GEOPHYSICS, VOL. 74, NO. 5 (SEPTEMBER-OCTOBER 2009); P. V89–V98, 9 FIGS.
10.1190/1.3157244

Random and coherent noise attenuation by empirical mode decomposition

Maïza Bekara[1] and Mirko van der Baan[2]

ABSTRACT

We have devised a new filtering technique for random and coherent noise attenuation in seismic data by applying empirical mode decomposition (EMD) on constant-frequency slices in the frequency-offset (f-x) domain and removing the first intrinsic mode function. The motivation behind this development is to overcome the potential low performance of f-x deconvolution for signal-to-noise enhancement when processing highly complex geologic sections, data acquired using irregular trace spacing, and/or data contaminated with steeply dipping coherent noise. The resulting f-x EMD method is equivalent to an autoadaptive f-k filter with a frequency-dependent, high-wavenumber cut filtering property. Removing both random and steeply dipping coherent noise in either prestack or stacked/migrated sections is useful and compares well with other noise-reduction methods, such as f-x deconvolution, median filtering, and local singular value decomposition. In its simplest implementation, f-x EMD is parameter-free and can be applied to entire data sets without user interaction.

INTRODUCTION

Spatial prediction filtering in the frequency-offset (f-x) domain is an effective method for random noise attenuation. Originally proposed by Canales (1984), the idea exploits signal predictability in the spatial direction. Noise-free events that are linear in the t-x domain manifest themselves as a superposition of harmonics in the f-x domain. These harmonics are perfectly predictable using an autoregressive (AR) filter. When the data are corrupted by random noise, the "signal" is considered to be the part that can be predicted by the AR filter and the "noise" is the rest.

In reality, seismic events do not follow Canales's assumptions exactly. They are spatially nonstationary.[3] Examples include a hyperbolic moveout or a linear event with an amplitude that varies with distance. The "signal" is no longer mapped to a superposition of simple harmonics but rather a superposition of nonstationary ones. More distortion is added to Canales's model when the seismic data are sampled irregularly in the spatial direction. The use of a recursion-type filter such as an AR filter is not necessarily optimal in this case because it implicitly assumes regular spacing.

Standard spatial filtering techniques such as f-x deconvolution cope with nonlinearity and nonstationarity by filtering the data over a short spatial window in which the data are assumed to be piecewise stationary and linear. This leaves the choice of finding optimal parameters for the window size and the filter length to the processing specialist. The selection of these parameters depends strongly on the smoothness of the data and should, in the optimum case, vary with frequency.

In this paper, we propose an alternative for seismic noise attenuation in the f-x domain using a unique technique for time-frequency analysis. Empirical mode decomposition (EMD) was developed by researchers at NASA with the specific aim of analyzing nonlinear and nonstationary data (Huang et al., 1998). Therefore, it constitutes an interesting and unique domain to design data-adaptive filters for the reduction of random and coherent noise.

We start this paper with a description of f-x deconvolution. Then, we describe EMD and show how it can be used in the f-x domain for coherent and random noise attenuation. Next, we compare its performance with those of other noise-suppression techniques such as f-x deconvolution, local median filtering, and local singular-value decomposition (Bekara and van der Baan, 2007) on a variety of real data sets. We give a physical interpretation of the filtering properties of f-x domain EMD and discuss its strengths and weaknesses compared to the other methods.

Manuscript received by the Editor 2 November 2008; revised manuscript received 6 March 2009; published online 14 August 2009.
[1]Formerly at University of Leeds, School of Earth and Environment: Earth Sciences, Leeds, U. K., presently at PGS, Brooklands Weybridge, U. K. E-mail: Maiza.Bekara@pgs.com.
[2]Formerly at University of Leeds, School of Earth and Environment: Earth Sciences, Leeds, U. K., presently at University of Alberta, Department of Physics, Edmonton, Alberta, Canada. E-mail: Mirko.VanderBaan@ualberta.ca.

[3]By nonstationary harmonic we mean a harmonic whose amplitude, wavenumber, or phase changes with distance.

f-x DOMAIN FILTERING

Signal model

Consider a noise-free seismic section $s(t,x)$ at time t and offset x. For simplicity, assume the section contains a single linear event with velocity V and constant amplitude so that

$$s(t,x) = w(t - x/V), \tag{1}$$

where $w(t)$ is the source wavelet. We denote its Fourier transform by $W(f)$. The f-x domain representation of $s(t,x)$ is obtained by taking the Fourier transform of each trace so that

$$S(f,x) = W(f)e^{i2\pi fx/V}. \tag{2}$$

We assume the trace spacing is regular, i.e., $x = m\Delta x$, where $m = 1,2,\ldots,M$, with M being the number of traces in the section. At each frequency f we consider the complex sequence

$$S_f(m) \equiv S(f,m\Delta x), \quad m = 1,2,\ldots,M. \tag{3}$$

It is easy to show that this sequence defines a linear recursion along offset

$$S_f(m) = a_1(f)S_f(m-1), \quad m \geq 2, \tag{4}$$

where $a_1(f) = \exp(i2\pi f\Delta x/V)$. This recursion is a first-order differential equation known also as an autoregressive (AR) model of order 1 and represents a single complex-valued harmonic.

A noise-free linear event in the t-x domain is therefore perfectly predictable in the f-x domain by finding the recursion coefficient $a_1(f)$. Similarly, it can be shown that the superposition of p linear events in the t-x domain is equivalent to the superposition of p complex harmonics in the f-x domain; also, it can be predicted perfectly by an AR filter of order p (Tufts and Kumaresan, 1980; Harris and White, 1997).

Linear prediction filtering using an AR model: *f-x* deconvolution

So far, the recursive equation has been derived to predict a noise-free superposition of harmonics. Additive noise corrupts the data in practice. Therefore,

$$Y_f(m) = S_f(m) + \epsilon_f(m), \tag{5}$$

where $\epsilon_f(m)$ represents a spatially varying complex noise sequence. Filtering is performed on the observed data $Y_f(m)$ to estimate the noise-free signal $S_f(m)$. Canales (1984) argues a good estimate of signal $S_f(m)$ is the predictable part of data $Y_f(m)$ obtained by an AR model. Consequently, the data $Y_f(m)$ are filtered using an AR filter with coefficients estimated directly from the data. This operation is similar to a deconvolution process and the standard name of this technique is f-x deconvolution.

In practice, f-x deconvolution is performed over a short sliding window in space and time to cope with the nonstationarity of the seismic record and to reduce artifacts such as ghost events and high-frequency dispersion (Galbraith, 1991). The AR model is fitted over a short spatial window and then used to predict one sample ahead. The operation is repeated by sliding the window along the offsets one sample at a time. Prediction can either be done in a single direction, e.g., with increasing offset, or by averaging forward and backward predictions thereby improving the filtering quality.

Limitations of *f-x* deconvolution

Real seismic data often include events more complicated than the basic model in equation 1. Examples include a parabolic event or a linear event with amplitudes that vary with offset. The noise-free model, equation 4, is no longer a superposition of harmonics but rather a superposition of nonstationary signals in the sense that the noise-free model no longer has a sparse representation in a Fourier basis.

More distortion is added to the noise-free model when the seismic data are not uniformly sampled in the spatial direction. This can be common in land surveys. In this case, even for a linear event, the corresponding f-x domain recursion defined in equation 4 loses its simplicity with coefficient a_1 varying from trace to trace. This leads to nonstationary behavior of the AR filter if regular trace spacing is assumed.

Standard spatial filtering techniques such as f-x deconvolution use short temporal and spatial analysis windows in which piecewise linearity and stationarity are assumed. This is why f-x deconvolution can fail to deliver satisfactory results when processing data acquired over complex geologies where signals are always nonstationary and acquisition grids can be irregular. In addition, an ideal application of f-x deconvolution would use a variable filter length that increases with frequency because low-frequency wavefields are less complex and easier to predict than high-frequency ones. This is done rarely in practice.

What alternative can be used for AR filtering in the f-x domain? A convenient solution consists of using an adaptive and nonlinear filtering method that can handle nonstationarity and nonlinearity better. In addition, this method should be nonrecursive to make it less sensitive to irregular sampling. A unique filtering method that satisfies these requirements can be devised using EMD.

EMPIRICAL MODE DECOMPOSITION

Background

Empirical mode decomposition decomposes a data series into a finite set of signals called intrinsic mode functions (IMFs). They represent the different oscillations embedded in the data. They are constructed to satisfy two conditions: (1) The number of extrema and the number of zero crossings must be equal or differ at most by one, and (2) at any point the mean value of the envelope defined by the local maxima and the envelope defined by the local minima must be zero.

These conditions are necessary to ensure that each IMF has a localized frequency content by preventing frequency spreading because of asymmetric waveforms. A Fourier transform decomposes a signal into a sum of single-frequency constant-amplitude harmonics, whereas the IMFs are elementary amplitude/frequency-modulated harmonics that can capture the nonstationary and nonlinear variations in the signal (Huang et al., 1998).

The IMFs are computed iteratively starting with the most oscillatory one. The decomposition method uses the envelopes defined by the local maxima and the local minima of the data series. Once the extrema are identified, all the local maxima are interpolated by a cubic spline to construct the upper envelope. The procedure is repeated for local minima to produce the lower envelope. At every point in the series, the mean of the upper and lower envelopes is calculated and subtracted from the initial data and the same interpolation scheme is

reiterated on the remainder. This sifting process terminates when the mean envelope is reasonably close to zero everywhere and the resultant signal is designated as the first IMF.

The first IMF is subtracted from the original signal and the difference is treated as a new signal on which the same sifting procedure is applied to obtain the next IMF. The decomposition is stopped when the last IMF has a small amplitude or becomes monotonic.

Empirical mode decomposition has interesting properties that make it an attractive tool for signal analysis. It results in complete signal decomposition, i.e., the original signal is reconstructed by summing all IMFs. No loss of information is incurred. The EMD is a quasi-orthogonal decomposition in that the crosscorrelation coefficients between the different IMFs are always close to zero. This minimizes energy leakage between the IMFs. The EMD acts similarly to a dyadic filter bank (Flandrin et al., 2005), i.e., the IMFs correspond to the output of a time-varying subband filtering of the original signal and have partially overlapping frequency contents.

Some important features distinguish EMD from other decomposition methods such as Fourier or wavelet transforms. The latter transforms decompose the signal using a predefined basis (e.g., sines and cosines or a mother wavelet) whereas no a priori basis is used by EMD; the basis is derived adaptively from the data. Furthermore, EMD uses spline interpolations and is therefore less affected by irregular sample spacing in contrast to Fourier and Wavelet transforms, which require regular sampling for efficient implementations.

An example of applying EMD on a real signal is shown in Figure 1. The original signal (Figure 1a) is nonstationary and oscillating. The IMFs are derived iteratively starting with the fastest component, IMF 1, to the slowest one, IMF 7 (Figure 1b-h). The IMF 1 captures the highest-frequency oscillations in the data and the IMFs become subsequently smoother. The last IMF represents the general trend in the data. The shapes of IMFs 2 and 3 suggest partially overlapping frequency contents although the common frequencies occur at different times. Empirical mode decomposition is therefore different from simple band-pass filtering.

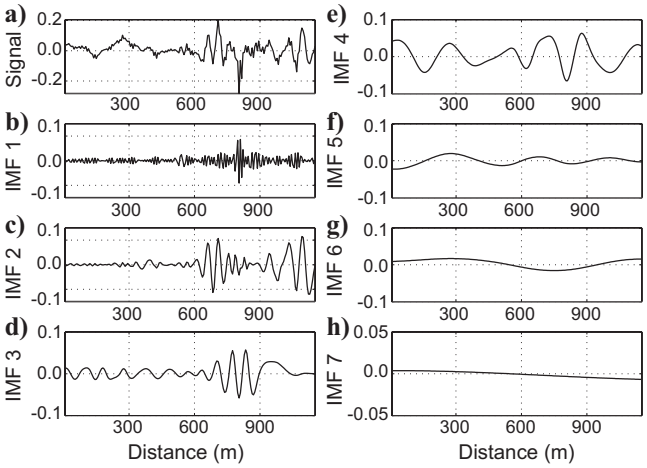

Figure 1. Empirical mode decomposition of a signal into IMFs. (a) The signal exhibits clear nonstationary behavior. (b-h) The decomposition performed by EMD yields 7 IMFs here. The number of oscillations decreases with increasing IMF number. (h) IMF 7 represents the trend in the data. Empirical mode decomposition is different from simple band-pass filtering in that different IMFs can have partially overlapping frequency content (e.g., IMFs 2 and 3). Note the vertical scales are not all the same.

f-x domain EMD

How can EMD be used to remove seismic noise? For many data sets, the random noise and any steeply dipping coherent noise make a significantly larger contribution to the high-wavenumber energy in the *f-x* domain than any desired signal. The IMF 1 represents the fastest oscillations in the data, i.e., it contains the largest wavenumber components in a constant-frequency slice in the *f-x* domain. Therefore, signal-to-noise enhancement can be achieved by subtracting IMF 1 from the data.

To process a whole seismic section, *f-x* EMD filtering is implemented in a similar way to *f-x* deconvolution using the following scheme:

1) Select a time window and transform the data to the *f-x* domain.
2) For every frequency,
   a) separate real and imaginary parts in the spatial sequence
   b) compute IMF 1 for the real signal and subtract to obtain the filtered real signal
   c) repeat for the imaginary part
   d) combine to create the filtered complex signal
3) Transform data back to the *t-x* domain.
4) Repeat for the next time window.

Unlike *f-x* deconvolution, which uses a fixed filter length for all frequencies, EMD adaptively matches its decomposition to the smoothness of the data. This offers the opportunity to implement a different filtering scheme for each frequency. It is worth emphasizing that removing only IMF 1 at each frequency is a single possibility among many. This scheme is the simplest one and has led to good performance on nearly all data sets we have tested to date.

Figure 2 illustrates why such a scheme can be a viable alternative to *f-x* deconvolution. It shows an example of applying EMD to a real signal. The data in Figure 1a represent the real part of a spatial se-

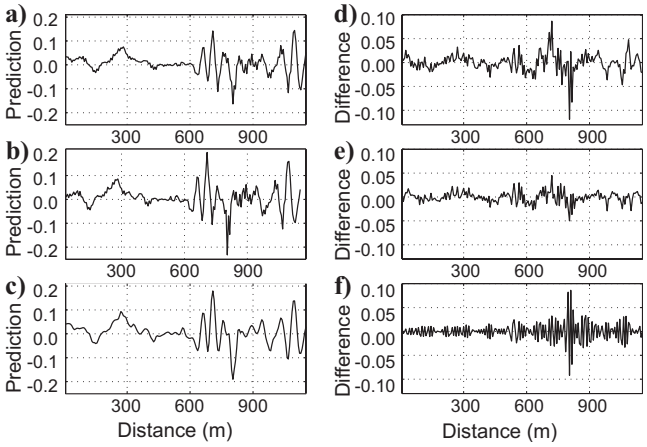

Figure 2. Comparison of EMD and AR filtering on the input signal shown in Figure 1a. (a) AR order = 2 and (d) its residual; (b) AR order = 4 and (e) its residual; (c) EMD filtering by removing the first IMF and (f) its residual, i.e., the first IMF (Figure 1b). A low AR order reproduces most of the data trend yet some slow spatial variations remain visible in the residual plot. A higher AR order captures more signal variations. Empirical mode decomposition filtering recovers all low wavenumber variations and produces smooth output. Note different vertical scales in the left and right columns.

quence in the f-x domain at $f = 42$ Hz. We consider linear prediction filtering using AR models of order 2 and 4, respectively, and compare this to the removal of the first IMF. Results of the different filtering methods are shown in Figure 2.

A low AR order produces a smooth filter output and only recovers the dominant spatial variations (Figure 2a). Some coherent signals including low wavenumber components remain visible in the difference plot (i.e., data minus filtered signal) as shown in Figure 2d. Increasing the filter order copes well with the rapid variations in the data (Figure 2b) and produces a better prediction of the low wavenumber trend and smoother residuals (Figure 2e). Increasing the filter order further would predict perfectly all variations in the data and very little filtering would be achieved.

The signal obtained by removing the first IMF from the original signal is displayed in Figure 2c. It captures all the low-wavenumber variations in the original signal. This signal could be obtained by summing IMFs 2 to 7 in Figure 1. For completeness, IMF 1 is again shown in Figure 2f. A comparison of the residuals in Figure 2d-f shows the f-x EMD filter does not contain any low-wavenumber information, and its residual is very localized without oversmoothing the output. Similar results are obtained for the imaginary part.

REAL DATA APPLICATIONS

We compare the performance of f-x deconvolution and f-x EMD for signal-to-noise enhancement on four real data sets resulting from all stages of the processing sequence. We also show the outcome of two noise attenuation methods based on local singular value decomposition and local median filtering (Bekara and van der Baan, 2007). All f-x domain-filtering techniques use a short-time Fourier transform with a sliding temporal window of length 512 ms and an overlap of 50% to remove edge effects. Frequencies beyond 60% of the Nyquist frequency are not processed and are damped to zero.

Data set 1: Shot gather

A shot gather is displayed in Figure 3a. The data contain interesting features such as shallow backscattered energy (A), a linear right-dipping event (B), ground roll (C), nearly flat reflections (D), and linear left-dipping events (E). The primary objective in processing this gather is to preserve the target reflections (D) and to attenuate all other coherent and random events. Frequency-offset deconvolution is implemented using an AR filter of length 4 designed using 20 spatial samples to estimate the filter coefficients.

Frequency-offset deconvolution emphasizes the signal-to-noise ratio of all coherent events including the unwanted ones, such as the left-dipping events (E) and the ground roll (C) (Figure 3b). The difference section for f-x deconvolution (Figure 3c) shows that the backscattered energy (A) and the right-dipping event (B) are removed partially yet other events are emphasized (e.g., the ground roll). Frequency-offset deconvolution is capable of interpolating energy. This can be an advantage or a disadvantage depending on the event considered. It is a clear advantage if we consider the target events (D) but a disadvantage if we consider the left-dipping events (E).

Frequency-offset EMD emphasizes the signal-to-noise ratio of the target reflections (D) and filters out the backscattered energy (A) and the ground roll (C) very effectively (Figure 3d). It also removes the right-dipping events (B) (Figure 3e). Frequency-offset EMD has less interpolation power compared to f-x deconvolution. The EMD is not a recursive spatial filtering method so no signal energy is passed to the next sample.

Inspection of the difference sections (new minus old) shows that f-x EMD performs much better than f-x deconvolution on this shot gather. Frequency-offset deconvolution enhances the signal-to-noise ratio of any coherent energy and is therefore less appropriate for this data set. Changing its parameter settings does not lead to significantly better results in this case.

We analyze the f-k spectra of the original and filtered data as shown in Figure 4 to understand the filtering behavior of f-x EMD better. The wavenumber axis is normalized by the Nyquist value. Standard f-k transforms implicitly assume the data are sampled regularly both in time and space. This is not the case here (Figure 5). The trace spacing of this shot gather is highly irregular and alternates between 5 and 7 m, leading to several aliasing artifacts visible in Figure 4a. For instance, the upside-down half-cones centered at normalized wavenumber of ± 1 are artifacts caused by the irregularity in the spatial sampling. They disappear if only regularly spaced traces are extracted. Despite these artifacts, the f-k spectra reveal many interesting features of f-x EMD versus f-x deconvolution and give a physical interpretation of its filtering behavior.

The ground roll (B) is aliased and mirrored spatially in (B1). The refraction (C) dominates the signal energy while the background noise (D) is spread out over the high-frequency area of the spectrum. The target reflections (E) are located around the zero wavenumber (they are predominantly horizontal in Figure 3).

Frequency-offset deconvolution removes some of the background noise (D) thereby enhancing the reflections (E) but leaving the ground roll (B, B1) unaffected (Figure 4b). Frequency-offset EMD, on the other hand, enhances the reflections (E) compared to the background noise while largely attenuating the ground roll (B, B1), its aliased energy (B1), and the high-frequency components (typically above 60 Hz) of the refractions (C).

Frequency-offset EMD acts as an adaptive high-cut wavenumber filter in the f-k domain (Figure 4c). At the lower-frequency end, the ground roll has been removed. At the middle- to high-frequency spectrum, all energy outside the normalized wavenumber $[-1/3, 1/3]$ has been dampened leading to the automatic suppression of background noise (D) and much of the aliased energy. The algorithm determines from the data what wavenumbers are to be suppressed as a function of frequency.

For completeness, we compare the performance of the f-x techniques with two other methods for signal-to-noise enhancement, namely local singular-value decomposition (SVD) and local median filtering (Bekara and van der Baan, 2007). Local methods implement dip steering over a sliding window of length 32 time samples, width 10 traces, and overlap of 50%. In local SVD, only one eigenimage is retained in the enhanced local window. The median filter is set up in two sequential steps: A median filter of length 3 and then 5 is applied consecutively. The outcome is displayed in Figure 6.

Both local methods show a greater ability to suppress background noise (i.e., noncoherent energy) as compared with the f-x techniques. This is particularly true for local SVD (Figure 6a). They are also able to partially suppress the backscattered energy (A), the right-dipping event (D), and the ground roll (C). This is because of the dip steering implemented in local methods which allows for dip selection. By restraining the maximum dip to be aligned, steeper events such as the ground roll or the backscattered events can be eliminated.

Data set 2: Common midpoint gathers

Next, we consider a second data set displayed in Figure 7. It consists of two moveout-corrected common midpoint (CMP) gathers that contain a mixture of shallow horizontal events (before 2.5 s) and deeper hyperbolic events (strongest one at approximately 3.8 s) superposed with different quasilinear events, which are likely multiples. Frequency-offset deconvolution is implemented using an AR filter of length 3 designed over 20 spatial samples.

Frequency-offset deconvolution (Figure 7b) boosts all the reflections including the multiples yet suppresses some of the background noise as shown in the difference section (Figure 7c). It also distorts reflector amplitudes and removes some useful energy in the shallow reflectors.

Frequency-offset EMD enhances the target reflectors, compared to the background noise, and removes some of the multiples (Figure 7d). There is no noticeable amplitude distortion (Figure 7e). The multiples are characterized by relatively high wavenumbers compared to the other events and therefore have been removed.

Data set 3: Stacked section

Next, we consider a stacked section containing shallow horizontal and marginally dipping reflectors in the middle (Figure 8a). Some crossing artifacts are present in the bottom of the section, probably resulting from previous data processing. We apply *f-x* EMD and *f-x* deconvolution with the same parameter values as for the previous example.

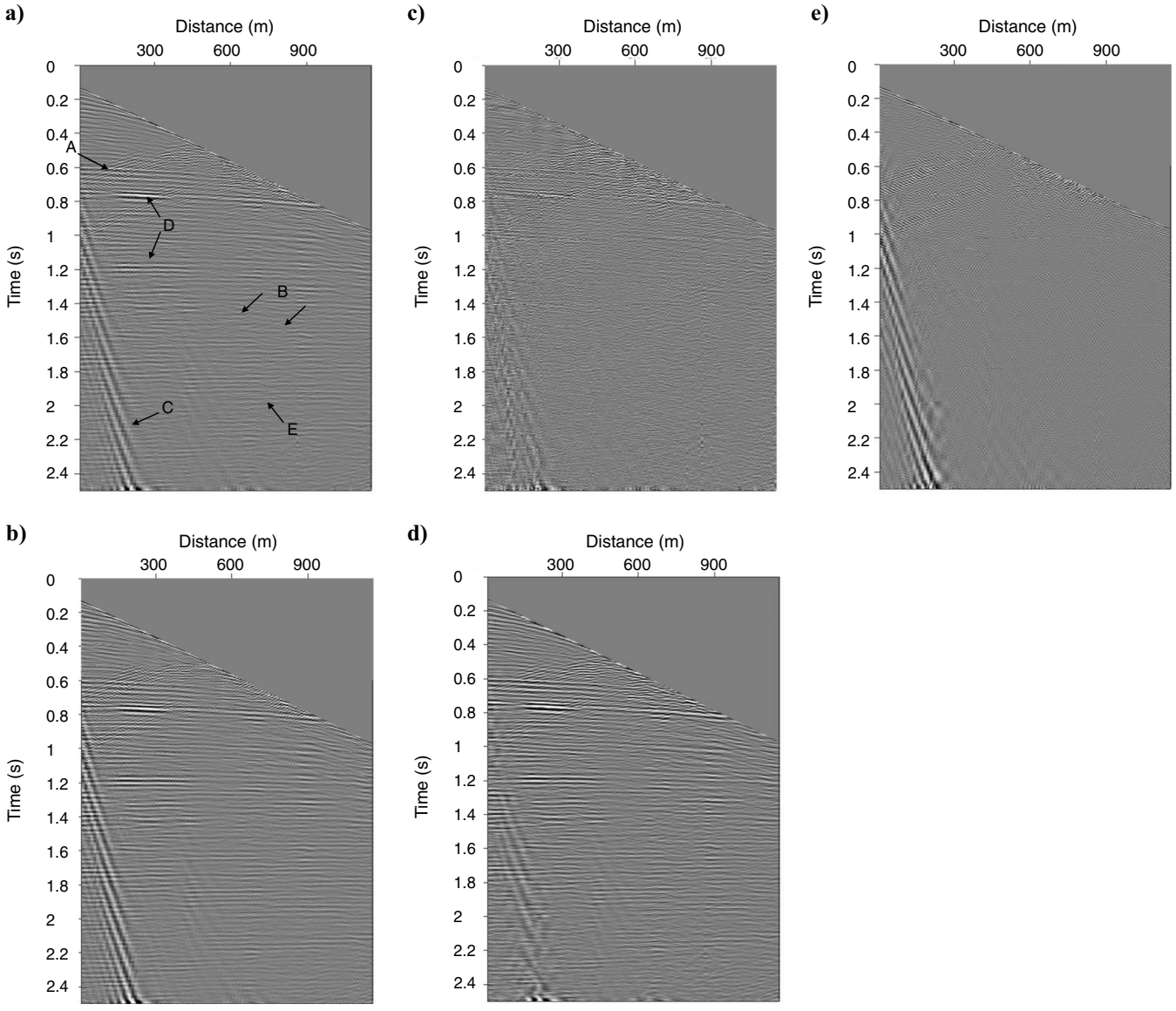

Figure 3. Data set 1: Shot gather. (a) Original data, (b) result of *f-x* deconvolution, (c) *f-x* deconvolution difference section, (d) result of *f-x* EMD, and (e) associated difference section. The gather contains a variety of signals such as (A) shallow backscattered waves, (B) a right-dipping event, (C) ground roll, (D) reflections, and (E) linear left-dipping events. Frequency-offset deconvolution reduces the background noise but enhances all coherent energy, whereas *f-x* EMD removes most background noise, backscattered energy, most remnant ground roll, and the linear dipping events. Data courtesy of BP.

Frequency-offset deconvolution attenuates some of the background noise but leaves the crossing artifacts untouched (Figure 8b and c). It also causes amplitude distortion by partially removing useful reflector energy (particularly at 600 ms) as shown in the difference section (Figure 8c). Frequency-offset EMD (Figure 8d) also attenuates some background noise but very little amplitude distortion occurs because no obvious reflector energy is visible in the difference section (Figure 8e). More importantly, *f-x* EMD is able to remove the crossing artifacts leading to a superior result over *f-x* deconvolution.

The effects of applying local SVD and median filtering on this data set are shown in Figure 7 of Bekara and van der Baan (2007). Both local techniques remove more of the random background noise in the filtered output, yet they leave more crisscrossing events than *f-x* EMD. For this particular data set, the local techniques are therefore better suited than *f-x* deconvolution because both random and coherent noise is present. We ultimately prefer the result of *f-x* EMD because it provides a better compromise between the reduction of background noise and the removal of the crisscrossing artifacts.

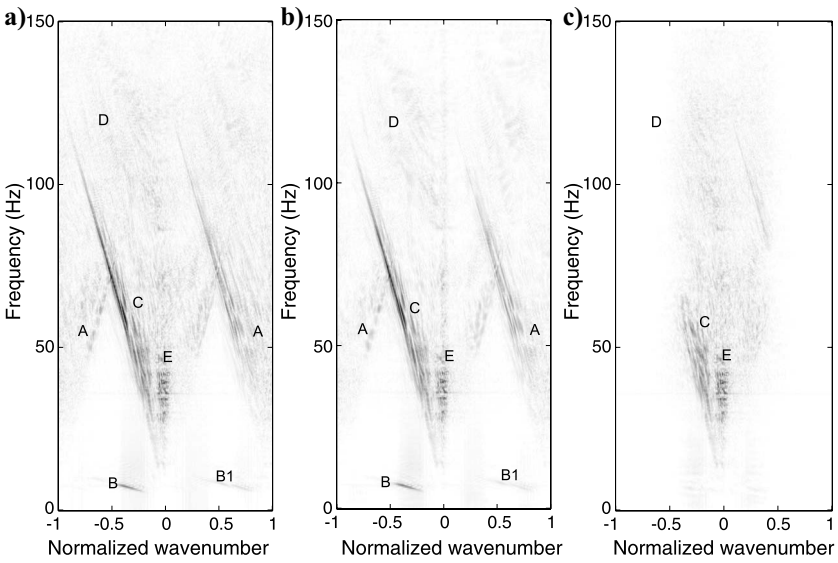

Figure 4. *F-k* spectra related to data sections in Figure 3. (a) Data and results of (b) *f-x* deconvolution, and (c) *f-x* EMD filtering. Frequency-offset deconvolution (D) removes some of the background noise, (E) enhances the reflections, and (B) slightly attenuates the ground roll. Frequency-offset EMD (B, B1) largely attenuates the ground roll and (C) the high-frequency components of the refractions (typically above 60 Hz) while (E) enhancing the reflections. It also removes all the energy outside the normalized wavenumber range $(-1/2, 1/2)$.

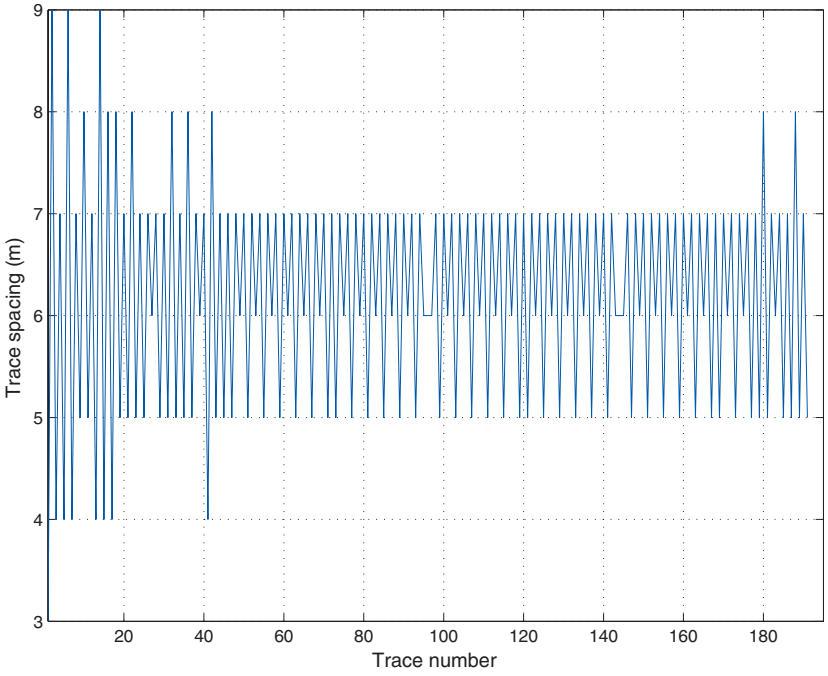

Figure 5. Trace spacing in data set 1. The data set has an irregular trace spacing that causes the artifacts A in the computation of the *f-k* spectra (Figure 4).

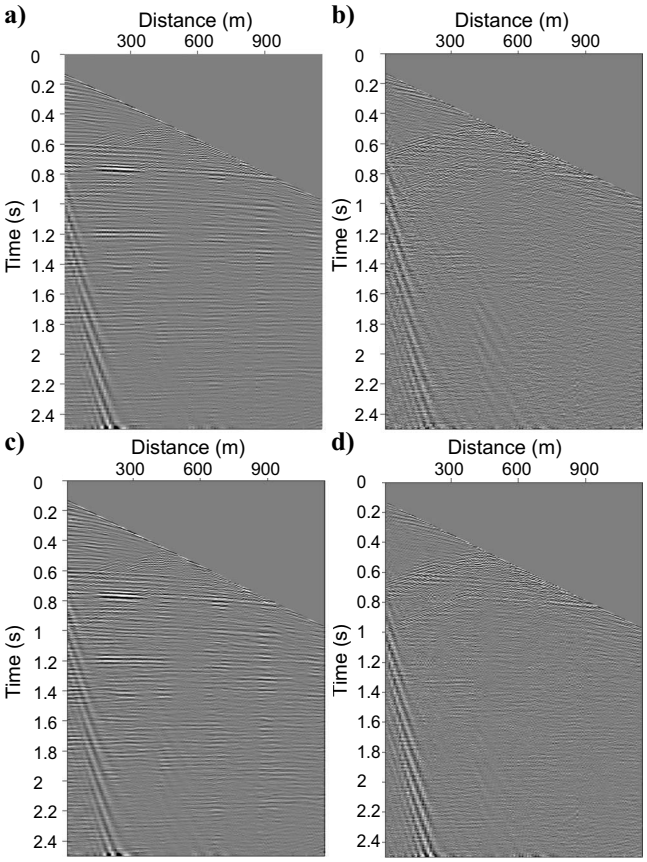

Figure 6. Alternative filtering results for data set 1 shown in Figure 3 using local SVD and local median filtering. (a) Local SVD and (b) its difference section; (c) local median filtering and (d) its difference section. Local SVD and local median filtering have the advantage over *f-x* techniques in that they allow for dip selection because they emphasize only those events that are aligned laterally.

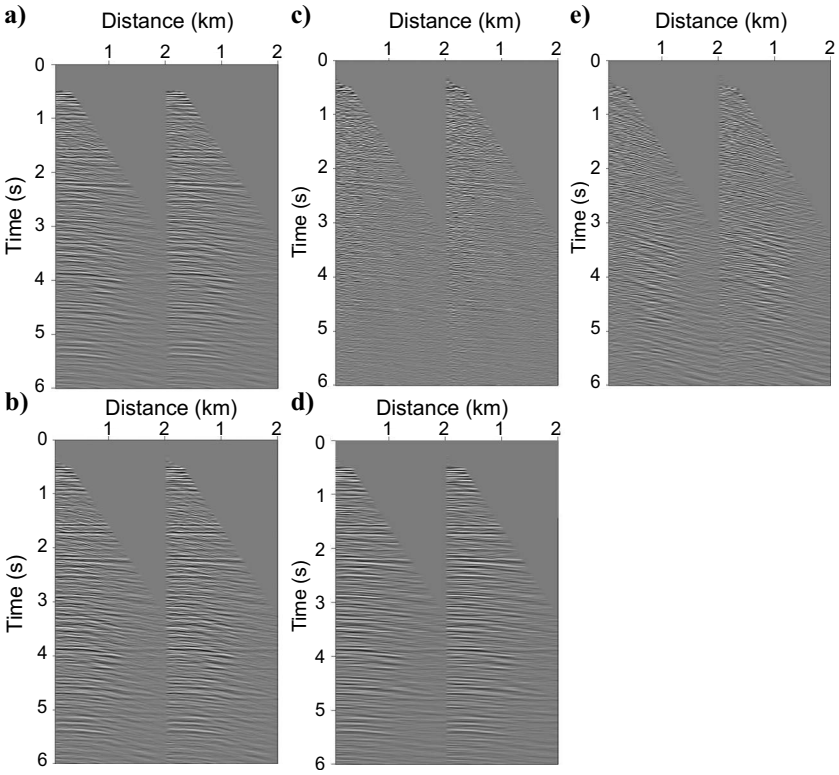

Figure 7. Data set 2: Two moveout-corrected CMP gathers. (a) Original data, (b) result of *f-x* deconvolution, (c) *f-x* deconvolution difference section, (d) result of *f-x* EMD, and (e) associated difference section. The CMP gathers contain a mixture of shallow horizontal events and deeper hyperbolic ones, superposed with different quasilinear events that are probably multiples. Frequency-offset deconvolution emphasizes all the reflections including multiples. It suppresses not only some background noise but also some useful energy related to the shallow reflectors. Frequency-offset EMD enhances the target reflectors and removes some of the multiples in addition to the background noise.

Data set 4: Migrated section

Finally, we consider a migrated section as shown in Figure 9a. It displays shallowly dipping reflections over an undulating interface. Again, we apply *f-x* EMD and *f-x* deconvolution with the same parameter values as the last example. The results are displayed in Figure 9b-e.

Frequency-offset EMD and *f-x* deconvolution have different strengths and weaknesses in this example. Frequency-offset EMD seems to remove more background noise and less of the shallowly dipping events between 0.6 and 0.8 s than *f-x* deconvolution. On the other hand, *f-x* EMD removes the dipping part of the undulating interface. Frequency-offset deconvolution, however, does not handle the amplitude fluctuations along the horizontal reflectors as well. This last shortcoming of *f-x* deconvolution is also noted in Bekara and van der Baan (2007).

DISCUSSION

Applications of EMD in geophysics are few, even though it offers many promising features for analyzing and processing geophysical data. It has been used for seismic attribute analysis (Magrin-Chagnolleau and Baraniuk, 1999), for the analysis of gravity data (Hassan, 2005), and recently as a tool to remove cable strum noise (Battista et al., 2007). We demonstrate its capabilities as an alternative technique to *f-x* deconvolution for random and coherent noise attenuation.

Frequency-offset EMD acts as an adaptive soft high-wavenumber cut filter. The cut-off wavenumbers are determined automatically from the data and vary as a function of frequency. For example, at lower frequencies most ground roll in Figure 3 is removed whereas at higher frequencies only refractions and background noise are eliminated (see also Figure 4). This is an advantage over *f-x* deconvolution, which uses in practice a single filter length for all frequencies. This is not necessarily optimal because spatial signals at low frequencies are less complex and easier to predict than their high-frequency counterparts.

Another advantage of *f-x* EMD over *f-x* deconvolution is trace spacing need not be perfectly regular because no convolutional operators are used, as is the case for standard linear prediction filtering. It shares this advantage with local SVD and local median filtering (Bekara and van der Baan, 2007).

Frequency-offset EMD compares favorably with *f-x* deconvolution, local singular-value decomposition (SVD), and local median filtering as shown in Figures 3 and 6–9, and Figure 7 in Bekara and van der Baan (2007). Unlike the other approaches, *f-x* deconvolution can remove only random noise; however, it is very useful for event interpolation (i.e., reconstruction of missing energy). Local SVD and local median filtering use dip steering, thus providing the opportunity to attack dipping coherent noise just as *f-x* EMD does.

Frequency-offset EMD produces less amplitude distortion and removes more background noise compared to *f-x* deconvolution as shown in Figures 8 and 9. Unfortunately, not all steeply dipping energy is unwanted and removal of the first IMF in the *f-x* domain could eliminate desired reflections, such as the flanks in Figure 9.

Thus, *f-x* EMD has a tendency to remove the largest wavenumbers. Often these are associated with random or steeply dipping coherent noise. Sometimes, however, random noise levels are very low and *f-x* EMD might remove steeply dipping reflections (e.g., Figure 9). Empirical mode decomposition acts as a dyadic filter bank. For white signals with Gaussian density distribution, the first IMF is the output of a high-pass filter whose cut-off wavenumber is about half the Nyquist value (Flandrin et al., 2005). This explains why *f-x* EMD in the current implementation removes all energy outside the normalized wavenumber range $(-1/2,1/2)$, especially at high frequencies. In most data sets tested to date, we only occasionally found this to be an issue. We recommend routine inspection of difference sections to determine if any useful signal had been removed. We anticipate this last weakness could be overcome by the proper selection

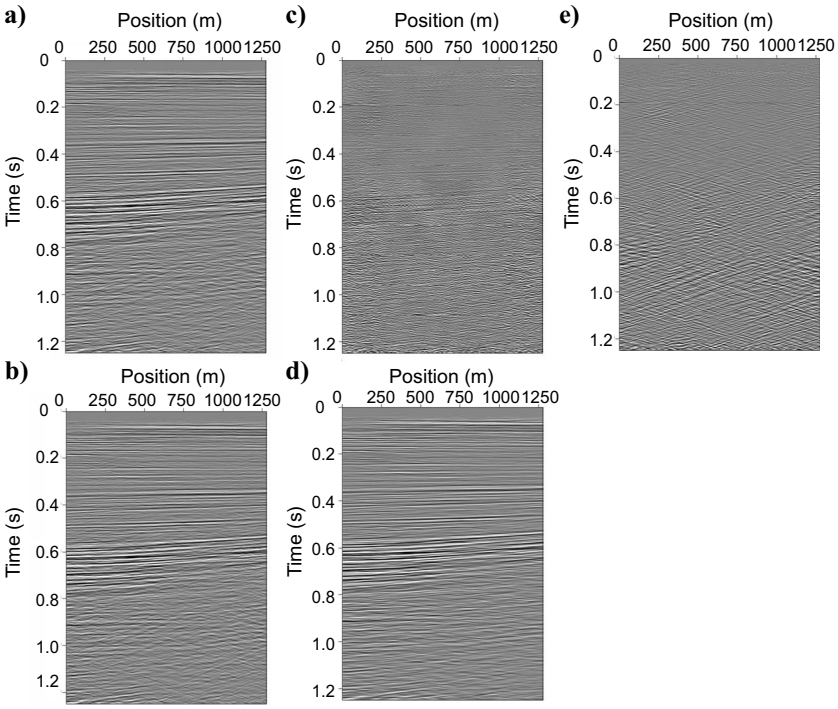

Figure 8. Data set 3: A stacked section. (a) Original data, (b) result of *f-x* deconvolution, (c) *f-x* deconvolution difference section, (d) result of *f-x* EMD, and (e) associated difference section. Both *f-x* deconvolution and *f-x* EMD reduce the background noise but only *f-x* EMD removes the crisscrossing artifacts. Frequency-offset deconvolution also removes some reflection energy. Data courtesy of Shell.

of the IMFs to be removed or by injecting small amounts of noise into the data.

The *f-x* EMD results can be recreated by judicious tapering in the *f-k* domain. The chosen *f-x* EMD approach is, however, automatic, and several weak noise types in the displayed data sets were detected only after the inspection of the obtained difference sections. Frequency-offset EMD thus can be used not only for noise reduction but also for data analysis.

The performance of *f-x* EMD is sensitive to highly irregular acquisition geometries with variable trace spacing and possibly large gaps. A detailed assessment is beyond the scope of this paper. The EMD is based on spline interpolation, thus greatly reducing the need for regular sample intervals. The EMD performance is, however, also governed by the accuracy of extrema detection. This requires that all relevant wavelengths are sampled preferably by many points. Correct identification of the extrema of the largest wavenumbers might thus require a fair amount of oversampling. See Rilling and Flandrin (2009) for further details.

Eliminating only the first IMF in *f-x* EMD results in a fast processing algorithm with comparable computation cost to *f-x* deconvolution. It also has the advantage that no parameter is given by the processing analyst. This choice is known to be too simplistic for some data sets (e.g., Figure 9). On the other hand, it has led to satisfactory results in nearly all data sets considered to date. Other IMFs could better capture some coherent or noncoherent noise signals in specific situations. Case-dependent selection of useful IMFs will result in a more flexible processing scheme, which would open up new avenues for both data analysis and processing.

Finally, a 2D EMD implementation exists (Linderhed, 2002). The described method can therefore be extended to attenuate random and coherent noise in 3D data by means of *f-x-y* EMD.

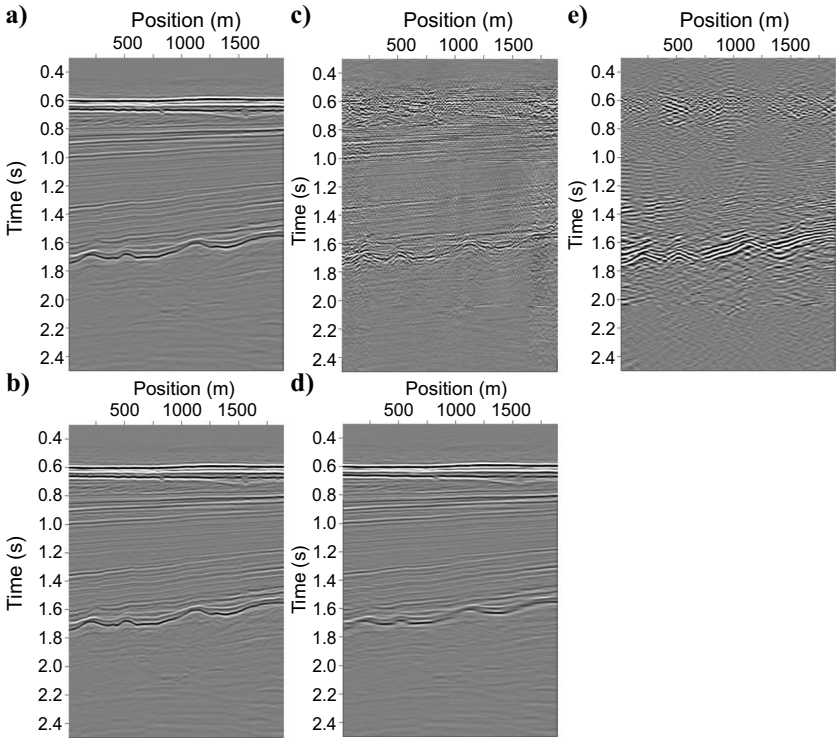

Figure 9. Data set 4: A migrated section. (a) Original data, (b) result of *f-x* deconvolution, (c) *f-x* deconvolution difference section, (d) result of *f-x* EMD, and (e) associated difference section. The data consist of shallow reflections over an undulating interface. Frequency-offset EMD removes more background noise and does not harm the shallowly dipping events between 150 and 200 ms. However, it also removes the dipping parts of the undulating interface. Frequency-offset deconvolution shows the opposite behavior: It removes less background noise and leaves the flanks intact but does not handle the amplitude fluctuations along the horizontal reflectors as well. Data courtesy of BP.

CONCLUSION

Frequency-offset EMD corresponds to an auto-adaptive wavenumber filter that determines which wavenumbers are to be removed from the data for each individual frequency to attenuate both random and steeply dipping coherent noise.

Contrary to many alternative noise-reduction tools such as *f-x* deconvolution, *f-x* EMD invokes no piecewise-stationarity assumption due to its origin as a time-frequency analysis tool. It can thus overcome the potentially low performance of *f-x* deconvolution that arises with processing structurally complex data or data contaminated by coherent noise. It is also less sensitive to irregular spatial sampling.

An interesting aspect of the EMD is that it is parameter-free in its simplest implementation in which only the first intrinsic mode function is removed. Other schemes are possible, adding flexibility at the expense of requiring interaction.

ACKNOWLEDGMENTS

We thank the BG group, BP, Chevron, the Department of Trade and Industry, and Shell for financial support of the project Blind Identification of Seismic Signals, and BP, an anonymous company, and Shell EP Europe for permission to use the data. The constructive and very detailed reviews of Mike Galbraith, Eric Verschuur, and three anonymous reviewers are much appreciated.

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.

Bekara, M., and M. van der Baan, 2007, Local singular value decomposition for signal enhancement of seismic data: Geophysics, **72**, no. 2, V59–V65.

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

Flandrin, P., G. Rilling, and P. Goncalves, 2005, Empirical mode decomposition as a filter bank: IEEE Signal Processing Letters, **11**, 112–114.

Galbraith, M., 1991, Random noise attenuation by *f-x* prediction: A tutorial: 61st Annual International Meeting, SEG, Expanded Abstracts, 1428–1431.

Harris, P. E., and R. E. White, 1997, Improving the performance of *f-x* prediction at low signal-to-noise ratios: Geophysical Prospecting, **45**, 269–302.

Hassan, H., 2005, Empirical mode decomposition (EMD) of potential field data: Airborne gravity data as an example: 75th Annual International Meeting, SEG, Expanded Abstracts, 704–707.

Huang, N. E., Z. Shen, S. R. Long, M. L. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, 1998, The empirical mode decomposition and Hil-

bert spectrum for nonlinear and nonstationary time series analysis: Proceedings of the Royal Society of London Series A — Mathematical Physical and Engineering Sciences, **454**, 903–995.
Linderhed, A., 2002, 2D empirical mode decompositions in the spirit of image compression: Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 1–8.
Magrin-Chagnolleau, I., and R. G. Baraniuk, 1999, Empirical mode decomposition based time-frequency attributes: 69th Annual International Meeting, SEG, Expanded Abstracts, 1949–1953.
Rilling, G., and P. Flandrin, 2009, Sampling effects on the empirical mode decomposition: Advanced adaptive data analysis, **1**, 43–59.
Tufts, D. W., and R. Kumaresan, 1980, Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood: Proceedings of the IEEE, **70**, 975–989.