Adaptive Denoising Approach for High-Rate GNSS Seismic Waveform Preservation: Application to the 2010 EI Mayor-Cucapah Earthquake and 2012 Brawley Seismic Swarm

YANYAN LI [1], **CAIJUN XU** [2], **CHUANFA CHEN** [1], **HAITAO YIN** [3], **LEI YI** [4,5], **AND XIAOXING HE** [6]

[1] Key Laboratory of Geomatics and Digital Technology of Shandong Province, Shandong University of Science and Technology, Qingdao 266590, China
[2] School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China
[3] Shandong Earthquake Agency, Jinan 250014, China
[4] Key Laboratory of Comprehensive and Highly Efficient Utilization of Salt Lake Resources, Qinghai Institute of Salt Lakes, Chinese Academy of Sciences, Xining 810008, China
[5] Qinghai Provincial Key Laboratory of Geology and Environment of Salt Lakes, Qinghai Institute of Salt Lakes, Chinese Academy of Sciences, Xining 810008, China
[6] School of Civil Engineering and Architecture, East China Jiaotong University, Nanchang 330013, China

Corresponding author: Yanyan Li (yylee@whu.edu.cn)

ABSTRACT The high noise levels of high-rate Global Navigation Satellite System (GNSS) solutions limit their seismological applications, including capturing earthquake-induced coseismic displacements. In this study, we developed a new adaptive denoising approach for high-rate GNSS observations to improve the precision of seismic displacements and preserve seismic waveforms in GNSS coseismic signals. The performance of the proposed method was evaluated using high-rate (5 Hz) GNSS data acquired from the moderate EI Mayor–Cucapah earthquake (M_w 7.2) on 4 April 2010 and the small Brawley seismic swarm (M_w 4.6–5.5) on 26 August 2012. The performance of the proposed method was compared with those of modified sidereal filtering, Stacking filtering, and MSF plus Stacking. The comparison showed that the proposed method is more precise than other widely used method. It can significantly remove high-frequency white noise and low-frequency colored noise caused by CME, multipath errors, and/or other unmodeled systematic errors in high-rate GNSS displacements. The results were also compared with collocated strong motion data (50 and 200 Hz). The high precision of the proposed method was mainly afforded by the high performance of complete ensemble empirical mode decomposition, which was used to decompose the GNSS signal into different frequency modes. However, the normalized autocorrelation function and correlation coefficients used to determine the noise-dominated high-frequency modes and the "wavelet-like" soft-threshold used for direct denoising of the noise-dominated high-frequency modes also contributed. Despite the high noise levels of GNSS solutions, especially regarding the vertical displacement components, some small-amplitude details, which are usually only detectable by seismic instruments, could be observed in the denoised displacements in this study. The results reported herein indicate that the proposed method significantly improves the precision and reliability of GNSS displacements and the effectiveness of seismic signal detection, which is particularly critical for the measurement of earthquake-induced coseismic displacements.

INDEX TERMS Coseismic displacements, empirical mode decomposition, high-rate GNSS.

The associate editor coordinating the review of this manuscript and approving it for publication was Venkata Ratnam Devanaboyina.

I. INTRODUCTION

Global Navigation Satellite Systems (GNSSs) play an important role in structure health monitoring [1], [2], seismic

wave detection, and moderate and large earthquake early warning (EEW) systems [3], [4]. The true dynamic displacements estimated by high-rate GNSS observations enable the generation of precise coseismic waveforms. Coseismic displacement can be used to quickly estimate the magnitudes of large earthquakes for purposes such as issuing tsunami warnings [5], [6], determining coseismic fields for earthquake source model inversion [7], [8], producing centroid moment tensor solutions [9], generating inverse earthquake focal mechanisms [10], [11], and modelling finite fault slips [12], [13].

With respect to the use of broadband seismometers and strong motions, GNSS observations have the significant advantages of the no-clipping, no-tilting, and no-integrating features [14]. Thus, GNSS data are suitable for use close to the source of a large ($>M_w$ 7) seismic event. This enables the acquisition of reliable displacement information for determination of the fault rupture process, identification of the focal mechanism, and EEW monitoring within minutes of earthquake initiation [15], [16]. For typical EEW systems [17], [18], the key issue is detection of the arrival time of the P-wave in order to predict the arrival time and intensity of the destructive S-wave and surface wave [19], [20]. However, the precision and reliability of high-rate GNSS solutions are strongly affected by systematic errors, such as common mode error (CME) caused by spatial correlations between GNSS stations [21], [22] and the multipath effects of the coordinate time series [23]–[25]. High-rate GNSS dynamic displacements are sufficiently low to resolve most of the surface wave spectrum of a moderate event within near-source range or a large event at teleseismic distance [26]. In addition, although high-rate GNSS observations can be directly used to measure static offsets and dynamic coseismic displacements, the latter measurements are not sufficiently accurate for identifying P-waves with millimeter-level or smaller amplitudes [27]. Furthermore, GNSS-based displacement waveforms are not sufficiently precise in the vertical direction, in which lies most of the energy of an arriving P-wave. This makes the detection of P-wave arrival more difficult [28]. Boosting signals in seismic data is especially crucial for seismic imaging [29], inversion [30], [31], and interpretation [32]. Therefore, to improve the precision and reliability of high-rate GNSS displacements and enable the detection of P-wave arrival, denoising algorithms are required to reduce or eliminate inherent noise and preserve seismic waveforms in measured signals.

Many studies have revealed significant nonlinear variations in the dynamic displacements measured by high-rate GNSSs [33]. Time-varying characteristics have also been observed in short-period geophysical motions such as the seismic waves generated by an earthquake. Because GNSS time series are essentially nonlinear and multiscale, it is often beneficial to have a method for decomposing signals into different temporal scales. A high-rate GNSS-PPP method combining a wavelet-package transformation (WPT) denoising and neural network prediction (NN) was proposed

in [34]. The method was proposed to denoise GNSS-derived vertical displacements to improve the dynamic behavior of structures. The results of experimental tests revealed that the WPT-NN model improved the GNSS-PPP displacements and can be used to accurately detect the dynamic behavior of engineering structures. A data-driven method referred to as complete ensemble empirical mode decomposition (CEEMD)-based multiscale multiway principal component analysis (C-MSMPCA) was proposed in [35] for denoising daily GNSS position time series. C-MSMPCA combines the ability of MSMPCA to extract the linear relationship between signals with that of CEEMD to adaptively decompose nonlinear and time-varying signals into various intrinsic mode functions (IMFs). Daily and high-rate GNSS position results indicate that C-MSMPCA outperforms other methods commonly used for denoising non-coseismic GNSS time series [36]. However, the performance of C-MSMPCA for high-rate GNSS coseismic displacements during an earthquake is unknown.

Seismic waves are generally non-linear signals originating from a non-stationary process [37]. Reference [38] proposed empirical mode decomposition (EMD), which is an innovative signal analysis method for adaptively decomposing signals into IMFs and residue from high-frequency to low-frequency depending on the inherent characteristic of the signals. Reference [39] reported a numerical study on uniformly distributed white noise, finding that EMD is a dyadic filter that can be used to effectively separate white noise into IMF components. Reference [40] also numerically investigated the statistical characteristics of fractional Gaussian noise (fGn) [41] and found that EMD could be used as a data-driven wavelet-like filter bank. Reference [42] reported findings that agree with the results presented in [39] for the case of white noise. The EMD procedure is adaptive, data-driven, and very suitable for analyzing non-stationary and non-linear data in various fields [43]–[48].

The high-frequency IMF signals decomposed by EMD mainly contain stochastic white noise and the low-frequency IMFs usually contain the useful signal. The first few high-frequency IMFs, considered to be composed of only noise, can be directly removed. The denoised signals can then be reconstructed using the remaining IMFs and the residue [49]. In the denoising of white noise by abandoning the first IMFs, some useful signals will also be lost, particularly those containing sharper information, such as seismic signals. Moreover, all data analysis tools are susceptible to noise corruption. Reference [50] reported the noise corruption of EMD and analyzed the characteristics of IMFs. They found that there are "transition" IMFs, which capture both white noise and signal in the same mode. Three categories of IMFs decomposed by EMD of a noisy signal can be identified [50]: 1) noisy IMFs, which are noise-dominated, 2) transition IMFs, which contain both white noise and signal, and 3) monochromatic IMFs, which are signal-dominated. Therefore, as in the CEEMD-based denoising method proposed in [51] for solving the mode mixing problem and

achieving residue noise toleration, the key is determining the cut-off point for the transition between noise-dominated and signal-dominated IMFs.

To improve precision and reliability and preserve precise seismic waveforms in high-rate GNSS seismic displacements, the present study implemented a new improved CEEMD-MPCA denoising approach based on a correlation coefficient, a normalized autocorrelation function, and "wavelet-like" threshold denoising. The performance of the proposed method was validated using high-rate (5 Hz) GNSS data collected during the moderate EI Mayor-Cucapah earthquake (M_w 7.2) on 4 April, 2010 and the small Brawley seismic swarm (M_w 4.6–5.5) on 26 August, 2012. The results were compared with collocated strong motion data (50 and 200 Hz).

II. THEORY

Assume that n-dimensional noisy seismic signals can be modeled as

$$X(t) = x(t) + c(t) + w(t), \quad t = 1, 2, \ldots, m \quad (1)$$

where n and m are the numbers of GNSS stations and epochs, respectively; $X(t)$ and $x(t)$ (with dimensions $m \times n$) denote the original and denoised GNSS displacements, respectively; $c(t)$ represents a low-frequency colored noise; and $w(t)$ represents a Gaussian white noise. Given noisy displacements $X(t)$, a denoising method aims to recover $x(t)$ by removing $c(t)$ and $w(t)$.

In the remainder of this section, the statistical properties of the normalized autocorrelation function of white noise will first be reviewed. Secondly, the correlation coefficient between the original signal and each IMF decomposed by CEEMD and the wavelet-like threshold denoising method will be introduced. Finally, the proposed denoising method will be described, along with its application to high-rate GNSS seismic displacements.

A. NORMALIZED AUTOCORRELATION FUNCTION OF EACH IMF

The autocorrelation function of white noise $w(t)$ is an average measure of its time-domain characteristics and reflects the correlation degree of the white noise at two different times. The autocorrelation function of white noise can be expressed as

$$R_w(t_1, t_2) = E[w(t_1)w(t_2)] \quad (2)$$

where the mean value of $w(t)$ is 0.

To accurately represent the correlation degree of white noise at different times, the normalized autocorrelation function is expressed as

$$\rho_w(\tau) = \frac{R_w(\tau)}{R_w(0)} \quad (3)$$

where $\tau = t_2 - t_1$ represents the time lag. According to (3), the normalized autocorrelation function of white noise $n(t)$ is as shown in Fig. 1.
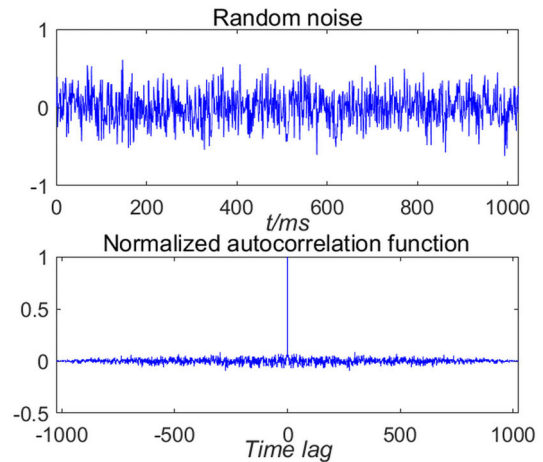

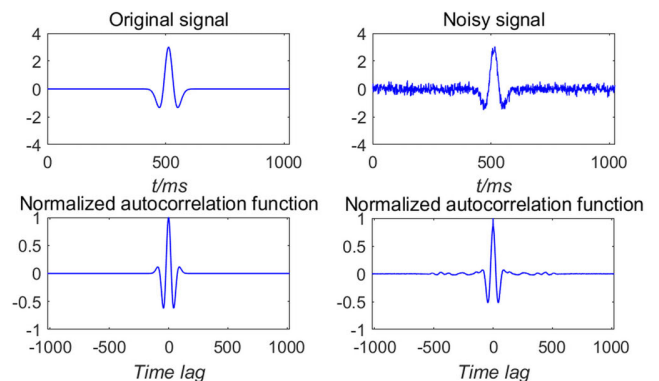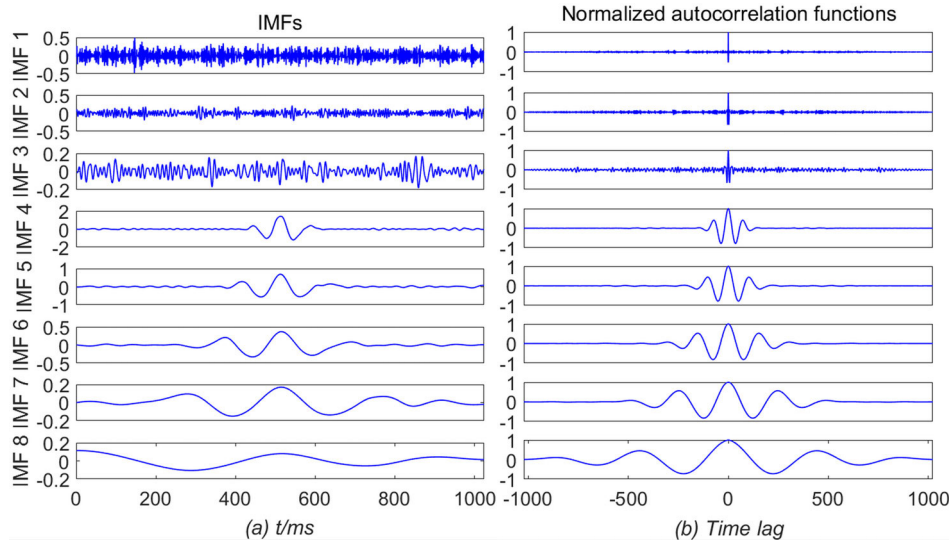**FIGURE 1.** White noise and its normalized autocorrelation function.

FIGURE 2. Original signal, noisy signal, and corresponding normalized autocorrelation functions.

It can be observed from Fig. 1 that the correlation of white noise is statistically characterized by randomness and is very weak at any given time. The maximum autocorrelation value at time zero is 1, while the values of the normalized autocorrelation function at any other time decrease rapidly to 0.

The original and noisy signals and their normalized autocorrelation functions are shown in Fig. 2. The original signal is a 100-Hz Ricker wavelet with a peak value of 1, sampling frequency of 1 ms, and data length of 1024. The frequency of the effective signal of a Ricker wavelet is mainly within the range 0–0.03 Hz. Gaussian white noise was then added to the original signal. It can be seen from Fig. 2 that the normalized autocorrelation function of the general signal has its maximum value at the zero point. Due to the relativity of the signal itself, the value of the normalized autocorrelation function does not rapidly decay at other times but varies with the time difference. This variation is obviously determined by the difference relative to the normalized autocorrelation function of the white noise.

The noisy Ricker wavelet was first decomposed into various IMFs by CEEMD [51]. As can be observed from Fig. 3(a), CEEMD actually involves adaptive decomposition,

FIGURE 3. (a) IMFs obtained by CEEMD of the noisy Ricker wavelet and (b) the corresponding normalized autocorrelation functions.

TABLE 1. Correlation coefficients between the noisy Ricker signal and the IMFs.

IMF	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8
r	0.31	0.21	0.19	0.81	0.80	0.54	0.23	0.08

producing high-frequency to low-frequency IMFs depending on the inherent signal characteristics. The corresponding normalized autocorrelation function of each IMF is shown in Fig. 3(b). According to the definition of EMD, an IMF is any function that has the properties of forced symmetry and adaptability of the envelopes separately defined by the local maxima and minima. From Fig. 3(b), it can be seen that, the first few high-frequency IMF components mainly contain white noise. The characteristics of this white noise are destroyed, resulting in it no longer being white noise in the real sense. However, the statistical characteristics of the white noise are still approximately retained. This means the normalized autocorrelation function's maximum at time is zero, and the values at other times rapidly decrease. Regarding the other low-frequency IMF components, which occur as general signals, their normalized autocorrelation functions maximize at the zero point, and the values at other times do not rapidly decay but vary with the time difference.

Fig. 3(b) shows that the third IMF is mainly composed of noise, while the fourth is mainly composed of signals. However, the signals and noise in IMF3 and IMF4 could not be effectively separated. IMF1–IMF3 are noise-dominated IMFs. IMF5–IMF8 are signal-dominated IMFs. IMF4 is thus the cut-off point component for IMF transition. The cut-off point between the noise-dominated IMF and signal- dominated IMF could be estimated based on the statistical characteristics of the normalized autocorrelation function of the white noise.

B. CORRELATION COEFFICIENT BETWEEN ORIGINAL SIGNAL AND EACH IMF DECOMPOSED BY CEEMD

From Figs. 1–3, in the case of the first few high-frequency IMF components, it was determined that there was a gradual decrease in the primary function of the white noise in each IMF, and a gradual increase in that of the function of the signal. Therefore, the correlation coefficient between the original signal and each IMF can be used to determine the cut-off point between the noise-dominated IMF and signal-dominated IMF. The correlation coefficient between the original signal and each IMF can be expressed as

$$r_{i,j}(X_j, c_{i,j}) = \frac{\sum\limits_{m}(X_i - \bar{X}_i)(c_{i,j} - \bar{c}_{i,j})}{\sqrt{\sum\limits_{m}(X_i - \bar{X}_i)^2 \sum\limits_{m}(c_{i,j} - \bar{c}_{i,j})^2}} \quad (4)$$

where X_j denotes the first j signals, $c_{i,j}$ denotes the first iIMF components of the j signals, and m denotes the length of the signal.

According to the correlation coefficient between the noisy signal and IMFs of the noisy signal, the cut-off point for the transition IMF can be defined as follows: Beginning with the first IMF component, the IMF corresponding to the first extremum of the correlation coefficients can be determined, and the latter IMF can be determined as the cut-off transition IMF. The correlation coefficients between the noisy Ricker signal and the IMFs decomposed by CEEMD are listed in Table 1. By comparing the normalized autocorrelation functions of the IMF components (Fig.3) with the

correlation coefficients between the IMF components and the noisy signal (Table 1), IMF4 was confirmed to be the cut-off point component. The cut-off point for IMF transition determined by the correlation coefficient and the normalized autocorrelation function could thus be compared and verified, respectively.

C. WAVELET-LIKE' SOFT-THRESHOLD DENOISING

For a given orthogonal wavelet basis, the discrete wavelet transform (DWT) decomposes the original noisy signal $X(t)$ into wavelet coefficients:

$$c(t) = WX(t) \qquad (5)$$

where matrix W is an orthogonal $(J+1) \times (J+1)$ matrix, $c(t)$ contains J detailed components and an approximate component, and J is the chosen decomposition level.

A simple wavelet de-noising algorithm involves three steps: 1) wavelet coefficient decomposition, which includes selection of the wavelet basis and the decomposition level J, 2) threshold quantization processing of the wavelet detail coefficients, including determination of the threshold method and rule to reduce the noise in each level based on the signal-to-noise ratio (SNR), and 3) reconstruction of the denoised original signal using the inverse wavelet transform obtained from the threshold detail coefficients and approximation coefficients.

The fundamental principle of soft wavelet thresholding is to assign a value of zero to all wavelet coefficients lower than a threshold related to the noise level and to appropriately shrink the other wavelet coefficients by an amount equal to the threshold [52]. The soft thresholding operator is defined by

$$\hat{c}_j = \begin{cases} \text{sgn}\,(c_j)\,(|c_j| - \delta_j) & |c_j| > \delta_j \\ 0 & |c_j| \le \delta_j \end{cases} \qquad (6)$$

where *sgn* denotes the sign function, c_j denotes the wavelet coefficients at level j, and δ_j denotes the threshold at level j, determined by any of several available methods [53]. For removal of the added Gaussian white noise, a universal threshold δ_j was proposed by Donoho and Johnstone [54], given by

$$\delta_j = \sigma_j \sqrt{2\ln(m)} \qquad (7)$$

$$\sigma_j = MAD_j/0.6745 \qquad (8)$$

where σ_j is the noise level of the jth wavelet coefficient, and MAD_j denotes the absolute median deviation of the jth wavelet coefficient.

Instead of thresholding the wavelet coefficient, a wavelet-like soft threshold denoising method can be used to directly threshold the noise-dominated IMFs as follows [55]:

$$C'_j = \begin{cases} \text{sgn}\,(C_j)\,(|C_j| - \beta_j) & |C_j| > \beta_j \\ 0 & |C_j| \le \beta_j \end{cases} \qquad (9)$$

where C_j denotes the jth noise + signal mixed IMF component, and β_j is the threshold value of the jth noise + signal

mixed IMF component, given by

$$\beta_j = \sigma_j \sqrt{2\ln(m)} = \frac{Median\,(abs\,(C_j))}{0.6745} \times \sqrt{2\ln(m)} \qquad (10)$$

The selection of the wavelet basis type and decomposition level in the application of DWT is difficult and requires experience. This significantly contributes to the achieved denoising effect of the wavelet denoising algorithm. Without the need for selection of the wavelet basis type and decomposition level required for DWT, CEEMD can be used to adaptively decompose the signals into various IMFs from high-frequency to low-frequency based on the inherent characteristic of the signals. The first few noise + signal mixed IMFs can then be directly shrunk by the 'wavelet-like' soft threshold method through the use of (9) to remove the added Gaussian white noise. Especially in the case of a seismic wave signal, the first few high-frequency IMFs contain sharper information, which can be effectively retained by applying the proposed CEEMD-based 'wavelet-like' soft thresholding method.

D. PROPOSED ADAPTIVE THRESHOLD DENOISING APPROACH

Based on the above methods, we developed an adaptive CEEMD-MPCA denoising method. This enables the use of the normalized autocorrelation function and correlation coefficient to effectively remove the high-frequency white noise and low frequency colored noise in a GNSS coseismic signal.

In the proposed method, CEEMD is first used to decompose the signal into various IMFs. The cut-off point s is determined using the normalized autocorrelation function and correlation coefficient. Then, the IMFs are classified as "noise + signal" IMFs and "signal-dominated" IMFs. The former are noisy IMFs and transition IMFs, while the latter are almost monochromatic. The noise + signal IMF modes are subsequently directly denoised using the wavelet-like soft threshold. Finally, the processed noise + signal IMFs and signal-dominated IMFs are grouped based on their frequency bands (FBs) and processed by multiway PCA (MPCA) to obtain the denoised signal. The detailed procedure of the proposed method is as follows:

(i) Each variable of $X(t)$ is decomposed by CEEMD [53] into K different time scales, $\{C_1, C_2, \ldots, C_K\}$.
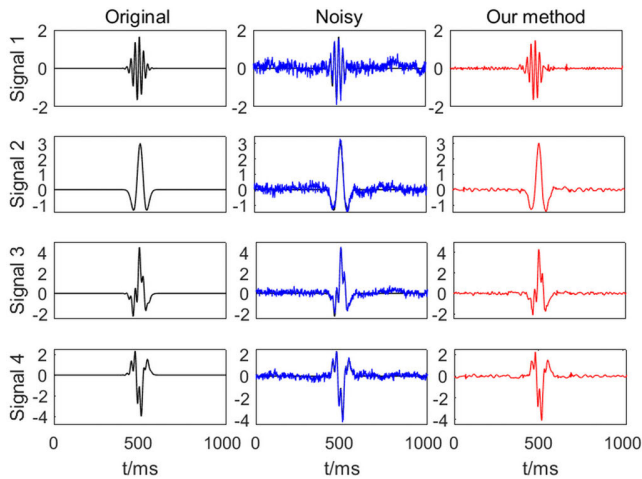
(ii) The normalized autocorrelation function of each IMF and the correlation coefficient between the original signal and each IMF (see above) is used to determine the cut-off point s. The IMFs are then classified as noise + signal or signal-dominated.

(iii) The first noise + signal IMFs are directly denoised using the wavelet-like soft threshold to obtain the denoised "IMFs" $\{C'_1, C'_2, \ldots, C'_s\}$ $(s < K)$.

(iv) The denoised "IMFs" and signal-dominated IMFs are grouped based on their frequency bands (FBs), which were estimated by S-transform [56]. Each group of modes with the same FB is then denoised using MPCA.

TABLE 2. Correlation coefficients between signal 2 and the IMFs decomposed by CEEMD.

IMF	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8
r	0.20	0.19	0.16	0.81	0.83	0.53	0.26	0.03

FIGURE 4. Simulation results for the original and noisy versions of the Damped-Sinusoids and Ricker signals and their sum and difference. The denoised signals are obtained by the proposed method.

(v) A new matrix is reconstructed from the processed IMFs.

(vi) The new matrix is processed by the MPCA method to obtain the denoised signal.

III. SYNTHETIC EXAMPLE

A four-dimensional model with $m = 1024$ was used as a synthetic dataset to verify the proposed method for seismic signals. In the synthetic dataset test, a 1024×4 original data matrix X was simulated with two known true signals: "Damped-Sinusoids" and "Ricker". The Damped-Sinusoids signal is given by

$$S(t) = 0.8 \times rect(t) \times \sin(2\pi f t) \times (a/\pi)^{1/8} \times \exp(-at^2/2)$$

(11)

where $f = 40$ Hz. If $|t| \leqslant 0.15$, rect $(t) = 1$; otherwise, rect$(t) = 0$. The second signal was simulated by a 100-Hz Ricker wavelet with a peak value of 3. The sampling frequencies of the first two signals were 1 ms, and their data lengths were 1024. The third and fourth signals were defined by the sum and difference of the first two signals (Fig. 4), respectively. Gaussian white noise and colored noise with a strong spatiotemporal correlation were added to the original signals. The Gaussian white noise was produced by an fGn model with $H = 0.5$, and the time-correlated colored noise was produced by an fGn model [42] with a Hurst exponent H of 0.9. The detailed procedure of the noise production is available in [35].

The second Ricker wavelet was used as an example to analyze the performance of the proposed method. The noisy

TABLE 3. Root mean square errors (RMSEs) and signal-to-noise ratios (SNRs) for the different methods used to denoise the simulated noisy signals.

Signal	Noisy signal		Proposed method	
	RMSE	SNR	RMSE	SNR
Signal 1	0.21	3.9	0.07	11.7
Signal 2	0.21	7.8	0.08	16.4
Signal 3	0.22	8.4	0.09	16.0
Signal 4	0.21	8.9	0.10	15.4

signal 2 was first decomposed using CEEMD (Fig. 5a). Fig. 5b shows the corresponding normalized autocorrelation functions, and Table 2 lists the correlation coefficients between signal 2 and the IMFs decomposed by CEEMD. From Table 2 and Fig. 5b, it can be seen that the higher-frequency IMFs (IMF1–IMF3) are mainly composed of noise, while signals are the main components of IMF4. However, the signals and noise in IMF3 and IMF4 could not be effectively separated. IMF4 was thus the cut-off point component. IMF1–IMF4 were therefore classified as noise + signal modes and directly denoised using the wavelet-like soft threshold. Finally, the processed high-frequency and low-frequency IMFs were grouped based on their FBs and handled using MPCA to obtain the denoised signal.

The original simulated signal and the denoised signals obtained by the proposed method are shown in Fig. 6. The root mean square errors (RMSEs) and signal-to-noise ratios (SNRs) of the four noisy signals are listed in Table 3. From Fig. 6, it can be seen that the four signals are well recovered.

The proposed method can be used to remove much of the noise inherent in the sample dataset and preserve useful signals. Compared with the noisy signals, the proposed method reduced the RMSEs of the four signals by 66.7%, 61.9%, 59.1%, and 52.4%, respectively, and improved the SNRs by 66.7%, 52.4%, 45.8%, and 42.2%, respectively. Overall, Fig. 6 and Table 3 illustrate the satisfactory performance of the proposed method.

IV. HIGH-RATE GNSS SEISMIC DATASETS AND DATA PROCESSING

The performance of the proposed method was tested on two high-rate (5 Hz) GNSS seismic datasets, which were compared with strong motion accelerometer data with high-rate sampling at 50-200 Hz and high precision. The first GNSS seismic dataset was for the M_w 7.2 moderate El Mayor–Cucapah earthquake. It occurred at 22:40:42 UTC on 4 April 2010 in northern Baja, California, approximately 40 km south of the United States-Mexico border (epicenter location: 32.286°N, 115.295°W,
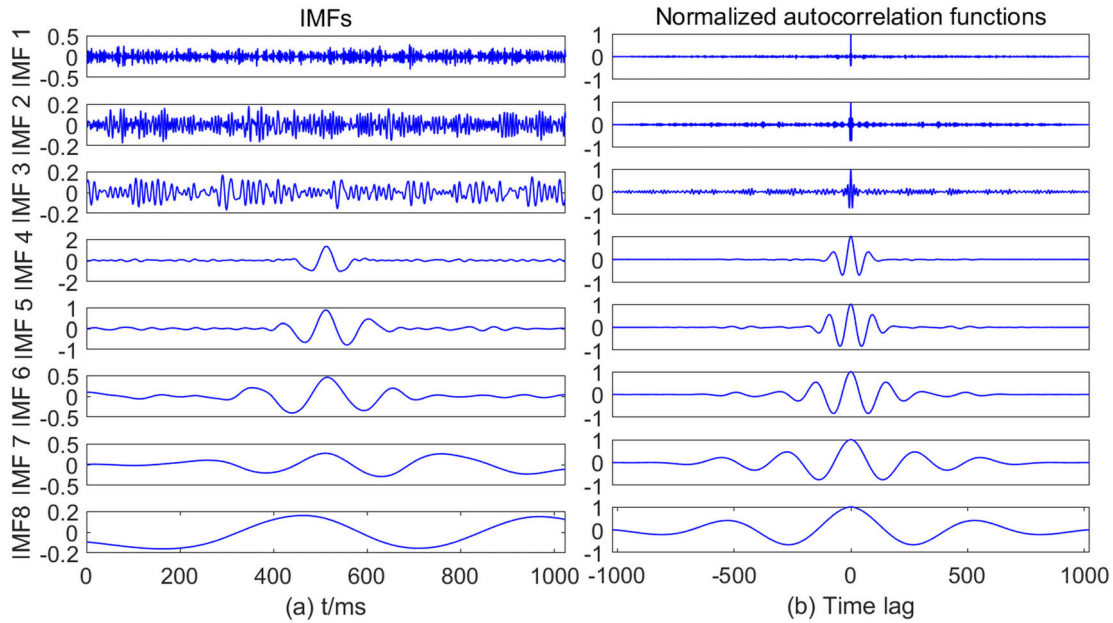
FIGURE 5. (a) IMFs obtained by CEEMD for signal 2 and (b) the corresponding normalized autocorrelation functions.
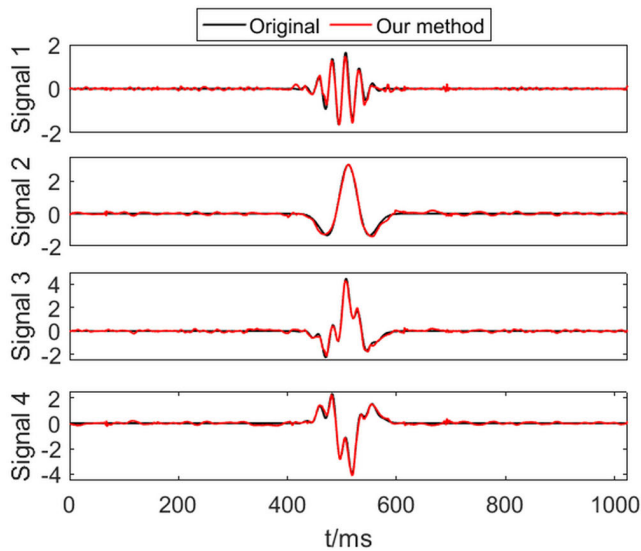

FIGURE 6. Original and denoised signals. First to fourth rows: Damped-Sinusoids and Ricker signals and their sum and difference.

FIGURE 7. Distribution of the 45 GNSS stations of the California Real-time GNSS Network (CRTN) and 18 strong motion stations in southern California. The green triangles represent the high-rate GNSS stations, the red circles represent the strong motion stations, and the red star indicates the location of the M_w 7.2 El Mayor–Cucapah earthquake.

focal depth: 10 km) (Fig. 7). The high-rate (5 Hz) GNSS data obtained through high-rate GNSSs in southern California were available from UNAVCO (ftp://data-out.unavco.org/pub/highrate/). Very high-rate (200 and 50 Hz) strong motion data were obtained from CESMD (http://www.strongmotioncenter.org/cgi-bin/CESMD/archive .pl). The GNSS stations used to collect the data were generally not collocated with the strong motion stations in the observation network. Fig. 7 shows the epicenter of the earthquake, the distribution of the 45 GNSS stations in the California Real-time GNSS Network (CRTN), and the 18 strong motion stations in southern California used to monitor the
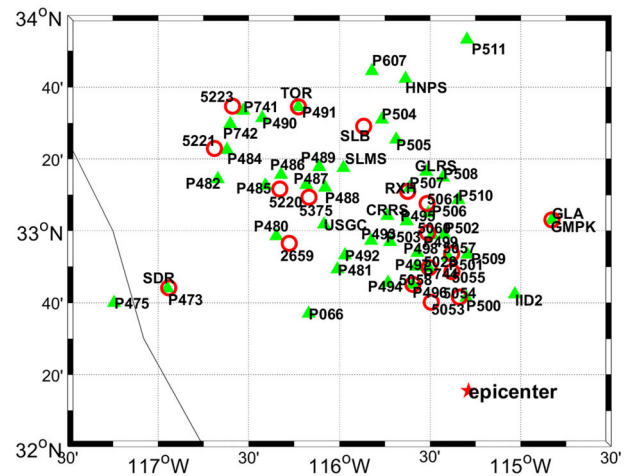
El Mayor–Cucapah earthquake. Table 4 gives the station names, locations, and distances to the epicenter, as well as the separation distances of five collocated stations.

The second GNSS seismic dataset was for the small Brawley seismic swarm that started at approximately 15:30 UTC on 26 August 2012, comprised of six events with M_w < 2.0 and three M_w 2.5 events. All the events occurred within a few minutes. The four largest events had magnitudes of 4.6–5.5. The earthquakes occurred in a northeast-striking fault zone located approximately 6 km north of the northeast end of the Imperial fault (http://www.scsn.org/2012Brawley. html). This area has a history of seismic swarms including an

TABLE 4. Collocated high-rate GNSS and strong motion (SM) stations of the Mw 7.2 El Mayor–Cucapah earthquake.

Station (GNSS/SM)	Latitude (˚ N)	Longitude (˚ W)	Distance to Epicenter (km)	Separation (km)
P496	32.751	115.596	61.9	0.178
5058	32.752	115.595		
P744	32.829	115.508	66.7	0.320
5028	32.829	115.505		
GMPK	33.051	114.827	98.0	0.034
GLA	33.051	114.827		
P491	33.575	116.227	170.6	0.084
TOR	33.575	116.226		
P473	32.734	116.950	164.6	0.745
SDR	32.736	116.942		

TABLE 5. Locations and magnitudes of the analyzed events of the 26 August 2012 Brawley seismic swarm [A].

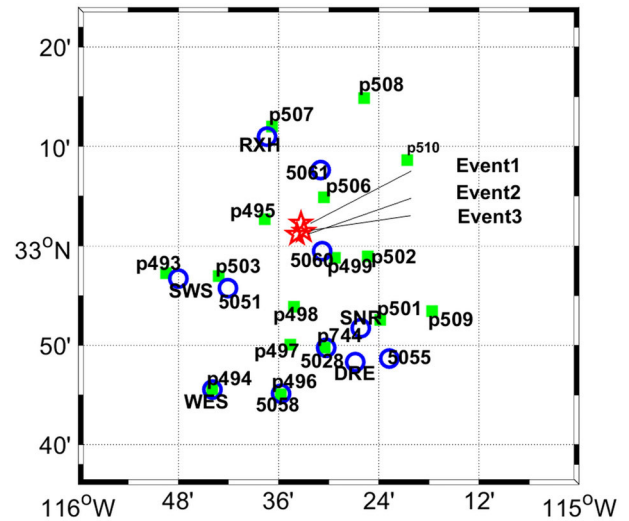| Event (GNSS/SM)	M_w	Latitude (˚ N)	Longitude (˚ W)	Depth (km)	Time (UTC)
1	4.6	33.016	115.536	4.8	19:20:04
2	5.4	33.017	115.554	8.3	19:31:23
2-doublet	4.9	33.014	115.559	4.4	19:33:01
3	5.5	33.019	115.540	8.3	20:57:58

[A]Source: https://www.globalcmt.org/CMTsearch.html.

TABLE 6. Collocated high-rate GNSS and strong motion (SM) stations of the 26 August 2012 Brawley seismic swarm.

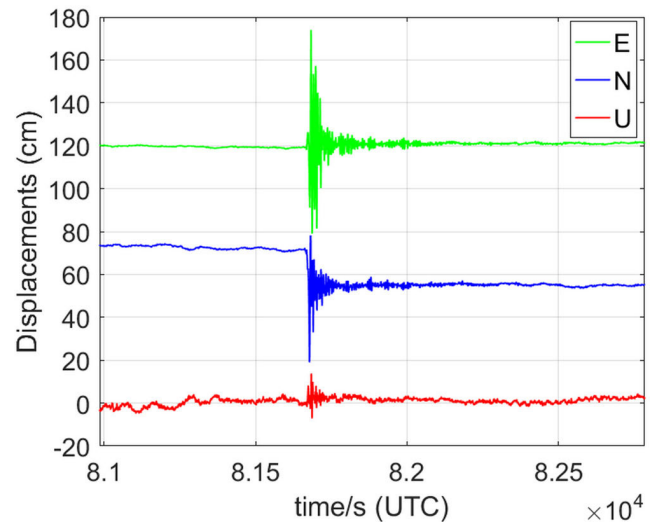| Station (GNSS/SM)	Latitude (˚ N)	Longitude (˚ W)	Distance to Epicenter (m) (Event 1)	Separation (km)
P744	32.829	115.508	20.0	0.320
5028	32.829	115.505		
P496	32.751	115.596	30.7	0.178
5058	32.752	115.595		
P494	32.760	115.732	34.0	0.745
WES	32.759	116.732		

earthquake with a maximum magnitude of 5.1 in 2005, specifically at Obsidian Buttes [57], and another in June 2008 in the same general location as the 2012 event [58]. The three largest events among those considered herein (Table 5) occurred at 19:20:04 UTC (M_w = 4.6; Event 1), 19:31:23 UTC (M_w = 5.4; Event 2), 19:33:01 UTC (M_w = 4.9; Event 2-doublet), and 20:57:58 UTC (M_w = 5.5; Event 3) (USGS Earthquake Database, https://earthquake.usgs.gov/earthquakes/browse/).

The high-rate (5 Hz) GNSS data and strong motion data of the second seismic swarm dataset used in this study were also obtained from UNAVCO and CESMD, respectively. Fig. 8 shows the epicenter location of the three largest events and the distribution of the 16 GNSS stations and 11 strong motion stations used to monitor the Brawley swarm. Table 6 lists the station names, locations, and distances to the epicenter, as well as the separation distances of three collocated stations.

Both considered GNSS datasets were analyzed using the Positioning and Navigation Data Analyst (PANDA) software developed by the GNSS Research Center, Wuhan University [59]. The post-processing kinematic Precise Point

FIGURE 8. Distribution of the 16 GNSS stations of the California Real-time GNSS Network (CRTN) and 11 strong motion stations in southern California. The green squares represent the high-rate GNSS stations, the blue circles represent the strong motion stations, and the red stars indicate the locations of the three largest events of the 26 August 2012 Brawley seismic swarm.
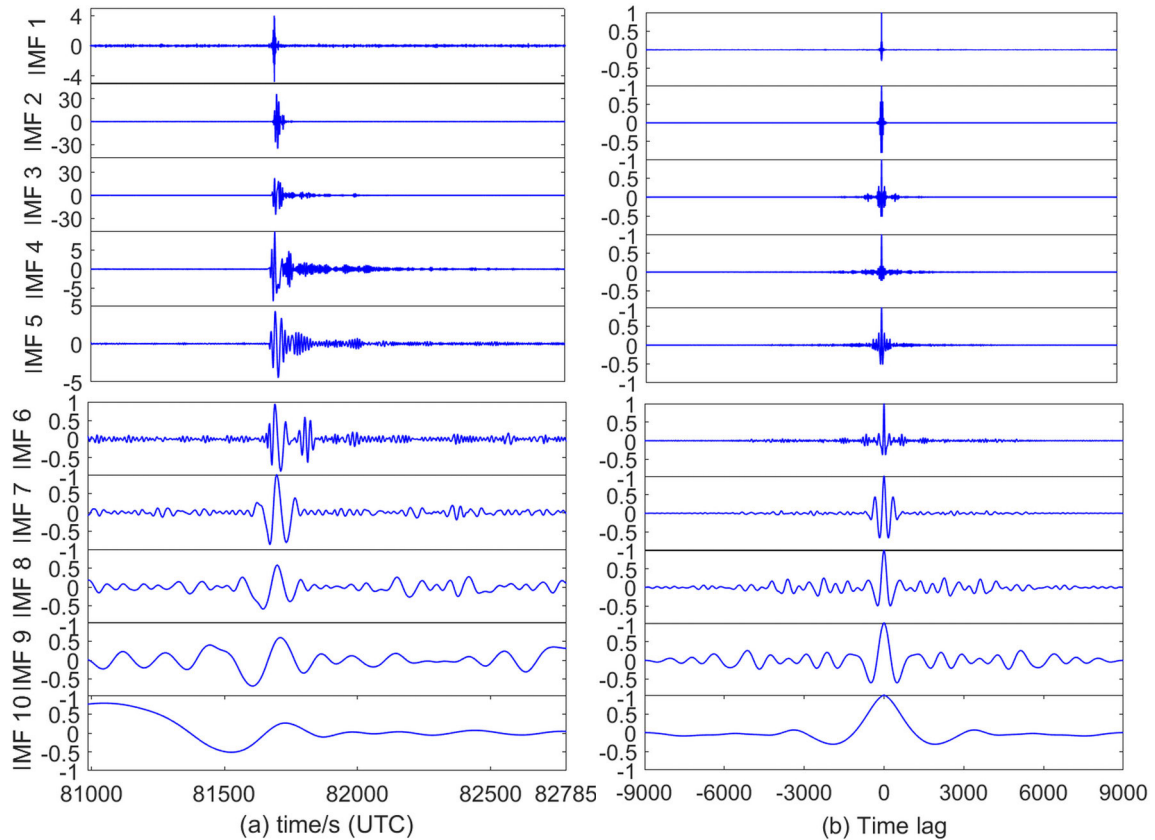

FIGURE 9. East (green), north (blue), and vertical (red) components of the GNSS dynamic seismic displacements at station P496 during the 4 April 2010 M_w 7.2 El Mayor–Cucapah earthquake.

Positioning (PPP) mode of the software was employed. The procedure of GNSS data processing using PPP is described in [36]. The time series for the east, north, and vertical coordinate components were combined in a 3-D data matrix $\bar{X}(m, n, p)$, where m, n, and p represent the numbers of epochs, stations, and coordinate components. The 3-D data matrix $\bar{X}(m, n, p)$ was transformed into a new 2-D data matrix $X(m, q)$, where $q = n \times p$.

V. RESULTS AND DISCUSSION

We first analyze the results of the application of the proposed method to the 5-Hz GNSS seismic time series, compared with those of modified sidereal filtering (MSF) [23], Stacking filtering [60] and MSF plus Stacking. Then, we further

FIGURE 10. (a) IMFs obtained by CEEMD for the *E* components of the GNSS seismic displacements at station P496 and (b) the corresponding normalized autocorrelation functions.

TABLE 7. Correlation coefficients between the E components of the GNSS seismic displacements at station P496 and the IMFs obtained by CEEMD.

IMF	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	IMF10
r	0.11	0.75	0.71	0.36	0.12	0.03	0.02	0.05	0.07	0.00

analyze the performance of the proposed method by comparison with the collocated strong motion data (200 and 50 Hz) for the moderate 2010 EI Mayor–Cucapah earthquake. Due to the relatively small magnitudes of the earthquakes and the excellent distribution of the nearby GNSS stations, we also analyze the results of the proposed method for the 5-Hz GNSS seismic time series and compare the GNSS data with the collocated strong motion data (200 Hz) collected during the small 2012 Brawley seismic swarm.

A. MODERATE EARTHQUAKE EXAMPLE: 2010 EI MAYOR–CUCAPAH EARTHQUAKE

As examples for analysis of the performance of the proposed method, we firstly take the seismic displacements of the station "P496" as an example to analyze the performance of the proposed method compared with the widely used methods. Then, we further analyze the performance of the propose method and compare with the seismic displacements of the collocated stations P496 (GNSS)/5058 (strong motion), P744/5028, and GMPK/GLA. The three pairs of

the collocated stations are located approximately 60, 67, and 98 km from the earthquake epicenter and approximately 0.18, 0.32 and 0.03 km from each other, respectively.

1) GNSS SEISMIC DISPLACEMENT ANALYSIS
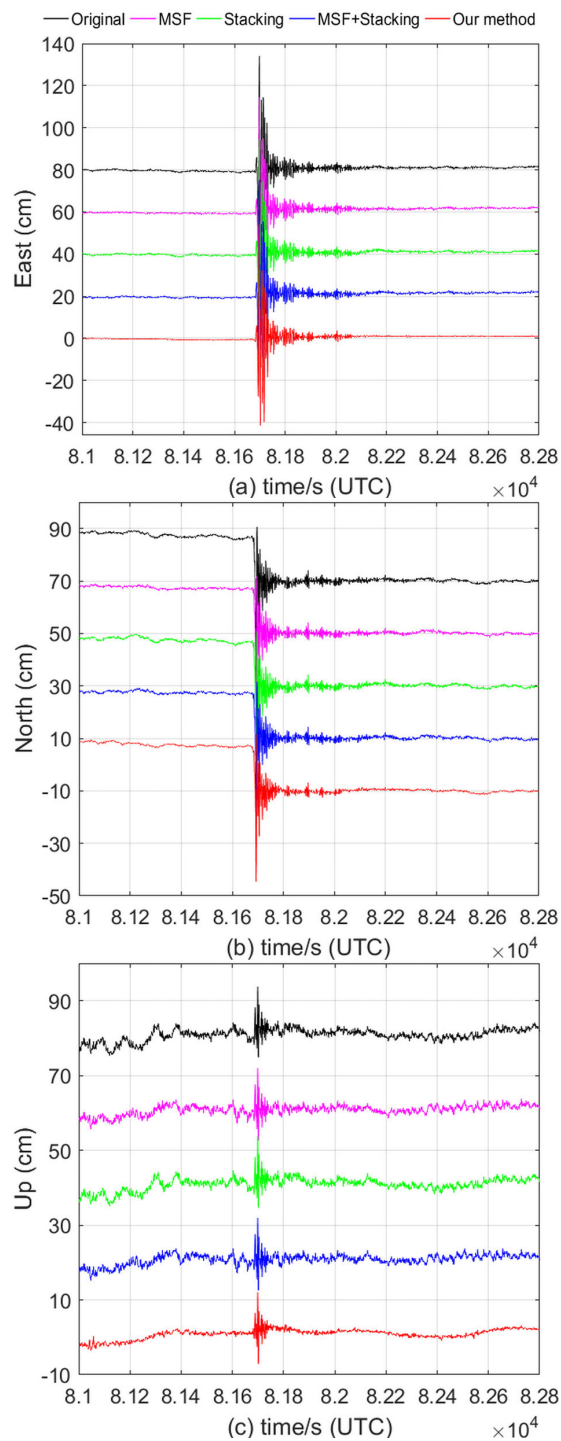
Fig. 9 shows the east (E), north (N), and vertical (U) components of the 5-Hz GNSS dynamic seismic displacements of station P496 during the EI Mayor-Cucapah earthquake. It can be seen from the figure that the GNSS seismic displacements estimated by instantaneous GNSS positioning in the PPP mode contain various errors. To obtain pure GNSS seismic signals, it was necessary to first estimate the noise in the dynamic seismic displacements. The 5-Hz GNSS seismic displacements at station P496 were therefore denoised using the proposed method.

First, the dynamic seismic displacements were decomposed into different IMFs by CEEMD based on the inherent characteristics of the seismic signals. Considering the IMFs in the *E* components at station P496 (Fig. 10a) as an example, the seismic signal is composed of 10 IMFs and a

residue. Because of the large amplitude of the seismic signal, IMF1 still contains a part of the signal and not only pure white noise components. In addition, it can be observed that the noise plays a dominant role in IMF1, while the signal plays the dominant role in the other IMFs. It can also be seen from Fig. 10(b) and Table 7 that the noise plays dominant roles in IMF1 and IMF2 and that IMF3 is the cut-off IMF component. Therefore, during the denoising, IMF1–IMF3, which were considered as the noise + signal modes, were directly denoised using the wavelet-like soft threshold. The processed IMF1–IMF3 and the remaining IMFs were then grouped based on their FBs and handled by MPCA to obtain the denoised signal.

To analyze the performance of the proposed method, the results of the proposed method are compared with those of MSF [23], Stacking filtering and MSF plus Stacking. For MSF, the original displacements for the two days (DOY 092 and 093) before the El Mayor–Cucapah earthquake were timed shifted by the orbital repeat period (23 h, 55 min, 54 s). Then they were point-wised averaged and smoothed with 3-level wavelet denoising to form a profile on the earthquake day (DOY 094). MSF can remove station-specific errors caused by repeatable errors such as multipath, resulting in a filtered coordinate time series. For Stacking filtering, positions from several nearest stations outside the region of active deformation were combined to form a profile of the common-mode error. To apply Stacking filtering, the profile was subtracted from positions of the station of interest. Stacking can remove systematic errors caused by un-modeled errors affecting all stations simultaneously [3]. For the majority of results in this study, because the available nearest stations with high-rate sampling outside the region of active deformation is limited, two IGS stations BREW and AMC2 were chosen for Stacking spatial filtering. The BREW and AMC2 stations with sampling at 1 Hz are located at (48.1333° N, 119.6833° W) and (38.8031° N, 104.5246° W), respectively. The two IGS data at 1Hz used for Stacking were obtained from IGS (ftp://cddis.gsfc.nasa.gov/pub/gps/products/). For MSF plus Stacking, after MSF, Stacking filtering was further used for the filtered displacements by MSF.

The two stations used for Stacking spatial filtering were at 1 Hz. Thus, before analyzing the results of the original and denoised seismic displacements by MSF, Stacking and MSF plus Stacking, the 5Hz original GNSS data were resampled to 1 Hz. Fig. 11 presents examples of the original and denoised GNSS seismic displacements with resampling rate 1 Hz for the three components of the ''P496'' station. From Fig. 11(a-b), it can be seen that, after denoising the horizontal components of the seismic displacements at station P496, in two periods before (80985–81642 s) and after (82001–82785 s) the earthquake, the denoised displacements using the proposed method were smoother than the original displacements and those of the MSF, Stacking and MSF plus Stacking methods. Moreover, the seismic wave signals during the occurrence of the earthquake (81642–82000 s) were well preserved. The results show that the proposed method and

FIGURE 11. Original and GNSS denoised (a) east, (b) north, and (c) vertical components of the seismic displacements at station P496 with resample rate 1Hz for the 4 April 2010 El Mayor–Cucapah earthquake (UTC time). The displacements of the original GNSS, MSF, Stacking, and MSF plus Stacking are represented by the black, pink, green, blue, and red solid lines, respectively. For applying MSF, the original seismic displacements on the day of the earthquake subtracted the MSF filter profile. For Stacking, the distant stations over 1200km from P496. The original seismic displacements subtracted the Stacking filter profile form P496. For MSF plus Stacking, the filtered displacements after MSF subtracted the Stacking filter profile from P496, resulting in the final GNSS displacements. Note: For better distinguishing the performance of each method, the seismic displacements are plotted with respect to nominal offsets.

other methods can eliminate the low-frequency noise while effectively retaining the seismic wave signals.

Owing to a number of factors such as satellite constellation distribution, receiver clock error, and tropospheric delay error, the precision of the vertical components of the GNSS solutions was usually lower than that of the horizontal components by a factor of 2–3 [61]. It can also be observed from Fig. 11(c) that the vertical components of the seismic displacements contain centimeter-level fluctuations and high-frequency random fluctuations. This indicates that the original signals contained low-frequency colored noise and high-frequency random noise. Similarly, after denoising the vertical components of the seismic displacements, during the two periods before (80985–81642 s) and after (82001–82785 s) the earthquake, the denoised seismic displacements using the proposed method were significantly smoother than those of the original, MSF, Stacking and MSF plus Stacking methods. Meanwhile, the seismic wave signal was also well preserved during the earthquake occurrence period (81642–82000 s). The results also show that, compared with other methods, the proposed method can significantly eliminate more low-frequency colored noise and high-frequency white noise while more effectively retaining the seismic wave signals.
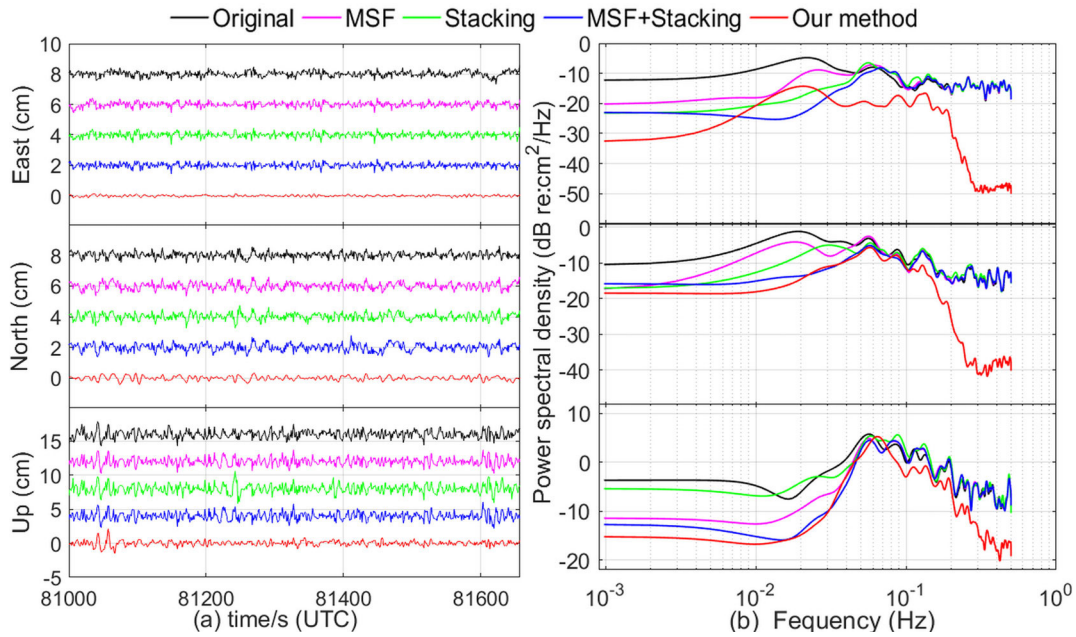
To accurately analyze the performance of the proposed method, Fig. 12(a) shows examples of the residual displacement time series of the original, MSF, Stacking, MSF plus Stacking at P496 station, as observed in the 10 min before the EI Mayor–Cucapah earthquake began. Fig. 12(b) shows the power spectral density (PSD) of the original displacement time series and those of other methods. The PSD of each time series was computed using the Welch method [62]. Each spectrum can be divided into two parts: high-frequency white noise over 0.1–0.5 Hz (2–10 s periods) and low-frequency colored noise over 0.001–0.1 Hz (10–1000 s periods) [3]. By comparing the PSDs of the original time series with the PSDs of the MSF, Stacking, MSF plus Stacking and the proposed method, these results indicated that the high-rate GNSS data suffer from high-frequency white noise and low-frequency colored noise. MSF, Stacking and MSF plus Stacking captured the low-frequency variability in the range of 80–1000 seconds. Because their results are still subject to high-frequency noise, these commonly used methods seem to have limited ability to deal with high-frequency white noise in the range of 2–10s. This may be due to the fact that they do not adopt the advantage of wavelet transform to decrease the influence of high-frequency white noise [36]. The proposed method can significantly remove high-frequency white noise and low-frequency colored noise, caused by CME, multipath errors, and/or other unmodeled systematic errors in the high-rate GNSS displacements. In comparison, for removing the high- and low-frequency noises, the proposed method is more effective and accurate than the aforementioned methods. This is mainly attributed to the combination of the advantages of wavelet-like soft-threshold denoising method, CEEMD and MSMPCA.

Table 8 lists the average standard deviations of the original and denoised residual GNSS time series (10 min pre-event displacement time series) of the 45 GNSS stations. The average standard deviations of the original residual GNSS time series were determined to be 1.79, 2.62, and 7.28 mm for the east, north, and vertical components, respectively. The results indicate that the average original standard deviations on the horizontal component are much less than those on the vertical component, and those on the east component are less than those on the north component. This may be caused by incomplete integer-cycle phase ambiguity resolution in this network. This agrees with the results of the CRTN in south California from [36]. The average standard deviations obtained by MSF plus Stacking are both less than those by MSF and Stacking. MSF reduced the nonrandom error caused by repeatable errors, such as multipath error. After MSF, some nonwhite systematic errors remained, which are caused by unmodeled common mode error affecting all stations simultaneously. Thus, MSF plus Stacking is superior to MSF and Stacking. The average RMSEs obtained by the proposed method were 0.70, 0.91, and 4.19 mm, respectively, representing RMSE reductions of 60.9%, 65.3%, and 42.4%, respectively. The results of the proposed method reflect submillimeter-level accuracy for the horizontal component, and millimeter-level accuracy for the vertical component. The proposed method is superior to MSF, Stacking, and MSF plus Stacking. The high accuracy of the proposed method can be mainly attributed to the following: 1) the high performance of CEEMD for decomposing GNSS signals into IMFs of differing frequencies, 2) the use of the normalized autocorrelation function and correlation coefficients to determine the noise-dominated high-frequency IMFs, and 3) the use of the wavelet-like soft threshold to directly denoise the noise-dominated high-frequency IMFs.

2) COMPARISON OF GNSS DATA AND STRONG MOTION DATA

It is generally accepted that seismic seismometers-derived displacements with high-rates (50-200 Hz) have extremely high precision and are reliable [3], [26]. For further performance analysis of the proposed method, the seismic displacements of the collocated stations are analyzed. GNSS stations are generally not collocated with strong motion stations in an observation network. The collocation details of the stations for the present study are presented in Table 4. The coseismic displacements of the collocated stations P496 (GNSS)/5058 (strong motion), P744/5028 and GMPK/GLA, were compared to analyze the performance of the proposed method. Fig. 13 compares the east, north, and vertical components of the original (black dots), GNSS denoised (blue lines), and seismic displacements (red lines) on the collocated stations P496 (5 Hz GNSS) and 5058 (200 Hz strong motion).

Fig. 13(a) shows the east component of the seismic shaking over its entire period. The original and seismic displacements can be observed to agree very well, almost coinciding. This indicates that the GNSS positioning solutions have very high

FIGURE 12. (a) Original and GNSS signals denoised for all three components of the residual time series of the original and other methods at station P496 with resampling rate 1 Hz, as observed in the 10 min before the 4 April 2010 EI Mayor–Cucapah earthquake began. The residual time series of the original GNSS, MSF, Stacking, and MSF plus Stacking are represented by the black, pink, green, blue, and red solid lines, respectively. (b) The corresponding power spectral density for all three components of the residual time series of the original and other methods at station P496, as observed in the 10 min before the 4 April 2010 EI Mayor–Cucapah earthquake began. Note: For better distinguishing the performance of each method, the residual time series are plotted with respect to nominal offsets.

TABLE 8. Average standard deviations of the original and denoised residual time series (10-min pre-event time series) with resampling rate 1 Hz of the 45 GNSS stations for the 4 April 2010 EI Mayor–Cucapah earthquake (units: mm).

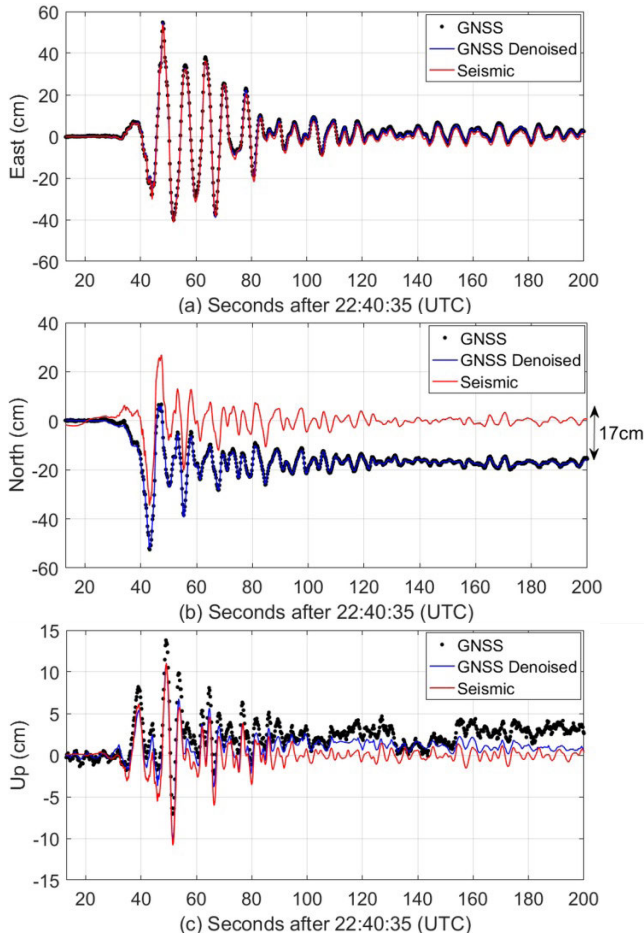| Direction	Original	MSF	Stacking	MSF+Stacking	Our method
East	1.79	1.58	1.43	1.26	0.70
North	2.62	2.26	2.14	1.98	0.91
Vertical	7.28	6.32	6.85	5.72	4.19

precision on the east component. The denoised and seismic displacements similarly agree very well, indicating that the proposed method effectively retains the seismic signal.

As can be further observed from Fig. 13(b), the original and seismic displacements also exhibit a high degree of similarity between their north components. The major difference exists over approximately the first 25 s, where a permanent coseismic offset of 17 cm can be seen. Because the location of the GNSS receiver antenna appears to have a permanent offset [63] and a seismic measurement instrument uses gravity as the datum, the permanent coseismic offset in seismic displacements during an earthquake is lost. The good agreement between the present original and denoised displacements indicates a low noise level in the north component, and that the proposed method effectively retains the seismic signal.

Expectedly, the vertical components of the original seismic displacements had the highest noise level, due to the effects of the satellite constellation distribution, receiver clock error, and tropospheric delay error. However, the denoised seismic displacements were smoother. Compared with the original

results, the denoised and seismic displacements exhibited a higher degree of similarity between their vertical dynamic components. This indicates that the proposed method produced a more precise seismic waveform. Despite the high noise level in the vertical components of the GNSS solutions for moderate earthquakes, the small amplitudes of the coseismic signal could not be detected in the original displacements. However, they were detected in the denoised vertical displacements obtained by the propose method (e.g. between 13 and 30 s in Fig. 13c). This represents a significant improvement for cases in which the seismic signal can only be detected in strong events with significant shaking. Hence, the proposed method not only significantly eliminates the noise in the GNSS seismic displacements, but also effectively retains the seismic signal.
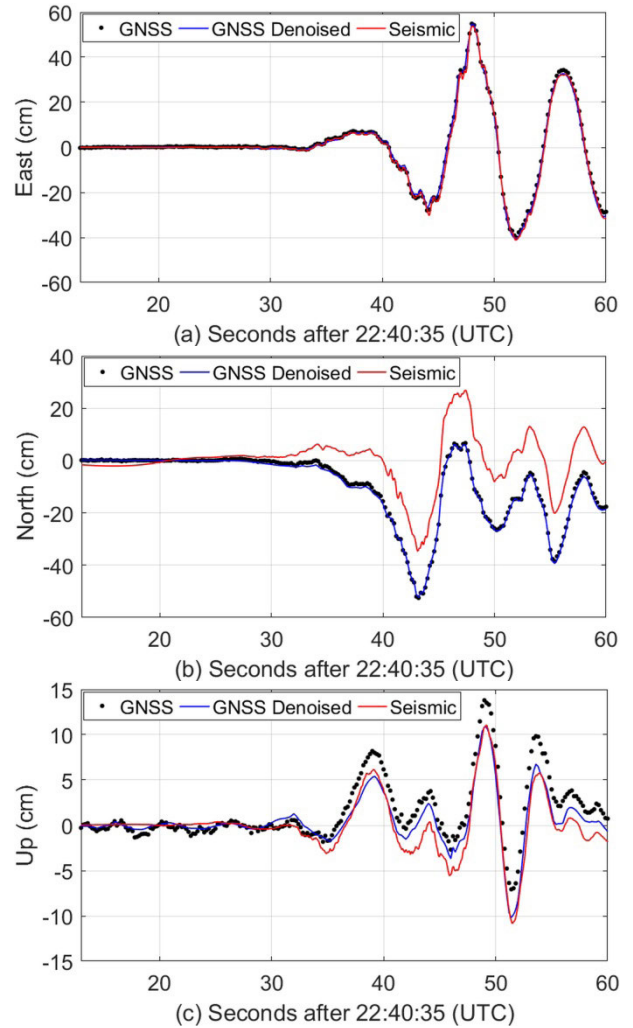
For clarity, Fig. 14 shows magnifications of the coseismic displacement time series over the first 47 s. It can be seen that, in terms of the peak displacements and the long-term stability of the horizontal components (Fig. 14a-b), the denoised displacements are in good agreement with the original solutions. The precision of the displacement time series was

FIGURE 13. Comparison of the original, GNSS denoised, and seismic displacements on the collocated stations P496 (5-Hz GNSS) and 5058 (200-Hz strong motion) during the M_w 7.2 El Mayor–Cucapah earthquake on 4 April 2010. Panels (a)–(c) respectively show the east, north, and vertical components of the seismic shaking over its entire period. The GNSS, GNSS denoised, and 200-Hz strong motion displacements are represented by the black dotted, blue solid, and red solid lines, respectively.

FIGURE 14. Magnifications of the first 47 s of the (a) east, (b) north, and (c) vertical components of the coseismic displacements of stations P496 (GNSS) and 5058 (strong motion) during the El Mayor–Cucapah earthquake on 4 April 2010.

also improved by the proposed method. It is well known that the precision of the vertical component of a GNSS positioning solution is significantly lower than those of the horizontal components. Therefore, compared with the vertical components of the original displacements (Fig. 14c), the denoised and seismic displacements exhibit a higher degree of similarity in their dynamic components.
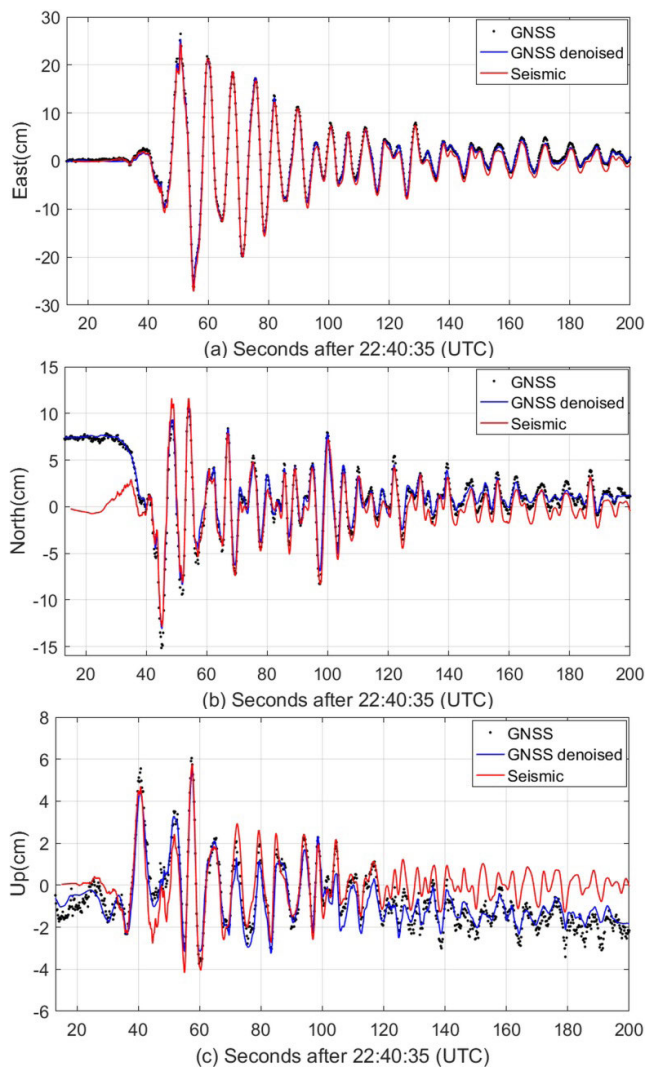
Figs. 15 and 16 compare the east, north, and vertical components of the original (black dots), GNSS denoised (blue lines), and seismic displacements (red lines) on the collocated stations P744/5028 (200 Hz strong motion) and GMPK /GLA (50 Hz strong motion). It can be observed from Fig. 15(a-b) that, the original and seismic displacements on the collocated stations P744/5028 also exhibit a high degree of similarity between their horizontal components. The GNSS denoised and seismic displacements similarly agree very well. For the vertical component (Fig. 15c), the original and seismic displacements only exhibit a certain degree of similarity. However, the denoised GNSS and seismic displacements

exhibit a higher degree of similarity, indicating that the proposed method produces a more precise seismic waveform. It can be seen from Fig. 16(a-b) that, the original and seismic waveforms on the collocated stations GMPK /GLA have a low degree of similarity. However, the denoised GNSS and seismic displacements exhibit a higher degree of similarity in their dynamic components. Due to the expected high noise level in the vertical component, the coseismic displacements of the original and denoised GNSS do not agree with that of the seismic displacements, except for the time segments during 60–80s (Fig. 16c).

Compared with the amplitudes of seismic signals of the collocated stations P496/5058 (Fig. 13), and P744/5028 (Fig. 15), those on the collocated stations GMPK/GLA (Fig. 16) are obviously smaller. Based on the analysis of the seismic displacements of the three pairs, the visual performance of the proposed method on the El Mayor earthquake (M_w 7.2) depends on the amplitudes of seismic signals,
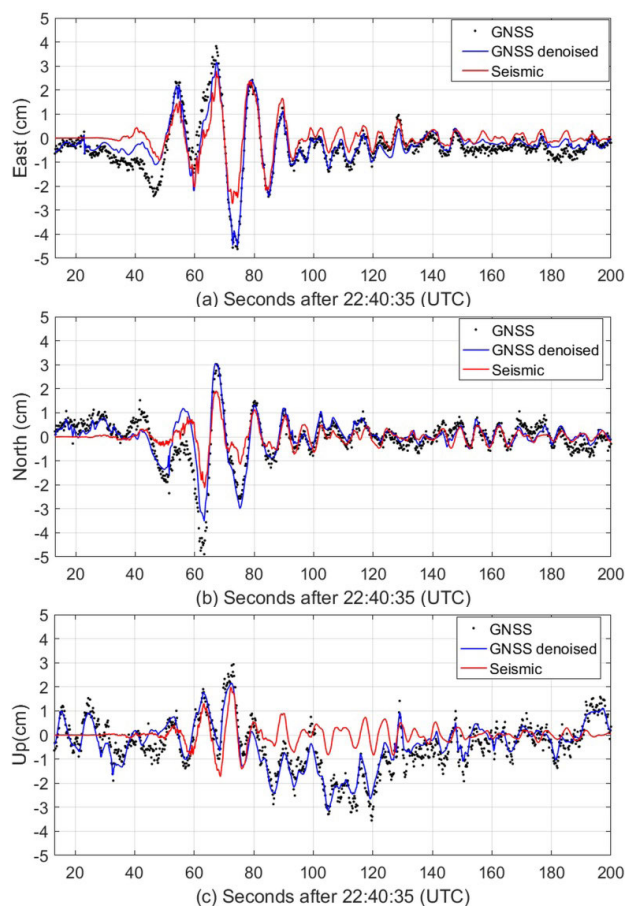
FIGURE 15. Comparison of the original, GNSS denoised, and seismic displacements on the collocated stations P744 (5-Hz GNSS) and 5028 (200-Hz strong motion) during the M_w 7.2 EI Mayor–Cucapah earthquake on 4 April 2010. Panels (a)–(c) respectively show the east, north, and vertical components of the seismic shaking over its entire period. The GNSS, GNSS denoised, and 200-Hz strong motion displacements are represented by the black dotted, blue solid and red solid lines, respectively.

FIGURE 16. Comparison of the original, GNSS denoised, and seismic displacements on the collocated stations GMPK (5-Hz GNSS) and GLA (50-Hz strong motion) during the M_w 7.2 EI Mayor–Cucapah earthquake on 4 April 2010. Panels (a)–(c) respectively show the east, north, and vertical components of the seismic shaking over its entire period. The GNSS, GNSS denoised, and 50-Hz strong motion displacements are represented by the black dotted, blue solid and red solid lines, respectively.

which are related to the earthquake magnitude and epicentral distance. As the amplitudes of seismic signals become smaller and the performance of our approach becomes more remarkable. Therefore, the proposed method can significantly improve the precision of GNSS solutions and enable effective detection of seismic signals for a moderate earthquake.

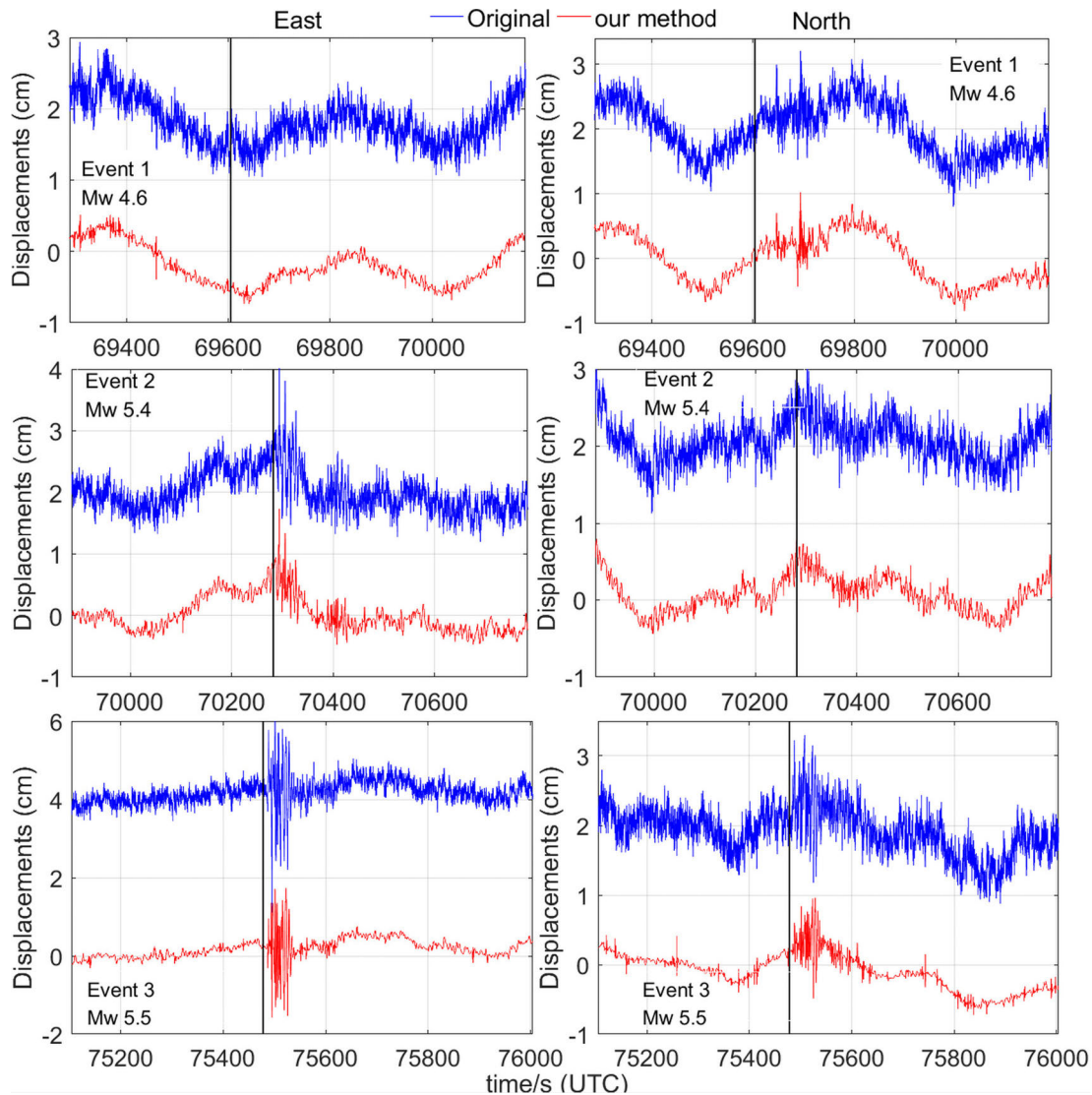## B. MODERATE EARTHQUAKE EXAMPLE: 2010 EI MAYOR–CUCAPAH EARTHQUAKE

Due to the relatively small magnitudes of the earthquakes and excellent distribution of the nearby GNSS stations, the 2012 Brawley seismic swarm produced unique seismic waveforms. They were estimated from the high-rate GNSS data with collocated strong motion data for further

validation of the performance of the proposed denoising method. As example, we considered the seismic displacements of the GNSS station P744 and the collocated seismic station 5028. The collocated stations were located approximately 20 km from the epicenter of Event 1 and approximately 0.32 km from each other.

Fig. 17 shows the horizontal components of the original and denoised seismic displacements at station P744 during the three largest events. It can be seen that the horizontal components of the seismic displacement time series for the three events have obvious centimeter-level fluctuations and high-frequency random fluctuations. This indicates that the original displacements contain low-frequency colored noise and high-frequency random noise. The coseismic displacements of Event 1 (M_w 4.6) do not have any significant horizontal (east and north) component. In the case of Event 2 (M_w 5.4), the coseismic displacements only have significant east components, while those of Event 3 (M_w 5.5) have significant east and north components. Expectedly, no significant vertical components of the coseismic displacements were
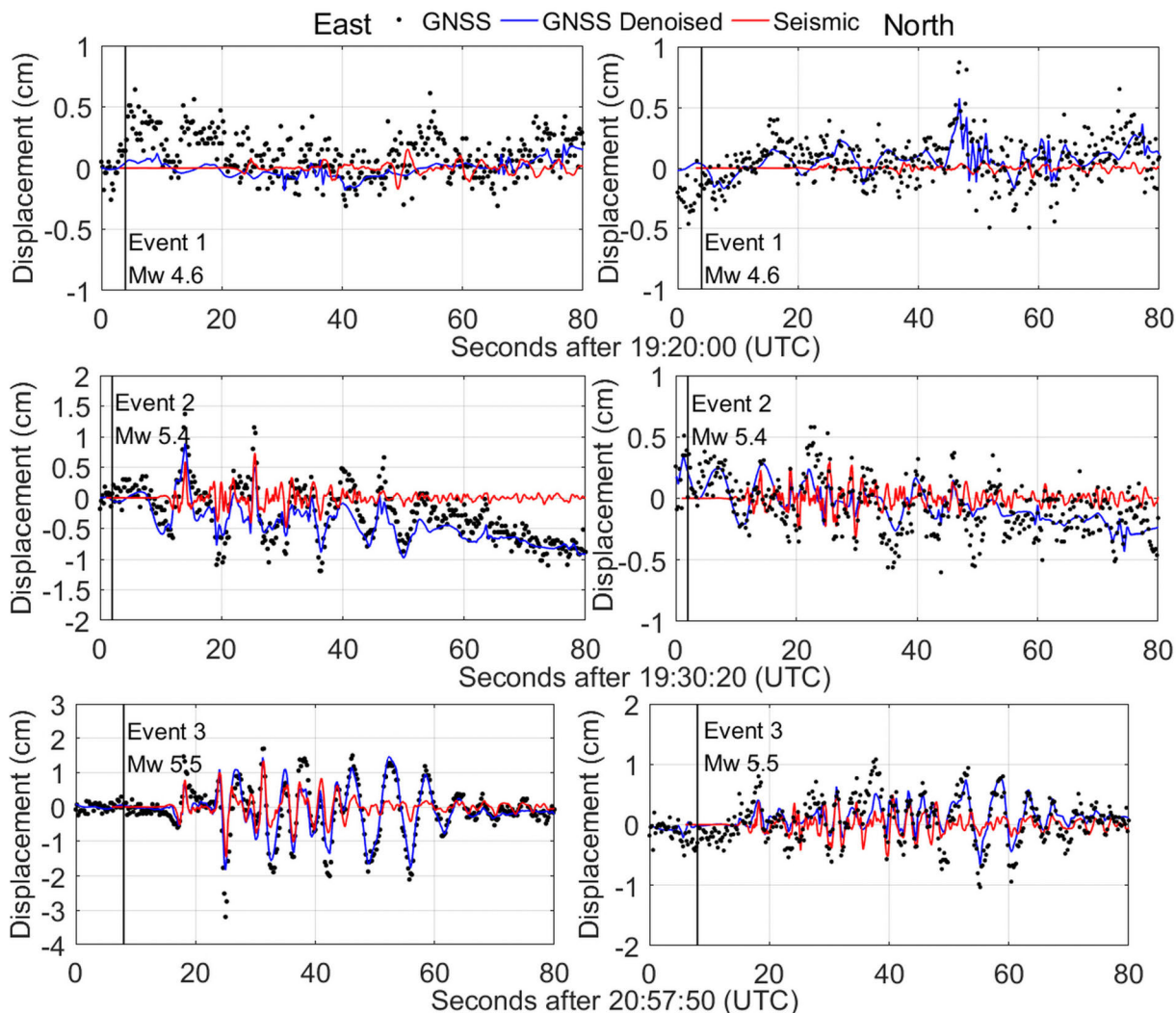
FIGURE 17. Original and GNSS denoised horizontal components of the seismic displacements at station P744 during the three largest events of the Brawley Swarm on 26 August 2012. It should be noted that Event 2 involved only the seismic displacements of the first earthquake ($M_w = 5.4$). The displacements of the original and our method are represented by the blue and red solid lines, respectively. The vertical grey lines indicate the origin time of the corresponding earthquake.

observed in any of the three events, including at the stations closest to the epicenters. After denoising using the proposed method, the displacements for the three events were smoother and the observed coseismic signals were well preserved, especially the horizontal components of Event 3. The present denoising results show that the proposed method significantly eliminates high-frequency white noise and low-frequency colored noise while effectively retaining the seismic wave signals.

For further performance analysis of the proposed method, the seismic displacements of collocated stations P744 and 5028 were compared. Fig. 18 compares the horizontal components of the original, denoised, and seismic displacements at these collocated stations during the Brawley seismic swarm on 26 August 2012. It can be seen from the figure that for

Event 1 ($M_w = 4.6$), the horizontal components of the original and denoised coseismic displacements do not agree with those of the seismic seismometer-derived displacements. This indicates that, even after denoising, no significant horizontal component of the coseismic displacements could be found. Similarly, the horizontal components of the original displacements during Event 2 ($M_w = 5.4$) do not match those of the seismic displacements, and there are also clear static offsets in the east and north components (0.8 and 0.3 cm, respectively) after the earthquake. However, denoising using the proposed method produces some similarities between the denoised coseismic displacements and the seismic displacements over the first 13–45 s, especially in the east component. This shows that the proposed method not only significantly eliminates noise from the GNSS seismic dis-

FIGURE 18. Comparison of the horizontal components of the original, GNSS denoised, and seismic displacements at collocated stations P744 (GNSS) and 5028 (strong motion) during the three largest events of the Brawley seismic swarm on 26 August 2012. The GNSS, GNSS denoised, and 200-Hz strong-motion displacements are represented by the black dotted, blue solid, and red solid lines, respectively. The vertical black lines indicate the origin time of the corresponding earthquake. It should be noted that Event 2 involved only the seismic displacements of the first earthquake ($M_W = 5.4$).

placements but also effectively distinguishes the seismic signal. In the case of Event 3 ($M_w = 5.5$), the horizontal components of the original displacements do not match those of the seismic displacements. However, denoising using the proposed method produces significant similarity, especially during the first 8–45 s. This shows that the proposed method affords a more precise seismic waveform.

Despite the high noise level in the horizontal components of the GNSS solutions for the small earthquakes, the small amplitudes of the coseismic signal are not detectable in the original displacements. However, they could be detected from the denoised displacements obtained by the proposed method (e.g. between 8 and 15 s in Event 3 in Fig. 18). Compared with the original GNSS displacements, this represents a significant improvement for cases in which the seismic signal can only be detected for strong events with significant shaking.

The results indicate that the proposed method significantly improves the precision of the GNSS solutions and effectively preserves the seismic wave signals.

Table 9 gives the average standard deviations of the original and denoised displacements (10-min pre-event time series) of the 16 GNSS stations for the three largest earthquakes during the Brawley seismic swarm on 26 August 2012. It can be seen that, for the three earthquake events, the average standard deviations of the horizontal components of the original and denoised high-rate GNSS displacement time series are smaller than those of the vertical components. As also indicated in Table 9, the averages standard deviations of the original GNSS solutions for the three earthquake events are 1.60, 1.57, and 4.49 mm for the east, north, and vertical components, respectively, while the corresponding values for the denoising results are 0.60, 0.61,

TABLE 9. Average standard deviations of the original and GNSS denoised displacements (10-min pre-event time series) of the 16 GNSS stations at 5 Hz for the three largest earthquakes during the Brawley swarm on 26 August 2012 (units: mm).

Event	East		North		Vertical	
	O	P	O	P	O	P
Event 1	1.52	0.56	1.52	0.62	4.47	2.26
Event 2	1.67	0.64	1.61	0.71	4.51	2.38
Event 3	1.60	0.61	1.58	0.51	4.49	2.87
Mean	1.60	0.60	1.57	0.61	4.49	2.50

O denotes the original GNSS data; P denotes the proposed method.

and 2.50 mm, respectively. This represents standard deviation reductions of 62.5%, 61.1%, and 44.3%, respectively. The proposed method offers submillimeter-level accuracy for the horizontal components, and millimeter-level accuracy for the vertical component. This high accuracy is mainly afforded by the high performance of CEEMD, the use of the normalized autocorrelation function and correlation coefficients, and the use of the wavelet-like soft threshold to directly denoise the noise-dominated high-frequency IMFs. The proposed method thus avoids the theoretical shortcoming of the excessive sensitivity of the wavelet transform to the modulation of the wavelet parameters, and thus produces more precise seismic waveforms for moderate and small earthquakes.

VI. CONCLUSION

We presented a new adaptive denoising method for preserving more accurate high-rate GNSS seismic waveforms. The proposed method utilizes the correlation coefficients between the original signal and the IMFs obtained from it by CEEMD, the normalized autocorrelation functions of the IMFs, and 'wavelet-like' threshold denoising. The performance of the proposed method was validated using the high-rate (5 Hz) GNSS data obtained during the moderate EI Mayor–Cucapah earthquake (M_w 7.2) of 4 April 2010 and the small Brawley seismic swarm (M_w 4.6–5.5) of 26 August 2012.

The results showed that for denoising the high-rate GNSS displacements, the proposed method is more precise than the other widely used method, such as MSF, Stacking, and MSF plus Stacking. The proposed method can significantly remove high-frequency white noise and low-frequency colored noise, caused by CME, multipath errors, and/or other unmodeled systematic errors in high-rate GNSS displacements. Further, the standard deviations of the GNSS original data indicate millimeter-level and centimeter-level accuracies for the horizontal and vertical displacements components, respectively. While denoised GNSS results obtained by the proposed method reveal submillimeter-level and millimeter-level accuracies, respectively. In particularly, the GNSS denoised coseismic displacements obtained by the proposed method were well preserved to match the seismic displacements measured by seismic instruments, better than the original GNSS solutions. Some small-amplitude details that were undetected in the original seismic displacements were detected in the GNSS denoised seismic displacements. This high accuracy of

the proposed method is mainly attributable to the following: 1) the high performance of CEEMD for decomposing GNSS signals into IMFs of different frequencies, 2) the use of the normalized autocorrelation function and the correlation coefficients to determine the noise-dominated high-frequency IMFs, and 3) the use of the wavelet-like soft threshold to directly denoise noise-dominated high-frequency IMFs. The above attributes confirmed that the proposed method significantly improved the precision of GNSS solutions and enabled effective detection of seismic signals for moderate and small earthquakes.

For future prospects and challenges, three main issues must be considered. First, denosied GNSS coseismic displacements require further assessments of the lower bound of sensitivity of GNSS-only or combined GNSS/accelerometer seismogeodetic waveforms. Secondly, further studies will be built on the characteristics of GNSS noise and the denoising techniques introduced in this paper to improve the precision of GNSS solutions. Finally, the real-time coseismic estimation of GNSS-only and combined GNSS/ accelerometer solutions with the proposed method must be additionally checked to accurately detect P-wave arrival and quickly estimate the earthquake magnitude for EEW systems.

REFERENCES

[1] C. O. Yigit, "Experimental assessment of post-processed kinematic precise point positioning method for structural health monitoring," *Geomatics, Natural Hazards Risk*, vol. 7, no. 1, pp. 360–383, 2016.

[2] C. O. Yigit and E. Gurlek, "Experimental testing of high-rate GNSS precise point positioning (PPP) method for detecting dynamic vertical displacement response of engineering structures," *Geomatics, Natural Hazards Risk*, vol. 8, no. 2, pp. 893–904, Feb. 2017.

[3] A. Bilich, J. F. Cassidy, and K. M. Larson, "GPS seismology: Application to the 2002 Mw 7.9 denali fault earthquake," *Bull. Seismolog. Soc. Amer.*, vol. 98, no. 2, pp. 593–606, Apr. 2008.

[4] B. W. Crowell, Y. Bock, and M. B. Squibb, "Demonstration of earthquake early warning using total displacement waveforms from real-time GPS networks," *Seismolog. Res. Lett.*, vol. 80, no. 5, pp. 772–783, Sep/Oct. 2009.

[5] Z. Duputel, L. Rivera, H. Kanamori, G. P. Hayes, B. Hirshorn, and S. Weinstein, "Real-time W phase inversion during the 2011 off the Pacific coast of Tohoku earthquake," *Earth Planets Space*, vol. 63, no. 7, pp. 535–539, Jul. 2011.

[6] G. Blewitt, C. Kreemer, W. C. Hammond, H.-P. Plag, S. Stein, and E. Okal, "Rapid determination of earthquake magnitude using GPS for tsunami warning systems," *Geophys. Res. Lett.*, vol. 33, no. 11, pp. L11309–L11312, Jun. 2006.
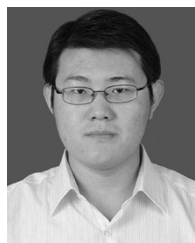
[7] D. Melgar, Y. Bock, and B. W. Crowell, "Real-time centroid moment tensor determination for large earthquakes from local and regional displacement records," *Geophys. J. Int.*, vol. 188, no. 2, pp. 703–718, Feb. 2012.

[8] D. Melgar, B. W. Crowell, Y. Bock, and J. S. Haase, "Rapid modeling of the 2011 Mw 9.0 Tohoku-oki earthquake with seismogeodesy," *Geophys. Res. Lett.*, vol. 40, no. 12, pp. 2963–2968, Jun. 2013.

[9] B. W. Crowell, Y. Bock, and D. Melgar, "Real-time inversion of GPS data for finite fault modeling and rapid hazard assessment," *Geophys. Res. Lett.*, vol. 39, no. 9, pp. 9305–9310, May 2012.

[10] A. Guo, S. Ni, W. Chen, J. T. Freymueller, and Z. Shen, "Rapid earthquake focal mechanism inversion using high-rate GPS velometers in sparse network," *Sci. China Earth Sci.*, vol. 58, no. 11, pp. 1970–1981, Nov. 2015.

[11] S. Zhong, C. Xu, L. Yi, and Y. Li, "Focal mechanisms of the 2016 central italy earthquake sequence inferred from high-rate GPS and broadband seismic waveforms," *Remote Sens.*, vol. 10, no. 4, p. 512, Mar. 2018.

[12] S. E. Minson, J. R. Murray, J. O. Langbein, and J. S. Gomberg, "Real-time inversions for finite fault slip models and rupture geometry based on high-rate GPS data," *J. Geophys. Res., Solid Earth*, vol. 119, no. 4, pp. 3201–3231, Apr. 2014.

[13] T. J. Wright, N. Houlié, M. Hildyard, and T. Iwabuchi, "Real-time, reliable magnitudes for large earthquakes from 1 Hz GPS precise point positioning: The 2011 Tohoku-Oki (Japan) earthquake," *Geophys. Res. Lett.*, vol. 39, no. 12, pp. 12302–12306, Jun. 2012.

[14] Y. Zheng, J. Li, Z. Xie, and M. H. Ritzwoller, "5Hz GPS seismology of the El Mayor—Cucapah earthquake: Estimating the earthquake focal mechanism," *Geophys. J. Int.*, vol. 190, no. 3, pp. 1723–1732, Jun. 2012.

[15] G. Blewitt, W. C. Hammond, C. Kreemer, H.-P. Plag, S. Stein, and E. Okal, "GPS for real-time earthquake source determination and tsunami warning systems," *J. Geodesy*, vol. 83, nos. 3–4, pp. 335–343, Mar. 2009.

[16] D. Melgar, Y. Bock, D. Sanchez, and B. W. Crowell, "On robust and reliable automated baseline corrections for strong motion seismology," *J. Geophys. Res., Solid Earth*, vol. 118, no. 3, pp. 1177–1187, Mar. 2013.

[17] R. M. Allen, P. Gasparini, O. Kamigaichi, and M. Bose, "The status of earthquake early warning around the world: An introductory overview," *Seismolog. Res. Lett.*, vol. 80, no. 5, pp. 682–693, Sep. 2009.

[18] Y. Nakamura and J. Saita, "UrEDAS, the earthquake warning system: Today and tomorrow," in *Earthquake Early Warning Systems*. Berlin, Germany: Springer, 2007, pp. 249–281.

[19] R. M. Allen and H. Kanamori, "The potential for earthquake early warning in southern California," *Science*, vol. 300, no. 5620, pp. 786–789, May 2003.

[20] T. H. Heaton, "A model for a seismic computerized alert network," *Science*, vol. 228, no. 4702, pp. 987–990, May 1985.

[21] D. Dong, P. Fang, Y. Bock, F. Webb, L. Prawirodirdjo, S. Kedar, and P. Jamason, "Spatiotemporal filtering using principal component analysis and Karhunen-Loeve expansion approaches for regional GPS network analysis," *J. Geophys. Res.*, vol. 111, no. B3, pp. 405–420, Mar. 2006.

[22] X. He, X. Hua, K. Yu, W. Xuan, T. Lu, W. Zhang, and X. Chen, "Accuracy enhancement of GPS time series using principal component analysis and block spatial filtering," *Adv. Space Res.*, vol. 55, no. 5, pp. 1316–1327, Mar. 2015.

[23] K. Choi, A. Bilich, K. M. Larson, and P. Axelrad, "Modified sidereal filtering: Implications for high-rate GPS positioning," *Geophys. Res. Lett.*, vol. 31, no. 22, pp. L22608–L22611, Nov. 2004.

[24] K. M. Larson, A. Bilich, and P. Axelrad, "Improving the precision of high-rate GPS," *J. Geophys. Res., Solid Earth*, vol. 112, no. B5, pp. 422–432, May 2007.

[25] M. Wang, J. Wang, D. Dong, W. Chen, H. Li, and Z. Wang, "Advanced sidereal filtering for mitigating multipath effects in GNSS short baseline positioning," *ISPRS Int. J. Geo-Inf.*, vol. 7, no. 6, p. 228, Jun. 2018.

[26] Y. Bock and D. Melgar, "Physical applications of GPS geodesy: A review," *Rep. Prog. Phys.*, vol. 79, no. 10, pp. 106801–106919, Aug. 2016.

[27] T. Rui, M. Ge, R. Wang, and T. R. Walter, "A new algorithm for tight integration of real-time GPS and strong-motion records, demonstrated on simulated, experimental, and real seismic data," *J. Seismol.*, vol. 18, no. 1, pp. 151–161, Jan. 2014.

[28] Y. Bock, R. M. Nikolaidis, P. J. de Jonge, and M. Bevis, "Instantaneous geodetic positioning at medium distances with the global positioning system," *J. Geophys. Res., Solid Earth*, vol. 105, no. B12, pp. 28223–28253, Dec. 2000.

[29] Y. Chen, Y. Jiang, S. Zu, S. Qu, and S. Gan, "Seismic imaging of simultaneous-source data using constrained least-squares reverse time migration," *J. Appl. Geophys.*, vol. 114, pp. 32–35, Mar. 2015.

[30] Y. Chen, H. Chen, K. Xiang, and X. Chen, "Geological structure guided well log interpolation for high-fidelity full waveform inversion," *Geophys. J. Int.*, vol. 207, no. 2, pp. 1313–1331, Nov. 2016.

[31] S. Gan, S. Wang, Y. Chen, S. Qu, and S. Zu, "Velocity analysis of simultaneous-source data using high-resolution semblance-coping with the strong noise," *Geophys. J. Int.*, vol. 204, no. 2, pp. 768–779, Dec. 2016.

[32] Y. Chen, "Probing the subsurface karst features using time-frequency decomposition," *Interpretation*, vol. 4, no. 4, pp. T533–T542, Sep. 2017.

[33] C. Xu, Z. Gong, and J. Niu, "Recent developments in seismological geodesy," *Geodesy Geodyn.*, vol. 7, no. 3, pp. 157–164, May 2016.

[34] M. R. Kaloop, C. O. Yigit, and J. W. Hu, "Analysis of the dynamic behavior of structures using the high-rate GNSS-PPP method combined with a wavelet-neural model: Numerical simulation and experimental tests," *Adv. Space Res.*, vol. 61, no. 6, pp. 1512–1524, Mar. 2018.

[35] Y. Li, C. Xu, L. Yi, and R. Fang, "A data-driven approach for denoising GNSS position time series," *J. Geodesy*, vol. 92, no. 8, pp. 905–922, Aug. 2018.

[36] Y. Li, C. Chen, R. Fang, and L. Yi, "Accuracy enhancement of high-rate GNSS positions using a complete ensemble empirical mode decomposition-based multiscale multiway PCA," *J. Asian Earth Sci.*, vol. 169, pp. 67–78, Jan. 2019.

[37] J. Sun and J. Wen, "Target location method for pipeline pre-warning system based on HHT and time difference of arrival," *Measurement*, vol. 46, no. 8, pp. 2716–2725, Oct. 2013.

[38] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. London A, Math., Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.

[39] Z. Wu and N. E. Huang, "A study of the characteristics of white noise using the empirical mode decomposition method," *Proc. Roy. Soc. London A, Math., Phys. Eng. Sci.*, vol. 460, no. 2046, pp. 1597–1611, Jun. 2004.

[40] P. Flandrin and P. Gonçalvès, "Empirical mode decompositions as data-driven wavelet-like expansions," *Int. J. Wavelets, Multiresolution Inf. Process.*, vol. 2, no. 4, pp. 477–496, 2004.

[41] B. B. Mandelbrot and J. W. Van Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM Review*, vol. 10, no. 4, pp. 422–437, Oct. 1968.

[42] P. Flandrin, P. Gonçalvès, and G. Rilling, "EMD equivalent filter banks, from interpretation to applications," in *Hilbert-Huang Transform and Its Applications*. Singapore: World Scientific, 2005, pp. 57–74.

[43] M. A. Colominas, G. Schlotthauer, and M. E. Torres, "Improved complete ensemble EMD: A suitable tool for biomedical signal processing," *Biomed. Signal Process. Control*, vol. 14, pp. 19–29, Nov. 2014.

[44] J. Han and M. van der Baan, "Empirical mode decomposition for seismic time-frequency analysis," *Geophysics*, vol. 78, no. 2, pp. O9–O19, Feb. 2013.

[45] H. Jiang, C. Li, and H. Li, "An improved EEMD with multiwavelet packet for rotating machinery multi-fault diagnosis," *Mech. Syst. Signal Process.*, vol. 36, no. 2, pp. 225–239, Apr. 2013.

[46] M. Žvokelj, S. Zupan, and I. Prebil, "EEMD-based multiscale ICA method for slewing bearing fault detection and diagnosis," *J. Sound Vib.*, vol. 370, pp. 394–423, May 2016.

[47] J. Wang, J. Wang, and C. Roberts, "Reducing GPS carrier phase errors with EMD-wavelet for precise static positioning," *Surv. Rev.*, vol. 41, no. 312, pp. 152–161, Jul. 2009.

[48] S. Baykut, T. Akgul, and S. Ergintav, "EMD-based analysis and denoising of GPS data," in *Proc. IEEE Signal Process. Commun. Appl. Conf.*, Apr. 2009, pp. 644–647.

[49] P. Flandrin, P. Gonçalvès, and G. Rilling, "Detrending and denoising with empirical mode decompositions," in *Proc. 12th Eur. Signal Process. Conf.*, Sep. 2004, pp. 1581–1584.

[50] D. N. Kaslovsky and F. G. Meyer, "Noise corruption of empirical mode decomposition and its effect on instantaneous frequency," *Adv. Adapt. Data Anal.*, vol. 2, no. 3, pp. 373–396, 2010.

[51] J. R. Yeh, J. S. Shieh, and N. E. Huang, "Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method," *Adv. Adapt. Data Anal.*, vol. 2, no. 2, pp. 135–156, Apr. 2010.

[52] Y. Kopsinis and S. McLaughlin, "Empirical mode decomposition based soft-thresholding," in *Proc. Eur. Signal Process. Conf.*, Aug. 2008, pp. 1–5.

[53] U. Amato and A. Antoniadis, "Adaptive wavelet series estimation in separable nonparametric regression models," *Statist. Comput.*, vol. 11, no. 4, pp. 373–394, Oct. 2001.

[54] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, Aug. 1994.

[55] A. O. Boudraa, J. C. Cexus, and Z. Saidi, "EMD-based signal noise reduction," *Int. J. Signal Process.*, vol. 1, no. 1, pp. 33–37, Jan. 2004.

[56] N. Ricker, "Further developments in the wavelet theory of seismogram structure," *Bull. Seismolog. Soc. Amer.*, vol. 33, no. 3, pp. 197–228, 1943.

[57] R. B. Lohman and J. J. Mcguire, "Earthquake swarms driven by aseismic creep in the Salton Trough, California," *J. Geophys. Res., Solid Earth*, vol. 112, no. B4, pp. 1–10, Apr. 2007.

[58] J. Geng, Y. Bock, D. Melgar, B. W. Crowell, and J. S. Haase, "A new seismogeodetic approach applied to GPS and accelerometer observations of the 2012 Brawley seismic swarm: Implications for earthquake early warning," *Geochem. Geophys. Geosys.*, vol. 14, no. 7, pp. 2124–2142, Apr. 2013.

[59] C. Shi, Q. Zhao, J. Geng, Y. Lou, M. Ge, and J. Liu, "Recent development of PANDA software in GNSS data processing," *Proc. SPIE*, vol. 72851, Dec. 2008, Art. no. 72851S.

[60] S. Wdowinski, Y. Bock, J. Zhang, P. Fang, and J. Genrich, "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*, vol. 102, no. B8, pp. 18057–18070, Aug. 1997.

[61] S. D. P. Williams, Y. Bock, P. Fang, P. Jamason, R. M. Nikolaidis, L. Prawirodirdjo, M. Miller, and D. J. Johnson, "Error analysis of continuous GPS position time series," *J. Geophys. Res.*, vol. 109, no. B3, pp. 142–160, Mar. 2004.

[62] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*. Upper Saddle River, NJ, USA: Pearson, 1997.

[63] R. Nikolaidis, "Observation of geodetic and seismic deformation with the global positioning system," *Cancer Res.*, vol. 71, no. 8, p. 714, Jan. 2002.

CHUANFA CHEN was born in Zibo, Shandong, China, in 1982. He received the Ph.D. degree in cartography and geography information system from the Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China, in 2010.

He is currently a Professor with the Shandong University of Science and Technology. He is major in digital terrain modeling and lidar point cloud classification.

HAITAO YIN was born in Liaocheng, Shandong, China, in 1980. He received the Ph.D. degree in cartography and geography information system from Southwest Jiaotong University, Chengdu, China, in 2007.

He is currently an Associate Researcher with the Earthquake Administration of Shandong Province, China. His work has focused on the high-rate GPS data processing and strong ground movement analyzing.

YANYAN LI was born in Weifang, Shandong, China, in 1987. She received the Ph.D. degree in geodesy and geomatics from Wuhan University, Wuhan, China, in 2017.

She is currently a Lecture with the College of Geomatics, Shandong University of Science and Technology. Her current research interest includes high-rate GPS data processing and its application on seismology.

LEI YI was born in Wuhan, China, in 1989. He received the Ph.D. degree in solid geophysics from the School of Geodesy and Geomatics, Wuhan University, in 2017. pt He is currently a Faculty with the Qinghai Institute of Salt Lakes, Chinese Academy of Sciences. His current research interests include tectonic earthquake and joint inversion in rupture process.

CAIJUN XU was born in Jiangyin, Jiangsu, China, in 1964. He received the degree from the Wuhan Technical University of Surveying and Mapping Geodesy, in 1985, and the master's and the Ph.D. degree's in 1988 and 1994, respectively.

He is the specially-appointed Professor of the Cheung Kong Scholars Program, the Vice Dean of the Geodesy and Geomatics, Wuhan University. He is currently a member of the Sub-Committee of Applications of Satellite and Airborne Imaging Systems of International Association of Geodesy (IAG). His current research interests are space geodesy, geophysical geodesy, and geodynamics. He is an Editorial Board Member of *Journal of Applied Geodesy*, *The Journal of Global Positioning Systems*, and *Journal of Geodesy and Geodynamics*.

XIAOXING HE received the Ph.D. degree in geodesy and surveying engineering from the School of Geodesy and Geomatics, Wuhan University, China, in 2016.

He has been a Postdoctoral Researcher at the GNSS Research Center, Wuhan University, since 2017. He is currently a Lecturer with the School of Civil Engineering and Architecture, East China JiaoTong University, Nanchang, China. His research interests include the theory of satellite Geodesy and its applications, assessment of noise characteristics, and analysis of geodetic time series.

• • •