APPLICATION OF THE EMPIRICAL MODE DECOMPOSITION METHOD TO GROUND-ROLL NOISE ATTENUATION IN SEISMIC DATA

Luiz Eduardo Soares Ferreira 1, Milton José Porsani 1, Michelângelo G. da Silva 1 and Giovani Lopes Vasconcelos 2

ABSTRACT. Seismic processing aims to provide an adequate image of the subsurface geology. During seismic processing, the filtering of signals considered noise is of utmost importance. Among these signals is the surface rolling noise, better known as ground-roll. Ground-roll occurs mainly in land seismic data, masking reflections, and this roll has the following main features: high amplitude, low frequency and low speed. The attenuation of this noise is generally performed through so-called conventional methods using 1-D or 2-D frequency filters in the $f-k$ domain. This study uses the empirical mode decomposition (EMD) method for ground-roll attenuation. The EMD method was implemented in the programming language FORTRAN 90 and applied in the time and frequency domains. The application of this method to the processing of land seismic line 204-RL-247 in Tacutu Basin resulted in stacked seismic sections that were of similar or sometimes better quality compared with those obtained using the $f-k$ and high-pass filtering methods.

Keywords: seismic processing, empirical mode decomposition, seismic data filtering, ground-roll.

RESUMO. O processamento sísmico tem como principal objetivo fornecer uma imagem adequada da geologia da subsuperfície. Nas etapas do processamento sísmico a filtragem de sinais considerados como ruídos é de fundamental importância. Dentre esses ruídos encontramos o ruído de rolamento superficial, mais conhecido como ground-roll. O ground-roll ocorre principalmente em dados sísmicos terrestres, mascarando as reflexões e possui como principais características: alta amplitude, baixa frequência e baixa velocidade. A atenuação desse ruído é geralmente realizada através de métodos de filtragem ditos convencionais, que utilizam filtros de frequência 1D ou filtro 2D no domínio $f-k$. Este trabalho utiliza o método de Decomposição em Modos Empíricos (EMD) para a atenuação do ground-roll. O método EMD foi implementado em linguagem de programação FORTRAN 90, e foi aplicado no domínio do tempo e da frequência. Sua aplicação no processamento da linha sísmica terrestre 204-RL-247 da Bacia do Tacutu gerou como resultados, seções sísmicas empilhadas de qualidade semelhante e por vezes melhor, quando comparadas as obtidas com os métodos de filtragem $f-k$ e pasa-alta.

Palavras-chave: processamento sísmico, decomposição em modos empíricos, filtragem dados sísmicos, atenuação do ground-roll.

1Centro de Pesquisa em Geofísica e Geologia, Instituto de Geociências, Universidade Federal da Bahia. Campus Universitário da Ferecção Salvador, Bahia, Brasil. Phone: +55(71) 3283-8500 – E-mails: luizesferreira@gmail.com; porsani@cpgg.ufba.br; mgs@cpgg.ufba.br

2Departamento de Física, Universidade Federal de Pernambuco. Av. Prof. Moraes Rego, 1235 – Cidade Universitária, 50670-901 Recife, PE. Phone: +55(81) 2126-8000 – E-mail: giovani@df.ufpe.br
INTRODUCTION

Seismic processing has as its main goal the acquisition of an image of the subsurface geology. This acquisition is quite sensitive to the presence of noise in the seismic records. Surface rolling noise, also known as ground-roll, is among the different types of noise.

Ground-roll is a particular type of Rayleigh wave with high amplitude, low frequency, low speed and a dispersive nature, and this noise usually masks the reflections of interest. Its dispersive nature interferes with the shallow reflections in the short offsets (Henley, 2003). Many authors have shown that ground-roll can be attenuated during data acquisition through the appropriate arrangement of geophones (Harlan et al., 1984; Anstey, 1986; Shieh & Herrman, 1990; Pritchett, 1991). This strategy cannot always be applied and is useless when the data have already been acquired. There are several available filtering methods in commercial seismic processing software that can be used for ground-roll attenuation. The most commonly used filters employ the 1-D Fourier transform (high-pass filter) and the 2-D Fourier transform (f-k filter) (Yilmaz, 2001). Because the seismic trace is a nonstationary signal, ground-roll attenuation using the Fourier transform is somewhat inefficient. This inefficiency arises because this transform uses sines and cosines in a base function that, among other shortcomings, is a stationary function in the sense that the function has the same frequency components regardless of the time instant considered.

The empirical mode decomposition (EMD) method, developed by Huang et al. (1998), decomposes the signal independently of the phase characteristics or statistical properties that vary along the trace. The method's principle is to decompose a signal into a number of functions that have the same number of zero-crossings and extrema and are symmetric with respect to the local mean. These functions are called "intrinsic mode functions" (IMFs). The IMFs are extracted directly from the signals and are equivalent to sinusoidal shapes locally modified by modulation in frequency and amplitude. This method is a promising technique in ground-roll attenuation (Bekara & Baa, 2008).

The method was applied on a land seismic line, 204-RL-247, of the Tacutu basin. The results for the stacked sections after application of the EMD method were shown in the time and frequency domains and compared with those obtained with the f-k and high-pass frequency filters.

EMPIRICAL MODE DECOMPOSITION METHOD

The EMD method was first presented by Huang et al. (1998) and is based on the simple hypothesis that any data consist of different and simple intrinsic oscillatory modes. Each intrinsic mode, linear or nonlinear, represents a simple oscillation with the same number of zero-crossings and extrema. Furthermore, the oscillation is also symmetric in relation to the "local mean". At any moment, the data can have different coexisting oscillatory modes, one superimposed on the other. Each one of these oscillatory modes is represented by an IMF, with the following definition:

(i) the number of zero-crossings and extrema are equal or differ by no more than one, and
(ii) the modes are symmetric regarding the mean, that is, for any point, the sum of the maximum and minimum envelope values is equal to zero.

These two conditions ensure that the functions are symmetric regarding the local mean and that the instantaneous frequency does not have unwanted fluctuations.

To evaluate the most effective use of the EMD method, we tested its application in the domain of time, frequency and bandwidth-limited frequency.

DME in the time domain

The EMD method in the time domain acts in a direct way on the seismic traces. The seismic trace $x(t)$ (Fig. 1) is decomposed into IMFs, which are obtained through the following steps:

(i) All local extrema (maximum or minimum) of the data are identified (Fig. 1);
(ii) After determining the local maximum and minimum, the interpolation of points is conducted, thus creating an envelope associated with the maxima and another one associated with the minima (Fig. 1). This step uses the 1-D interpolation subroutine (Porsani, 2009), developed based on the inverse of the distances (Shepard, 1968);
(iii) For each point, the mean $(m_1(t))$ of the maximum and minimum envelopes is calculated (Fig. 1), and subtracted from the input data $x(t)$, as indicated in Equation 1. The result of the difference is the first component $h_1(t)$, an IMF candidate (Eq. 2).

$$m_1(t) = \frac{e_{\text{sup}}(t) + e_{\text{inf}}(t)}{2}$$  \hspace{0.5cm} (1)

$$h_1(t) = x(t) - m_1(t)$$  \hspace{0.5cm} (2)

To be considered an IMF, $h_1(t)$ must satisfy conditions (i) and (ii) of the three immediately aforementioned conditions. However, if these conditions are not satisfied, $h_1(t)$ is treated as a new input data point, and the separation
process is repeated \( k \) times as follows until an IMF is obtained:

\[
h_{1,j}(t) = h_{1,j-1}(t) - m_{1,j-1}(t), \quad j = 1, \ldots, k,
\]

where \( h_{1,0}(t) = h_1(t) \) and \( m_{1,j-1}(t) \) are the mean of the maximum and minimum envelopes of \( h_{1,j-1}(t) \). The function \( h_{1,k}(t) \) thus obtained satisfies both conditions (i) and (ii) mentioned above and corresponds to the first IMF of the data, denoted by \( c_1(t) \):

\[
c_1(t) = h_{1,k}(t).
\] (3)

Assuming that the first IMF, \( c_1(t) \), was found, it can be removed from the data, thus obtaining a residue that is treated as the new input data:

\[
r_1(t) = x(t) - c_1(t).
\] (4)

This process is repeated on all subsequent residues and stops when a predetermined number \( N \) of IMFs is reached or when \( r_N \), for some \( N \), becomes a function from which no IMFs can be extracted, that is, a function with a maximum number of extrema equal to three, and no other envelopes can be obtained:

\[
r_j(t) = r_{j-1}(t) - c_j(t), \quad j = 2, \ldots, N.
\] (5)

According to equations (5) and (6), the original data can be obtained as follows:

\[
x(t) = \sum_{j=1}^{N} c_j(t) + r_N(t),
\] (6)

where \( N \) is the number of IMFs into which the data have been decomposed and \( r_N(t) \) is the final residue (Fig. 2).

The IMFs extracted from the data represent different frequency modes, and in the time domain, the first IMF has the highest frequencies, whereas the last IMF has the lowest ones (Fig. 2). According to the desired frequency bandwidth, the corresponding IMFs can be properly subtracted from the data, as discussed later.

Two stopping criteria were considered for the determination of the IMFs.

So that the process of searching for an IMF is not overly time-consuming, it is required that the normalized squared difference between the two IMF candidates (the current one, \( h_i(t) \), and the previous one, \( h_{i-1}(t) \)) is less than a given tolerance (tol), as discussed by Huang et al. (1998):

\[
\sum_t \left( \frac{h_i(t) - h_{i-1}(t)}{h_{i-1}(t)} \right)^2 < \text{tol.}
\] (7)

In the determination of the number of IMFs into which a datum must be decomposed, Huang et al. (2003), after a few tests, established that the number of IMFs obtained from a datum is between 4 and 8. The number of IMFs chosen in the present study was 5, plus the residue. In the present case, although the “residue” may still contain some IMFs, it is not necessary to continue the procedure, and it is sufficient to consider the first five IMFs and the corresponding residue.

Figure 1 – Trace 56 of shot 64, shown in samples 200 to 450, with maximum and minimum envelopes and the mean envelope curve.
EMD in the frequency domain

For this domain, the EMD method is applied for each frequency of the set of traces from a common shot-point gather. To use the same algorithm that was developed to work with real data, we used a property of the Fourier transform (FT), for which the FT of a signal that is real and even is also real and even. Thus, each trace \( x(t) \) with \( N_s \) samples is “mirrored” to represent a real and even signal \( y(t) \) to obtain a panel of traces that has only real representation in the frequency domain. Moreover, the panel of traces in the \( tx \) (time-distance) domain is transformed to the \( fk \) (frequency – distance) domain by a discrete Fourier transform (DFT) of each trace (Fig. 3). Thus, we have the data represented in the \( fk \) domain, and the EMD method is applied to each frequency, i.e., for each frequency \( f \), the signal \( Y_f(x) \) is taken as the input data for the EMD, corresponding to the amplitudes in the frequency \( f \) for each one of the traces.

The four procedures for the method are summarized below.

1. Procedure for obtaining the real and even vector \( y_k(t) \).

(a) The Fast Fourier Transform subroutine, used to perform the DFT, assumes that the number of samples of the signal \( y_k(t) \) is

\[ nft = 2^n > 2N_s \] (8)

(b) Given that the trace

\[ x_{k}(t) = x_{k1}, x_{k2}, x_{k3}, \ldots, x_{kN_s} \],

we obtain the following vector:

\[ y_{k}(t) = x_{k1}, x_{k2}, x_{k3}, \ldots, x_{kN_s}, 0, 0, \ldots, 0, 0, x_{kN_s+1}, x_{kN_s+2}, \ldots, x_{kN_s}, x_{kN_s} \] (9)

It is important to highlight the following points:

- \( y_k(t) \) is obtained by mirroring the trace \( x_k(t) \) relative to the point \( nft/2 \), so that periodicity is ensured, the last sample of vector \( y_k(t) \) must be \( y_k(nft) = x_{k2} \), and
- the intermediate samples

\[ y_k(N_s+1), y_k(N_s+2), \ldots, y_k(nft-N_s+1) \]

should be set to zero, and this number of zeroed samples is always odd.

2. Procedure for obtaining the data in the frequency domain \( y_k(f) \).

The input data in the frequency domain \( y_k(f) \) for application of the EMD method is obtained by applying the DFT in vector \( y(t) \):

\[ y_k(f) = \mathcal{F} \{ y_k(t) \} \] (10)

Figure 2 – The five first IMFs plus residue of trace 56 and their respective frequency bandwidths.
3. Procedure for application of the EMD method to the data \( y_k(f) \). For each frequency \( f \) up to the Nyquist frequency, the following vector is formed:

\[
Y_f(x) = y_1(f), y_2(f), \ldots, y_{N_t}(f),
\]  

where \( N_t \) is the number of traces of a given common shot-point gather. The EMD method (described in the item above) is applied to the datum \( Y_f(x) \), decomposing it into the desired number of IMFs (Fig. 4). Observe that the IMFs extracted from the data represent different wavenumbers, and the first IMF has the elements present in the data with the highest wavenumbers, whereas the last IMF has the elements with the lowest wavenumbers (Fig. 4). Then, the desired IMFs are subtracted from the input data, finally locating the filtered data \( z_k(f) \) for \( k = 1, \ldots, N_t \).

4. Application of the inverse Fourier transform to datum \( z_k(f) \). This application thus obtains the filtered trace in the time domain \( z_k(t) \):

\[
z_k(t) = \mathcal{F}^{-1}\{z_k(f)\}
\]

**EMD in the bandwidth-limited frequency domain**

In the previous section, the implementation of the EMD method was discussed for the frequency domain, where the vector \( Y_f(x) \) for each frequency \( f \) up to the Nyquist frequency is decomposed into a certain number of IMFs plus a residue. In another variant of this procedure, we modified the method to operate in a frequency bandwidth between the initial frequency \( f_1 \) and the final frequency \( f_2 \), directing the operation of the filter within the frequency spectrum to the bandwidth of interest, where the noise to be attenuated is located.
APPLICATION OF THE EMPIRICAL MODE DECOMPOSITION METHOD TO GROUND-ROLL NOISE ATTENUATION IN SEISMIC DATA

RESULTS ON REAL DATA

To demonstrate the effectiveness of the methods, we used the seismic reflection line 204-RL-247 from the Tacutu basin (Table 1). All the seismic processing steps, except for the EMD filtering, were performed with FOCUS, a commercial software program (Paradigm Geophysical Corporation).

The \( f_k \), high-pass (low-cut) and EMD filtering methods were tested in the time domain, and the EMD was tested in the frequency domain and in the bandwidth-limited frequency domain.

**Table 1** — Parameters for the acquisition of seismic line 204-RL-247.

<table>
<thead>
<tr>
<th>LINHA 204 – RL – 247</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Basin</td>
<td>Tacutu</td>
</tr>
<tr>
<td>Location</td>
<td>Roraima, Brazil</td>
</tr>
<tr>
<td>Segment</td>
<td>3850 – 100 – 0 – 100 – 1050 m</td>
</tr>
<tr>
<td>Coverage</td>
<td>4800%</td>
</tr>
<tr>
<td>Number of channels</td>
<td>96</td>
</tr>
<tr>
<td>Interval between stations ( \Delta G )</td>
<td>50 m</td>
</tr>
<tr>
<td>Interval between shot points ( \Delta S )</td>
<td>50 m</td>
</tr>
<tr>
<td>Sampling time interval</td>
<td>4 ms</td>
</tr>
<tr>
<td>Recording time</td>
<td>4 s</td>
</tr>
<tr>
<td>Arrangement</td>
<td>L-10/50 m</td>
</tr>
</tbody>
</table>

\( f_k \) filtering

Filtering in the \( f_k \) domain basically consists of choosing the area in which to extract the amplitudes. To define this area, the preservation of the signal and the attenuation of linear events in the \( f_k \) domain are taken into account. For this purpose, a shot was used to define the filter to be applied, as discussed by Ferreira (2011). Figure 5b shows shot 64 after the application of the \( f_k \) filtering, resulting in good ground-roll attenuation compared with the original data shown in Figure 5a.

Figure 9b shows the stacked seismic section obtained after the \( f_k \) filtering. Compared with the seismic section obtained without any filtering (Fig. 9a), we note that the good ground-roll noise attenuation resulted in reflectors with great continuity and superior time and spatial resolution.

High-pass filter (low-cut)

The high-pass filter was used with cutoff and pass frequencies of \( f_1 = 10 \) Hz and \( f_2 = 15 \) Hz, respectively. Figure 5c shows shot 64 after the application of this high-pass filter, which resulted in good ground-roll attenuation, although there was also attenuation in the low frequencies that are not a part of the ground-roll.

The stacked seismic section obtained with this filtering method (Fig. 9c) showed an improvement over the raw section (Fig. 9a). The vertical and horizontal resolutions were significantly improved because of the ground-roll noise attenuation. However, in the stacked section, due to the attenuation of the signal with low frequencies, the reflectors contained in the stacked section obtained with the \( f_k \) filtering method were not observed (Fig. 9b).

EMD in the time domain

The separation of each shot into five IMFs plus residue was determined in the time domain through the EMD method (Fig. 6). Each IMF obtained in the time domain contains a frequency bandwidth of the signal, and the first IMF contains a bandwidth with the highest frequencies, whereas the residue has the lowest frequencies. Seeking to maximize the ground-roll noise attenuation, only the first and second IMF of each shot were preserved (Fig. 5d).

Figure 10a shows the stacked seismic section using the data filtered with the EMD method in the time domain. A better definition of the reflectors is observed compared with the original seismic section with no filtering (Fig. 9a). Moreover, the filtering of the section significantly reduces the presence of linearly sloping events derived from ground-roll noise.
Figure 5 – For shot 64, original datum in (a) and after the application of the $f_k$ filter in (b), the high-pass filter in (c), the EMD method in the time domain in (d), the EMD method in the frequency domain in (e) and the EMD method in the frequency domain for the 1 to 15 Hz bandwidth in (f).

EMD in the frequency domain

Applying the EMD method in the frequency domain, each shot of the line was decomposed into five IMFs plus residue (Fig. 7). The IMFs obtained in the frequency domain contain intervals according to the wavenumber of the datum, with the first IMF containing the elements with the highest wavenumbers (most of the groundroll noise and direct wave), whereas the residue has the elements.
with the lowest wavenumbers (Bekara & Baan, 2008). The first IMF was extracted from the remaining IMFs, and the filtered datum obtained was the sum of the four remaining IMFs with the residue (Fig. 5e), where good ground-roll noise attenuation is observed, making visible the reflections previously masked by the noise.

The stacked seismic section was generated from the filtered data with the EMD method in the frequency domain, extracting the first IMF (Fig. 10b). There was an improvement in the continuity of the reflectors and increased time resolution compared with the raw section (Fig. 9a). This result is similar to that obtained with the $f_k$ filtering method (Fig. 9b).

**EMD in the bandwidth-limited frequency domain**

The performance of this method was determined for the closed interval $f_1 = 1$ Hz and $f_2 = 1.5$ Hz, where each frequency was
decomposed into five IMFs plus residue, and all these components were extracted from the data, resulting in an almost complete removal of the ground-roll noise (Fig. 5f). This excellent result for ground-roll attenuation generated a stacked seismic section in which the linear events are almost eliminated (Fig. 10c). This result is similar to that obtained with the high-pass filter, except that the deepest reflectors were better preserved, as confirmed by the amplitude spectrum (Fig. 8).

Figure 8 – Amplitude spectrum of the original data in blue and after the application of the $f_k$ filter in red, the high-pass filter in black and the EMD method in the frequency domain for the 1 to 15 Hz bandwidth in green.

Figure 9 – Stacked seismic section obtained with no filtering method in (a), with the $f_k$ filtering method in (b) and with the high-pass filtering method in (c).

Figure 10 – Stacked seismic section obtained by preserving the first and second IMFs in the time domain in (a), extracting the first IMF in the frequency domain in (b) and extracting the IMFs and the residue obtained in the frequency bandwidth from 1 to 15 Hz in (c).

CONCLUSIONS

The EMD method developed by Huang (1998) was used for ground-roll noise attenuation and implemented in the following domains: time, frequency and bandwidth-limited frequency. The EMD method was applied to land seismic line 204-RL-247 of the Tacutu basin, and the results obtained were superior to those obtained with the high-pass and $f_k$ conventional filtering methods.

The application of the EMD method in the time domain, preserving the first and second IMFs, showed good results for
ground-roll noise attenuation, as confirmed in the stacked seismic section, in which there was a decrease in the linear events associated with ground-roll.

The application of the EMD method in the frequency domain showed excellent results in ground-roll attenuation in the shot domain and seismic section, with good time resolution and lateral continuity of the reflectors, a similar result to that obtained with the FK filter.

The EMD method applied in the frequency domain limited to the frequency bandwidth of the ground-roll was more effective than the other methods, generating a high-quality stacked seismic section, in which the reflectors show superior continuity. Although the result is similar to that obtained with the high-pass filter, better preservation of certain deep reflectors can be observed. The EMD method can be applied in the time or frequency domain, it has low computational cost, and its implementation is relatively simple. Thus, EMD is an especially promising method for seismic data processing.

ACKNOWLEDGMENTS

The authors thank the Brazilian energy company PETROBRAS, the Funding Authority for Studies and Projects (Financiadora de Estudos e Projetos – FINEP), the Bahia State Research Foundation (Fundação de Amparo à Pesquisa do Estado da Bahia – FAPESB), the National Agency of Oil, Natural Gas and Biofuel (Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP) and the National Council of Technological and Scientific Development, National Institute of Science and Technology of Oil Geophysics (Conselho Nacional de Desenvolvimento Científico e Tecnológico, Instituto Nacional de Ciência e Tecnologia de Geofísica do Petróleo – CNPq/INCT-GP) for financially supporting the research. The authors are also grateful to Paradigm and Landmark for the processing software licenses made available to the LAGEP-CPGG-UFBA.

REFERENCES

NOTES ABOUT THE AUTHORS

Luiz Eduardo Soares Ferreira holds B.S. (2009) and M.S. (2011) degrees in geophysics from UFBA.

Milton José Porsani received a B.S. (1976) in geology from USP and an M.S. (1981) and Ph.D (1986) in geophysics from UFBA. He has been a Researcher at CPGG-UFBA since 1986 and completed a one-year postdoctoral appointment in geophysics at the Institute of Geophysics, University of Texas, Austin, USA, in October 1993. He additionally served as Professor, Department of Geology and Applied Geophysics, Geosciences Institute (Instituto de Geociências – IGEO) at UFBA (1996), Full Professor for the subject Oil Exploration (2000), Researcher IA, CNPq, and Coordinator of the Oil Exploration Program, CPGG-UFBA. His areas of interest include the development of methods and algorithms for mono- and multichannel filtering, seismic data processing, inversion of geophysical data and geophysics applied to groundwater exploration.

Michelangelo G. da Silva received his B.S. in surveying engineering, School of Surveying Engineering (Escola de Engenharia de Agrimensura – EEA) in 2001 and M.S. in oil exploration geophysics in 2004 from UFBA, where he is currently a Ph.D. student in the Graduate Program in Geophysics. His main areas of interest are filtering methods, processing of seismic reflection data and seismic interpretation.

Giovani Lopes Vasconcelos received an electrical engineering degree (1985) and an M.S. degree in physics (1987) from UFPE and a Ph.D. in physics from the University of Chicago in 1993. He held a postdoctoral position in applied mathematics at Ohio State University from July 1993 until June 1994 a position as Fellow Researcher at the Alexander von Humboldt Foundation, Fachbereich Physik, Freie Universität Berlin from August 1994 to July 1995 and a position as Professor beginning in October 1995 in the Physics Department of UFPE, where he is currently Associate Professor 1. His areas of interest encompass fluid dynamics, complex systems and statistical physics.