Assessment of spatiotemporal filtering methods towards optimising crustal movement observation network of China (CMONOC) GNSS data processing at different spatial scales

Hao Wang, Wenhao Li, Chanfang Shu, C. K. Shum, Fei Li, Shengkai Zhang and Zikang Zhang

School of Geomatics Science and Technology, Nanjing Tech University, Nanjing, China; Division of Geodetic Science, School of Earth Sciences, Ohio State University, Columbus, OH, USA; State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China; Chinese Antarctic Center of Surveying and Mapping, Wuhan University, Wuhan, China

ABSTRACT

Spatiotemporal filtering can effectively remove the common mode error (CME) which significantly affects the accuracy of the Global Navigation Satellite System (GNSS) coordinate time series. This contribution explores the performance of different spatiotemporal filtering methods applied to GNSS networks at different spatial scales. We selected small-scale (<500 km) and large-scale (>2000 km) GNSS networks from the Crustal Movement Observation Network of China (CMONOC) for the focus of the study. To remove or mitigate CME from the different-scale GNSS networks, principal component analysis (PCA), independent component analysis (ICA) and correlation-weighted spatial filtering (CWSF) are compared. In addition, we investigate the correlations between each of the GNSS station residual time series to examine the effectiveness of the novel CME filter. When compared with PCA and ICA results, we find that CWSF is less intrusive in reducing the CME in the different-scale GNSS networks, and thus the preferred the filtering methodology. We conclude that this study could provide an important reference to remove CME from GNSS coordinate time series.

1. Introduction

For regional GNSS networks, spatial correlation between stations is the largest source of error, which is often referred to as the common mode error (CME; Wdowinski et al., 1997). There is not yet an exact or complete characterisation of the physical origins of CME, but it is widely assumed that CME in GNSS coordinate time series is caused primarily by unmodelled systematic errors, including GNSS satellite orbital errors, the influence of the stations’ surrounding environment, cryospheric-atmospheric-hydrologic regimes, instrument errors (masts temperature variations, antenna, multi-path), clock errors, geophysical forward model correlation errors, errors in the International Terrestrial Reference Frame (ITRF), and other, which manifested as the spatiotemporal correlation between GNSS stations (Dong et al., 2002, 2006). The accuracy and reliability of GNSS solutions are greatly affected by CME which must be removed to obtain high accuracy coordinate time series. Several filtering methods have been applied to remove or reduce CME over the last few decades. Wdowinski et al. (1997) first introduced stacking to remove CME in Southern California. A single-day solution variance was used as a weight for spatial filtering of the stacking method (Nikolaidis, 2002). However, the stacking method assumes that CME is uniformly distributed, thus unsuitable for larger spatial scales (spanning more than 1000 km); to overcome the limitation of spatial scale, Márquez-Azuá and DeMets (2003) used baseline lengths as stacking weight to remove/reduce CME on large spatial-scale GPS networks (improved weighted stacking).

Dong et al. (2006) used principal component analysis (PCA) to filter the GPS coordinate time series in Southern California. In recent years, modified PCA methods have been widely used in spatial filtering (He et al., 2015; Li et al., 2015; MA et al., 2016; Shen et al., 2014). Independent component analysis (ICA) is also used to analyse the spatial and temporal characteristics of GPS coordinate time series. Ming et al. (2017) used ICA to extract the CME of continuous observations obtained from the present 259 stations comprising the Crustal Movement Observation Network of China (CMONOC). Liu et al. (2018) applied ICA to GPS vertical coordinate time series in Antarctica from 2010 to 2014 and conducted a comparison with the PCA. Compared with PCA, the CME extracted using the ICA method demonstrated very good consistency with the soil moisture and atmospheric mass loading, implying that ICA is superior in spatiotemporal analysis (Liu et al., 2015), and the results are more reasonable (Ming et al., 2017).
However, PCA depends on second-order statistics (covariance and variance), which do not make full use of the high-order statistical information. The CME contains coloured noise that does not satisfy the normal distribution. Therefore, PCA may lead to the confounding of principal components (PCs) and other physical modes (Ming et al., 2017). Traditional PCA methods have certain limitations when dealing with changing CME and large-scale spatial filtering because PCA can reveal spatial variation in the most crucial PC of each station, but it is usually systematic in its overall spatial variation. If there are a few stations with other relatively small regional effects or anomalous variations at the same time, then PCA would usually ignore them as part of the CME algorithm (Kreemer & Blewitt, 2021). PCA/ICA requires that the time series has no data gaps. However, GNSS coordinate time series commonly have missing data, and thus the quality of the interpolation method directly affects the results of the estimates of the principal/independent components. Another important aspect is that the spatial responses (SRs) of PCA/ICA to some or all principal/independent components are significantly different over larger study areas (Kreemer & Blewitt, 2021; Liu et al., 2015, 2018; Serpelloni et al., 2013; Yuan et al., 2018), and the regional filtering results of small spatial scales are better than the results of large spatial scales (Wu et al., 2019). Dividing subregions will increase the workload exponentially, but no unified division standard is currently available. Previous studies also found that PCA/ICA filtering results on GNSS network CME estimates differ significantly between the Antarctica Peninsula and the whole Antarctic. Despite the use of nearly identical GNSS station data, the RMS (Root Mean Square) reductions for the E, N and U components of the Antarctic Peninsula were 44.69%, 26.94%, and 34.87%, respectively, which are significantly better than the 14.45%, 8.97%, and 13.27% for the whole Antarctica (Li et al., 2019).

Tian and Shen (2016) proposed a spatial filtering method called correlation-weighted spatial filtering (CWSF) to remove CME for a small- to medium-scale GPS network in western North America. The technique utilises a weighting scheme that incorporates two factors: distances between neighbouring sites and their correlations in long-term residual position time series. The CWSF does not require much intervention on the data and is applicable to GNSS networks at any scale. This method can also be used to detect CME irrespective of its origins. By varying the filtering distance range, the long-range CME associated with atmospheric disturbances can be filtered out, and the CME associated with transient tectonic deformations can be found.

In this paper, given that CME has a spatially correlated error, which may change as the spatial scale of the GPS network increases, as will the applicability of spatial filtering methods, so we divided the GPS network into different scales. We use the CWSF method to filter the coordinate time series in GPS networks of different scales, and we compare the results of PCA and ICA at the same time. In our study, the CMONOC’s 10 GPS stations were chosen to form a small-scale GPS network, while the 221 GPS stations covering the entire country were chosen to form a large-scale GPS network.

2. Methods

2.1. Principal component analysis (PCA)

PCA is a statistical method for dimensionality reduction. Suppose $n$ GPS stations are present in an area with a time span of $m$. The matrix $X(m \times n)$ is a residual time series for each column of one site, $x(t, s_j)$ represents the residuals of the jth station on the day i. Its covariance array $B$ is defined as

$$b_{ij} = \frac{1}{m-1} \sum_{l=1}^{m} x(t_l, s_j)x(t_l', s_j)$$

(1)

$B$ is an $n$-dimensional symmetric matrix, for which an orthogonal decomposition is performed

$$B = V \Lambda V^T$$

(2)

where $\Lambda$ is a diagonal matrix with $r$ non-zero eigenvalues, and $V^T$ is an $n \times n$ eigenvector matrix with row vectors orthogonal to the matrix $X$. Expanding the matrix $X$ with the orthogonal basis $V$ yields

$$X = AV^T$$

(3)

Based on the properties of orthogonal matrices, Equation (3) can be transformed to

$$A = XV$$

(4)

Namely,

$$a_k(t_i) = \sum_{j=1}^{n} x(t_i, s_j) v_k(s_j)$$

(5)

Equation (5) $a_k(t_i)$ is called the jth PC of the matrix $X$, which represents the time change; and $v_k(s)$ represents the SR of the PC $a_k(t_i)$. The eigenvectors are arranged in descending order according to the size of the eigenvalues, and the PC with a larger variance in the front row represents the CME signal in the $v_k(s)$ residual time series $X$. Assuming that the first $p$ PCs represent the CME, the CME of the PCA method can be defined as

$$\epsilon(t_i, s_j) = \sum_{k=1}^{p} a_k(t_i) v_k(s_j)$$

(6)

PCA decomposes the GPS residual time series into principal components in the time domain and eigenvectors (spatial response) in the spatial domain. Since the PCA decomposition is ordered by the contribution rate of each principal component, the CME of station time series in the regional GPS network is frequently best represented by the variation characteristics of the
first few principal components, and the eigenvectors corresponding to these principal components reflect the spatial distribution of these CME.

### 2.2. Independent component analysis (ICA)

ICA uses high-order statistical properties of source signals to separate independent signals from mixed signals. This is a PCA extension (Comon, 1994). Its model is

\[
X = AS
\]

(7)

where \(X = [x_1 \ldots x_n]\) is the observation data, \(A\) is the mixing matrix of \(n \times n\), and \(S = [s_1 \ldots s_n]\) is the original signal. To recover the original signal, the inverse of \(A\), also known as the unmixing matrix \(W\), needs to be found

\[
W = A^{-1}
\]

(8)

In this paper, the FastICA algorithm (Hyvärinen, 1999; Hyvärinen & Oja, 2000) is used to complete the ICA analysis, and the CME can be defined as

\[
\text{CME}_{\text{ica}} = \sum_{j \in R} A_j S_j
\]

(9)

Equation (9) \(S\) is a subset of independent components (ICs) with common spatial features in \(R\) and is the corresponding SR.

ICA is a blind signal separation method developed in the last two decades that can separate the original mutually statistically independent signals from the observed mixed signals. Separate geophysical processes will generate statistically distinct signals that will be superimposed on the GPS time series. As a result, the CME extraction of the GPS network is regarded as a blind signal separation problem, and the GPS residual time series are decomposed into maximally statistically independent components using the ICA method, and the ICs with almost uniform SRs are regarded as the CME source signals.

### 2.3. Correlation-weighted spatial filtering (CWSF)

CWSF introduces the correlation coefficient between stations as a weighting factor, thus having a good SR and the ability to overcome the distance limitation of superposition filtering. Its CME can be expressed as

\[
e_{i,k} = \sum_{j=1}^{N_s} \frac{V_{j,k}}{\sigma_{j,k}^2} \times W_{i,j}
\]

(10)

In Equation (10), \(e_{i,k}\) is the CME value of the ith station on the day \(k\); \(N_s\) is the number of stations involved in the CME calculation; \(V_{j,k}\) and \(\sigma_{j,k}\) are the residual and standard deviation of a component of the jth station on the day \(k\), respectively; and \(W_{i,j}\) is the weight, which is calculated as

\[
W_{i,j} = \frac{r_{ij}}{\sum_{j=1}^{N_s} W_{i,j}} < \frac{d_{ij}}{r_{ij}} \quad (11)
\]

where \(d_{ij}\) is the radial distance between station \(i\) and station \(j\); \(r_{ij}\) is a distance weighting constant that adjusts the weights by the distance between stations; \(W_{i,j}\) is a series in descending order according to \(W_{i,j}\), where \(W_i\) is a parameter to be determined; and \(r_{ij}\) is the correlation coefficient between stations \(i\) and \(j\)

\[
r_{xy} = \frac{\sum_{k=1}^{N} (x_k - \bar{x})(y_k - \bar{y})}{\sqrt{\sum_{k=1}^{N} (x_k - \bar{x})^2} \sqrt{\sum_{k=1}^{N} (y_k - \bar{y})^2}}
\]

(12)

where \(N\) is the number of common ephemerides, \(x\) and \(y\) are the residual coordinate time series of the two GPS stations, \(\bar{x}\) and \(\bar{y}\) are the mean values of their residual coordinate time series (Tian & Shen, 2016).

The CWSF method is improved based on the stacking method (Wdowinski et al., 1997). The CME extracted by stacking is a weighted average of the residual time series of each station. On this basis, CWSF introduces the correlation coefficient and distance between stations as the weight of CME calculation. Unlike existing filtering methods, the CWSF does not require the assumption of uniform spatial distribution.

### 3. Data

#### 3.1. Data sources

We use data from the CMONOC (http://www.cgps.ac.cn), a national geoscience observation network comprised of 260 reference stations for continuous monitoring and 2000 regional stations for irregular monitoring. CMONOC has a spatial scale of about 5000 km, which can accurately reflect the CME distribution pattern. The station selection is based on the following principles: (1) the missing rate of station data is low, and (2) the coordinate time series has a large span and high stability. Then, 10 stations distributed across four provinces in China were formed into a small spatial-scale GPS network, while the 221 stations across the entire China region were formed into a large spatial-scale GPS network, with a time span spanning from June 2012 to March 2020. Figure 1 shows the distribution of GPS stations.

#### 3.2. Acquisition of residual time series

The extraction of the CME is based on the residual time series, which is obtained for each station by removing outliers, offsets, trends, and periods from the raw
3.3 Outlier removal and data interpolation

Due to multipath effects, station-related errors (e.g. electromagnetic interference), and orbit anomalies, outliers in GPS coordinate time series are unavoidable and need to be removed before spatial filtering. Methods such as the 3σ criterion and interquartile range (IQR) have been widely used in outlier detection. In this paper, we use a more robust IQR method to remove outliers (Beavan, 2005; Nikolaidis, 2002). Calculate IQR using the coordinate component’s first and third quartiles, then define the normal data range, and any value outside this range is considered outliers. In this paper, outliers are defined as data that exceed the 3IQR.

In the preprocessing process of GPS coordinate time series, the effect of the offsets should also be considered in addition to the outliers. Factors such as earthquakes, equipment failures, and environmental changes may cause offsets. We correct the offsets of relevant stations published by CMONOC, as well as some unpublished but clearly observable ones, using the open-source GNSS coordinate time-series analysis package (iGPS) (http://www.ngs.noaa.gov/gpsto). The ‘cleaned’ time series obtained by removing outliers is fitted according to equation (13) to obtain the residual time series.

Filtering of time series using PCA and ICA methods requires no missing data in the time series. However, certain GPS coordinate time series are unavoidably missing data due to a variety of factors (GPS receiver hardware or software, satellite signal propagation). In this paper, we use a widely used interpolation method in meteorology, regularised expectation maximisation (RegEM), to supplement the data for the time periods in the time series that have missing data. The method considers the physical context of the time series and the correlation between stations, and it does not rely on data models or introduce a priori information. RegEM calculates the mean and covariance matrices iteratively from incomplete data sets and then fills in the missing data using the resulting covariance arrays.

4. Results

4.1. PCA results

GNSS coordinates time series are decomposed into PCs/ICs in the time domain and eigenvectors (spatial responses) in the spatial domain by PCA/ICA. The variation in the time domain is represented by PCs/ICs, and the eigenvectors (spatial responses) corresponding to PC/IC reflect the spatial distribution characteristics of the CME.

PCA is applied to the north (N), east (E), and vertical (U) components of the coordinate time series of 10 stations in the small-scale GPS network. The SRs of the first three PCs of the N, E, and U components are shown in Figure 3, and the upward black arrows indicate the proportion of positive SRs, while the downward red arrows indicate the proportion of negative SRs, with the legend indicating 100% SR. Singular value decomposition (SVD) applied to the covariance matrix of N, E, and U component residual series, and the
cumulative variance is obtained, as shown in Figure 2 (a), which shows that the sum of the variances of the first three PCs of N, E, and U components accounts for 81.35%, 82.15% and 87.66% of the total variance, respectively. The variance of PC1 contributes 59.66%, 55.16%, and 63.67%, respectively. This indicates that PC1 can explain the majority of the CME. According to Dong et al. (2006), most stations (50%) have significant

![Figure 2](image_url)

**Figure 2.** (a) Cumulative percentage of the variance of 10 PCs derived from PCA method in small-scale GPS network. (b) Cumulative percentage of the variance of the first 30 PCs derived from PCA method in large-scale GPS network.

![Figure 3](image_url)

**Figure 3.** The first three PCs of the north(N), east(E), and vertical(U) components were obtained using PCA in the small-scale GPS network (the upward arrows are positive SRs, and the downward ones are negative SRs).
(>25%) normalised SRs according to a PC model, and the eigenvalue of this mode exceeds 1% of the sum of all eigenvalues and can be considered a common mode. Figure 3 shows that the SRs of PC1 exhibit obvious spatially uniform localised patterns and PC1 of the three coordinate components has the largest contribution rate, so we attempted to extract CME using the PC1 of the 10 stations based on this criterion.

Then, we apply the PCA method to the residual time series of the large-scale GPS network. Figure 4 shows the first three PCs and corresponding SRs of the N, E, and U components in the large-scale GPS network, which can be seen that PC1 has almost uniform spatial patterns. The cumulative variances of the first 30 PCs are obtained as shown in Figure 2(b), which shows that the sum of the variances of the first three PCs of N, E and U components accounts for 55.05%, 53.03%, and 65.18% of the total variance, respectively, and the contribution of the variance of PC1 to the total variance is 32.37%, 32.08%, and 53.05%, respectively, PC1 of three coordinate components has the largest contribution rate. According to the above criteria, we used the PC1 of 221 stations to extract the CME.

4.2. ICA results

Unlike PCA, ICA can maximise the non-Gaussian component and separate the abnormal signal as an independent component. Figure 5 shows IC9 of the N component, IC10 of the E component and IC10 of the U component for each station in the small-scale GPS network. The above ICs have almost uniform spatial patterns, while the SRs of the other ICs have no obvious spatially uniform localised pattern (Figure 6). CME is a spatially correlated error, and its definition is based on its spatial characteristics; we also define CME according to Dong et al. (2006). The SRs of these ICs with uniform spatial patterns represent the corresponding uniform spatial patterns of the CME in the N, E, and U components. Based on these results, we used IC9 to calculate the CME for the

![Figure 4](image-url)  
**Figure 4.** The first three PCs of the north(N), east(E), and vertical(U) components were obtained using PCA in the large-scale GPS network (the upward arrows are positive SRs, and the downward ones are negative SRs).
N component, IC10 to calculate the CME for the E component, and IC10 to calculate the CME for the U component.

Then, we used the ICA method to filter the residual time series of 221 stations in the large-scale GPS network. Figure 7 shows the results of the large-scale GPS network for N, E, and U components. The IC8, IC18, and IC19 of the N component, IC9, IC16, and IC20 of the E component, and IC18, IC19, and IC20 of the U component exhibit nearly uniform spatial patterns, whereas the SRs of the other ICs exhibit neither an obvious spatially uniform localised pattern nor a high degree of spatial coherence. These ICs with uniform spatial patterns constitute CME in N, E, and U components. Based on these results, we calculate the CME for the N, E, and U components.

ICA is a blind source separation algorithm, which can decompose the mixed signal into several independent source signals. The responses of GPS stations to non-tectonic movement are not identical in different-scale GPS networks, and the CME is also different. After using ICA to decompose the time series, we discovered that the CME in each coordinate component had one significant IC in the small-scale GPS network, and the CME in each coordinate component had three significant ICs in the large-scale GPS network. As a result, the number of ICs used to calculate CME in different-scale networks varies.

4.3. CWSF results

4.3.1. Results for small-scale network

To further understand the effect of the CWSF method in removing CME, we used this method to filter the small-scale GPS network and the large-scale GPS network above, comparing the effect with PCA and ICA. Figure 8 shows the comparison of the JXHK station residual time series after filtering by PCA, ICA, and CWSF methods. The grey line is the original GPS residual time series, and the blue line is the filtered residual time series. As shown in Figure 8(a), the trend of the residual time series of the U component of the JXHK station after PCA filtering changes from the original residual series, and some non-linear quantities are removed. Figure 8(b, c) compared the residual time series of the JXHK station after ICA and CWSF filtering, respectively, and found that neither method removes the non-linear signal from the JXHK station’s U component.

We use equation (13) to fit the raw time series after removing the outliers and correcting the offsets, and then remove the linear trend and periodic term to obtain the residual time series containing CME. As shown in Figure 8(a), the non-linear signal in the residual time series of the U component is not accurately identified and separated, so we include this non-linear signal in the residuals and then filter them using three methods. As can be clearly seen from Figure 8(a), the PCA method removes the non-linear signal and part of

Figure 5. The results of the small-scale GPS network for north(N), east(E), and vertical(U) components using the ICA method (the upward arrows are positive SRs, the downward are negative SRs).

Figure 6. SR ratios of N, E, and U components in the small-scale GPS network.
CME, which may result in some original information in the residuals being removed. However, ICA and CWSF can identify CME well, and the non-linear signal is not considered as CME being removed together.

As shown in Figure 8(a,b) after the PCA and ICA methods were used, the original residual time series after data interpolation has a significant offset during 2015–2016. This offset may be caused by the interpolation method, but there are inevitable introducing errors into the data regardless of which interpolation method is used. However, the CWSF method does not require interpolation, thus avoiding the error effect caused by data interpolation.

The changes in RMS of the residual time series before and after filtering by the three methods are shown in Table 1. After PCA filtering, the RMS of the residual time series in the N, E, and U components was reduced by 33.68%, 33.02%, and 34.13%, respectively. The RMS of the residual time series in the N, E, and U components was reduced by an average of 26.94%, 32.56%, and 29.79% after ICA filtering, respectively. The RMS in the N, E, and U components of the residual time series decreased by 38.34%, 35.35% and 33.26%, respectively, after CWSF. The PCA method reduced the RMS of the JXHK station in the U component by 59.74%, which was much higher than its average value, confirming that the above PCA removed some of the station’s original information.

From Figure 3, we can see that the SR of the U-component at JXHK station is much larger than other stations, and considering it is near Poyang Lake and Yangtze River, we calculated the displacement of the U-component at 10 stations in the small-scale network due to hydrology loading (Figure 9). Then, we find that the difference in hydrology loading on these 10 stations is not significant. Since the composition of the CME cannot be clarified at present, we then tried to re-perform the PCA analysis after removing the JXHK station in the U component alone, choosing the first principal component to calculate the CME using the average principal component of other stations as the JXHK principal component, and the results of the principal component analysis are shown in Figure 10(a), and Figure 10.
(b) is a comparison of the residual time series before and after PCA filtering at JXHK station, which shows that the trend of the filtered U-component residuals does not change compared to the original residuals after removing the anomalous component. This result illustrates that the PCA method cannot accurately identify the anomalous signal, further verifying that the ICA and CWSF methods are better than PCA.

In conclusion, the PCA method may remove some of the original information from the GPS time series, and the ICA method may be negatively affected by the interpolation method. The CWSF method, on the other hand, avoids these effects and is also found to be effective as evidenced by statistical testing in terms of RMS computation. As a result, we find that the CWSF method is optimal for addressing the CME retrieval for this study’s example of a small-scale GNSS network.

### 4.3.2. Results for large-scale network

Figure 11 shows the RMS of residual time series before and after the PCA, ICA, and CWSF filtering in the large-scale GPS network. The RMS values of the residual time series in all three components of the GPS stations in each region improved significantly after using the CWSF method, as shown in the figure. The majority of the RMS improvement percentages for this method ranged from 20% to 50%, with mean values of 41.99%, 40.86%, and 46.79% for the N, E, and U components, respectively, with the maximum improvement occurring in the U component at XJML station (69.54%) and the minimum occurring in the E component at XIAG station (7.43%). The mean values for the percentage improvement in RMS for the three components of the PCA method were 20.94%, 20.12%, and 33.80%, respectively. The mean RMS improvement percentage values for the three ICA method components were 23.65%, 23.86%, and 23.42% (Table 2), respectively. CWSF has a higher RMS reduction percentage of residual time series than the other two methods.

### 5. Discussion

We select five stations in the CMONOC (Figure 12) and use stacking and CWSF to remove CME to determine the effect of conventional stacking on the extraction of CME from large-scale networks. Table 3 displays the correlation coefficients of the U-component residual time series of the BJFS station and the other five stations before and after filtering. The CME is based on the spatial correlation between GPS stations, and the CME is larger for the regional network with a strong spatial correlation. The spatial filtering method weakens the correlation between stations, and the correlation coefficient decreases accordingly. Table 3 shows that after stacking filtering, the correlation coefficient between stations increases, indicating that stacking is unsuitable for large-scale networks. Except for the LHAS station, the correlation between all stations and the BJFS station’s U-component residual time series was reduced after CWSF filtering.

### 6. Conclusions

CME has significantly affected time-series accuracy in regional GNSS networks. Multiple spatiotemporal filtering methods can effectively remove CME and improve
the accuracy of GNSS coordinate time series. For CMONOC, there is a lack of research to demonstrate which method, ICA or CWSF, is better at removing CME. In this paper, we apply three filtering methods (PCA, ICA, and CWSF) to GPS networks of different scales to compare the most effective method for CMONOC.

Applying PCA to the small-scale GPS network, we found a strong local response in the U component for the JXHK station, we then calculated the CME using PC1 of these 10 stations and found that the trend of the residual time series changed after removing the CME. When we tried to re-perform the PCA after removing the U component of the JXHK station, the results returned to normal. However, both ICA and CWSF identified this abnormality and showed an excellent filtering effect without removing the abnormal component, the RMS of the residual time series in the N, E, and U components were reduced by an average of 26.94%, 32.56%, and 29.79% after ICA filtering, the average reduction in RMS after CWSF is 38.34%, 35.35%, and 33.26%, respectively.

We apply the three filtering methods to a large-scale GPS network that does not include the abnormal station mentioned above. The RMS of the N, E, and U components of the residual time series were reduced by 41.99%, 40.86%, and 46.79% on average, respectively, which is much more effective in removing CME than PCA or ICA. CWSF method interferes little with the data, which is one of its advantages. From the above analysis, we found that it is more advantageous to apply the CWSF method to CMONOC for spatial filtering. Using correlation analysis, we have proved the effectiveness of the CWSF method.

The results could provide an important reference to removing CME from GNSS coordinate time series, but this study also contains some limitations: (1) We did
Figure 11. The RMS of residual time series before and after the PCA, ICA, and CWSF filtering (%); the colour bar represents the percentage of RMS reduction.

Table 2. RMS variations of residual series before and after filtering using three methods for large-scale GPS network (unit: mm).

|     | N   | Filter | Reduce | E   | Filter | Reduce | U   | Filter | Reduce |
|-----|-----|--------|--------|-----|--------|--------|-----|--------|--------|
| PCA | 1.96| 1.55   | 20.94% | 1.67| 20.12% |        | 5.22| 33.80% |
| ICA | 1.49| 1.49   | 23.65% | 2.09| 23.42% |        | 7.89| 23.42% |
| CWSF| 1.14| 1.14   | 41.99% | 1.24| 40.86% |        | 4.20| 46.79% |

Figure 12. Location of the 5 stations used for correlation analysis.
Table 3. Comparison of correlation coefficients before and after filtering of U-component residual time series between BJFS station and five other stations.

| Site       | Correlation | Stacking | CWSF | Distance/km |
|------------|-------------|----------|------|-------------|
| DLHA-BJFS  | 0.54        | −0.58    | 0.06 | 1629        |
| GUAN-BJFS  | 0.11        | −0.32    | 0.02 | 1844        |
| LHAS-BJFS  | 0.04        | −0.26    | −0.27| 2513        |
| WUHN-BJFS  | 0.24        | −0.14    | 0.02 | 1020        |
| XIAG-BJFS  | 0.15        | 0.25     | −0.04| 2133        |

not verify the results obtained by using different interpolation methods; (2) We did not extract CME for the three methods for further analysis. CWSF filtering can also be applied to other GNSS networks to detect and remove CME, and analysing the CME will help to investigate its physical origins.

Acknowledgments

We would like to acknowledge the Crustal Movement Observation Network of China (CMONOC). Thanks to GPS Toolbox (http://www.ngs.noaa.gov/gpstoolbox) for providing the iGPS tool package (developed by Tian) for GPS coordinate time-series analysis. The Generic Mapping Tool (GMT) software package was used to plot the figures. The authors would also like to thank two anonymous reviewers and the Editor, Professor Yuyu Zhou for their insightful comments and suggestions, which help to improve the manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Data availability statement

These data were derived from the following resources available in the public domain:CMONOC time-series products are available from http://www.cgps.ac.cn; The iGPS code for GNSS coordinate time-series analysis is available from the GPS Toolbox http://www.ngs.noaa.gov/gpstoolbox. Environmental loading products are available from http://esmdata.gfz-potsdam.de:8080/repository

Abbreviations

CME Common Mode Error
CMONOC Crustal Movement Observation Network of China
CWSF Correlation-Weighted Spatial Filtering
GNSS Global Navigation Satellite System
GPS Global Positioning System
IC Independent Component
ICA Independent Component Analysis
IGS International GNSS Service
IQR Interquartile Range
PC Principal Component
PCA Principal Component Analysis
RegEM Regularized Expectation Maximization
RMS Root Mean Square
SVD Singular Value Decomposition
SOPAC Scripps Orbit and Permanent Array Center
SR Spatial Response

Authors’ contributions

HW performed the spatiotemporal filtering analysis and drafted the manuscript. Li and Shu planned the study and participated in the design and drafting of the manuscript. Other authors participated in the revision of the manuscript. All authors read and approved the final manuscript.

Funding

This research was funded by the National Natural Science Foundation of China [42074006], the State Key Laboratory of Geo-Information Engineering and Key Laboratory of Surveying and Mapping Science and Geospatial Information Technology of the Ministry of Natural Resources (MNR), CASM [no. 2021-00-05], the National Key Research and Development Program of China [2017YFA0603104], The State Key Program of the National Natural Science Foundation of China [41531069].

ORCID

C. K. Shum http://orcid.org/0000-0001-9378-4067
Fei Li http://orcid.org/0000-0002-4497-5868

References

Beavan, J. (2005). Noise properties of continuous GPS data from concrete pillar geometric monuments in New Zealand and comparison with data from US deep drilled branded monuments. Journal of Geophysical Research: Solid Earth, 110(B8), 642. https://doi.org/10.1029/2005JB003642
Comon, P. (1994). Independent component analysis, a new concept? Signal Processing, 36(3), 287–314. https://doi.org/10.1016/0165-1684(94)90029-9
Dong, D., Fang, P., Bock, Y., Cheng, M., & Miyazaki, S. (2002). Anatomy of apparent seasonal variations from GPS-derived site position time series. Journal of Geophysical Research: Solid Earth, 107(B4), ETG 9-1-ETG 9–16. https://doi.org/10.1029/2001JB000573
Dong, D., 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. Journal of Geophysical Research: Solid Earth, 111(B3), 806. https://doi.org/10.1029/2005JB003806
He, X., Hua, X., Yu, K., Xuan, W., Lu, T., Zhang, W., & Chen, X. (2015). Accuracy enhancement of GPS time series using principal component analysis and block spatial filtering. Advances in Space Research, 55(5), 1316–1327. https://doi.org/10.1016/j.asr.2014.12.016
Hyrvarinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. IEEE Transactions on Neural Networks, 10(3), 626–634. https://doi.org/10.1109/72.761722
Hyrvarinen, A., & Oja, E. (2000). Independent component analysis: Algorithms and applications. Neural Networks, 13(4–5), 411–430. https://doi.org/10.1016/S0893-6080(00)00026-5
Kreemer, C., & Blewitt, G. (2021). Robust estimation of spatially varying common-mode components in GPS time-series. Journal of Geodesy, 95(1), 1–19. https://doi.org/10.1007/s00190-020-01466-5
Li, W., Shen, Y., & Li, B. (2015). Weighted spatiotemporal filtering using principal component analysis for analyzing regional GNSS position time series. *Acta Geodaetica Et Geophysica, 50*(4), 419–436. https://doi.org/10.1007/s40328-015-0100-1

Li, W., Li, F., Zhang, S., Lei, J., Zhang, Q., & Yuan, L. (2019). Spatiotemporal filtering and noise analysis for regional GNSS network in Antarctica using independent component analysis. *Remote Sensing, 11*(4), 386. https://doi.org/10.3390/rs11040386

Liu, B., Dai, W., Peng, W., & Meng, X. (2015). Spatiotemporal analysis of GPS time series in vertical direction using independent component analysis. *Earth, Planets and Space, 67*(1), 1–10. https://doi.org/10.1186/s40623-015-0357-1

Liu, B., King, M., & Dai, W. (2018). Common mode error in Antarctic GPS coordinate time-series on its effect on bedrock-uplift estimates. *Geophysical Journal International, 214*(3), 1652–1664. https://doi.org/10.1093/gji/ggy217

MA, C., Li, F., ZHANG, S.-K., LEI, J.-T., Dong-Chen, E., HAO, W.-F., & ZHANG, Q.-C. (2016). The coordinate time series analysis of continuous GPS stations in the Antarctic Peninsula with consideration of common mode error. *Chinese Journal of Geophysics, 59*(8), 2783–2795. https://doi.org/10.6038/cjg20160806

Márquez-Azúa, B., & DeMets, C. (2003). Crustal velocity field of Mexico from continuous GPS measurements, 1993 to June 2001: Implications for the neotectonics of Mexico. *Journal of Geophysical Research: Solid Earth, 108*(B9), 241. https://doi.org/10.1029/2002JB002241

Ming, F., Yang, Y., Zeng, A., & Zhao, B. (2017). Spatiotemporal filtering for regional GPS network in China using independent component analysis. *Journal of Geodesy, 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.

Serpelloni, E., Faccenna, C., Spada, G., Dong, D., & Williams, S. D. (2013). Vertical GPS ground motion rates in the Euro-Mediterranean region: New evidence of velocity gradients at different spatial scales along the Nubia-Eurasia plate boundary. *Journal of Geophysical Research: Solid Earth, 118*(11), 6003–6024. https://doi.org/10.1002/2013JB010102

Shen, Y., Li, W., Xu, G., & Li, B. (2014). Spatiotemporal filtering of regional GNSS network’s position time series with missing data using principle component analysis. *Journal of Geodesy, 88*(1), 1–12. https://doi.org/10.1007/s00190-013-0663-y

Tian, Y., & Shen, Z. K. (2016). Extracting the regional common-mode component of GPS station position time series from dense continuous network. *Journal of Geophysical Research: Solid Earth, 121*(2), 1080–1096. https://doi.org/10.1002/2015JB012253

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. *Journal of Geophysical Research: Solid Earth, 102*(B8), 18057–18070. https://doi.org/10.1029/97JB01378

Wu, S., Nie, G., Liu, J., Wang, K., Xue, C., Wang, J., Li, H., Peng, F., & Ren, X. (2019). A sub-regional extraction method of common mode components from IGS and CMONOC stations in China. *Remote Sensing, 11*(11), 1389. https://doi.org/10.3390/rs11111389

Yuan, P., Jiang, W., Wang, K., & Sneeuw, N. (2018). Effects of spatiotemporal filtering on the periodic signals and noise in the GPS position time series of the crustal movement observation network of China. *Remote Sensing, 10*(9), 1472. https://doi.org/10.3390/rs10091472