GROUNDROLL ATTENUATION BY WAVELET TRANSFORMED AMPLITUDE THRESHOLD

Ítalo Cley B. de S. Maurício 1,2, Michelângelo G. da Silva 1,2, Milton J. Porsani 1,2

ABSTRACT. In land seismic data processing, the groundroll is a recurring problem. It is a coherent noise caused by Rayleigh surface waves, an association between P and S waves. This noise in the seismogram is described as a vertical cone, that contaminates seismic data by overlapping reflections that may be of interest in the seismic survey. This work presents a filtering method using 1D wavelet transform to attenuate the groundroll. This filtering method is unprecedented in land seismic processing; however, it is used efficiently in filtering electrocardiograms - ECG and audio signals. The wavelet transform is used to represent a signal at different resolutions and various time and frequency contents. In this way, it is possible to isolate the frequency band in which the noise is located and then attenuate it, improving the signal-to-noise ratio of the data. The data filtering is done by estimating a cut-off threshold for the noise-related signal amplitudes. The determination of \( \lambda \) takes into account the number of samples and the standard deviation of the noisy signal. In this work, the discrete Meyer wavelet was selected, according to the Shannon entropy criterion to perform the spectral decomposition of the seismic data. This wavelet performed better than the wavelet families: daubechies, coiflets, symlets and haar wavelet. The results of the filtering performed in a set of seismic traces in the shot domain, the supergather and later in the complete seismic line were relevant. Thanks to the attenuation of the groundroll, it was noticed significant improvements in the resolution of the seismic sections and on the velocity spectrum panel.

Keywords: groundroll attenuation; wavelet transform; filter banks; multiresolution analysis.

RESUMO. No processamento de dados sísmicos terrestres, o groundroll é um problema recorrente, tratando-se de um ruído coerente ocasionado por ondas superficiais Rayleigh, combinação entre as ondas P e S. No sismograma o groundroll é caracterizado por um cone vertical que se sobrepõe às reflexões de interesse no levantamento sísmico. Neste trabalho é apresentado um método de filtragem traço a traço utilizando a transformada wavelet 1D para atenuar o groundroll. Este método de filtragem é inovador no âmbito do processamento sísmico terrestre. No entanto, é utilizado com eficiência em filtragem de eletrocardiogramas - EGC e sinais de áudio. A transformada wavelet é utilizada para representar um sinal em distintas resoluções e variados conteúdos de tempo e frequência. Desta maneira, é possível isolar a banda de frequência na qual o ruído está localizado e em seguida atenuá-lo, melhorando a razão sinal/ruído do dado. A filtragem do dado é feita estimando um limiar de corte, definindo como threshold \( \lambda \), para as amplitudes do sinal que estão relacionadas ao ruído. A determinação do \( \lambda \) leva em conta o número de amostras e o desvio-padrão \( \sigma \) do sinal ruidoso. Neste trabalho, a wavelet discreta de Meyer \( dmey \) foi selecionada segundo o critério da entropia de Shannon para realizar a decomposição espectral do dado sísmico. Esta wavelet obteve um melhor desempenho frente às famílias de wavelets: daubechies, coiflets, symlets e a wavelet de haar. Os resultados da filtragem aplicada em um conjunto de traços sísmicos no domínio do tiro, do supergather e posteriormente na linha sísmica completa foram relevantes. Graças à atenuação do groundroll, notou-se melhorias significativas na solução das seções sísmicas e no painel do espectro de velocidade.

Palavras-chave: atenuação do groundroll; transformadas wavelets; banco de filtros; análise multiresolução.
INTRODUCTION

Reflection seismology is the most common method in hydrocarbon exploration. This method consists in imaging the subsurface layers of the sedimentary basins that are at great depths (Yilmaz, 1987). In onshore seismic data acquisition, the groundroll is the main problem in the representation of the information recorded in the seismograms, and the attenuation of this noise is the goal of many geophysical applications. In ground seismic data processing, groundroll is a recurring problem. The groundroll is a coherent noise originated by Rayleigh surface waves, a combination between the P and S waves. This noise in the seismogram is described by a vertical cone, containing low frequencies and high amplitudes, that contaminates the seismic data, overlapping with the reflections that are of interest for the interpreter.

In this work we present a new method based on Wavelet Denoising Thresholding - WDT, for attenuation of groundroll in the frequency domain. Another 1D attenuation technique can be found in Deighan and Watts (1997). It is worth noting that other methods using wavelet transformations have already been proposed for groundroll attenuation; however, they were based on the analysis of the signal in two dimensions-2D (de Matos and Osório, 2002; de Almeida, 2015). Widely used by the electronics engineering community, WDT has numerous applications in the filtering of ECG (Perin and Kozakevicius, 2009) and voice signals (da Silveira and Faceroli, 2014). These filters are directed to high frequency noises, and in these cases the results are excellent. In seismic processing, this filtering method has not previously been used to attenuate groundroll. Melo et al. (2009) address the derivative filter in groundroll attenuation; methods like Singular Value Decomposition - SVD, F-K and the spectral balancing technique in filtering this noise. It is important to note that the WDT was initially developed for high frequency filtration. In this work it was adapted to attenuate low frequencies associated to groundroll. This paper summarizes the methodology and basic concepts of filter banks, wavelet transform, multiresolution analysis and threshold that allow the filtering of the groundroll.

THE METHOD

The method proposed in this work aims at eliminating the groundroll without changing the frequency range of the signal in order to preserve as much information as possible. Thus we can significantly increase the signal-to-noise ratio, the main purpose of seismic data processing.

The attenuation method is characterized by decomposing the signal at various levels based on the frequency ranges (wavelet decomposition tree) using the Discrete Wavelet Transform – DWT, and the amplitude information is used to attenuate the noise. A threshold \( \lambda \) is calculated based on the standard deviation and sample size of each seismogram trace. The fixed thresholding function is used in this work, due to the low signal-to-noise ratio of the signal, and the Meyer’s wavelet was used to decompose the signal into several frequency ranges. The analysis of the signal with the proposed method was made trace by trace, processing faster and with fewer samples.

Filter banks, multiresolution and wavelets

Filter banks

A filter bank is a set of associated filters. In signal processing, these banks promote analysis operations and/or synthesis. In general, the analysis bank has two filters, low pass (\( L P F \)) and high pass (\( H P F \)). They separate the input signal by frequency bands. These subsignals are processed more efficiently than the original signal. At any time these subsignals can be recombined by a bank of synthesis (Diniz et al., 2014).

It is not necessary to preserve all samples of the output signals from the analysis filters, since the \( L P F \) and \( H P F \) outputs have the same number of samples as the input. Normally, the outputs undergo a decimation or downsampling process, i.e. only the even components of the outputs are preserved.

Let \( h_0 = h_0(n) \) and \( h_1 = h_1(n) \) be the impulse responses of the \( L P F \) and \( H P F \), respectively. The synthesis filters \( F_0 \) and \( F_1 \) must be specially adapted to the analysis filters \( h_0 \) and \( h_1 \) in order to cancel the errors in this analysis bank. The downsampling (\( \uparrow 2 \)) and expander or upsampling (\( \downarrow 2 \)) process, an analysis and a synthesis bank are represented in Figure 1:

![Figure 1. Analysis and Synthesis filter bank, called quadrature mirror filters - QMF.](image)

\[ y_0(n) \] and \( y_1(n) \), which are the outputs of the low and high pass branches respectively, have half the original number of samples. Obviously, \( u_0(n) \) and \( u_1(n) \) are versions of \( y_0(n) \) and \( y_1(n) \) after the process of upsampling, increasing the length of the signal component by inserting zeros between samples (Strang and Nguyen, 1996).

Multiresolution

The concept of multiresolution refers to the division of a signal into different scales of resolution in contrast to division into different frequencies (Strang and Nguyen, 1996). We can observe a signal in several resolutions \( (r) \) from a complete space \( \cup V_r \):

\[ \ldots \subset V_{-1} \subset V_0 \subset V_1 \subset \ldots \subset V_r \subset V_{r+1} \subset \ldots \quad (1) \]

Graphically we can represent, for example, two functions associated to subspaces \( V_0 \) and \( V_1 \) for a simple
that, to the other, as we can observe in Figure 2. It is noted structure of the signal, while

generate \( V \) traction between functions associated with subspaces

\[ \psi(t) = \phi(2t) - \phi(2t - 1), \quad (8) \]

Generalizing, we have the expression:

\[ \psi(t) = 2 \sum_{\kappa} f_1(\kappa) \phi(2t - \kappa), \quad (9) \]

which is called the wavelet equation; \( f_1(\kappa) \) is the coefficient of the high pass filter (Fernandes, 2015).

Wavelet Transform

The wavelet transform operates functions \( f(t) \) in continuous time and vectors \( x[n] \) in discrete time. Therefore, in order to understand its fundamentals, it is essential to use the two previous approaches: the theory of multiresolution (in continuous time) and the theory of filter banks (in discrete time) (Fernandes, 2015). This transform is based on the dot product between a signal \( x(t) \) and a basis of oscillating functions \( \psi_{rk}(t) \) located in a given time interval which are staggered and displaced along the time axis:

\[ a_{rk} = \langle x(t), \psi_{rk}(t) \rangle \quad (10) \]

\[ b_{rk} = \langle x(t), \psi_{rk}(t) \rangle \quad (11) \]

where \( r \) represents the scale; \( k \) represents the displacement \( \phi_{rk} = \phi(2^{r-k}) \); and \( \psi_{rk} = \psi(2^{r-k}) \) are the scaled and displaced versions of the scale function and of a mother wavelet, \( \psi(t) \). These are the analysis equations that generate the coefficients \( a_{rk} \) and \( b_{rk} \).

The main difference between the function basis of the Wavelet Transform and the Fourier Transform is based on the fact that the wavelets are, in most applications, of compact support, i.e., restricted to a finite time interval, while the Fourier base oscillates infinitely. This makes the wavelet transform a good tool to find events in time. Another difference comes from the process of representing a signal at various scales. Through the scaling of wavelets, the same signal can be seen with more or less detail (Strang and Nguyen, 1996). From the basis of continuous time functions \( \psi_{rk}(t) \), it is possible to produce:

\[ x(t) = \sum_{r,k} \psi_{rk}(t) \quad (12) \]

where \( x(t) \) is expanded in the wavelet base. This basis of functions is all built from a mother wavelet. Using the results of the multiresolution analysis of the equations (3) and (9), it can be seen that a mother wavelet can be written as a function of \( \phi(t) \) (scale) since the coefficients \( f_1(k) \) are known. In turn, to denote the scale function is equivalent to knowing the coefficients \( f(k) \), where \( f(k) \) and \( f_1(k) \) are coefficients of the \( LPF \) and \( HPF \), respectively. In this way, observing again, equation (7) indicates that a given scale can be expanded by

![Figure 2. Example for functions belonging to subspaces \( V_0 \) and \( V_1 \).](image)

![Figure 3. Example of a function that represents the subtraction between functions associated with subspaces \( V_0 \) and \( V_1 \) of Figure 2.](image)
another added scale of wavelets. It is possible to prove that, in discrete time, to perform this expansion is to input the signal in a filter bank with the topology indicated in Figure 4, where \( b_0 \) and \( b_1 \) are the vectors representing the wavelet coefficients or the detail and the vector \( a_0 \) coefficient scale or approximation.

![Figure 4. Discrete filter bank equivalent to the wavelet transform with two resolution levels.](image)

Finally, it is noticed that, for the analysis of discrete signals, there is no need to define the functions \( \phi(t) \) and \( \psi(t) \); we only need to determine the coefficients of the appropriate filters. Several types of filters are possible, but the most useful are those related to orthogonal and bi-orthogonal wavelets that have, normally, compact support (time-limited), that is, they are associated to a FIR filter (Fernandes, 2015).

**Wavelet Function Selection**

There are countless wavelet functions and, given such a variety, several studies have sought a criterion to select the most appropriate function to analyze the signal of interest. The wavelet function that most closely resembles the signal of interest is called optimal base (Misiti et al., 2009).

Decomposing a signal through the DWT means separating it into distinct energy and frequency levels. The energy is captured mainly in the detail components, high frequency, which have smaller amplitudes. On the other hand the energy of the low frequency components of the signal have larger amplitudes. This imbalance between the energy of the components can be measured through Shannon’s entropy (Misiti et al., 2009).

The Shannon entropy is very useful for calculating the degree of similarity between the wavelet and the waveform to be decomposed. That is, when a base function produces a relatively low entropy measure, the wavelet used is able to capture more efficiently the components of the wavelet signal and noise. Therefore, this base function is the most appropriate and defines the best choice, since the ability to separate such components is maximized. In other words, the greater the similarity between the wavelet used and the analyzed signal, the better the noise attenuation process (Honório, 2011).

The Shannon entropy can be calculated by the following expression (da Silveira and Faceroli, 2014):

\[
H_s = - \sum_i |s_i|^2 \log_2 \frac{|s_i|^2}{|s|^2},
\]

where \( |s|^2 = \sum_i |s_i|^2 \), \( s \) is the signal we want to calculate the entropy of, and \( s_i \) is the set of wavelet coefficients obtained by the wavelet transform of signal \( s \).

Table 1 (the complete table can be found in Maurício (2017), represents the mean Shannon entropies of the main orthogonal wavelet for the shot 138 of the seismic line 31-81. The wavelets in bold have generated the best results with the highest degree of similarity with the seismic trace.

Table 1 (the complete table can be found in Maurício (2017), represents the mean Shannon entropies of the main orthogonal wavelet for the shot 138 of the seismic line 31-81. The wavelets in bold have generated the best results with the highest degree of similarity with the seismic trace.

Note that Meyer’s Discrete Wavelet (dmey) obtained the best result so it was chosen to carry out the present work. Some wavelets of the families: Daubechies, Symlet and Coiflets had relevant results (Maurício, 2017). However, not enough to overcome the performance of dmey, especially considering the expenses with computational cost.

The Meyer’s wavelet in Figure 5, is an orthogonal wavelet proposed by Yves Meyer. The FIR based approximation is infinitely differentiable with infinite support (Misiti et al., 2009).

![Figure 5. Meyer’s Wavelet (Misiti et al., 2009).](image)

**Thresholding**

The filtering process involves the manipulation of a time series in order to change the spectral characteristics of the data. In general terms, filtering modes can be divided into frequency-selective filtering (FSF), threshold filtering (TF) and wiener filtering (WF) filters. In each case, a different procedure is used to isolate the noise from signal in the time series (Honório, 2011). In the present work the threshold filtering approach was used, equation (14), specifically the soft threshold, Figure 6.

\[
\delta^*_\lambda = \begin{cases} 
\text{Sign}(w)(|w| - \lambda) & \forall |w| > \lambda \\
0 & \text{c.c.}
\end{cases}
\]

(14)

The concept of wavelet thresholding applied to the noise attenuation process was introduced by Donoho and Johnstone (1994). The FSFs method involves the removal of unwanted frequency components. The TF approach removes all information related to variations below a certain threshold (or amplitude level). The cut-off threshold, \( \lambda \), is chosen based on the signal energy.
Table 1. Table of the Shannon entropies of the main orthogonal wavelets for the shot 138 of the seismic line 31-81 of the National Petroleum Reserve-Alaska.

<table>
<thead>
<tr>
<th>Wavelet</th>
<th>Entropy</th>
<th>Wavelet</th>
<th>Entropy</th>
<th>Wavelet</th>
<th>Entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>db37</td>
<td>-926,3514</td>
<td>sym5</td>
<td>-958,3544</td>
<td>sym17</td>
<td>-1073,400</td>
</tr>
<tr>
<td>db38</td>
<td>-1012,500</td>
<td>sym6</td>
<td>-891,9574</td>
<td>sym18</td>
<td>-955,3478</td>
</tr>
<tr>
<td>db39</td>
<td>-878,0149</td>
<td>sym7</td>
<td>-950,8834</td>
<td>sym19</td>
<td>-981,7670</td>
</tr>
<tr>
<td>db40</td>
<td>-947,7248</td>
<td>sym8</td>
<td>-869,5652</td>
<td>sym20</td>
<td>-963,3817</td>
</tr>
<tr>
<td>db41</td>
<td>-933,6557</td>
<td>sym9</td>
<td>-1081,500</td>
<td>sym21</td>
<td>-906,9898</td>
</tr>
<tr>
<td>db42</td>
<td>-998,7749</td>
<td>sym10</td>
<td>-962,1787</td>
<td>coif1</td>
<td>-916,7691</td>
</tr>
<tr>
<td>db43</td>
<td>-1078,200</td>
<td>sym11</td>
<td>-1014,300</td>
<td>coif2</td>
<td>-988,8653</td>
</tr>
<tr>
<td>db44</td>
<td>-1052,500</td>
<td>sym12</td>
<td>-974,1979</td>
<td>coif3</td>
<td>-928,4050</td>
</tr>
<tr>
<td>db45</td>
<td>-1044,800</td>
<td>sym13</td>
<td>-934,2484</td>
<td>coif4</td>
<td>-952,3634</td>
</tr>
<tr>
<td>sym2</td>
<td>-832,5636</td>
<td>sym14</td>
<td>-955,1905</td>
<td>coif5</td>
<td>-1031,500</td>
</tr>
<tr>
<td>sym3</td>
<td>-737,6859</td>
<td>sym15</td>
<td>-952,9619</td>
<td>haar</td>
<td>-598,0637</td>
</tr>
<tr>
<td>sym4</td>
<td>-960,8373</td>
<td>sym16</td>
<td>-909,5691</td>
<td>dmey</td>
<td>-1234,600</td>
</tr>
</tbody>
</table>

Figure 6. Soft threshold.

and the noise standard deviation. If the wavelet coefficient is greater than λ, it is assumed that its contribution to the signal is significant and therefore used in the reconstruction. Otherwise, it is considered as coming from noise and then discarded. Determining the threshold value λ is an important part in the noise elimination process. A low threshold may result in a signal quite similar to the input data, however still with the presence of noise. At the other extreme, coefficients that have relevant information can be overridden, making the output signal excessively soft (Honório, 2011).

Donoho and Johnstone (1994) have shown that for \( n \) independent and identically distributed variables, the expected maximum value is \( \sqrt{2 \log_e(n)} \), which leads to universal thresholding, also known as fixed form. This is one of the first proposed rules and provides a quick, automatic and easy threshold (Katul and Vidakovic, 1995). The universal threshold expression is given by:

\[
\lambda = \sigma \sqrt{2 \log_e(n)},
\]

where \( \sigma \) is the standard deviation of the signal samples and \( n \) is the number of samples.

NUMERICAL RESULTS

The acquisition of the seismic reflection line 31-81 of the National Petroleum Reserve-Alaska was carried out by the company Geophysical Service in 1981 at the request of the United States Geological Service - USGS. The survey was performed using a symmetric split-spread. Table 2 shows the main information regarding the parameters of the acquisition. Additional information on the seismic data of the line is available (USGS).

The flow chart in Figure 7 summarizes the groundroll filtering process. As can be seen, the cut-off threshold was calculated from each seismic trace. Next, each trace of the seismogram was decomposed up to level 4, with Meyer’s discrete wavelet, given \( \{A_j, B_j\} \), \( j = 1, \ldots, 4 \), which corresponds to the frequency range of the groundroll (2-18 Hz). Then, the Soft Thresholding function was used to eliminate the high amplitudes that correspond to the groundroll. Finally, the Inverse Discrete Wavelet Transform (IDWT) was applied to the signal with the levels of approximation \( A_4 \) and details
Table 2. Parameters of acquisition of the seismic reflection line 31-81 of the National Petroleum Reserve-Alaska.

<table>
<thead>
<tr>
<th>Line</th>
<th>31-81</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spread (ft)</td>
<td>5525-55-0-55-5525</td>
</tr>
<tr>
<td>Registration time (s)</td>
<td>2</td>
</tr>
<tr>
<td>Sampling interval (ms)</td>
<td>4</td>
</tr>
<tr>
<td>Degree of coverage (%)</td>
<td>1200</td>
</tr>
<tr>
<td>Interval between shooting points (ft)</td>
<td>200</td>
</tr>
<tr>
<td>Distance between geophones (ft)</td>
<td>110</td>
</tr>
<tr>
<td>Number of channels</td>
<td>96</td>
</tr>
<tr>
<td>Number of samples per trace</td>
<td>571</td>
</tr>
</tbody>
</table>

![Flowchart](image)

Figure 7. Flowchart of the seismic signal filtering.

$\lambda$ threshold to reconstitute the original signal without the marked presence of noise.

Figure 8 shows the effectiveness of the Wavelet-Thresholding method. The original shot-gather, number 138 of the line 31-81, is presented in Figure 8a, the filtered data are shown in Figure 8b and the residue is in Figure 8c. Note that the filtered data preserve the reflections of interest, especially the reflections that were superimposed by the groundroll. Consequently, the residue concentrates mostly groundroll related events, as can be seen in Figure 8c.

Figure 9 shows the average amplitude spectra of data in Figure 8. In the filtered data (blue curve), it can be noted that all the frequency bands of the original signal were preserved and the amplitudes related to the frequency band of the groundroll were attenuated. The red curve for the residue shows that the frequency range that brings together the largest amplitudes is exactly the band that characterizes the groundroll. In short, the residue consists essentially of low frequencies and high amplitudes.

The velocity spectrum obtained by the supergather’s filtration, in Figure 10, reveals that the attenuation of the groundroll significantly improved the semblance panel and subsequently the velocity analysis. The existing points of coherence became more defined, their values increased, and new points could be observed, especially between the times of 0.5 - 1.0s and 1.5 - 2.0s.

Finally, the stacked section, in Figure 12, obtained through the flowchart (Fig. 11) (preprocessed data in flowchart are: data with geometry, edition, mute and Automatic Control Gain - AGC), and its correspondent spectrum frequency. In Figure 13, the line with the filtering proposed in this work presents good results for the subsurface interpretation. In Figure 12a, we have the original data; note that there is a strong presence of groundroll, especially on the upper left edge. In Figure 12b, we have the filtered data in shot domain, highlighted in Figure 11 by dashed lines, see that there is still groundroll in lesser quantity. Nevertheless, there is a significant improvement in the visualization of the reflections. In Figure 12c, we have the filtered data in supergather domain, highlighted in Figure 11 by red lines. The groundroll was eliminated completely and the reflections were preserved. Compare the upper left margin of Figures 12a and 12c for better understanding.

The observations made for the frequency spectrum of the shot are also valid for the spectrum of the seismic sections of Figure 12.

**CONCLUSION**

The study of groundroll attenuation using the Wavelets Thresholding method is promising, given the satisfactory results presented in the seismic filtering of a field data example. Because it does not depend on the spatial and temporal distribution of groundroll, the method is useful for filtering in cases where the noise in question is not well defined in time and/or space. We can observe that the method used has a relevant application for performing the velocity analysis as seen in the velocity spectrum panels. The method has simple implementation and low computational cost.

Braz. J. Geophys., 39(4), 2021
Figure 8. Result of the wavelet-thresholding method. The original shot gather in (a), the filtered in (b) and the residue in (c).

Figure 9. Average amplitude spectra of the data in Figure 8.
Figure 10. Supergather 549 of line 31-81. Original supergather in (a), velocity spectra of the supergather in (b), filtered supergather in (c) and velocity spectra of the filtered supergather in (d).

Figure 11. Flowchart to obtain the seismic sections.

Figure 12. Stacked section of the original data in (a). Stacked section of the data filtered in the shotgather domain in (b) and stacked section of the data filtered in the supergather domain in (c).

Figure 13. Average amplitude spectra of data in Figures 12a (black curve) and 12b (blue curve), and residue (red curve).
ACKNOWLEDGMENTS

The authors wish to express their gratitude to INCT-GP/CNPq/MCT, PETROBRAS, ANP, and FINEP, for financial support. We also thank LANDMARK for the licenses granted to CPGG/IGEO/UFBA.

REFERENCES


I.C.B.S.M.: responsible for the development, implementation, and testing of the algorithms. Also generated the figures and final results presented in the article; M.J.P and M.G.S.: contributed to the critique and evaluation of the results, and also helped in the writing and final organization of the article.

Received on May 12, 2021/ Accepted on March 21, 2022.

Creative Commons attribution-type BY