AN IMPROVED ROBUST THRESHOLD FOR VARIATIONAL MODE DECOMPOSITION BASED DENOISING IN THE FREQUENCY-OFFSET DOMAIN

YUANYUAN MA and SIYUAN CAO

The State Key Laboratory of Petroleum Resources and Engineering, China University of Petroleum, Changping 102249, Beijing, P.R. China. myy_upc@163.com ; siyuan.caoff@gmail.com

(Received June 4, 2018; revised version accepted April 15, 2019)

ABSTRACT

Ma, Y.Y. and Cao, S.Y., 2019. An improved robust threshold for variational mode decomposition based denoising in the frequency-offset domain. *Journal of Seismic Exploration*, 28: 277-305.

We proposed a novel robust denoising method using variational-mode decomposition (VMD) and the detrended fluctuation analysis (DFA) in the frequency-offset (*f-x*) domain, named robust DFA-VMD. DFA is mainly introduced to solve the problem that VMD requires the number of modes to be predefined. The scaling exponent obtained by DFA is a robust metric to measure the long-range correlations and can be used to adjust the number of intrinsic mode functions (IMFs) automatically. To reconstruct the denoised signal, a scaling exponent is also used as a threshold to identify and remove the noisy modes. We define a novel robust threshold of random noise in seismic data, because the predefined noise boundaries for other time series cannot perform perfectly when dealing with seismic data. The proposed robust DFA-VMD is an almost parameters-free denoising approach and we apply it in the (*f-x*) domain for seismic denoising. We have verified its performance by comparing it with the results from several other methods including (*f-x*) deconvolution and the conventional DFA-VMD. Two synthetic examples and three field-data examples revealed the effectiveness of the proposed approach in applications to random and coherent noise attenuation.

KEY WORDS: denoising, frequency-offset domain, adaptive filtering, variational mode decomposition (VMD), detrended fluctuation analysis (DFA).

0963-0651/19/$5.00

INTRODUCTION

Signal decomposition such as short-time Fourier transform (Allen, 1977), wavelet transform (Grossmann and Morlet, 1984), and S-transform (Stockwell et al., 1996) have been successfully applied in signal denoising. Empirical mode decomposition (EMD) developed by Huang et al. (1998) is an entirely data-driven method for signal decomposition. It has been widely used in seismic attributes analysis (Magrin-Chagnolleau and Baraniuk, 1999) and noise attenuation (Chen et al., 2014). However, EMD suffers from several drawbacks (Mandic et al., 2013). Mode mixing is one of the most severe problems of the EMD algorithm (Huang and Wu, 2008). Different frequency components are mixed in one or more intrinsic mode functions (IMFs), which is the main obstacle to interpret the decompositions. Many approaches are introduced to overcome this negative feature, such as ensemble EMD (EEMD) (Wu and Huang, 2009) and complete ensemble EMD (CEEMD) (Torres et al., 2011). Moreover, it is a crucial challenge to explain the meaning of each IMF or determine which IMFs correspond to noisy oscillation (Wu and Huang, 2004). The theoretical foundation of EMD-based denoising need to be further improved.

Variational Mode Decomposition (VMD) (Dragomiretskiy and Zosso, 2014) is a recently proposed non-recursive signal decomposition technique to analyze non-linear and non-stationary time series. VMD exhibits advanced features compared with the classic EMD. Firstly, in contrast to EMD, VMD adaptively decomposes a signal into an ensemble of band-limited IMFs and all modes are extracted concurrently. Each mode resulting from VMD is considered almost compact around a corresponding center frequency (Upadhyay and Pachori, 2015). In essence, VMD is a generalization of the classic Wiener filter into multiple and adaptive bands. Secondly, the modes from VMD are less sensitive to noise than those from EMD, in which the first two modes always contain more residual noise. Thirdly, the VMD has lower computational cost compared with CEEMD. These features are helpful to accurately capture components and reconstruct the filtered signal. VMD has been successfully used in time-frequency analysis (Liu et al., 2016), ground-roll attenuation (Liu et al., 2015) and sedimentary pattern characterization (Li et al., 2016) to meet the challenges of seismic application. Recently, VMD has been used to attenuate seismic random noise (Liu et al., 2017; Yu and Ma, 2018). In spite of its considerable success, the selection of predefined parameters for VMD and the determination of useful modes are two challenges, which are the motivation of our research.

The VMD algorithm depends on several essential parameters, in which the number of modes to be extracted is one of the most important parameters (Dragomiretskiy and Zosso, 2014). The number of modes needs to be predefined and its value has a strong influence on the efficiency of decomposition. An inappropriate value (underbinning or overbinning) can be detected and corrected in post-processing by checking the spectral overlap or orthogonality between modes, or by looking at the residuals not accounted for by any mode (Dragomiretskiy and Zosso, 2014). Unfortunately, this solution is beyond the scope of the VMD algorithm. In the actual application of VMD, the number is usually selected based on experience. Consequently, an automatic adjustment method needs to be developed to overcome this classical limitation.

To select the IMFs automatically, the detrended fluctuation analysis (DFA) (Peng et al., 1994) was used as a threshold to adjust VMD automatically. Liu et al (2016) proposed a denoising method named DFA-VMD to remove white Gaussian noise (WGN) from electrocardiograph (ECG) and electrocardiogram (EEG) signal. DFA is a powerful tool to measure long-range correlations and fractal scaling properties for non-stationary time series. It has been successfully applied to different fields such as biomedicine (Peng et al., 1995; Bryce and Sprague, 2012), meteorology (Ivanova and Ausloos, 1999), economics (Kantelhardt et al., 2002), and ethnology (Telesca et al., 2007). One significant advantage of DFA is that it avoids the spurious detection of apparent long-range correlations that are artifacts of non-stationary series (Peng et al., 1995). Li et al. (2017) applied DFA-VMD in seismic denoising. DFA can be used to characterize different components because the seismic data is a combination of harmonic components, while the noise component is random and uncorrelated. However, the predefined noise boundaries (Peng et al., 1994) for other time series cannot perform perfectly when dealing with seismic data.

We propose a novel robust threshold for DFA-VMD denoising in the frequency-offset (f-x) domain. DFA is mainly introduced to adjust the number of IMFs automatically. To reconstruct the denoised signal, a scaling exponent is also used as a threshold to identify and remove the noisy modes. We defined a robust threshold of random noise in seismic data and verified it using a lot of synthetic-models. The novel threshold is approximately equal to the scaling exponent of the upper effective frequency of seismic data. The robust DFA-VMD is an almost parameters-free approach, and we apply it on constant-frequency slices in the (f-x) domain for seismic denoising. We have compared it to several other methods including (f-x) deconvolution and conventional DFA-VMD. Two synthetic examples and three field-data

examples revealed the excellent performance of the proposed approach in applications for random and coherent noise attenuation.

The paper is organized as follows: we first give a brief introduction about the VMD and its limitations. Then the VMD and DFA method are reviewed. Next, we defined a novel threshold denoising method named robust DFA-VMD and verified it using a lot of synthetic-models. Then we apply the improved denoising method in the (*f-x*) domain to remove seismic random and coherent noise. Two synthetic examples and three field examples are used to verify the superior performance of the proposed approach over (*f-x*) deconvolution and conventional (*f-x*) DFA-VMD.

METHOD

Variational mode decomposition

VMD is an adaptive and non-recursive signal-decomposition approach. The aim of VMD is to decompose an input signal $f(t)$ into a discrete number of band-limited sub-signals or modes $\mu_k(t)$, where each mode is mostly compact around their respective center frequency w_k (Upadhyay and Pachori, 2015). The bandwidth of each mode can be estimated by solving the constrained variational optimization problem with the following scheme:

1) Compute the related analytic signal for each mode through a Hilbert transform to build a unilateral frequency spectrum.

2) Shift the frequency spectrum of each mode to the baseband by mixing with an exponential tuned to the respective estimated center frequency.

3) The bandwidth of each mode can be estimated through the H^1 Gaussian smoothness of the demodulated signal, i.e., the squared $L2$-norm of the gradient.

The resulting constrained variational problem can be describe as follows (Dragomiretskiy and Zosso, 2014):

$$\min_{\{u_k\},\{\omega_k\}}\left\{\sum_k\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right]e^{-j\omega_k t}\right\|_2^2\right\}$$
$$s.t.\quad \sum_k u_k(t)=f(t) \qquad (1)$$

where $\{u_k\}=\{u_1,u_2,\mathrm{L}\ ,u_k\}$ represent the set of all modes and $\{\omega_k\}=\{\omega_1,\omega_2,\mathrm{L}\ ,\omega_k\}$ are their center frequencies, respectively. δ is the Dirac distribution and $\sum_k=\sum_{k=1}^{K}$ is the summation over all modes.

In order to render the problem unconstrained, a quadratic penalty term and Lagrangian multipliers, λ are introduced. The augmented Lagrangian, L is given by

$$L(\{u_k\},\{\omega_k\},\lambda)=\alpha\sum_k\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right]e^{-j\omega_k t}\right\|_2^2 + \left\|f(t)-\sum_k u_k(t)\right\|_2^2+\left\langle\lambda(t),f(t)-\sum_k u_k(t)\right\rangle, \qquad (2)$$

where α denotes the balancing parameter of the data-fidelity constraint. Eq.(2) is then solved with the alternate direction method of multipliers (ADMM) (Hestenes, 1969). All the modes in the frequency and time domain are obtained as follows:

$$\hat{u}_k^{n+1}(\omega)=\frac{\hat{f}(\omega)-\sum_{i=1,i\neq k}^{K}\hat{u}_i(\omega)+\frac{\hat{\lambda}(\omega)}{2}}{1+2\alpha(\omega-\omega_k)^2}, \qquad (5)$$

$$\hat{u}_k^{n+1}(t)=\Re\left\{F^{-1}\left(\hat{u}_k^{n+1}(\omega)\right)\right\}, \qquad (4)$$

where $\hat{f}(\omega),\hat{u}_i(\omega)$, and $\hat{\lambda}(\omega)$ are the Fourier transforms of $f(t),u_i(t)$, and $\lambda(t)$, respectively. It is clear that the optimal $\hat{u}_i(\omega)$ is directly renewed by Wiener filtering in the Fourier domain, which makes VMD much more robust to sampling and noise. $F^{-1}(\bullet)$ represents the inverse Fourier transform and $\Re(\bullet)$ denotes the real part of the signal. The modes in the time domain can be

obtained as the real part of the inverse Fourier transform of this filtered analytic signal.

The center frequency ω_k of each mode can be obtained as follows, its optimization also takes place in Fourier domain:

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega \left|\hat{u}_k(\omega)\right|^2 d\omega}{\int_0^\infty \left|\hat{u}_k(\omega)\right|^2 d\omega}, \tag{5}$$

which locates the new ω_k at the center of gravity of the corresponding mode's power spectrum.

The VMD algorithm depends on several parameters, especially the number of modes to be extracted, K. A larger value of K possibly results in frequency mixing, whereas a lower value decreases the focus of the time-frequency representation. Usually, the number K is selected based on practical experience. In addition, the main principle of the existing VMD-based or EMD-based denoising methods is to distinguish and exclude the noise-only IMFs. The DFA algorithm is introduced to solve these problems in this paper.

The robust DFA thresholding

The autocorrelation function determines the dependency of two sample points in a time series, and it can measure the correlation between a signal and its time-shifted version. The scaling exponent, also known as Hurst exponent or self-similarity, is an important way to estimate the strength of the autocorrelation, which can indicate long-range correlations, mild or wild randomness (Hurst, 1951). Unfortunately, the Hurst-exponent calculation causes spurious detection when the signal has non-stationary properties. Thus, DFA is developed to estimate the long-range correlations for non-stationary time series.

For a given time series $\{x(t), t = 1,2,...,N\}$, the DFA involves the following steps:

1) The first step of the algorithm is to find the integrated time series $y(k)$ by removing the average of $x(t)$ as follows:

$$y(k)=\sum_{t=1}^{k}\left[x(t)-\langle x\rangle\right],1\le k\le N, \tag{6}$$

where

$$\langle x\rangle=\frac{1}{N}\sum_{t=1}^{N}x(t).$$

2) The integrated series $y(k)$ is then divided into a set of non-overlapping boxes of equal length *n*. In each box, a least-squares fitting is applied to obtain the local trend $y_n(k)$. If the order of the polynomial fitting is chosen to be $l=2$, $y_n(k)$ can be defined as follows:

$$y_n(k)=a_n k^2+b_n k+c_n, \tag{7}$$

3) Next, we detrend the integrated time series by subtracting the local trend in each box. The root-mean-square fluctuation function $F(n)$ is calculated as follows:

$$F(n)=\sqrt{\frac{1}{N}\sum_{k=1}^{N}\left[y(k)-y_n(k)\right]^2}, \tag{8}$$

4) If the series $x(t)$ is long-range correlated, $F(n)$ will increase with the box size *n*. Under such conditions, the fluctuations can be characterized by a scaling exponent h with a power-law:

$$F(n)\propto n^h, \tag{9}$$

Subsequently, the slope *h* of the line relating $\log(F(n))$ to $\log(n)$ is calculated by linear least-squares regression. It is determined by

$$\ln(F(n))=h\ln(n)+C \tag{10}$$

The scaling exponent *h* is a measure for long-range correlations of non-stationary signals (Peng et al., 1994). When $0.5<h<1.0$, the signal presents persistent long-range power-law correlations such that large (compared to the average) fluctuations are followed by large ones. In contrast, $0<h<0.5$ is called "anti-correlated", in which large fluctuations are likely to be followed by small ones. For $1.0<h$, the long-range correlations exist but do not reveal a power-law form. The most important reason which makes the scaling exponent a reliable metric is that the values $h=0.5, h=1.0$ and $h=1.5$ indicate completely uncorrelated white noise, pink

noise and Brownian noise, respectively (Peng et al., 1994). The conventional DFA-VMD denoising method is theoretically based on these special cases and the workflow is shown in Fig. 1. The DFA-VMD denoising method can be achieved as follows:

1) Calculate the scaling exponent h_0 of the input data, and obtain the expected number J of useful IMFs according to the following equation:

$$J = \begin{cases} 1, & h_0 \le 0.8 \\ 2, & 0.8 < h_0 \le 1.0 \\ 3, & 1.0 < h_0 \le 1.2 \\ 4, & 1.2 \le h_0 \end{cases} \tag{11}$$

2) Decompose the input data from K=1, and calculate the scaling exponent $h_{1:K}$, K=1,2,3,... of each IMF. Record the real number C of useful IMFs whose scaling exponents are bigger than the threshold θ.

$$C = \left\{ IMF \middle| h_{1:K} \ge \theta, K = 1,2,3,... \right\} \tag{12}$$

3) Repeat step 2 until the number of useful IMFs satisfies the expectation, i.e., the best K for VMD can be obtained when C equals to the expected number J as follows:

$$K = \left\{ K \middle| C = J, K = 1,2,3,... \right\}, \tag{13}$$

4) Reconstruct the denoised signal S using the useful IMFs as follows:

$$s = \sum IMF_i, i = \left\{ i \middle| h_i \ge \theta \right\} \tag{14}$$

The DFA algorithm depends on the selection of the box size n, which is data dependent $l + 2 \le n \le N/4$ (Peng et al., 1995). However, the ranges $4 \le n \le 16$ and $16 \le n \le 200$ are the most popular and reliable linear regions for h calculation, which was successfully studied in biomedical signal processing (Jospin et al., 2007; Leistedt et al., 2007; Leitet et al., 2010). In this paper, we apply the DFA with $4 \le n \le 16$.

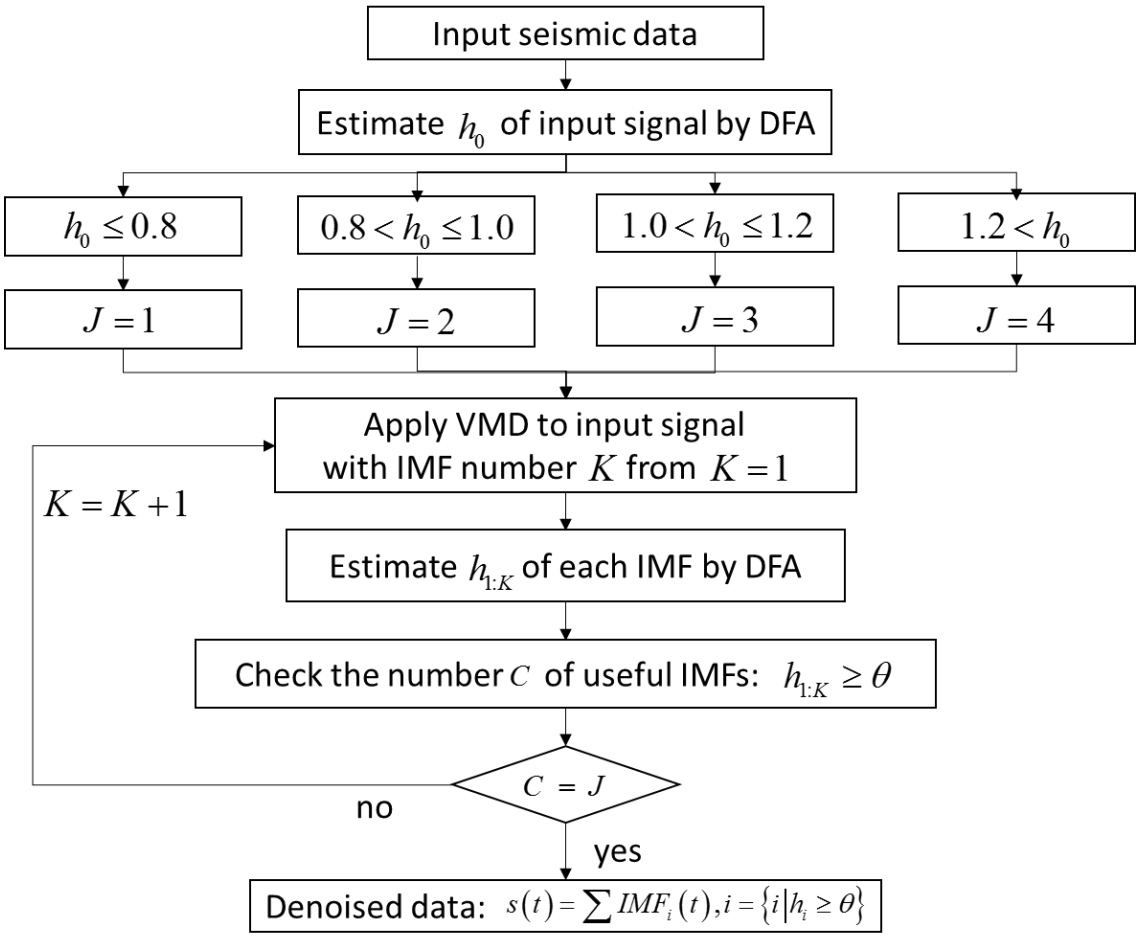
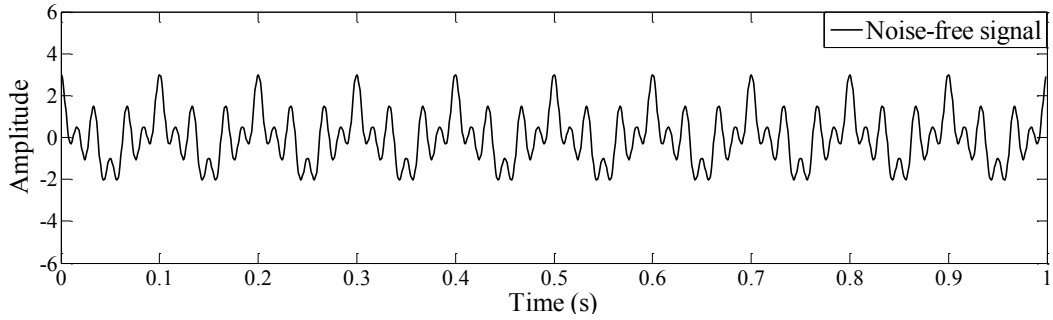


Fig. 1. The workflow of DFA-VMD method.

We first used a synthetic example to show the theoretical fundamental of conventional DFA-VMD denoising. Fig. 2 (a) shows a synthetic signal that comprises three monochromatic components $\cos(20\pi t)$, $\cos(60\pi t)$, and $\cos(120\pi t), t \in [0\ 1.0\,s]$. Their dominant frequencies are 10 Hz, 30 Hz, and 60 Hz, respectively. Fig. 2(b) is the noisy signal that is contaminated by WGN with 1.74 dB SNR. To analyze the influence of *K* on *h*, DFA of each IMF with changing *K* is demonstrated in Fig. 3. Although the scaling exponent *h* of each IMF changes with changing *K*, it decreases with the growing main frequency of each IMF. Generally, the front modes are chosen to reconstruct the filtered signal for VMD-based denoising. Therefore, a short conclusion can be formed-the scaling exponents *h* of the chosen IMFs are obviously greater than the IMFs that are not selected. In other words, the scaling exponent *h* can be used to adjust the modes number and determine the valuable IMFs.

a)



b)

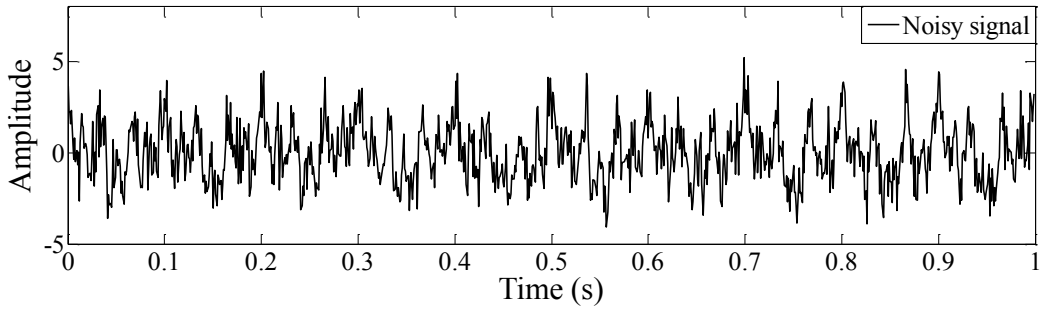


Fig. 2. (a) The synthetic noise-free signal with three frequency components at 10 Hz, 30 Hz, and 60 Hz. (b) Noisy signal with SNR=1.74 dB, $h_0 = 1.02$.

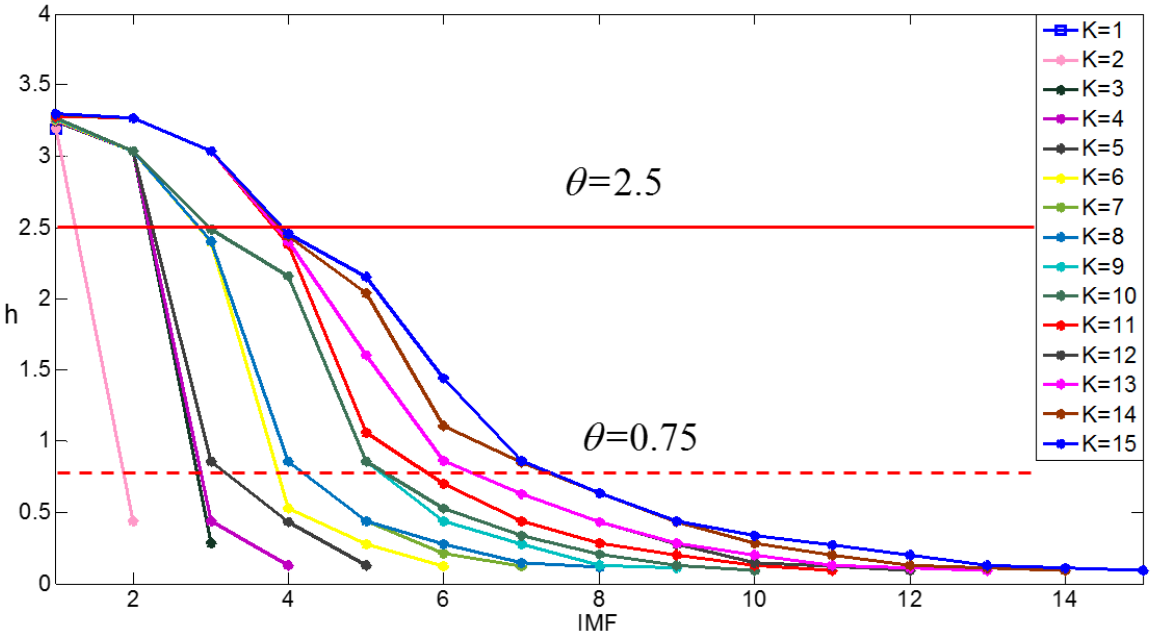
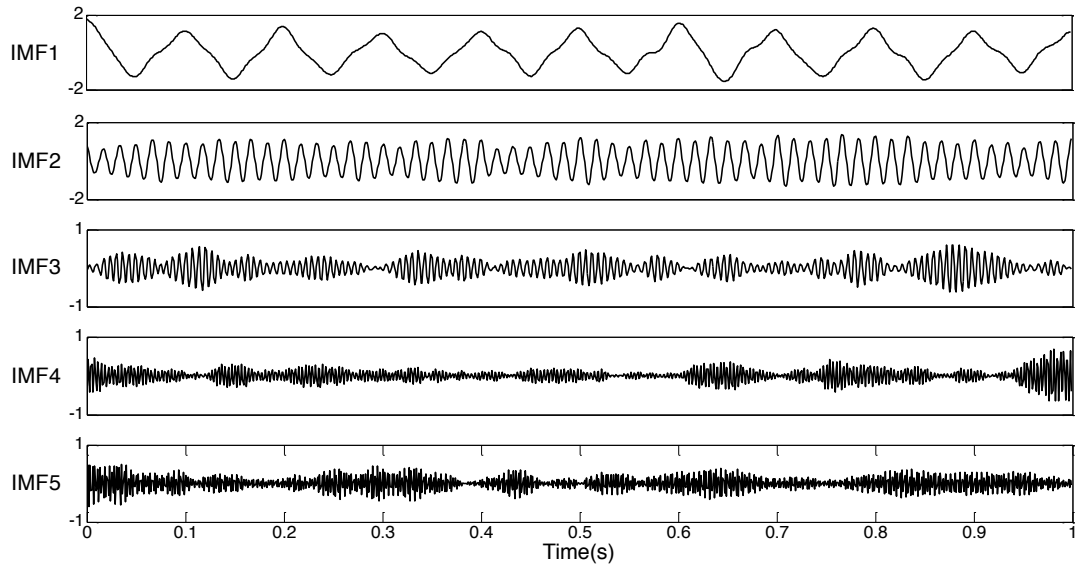


Fig. 3. DFA of each IMFs with different K from K=1 to K=15.

Considering the mode mixing and the calculation of the scaling exponent, the threshold for WGN is generally defined as $\theta = h + 0.25$ ($h = 0.5$ for WGN) for EEG or ECG signal (Mert and Akan, 2014; Liu et al., 2016). The IMFs whose scaling exponents are bigger than θ are chosen to reconstruct the filtered data, and the IMFs with smaller h are considered as noise. However, the predefined noise boundaries (Peng et al., 1994) for other time series cannot perform perfectly when dealing with seismic data. The ECG signal has obvious morphological features in the time domain, and the dominant frequency usually is smaller than 1 Hz. Unlike biomedical signals, the effective bandwidth of seismic data is about 0-120 Hz. The earth's absorption and attenuation effect is one of the most important factors that affect the resolution of seismic survey. Therefore, the useful information in seismic data is band-limited, and the special case that h = 0.5 stands for WGN is not suitable for seismic denoising. In this paper, we defined a robust threshold for random noise in seismic data as $\theta = 2.5$, which is approximately equal to the scaling exponent of 120 Hz components (i.e., the upper effective frequency of seismic data).

We verified the novel threshold using a lot of synthetic-models. For the noisy signal (h_0 = 1.02) in Fig. 2(b), if we set θ=0.75, the number of IMFs will satisfies the expectation when K = 5 according to the workflow of DFA-VMD. The five IMFs and their Fourier spectra are shown in Fig. 4. The mode mixing is apparently shown in Fig. 4(b). Although the IMF1 and IMF2 can be easily interpreted as 10 Hz and 60 Hz, the 30 Hz component is visibly mixed in the first two IMFs. This problem is caused by inappropriate K value and it is consistent with IMFs in the time domain [Fig. 4(a)]. The predefined threshold θ=0.75 for WGN in biomedical data cannot perform robustly when dealing with seismic data. Fig. 5 provides better decomposition using VMD with θ=2.5，K=11. The 10 Hz, 30 Hz and 60 Hz components are captured exactly by IMF1, IMF2, and IMF3 [Fig. 5(b)]. More importantly, the IMFs whose scaling exponents are in 0.75-2.5 still contain considerable noise. Fig. 6 shows the denoised outputs using conventional DFA-VMD and our proposed robust DFA-VMD. After filtered using the robust threshold filtering (SNR = 12.63), the signal-to-noise ratio (SNR) improves remarkably compared with denoised signal (SNR = 6.78) using conventional DFA-VMD. We further use the synthetic data with different h_0 and SNR to test the applicable range of this novel threshold. The results are shown in Table 1 (as Appendix), where the numbers in bold are the relatively better outputs. Table 1 shows that our improved threshold (θ=2.5) obtained superior outputs with higher calculation cost, because it requires decomposing the input data more times to get enough valuable IMFs to reconstruct the filtered signal.

a)



b)

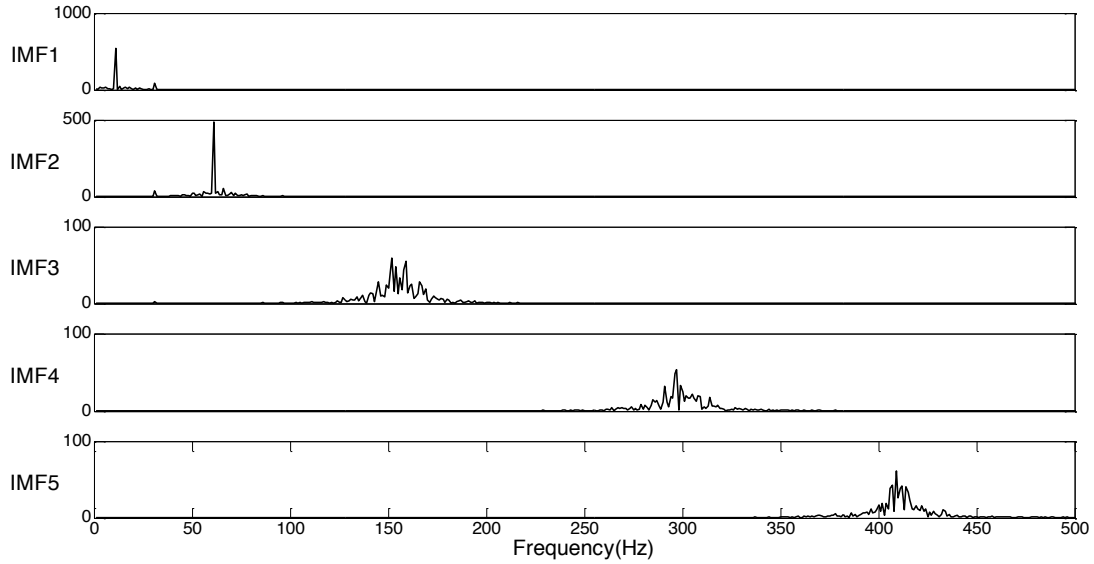
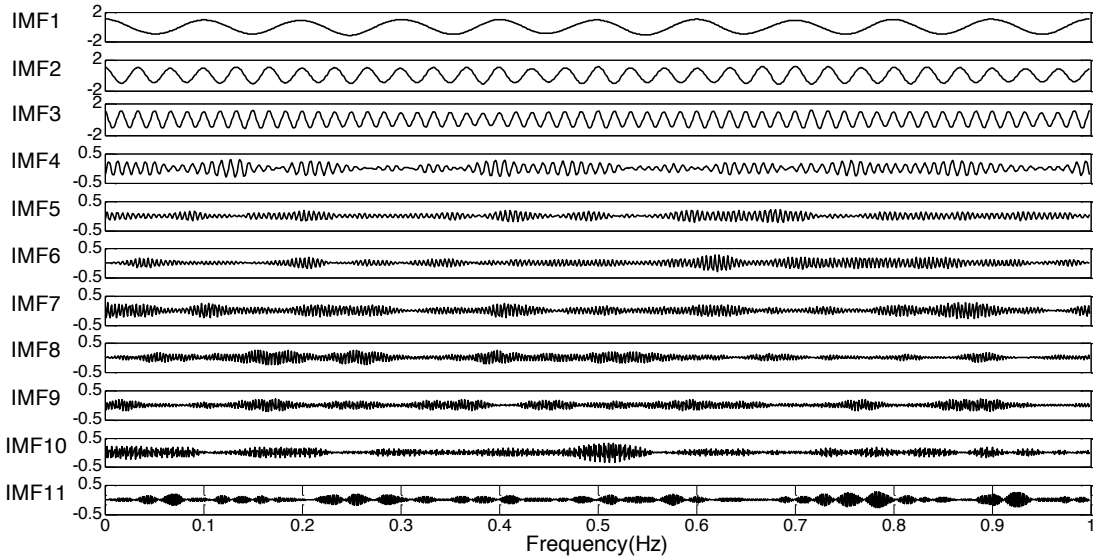


Fig. 4. The IMFs extracted from VMD with $K = 5$ (a) and their Fourier spectra (b). Mode mixing caused by inappropriate K value.

a)



b)

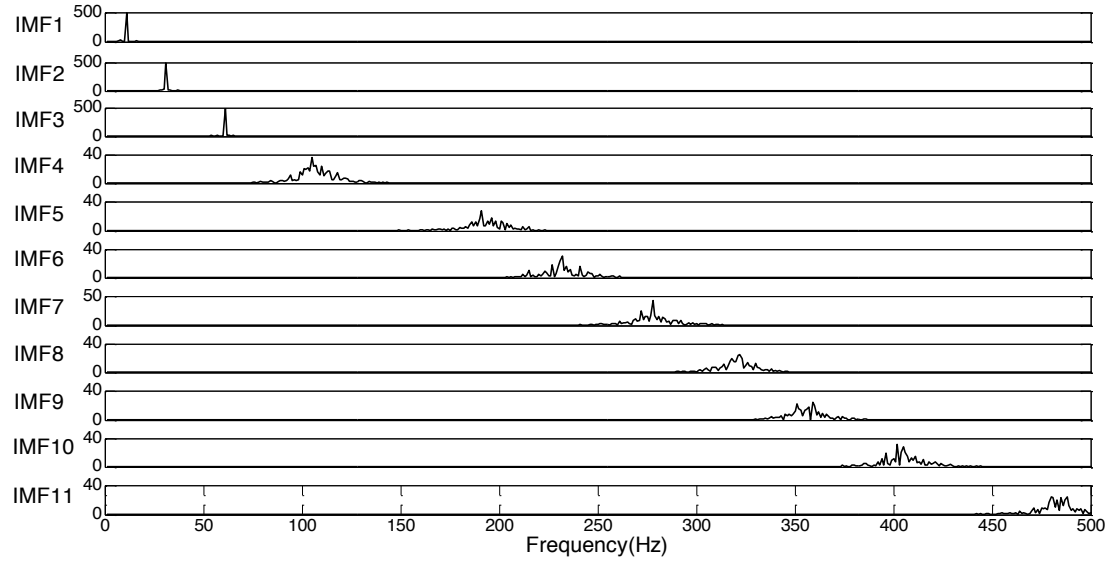
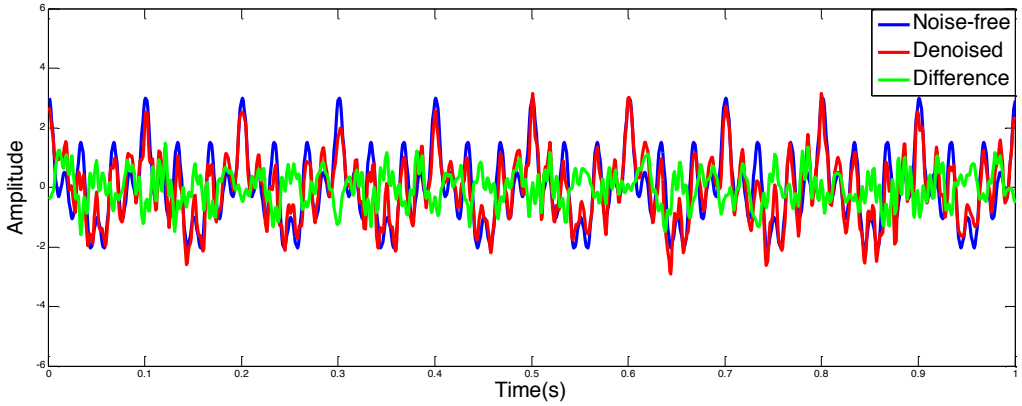


Fig. 5. The IMFs extracted from VMD with $K = 11$ (a) and their Fourier spectra (b). Mode mixing caused by inappropriate K value. The 10 Hz, 30 Hz and 60 Hz components are captured exactly by IMF1, IMF2, and IMF3.

a)



b)

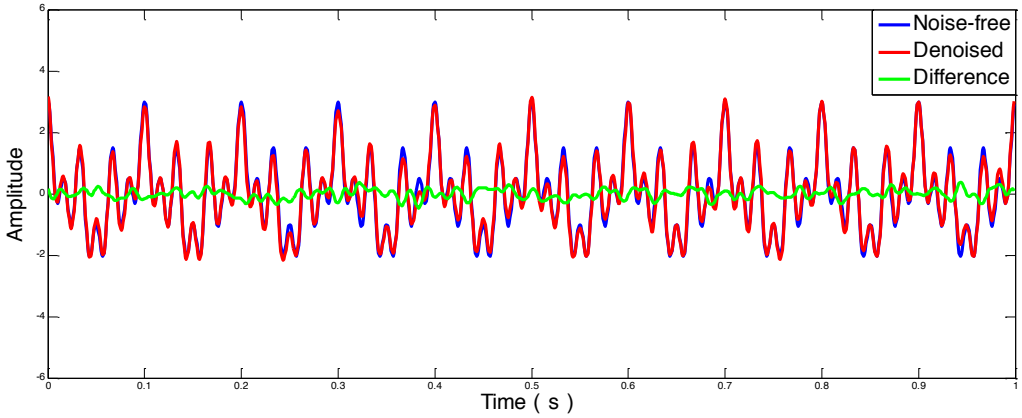


Fig. 6. Comparison of DFA-VMD denoising using different threshold (a) Denoising using conventional DFA-VMD with $\theta = 0.75$, $K = 5$ (SNR=6.79). (b) Denoising using the robust DFA-VMD with $\theta = 2.5$, $K = 11$ (SNR=12.63). The original noise-free data (blue), the denoised data using DFA-VMD (red), and the difference (green) are provided.

The robust DFA-VMD denoising in the (*f-x*) domain

We apply the robust DFA-VMD in the (*f-x*) domain to remove the random and coherent noise in seismic data. The *f-x* predictive filtering was first used to suppress random noise by Canales (Canales, 1984). Linear or quasilinear events in the (*f-x*) domain manifest as a superposition of harmonics in the (*f-x*) domain. These harmonics can be perfectly predicted by an autoregressive (AR) filter. In other words, the signal can be predicted by AR filter and the rest is noise. However, in reality, seismic events do not follow Canales's assumptions exactly and show nonlinear and non-stationary spatial behavior. A hyperbolic or a linear event with amplitude

versus offset does not display as a superposition of simple harmonics. The AR filter is not the optimal choice in this case. We apply the robust DFA-VMD on constant-frequency slices in the (*f-x*) domain to overcome the low performance of (*f-x*) deconvolution when processing highly complex geologic sections. The scaling exponent *h* obtained by DFA can also help describe the "roughness" of time series. A larger value of *h* represents a smoother time series. Conversely, a small value of *h* represents more rapid fluctuations. IMFs with rapid oscillations in the data means large wavenumber components. In that case, *f-x* DFA-VMD approach can suppress noise by subtracting IMFs with lower scaling exponent. The proposed approach is implemented in a similar way to (*f-x*) deconvolution using the following scheme:

1) Select a time window and transform the data into the (*f-x*) domain by Fourier transform. In this paper, the length of the time window is chosen to be 256 ms.
2) For every frequency,
    a) separate real and imaginary parts in the spatial sequence,
    b) compute IMF resulting from VMD,
    c) discriminate noisy IMFs and subtract to obtain the filtered real signal,
    d) repeat for the imaginary part,
    e) combine to create the filtered complex signal.
3) Transform data back to the (*f-x*) domain.
4) Repeat for the next time window.

SYNTHETIC EXAMPLES

We use a synthetic 2D example to test the proposed denoising algorithm. The synthetic data [Fig. 7(a)] contains two horizontal events, two hyperbolic events, and a varying-energy event with polarity reversal at 0.27s. After adding steeply dipping coherent and band-limited random noise (the same band as the noise-free data), we obtained the noisy data with 0.2 dB SNR [Fig. 7(b)]. The *f-x* deconvolution and conventional *f-x* DFA-VMD denoising are shown to make a comparison. All three methods in the *f-x* domain are implemented between 0 Hz and 50% (125 Hz) of the Nyquist frequency. The *f-x* deconvolution uses the length of the autoregressive operator as 10. The parameters for VMD are kept at $\alpha = 5000, \tau = 0, DC = Init = 0$, and $Tol = 1e-07$ in all cases considered. Fig. 8 shows the outputs of different three methods. The denoised section using the *f-x* deconvolution [Fig. 8(a), SNR = 3.5 dB] is not satisfactory and the corresponding difference section [Fig. 8(d)] reveals great damage to useful

events. The conventional thresholded VMD removes more noise, and retains useful energy better [Fig. 8(b), SNR = 4.1 dB] than the *f-x* deconvolution, but the difference section [Fig. 8 (e)] contains a visible leakage of the hyperbolic events. Using the robust *f-x* DFA-VMD, we obtained a much cleaner denoised section [Fig. 8(c), SNR = 9.0 dB] with stronger amplitude events. In the difference section [Fig. 8(f)], the energy of two hyperbolic events are extremely weak and the steeply dipping coherent noise is obvious, suggesting the proposed approach remove more noise, and increase the preservation of useful signal. The variant energy event at 0.27s is amplitude-preserved remarkably.

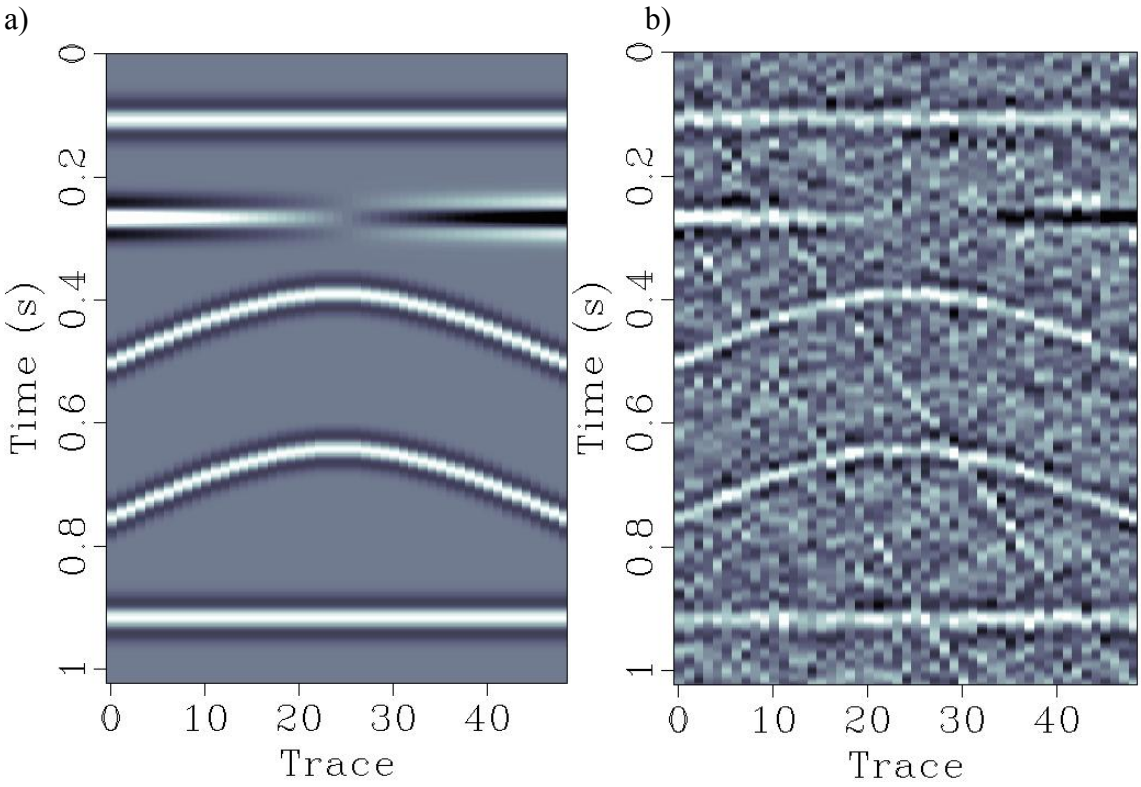


Fig. 7. (a) Synthetic clean signal. (b) Noisy section (SNR = 0.2 dB)

FIELD EXAMPLES

In this section, we compare the performance of *f-x* deconvolution, conventional *f-x* DFA-VMD and the robust *f-x* DFA-VMD using three field-data sets. All denoising methods in the (*f-x*) domain use a short-time Fourier transform with a sliding temporal window of length 256 ms, and Frequencies beyond 80% of the Nyquist frequency are not processed and are damped to zero.

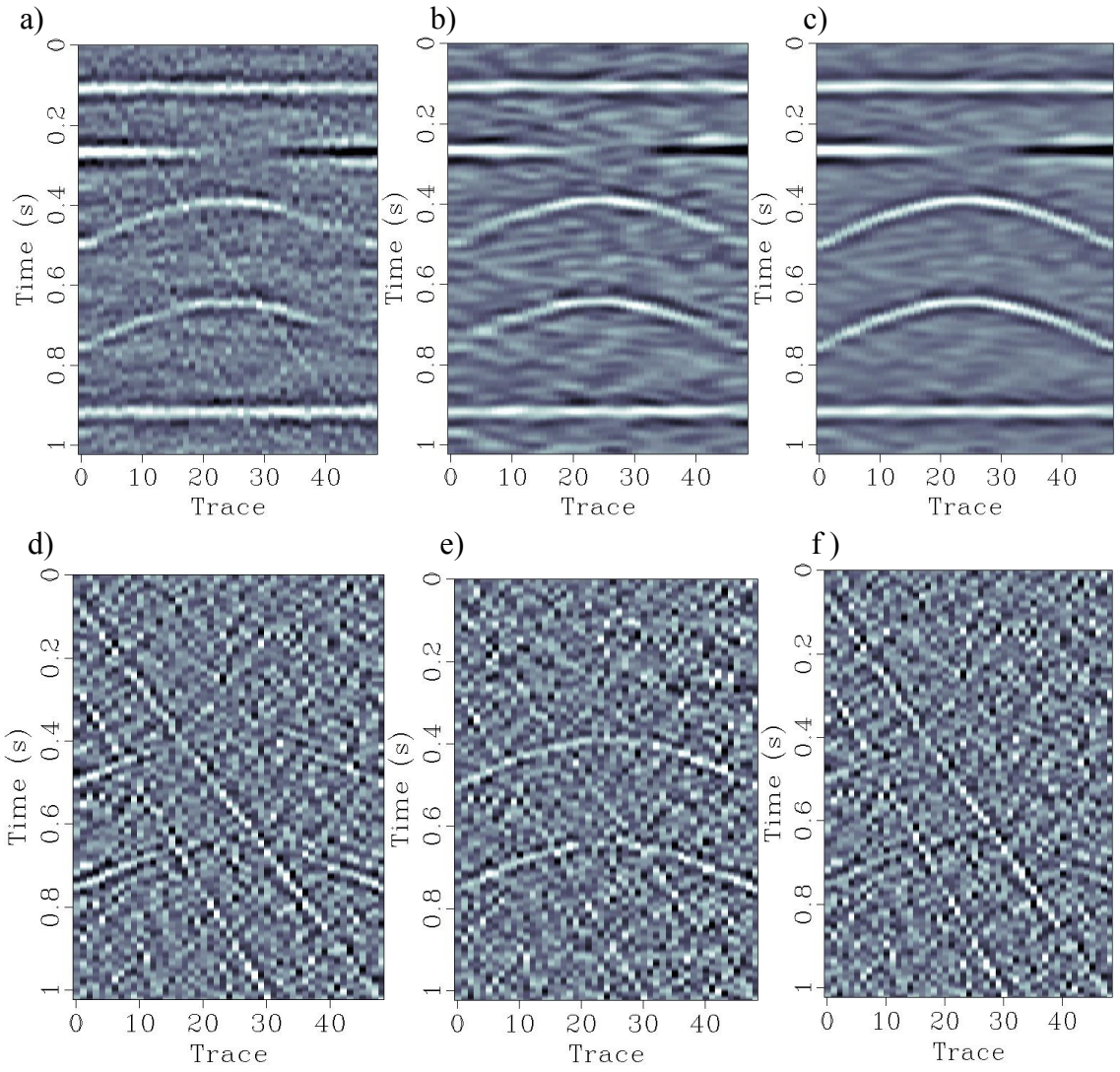


Fig. 8. (a) Denoised section using f-x deconvolution (SNR=3.5 dB), (b) denoised section using f-x DFA-VMD (SNR = 4.1 dB), (c) denoised section using the robust f-x DFA-VMD (SNR = 9.0 dB), (d) difference section using f-x deconvolution, (e) difference section using f-x DFA-VMD, and (f) difference section using the robust f-x DFA-VMD.

Data set 1: Shot gather

A shot gather is displayed in Fig. 9(a). The original data contain ground-roll and random noise. For the (f-x) deconvolution, the length of the autoregressive operator is 10. After using the f-x deconvolution [Fig. 9 (b)], only tiny amounts of noise are suppressed and the difference section [Fig. 9(c)] contains a visible loss of useful reflections at around 0.5 s. Changing its predefined parameters does not lead to significantly better outputs in this case. The conventional DFA-VMD provides a much clearer shot gather

[Fig. 9 (d)], but the useful events in far-offset are blurred and destroyed (marked by the rectangle). The events at around 0.5 s are weak but still present in the difference section [Fig. 9 (e)]. The improved thresholded VMD [Fig. 9 (f)] performs much better than the previous two techniques on this shot gather. There is no noticeable leakage of useful reflectors even in far-offset [Fig. 9(g)].

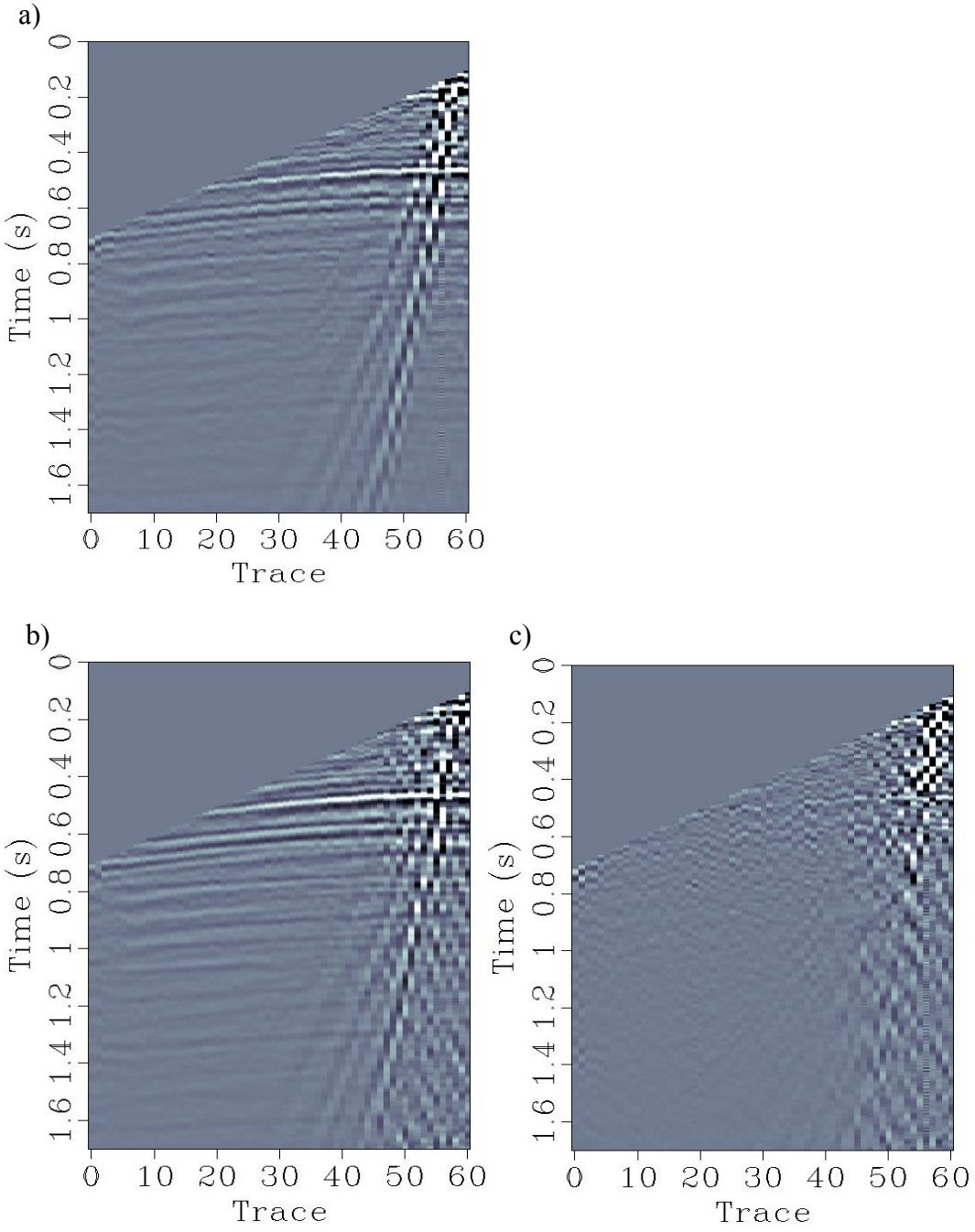


Fig. 9. Data set 1: Shot gather. (a) Original data, (b) denoised section using *f-x* deconvolution, (c) difference section using *f-x* deconvolution.

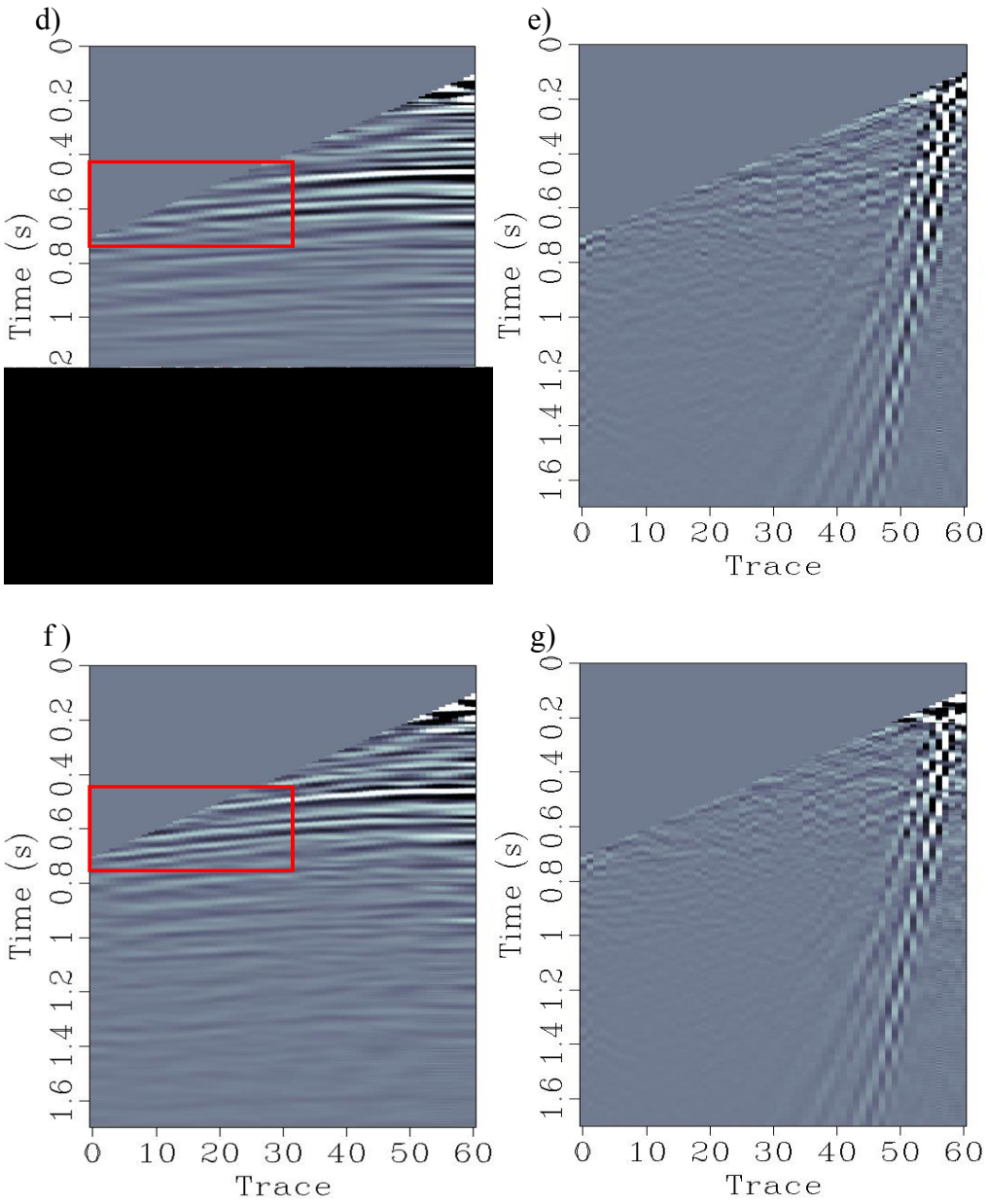


Fig. 9. Data set 1: Shot gather. (d) denoised section using *f-x* DFA-VMD, (e) difference section using *f-x* DFA-VMD, (f) denoised section using the robust *f-x* DFA-VMD, and (g) difference section using the robust *f-x* DFA-VMD.

Data set 2: Common midpoint gathers

Next, we consider a moveout-corrected common midpoint (CMP) gathers [Fig. 10(a)] that contain a mixture of horizontal events (before 5.0 s), hyperbolic events (strongest one at approximately 5.8 s), overwhelmed by steeply dipping linear noise, ground-roll, random noise, and coherent noise. The *f-x* deconvolution uses the autoregressive operator as 20.

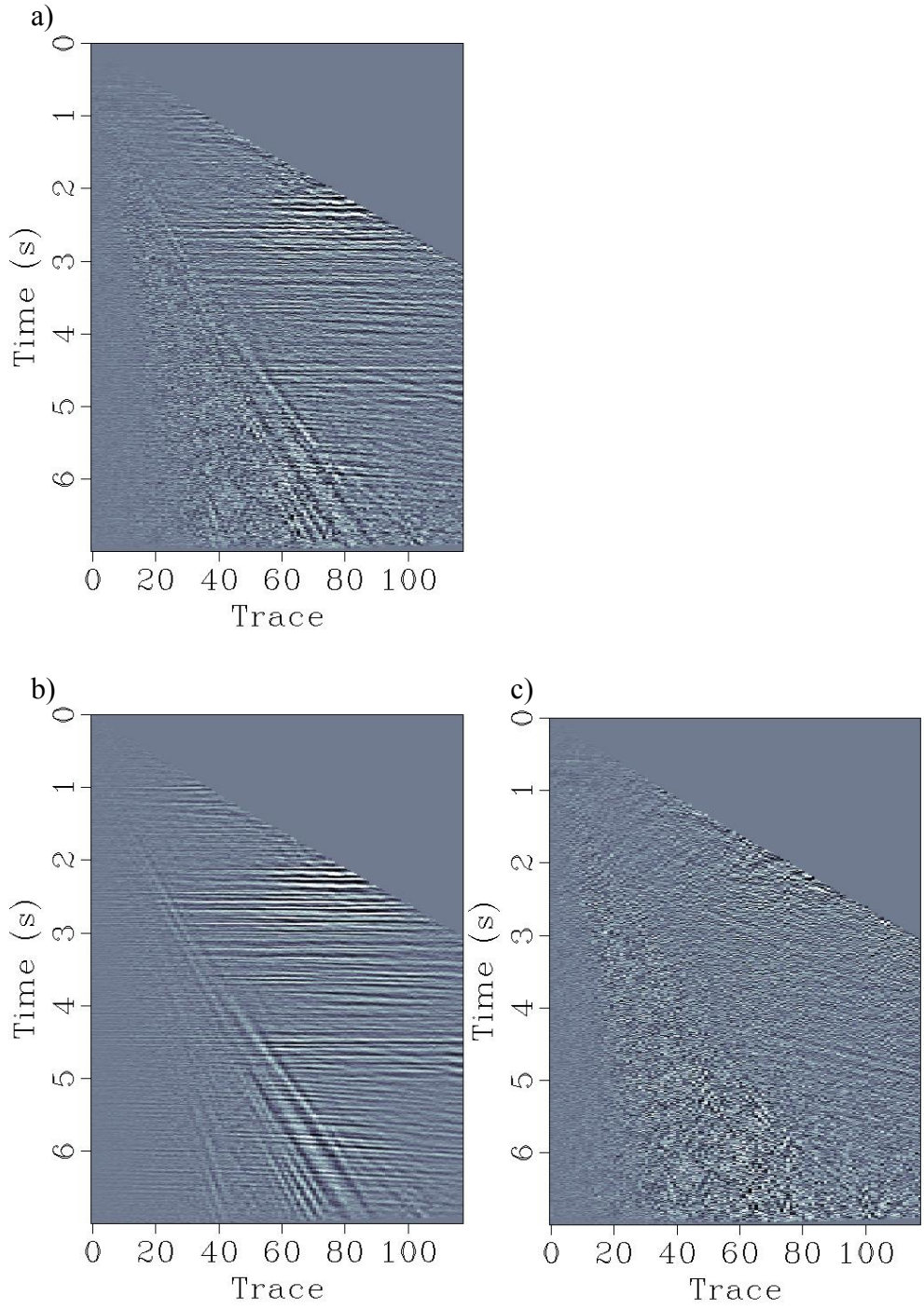


Fig. 10. Data set 2: A moveout-corrected CMP gathers. (a) Original data, (b) denoised section using *f-x* deconvolution, (c) difference section using *f-x* deconvolution.

The *f-x* deconvolution [Fig. 10 (b)] enhances the SNR of all coherent events including the linear noise and useful reflection, yet suppresses some of the random noise as shown in the difference section [Fig. 10(c)]. The *f-x*

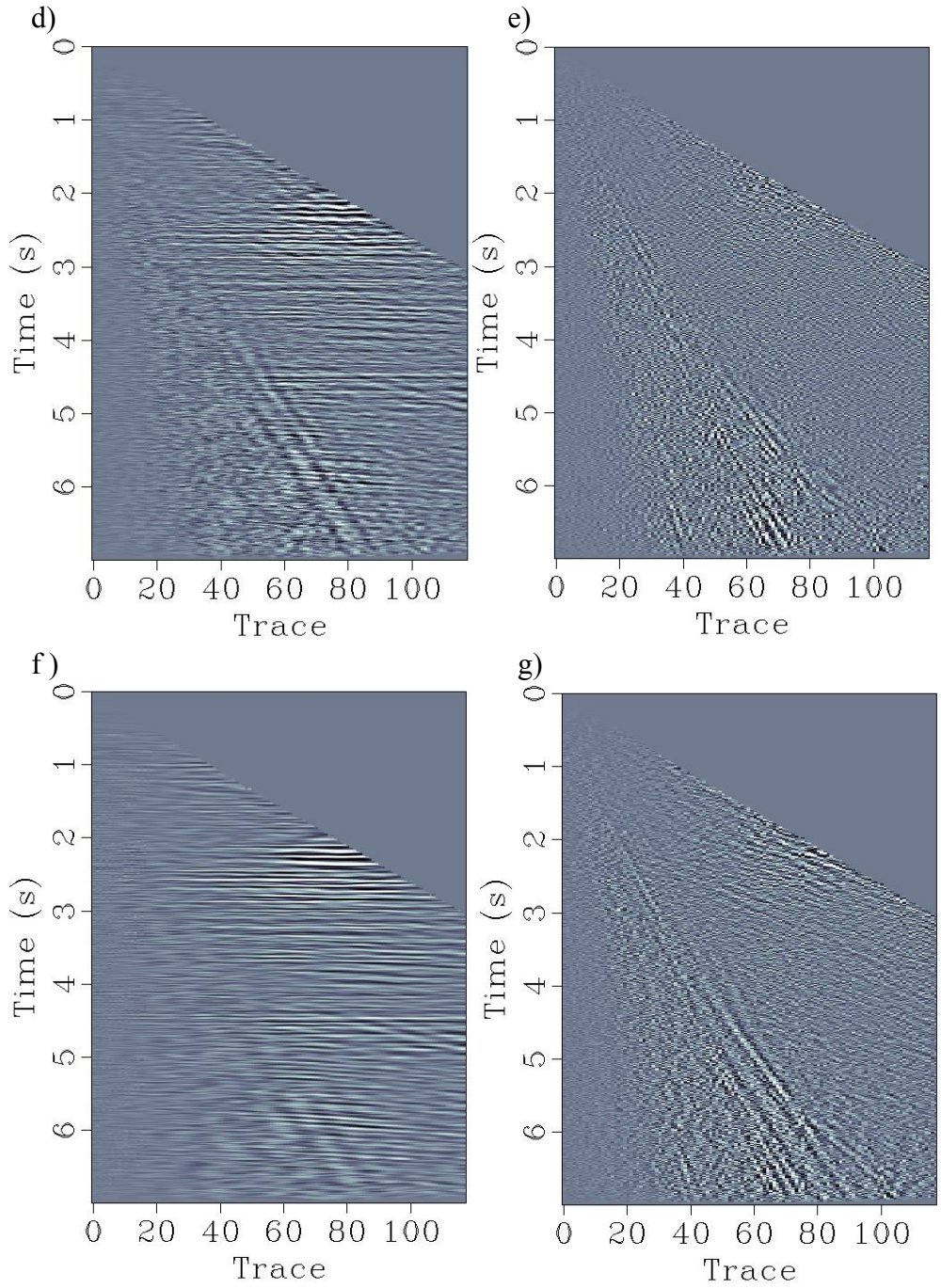


Fig. 10. Data set 2: A moveout-corrected CMP gathers. (d) denoised section using the robust *f-x* DFA-VMD, (e) difference section using *f-x* DFA-VMD, (f) denoised section using the proposed approach, and (g) difference section using the robust *f-x* DFA-VMD.

deconvolution can improve the SNR of all coherent events. This is an advantage if we consider the useful events but a disadvantage if we consider the linear noise and ground-roll. The conventional *f-x* DFA-VMD [Figs. 10(d) and 10(e)] has almost no influence on the near-offset noise. The robust

f-x DFA-VMD [Fig. 10(f)] boosts the target reflection compared to the noise and suppresses the random- noise. It also removes most linear noise and partially ground-roll in this case. No noticeable distortion of useful information in the difference section [Fig. 10 (g)].

Data set 3: Stacked section

The third example uses the stacked field-data from Alaska (Geological Survey, 1981), which was used previously by Han (2015). The post-stack section is shown in Fig. 11(a) and a deeper zoomed section (the time interval is 3.8-5.2 s and trace number is 250-400) is shown in Fig. 11(b). Random and coherent noise still exist after stacking. Some crossing artifacts are shown in the deep section, probably caused by previous data processing. We apply three approaches with the same parameters as for the previous example.

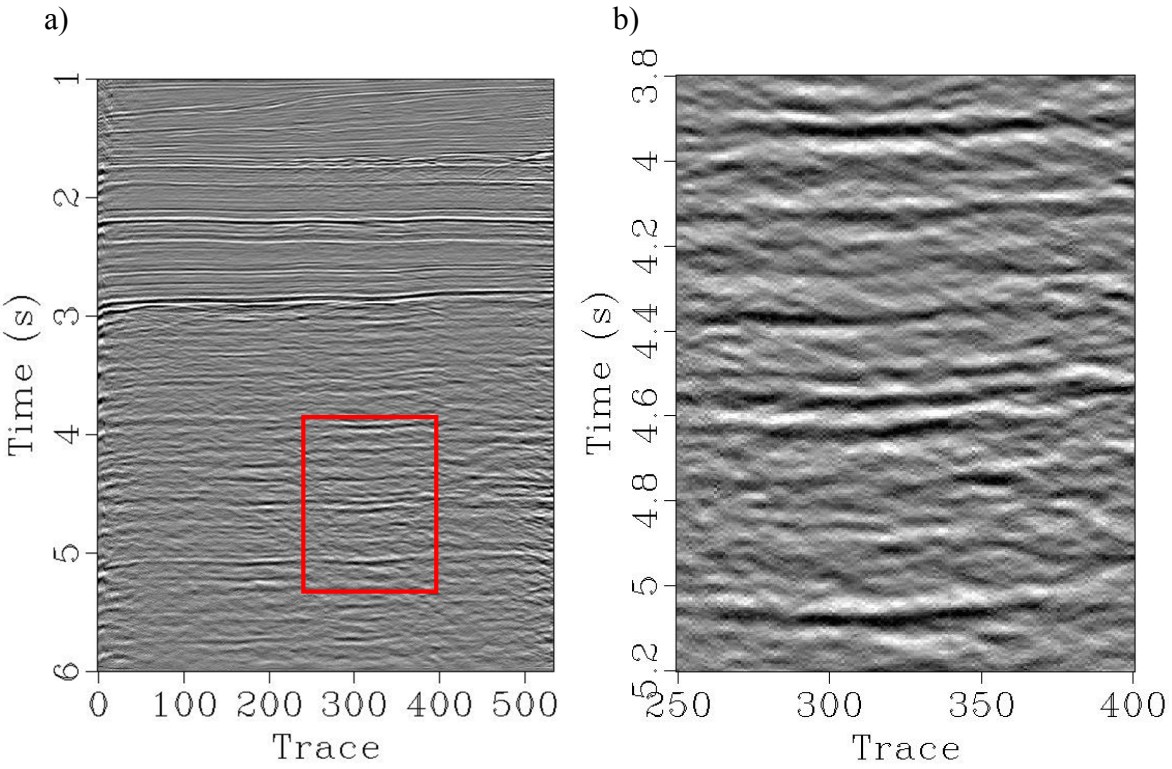


Fig. 11. Data set 3: (a) a stacked section from Alaska, (b) the zoomed section from time 3.8-5.2 s and trace number is 250-400 (marked by the rectangle). Random noise, coherent noise, and some crossing artifacts are still exist after stacking.

All of the three methods make the events cleaner and improve the quality of the input data especially at the deeper part. The *f-x* deconvolution attenuates some of the random-noise but leaves the crossing artifacts unaffected [Fig. 12 (a)]. When checking the noise section in Fig. 12(d), we found a considerable leakage of useful reflection information especially the events before 3.0 s. The denoised section after the use of conventional DFA-VMD is shown in Fig. 12(b). From the noise section in Fig. 12(e), we can see that the denoising method using threshold θ=0.75 only eliminates a small amount of the random noise yet decreases the leakage of useful reflectors. This is because the seismic data cannot be decomposed suitably with inappropriate K value. The best output is given by our method (Fig. 12(c)), which removes more random noise and the crossing artifacts. It is also characterized by significant preservation of the useful signal [Fig. 12(f)].

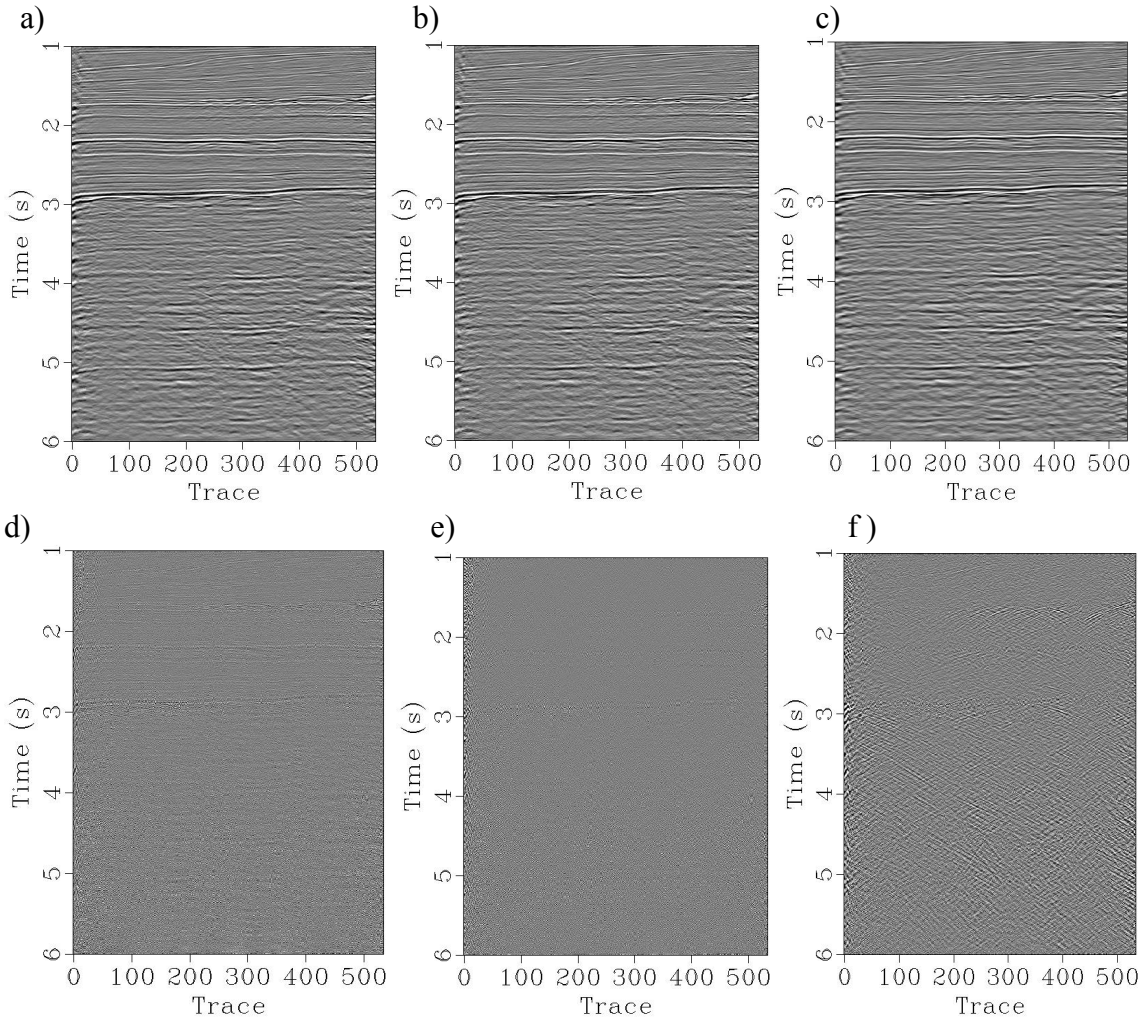


Fig. 12. Data set 3: A stacked section. (a) denoised section using f-x deconvolution, (b) denoised section using f-x DFA-VMD, (c) denoised section using the robust f-x DFA-VMD, (d) difference section using f-x deconvolution, (e) difference section using f-x DFA-VMD, and (f) difference section using the robust f-x DFA-VMD.

Fig. 13 shows the same zoomed parts as Fig. 11(b) of three denoising outputs. The zoomed section clearly shows that a lot of crossing artifacts interfere with the useful signal in this deeper section. The *f-x* deconvolution [Fig. 13(a)] and conventional threshold method [Fig. 13(b)] are only valid for random-noise in the deep part. The filtered sections are nearly no changes compared to the unprocessed data. Our robust thresholding method [Fig. 13(c)] is able to remove the crossing artifacts leading to a superior result over the first two techniques. When checking the difference sections, the *f-x* deconvolution partially removes useful energy [Fig. 13(d)]. There is no apparent useful information in the difference section after using VMD based denoising [Figs. 13(e) and 13(f)].

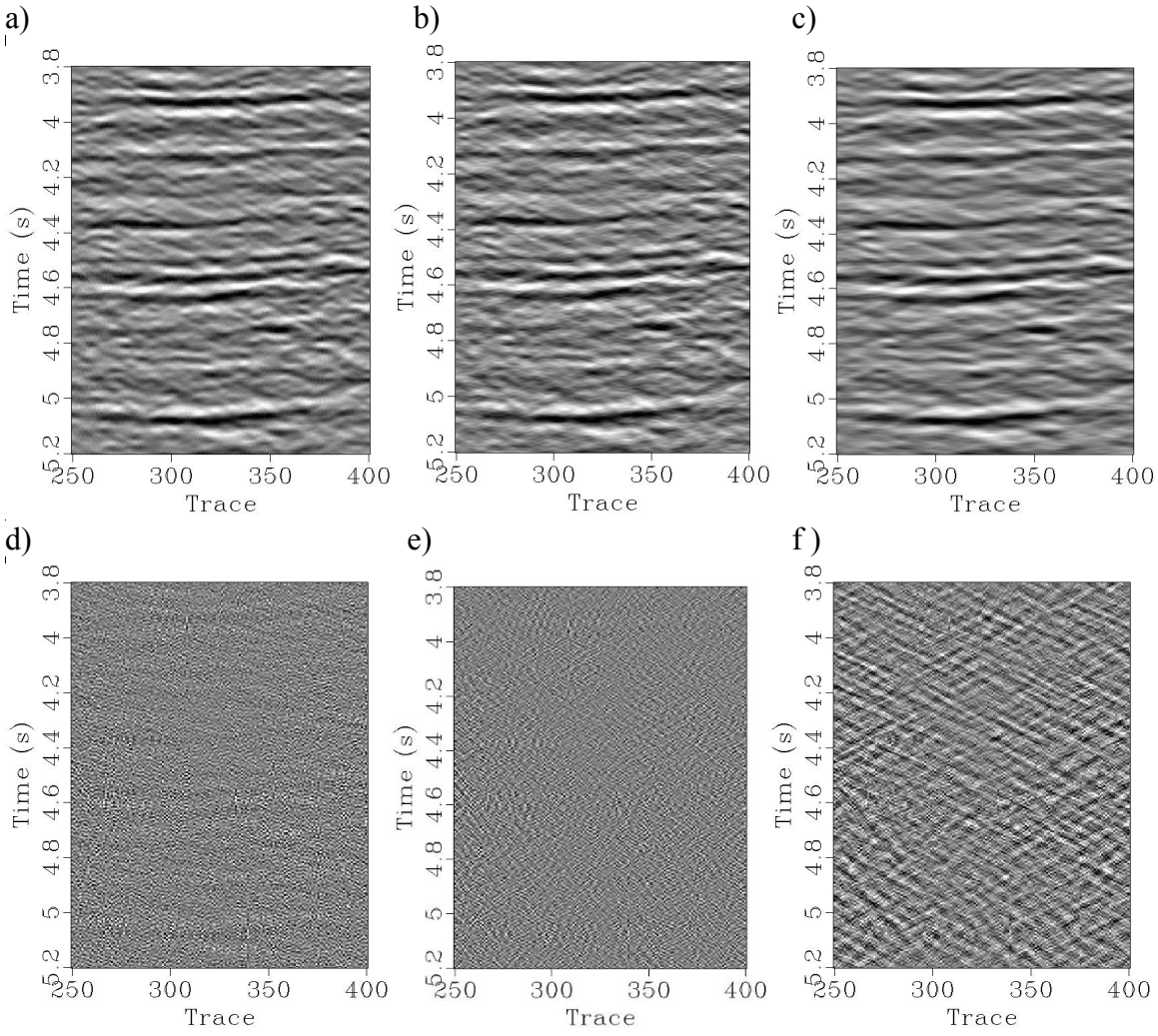


Fig. 13. Comparison of three denoising method in the zoomed section (a) Denoised section using *f-x* deconvolution, (b) denoised section using *f-x* DFA-VMD, (c) denoised section using the robust *f-x* DFA-VMD, (d) noise section using *f-x* deconvolution, (e) noise section using *f-x* DFA-VMD, and (f) noise section using the robust *f-x* DFA-VMD.

DISCUSSION

As is known, VMD is a novel decomposition technique that can non-recursively decompose time series into an ensemble of band-limited IMFs. The modes from VMD are less sensitive to noise than those from EMD, in which the first two modes always contain more residual noise. Furthermore, VMD has lower computational cost than the CEEMD-based methods. These features are helpful to accurately capture components and reconstruct the filtered signal. However, the number of modes resulting from VMD is required to be predefined and its value has expected influence on the efficiency of VMD. Both overbinning and underbinning have impact on the nature of the detected modes. In practice, prior information about the expected number of modes is rarely available. DFA is a powerful tool to solve this problem.

The core of the proposed DFA-VMD approach is that the scaling exponent of the input data can be used to adjust the number for VMD automatically. The relationship between the number of IMFs and the scaling exponent for seismic data is obtained from a lot of synthetic-models and field examples. A completely robust scaling exponent makes VMD an almost parameters-free denoising approach. However, DFA-VMD requires to decompose the signal repeatedly, which increase the computational cost. Many extensions to DFA, such as multifractal DFA (Kantelhardt et al., 2002), multivariate DFA (Xiong and Shang, 2017), are developed to give a deeper insight into time series in higher dimensions. Therefore, a criterion for more accurate computation of scaling exponents or other automated techniques to ease the selection of the IMFs should be also devised. The authors guess that it is possible to research a simpler and clearer relationship between the number K and frequency characteristic of input data. A short conclusion can be drawn from our examples that the useful IMFs with seismic effective frequency bands have much bigger scaling exponents, because signals with low dominant frequency have less fluctuation and more gentle trends. The IMFs whose scaling exponents are between two thresholds ($0.5 \le \theta \le 2.5$) are beyond the effective bands of seismic data and contain a lot of random noise. The new threshold $\theta = 2.5$ is approximately equal to the scaling exponent of the upper effective frequency of seismic data. Although the above conclusion has been well proved by a large number of examples, we are still working on a strict mathematical theory to explain the relationship between the scaling exponents and the number of IMFs.

In field-data example, the number K is different for each trace but the number of valuable modes is almost the same, suggesting similar noise density. It should be mentioned that DFA only improve the parameters

selection, rather than the decomposition algorithm itself. The VMD algorithm also depends on several other parameters, such as the balancing parameter of the data-fidelity constraint α , the step on the recursion that updates the Lagrangian multipliers τ , the tolerance of the convergence criterion Tol , the number of DC components, and the initialization of center frequencies $Init$. These parameters are beyond the scope of this paper but easier to be predefined than the number K .

CONCLUSIONS

We propose a robust DFA-VMD denoising method and apply it in the (*f-x*) domain for seismic denoising. The VMD depends on several predefined parameters. The number of IMFs is extremely crucial and its value has serious influence on the efficiency of VMD. DFA is a powerful tool to solve this problem. The scaling exponent obtained by DFA can also help to identify and remove the noisy IMFs. To ease the application of VMD based denoising in practice, we proposed a more robust threshold according to the effective frequency of seismic data, because the threshold for other time series denoising cannot perform perfectly when dealing with seismic data. The proposed approach is an almost parameters-free denoising approach, which decreases the difficult of preferences and provides a theory fundamental for VMD-based denoising. We apply it in the (*f-x*) domain to remove the random and coherent noise in seismic data. Using two synthetic examples and three field examples, we illustrated the superior performance of the method we proposed over (*f-x*) deconvolution and conventional DFA-VMD.

ACKNOWLEDGMENTS

The authors would like to thank the United States Geological Survey for sharing the Alaska data set. The authors also thank the editor and an anonymous reviewer for their helpful suggests that greatly improved my manuscript. This research is partially supported by the National Natural Science Foundation of China (No. 41674128), the National Science and Technology Major Project of China (No. 2016A-33), and the Technology Innovation Project of CNPC (No. 2017D-5007-0302).

REFERENCES

Allen, J.B., 1977. Short term spectral analysis, synthetic and modification by discrete Fourier transform. IEEE Transact. Acoust., Speech Sign. Proces., 25: 235-238.

Bryce, R.M. and Sprague, K.B, 2012. Revisiting detrended fluctuation analysis. Scientific Reports, 2(3): 315-320.

Canales, L., 1984. Random noise reduction. Expanded Abstr., 54th Ann. Internat. SEG Mtg., Atlanta: 525-527.

Chen, Y. and Ma, J., 2014. Random noise attenuation by *f-x* empirical-mode decomposition predictive filtering. Geophysics, 79(3): V81-V91.

Dragomiretskiy, K. and Zosso, D., 2014, Variational mode decomposition. IEEE Transact. Sign. Process., 62: 531-544.

Geological Survey U.S., 1981. http://wiki.seg.org/wiki/ALASKA_2D_LAND_LINE_31-81, accessed 2 September 2017.

Grossmann, A. and Morlet, J., 1984. Decomposition of Hardy functions into square integrable wavelets of constant shape. SIAM J. Mathemat. Analys., 15: 723-736.

Han, J. and van der Baan, M., 2015. Microseismic and seismic denoising via ensemble empirical mode decomposition and adaptive thresholding. Geophysics, 80(6): KS69-KS80.

Hestenes, M.R., 1969. Multiplier and gradient methods. J. Optimizat. Theory Applicat., 4: 303-320.

Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.-C., Tung, C.C. and Liu, H.H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. Mathemat. Physic. Engineer. Sci., 454: 903-995.

Huang, N. and Wu, Z., 2008. A review on Hilbert-Huang transform: Method and its applications to geophysical studies. Rev. Geophys., 46: RG2006.

Hurst, H.E., 1951. Long-term storage capacity of reservoirs. Transact. Am. Soc. Civil Engin., 116: 770-779.

Ivanova, K. and Ausloos, M, 1999. Application of the detrended fluctuation analysis (DFA) method for describing cloud breaking. Physica A - Statist. Mechan. Applicat., 274: 349-354.

Jospin, M., Caminal, P., Jensen, E.W., Litvan, H., Vallverdú, M., Struys, M.M., Vereecke, H.E. and Kaplan, D.T., 2007. Detrended fluctuation analysis of EEG as a measure of depth of anesthesia:. IEEE Transact. Biomed. Engineer., 54: 840-846.

Kantelhardt, J.W., Zschiegner, S.A., Koscielny-Bunde, E., Havlin, S., Bunde, A. and Stanley, H.E, 2002, Multifractal detrended fluctuation analysis of nonstationary time series. Physica A – Statist. Mechan. Applicat., 316: 87-114.

Leistedt, S., Dumont, M., Lanquart, J.P., Jurysta, F. and Linkowski, P., 2007. Characterization of the sleep EEG in acutely depressed men using detrended fluctuation analysis. Clinic. Neurophysiol., 118: 940-950.

Leite, F.S., Rocha, A.F. and Carvalho, J.L., 2010. Matlab software for detrended fluctuation analysis of heart rate variability. Biosignals 2010 – Internat. Conf. Bio-Insp. Syst. Sign. Process.: 225-229.

Li, F., Zhang, B., Verma, S. and Marfurt, K.J., 2017. Seismic signal denoising using threshold variational mode decomposition. Explor. Geophys. doi.org/10.1071/EG17004

Li, F., Zhao, T., Qi, X., Marfurt, K.J. and Zhang, B, 2016. Lateral consistency preserved Variational Mode Decomposition (VMD). Expanded Abstr., 86th Ann. Internat. SEG Mtg., Dallas: 1717-1721.

Liu, W., Cao, S. and Chen, Y., 2016. Applications of variational mode decomposition in seismic time-frequency analysis. Geophysics, 81(5): V365-V378.

Liu, W., Cao, S. and He, Y., 2015. Ground roll attenuation using variational mode decomposition. Extended Abstr., 77th EAGE Conf., Madrid: Th P6 06.

Liu, W., Cao, S. and Wang, Z., 2017. Application of variational mode decomposition to seismic random noise reduction. J. Geophys. Engineer., 14: 888-899.

Liu, Y., Yang, G., Li, M. and Yin, H,, 2016. Variational mode decomposition denoising combined the detrended fluctuation analysis. Sign. Process., 125: 349-364.

Magrin-Chagnolleau, I. and Baraniuk, R., 1999. Empirical mode decomposition based time-frequency attributes. Expanded Abstr., 69th Ann. Internat. SEG Mtg., Houston: 1949-1952.

Mandic, D., Wu, Z. and Huang, N.E., 2013. Empirical mode decomposition based time-frequency analysis of multivariate signals: The power of adaptive data analysis. IEEE Sign. Process., 30(6): 74-86.

Mert, A. and Akan, A, 2014. Detrended fluctuation thresholding for empirical mode decomposition based denoising. Digital Signal Processing, 32, no.2, 48-56.

Peng, C.K., Buldyrev, S.V., Havlin, S., Simons, M., Stanley, H.E. and Goldberger, A.L, 1994. Mosaic organization of DNA nucleotides. Phys. Rev. E Statist. Phys. Plasm. Fluids Relat. Interdiscipl. Topics, 49: 1685-1689.

Peng, C.K., Havlin, S., Stanley, H.E. and Goldberger, A.L, 1995. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. Chaos, 5: 82-87.

Stockwell, R.G., Mansinha, L. and Lowe, R.P., 1996. Localization of the complex spectrum. The S-transform. IEEE Transact. Sign. Process., 44: 998-1001.

Telesca, L., Lapenna, V., Macchiato, M. and Hattori, K, 2007. Nonuniform scaling behavior in ultralow-frequency geomagnetic data in relationship with seismicity. Expanded Abstr., 77th Ann. Internat. SEG Mtg., San Antonio: 1167-1171.

Torres, M., Colominas, M., Schlotthauer, G. and Flandrin, P., 2011. A complete ensemble empirical mode decomposition with adaptive noise. IEEE Internat. Conf. Acoust., Speech Sign. Process.: 4144-4147.

Upadhyay, A. and Pachori, R.B., 2015. Instantaneous voiced/non-voiced detection in speech signals based on variational mode decomposition. J. Franklin Inst., 352: 2679-2707.

Wu, Z. and Huang, N.E., 2004. A study of the characteristics of white noise using the empirical mode decomposition method. Proc. Mathemat., Phys. Engineer. Sci, 460: 1597-1611.

Wu, Z. and Huang, N.E., 2009. Ensemble empirical mode decomposition: A noise-assisted data analysis method. Adv. Adapt. Data Analys., 1: 1-41.

Xiong, H. and Shang, P., 2017. Detrended fluctuation analysis of multivariate time series. Communicat. Nonlin. Sci. Numer. Simulat., 42: 12-21.

Yu, S. and Ma, J., 2018. Complex variational mode decomposition for slop-preserving denoising. IEEE Transact. Geosci. Remote Sens., 56: 586-897.

APPENDIX

Table 1. Comparison of the DFA-VMD denoising with different threshold (The rest parameters for VMD are kept at $\alpha = 5000, \tau = 0, DC = Init = 0$, and $Tol = 1e-07$ in all cases considered).

SNR(dB)	-4.67		-1.53		0.48		1.76		6.55	
h_0	0.76		0.83		0.91		1.09		1.37	
θ	2.5	0.75	2.5	0.75	2.5	0.75	2.5	0.75	2.5	0.75
$K = 1$	2.81	**2.81**	2.42	2.42	3.31	3.31	4.08	4.08	7.91	7.91
$K = 2$	2.81	**2.81**	2.41	2.41	3.39	3.39	4.08	4.08	7.86	7.86
$K = 3$	2.80	**2.80**	2.40	2.40	3.22	3.22	4.06	4.06	12.18	9.53
$K = 4$	**4.24**	2.76	4.55	**4.37**	6.49	3.24	6.80	**8.34**	12.16	9.51
$K = 5$	**4.23**	2.77	4.55	**4.37**	6.52	**7.80**	6.86	**8.32**	13.35	9.48
$K = 6$	N/A	N/A	4.56	**4.37**	6.52	**7.81**	6.89	**8.51**	13.35	9.47
$K = 7$	N/A	N/A	**7.32**	**4.22**	6.55	**7.88**	6.89	**8.48**	13.35	**12.25**
$K = 8$	N/A	N/A	**7.32**	**4.26**	**10.47**	**7.88**	**12.16**	**8.43**	13.34	**12.26**
$K = 9$	N/A	N/A	**7.33**	**4.25**	**10.49**	7.79	**12.16**	6.06	13.36	**12.25**
$K = 10$	N/A	N/A	N/A	N/A	N/A	N/A	**12.17**	6.32	15.30	**12.25**
$K = 11$	N/A	N/A	N/A	N/A	N/A	N/A	**12.18**	6.22	15.31	9.24
$K = 12$	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	**17.25**	9.34
$K = 13$	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	**17.26**	9.32