Seismic Attributes and Time-Frequency analysis by using Wigner-Ville distribution combined with the Maximum Entropy Method

Leonardo M. Batista*, CPGG/UFBA, Milton J. Porsani, CPGG/UFBA

This paper was prepared for presentation at the 15th International Congress of the Brazilian Geophysical Society, held in Rio de Janeiro, Brazil, 31 July to 3 August 2017, 2017.

Abstract

The growing need for more defined images of seismic attributes for analysis and description of reservoirs encourages the development of new methods that are capable of displaying high resolution images of seismic attributes. Therewith, we propose the use of the method called Maximum Entropy of Wigner-Ville (MEM-WV), that combines the distribution of Wigner-Ville with the Burg's Maximum Entropy method, in order to get the high resolution power spectrum of the seismic signal. In this method, the power spectrum is obtained by doing Fourier Transform (FT) of each Kernel term of Wigner-Ville, and then by estimating and extending each Kernel sequence. In order to better describe hydrocarbon reservoirs, this work will favor the achievement of structural, stratigraphic and geomorphological attributes with higher resolution.

Introduction

The subject titled Time-Frequency Analysis (TFA, T-F), or Time / Scale-Frequency Analysis (TSFA) usually combines principles of signal analysis with those of differential equations. This means that the harmonic analysis is described in the time-space domain and in the associated spectral content. Thus, the different techniques retain a direct relationship with the Fourier transform to decompose and represent a function (Leite, 2015).

Techniques for decomposing a signal in time and frequency can be found in many areas of science, including the analysis of the volatility and the correlation of financial rates (Pimentel and da Silva, 2011), the analysis of disturbances in electrical networks (Soares, 2013). In seismic studies, the list related to T-F Analysis is long and constantly evolving.

In this work, we use the maximum entropy method of Burg (1967) to estimate the spectrum of the data with high resolution, avoiding the crossed terms. This method is based on the estimation of a linear predictive operator and on the estimation and extension of the autocorrelation function, called (Zoukaneri and Porsani, 2015) maximum entropy of Wigner-Ville method (MEM-WV). After generating the unknown coefficients of the self-correlation, the Kernel terms of the Wigner-Ville Maximum Entropy can be generated, which, once applied to the Fourier Transform, generate the maximum entropy spectrum.

In the last step, we extracted the mean instantaneous frequency, mean variance, and Skewness and Kurtosis attributes of the WV-MEM, which were essential for the characterization of reservoirs based on seismic attributes.

Theory

For an uniformly sampled signal $z(t)$ with Δ t interval, the discrete Wigner-Ville can be represented as follows:

$$W(t,f) = 2\sum_{k=0}^{N-1} z(t-k)z*(t+k)e^{(-2j\pi fk)} \quad (1)$$

where $z(n)$ is the analytical signal corresponding to $x(n)$ and represented by:

$$z(n) = x(n) + jH[x(n)] \quad (2)$$

where $H[x(n)]$ represents the Hilbert transform of the signal $x(n), n = 0,...,N_s-1$, and N_s, the number of samples. The covariance matrix associated to the analytic signal can be represented by:

$$C = zz* \quad (3)$$

Where " * " represents the conjugate transpose of the matrix. The Kernel of the Wigner-Ville distribution is represented by the set of sequences of terms along the secondary diagonals of the associated covariance matrix, represented by:

$$K(n) = \{k_n(-l),...,k_n(0),...,k_n(l)\} \quad (4)$$

where each term in equation 4 is:

$$\begin{matrix} z(n-l)z^*(n+l), & |l| \leq min\{n, N_s-n\} \\ 0, & |l| \leq min\{n, N_s-n\} \end{matrix}$$

The central term $k_n(0) = z(n)z^*(n)$ is associated with the sample $z(n)$.

To facilitate the understanding of equation 1, the Figure 1 represents a graphic illustration of discrete Wigner-Ville. The covariance matrix is formed by z and z^*, the kernel sequences are represented by the blue lines, and the $k_n(.)$ terms, by black circles.

The power spectrum corresponding to the Fourier transform of the Kernel $K(n)$ of Wigner-Ville can also be seen in Figure 1 and can be analyzed by means of the instantaneous power spectrum of the signal $z(n)$ and represented as:

$$P(n) = W(n) = \left\{ w_n(-\frac{N-1}{2}),...,w_n(0),...,w_n(\frac{N-1}{2}) \right\}, \quad (5)$$

where each coefficient is given by:

$$w(m) = \frac{1}{N} \sum_{l=-(N-1)/2}^{(N-1)/2} k_n(l) W_4^{ml} \quad (6)$$

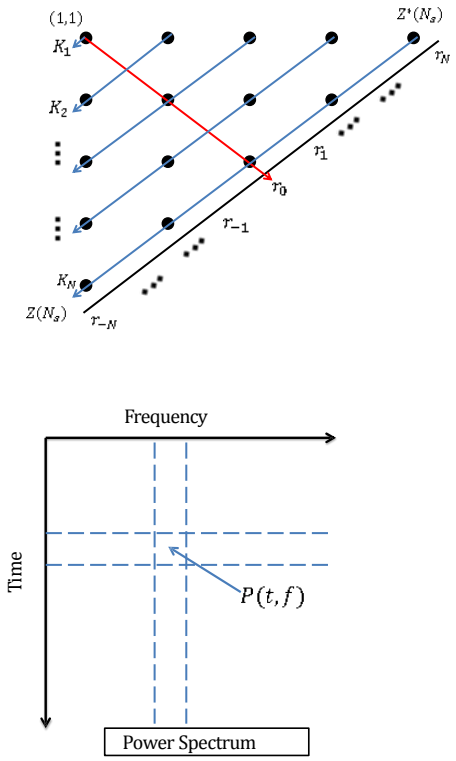

Figure 1: Modified image from Zoukaneri (2014). Schematic discrete distribution of Wigner-Ville. The blue diagonals represent the Kernel $K(n)$ sequences. The red arrow represents the sum of the autocorrelation terms. The image of the instantaneous power spectrum $P(t,f)$ is generated by Fourier transforming the kernel sequences.

N is the number of terms used in the Discrete Fourier Transform except that the so called twiddle factor is normally defined as $W_2 = exp[-j2\pi/N]$. The additional power of 2 represents a scaling of the frequency axis by a factor of 2. Equation 6 can be evaluated efficiently using the standard fast Fourier transform (FFT) algorithms (Zoukaneri and Porsani, 2015).

Maximum Entropy Method

Usually the power spectrum is estimated using the Short Window Fourier Transform (SWFT) of the autocorrelation coefficients (ACC). However, the leakage effect when the data is truncated makes the Fourier transform limited. To obtain good resolution from a limited series of data, Burg (1995) formulated the Maximum Entropy Method, constituting a linear predictive filter, which gives the maximum quadratic error between the data and the predicted value of the data. Marple (1978) presents the entropy of a Gaussian process proportional to:

$$\int_{-f_n}^{f_n} logP(f)df \tag{7}$$

where $P(f)$ is the power spectrum, and f_N is the Nyquist frequency. Burg maximizes the entropy with the following condition:

$$R_n = \int_{-f_N}^{f_N} P(f)e^{(i2\pi fn\Delta t)}df \tag{8}$$

where R_n is self covariance, $-N < n < N$. With the help of Lagrange coefficients, we can obtain the solution λ_k $k = 1,...,N$, such that:

$$\frac{\partial}{\partial P(f))}\left(\int_{-f_n}^{f_n} logP(f)df - \lambda_k\left(\sum_{-M}^{M}\int_{-f_N}^{f_N} P(f)e^{(i2\pi fn\Delta t)}df\right)\right) = 0 \tag{9}$$

The solution of equation 9 is given by:

$$P(f) = \frac{E_{N_c}\Delta t}{\left|\sum_{n=0}^{N_c-1} c_n e^{-j2\pi fn\Delta t}\right|^2}, \tag{10}$$

where $P(f)$ is the power spectrum, c_n, $n = 0,...,N_c-1$, $(c_0 = 1)$, represents the coefficients of the prediction error operator (PEO) of order N_c, and E_c is the corresponding energy of the error. f is limited by the range of Nyquist $-1/(2\Delta t) \leqslant f \leqslant 1/(2\Delta t)$.
The basic form of the maximum entropy method is given by the 10 equation. Thus, the power spectrum $P(t)$ is completely defined if the coefficients c_n and energy E_{N_c} are known. Among the methods used to determine the coefficients of the PEO of order N_c and the corresponding energy E_{N_c}, the most used are the Yulle-Walker equation and the Burg algorithm, which was used in this work.

Solution of the Maximum Entropy equation via Burg equation

This method, originally proposed by Burg (1967) as a maximum entropy algorithm, was later interpreted as a restrictive least squares minimization algorithm. Estimates of the model parameters c_n are obtained by minimizing, for each model of order N_c, the arithmetic mean of the power of direct and reverse linear predictive errors, with the limitation that the parameters c_n satisfy Levinson's recursion:

$$c_{j,i} = c_{j-1,i} + c_{j,i}c_{j-1,j-1} \tag{11}$$

This limitation is made to ensure stability of the Kay and Marple operator (1981).
Considering a signal $x_n, n = 1,...,N$ and some coefficients $c_i, i = 1,...,k$, the direct and reverse linear prediction y_n and Z_n of the original signal X_n can be obtained, respectively, by:

$$y_n = -\sum_{i=1}^{k} c_i x_{n-1} \tag{12}$$

and

$$z_n = -\sum_{i=1}^{k} c_i x_{n+1} \tag{13}$$

where c_i $i = 1,...,k$ is the prediction operator coefficient of order K, y_n is the linear combination of the previously known weighted coefficients (direct prediction), and z_n is the weighted linear combination of the next known coefficients (reverse prediction) (Zoukaneri, 2014).
Burg (1967) proposes to minimize simultaneously the sum of squared errors of direct and reverse prediction subjected to a constraint.
By calling F_k the direct prediction error and B_k the reverse prediction error, we have:

$$F_k = \sum_{n=1}^{N}(x_n - y_n)^2 = \sum_{n=1}^{N}(x_n - (-\sum_{i=1}^{k} c_i x_{n-1}))^2 \tag{14}$$

and

$$B_k = \sum_{n=1}^{N} (x_n - z_n)^2 = \sum_{n=1}^{N} (x_n - (-\sum_{i=1}^{k} c_i x_{n+1}))^2 \quad (15)$$

Equation 11 can be rewritten as:

$$c_{n+1} = c_n + \mu c_{k+1-n}, \quad (16)$$

where $\mu = c_{j,j}$, $n = 1,...,k$, and k is the number of coefficients. For $c_0 = 1$, equation 14 can be rewritten as:

$$F_k = \sum_{n=1}^{N} (c_0 x_n + \sum_{i=1}^{k} c_i x_{n-1})^2 = \sum_{n=1}^{N} (f_k(n))^2 \quad (17)$$

where

$$f_k(n) = \sum_{i=0}^{k} c_i x_{n-1} \quad (18)$$

According to this, we can rewrite equation 15 in the same way:

$$B_k = \sum_{n=1}^{N} (c_0 x_n + \sum_{i=1}^{k} c_i x_{n+1})^2 = \sum_{n=1}^{N} (b_k(n))^2 \quad (19)$$

and

$$b_k(n) = \sum_{i=0}^{k} c_i x_{n+1} \quad (20)$$

Therefore, we can obtain:

$$F_{k+1} + B_{k+1} = \sum_{n=1}^{N} (f_{k+1}(n))^2 + \sum_{n=1}^{N} (b_{k+1}(n))^2 \quad (21)$$

Using the equations 16, 18, and 20, we can derive:

$$f_{k+1}(n) = f_k(n) + \mu b_k(n-k-1) \quad (22)$$

$$b_{k+1}(n) = b_k(n) + \mu f_k(n+k+1) \quad (23)$$

Using 22 and 23 in 21 and minimizing with respect to μ, we have:

$$\mu = \frac{-2\sum_{n=0}^{N-k-1} f_k(n+k+1) b_k(n)}{\sum_{n=0}^{N} f_k(n)^2 + \sum_{n=0}^{N} b_k(n)^2} \quad (24)$$

where the Burg algorithm is summarized by:

$$\begin{cases} \frac{\partial (f_{k+1} + b_{k+1})}{\partial \mu} = 0, \\ c_{n+1} = c_n + \mu c_{k+1-n}. \end{cases} \quad (25)$$

It is possible to estimate the coefficients $c_{j,i}$ of the prediction error operator from the discrete signal using the Burg algorithm. This algorithm does not impose zeros outside the window, does not require previous coefficients of the autocorrelation function (ACF), and produces a minimum phase operator (Ulrych and Bishop, 1975), (Ulrych and Clayton, 1976), (Marple, 1978), (Barrodale and Erickson, 1980), Porsani (1986), (Zoukaneri, 2014)
In addition to the method for estimating PEO coefficients, Burg established the relationship for predicting the coefficients of autocorrelation with known coefficients $c_{j,i}$. We have:

$$r_z(j) = -\sum_{i=1}^{j-1} r_z(j-i) c(j-1,i) - c(j,j) E_{j-1} \quad (26)$$

The equation 26 guarantees the recursive obtaining of the AC coefficients associated to the maximum entropy spectrum. The Burg method allows to estimate directly j coefficients of the autocorrelation, while in the Yule-Walker method is necessary to know previously the coefficients of autocorrelation to estimate the PEO (Zoukaneri and Porsani, 2015).

Maximum entropy method applied to WVD

The Burg method mentioned above is used to compute the PEO and then using the PEO coefficients to compute and extend the power spectrum of each Wigner-Ville Kernel sequence $K(n), n = 0,...,N_s - 1$.
Each sequence $K(n)$ can be associated with an analytic signal $\tilde{z}(n)$ whose boundaries are associated with window size L.

$$\tilde{z}(n) = \left\{ z(n - \frac{L}{2}),...,z(n),...,z(n + \frac{L}{2}) \right\} \quad (27)$$

The window size L and the number of filter coefficients N_c control the resolution of the decomposition in the FT plane. L is an odd number related to an odd symmetric time window centered on $Z(n)$. We can then write the sequence of the autocorrelation coefficients related to the signal $\tilde{z}(n)$ as:

$$\tilde{K}(n) = \left\{ k(n - \frac{L}{2}),...,k(n),...,k(n + \frac{L}{2}) \right\} \quad (28)$$

The coefficients of the autocorrelation function ACF associated to the maximum entropy spectrum can be obtained using the Burg algorithm. The coefficients are extended as:

$$\tilde{k}_n(j) = -\sum_{i=1}^{j-1} \tilde{k}_n(j-1) c(j-1,i) - c(j,j) E_{j-1} \quad (29)$$

By making the Fourier transform of 29, we can obtain the instantaneous power of the Wigner-Ville distribution. By sliding the window and repeating the process for all kernel sequences, the time-frequency representation of Wigner-Ville maximum entropy is obtained without the influence of the crossed terms.

Extraction of instant attributes from WV-MEM

Average instantaneous frequency

According to Boashash (1992), the average instantaneous frequency can be obtained by computing the first moment of the Wigner-ville distribution, expressed as:

$$\hat{f}(t) = \frac{\int_{-\infty}^{+\infty} f W(t,f) df}{\int_{-\infty}^{+\infty} W(t,f) df} \quad (30)$$

where $\hat{f}$ is the instantaneous average frequency, f is the frequency, $W(f,t)$ is the Wigner-Ville distribution obtained with WV-MEM, and $\int_{-\infty}^{+\infty} W(t,f) df$ corresponds to the time-boundary condition of the Wigner-Ville distribution.
The frequency calculated using the first moment of the MEM-WV is robust, which is why the presence of noise causes low interference to the results (Fomel and Backus, 2003). The robustness evaluation is solved in Zoukaneri (2014).

Variance

The second moment of MEM-WV is related to the local deviation of the frequencies with respect to the average frequency. This deviation is called bandwidth and / or variance, according to Barnes (1993), and is given by:

$$\sigma^2(t) = \frac{\int_{-\infty}^{+\infty} (f - \hat{f}(t))^2 W(t,f) df}{\int_{-\infty}^{+\infty} W(t,f) df} \quad (31)$$

Skewness

The Skewness attribute is related to the third moment of WV-MEM. This attribute describes the deviation of the density function with respect to the normal:

$$S(t) = \frac{\int_{-\infty}^{+\infty}(f-\hat{f}(t))^3 W(t,f)df}{\sigma^3(t)\int_{-\infty}^{+\infty}W(t,f)df} \quad (32)$$

Kurtosis

The Kurtosis attribute is related to the fourth moment of WV-MEM and reflects how close the distribution is to a delta. It is given by:

$$K(t) = \frac{\int_{-\infty}^{+\infty}(f-\hat{f}(t))^4 W(t,f)df}{\sigma^4(t)\int_{-\infty}^{+\infty}W(t,f)df} - 3 \quad (33)$$

According to Steeghs and Drijkoningen (2001), the term -3 causes $K(t)$ to assume zero value for the case of normal distribution.

Results

The WV-MEM method and its relation to the window size L and the PEO operator order

In the previously reported information, the method resolution (WV-MEM) is controlled by the window size L and the number of coefficients Nc of the PEO operator. Now, we will understand how the resolution is conditioned by the effect of the operator order and the size of the window.

Wigner-Ville method and the effect of the window size L

It is possible to observe the effect of window size in Figure 2, that shows a synthetic time-frequency seismic trace using the Wigner-Ville method for different window sizes L. Through a thorough observation, it can be concluded that large windows make the surface of the energy smooth while small windows allow the observation of small details. Of course, to characterize the stratigraphy, we should use short size windows. This is contrary to what is observed in traditional methods such as the short-window Fourier Transform (SWFT), where a larger window is needed to achieve a frequency resolution. The explanation is that the SWFT method aims to solve a non-stationary problem based on the window size, and on the proposed approach; besides, it becomes the non-stationary problem in small stationary problems. In this way, the short window is necessary to achieve better solution. Each short window is characterized by its center frequency (average), such that, in MEM-WV, only the center frequency and its associated energy for each window considered (Zoukaneri, 2014) are represented.

Wigner-Ville method and the PEO effect

It is possible to observe how the order of the PEO operator in Figure 3 influences the time-frequency spectrum. However, it is clear that the operator of the order Nc = 1 is sufficient. According to Zoukaneri (2014), the operator of order $Nc = 1$ corresponds to a PEO of single reflection coefficient $(1, c_{1,1})$ and can be associated with the single wave plane propagating indefinitely with center frequency equal to the average frequency of the wave. In this way,

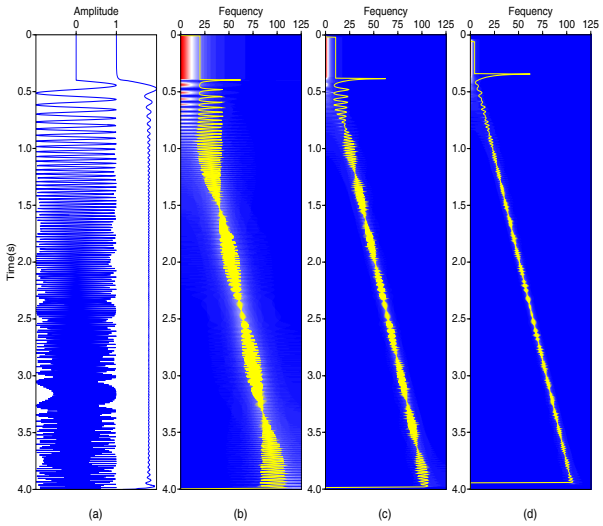

Figure 2: (A) Sinusoidal signal and the corresponding envelope (B) Time-frequency decomposition using window equal to 5, (c) DTF using window equal to 15 and (d) using window equal to 25

the WV-MEM spectrum with $Nc = 1$ represents the energy distribution around the instantaneous average frequency. Similarly, the operator of order $Nc = 3$ $(1, c_{1,1}, c_{1,2})$ is associated with two wave planes, where each wave plane in the Wigner- Ville spectrum is represented by the energy distribution around the mean frequency corresponding to the wave plane.

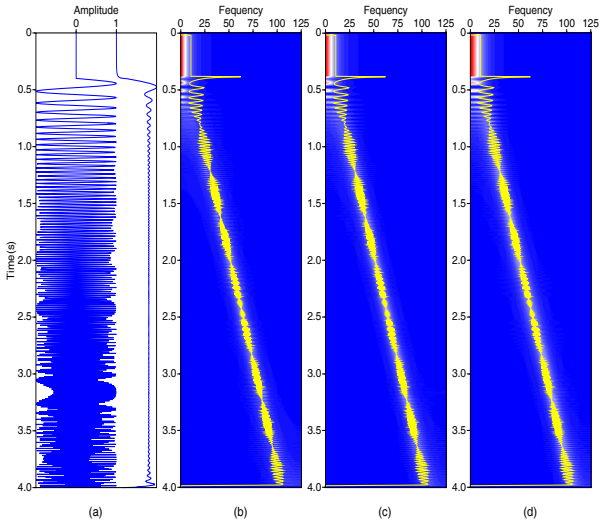

Figure 3: (A) Sine signal and the corresponding envelope, (b) Time-frequency decomposition using $N_c = 2$ (c) DTF using $N_c = 3$ (d) used $N_c = 4$

According to Zoukaneri (2014), the choice of the order is of great importance to identify the components that a signal carries, especially for multicomponent signals and in noise environments. The best choice for the order of the model N_c is not usually known at the beginning, so it is necessary to perform experiments with several orders. Two of these criteria are the final prediction error (FPE) described by

Akaike (1969) and the maximum likelihood approach (AIC). However, for applications in geophysics, where it is desired to analyze the energy variation of the signal, an operator of order $N_c = 1$ is sufficient.
Another transparent observation in the Figure 3, which justifies the used method, concerns the instantaneous average frequency curve. Note that for different values of N_c the instantaneous average frequency curve remains constant, affirming the robustness of the method in noisy environments.

Extraction of Attributes and Characterization of Reservoirs in Real Data

Most of the extracted attributes were obtained using the WV-MEM spectrum in post-stacked data. As a geomorphological attribute, we have the instantaneous average frequency presented in Figure 5, extracted from the Figure 4. The vertical resolution of the spectral section is observed due to the solution of the WV-MEM method. Thus, it is possible to observe a decrease in the frequencies as the depth increases and more important regions of low frequencies indicating a probable accumulation of hydrocarbons in the time of seven seconds for the CMPs 300 and 600.
One of the direct indicators of hydrocarbons is the low frequency anomaly, since low frequency energy is highlighted below the reservoirs. Spectral decomposition has been used to study the energy variation with frequency in the data.
Figure 6 refers to the variant attribute and corresponds to local deviation and frequency with respect to average frequency. Note that this attribute also reaffirms the reservoir for CMPs 300 and 600 in the time of seven seconds.
For the same CMPs and time, the attributes Skewness, Figure 7, and Kuttosis, Figure 8, guarantee information related to the reservoirs. Note in Figure 7 that the region of the possible reservoir in the Skewness attribute has a high degree of asymmetry. Figure 8 corresponds to the Kurtosis seismic attribute. Observe that the attribute is efficient in the identification of the reservoir, once it reflects the pulse distortion due to the presence of fluid in the region (note lower values when compared to the host medium).
Another interesting attribute is the energy of the Prediction Error, Figure 9, which represents how close the calculated value is to the predicted data.

Conclusion

The Wigner-Ville maximum entropy distribution was obtained using the Burg method to extend the Wigner-Ville kernel sequences using the prediction error operator and applying to each extended Kernel sequence the Fourier transform.
The solution of the method is controlled by the operator order N_c and the size of the window used to estimate the prediction operator. The number of components of the spectrum is directly related to the number of coefficients of the predictive error operator. In the signal analysis, the $N_c = 2$ proved to be satisfactory. Small window allows to capture details of the data while a large window manifests the soft spectrum.
From the WV-MEM spectrum, it was possible to obtain the attributes instantaneous average frequency, mean variance, Skewness and Kurtosis attributes with high resolution for the characterization of reservoirs.

Acknowledgements

This research was supported by CNPq and INCT-GP/CNPq. The facility support from CPGG/UFBA is also acknowledged.

References

Akaike, H., 1969, Fitting autoregressive models for prediction: Annals of the institute of Statistical Mathematics, **21**, 243–247.
Barnes, A. E., 1993, Instantaneous spectral bandwidth and dominant frequency with applications to seismic reflection data: Geophysics, **58**, 419–428.
Barrodale, I. and R. Erickson, 1980, Algorithms for least-squares linear prediction and maximum entropy spectral analysisâpart i: Theory: Geophysics, **45**, 420–432.
Boashash, B., 1992, Estimating and interpreting the instantaneous frequency of a signal. i. fundamentals: Proceedings of the IEEE, **80**, 520–538.
Burg, J. P., 1967, Maximum entropy spectral analysis.: Presented at the 37th Annual International Meeting.
Fomel, S. and M. M. Backus, 2003, Multicomponent seismic data registration by least squares, *in* SEG Technical Program Expanded Abstracts 2003, 781–784, Society of Exploration Geophysicists.
Leite, L. W., 2015, Conceitos da anÃ¡lise espectral de sinais em geofÃsica: Presented at the .
Marple, S., 1978, Frequency resolution of high resolution spectrum analysis technique: Proceedings of the 1st RADC Spectrum Estimation Workshop, 19–35.
Pimentel, E. A. and J. F. da Silva, 2011, Decomposição de ondaletas, análise de volatilidade e correlação para índices financeiros: Estudos Econômicos (São Paulo), **41**, 441–462.
Porsani, M. J., 1986, Desenvolvimento de algoritmos tipo-levinson para o processamento de dados sismicos: Ph. D. Dissertation, Universidade Federal da Bahia, Salvador, Brasil.
Soares, L. B., 2013, Proposta e avaliação de técnicas para compressão de transitórios rápidos e análise tempo-frequência de distúrbios em redes elétricas ac.
Steeghs, P. and G. Drijkoningen, 2001, Seismic sequence analysis and attribute extraction using quadratic time-frequency representations: Geophysics, **66**, 1947–1959.
Ulrych, T. J. and T. N. Bishop, 1975, Maximum entropy spectral analysis and autoregressive decomposition: Reviews of Geophysics, **13**, 183–200.
Ulrych, T. J. and R. W. Clayton, 1976, Time series modelling and maximum entropy: Physics of the Earth and Planetary Interiors, **12**, 188–200.
Zoukaneri, I., 2014, Análise tempo-frequência do sinal sísmico utilizando a distribuição wigner-ville e o método de máxima entropia: Aplicada para a estimativa do fator q e de atributos: Tese de Doutorado.
Zoukaneri, I. and M. J. Porsani, 2015, A combined wigner-ville and maximum entropy method for high-resolution time-frequency analysis of seismic data: Geophysics, **80**, O1–O11.

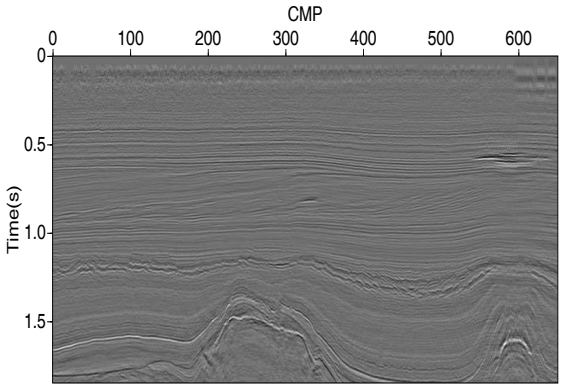

Figure 4: Stacked opendtect seismic section,

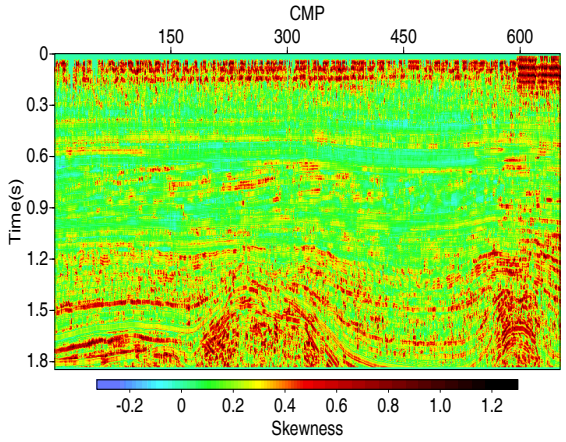

Figure 7: Seismic attribute Skewness, Nc=1 e L=7

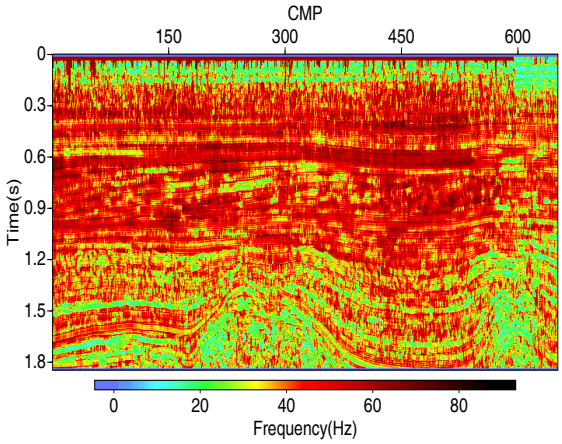

Figure 5: Attribute average instantaneous frequency, Nc=1 e L=7

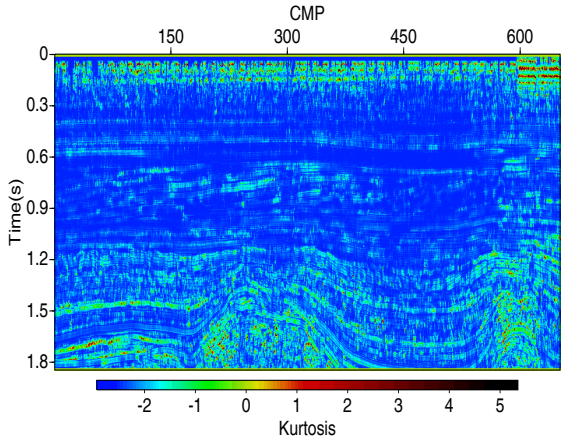

Figure 8: Seismic attribute Kurtosis, Nc=1 e L=7

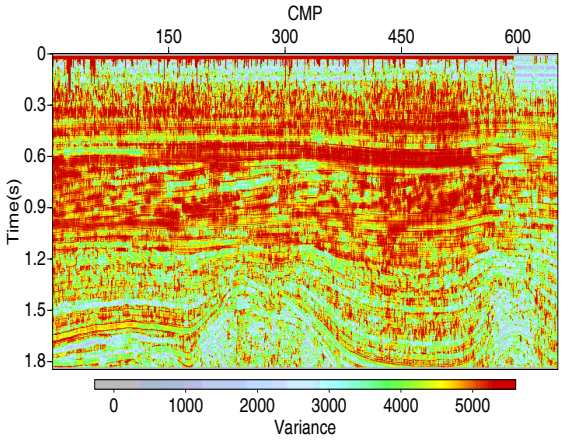

Figure 6: Attribute mean variance, Nc=1 e L=7

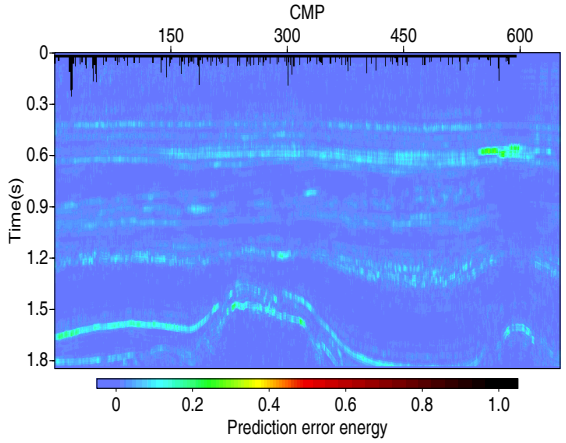

Figure 9: Seismic attribute energy prediction error., Nc=1 e L=7