Denoising controlled-source electromagnetic data using least-squares inversion

Yang, Yang; Li, Diquan; Tong, Tiegang; Zhang, Dong; Zhou, Yatong; Chen, Yangkang

DOI
10.1190/geo2016-0659.1

Publication date
2018

Document Version
Final published version

Published in
Geophysics

Citation (APA)
Yang, Y., Li, D., Tong, T., Zhang, D., Zhou, Y., & Chen, Y. (2018). Denoising controlled-source electromagnetic data using least-squares inversion. *Geophysics*, *83*(4), E229-E244. https://doi.org/10.1190/geo2016-0659.1

GEOPHYSICS, VOL. 83, NO. 4 (JULY-AUGUST 2018); P. E229–E244, 26 FIGS.
10.1190/GEO2016-0659.1

Denoising controlled-source electromagnetic data using least-squares inversion

Yang Yang[1], Diquan Li[2], Tiegang Tong[2], Dong Zhang[3], Yatong Zhou[4], and Yangkang Chen[5]

ABSTRACT

Strong noise is one of the toughest problems in the controlled-source electromagnetic (CSEM) method, which highly affects the quality of recorded data. The three main types of noise existing in CSEM data are periodic noise, Gaussian white noise, and nonperiodic noise, among which the nonperiodic noise is thought to be the most difficult to remove. We have developed a novel and effective method for removing such nonperiodic noise by formulating an inverse problem that is based on inverse discrete Fourier transform and several time windows in which only Gaussian white noise exists. These critical locations, which we call reconstruction locations, can be found by taking advantage of the continuous wavelet transform (CWT) and the temporal derivative of the scalogram generated by CWT. The coefficients of the nonperiodic noise are first estimated using the new least-squares method, and then they are subtracted from the coefficients of the raw data to produce denoised data. Together with the nonperiodic noise, we also remove Gaussian noise using the proposed method. We validate the methodology using real-world CSEM data.

INTRODUCTION

The controlled-source electromagnetic (CSEM) method is an important exploration technique for engineering and environmental geophysics (Goldstein and Strangway, 1975), and it is also supplemented with seismic methods in the field of oil and gas exploration (Seigel, 1959; Goldstein and Strangway, 1975; He, 2010; Ziolkowski and Wright, 2012). By applying inversion algorithms on these CSEM data, we can infer geologic structures (Goldstein and Strangway, 1975; Strang, 2007). For all electromagnetic (EM) exploration methods, denoising the recorded data is always a challenging problem (Streich et al., 2013; Yang, 2016). There exist many kinds of electromagnetic noise, some of which originate from cultural sources and others are from natural sources (Reninger et al., 2011). Especially for land-based CSEM, strong noise from industrial, communication, mining, and civil sources can severely pollute the CSEM data (Tang et al., 2012).

Considerable research has been carried out focusing on attenuating noise in EM data. Strack et al. (1989) propose a framework for processing the long-offset transient electromagnetic method (TEM) (LOTEM) data with high cultural noise levels. A robust least-squares stacking method was proposed by Streich et al. (2013) to process CSEM data. Another method based on an equivalent source processing procedure was used to denoise the static shift noise for land-based frequency-domain CSEM data (Maclennan and Li, 2013). A discrete wavelet transform method was used to denoise CSEM data (Willen, 2010). Decomposing signal by singular-value decomposition has been used in airborne TEM data processing (Reninger et al., 2011). Chen et al. (2012) propose empirical-mode decomposition (Huang et al., 1998; Chen and Ma, 2014; Chen, 2016) based on a denoising algorithm for marine magnetotelluric data and obtain acceptable results. Denoising methods that are used in seismic data can also be applied in processing CSEM data, such as some decomposition methods (Chen, 2017, 2018). All the methods mentioned above could achieve good results in some cases. However, when the noise becomes more complicated, it is hard to get a satisfactory result using only one method.

Also, because most geophysical signals are nonstationary, there will be some errors when applying traditional denoising methods. In

Manuscript received by the Editor 8 December 2016; revised manuscript received 4 April 2018; published ahead of production 28 April 2018; published online 28 June 2018.
[1]Central South University, School of Geoscience and Info-physics, Changsha, China and Shandong University, Research Center of Geotechnical and Structural Engineering, Jinan, China. E-mail: noon.y.yang@gmail.com.
[2]Central South University, School of Geoscience and Info-physics, Changsha, China and Central South University, Key Laboratory of Metallogenic Prediction of Non-Ferrous Metals and Geological Environment Monitor, Ministry of Education, Changsha, China. E-mail: lidiquan@126.com; 76399509@qq.com; ttg75@126.com.
[3]Delft University of Technology, Department of Imaging Physics, Mekelweg 2, 2628 CD Delft, Netherlands. E-mail: d.zhang-3@tudelft.nl.
[4]Hebei University of Technology, School of Electronic and Information Engineering, Tianjin 300401, China. E-mail: zyt@hebut.edu.cn.
[5]Zhejiang University, School of Earth Sciences, Hangzhou, Zhejiang 310027, China. E-mail: yangkang.chen@zju.edu.cn.

addition, many filters require a convolution step in the time domain, which corresponds to a multiplication step in the frequency domain (Strang, 2007). The filtering operator can cause different impacts on different frequencies, but it is difficult to distinguish between two different components in a single frequency. When attenuating noise in CSEM data, we aim to derive the true coefficients at those CSEM frequencies, which are usually complex valued. Because of the special features of CSEM data, most of the seismic denoising methods tend to cause more or less damage to useful CSEM signals.

Under such circumstances, we develop an effective denoising methodology designed specifically for CSEM in this paper. As is known, Fourier basis is the best for representing periodic signals (Bracewell, 2000). The useful signal is transmitted by a controlled (periodic) source; hence, the spectrum of it (including main frequencies and harmonics) is a set of discrete, finite spikes in the frequency domain (Welch, 1967; Bracewell, 2000). The spectrum of noise is continuous, which means that the energy of the noise is distributed everywhere in the frequency domain. Besides, we have a direct measurement of noise data for a wide frequency spectrum, i.e., for all frequencies except for those that contain noise and CSEM signal. In this paper, we estimate the noise component first and then subtract it from the raw signal to output the denoised signal component. We formulate an inverse problem for reconstructing the noise energy at several frequency locations where useful energy is located, and we solve it using a least-squares method. The inverse problem takes advantage of the inverse discrete Fourier transform (IDFT) and those locations where only Gaussian noise exists. To accurately find the Gaussian-noise-only locations, which we call reconstruction locations, we use the locating properties of the continuous wavelet transform (CWT) and the temporal derivative of the scalogram generated by CWT.

We organize the paper as follows: First, we introduce in detail the formulation of the inverse problem for inverting the nonperiodic component (NPC), which can be easily solved by a least-squares method. Then, we introduce the approach that we used to find the time domain locations where only Gaussian noise exists using a benchmark synthetic example that is composed of different types of noise. Finally, we apply the proposed method to real-world CSEM data and demonstrate the excellent result we have obtained. Several key conclusions are drawn at the end of the paper.

METHOD

NPC estimation via least-squares inversion

Signal recorded by CSEM receivers can be divided into two components: periodic component (PC) and NPC. Figure 1 shows a clear

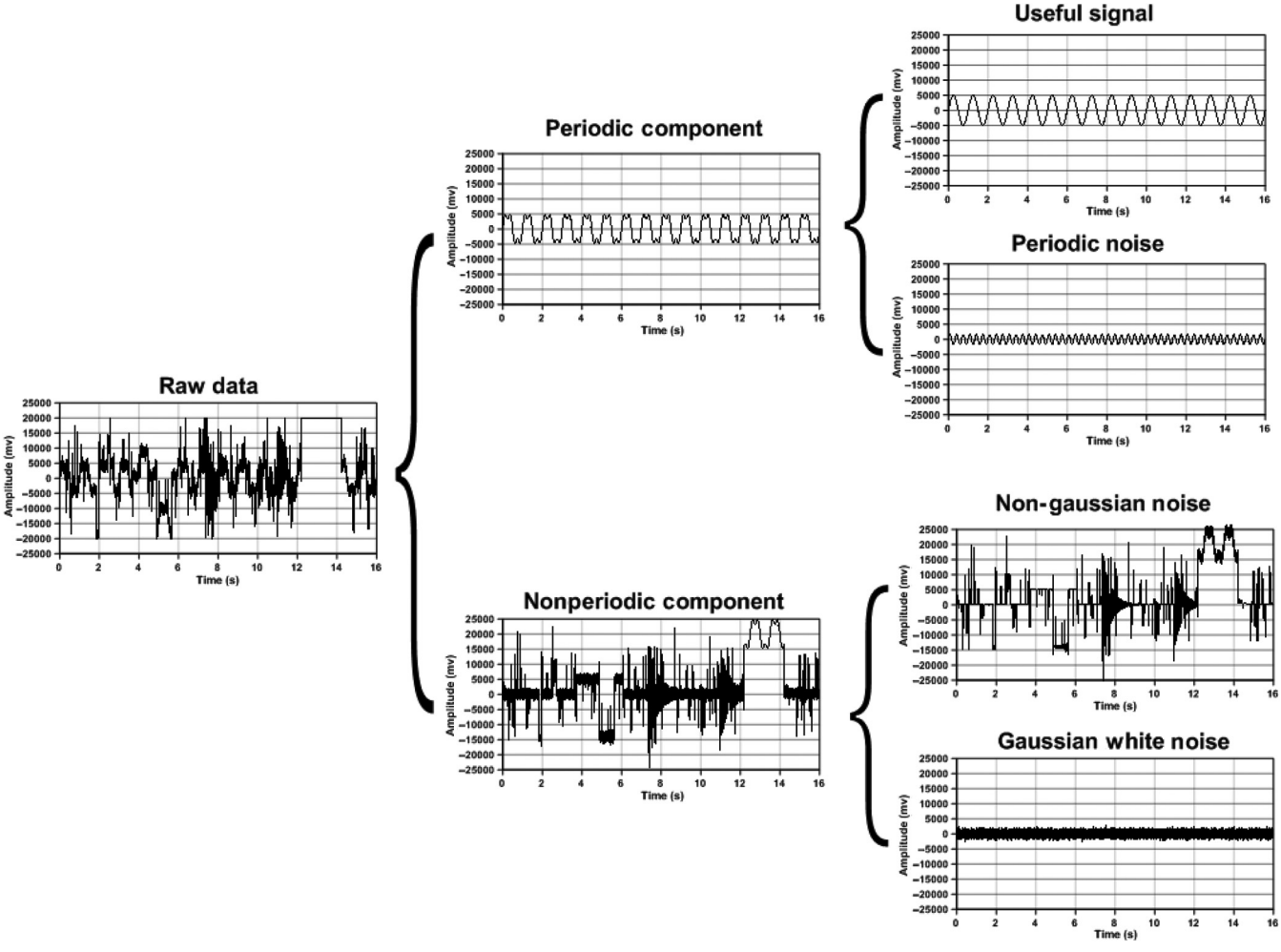

Figure 1. Classifications of raw CSEM data with different components.

classification of signal components in the CSEM data. NPC includes nonperiodic noise and Gaussian white noise that are the noise components of the CSEM data. PC is composed of useful signal and periodic noise. Periodic noise can be rejected by applying a narrow (notch) filter that removes a narrow band around those frequencies of interest. Recorded CSEM data contain periodic noise, Gaussian white noise, and nonperiodic noise. The nonperiodic noise, however, is extremely difficult to attenuate because of its highly nonstationary and unpredictable features. We focus on attenuating nonperiodic noise in this paper. More specifically, we estimate NPC by constructing an over-determined equation.

The assumption of the proposed method is that PCs and Gaussian white noise are spreading all over the signal, whereas the nonperiodic noise is time limited, which indicates that we can pick some parts of the data that only contain Gaussian white noise.

The proposed method is based on IDFT:

$$f(m) = \frac{1}{N}\sum_{k=0}^{N-1} F(k) w^{km}, \quad m = 0, 1, \cdots, N-1, \tag{1}$$

where $f(m)$ is the time-domain signal of NPC at mth location, $F(k)$ is the frequency-domain signal of NPC at kth location, N is the total length of signal, and $w = e^{2\pi i/N}$. The detailed matrix form of equation 1 can be expressed as

$$\begin{bmatrix} f(0) \\ f(1) \\ f(2) \\ \vdots \\ \vdots \\ f(N-2) \\ f(N-1) \end{bmatrix} = \frac{1}{N} \begin{bmatrix} 1 & 1 & 1 & \cdots & \cdots & \cdots & 1 \\ 1 & w^1 & w^2 & \cdots & \cdots & \cdots & w^{N-1} \\ 1 & w^2 & w^4 & \cdots & \cdots & \cdots & w^{2(N-1)} \\ \vdots & \vdots & \vdots & \cdots & \cdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & \cdots & \cdots & \vdots \\ 1 & w^{N-2} & w^{2(N-2)} & \cdots & \cdots & \cdots & w^{(N-1)(N-2)} \\ 1 & w^{N-1} & w^{2(N-1)} & \cdots & \cdots & \cdots & w^{(N-1)^2} \end{bmatrix} \times \begin{bmatrix} F(0) \\ F(1) \\ F(2) \\ \vdots \\ \vdots \\ F(N-2) \\ F(N-1) \end{bmatrix}. \tag{2}$$

If there are four desired frequencies in the frequency domain, we can solve them by choosing any four points with a known value in the time domain to construct a linear equation. The "desired frequencies" mean those frequencies that contain noisy CSEM signal. In equation 3, the red entries are the desired frequencies and the blue entries are those frequencies contain only noise and no CSEM signal that is known. The blue entries in the left are the chosen points in the time domain.

$$\begin{bmatrix} f(0) \\ f(1) \\ f(2) \\ f(m) \\ f(n) \\ f(p) \\ f(q) \\ \vdots \\ f(N-1) \end{bmatrix} = \frac{1}{N} \begin{bmatrix} 1 & 1 & 1 & \cdots & \cdots & \cdots & 1 \\ 1 & w^1 & w^2 & \cdots & \cdots & \cdots & w^{N-1} \\ 1 & w^2 & w^4 & \cdots & \cdots & \cdots & w^{2(N-1)} \\ \vdots & \vdots & \vdots & \cdots & \cdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & \cdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & \cdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & \cdots & \cdots & \vdots \\ 1 & w^{N-2} & w^{2(N-2)} & \cdots & \cdots & \cdots & w^{(N-1)(N-2)} \\ 1 & w^{N-1} & w^{2(N-1)} & \cdots & \cdots & \cdots & w^{(N-1)^2} \end{bmatrix} \begin{bmatrix} F(0) \\ F(1) \\ F(2) \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ F(N-2) \\ F(N-1) \end{bmatrix} \tag{3}$$

Because there exists Gaussian white noise in the data, more points in the time domain should be used to construct an over-determined linear equation to solve the desired frequency components.

If we can find certain locations in the time domain where only Gaussian noise exists, a set of over-determined equations will be constructed to reconstruct the spectrum of NPC at those CSEM frequency locations. Note that here NPC includes nonperiodic noise and Gaussian white noise. Although we only select the locations containing Gaussian white noise in the time domain, the spectrum that requires reconstruction is a mixture of nonperiodic noise and Gaussian white noise.

Assume that there are six time-domain points selected, and there are four frequency locations. The over-determined linear equations can then be expressed as

$$
\begin{aligned}
g(m) &= \frac{1}{N}\sum_{k\neq l_1,l_2,\tilde{l}_1,\tilde{l}_2}^{N-1} F(k)w^{km} + \frac{1}{N}(F(l_1)w^{l_1 m} \\
&\quad + F(\tilde{l}_1)w^{\tilde{l}_1 m} + F(l_2)w^{l_2 m} + F(\tilde{l}_2)w^{\tilde{l}_2 m}) \\
g(n) &= \frac{1}{N}\sum_{k\neq l_1,l_2,\tilde{l}_1,\tilde{l}_2}^{N-1} F(k)w^{kn} + \frac{1}{N}(F(l_1)w^{l_1 n} \\
&\quad + F(\tilde{l}_1)w^{\tilde{l}_1 n} + F(l_2)w^{l_2 n} + F(\tilde{l}_2)w^{\tilde{l}_2 n}) \\
g(p) &= \frac{1}{N}\sum_{k\neq l_1,l_2,\tilde{l}_1,\tilde{l}_2}^{N-1} F(k)w^{kp} + \frac{1}{N}(F(l_1)w^{l_1 p} \\
&\quad + F(\tilde{l}_1)w^{\tilde{l}_1 p} + F(l_2)w^{l_2 p} + F(\tilde{l}_2)w^{\tilde{l}_2 p}) \\
g(q) &= \frac{1}{N}\sum_{k\neq l_1,l_2,\tilde{l}_1,\tilde{l}_2}^{N-1} F(k)w^{kq} + \frac{1}{N}(F(l_1)w^{l_1 q} \\
&\quad + F(\tilde{l}_1)w^{\tilde{l}_1 q} + F(l_2)w^{l_2 q} + F(\tilde{l}_2)w^{\tilde{l}_2 q}) \\
g(r) &= \frac{1}{N}\sum_{k\neq l_1,l_2,\tilde{l}_1,\tilde{l}_2}^{N-1} F(k)w^{kr} + \frac{1}{N}(F(l_1)w^{l_1 r} \\
&\quad + F(\tilde{l}_1)w^{\tilde{l}_1 r} + F(l_2)w^{l_2 r} + F(\tilde{l}_2)w^{\tilde{l}_2 r}) \\
g(t) &= \frac{1}{N}\sum_{k\neq l_1,l_2,\tilde{l}_1,\tilde{l}_2}^{N-1} F(k)w^{kt} + \frac{1}{N}(F(l_1)w^{l_1 t} \\
&\quad + F(\tilde{l}_1)w^{\tilde{l}_1 t} + F(l_2)w^{l_2 t} + F(\tilde{l}_2)w^{\tilde{l}_2 t}), \qquad (4)
\end{aligned}
$$

where $\tilde{l}_1 = N - l_1$ and $\tilde{l}_2 = N - l_2$. Note that $F(1)$ and $F(N-1)$ are the complex conjugate in equations 1 and 4; thus, there are actually two frequency components missing in the above equation. The term $g(j), j = m, n, p, q, r, t$, in equation 4, denotes the time-domain signal (i.e., the Gaussian white noise) at different locations (m, n, p, q, r, t).

Rearranging equation 4, we obtain a clear structure of an inverse problem:

$$
\begin{aligned}
&\frac{1}{N}(F(l_1)w^{l_1 m} + F(\tilde{l}_1)w^{\tilde{l}_1 m} + F(l_2)w^{l_2 m} \\
&+ F(\tilde{l}_2)w^{\tilde{l}_2 m}) = -\frac{1}{N}\sum_{k\neq l_1,l_2,\tilde{l}_1,\tilde{l}_2}^{N-1} F(k)w^{km} + g(m) \\
&\frac{1}{N}(F(l_1)w^{l_1 n} + F(\tilde{l}_1)w^{\tilde{l}_1 n} + F(l_2)w^{l_2 n} \\
&+ F(\tilde{l}_2)w^{\tilde{l}_2 n}) = -\frac{1}{N}\sum_{k\neq l_1,l_2,\tilde{l}_1,\tilde{l}_2}^{N-1} F(k)w^{kn} + g(n) \\
&\frac{1}{N}(F(l_1)w^{l_1 p} + F(\tilde{l}_1)w^{\tilde{l}_1 p} + F(l_2)w^{l_2 p} \\
&+ F(\tilde{l}_2)w^{\tilde{l}_2 p}) = -\frac{1}{N}\sum_{k\neq l_1,l_2,\tilde{l}_1,\tilde{l}_2}^{N-1} F(k)w^{kp} + g(p) \\
&\frac{1}{N}(F(l_1)w^{l_1 q} + F(\tilde{l}_1)w^{\tilde{l}_1 q} + F(l_2)w^{l_2 q} \\
&+ F(\tilde{l}_2)w^{\tilde{l}_2 q}) = -\frac{1}{N}\sum_{k\neq l_1,l_2,\tilde{l}_1,\tilde{l}_2}^{N-1} F(k)w^{kq} + g(q) \\
&\frac{1}{N}(F(l_1)w^{l_1 r} + F(\tilde{l}_1)w^{\tilde{l}_1 r} + F(l_2)w^{l_2 r} \\
&+ F(\tilde{l}_2)w^{\tilde{l}_2 r}) = -\frac{1}{N}\sum_{k\neq l_1,l_2,\tilde{l}_1,\tilde{l}_2}^{N-1} F(k)w^{kr} + g(r) \\
&\frac{1}{N}(F(l_1)w^{l_1 t} + F(\tilde{l}_1)w^{\tilde{l}_1 t} + F(l_2)w^{l_2 t} \\
&+ F(\tilde{l}_2)w^{\tilde{l}_2 t}) = -\frac{1}{N}\sum_{k\neq l_1,l_2,\tilde{l}_1,\tilde{l}_2}^{N-1} F(k)w^{kt} + g(t), \qquad (5)
\end{aligned}
$$

which can be written in a matrix vector formulation as

$$
\mathbf{A}\mathbf{x} = \mathbf{b} + \mathbf{g}. \qquad (6)
$$

Here,

$$
\mathbf{A} = \begin{bmatrix}
w^{l_1 m} & w^{\tilde{l}_1 m} & w^{l_2 m} & w^{\tilde{l}_2 m} \\
w^{l_1 n} & w^{\tilde{l}_1 n} & w^{l_2 n} & w^{\tilde{l}_2 n} \\
w^{l_1 p} & w^{\tilde{l}_1 p} & w^{l_2 p} & w^{\tilde{l}_2 p} \\
w^{l_1 q} & w^{\tilde{l}_1 q} & w^{l_2 q} & w^{\tilde{l}_2 q} \\
w^{l_1 r} & w^{\tilde{l}_1 r} & w^{l_2 r} & w^{\tilde{l}_2 r} \\
w^{l_1 t} & w^{\tilde{l}_1 t} & w^{l_2 t} & w^{\tilde{l}_2 t}
\end{bmatrix}, \quad
\mathbf{x} = \begin{bmatrix}
F(l_1) \\ F(\tilde{l}_1) \\ F(l_2) \\ F(\tilde{l}_2)
\end{bmatrix}, \qquad (7)
$$

and

$$
\mathbf{b} = -\begin{bmatrix}
\sum_{k\neq l_1,l_2,\tilde{l}_1,\tilde{l}_2}^{N-1} F(k)w^{km} \\
\sum_{k\neq l_1,l_2,\tilde{l}_1,\tilde{l}_2}^{N-1} F(k)w^{kn} \\
\sum_{k\neq l_1,l_2,\tilde{l}_1,\tilde{l}_2}^{N-1} F(k)w^{kp} \\
\sum_{k\neq l_1,l_2,\tilde{l}_1,\tilde{l}_2}^{N-1} F(k)w^{kq} \\
\sum_{k\neq l_1,l_2,\tilde{l}_1,\tilde{l}_2}^{N-1} F(k)w^{kr} \\
\sum_{k\neq l_1,l_2,\tilde{l}_1,\tilde{l}_2}^{N-1} F(k)w^{kt}
\end{bmatrix}, \quad
\mathbf{g} = N\begin{bmatrix}
g(m) \\ g(n) \\ g(p) \\ g(q) \\ g(r) \\ g(t)
\end{bmatrix}. \qquad (8)
$$

Equation 6 can be solved via the least-squares method

$$
\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2, \qquad (9)
$$

and the solution is expressed as

$$
\hat{\mathbf{x}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}, \qquad (10)
$$

where $[\cdot]^T$ denotes the transpose and $(\cdot)^{-1}$ denotes the inverse.

The output error caused by the Gaussian white noise at reconstruction locations can be expressed as

$$
\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}. \qquad (11)
$$

Inserting equation 10 into equation 11, it is easy to derive that

$$\begin{aligned}\mathbf{e} &= \mathbf{x} - \hat{\mathbf{x}} \\ &= \mathbf{x} - (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b} \\ &= (\mathbf{A}^T\mathbf{A})^{-1}((\mathbf{A}^T\mathbf{A})\mathbf{x} - \mathbf{A}^T\mathbf{b}) \\ &= (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T(\mathbf{A}\mathbf{x} - \mathbf{b}) \\ &= (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{g} \\ &= \mathbf{L}\mathbf{g}, \end{aligned} \tag{12}$$

where $\mathbf{L} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$. Because the Gaussian noise vector $\mathbf{g}$ is unknown, it is impossible to calculate $\mathbf{e}$ directly. Fortunately, if there is only Gaussian noise at these locations, it is possible to estimate output error covariance as

$$\begin{aligned} E(\mathbf{e}\mathbf{e}^T) &= E(\mathbf{L}\mathbf{g}\mathbf{g}^T\mathbf{L}^T) = \mathbf{L}E(\mathbf{g}\mathbf{g}^T)\mathbf{L}^T \\ &= \sigma^2\mathbf{L}\mathbf{L}^T = \sigma^2(\mathbf{A}^T\mathbf{A})^{-1}, \end{aligned} \tag{13}$$

where $E(\mathbf{g}\mathbf{g}^T) = \sigma^2\mathbf{I}$, σ is the variance of Gaussian noise. It is much easier to estimate σ of the Gaussian noise than its accurate value. The term σ is almost the same as the variance of the residual error at those reconstruction locations. If more reconstruction locations are chosen; it adds new rows to $\mathbf{A}$ and increases $\mathbf{A}^T\mathbf{A}$. Then, the output covariance becomes smaller, which means that the solution becomes more stable and accurate. Each element on the principal diagonal of the covariance matrix $\sigma^2(\mathbf{A}^T\mathbf{A})^{-1}$ is the variance of one of the desired CSEM frequency coefficients.

Actually, the stability of the solution depends on two parts. One is the σ^2 of the Gaussian noise at the reconstruction locations. If the variance at the reconstruction locations is smaller, the variance of one single frequency coefficient is smaller. The other one is the value of $(\mathbf{A}^T\mathbf{A})^{-1}$. If more reconstruction locations are chosen, more rows are added to $\mathbf{A}$ and the values on the diagonal of $(\mathbf{A}^T\mathbf{A})^{-1}$ decrease. The variances of the frequency coefficients are reduced, and thus the solution becomes more stable. This makes sense especially when the condition number of $\mathbf{A}^T\mathbf{A}$ is small. When the condition number is small, the matrix is almost diagonal and the off-diagonal values are almost zero, which indicates that the coefficient is not correlated with other frequency coefficients and the diagonal can describe the stability of the coefficient perfectly. Let V be the diagonal values of the covariance matrix or the variance of each CSEM frequency coefficient:

$$V = \mathrm{diag}(\sigma^2(\mathbf{A}^T\mathbf{A})^{-1}). \tag{14}$$

Let k_f be the relative standard deviation (RSD) of the coefficients at frequency f, then

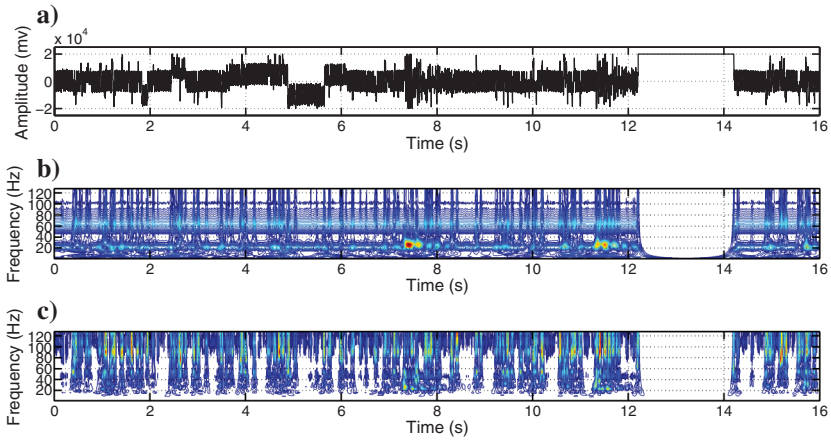

Figure 2. (a) Raw data (time domain) with a high-amplitude periodic signal in 20 and 60 Hz. (b) Continuous complex Morlet wavelet (fc = 1 Hz, fb = 1.5 Hz) scalogram. (c) Temporal derivation of the scalogram. The frequency range is 0.02726–128 Hz.

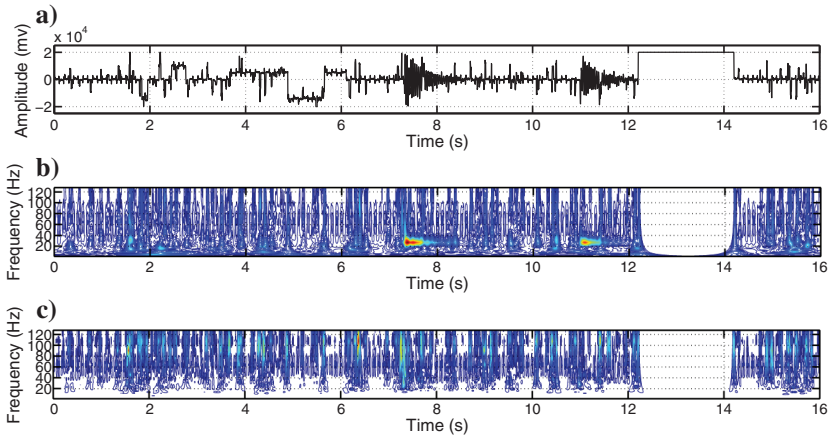

Figure 3. (a) Raw data (time domain) with periodic signal in many frequencies. (b) Continuous complex Morlet wavelet (fc = 1 Hz, fb = 1.5 Hz) scalogram. (c) Temporal derivation of the scalogram. The frequency range is 0.2726–128 Hz.

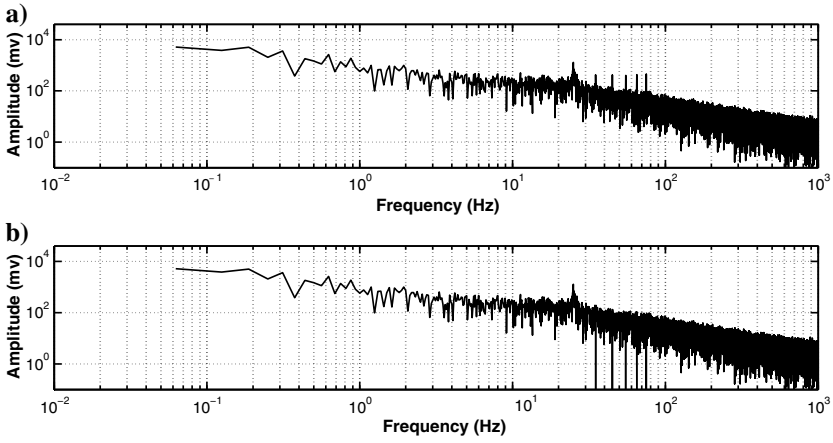

Figure 4. (a) Raw spectrum and (b) initial NPC spectrum after zeroing the desired CSEM frequencies.

$$k_f = \frac{\sqrt{V_f}}{S_f}, \tag{15}$$

where S_f is the processed spectrum at frequency f.

Once the reconstruction locations have been chosen, the matrix $(\mathbf{A}^T\mathbf{A})^{-1}$ is fixed. Because we can get the estimated variance σ^2 of the Gaussian noise at the reconstruction locations, the total covariance matrix is obtained. Then, the stability of the solution can be calculated by equation 15. The RSD of every frequency can be obtained to evaluate the accuracy of the processed result. An RSD threshold can be set (e.g., 5%) for accepting the processed data.

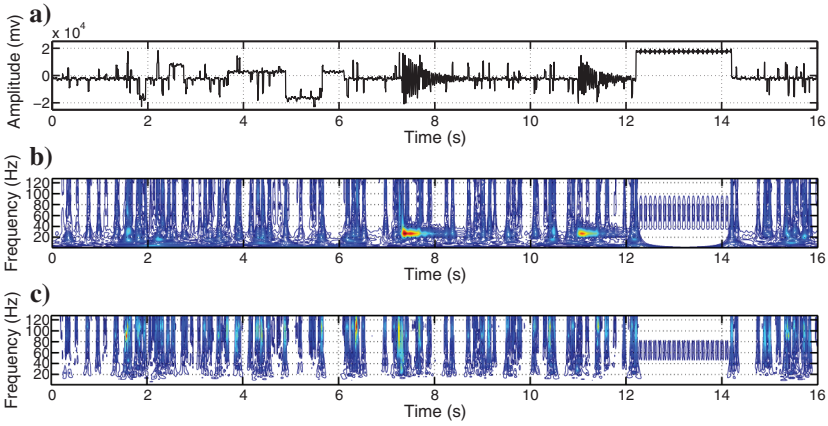

Figure 5. (a) Initial NPC in the time domain, (b) continuous complex Morlet wavelet (fc = 1 Hz, fb = 1.5 Hz) scalogram, and (c) temporal derivation of the scalogram. The frequency range is 0.2726–128 Hz.

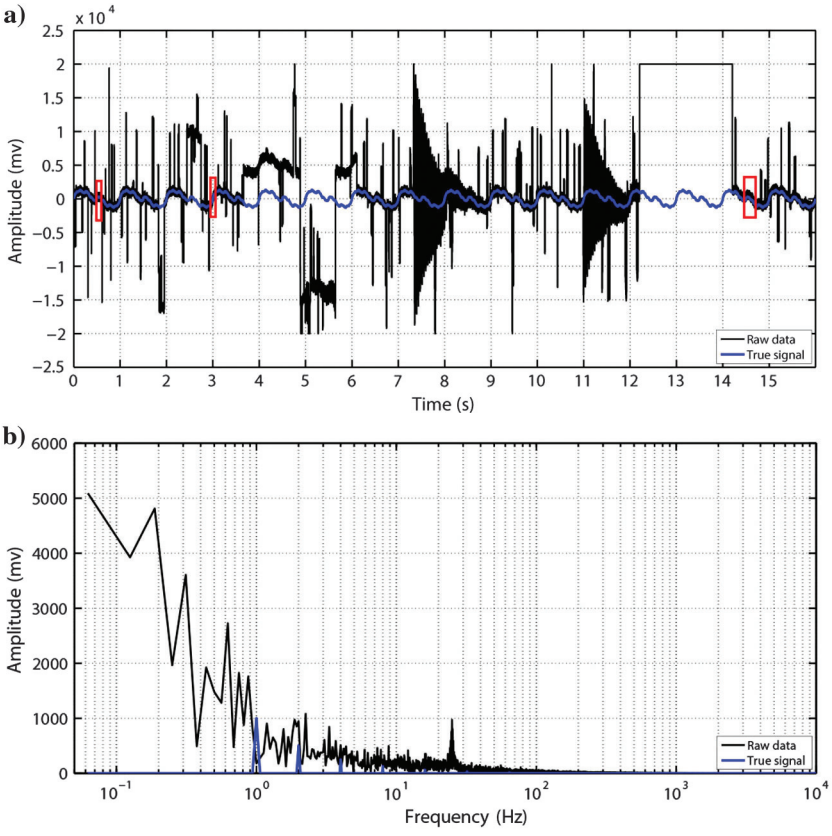

Figure 6. (a) Raw data overlapped by the true signal (blue lines). The red boxes are the candidate reconstruction locations. (b) The spectrum of the raw data overlapped by the spectrum of the true signal (blue line).

Picking the reconstruction locations via CWT

In this part, we will introduce a method for picking the time-domain locations that only contain predominantly Gaussian white noise, or what we call "reconstruction locations." The method is based on the CWT:

$$Wf(u,s) = \int_{-\infty}^{+\infty} f(t)\frac{1}{\sqrt{s}}\phi^*\left(\frac{t-u}{s}\right)dt, \tag{16}$$

where ϕ is the mother wavelet, $f(t)$ is the input 1D signal, $Wf(u,s)$ is the 2D time-frequency spectrum, $[\cdot]^*$ denotes the complex conjugate, u and s denote the position and scale, respectively. We use the complex Morlet wavelet as the mother wavelet because properties of the Morlet wavelet are favorable for the method.

In principle, CWT analysis should be applied to the NPC only, which just includes nonperiodic noise and Gaussian white noise, for finding reconstruction locations because what we need to know is the distribution of NPC, not PC. However, we do not know the NPC at the beginning. What we have is just raw data. Sometimes, especially when high-energy periodic signal or noise exists, it will be difficult to find those locations directly in the wavelet spectrum as shown in Figure 2b. Under this condition, we take the temporal derivative of the wavelet spectrum, by which we get the temporal inhomogeneity information as shown in Figure 2c. Those white areas in Figure 2c are more easily recognizable as potential reconstruction locations than in Figure 2b.

But that is not accurate enough to find the reconstruction locations. When the number of CSEM frequencies increases, it is not easy to find the Gaussian-noise-only places directly from the CWT spectrum and its temporal derivative. A synthetic example is shown below. The CSEM frequencies are 5, 15, 25, ..., 75 Hz.

The example indicates that the periodic signal with strong energy masks the spectrum of the NPC. So an initial guess of NPC should be estimated to reduce the influence from PC. It is necessary to modify the spectrum and construct a modified time-domain signal of initial NPC for CWT analysis to reduce the influence of strong periodic signal. By CWT analysis on this initial NPC, we can find some Gaussian-only places. In

this paper, we give a criterion for zeroing the desired CSEM frequency coefficients, and then we reconstruct the modified coefficients back to the time domain to get the initial guess of NPC. As shown in equation 17, we create a modified spectrum, by setting the spectrum at the desired CSEM frequencies to the average of the nearest two frequency amplitudes, in which f_s are the CSEM frequencies:

$$\hat{S}(f) = \begin{cases} S(f) & f \neq f_s \\ \frac{S(f-df)+S(f+df)}{2} & f = f_s \end{cases}. \quad (17)$$

By using equation 18, we can estimate a noise to signal ratio k in those CSEM frequencies. If k is smaller than 50%, the coefficient of this frequency will be set to zero; otherwise, the coefficient will be kept unchanged. Following the criterion above, comparing to Figure 3, the coefficients of 35, 45, 55, 65, and 75 Hz have been set to zero, whereas the coefficients of 5, 15, and 25 Hz are kept unchanged as shown in Figure 4b because at these frequencies the ratio is larger than 50%, which means that the noise in these frequencies is very strong compared with the CSEM signal:

$$k_{f_s} = \frac{\hat{S}(f_s)}{S(f_s)}. \quad (18)$$

By inverse Fourier transform, we get the modified signal in the time domain. The CWT analysis on this modified signal as shown in Figure 5 is quite clear compared with the result in Figure 3.

Using the CWT spectrum from the modified signal, it is much easier for us to find those Gaussian-noise-only places in Figure 5. We do CWT analysis on initial NPC here, rather than raw data. By solving the over-determined equation, we can get the first least-squares solution, and at the same time, we can get a new NPC. This new NPC usually is more accurate than the initial one. Then, a second CWT analysis may be applied and a new set of reconstruction locations could be obtained, then another calculation may be involved. In this way, the denoising result can be improved iteratively.

After zeroing some PCs, the three criteria for picking the time-domain reconstruction locations are as follows:

1) The amplitude of the reconstruction locations should be close to zero.
2) The CWT spectrum map and the temporal derivative map should have negligible energy at the reconstruction locations.
3) If the CWT spectrum is influenced by periodic signal or noise and there exists locations

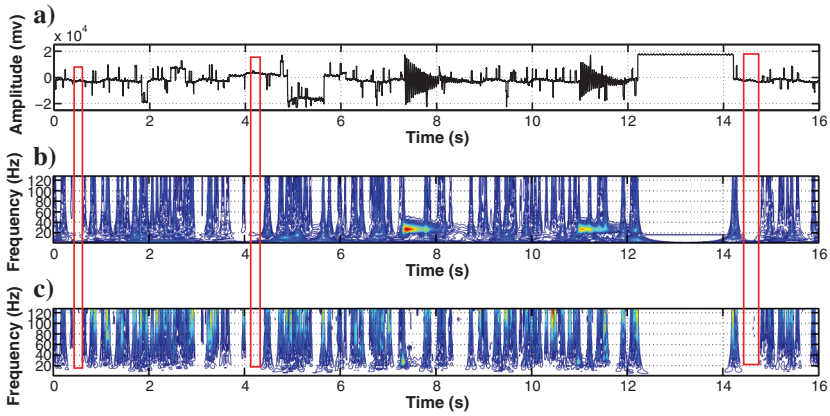

Figure 7. (a) Initial NPC, (b) time-frequency spectrum using CWT, and (c) temporal derivation of the scalogram. The frequency range is 0.02726–128 Hz. The red boxes are the chosen reconstructed locations.

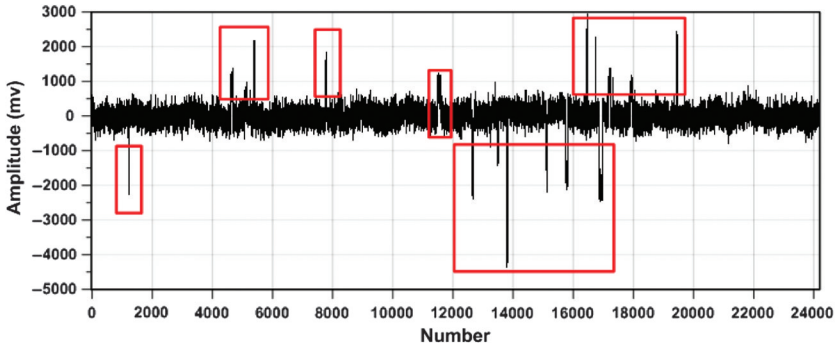

Figure 8. The NPC at reconstructed locations after the first computation. After the computation, we can check the characters of the NPC at the reconstructed locations.

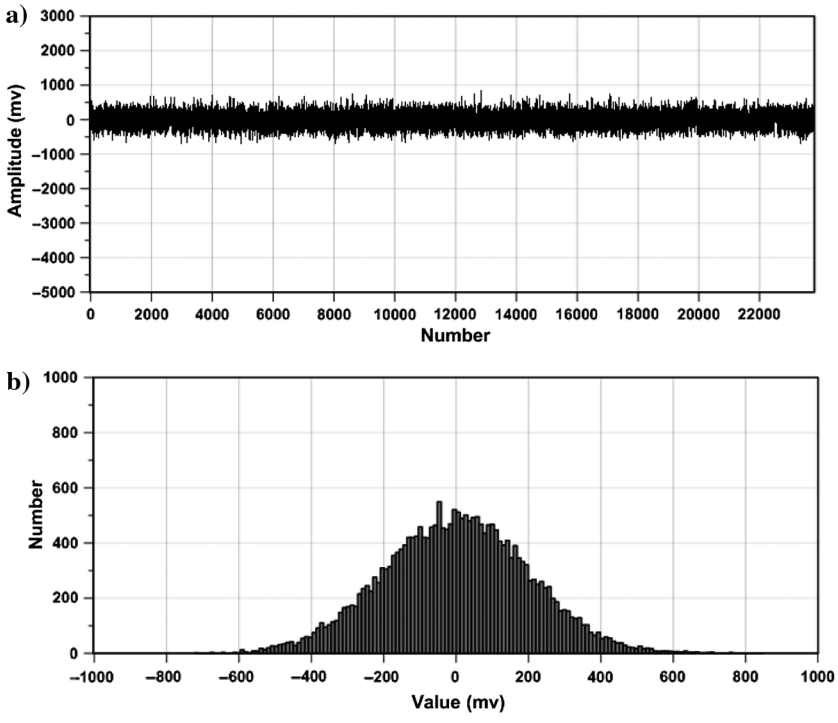

Figure 9. (a) The NPC at the reconstructed locations after the last computation and (b) distribution diagram of the noise residual.

in temporal derivative map with negligible energy (as shown in Figure 2), then these locations will be chosen.

Another synthetic example is shown below to demonstrate the denoising performance of the proposed algorithm. Actually, this one can be considered as a preprocessed signal of the previous simulated signal because they share almost the same nonperiodic noise. The sampling frequency is 65,536 Hz, and the total time length is 16 s. It contains seven frequencies, which are 1, 2, 4, 8, 16, 32, and 64 Hz, with amplitude of 1000, 500, 250, 125, 62, 31, and 15 mV, respectively. It also contains Gaussian noise with a mean square error of 200 mV and other nonperiodic noise. For land-based CSEM, strong noise, such as square, triangle, step, impulse, charging and discharging noise, or a mix of them, often exists in raw data (Tang et al., 2012). Here, we simulated different sporadic noise as shown in Figure 6.

Applying CWT analysis, we can find those candidate reconstruction locations as shown in Figure 7. We combine all these locations and construct a large over-determined system of equations. This means that what our concern is the character of the large matrix $\mathbf{A}$, or more accurately $\mathbf{A}^T\mathbf{A}$. If $\mathbf{A}^T\mathbf{A}$ is a full-rank matrix, then the solution will be stable. Besides, another parameter for stability is

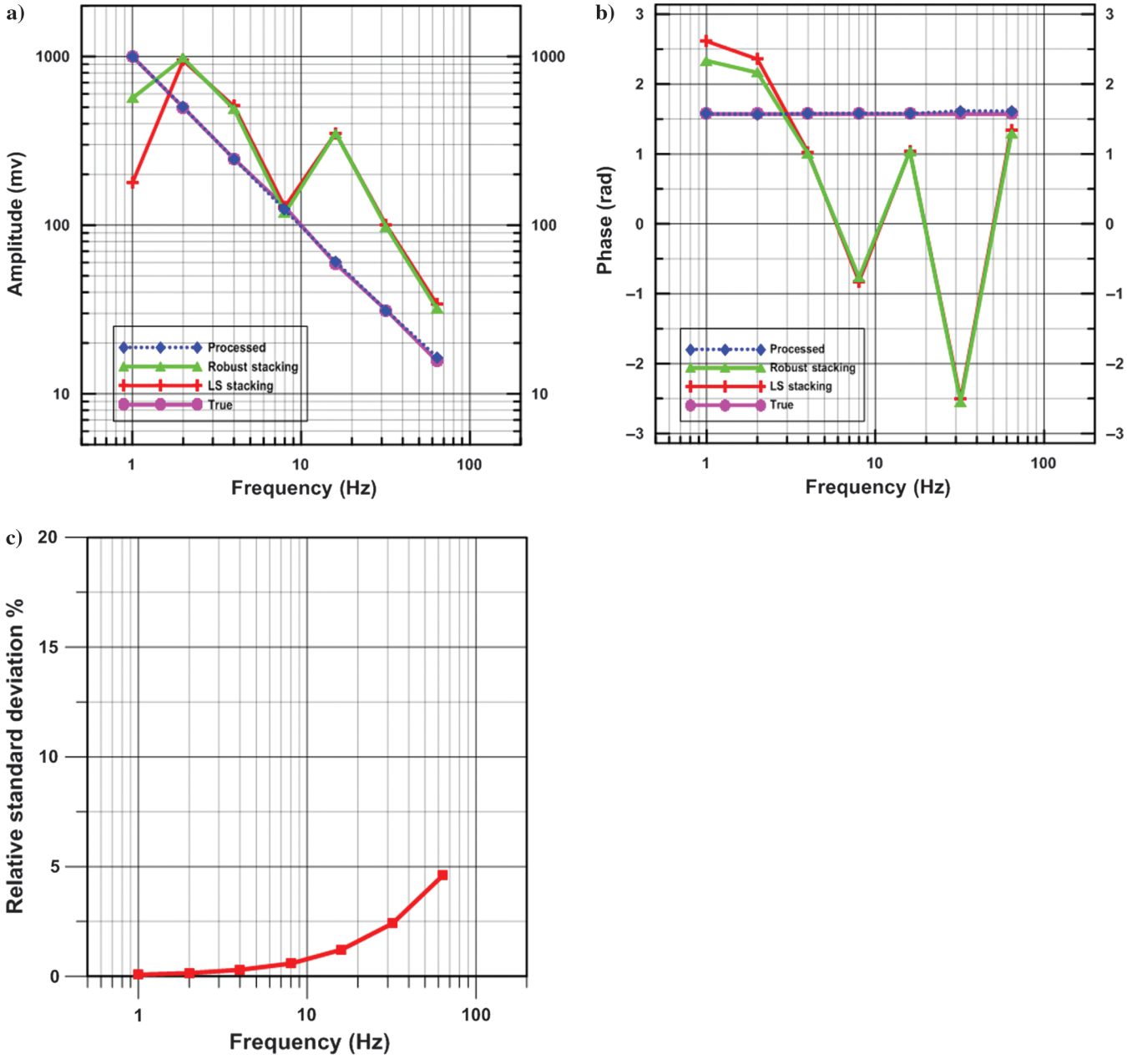

Figure 10. (a) Amplitude of the processed, least-squares stacking, robust stacking, and true data in the frequency domain. (b) Phase of the processed, least-squares stacking, robust stacking, and true data in the frequency domain. (c) RSD of the processed result in CSEM frequencies.

the variance of noise in reconstruction locations. If the chosen locations have little energy and a small variance, we can get quite good solutions for the over-determined system. Figure 8 shows the noise residuals at the reconstruction locations after the first calculation. As can be seen, there exist some bad points including impulse noise. We delete these locations and use the remaining locations.

After deleting those bad locations, another calculation is done. The noise residual at the reconstruction locations is shown in Figure 9a, which is better than the noise residual of the first calculation. As shown in Figure 9b, the distribution of noise residual follows the assumption of the Gaussian well. We also show the results with other methods as benchmarks:, which are the least-squares stacking method and the robust stacking method as shown in Figure 10a. The robust method uses the Huber psi-function. Figure 10a shows a comparison of amplitude at the CSEM frequencies. Figure 10b shows the comparison of phase in CSEM frequencies. The low number of samples is the most possible reason why the robust result looks bad in this case. Figure 10c shows the error estimation of the corresponding frequencies, which is calculated in equation 15. Figure 11a shows the result with different lengths of reconstruction locations. Here, we use same densely sampled signal, but we use shorter reconstruction time windows each time. Figure 11c shows the absolute error amplitude of different numbers of reconstruction

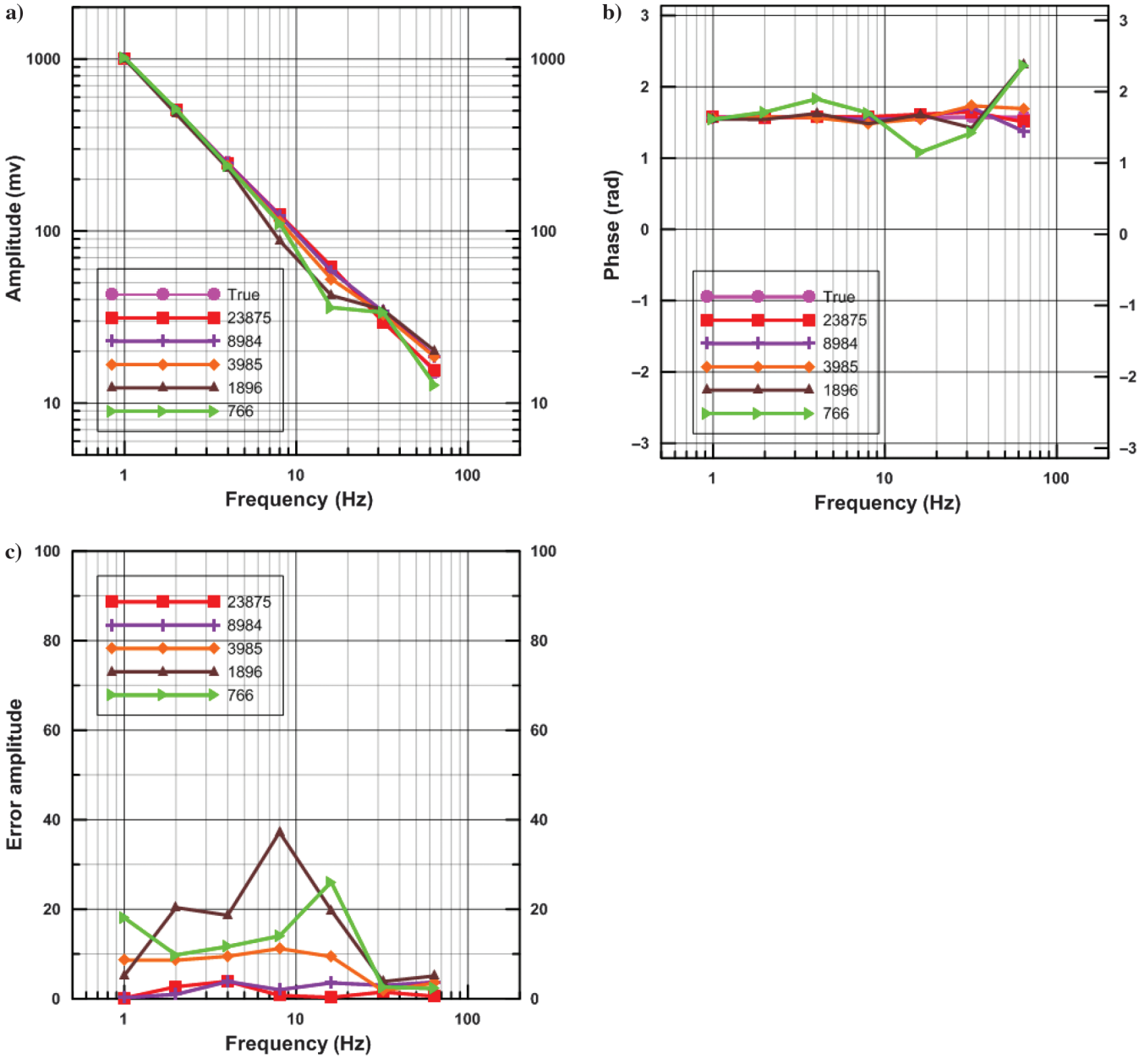

Figure 11. (a) Error amplitude with different number of reconstruction locations, (b) phase of processed results with different number of reconstruction locations, and (c) error amplitude with a different number of reconstruction locations.

points. Figure 11b shows the processed phase result of different numbers of reconstruction points. Figure 12 shows the comparison result in the time domain. It is clear that errors gradually decrease with the increasing number of reconstruction locations. The increasing number of locations can be certain to improve the variance of the solution, but it does not make certain that their error value is always smaller than lower number of locations. This is why in some frequencies there is little benefit of going from 8984 to 23,875 locations.

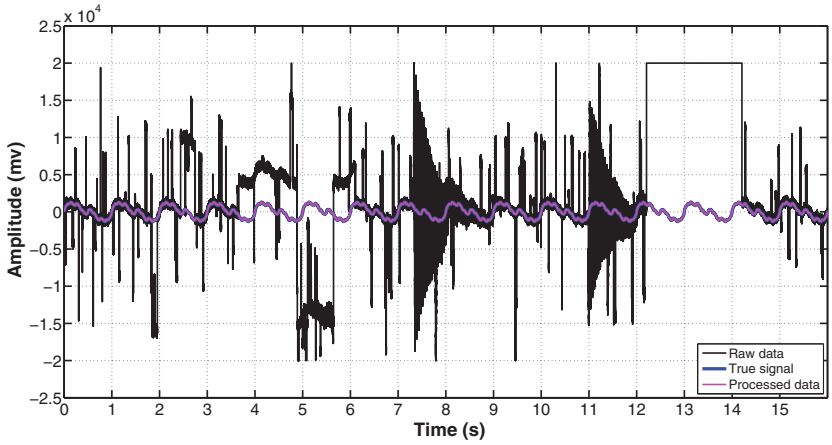

Figure 12. Comparison among raw data, true signal, and denoised signal in the time domain. Note that the blue line and the pink line are almost the same.

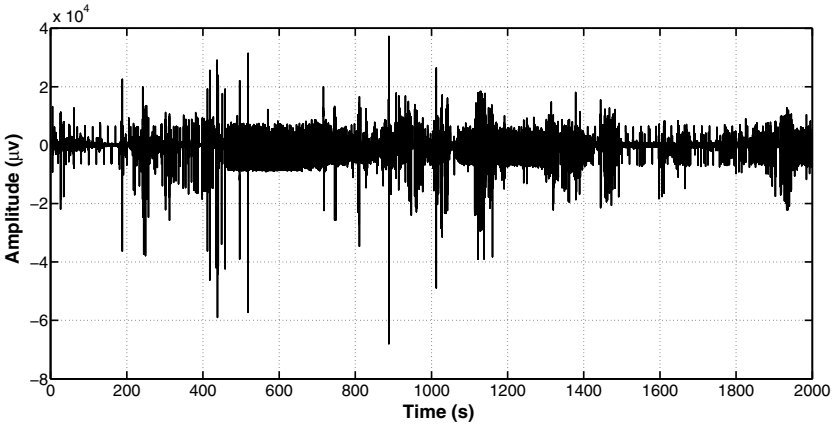

Figure 13. Real CSEM data in the time domain.

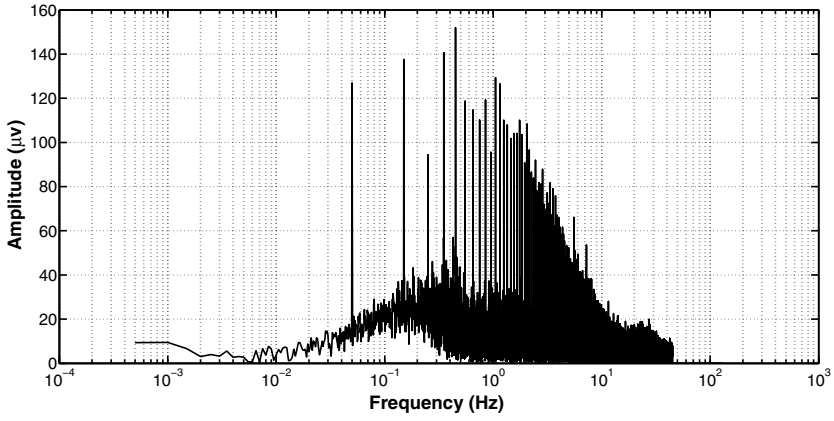

Figure 14. Spectrum of the real data.

An extreme case is that if there is only one period of source signal, it is still possible for this method to denoise the signal as long as it is sufficiently densely sampled. If we can get enough locations in the time domain, we can denoise the signal effectively because there are only seven frequencies of interest. In contrast, no statistical method that relies on multiple samples of the desired frequency components (such as robust stacking) can be applied in this case, because there is only one period of signal.

EXAMPLES

We apply the proposed method to a real CSEM data, as shown in Figure 13. It is a LOTEM data set acquired in the northern part of China for oil-gas exploration, with the source-receiver offset at 5 km. The transmitting signal is a reversing-polarity square wave with a duty cycle of 50%, whose signal type is +0-0. We concatenate all the transients to one signal in the time series because the signal is only recorded whereas the source signal is zero. It has two transients in one period, including one positive and one negative transient. The sampling frequency is 500 Hz. The length of the data is 2000 s, containing a total of 100 signal periods. The main frequency is 0.05 Hz. The frequencies of interest are 0.05 Hz and its harmonics, approximately 400 frequencies (all frequencies higher than 40 Hz have been eliminated for computation convenience). The energy of all the other frequencies is noise.

From Figure 13, we can find some periodic structures and strong noise. The spectrum of the raw data is shown in Figure 14, from which we can see that the signal is seriously corrupted by noise.

To use the proposed denoising method, we first apply CWT analysis to the raw data. As we can see in Figure 15, the CWT spectrum has been influenced by a periodic signal or in other words the CSEM signal. Therefore, we apply the zeroing criteria in equations 17 and 18 to the raw data, then we obtain a modified time-domain signal as shown in Figure 16a, in which the coefficients of 189 CSEM frequencies have been put zero. Applying CWT analysis to this modified signal, we can get the CWT spectrum and temporal derivative from which we can find some candidate reconstruction locations. This is just the first set of reconstruction locations. Actually, we can get a new NPC after every iteration. If we could find some good reconstruction locations the first time, it will be quite fast to get a satis-

factory result. In this case, we obtained the final result after only two iterations.

Based on CWT analysis, the selection of reconstruction locations can be automated by defining a certain threshold in a CWT spectrum or the sum of some frequencies' spectrum (e.g., calculate the sum of the 10–40 Hz wavelet spectra as shown in Figure 17b, then show a curve based on these sums as shown in Figure 17c). In this case, we first select some locations manually (the red boxes in Figure 16), and then we apply this criterion to the whole signal length, which is the automatic part. By statistical analysis of the sum of 10–40 Hz wavelet spectra at manually selected locations, we could find the value in those manually selected locations to show how low the value will be qualified for reconstruction locations. Then, by giving a certain threshold value as shown in Figure 17c, in which the locations with a value smaller than the given threshold are qualified locations, we can select a lot of candidate reconstruction locations automatically. As shown in Figures 18 and 19, a set of reconstruction locations, particularly many very short time windows

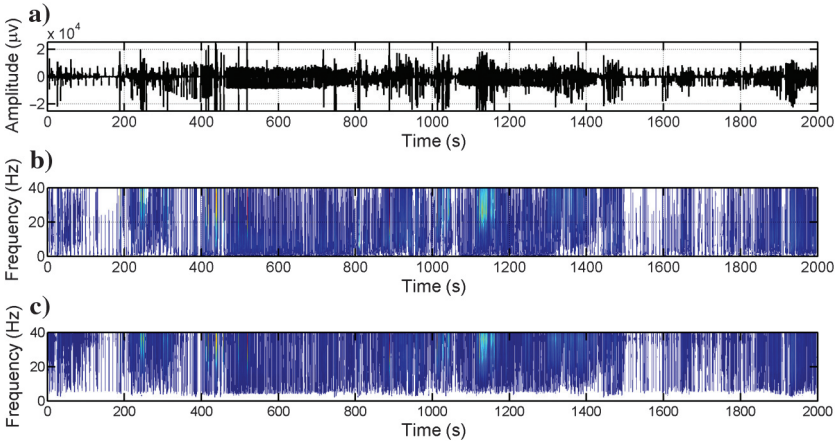

Figure 15. (a) Raw data, (b) time-frequency spectrum using CWT, and (c) temporal derivation of the scalogram. The frequency range is 0.05–40 Hz.

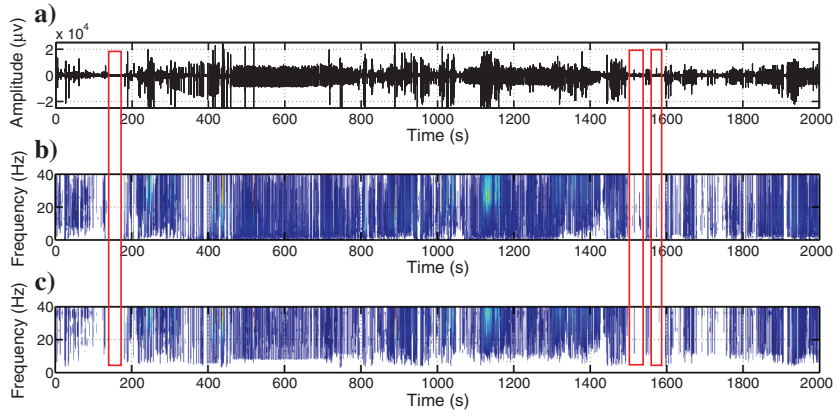

Figure 16. (a) Initial NPC after zero operating, (b) time-frequency spectrum using CWT, and (c) temporal derivation of the scalogram. The frequency range is 0.05–40 Hz. The red boxes are the candidate reconstruction locations by manually selection.

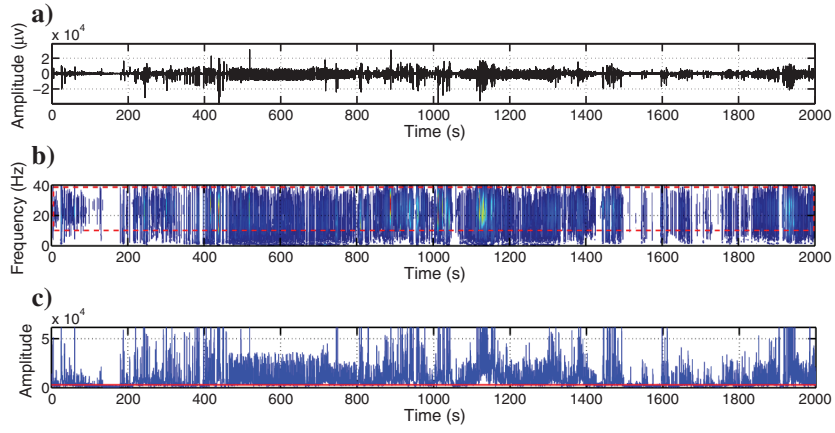

Figure 17. (a) Initial NPC after zero operating. (b) Time-frequency spectrum using CWT, frequency range is 0.05–40 Hz, the red slashed box is the area for summing wavelet spectra (10–40 Hz). (c) The curve of the sum for the 10–40 Hz wavelet spectra, the red line is a candidate threshold to select reconstruction locations.

that would not have been identified manually, are obtained by this automatic selecting process.

Figure 20a shows the processed result of amplitude at CSEM frequencies. Figure 20b shows the processed result of phase at CSEM frequencies. Figure 20c shows the error estimation of the corresponding frequencies. Figure 21 shows the comparison of the least-squares stacking, robust stacking, and processed result in the time domain. In this case, the result of robust stacking does improve some, but not quite so well. Figure 22 shows the coefficients of 0.75 Hz for 100 periods, which are quite scattered. This is a possible reason why the result for the robust method is not so good. In Figure 23a, the noise residual at the reconstruction locations is shown. We suppose in these locations, there exists only Gaussian noise or small-amplitude noise. It is obvious that the noise in these reconstruction locations follows the assumption quite well as shown in Figure 23b.

Using the same procedure, we denoise the data with 4000s length (200 periods) and 1000s length (50 periods). The processed results as shown in Figures 24 and 25 show that the proposed denoising method is quite effective.

In Figure 26a, we compare the denoising results in one figure for data with different lengths. The respective relative estimated standard deviations at CSEM frequencies are shown in Figure 26b. In Figure 26a, the processed amplitude curve of 4000s is smoother than 2000s, and the curve of 2000s is smoother than that of 1000s. By applying the same threshold criterion to the 4000s, 2000s, and 1000s data, the numbers of reconstruction locations are 34,615, 21,304, and 8328, respectively. It is clear that more qualified reconstruction locations can help to get better results. If more locations are involved, the RSD decreases, which indicates that the error in the corresponding frequency is smaller.

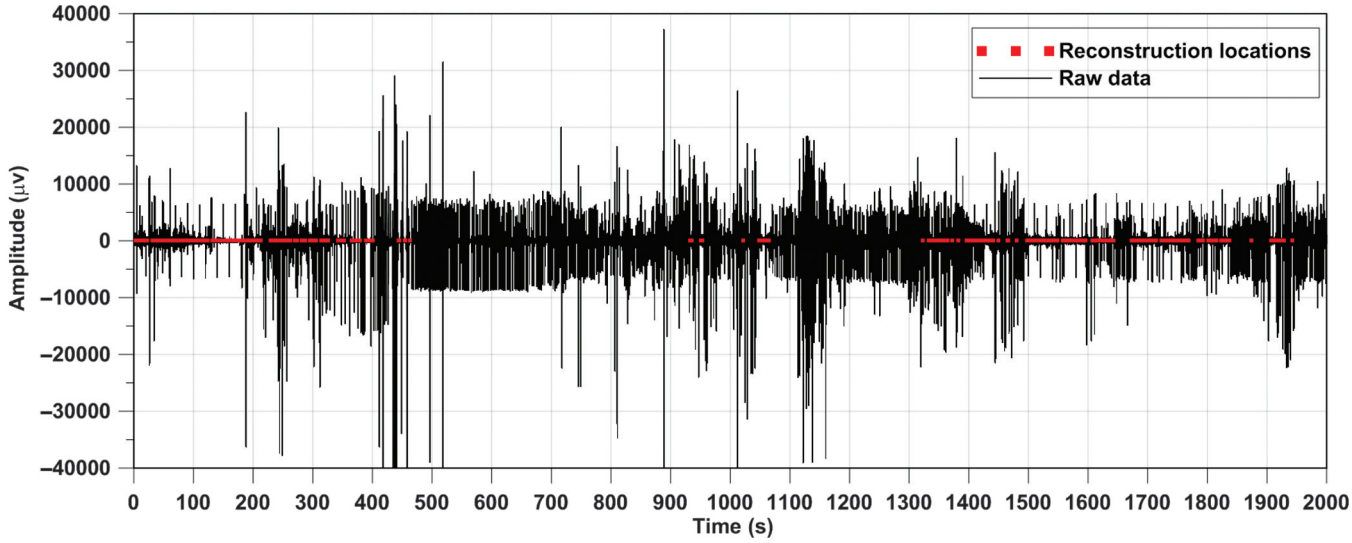

Figure 18. Reconstruction locations (in red) and raw data (in black) in the time domain.

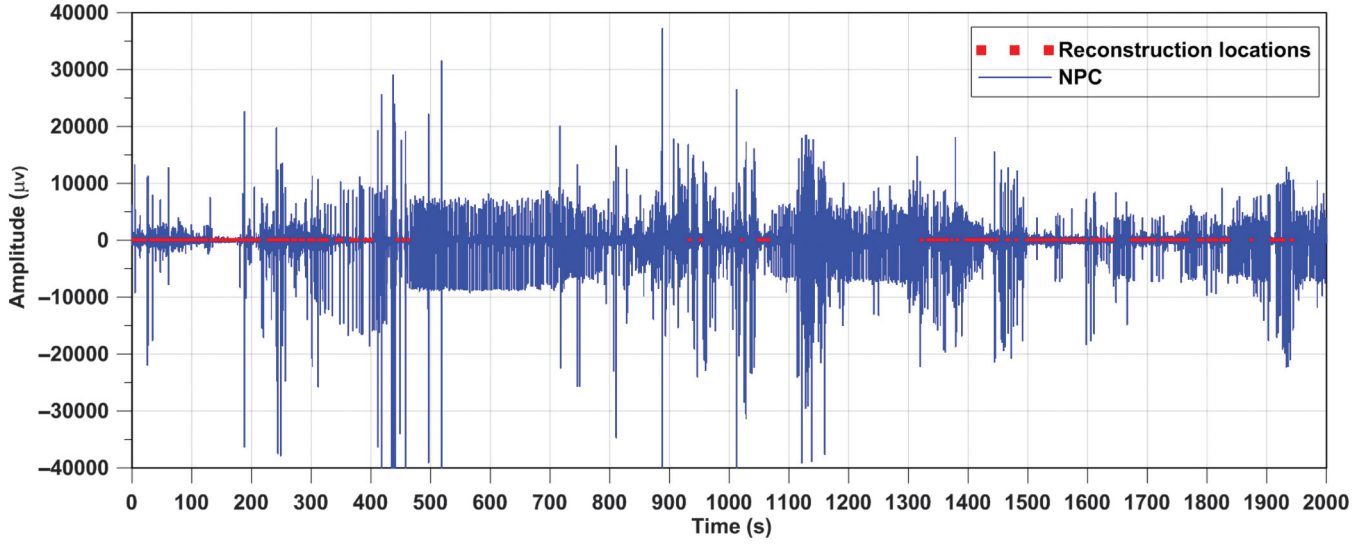

Figure 19. Reconstruction locations (in red) and NPC (in blue) in the time domain.

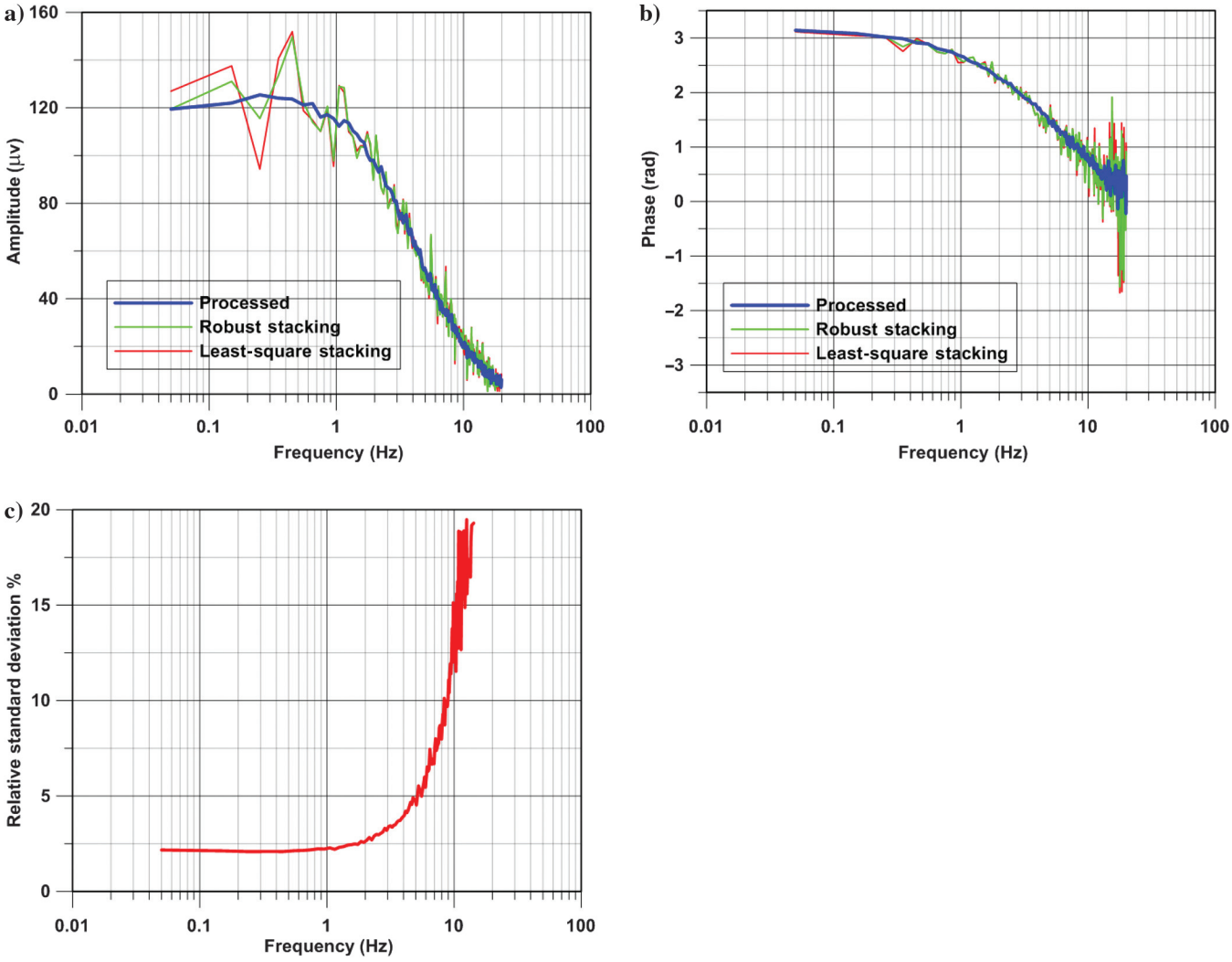

Figure 20. (a) Amplitude of the processed, least-squares stacking, and robust stacking in the frequency domain for 2000s data. (b) Phase of the processed, least-squares stacking, and robust stacking in the frequency domain. (c) RSD of the processed result in CSEM frequencies.

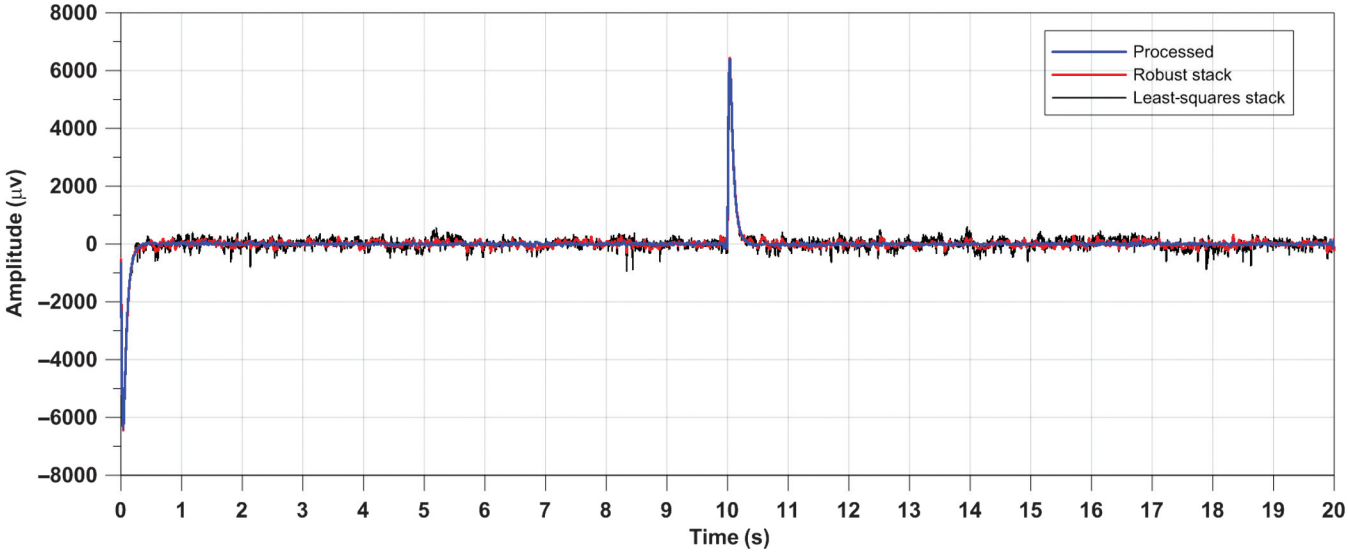

Figure 21. Processed, robust stacking, and least-squares stacking result in the time domain (shown in 20 s length).

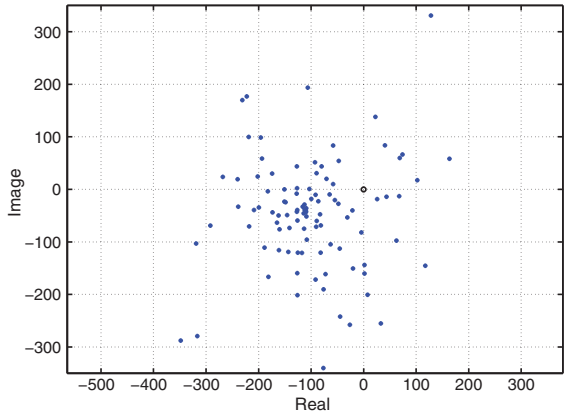

Figure 22. Coefficients of 0.75 Hz for 100 periods. The horizontal axis is real, and the vertical axis is imaginary.

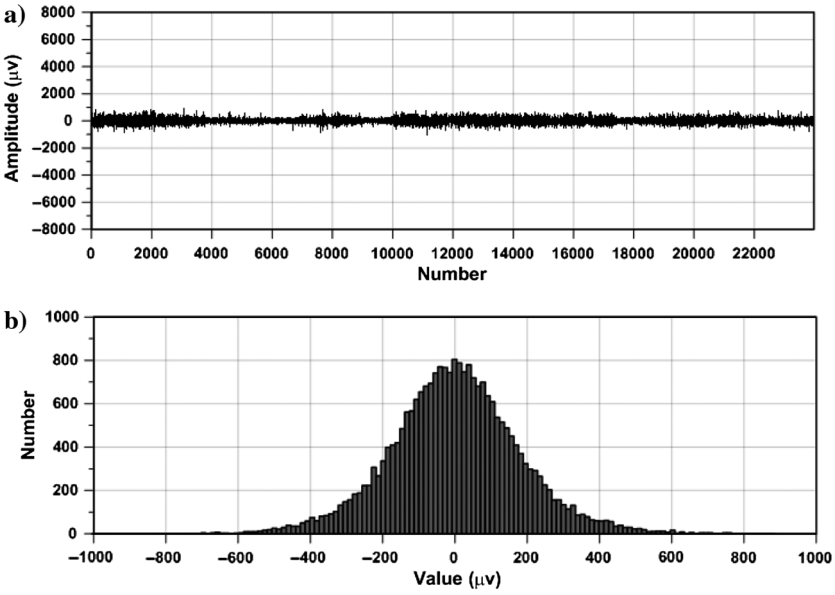

Figure 23. (a) NPC value in the reconstruction locations after processing and (b) distribution diagram of the noise residual.

DISCUSSIONS

The reason why we put most of the desired frequencies to zero for the CWT analysis is that the CWT analysis is based on NPCs. In many cases, the periodic CSEM signal can mask the structure of the NPC (including nonperiodic noise and Gaussian noise). Zeroing the signal at the frequencies of interest aims to get an initial guess of NPC. Besides, we have given a criterion for zeroing CSEM frequency coefficients. After each computation, a new NPC can be obtained. The CWT analysis is then applied to the new NPC. Under such conditions, we can find those good locations more accurately. Our main goal is to remove NPC from the data (although as a by-product, we also estimate the spectrum of Gaussian white noise at the frequencies of interest during the inversion). The method works if the influence of the NPC is inhomogeneous. The key principle of this method is to gather all the time-domain locations with low-amplitude noise (qualified reconstruction locations), and then by the presented inversion algorithm, to transform all these information into frequency coefficients.

For the robust stacking method, different weighting functions may lead to different results. The result is quite related to the weighting function that we choose. In our method, the problem has been transferred to find a time of low noise because what we do in this method is to find small-value places of NPC. Besides, for the statistic method, it is usually to operate on time windows of data that are an integer multiple of the signal period, whereas in our method, the signal is divided into much smaller pieces. It is unnecessary to operate on time windows of fixed length related to the signal period, in which we can construct the inversion problem by combining all fragmented locations together into a large over-determined equation. Thus, it is more flexible to process CSEM data via the presented framework. Furthermore, over-sampling the signal without adding recording time (i.e., sampling much more densely than what is required for correctly recording the frequencies of interest) can

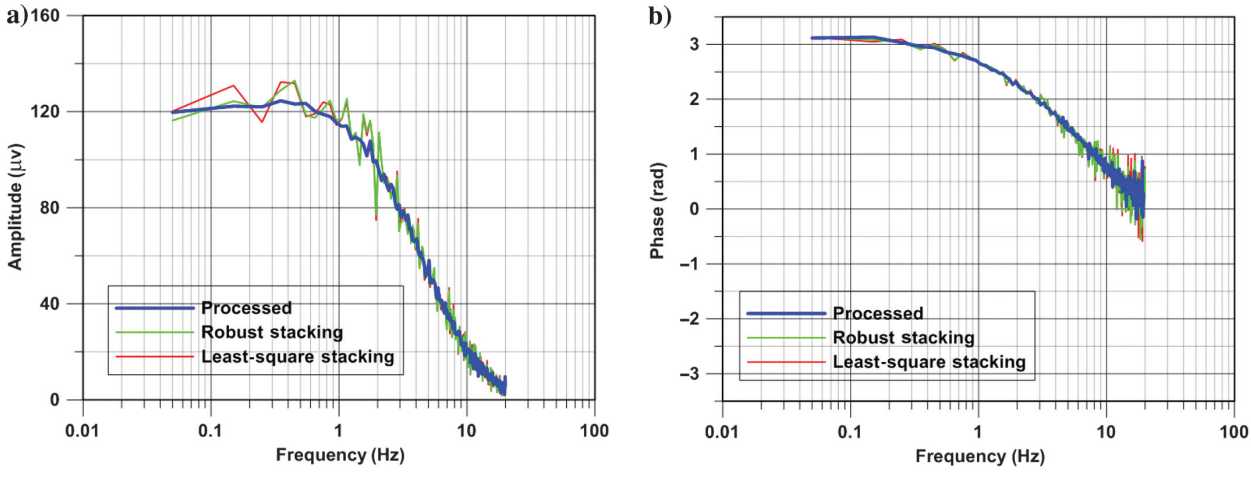

Figure 24. (a) Amplitude of the processed, least-squares stacking, and robust stacking in the frequency domain for 4000s data. (b) Phase of the processed, least-squares stacking, and robust stacking in the frequency domain.

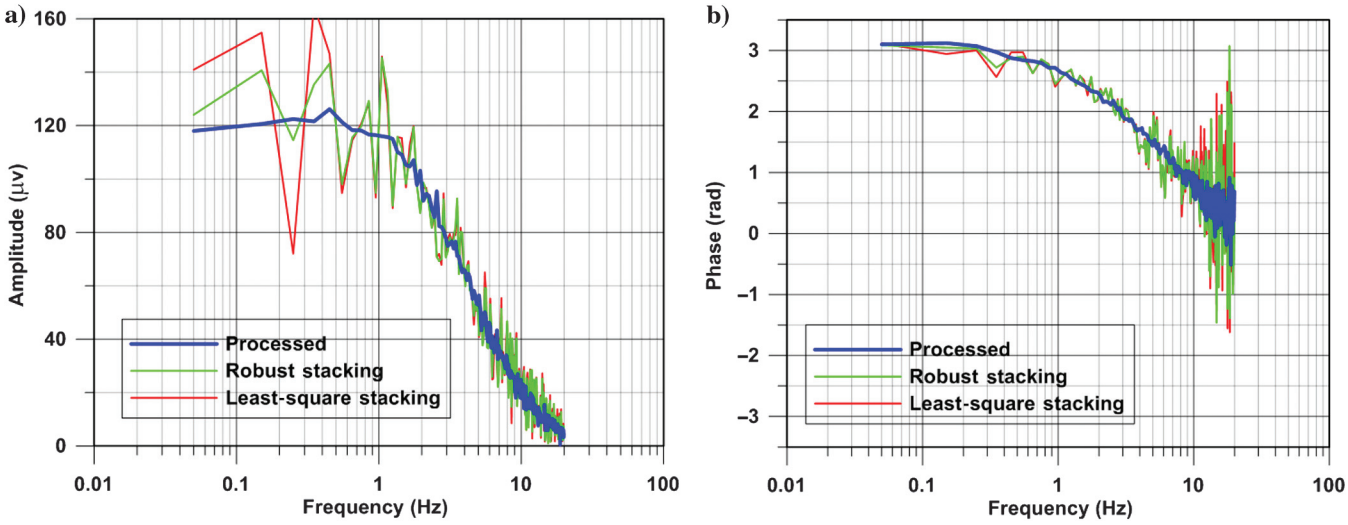

Figure 25. (a) Amplitude of the processed, least-squares stacking, and robust stacking in the frequency domain for 1000s data. (b) Phase of the processed, least-squares stacking, and robust stacking in the frequency domain.

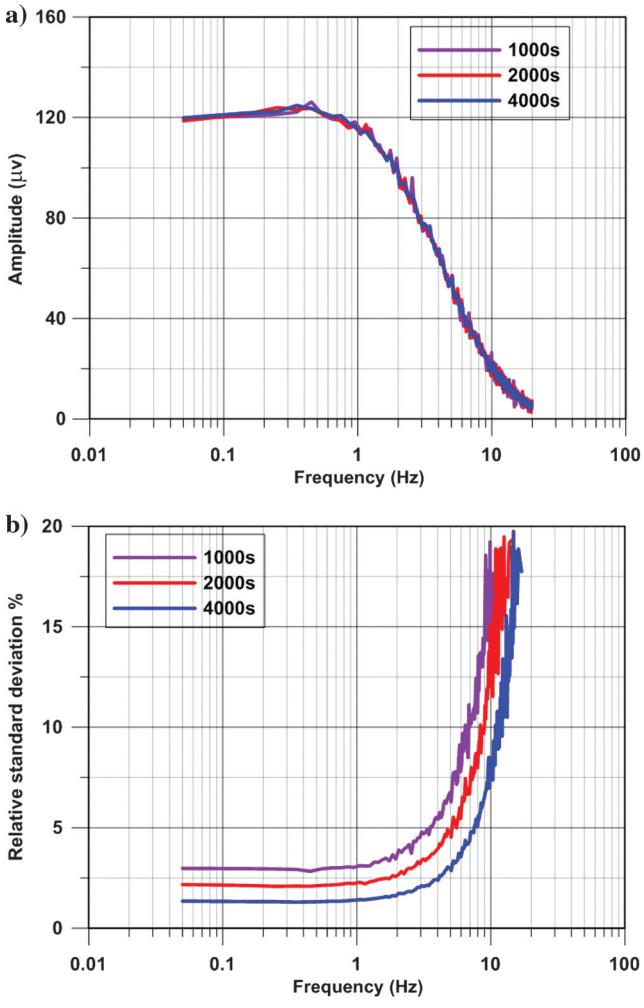

Figure 26. (a) Amplitude of the processed result for 1000s, 2000s, and 4000s data. (b) RSD of the processed result for 1000s, 2000s, and 4000s at CSEM frequencies.

help to increase the accuracy of the solution and improve the denoising result.

CONCLUSION

The NPC is a type of strong and irregular noise existing in CSEM data. It is difficult to remove such noise using conventional algorithms, which hinders the wide application of the CSEM method due to the low signal-to-noise ratio of many data sets. We have proposed a novel method for effectively attenuating the NPC that includes nonperiodic noise and Gaussian noise. The NPC can be accurately estimated by least-squares inversion. To formulate the inverse problem, we need to first detect a set of time-domain locations where only Gaussian white noise and PCs are presented. The zero preprocessing is helpful to reduce the influence of periodic signal on CWT analysis. The temporal derivative of the time-frequency spectrum that is generated using CWT can serve as an effective indicator for picking the time-domain reconstruction locations. The real CSEM data demonstrate the very successful performance of the proposed methodology. Besides, many harmonics of signal are also denoised in this method, which can be further used.

ACKNOWLEDGMENTS

We would like to thank S. Fomel and J. Tang for their technical advice on the proposed method, R. Streich, K. MacLennan, N. Cuevas, and one anonymous reviewer for the constructive suggestions that greatly improved the manuscript, and J. He for the endless help throughout this research. We are also grateful to Z. He for providing the real data. The research is supported by the National Science Foundation of China (grant no. 41227803) and Guangxi Key Laboratory of Hidden Metallic Ore Deposits Exploration (grant no. 13-051-19). Y. Chen was partially supported by the starting fund at Zhejiang University.

REFERENCES

Bracewell, R., 2000, The Fourier transform & its applications: McGraw Hill.
Chen, J., B. Heincke, M. Jegen, and M. Moorkamp, 2012, Using empirical mode decomposition to process marine magnetotelluric data: Geophysical

Journal International, **190**, 293–309, doi: 10.1111/j.1365-246X.2012.05470.x.
Chen, Y., 2016, Dip-separated structural filtering using seislet thresholding and adaptive empirical mode decomposition based dip filter: Geophysical Journal International, **206**, 457–469, doi: 10.1093/gji/ggw165.
Chen, Y., 2017, Fast dictionary learning for noise attenuation of multidimensional seismic data: Geophysical Journal International, **209**, 21–31, doi: 10.1093/gji/ggw492.
Chen, Y., 2018, Non-stationary least-squares complex decomposition for microseismic noise attenuation: Geophysical Journal International, **213**, 1572–1585, doi: 10.1093/gji/ggy079.
Chen, Y., and J. Ma, 2014, Random noise attenuation by f-x empirical mode decomposition predictive filtering: Geophysics, **79**, no. 3, V81–V91, doi: 10.1190/geo2013-0080.1.
Goldstein, M. A., and D. W. Strangway, 1975, Audio-frequency magnetotellurics with a grounded electric dipole source: Geophysics, **40**, 669–683, doi: 10.1190/1.1440558.
He, J., 2010, Wide field electromagnetic sounding methods: Journal of Central South University (Science and Technology), **41**, 1055–1072.
Huang, N. E., Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, 1998, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis: Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, **454**, 903–995, doi: 10.1098/rspa.1998.0193.
Maclennan, K., and Y. Li, 2013, Denoising multicomponent CSEM data with equivalent source processing techniques: Geophysics, **78**, no. 3, E125–E135, doi: 10.1190/geo2012-0226.1.
Reninger, P. A., G. Martelet, J. Perrin, and Y. Chen, 2011, Singular value decomposition as a denoising tool for airborne time domain electromagnetic data: Journal of Applied Geophysics, **75**, 264–276, doi: 10.1016/j.jappgeo.2011.06.034.
Seigel, H. O., 1959, Mathematical formulation and type curves for induced polarization: Geophysics, **24**, 547–565, doi: 10.1190/1.1438625.
Strack, K. M., T. H. Hanstein, and H. N. Eilenz, 1989, LOTEM data processing for areas with high cultural noise levels: Physics of the Earth and Planetary Interiors, **53**, 261–269, doi: 10.1016/0031-9201(89)90010-1.
Strang, G., 2007, Computational science and engineering: Wellesley-Cambridge Press.
Streich, R., M. Bechen, and O. Ritter, 2013, Robust processing of noisy land-based controlled-source electromagnetic data: Geophysics, **78**, no. 5, E237–E247, doi: 10.1190/geo2013-0026.1.
Tang, J., Z. Xu, X. Xiao, and J. Li, 2012, Effect rules of strong noise on magnetotelluric (MT) sounding in the Luzong ore cluster area: Chinese Journal of Geophysics, **55**, 4147–4159, doi: 10.6038/j.issn.0001-5733.2012.12.027.
Welch, P. D., 1967, The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms: IEEE Transactions on Audio and Electroacoustics, **15**, 70–73, doi: 10.1109/TAU.1967.1161901.
Willen, D. W., 2010, Method for wavelet denoising of controlled source electromagnetic survey data: E.P. Patent 1, 922, 567.
Yang, Y., 2016, A de-noising method for periodic CSEM data based on inverse discrete Fourier transform (IDFT), continuous wavelet transform (CWT) and over-determined equations (ODEs): 86th Annual International Meeting, SEG, Expanded Abstracts, 1039–1042.
Ziolkowski, A., and D. Wright, 2012, The potential of the controlled source electromagnetic method: A powerful tool for hydrocarbon exploration, appraisal, and reservoir characterization: IEEE Signal Processing Magazine, **29**, 36–52, doi: 10.1109/MSP.2012.2192529.