A simple method inspired by empirical mode decomposition for denoising seismic data

Julián L. Gómez[1] and Danilo R. Velis[1]

ABSTRACT

We developed a new and simple method for denoising seismic data, which was inspired by data-driven empirical mode decomposition (EMD) algorithms. The method, which can be applied either as a trace-by-trace process or in the f-x domain, replaces the use of the cubic interpolation scheme, which is required to calculate the mean envelopes of the signal and the residues, by window averaging. The resulting strategy is not viewed as an EMD per se, but a user-friendly version of EMD-based algorithms that permits us to attain, in a fraction of the time, the same level of noise cancellation as standard EMD implementations. Furthermore, the proposed method requires less user intervention and easily processes millions of traces in minutes rather than in hours as required by conventional EMD-based techniques on a standard PC. We compared the performance of the new method against standard EMD methods in terms of computational cost and signal preservation and applied them to denoise synthetic and field (microseismic and poststack) data containing random, erratic, and coherent noise. The corresponding f-x EMDs implementations for lateral continuity enhancement were analyzed and compared against the classical f-x deconvolution to test the method.

INTRODUCTION

The empirical mode decomposition (EMD) algorithms were first introduced to the signal processing community by Huang et al. (1998). These methods decompose a digital signal s_n of length N, where the subscript n denotes the sample number, into a collection of modes, or intrinsic mode functions (IMFs), without the need for any predefined basis functions, thus allowing for nonlinear

and nonstationary adaptivity to the intrinsic signal characteristics (Huang and Wu, 2008; Mandic et al., 2013).

The modes IMF_n^k, where the superscript $k = 1, \ldots, K$ denotes the order, have a frequency content that decreases with k and, like amplitude and frequency modulated sinusoids, have zero mean and several extrema equal (or different at most by one) to the number of zero crossings. This means that all their maxima and minima are positive and negative, respectively (Kopsinis and McLaughlin, 2009; Mandic et al., 2013). It is these extrema and zero mean conditions that allow EMD to provide a finite set of functions that capture the nonstationary characteristics of the signal (Han and van der Baan, 2015).

The IMFs obtained by the EMD algorithm are prone to suffer from the so-called mode mixing/splitting problem, which means that a single IMF can display oscillatory signals of different scales (mixing) or that a single oscillatory signal may leak into several modes (splitting). Throughout the years, many EMD improvements and strategies have been proposed to alleviate this problem (see Han and van der Baan, 2015), such as the intermittency check (Huang et al., 1999) and the noise-assisted EMD extensions, which decompose the original signal with added zero-mean Gaussian white noise. The noise-assisted repertoire includes the ensemble EMD (Wu and Huang, 2009; Han and van der Baan, 2015), the modified ensemble EMD (Zhang et al., 2010), the complementary ensemble EMD (Yeh et al., 2010), the complete ensemble EMD (CEEMD) with adaptive noise (Torres et al., 2011) and recently, the improved CEEMD (ICEEMD) (Colominas et al., 2014), which enhances the noise cancellation property and reduces the number of spurious modes often observed in CEEMD. The noise-assisted algorithms are based on the hypothesis that the contribution of the added noise will cancel out after stacking a large number of realizations, for each realization contains a uniform background in the time-frequency space where different scales of the original signal can be projected (Huang and Wu, 2008).

Apart from empirically decomposing an input signal into a series of IMFs, one of the main applications of the various EMD algo-

rithms is to separate signal from noise. Over the past few decades, several authors have proposed different strategies for noise attenuation of seismic data using various EMD-based approaches. For example, Battista et al. (2007) present a method to significantly attenuate low-frequency cable strum in reflection data. Bekara and van der Baan (2009) propose a frequency-offset (f-x) domain EMD filtering and successfully apply it to shot- and common-midpoint gathers as well as to stacked and migrated sections. Hooshmand et al. (2012) apply a CEEMD-based filter in the time-offset (t-x) and f-x domains to remove random and coherent noise from shot gathers. In addition, thresholding EMD offers a very attractive variant especially suited for improved denoising. The interval thresholded EMD, for example, was first introduced by Kopsinis and McLaughlin (2009), and recently has been applied to denoise seismic data (Han and van der Baan, 2015). Han and van der Baan develop an adaptive-thresholding ensemble EMD scheme to enhance the signal-to-noise ratio (S/N) of microseismic data and further extend the method into the f-x domain to denoise stacked seismic reflection data.

The results shown in these works are very encouraging because they demonstrate the usefulness of EMD to suppress random and/or coherent noise not only in the t-x domain but also in the f-x domain. Consequently, EMD-based denoising represents an interesting alternative to standard denoising schemes based on band-pass filtering, wavelet analysis, or f-x deconvolution (Bekara and van der Baan, 2009; Han and van der Baan, 2015). However, there is a main drawback shared by all these empirical methods: the computational cost that arises when the volume of the seismic data acquires real-life proportions. For example, processing a relatively small seismic volume of one million traces may take up to several hours on a standard workstation, as we will show later.

This computational disadvantage is mainly due to the interpolation step that is used in the EMD algorithm to obtain the mean envelope of the data, a process that is repeated several times for each IMF that is calculated. This drawback is even more noticeable when using the noise-assisted EMD methods because of the usually large number of signal realizations that are needed to reduce the inherent mode mixing/splitting problem. The envelopes are usually calculated by cubic spline interpolation, although a simpler and faster linear interpolation could be used, but at the expense of serious mode mixing and splitting (Rilling et al., 2003). The best results are obtained when the envelopes are accurately computed by Hermite polynomials (Meignen and Perrier, 2007) or through optimization (Fomel, 2008, 2013; Yang et al., 2014). However, none of these alternatives is faster than the spline interpolation because they add more computational complexity to the EMD algorithm.

A fast scheme to provide appropriate envelopes is then desirable, especially when dealing with large volumes of data. The strategy that we propose as an alternative to the interpolation step and to significantly reduce the computational cost of the EMD approaches is based on a window averaging (WA) similar to the one used by Li et al. (2011). Our strategy also retains the simplicity of the original EMD algorithm, a feature that is very valuable because it minimizes the user intervention that is often required to set the various parameters in the other EMD methods.

It is important to stress that the focus of this work is not on displaying the benefits of denoising seismic data using the various EMD flavors, but on showing the advantages of our scheme once the user is considering using EMD for noise attenuation. In a way,

our method is not an EMD decomposition per se because it only adopts from EMD the sifting idea, but it is a simple user-friendly algorithm that can be used to denoise large volumes of seismic data in an efficient and effective way. In fact, this new strategy, which will be termed window averaged sifting method (WASM), decomposes the data empirically into two modes or components: one containing the noise and the second, which is equal to the residue, containing the signal. As a result, the processing of moderate volumes of data may be reduced from hours (using the standard techniques) to minutes (using the proposed method) on a standard PC. In addition, results obtained using synthetic and field data show that the new procedure attains, in a fraction of the time, the same noise attenuation of standard EMD algorithms.

We analyze the performance of the proposed method in terms of efficiency and signal preservation, and contrast the results against the standard EMD and ICEEMD algorithms. The field data examples include a microseismic data set that contains random and erratic noise, and a seismic poststack line containing random and coherent noise. In the case of seismic 2D data, and to preserve and enhance the lateral continuity of the events, we implement the f-x version of the algorithm and compare the results with the classical f-x deconvolution (Canales, 1984; Gulunay, 1986) and the "standard" f-x EMD-based approaches.

BACKGROUND

Given an input signal s_n, the first IMF is obtained in the original EMD algorithm by iterating over the following steps:

1) Obtain the *mean envelope* $m_n = (\text{env}_n^+ + \text{env}_n^-)/2$, where env_n^+ and env_n^- are the upper and lower envelopes that pass through the maxima and minima of s_n, respectively.
2) Set the residue $r_n = s_n - m_n$ as the new input and repeat step 1.

This two-step iterative process, which is known as "sifting," is repeated until a stoppage criterion is satisfied. The resulting residue of this iteration can be regarded as the first IMF, that is IMF_n^1. The next IMFs are derived by subtracting this IMF from the original signal s_n and repeating the sifting iteration. As mentioned before, the upper and lower envelopes are calculated by means of cubic spline interpolation.

This might be considered as a deflationary process that continues until the residue is either another IMF or becomes a monotonic series, having less than three extrema. The input signal can then be expressed by the following sum:

$$s_n = \sum_{k=1}^{K} \text{IMF}_n^k + R_n, \qquad (1)$$

where R_n is the final residual. Common stoppage criteria for the sifting process, ordered from faster to slower ones, include: (1) to sift a low but fixed number of times (fixed criterion; Wu and Huang, 2009), (2) to sift a low and fixed number of times after the IMF conditions are satisfied for the residue (Huang and Wu, 2008), and (3) the two-threshold criterion (Rilling et al., 2003), which works in most practical cases, but requires more sifting iterations to reach mean amplitudes that are globally small. One key issue of the EMD algorithms is that if the sifting iterations are excessive, the resulting IMFs may fail to contain useful information, and if

they are too few, the resulting modes may display partial or incomplete characteristics of the original signal.

As an example, Figure 1a shows a signal that is composed of three sinusoids of different frequencies, and the corresponding EMD. In this case, because the signal contains no noise and the scales of the sinusoids are markedly different, the EMD is able to capture the tones in the first three IMFs. When the signal is contaminated with uniform random noise of amplitude 0.3 of the maximum signal value, the decomposition displays mode mixing and splitting, as shown in Figure 2a.

The corresponding decompositions using CEEMD and ICEEMD are shown in Figures 1b and 1c and 2b and 2c, respectively. The improvement is noticeable in the noisy case. These results will be analyzed in the "Examples" section.

So far, ICEEMD is the most effective EMD approach when applied for mode decomposition. The noise cancellation is improved by calculating the residues as averages of the mean envelopes resulting from the EMD sifting. The number of spurious modes is reduced by using only EMD modes of the added white noise in the realizations (Colominas et al., 2014). Unlike EMD, the ICEEMD algorithm requires the user to set not only the proper sifting stoppage criterion but also the number of noise realizations N_n and the amplitude of the white noise ε_0 to initiate the decomposition of a given signal. For more details of the algorithm, the reader is referred to the work by Colominas et al. (2014).

EMD DENOISING

Denoising a signal using any of the standard EMD algorithms can be carried out in two steps: (1) decomposing the input data into a complete set of IMFs and (2) reconstructing the signal using either an appropriate subset of IMFs that contain signal and no noise or by band modulating the EMD via mode weighting without exclusion (Macelloni et al., 2011). For example, a typical EMD-based mode exclusion denoising algorithm may consist of subtracting low-order modes from the input signal (Bekara and van der Baan, 2009) or removal of undesired energetic high-order modes related to coherent noise (Battista et al., 2007). In this context, the denoised signal s_n^d of a given seismic trace s_n can be obtained by

$$s_n^d = s_n - \sum_{k=M_1}^{M_2} \text{IMF}_n^k, \qquad (2)$$

with $M_1 \leq M_2$. Recalling that the IMFs have a frequency content that decreases with k, then, to remove high-frequency noise, one may set $M_1 = 1$ and $M_2 \ll K$. But to suppress low-frequency coherent components, one should set $M_1 > 1$ and $M_2 \leq K$. A logical combination of the standard denoising expressed in equation 2 is

$$s_n^d = \sum_{k=M_1}^{M_2} \text{IMF}_n^k, \qquad (3)$$

with $1 < M_1 \leq M_2 \leq K$. In this case, the modes whose order are lower than M_1 are associated with the high-frequency noise component of the signal, whereas the modes of order greater than M_2 are associated with coherent and noisy low-frequency components. It can be shown that the denoising schemes expressed in equations 2 and 3 are similar in applying a series of adaptive band-pass filters, but without the need to specify a priori how many filters and which cutoff frequencies to use (Colominas et al., 2014).

The parameters M_1 and M_2 need to be selected by the user. Usually, $M_1 = M_2 = 1$ is enough for most seismic applications when the original EMD is the chosen algorithm in conjunction with equation 2. In this case, due to the fact that only the first IMF is taken into account, the mode splitting is assumed to play a minor role, otherwise significant high-frequency noise could evade the denoising by mixing with the second and higher modes.
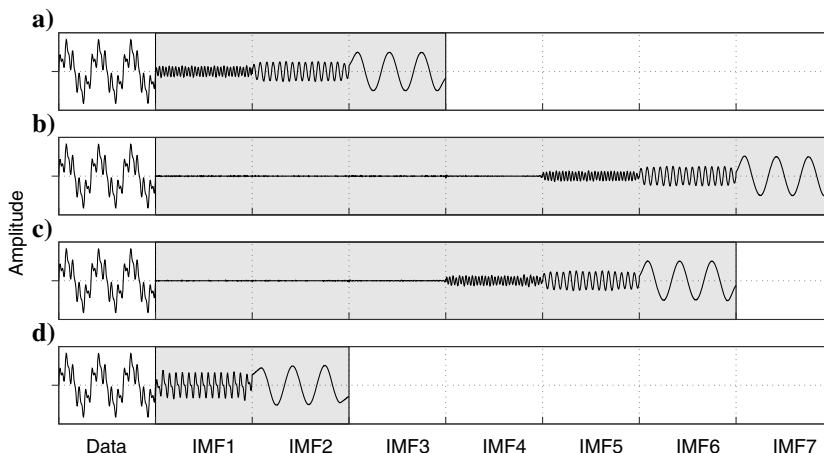

Figure 1. Decomposition of a noise-free three-tone sinusoidal into IMFs using (a) EMD, (b) CEEMD, (c) ICEEMD, and (d) WASM.
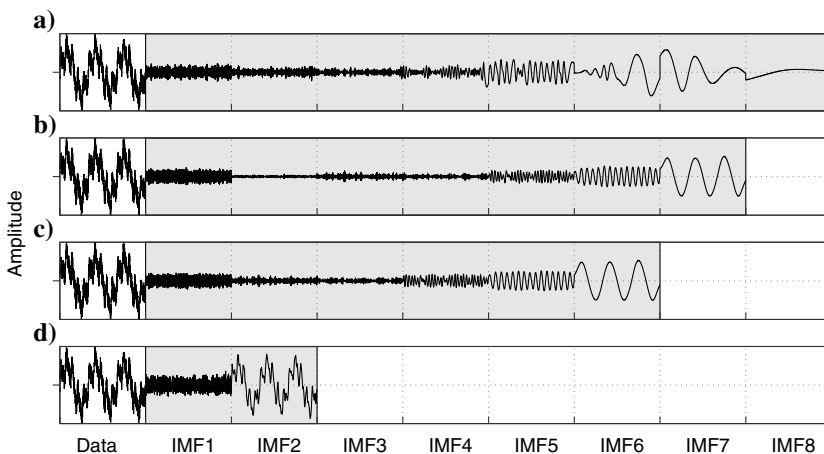

Figure 2. Decomposition of a noisy three-tone sinusoidal into IMFs using (a) EMD, (b) CEEMD, (c) ICEEMD, and (d) WASM.

When using equation 3, instead, the noise-assisted algorithms should be applied because it is desirable that the IMFs are as free of mode mixing as possible so that the selection of M_1 and M_2 is more robust.

Here, the first modes may be contaminated by residual noise inherent in the decomposition algorithm, and may also contain spurious modes (as is the case of CEEMD and shown in Figures 1b and 2b). Hence, selecting $M_1 = 1$ is not always optimal for noise reduction if the noise-assisted methods are applied. Following Kopsinis and McLaughlin (2009), a way to select parameters M_1 and M_2 appropriately consists of decomposing the signal s_n into all its modes and using a measure of their energy for mode inspection and thresholding (Han and van der Baan, 2015). However, a full mode decomposition for energy analysis of each trace with spline interpolation seems unfeasible when terabytes of data are considered.

WASM DENOISING

The methodology we will describe and propose in this paper is WASM. WASM has two advantages over the standard EMD denoising. The first advantage is the removal of random noise that is done by selecting a small value for the parameter α that is used in the method and is achieved in a fraction of the time required by the standard EMD and ICEEMD. The second one is, by increasing α, low-frequency signal attenuation is possible without the need to estimate the parameter M_2 in equation 3. This can be explained by the fact that after scanning the input data with the moving window of length N_w, the signal information of the corresponding scale is captured almost entirely in the first IMF. On the contrary, the standard EMD and ICEEMD would require a full decomposition into several modes and inspection of the higher order IMFs to estimate the appropriate M_1 and M_2 values to be used in equation 3. This full mode calculation for each trace makes the removal of low-frequency content a much slower process than the proposed WASM approach.

The cubic spline interpolation of the EMD may become a speed limitation when processing large volumes of data or when real-time outputs are required. The standard EMD is computationally efficient, but after processing thousands or millions of traces, the impact of the spline interpolations inside the sifting loop may become very significant. In addition, often the mode mixing/splitting leads to situations where most high-frequency noise energy is not trapped in the first mode, so more than one mode may have to be calculated for each trace. The noise-assisted EMDs will be more computationally intensive due to the noise realizations of the signal and their residues. In addition, the existence of spurious modes usually requires that more than one mode has to be calculated for properly denoising the input signal. As noticed by Macelloni et al. (2011), EMD with partial sifting or without a strict focus on producing a complete set of IMFs can improve the computational efficiency of the algorithm.

To speed up the denoising process while keeping the implementation and the parameter selection as simple as possible, we propose using a moving window average to compute in one step the mean local envelope required in the EMD sifting loop. The WA is calculated by convolving the signal with a normalized Hanning window of length N_w. A similar idea was proposed by Li et al. (2011) for signal decomposition using an adaptive window, but their approach requires extra user-defined parameters and more rounds of sifting.

In our implementation, the goal is filtering out the noise and not to decompose the signal into modes. This incomplete decomposition allows us to significantly improve the computational costs because a single mode or component needs to be obtained and the denoised signal is

$$s_n^d = s_n - \text{IMF}_n^1. \tag{4}$$

The described algorithm is called WASM because it is neither a complete nor a partial EMD. Note that the moving window average of the new algorithm can be applied in the sifting loops of the original EMD and ICEEMD methods for denoising either in the time domain (i.e., on a trace-by-trace process) or in the f-x domain. The latter follows the same objectives of the classical f-x deconvolution, that is, attenuating noncoherent noise in the frequency-space domain to preserve and enhance the lateral continuity of the events.

The choice of the window length N_w can be automated in a straightforward and yet effective way as follows. Because of the information of the zero crossings of the input signal gathered by the original EMD algorithm, we compute the mean zero-crossing distance D of the signal as the average of the number of samples between all successive pairs of zero crossings of the data. Then, the number of coefficients in the window is set to

$$N_w = \alpha D, \tag{5}$$

where α is a user-defined parameter. For data comprising several traces, we can choose to calculate the mean of the resulting N_w of each trace $\langle N_w \rangle$, and then use that value for processing all traces. As for the stoppage criterion required in the sifting process, we prefer to use the fixed criterion, where approximately 10 sifting iterations are usually enough for a good performance.

NUMERICAL EXAMPLES

Synthetic data examples

To illustrate the EMD methods and their denoising capabilities as compared with the proposed WASM approach, we consider a simple sinusoidal signal sampled at 1 ms for 1 s and composed of three tones: $s_n = \sin(2\pi f_0 t_n) + 0.5 \sin(2\pi f_1 t_n) + 0.25 \sin(2\pi f_2 t_n)$ with frequencies $f_0 = 3$, $f_1 = 15$, and $f_2 = 30$ Hz (Figure 1). We note that a pure sinusoid is itself an IMF because it has zero mean and is symmetric, so, ideally, any EMD should lead to only three IMFs. In the example we apply various EMDs with cubic spline interpolation and the two-threshold criterion for the sifting loop, $N_n = 20$ and $\epsilon_0 = 0.2$ for the noise-assisted methods. For the WASM scheme, we set $\alpha = 1$ and the fixed stopping criterion with 10 iterations. Because the data contain no noise, the conventional EMD decomposes the signal spontaneously in their three tones without mode mixing or splitting (Figure 1a). Figure 1b and 1c shows the CEEMD and ICEEMD results, respectively. Despite the fact that they show low-order spurious modes due to the residual left by the added noise during the algorithm, the original tones are well-isolated in the three last IMFs. Notice that the ICEEMD decreases by one the number of spurious modes of the CEEMD and exhibits a smaller amount of noise residual due to the more effective noise cancellation. The Hanning window length calculated in WASM N_w is 335 (which is practically the number of

samples entailed by one period of the lowest frequency sinusoid: $(1/f_0)/(\Delta t) \approx 333)$ focusing the algorithm for frequencies higher than f_0 and collecting the two higher frequency tones into its first mode (Figure 1d). Because the concern of the WASM algorithm is not mode decomposition, this result can be interpreted as the isolation of the higher frequency tones from the sinusoid of frequency f_0. Application of WASM to ICEEMD shows a similar behavior, and for this reason the corresponding results are not shown.

Next, we add uniform random noise to the signal of amplitude $0.3 \max(|s_n|)$ (Figure 2). Because of the noise, the original EMD now exhibits an evident mode mixing and splitting as shown in Figure 2a. As a consequence of the added noise, the isolation of the three tones in separate modes using the EMD algorithm is less effective than the results obtained from the noise-assisted approaches CEEMD and ICEEMD (Figure 2b and 2c, respectively). The WASM gives in this case $N_w = 25$ due to the more frequent zero crossings of the added random noise; the result is displayed in Figure 2d. In this case, the algorithm is focused on frequencies higher than 40 Hz $(f \approx (1/N_w)/(\Delta t))$, and as opposed to the standard EMD and ICEEMD algorithms, it traps most of the random noise energy in the first mode, leaving the three tones in the residual. Figure 3 shows the denoised signals, their differences with respect to the original noise-free data and the corresponding amplitude spectra. For simplicity, in all cases the denoising was carried out by removing only the first IMF. Even though the amplitude spectra do not clearly show that the WASM works better at removing high-frequency noise, the improvement can be appreciated by inspecting the respective differences shown on the second column of Figure 3.

As a measure of signal preservation we use the quality factor (Chen and Sacchi, 2015)

$$Q = 10 \log \frac{\|\mathbf{S}^0\|_F^2}{\|\mathbf{S}^0 - \mathbf{S}^d\|_F^2}, \qquad (6)$$

where $\mathbf{S}^0$ is the noise-free seismic data matrix of N time samples and M traces, $\mathbf{S}^d$ is the denoised result, and $\| \cdot \|_F$ is the Frobenius norm, which is the rms value of the matrix elements. When dealing with a single trace, $\mathbf{S}^0 = s_n^0$ and $\mathbf{S}^d = s_n^d$. A larger value of the factor Q indicates better denoising performance. In this particular example, the largest Q and the fastest result are attained when using the WASM (Table 1).

A second example consists of a synthetic trace composed of three Ricker wavelets of unit amplitude sampled at 2 ms with peak frequencies of 30, 25, and 20 Hz, respectively (Figure 4, first row). The trace is then contaminated with zero-mean uniform random noise of amplitude 0.2 plus a coherent chirp of amplitude 0.8 with a linear frequency sweep that goes from 250 to 200 Hz within the time window considered (Figure 4, second row). For simplicity, all EMDs are performed under suboptimal conditions for the stoppage criterion and the number of noise realizations to simulate the computational requirements of processing seismic data and also to

obtain comparable computational performances and results. Thus, we set again $N_n = 20$ and a fixed number of 10 sifting iterations in all cases. We recall that the WASM scheme runs with the fixed criterion as default. In addition, we set $\alpha = 4$, which leads to $N_w = 17$.

Rows 3–5 of Figure 4 show the results of the EMD denoising when using the EMD, ICEEMD, and WASM. Notice how the denoising is less effective from left to right due to the frequency content of the chirp. The results show comparable Q values for EMD and ICEEMD and a better denoising by the WASM algorithm (Table 2). The differences between the original noise-free synthetic and the denoised result as well as the amplitude spectra of the denoised signals clearly demonstrate the advantages of denoising by means of the WASM scheme. Not only does this approach show
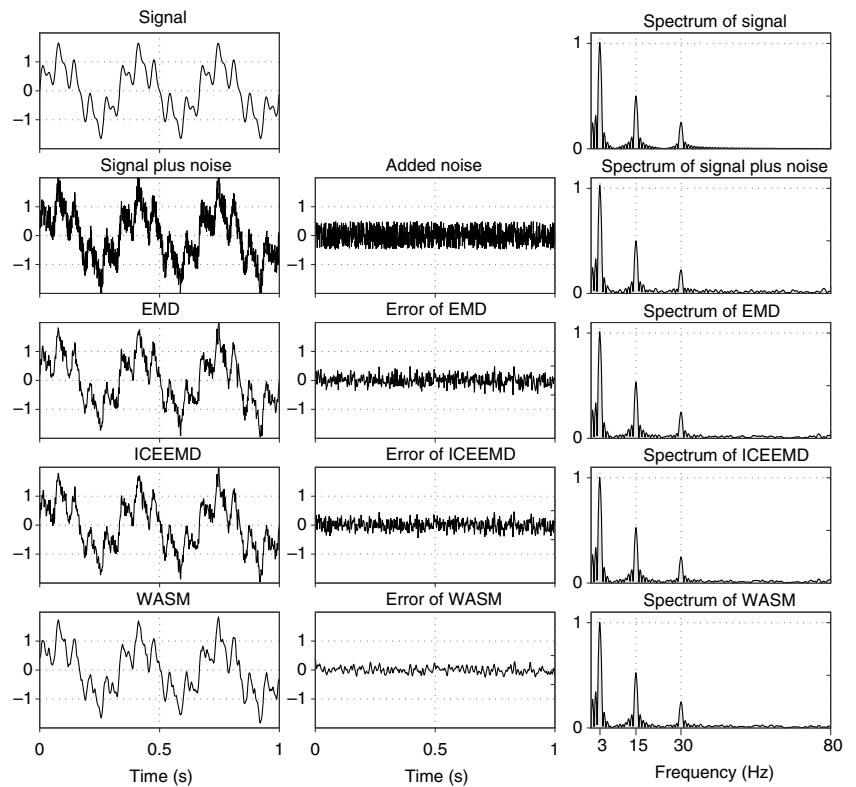

Figure 3. Denoising of a three-tone sinusoidal, the corresponding errors, and amplitude spectra. First row: noise-free signal. Second row: noisy signal. Rows 3–5: denoised signals after EMD, ICEEMD, and WASM, respectively. In all cases, the denoising was carried out by removing the first IMF only.

Table 1. Quality factor Q and computational time required in denoising the sinusoidal data of Figure 2 for different algorithms. The CPU is a Intel Core i3 at 3.10 GHz with 3 GB of RAM.

Algorithm	Q	Time (s)
EMD	8.2	0.017
ICEEMD	9.7	0.887
WASM	13.5	0.0013

a lower residual and a better high-frequency noise attenuation, but also it requires a lower computational cost than the standard methods (Table 2).

Figure 5 shows the average Q factor for the noise-free synthetic trace of Figure 4 plus uniform random noise with amplitudes ranging from zero to one in steps of 0.1 after 100 realizations per step. In all cases, the denoising was carried out by removing the first IMF only. The results indicate that a similar denoising to that of the standard EMDs algorithms can be expected from the WASM alternatives.

The previous examples demonstrate the performance of the various EMDs implementations when applied on a trace-by-trace basis. However, when processing 2D seismic data it is usually desired to carry out the denoising by enhancing the lateral continuity of the events. Bekara and van der Baan (2009), Chen and Ma (2014),

and Han and van der Baan (2015) show that the f-x EMD methods can significantly improve the noise suppression of the classical f-x deconvolution, specially when coherent noise is present. In its simplest form, the f-x EMD consists of taking the 1D Fourier transform of the t-x data on a selected time window, and applying the EMD to the real and imaginary parts in the offset direction, one frequency at a time. Finally, the first IMFs are removed and the data are transformed back to the original domain to produce the filtered data (Bekara and van der Baan, 2009).

To analyze the results of processing 2D seismic data using the proposed method, we implemented the f-x versions of the algorithms and compared their performances against the classical f-x deconvolution. For this purpose, we consider a portion of the post-stack data set of the well-known 2D Marmousi model (Versteeg and Grau, 1991; Figure 6, first row), and add band-limited Gaussian random noise as well as erratic noise with different S/Ns. The erratic noise is also generated as random noise, but it is multiplied by a factor of five at certain traces. We then denoise this data set using f-x EMD, f-x WASM, and the classic f-x deconvolution (Canales, 1984; Gulunay, 1986).

The results are shown in Figure 6 for S/N = 7. The quality factors are summarized in Table 3. These results indicate that the signal preservation of the f-x WASM is similar to the other f-x implementations. Comparable amounts of random noise are attenuated by the three methods, and they all preserve the reflections quite well even for those traces that contain the erratic noise. In terms of computational cost, the WASM approach obtains the results in less time. For the case depicted in Figure 6 (300 traces), f-x ICEEMD takes 33 s, f-x EMD takes 0.53 s, and f-x WASM takes only 0.03 s. The WASM scheme is 18 times faster than the EMD implementation. Note that in this example, and to obtain Q values comparable with those for the WASM, we subtracted two IMFs for f-x EMD and ICEEMD methods. In addition, we set $N_n = 10$ and $N_w = 29$ for ICEEMD and WASM, respectively.

Field data examples

We provide two examples to test the proposed methods on field data: (1) a microseismic data set containing high levels of different kinds of noises and (2) a 2D seismic line that contains random and coherent noise. In the first case, we apply the trace-by-trace EMDs implementations, whereas in the second case we use the corresponding f-x EMDs approaches.

As is well-known, microseismic records are often contaminated by random and coherent noise, such as ground roll, geophone reverberations due to tube waves and waves propagating in the well casing (Tary and van der Baan, 2012), sensor-induced electrical noise from incorrectly placed receivers and cabling (Maxwell, 2014), and background noise from human activities (Tselentis et al., 2012). In addition, data sets are usually very large and the S/N is very low. Thus, a fast and simple denoising algorithm may be
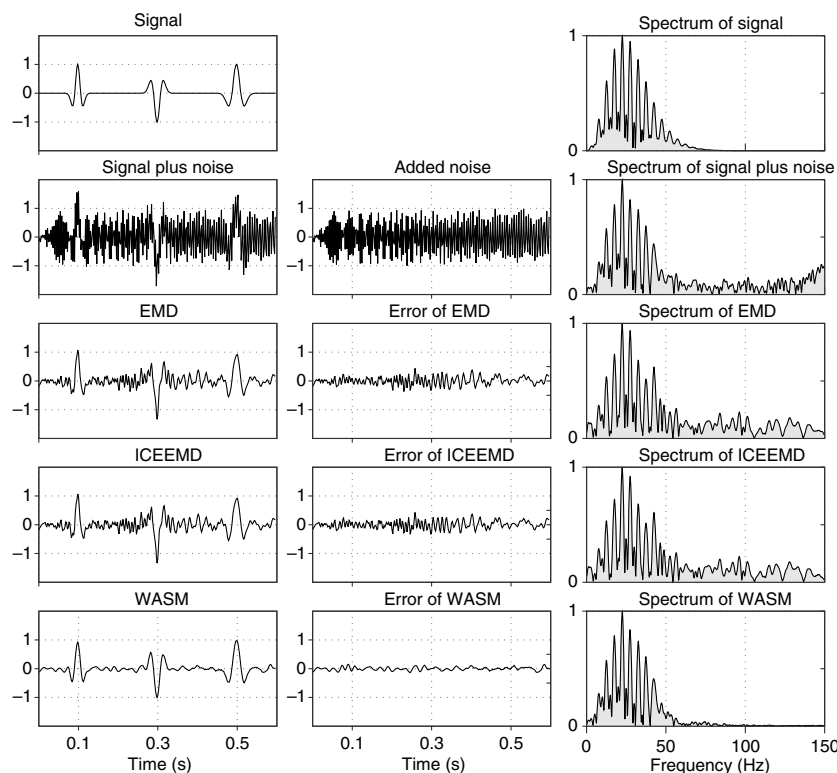

Figure 4. Denoising of a synthetic trace containing three Ricker wavelets, the corresponding errors, and amplitude spectra. First row: noise-free signal. Second row: noisy signal. Rows 3–5: denoised signals after EMD, ICEEMD, and WASM, respectively. In all cases, the denoising was carried out by removing the first IMF only.

Table 2. Quality factor Q and computational time required in denoising the synthetic trace of Figure 4 for different algorithms.

Algorithm	Q	Time (s)
EMD	−0.9	0.0023
ICEEMD	−0.6	0.357
WASM	8.2	0.0011

necessary or mandatory for their processing, especially when real-time decisions are to be made on the injection site.

The first panel of Figure 7 shows a small portion of a microseismic data set (1C only) that was recently acquired in South America during a downhole hydraulic fracturing monitoring survey. Note that the data set exhibits random, erratic, and coherent noise that masks the continuity of two microseismic events approximately at 3.3 and 4.1 s. The second panel shows the denoised traces using

the WASM, whereas the third panel shows the corresponding residual, in which the noise components have been mostly trapped. In this example, we obtained a good noise attenuation by setting $\alpha = 3$, which yields for the whole gather a value of $\langle N_w \rangle = 49$. For comparison, Figure 7 also shows the denoised traces and residues that result from the standard EMD and ICEEMD implementations. In all cases, we used the fixed stoppage criterion method and 10 sifting iterations. We observe that the amount of noise at-
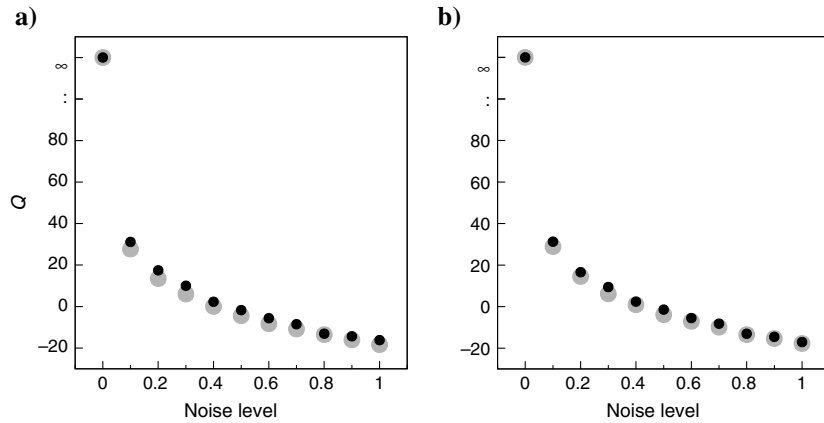
Figure 5. Mean quality factor Q after denoising 100 realizations of the synthetic trace of Figure 4 with random uniform noise of variable amplitude. (a) EMD (gray circles) and WASM (black circles). (b) ICEEMD (gray circles) and WASM in ICE-EMD (black circles).
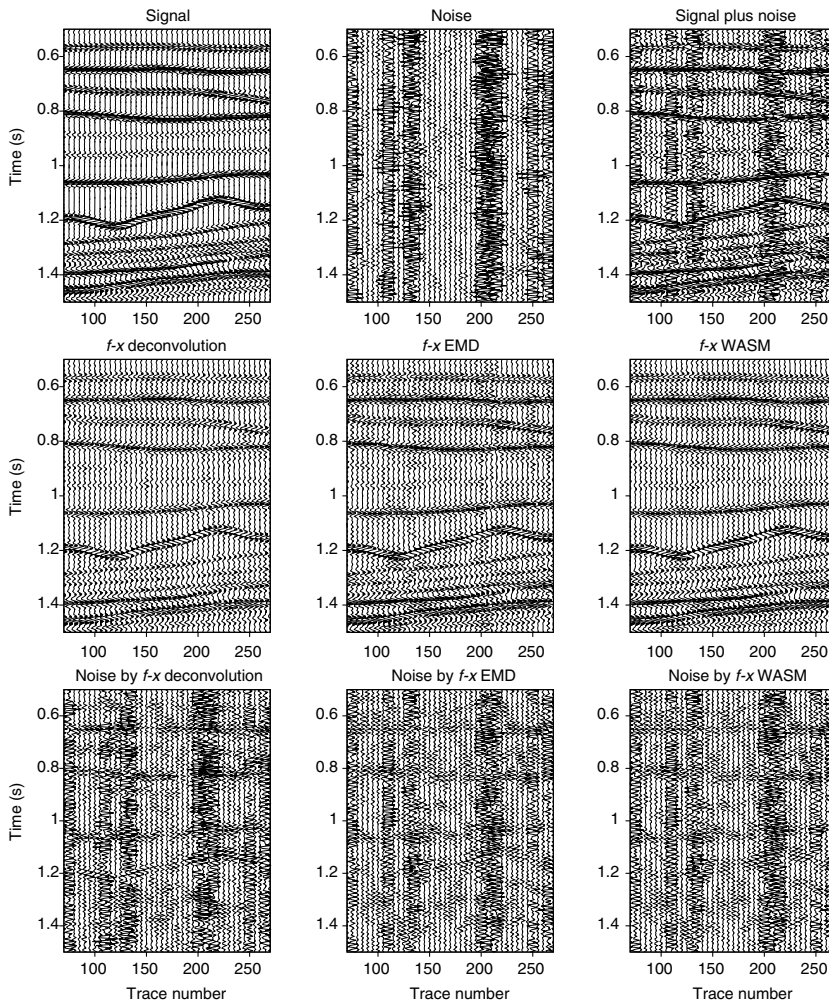

Figure 6. Marmousi data example. First row: portion of the Marmousi poststack data set, added noise and data plus noise. Second row: denoising using f-x deconvolution, f-x EMD, and f-x WASM. Third row: corresponding difference sections. The processed data set contains 300 traces, but for simplicity we show one of every five traces.

tenuation is very similar in all cases. In the case of the standard EMD and ICEEMD, and after a visual inspection of the denoised data, we decided to remove the first two IMFs for optimum results.

The second field data example illustrates the denoising in the f-x domain of the 2D Alaska North Slope Land Line 31–81 (Geological Survey, 1981). This data set, shown in Figure 8, consists of 536 traces and 3000 time samples per trace with a sampling interval of 2 ms. The WASM algorithm gives a denoised result similar or better than the classical f-x deconvolution and the standard f-x EMD. For the f-x deconvolution, an operator length of 20 traces is used for every frequency slice and for every window comprising 100 traces in the offset direction. Similar to Han and van der Baan (2015), the algorithm is implemented between 0 Hz and 60% of the Nyquist frequency, zeroing out the frequencies beyond this limit.

Figure 9 shows an enlarged portion of the data, the denoised results, and the corresponding difference sections. We observe that while the f-x deconvolution subtracts mostly random noise, the EMD-based f-x approaches remove random and coherent noise without harming signal integrity. Inspection from the results depicted in Figures 8 and 9 show that the standard f-x EMD and

the f-x WASM approaches provide very similar results, with the WASM alternative suppressing more coherent noise. In contrast to the f-x EMD, the f-x WASM algorithm obtains the results in a fraction of the time. In effect, the f-x EMD takes approximately 19 s (removing two IMFs), whereas the f-x WASM takes 2.1 s with a window length of 45 samples.

DISCUSSION

A brief analysis of the computational complexity of the algorithms shows that the WASM approach is able to obtain the first mode required for the denoising using a significantly smaller number of operations than the standard EMD technique. In effect, the WASM algorithm avoids the step of cubic interpolation for both envelopes to obtain the mean envelope for each sifting iteration of the standard EMD. For a single cubic natural spline interpolation, the computational complexity for a signal of length N is $\mathcal{O}(N + \log_2 N)$: $\mathcal{O}(N)$ to generate the spline once and $\mathcal{O}(\log_2 N)$ to evaluate the spline at a single point using bisection (Press et al., 1996). For a whole trace, this last step involves $\mathcal{O}(N \log_2 N)$ operations for each envelope. Then, for one sifting iteration, the computational complexity involved in the calculation of the mean envelope in EMD is $\mathcal{O}(N + N \log_2 N)$. For a window of length N_w, the proposed moving WASM requires $\mathcal{O}(N_w)$ evaluations to obtain the mean envelope at a single point with $N_w \ll N$ in the practical applications. For the N points and one sifting iteration, this scheme requires $\mathcal{O}(NN_w)$ calculations or $\mathcal{O}(N \log_2 N)$, with the aid of an FFT.

The mode mixing/splitting problem inherent to the standard EMD method often demands more than one mode for properly denoising the high-frequency noise content. The CEEMD and ICEEMD methods greatly reduce the mode mixing and splitting and they also require more than one mode for the same task because of the presence of spurious modes. Unlike those methods, the WASM scheme is more efficient in terms of computational cost and is able to attenuate the same or even better levels of noise en-

Table 3. Quality factor Q versus S/N using f-x deconvolution, f-x EMD, f-x ICEEMD, and f-x WASM for a portion of the Marmousi data set.

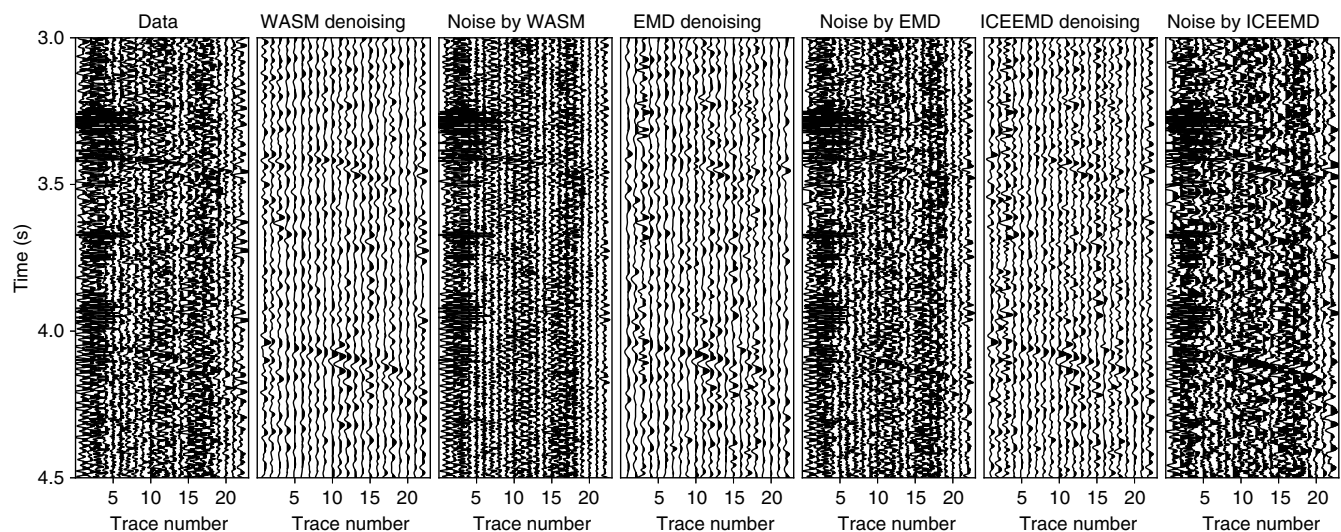| Algorithm	S/N				
	10	7	5	2	1
f-x deconvolution	14.5	13.8	12.7	6.5	−0.4
f-x EMD	14.9	14.1	12.5	2.7	−8.5
f-x ICEEMD	15.3	14.6	12.9	3.3	−7.6
f-x WASM	15.5	15.1	14.3	8.7	−0.8

Figure 7. Field data example 1: the first panel shows the raw microseismic data. The remaining panels show the denoising results and the corresponding residuals using WASM, EMD, and ICEEMD, respectively.

ergy using a single mode, as demonstrated by the various synthetic and field data examples.

Table 4 further shows the expected speedup when processing a microseismic data set. The table depicts the computation times after running 10 times the denoising algorithms using the cubic and the WASM schemes associated with the EMD and ICEEMD approaches on the whole 10 s long microseismic record, of which a 1.5 s section was shown in Figure 7. In all cases the alternative algorithm is faster by at least one order of magnitude, being the relative decrease in computational time more noticeably for the ICEEMD algorithm, as expected. The WASM is more than 30 times faster than the standard EMD, whereas the WASM in ICEEMD is approximately 100 times faster than ICEEMD. This significant drop in computational time then allowed for the processing of the 3000 traces of the whole data set in a reasonable time frame (of seconds), which would have been of several minutes or hours, respectively, if the standard EMD and ICEEMD implementations were used.

The results from the f-x denoising suggest that the WASM alternative provides comparable or better results than those of the other f-x empirical implementations with less computational effort. If a medium size seismic data volume of approximately one million traces was to be processed, a simple estimate based on the measured processing times required to process the Alaska data set depicted in Figure 8 indicates that the standard f-x approach would take approximately 10 h, whereas the f-x WASM approach only 1 h on the standard PC we used for this analysis (Intel® Core™ i3 at 3.10 GHz with 3 GB of RAM). This speedup of approximately 10 (in the Marmousi example, the speedup was approximately 18) is a rough estimate because the computation time strongly depends on the selected parameters, especially the number of IMFs removed in the standard EMDs implementations (recall that the WASM algorithm always selects one), or the number of sifting iterations.

As one would expect, when large volumes of seismic data are considered, a full or partial decomposition into several IMFs for each trace or frequency slice seems unfeasible for mode inspection before denoising. The sequential character of the IMFs calculations of the EMDs implies that this processing is not parallelizable for a single trace. For noise attenuation of different scales, the proposed WASM denoising can bypass this limitation if different window lengths are supplied (Gómez and Velis, 2015).

In this sense, the f-x WASM algorithm permits selection of an adaptive window length for each frequency slice. This offers the analyst the option of controlling the smoothness of the final result. When using one representative value for

all traces (as done in the second field data example), smoother results are obtained that are comparable with the removal of more than one or two IMFs when using the standard f-x EMD and f-x ICE-EMD approaches. This analysis is beyond the scope of this work.
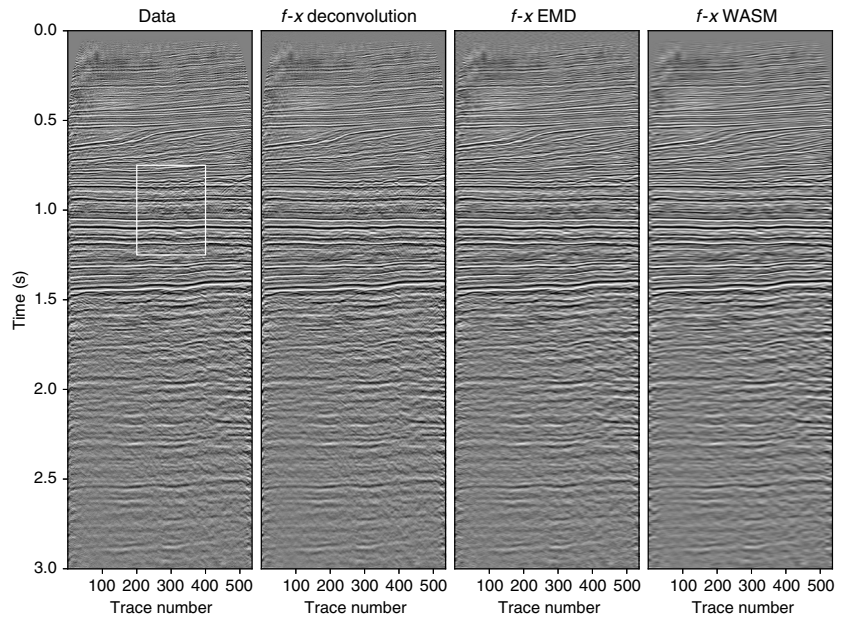

Figure 8. Field data example 2: the 2D Alaska field data set, in which the white box indicates a section to be enlarged in Figure 9. Denoising using f-x deconvolution, f-x EMD, and f-x WASM are shown.

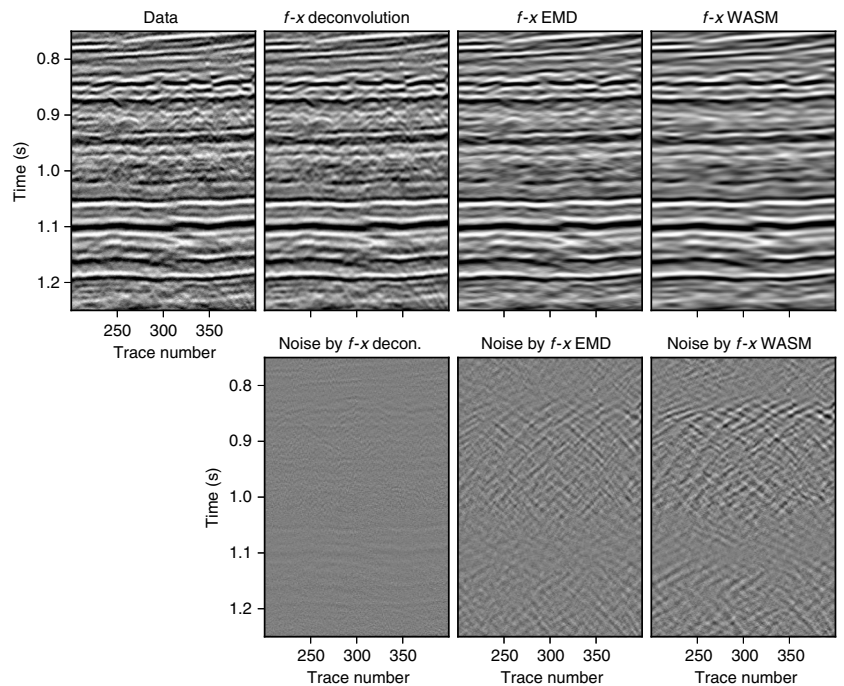
Figure 9. Field data example 2: first row: enlarged section from the Alaska data set and denoising using f-x deconvolution, f-x EMD, and f-x WASM, respectively. Second row: corresponding difference sections.

Table 4. Computational time (in s) required to denoise a 10 s microseismic record using EMD, ICEEMD, and WASM. The last row shows the corresponding speedups.

Algorithm	EMD (s)	ICEEMD (s)
Cubic	8.25	618.70
WASM	0.21	6.06
Cubic/WASM	38	102

CONCLUSIONS

The proposed WASM technique speeds up computation time to a range that EMD-like denoising can be efficiently applied to large volumes of data on a standard workstation. The bottleneck in all previous EMD-based denoising implementations is in the polynomial interpolation step used to find the mean local envelope. We have simplified this costly step by a simple, yet elegant solution, using WASM. This translates into reducing the computational complexity of the original algorithm and keeping to a minimum the number of user-defined parameters required to filter the data.

The WASM implementations inspired by the EMD and ICEEMD sifting algorithms were successfully applied for suppressing coherent, erratic, and random noise in synthetic and field data examples. These include algorithms for trace-by-trace processing, which were applied for denoising a microseismic data set containing noise of a different nature, and to process seismic data in the f-x domain to enhance the lateral continuity of the events, as in the case of the classical f-x deconvolution. This implementation was applied to denoise a Marmousi section containing random and erratic noise and to the Alaska field data set that contained random and coherent noise.

Denoised signals by means of WASM are equivalent and often superior to those obtained by using the standard algorithms in terms of signal preservation and denoising capabilities. In addition, WASM led to speedups of one order of magnitude or more in some cases, allowing processing of relatively large volumes of data on a standard PC within a reasonable time frame in minutes, and with minimum user intervention. We believe that WASM is a good complement to the standard empirical decomposition techniques and f-x deconvolution.

ACKNOWLEDGMENTS

REFERENCES

Battista, B. M., C. Knapp, T. McGee, and V. Goebel, 2007, Application of the empirical mode decomposition and Hilbert-Huang transform to seismic reflection data: Geophysics, **72**, no. 2, H29–H37, doi: 10.1190/1.2437700.

Bekara, M., and M. van der Baan, 2009, Random and coherent noise attenuation by empirical mode decomposition: Geophysics, **74**, no. 5, V89–V98, doi: 10.1190/1.3157244.

Canales, L., 1984, Random noise reduction: 54th Annual International Meeting, SEG, Expanded Abstracts, 525–527.

Chen, K., and M. D. Sacchi, 2015, Robust reduced-rank filtering for erratic seismic noise attenuation: Geophysics, **80**, no. 1, V1–V11, doi: 10.1190/geo2014-0116.1.

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.

Colominas, M. A., G. Schlotthauer, and M. E. Torres, 2014, Improved complete ensemble EMD: A suitable tool for biomedical signal processing: Biomedical Signal Processing and Control, **14**, 19–29, doi: 10.1016/j.bspc.2014.06.009.

Fomel, S., 2008, Nonlinear shaping regularization in geophysical inverse problems: 78th Annual International Meeting, SEG, Expanded Abstracts, 2046–2051.

Fomel, S., 2013, Seismic data decomposition into spectral components using regularized nonstationary autoregression: Geophysics, **78**, no. 6, O69–O76, doi: 10.1190/geo2013-0221.1.

Gómez, J. L., and D. R. Velis, 2015, A fast empirical mode decomposition for noise attenuation of seismic data: 85th Annual International Meeting, SEG, Expanded Abstracts, 4760–4764.

Gulunay, N., 1986, FXDECON and complex Wiener prediction filter: 56th Annual International Meeting, SEG, Expanded Abstracts, 279–281.

Han, J., and M. van der Baan, 2015, Microseismic and seismic denoising via ensemble empirical mode decomposition and adaptive thresholding: Geophysics, **80**, no. 6, KS69–KS80, doi: 10.1190/geo2014-0423.1.

Hooshmand, A., J. Nasseri, and H. R. Siahkoohi, 2012, Seismic data denoising based on the complete ensemble empirical mode decomposition: International Geophysical Conference and Oil & Gas Exhibition, SEG and UCTEA Chamber of Geophysical Engineers of Turkey, Expanded Abstracts, doi: 10.1190/IST092012-001.28.

Huang, N. E., Z. Shen, and S. R. Long, 1999, A new view of nonlinear water waves: The Hilbert spectrum 1: Annual Review of Fluid Mechanics, **31**, 417–457, doi: 10.1146/annurev.fluid.31.1.417.

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 Series A: Mathematical, Physical and Engineering Sciences, **454**, 903–995.

Huang, N. E., and Z. Wu, 2008, A review on Hilbert-Huang transform: Method and its applications to geophysical studies: Reviews of Geophysics, **46**, 1–23, doi: 10.1029/2007RG000228.

Kopsinis, Y., and S. McLaughlin, 2009, Development of EMD-based denoising methods inspired by wavelet thresholding: IEEE Transactions on Signal Processing, **57**, 1351–1362, doi: 10.1109/TSP.2009.2013885.

Li, C., X. Wang, Z. Tao, Q. Wang, and S. Du, 2011, Extraction of time varying information from noisy signals: An approach based on the empirical mode decomposition: Mechanical Systems and Signal Processing, **25**, 812–820, doi: 10.1016/j.ymssp.2010.10.007.

Macelloni, L., B. M. Battista, and C. Knapp, 2011, Optimal filtering high-resolution seismic reflection data using a weighted-mode empirical mode decomposition operator: Journal of Applied Geophysics, **75**, 603–614, doi: 10.1016/j.jappgeo.2011.09.018.

Mandic, D. P., N. U. Rehman, Z. Wu, and N. E. Huang, 2013, Empirical mode decomposition-based time-frequency analysis of multivariate signals: The power of adaptive data analysis: IEEE Signal Processing Magazine, **30**, 74–86, doi: 10.1109/MSP.2013.2267931.

Maxwell, S., 2014, Microseismic imaging of hydraulic fracturing: SEG Distinguished Instructor Series 17.

Meignen, S., and V. Perrier, 2007, A new formulation for empirical mode decomposition based on constrained optimization: IEEE Signal Processing Letters, **14**, 932–935, doi: 10.1109/LSP.2007.904706.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1996, Numerical recipes in Fortran 77 and Fortran 90: Cambridge University Press.

Rilling, G., P. Flandrin, and P. Gonçalves, 2003, On empirical mode decomposition and its algorithms: IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing, NSIP-03, 8–11.

Tary, J. B., and M. van der Baan, 2012, Potential use of resonance frequencies in microseismic interpretation: The Leading Edge, **31**, 1338–1346, doi: 10.1190/tle31111338.1.

Torres, M. E., M. A. Colominas, G. Schlotthauer, and P. Flandrin, 2011, A complete ensemble empirical mode decomposition with adaptive noise: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 4144–4147.

Tselentis, G.-A., N. Martakis, P. Paraskevopoulos, A. Lois, and E. Sokos, 2012, Strategy for automated analysis of passive microseismic data based on S-transform, Otsu's thresholding, and higher order statistics: Geophysics, **77**, no. 6, KS43–KS54, doi: 10.1190/geo2011-0301.1.

U.S. Geological Survey, 1981, Alaska 2D land line 31-81, http://wiki.seg.org/wiki/ALASKA_2D_LAND_LINE_31–81, accessed 20 January 2015.

Versteeg, R. J., and G. Grau, eds., 1991, The Marmousi experience: 1990 EAGE workshop on practical aspects of seismic data inversion: EAGE.

Wu, Z., and N. E. Huang, 2009, Ensemble empirical mode decomposition: A noise-assisted data analysis method: Advances in Adaptive Data Analysis, **1**, 1–41, doi: 10.1142/S1793536909000047.

Yang, W., R. Wang, J. Wu, and Y. Chen, 2014, Random noise attenuation using a new spectral decomposition method: 84th Annual International Meeting, SEG, Expanded Abstracts, 4366–4370.

Yeh, J.-R., J.-S. Shieh, and N. E. Huang, 2010, Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method: Advances in Adaptive Data Analysis, **2**, 135–156, doi: 10.1142/S1793536910000422.

Zhang, J., R. Yan, R. X. Gao, and Z. Feng, 2010, Performance enhancement of ensemble empirical mode decomposition: Mechanical Systems and Signal Processing, **24**, 2104–2123, doi: 10.1016/j.ymssp.2010.03.003.