Denoising of the groundroll using *Wavelets Thresholding*

Ítalo Maurício, Michelângelo G. da Silva and Milton J. Porsani CPGG/IGEO/UFBA and INCT-GP/CNPq/Brazil

Abstract

In this work the method is based on the spectral decomposition of the seismic data using Wavelet Transforms (WT) to attenuate the groundroll. The WT is used to represent a signal at different resolutions and various time and frequency contents. The filtering method separates the signal by frequency range and the band of coefficients related to the noise, then signal-to-noise ratio of the seismic data is improved by attenuating the groundroll. The signal filtering is done by estimating a cut-off, threshold (λ), for signal amplitudes which are related to noise. The determination of λ takes into account the standard deviation and the size of the signal sample. In order to test the efficiency of the method in question, the land seismic line (RL5090) of the Tacutu basin was processed and good results were obtained.

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 ground acquisitions groundroll is the main problem in the representation of the information recorded in the seismograms, the attenuation of this noise is the goal of many geophysicists. 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. Some already established methods eliminate the groundroll (FK and $Fourier$) however, some of the reflections are also eliminated because they are in the same frequency range of the noise.

In this work we present a new method based on *wavelet denoising thresholding* for attenuation of groundroll in the frequency domain. It is worth noting that other methods using wavelet transformations have already been proposed for groundroll attenuation, however, these were based on the analysis of the signal in two dimensions-$2D$ (de Almeida (2015)).

Method

The method proposed in this work aims to eliminate 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 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 amplitudes information is used to attenuate the noise. A threshold λ is calculated based on the variance 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. Each seismic trace has 1001 samples.

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 (LPF) and high pass (HPF). 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 (Paulo Sergio R. Diniz and Netto (2014)). It is not necessary to preserve all samples of the output signals from the analysis filters, since the LPF and HPF outputs have the same number of samples as the input, ie in total, twice as many samples have been generated. 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 LPF and HPF, 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* $(\downarrow 2)$ and expanders or *upsampling* $(\uparrow 2)$ process, an analysis and a synthesis bank are represented by Figure 1:

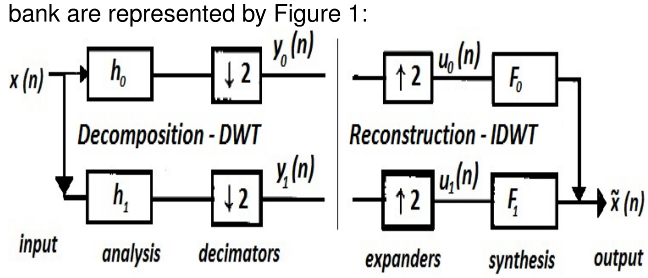


Figure 1: Analysis and Synthesis filter bank, called *quadrature mirror filters - QMF*

Where 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 and u_1 are versions of y_0 (n) and y_1 (n) after the process of *upsampling*, lengthening 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 $\cup V_r$:

$$\ldots \subset V_{-}1 \subset V_0 \subset V_1 \subset \ldots \subset V_r \subset V_r+1 \subset \ldots \tag{1}$$

Graphically we can represent, for example, two functions associated to subspaces V_0 and V_1 for a simple and specific case:

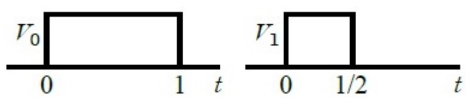


Figure 2: Example for functions belonging to subspaces V_0 and V_1

It is possible to associate to the subspace V_0 a set of functions $\phi(t-\kappa)$ and to the subspace V_1 another set $\phi(2t-\kappa)$, that is, one version staggered with respect to the other, as we can observe in Figure 2. It is noted that,

$$\phi(t) = \phi(2t) + \phi(2t-1) \tag{2}$$

In this way, it is possible to generalize to the following equation

$$\phi(t) = 2\sum_{\kappa} f(\kappa)\phi(2t-\kappa) \tag{3}$$

Where $f(\kappa)$ are weighting coefficients and the expression is known as the refinement equation (or dilation eq). The function $\phi(t)$ is called the scale function. We can now use the information that results from the subtraction between two scales represented by the function $\psi(t)$ belonging to a subspace W_0:

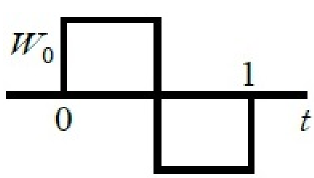


Figure 3: Example of a function that represents the subtraction between functions associated with subspaces V_0 and V_1 of the figure 2

It is possible to observe that the function $\psi(t)$, associated with subspace W_0 (Figure 3) together with V_0 can generate V_1, according to Figure 2, in this way:

$$\psi(t) = 2\phi(2t) - \phi(t) \tag{4}$$

that is,

$$V_1 = V_0 + W_0 \tag{5}$$

If we increase the resolution, it is possible to write $V_2 = V_1 + W_1$, and replacing in this expression the equation (5), observe that:

$$V_2 = V_0 + W_0 + W_1, \tag{6}$$

We can conclude that V_0 refers to an approximate structure of the signal while W_0 and W_1 are the details. Generalizing the above results, we have:

$$V_{R+1} = V_0 + \sum_{r=0}^{R} W_r \tag{7}$$

The subspace W_r has an oscillatory nature because it express differences, and in transform theory the signals belonging to this subspace are called *wavelets*. Therefore, it is possible to obtain, substituting equation (2) in equation (4),

$$\psi(t) = \phi(2t) - \phi(2t-1), \tag{8}$$

generalizing, we have the expression:

$$\psi(t) = 2\sum_{\kappa} f_1(\kappa)\phi(2t-\kappa), \tag{9}$$

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

Wavelets

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 $\psi_{r\kappa}(t)$ located in a given time interval which are staggered and displaced along the time axis:

$$a_{r\kappa} = < x(t), \phi_{r\kappa}(t) > \tag{10}$$

$$b_{r\kappa} = < x(t), \psi_{r\kappa}(t) > \tag{11}$$

Where r represents the scale, k represents the displacement, $\phi_{r\kappa} = \phi(2^{rt-\kappa})$ and $\psi_{r\kappa} = \psi(2^{rt-\kappa})$ 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 $a_{r\kappa}$ and $b_{r\kappa}$. 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 $\psi_{r\kappa}(t)$, it is possible to produce:

$$x(t) = \sum_{r,\kappa} \psi_{r\kappa}(t) \tag{12}$$

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 $g(k)$ are known. In turn, to denote the scale function means to know the coefficients $f(k)$, where $f(k)$ and $g(k)$ are coefficients of the LPF and HPF, respectively. In this way, again observing the equation (7), which indicates that a given scale can be expanded by 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 *detail* and the vector a_0 coefficient scale or *approximation*.

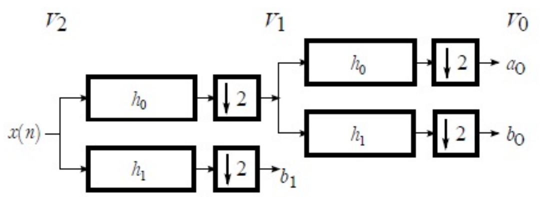


Figure 4: Discrete filter bank equivalent to the wavelet transform with two resolution levels

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 biortogonal wavelets that have, normally, compact support (time-limited) , that is, they are associated with a FIR filter (Fernandes (2015)).

Meyer's Wavelet

The Meyer's wavelet is an orthogonal wavelet proposed by *Yves Meyer*, the FIR based approximation is infinitely differentiable with infinite suppor (Michel Misiti and Poggi (2009)).

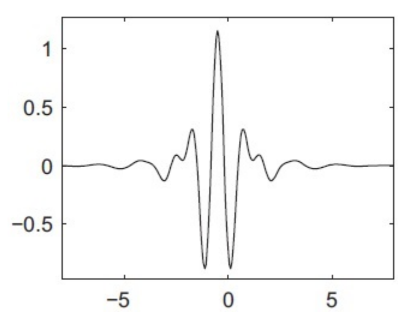


Figure 5: Meyer's Wavelet (Toolbox MatLab,2009)

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 signal in the time series (Honório (2009)). In the present work, the threshold filtering approach was used.

$$\delta_\lambda^s = \begin{cases} Sign(w)(|w| - \lambda) & \forall \quad |w|>\lambda \\ 0 & c.c. \end{cases} \tag{13}$$

The concept of wavelet thresholding applied to the noise attenuation process was introduced by Donoho (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, λ, is chosen based on the signal energy and the noise variance. 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 (2009)).

Donoho (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 (1995)). The universal threshold expression is given by:

$$\lambda = \sigma\sqrt{2\log_e(n)}, \tag{14}$$

where σ is the standard deviation of the signal samples and n sample numbers.

Attenuation of groundroll

The flow chart in Figure 6 summarizes the groundroll filtering process. As can be seen, each seismic trace of the seismogram was decomposed up to level 4, which corresponds to the frequency range of the groundroll (0-18Hz), with Meyer's discrete wavelet. Then, the cut threshold calculated from the seismic trace was used by the *Soft Thresholding* function to eliminate the high amplitudes that correspond to the groundroll. Then, the Inverse Discrete Wavelet Transform (IDWT) was applied to the signal with the levels of approximation (A4) and details (B4) thresholded to reconstitute the original signal without the marked presence of noise.

Numerical Results

The results below show the effectiveness of the Wavelet-Thresholding method. In Figure 7 we can observe the (A) shot 128 of the $LR5090$ line of the Tacutu basin, (b) the filtered shot and (c) the respective residue. The observed frequency spectrum (Figure 8) reaffirms the efficacy of the method. It can be noted that all the

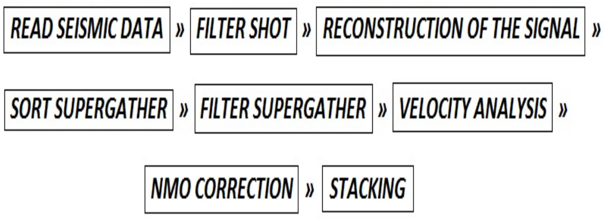


Figure 6: Seismic signal filtering flowchart

frequency bands of the original signal were preserved and only the amplitudes related to the frequency band of the groundroll were attenuated in relation to the signal of the original shot. The velocity spectrum obtained by the supergather filtration (Figure 9) reveals that the attenuation of the groundroll significantly improved the semblance panel and subsequently the velocity analysis. Finally, the stacked section (Figure 11) and its correspondent spectrum frequency (Figure 10) of the $LR5090$ line of the Tacutu basin with the filtering proposed in this work presents good results for the subsurface interpretation.

Conclusion

The introductory study of groundroll attenuation using the *Wavelets Thresholding* method is promising given the satisfactory results presented in the seismic filtering of the line $RL5090$ of the Tacutu basin. 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 spatially. 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.

Future Directions

Studies with other families of wavelets and the windowing of the seismic trace in the filtering process should be carried out with the intention of refining the presented method in search of better results.

References

de Almeida, L., 2015, Atenuação de ruídos coerentes e interpolação de dados sísmicos utilizando a an´alise de multiresolução: Salvador, UFBA.

Donoho, D., . J. I., 1994, Ideal spatial adaptation by wavelet shrinkage: Biometrika, **81**, 425–455.

Fernandes, A., 2015, Contribuições ao problema de extração de tempo musical: Unicamp, Campinas.

Honório, B. C. Z., 2009, Aplicação da transformada wavelet na atenuação de ruídos e determinação de ciclos sedimentares em perfis geofísicos, dissertação de mestrado: Campinas, Unicamp.

Katul, G., . V. B., 1995, The partitioning of attached and detached eddy motion in the atmospheric surface layer using lorentz wavelet filtering: Boundary-Layer Meteorology, **77**, 153–172.

Michel Misiti, Yves Misiti, G. O. and J.-M. Poggi, 2009, Wavelet toolbox, matlab: Mathworks.

Paulo Sergio R. Diniz, E. d. S. and S. L. Netto, 2014, Processamento digital de sinais projeto e análise de sistemas: Bookman,Porto Alegre.

Strang, G. and T. Nguyen, 1996, Wavelets and filter banks: Wellesley-Cambridge Press.

Yilmaz, O., 1987, Seismic data processing: Society of Exploration Geophysicists, Tulsa - Oklahoma.

Acknowledgements

Í. Maurício thanks the fellow student D. Montenegro for the text revision. The authors also wish to express their gratitude to brazilian agencies, (INCT-GP/CNPq/MCT, CAPES, PETROBRAS, ANP, FINEP) for financial support.

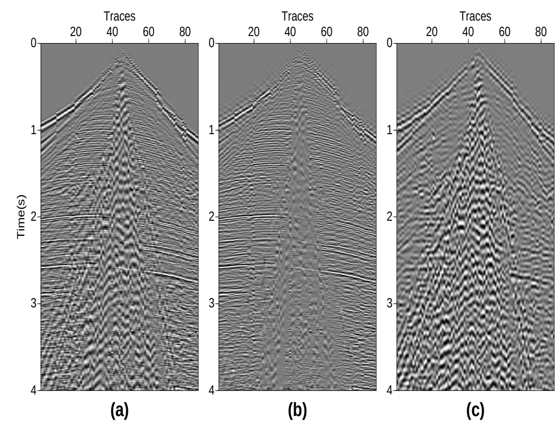


Figure 7: The result of filtering. The original shot gather in (a),The filtered in (b) and the residue in (c)

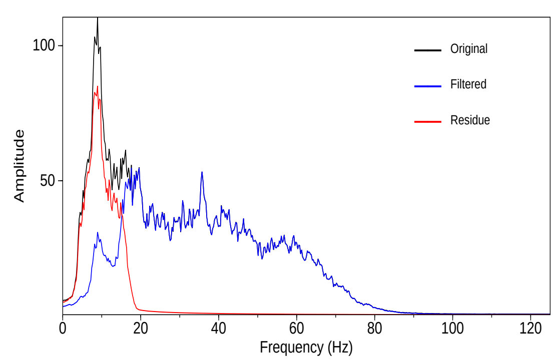


Figure 8: Average amplitude spectra of the data in Fig.7

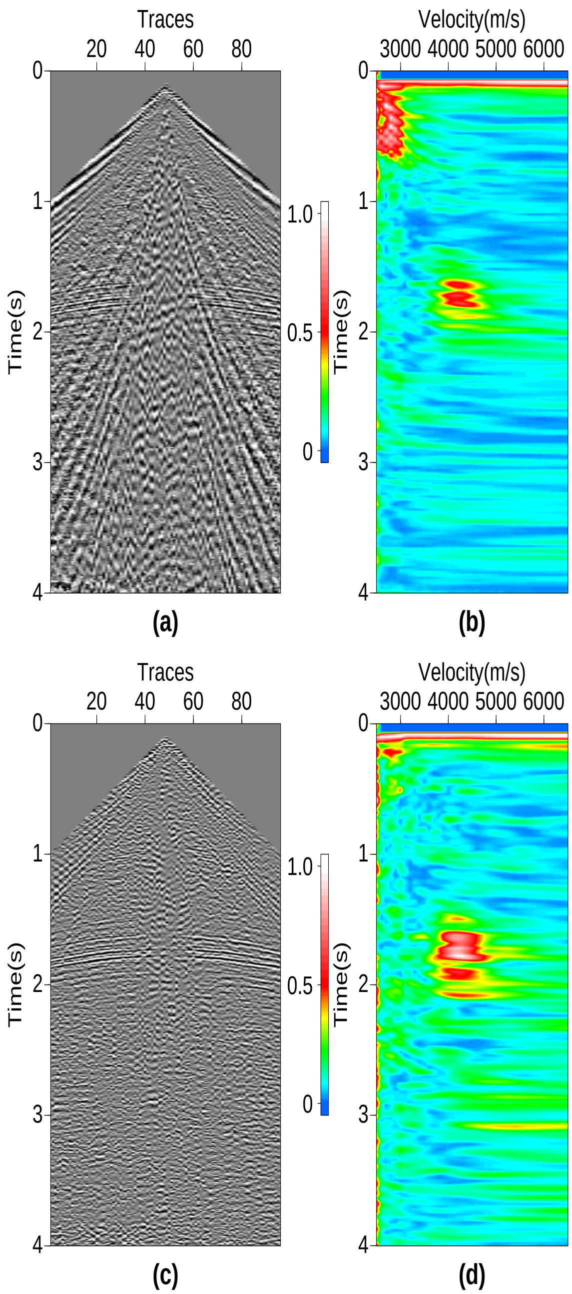


Figure 9: Original supergather in (a), Velocity spectra of the supergather in (b), Filtered supergather in (c) and Velocity spectra of the filtered supergather in (d)

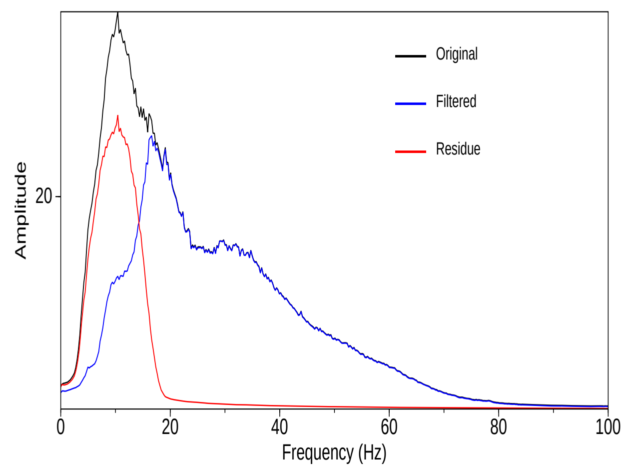
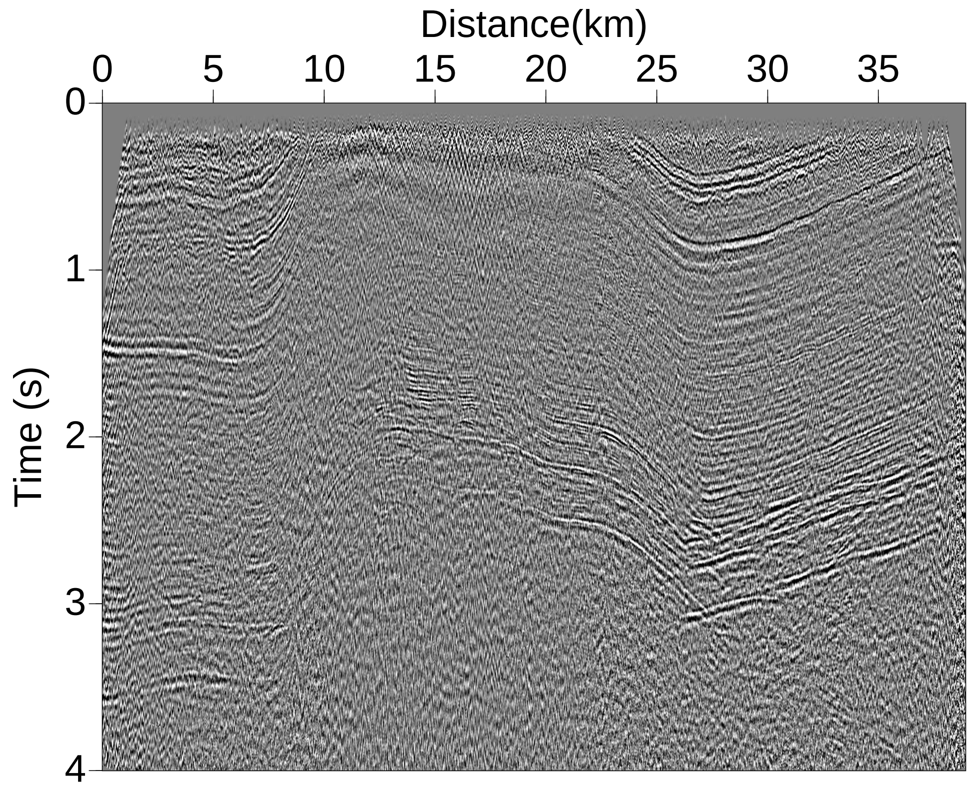
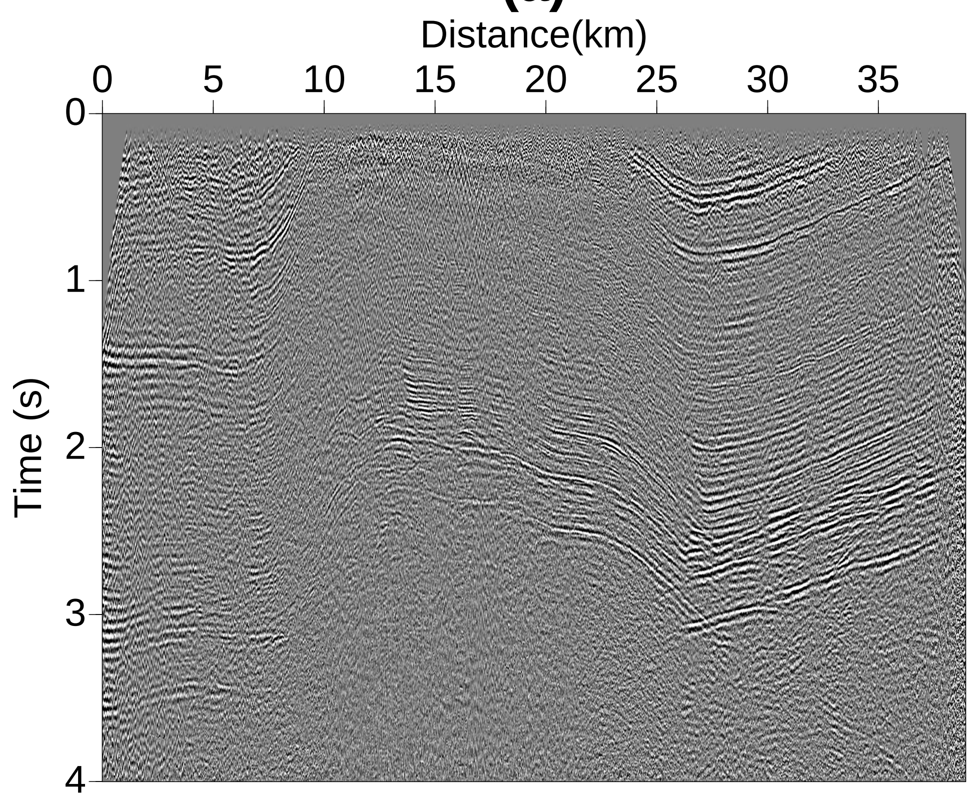


Figure 10: Averange amplitude spectra of original and filtered stacked sections in Fig.11



(a)



(b)

Figure 11: Stacked section of the original data in (a) and Stacked section of the filtered data in (b).