Separation of magnetotelluric signals based on refined composite multiscale dispersion entropy and orthogonal matching pursuit

Xian Zhang
Central South University School of Geosciences and Info Physics https://orcid.org/0000-0001-9156-9742

Jin Li (geologylj@163.com)
https://orcid.org/0000-0001-6606-6967

Diquan Li
Central South University School of Geosciences and Info Physics

Yong Li
Chinese Academy of Geological Sciences Institute of Geophysical and Geochemical Exploration

Bei Liu
Hunan University Art and Science College of Mathematics and Physics

Yanfang Hu
Central South University School of Geosciences and Info Physics

Full paper

Keywords: magnetotelluric (MT), refined composite multiscale dispersion entropy (RCMDE), orthogonal matching pursuit (OMP), noise separation

Posted Date: January 7th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-53570/v2

Version of Record: A version of this preprint was published on March 23rd, 2021. See the published version at https://doi.org/10.1186/s40623-021-01399-z.

Separation of magnetotelluric signals based on refined composite multiscale dispersion entropy and orthogonal matching pursuit

Xian Zhang [1,2,3,4], Jin Li [2,3,*], Diquan Li [1,4,*], Yong Li [3], Bei Liu [5], Yanfang Hu [1,4]

[1] Key Laboratory of Metallogenic Prediction of Nonferrous Metals and Geological Environment, Monitoring Ministry of Education, School of Geoscience and Info-Physics, Central South University, Changsha 410083, China.
[2] Hunan Provincial Key Laboratory of Intelligent Computing and Language Information Processing, College of Information Science and Engineering, Hunan Normal University, Changsha 410081, China.
[3] Key Laboratory of Geophysical Electromagnetic Probing Technologies of Ministry of Natural Resources, Institute of Geophysical and Geochemical Exploration, Chinese Academy of Geological Science, Langfang 065000, China.
[4] Hunan Key Laboratory of Nonferrous Resources and Geological Hazard Exploration, Changsha 410083, China.
[5] College of Mathematics and Physics, Hunan University Arts and Science, Changde 415000, China.

***Correspondence**: geologylj@163.com; lidiquan@csu.edu.cn

Abstract: Magnetotelluric (MT) data processing can increase the reliability of measured data. Traditional MT de-noising methods are usually filtered in entire MT time-series sequence, which result in losing of useful MT signals and the decrease of imaging accuracy of electromagnetic inversion. However, targeted MT noise separation can retain the part of data not affected by strong noise, and enhance the quality of MT data. Thus, we proposed a novel method for MT noise separation, which using refined composite multiscale dispersion entropy (RCMDE) and orthogonal matching pursuit (OMP). Firstly, the RCMDE characteristic parameters are extracted from each segment of the MT time-series. Then, the characteristic parameters are input to the fuzzy c-mean (FCM) clustering for automatic identification of MT signal and noise. Next, OMP method is utilized to remove the identified noise segments independently. Finally, the reconstructed signal consists of the denoised data segments and the identified useful signal segments. We conducted the simulation experiments and algorithm evaluation on the EMTF data, simulated data and measured sites. The results indicate that the RCMDE can improve the stability of multiscale dispersion entropy (MDE) and multiscale entropy (MSE) by analyzing the characteristics of the signal samples library, effectively dividing MT signals and noise. Compared with the existing techniques of the entire time domain de-noising and signal-noise identification, the proposed method used RCMDE and OMP as characteristic parameter and noise separation, simplified the multi-features fusion, and improved the accuracy of signal-noise identification. Moreover, the de-noising efficiency has accelerated, and the MT data quality of low-frequency band has improved greatly.

Keywords: magnetotelluric (MT); refined composite multiscale dispersion entropy (RCMDE); orthogonal matching pursuit (OMP); noise separation

Introduction

Magnetotelluric (MT) sounding is one of the most mature electrical exploration techniques in recent years (Tikhonov 1950; Cagniard 1953), which is mainly used to measure the orthogonal electric and magnetic field at the Earth's surface of the geoelectrical structures, mineral electrical

exploration, and electromagnetic fracture monitoring (Becher and sharpe 1969; Vallianatos 1996). Due to the wide frequency range of natural MT signal, it is vulnerable to interfere from an increasingly wide range of natural and artificial environmental noise. Therefore, effectively suppressing noise can improve the signal-noise ratio (SNR), and ensure the quality of MT data. However, the characteristics of MT signals are non-linear, non-stationary and non-minimum phase, which does not comply with the conditions of Fourier transform (Hermance 1973). Thus, strong interference of MT data will cause the excessive distortion of the apparent resistivity-phase curve and excessive concentration of phase angle in polarization direction. The high quality data obtained after de-noising will provide technical support for the following inversion interpretation (Qi et al. 2020; Li et al. 2020a).

In order to suppress the MT strong interference and improve the quality of MT data, numerous de-noising techniques have been applied. Remote reference method (Gamble et al. 1978; Gamble at al. 1979) and robust impedance estimation method (Egbert and booker 1986; Jones et al. 1984) are the most commonly used MT data processing method. Ritter uses indicators such as the transfer function between the magnetic field at the measuring sites and reference site to judge the noise of each data segment, and removes the noisy data segment, which does not participate in the next impedance estimation (Ritter et al. 1998). Varentsov proposed the "remote reference magnetic field control" criterion, which uses the magnetic field transfer function to control the impedance estimation values (Varentsov 2006). Although the remote reference method can eliminate the related noise, it relies on the selection of the reference sites, and the robust impedance estimation method frequently fails the effect of the correlated noise and persistent strong interference. The robust impedance estimation method uses the measured field values and theoretical values to estimate impedance, reducing the weight of the flight points and aligning the measured value with the estimated value, thereby achieving better impedance estimate the effect (Chave and Thomson 1989; Chave and Thomson 2004). Although the robust method can effectively reduce the dispersion of the apparent resistivity-phase curve and eliminate the non-Gaussian normal distribution noise in the MT data, the robust method is incapable of noise caused by the input, which cannot eliminate the near-source interference with strong energy. Some novel MT data processing methods have been applied to the MT noise suppression. For example, wavelet transform can effectively suppress the local electromagnetic related noise (Trad and Travassos 2000), which relies on the selection of the wavelet basis function. However, with the increase of the scales, the spectral localization of the corresponding orthogonal basis function deteriorates has limit the fine decomposition of MT signals. Hilbert-Huang transform (Huang et al. 1998) has been applied to the electrical method exploration, which can effectively suppress the MT signal with power frequency interference (Cai et al. 2009). Compared with wavelet transform, the choice of basis function has stronger time-frequency characterization capabilities. Mathematical morphological filtering can effectively suppress large-scale interference and baseline offset in the MT signal, and maintain the local characteristics of the target signal (Tang et al. 2012), but it difficult to select the types and sizes of structural elements. By analyzing the strong interference of the MT data, which mainly comes from artificial facilities with certain regular, such as power lines, engineering machinery and equipment. Wang et al. can treat the electric and magnetic field of time-series independently, which can replace windows of the noisy time-series. Besides, he proposed synthesis time-series method based on interstation transfer function, which eliminates the influence of anthropogenic noise (Wang et al. 2017). Variational mode decomposition (VMD) is a novel mode decomposition algorithm, which has been applied for MT noise suppression. Since the K value in the VMD algorithm can only be manually selected, Li et al. combined VMD with detrended fluctuation analysis (DFA) to adaptively select K value and the number of the reconstructed modals, improving the de-noising effect (Li et al. 2020b). By removing noise in time domain, the statistical analysis and time-series editing methods can directly and effectively improve the quality of MT data, but they could also damage the effective signal contained in the noisy segments.

Here, entropy shows the degree of chaos. It was originally used by physics to define the chaos degree of material, and later extended to thermodynamics and informatics to describe an event as the dynamic process of chaos. Entropy is now widely used, such as sample entropy

(Richman et al. 2004), fuzzy entropy (Kosko 1986) and approximate entropy (Pincus 1991), which can be used to assess the complexity of the system. On the basis of multi-scale analysis, multiscale entropy (MSE) is proposed to quantify the complexity of signals on multi-scales. Multiscale dispersion entropy (MDE) is a parameter to evaluate the multiple scales dynamic complexity of time-series (Zhang et al. 2018). The refined composite MDE (RCMDE) can increase the accuracy of entropy estimation and decrease the probability of facing situations where entropy is undefined (Azami et al. 2017). RCMDE reduces the sensitivity of MSE and MDE to the signal length for the time-series. It can combine the information of multiple coarse-grained sequences, reduce the standard deviation of entropy, and improve the stability of numerical results. The Fuzzy c-mean (FCM) clustering algorithm is an unsupervised method for data analysis and modeling, which is widely used in data classification and pattern recognition (Xu et al. 2019; Zhang et al. 2019). Through using the input features to generate the clustering center, calculating the Euclidean distance between the clustering points and the clustering centers, obtaining the membership degree of the clustering centers to divide the types of input features automatically. The sparse representation is to use as few atoms as possible to represent the signal in a given over-complete dictionary, so as to obtain a more concise representation of signal and acquire the contained information and process signal more conveniently. Based on matching pursuit (MP) algorithm(Mallat and Zhang 1993), the orthogonal matching pursuit (OMP) algorithm is a kind of the classic greedy algorithm (Pati et al. 1993). By using the Gram-Schmidt orthogonalization method to orthogonal the selected optimal atom and atomic set, so that the selected atom is orthogonal to the residual in each iteration, thereby accelerating the convergence speed of the algorithm. In view of the fact that the signals usually contain the stationary and non-stationary components, an atomic library, namely an over-complete dictionary (Cai and Wang 2011; Needell and Vershynin 2010), composed of the sine, cosine atoms and wavelet atoms are designed to realize the adaptive and accurate matching of signals.

In the paper, based on signal-noise inherent characteristics in the respective time-series, we used the RCMDE and OMP to realize the MT noise identification and separation. Among them, the RCMDE characteristic is associated with the FCM clustering for MT signal-noise identification, and using orthogonal matching pursuit (OMP) algorithm for MT noise separation. Firstly, we verified the stability of RCMDE and the de-noising effect of OMP. Then, the proposed method is applied to EMTF simulation and measured signal analysis. Compared with the remote reference method and the OMP-based overall method, the proposed method can purposefully remove the identified noise and retain the low-frequency useful MT signal. Compared with the fractal-entropy and clustering method (Li et al. 2018) to analyze the multiple characteristics of MT data, it makes the features extraction of MT data more precise and efficient of MT signal-noise identification. Moreover, the experimental results of apparent resistivity-phase curves and polarization direction at the measured sites demonstrated that the proposed method can approach to the natural MT field, and more truly reflect the underground electrical structural information.

Methods

In view of the long acquisition time of MT data, the degree of interference is not completely consistent. We search for the robust characteristic parameters to quantitatively identify the data segments that are not interfered with noise, and retain the MT useful signal from being over-processed. Only by eliminating the abnormal waveforms can the accuracy of MT noise suppression be improved and high-resolution MT signals can be obtained. Thus, compared with the MSE and MDE, the proposed method extracts RCMDE feature from time-series waveforms and inputs to FCM cluster analysis, then using OMP algorithm for noise separation. Simultaneously, the following respectively introduces RCMDE as unique feature extraction parameter and OMP as de-noising method.

Dispersion Entropy (DE)

Dispersion entropy (DE) is proposed by Rostaghi and Azami in 2016, which is a nonlinear dynamics method to characterize the complexity and irregularity of the time-series (Rostaghi and Azami 2016; Mitiche et al. 2018).

(1) Suppose time-series $x_j, \left(j=1,2,...,N\right)$ is mapped to c classes with integer indices from 1 to c. To realize the purpose, the normal cumulative distribution function (NCDF) maps x into $y=\left\{y_1, y_2, ...y_N\right\}$ from 0 to 1 as follows:

$$y_j = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x_j} e^{\frac{-(t-\mu)^2}{2\sigma^2}} dt \tag{1}$$

where σ and μ are the standard deviation (SD) and mean of time-series x, respectively. Then, a linear algorithm is employed to each y_i to an integer which ranges from 1 to c. For each member of the mapped signal, we use $z_j^c = round\left(c \times y_j + 0.5\right)$, where z_j^c denotes the j^{th} member of the classified time-series and rounding involves either increasing or decreasing a number to the next digit. Note that, although this part is linear, the entire mapping method is non-linear due to the use of the NCDF.

(2) Each embedding vector $z_i^{m,c}$ are made with embedding dimension m and time delay d, according to construct the time-series $z_i^{m,c} = \left\{z_i^c, z_{i+d}^c, ..., z_{i+(m-1)d}^c\right\}, i = 1,2,...,N-\left(m-1\right)d$ (Bandt and Pompe 2002) which is mapped to a dispersion pattern $\pi_{v_0 v_1 ... v_{m-1}}$, where $z_i^c = v_0$, $z_{i+d}^c = v_1, ..., z_{i+(m-1)d}^c = v_{m-1}$. The number of possible dispersion patterns that can be assigned to each time-series $z_i^{m,c}$ is equal to c^m, since the signal has m members and each member can be one of the integers from 1 to c.

(3) For each of c^m potential dispersion patterns, relative frequency is obtained as follows:

$$p\left(\pi_{v_0 ... v_{m-1}}\right) = \frac{Number\left\{i \middle| i \le N-\left(m-1\right)d, z_i^{m,c} \quad has \quad type \quad \pi_{v_0 v_1 ... v_{m-1}}\right\}}{N-\left(m-1\right)d} \tag{2}$$

where $p\left(\pi_{v_0 ... v_{m-1}}\right)$ denotes the number of dispersion patterns $\pi_{v_0 v_1 ... v_{m-1}}$ that are assigned to $z_i^{m,c}$, divided by the total number of embedded signals with embedding dimension m.

(4) The DE value is derived from the definition of Shannon's entropy and is defined as follows:

$$DE\left(x,m,c,d\right) = -\sum_{\pi=1}^{c^m} p\left(\pi_{v_0 v_1 ... v_{m-1}}\right) \cdot \ln\left(p\left(\pi_{v_0 v_1 ... v_{m-1}}\right)\right) \tag{3}$$

Refined Composite Multiscale Dispersion Entropy (RCMDE)

Multiscale dispersion entropy (MDE) is the combination of the coarse-graining (Costa et al. 2005) with DE, and then calculate the DE value of the coarse-graining sequence to obtain the DE at different scale. Instead, the mapping based on the normal cumulative distribution function (NCDF) used in the calculation of DE for the first temporal scale is maintained across all scales. RCMDE is improved MDE algorithm as follows:

For scale factor τ which has different time-series, corresponding to different starting points of the coarse-grained process are created, and the RCMDE value is defined as the Shannon entropy value of the average of the dispersion patterns of those shifted sequences (Azami et al. 2017). The k^{th} coarse-grained time-series $x_k^{(\tau)} = \left\{x_{k,1}^{(\tau)}, x_{k,2}^{(\tau)}, ...\right\}$ of u is as follows:

$$x_{k,j}^{(\tau)} = \frac{1}{\tau} \sum_{b=k+\tau(j-1)}^{k+\tau j-1} u_b, 1 \le j \le N, 1 \le k \le \tau \tag{4}$$

Then, for each scale factor, RCMDE is defined as follows:

$$RCMDE(x,m,c,d,\tau) = -\sum_{\pi=1}^{c^m} \overline{p}\left(\pi_{v_0 v_1 \ldots v_{m-1}}\right) \cdot \ln\left(\overline{p}\left(\pi_{v_0 v_1 \ldots v_{m-1}}\right)\right) \tag{5}$$

where $\overline{p}\left(\pi_{v_0 v_1 \ldots v_{m-1}}\right) = \frac{1}{\tau} \sum_1^{\tau} p_k^{(\tau)}$ with the relative frequency of the dispersion pattern π in the time-series $x_k^{(\tau)}\left(1 \le k \le \tau\right)$.

According to the entropy differences of the multiscale analysis in the sample library signal, which include the 150 sets of actual MT time-series (Li et al. 2018). Among them, 50 sets of actual MT time-series without interference are from the unmanned areas, and the remaining 100 sets of actual MT time-series (square wave interference and triangle wave interference) are collected from the strong electromagnetic interference areas. Thus, we predefined the parameter values (Rostaghi and Azami 2016), such as embedding dimension m is 2, number of classes c is 6, time lag d is 1, scale factor τ, that is corresponding to τ sets of characteristic parameter values in this paper.

Figure 1 shows the results of MSE, MDE and RCMDE using a set of sample library signals with different scale factor. The sample library signals are used for feature extraction and FCM clustering analysis to distinguish signal and noise.

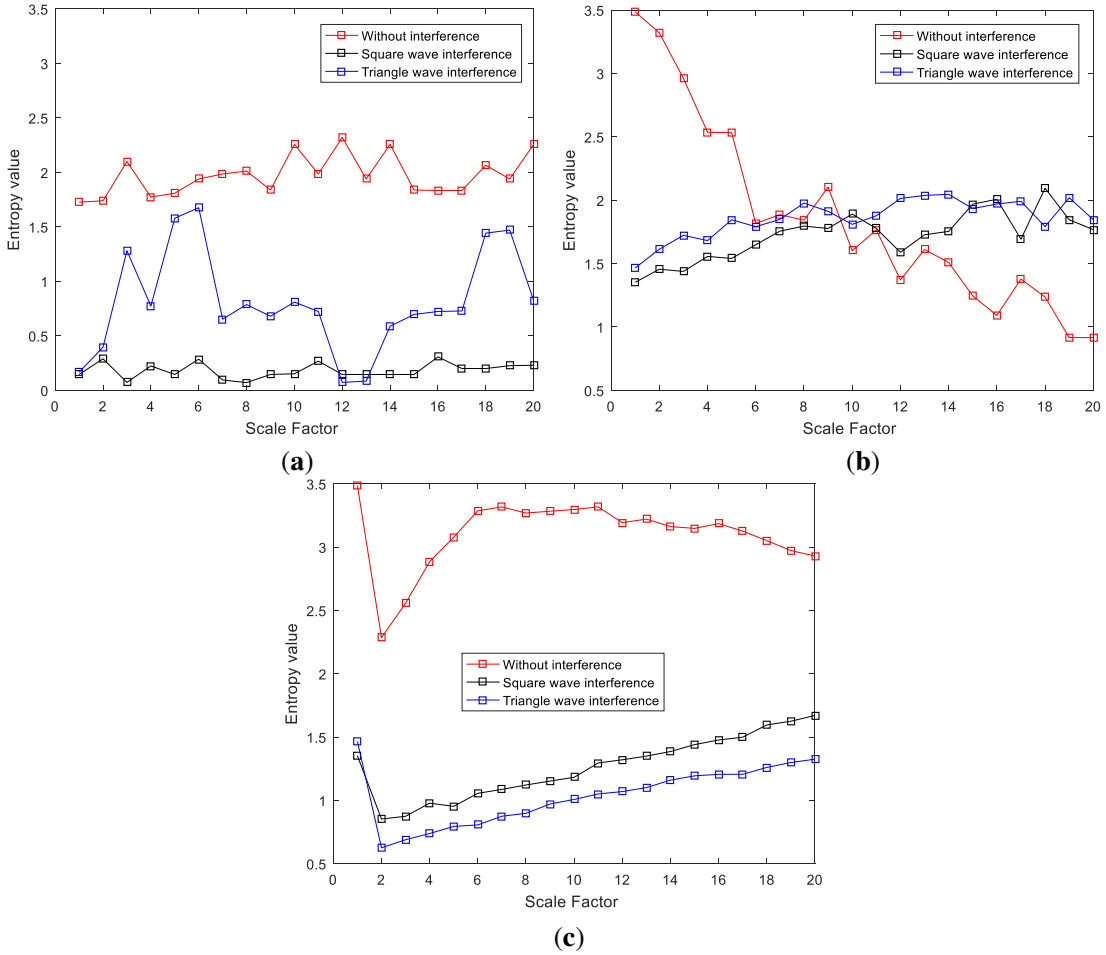


Fig.1 The results obtained for (**a**)MSE, (**b**)MDE and (**c**)RCMDE using a set of sample library signals at different scale factor, among them, the abscissa represents the scale factor τ, and the ordinate represents the entropy value at the corresponding scale.

We extract the characteristic parameters using only MT data without interference and MT data with interference (square wave interference and triangle wave interference). From Figure 1, a set of sample library signals have different entropy values at different scales. When the scale factor increased, the difference between MSE and MDE of a set of sample library signals is lower, crossover phenomenon and higher oscillations of the different interference curves, which are not conducive to classify noise and signal. However, the RCMDE shows the importance of the refined composite technique to improve the stability of results. Moreover, extracting appropriate characteristic parameters to describe the features of the MT data is helpful to improve the FCM clustering effect. As can be seen from Figure 1(c), when the scale factor is 1, the RCMDE feature value is the largest, and the single characteristic value will be meaningless, which cannot reflect the scale characteristic of the MT data. With the scale increases, these curves become more stable, the scale factor will determine the number of characteristic parameter values. The larger the scale factor is, the more characteristic parameters values will be generated, and the accuracy and efficiency of signal-noise identification of the massive measured data also will decline. In the paper, the RCMDE can generate multiple feature parameter values to represent MT signal-noise at different scale factor, while many characteristic parameters will reduce the clustering accuracy and consume more time. Thus, we only select the RCMDE as the feature vector analysis.

De-noising Method

Traditional matching pursuit (MP) algorithm at each iteration only ensures that matching residual data is orthogonal with an atom, which is prone to local optimization and result in a low matching accuracy and a large amount of computation (Huang and Makur 2011; Jin et al. 2014). The orthogonal matching pursuit (OMP) algorithm is based on MP algorithm, and also an improved version of MP by ensuring full backward orthogonality between the matching residual and the selected waveforms at each iteration, which ensures the optimal approximation with regard to all the selected subset of the dictionary after any finite number of iterations (Wang et al. 2013).

Given time-series x, length is N. $D=\left(g_\gamma\right)_{\gamma\in\Gamma}$ is an over-complete dictionary, that is the Fourier atomic library and wavelet atomic library, which for signal sparse decomposition. g_γ is the γ_{th} atom in the dictionary set Γ, and $\left\|g_\gamma\right\|=1$.

Initialized signal residual $R^0=x$, and reconstructed signal $\overline{x}_m=0$, we selected the atom set ψ_0 is empty set, the number of iterations $m=1$, the maximum number of iterations M. When the iteration stop condition is not met, repeat the following steps (Tropp and Gilbert 2007).

(1) We select the atom g_γ that most closely matches the analysis signal f from the dictionary. They meet the following conditions:

$$\left|\left\langle f,g_\gamma^m\right\rangle\right|=\sup_{\gamma\in\Gamma}\left|\left\langle f,g_\gamma^m\right\rangle\right| \tag{6}$$

(2) Update the selected atom set $\psi_m=\psi_{m-1}\cup\left\{g_\gamma^m\right\}$.

(3) Find the projection coefficient according to the least square method $u_m=\left(\psi_m^T\psi_m\right)^{-1}\cdot\psi_m^T x$. Thus, the reconstructed signal is $\overline{x}_m=\psi_m x_m$, and residual signal is $R^m=x-\overline{x}$.

(4) Update the number of iterations $m=m+1$. Judge whether the energy ratio $\left\|R^m\right\|_2/\left\|x\right\|_2$ of the residual signal to the original signal is less than the given value. If not satisfied, return to step (1). If satisfied, the reconstructed signal is obtained $x=\overline{x}_m$ and residual signal is $R=R^m$.

Figure 2 shows the de-noising effect of MP and OMP methods for the noisy signal with 20dB white Gaussian noise, and the length of the noisy signal is 2048. Among them, the mentioned atomic library consists of sine (sin) and discrete cosine (dct) atoms, and the symlets (sym) and daubechies (db) wavelet atoms.

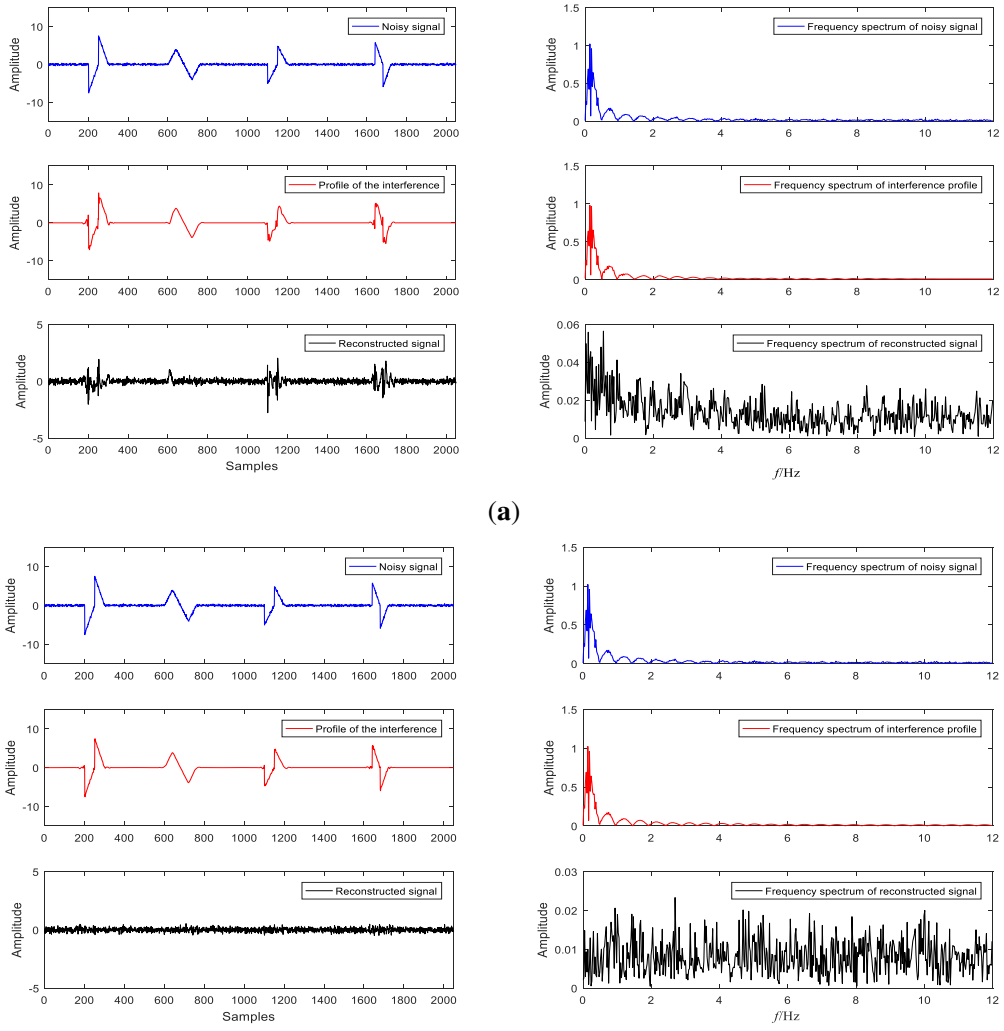


Fig. 2 The de-noising effect and frequency spectrum analysis of the noisy signal with (**a**) matching pursuit (MP) and (**b**) orthogonal matching pursuit (OMP).

To estimate the de-noising effects of MP and OMP methods for the above noisy signal, the normalized cross-correlation (NCC), signal-noise ratio (SNR), mean square error (MSE) and the time factor (T) were used for the quantitative analysis. Definitions of these parameters are as refer to Li et al. 2020b.

Table 1 is the comparison between the de-noising performance of MP and OMP methods.

Table 1 The de-noising performance of MP and OMP methods.

Methods	**NCC**	**SNR(dB)**	**MSE**	**T(s)**
MP	0.9848	15.1976	0.0651	2.45
OMP	0.9989	26.3956	0.0049	0.59

As can be seen from Figure 2 and Table 1, by comparing the de-noising effects of MP and OMP methods with the same number of atoms and iterations, we found that the MP method still has residual noise in the reconstructed signal, while the OMP method shows the excellent characteristics in the NCC, SNR, MSE and Time, and its frequency spectrum also strictly corresponds to the left signal. Therefore, OMP method ensures that the residue is orthogonal to the all selected atoms, and converge speed is faster than MP method. The OMP method is more suitable for subsequent MT noise separation, and also requires much less memory and computation time.

Experiments

Algorithm steps

The algorithm steps of the proposed method are as follows:

Step 1: Input MT data, that is using 240 as the equal intervals for the MT data segmentation;

Step 2: Extracting RCMDE feature with scale factor 2 of each segment MT data, that is, each segment data has two feature parameter values;

Step 3: The all feature parameter values input into the FCM clustering, and automatically identify MT signal and noise, that is, one type is MT signal, the other is noise;

Step 4: The part identified as MT signal should be retained, and identified as noise is targeted denoised by OMP method;

Step 5: Reconstructed MT data consists of the superposition of the retained MT signal and the denoised MT signal.

Clustering Analysis of the Sample Library

In this section, we applied the FCM clustering analysis for sample library signals through feature extraction based on RCMDE with scale factor 2. The FCM clustering algorithm obtains the membership degree of each sample to all clustering centers by optimizing the objective function, thereby determining the type of the samples to achieve the purpose of automatic classification.

Figure 3 is the FCM clustering effect of library signals.

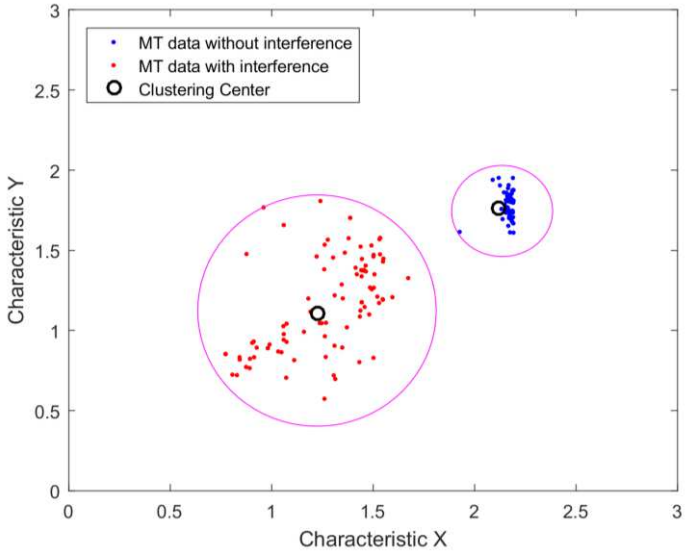


Fig. 3 FCM clustering effect of sample library signals. Among them, characteristic X and Y in the coordinate axis represent the refined composite multiscale dispersion entropy (RCMDE) value when the scale factor is 1 and 2, respectively.

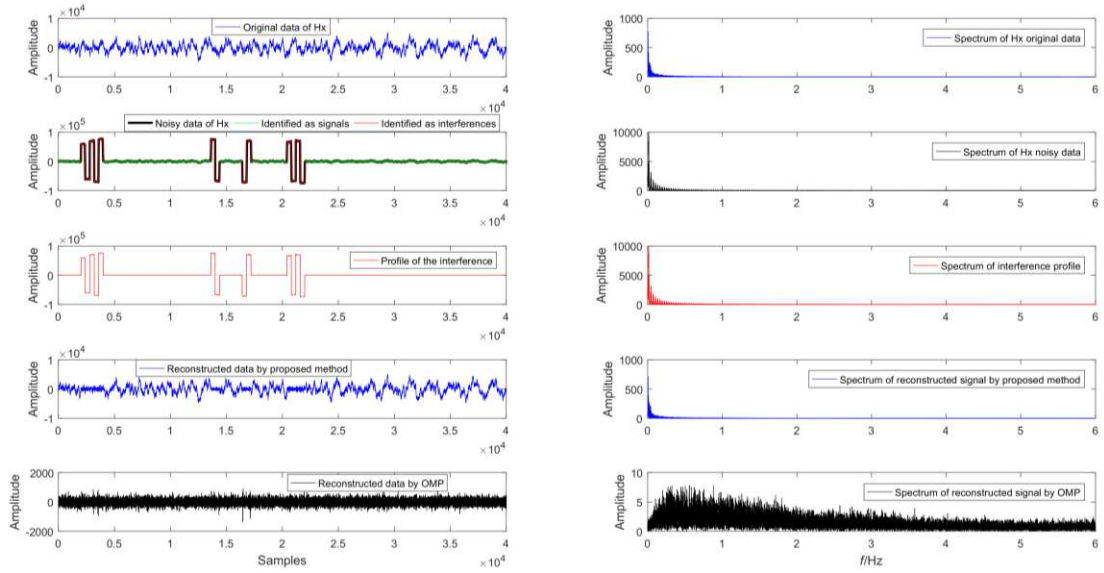
We calculate the Euclidean distance from each sample point to the cluster center, and select the length of the farthest distance point as the radius of the pink circle, including all the longest and

shortest distance sample points in same type of sample point. Figure 3 shows the FCM clustering algorithm effect that divides MT data without interference and with interference. Specially, two pink circles indicate the shortest and longest distance from the samples to the cluster center in the same type, which can effectively divide the sample signals into different groups, thereby accurately identifying the MT signals without or with interference. According to verify with the FCM clustering effect, MT data without interference are classified, that is represented by blue points. The MT data with interference are classified which can be purposefully removed.

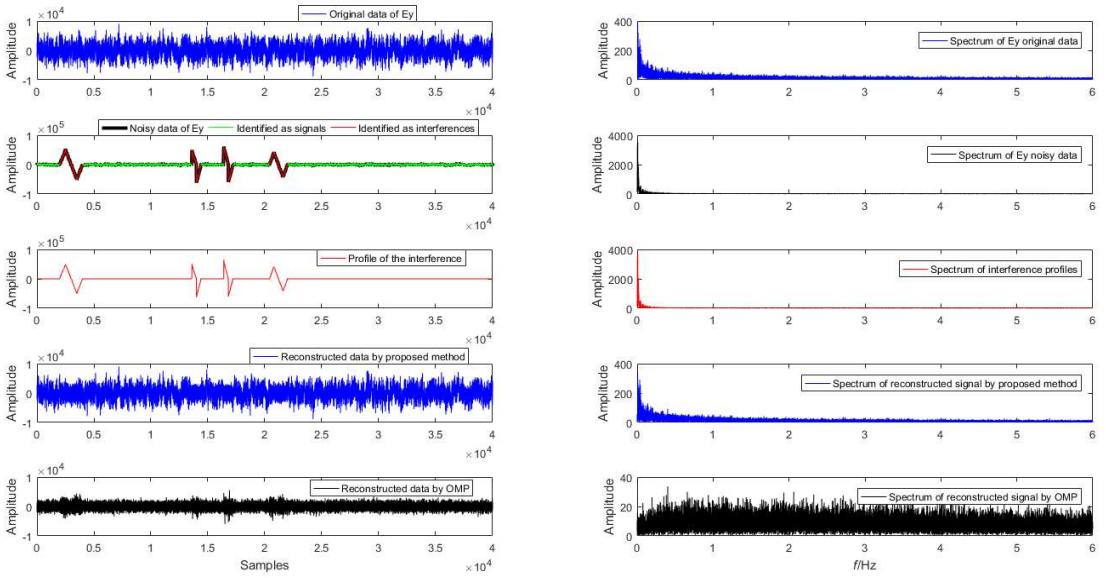
Numerical Simulation Analysis of the EMTF Data

The numerical simulation analysis is based on the open source software EMTF, which is an open source code for the MT single-site robust estimation and remote reference analysis written by Eisel and Egbert (Eisel and Egbert 2001). The open source code package provides two 100 $\Omega \cdot m$ uniform half-space time-series data (test1.asc and test2.asc). Each time-series data has five columns of data with a length of 40000 and sample rate is 1Hz. The five columns of data represent the x, y, z direction of the magnetic field and the x, y direction of the electric field for the observed data. The correlation between the signals of the two time-series data provided by the open source package is close to 1 (Egbert and livelybrooks 1996). Furthermore, typical strong interference is added to the original signal of EMTF data (test1.asc) to assess the performance of the OMP method and the proposed method on the time-series.

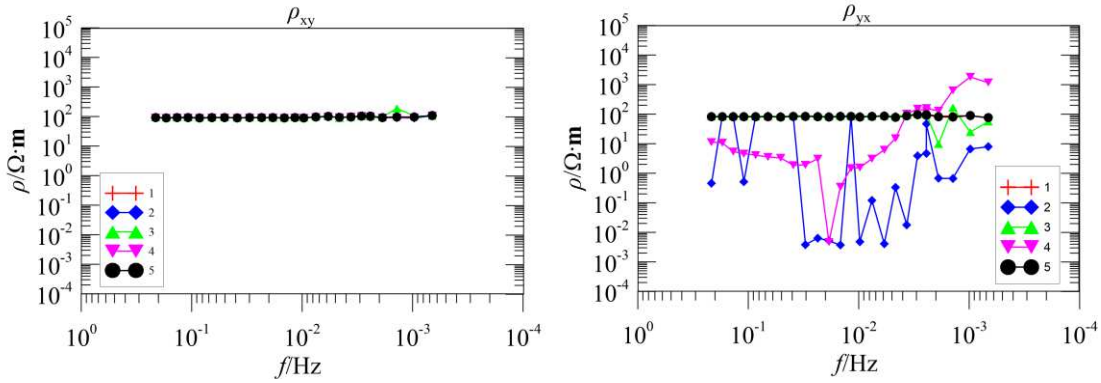
Figure 4(a) and (b) show the addition of square wave interference on the original time-series of Hx, and the addition of the triangle wave interference on the original time-series of Ey of the EMTF data for signal-noise separation. At the same time, Figure 4(c) shows the comparison of the apparent resistivity curves obtained by the noisy data, the remote reference method, OMP-based overall method and the proposed method.



(**a**)



(**b**)



(**c**)

Fig. 4 The EMTF data is disturbed by a noisy signal with (**a**) square wave interference of Hx and (**b**) charge and discharge triangle wave interference of Ey; (**c**) is a comparison of the apparent resistivity-phase curves, left figure curve is ρ_{xy} and right figure curve is ρ_{yx}. Among them, curve 1 is original data, curve 2 is noisy data, curve 3 is remote reference method, curve 4 is OMP-based overall method, and curve 5 is the proposed method.

Considering that Ey and Hx data affects the change of the ρ_{yx} curve, we add the strong interference to the same position for Ey and Hx data, that is, the strong interference is added at the same time period. From Figure 4(a) and 4(b), it can be known that the interference and signals are accurately identified, and the proposed method can extract the profile of the interference. By comparing the original data (Hx and Ey) with the reconstructed data (Hx and Ey) of the proposed method, the NCC is 0.9524 and 0.9683 respectively, and the frequency spectrum of original data and reconstructed data are also very close. While the NCC is respectively 0.1162 and 0.4325 on the reconstructed data by OMP method, low-frequency useful signal is basically lost. From Figure 4(c), the apparent resistivity curve 1 of the original data is smooth. Curve 2 is obtained by adding noise to the original data (Hx and Ey), because of the noise are added to the Hx and Ey channel data, it only

causes the apparent resistivity curve of ρ_{yx} appears a variety of change. While curve 3 is obtained from the remote reference method, and the EMTF time-series data (test2.asc) as reference data is used suppress noise-added data. Since the noise amplitude is high and added to the relevant noisy data, this method makes a few frequency jumps in the low frequency of ρ_{yx}, which cannot achieve an ideal effect. Curve 4 is obtained by the OMP-based overall method that can suppress the strong interference in the time-series, which filters the entire data blindly, causing to the loss of the useful data, and the entire apparent resistivity curves still disorder. According to the time-series of signal-noise identification and separation, through NCC analysis and the apparent resistivity curves, the shape and smoothness of the apparent resistivity curves obtained by the proposed method that is the curve 5 is close to the original data. Thus, the proposed method is more suitable for subsequent measured MT noise separation.

Noise Separation of the Measured MT Data

To verify the effectiveness of the proposed method, Figure 5 shows the signal-noise identification and targeted de-noising effect by comparing with the OMP-based overall method and the proposed method for the MT signal with square wave and triangle wave interference.

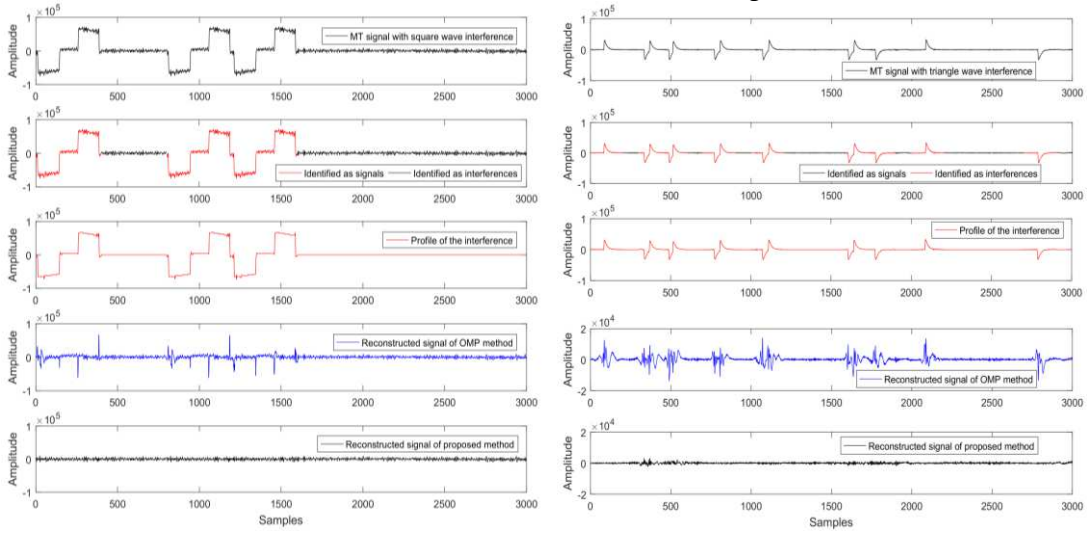


Fig. 5 The signal-noise identification and targeted de-noising for the measured MT data, and compared with the OMP method-based overall method and the proposed method, (a) square wave interference and (b) triangle wave interference.

In the case of equivalent atoms and the number of iterations, the OMP-based overall method can remove the large-scale strong interference, but some interference still remains. Thus, the useful MT signal segment is also filtered, resulting in the reconstructed signal cannot be close to the real MT signal. The proposed method carries out signal-noise identification and noise separation, which improves the authenticity of the reconstructed signal, thereby avoiding the under-processing and over-processing of the OMP method.

Results

Apparent Resistivity-Phase Curve Analysis of the Measured Sites

In this section, we compared with the apparent resistivity-phase curves of the original data, the remote reference (RR) method, the OMP-based overall method, the fractal-entropy and clustering method (Li et al. 2018) and the proposed method. These measured sites (D37890, EL22189 and EL22174) were affected from the square wave and triangle wave interference types in time-series sequence.

Figure 6 shows the comparison of the apparent resistivity-phase curves of the measured MT site D37890.

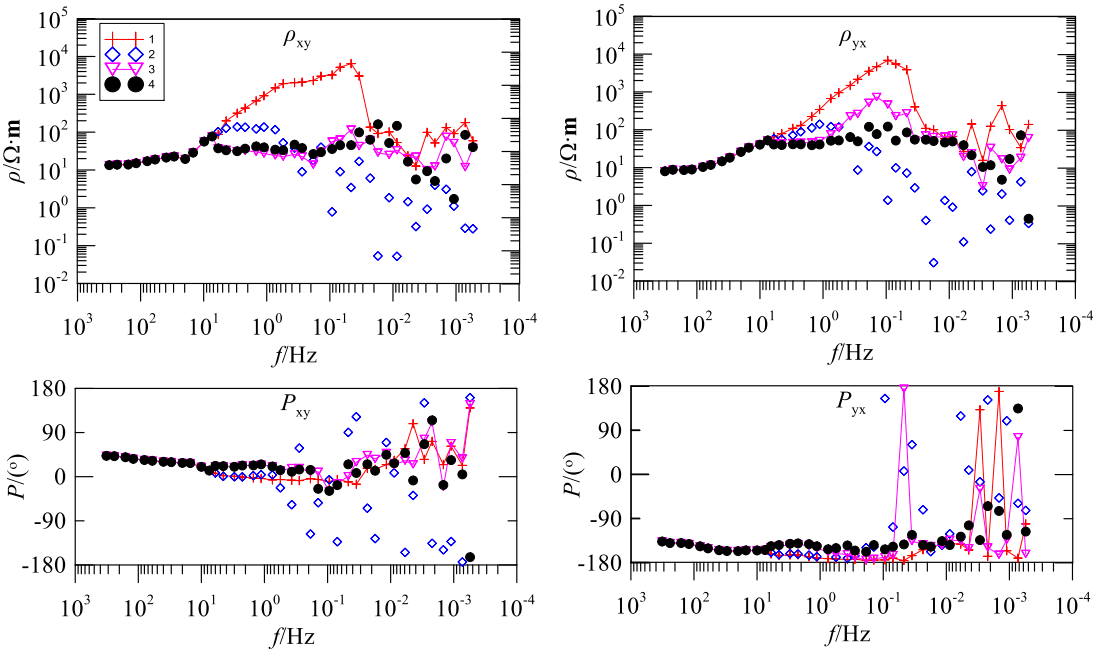


Fig. 6 Comparison of the apparent resistivity-phase curves of the measured MT site D37890, among them, curve 1 is original data, curve 2 is the data filtered by the OMP method-based overall method, curve 3 is the result derived from the fractal-entropy and clustering method, and curve 4 is the result derived from the proposed method.

In Figure 6, the apparent resistivity-phase curve of the original data (curve 1) gradually rises to 10^4 $\Omega \cdot m$ at 10-0.03 Hz, and the phase curve of the corresponding frequency band is attached to 0 °. The curve presents a downward trend at 0.03 Hz, and its corresponding phase jump becomes larger. Therefore, the data quality is obviously poor, and shows the near-source effect. By observing the time-series sequence of original data, the strong interference almost submerges the entire low-frequency useful MT signal. As a result, it cannot characterize the underground electrical structure information.

According to the data filtered by the OMP method-based overall method, the entire low-frequency band of the curve 2 has been declining, and the corresponding phase curve is seriously disturbed. Although the de-noising result of time-series waveforms shows that large-scale strong interference is also eliminated, it lacks the retention of useful signals and incomplete removal of complex interference, resulting in the failure to provide valid data. The fractal-entropy and clustering method (curve 3) extracts four types of the feature parameters, and has a relatively stable discrimination for the MT signal-noise. Then, the traditional wavelet is used to carry out targeted de-noising. However, the curve 3 in the ρ_{yx} of 3-0.03 Hz has not been alleviated, which shows a convex state. The difference of characteristic parameters has led to the inability of identify MT signal and noise in this frequency band. Moreover, this method still takes longer time to calculate the feature parameters, and needs to predefine different wavelet bases and decomposition layers for wavelet de-noising. The proposed method (curve 4) can purposefully suppress the identified as the MT interference segments and preserve the useful MT signals segments, and reconstructed MT signals is close to the natural field MT data. The apparent resistivity curve becomes stable and continuous, and the phase curve has also been significantly

improved. The result indicated that the denoised data by the proposed method can be more truly reflect the underground electrical structure information.

Figure 7 shows the comparison of the apparent resistivity-phase curves of the measured AMT site (EL22189 and EL22174).

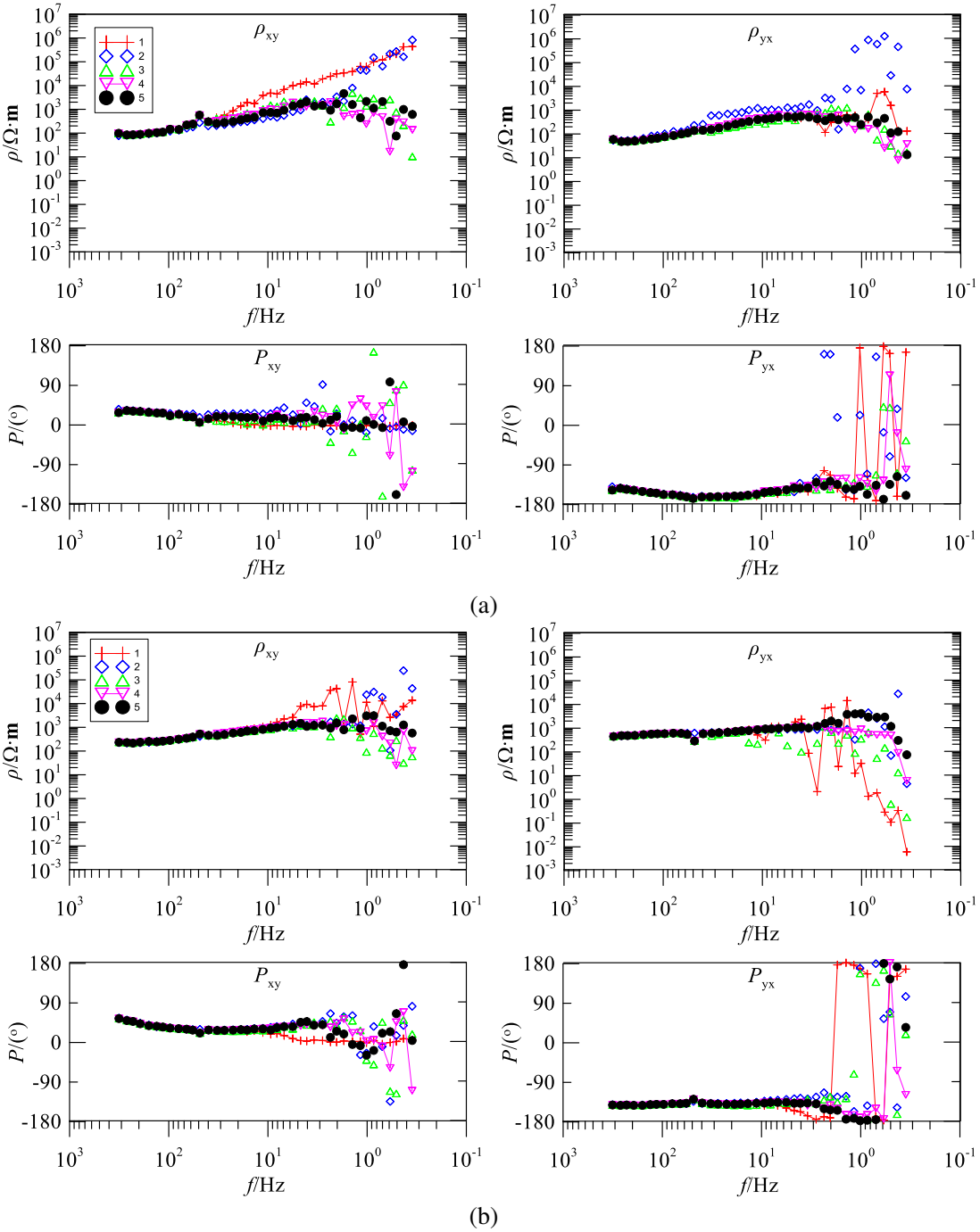


Fig. 7 Comparison of the apparent resistivity-phase curves of the measured AMT site (a) EL22189 and (b) EL22174, among them, curve 1 is original data, curve 2 is the result obtained by the remote reference (RR) method, curve 3 is the data filtered by the OMP method-based overall method, curve 4 is the result derived from the fractal-entropy and clustering method, and curve 5 is the result derived from the proposed method.

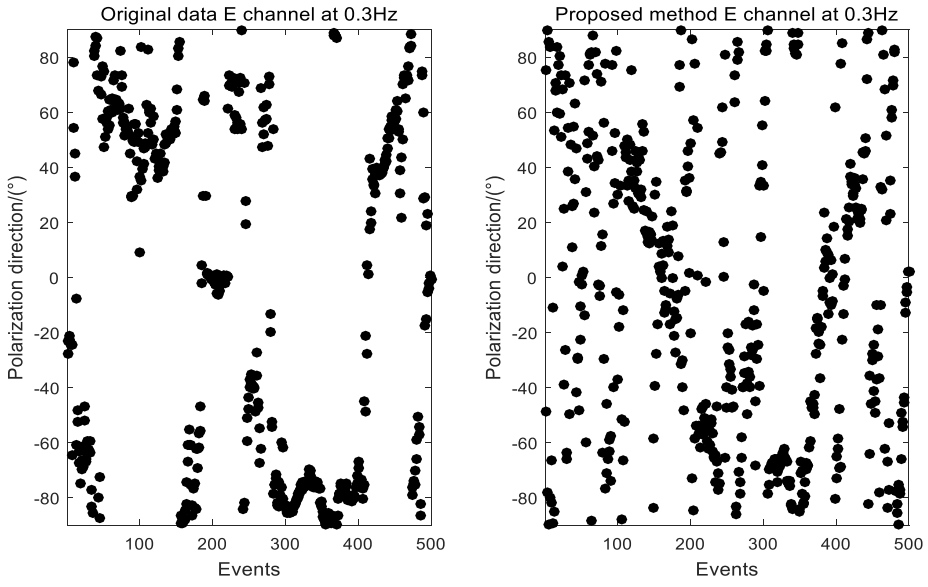
In Figure 7 (a), the apparent resistivity curve ρ_{xy} keeps rising and curve ρ_{yx} has several frequency fluctuations in the low frequency band. In Figure 7 (b), the apparent resistivity curve rises

in the ρ_{xy} and drops in the ρ_{yx}, and the apparent resistivity curve becomes confusion under the 3 Hz frequency band. These corresponding phase curves are based on 0° and ±180°. These results show that the data quality of the low frequency band is seriously reduced.

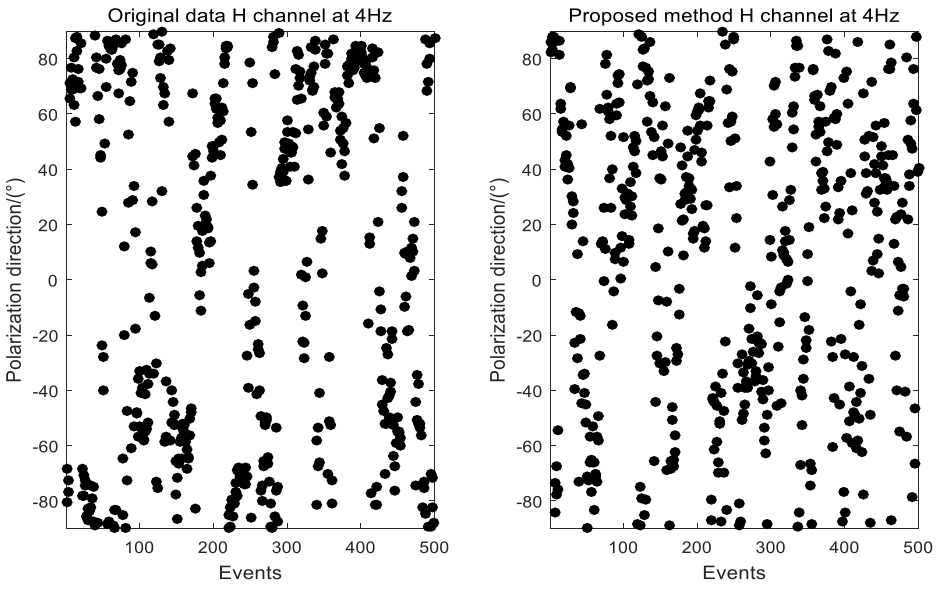
Compared to the curve 2 for the result obtained by the remote reference method, the apparent resistivity-phase curves become smoother, there is still frequency fluctuation at 3-0.3Hz in the low frequency band. Since the remote reference method depends on the selection of the reference site and the measured site distance, it is difficult to achieve the desired effect. Curve 3 is the data filtered by the OMP method-based overall method that caused to the low-frequency band of the apparent resistivity curve fall seriously, and the corresponding phase curve is more chaotic, the reason is that the time-series has been over-processed and useful MT signals are lost. Meanwhile, the result cannot provide the reasonable AMT data. Curve 4 is the result derived from the fractal-entropy and clustering method which obtained the smoother of the entire apparent resistivity curve (Li et al. 2018). Since four characteristic parameters describe the AMT signal and noise at different angles, and the identification of weak noise as a useful signal and a large number of human factors, resulted in the low of AMT data quality in the low frequency band. The proposed method (curve 5) can remove the segment identified as AMT interference, and the denoised data combine with segment identified as AMT signals for the precisely reconstructed data. Moreover, the apparent resistivity-phase curve is more continuous and smoother. Thus, the result can provide more reliable and reasonable AMT data for subsequent electromagnetic inversion.

Polarization Direction Analysis

According to the definition of polarization direction of electromagnetic field (Weckmann et al. 2005), it is also an indicator to evaluate the quality of MT or AMT data, and introduced to verify the effectiveness of the proposed method. Figure 8 shows the comparison of electromagnetic polarization direction at 0.3Hz and 4Hz for the measured site EL22174 in the electric and magnetic fields, respectively.



(**a**)



(**b**)

Fig. 8 Comparison of the polarization direction for site EL22174, (**a**) electric field data at 0.3 Hz, and (**b**) magnetic field data at 4 Hz.

From Figure 8, the polarization direction of the electric field is derived from the original data, which is concentrated on -80° and 50° at 0.3 Hz. The polarization direction of the original data in the magnetic field is concentrated on -60° and 80° at 4 Hz. Because the time-series of original data contained persistent strong interference, analyzing the polarization direction of the proposed method, the randomness of the polarization direction of the electromagnetic field increases, and the polarization angle is disordered, which is consistent with the polarization characteristics of the natural field. Combined with the apparent resistivity-phase curve, the proposed method can purposefully suppress the strong interference and the reconstructed data is close to natural MT field data.

Discussions

Since the theory of MT sounding was proposed, the problem of noise has been bothering the majority of MT researchers. Typical noise is often full of time domain waveforms, which affects the degradation of data quality in the low frequency band. The apparent resistivity is chaotic and discontinuous, and the polarization direction is highly concentrated in a certain direction. Therefore, the proposed method purposefully eliminates MT strong interference, and improves the quality of low-frequency data in the final results.

Based on the existing technology in the MT noise suppression, the reliable MT sounding data can be obtained through editing, filtering and signal-noise identification in the time-frequency domain. These techniques can provide effective data for the geological exploration and interpretation. The signal-noise identification techniques have also been continuously derived. By extracting the characteristic parameters of MT data for clustering and classification, identifying signals and noise, while the information of low-frequency slow-change of MT is kept abundantly and improved the data quality. At present, the signal-noise identification method has extracted a large number of the feature parameters and analyzed the essential characteristics of MT data from different angles, such as recurrence analysis, sample entropy, fuzzy entropy, approximate entropy, LZ complexity, fractal dimension, etc. These characteristic parameters are not scale features of the analysis, only through the fusion of a large number of characteristic parameters to judge whether a segment of MT signal is affected by noise, the multiple characteristic parameters are differ greatly, resulting in the reduction of the accuracy and efficiency of clustering and classification. At the

same time, when the MT signal-noise is gradually blurred, the ability of feature recognition is also ineffective, and the apparent resistivity curve shows a downward trend in the low frequency band, which cannot accurately interpret the underground electrical structure information.

In this study, when the time scale is uncertain, MSE will extend the sample entropy to multiple time scales, so as to provide additional observation angles. MSE involves the process of coarse-graining time-series, which is mainly used to analyze time-series with increasingly coarse time resolution. The process of MDE is similar to MSE, which calculating DE value at different scales. Meanwhile, RCMDE is an improvement of MDE algorithm, using refined composite technology to obtain the better consistency, stability of entropy value and faster calculation speed. As shown in Figure 1, the stability of RCMDE is higher than other features, and using a scale factor 2 for FCM clustering can improve the accuracy and efficiency. In the MP algorithm, if the vertical projection of the signal (that is residual signal) on the selected atoms is non-orthogonal, it will trim the result of each iteration, and requires many iterations to converge. The OMP algorithm means that all the selected atoms are orthogonal at each step of the decomposition, which makes the convergence speed of OMP algorithm faster under the requirement of the same accuracy. With the same number of atoms and iterations in Figure 2, the OMP algorithm is superior to MP algorithm in the signal-noise separation, and can obtain the high-quality of the denoised signal more quickly. Therefore, we mined the robust characteristic RCMDE for FCM clustering analysis (Figure 3), and divided the MT signal and noise through its scale characteristics with high precision, and used OMP algorithm to separate noise. In order to verify the feasibility of the proposed method, and experimental results based on the different signal are discussed (Figure 2-Figure 5). The proposed feature parameter was compared with MSE and MDE, and the denoising method was compared with MP. The proposed method improves the stability of MDE and the efficiency of MT signal noise identification. Besides, two feature parameter values are generated by RCMDE characteristic with the scale factor 2. Using simple characteristic parameter for FCM clustering and rapid de-noising method are applied to EMTF simulation data and measured MT data, which can improve the effect and efficiency of the existing technology in processing MT noisy data as shown in Figure 4 and 5. By comparing the remote reference method, OMP-based overall de-noising method and signal-noise identification and separation method, it can be seen from the Figure 6 and 7 that the proposed method effectively improves the multiple frequency information in the low frequency band, and the entire low-frequency curve becomes smoother and more stable. The polarization direction is shown in Figure 8, which further illustrates the poor quality of the original data, and the obtained data by proposed method is closer to the real MT data.

In short, many MT signal processing methods are from the perspective of time-frequency domain to eliminate the MT strong interference. The common purpose is to obtain high-quality MT data and provide reliable data for subsequent electromagnetic inversion. However, the pulse interference exists in a sudden change, it is impossible to accurately identify some pulse interference. The limitation of the proposed method is that the noise types are complex and the length of noisy data cannot be determined, which leads to inaccurate segmentation and recognition effect. When the signal and noise were gradually blurred, the clustering precision will also drop sharply. In addition, it is still to be studied how to adaptively segment to keep more useful signals from being over-recognized, and refine the multiscale analysis of characteristic parameters, how to optimize the atomic matching in the de-noising method. Besides, the proposed method can further research the preset parameters values or parameters selection will improve MT data processing efficiency and accuracy, and provide richer technical support for the massive MT data, which will inevitably improve the universality of the methods.

Conclusions

We have proposed a novel noise separation method for MT data by using RCMDE and OMP. As a robust feature parameter, RCMDE can generate multiple feature parameter values to distinguish MT signal and noise. OMP is used as a rapid de-noising method. Combined RCMDE and OMP, we improved the efficiency and accuracy of MT feature extraction, identification and

noise separation. The experimental results show that the identified strong interference are purposefully eliminated, the useful MT signals are bounteously reserved, and the quality of MT data is improved. The apparent resistivity-phase curve by using the proposed method becomes more continuous and smooth, and the polarization direction becomes more scattered and random. It will further provide an innovative technology route for MT data processing, and obtain high-precision MT data for subsequent electromagnetic inversion.

Abbreviations

MT: Magnetotelluric; RCMDE: Refined composite multiscale dispersion entropy; MP: matching pursuit; OMP: Orthogonal MP; VMD: Variational mode decomposition; SNR: Signal-noise ratio; MSE: Mean square error; NCC: Normalized cross-correlation; T: Time factor; SE: Sample entropy; FE: Fuzzy entropy; AE: Approximate entropy; DE: Dispersion entropy; RR: Remote reference; FCM: Fuzzy c-mean; MSE: Multiscale entropy; MDE: Multiscale dispersion entropy; DFA: detrended fluctuation analysis; SD: Standard deviation; NCDF: Normal cumulative distribution function.

Authors' contributions

Zhang X wrote the manuscript, designed and analyzed the experiments; Li J and Li D Q conceived the idea and helped revise the manuscript; Li Y helped analyze the experimental results; Liu B and Hu Y F helped discuss the algorithm and experiment. All authors have read, revised, and approved the final manuscript.

Acknowledgements

The authors would like to thank the anonymous reviewers for providing critical comments that improve the manuscript greatly. This research was supported by the National Key R&D Program of China (No.2018YFC0807802, No.2018YFE0208300), the National Natural Science Foundation of China (No.42074084, No.41874081), the Open Research Fund Program of Key Laboratory of Metallogenic Prediction of Nonferrous Metals and Geological Environment Monitoring (Central South University), Ministry of Education (No.2020YSJS06), the Key Laboratory of Geophysical Electromagnetic Probing Technologies of Ministry of Natural Resources (No.KLGEPT201905), and the Natural Science Foundation of Hunan Province (No.2018JJ2258).

Availability of data and materials

The datasets and MATLAB code used during the current study are available from the corresponding authors on reasonable request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Conflicts of interest

The authors declare no conflict of interest.

Author details

[1] Key Laboratory of Metallogenic Prediction of Nonferrous Metals and Geological Environment, Monitoring Ministry of Education, School of Geoscience and Info-Physics, Central South University, Changsha 410083, China. [2] Hunan Provincial Key Laboratory of Intelligent Computing and Language Information Processing, College of Information Science and Engineering, Hunan Normal University, Changsha 410081, China. [3] Key Laboratory of Geophysical Electromagnetic Probing Technologies of Ministry of Natural Resources, Institute of Geophysical and Geochemical Exploration, Chinese Academy of Geological Science, Langfang 065000, China. [4] Hunan Key Laboratory of Nonferrous Resources and Geological Hazard Exploration, Changsha 410083, China.

[5] College of Mathematics and Physics, Hunan University Arts and Science, Changde 415000, China.

References

Azami H, Rostaghi M, Abásolo D, Escudero J (2017) Refined composite multiscale dispersion entropy and its application to biomedical signals. IEEE T Bio-Med Eng 64: 2872–2879.

Bandt C, Pompe B (2002) Permutation entropy: a natural complexity measure for time series. Phys Rev let 88: 174102.

Becher WD, Sharpe CB (1969) A synthesis approach to magnetotelluric exploration. Radio Sci 4(11): 1089–1094.

Cagniard L (1953) Basic theory of the magnetotelluric method of geophysical prospecting. Geophysics 18(3): 605–635.

Cai JH, Tang JT, Hua XR, Gong YR (2009) An analysis method for magnetotelluric data based on the Hilbert-Huang transform. Explor Geophys 40(2): 197–205.

Cai T T, Wang L (2011) orthogonal matching pursuit for sparse signal recovery with noise. IEEE T Inform theory 57(7):4680–4688.

Chave AD, Thomson DJ (1989) Some comments on magnetotelluric response function estimation. J Geophys Res 94(10): 14215–14225.

Chave AD, Thomson DJ (2004) Bounded influence magnetotelluric response function estimation. Geophys J Int 157(3): 988–1006.

Costa M, Goldberger AL, Peng CK (2005) Multiscale entropy analysis of biological signals. Phys Rev E 71: 021906.

Egbert GD, Booker JR (1986) Robust estimation of geomagnetic transfer functions. Geophys J Roy Astr Soc 87(1): 173–194.

Egbert GD, Livelybrooks DW (1996) Single station magnetotelluric impedance estimation: coherence weighting and the regression M-estimate. Geophysics 61(4): 964–970.

Eisel M, Egbert GD (2001) On the stability of magnetotelluric transfer function estimates and the reliability of their variances. Geophys J Int 144: 65–82.

Gamble TM, Goubau WM, Clarke J (1978) Magnetotelluric data analysis: removal of bias. Geophysics 43(6): 1157–1169.

Gamble TM, Goubau WM, Clarke J (1979) Magnetotelluric with a remote magnetic reference. Geophysics 44(1): 53–68.

Hermance JF (1973) Processing of magnetotelluric data. Phys Earth Planel Interiors 7(3): 349–364.

Huang H, Makur A (2011) Backtracking-based matching pursuit method for sparse signal reconstruction. IEEE Signal Proc Let 18(7):391–394.

Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, Yen NC, Tung CC, Liu HH (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc A Math Phys Eng Sci 454: 903–995.

Jin W, Wang L, Zeng X, Liu Z, Fu R (2014) Classification of clouds in satellite imagery using over-complete dictionary via sparse representation. Pattern Recogn Lett 49(1): 193–200.

Jones AG, Chave AD, Egbert GD, Auld D, Bahr K (1989) A comparison of techniques for magnetotelluric impedance estimation. J Geophys Res 94(10): 14201–14213.

Kosko B (1986) Fuzzy entropy and conditioning. Inform. Sciences. 40(2): 165–174.

Li G, Liu XQ, Tang JT, Deng JZ, Hu SG, Zhou C, Chen CJ, Tang WW (2020a) Improved shift-invariant sparse coding for noise attenuation of magnetotelluric data. Earth Planets Space 72: 45.

Li J, Zhang X, Gong JZ, Tang JT, Ren ZY, Li G, Deng YL, Cai J (2018) Signal-noise identification of magnetotelluric signals using fractal-entropy and clustering algorithm for targeted de-noising. Fractals 26(2): 1840011.

Li J, Zhang X, Tang JT (2020b) Noise suppression for magnetotelluric using variational mode decomposition and detrended fluctuation analysis. J Appl Geophys 180: 104127.

Mallat SG, Zhang Z (1993) Matching pursuit with time-frequency dictionaries. IEEE Trans Signal Process 41(12): 3397–3415.

Mitiche I, Morison G, Nesbitt A, Hughes-Narborough M, Stewart BG, Boreham P (2018) Classification of partial discharge signals by combining adaptive local iterative filtering and entropy features. Sensors 18: 406.

Needell D, Vershynin R (2010) Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit. IEEE J-STSP 4(2):310–316.

Pati YC, Rezaiifa R, Krishnaprasad PS (1993) Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. Proceedings of 27th Asilomar Conference on Signals Systems and Computers 40–44.

Pincus SM (1991) Approximate entropy as a measure of system complexity. Proc. Natl. Acad. Sci. 88:2297–2301.

Qi J, Zhang L, Zhang K, Li L, Sun J (2020) The application of improved differential evolution algorithm in electromagnetic fracture monitoring. Adv Geo-energ Res 4: 233–246.

Richman JS, Lake DE, and Moorman J R (2004) Sample entropy. Numerical Computer Method, Part E. 172–184

Ritter O, Junge A, Dawes G (1998) New equipment and processing for magnetotelluric remote reference observations. Geophys J Int 132(3): 535–548.

Rostaghi M, Azami H (2016) Dispersion entropy: A measure for time series analysis. IEEE Signal Proc Let 23: 610–614.

Tang JT, Li J, Xiao X, Zhang LC, Lv QT (2012) Mathematical morphology filtering and noise suppression of magnetotelluric sounding data. Chin J Geophys 55(5): 1784–1793.

Tikhonov AN (1950) On determining electrical characteristics of the deep layers of the Earth's crust. Dokl. Akad. Nauk SSSR 73: 295–297.

Trad DO, Travassos JM (2000) Wavelet filtering of magnetotelluric data. Geophysics 65(2): 482–491.

Tropp JA, Gilbert AC (2007) Signal recovery from random measurements via orthogonal matching pursuit. IEEE Trans Inform Theory 53(12): 4655–4666.

Vallianatos F (1996) Magnetotelluric response of a randomly layered earth. Geophys J Int 125(2): 577–583.

Varentsov IM (2006) Arrays of simultaneous electromagnetic sounding: design, data processing and analysis. Methods in Geochemistry and Geophysics 40: 259–273.

Wang H, Campanya J, Cheng JL, Zhu GW, Wei WB, Jin S, Ye GF (2017) Synthesis of natural electric and magnet Time-series using Inter-station transfer functions and time-series from a Neighboring site (STIN): Applications for processing MT data. J Geophys Res-Sol Ea 122(8): 5835–5851.

Wang JB, Wang SX, Yin HJ, Zhang R (2013) A self-adaption denoising method using orthogonal matching pursuit. SEG Technical Program Expanded Abstracts.

Weckmann U, Magunia A, Ritter O (2005) Effective noise separation for magnetotelluric single site data processing using a frequency domain selection scheme. Geophys J Int 161(3): 635–652.

Xu Y, Chen R, Li Y, Zhang P, Yang J, Zhao X, Liu M, Wu D (2019) Multispectral image segmentation based on a fuzzy clustering algorithm combined with tsallis entropy and a gaussian mixture model. Remote Sens 11: 2772.

Zhang H, Liu J, Chen L, Chen N, Yang X (2019) Fuzzy Clustering algorithm with non-neighborhood spatial information for surface roughness measurement based on the reflected aliasing images. Sensors 19: 3285.

Zhang YD, Tong SG, Cong FY, Xu J (2018) Research of feature extraction method based on sparse reconstruction and multiscale dispersion entropy. Appl Sci 8: 888.

Figures

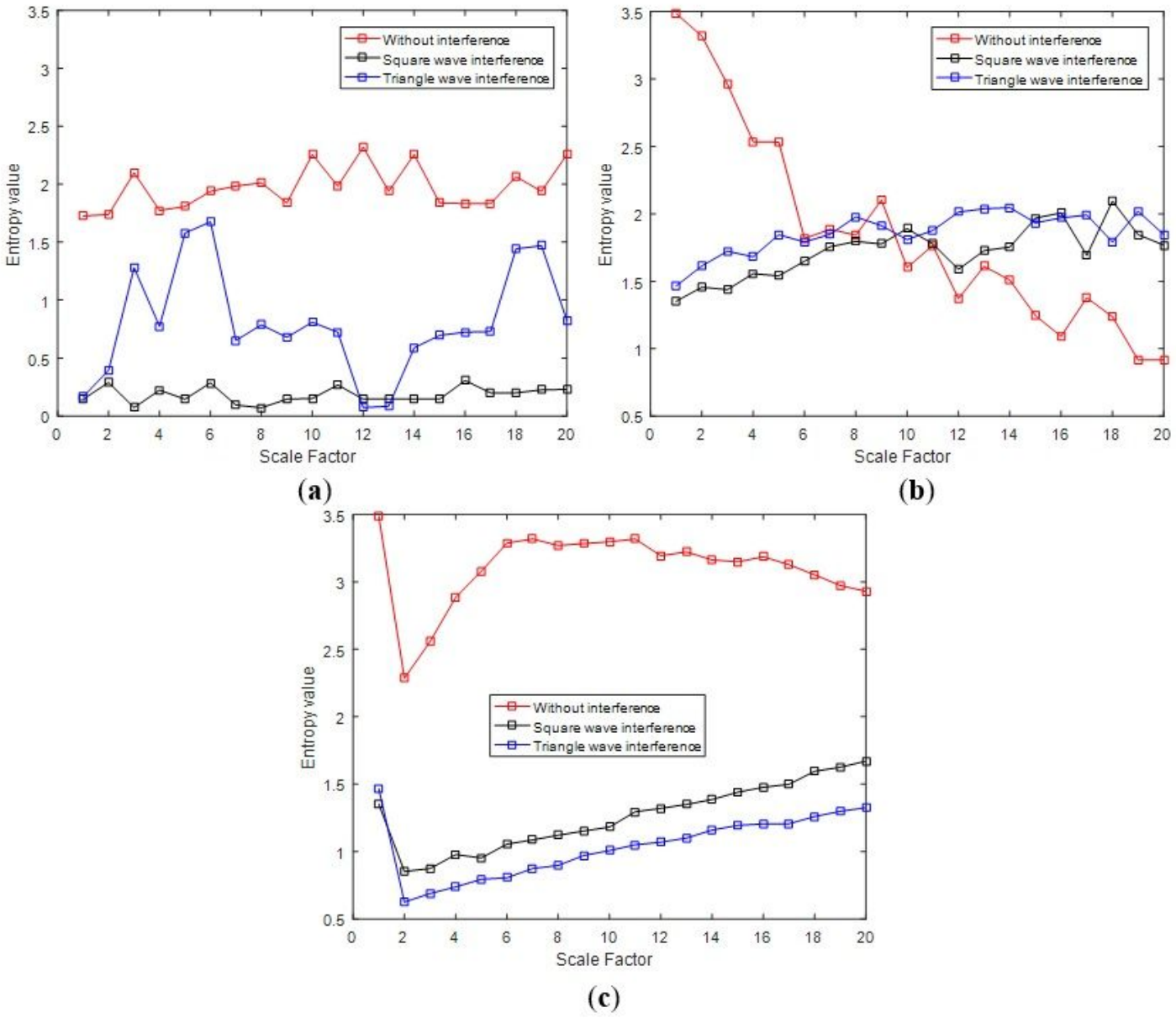


Figure 1

The results obtained for (a)MSE, (b)MDE and (c)RCMDE using a set of sample library signals at different scale factor, among them, the abscissa represents the scale factor , and the ordinate represents the entropy value at the corresponding scale.

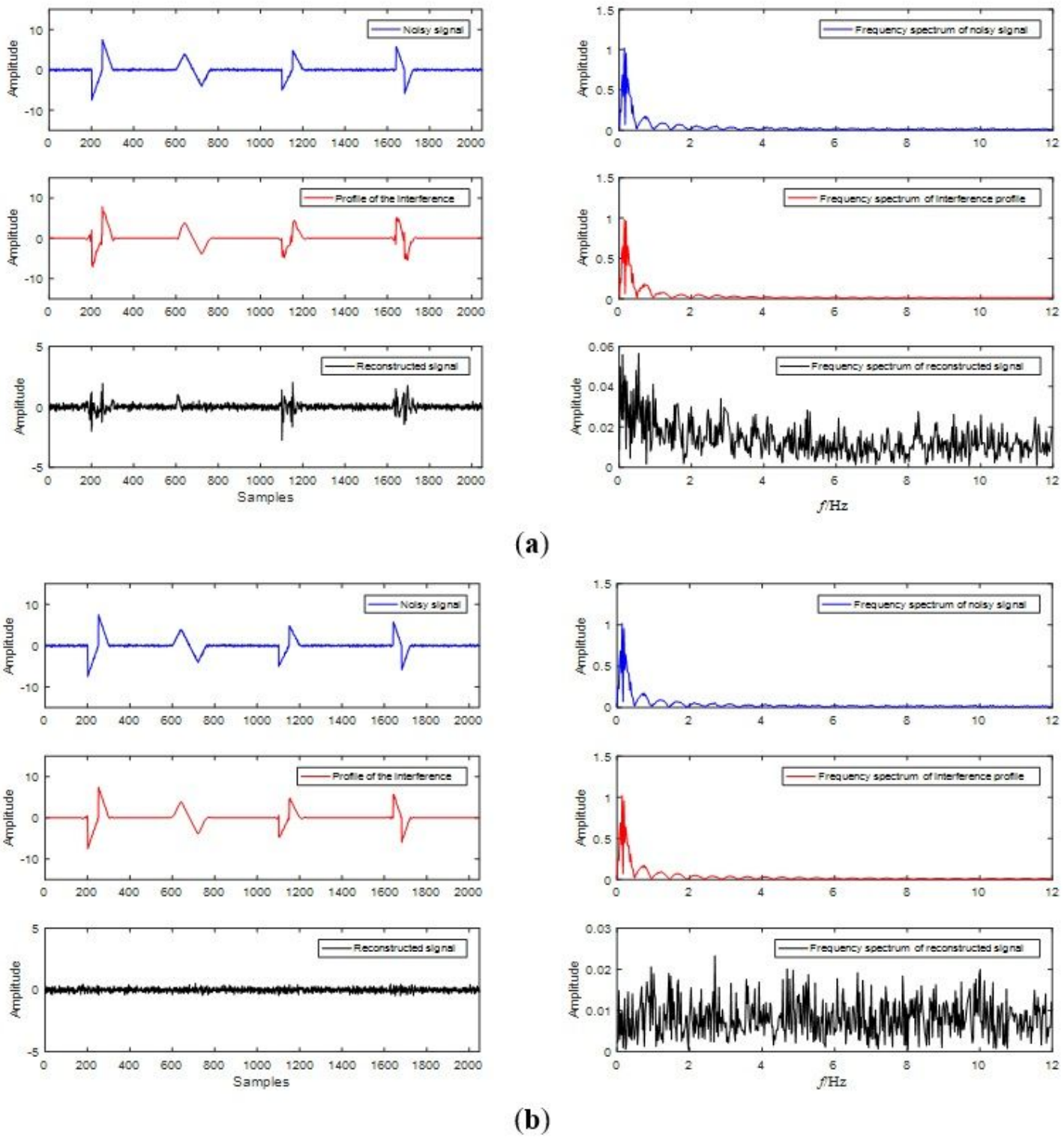


Figure 2

The de-noising effect and frequency spectrum analysis of the noisy signal with (a) matching pursuit (MP) and (b) orthogonal matching pursuit (OMP).

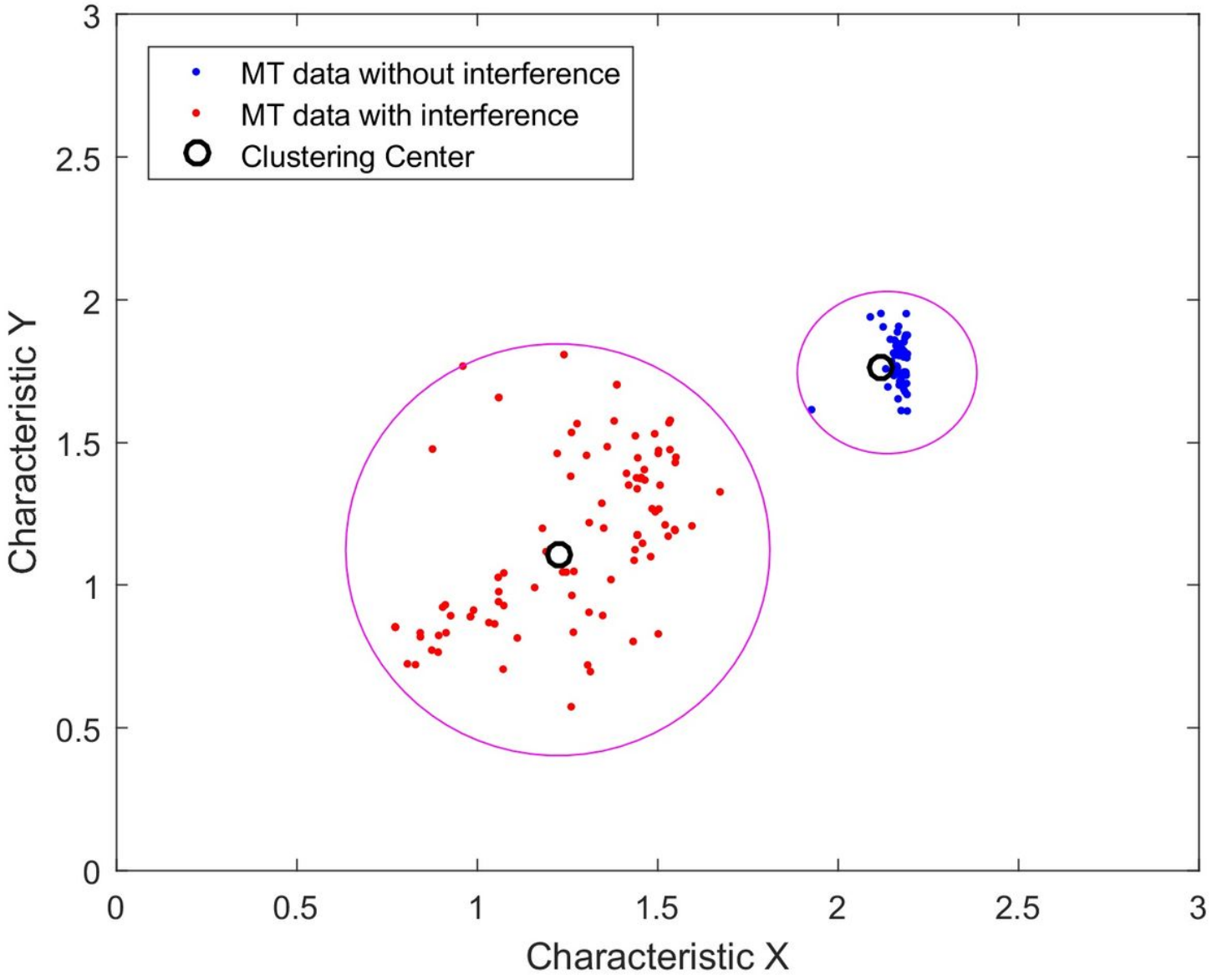


Figure 3

FCM clustering effect of sample library signals. Among them, characteristic X and Y in the coordinate axis represent the refined composite multiscale dispersion entropy (RCMDE) value when the scale factor is 1 and 2, respectively.

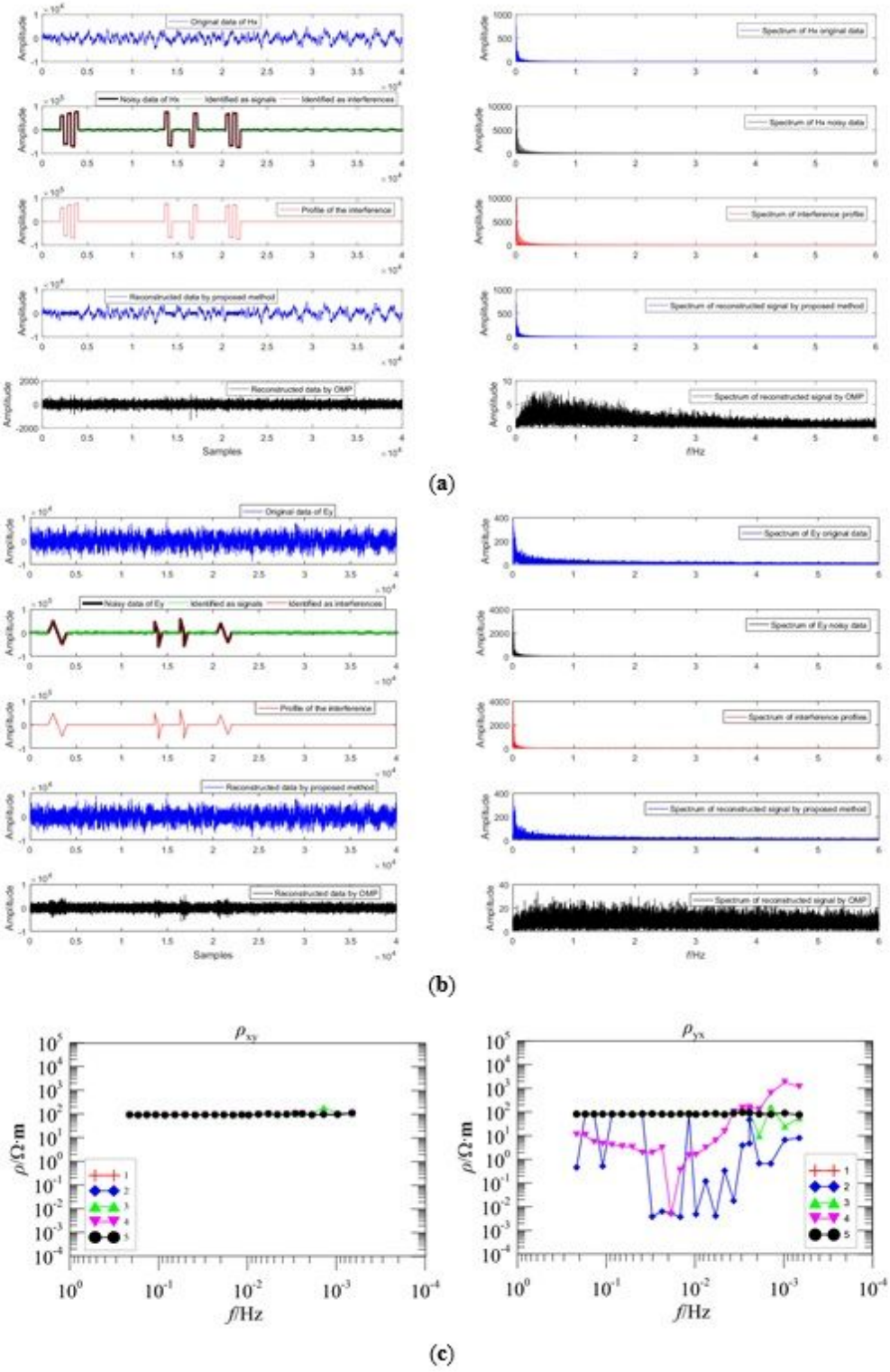


Figure 4

The EMTF data is disturbed by a noisy signal with (a) square wave interference of Hx and (b) charge and discharge triangle wave interference of Ey; (c) is a comparison of the apparent resistivity-phase curves, left figure curve is and right figure curve is . Among them, curve 1 is original data, curve 2 is noisy data, curve 3 is remote reference method, curve 4 is OMP-based overall method, and curve 5 is the proposed method.

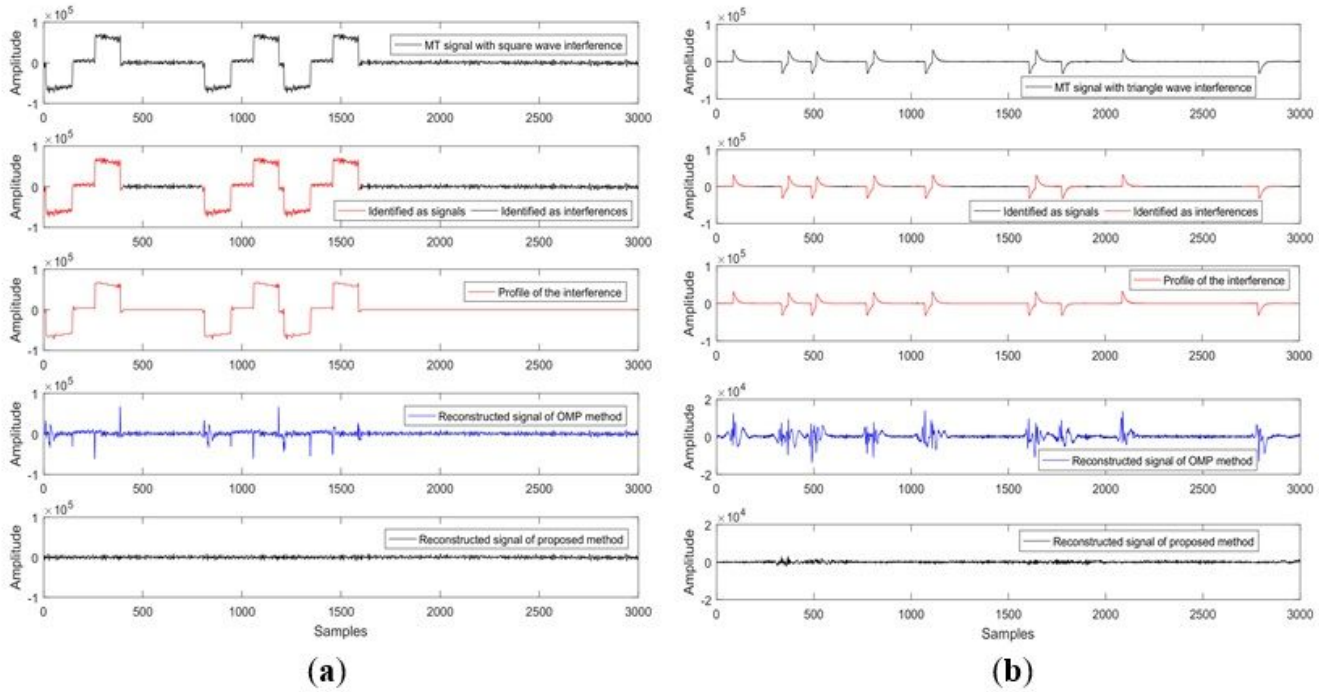


Figure 5

The signal-noise identification and targeted de-noising for the measured MT data, and compared with the OMP method-based overall method and the proposed method, (a) square wave interference and (b) triangle wave interference.

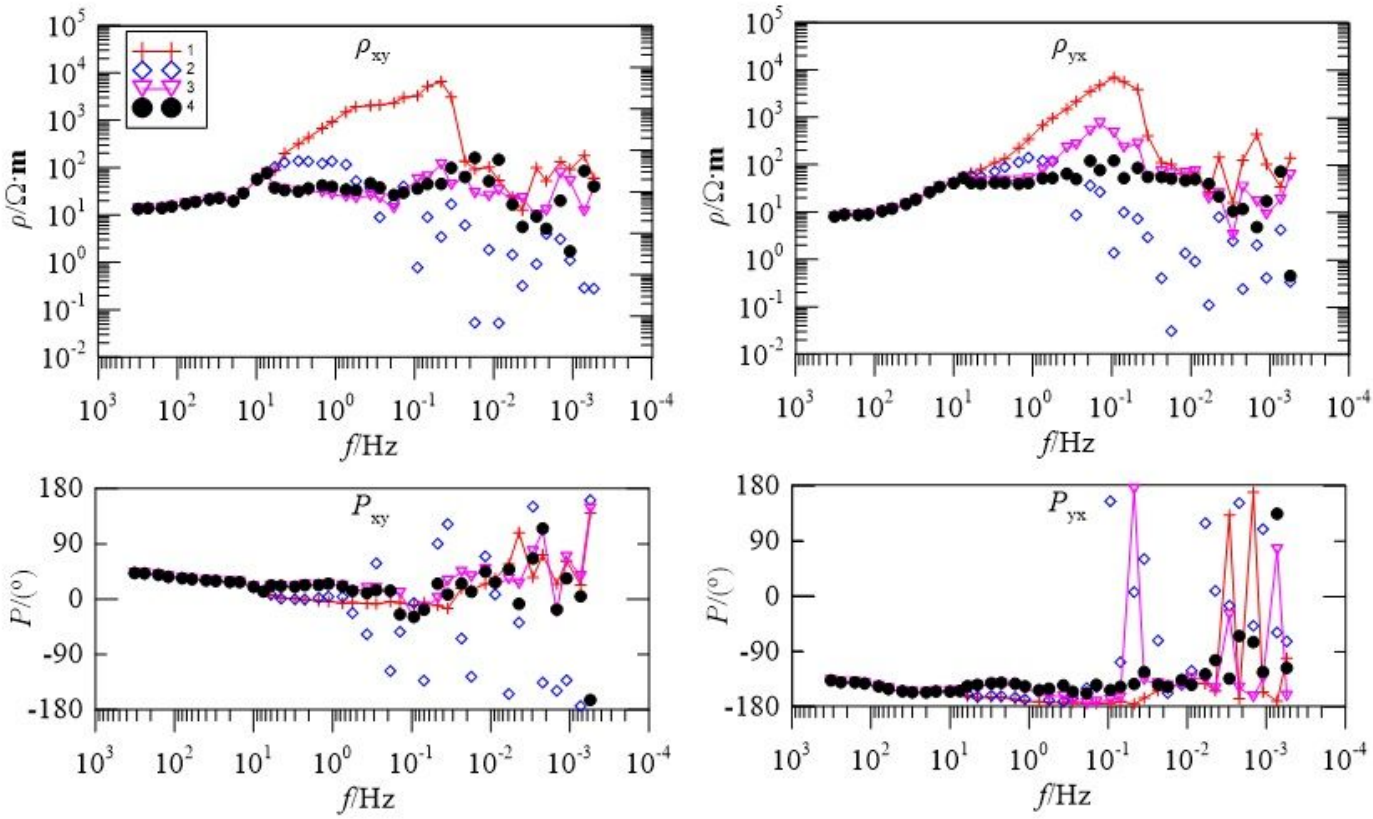


Figure 6

Comparison of the apparent resistivity-phase curves of the measured MT site D37890, among them, curve 1 is original data, curve 2 is the data filtered by the OMP method-based overall method, curve 3 is the result derived from the fractal-entropy and clustering method, and curve 4 is the result derived from the proposed method.

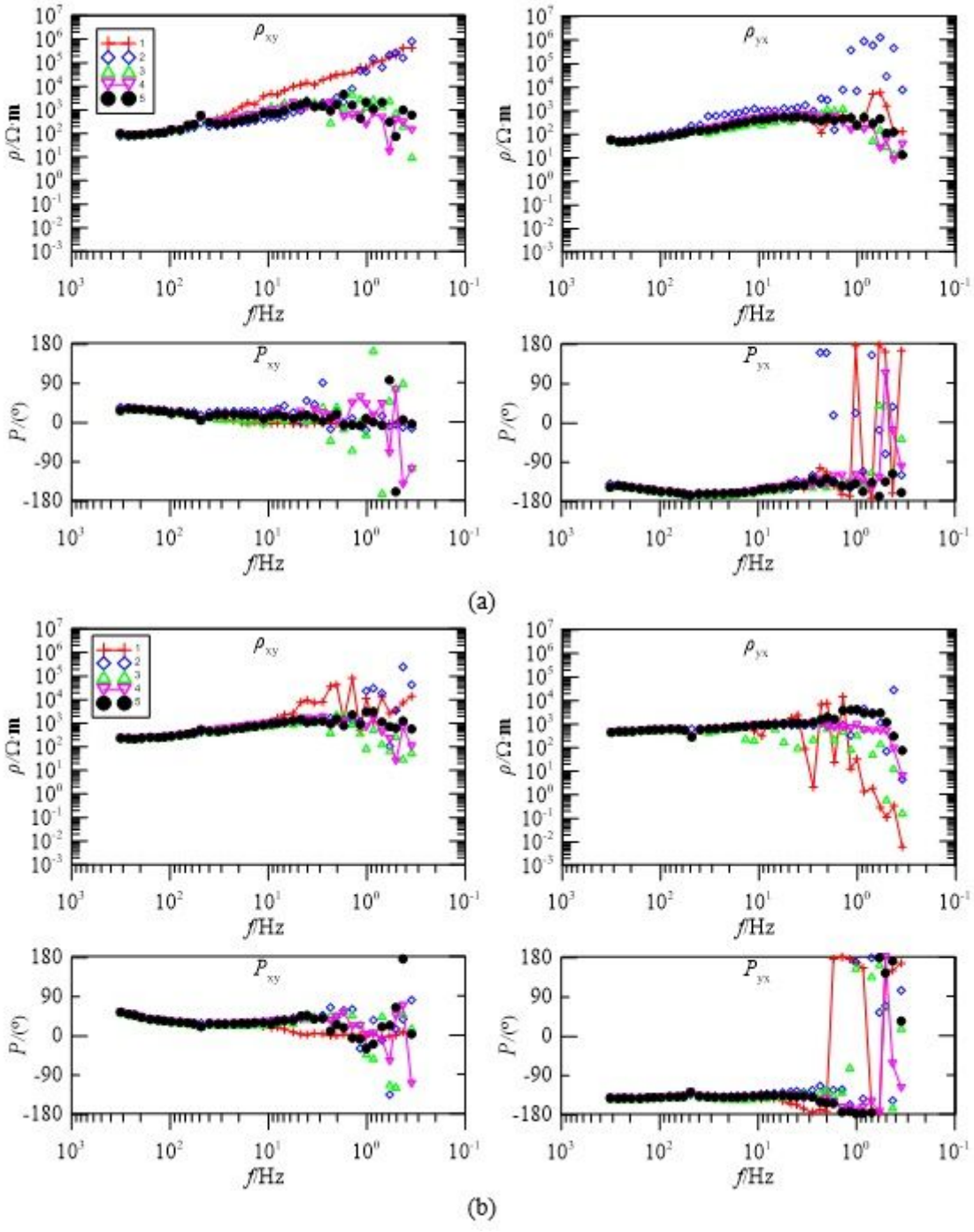


Figure 7

Comparison of the apparent resistivity-phase curves of the measured AMT site (a) EL22189 and (b) EL22174, among them, curve 1 is original data, curve 2 is the result obtained by the remote reference (RR) method, curve 3 is the data filtered by the OMP method-based overall method, curve 4 is the result derived from the fractal-entropy and clustering method, and curve 5 is the result derived from the proposed method.

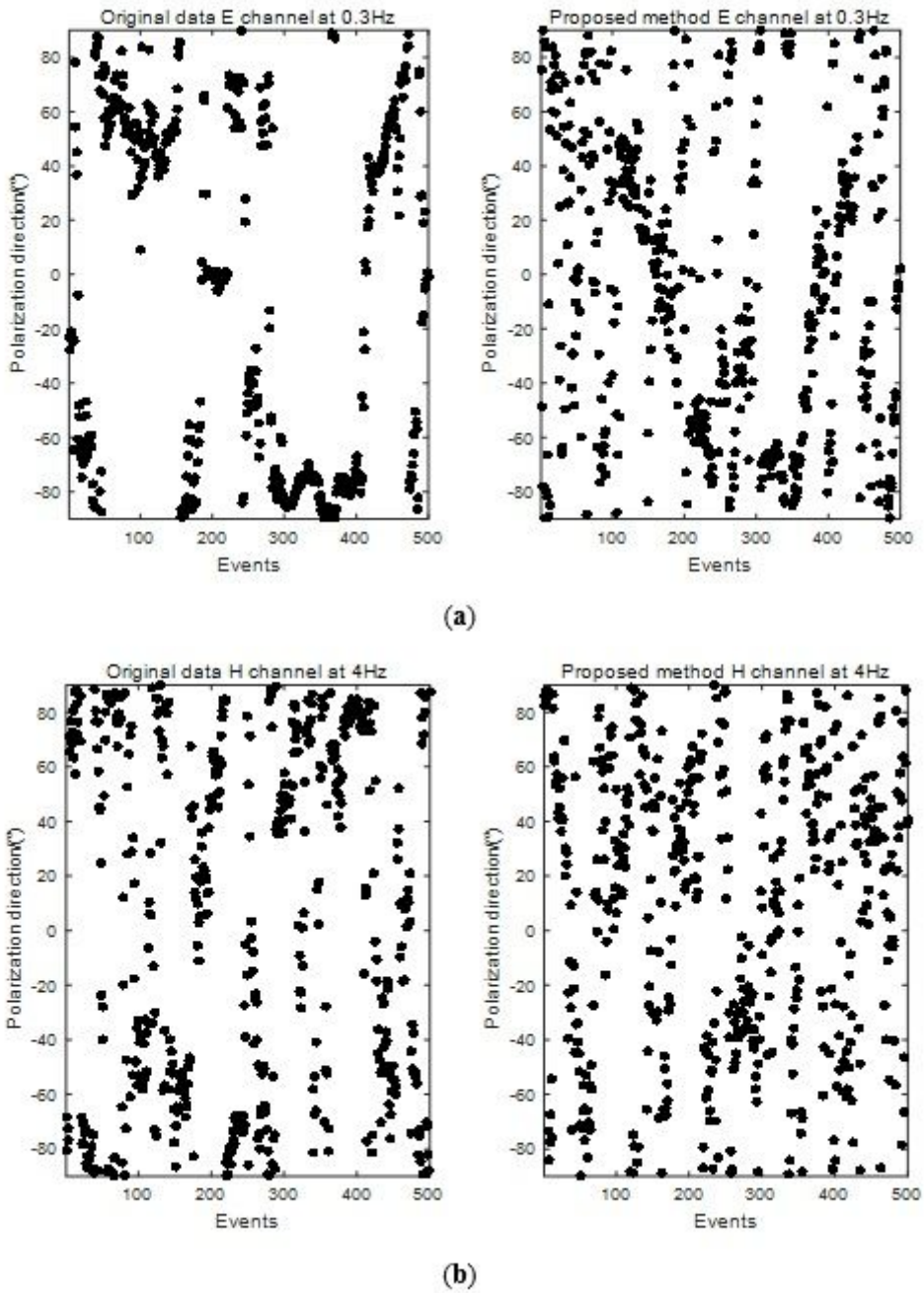


Figure 8

Comparison of the polarization direction for site EL22174, (a) electric field data at 0.3 Hz, and (b) magnetic field data at 4 Hz.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- GraphicalAbstract.jpg