De-noising low-frequency magnetotelluric data using mathematical morphology filtering and sparse representation

Guang Li a,b, Xiaqiong Liu c, Jingtian Tang b,*, Jin Li c,*, Zhengyong Ren b, Chaojian Chen b

a State Key Laboratory of Nuclear Resources and Environment, School of Geophysics and Measurement-control Technology, East China University of Technology, Nanchang, Jiangxi 330013, China
b Key Laboratory of Metallogenic Prediction of Non-Ferrous Metals and Geological Environment Monitor, Ministry of Education, Central South University, Changsha, Hunan 410083, China
c College of Information Science and Engineering, Hunan Normal University, Changsha, Hunan 410081, China

A R T I C L E   I N F O

Article history:
Received 11 January 2019
Received in revised form 5 December 2019
Accepted 6 December 2019
Available online 9 December 2019

Keywords:
Magnetotelluric data
Noise attenuation
Sparse representation
Over-complete dictionary
Low-frequency signal
Mathematical morphology filtering

A B S T R A C T

De-noising the magnetotelluric (MT) data using the conventional time-series editing methods is at the risk of losing low-frequency signals, especially the signal below 1 Hz. To overcome this deficiency, we propose a combinatorial method based on sparse representation and mathematical morphology filtering. First, the effective low-frequency signal is reconstructed using the mathematical morphological filtering (MMF) method and protected. Then, the residual noisy signal of high frequency is sparsely decomposed using the subspace pursuit (SP) algorithm to obtain noise-free high-frequency MT signals. Finally, the effective low-frequency signal is added to the de-noised high-frequency signal to get the full-band MT data. We evaluate the proposed method using a synthetic data set and two real data sets collected in Qiadam Basin, the northeastern part of the Tibetan Plateau. Experimental results demonstrate that the presented approach can be used to remove different kinds of cultural noises while preserve the low-frequency signal below 1 Hz. The evaluation results also indicate that the proposed method is superior to the conventional methods in terms of the signal-to-noise ratio (SNR), reconstruction error (E) and normalized cross-correlation (NCC).

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

Magnetotelluric (MT) method is one of the most commonly used methods for the exploration of deep electrical structure (Zhao et al., 2017; Nittinger and Becken, 2018; Tang et al., 2018). Nevertheless, rapid socioeconomic development increases the difficulty of obtaining high-quality MT data. Strong cultural noise has become one of the key obstacles to the application of MT method (Escalas et al., 2013; Larnier et al., 2016; Giuseppe et al., 2018).

To improve the quality of MT data, numerous approaches have been developed. Meanwhile, the remote reference method (Goubau et al., 1978; Gamble et al., 1979) and robust statistic method (Egbert and Booker, 1986) are among the most commonly applied methods. However, the results of the remote reference method are critically dependent on the quality of data collected in the remote reference station. In addition, the robust statistic method is often ineffective when the noise is relevant or persistent (Weckmann et al., 2005). To remove cultural noises, several modern technologies in the field of signal processing have been applied to process MT time-series, such as wavelet decomposition (Zhang and Paulson, 1997; Trad and Travassos, 2000; Garcia and Jones, 2008), empirical mode decomposition (EMD) (Neukirch and Garcia, 2014; Cai, 2016), and mathematical morphology filtering (MMF) (Tang et al., 2012). These time-domain methods are effective for the removal of relevant or persistent noises as they directly eliminate those structures with abnormally large amplitude. However, the direct extraction of noise from time-series has the danger of discarding effective signals, especially the low-frequency signal (or called the effective slow portion because it varies slowly with time) below 1 Hz (Ling et al., 2015).

To preserve the low-frequency signal, several methods have been proposed based on sparse decomposition and greedy algorithms. A redundant dictionary, which matches the cultural noise but is insensitive to the effective MT signal, was developed to merely process the noisy segments (Li et al., 2017; Tang et al., 2017). In addition, a signal-noise identification method based on multifractal spectrum was developed to preserve the noise-free segments before sparse decomposition denoising (Li et al., 2019). Although these methods could effectively preserve noise-free segments, they could also damage the effective slow portion in the noisy segments.

To overcome the above-mentioned issue, we propose a new time-series editing strategy by combining the mathematical morphological
filtering and sparse representation. In our new scheme, the effective low-frequency MT signal is firstly reconstructed (or called extracted) using MMF and protected. Then the residual noisy high-frequency signal is de-noised using a predesigned redundant dictionary and the sub-space pursuit (SP) algorithm (i.e., a signal reconstruction algorithm for compressed sensing). The main difference between our new scheme and other existing time-domain methods is that we do not directly extract the noise from the raw data, but first reconstruct the effective slow portion. In this way, there is little risk of discarding the effective low-frequency signal. Synthetic and measured MT data sets are tested in order to evaluate the performance of the presented method. In addition, the newly proposed strategy is compared with other time-series editing methods to verify the feasibility and effectiveness.

This paper is organized as follows. To begin with, the basic principles of MMF and sparse decomposition are described in detail (Section 2). Subsequently, several numerical experiments are conducted to illustrate the accuracy of slow portion extraction and the effectiveness of signal-noise separation (Section 3). Then, the new scheme is applied for the noise attenuation of the field MT data collected from Qaidam Basin, the northeastern part of the Qinghai-Tibet Plateau (Section 4). Finally, conclusions are given and the disadvantages of the new scheme are discussed (Section 5).

2. Methods and theories

2.1. Workflow of the proposed method

Here, we present a brief workflow of the proposed de-noising strategy. As illustrated in Fig. 1, first, the noisy MT time-series \( A \) is decomposed into the effective low-frequency signal \( B \) and the noisy high-frequency signal \( X \) using MMF. Subsequently, the noisy high-frequency signal \( X \) is decomposed into the effective high-frequency signal \( Y_h \) and cultural noise \( Y_c \) using sparse decomposition (i.e., using a predefined redundant dictionary and the SP algorithm). Eventually, the low-frequency signal \( B \) and high-frequency signal \( Y_h \) are merged to obtain the full-band effective MT signal \( Y \) (i.e., the de-noised MT signal).

2.2. Mathematical morphology filtering

Mathematical morphology is derived from the study of the geometry of binary porous media. Mathematical morphological filtering is a nonlinear filtering technique developed from mathematical morphology, which is widely adopted for image segmentation (Pal and Pal, 1993), noise reduction of EEG signals (Nishida et al., 1999), microseismic monitoring (Li et al., 2016), noise suppression of seismic data (Huang et al., 2017), and many other applications (Soille and Pesaresi, 2002; Zhang et al., 2003). There are two fundamental operations in MMF, i.e., erosion and dilation. Briefly, the operation of erosion (dilation) is equivalent to the minimum (maximum) filtering of the discrete function within the sliding filter window (equivalent to the structural element). Let \( A = [a_1, a_2, \ldots, a_N]^T \) be the raw MT signal to be processed, and \( g = [b_1, b_2, \ldots, b_M]^T \) be a structural element; then the erosion and dilation operations can be defined as:

\[
(A \ominus g) = \min_{m=1, 2, \ldots, M} \left\{a(n+m) - g(m)\right\}, \quad n = 1, 2, \ldots, (N+M-1),
\]

\[
(A \oplus g) = \max_{m=1, 2, \ldots, M} \left\{a(n-m) + g(m)\right\}, \quad n = 1, 2, \ldots, (N-M+1)
\]

where “\(T\)” stands for matrix transposition. “\(\ominus\)” denotes the operation of erosion, “\(\oplus\)” represents the operation of dilation. The operation of erosion is a process of shrinking, which can be used to eliminate the boundary points of the object and reduce the spikes. The operation of dilation is a process of expansion, which can be applied to fill the concave portion when the boundary is not smooth (Tang et al., 2012).

New operations can be produced by combining erosion and dilation. The operation of opening can be accomplished by carrying out an erosion of the target signal, followed by a dilation; while the operation of closing can be generated by performing a dilation first and then an erosion:

\[
(A \ominus g) \oplus g,
\]

\[
(A \oplus g) \ominus g
\]

where “\(-\)” and “\(+\)” are the operations of opening and closing, respectively. The operations of opening and closing constitute the most basic filters in mathematical morphology. The filter of opening-closing (OC) is achieved by performing the opening operation and closing operation in sequence. The filter of closing-opening (CO) is accomplished by

![Fig. 1. Workflow of the proposed MMF-SP method.](image-url)
performing a closing operation and then an opening operation:

\[
\text{OC}(A) = A \ast g \ast g.
\]

(5)

\[
\text{CO}(A) = A \ast g \ast g.
\]

(6)

Both OC and CO can filter out the noise in the target signal. Nevertheless, the OC filter often results in a smaller output, while the CO filter amplifies the output. Hence, an average of these outputs is often used to approximate the original signal, which can be expressed as:

\[
B = \frac{\text{OC}(A) + \text{CO}(A)}{2}.
\]

(7)

In the community of magnetotelluric data processing, mathematical morphological filtering was originally used for the reconstruction of cultural noises (Tang et al., 2012; Li et al., 2017). However, the low-frequency MT signal has large amplitude and large time scale (or called data points). That is to say, the effective low-frequency component of the MT data below 1 Hz, which causes the output signal \(B\) in Eq. (7) to be predominantly low-frequency effective signal instead of noise. The aforementioned MMF-based methods treat the output signal \(B\) as noise and discard it, resulting in the loss of low-frequency signals. For this reason, we manage to reconstruct and protect the effective low-frequency MT component using MMF. In our scheme, the output signal \(B\) represents the effective low-frequency MT signal, \(X\) stands for the noisy high-frequency signal obtained by subtracting the effective low-frequency signal \(B\) from the raw data \(A\):

\[
X = A - B.
\]

(8)

### 2.3. Sparse decomposition and the subspace pursuit (SP) algorithm

Sparse decomposition (Mallat and Zhang, 1993) is also known as sparse representation. Being one of the prerequisites of compressed sensing, sparse representation has gained widespread attention recently and has been extensively applied for the analysis of speech signals (Jafari and Plumbley, 2011), fault diagnosis of mechanical sensing, and sparse representation has gained widespread attention recently (Zhou et al., 2014; Chen, 2017; Zhang et al., 2018), two-dimensional magnetic field inversion (Nittinger and Becken, 2018), etc. According to the principle of sparse representation, signal \(X = [x_1, x_2, ..., x_N]^T\) can be expressed as a linear combination of a redundant dictionary \(D^{K \times L} = [d_1, d_2, ..., d_K]\) and the corresponding linear coefficients \(s\):

\[
X = Ds + e,
\]

(9)

where \(d_i = [d_{i1}, d_{i2}, ..., d_{iL}]^T\) is an atom (or called a basis) in \(D\), \(s = [s_1, s_2, ..., s_K]^T\) is a vector of weight coefficients with all but the largest \(K\) coefficients setting to zero, \(K\) is the sparsity of signal \(X\), and \(e\) is the error of approximation.

As mentioned before, the sparse decomposition mainly includes two parts, i.e., the design of a redundant dictionary and the reconstruction of the target signal. Commonly used redundant dictionaries are the Gabor dictionary (Mallat and Zhang, 1993; Zibulski and Zeevi, 1997), impulsive dictionary (Wang et al., 2013), Fourier dictionary (or called harmonic wave dictionary, Zhu et al., 2015), square-wave dictionary (Tang et al., 2017), wavelet dictionary (Nittinger and Becken, 2018), etc. Notably, the square-wave dictionary is suitable for extracting structures with strong instantaneous energy such as square-waves and spikes. In our method, we choose the square-wave dictionary because the predominant interferences in MT data are impulsive and square-wave noises. The dictionary is defined as follows:

\[
d_y = \begin{bmatrix} 0, & \tau < \tau_1 \\ a, & \tau_1 < \tau \leq \tau_2 \\ 0, & \tau > \tau_2 \end{bmatrix}.
\]

(10)

where \(d_y\) is a normalized atom defined by the parameters \(\tau\) and \(a\), \(\tau \in (1, N)\) is the duration, \(a\) is the amplitude. A large number of atoms with different shapes can be produced by changing the parameters \(\tau\) and \(a\).

There are mainly three types of algorithms for signal reconstruction, i.e., greedy pursuit, convex relaxation and combinatorial algorithms (Needell and Tropp, 2009; Fang et al., 2016). Due to low complexity and simple geometric interpretation (Tropp and Gilbert, 2007; Dai and Milenkovic, 2009), greedy algorithms are widely applied in many fields. With the development of compressed sensing, plenty of optimized algorithms have been proposed based on the standard matching pursuit (MP) algorithm (Mallat and Zhang, 1993) and orthogonal MP (OMP) algorithm (Pati et al., 1993), for instance, compressive sampling MP (CoSaMP, Needell and Tropp, 2009) and subspace pursuit (SP, Dai and Milenkovic, 2009). The standard OMP algorithm selects only one atom per iteration, and the number of atoms in the redundant dictionary is huge. Therefore, the reconstruction of OMP is quite time-consuming. Besides, once an atom is selected, it is unchangeable. The SP algorithm adds \(K\) new atoms to the candidate set per iteration, and then \(K\) atoms are selected from the candidate set. This allows the initial atom to be replaced with a more suitable one in any subsequent iteration. In most cases, SP is more accurate and efficient than OMP. Whilst the SP and CoSaMP algorithms have comparable accuracy and ratio of successful reconstruction, they select the candidate atoms differently. In each iteration, the CoSaMP selects 2 \(K\) candidate atoms, while the SP selects only \(K\). This makes the SP algorithm more efficient (Dai and Milenkovic, 2009; Marques et al., 2019).

Supposing \(D = \{d_y\}_{y=1}^\gamma\) is the designed redundant dictionary, \(d_y\) is an atom in \(D\), and \(\|d_y\| = 1\), \(\gamma\) is the order of atom \(d_y\), \(\Gamma\) is the set of \(\gamma\), \(K\) is the degree of sparsity, \(R\) is the residual after iterating \(l\) times, \(A_i\) is the candidate set of atoms, \(\psi_i\) is the selected set of atoms, and \(X\) is the signal to be processed. The steps of the SP algorithm can be expressed as follows:

**Input:** \(K, D, X, l\), the candidate set of atoms \(A_i\), the selected set of atoms \(\psi_i\), the residual \(R\).

**Initialization:** \(l = 1, R^0 = X, \psi_0 = \emptyset, A_0 = \emptyset\).

**Iteration:**

1. Update the candidate set of atoms \(A_i\):

\[
A_l = A_{l-1} \cup \{K \text{ atoms corresponding to the largest value of } ||R^l D||\}
\]

(11)

2. Calculate the projection coefficient \(u_i\):

\[
u_i = \left(A_l^T A_l\right)^{-1} \cdot A_l^T X
\]

(12)

3. Update the selected set of atoms \(\psi_i\):

\[
\psi_i = \{K \text{ atoms lead to the largest value of } u_i\}
\]

(13)

4. Update the reconstructed signal \(Y_i\) and residual \(R^l\):

\[
Y_i = \psi_i (\psi_i^T \psi_i)^{-1} \cdot \psi_i^T X
\]

(14)

\[
R^l = X - Y_i
\]

(15)
If $\| R \|^2 > \| R^{-1} \|^2$, let $\psi = \psi_{l-1}$ and stop iteration; otherwise, $l = l + 1$ and return to step 1.

Output: the reconstructed signal $Y_c = X - Y_c$.

The de-noised high-frequency MT signal $Y_h$ can be obtained as:

$$ Y_h = X - Y_c. \tag{16} $$

Finally, the de-noised full-band MT data $Y$ is obtained by merging the effective low-frequency signal $B$ and high-frequency portion $Y_h$, given by:

$$ Y = B + Y_h. \tag{17} $$

### 3. Case study of synthetic data sets

In this section, a series of experiments are conducted to quantitatively analyze the feasibility and effectiveness of the proposed method. A noise-free MT data set (site QH40150, as shown in Fig. 2) is selected from a large number of measured MT data gathered from Qaidam Basin, the northeastern part of Qinghai-Tibet Plateau. The MT data were recorded using MTU (Phoenix Geophysics Ltd) in 2012. At each site 21 h of recording was made covering a frequency range from 320 Hz to 0.001 Hz. The synthetic data were obtained by adding the simulated noise to good real data set of TS5, which was recorded at a sampling rate of 15 Hz.

#### 3.1. Analysis of time-series

We randomly selected four synchronous data segments of 1000 s from the real data QH40150 (see Fig. 3). Obviously, MT time-series recorded at this site are irregular, random and alike nonstationary noises. There is no abnormally large-amplitude structure, and the amplitudes of the signals vary slowly with time. Overall, MT data recorded at site QH40150 has a good signal-to-noise ratio. In addition, very low-frequency components can be found in each channel, especially in the electric channel of $E_y$. Nonetheless, it is difficult to separate cultural noise from the effective low-frequency component using traditional time-series editing methods because both of them have large energy and time scale.

The key step of our method is the extraction of the effective low-frequency signal. Since wavelet decomposition, MMF, and complementary ensemble EMD (CEEMD) are commonly applied for the extraction of low-frequency components (Li et al., 2016; Huang et al., 2017; Liu et al., 2019), an experiment of low-frequency signal extraction on synthetic data was performed using these approaches. In this experiment, the size of the structural elements $M$ in MMF is set to 80. Wavelet decomposition is performed using bior 4.4 with a 5-layer decomposition. The effective low-frequency signal obtained by CEEMD is realized by...
discarding the four highest frequency intrinsic mode functions (IMF) of the ten components. Following Yeh et al. (2010), the ratio of the standard deviation of the added noise in CEEMD is set to 0.2.

As shown in Fig. 4a, the raw noise-free data are a segment of the Ey component. Fig. 4b shows the simulated noises, including a spike and three intermittent square-wave noises with durations of 1 s (15 data points), 3 s (45 data points), and 5 s (75 data points). Fig. 4c shows the synthetic noisy data and effective slow portions extracted by different methods. It is clear that the slow portion extracted using CEEMD is inaccurate as severe distortion can be found at the locations of noises. This is because the strong noise affects the maximum and minimum values, resulting in inaccurate envelopes extracted during EMD decomposition. Similarly, the wavelet decomposition did not yield good results. Only MMF accurately extracted the effective low-frequency signal.

The size of the structural element is the most important parameter that affects the extraction of the slow portion. In this experiment, we keep other parameters constant, and merely alter the size of the structural elements \( M \) in order to investigate the effect of different structural element sizes on the results. Fig. 5 displays the effective slow portions extracted from the same synthetic data using linear structural elements with different sizes (e.g., 30, 50, and 80). When the size of the structural element is 30, noises with data points greater than 30 are misjudged as effective low-frequency signals. When \( M \) is set to 50, square-wave noise with a duration of 5 s is misjudged as slow portion. Misjudgment does not occur only when the size of the structural element is greater than the data points of noise. In fact, since the value of \( M \) cannot be too large, the ultra-low noises will be treated as effective signals. That is to say, the de-noising strategy described in this paper cannot eliminate ultra-low noises.

Fig. 6 shows the experimental results of MMF decomposition, where Fig. 6a is the synthetic noisy data obtained by adding simulated noises (as shown in Fig. 4b) to good real data; Fig. 6b is the slow portion of MT signal extracted by MMF; Fig. 6c is the residual noisy signal, mainly composed of cultural noise and effective high-frequency MT signal. It is much easier to perform noise attenuation on the signal shown in Fig. 6c than on the signal displayed in Fig. 6a. In our scheme, sparse decomposition is selected because of its superior ability in signal-noise identification (Li et al., 2017; Tang et al., 2017).

The performance of the proposed scheme is compared with EMD-ICA, SP, MMF-Wavelet, and MMF-SP (Harmonic). Figs. 7 and 8 present the time-domain and frequency-domain results, respectively. EMD-ICA means the combinatorial filtering method using EMD and independent component analysis (ICA) (Cai, 2016). SP represents the de-noising method using the SP algorithm and square-wave dictionary (Tang et al., 2017); MMF-Wavelet represents the de-noising method using MMF and wavelet threshold decomposition. MMF-SP (Harmonic) represents the de-noising method using MMF, SP algorithm, and harmonic dictionary (Zhu et al., 2015); MMF-SP (Square) is the new method proposed in this paper. Similar to our new scheme, MMF is used to reconstruct the effective low-frequency signal in the procedures MMF-Wavelet and

---

Fig. 4. Synthetic data and low-frequency signals extracted using different methods. (a) Real noise-free data; (b) simulated noise; (c) synthetic noisy data (green line) and effective slow portions extracted using MMF (purple line), CEEMD (red line), and wavelet (blue line), respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
Fig. 5. Synthetic noisy data and effective slow portions extracted using different sizes (M) of structural elements: (a) M = 30; (b) M = 50; and (c) M = 80.

Fig. 6. Results of MMF decomposition: (a) Synthetic noisy data; (b) effective low-frequency signal extracted using MMF; and (c) residual noisy high-frequency signal.
MMF-SP (Harmonic). That is to say, the output signal $B$ in Eq. (7) is protected as useful slow portion.

As shown in Fig. 7c, the slow portion is discarded together with the noises and large spikes remains after the application of the EMD-ICA method.

As illustrated in Fig. 7d, the SP-based method accurately identifies and attenuates all added noises. However, the added impulsive noise is not completely eliminated, and some useful signals are removed at the locations of the 150 and 200 s. The principle of the greedy algorithm shows that the basis for identifying noise is the largest inner product of noise and atoms. Therefore, the single SP-based method sometimes regards the effective signal with a large amplitude as noise. On the other side, some weak noise might be misjudged as useful signals. A direct application of SP-based method for noise reduction might lead to over-processing or under-processing.

The result obtained using MMF-Wavelet appears reasonable as it successfully eliminates all noises and preserves the slow portion. However, the signal obtained using MMF-Wavelet method becomes slimmer compared with Fig. 7a, which indicates the loss of some effective signals.

Fig. 7 reflects that the MMF-SP(Harmonic) method can preserve the slow portion but can not fully eliminate noises. This is because the atoms in the harmonic dictionary are clearly different from square-wave and impulsive noises. Therefore MMF-SP(Harmonic) based method can only remove harmonic components from the noise.

Similar to the MMF-SP (harmonic) method, we replaced the harmonic dictionary with a square wave dictionary and obtained satisfactory results. Fig. 7g clearly demonstrates that the noise is completely removed and the effective slow portion is well preserved.

As shown in Fig. 8, the simulated noise causes severe distortion of the amplitude-frequency curve. Not surprisingly, the loss of low-frequency signals caused by EMD-ICA is more obvious in the frequency-domain. The processing of the SP and MMF-SP (harmonic) results in an obvious improvement over the previous. However, significant errors can be observed compared to the noise-free curve. Compared with SP and MMF-SP (harmonic), MMF-Wavelet achieved better results. But there are still visible deviations between 0.02 and 0.1 Hz. Only the result obtained by MMF-SP(Square) is almost the same as the original amplitude-frequency curve.
As shown in Table 1, the signal-to-noise ratio (SNR), reconstruction error ($E$) and normalized cross-correlation (NCC) are used for objective comparison of performance (Li et al., 2017). The EMD-ICA improves the signal-to-noise ratio and reconstruction error, but decreases the normalized cross-correlation due to the loss of low-frequency signals. The SP method achieved better results because of its capability in signal-noise identification. The NCC increases from 0.4710 to 0.9779; SNR increases from $-5.5008$ dB to 13.5862 dB; and the reconstruction error $E$ decreases from 1.8838 to 0.2093. The MMF-Wavelet, MMF-SP (Harmonic), and MMF-SP(Square) all demonstrated substantial improvement because they protected the slow portion before de-noising. It is clear that MMF-SP (Square) outperforms all other methods in terms of the SNR, $E$ and NCC (with an NCC value of 0.9992; an SNR of 27.6887; and a reconstruction error of only 0.0413). Experimental results from Fig. 7, Fig. 8, and Table 1 have demonstrated the superiority of our method in terms of both objective and subjective evaluation criteria.

### 3.2. Results of MT response

To further evaluate the effectiveness of the proposed method, we added the same simulated noises (as shown in Fig. 4b) to the time-series of QH40150 repeatedly. In other words, all MT time-series are contaminated by simulated cultural noises. Then, the proposed method and the EMD-ICA method are used to filter the synthetic data. As shown in Fig. 9a, the apparent resistivity and phase curves calculated using the raw noise-free data vary smoothly and continuous with the frequencies except for some little bias below 0.001 Hz. As shown in Fig. 9b, apparent resistivity and phase curves obtained using the synthetic noisy data are severely distorted between 1 Hz and 0.003 Hz. In particular, the variance is huge for the curves around 0.1 Hz. As shown in Fig. 9c, the processing of the EMD-ICA method results in an obvious improvement over the previous between 1 Hz and 0.05 Hz. However, EMD-ICA also results in severe distortion of all curves below 0.05 Hz, especially the xy component. We applied our new procedure to the

<table>
<thead>
<tr>
<th>Methods</th>
<th>SNR</th>
<th>$E$</th>
<th>NCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unprocessed</td>
<td>$-5.5008$</td>
<td>1.8838</td>
<td>0.4710</td>
</tr>
<tr>
<td>EMD-ICA</td>
<td>$-0.3592$</td>
<td>1.0422</td>
<td>0.0231</td>
</tr>
<tr>
<td>SP(Square)</td>
<td>13.5862</td>
<td>0.2093</td>
<td>0.9779</td>
</tr>
<tr>
<td>MMF-Wavelet</td>
<td>22.3376</td>
<td>0.0764</td>
<td>0.9971</td>
</tr>
<tr>
<td>MMF-SP(Harmonic)</td>
<td>5.5028</td>
<td>0.5252</td>
<td>0.8847</td>
</tr>
<tr>
<td>MMF-SP(Square)</td>
<td>27.6887</td>
<td>0.0413</td>
<td>0.9992</td>
</tr>
</tbody>
</table>

Table 1: Quantitative evaluation of the simulated results.
synthetic noisy data and obtained the apparent resistivity and phase curves as shown in Fig. 9d. Except for a small deviation near 0.04 Hz, the results achieved by our scheme are well consistent with Fig. 9a. These results demonstrate the effectiveness and accuracy of the proposed method.

4. Real case study

4.1. Results of time-series

The low-frequency MT data recorded in noisy environment are often contaminated by square-wave noise, impulsive noise, and charge-discharge-like noise. In this section, three typical noisy time-series segments measured at a sampling rate of 15 Hz are used to evaluate the feasibility and effectiveness of the proposed method. Figs. 10, 11, and 12 show the results of noise attenuation for square-wave noise, impulsive noise, and charge-discharge-like noise, respectively. It is easy to find that the noises are significantly stronger than the effective signal. As expected, all kinds of cultural noises are completely eliminated after applying the proposed noise reduction method. It is worth noting that there is no visible damage to the effective slow portion. The results of real time-series editing show that the proposed method is effective for different kinds of noises.

4.2. Results of MT response

Measured MT data sets QH40180 and QH40200 were also recorded in Qiadam Basin (see the distribution of stations in Fig. 2). Fig. 13 shows the apparent resistivity and phase curves of the real station QH40180. As shown in Fig. 13a, apparent resistivity and phase curves...
obtained using the raw data are severely distorted between 5 Hz and 0.05 Hz. As shown in Fig. 13b, the MT response obtained using the presented method yields a great improvement over the previous. All the apparent resistivity and phase curves calculated using the de-noised data vary smoothly and continuous with the frequencies.

Fig. 14 shows the apparent resistivity and phase curves of the real station QH40200. These results demonstrate that all the apparent resistivity and phase curves obtained using the raw data are severely distorted between 5 Hz and 0.05 Hz. The curves of xy component below 0.01 Hz also reveal visible distortion. Obviously, the proposed method achieved more reasonable and reliable results. The bias between 5 Hz and 0.05 Hz is completely eliminated. Nonetheless, the improvement below 0.01 Hz is not obvious. This is the limitation of the proposed method – it is not effective for the removal of noises whose data points are greater than the size of structure element.

5. Conclusion

We have proposed a combinatorial time-series editing method named MMF-SP for noise reduction of MT data. The vital difference between our new scheme and other time-series editing methods is that the effective slow portion is protected before de-noising. The sparse de-composition is then adopted to perform the targeted noise suppression on the remaining noisy signal. Case studies of synthetic and measured MT data sets illustrate that:

Fig. 10. Noise attenuation for real data which are polluted by square-wave noise: (a) Raw time-series segments; (b) de-noised using the proposed method.

Fig. 11. Noise attenuation for real data which are polluted by impulsive noise: (a) Raw time-series segments; and (b) de-noised using the proposed method.
(a) The proposed method can remove different types of cultural noises while preserve the low-frequency signal. The result obtained using the new method is superior to the conventional methods because there is little danger of discarding useful low-frequency signals. This renders the new scheme effective for noise reduction of MT data below 1 Hz. Low-frequency signals are critical for large depth and high precision electromagnetic exploration. That is to say, our new strategy may improve the results of deep electrical structure exploration.

(b) The quality of the apparent resistivity and phase curves below 0.01 Hz is difficult to be improved using the proposed method. This is due to the fact that it is not effective to remove the noise when its time scale is greater than the length of the structural element. The larger the time scale is, the lower frequency of the noise has. These kinds of noises have ultra-low frequency and usually result in distortion at the frequencies below 0.01 Hz. The reduction of ultra-low frequency cultural noises in the MT signal needs further investigations.

Author statement

Li G proposed the idea, performed the experiments and wrote most of the manuscript. Liu XQ assisted in processing MT data and writing part of the manuscript. Tang JT and Li J are administrators of the projects and provided technical guidance. Ren ZY and Chen CJ assisted in analyzing MT data, and revised the manuscript. All authors read and approved the final manuscript.
Declarations of Competing Interest

The authors declare that they have no conflict of interests.

Acknowledgment

This work is financially supported by the National Natural Science Foundation of China (Nos. 41904076, 41904072, 41904073 and 41830107), Natural Science Foundation of Jiangxi Province (No. 20192BAB212009), Science and Technology Project of Jiangxi Provincial Education Department (No. GJJ180368) and National Key R&D Program of China (No. 2018YFC0603202). The authors acknowledge Professor Jianhua Cai for use of his codes and for helpful discussions. The authors also acknowledge the anonymous reviewers for their valuable comments helpful in improving this paper.

References