Matching pursuit spectral decomposition of seismic data

By Let it Wave.
For BHP Billiton.

July 2007

Contents

List of Figures

Introduction

This report presents Let It Wave's matching pursuit tool, adapted to seismic data analysis. This is an intermediate report, that details the progress made so far.

Matching pursuit is a spectral decomposition method, which provides a high-resolution and more flexibility than other traditional methods like the Continuous Wavelet Transform (CWT) or the Short-Time Fourier Transform (STFT). The principle is to represent a signal with a certain number of relevant atoms, waveforms characterized by their position in time, scale and modulation frequency. Our goal is to build such a decomposition adapted to seismic data and analysis, such that it could be understood in a global approach, and extract relevant geophysical attributes

We will first present the theory of matching pursuit in further detail, and that of the Wigner-Ville transform, a classic display commonly associated. Then we will go in further to explain the changes and adjustments made to fit this technique to the specific case of seismic data, and will conclude with our results so far displayed with two different methods to try and emphasize the information provided by the matching pursuit.

1 The matching pursuit spectral decomposition

Matching pursuit (later referred to as MP) is an algorithm that was first introduced by Stephane Mallat (see [2]). It provides a spectral decomposition of a signal in a very flexible and efficient manner. The signal is thus represented by a series of 'atoms' and associated coefficients, which are Gabor functions, and their weights. Let us present this decomposition in further details.

1.1 Gabor transform

The basis for the matching pursuit algorithm is what is commonly referred to as the Gabor transform. Its principle is similar to a wavelet transform, only with one more degree of liberty. Indeed, the Gabor transform is the collection of coefficients resulting from inner products of a signal and a family of waveforms. However, when the CWT family is generated from a mother wavelet, dilated and translated for the wavelet transform, in this case the family is generated from a mother waveform, dilated, translated *and* modulated.

The initial waveform chosen is usually a suitably normalized Gaussian that will be called $g(t)$.

$$g(t) = \frac{1}{\sqrt{2}} \exp[-2\pi t^2] \tag{1}$$

An atom $g_{(u,s,\xi)}$ is now generated from this function and can be identified by three parameters:

- position u
- scale s or variance $\sigma = 2^s$
- modulation frequency ξ

From now on this triplet will be called $\gamma = (u, s, \xi)$ It is defined as,

$$g_\gamma(t) = K_\gamma \exp\left[-2\pi \left(\frac{t-u}{2^s}\right)^2\right] e^{-i\xi t} \tag{2}$$

Various examples of atoms, including the unmodulated Gabor functions are represented in figure 1. The main difference with the CWT to notice, is how scale and frequency can be independently chosen to characterize an atom. Indeed, the figure displays two atoms of same scale but different frequencies (second and third atoms), likewise there are two atoms modulated at the same frequency but with different scales (third and fourth atoms). Also, the second atom from the left has a small scale and a low frequency, while the fourth atom has a large scale and high frequency, both configurations that don't exit in a classic continuous wavelet transform.

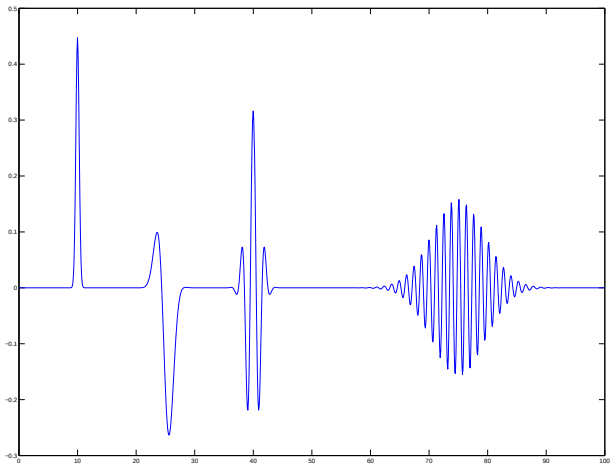


Figure 1: Various Gabor atoms. From left to right: the Gabor function, scale=0, frequency=0; scale=2 frequency=$\frac{8\omega_N}{30}$; scale=2 frequency=ω_N; scale=4 frequency=ω_N, where ω_N is the Nyquist frequency .

Note that in this case the family of waveforms will be redundant for the purpose of the pursuit. Nevertheless, a reconstruction of signal f from the transform can be written as:

$$f = \sum_{n=0}^{+\infty} \langle R^n f, g_{\gamma_n} \rangle g_{\gamma_n} \tag{3}$$

Of course, in the cases of matching pursuit and continuous wavelet transform alike, atoms are restricted by the Heisenberg inequality in terms of shape and spectral support(see [1]). Heisenberg's inequality states that the variance in time σ_t^2 and the variance in frequency σ_ω^2 of a function $f \in \mathbf{L}(\mathbb{R}^2)$ verify:

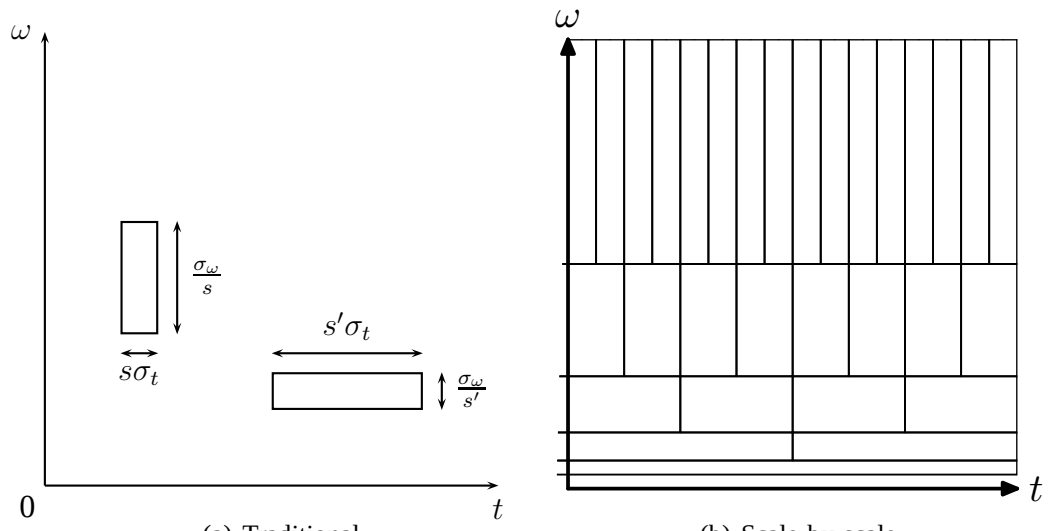
$$\sigma_t^2 \sigma_\omega^2 \geq \frac{1}{4}$$

it means that a function can't be both perfectly localized in time and frequency. This is better represented in figure 2. On the right is the tessellation of the spectral domain by the CWT atoms (as used by BHP-Billiton), and on the left two boxes showing potential atoms' support for the matching pursuit.

Note. It is essential to keep in mind that, though similar in approach, the CWT and the Gabor transform lead to very different representations and have to be thought of as two separate concepts. In particular, it is important to make the distinction between scale and frequency during the analysis of a Gabor transform or a matching pursuit.

1.2 Pursuit mechanism

The matching pursuit goes farther than the Gabor transform. Its objective is to select atoms that best describe the signal among all possible waveforms. The process is iterative and lead to a very efficient representation of the signal. Starting from the Gabor transform, the principle



(a) Traditional (b) Scale-by-scale

Figure 2: Those figures are intended to show the tessellation of the time-frequency plane by atoms. On the left, is the representation of two atoms, constrained by the Heisenberg inequality, the width of an atom determines it height. In the case of the Gabor transform, atoms can be positioned wherever in this plane, as long as respect this inequality. However, in the case of the CWT, atoms are positioned in a q-determined fashion as can be seen on the right figure. Of course, Gabor transform atoms can assume all the positions of that of the CWT and more.

is at each step to select the atom that best matches the signal, then subtract it from the signal, and start again. This requires to know the Gabor transform coefficients of each modified signal. Because computing the remainder, then its Gabor transform, at each step would be very costly, the trick is to compute the Gabor transform only once, then deduce the new coefficients for each step from the previous one, we will refer to this step as coefficients updating.

The pursuit can be summed up in a few step. Given the Gabor transform defined by equation 3. Define g_{γ_0}, such that:

$$\langle f, g_{\gamma_0} \rangle = \max_{k \in \mathbb{N}} \langle f, g_{\gamma_k} \rangle$$

Then, if $R^0 f$ is the remainder of a signal minus its most important atom:

$$R^0 f = f - \langle f, g_{\gamma_0} \rangle g_{\gamma_0} \tag{4}$$

From there, the process has to be repeated on the remainder: take its transform, then find the maximum and subtract, and so on, recursively. However, computing once again a Gabor transform is out of question. To avoid the computational cost, just note that from equation 4, we can derive:

$$\langle R^0 f, g_\gamma \rangle = \langle f, g_\gamma \rangle - \langle f, g_{\gamma_0} \rangle \langle g_{\gamma_0}, g_\gamma \rangle \tag{5}$$

and recursively, we obtain the equation for updating inner products:

$$\langle R^{n+1} f, g_\gamma \rangle = \langle R^n f, g_\gamma \rangle - \langle R^n f, g_{\gamma_n} \rangle \langle g_{\gamma_n}, g_\gamma \rangle \tag{6}$$

Note. About real and complex atoms:

So far the atoms used in this process were all complex valued. Of course, in terms of display and analysis, this is neither practical, nor realistic. The atoms we are working with eventually are real atoms obtained from the usual combination of a complex atom and its complex conjugate. Let $g_{\gamma,\phi}$ be the real atom, then:

$$g_{\gamma,\phi} = K_{\gamma,\phi}(e^{i\phi} g_\gamma + e^{-i\phi} g_\gamma^*) \tag{7}$$

(Note that $g^*_{(u,s,\xi)} = g_{(u,s,-\xi)}$)

Even though most intermediate computations are made using complex atoms, the resulting process is kept in the real domain, by subtracting a real atom from the signal at each step. Hence the true updating equation is:

$$\langle R^{n+1} f, g_\gamma \rangle = \langle R^n f, g_\gamma \rangle - \langle R^n f, g_{\gamma_n,\phi} \rangle \langle g_{\gamma_n,\phi}, g_\gamma \rangle \tag{8}$$

1.3 Matching pursuit lexicon

Matching pursuit comes with a small lexicon very similar to that of coding. This is just a small summary to clarify the terms employed in this report and our softwares.

- **Atom** : Any waveform obtained from the original Gaussian through translation, dilatation and modulation, and following equation 2.
- **Dictionary** : The collection of all atoms considered in the pursuit, from which the matched atoms can be selected.
- **Book** : This is the output of the matching pursuit, it is the collection of selected atoms paired with corresponding coefficient.

1.4 Wigner-Ville transform

In praxis, the output of the matching pursuit is a collection of atoms defined by three parameters, position, scale and frequency. This means that a basic approach would use a three-dimensional object to describe the decomposition of a one-dimensional signal, like a single trace. We are aiming for (and have designed) a same-D representation, however, a common solution was devised by J.Ville in [3], following the work of E.P. Wigner ([4]), with the Wigner-Ville transform. It consists in representing the addition of all atoms Fourier transform. The Wigner-Ville distribution of a function f is defined as follows:

$$\begin{aligned} WVf(u,\xi) &= \int_{-\infty}^{+\infty} f\left(u + \frac{\tau}{2}\right) f^*\left(u - \frac{\tau}{2}\right) e^{-i\tau\xi} d\tau \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}\left(\xi + \frac{t}{2}\right) \hat{f}^*\left(\xi - \frac{t}{2}\right) e^{-itu} dt \end{aligned} \tag{9}$$

One advantage of this distribution is that it doesn't cause spreading of atoms support, i.e. the time support of a Dirac function is still a Dirac function. In our particular case, the distribution becomes:

$$WVf[n,k] = \sum_{m=0}^{+\infty} |\langle R^m f, g_{\gamma_m,\phi}\rangle|^2 WVg_\gamma[n,k] \tag{10}$$

where $WVg_\gamma[n,k]$ is a bi-dimensional Gaussian centered in (u_m, ξ_m) and dilated by 2^{s_m} in time and 2^{-s_m} in frequency. Figure 3 shows the Wigner-Ville representation of a matching pursuit completed on the signal shown in figure 1.

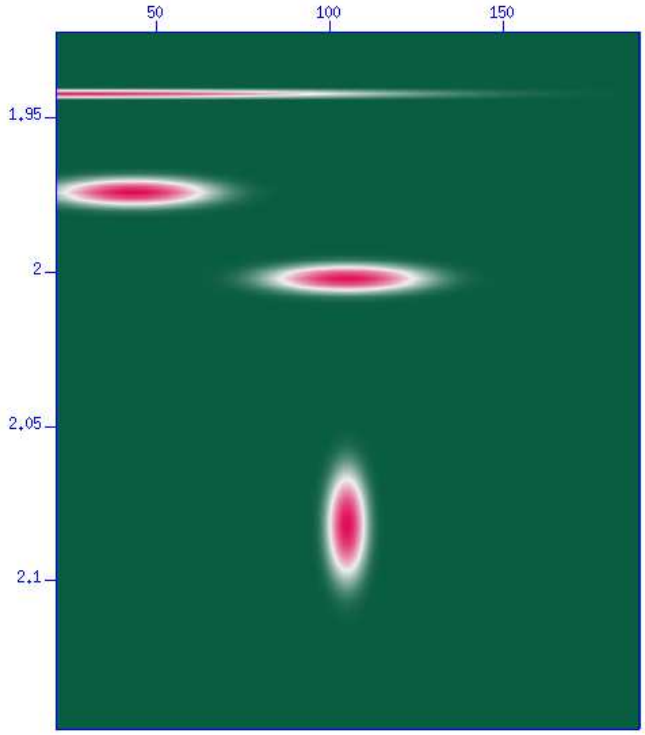


Figure 3: A simple example of the Wigner-Ville representation of a matching pursuit on the signal displayed in figure 1. All atoms were chosen to have the same energy, hence their similar colors. The horizontal axis is the frequency axis, the vertical one represent time. The shape of atoms is determined by their scale, notice how the first atom, scale 0 is very elongated in the horizontal direction (like a Dirac) and how the last atom, scale 4 is elongated in the other direction.

2 Scale-by-scale matching pursuit

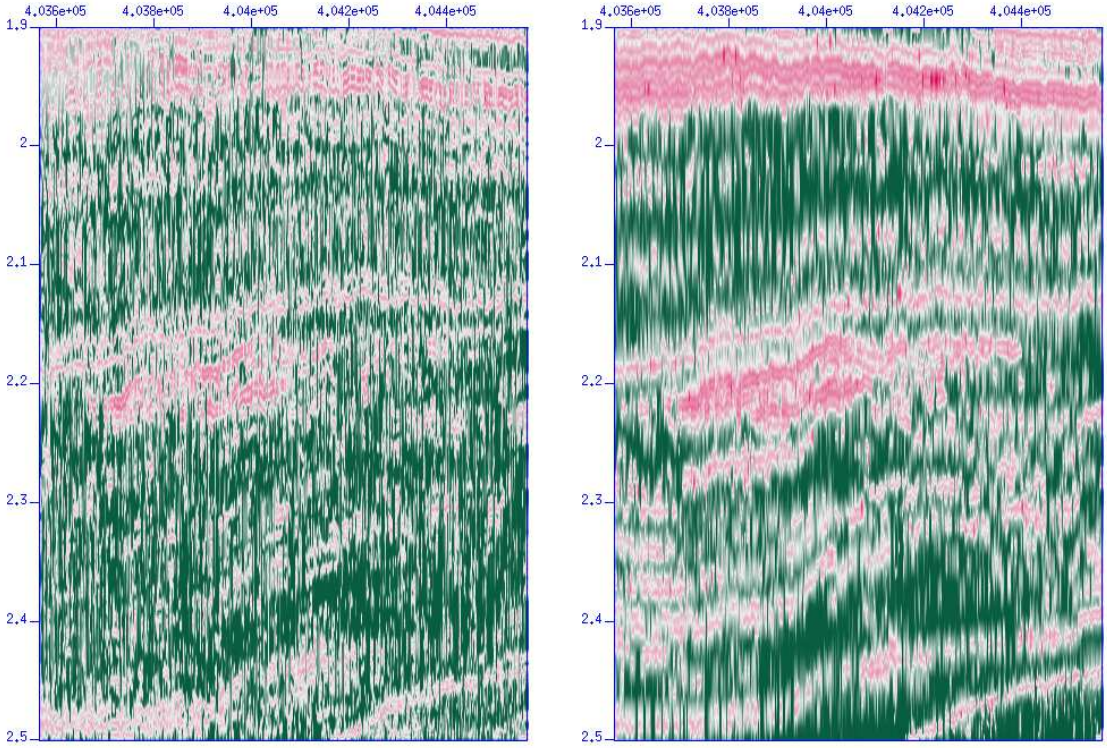
The original matching pursuit algorithm wasn't particularly intended for seismic data analysis. Our mains concerns here were the fact that we needed to ensure spatial continuity among traces for coherent interpretation and that we needed to separate the influence of the measuring wavelet from the real signal information since our data are convolved. We have hence brought some changes to the traditional algorithms and developed some other tools to suit our needs.

2.1 Limits of regular matching pursuit

As mentioned previously, spatial continuity is a prerequisite to any coherent analysis on seismic data.

Matching pursuit picks out atoms best suited to describe structures. Hence the choice of an atom affects the rest of the process, creating a new residual signal, basis for the following steps. Also, if slightly altered, a structure may be, for example, best fitted by a large atom instead of a series of smaller ones. Because of all this, the result of the decomposition may be different from one trace to the other. Of course, it won't be altogether different, and the discrepancies usually concerns larger scale modulated atoms which are harder to fit. To alleviate this problem, one solution consists in doing 'scale-by-scale' matching pursuit. The principle is to compute the pursuit several time on sub-dictionaries, each of them containing atoms corresponding to one fixed scale. Thus, we can ensure to control atoms at each scale, which provides lateral continuity from one trace to the other.

Figure 4 shows the differences between a regular matching pursuit and a scale-by-scale treatment, on Scarborough's reservoir, at a fixed frequency. Imposing a scale by scale analysis, makes sure that there is an even number of atoms at each scale.



(a) Traditional (b) Scale-by-scale

Figure 4: Figure on the left shows the Wigner-Ville of a matching pursuit made on Scarborough's reservoir (at fixed zero frequency). The computation is made on scale 0 to 5 and with 50 atoms per trace. On the right, the same Wigner-Ville on the matching pursuit scale-by-scale of the same data set. Atoms are still between scale 0 to 5, and with 20 atoms per trace and per scale.

2.2 Algorithm

This is the algorithm implemented for the scale-by-scale matching pursuit in `liwMPS` only. It is a scale-by-scale algorithm and represents only the skeleton of our process.

For each trace and each scale:

1. Compute the Gabor transform (over the appropriate dictionary)
2. Find the max over all coefficient and the corresponding atom g_{γ_0}
3. Add g_{γ_0} to the book
4. From (n) to $(n+1)$
   (a) Update the dictionary
   (b) Update coefficients
   (c) Determine the new max over all coefficients, and corresponding atom g_{γ_n}
   (d) Add g_{γ_n} to the book
   (e) Test the energy corresponding to the word, if it is less than the threshold, stop, otherwise start step (4) again

2.3 Examples and displays

Here our some examples of what this software can do. Following are representations of decompositions of reflectivity. Reflectivities are the best signal to appraise the matching pursuit since they are the only signal free for the interference of the measuring wavelet.

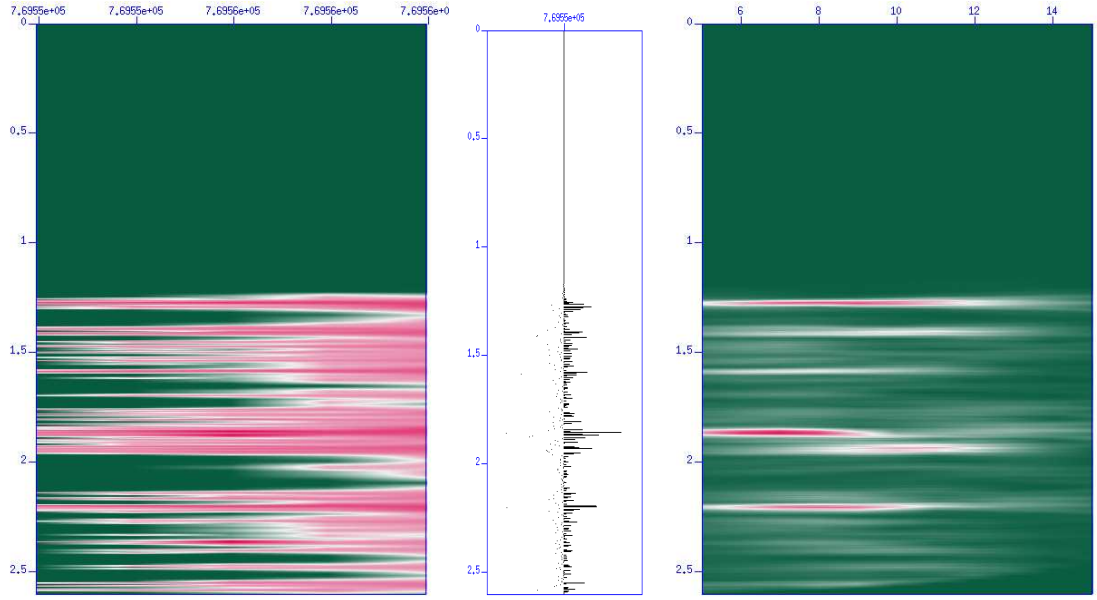
The matching pursuit is a technique that provides a wealth of information on the signal with a decomposition on position, scale and frequency. The traditional display method is the Wigner-Ville display as mentioned earlier. However, this display is hardly comparable to that of a CWT. Hence, LIW has designed a simplified display, showing the spread of atoms according to their position and scale, the later referred to as, scale-by-scale display. It is computed with `liwSScaleDisplay`. Figure 5 shows the matching pursuit of a trace compared to its CWT decomposition.

For a thorough analysis, the Wigner-Ville display is however more recommended. Below in figure 6 examples of what can be done using BHPViewer to take fully advantage of it.

Both `liwSScaleDisplay` and `liwWVS` include options to filter out data along scale and frequency.

3 Normalization

Common seismic signals result from geological measures and hence have been convolved by the measuring wavelet. This convolution 'taints' the result shown by any spectral decomposition if no additional processing is performed. It is essential to be able to recover the original information present in the signal and to discriminate it from information of the measuring wavelet. Two softwares have been developed in order to represent the matching pursuit as it would be before convolution.



(a) Matching Pursuit (b) Seismic trace (c) CWT

Figure 5: Comparison between a matching pursuit and a CWT. Left the matching pursuit decomposition of trace from Scarborough in a time vs scale display, center, said trace and right the CWT after smoothing.

3.1 Main ideas

The problem of deconvolution on seismic data is a very delicate, non-trivial one. First of all, it is extremely difficult to extract the measuring wavelet from data or plan its behavior. That means that it is impossible to just 'inverse' the filter. With next to no reliable information on the wavelet, we need to perform that is generally referred to as a 'blind deconvolution'.

The principle of such a method, is to try to find a way to gather a model of the filter directly from the observed data, than to try and inverse it. In order to explain this procedure, let us go back to classic representation of a matching pursuit.

The matching pursuit decomposes a seismic signal S as a sum of Gabor atoms g_γ adaptively chosen :

$$S = \sum_\gamma c_\gamma \, g_\gamma \quad .$$

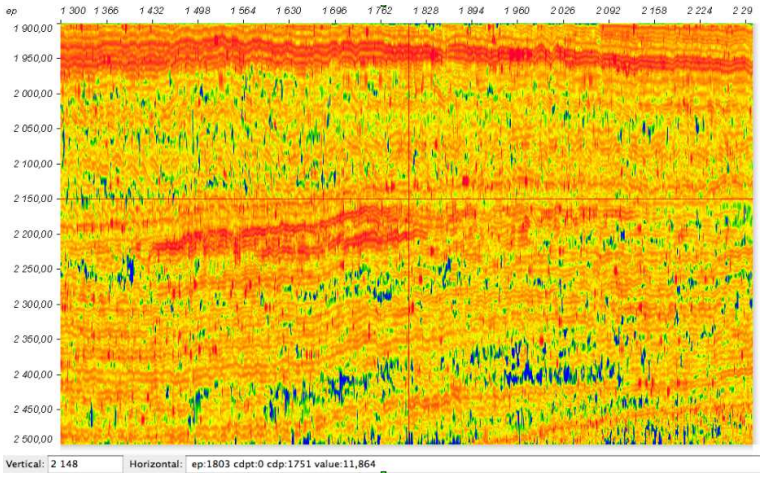
Its energetic representation is obtained by summing the energy of the different atoms :

$$E(S) = \sum_\gamma |c_\gamma|^2 \, E(g_\gamma) \quad .$$

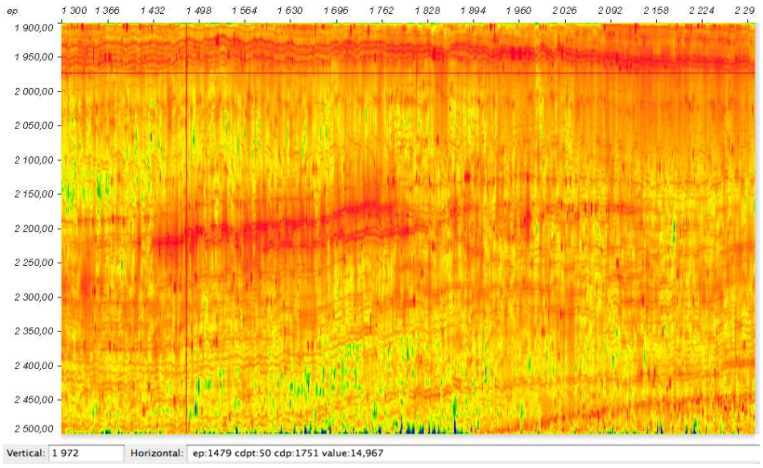
where $E(g)$ is either the Wigner-Ville energy spectrum of the atom of its scale by scale spectrum representation.

The seismic signal S is the convolution of the reflectivity R with a seismic wavelet Ψ (we will for now overlook any additive noise) and thus

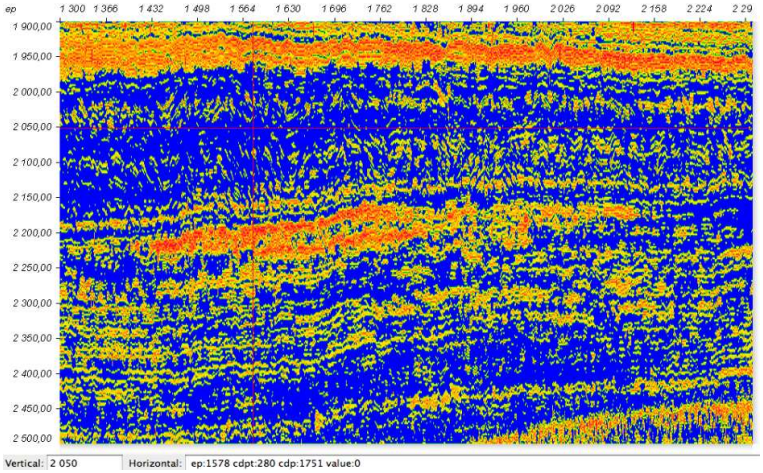
$$R \star \Psi = S = \sum_\gamma c_\gamma \, g_\gamma \quad .$$



(a) Frequency=0



(b) Frequency=$\frac{\omega_N}{6}$



(c) Frequency=ω_N

Figure 6: Three slices of a Wigner-Ville visualized with BHP Viewer at three different frequencies, $0, \frac{\omega_N}{6}$, ω_N, where ω_N is once the again the Nyquist frequency. The original dataset is once again a close-up on Scarborough's reservoir, the exact same one that was displayed in figure 4

If the wavelet Ψ was known, the reflectivity could be obtained from the seismic trace through a simple convolution with Ψ^{-1} :

$$R = R \star \Psi \star \Psi^{-1} = S \star \Psi^{-1} = \sum_{\gamma} c_\gamma \, g_\gamma \star \Psi^{-1} \quad .$$

The energetic representation of R would thus be approximated by :

$$E(R) = \sum_{\gamma} |c_\gamma|^2 \, E(g_\gamma \star \Psi^{-1}) \quad .$$

As stated before, this is an ideal representation, one it is very difficult to have access to, but one that can be approximated.

To stabilize the display with respect to the wavelet, we choose to approximate $E(g_\gamma \star \Psi^{-1})$ by $\lambda_\gamma^2 E(g_\gamma)$. This is equivalent to saying that the Gabor atoms are approximate eigenfunctions of the convolution operator. This is very similar to the underlying technique used in BHP's renormalization for the CWT.

The main issue lies in the choice of the parameters λ_γ : only the seismic trace S is observed and the wavelet Ψ is unknown. Fortunately, the classical hypothesis that the reflectivity R follows the white noise model gives a way to estimate this λ_γ. This hypothesis can be reinforced by considering a spatial average of the reflectivity instead of the reflectivity.

Indeed for any γ, a local average in time yields over the $g_{\gamma'}$ that are translate of g_γ

$$\mathbb{E}(|\langle S, g_{\gamma'} \rangle|^2) = \mathbb{E}(|\langle R \star \Psi, g_{\gamma'} \rangle|^2) = \mathbb{E}(|\langle R, g_{\gamma'} \star \tilde{\Psi} \rangle|^2)$$

with $\tilde{\Psi}$ the conjugate reverse of Ψ.

Note now that if R is a white noise, this quantity is a good estimate of $\|g_\gamma \star \Psi\|_2^2$ up to a constant factor. As the power spectrum of a white noise is constant, this is indeed the variance of $\langle R, g_\gamma \star \Psi \rangle$ which is the same as $\langle R, g_\gamma \star \tilde{\Psi} \rangle$.

Assume $g_\gamma \star \Psi^{-1} = \lambda_\gamma g_\gamma$, which is true for any eigenvalue of Ψ and supposed approximately true for g_γ, then $g_\gamma \star \Psi = \frac{1}{\lambda_\gamma} g_\gamma$,

$$\|g_\gamma \star \Psi\|^2 = \frac{1}{\lambda_\gamma^2} \|g_\gamma\|^2 \quad ,$$

and as the g_γ are normalized :

$$\|g_\gamma \star \Psi\|^2 = \frac{1}{\lambda_\gamma^2} \quad .$$

$\mathbb{E}(|\langle S, g_{\gamma'} \rangle|^2)$ is thus a good candidate estimate for $\frac{1}{\lambda_\gamma^2}$.

The proposed renormalization is then obtained by the following formula

$$E(R) = \sum_{\gamma} \frac{|c_\gamma|^2}{\mathbb{E}(|\langle \tilde{S}, g_{\gamma'} \rangle|^2)} E(g_\gamma)$$

where $\tilde{S}$ is a spatial average of the seismic traces and $\mathbb{E}$ is a local average over the Gabor atoms that are translates of g_γ.

One should note that using a local time average instead of a global time average makes the whole process robust to a deformation of the seismic wavelet along the trace.

3.2 Implementation : liwmpcoeffs and liwmpnorm

The normalization is implemented in two steps. First, coefficients $\mathbb{E}(|\langle \tilde{S}, g_{\gamma'} \rangle|^2)$ are computed over a dataset, then each trace is normalized using those coefficients. Because coefficients only need to be computed once over a dataset and not for each trace, the two processes were separated into two softwares for more efficiency.

Let us define α_γ as the correction coefficients for each atom g_γ. The α_γ coefficients are computed by `liwmpcoeffs` over a whole block of traces, then, it is a simple matter performed by `liwmpnorm` to normalize the book, the new coefficients c'_γ are obtained simply by multiplying each coefficient:

$$c'_\gamma = \alpha_\gamma * c_\gamma$$

If there were no additional noise, the correction coefficients would be:

$$\alpha_\gamma = \frac{1}{\sqrt{\mathbb{E}(|\langle \tilde{S}, g_{\gamma'} \rangle|^2)}}$$

However, since it is a rather unreasonable assumption, we can estimate the noise level over the signal and represent by its variance σ^2. The corrections coefficients then become:

$$\alpha_\gamma = \frac{\sqrt{\mathbb{E}(|\langle \tilde{S}, g_{\gamma'} \rangle|^2) - \sigma^2}}{\mathbb{E}(|\langle \tilde{S}, g_{\gamma'} \rangle|^2)}$$

Finally, because the previous expression is not always defined, we will introduce parameter ϵ and rewrite the coefficients in their final form as follows.

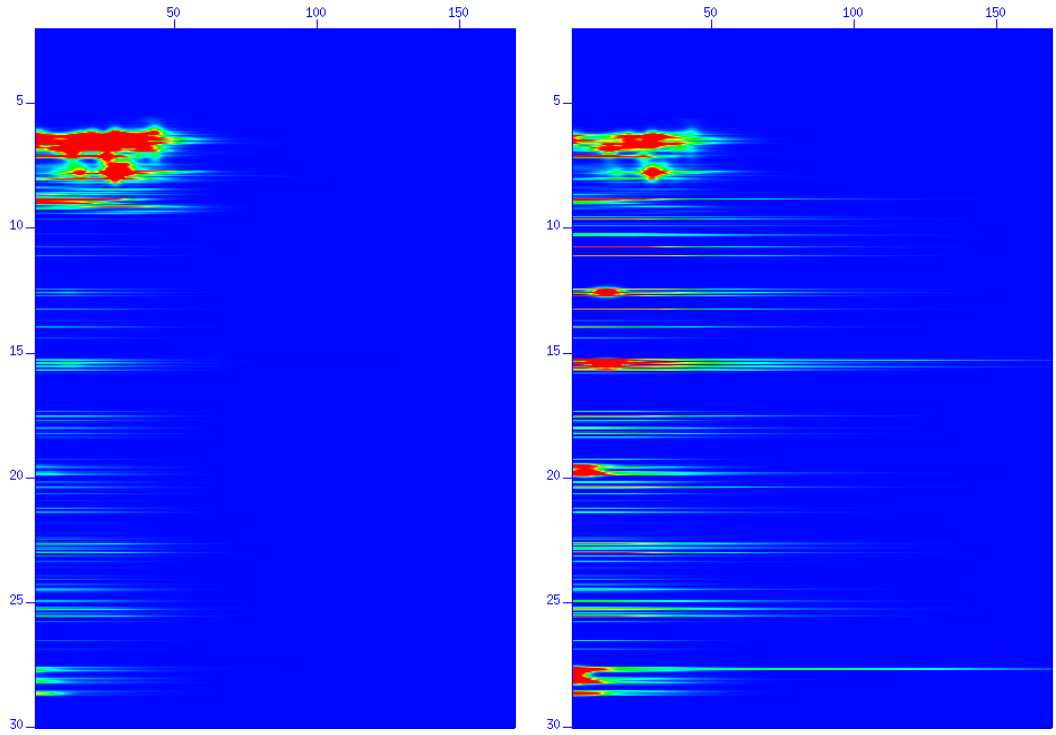
$$\alpha_\gamma = \min \left(\frac{\sqrt{\max(\mathbb{E}(|\langle \tilde{S}, g_{\gamma'} \rangle|^2) - \sigma^2, 10^{-6})}}{\mathbb{E}(|\langle \tilde{S}, g_{\gamma'} \rangle|^2)}, \frac{1}{\epsilon} \right)$$

3.3 Pre-normalization

The normalization is performed on the book of atoms and not the signal itself. An important advantage of this method is that while there is always a risk of amplifying noise while performing a deconvolution, the risk is minimum here as the matching pursuit performs inherent denoising. In other words, the book of atoms is devoid of noise, hence normalization should induce no noisy phenomenon.

However, 'post-normalization' has some drawbacks. After convolution with the wavelet, atoms are likely to be slightly shifted if not always in position, at least in frequency (scale is relatively stable). The post-renormalization can put emphasis on some atoms or decrease the importance of others, but it won't change their coordinates. It would be really good to be able to normalize the convoluted signal directly to reveal the 'real' atoms. However, given the non negligible presence of noise in seismic signal, this strategy causes great difficulty as the dreaded phenomenon of noise amplification occurs.

The solution is to pre-normalize, but 'not too much'. One way to realize that is to use the ϵ parameter defined in section 3.2. This parameter was introduced to avoid divide-by-zero cases,



(a) Before normalization (b) After normalization

Figure 7: On the left, the Wigner-Ville representation of the matching pursuit decomposition of a seismic trace taken from Neptune. On the right, that of the normalized book of atoms.

but it can be used to moderate amplification through normalization. Indeed, only atoms associated with α_γ coefficients greater than 1 are amplified. Obviously this is approximately realized when $\max(\mathbb{E}(|\langle \tilde{S}, g_{\gamma'} \rangle|^2), \epsilon) < 1$, thus one way to ensure that there will be no amplification is to set ϵ to 0. This causes the pre-normalization process to eliminate some atoms by reducing their weight, and will not amplify noise.

In praxis, it is not necessary to set ϵ to 0, values like 0.1 or 0.01 do a reasonable job and generally do not cause unwanted noise amplification. However, since this is performing a 'full' normalization it is best to combine it with a pass of post normalization, applied with a much smaller value of ϵ.

4 Inverse transform

The Gabor transform can be inverted, we can reconstruct the signal from the book of atoms. This provides many interesting applications. The most simple being de-noising, by selecting only relevant atoms the matching pursuit gets rid of the noise contained in the signal, the reconstructed signal will hence be free of noise. Others applications come from filtering. It is possible to process or filter the book atoms before reconstructing for analysis purpose. Some basic functions are implemented in our software since it is possible to filter atoms according to their scale or frequency.

4.1 Reconstruction

Let us go back on the general equation of the regular matching pursuit.

$$f = \sum_{n=0}^{+\infty} \langle R^n f, g_{\gamma_n} \rangle g_{\gamma_n} \tag{11}$$

After N iterations of the algorithm, we can partially reconstruct the $\tilde{f}$.

$$\tilde{f} = \sum_{n=0}^{N} \langle R^n f, g_{\gamma_n} \rangle g_{\gamma_n}$$

In the particular case of scale-by-scale matching pursuit, we can derive a very similar reconstruction formula. Indeed this reconstruction holds true for each scale individually, thus reconstructing the signal from the whole book, is a matter of averaging the reconstructions from each scale.

$$\tilde{f} = \frac{1}{S} \sum_{s=0}^{S-1} \Big(\sum_{n=0}^{N_s} \langle R^n f, g_{\gamma_n} \rangle g_{\gamma_n} \Big) \tag{12}$$

where S is the number of scales.

4.2 Examples

Let's start with the reconstruction of a single trace, extracted form Pyrenees. Figure 8 shows the type of reconstruction that can be obtained by performing a matching pursuit using 10, 50, 100 and 250 atoms per scale. A very similar comparison could be made using smaller and smaller thresholds.

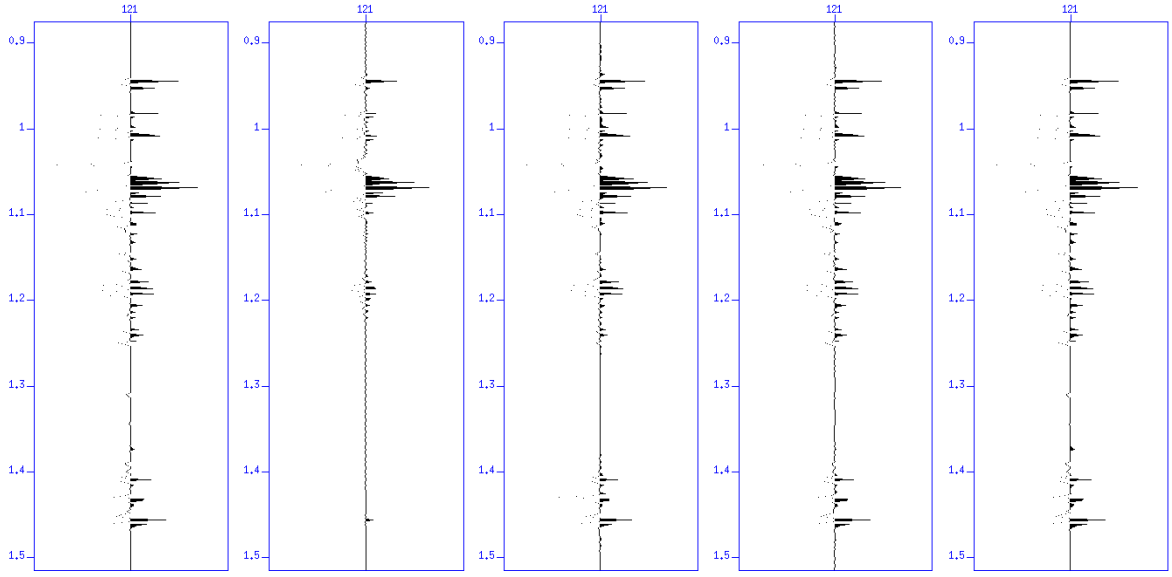
Figure 9 shows the reconstruction of a whole slice of dataset. This was obtained with a scale-by-scale matching pursuit performed on five scales and counting 500 iterations for traces of 751 points. This is a very high number of iterations to consider, and only used here for demonstration purpose. A smaller number of iterations would provide almost the same result. Still, it is interesting to note, that although there are differences, the reconstruction is almost perfect and not suffering much form lateral discontinuity problems.

5 Workflow

5.1 Block diagram

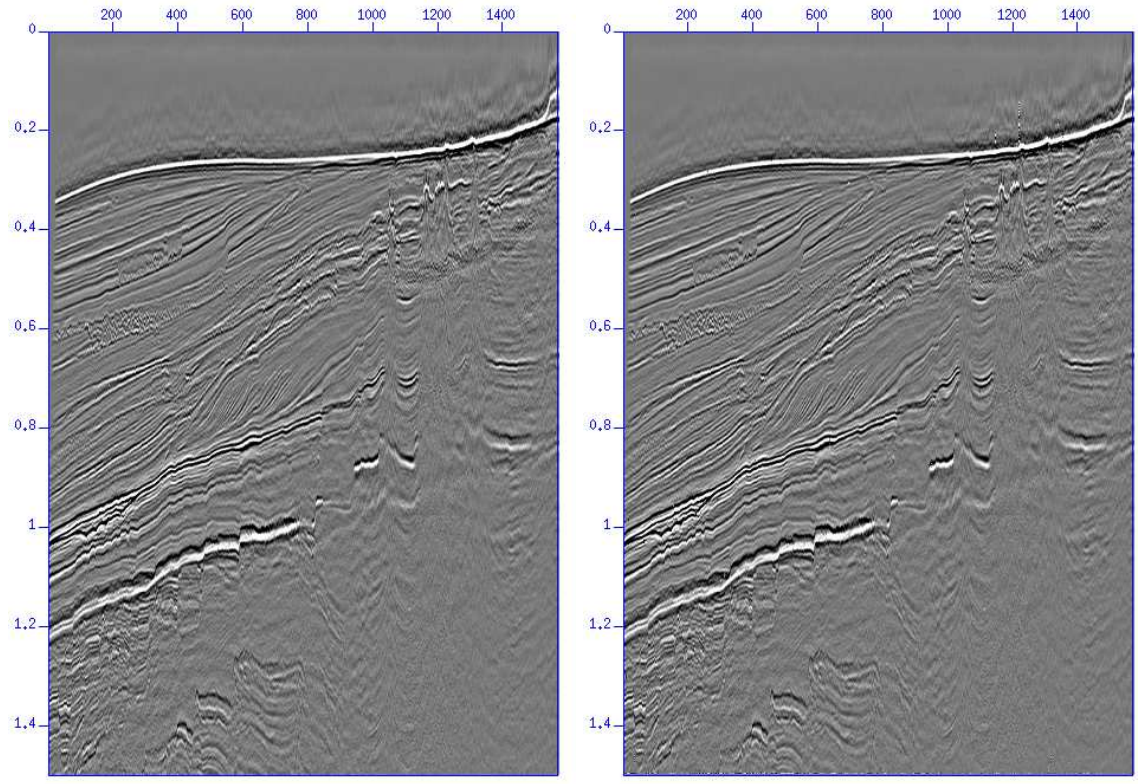
Figure 10 shows a block diagram of possibilities in using all softwares. The choice of workflow really just depends on the type of data that is processed, whether it might be simple reflectivities or real seismic traces.

Note that in case the aim is to obtain a perfect reconstruction, the book of atoms should never be normalized. However, the reconstruction of a normalized dataset can offer insights on what 'should be'.



(a) Original (b) 10 iterations (c) 50 iterations (d) 100 iterations (e) 250 iterations

Figure 8: On the far left, the original wigb depiction of the signal. Remaining figures show the reconstruction of this signal after a matching pursuit decomposition using various numbers of iterations. Figure b and c already show how it is possible to extract relevant information from the signal with few iterations, going to much higher numbers, the reconstruction ends up near perfect.



(a) Original (b) Reconstruction

Figure 9: On the left the original dataset, on the right its reconstruction, minor differences can be seen especially in the upper right corner of the image.

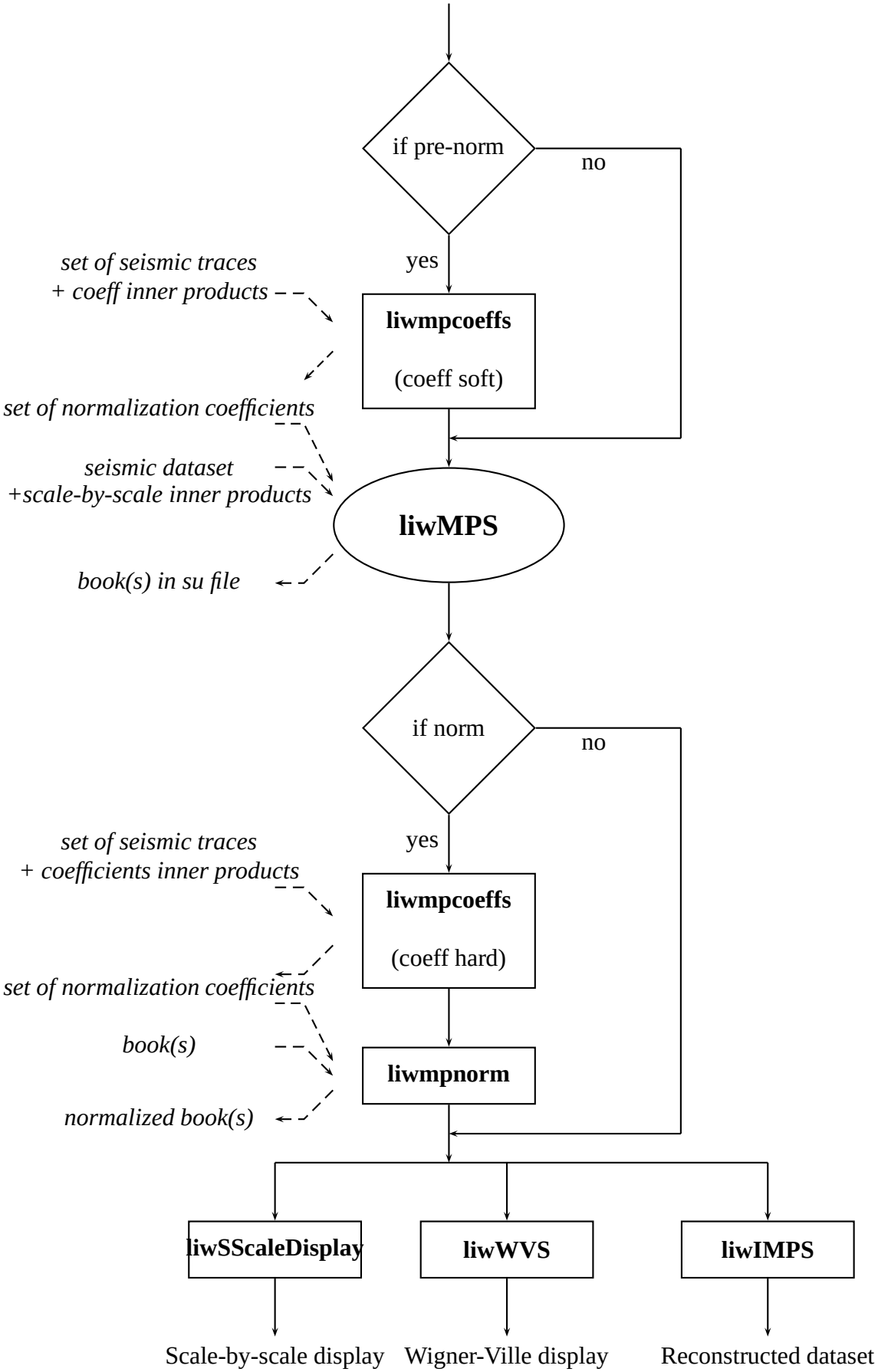
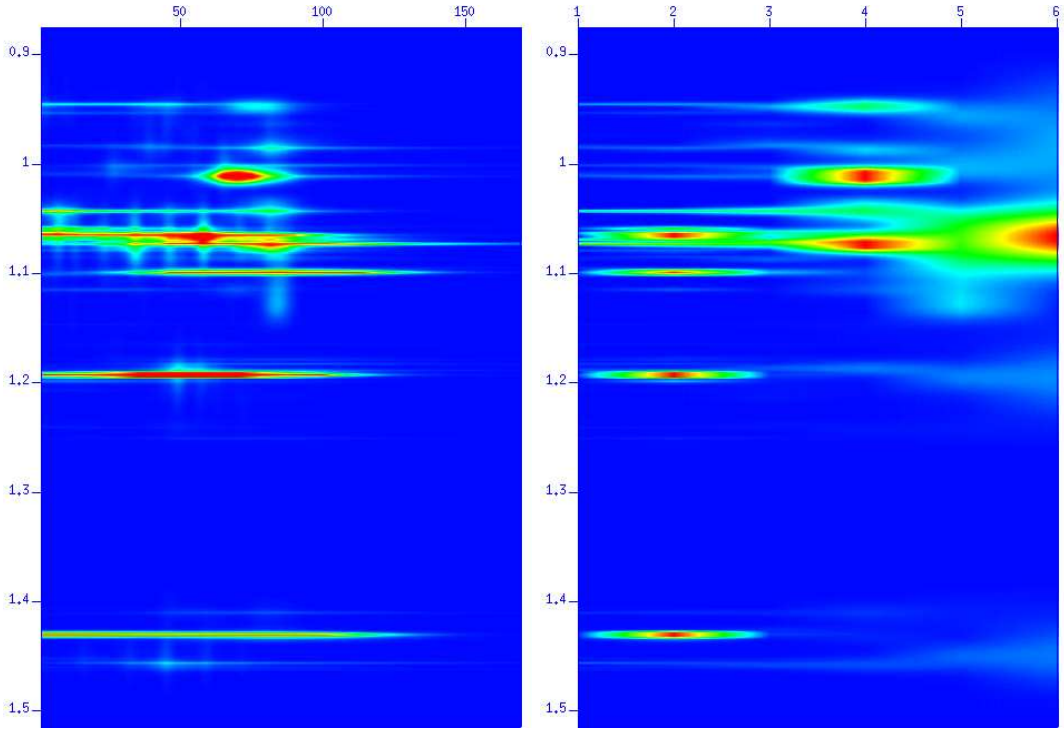


Figure 10: Block diagram of a typical workflow.

5.2 Examples of script

Following are two examples of scripts used on seismic data, whether it would be reflectivities or noisy, convolved, seismic data. The examples were taken from Pyrenees. We will first show how to proceed on the reflectivity.The results are displayed in figure 11.

```
liwMPS <rfc_Crosby-2_full_little.su  scale_max=5 nb_iter=500
product_dir=/storage/lannes/Seismic   >rfc_bk.su

liwWVS <rfc_bk.su >rfc_WV.su
liwSScaleDisplay <rfc_bk.su >rfc_WV.su
```



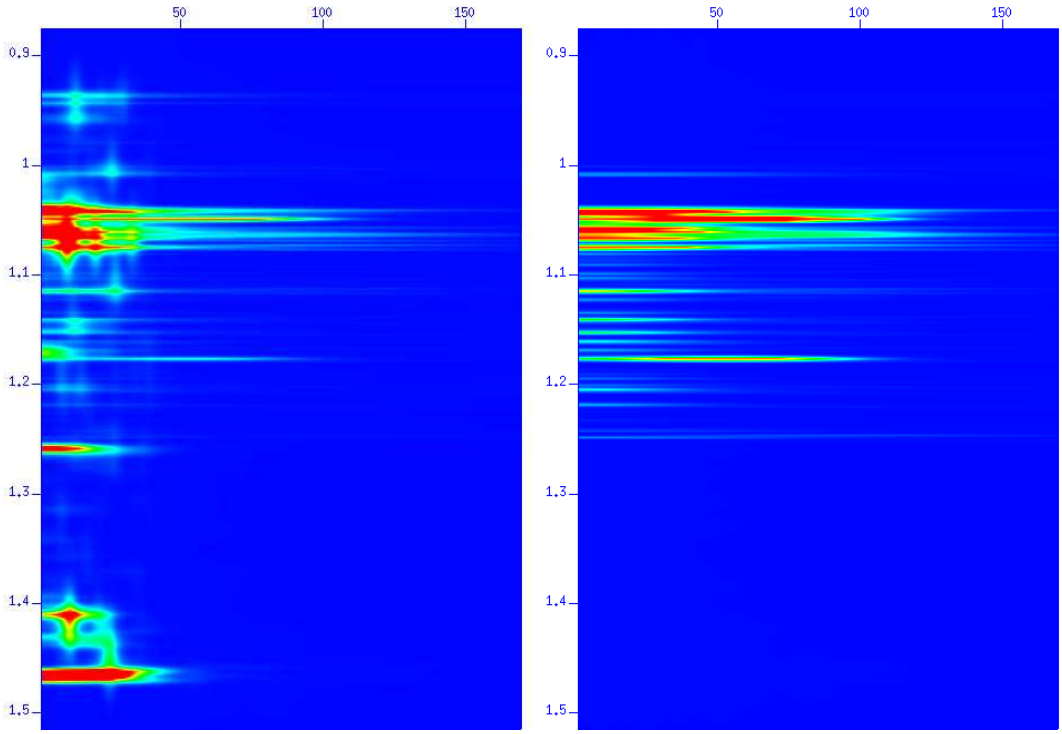
(a) WV (b) SD

Figure 11: The results of the matching pursuit, let the Wigner-Ville representation, right the scale-by-scale display.

On to the 'real' case. This includes pre-normalization and renormalization. Figure 12 show the results with and without normalization.

```
liwmpcoeffs <mig16_pst_full_il1434.su scale_max=5 window=100
product_dir=/storage/lannes/Seismic enhance_level=0.01   >coeffs_soft.su


liwMPS <seismic_well_Crosby-2_full_little.su  scale_max=5 nb_iter=100
product_dir=/storage/lannes/Seismic norm=1 input_coeffs=coeffs_soft.su
>seismic_bk.su

liwmpcoeffs <mig16_pst_full_il1434.su scale_max=5 window=100
```

```
product_dir=/storage/lannes/Seismic enhance_level=0.000001
>coeffs_hard.su

liwmpnorm <seismic_bk.su input_coeffs=coeffs_hard.su >seismic_norm.su

liwWVS <seismic_norm.su >norm_WV.su
liwSScaleDisplay <seismic_norm.su >norm_WV.su
```



(a) Before normalization (b) After normalization

Figure 12: On the left, the Wigner-Ville representation of the matching pursuit decomposition of the seismic trace. On the right, that of the normalized book of atoms. When comparing with the previous figure, it is obvious that the high-frequency content has mostly disappear, but some structures are partially recreated.

References

[1] S. Mallat. *Une exploration des signaux en ondelettes*. Les Editions de l'Ecole Polytechnique, 2000.

[2] S. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12):3397–3415, December 1993.

[3] J. Ville. Theorie et applications de la notion de signal analytique. *Cables et Transm.*, 2A(1):61–74, 1948.

[4] E. P. Wigner. On the quantum correction for thermodynamic equilibrium. *Phys. Rev.*, 40:749–759, 1932.

A liwMPS

```
[lannes@liw67 bhp]$ liwMPS

 LIWMPS -


 liwMPS <stdin  scale_max=  nb_iter= compute= write=  >stdout

 stdin          Input, dataset to analyse
 stdout         Input, book of atoms

 Required parameters:

 scale_max=     Input, maximum logarithmic scale of the decomposition
                Scales go from 0 to sclae_max, or 3*dt to
                (2^(scale_max+1))*dt
 nb_iter=       Input, maximum number of atoms to compute per trace
 product_dir=   Input, name of the directory containing atoms
                inner products


 Optional parameter:

 compute=0      =1 to re-compute inner products between atoms
 write=0        =1 to write the table of inner product in a RAW file
 threshold=0    proportional threshold on atoms' energy
                Algorithm stops either when the energy falls below
                the threshold*singal_energy or when nb_iter is reached.
                e.g. 0.01, means the algorithms stops when the
                energy of the residual signal represent 1%  of
                the energy of the original signal or after
                nb_iter iterations, whichever comes first
 norm=0         =1 to pre-normalize
 grid=1         factor of grid refinement, 1 is minimum,
                2 is twice finer and so on.
                default is 1, corresponding to a 1/2 overlap
                between atoms
 input_coeffs   filename containing coefficients needed for
                normalization
 verbose=0      =1 to allow advisory message
                =2 to display details about every atoms


 Header words:

 ns             5*nb_iter*(scale_max+1)
 mark           scale_max
 cdpt           length of the original trace

 For each atom, the output book of a trace stores in the following
 order:
```

```
                - its scale, as an integer between 0 and scale_max
                included
                - its position, as an integer corresponding to the
                sample number
                - its frequency as an integer corresponding to its
                index representation
                - its coefficient as two floats, the real and
                imaginary part

 Example

 liwMPS <nept_depth_reftrace_long.su  scale_max=5 nb_iter=100
 threshold=0.01 product_dir=/storage/lannes/Seismic  norm=1
 input_coeffs=coeffs_soft.su  >nept_pre_bk.su
```

Example of command line, to perform matching pursuit with pre-normalization.

```
liwMPS <nept_depth_reftrace_long.su  scale_max=5 nb_iter=100
threshold=0.01 product_dir=/storage/lannes/Seismic norm=1
input_coeffs=coeffs_soft.su >nept_pre_bk.su
```

B liwmpcoeffs

```
[lannes@liw67 bhp]$ liwmpcoeffs

 LIWMPCOEFFS -


 liwmpcoeffs <stdin   scale_max= window= product_dir= >coeffs.su

 Required parameters:

 scale_max      Input, maximum scale considered
 product_dir    Input, directory containing atoms inner products   .


 Optional parameter:

 window=0       Size of the time-averaging window
                Default at 0 automatically sets the length of the
                windows to be the length of the signal, hence assuming
                the wavelet is time-invariant over the length
                of the signal.
 verbose=0      Default 0 is no messages                      .
                =1 is estimated energy of the signal
                =2 is general information (including estimated noise)
 compute=0      =1 to compute inner products
 write=0        =1 to write the table of inner product in a RAW file
 noise_level    impose noise level in dB, otherwise compute
                if E is an estimate of the signal’s energy than the
```

```
               noise level in dB is written as
               N_dB = 10 log( Sigma^2 / E)
 enhance_level coefficient to limit amplification,
               1 is no amplification.
               0.1 or 0.01 are recommended for pre-normalization
               0.000001 for full normalization
               default is 0.001
               correction coefficients belong to
               [enhance_level, 1/enhance_level]
 corr_min      imposed minimum on correction coefficients
               default is equal to enhance_level
               correction coefficients belong to
               [corr_min, 1/enhance_level]

 Note that the energy of the signal is estimated and can be output
 through verbose

 About header words:

 ep             logarithmic scale of atoms
 cdp            frequency index of atoms

 Examples

 liwmpcoeffs <nept_depth_refline_long.su  scale_max=5 window=100
 product_dir=/storage/lannes/Seismic enhance_level=0.1
 >coeffs_soft.su

 liwmpcoeffs <nept_depth_refline_long.su  scale_max=5 window=100
 product_dir=/storage/lannes/Seismic enhance_level=0.000001
 >coeffs_hard.su
```

Two examples of command line, depending on whether the coefficients are to be used for pre-normalization or for the regular normalization.

```
liwmpcoeffs <nept_depth_refline_long.su  scale_max=5 window=100
product_dir=/storage/lannes/Seismic enhance_level=0.1
>coeffs_soft.su

liwmpcoeffs <nept_depth_refline_long.su  scale_max=5 window=100
product_dir=/storage/lannes/Seismic enhance_level=0.000001
>coeffs_hard.su
```

C liwmpnorm

```
[lannes@liw67 bhp]$ liwmpnorm

 LIWMPNORM -
```

```
liwmpnorm <stdin  input_coeffs= >stdout

Required parameters:

input_coeffs=  Input, sufile containing correction coeffcients


Optional parameter:

verbose=0     =1 to allow advisory message                          .
window=ns     length of the averaging window, defaults
              is whole length of the signal
              a reasonable choice is to pick same value
              as used for liwmpcoeffs
              (a sum along the time axis using a sliding window
              is performed to ensure all scales have the same
              total energy, this window represents the number of
              samples to consider)


Example

liwmpnorm <nept_pre_bk.su window=100 input_coeffs=coeffs_hard.su
>nept_norm.su
```

Example:

```
liwmpnorm <nept_pre_bk.su  input_coeffs=coeffs_hard.su >nept_norm.su
```

D liwWVS

```
[lannes@liw67 bhp]$ liwWVS

LIWWVS -


liwWVS <stdin   >stdout


stdin          Input, su file containing book of atoms
stdout         Output, su file containing display

Optional parameter:

scale_min=0    Displays only atoms of scale above scale_min
scale_max=100  Displays only atoms of scale under scale_max
               Scales are logarithmic and start from 0.
               Scale n corresponds to n atom of size ( 2^(n+1)+1 )*dt
freq_min=0     Displays only atoms of frequency above freq_min
```

```
 freq_max=3.14  Displays only atoms of frequency under freq_max
                Frequencies go from 0 to pi=3.14 where pi is 1/dt
                Frequencies are k*pi/N, k from 0 to N where N is
                max number of frequencies
                and corresponds to k/(N*dt) is Hertz
 nb_freq        Number of frequencies considered for the display
                (sampling of the display)
 velocity=2500  Velocity used to compute bed thickness from scale
 verbose=0      =1 to allow advisory message

 Header words:

 f1             Frequencies in Hz
 f2             Wavelength in m
 cdpt           Frequencies' indices k as in k*pi/N

 Example:

 liwWVS <nept_norm.su >nept_WV.su
```

Example:

```
liwWVS <nept_norm.su | suximage perc=99
```

E liwSScaleDisplay

```
[lannes@liw67 Neptune]$ liwSScaleDisplay

 LIWSSCALEDISPLAY -


 liwSScaleDisplay <stdin   >stdout


 stdin          Input, su file containing book of atoms
 stdout         Output, su file containing display

 Optional parameter:

 scale_min=0    Displays only atoms of scale above scale_min
 scale_max=100  Displays only atoms of scale under scale_max
                Scales are logarithmic and start from 0.
                Scale n corresponds to n atom of size ( 2^(n+1)+1 )*dt
 freq_min=0     Displays only atoms of frequency above freq_min
 freq_max=3.14  Displays only atoms of frequency under freq_max
                Frequencies go from 0 to pi=3.14 where pi is 1/dt
                Frequencies are k*pi/N, k from 0 to N where N is
                max number of frequencies
                and corresponds to k/(N*dt) is Hertz
```

```
 velocity=2500  Velocity used to compute bed thickness from scale
 verbose=0      =1 to allow advisory message

 Header words:

 f2             Bed thickness in m
 cdpt           Log scale (+1)

 Example:

 liwSScaleDisplay  <nept_norm.su >nept_SD.su      .
```

Example:

```
liwSScaleDisplay <nept_norm.su | suximage perc=99
```

F liwIMPS

```
[lannes@liw67 Neptune]$ liwIMPS

 LIWIMPS -


 liwIMPS <stdin   >stdout


 Optional parameter:

 verbose=0      =1 to allow advisory message
 scale_min=0    Displays only atoms of scale above scale_min
 scale_max=100  Displays only atoms of scale under scale_max
                Scales are logarithmic and start from 0.
                Scale n corresponds to n atom of size
                ( 2^(n+1)+1 )*dt
 freq_min=0     Displays only atoms of frequency above freq_min
 freq_max=3.14  Displays only atoms of frequency under freq_max
                Frequencies go from 0 to pi=3.14 where pi is 1/dt
                Frequencies are k*pi/N, k from 0 to N where N is
                max number of frequencies
                and corresponds to k/(N*dt) is Hertz

 Example:

 liwIMPS <rfc_bk.su >rfc_rec.su
```

Example:

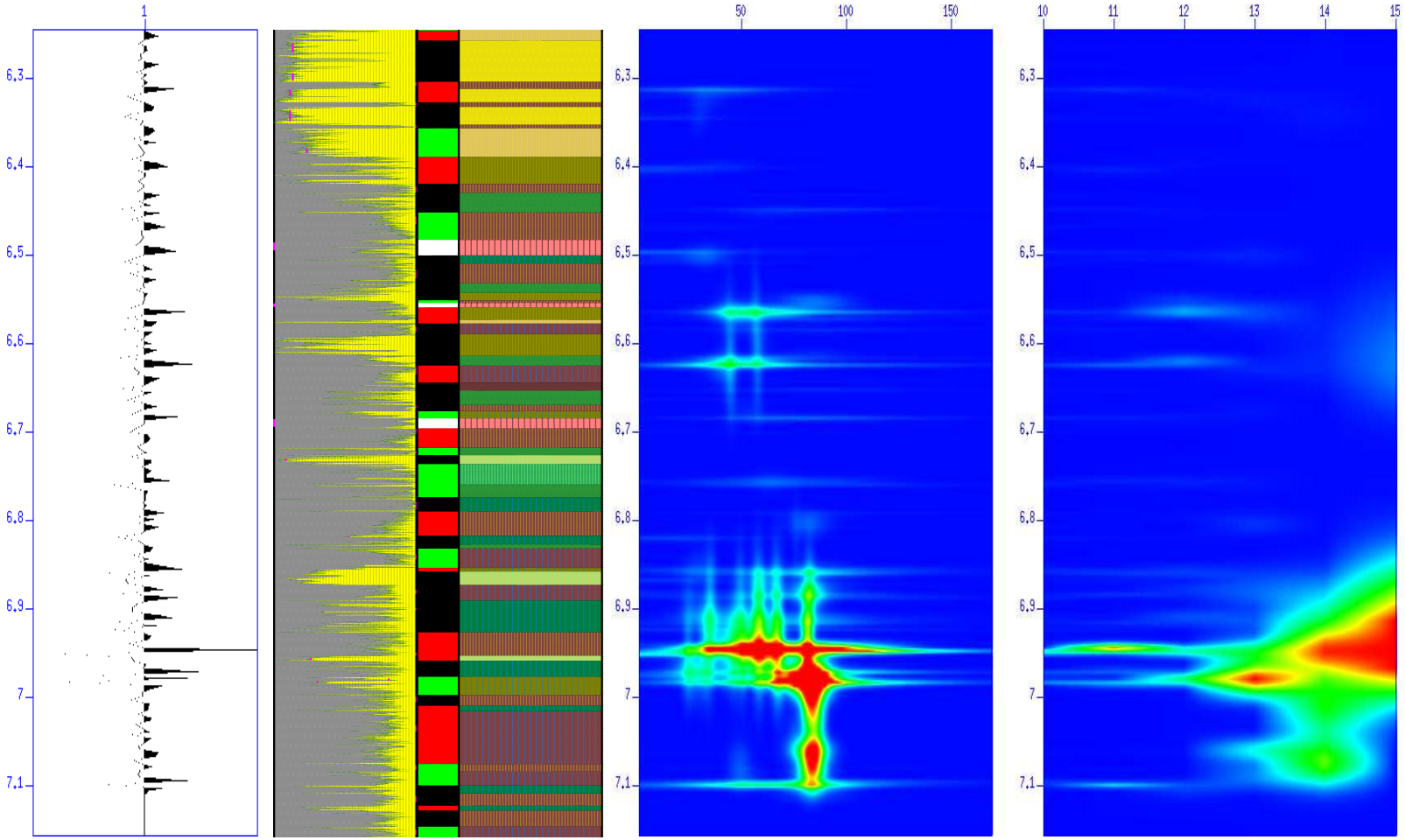
```
liwIMPS <rfc_bk.su >rfc_rec.su
```

G Lithofacies

Comparison with lithofacies are shown in figure 13 to 17.

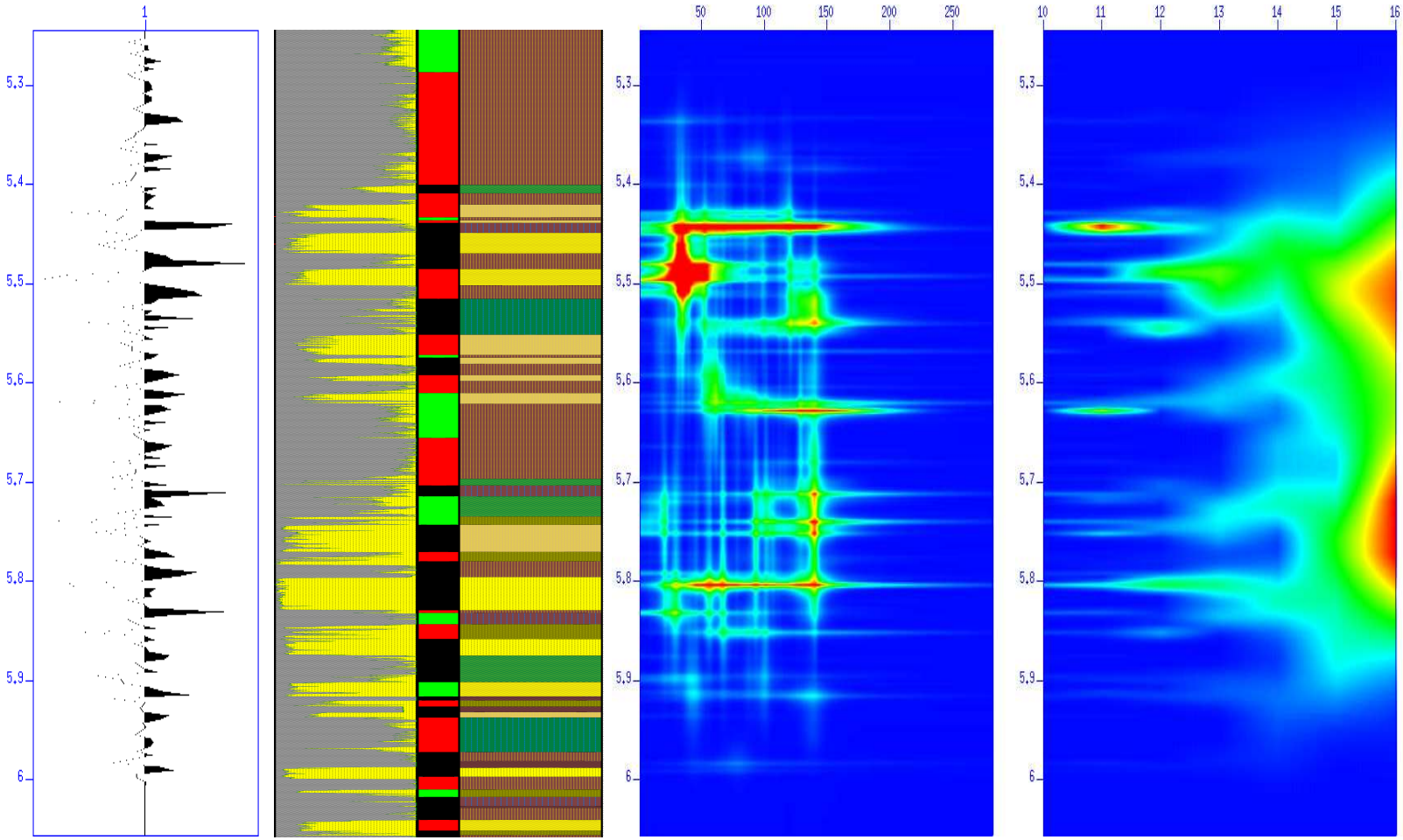
H Let It Wave

Address:	Let It Wave SA 8-16 rue Paul Vaillant-Couturier 92240 MALAKOFF, FRANCE
Phone numbers:	+33 1 40 92 54 43 +33 1 40 92 54 54
Fax:	+33 1 40 92 54 41
Web:	www.letitwave.fr



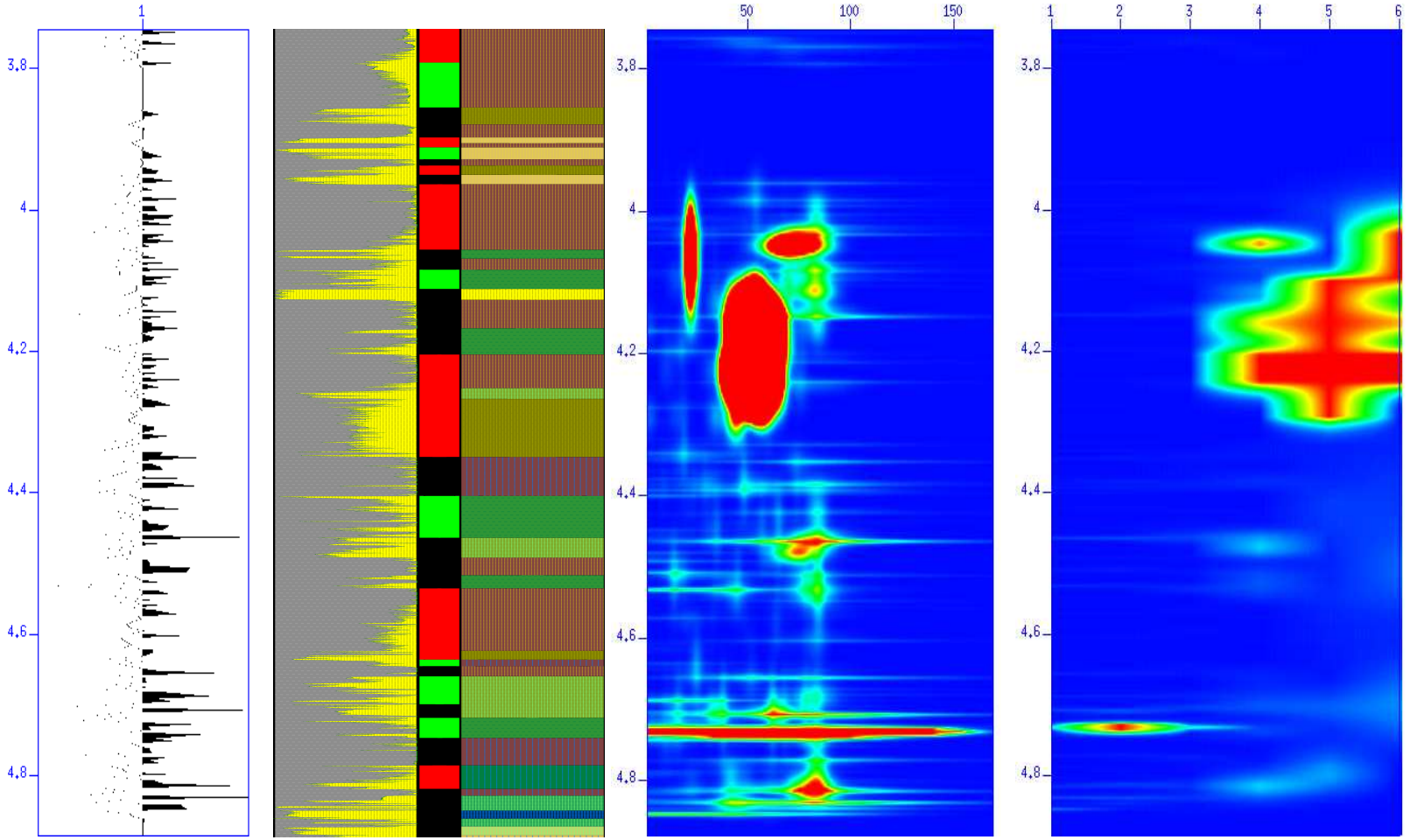
(a) Trace (b) lithofacies (c) WV (d) SD

Figure 13: Chinook



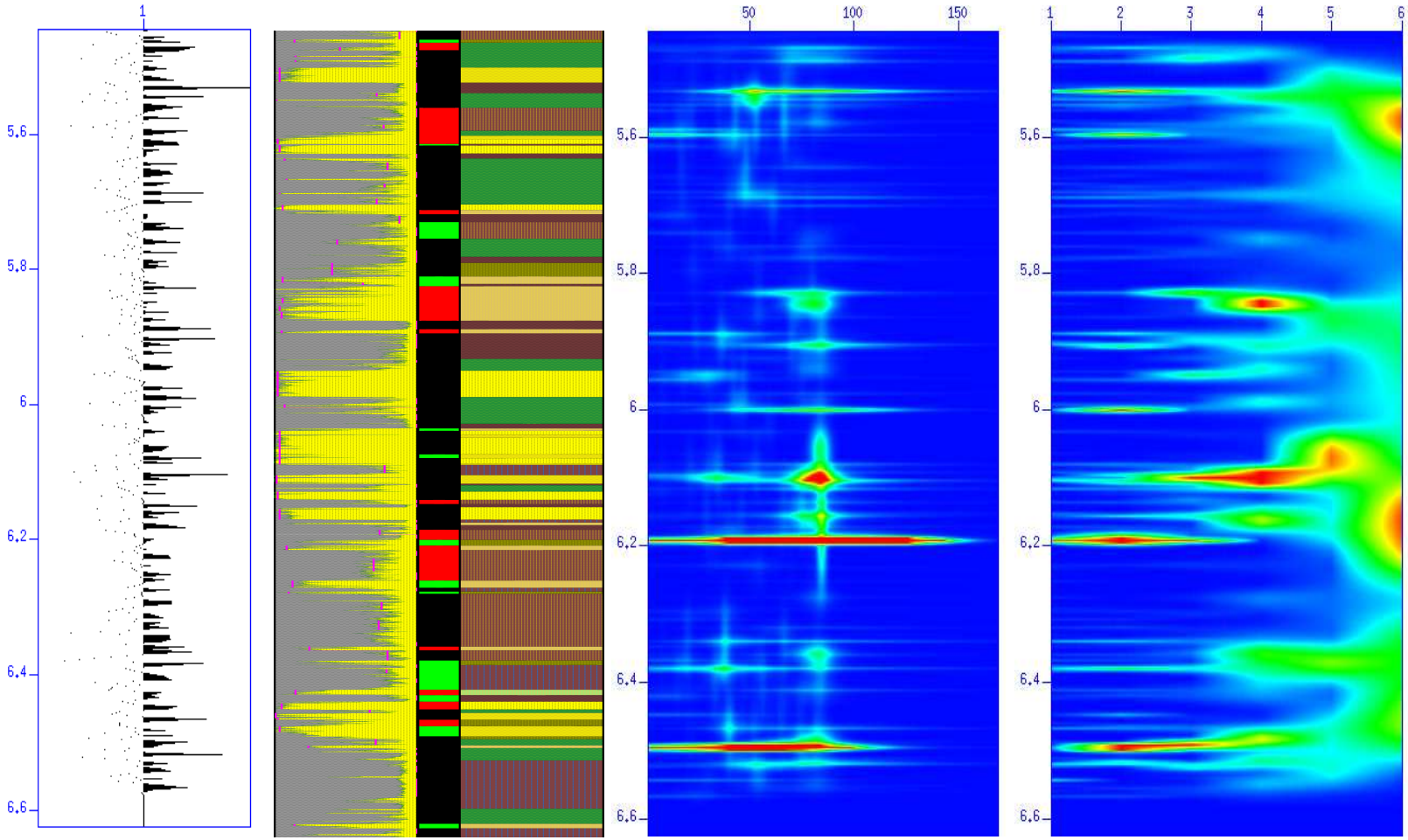
(a) Trace (b) lithofacies (c) WV (d) SD

Figure 14: Atlantis



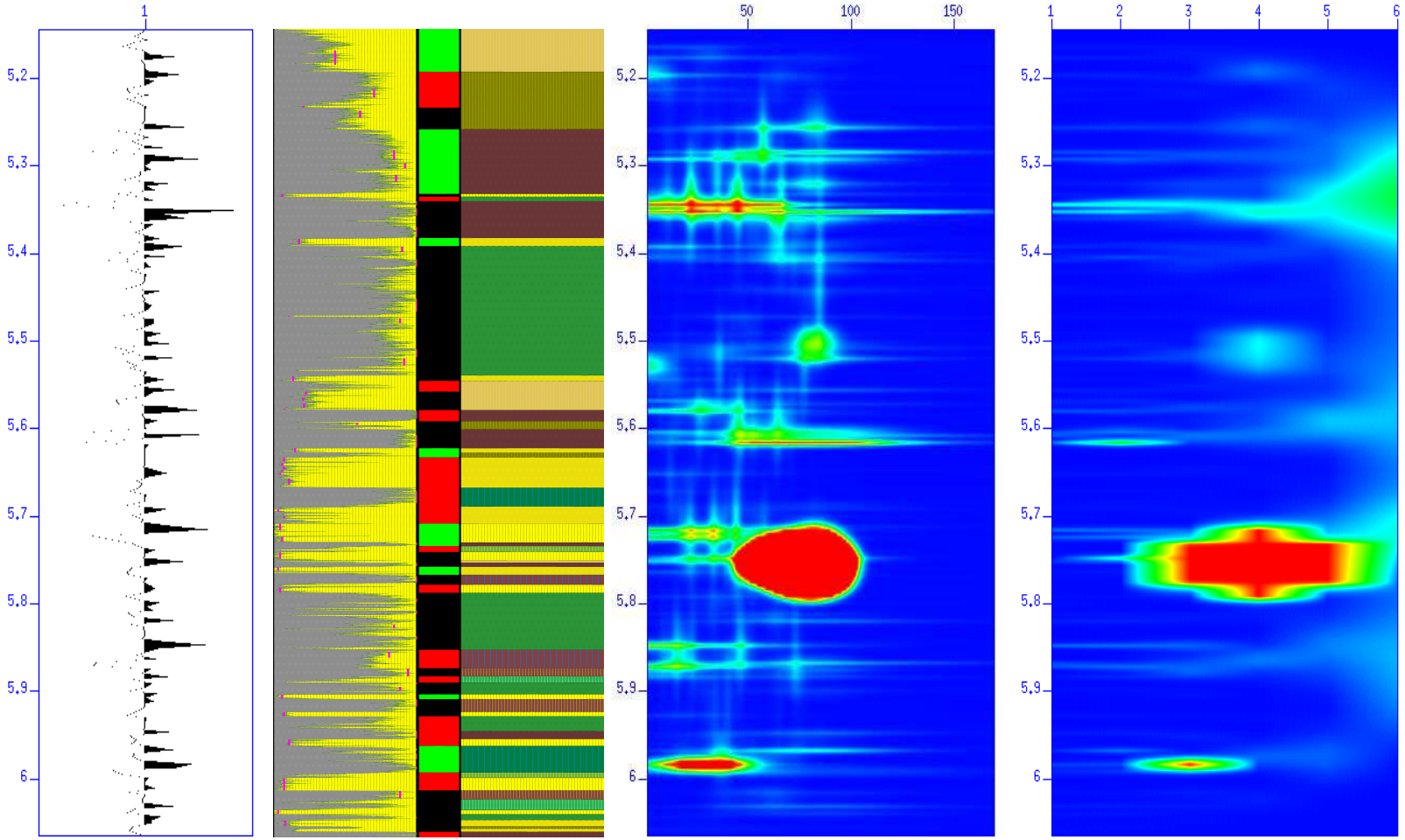
(a) Trace (b) lithofacies (c) WV (d) SD

Figure 15: Blackjack



(a) Trace (b) lithofacies (c) WV (d) SD

Figure 16: Dendara



(a) Trace (b) lithofacies (c) WV (d) SD

Figure 17: Frampton