-down menu to execute it. (2) Add the source code folder into the Matlab search path, enter "GTS_CME" in the command line window, and await the appearance of GTS_CME software's primary interface.

Features of GTS_CME

GTS_CME offers functionalities grouped into three main modules: the time series preprocessing module, the spatiotemporal filtering and common mode error analysis module, and the auxiliary function module. Details about the auxiliary module will be covered in the user's manual.
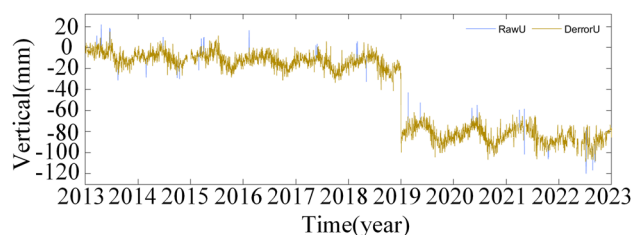
Fig.1 Outliers removal results. The light yellow line depicts the coordinate time series after removing outliers
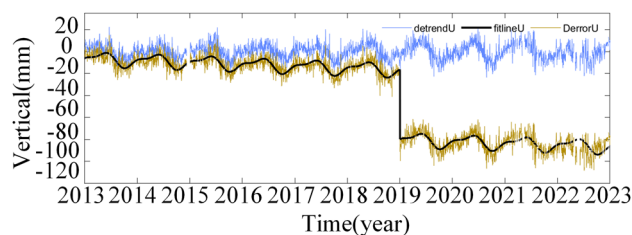
Fig.2 The black line represents the fitting curve, and the blue line represents the residual series with period term remaining

Time series preprocessing

The GNSS time series preprocessing includes commonly used preprocessing features such as outlier removal, offset detection and removal, and missing value interpolation. Here, we use the time series of 46 GNSS stations in the Sichuan Yunnan region of China as the experimental object. Next, we will take the SCYX station as an example to introduce the preprocessing functions of GTS_CME in detail.

Outliers removal

The existence of outliers can significantly impact the findings of spatiotemporal filtering analyses in regional GNSS networks, leading to anomalous temporal and spatial responses in the filtering process and impairing the separation of common mode components at other sites (Dong et al. 2006). GTS_CME introduces five outlier removal methods: 3sigma, interquartile range (IQR), median absolute deviation (MAD), moving window-3sigma (Win-3sigma), and moving window median absolute deviation (Win-MAD). In Fig. 1, taking the vertical time series of SCYX is used as an example, we employed the Win-3sigma detection method with a moving window size of 60 days to detect the outliers.

Offset removal

Offset in GNSS coordinate time series mainly results from antenna changes and geophysical phenomena such as earthquakes. In GTS_CME, we employ the JUST offset detection algorithm in conjunction with an artificial visual inspection (Ghaderpour 2021). GTS_CME allows users to change the additional offset data in the artificial visual inspection based on the site record files given by the specific data center. GTS_CME provides batch-adding offsets, allowing users to make empirical and repeated decisions. Using the vertical GNSS time series of the site SCYX as an example, we remove both linear and offset terms in the deterministic model, as shown in Fig. 2.

Users are allowed to remove the linear, offset, annual period and semi-annual period terms selectively from the GNSS coordinate time series within GTS_CME, followed by an analysis of the residual series. We applied a deterministic model commonly used for fitting GNSS coordinate time series (Nikolaidis 2002):

$$y(t_i) = a + bt_i + c\sin(2\pi t_i) + d\cos(2\pi t_i) + e\sin(4\pi t_i)$$
$$+ f\cos(4\pi t_i) + \sum_{j=1}^{n_g} g_i H(t_i - T_{gi}) + \varepsilon_i \quad (1)$$

where $t_i(i = 1,2,\cdots,N)$ is the daily solution epoch in units of years, a and b are the intercept and linear rate, c and d are annual term coefficients, e and f are semi-annual coefficients, $H(t_i - T_{gi})$ is a Heaviside step function (when $t_i \geq T_{gi}$, its value is 1, otherwise it is 0), g_i is the change of rate from $t_i = T_{gi}$, and ε_i is the model residual.

Missing values interpolation

Missing data in the GNSS coordinate time series is unavoidable due to instrument signal interruption, outlier removal, or other unidentified causes. Simultaneously, it constitutes an essential step in data preprocessing prior to applying PCA and ICA spatiotemporal filtering. GTS_CME's data interpolation module employs three methods
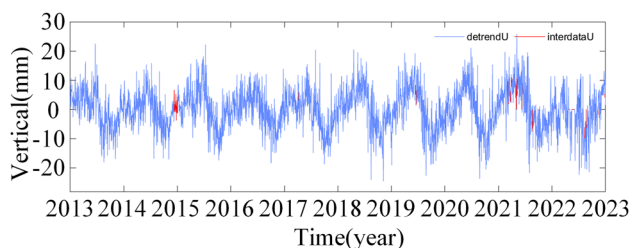
Fig.3 The data results are interpolated using PCA iteration

to interpolate missing data in GNSS coordinate time series: PCA iteration (Liu et al. 2018), DINEOF (Alvera-Azcárate et al. 2005), and RegEM (Schneider 2001). Using the vertical data from the site SCYX as an example, the time series was interpolated using the top 10 principal components with a cumulative contribution rate over 79% through the PCA iterative method, as shown in Fig. 3.

## Spatiotemporal filtering and common mode error analysis

The spatiotemporal filtering and common mode error analysis module combines the M_map tool (Wessel et al. 2019) for automatic plotting, PCA and ICA spatiotemporal filtering, power spectrum analysis, and Comparative analysis of CME and mass loading effects to effectively extract and analyze common mode errors.

Principal component analysis

PCA converts a set of various and possibly related vectors into a small number of unrelated main components, each of which reflects the feature information in the original sequence to varying degrees. PCA is then employed in the regional GNSS coordinate time series analysis (Dong et al. 2006). Taking the vertical data from 46 GNSS stations in the Sichuan Yunnan as a running example, we applied the PCA method to perform spatiotemporal filtering and extract the common mode components from the data, and the results are shown in Fig. 4. Here, we show the results of the first three principal components, whose variance were 54.7%, 4.9% and 3.2%, respectively. It can be seen from Fig. 4 that PC1 exhibit obvious spatial consistency, PC2 presents a local spatial response, and PC3 shows opposite responses in different regions.

Independent component analysis

ICA is designed to extract statistically independent unknown source signals from mixed observed variables even without prior information (Hyvärinen and Oja 2000). ICA has been used in some studies to extract the potential geophysical signals embedded in the GNSS coordinate time series, and their corresponding spatial response can be used to quantitatively characterize the contribution of geophysical sources in various spatial locations. For example, Liu et al. (2018), Yan et al. (2019), and others employ ICA to extract common mode components from GNSS coordinate time series, comparing them with mass loading displacements for analysis. We applied the ICA method to perform spatiotemporal filtering using the same vertical data as the running example. We
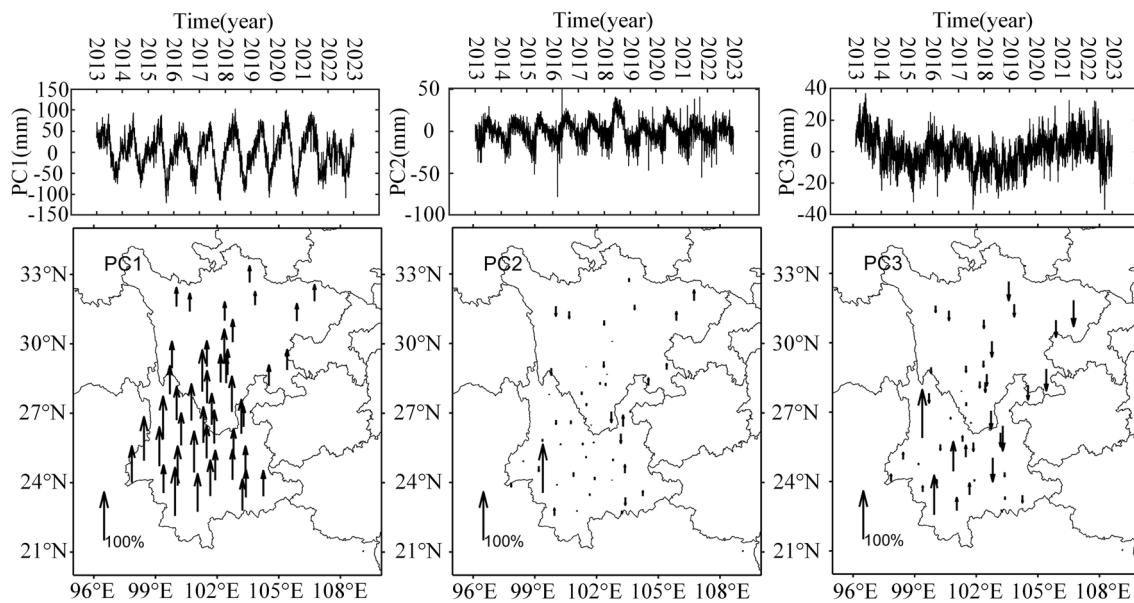
Fig.4 The first three scaled principal components and their normalized spatial responses. The up arrows represent positive spatial responses, and the down arrows represent negative responses. The legend represents a spatial response with 100%
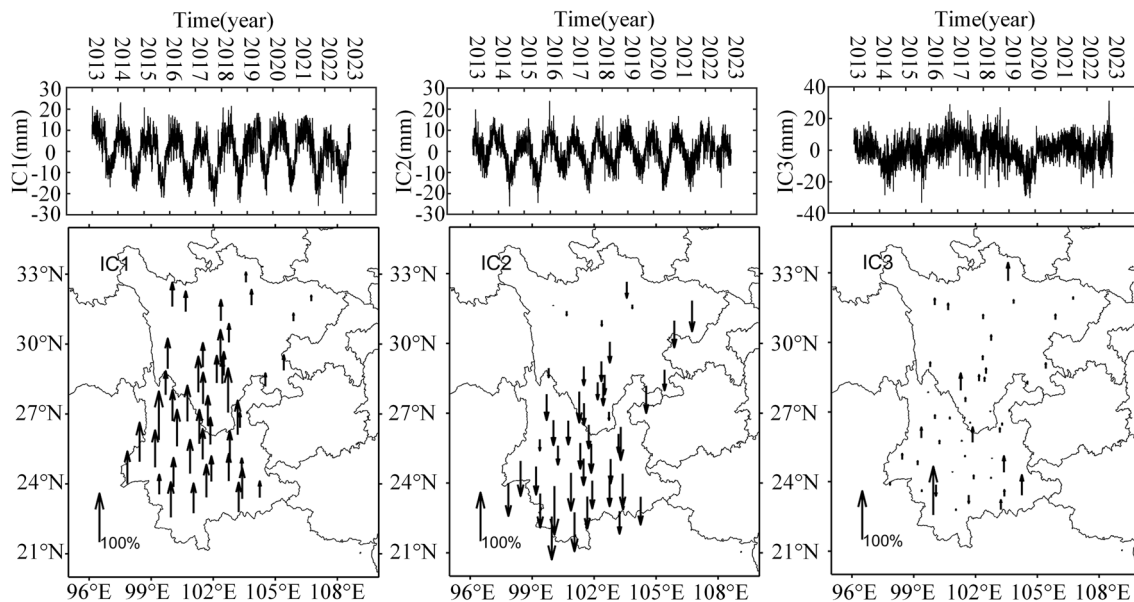
Fig.5 The first three scaled independent components and their normalized spatial responses. The up arrows represent positive spatial responses, and the down arrows represent negative responses. The legend represents a spatial response with 100%

used the top six principal components to perform ICA, and their cumulative contribution percentage exceeded 71%. The independent components are sorted based on the Ratio values (Liu et al. 2017a). The first three ICs and their responses are shown in Fig. 5. It is evident that IC1 and IC2 have obvious periodic properties, and their normalized spatial responses exhibit obvious spatial consistency. Compared to PCA, ICA has separated two mutually maximally independent components.
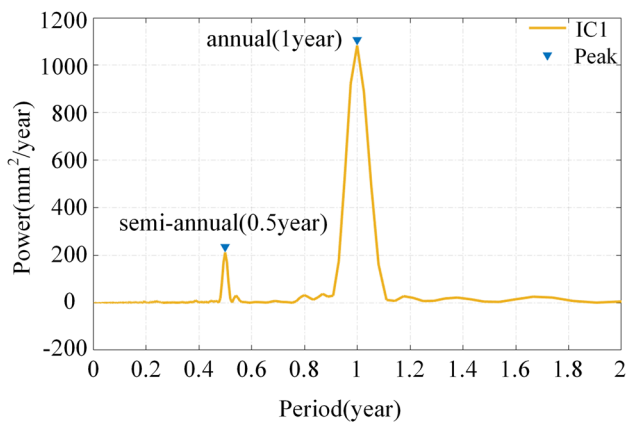
Fig.6 The power spectrum analysis result of IC1
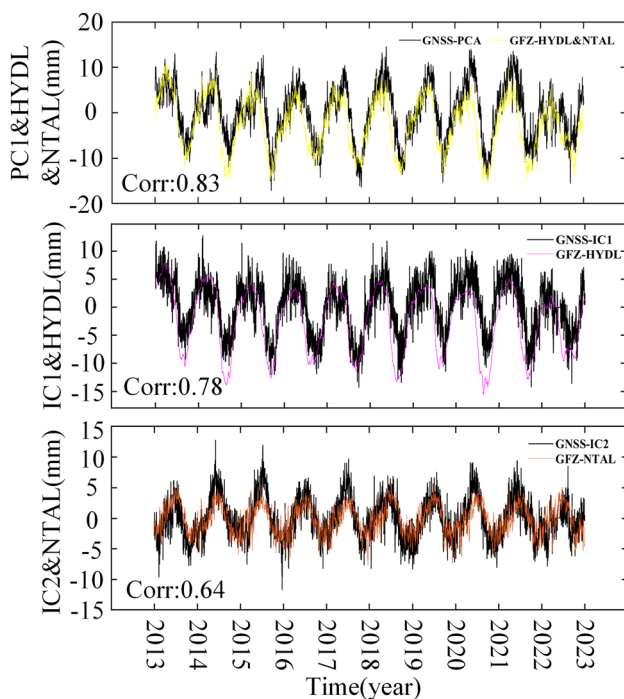
Fig. 7 The black line represents the PCs/ICs derived average displacement across the 46 GNSS stations, and the yellow/pink/red line denotes the average displacement resulting from loading data

Power spectral analysis

Power spectral analysis is a valuable technique in GNSS coordinate time series analysis for examining periodic variations and identifying associated cycles. These variations may arise from a variety of factors. In GTS_CME, the "fastlomb" function is employed to compute the Lomb-Scargle power spectrum (Press 1992), an efficient method

suitable for uniformly and non-uniformly sampled time series. As an illustrative example, the extracted vertical common mode component IC1 from 46 GNSS stations in the Sichuan Yunnan region is analyzed, and the result is presented in Fig. 6.

Comparative analysis of CME and mass loading effects

In the spatiotemporal analysis of GNSS coordinate residual series with periodic terms remaining, the CME is usually contains an obvious periodic term. For example, Yan et al. (2019) separated GNSS coordinate time series using the ICA to investigate the source of seasonal uplift in mainland China and demonstrated that surface mass loadings could address independent components extracted from GNSS data.

As an illustrative example, we employed PCA and ICA to extract vertical common mode components PC1, IC1, and IC2 from 46 GNSS stations in the Sichuan-Yunnan region in China and compare them with non-tidal atmospheric and hydrological loading displacements obtained from ESMGFZ (Fig. 7). The findings revealed a significant correlation between common mode displacements and loading displacements. As shown in Fig. 7, the PC1-derived displacement is highly correlated with the cumulative impact of two primary mass loading effects in the Sichuan-Yunnan region. In comparison, the two common mode independent displacements exhibit high consistency with hydrological and atmospheric loading displacements, respectively. In contrast to PCA, ICA is characterized by its capability to further segregate independent signals based on PCA, which usually represent the potential geophysical signals hidden in GNSS datasets (Liu et al. 2018).

Auxiliary functions module

The auxiliary module of GTS_CME encompasses a format conversion tool, a surface mass loading effect data acquisition tool, and a Hector calling tool. The format conversion tool can convert time series data from different institutions to the *.gta format. The surface mass loading effect data acquisition tool can download mass loading displacement data from EMSGFZ (Earth-System-Modelling group GFZ, http://rz-vm115.gfz-potsdam.de:8080/) for subsequent analysis. The Hector calling tool offers a convenient means to invoke Hector's application program, enabling further analysis of the GNSS coordinate time series. Additional details about the auxiliary functions of GTS_CME can be found in the manual.

Conclusions

The GTS_CME software, developed within an open-source framework, is designed to help users in spatiotemporal filtering and common mode error analysis of GNSS coordinate time series.

We provide an intuitive CME analysis software that integrates the ICA method for the first time, offering valuable functions for ICA application in CME analysis, such as independent component contribution analysis and component sorting.

Compared to existing GNSS analysis software, our software provides some auxiliary functions to enhance its comprehensive analysis capability, including format conversion, the easy download of mass load displacements, and the invocation of Hector software. These features enhance the software's capacity to support users in achieving a thorough analysis of GNSS data in their relevant studies.

Declarations

References

Alvera-Azcárate A, Barth A, Rixen M, Beckers JM (2005) Reconstruction of incomplete oceanographic data sets using empirical orthogonal functions: application to the Adriatic sea surface temperature. Ocean Model 9(4):325–346. https://doi.org/10.1016/j.ocemod.2004.08.001

Bock Y, Melgar D (2016) Physical applications of GPS geodesy: a review. Rep Progr Phys 79(10):106801. https://doi.org/10.1088/0034-4885/79/10/106801

Bos MS, Fernandes RMS, Williams SDP, Bastos L (2012) Fast error analysis of continuous GNSS observations with missing data. J Geodesy 87(4):351–360. https://doi.org/10.1007/s00190-012-0605-0

Dong DN, Fang P, Bock Y, Webb F, Prawirodirdjo L, Kedar S, Jamason P (2006) Spatiotemporal filtering using principal component analysis and Karhunen–Loeve expansion approaches for regional GPS network analysis. J Geophys Res Solid Earth. https://doi.org/10.1029/2005jb003806

Ghaderpour E (2021) JUST: MATLAB and python software for change detection and time series analysis. GPS Solut 25(3):85. https://doi.org/10.1007/s10291-021-01118-x

He XX, Bos MS, Montillet JP, Fernandes RMS (2019) Investigation of the noise properties at low frequencies in long GNSS time series. J Geodesy 93(9):1271–1282. https://doi.org/10.1007/s00190-019-01244-y

He XX, Yu KG, Montillet J-P, Xiong CL, Lu TD, Zhou SJ, Ma XP, Cui HC, Ming F (2020) GNSS-TS-NRS: an open-source MATLAB-Based GNSS time series noise reduction software. Remote Sens 12(21):3532. https://doi.org/10.3390/rs12213532

He XX, Bos MS, Montillet JP, Fernandes RMS, Melbourne T, Jiang WP, Li WD (2021) Spatial variations of stochastic noise properties in GPS time series. Remote Sens 13(22):4534. https://doi.org/10.3390/rs13224534

Hyvärinen A, Oja E (2000) Independent component analysis algorithms and applications. Neural Netw 13(4):411–430. https://doi.org/10.1016/S0893-6080(00)00026-5

Jiang ZS, Hsu Y-J, Yuan LG, Feng W, Yang XH, Tang M (2022) GNSS2TWS: an open-source MATLAB-based tool for inferring daily terrestrial water storage changes using GNSS vertical data. GPS Solut 26(4):114. https://doi.org/10.1007/s10291-022-01301-8

Liu B, Dai W, Liu N (2017a) Extracting seasonal deformations of the Nepal Himalaya region from vertical GPS position time series using independent component analysis. Adv Space Res 60(12):2910–2917. https://doi.org/10.1016/j.asr.2017.02.028

Liu N, Dai WJ, Santerre R, Kuang CL (2017b) A MATLAB-based Kriged Kalman filter software for interpolating missing data in GNSS coordinate time series. GPS Solut. https://doi.org/10.1007/s10291-017-0689-3

Liu B, King M, Dai WJ (2018) Common mode error in Antarctic GPS coordinate time-series on its effect on bedrock-uplift estimates. Geophys J Int 214(3):1652–1664. https://doi.org/10.1093/gji/ggy217

Ming F, Yang YX, Zeng AM, Zhao B (2016) Spatiotemporal filtering for regional GPS network in China using independent component analysis. J Geod 91(4):419–440. https://doi.org/10.1007/s00190-016-0973-y

Nikolaidis R (2002) Observation of geodetic and seismic deformation with the global positioning system. University of California, San Diego

Press W (1992) Numerical recipes in FORTRAN: the art of scientific computing. Cambridge University Press, Cambridge

Schneider T (2001) Analysis of incomplete climate data: estimation of mean values and covariance matrices and imputation of missing values. J Clim 14(5):853–871

Tian YF (2011) iGPS: IDL tool package for GPS position time series analysis. GPS Solut 15(3):299–303. https://doi.org/10.1007/s10291-011-0219-7

Wdowinski S, Bock Y, Zhang J, Fang P, Genrich J (1997) Southern California permanent GPS geodetic array: Spatial filtering of daily positions for estimating coseismic and postseismic displacements induced by the 1992 Landers earthquake. J Geophys Res: Solid Earth 102(B8):18057–18070. https://doi.org/10.1029/97jb01378

Wessel P, Luis JF, Uieda L, Scharroo R, Wobbe F, Smith WHF, Tian D (2019) The generic mapping tools version 6. Geochem Geophys Geosyst 20(11):5556–5564. https://doi.org/10.1029/2019gc008515

White AM, Gardner WP, Borsa AA, Argus DF, Martens HR (2022) A review of GNSS/GPS in hydrogeodesy: hydrologic loading applications and their implications for water resource research. Water Resour Res. https://doi.org/10.1029/2022wr032078

Williams S (2008) CATS: GPS coordinate time series analysis software. GPS Solut 12:147–153

Wu DC, Yan HM, Shen YC (2017) TSAnalyzer, a GNSS time series analysis software. GPS Solut 21(3):1389–1394. https://doi.org/10.1007/s10291-017-0637-2

Yan J, Dong DN, Bürgmann R, Materna K, Tan W, Peng Y, Chen J (2019) Separation of sources of seasonal uplift in China using independent component analysis of GNSS time series. J Geophys Res Solid Earth 124(11):11951–11971. https://doi.org/10.1029/2019jb018139

Wujiao Dai is a professor in the School of Geosciences and Info-Physics at Central South University, China. He received his Ph.D. from CSU in 2007 and worked as a research assistant at Hong Kong Polytechnic University for 3 years. His research mainly focuses on GNSS data processing and deformation monitoring.

Xiaojun Ma is a master student in the School of Traffic and Transportation Engineering at the Changsha University of Science and Technology, China. His research focuses on the GNSS coordinate time series analysis.

Zien Xiao is a master student in the School of Traffic and Transportation Engineering at the Changsha University of Science and Technology, China. His research focuses on the GNSS coordinate time series analysis.

Xuemin Xing is a professor in the School of Traffic and Transportation Engineering at the Changsha University of Science and Technology, China. She received her Ph.D. degree from Central South University in 2011. Her research focuses on satellite geodesy.
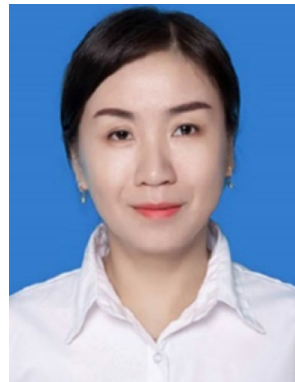

Bin Liu is an associate professor at the School of Traffic and Transportation Engineering at the Changsha University of Science and Technology, China. He received his Ph.D. degree at Central South University in 2018. His research focuses on the application of GNSS to the fields of geodesy and geophysics.

Yabo Luo is an associate professor in the School of Traffic and Transportation Engineering at the Changsha University of Science and Technology, China. He received his Ph.D. degree from the University of Information Engineering. His research focuses on GNSS data processing and deformation monitoring.