C F A / V I S H N O 2 0 1 6

Extraction des ondes sismiques sur la faille de San Jacinto via le bruit sismique ambiant et un réseau dense de capteurs

C. Gradon[a], L. Moreau[a], P. Roux[a], A. Lecointre[a] et Y. Ben Zion[b]

[a]Institut des Sciences de la Terre, ISTerre - Maison des Géosciences, 1381 rue de la Piscine BP 53, 38041 Grenoble Cedex 9, France

[b]Department of Earth Sciences, University of Southern California, Los Angeles, 90089-0740, USA

chloe.gradon@ujf-grenoble.fr

LE MANS

L'objectif de cette étude est de profiter d'un jeu de données sismiques multi-échelles observées en surface de la Faille de San Jacinto au Sud de la Californie. Ces données ont été mesurées en juin 2014 par un réseau très dense de capteurs, constituée de 1108 sismomètres espacés de 10 à 30 m dans une zone carrée (600m x 600m) centrée sur la partie endommagée de la faille. L'inter-corrélation du bruit sismique enregistré sur chaque capteur a permis de révéler un champ sismique très complexe : il inclut des ondes de surface et de volume fortement corrélées d'une station à l'autre, mais aussi leurs réflexions sur les différentes interfaces de la faille. Dans un premier temps, nous appliquons des méthodes de traitement d'antennes aux corrélations de bruit afin d'extraire et d'identifier ces différentes ondes. Dans un second temps, nous utilisons les ondes extraites dans le problème inverse afin de relocaliser des sources sismiques situées en profondeur sous le réseau. Le suivi de l'évolution de ces sources sur la durée des mesures devrait permettre de mieux appréhender le comportement de la faille.

The objective of the study of fault is to monitor the mechanics and structural changes of the crust, in order to find answers to the questions related to the triggering of earthquakes. The goal of this study is to make use of a multi-scale seismic dataset collected at the surface of the San Jacinto Fault in Southern California. The data were recorded in June 2014 by a dense array of sensors, comprising 1108 seismometers set from 10 to 30 meters apart in a square area (600m x 600m) centered on the damaged part of the fault. The cross-correlations of the seismic noise recorded on each sensor reveals a very complex seismic wave field: it includes surface and body waves that are highly correlated throughout the array, but also their reflexion at the fault interfaces. First, array processing methods have been applied to the noise correlations to extract and identify these different waves. The extracted waves will then be used in the inverse problem to relocalize seismic sources located deep under the array. The monitoring of these sources for the duration of the experiment should provide a better understanding of the behavior of the fault.

1 Introduction

Faults are indicators of the mechanical and structural changes of the crust. They are critical areas that are of significant importance in Earth sciences because their internal structure is related to the behavior of the crust [1]. The goal of this study is to use a seismic dataset acquired on the active part of the San Jacinto Fault Zone (SJFZ) in South California by a dense array of sensors. This density is a favorable condition for ambient noise processing through cross-correlations processing [2].

Because of its continuous activity and localisation, the SJFZ in California has been the stage of many studies, both in regard to the history recorded on its surface [3] or its deeper structure and behavior as a propagative medium [4] [5]. However, the subsurface characterization of the fault is still an on-going research topic [1].

The use of ambient noise in seismology as well as array processing has helped overcome the need for active sources in seismic prospection [6]. It has also given access to measurements of wave propagation at higher frequencies, thanks to different natures of noise sources. Because of their smaller wavelength, those higher frequencies can help gain better resolution on subsurface imaging.

A recurrent problematic in seismology is the ability to separate and identify the different waves that compose the seismic field. At the surface of the SJFZ, the recorded data revealed a complex seismic field including surface and body waves that are highly correlated throughout the array, but also their reflexion at the fault interface as well as trapped waves [7].

There are several ways to separate different waves from their original wave field, such as a Fourier-domain analysis of the wavenumbers or a projection on the time-frequency domain. In this study, we take advantage of the density of the array to apply antenna processing methods, i.e. a beamforming approach which consists in applying calculated delays to a signal to deduce either the location of a wave or its velocity.

Beamforming is used in oceanography [8] as well as in seismology, [9] [10] and can also be applied to electromagnetic waves [11]. In this study, beamforming has been used as a wave extracting tool and a way to determine their phase and group velocities. In addition, an amplitude treatment has been applied to iterate the process, allowing the extraction of several waves. The cross-correlations used in this study were also filtered using a singular value decomposition-based process to gain better access to the surface waves in the data.

The SVD is a mathematical tool for factorizing a matrix into unitary matrices. It has applications in many different fields such as image processing and big data exploitation. One of its application in Geophysics is the separation of the data into a signal and a noise subspace, for denoising purposes [12]. In the present study, it is used in combination to a spatial Fourier transform in order to design a filter for i) denoising the data and ii) removing unwanted waves that may prevent a clean extraction of the surface and body waves.

In the first section, the wave extraction procedure is described. In the second section, the method is applied to the extraction of surface and body waves on the SJFZ.

2 Methods

The study was carried out on the Clark branch of SJFZ. This fault is part of the Californian plate boundary system and accommodates a large portion of the plate motion in the area . A ~600 square meter array centered on the fault was instrumented with 1108 geophones distributed in twenty rows of about 55 sensors perpendicular to the fault (figure 1). In each row the inter-station spacing is 10 meters and the distance between each row is 30 m. For a detailed description of the configuration, the reader is invited to refer to [1] and [2].

Data were continuously recorded for a month between the 7th of May and the 13rd June of 2014. In order to compute ambient noise correlations, a specific cross-correlation pre-processing was applied to the data: first, a one-bit time domain normalization was used in combination with frequency
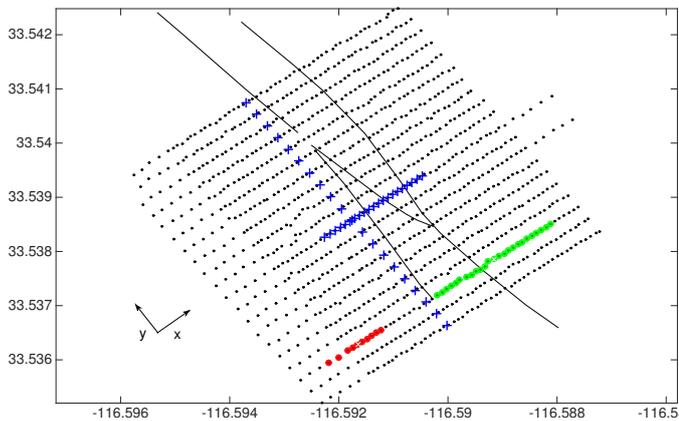
Figure 1: Disposition of the dense array: emission antenna in red and reception antenna in green. Point of extraction of the waves: line of blue crosses in the y-direction for the scan along the fault and in the x-direction for the scan across the fault.

whitening in frequency bands where the bandwidth is inferior to the two thirds of the central frequency. This was done in order to prevent signal harmonics caused by the one-bit clipping in the frequency band of interest. Then, the resulting data were cross-correlated for each station pair. Finally, correlations were averaged over one day time windows [2].

Each cross-correlation is representative of the green function between two sensors. Sorting each correlation with respect to the inter-station distances allows the reconstitution of the wave field propagating between any two sets of sub-arrays, which is useful for the application of antenna processing. While the grid of sensors used in the study is composed of passive sensors, the first sub-array, called here emission array, can be used as a sources array thanks to array processing. Its role is to add energy to the processed signal. It is smaller in size and in number of sensors than the second, receiving array. The latter has to cover enough distance to insure a satisfying resolution of the beamforming (section 2.2).

2.1 Filtering of the high velocity waves

Figure 2.a shows that part of the traces composing the signal are vertically aligned around t=0. This corresponds to a high energy wave recorded simultaneously by all the sensors. This can be interpreted as a plane wave generated by sources deep in the far field of the array. This study focuses on extracting waves for future subsequent near surface tomography. Such far field waves need to be suppressed in order to access the lower energy surface and body waves that will be used in the tomography.

Because this wave was recorded simultaneously by all stations, it has a very high apparent velocity, and a low wave number. To filter it out, a Fourier Transform was used along the position of the receiving array, creating a set of spatial Fourier-domain matrices for different wavenumber values. A singular value decomposition was then applied to each matrices. It consists in a projection of the signal onto its singular vectors, each vector being associated with a different level of energy in the signal subspace. It is possible to decompose A_k the spatial Fourier-domain matrix for the wavenumber k such that:

$$A_k = U S V^T \qquad (1)$$

Where U and V are N_e-by-N_e and N_r-by-N_r matrices. N_e and N_r are the number of sensors in the emission antenna and in the reception antenna respectively.. The columns of U and V are called left and right singular vectors for A_k respectively. The former are associated with the transmission subspace, and the latter to the reception subspace. S is a N_e-by-N_r diagonal matrix containing the singular values of A_k. Each singular value represents a different level of energy of the signal: the higher the singular value, the higher the energy.

One can see on figure 2.b that the first singular value has its highest values for small wave numbers, which corresponds to the wave that needs to be filtered out. Nullifying some of the singular values is equivalent to filtering the signal on the level of energy they are associated with. In order to eliminate the higher velocities on the highest energy waves of the signal, the first singular value was set to zero for wave numbers associated with a phase velocity greater than 4500 m/s. The lower singular values, corresponding to lower energy signal, were also suppressed to diminish the noise on the recordings, such that only four singular values remained. (figure 2.c).
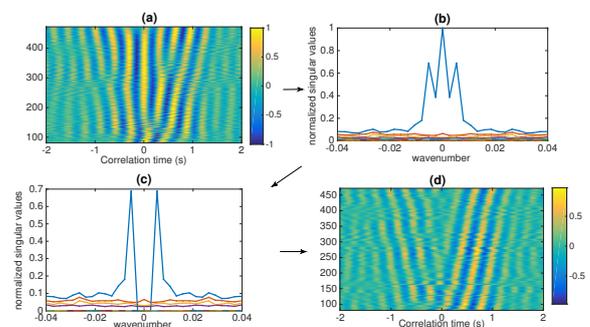
Figure 2: Normalized one-day correlations between the stations in subarrays 1 and 2 (see figure 1) with frequency whitening in the [2-3] Hz bandwidth (correlations are sorted w.r.t. inter-station distances). a) Raw data and b) corresponding wavenumber-dependent singular values. c) Singular values after a low-pass filter on the first one and separation of the noise subspace. d) Filtered correlations: the dominant wave with infinite phase velocity at t=0 has been removed, and noise is reduced.

2.2 Iterative beamforming wave extraction

The filtered correlations contain several wave paquets (figure 2.d). In order to identify the different waves, we introduce a procedure to extract them separately, namely a Beamforming approach.

It consists in decomposing the data according to the apparent slowness measured on the array. Delays are calculated by projecting a slowness vector onto the correlations matrix. This vector is then steered in order to align each trace. When the slowness in the steering vector matches that of the seismic waves, the projection reaches a maximum. This maximum can also be used to identify the relative amplitude of each wave as well as their arrival time. This projection is equivalent to a transform from the time-space plane to a "time-beam" plane according to the following formula:

$$b(t,v) = \sum_{i=1}^{N_e} \sum_{j=1}^{N_r} p(t - \tau(x_i^e, x_j^r, u), x_i^e, x_j^r) \qquad (2)$$

where b is the signal in the beam space, τ is the delay, x_i^{em} the position of the ith sensor of the emission antenna and x_j^{re} the position of the jth sensor of the reception antenna. u is the phase slowness of the extracted wave. τ can be expressed as:

$$\tau(x^{em}, x^{re}, u) = (x^{em} - x_c^{em})u - (x^{re} - x_c^{re})u \qquad (3)$$

Where x_c^{em} and x_c^{re} are the positions of the center of the transmitting and receiving arrays respectively.

Once the maximum is found and the slowness is extracted, a time window is applied to the aligned traces in order to suppress any secondary lobe effect. Subsequently, it is possible to recreate the associated wave by applying the opposite delay to every trace of the data.

$$b(t,v) = \sum_{i=1}^{N_{em}} \sum_{j=1}^{N_{re}} p(t + \tau(x_i^{em}, x_j^{re}, u_{max}), x_i^{em}, x_j^{re}) \qquad (4)$$

The phase and group velocities can be derived from the beamfoming process. The first is the inverse of the slowness u_{max}, for which all the traces have alined. All the traces are aligned in time according to a reference point, i.e. the center of each array. The group velocity can be derived from the arrival time lag of the wavelet, t_{ar}, on which the beamforming process has its maximum. Therefore, the group velocity is calculated such that

$$V_{gr} = \frac{d_{centers}}{t_{ar}} \qquad (5)$$

where d_{center} is the distance between the two center of the arrays.

During the beamforming process, the information about the amplitude variations along the stations of the arrays is lost, due to the summation of all traces. For a proper extraction of the wavelet from the correlations it is necessary to account for its distance-dependent amplitude. This is essential if the beamforming process is to be iterated for the successive extraction of several waves. Indeed, neglecting amplitude variations may result in improper wave extraction, which may corrupt the next beamforming process on the remaining data. In order to account for such variations, the wavelet is weighted by a five coefficients polynomial. Those coefficients are determined via a simulated annealing global optimization method.

Simulated annealing is an algorithm based on Markov-Chain Monte Carlo (MCMC) methods, the aim of which is to evaluate the probability density of a variable in multi-dimensional problems. MCMC methods operate via stochastic sampling of the variable space and evaluate the probability of the samples with a Bayesian criterion [13]. In the present problem, the variables are the 5 polynomial coefficients. More details about the simulated annealing optimization can be found in [13] and [14]

Once the most energetic wave is extracted with the beamforming process, the procedure can be iterated to extract the next most energetic wave from the remaining data. The next maximum in the beam space is found, and the inverse beamforming and amplitude ponderation are performed, allowing
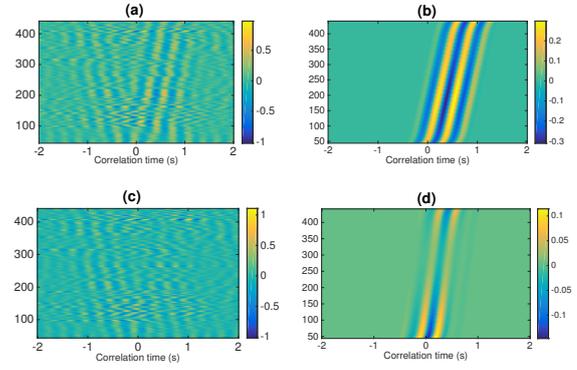
Figure 3: Extracting wave iterations : (a) Iteration 1 : Filtered correlations and b) the first extracted wave with phase velocity 574 m/s and group velocity 462 m/s, i.e. surface wave. (c) Remaining data after subtraction of the first wavelet and d) the second extracted wave with Vph=1113m/s and Vgr=1026m/s, i.e. body wave

the reconstruction of the new wave as well as the determination of its phase and group velocities. This process can be repeated as many times as there are waves in the data. However, the interpretation of each wave packet is not trivial. The order of extraction of the waves, depending on amplitude is not the same depending on the position of the sub-arrays. Therefore in the present study although four iterations have been used only two waves were extracted.

3 Variations of phase and group velocities of the waves across and along the SJF

To take advantage of the dense array of sensors, the wave extraction was performed for several positions of the transmitting and receiving arrays, which were displaced following either the x-oriented or y-oriented rows of the grid (figure1). This scanning can show the evolution of the phase and group velocities at different points of the grid and especially across the fault zone. Because of the noise sources distribution, [1] the extraction is more efficient with linear antennae perpendicular to the fault. To investigate the evolution of the extracted velocities along the fault, linear arrays aligned with the main direction of the noise were moved along the fault. On the other hand, to highlight the changes across the fault the scanning was done by sweeping the arrays on an entire row.

In order to extract relevant waves across the sensors grid, many configurations of antennae have been tested. Two relevant waves were extracted for each position of the antennae. They were sorted comparing their velocities and the shape of the maximum spot in the time-beam space for all iterations and positions used for the scan.

3.1 Evolution of the velocities in the direction parallel to the fault

In order to exploit the data, waves extracted with a certain configuration of sub-arrays are associated to a specific extraction point. It corresponds to the mid distance between the centers of the two arrays. The blue crosses in figure 1 repre-

sent those centers for each configuration used during the scan along and across the fault. The extraction of waves along the fault shows two different types of waves (figure 4). This differentiation was done by taking into account two criteria: the value of the velocity and the dispersion. Body waves have a higher velocity but are non-dispersive, meaning their phase and group velocity should be roughly the the same. There is no distinct dispersion for the extracted waves, but the slower wave in blue shows velocities of the same scale of the velocities of surface waves. The second wave, which exhibits higher velocities is interpreted here as a body wave.

There are only few positions at the northern part of the array for which the body wave was extracted. On the other hand, the first surface wave was detected all over the array. Thus, while it is possible to see that the wave speed along the fault remains stable for the surface waves, no conclusions can be made for the body wave.
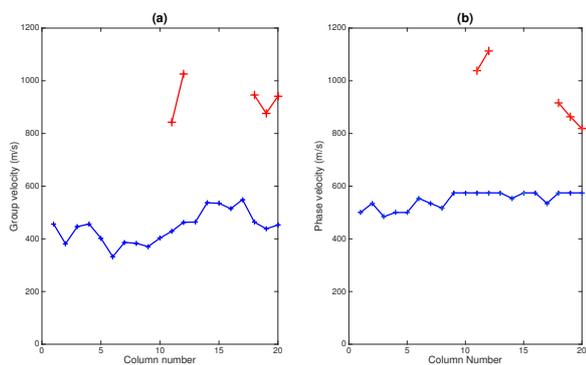
Figure 4: Waves extracted along the fault : (a) Group velocity for two extracted waves (b) Phase velocity for two extracted waves

3.2 Evolution of the velocities across the fault

A scan across the fault can give an idea of its effect on the speed of the waves that pass through it. The scanning was performed on a single column, the antennae being translated by one sensor along the column between two successive positions. The central position between the transmitting and receiving arrays covered the region represented by the blue crossesin the x-direction in figure 1. This is the portion over which the waves are extracted.

Similarly to the scan along the fault, two types of waves were detected (figure 5). Again, only a few positions reveal the presence of a body wave around 1100m/s and the low number of point makes establishing any clear evolution of the velocity difficult. Regarding the first surface wave, its group velocity is decreasing across the fault and its dispersive nature is visible.
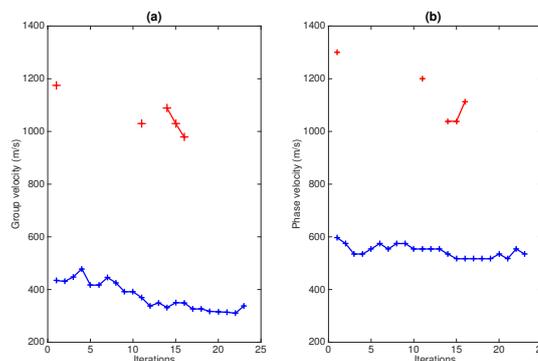
Figure 5: Waves extracted across the fault : (a) Group velocity for two extracted waves (b) Phase velocity for two extracted waves

4   Conclusion

Taking advantage of a dense array on the San Jacinto Fault, a procedure for extracting surface and body waves from ambient noise was introduced. The procedure relies on iterative beamforming between sub-arrays. The possibility of separating several waves, either surface or body waves, will permit the gathering of information on the fault. Indeed, because of their inherent nature and of the reflexions around the fault, the waves do not propagate at the same depth or along the same paths. For now, only one surface wave has been extracted successfully all around the array. However the very high number of available sub-arrays combinations increases the odds of finding the waves on part of the array where they have not been detected yet.

One of the other issue is the characterization of the extracted waves. Depending on the position and the shape of the antennae, the intensity of each wave is different and their order of extraction is not the same. Considering the fact that the wave velocities are likely to vary over the array, there is a need to identify each wave. The use of Double Beamforming [10], which takes into account the velocities and the azimuts of the wave at each array when calculating the delay, could also improve the separation of the waves from the original data. The waves would be distinguished by their speed but also their angle of arrival on the arrays, allowing the identification of waves with similar velocity. The use of pattern recognition algorithm could also help to automatize the characterization of the wave allowing a more efficient scanning of the array.

References

[1] Y. Ben-Zion, F. Vernon, Y. Ozakin, D. Zigone, Z. Ross, H. Meng, M. White, J. Reyes, D. Hollis and M. Barklage, Basic data features and results from a spacially-dense seismic array on the San Jacinto fault zone, *Geophys. J. Int.* **202**, (2015).

[2] P. Roux, L. Moreau, A. Lecointre, G. Hillers, M. Campillo, Y. Ben-Zion, D. Zigone and F. Vernon, A methodological approach toward high-resolution seismic imaging of the San Jacinto Fault Zone using continuous ambient noise recordings at a spatially-dense array, *Under Review*, (2016).

[3] T.K. Rockwell, T.E. Dawson, J. Young and G. Seitz, A 21 event, 4,000-year history of surface ruptures in the Anza Seismic Gap, San Jacinto Fault: Implications for long-term earthquake production on a major plate boundary fault, *Pure Appl. Geophys.* **172**, (2015).

[4] A.A. Allam and Y. Ben-Zion, Seismic velocity structures in the Southern California plate-boundary environment from double-difference tomography, *Geophys. J. Int.* **190**, (2012).

[5] D. Zigone, Y. Ben-Zion, M. Campillo and P. Roux, Seismic tomography of the Southern California plate boundary region from noise-based Rayleigh and Love Waves, *Pure Appl. Geophys.* **172**, (2015).

[6] N.M. Shapiro and M. Campillo, Emergence of broadband rayleigh waves from correlations of the ambient seismic noise , *Geophys. Res. Lett.* **31**, (2004).

[7] G. Hillers, M. Campillo and P. Roux, Seismic fault zone trapped noise, *J. Geophys. Res.* **119**, (2014).

[8] P. Roux, W. Kupermann and K.G. Sabra, Ocean acoustic noise and passive coherent array processing, *C. R. Geosci.* **343**, (2011).

[9] B. De Cacqueray, P. Roux, M. Campillo, S. Catheline and P. Boué, Elastic-wave identification and extraction through array processing:An experimental investigation at the laboratory scale, *J. Appl. Geophys.* **74**, (2011).

[10] P. Boué, P. Roux, M. Campillo and B. de Cacqueray, Double Beamforming in a seismic prospecting context, *Geophysics* **78**, (2013).

[11] S. Wendt, A. Chicot and M. Skrok, On beamforming performance in Wi-Fi outdoor networks, *11th International Symposium on Wireless Communications Systems (ISWCS)* (2014)

[12] V.D. Vrabie, J.I. Mars and J.L. Lacoum, Modified Singular Value Decomposition by means of Independent Component Analysis, *Signal Proces.* **84**, (2004).

[13] M. Sambridge and K. Mosegaard, Monte Carlo Methods in Geophysical Inverse Problems, *Reviews of Geophysics* **40**, (2002).

[14] S.C. Kirkpatrick, D. Gelatt and M.P. Vecchi, Optimisation by simulated Annealing, *Science* **220**, (1983).