Analysis of the regional GNSS coordinate time series by ICA-weighted spatio-temporal filtering

ZHENG HOU, ZENGZHANG GUO* and JIUSHENG DU

School of Surveying and Land Information Engineering, Henan Polytechnic University, Jiaozuo 454003, People's Republic of China.

**Corresponding author. e-mail: guohpu@163.com*

MS received 16 November 2017; revised 25 July 2018; accepted 25 April 2019; published online 5 July 2019

Independent component analysis (ICA) is a blind source signal separation method which can effectively estimate high-order information and thus can effectively extract the common mode errors (CMEs) of a regional global navigation satellite system (GNSS) observation network. In this paper, ICA is used for the weighted filtering (WICA) and the extraction of CMEs of a regional GNSS observation network with the root mean square error (RMSE) of daily solution taken as the weighting factor. Through an analysis of the observed data from 19 valid stations of the Crustal Movement Observation Network of China (CMO-NOC) in North China, it is shown that the coordinate series precision of 13, 16 and 12 stations in the N, E and U directions, respectively, after filtering by WICA is higher than that by the traditional ICA method. The average correlation coefficient of the coordinate time series for each station after filtering is obviously decreased. Two simulation experiments are designed to extract known CMEs. It is shown that CMEs can be recovered better by WICA and that the standard deviations of most stations after filtering are smaller than those by ICA. The results from the real data and simulation experiments suggest that the RMSE of coordinate series be considered in spatio-temporal filtering.

Keywords. Independent component analysis; common mode error; root mean square error; spatial response; correlation coefficient.

1. Introduction

Since the 1990s, China has continuously accelerated the construction of global navigation satellite system (GNSS) base stations. The establishment and continuous improvement of the Crustal Movement Observation Network of China (CMO-NOC) provides strong data support for scientific research such as crustal deformation monitoring in China (Yang *et al.* 2008; Wang *et al.* 2011a; Li *et al.* 2012) and seismic research (Wang *et al.* 2011b; Wu *et al.* 2013). The GNSS coordinate time series mainly consists of tectonic deformation and non-tectonic deformation. In a regional GNSS network,

the common mode error (CME) dominates, which is usually a non-structural error caused by satellite orbital parameters, terrestrial reference frames, atmospheric or soil water quality loads (van Dam *et al.* 2001). The accurate extraction of CMEs using spatio-temporal filtering is critical to the detection of nonlinear crustal micro-deformation.

Wdowinski *et al.* (1997) used the regional stacking filtering to extract CMEs in the Southern California region for the first time. This method assumes that the distribution of CMEs is uniform and applies only to a small regional scale (less than 500 km). Taking into account the inconsistency of the spatial response (SR) of stations in different

regions, Nikolaidis (2002) introduced the single-day solution variance as a weighting factor, and proposed weighted stacking filtering. MárquezAzúa and DeMets (2003) used the distance between stations and the length of the GPS coordinate time series as weighted factors for stacking filtering to overcome the limitations of the regional scale. Tian and Shen (2011) took the correlation coefficients between stations as the weighting factors for weighted stacking filtering with the influence of factors such as distances between stations taken into account, and analysed the ‘common mode component’ of different scales. In the principal component analysis (PCA) (Shum 1994; Blanz and Vetter 2002; Dong *et al.* 2006), the eigenvalue of covariance matrix can be decomposed based on the GNSS coordinate time series, and spatial and temporal responses (TRs) can be obtained. Shen *et al.* (2014) improved the PCA method, used principal components to restore the missing coordinate data, and then extracted the CMEs. Li *et al.* (2015) proposed the method of PCA weighted spatio-temporal filtering (WSF) with the positioning formal error of daily solution taken as the weighting factor, and achieved better filtering effects.

Independent component analysis (ICA) is a blind signal separation method (Comon 1994). ICA can be used to estimate the independent information using high-order statistics, and analyse non-Gaussian data effectively. It was used in the crustal deformation analysis (Gualandi *et al.* 2016), mass loading information separation (Liu *et al.* 2015; Peng *et al.* 2017), etc. Ming *et al.* (2017) analysed the continuous GNSS coordinate time series of 259 stations in the Crustal Movement Observation Network of China (CMONOC) by ICA, and extracted the independent components with ‘common spatial characteristics’ in Mainland China as the principal components of CMEs. Liu *et al.* (2015) analysed the spatial and temporal features of vertical GPS coordinate time series in the UK and in the Sichuan–Yunnan region of China by ICA, and evaluated the influence of atmospheric and soil moisture mass loading.

In this paper, an ICA-based WSF (WICA) approach is developed with the root mean square error (RMSE) of daily solution taken into account. The CMEs of 29 stations are extracted in the region of North China. The standard deviations and average correlation coefficients of these stations have been improved in different degrees after filtering by WICA, compared with those by the

traditional ICA. The rest of this paper is organised as follows. In section 2, the ICA and WICA approaches are described. Real-time series are processed and analysed in section 3. In section 4, the results of the analysis of the synthetic time series are presented. Section 5 comprises the discussion of this paper. Conclusions are summarised in section 6.

2. Method

2.1 Independent component analysis

ICA is a blind signal separation method used to separate mixed signals into multiple statistically independent source signals by high-order statistics. The principle of ICA is shown in figure 1. Assume that the observed data consist of the daily solution of n GPS continuous observation stations for m days. The ICA mathematical model can be expressed as

$$X_{n \times m} = A_{n \times n} s_{n \times m}. \quad (1)$$

where X , A and s represent the observation matrix, the mixture matrix and the source signal matrix, respectively. The column of the mixture matrix A represents the SR of the signal, and the row of the source signal matrix s represents the TR of the signal. ICA is used to separate the ICs by solving for the mixed matrix W ($W = A^{-1}$). W can be obtained in two steps based on the optimisation criteria:

- (1) Whiten the observed matrix X by a whitening matrix H , let $V = HX$ and $E(VV^T) = I$ (I is a unit matrix).
- (2) Calculate the transformation matrix by the specific optimisation rule, let $Y = GV$ where Y (equation 2) is the optimal approximation matrix s and $W = GH$:

$$Y_{n \times m} = W_{n \times n} X_{n \times m} = W_{n \times n} A_{n \times n} s_{n \times m} = s_{n \times m}. \quad (2)$$

FastICA based on negentropy is a fast optimisation iterative ICA algorithm. It is a common quantitative measure of the non-Gaussianity of a random variable and is used in the ICA in this paper. The detailed steps are as follows:

- (1) Centralise and whiten the observation matrix.
- (2) Choose an initial vector of the unit norm (random) w .

Figure 1. Fundamental principles of ICA.

- (3) Update w through $w(k+1) = E[xg(w^T(k)x)] - E[g'(w^T(k)x)]w$.
- (4) Normalise w by $w(k+1) = \frac{w(k+1)}{\|w(k+1)\|}$.
- (5) Go back to Step (3) if not converged.

ICA is used to estimate the source signal from the observation matrix under the premise that the mixture matrix and the source signal matrix are unknown, and thus multiple sets of matrices A and s will certainly be obtained. To ensure the uniqueness of solution, the following is usually assumed for the ICA model:

- (1) the source signals are statistically independent of each other;
- (2) at most, one of the source signals is subject to the Gaussian distribution;
- (3) the mixture matrix is an invertible square matrix.

2.2 Weighted ICA

In this paper, the filtering effects are verified by the weighted filtering using the ICA method with the RMSE of daily solution taken as the weighting factor:

$$x'_{ij} = \frac{x_{ij}}{\sigma_{ij}} (i = 1, 2, \dots, n, j = 1, 2, \dots, m), \quad (3)$$

$$X'_{n \times m} = A'_{n \times n} s'_{n \times m}. \quad (4)$$

In the equations above, x_{ij} and σ_{ij} are the coordinate residual and the RMSE of the epoch j for station i , respectively. A' and s' represent a weighted mixture matrix and a source signal matrix, respectively.

3. Real data processing and analysis

3.1 Data source

For this paper, three consecutive years of GNSS coordinate data from 29 continuous stations (figure 2) of CMONOC in North China were collected. The GAMIT software was used to solve the site coordinates and the parameters such as satellite

orbit, and GLOBK was used to merge the single day loose solution obtained and the H file of the IGS global solution provided by SOPAC through the seven-parameter conversion. After that, the merged file was constrained by ITRF2008 through the globally distributed points of international terrestrial reference frame. At the same time, the impact of the following factors on the GNSS stations were taken into account: the annual and semi-annual terms of GNSS stations, large earthquakes (the earthquake at the Kunlun Mountains in 2001, the earthquake of sea areas near the north of Sumatra in 2004, the earthquake at Wenchuan in 2008, the earthquake at Tokai in 2011, etc.), instrument change and antenna replacement (<http://www.cgps.ac.cn/index.action>).

Due to factors such as communication and instrument failure, some of the station data were missing or accompanied by abnormal individual observation data. Therefore, stations with a data loss rate of less than 28% were taken as valid stations, and the number of valid stations at the same epoch was required to be greater than 5% (Dong *et al.* 2006). For abnormal data to be detected, the threshold for the single day coordinate solution and the coordinate residual in the N, E and U directions were set to 50, 50, 100; 10, 10, 20 mm, respectively, with over-limit data discarded. As shown in figure 3, most stations have a data missing ratio of less than 5% after the removal of abnormal data while the three stations of SDQD, SDCY and LNDD have a relatively large data missing ratio, but all less than 20%. The gap can be recovered by linear interpolation when the number of consecutive missing data is less than 5. If the number is greater than 5, multipoint cubic spline interpolation can be used (Wu and Huang 2004). The continuous coordinate time series at 1095 epochs from 29 stations in North China were obtained finally.

3.2 Abnormal station removal

An abnormal station is affected by strong regional factors and exhibits an abnormal SR. The effective detection and removal of abnormal stations is

Figure 2. Distribution of 29 GNSS continuous stations in north China.

Figure 3. Data missing ratio of 29 stations.

important for the accurate extraction of CMEs and the exploration of the impact of geophysical parameters on regional crustal deformation. In this paper, correlation coefficients and quartile deviations are used to detect abnormal stations.

Correlation coefficients (equation 5) can reflect the degree of similarity between data from different stations. When the correlation coefficient $r_{xy} < 0.5$, the correlation between stations was weak (Chen *et al.* 2012; Cui 2017). To avoid the removal of valid stations, the threshold for correlation coefficients is set to 0.4. After the average correlation coefficient of each station was calculated, the five stations of TAIN, SXCZ, SDLY, LNSY and JSLY were removed (figure 4):

$$r_{xy} = \frac{\sum_{k=1}^N (x_k - \bar{x})(y_k - \bar{y})}{\sqrt{\sum_{k=1}^N (x_k - \bar{x})^2 \sum_{k=1}^N (y_k - \bar{y})^2}}. \quad (5)$$

The abnormal stations were further detected using quartile deviation (equation 6) (Ming *et al.* 2017). In the course of the experiment, it was found that different abnormal stations were detected by the methods of ICA and WICA. To ensure that the experimental environment is the same, the five abnormal stations of ZHNZ, SDRC, LNDD, HECX and HAJY detected by the two methods were removed at the same time:

Figure 4. Average correlation coefficients of 29 stations in the N, E and U directions.

$$sr_j^i - \text{med}(SR_i) > 3 * \text{IQR}(SR_i), \quad (6)$$

where sr_j^i is the SR value of the i th component of the j th station, and $\text{med}(SR_i)$ and $\text{IQR}(SR_i)$ are the median and quartile deviation of the SR of the i th component, respectively. When the inequality is valid, the specific station is considered as an abnormal station.

3.3 Non-Gaussian analysis

ICA is usually based on the non-Gaussianity of source signals. Gaussian distribution (normal distribution) is the distribution of random variables described by mathematical expectations and variances. The closer the variable value to its mathematical expectation, the greater the probability of the occurrence. Relative to Gaussian distribution, the kurtosis (equation 7) reflects the degree of concentration of the data on both sides of the mean. When $\text{Kurt} = 0$, the observed data have a Gaussian distribution. When $\text{Kurt} \neq 0$, the observed data have a steep distribution deviating from the Gaussian distribution curve, which is a non-Gaussian distribution. Many studies have shown that not only Gaussian white noise exists, but also coloured noise exists in the GNSS continuous coordinate time series (Williams 2004; Dmitrieva *et al.* 2015; Zhao *et al.* 2015; Amiri-Simkooei 2016). Therefore, it is necessary to determine whether the observed data are non-Gaussian by equation (7) before filtering (Ming *et al.* 2017). As shown in figure 5, the number of kurtosis values greater than zero in the N, E and U directions accounts for 73.7%, 94.7% and 68.4%, respectively, which are all higher than 60%. Therefore, it is determined that the coordinate time series of the selected region has high-order statistical properties, and the ICA method

can be used to extract the independent component theoretically:

$$\text{Kurt} = E(x^4) - 3E^2(x^2). \quad (7)$$

3.4 Definition of CME

The source of CME is unclear and there is no strict definition so far (Tian and Shen 2009). Dong *et al.* (2006) analysed the definition standard of CME by the PCA method. Considering the regional spatial characteristics of CME, this paper attempts to define the CME of ICA spatio-temporal filtering based on the standard analysis of Dong *et al.* (2006):

- (1) the SR values are in the same direction (the symbols are the same);
- (2) most stations (>50%) have significant SR (>25%).

The independent component satisfying the above two conditions is called a principal component, which is the main body for calculating CME. The formula is as follows:

$$\text{CME}_{\text{ICA}} = \sum_{k \in R} A_k S_k, \quad (8)$$

where A_k and S_k represent the SR and TR of the k th principal component, respectively.

3.5 Real-time series analysis

After eliminating abnormal stations, the average RMSEs of the 19 stations in the N, E and U directions are shown in figure 6. In this figure, the precisions of the coordinate series in the horizontal directions are basically the same, all between 1 and 3 mm, while the precisions in the U direction are

Figure 5. Kurtosis values of GNSS stations of the CMONOC in north China.

relatively low, between 5 and 8 mm for most stations. This is consistent with the property that the measurement precision of GNSS in the vertical direction is lower than that in the horizontal direction.

According to the definition of CME in section 3.4, ICA (N: IC19; E: IC15; U: IC19) and WICA (N: IC17; E: IC12, IC14, IC15, IC17; U: IC11, IC12, IC18) are obtained. Figure 7 (one principal component in each of the N, E and U directions is drawn) shows that the fluctuations of the time response are relatively stable in the N direction, but obvious in the E and U directions, and a step occurred in October 2014. In this figure, the SRs of the stations in the Beijing–Tianjin region are more consistent with higher responsiveness than those of other stations for the principal components extracted by WICA, which indicates that the CMEs are regional. The SR distributions for the principal components extracted by ICA are more uniform, compared with those by WICA, which shows that the length of the arrow shaft of each station is more consistent for ICA. However, this does not mean that the CME extracted by ICA is more accurate. For WICA, more principal components were extracted, but their geophysical significance needs further study.

After the CME was removed by ICA, the coordinate time series precision expression is as follows:

$$\hat{x}_{ij} = x_{ij} - \sum_{k \in R} a_i(k) s_j(k), \quad (9)$$

$$\sigma_I(i) = \sqrt{\frac{1}{m-1} \sum_{j=1}^m (\hat{x}_{ij} - \bar{\hat{x}}_i)^2}. \quad (10)$$

The calculation formula for the standard deviation of each station after filtering by WICA is

$$\hat{x}'_{ij} = x'_{ij} - \sum_{k \in R} a'_i(k) s'_j(k), \quad (11)$$

$$\sigma_W(i) = \sqrt{\frac{1}{m-1} \sum_{j=1}^m [(\hat{x}'_{ij} - \bar{\hat{x}}'_i) \sigma_{ij}]^2}. \quad (12)$$

In the equations above, $\hat{x}_{ij}$, $\hat{x}'_{ij}$, σ_I , σ_W , $\bar{\hat{x}}_i$ and $\bar{\hat{x}}'_i$ are the coordinate residual, standard deviations and average values of the coordinate residual time series, respectively, after the CMEs are removed by the ICA and the WICA methods.

Figure 6. Average RMSEs of 19 stations.

Figure 8 shows the calculated standard deviations of the coordinate series of the 19 stations after filtering. WICA has a higher overall precision after the removal of CMEs than ICA (table 1) and the average standard deviations in the N, E and U directions are reduced by 4.12%, 13.84% and 4.61%, respectively. However, the precision of 6, 3 and 7 stations in the three directions, respectively, is still lower than that for ICA after filtering for unclear reasons.

The CME is a spatially correlated error, and after it is removed, the correlation of the GNSS coordinate time series of each station should be reduced. Table 2 shows the average correlation coefficients of the 19 stations after the removal of the CMEs by the two methods ICA and WICA were collected. It can be seen that the correlation coefficients of all stations in the three directions are reduced significantly and that the correlation coefficients of 13, 17 and 12 stations in the N, E and U directions after filtering by WICA are less than the results by ICA. Therefore, it can be inferred that WICA is more effective in the extraction of CMEs.

4. Simulation experiment

To further verify the feasibility of the WICA method, two simulation experiments were designed. In the first simulation experiment, the extracted CMEs (figure 9) of the six stations (HELQ, HELY, HEZJ, LNYK, SDZB and SXLQ) in the N direction were selected as the simulation signals, while the noise and flicker noise were added (equation (13)) using MATLAB. The

signal-to-noise ratio (SNR) was set to 0.8:1, 1:1 and 1.2:1, respectively:

$$y(i, j) = g(i, j) + \xi(i, j) + \gamma(i, j), \quad (13)$$

where $g(i, j)$ and $\xi(i, j)$ are the CME and noise of the simulated station, respectively.

The CMEs of the six stations were extracted by ICA and WICA to calculate the standard deviation ε of the CMEs after recovery and the standard deviation σ of the coordinate series after the removal of the CMEs:

$$\varepsilon_I(i) = \sqrt{\frac{1}{m-1} \sum_{j=1}^m (g_I(i, j) - g(i, j))^2}, \quad (14)$$

$$\varepsilon_W(i) = \sqrt{\frac{1}{m-1} \sum_{j=1}^m (g_W(i, j) - g(i, j))^2}, \quad (15)$$

$$\hat{y}_I(i, j) = y_I(i, j) - g_I(i, j), \quad (16)$$

$$\sigma_I(i) = \sqrt{\frac{1}{m-1} \sum_{j=1}^m (\hat{y}_I(i, j) - \bar{\hat{y}}_I(i))^2}, \quad (17)$$

$$\hat{y}'_W(i, j) = y'_W(i, j) - g_W(i, j), \quad (18)$$

$$\sigma_W(i) = \sqrt{\frac{1}{m-1} \sum_{j=1}^m [(\hat{y}'_W(i, j) - \bar{\hat{y}}'_W(i, j)) \sigma_{ij}]^2}, \quad (19)$$

Figure 7. Spatio-TRs in the N, E, and U directions by ICA (left) and WICA (right).

Figure 8. Standard deviations of continuous coordinate time series of 19 stations filtered by ICA and WICA.

Table 1. Standard deviations of coordinate series after filtering by ICA, WICA (mm).

	Unfiltered	ICA	WICA	Percentage (%)
N	1.4582	1.1436	1.0965	4.12
E	1.5947	1.1500	0.9908	13.84
U	5.3752	3.9545	3.7721	4.61

where $g_I(i, j)$, $g_W(i, j)$, $\varepsilon_I(i)$, $\varepsilon_W(i)$, $\hat{y}_I(i, j)$ and $\hat{y}'_W(i, j)$ are the CMEs extracted by ICA and WICA in the simulation experiment, their precisions of recovery and the coordinate series after the CMEs are removed, respectively. Symbol ‘-’ denotes average.

The results of the simulation experiment (table 3) show that WICA is more effective in

Table 2. Average correlation coefficients of the 19 stations in the N, E and U directions before and after filtering by ICA and WICA.

Station	N			E			U		
	Unfiltered	ICA	WICA	Unfiltered	ICA	WICA	Unfiltered	ICA	WICA
TJBH	0.70	0.44	0.32	0.59	0.56	0.42	0.66	0.45	0.43
SXLQ	0.63	0.40	0.36	0.57	0.51	0.44	0.66	0.45	0.37
SDZB	0.70	0.19	0.20	0.44	0.45	0.48	0.67	0.42	0.38
SDQD	0.67	0.41	0.40	0.55	0.52	0.40	0.62	0.44	0.44
SDJX	0.74	0.32	0.32	0.61	0.52	0.42	0.65	0.55	0.41
SDCY	0.69	0.26	0.15	0.58	0.48	0.29	0.64	0.34	0.37
LNYK	0.71	0.39	0.28	0.57	0.42	0.30	0.68	0.49	0.54
JIXN	0.66	0.30	0.29	0.55	0.53	0.43	0.67	0.43	0.38
HEZJ	0.73	0.38	0.30	0.62	0.47	0.34	0.69	0.44	0.46
HEYY	0.71	0.38	0.27	0.61	0.41	0.37	0.64	0.40	0.29
HELY	0.74	0.28	0.32	0.61	0.53	0.53	0.67	0.49	0.42
HELQ	0.75	0.35	0.40	0.63	0.51	0.49	0.68	0.46	0.49
HECD	0.73	0.45	0.32	0.64	0.49	0.33	0.68	0.48	0.47
HECC	0.75	0.45	0.32	0.62	0.54	0.47	0.70	0.47	0.42
HAHB	0.74	0.35	0.36	0.63	0.49	0.39	0.61	0.48	0.33
BJYQ	0.72	0.46	0.26	0.61	0.50	0.43	0.68	0.35	0.39
BJSH	0.64	0.33	0.32	0.55	0.53	0.45	0.65	0.44	0.34
BJGB	0.72	0.45	0.29	0.61	0.55	0.45	0.70	0.50	0.44
BJFS	0.68	0.18	0.25	0.48	0.27	0.18	0.69	0.51	0.51

Figure 9. CMEs of six stations (10 mm offset per station).

filtering (σ_W) and restoring CMEs (ε_W) with different SNRs. After filtering by WICA and ICA, σ_W and ε_W are smaller than σ_I and ε_I , respectively, for most stations. Nonetheless, there are still a few stations with lower precision for WICA.

The second simulation experiment was repeated 50 times. The extracted CMEs of the same six stations in the N direction were selected as the simulation signals, while the noise were added. The results of σ_I , σ_W and ε_I , ε_W are presented in figures 10 and 11. As shown in these two figures, over 80% of σ_W (except SXLQ) and ε_W were

smaller than σ_I and ε_I , respectively, for each station. Therefore, CMEs extracted by WICA are more close to their true values, and the standard deviation of the residual series filtered by WICA is lower than that by ICA.

5. Discussion

Too many abnormal stations may affect the correct judgement due to the quartile deviation. In this paper, abnormal stations were detected by using

Table 3. Precisions of σ_I , σ_W and ε_I , ε_W with the SNRs of 0.8:1, 1:1 and 1:1.2, respectively.

	SNR (0.8:1)				SNR (1:1)				SNR (1:1.2)			
	σ_I	σ_W	ε_I	ε_W	σ_I	σ_W	ε_I	ε_W	σ_I	σ_W	ε_I	ε_W
SXLQ	1.57	1.25	0.97	0.71	1.42	1.11	0.86	0.59	1.21	1.03	0.66	0.54
SDZB	1.45	1.29	0.89	0.62	1.14	1.17	0.82	0.51	1.09	1.02	0.58	0.44
LNYK	1.05	1.27	0.96	0.68	0.64	0.46	0.87	0.91	0.43	0.69	0.70	0.72
HEZJ	0.98	0.53	1.17	0.97	0.91	0.50	1.10	0.91	0.78	0.54	0.90	0.77
HELY	1.28	0.90	0.77	0.77	0.79	0.75	0.87	0.69	0.61	0.74	0.76	0.45
HELQ	1.54	0.79	0.97	1.25	1.49	0.36	0.85	1.21	1.11	0.64	0.68	0.70

Figure 10. σ_I and σ_W of simulation experiments repeated 50 times.

Figure 11. ε_I and ε_W of simulation experiments repeated 50 times.

Table 4. Abnormal stations with correlation coefficients of 0.3, 0.4, and 0.5.

Corr+IQR	Station
0.3+IQR	LNSY, JSLY, SXCZ, HAJY
0.4+IQR	TAIN, SXCZ, LNSY, JSLY, SDRC, LNDD, HECX, HAJY, SDLY, ZHNZ
0.5+IQR	TAIN, SXCZ, LNSY, JSLY, SDRC, LNDD, HECX, SDZB, SDQD, SDLY, JIXN, HAJY, BJFS, ZHNZ
IQR	TAIN, SXCZ, LNSY, JSLY, SDRC, LNDD, HECX, LNYK

the correlation coefficient combined with the quartile deviation (table 4). The correlation coefficients between different stations were calculated before filtering, and the stations with poor data similarity were removed. After the calculation by ICA, quartile deviation was used to further detect abnormal stations. Finally, a total of 10 abnormal stations were removed. From figure 2, it can be seen that these stations are mostly located at the edge of the analysis area (HAJY, JSLY, LNDD, LNSY, SDLY, SDRC, SXCZ, ZHNZ) or in the coastal areas (SDRC, JSLY). Therefore, the

Figure 12. Time series with temporally heterogeneous noise (10 mm offset per station).

Table 5. Precisions of σ_I , σ_W and ε_I , ε_W with temporally heterogeneous noise.

	SXLQ	SDZB	LNYK	HEZJ	HELY	HELQ
σ_I	1.20	1.21	0.68	1.35	0.79	1.28
σ_W	1.15	1.10	0.91	0.76	0.50	0.39
ε_I	0.84	0.74	0.85	0.83	0.76	0.86
ε_W	0.60	0.52	0.55	0.65	0.64	1.07

removal of these stations is conducive to the accurate extraction of CMEs.

In addition, we simulated the situation that a temporally heterogeneous noise occurred. With the CMEs in the N direction for the six stations of HELQ, HELY, HEZJ, LNYK, SDZB and SXLQ taken as known signals, MATLAB software was used to generate noise. The noises of 50, 100, 100, 150, 100, 50 epochs for these six stations, respectively, were deleted sequentially to generate a temporally heterogeneous noise sequence (figure 12). After filtering by ICA and WICA, the results (table 5) show that σ_W , ε_W are less than σ_I , ε_I respectively, for most stations, which is similar to the results from true data. Therefore, it was indicated that WICA can still extract CMEs effectively when a temporally heterogeneous noise occurs.

An attempt was made in this paper to define the CME of ICA spatial analysis based on the analysis of CME standard using PCA by Dong *et al.* (2006). Because the physical origin of CME is not clear, different definition standards and the number of principal components can still affect the precise removal of CME.

6. Conclusion

Previous studies have shown that GNSS continuous coordinate time series contains both Gaussian white noise and non-Gaussian coloured noise.

ICA can be used to effectively estimate the high-order information of the coordinate series and extract CMEs. In this paper, the RMSE of daily solution is taken as the weighting factor, and then the ICA approach is adopted for the WSF and the extraction of CMEs of a regional GNSS observation network. By analysing the observation data of 19 valid stations of CMONOC in North China, it is shown that the coordinate precisions of 13, 16 and 12 stations in the N, E and U directions after filtering by WICA are higher than those by the traditional ICA method. The coordinate time series precisions in the three directions are increased by 4.12%, 13.84% and 4.61%, respectively. The average correlation coefficient of the coordinate time series for each station after filtering was significantly reduced, and the results by WICA were reduced more prominently, which proves the spatial-temporal correlation of the CMEs. In this paper, two simulation experiments were designed to further verify the feasibility of WICA. The results obtained show that it is more effective to extract the CMEs by WICA with the SNRs of 0.8:1, 1:1 and 1.2:1. With the simulation experiments repeated 50 times, it is found that over 80% of σ_W and ε_W are smaller than σ_I and ε_I , respectively, for most stations. Therefore, with the RMSE of daily solution taken into account, the WICA approach can achieve better performance in regional filtering.

Acknowledgements

This study was sponsored by the Natural Science Foundation of China (Grant No. 41774041). It was also supported by the key scientific and technological project of Henan province (Grant No.

172102210277) and the guiding programme for national coal association science and technology research (Grant No. MTKJ2016-212). We thank the China National Seismological Bureau Service Platform of Digital Products for providing GNSS continuous time series of the CMONOC. The reviewers and editors are acknowledged for their constructive comments. We are grateful to Keke Xu and Xiaoqi Wang from Henan Polytechnic University for the helpful discussions.

References

- Amiri-Simkooei A R 2016 Non-negative least-squares variance component estimation with application to GPS time series; *J. Geod.* **90** 451–466.
- Blanz V and Vetter T 2002 Reconstructing the complete 3D shape of faces from partial information (Rekonstruktion der dreidimensionalen Form von Gesichtern aus partieller information); *Technische Informatik* **44** 295–302.
- Chen R, Tang B and Lv Z 2012 Ensemble empirical mode decomposition de-noising method based on correlation coefficients for vibration signal of rotor system; *J. Vib. Meas. Diagn.* **32**(4) 542–546.
- Comon P 1994 Independent component analysis, a new concept? *Signal. Process.* **36** 287–314.
- Cui Y 2017 Tow ways of understanding the coefficient of correlation; *Coll. Math.* **33**(3) 114–117.
- Dmitrieva K, Segall P and DeMets C 2015 Network-based estimation of time-dependent noise in GPS position time series; *J. Geod.* **89** 591–606.
- Dong D, Fang P, Bock Y, Webb F, Prawirodirdjo L, Kedar S and Jamason P 2006 Spatiotemporal filtering using principal component analysis and Karhunen-Loeve expansion approaches for regional GPS network analysis; *J. Geophys. Res.-Solid Earth* **111** 3405–3421.
- Gualandi A, Serpelloni E and Belardinelli M E 2016 Blind source separation problem in GPS time series; *J. Geod.* **90** 323–341.
- Li Q, You X, Yang S, Du R, Qiao X, Zou R and Wang Q 2012 A precise velocity field of tectonic deformation in China as inferred from intensive GPS observations; *Sci. China Earth Sci.* **55** 695–698.
- Li W, Shen Y and Li B 2015 Weighted spatiotemporal filtering using principal component analysis for analyzing regional GNSS position time series; *Acta Geod. Geophys.* **50** 419–436.
- Liu B, Dai W, Peng W and Meng X 2015 Spatiotemporal analysis of GPS time series in vertical direction using independent component analysis; *Earth Planets Space* **67** 189.
- MárquezAzúa B and DeMets C 2003 Crustal velocity field of Mexico from continuous GPS measurements, 1993 to June 2001: Implications for the neotectonics of Mexico; *J. Geophys. Res.-Solid Earth* **108** 149–169.
- Ming F, Yang Y, Zeng A and Zhao B 2017 Spatiotemporal filtering for regional GPS network in China using independent component analysis; *J. Geod.* **91** 419–440.
- Nikolaidis R 2002 Observation of geodetic and seismic deformation with the global positioning system; *Cancer Res.* **71** 714.
- Peng W, Dai W, Santerre R, Cai C and Kuang C 2017 GNSS vertical coordinate time series analysis using single-channel independent component analysis method; *Pure Appl. Geophys.* **174** 265–278.
- Shen Y, Li W, Xu G and Li B 2014 Spatiotemporal filtering of regional GNSS network's position time series with missing data using principal component analysis; *J. Geod.* **88** 1–12.
- Shum H 1994 Principal component analysis with missing data and its application to object modeling; In: *Conference on computer vision and pattern recognition, 1994. Proceedings CVPR'94*, IEEE Computer Society, *IEEE Trans. Pattern Anal. Mach. Intell.* **17** 854–867.
- Tian Y and Shen Z 2009 Research progress on the removal of non-structural noise in GPS coordinate time series; *Acta Seismol. Sin.* **31** 68–81.
- Tian Y and Shen Z 2011 Correlation weighted stacking filtering of common mode components in GPS observation network; *Acta Seismol. Sin.* **33** 198–208.
- van Dam T, Wahr J, Milly P, Shmakin A B, Blewitt G, Lavallée D and Larson K M 2001 Crustal displacements due to continental water loading; *Geophys. Res. Lett.* **28** 651–654.
- Wang H, Liu M, Cao J, Shen X and Zhang G 2011a Slip rates and seismic moment deficits on major active faults in mainland China; *J. Geophys. Res.-Solid Earth* **116** 1161–1172.
- Wang M, Li Q, Wang F, Zhang R, Wang Y, Shi H, Zhang P and Shen Z 2011b Far-field coseismic displacements associated with the 2011 Tohoku-oki earthquake in Japan observed by global positioning system; *Chin. Sci. Bull.* **56** 2419–2424.
- Wdowski S, Bock Y, Zhang J, Fang P and 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** 18057–18070.
- Williams S D P 2004 Error analysis of continuous GPS position time series; *J. Geophys. Res.-Solid Earth* **109** B03412.
- Wu Y and Huang L 2004 A new interpolation method in time series analyzing; *J. Geod. Geodyn.* **24** 43–47.
- Wu Y, Jiang Z, Wang M, Che S, Liao H, Li Q, Li P, Yang Y, Xiang H, Shao Z, Wang W, Wei W and Liu X 2013 Preliminary results pertaining to coseismic displacement and preseismic strain accumulation of the Lushan M 7.0 earthquake, as reflected by GPS surveying; *Chin. Sci. Bull.* **58** 3460–3466.
- Yang S, Li J and Wang Q 2008 The deformation pattern and fault rate in the Tianshan Mountains inferred from GPS observations; *Sci. China Ser.* **D51** 1064–1080.
- Zhao B, Huang Y, Zhang C, Wang W, Tan K and Du R 2015 Crustal deformation on the Chinese mainland during 1998–2014 based on GPS data; *Geod. Geodyn.* **67** 15.