Temporal variations of seismic anisotropy in seismogenic areas: An application to the Parkfield earthquake, 2004, $Mw = 6$.

Stéphanie Durand *

Résumé

Des variations de paramètres physiques, avant et suite au séisme de Parkfield, 2004, Mw 6.0, et pouvant être associés à l'anisotropie sismique, ont été extraites des ondes de surface reconstruites par la méthode de cross-corrélation du bruit sismique. En effet, en cross-corrélant les trois composantes du bruit enregistré passivement à deux stations sismiques, on obtient le tenseur des corrélations, dont les termes non diagonaux contiennent les informations sur l'anisotropie du milieu. Pour les extraire, une rotation du tenseur, minimisant les composantes RT, TR, ZT et TZ, est réalisée de manière à obtenir un tenseur de Rayleigh. Cette opération est réalisée à l'aide de deux angles, l'azimut et le tilt. Le premier est un angle de rotation dans le plan horizontal, autour de l'axe vertical, et le second, un angle de rotation vertical, autour de l'axe radial porté par le grand cercle reliant les deux stations sismiques. Ces deux angles sont sensibles aux variations de l'angle d'incidence du bruit sismique et à l'anisotropie du milieu, notamment à la distribution des fissures. Une étape de "beamforming" confirme la composante variation de l'angle d'incidence du bruit, qui peut se résumer en terme de variations saisonnières. Une décomposition en valeurs singulières permet d'individualiser chacune des contributions et de ne reconstruire que les variations non saisonnières. Appliquée aux données de 2004 et 2005 du réseau HRSN (High Resolution Seismic Network) de Berkeley, situé sur la faille de San Andreas, à Parkfield, on met en évidence une phase de modifications des propriétes du milieu, précurseur au séisme de Parkfield, d'une durée de 20 à 100 jours, ainsi qu'une variation brusque lors du séisme. Confrontés aux résultats de Crampin and Peacock [2008], on trouve que cela correspond à un séisme de magnitude $5.5 - 6.3$, ce qui est en bon accord avec la magnitude réelle du séisme. De plus, la phase précédant le séisme est interprétée comme une phase de coalescence des fissures présentes dans le milieu.

*Stage effectué dans l'équipe de sismologie de l'IPGP, encadré par Jean-Paul Montagner (IPGP), Philippe Roux (LGIT) et Florent Brenguier (IPGP).

Contents

1 Introduction

One of the biggest challenge of seismology is to record spatial and temporal variations of the crustal properties in order to better understand seismicity. The most favorable case would be to observe the complete seismic cycle through changes of some relevant physical parameters, before, during and after earthquakes. Because the Earth behaves as an elastic body, seismic waves are generated when an earthquake occurs and they are dependent on the elastic properties of the medium. It means that their characteristics, such as velocity, polarization, propagation direction, contain informations about the crust structure. That is why many researchers try to monitor crustal properties through the propagation properties of seismic waves. In particular, seismic waves are used to probe crustal anisotropy. Anistropy means that physical parameters of the propagation medium vary with the direction. It must be distinguished from heterogenity, defined as a variation of physical parameters with the location. However, the difference between both can be reduced to a matter of scale. Indeed, considering a stack of parallel layers: such a medium is heterogeneous at a large scale, because of lithology contrasts for instance, but it is anisotropic at small scale, because seismic waves do not propagate at the same velocity in all the directions. Three kinds of anisotropy are distinguished: the stress induced anisotropy, the structural anisotropy and intrinsic anisotropy. The first is induced by alignement of cracks and is also called Extensive-Dilatancy Anisotropy (EDA) of microcracks [Crampin, 1987], while the second is linked to aligned planar features, such as fault zone fabrics, sedimentary bedding planes or aligned mineral/grains [Mueller, 1991, Leary et al., 1990, Kern and Wenk, 1990], and the last one to the anisotropy created by the orientation of the cristalline axis of minerals.

Seismologists use different methods to study spatial and temporal variations in the propagation properties of the seismic waves, one relies on the P and S wave velocity ratio, $\frac{v_P}{v_S}$, and another on the Shear Wave Splitting (SWS). Indeed, the P and S wave velocity ratio has been repeatedly mentioned as an important indicator of temporal variations in crustal properties such as stress-field or fluid distribution in cracks. Semenov [1969] observed a premonitory decrease of $\frac{v_P}{v_S}$ for moderate earthquakes in the Garm region of Tadzik. However, it has not been systematically observed. For instance, McEvilly and Johnson [1974], who did an experiment using artificial sources, did not observe anything. SWS in the crust surrounding active fault has also been extensively studied. It refers to the splits of a shear wave when it encounters a crack. Indeed, when the shear wave arrives in the crack area, it separates into two orthogonal waves, the leading and lagging ones, with different propagation velocities. It is akin to optical birefringence. Assuming weak anisotropy, both final waves are denoted quasi-Sv and quasi-Sh waves, as they slightly differ from the initial shear wave. SWS is then completely described by two parameters, the Leading Shear wave Polarization Direction (LSPD) and the Delay Time (DT) between the leading and lagging shear waves. LSPD permits to discuss the maximal horizontal compressional stress, σ_H, orientation since the LSPD of a vertically propagating shear wave is parallel to it [Crampin, 1987]. Concerning DT, it is proportional to the degree of anisotropy and the length of the ray path in the anisotropic medium. Temporal [Bokelmann and Harjec, 2000, Baisch and Bokelmann, 2001, Tadokoro and Ando, 2002, Saga et al., 2003, Ikuta and Yamaoka, 2004] and spatial [Tadokoro et al., 1999, Iidaka and Hiramatsu, 2009] changes has been observed before earthquakes. However, findings of temporal changes of SWS has been controversial and many studies do not report any temporal changes [Aster et al., 1990, Cochran et al., 2003, Liu and Teng, 2004, Teanby et al., 2004]. It has also been attempted to design an automatic method [Aster et al., 1990] to determine DT and LSPD but it led to more than 200% of errors [Crampin et al., 1991]. Finally, using SWS to monitor anisotropy has several drawbacks. First, it relies on earthquake occurrence so it cannot be done everywhere, neither continuously. Second, sufficiently accurate informations about the source are needed, which is

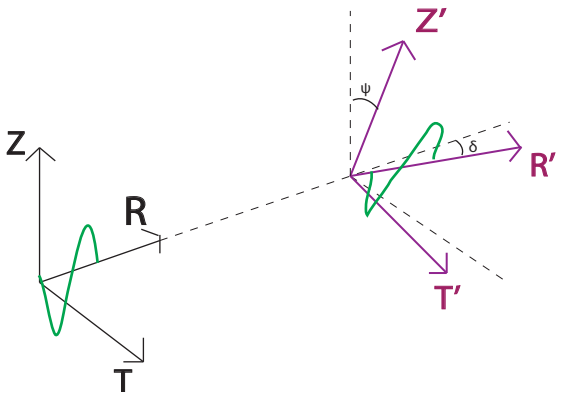

Figure 1: Rotation from the (R,T,Z) frame to a slightly different one (R',T',Z'). The final surface waves are called quasi-Rayleigh and quasi-Love waves.

not always possible. Third, direct surface waves only sample few directions, while many others remain unsampled. That is why, this study is an attempt to design a new method in order to monitor crustal properties but using surface waves reconstructed from cross-correlation of ambient seismic noise. Consequently, it is no longer dependent on earthquake occurrence and the number of sampled directions is significantly increased. Moreover, the anisotropy would be then horizontally integrated instead of vertically integrated when using shear waves. As a result, it should only record the crustal anisotropy and, since the anisotropy is accumulated along the ray path, it should be easier to detect any changes.

The emergence of the Green's function from cross-correlating random fields was first observed in heliosesimology [Duvall et al., 1993], but then theoretically [Draeger and Fink, 1999, Wapenaar and Fokkema, 2005] and numerically studied in ultrasonics [Weaver and Lobkis, 2001] and acoustics [Derode et al., 2003a,b, Larose et al., 2005, 2006, 2008]. Actually, whatever the kind of propagating wave is considered, it is a key point to recover the Green's function. That is why seismologists were also interested in this cross-correlation method since it permits to recover the impulse response of the Earth between two passive sensors. Since it demands to be provided with random fields, the application in seismology is only possible for coda waves [Campillo and Paul, 2003, Paul et al., 2005] and seismic noise [Shapiro and Campillo, 2004, Shapiro et al., 2005, Stehly et al., 2006, 2007, Bensen et al., 2007]. It was first used to make tomographic images of the crust. Indeed, since surface waves are reconstructed, it is then possible to measure group velocities for each interstation path so that an image of the upper crust is reconstructed. It has been a significant step in seismologic tomography: the method increases the fidelity and resolution of the crustal images obtained from surface wave analysis. Another important feature of the cross-correlation method is that when it is applied to the three seismic components, namely radial R, transverse T and vertical Z, it is possible to build the correlation tensor and the non-diagonal terms are then available. Considering an isotropic medium, the correlation tensor would be diagonal since polarization directions would merge with the (R, T, Z) frame. However, now considering an anisotropic medium, the correlation tensor would be full of non zero terms and, in particular, the anisotropic properties would be stored in the non diagonal terms. Indeed, they mean that energy is present on them so that the reconstructed wave is not propagating in the seismic frame (R, T, Z), but in a slightly different one. Assuming weak anisotropy, such surface waves are called quasi-Rayleigh and quasi-Love waves (figure (1)). Therefore, it appears that such a method would be able to monitor the crustal properties.

The method is applied to ambient seismic noise data, continuously recorded by the High

Resolution Seismic Network (HRSN) of Berkeley, located on the San Andreas Fault (SAF), since 2001. Moreover, the Parkfield earthquake, 2004, Mw 6.0, was one of the best recorded earthquake, with data of unprecedent quantity and quality. That is why, to test wether the method is able to monitor anisotropy, it has been applied to the 2004 data and to test the robustness it has also been applied to the 2005 data, where no significant earthquake occured. Anisotropy has already been studied in the Parkfield region. Zhang et al. [2007] produced a map of spatial anisotropy distribution by back projecting shear wave splitting DT along ray paths derived from 3D shear wave velocity model, assuming DT are accumulated along the ray path. They used data from the PASO and HRSN seismic networks of some 575 earthquakes. It appears that anisotropy largely varies in space and it seems to be quite complicated. SWS has also been extensively studied in the SAF region, and in particular in the Parkfield segment [Liu et al., 1997, 2008, Boness and Zoback, 2004, Cochran et al., 2006]. Large spatial variations of the LSPD has been reported [Liu et al., 1997, 2008] suggesting that cracks are σ_H parallel in the surrounding of the fault, but fault parallel very close to the fault. Concerning temporal changes, a recent study [Niu et al., 2008] reports preseismic changes in velocity from active source monitoring at SAFOD drill site. They showed stress dependence of the seismic wave velocity: the time needed for the wave to travel along a fixed path appeared to be anticorrelated with the barometric pressure. Moreover, two large excursions in travel time data were coincident with two earthquakes of small size, $Mw = 1$ and $Mw = 3$. However, concerning the Parkfield event, no precursory changes has been observed. Brenguier et al. [2008] reported variations in the ratio $\frac{\delta\tau}{\tau}$ just after the event so that changes should, at least, be seen after the event.

The report first presents the method. Then, a section is devoted to the data, their origin and the applied pre-processing. A third part descibes the processing, followed by one presenting the results. Finally, a dicussion of the method and the results is proposed before concluding.

2 The method

The method is made of two steps. First, the cross-correlation method is applied to ambient seismic noise. Second, parameters describing anisotropy are extracted. In this section, the theory of the emergence of the Green's function from cross-correlation of ambient seismic noise is exposed, followed by the description of the procedure designed to extract anisotropy properties.

2.1 On the emergence of the Green's function from correlation of ambient seismic noise

It has been first established in time-distance helioseismology [Duvall et al., 1993] that the position of the peak of the envelope of a cross-correlation function can be interpreted as the travel time of a wave packet between two spatially separated locations on the Sun. It means that the Green's function of a heterogeneous medium can be estimated between two passive sensors by cross-correlating the recorded random fields. It has been then theoretically demonstrated and numerically tested in ultrasonics [Weaver and Lobkis, 2001] before being widely studied in acoustics [Derode et al., 2003a,b, Larose et al., 2005, 2006, Wapenaar and Fokkema, 2005]. To begin with, let us try to interpret the emergence of the Green's function in the case of a homogeneous closed cavity, that is to say with reflecting boundaries. Two receiving points, A and B, and a source point C are considered. We will note $h_{IJ}(t)$ the scalar wave field sensed in I when a Dirac $\delta(t)$ is sent by J. Then, denoting $e(t)$ the excitation function at C, the wave fields, ϕ_A and ϕ_B, received in A and B, can be written:

$$\phi_A(t) = e(t) \otimes h_{AC}(t) \quad \text{and} \quad \phi_B(t) = e(t) \otimes h_{BC}(t),$$

where ($\otimes$) means convolution. Therefore, the cross-correlation C_{AB} of the two fields is:

$$C_{AB}(t) = \int \phi_A(t+\theta)\,\phi_B(\theta)\,d\theta \tag{1}$$

$$= h_{AC}(-t) \otimes h_{BC}(t) \otimes f(t), \tag{2}$$

where $f(t) = e(t) \otimes e(-t)$. The impulse response $h_{AB}(t)$ does not appear. However, the impulse response satisfies the following "cavity equation" [Draeger and Fink, 1999]:

$$h_{AC}(t) \otimes h_{BC}(t) = h_{AB}(t) \otimes h_{CC}(t).$$

The correlation function C_{AB} is then:

$$C_{AB}(t) = h_{AB}(t) \otimes h_{CC}(t) \otimes f(t). \tag{3}$$

Consequently, for a given h_{CC}, at least in the frequency domain imposed by the spectrum of $f(t)$, it is possible to extract the Green's function of the medium by simply deconvolving h_{CC}.

Now, is this still true for an open medium ? In this case, the "cavity equation" does not hold anymore, so that another physical argument is needed to make h_{AB} appear into the correlation function C_{AB}. Let us demonstrate that it still is possible to retrieve the Green's function and it requires to consider several sources judiciously distributed in the medium. To that end, the experiment is analysed in terms of time-reversal (TR) symmetry and spatial reciprocity. Indeed, as there is no flow into the medium, the propagation is reciprocal, $h_{IJ}(t) = h_{JI}(t)$. Now, cross-correlating the impulse response received in A and B and using the reciprocity property, we obtain that $h_{AC}(t) \otimes h_{BC}(-t) = h_{CB}(-t) \otimes h_{AC}(t)$. Then, imagine the following TR experiment is performed: B is no longer a receiver but a source, sending a pulse recorded at C as $h_{CB}(t)$, then time-reversed and sent back to A, so that the resulting wave field observed at A is $h_{CB}(-t) \otimes h_{AC}(t)$, which is exactly the same as $h_{AC}(t) \otimes h_{BC}(-t)$ because of the reciprocity. However, it generally differs from h_{AB}. Yet, let us go further and consider now several source points C, arranged in such a way that they form a perfect TR mirror, which means that they are distributed on a surface surrounding the receiving points A and B, according to the Helmholtz-Kirchhoff theorem. During the forward propagation, a pulse is sent by A for instance and propagates in the medium, including B, recording $h_{AB}(t)$, and all the sources C, recording $h_{AC}(t)$. When the field is time-reversed, it should exactly go backward and refocus in A at time $t = 0$ so that the field received in A after the TR is $h_{BA}(-t)$. But, it does not stop here, and the wave diverges again toward B and gives rise at times $t \gg 0$ to $h_{AB}(t)$. Consequently, such a distribution of sources leads to the following relationship:

$$\sum_C h_{AC}(-t) \otimes h_{CB}(t) = h_{AB}(t) + h_{BA}(-t).$$

The left-hand-side is referred to as source averaging and we showed that it actually contains the Green's function. Therefore, it still is possible to recover the Green's function from the cross-correlation function and it is even easier as we do not need any deconvolution by h_{CC}. A trade off exists between the number of sources and the level of correlation: the higher the number of sources, the higher the correlation coefficient. All these theoretical considerations have been done in addition to numerical modelling [Derode et al., 2003a, Larose et al., 2006], which permits to visualize the emergence of the Green's functions. It is noteworthy that, in the case of a homogeneous open medium, if the sources do not fulfill the condition of a TR mirror, then the Green's function cannot be rebuilt by cross-correlation function [Larose et al., 2006].

However, let us finally address the case of a heterogeneous open medium, containing a large number of scatterers randomly distributed all around the receiving points. Numerical experiments showed that in this case, even if the sources are not isotropically distributed within the medium, the Green's function still emerges from the correlation functions [Derode et al., 2003b, Larose et al., 2005, 2008]. It seems that the multiple scattering compensates the asymmetrical distribution of the sources. Let us note that, in a scattering medium, not only ballistic wave will be reconstructed, but also a coda.

Until now, the field measured in A and B originated from an active and coherent source. But, since the aim is to apply the theory to seismology, we must consider more realistic configurations, such as a medium with diffuse noise. Such a field exists in seismology, it is the ambient seismic noise [Shapiro et al., 2005, Shapiro and Campillo, 2004, Stehly et al., 2006, 2007, Bensen et al., 2007]. It is mostly made of surface waves, mainly excited by loads caused by pressure perturbations in the atmosphere and the ocean. Making the assumption that the seismic noise comes from virtual source points randomly distributed everytwhere in the medium and continuously generating uncorrelated white noise, that is to say seismic noise is a very good approximation of a fully diffuse field, then the Green's function still emerges from the cross-correlation functions [Campillo, 2006]. Indeed, a diffuse field in an elastic body can be written as:

$$\phi(t) = \sum_n a_n u_n(x) \exp^{i\omega_n t},$$

where x is the position, t the time, ω_n and u_n the eigenfrequencies and eigenfunctions of a real Earth, and a_n are the modal excitation coefficients. The latter are uncorrelated, as a diffuse field is considered, and it can be expressed as:

$$\langle a_n a_m^* \rangle = \delta_{nm} F(\omega_n),$$

where F is the spectral energy density. Finally, the correlation between two fields at location x and y is:

$$C(x, y, t) = \sum_n F(\omega_n) u_n(x) u_n(y) \exp^{-i\omega_n t}. \quad (4)$$

This is nearly the Green's function, it differs only by an amplitude factor, F. Therefore, it is possible to recover the Green's functions by simply cross-correlating seismic noise but let us remind that the assumption behind is that the seismic noise is a diffuse field. It means that the sources are randomly distributed everywhere and generate uncorrelated white noise. This condition is often respected by considering long time series which is the same as stacking the cross-correlation functions over a certain number of days, so that you randomize the noise sources. However, it was also mentioned that it does not need to be perfectly randomized because scattering helps in reconstructing the Green's function.

2.2 The Optimal Rotation Algorithm (ORA) procedure

Thanks to the cross-correlation method, it is possible to build the complete 9 components correlation tensor of a surface wave from the cross-correlation of ambient seismic noise. In the case of an isotropic medium, this tensor is diagonal, so that the wave is polarized into three perpendicular planes merged with the R, T and Z components. However, in the most general case of an anisotropic medium, the tensor also contains non zero transverse components. It means that the isotropic polarization frame is rotated into another one slightly different, so that the reconstructed surface waves are called quasi-Rayleigh and quasi-Love. Consequently, such a general tensor contains informations about the anisotropy. In order to recover them,

a strategy is to rotate the general obtained tensor in such a way that a Rayleigh tensor is obtained. A Rayleigh tensor has its RT, TR, ZT and TZ components null, because a Rayleigh wave is polarized in the vertical plane and has a retrograde elliptical displacement so there is no displacement on the T component. Therefore, the idea is to solve an inverse problem: to search for the orientations of the two stations of a receiver pair so that we cancel out the RT, TR, ZT and TZ components. To do so, four angles are used, two per station: the azimuth, ψ, and the tilt, δ. The first one refers to an angle of rotation in the horizontal plane and the second is an angle of rotation around the radial component. The principle is to explore the parameter space, constituted by the four angles belonging to $[0;\pi]$. In order to reduce the computing time the search begins by finding the best azimuths for both stations, then the stations being azimuthally rotated, it continues searching for the best tilts and finally, both stations are free to rotate around the four angles previously found. A parameter evaluating the success of the rotation is needed. It must quantify the reduction of energy present on the RT, TR, ZT and TZ components for each tested rotation. It is done via what is called the misfit parameter, M, definied as:

$$M\left(\psi_i,\psi_j,\delta_i,\delta_j\right)=\frac{\sum_{k=2}^{3}\sum_{l=2}^{3}\left[C_{ij}^2\right]_{kl}}{\sum_{k=1}^{3}\sum_{l=1}^{3}\left[C_{ij}^2\right]_{kl}}, \tag{5}$$

k and l denoting the R, T and Z components as following $1=Z$, $2=R$ and $3=T$, and $[C_{ij}]_{kl}$ being the correlation function between the component k of station i and l of station j. The less the misfit, the closer to a Rayleigh tensor the correlation tensor. Consequently, the optimal rotation is the one for which the misfit is minimum.

We designed an algorithm, called Optimal Rotation Algorithm (ORA) [Roux, 2009], realizing the inverse problem. Let us precise a little the procedure done by the ORA. First, considering each station separately, let us build the rotation matrices individualy before dealing with the total rotation of the correlation tensor. Considering a station i, the matrices are the following :

$$mat_\psi\left(i\right)=\begin{pmatrix}1 & 0 & 0\\ 0 & \cos\phi_i & \sin\psi_i\\ 0 & -\sin\psi_i & \cos\psi_i\end{pmatrix}\quad\text{and}\quad mat_\delta\left(i\right)=\begin{pmatrix}\cos\delta_i & 0 & -\sin\delta_i\\ 0 & 1 & 0\\ \sin\delta_i & 0 & \cos\delta_i\end{pmatrix}, \tag{6}$$

mat_ψ, respectively, mat_δ, refers to the rotation matrix for the azimth, respectively, for the tilt. As a result, the complete operation for the rotation of the correlation tensor involving the stations i and j is:

$$\begin{pmatrix}Z_iZ_j & Z_iR_j & Z_iT_j\\ R_iZ_j & R_iR_j & R_iT_j\\ T_iZ_j & T_iR_j & T_iT_j\end{pmatrix}=mat_\psi\left(i\right)*mat_\delta\left(i\right)*\begin{pmatrix}E_iE_j & E_iN_j & E_iZ_j\\ N_iE_j & N_iN_j & N_iZ_j\\ Z_iE_j & Z_iN_j & Z_iZ_j\end{pmatrix}\\ *\left(mat_\delta\left(j\right)\right)^T*\left(mat_\psi\left(j\right)\right)^T, \tag{7}$$

where (T) denotes the transpose matrix and $(*)$ the matricial product. It is noteworthy that $(*)$ is not commutative in this operation. Indeed, if the rotation is done in a different order, first in tilt and then in azimuth, the matrices involved in the product have to be reversed.

3 The Data

The method is applied to ambient seismic noise data recorded by HRSN. They cannot be directly used, a pre-processing phase is needed. It consistes in two steps: first the single station data preparation, including temporal normalization, spectral normalization and band pass filter, and second the cross-correlation calculation, including stacking and rotation of the tensor so that it is built in the (R,T,Z) frame.

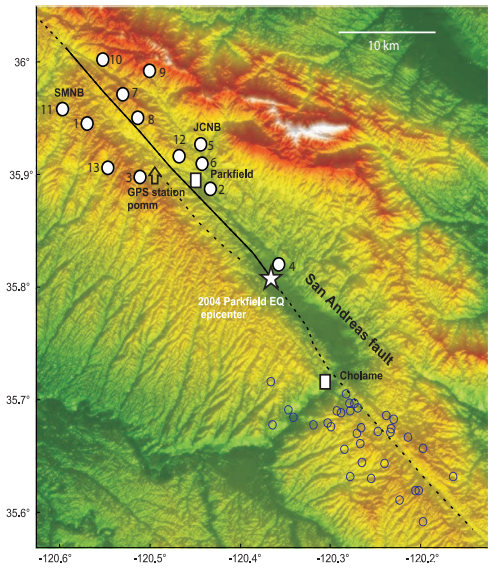

Figure 2: Location of the seismic stations of HRSN. The blue circles indicates to non volcanic tremor detected in the Cholame region. ¡The equivalence between the real name of the stations and their attributed number is as followed: $1 = CCRB$, $2 = EADB$, $3 = FROB$, $4 = GHIB$, $5 = JCNB$, $6 = JCSB$, $7 = LCCB$, $8 = MMNB$, $9 = RMNB$, $10 = SCYB$, $11 = SMNB$, $12 = VARB$, $13 = VCAB$.

3.1 The Parkfield, California, Earthquake Experiment

The Parkfield Experiment [Bakun and Lindh, 1985, Roeloffs and Langbein, 1994] is a comprehensive, long-term, earthquake research project on the SAF (135°N, 86°SE), a strike-slip plate boundary, accommodating the convergence between the Pacific and the North American lithospheric plates. Led by the USGS and the State of California, the purpose of the experiment is to better understand the physics of the earthquakes by observing the fault and surrounding crust at close range at the time before, during and after an earthquake. It led to a dense network of instruments poised to capture the anticipated earthquake and reveal the earthquake process in unprecedent details. In particular, the seismic network was enlarged with more short period, broad-band and strong-motion instruments. HRSN is one those seismic networks operated by Berkeley Seismological Laboratory. It is a borehole instrumentation array initially compounded of 10 boreholes. But, as it well recorded earthquake data, it was upgraded in 1987 with accurate timing and triggering capabilities to capture the approximately daily local microearthquakes. Since, 2001 the network was expanded to 13 stations with 3-component Mark Products L22 or Geospace HS1 or Litton 1023 sensor as equipment, recording at 20 samples per second. A map with the station repartition is given figure (2).

The town of Parkfield, located near to the SAF, experienced six moderate earthquakes of magnitude 6 in 1857, 1881, 1901, 1922, 1934, 1966 with similar characteristics and an average recurrence interval of 22 ± 3 years [Bakun and McEvilly, 1984]. The Parkfield seismic events are of similar magnitude and ruptur extend, which supports the concept of fault segmentation. Indeed, the peculiar fault segment of Parkfield is generally understood to be a transition zone between the 170 km long creeping portion to the northwest and the 300 km long locked portion to the southeast. The Parkfield segment is also charactarized by an abrupt cross-fault velocity gradient, with a 5 to 20 per cent lateral change in velocity in a 4 km wide zone parallel to the

fault trace. Such a series of similar periodic earthquakes is called a characteristic earthquake sequence. Considering the periodicity, Bakun and Lindh [1985] conjectured another event would occur before 1993. The Parkfield Earthquake Prediction Experiment began in 1985, after this conjecture was accepted by the National Earthquake Prediction Evaluation Council (CEPEC). The goals were to record geophysical signals accompanying the earthquake process at Parkfield, to capture effects of strong motion in the near field, and, most ambitiously, to attempt to issue a warning up to 3 days before the magnitude 6 event if the seismicity and/or the fault creep rate reach predefined threshold levels. However, the waited next earthquake only happened in 2004, 28^{TH} september [Langbein et al., 2005, Bakun et al., 2005]. Although the area was intensively monitored, no significant precursors were detected for the 2004 event. However, because of the density and diversity of the instrumentation located at Parkfield, the quality and quantity of data record from the 2004 event are unique and should provide new insights into the many aspects of the earthquake cycle.

Understanding the behaviour of the seismic activity along an active fault like the SAF system is a first step toward prediction. The Parkfield experiment shows that even in a situation where a characteristic earthquake occurs around a limited area of well studied fault, the prediction in time is cumbersome. Speculation on the reasons for the unsuccessful prediction include effect of several large earthquakes in the neighbourhood, influencing the stress pattern around Parkfield, and the assumption on the models behind the repeat times.

3.2 The pre-processing procedure

3.2.1 Single station data preparation

The purpose of this phase is to remove anything that tends to obscur the ambient seismic noise, namely earthquake signal and instrumental irregularities so that you finally get broad-band ambient noise. The obscuration caused by earthquakes is most severe above 15 s, so this step in data processing is important at periods longer than the microseism band ($\approx 5- \approx 17$ s). Moreover, as the spectral amplitude of the ambient noise peaks in the microseism band, we must use methods designed to extract longer period ambient noise from the seismic records.

A first step is the spectral normalization. It consists in band broadening the ambient noise but also to avoid degradations due to persistent monochromatic sources. Indeed, ambient noise is not flat in the frequency domain: it peaks near the microseisms periods (≈ 15 and ≈ 7.5 s) and rises at very long period above 50 s to form what is called the Earth "hum".

A second step, which is the most important, is the temporal normalization. It consists in reducing the effects on the cross-correlations of earthquakes, instrumental irregularities and non-stationary noise sources near stations. Several methods exist but only two were tried: 1-bit normalization and iterative clipping. The principle of the first one is quite simple: it consists in keeping only the information about the sign of the raw signal. To do so, all the positive amplitudes are replaced with a 1 and all the negative with a -1. Therefore, even if it is a drastic method, it really increases the signal-to-noise ratio (SNR). The second method, iterative clipping, is much sophisticated. For this one, any amplitude above a specified multiple of daily rms amplitude is down-weighted, and this is done iteratively until the entire waveform is below a certain level, fixed at six times the daily rms level. This method also gives good final SNR.

Finally, we applied a band-pass filter between 0.075 and 0.25 Hz.

First, iterative clipping, without any whitening neither band-pass filter, has been performed but it revealed, when looking at the cross-correlations, not to be sufficient. There were problems of monochromatic sources and the data needed to be filtered before computing the cross-correlations. Consequently, we chose to apply spectral whitening, 1-bit normalization, and to

filter between 0.075-0.25 Hz. This procedure is actually the basis of any study using cross-correlation of ambient seismic noise [Bensen et al., 2007].

3.2.2 Cross-correlation functions, rotation and stacking

The cross-correlations were performed between all the possible receiver pairs, representing 78 couples, and a selection of data is done later on. Due to the linearity of the cross-correlation procedure the stacking can either be done before or after the cross-correlation calculation. We prefered to do it after so that we realized stability tests before choosing the number of stacked days.

The nine-component noise correlation tensor $C_{ij}(t)$ is computed for each receiver pair $\{i, j\}$ as:

$$[C_{ij}(t)]_{kl} = \frac{\int_0^T S_{ik}(\tau)\, S_{jl}(t+\tau)\, d\tau}{\sqrt{\int_0^T S_{ik}^2(\tau) d\tau \int_0^T S_{jl}^2(\tau) d\tau}}, \tag{8}$$

where $\{k, l\}$ refers to R, T, and Z, components of the data seismograms $S_{ik}(t)$ and $S_{jl}(t)$. The choice to normalize the correlation function according to (8) implies that the amplitude indicates how well the components k and l are correlated, without taking into account the total energy of all the components. Moreover, the cross-correlation functions are two-sided time functions, typically stored between -120 s and 120 s. The positive, respectively the negative, part is called the causal, respectively the anticausal, signal.

A rotation of the cross-correlation tensor is then performed. Indeed, as they were computed between (E,N,Z) components and since we would like to recover the Green's functions between stations of each receiver pairs, it is necessary to change the coordinate frame to one related to the station couple, (R,T,Z). The latter is oriented such as the vertical axis is pointing up, the radial axis carried by the great circle linking both stations of the couple and the transverse one perpendicular to the two previous ones, such as to form a direct frame. The rotation of cross-correlation is not trivial and it is important to specify the considered angle in order to avoid any mistake of sign. First, let us note that this rotation does not affect the vertical axis at all, so we just need to turn the E and N axis to the R and T ones. Denoting ϕ_{ij} the azimuth of a receiver pairs including stations i and j, we can write the rotation of one station as the following operation :

$$\begin{pmatrix} Z_i \\ R_i \\ T_i \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ \sin\phi_{ij} & \cos\phi_{ij} & 0 \\ \cos\phi_{ij} & -\sin\phi_{ij} & 0 \end{pmatrix} \begin{pmatrix} E_i \\ N_i \\ Z_i s \end{pmatrix} \tag{9}$$

It is noteworthy that ϕ_{ij} is taken between 0° and 360° so that the orientation of the receiver pairs is kept. Indeed, to cross-correlate station i with station j, for instance, is not the same as to cross-correlate station j with i. Now, let us write the operation needed to rotate the entire cross-correlation tensors:

$$
\begin{pmatrix} Z_iZ_j & Z_iR_j & Z_iT_j \\ R_iZ_j & R_iR_j & R_iT_j \\ T_iZ_j & T_iR_j & T_iT_j \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ \sin\phi_{ij} & \cos\phi_{ij} & 0 \\ \cos\phi_{ij} & -\sin\phi_{ij} & 0 \end{pmatrix} \begin{pmatrix} E_i \\ N_i \\ Z_i \end{pmatrix}
$$
$$
\begin{pmatrix} E_j & N_j & Z_j \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ \sin\phi_{ij} & \cos\phi_{ij} & 0 \\ \cos\phi_{ij} & -\sin\phi_{ij} & 0 \end{pmatrix}^T \tag{10}
$$
$$
= \begin{pmatrix} 0 & 0 & 1 \\ \sin\phi_{ij} & \cos\phi_{ij} & 0 \\ \cos\phi_{ij} & -\sin\phi_{ij} & 0 \end{pmatrix}
$$
$$
\begin{pmatrix} E_iE_j & E_iN_j & E_iZ_j \\ N_iE_j & N_iN_j & N_iZ_j \\ Z_iE_j & Z_iN_j & Z_iZ_j \end{pmatrix} \begin{pmatrix} 0 & \sin\phi_{ij} & \cos\phi_{ij} \\ 0 & \cos\phi_{ij} & -\sin\phi_{ij} \\ 0 & 0 & 1 \end{pmatrix}.
$$

A cross-correlation tensor in the (R,T,Z) frame is obtained and figure (3) shows an example. It is important to note that the cross-correlation process leads to the reconstruction of a wave but it is also noticeable that the amplitude of the RZ component can be higher than that of the RR component, for instance. This is possible because of our normalization. Indeed, since the total energy is not taken into account when normalizing the correlation functions, we lost the relative importance of the various components and only the information about how well both involved components do correlate is kept. It is also noteworthy that the Green's function is not always very well reconstructed. Moreover, the correlation functions are not symmetric with respect to zero time. It means that the energy is coming from a particular direction, instead of being distributed on both sides of the array. For instance, looking at the tensor figure (3), the wave peak is in the negative times, so the energy is propagating from station 11 to station 10. According to the map figure (2), it means that the noise comes from west. This is confirmed by realizing beamforming tests. Indeed, we showed that the network mainly receives the energy from the Pacific ocean, such as the incidence angle of the noise typically varies in between 40 to 70°N. The beamforming results are discussed in more details later, in section 4.1. Being provided with those correlation tensors, a selection of data is done. Indeed, some stations revealed not to be reliable or even out of order. For instance, station 4 does not work and station 6 clearly has troubles in recording of the horizontal components.

Finally, in order to improve the SNR, the cross-correlation functions are stacked. The number of needed days is chosen so that a stable waveform is obtained. First, the daily cross-correlation functions are stacked day per day along the year, each time normalizing to the maximum amplitude, and we looked for the minimum number of days requiered to have a stable waveform. It shows 30 days are sufficient (figure (4)). Then, two tests of stability are made: one where the days are stacked along the year with a moving window of 30-days and a second where the 30 days are randomly chosen into the year (figure (5)). Whatever is the stacking method, the waveform is completely stable after stacking over 30 days. Very subtil amplitude variations can be observed when stacking along the year but it is negligible.

4 The data Processing

This section is devoted to the study of the origin of the ambient seismic noise recorded in California and, more locally, at Parkfield, and then to the ORA procedure, in particular to the meaning of its outputs.

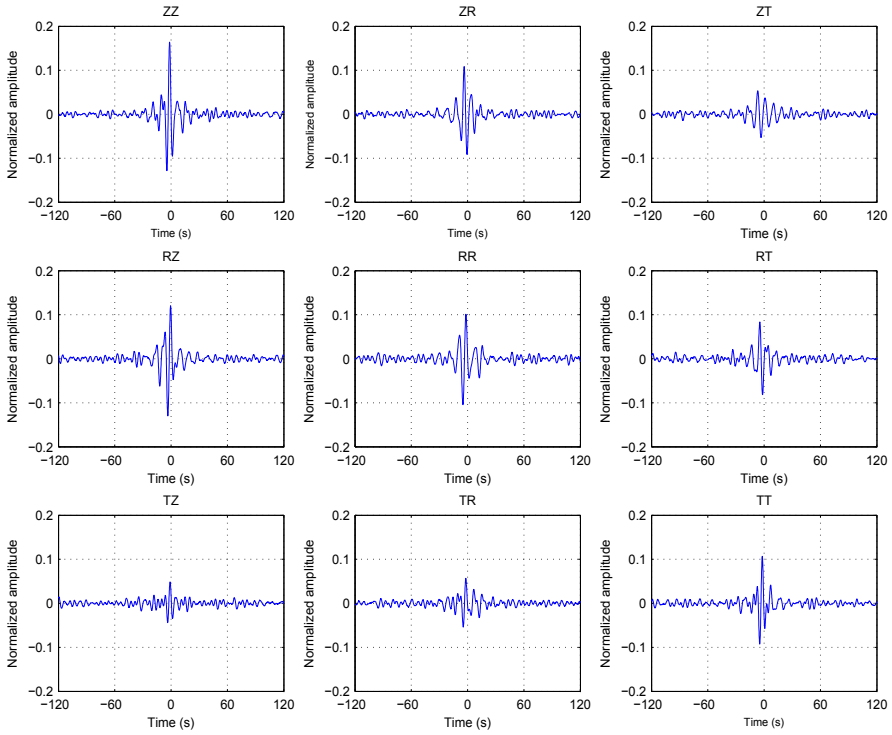

Figure 3: Correlation tensor of the station pair $10 - 11$.

4.1 The origin of the seismic noise

The correlation functions are not symmetric with respect to zero time and this means that the energy flux propagating between both receivers is not the same in both directions. In [Stehly et al., 2006], the authors studied the origin of the noise in California by simply taking 23 stations located in southern California and separated by few hundreds of kilometers. After stacking the correlation functions over 15 days along the year 2003 and normalizing by the square root of the interstation distance, they obtained diagrams of the amplitude of the correlation function along the year depending on the azimuth. The maxima that can be observed on such diagrams indicate the main direction of the normalized background energy flow across the array. They did this considering different period-bands: $5 - 10$, $10 - 20$, $20 - 40$ s. The results showed that the origin of the noise is not the same for all the period-bands. As we filtered our data between $0.075 - 0.25$ Hz let us look only at the results for the two first period-bands. It is noteworthy that between $5 - 10$ s the noise mainly comes from the Pacific ocean for a range of azimuth going from 220° to 225°N. However, the noise behaves quite differently in the period-band $10 - 20$ s. It still exhibits the principal direction of $220 - 225°$, but only during the summer, and it changes in winter to two other directions, 315° and 45°. The change between the two main regimes takes place in March and October. Therefore, this study already suggests that the noise sources are not randomly distributed around California.

Considering our network, a frequency-incoherent beamforming on the processed one day long seismic noise data, using the $N = 13$ stations of HRSN, is performed [Roux, 2009]. It permits to determine the average velocity c and direction θ_0 of the seismic noise. The beamforming $B(\theta, c)$

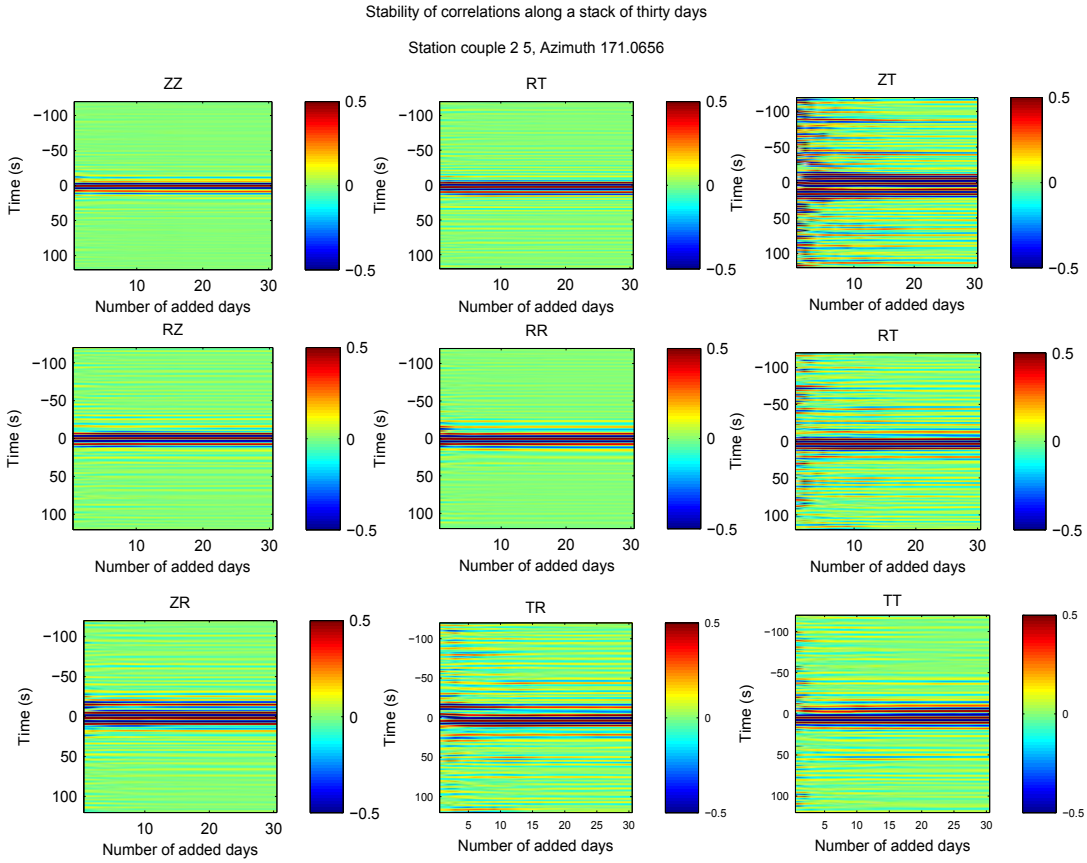

Figure 4: Stability test of all the components of the correlation tensor which permits to determine that 30 are sufficient to obtain a stable waveform. The colorbar represents the normalized amplitude.

can be written:

$$B(\theta, c) = \frac{1}{\Delta\omega} \int_{\omega_c - \Delta\omega/2}^{\omega_c + \Delta\omega/2} \left| \sum_{i=1}^{N} \tilde{S}_i(\omega) \exp\left[i \frac{\omega}{c} (x_i \sin\theta + y_i \cos\theta) \right] \right|^2 d\omega, \tag{11}$$

where ω_c is the central noise frequency and $\Delta\omega$ the frequency bandwidth, $\tilde{S}_i(\omega)$ is the complex Fourier component at frequency ω of the vertical component $S_i(t)$ on the i^{th} seismic station, with (x_i, y_i) the spatial coordinates of station number i. The results show that the azimuth θ_0 significantly varies day after day, it oscillates between 20° and 100° with some excursions to 160° (blue curve on figure (6)). In order to have less "noisy" variations, the signal is averaged using a 30-days large moving window. Then, the azimuths are between 40° and 70° and look like seasonal variations as it simply is an oscillation (red curve on figure (6)). This test shows that the noise source clearly comes from the Pacific ocean and whatever the number of stacked days is, it is nearly impossible to randomize the field.

4.2 The ORA and its outputs

The ORA performs the rotation of the station belonging to a receiver pair, minimizing the energy present on the transverse components of the correlation tensor. It does not rotate the whole correlation function, from -120 s to 120 s, to avoid a too important influence of the coda. Indeed, in the most general case, in addition of the emergence of a surface wave, we also get a coda from the cross-correlation calculation. The coda can be interpreted as the real coda waves recorded on seismograms, that is to say as emerging from the multiple scattering in the medium. Two points are relevant to highlight: on one hand this part of the signal could be used to randomize the noise field in order to reduce the asymmetrical distribution of the noise but on the other hand it also destabilizes the signal and disturbes the ORA when it tries to minimize the RT, TR, ZT and TZ components. It is due to the energy added which does not originate from a surface wave but from its multiple scattering. That is why, in order to get ride of the main part of the coda, the correlation functions are cut from -20 s to 20 s. This is arbitrarily chosen because it is not exactly known when the coda begins, however it is decided after having tested several time windows, retaining the one giving the best misfits.

Now, let us discuss about the meaning of the outputs of the ORA, namely the azimuth, ψ, and the tilt, δ. In particular, can they be used to monitor anisotropy ? The ORA lets the station of each receiver pair to turn themselves in such a way that a Rayleigh wave is reconstructed. In reality, the optimisation is not perfect and the obtained wave is still a quasi-Rayleigh wave, mixed with a quasi-Love wave. As noise source is directive, the stations should oriente themselves in this direction and so the azimuth should mostly represent the incidence angle of the noise.

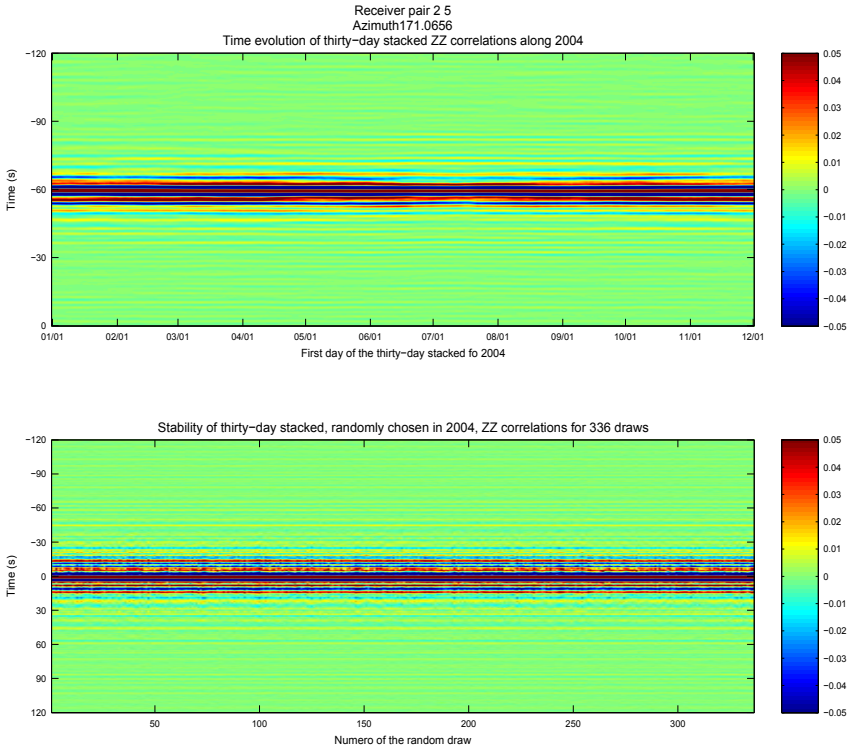

Figure 5: Stability tests confirming the stability of the signal ontained when 30 days are stacked. The colorbar indicates the mean amplitude.

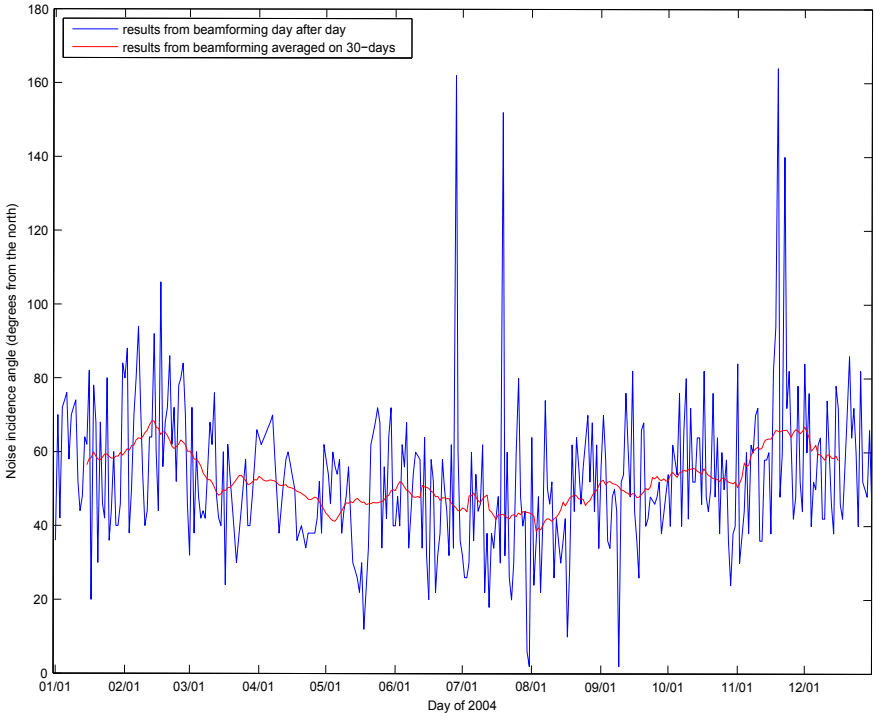

Figure 6: Time evolution of the noise incidence angle along 2004 obtained from beamforming. The blue curve is day per day and the red curve is the blue one averaged over 30 days using a moving windows along 2004.

But, another way how to interpret the azimuth is to consider azimuthal anisotropy. Indeed, if we consider that there is a certain crack distribution, different from the orientation of the receiver pairs, then even if the couple is aligned with the noise source, the final azimuth will be different. So, it seems that the azimuth is sensitive to both, incidence angle of the source and crack distribution.

Concerning, the tilt, it actually measures a coupling between the T and Z components. Therefore, it is also related to anisotropy since the coupling is a characteristic of waves crossing a complex anisotropic medium which affects the polarization of the wave. To illustrate this phenomenon, let us consider a medium crossed by some vertical crack planes, for instance. A wave propagating in such a medium would see its polarization modified so that the tilt records the anisotropy. However, like the azimuth, the tilt is also sensitive to the source incidence angle. Indeed, considering two waves of different incidence angle, they will not have the same final tilt after having crossed cracks.

To summarize, we have different cases:

- if the azimuth and the tilt are null, it means that we do not need any rotation to recover the Rayleigh tensor so that we are in the condition of application of the theory and the medium is isotropic.

- If the tilt is null and the azimuth is not, then it either means that the station pair is not aligned with the source or that there is a vertical crack plane oriented in a different way of the source or both.

- If the azimuth is null and the tilt not, it means that we have an anisotropic medium but its anisotropic feature is mainly oriented in the direction of the station pair, which is itself aligned with the source.
- Finally, the most general case is if the azimuth and the tilt are not null. It means that the wave is propagating in an anisotropic medium but we can distinguish two cases. If the azimuth is the same as the source incidence angle, the anistropy only creates a coupling of the Z and T components. But, if the azimuth differs from the incidence angle of the source, it means that there is also azimuthal anisotropy which can be linked to crack distribution within the upper crust.

It is important to note that a good approximation of the crack distribution within the uppermost crust surrounding the SAF is the one of a transversely isotropic medium with horizontal axis (HTI). Indeed, as the SAF is a strike fault the direction of the maximal stress is horizontal, oriented north/south. Therefore, the stress applied should mostly open NS vertical cracks, while closing EW vertical cracks. Since earthquakes occur when a stress threshold is reached and because ORA outputs are sensitive to the crack distribution, it is reasonnable to think that the method is adapted to the monitoring of the crustal properties. At least, without taking the absolute value of the azimuth and tilt into account, the temporal variations of those parameters must reflect temporal variations of the crustal properties.

5 The Results

5.1 The raw results from the ORA

First, a selection of the best stations pairs using the misfit (figure (7)) is needed. The maximum accepted mifit is arbitrarily chosen. As the misfit measures the remaining energy on the RT, TR, ZT, and TZ components, normalized by the total energy, a value of 0.1 means that 10% of the total energy is still present on the transversal components of the correlation tensor. Consequently, a reasonable extreme value for the misfit is 0.05, and the best couples are those with a misift inferior to it.

Once this is done, the temporal variations of the azimuth and the tilt for these best couples are available. A first observation is that it is difficult to see any clear variations close to the earthquake (figures (8)). However, considering the azimuth of receiver pairs involving station 12 only (figure (9)), an abrupt variation is visible at the end of September. The task is to discriminate between signal related to the earthquake, station troubles and local effect ? Before trying to answer, let us consider the azimuth of couples involving station 8 (figure (10)). All these station pairs exhibit a significant variation in April. It leads to the same question as before. The fact that both stations, when they are involved in a couple, exhibit large temporal variations that are nearly absent when considering other couples, is doubtful. It is not possible to reject the fact that it can be a local effect or even a station problem. However, HRSN published reports show that stations 8 and 12 had no problem during 2004 which enable us to reject the last hypothesis. So, there still are two possibilities: it is either earthquake related signal or a local effect. To avoid any ambiguities two sets of station pairs are formed, each containing the best couples but one, denoted (WHTOUT04), includes couples without stations 8 and 12, and another, denoted (WTH04), includes in addition station pairs involving stations 8 and 12 (figure (11)).

To continue, let us discuss the principal features of azimuth results. If one particular station pair is considered, the ORA does not find the same azimuths at both stations. It is a priori surprising but in reality it is normal. Indeed, it is reminded that there is a strong velocity

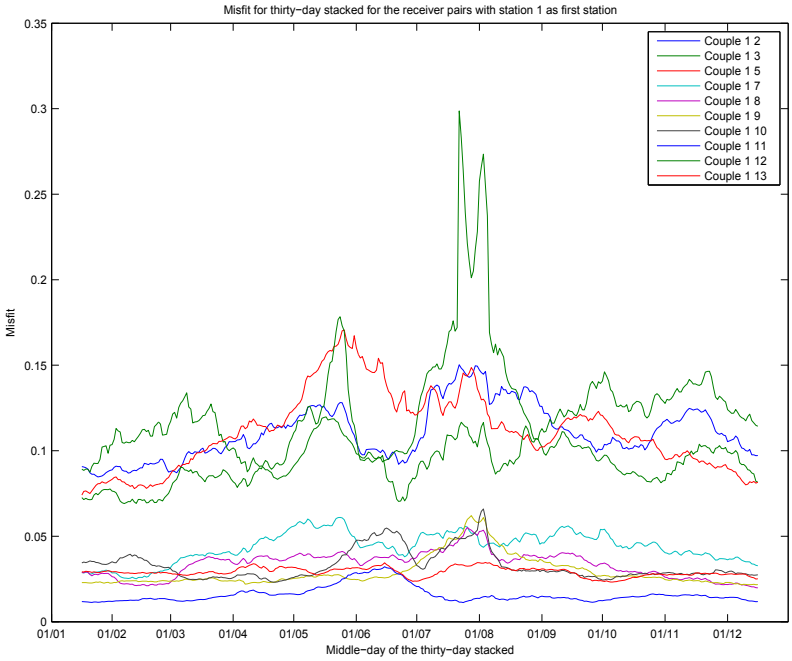

Figure 7: Misfit curves for all the station pairs containing station 1. Applying the selection, only the station pairs $1-7$, $1-8$, $1-9$, $1-10$, $1-11$ and $1-13$ will be kept.

gradient between the east and west part of the SAF. Consequently, if a station pair with each receiver on each side of the fault is considered, then it is not surprising to have different final azimuths. Moreover, even if they are on the same side of the fault, it is possible that some stations are much sensitive to another noise source than the one coming from the Pacific, for instance the noise originating from the fault. A way how to test if it is due to source effects would be to do the study in some several period-bands so that the different noise sources could be separated. Another striking feature of the azimuths found by ORA is that if you consider all the receiver pairs involving a particular station, it will not find the same azimuth. This can be explained considering two couples involving station i but with opposite orientation, so that the station i is not on the same side of the other station for both considered couples. Then, it is normal not to find the same azimuth as in one case it receives the noise in first but in the second one it receives it in second. It means that the noise can be perturbed so that it changes of orientation. In fact, since the aim is to monitor anistropy, it is quite reassuring to observe those features because it confirms that the azimth does not only record the source variations but also the azimuthal anisotropy. Therefore, it constitutes a mean to monitor the crustal properties of the uppermost crust as it is related to the crack distribution.

5.2 Single value decomposition (SVD) and temporal variations

It is difficult to see any clear variations close to the earthquake on the first results (figures (11). Therefore, in order to extract them a SVD is performed. To do so, two matrices are built. One is filled with the azimuth curves and another with the tilt curves, to which the mean value is supressed. It means that the matrices are filled with the relative temporal variations of the azimuth and the tilt. The SVD operation decomposes the matrices in three ones: a first one

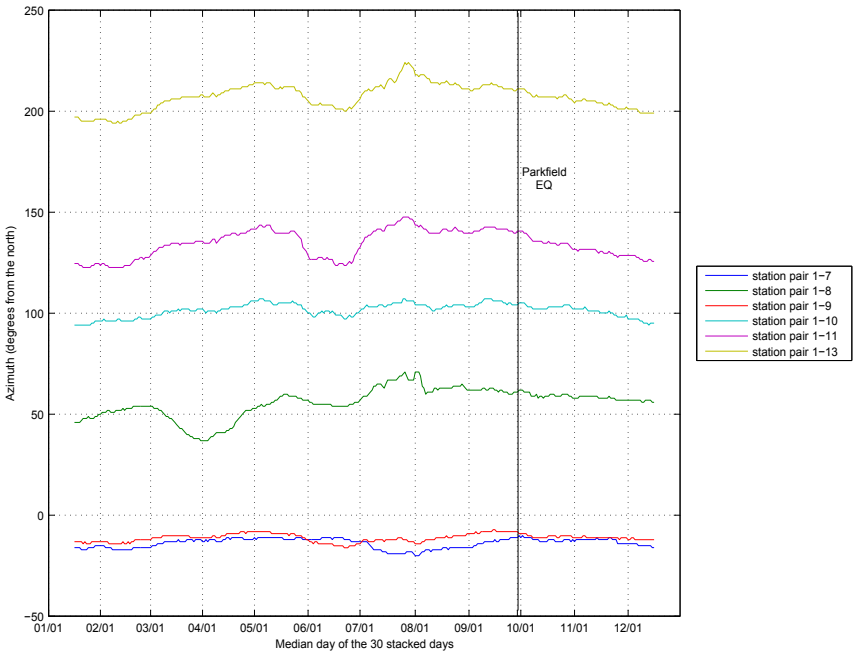

Figure 8: Time evolution of the azimuth obtained by running ORA along 2004, for station pairs containing station 1 and selected using the misfit results. No variation close to the Prakfield earthquake is seen.

where the temporal eigenvectors are stored, V, a second with the station eigenvectors, U, and a last one which is diagonal and contains the eigenvalues, sorted in decreasing order, S. The eigenvalues represent the weight of each eigenvectors. This operation can be summarized into the following expression:

$$M_{nm} = U_{nn} * S_{nm} * V^T_{mm} \quad , \tag{12}$$

where n and m refer to the matrix size and (T) denotes to the transpose matrix. The SVD was performed on the two sets of data and for both parameters, the azimuth and the tilt. Each time this operation is done, the most important temporal eigenvectors are ploted.

Finally, in order to get results which are not station pair depend, we passed from the amplitud to the energy, E, as:

$$E_\psi(t) = \sqrt{\sum_i \sum_j \psi^2_{ij}(t)} \quad \text{and} \quad E_\delta(t) = \sqrt{\sum_i \sum_j \delta^2_{ij}(t)},$$

where i and j refers to the station numbers. Now it is easier to interpret the results.

Looking more precisely at the temporal eigenvectors it is always possible to identify an eigenvector as representing seasonal variations. Indeed, this eigenvector exhibits only one oscillation over the whole year and is one of the most important (figures (12)). It is even more stiking when comparing for instance the first eigenvector of the azimuth and tilt of (WTHOUT04) with the variations of the source incidence angle obtained by beamforming (figure (13)). As the seasonal variations are always related to the first or the second eigenvector it explains why nothing can be seen on the previous curves: the seasonal variation hides everything.

Then, it is possible to rebuild the non seasonal variations by using the results from the SVD. To do so, the eigenvector carrying the seasonal variations is not taken into account in

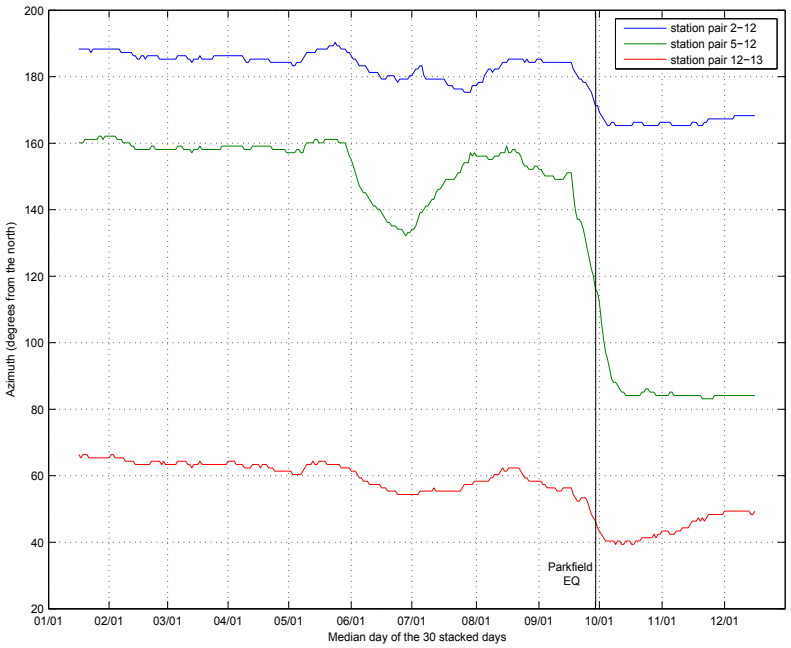

Figure 9: Time evolution of the azimuth obtained by running ORA along 2004, for station pairs containing station 12 and selected using the misfit results. A large variations at the moment of the earthquake is seen, but is it earthquake related signal, local effect or station disabilities ?

the reconstruction. The results do not look like to what we first obtained and the variations are clearer. Let us first consider the results from the set of data (WTHOUT04) concerning the azimuth (top left figure (14)). Before going further, it is needed to think about what is sought: a relationship between, the azimuth and tilt, and the Parkfield seismic event which was an earthquake of magnitude 6. The general view of the seismicity is based on the seismic cycle which contains a coseismic, post-seismic and inter-seismic phase. The coseismic one corresponds to the release of the stress which is expressed at the surface by displacements around the fault. The post-seismic part is the relaxation of the crust and generally the crust continues to deform in the same way as during the coseismic phase. On the contrary, the inter-seismic phase is seen as a period of accumulation of stress and it can be quite long. The idea being that the longer the inter-seismic period, the bigger the earthquake. So, coming back to the results, it means that for an earthquake of magnitude 6 the crustal properties must have begun to change long time before the earthquake. Looking at the energy curve for the azimuth (figure (14), top left), then the peak seen in June is perhaps the footprint of the beginning of the changes in the crustal properties before the earthquake. Going further, it is then noteworthy that three intermediate extrema seem to come after the other with a decreasing value until the earthquake, and they form a kind of decreasing exponential function from the June peak to the earthquake. This decrease permits to define a “precursory phase”, which lasted some hundreds days, and followed by an sharp decrease of the energy at the moment of the earthquake. Similar observations can be made looking at the results for the same data set but considering the tilt (figure (14) bottom left). They differ by the position of the peak, which is no longer in June but in August, so some 50 days before the earthquake.

Now, considering the results from the data set (WTH04), the June peak is still observed

for the azimuth (figure (15) top left), and it is also followed by decreasing extrema until the earthquake, where an abrupt decrease occurs. However, concerning the tilt, the peak is now in Septembre (figure (15) bottom left), so only 20 days before the earthquake.

In 2005, no significant earthquake has occured, so the comparison would test the reproducibility of the method, in particular checking if seasonal variations are still identified. We applied exactly the same method as before and the same kind of curves are obtained. To be able to compare the results two data sets are made, still with and without stations 8 and 12, respectively denoted (WTHOUT05) and (WTH05). Considering the results from the data set (WTHOUT05), it is still possible to identify a part a seasonal variations which really looks like those observed in 2004 (figure (16) bottom). This confirms our interpretation of the seasonal variations. Then, when supressing them, no significant variations are seen, which is in agreement with the fact that no significant earthquake occurred during 2005 (figure (14) right plots). The same conclusion is made when looking at the results for (WTH05) set (figure 15).

To be as rigorous as possible two sets of data were formed. It was to avoid any ambiguity that could come from the striking variations seen for stations 8 and 12. However, it revealed not to be necessary since, when the seasonal variations are removed, the energy signal is similar for both data sets. Finally, the observed variations on stations 8 and 12 is probably both, earthquake signal and local effect. Moreover, it is noteworthy that the large variations record by those stations are coincident with seasonal variations. Indeed, it has been mentioned in the part devoted to the origin of the noise, section 4.1, that the seasonal changes take place in March and October. Consequently, it is possible that a part of the signal recorded by those stations is seasonal variation related and another earthquake related.

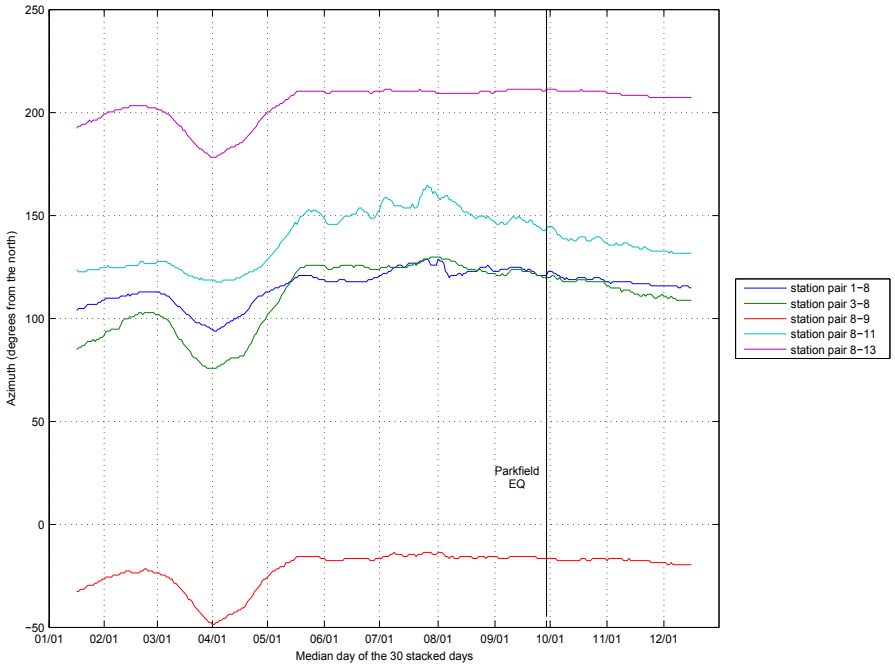

Figure 10: Time evolution of the azimuth obtained by running ORA along 2004, for station pairs containing station 8 and selected using the misfit results. A large variation is seen in April, which rises the same question as for the station couples containing station 12.

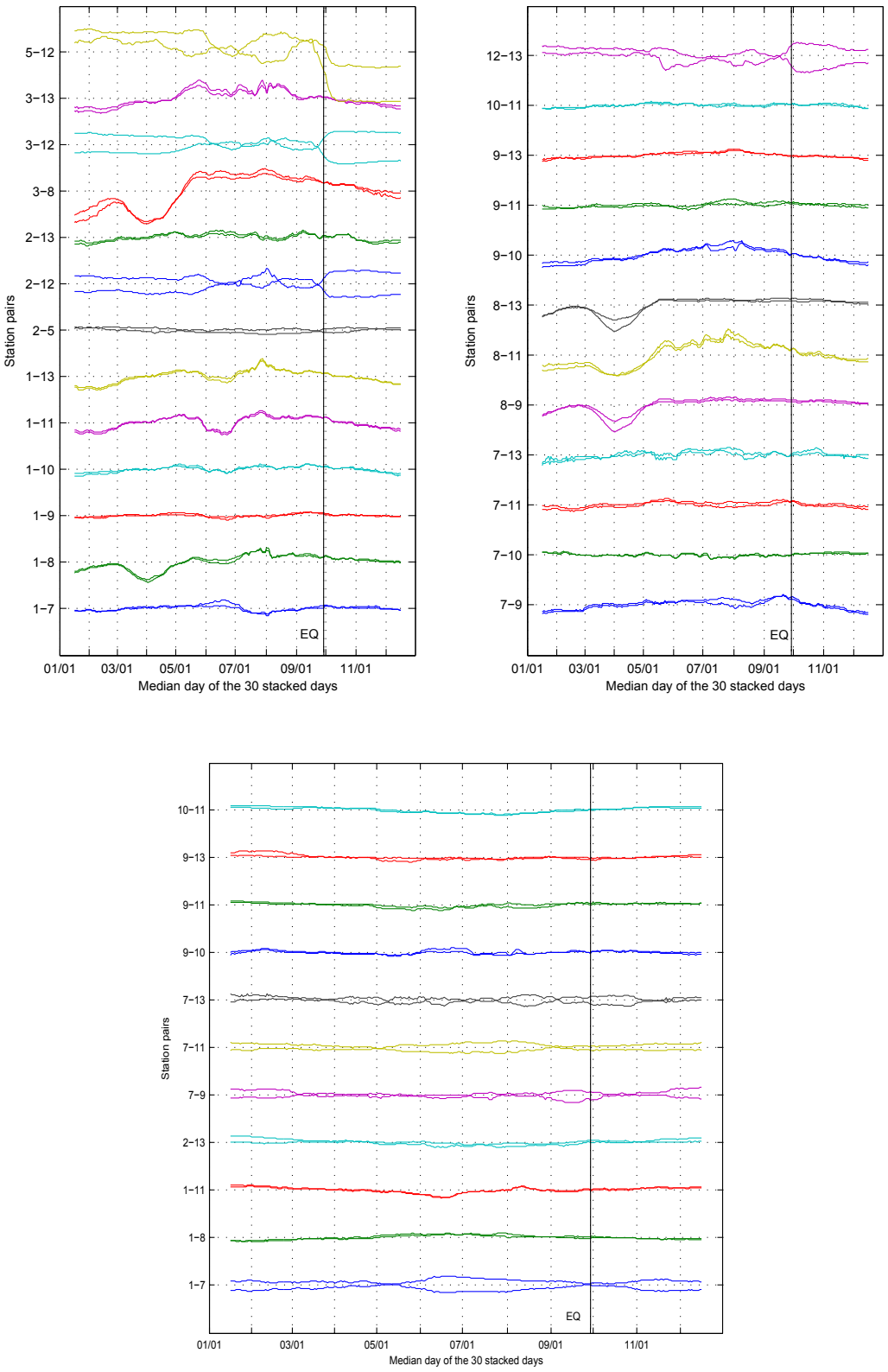

Figure 11: (WTH04) data set for the azimuth (top) and the tilt (bottom). (WTHOUT04) data set is the same but supressing stations pairs containing stations 8 and 12. Again, whatever the station couple considered is, no variation close to the earthquake are observed.

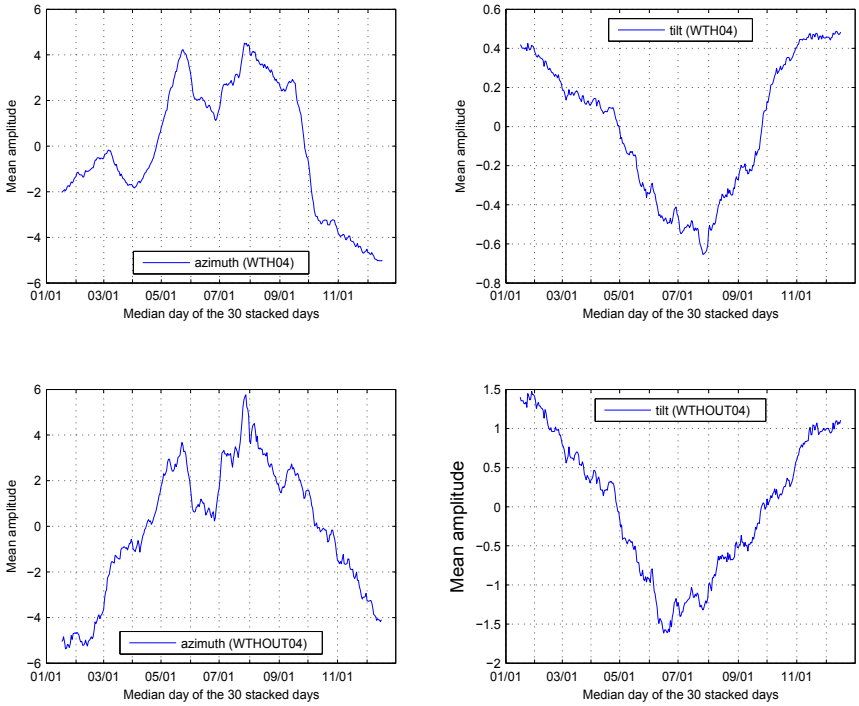

Figure 12: Seasonal variations obtained using the temporal eigenvectors found by SVD from data set (WTHOUT04) and (WTH04).

6 Discussion

The general idea of the study was to design a method able to monitor crustal anisotropy in order to check if any relationship exists between them and earthquakes. We extracted two parameters, azimuth and tilt, that can be related to anisotropy but not only. It is important to remind that they are sensitive to both, source incidence and azimuthal anisotropy. That is why the SVD helped to separate all the contributions.

How to interpret the precursory identified phase ? We can now remind the work of Crampin [Crampin and Peacock, 2008] who, by studying the S wave splitting and reviewing all the reported variations in DT before earthquakes, distinguished two phases. A first characterized by an increase of the DT, interpreted as a period of stress accumulation, and a second characterized by decreasing DT, interpreted as crack coalescence. Being provided with this, he proposed a logarithm law linking the magnitude of the earthquakes to the duration of each of these phase (figure (17)). It must be emphasized here the empirical character of this law and when looking back to the data he used, the extrapolation he did is not completely obvious. However, being conscious of this, let us confront the results to his law. With the duration of our phase varying from 20 to 100 days, which would correspond to the second phase of crack coalescence defined by Crampin, it gives an earthquake of $Mw = 5.5 - 6.3$. This must not be seen as a confirmation of the law proposed by Crampin since we only look to one earthquake. This interpretation can be discussed. However, the fact that it well works when comparing with the Crampin results suggests that it is reasonable. Moreover, it is not surprising to have comparable results for our azimuth with the DT of S wave of the Crampin's study because it represents an analogue

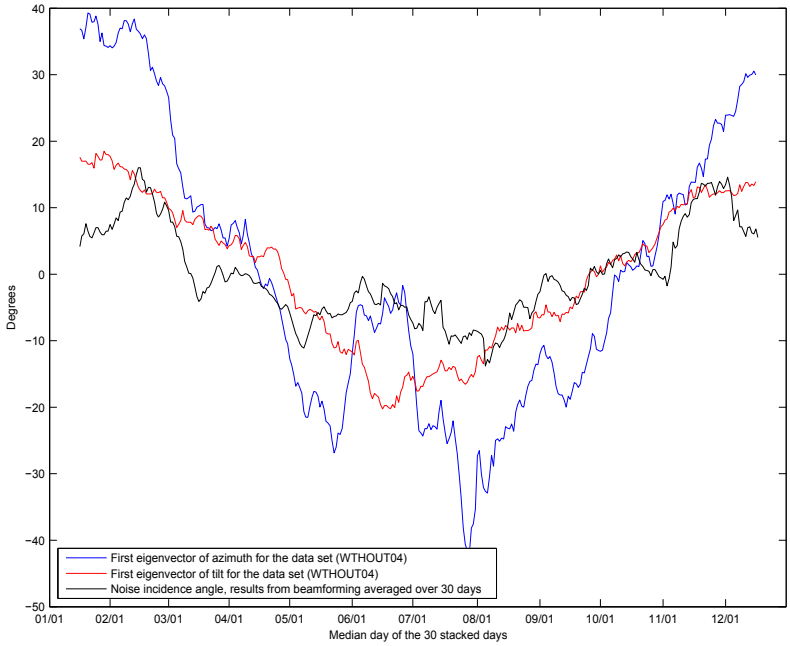

Figure 13: Comparison of the first temporal eigenvectors of the azimuth and the tilt with the noise incidence angle variations obtained by beamforming. It confirms the contribution of the seasonal variations of the noise incidence angle to the azimuth and tilt signals.

phenomenon. Indeed, the azimuth we found suggests, when it differs from the station pair orientation, which is generally the case, that the Rayleigh wave is propagating in the direction different from the original one. It is exactly the analogy of SWS phenomenon, which is the separation of an original S wave into two of orthogonal ones propagating into two directions, different from the original one.

Moreover, it is reminded that the correlation were stacked over a 30 days large time-window. It means that it cannot be more time resolutive than 30 days. We tried with only 8 days but the stability of the signal was lost so that it was harder to interpret the results.

It is also important to emphasize the fact that we were in trouble with the non isotropic distribution of the noise sources. To randomize the field, the correlation functions have been stacked but the source is so directive that it is impossible. Another way would have been to use the information contained in the coda. Indeed, the coda present on the correlation functions can be interpreted as emerging from multiple scattering. It means that all the scatters behave as secondary sources, which are probably widely dispersed in the medium, so that a symmetrical source repartition would have been achieved. In practice, it requiers to take the total energy of the coda and to send it back on the seismic array. It certainly would constitute a symmetrical noise source and have improved our results, but it is not trivial to perform.

Finally, a synthetic test really missed to the study. It would have consisted in modelling the propagation of a surface wave in a medium crossed by some crack planes considering several sources of noise, distributed arround the crack area. A first approximation would have been to considered a HTI medium and then, it would have been necessary to solve the propagation equations. Processing the cross-correlation functions between several receivers points randomly arranged in the medium, it would have been possible to run the ORA on them and look at the

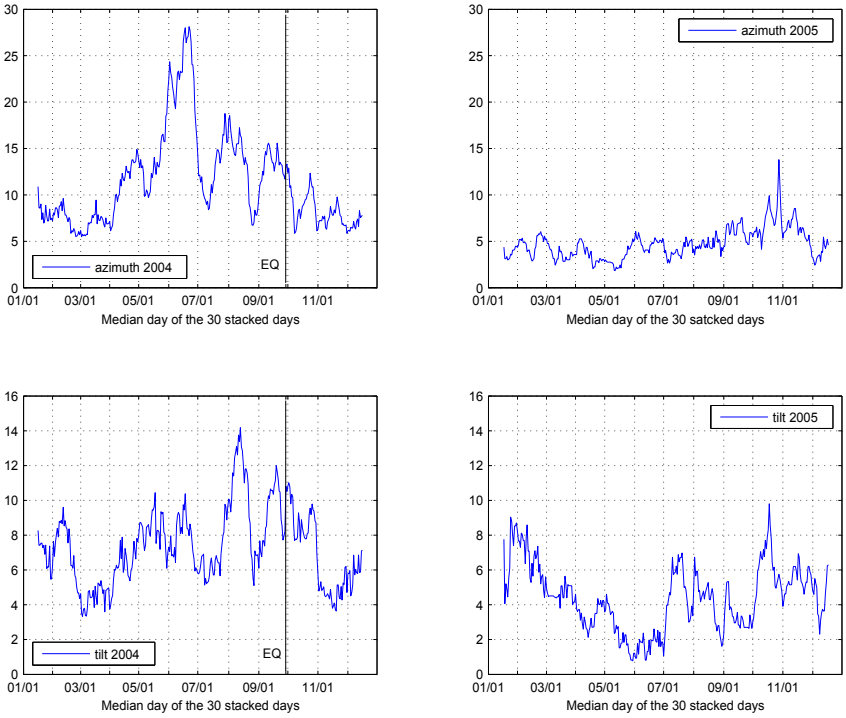

Figure 14: Non-seasonal variations reconstructed using SVD results from data set (WTHOUT04) and (WTHOUT05), by supressing the seasonal contribution.

results for the azimuth and the tilt. It would have certainly simplified our interpretation and the cause of the variations much simply identified. However, it is quite a big deal that could not be done in the time given for the study.

7 Conclusion

In order to monitor the crustal properties changes, we designed a method relying on the reconstruction of surface waves from cross-correlating ambient seismic noise. Solving an inverse problem, realized via ORA, we extracted anisotropy parameters, the azimuth and the tilt, both sensitive to the incidence source angle and the crack distribution. After, removing the seasonal variations by the mean of a SVD, the temporal variations of anisotropy has been obtained. Expressing the results in terms of energy, in order not to be station dependent, permits to highlight a precursory phase of decreasing energy of both azimuth and tilt, and which lasted from fifty to hundred days. This was confronted to Crampin's results and permited to determine a magnitude of $5.7 - 6.3$ for the Parkfield event, which is in agreement with the real magnitude, 6.0. Moreover, the absence of any significant variations in 2005 confirms the robustness of the method, since no moderate size earthquakes occured. The results confirm the fact that the method permits to record crustal properties changes and can be devoted to the understanding of earthquake processes.

Moreover, we used data from 1 Hz sensors so that we were in the limitation of the instrument. It is probably a reason of the necessity to stack over 30, instead of some ten days only. The

study should be applied to another seismic network equipied with broad-band seismometers. It should also be applied to a larger network and to an other seismogenic region.

It is also planed to realize a synthetic test so that the meaning of the azimuth and tilt parameters would be better understood and it would be easier to interpret their variations. The test must be done considering a box full of vertical crack in which a surface wave propagates.

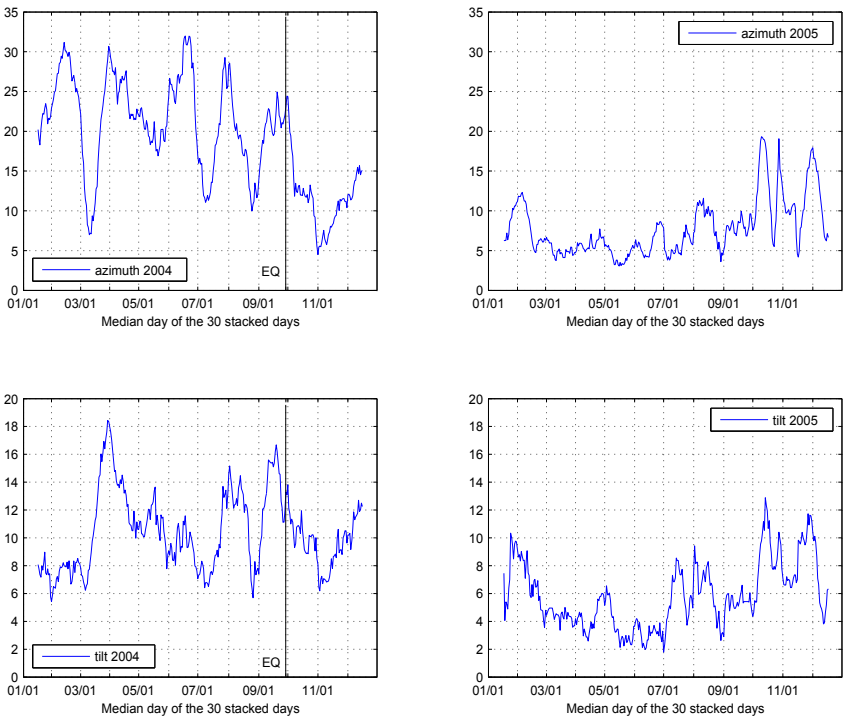

Figure 15: Non-seasonal variations reconstructed using SVD results from data set (WTH04) and (WTH05), by supressing the seasonal contribution.

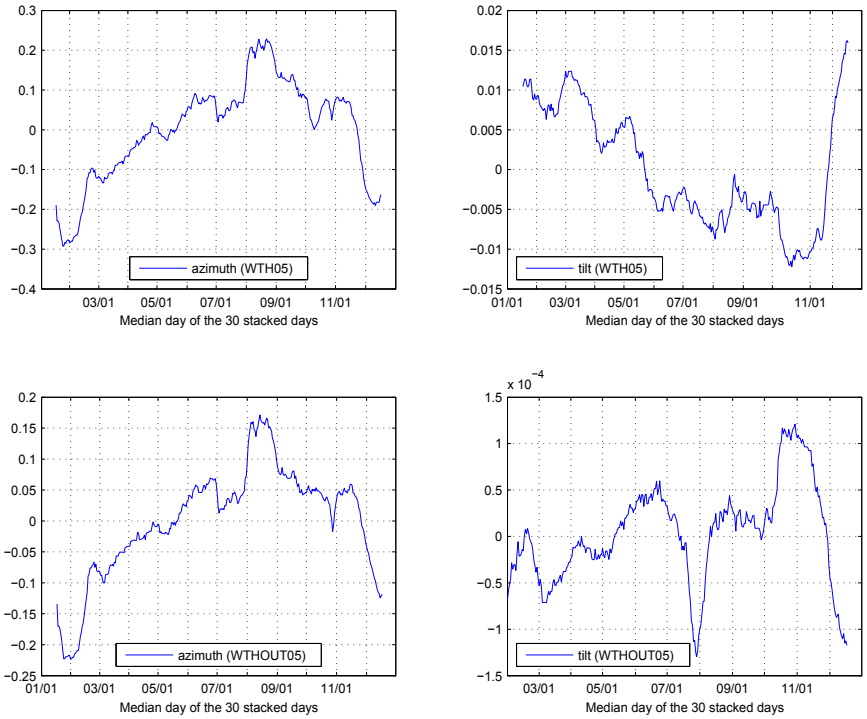

Figure 16: Seasonal variations obtained using the temporal eigenvectors found by SVD from data set (WTHOUT05) and (WTH05).

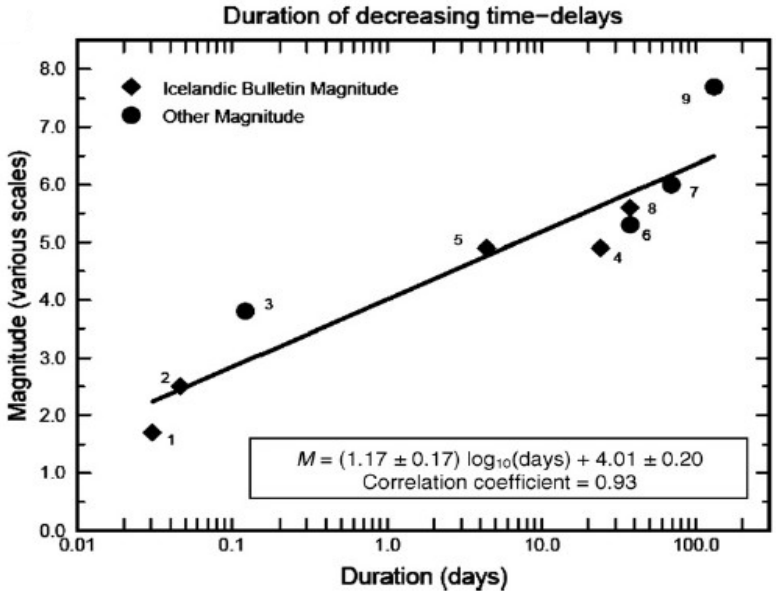

Figure 17: Crampin's law relating the duration of the crack coalescence period preceeding an earthquake to magnitude [Crampin and Peacock, 2008].

Remerciements. Je tiens à remercier le laboratoire de sismologie de l'IPGP pour son accueil. Je remercie aussi Jean Paul Montagner, Philippe Roux et Florent Brenguier pour leur encadrement et disponibilité. Ce stage m'a permis d'aborder une nouvelle branche de la géophysique, à savoir le traitement du signal pour extraire des informations des données. Ce fut très intéressant et instructif. Un petit merci aussi à Gaoutchi pour ses questions existentielles qui me font toujours avancées... Une pensée au cerf, Kikiben, JAO, Dublanch' qui m'ont accompagnés tout au long de ce stage de M2R.

References

R. C. Aster, P. H. Shearer, and T. Berger. Quantitative measurements of S wave polarizations at the Anza seismic network, Southern California: implications for S wave splitting and earthquake prediction. *J. Geophys. Res.*, 95, 1990.

S. Baisch and G. H. R. Bokelmann. Seismic waveform attributes before and after the Loma Prieta earthquake: scattering changes near the earthquake and temporal recovery. *J. Geophys. Res.*, 106, 2001.

W. H. Bakun and A. G. Lindh. The Parkfield, California, Earthquake Prediction Experiment. *Science*, 229(4714), 1985.

W. H. Bakun and T. V. McEvilly. Recurrence models and Parfield, California, earthquake. *J. Geophys. Res.*, 89(B5), 1984.

W. H. Bakun, B. Aagaard, B. Dost, W. L. Ellsworth, J. L. Hardebeck, R. A. Harris, C. Ji, M. J. S. Johnston, J. Langbein, J. J. Lienkaemper, A. J. Michael, J. R. Murray, R. M. Nadeau, P. A. Reasenberg, M. S. Reichle, E. A. Roeloffs, A. Shakal, R. W. Simpson, and F. Waldhauser. Implication for the prediction and hazard assessment from the 2004 Parkfield earthquake. *Nature*, 437(13), 2005.

G. D. Bensen, M. H. Ritzwoller, M. P. Barmin, A. L. Levshin, F. Lin, M. P. Moschetti, N. M. Shapiro, and Y. Yang. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Gepohys. J. Int.*, 169, 2007.

G. H. R. Bokelmann and H. P. Harjec. Evidence for temporal variations of seismic velocity within the upper continental crust. *J. Geophys. Res.*, 105, 2000.

N. Boness and M. Zoback. Stress induced seismic velocity anisotropy and physical properties in the SAFOD Pilot Hole in Parkfield, CA. *Geophys. Res. Lett.*, 31, 2004.

F. Brenguier, M. Campillo, C. Hadziioannou, N. M. Shapiro, R. M. Nadeau, and E. Larose. Postseismic relaxation along the San Andras fault at Parkfield from continuous seismological observations. *Science*, 321(1478), 2008.

M. Campillo. Phase and correlation in 'random' seismic fields and reconstuction of the Green function. *Pure Appl. Geophys.*, 163, 2006.

M. Campillo and A. Paul. Long-range correlations in the diffuse coda. *Science*, 299(5606), 2003.

E. Cochran, J. Vidal, and Y. G. Li. Near-fault anisotropy following the Hector Mine earthquake. *J. Geophys. Res.*, 10, 2003.

E. Cochran, Y. G. Li, and J. Vidal. Anisotropy in the shallow crust observed around the SAF before and after the 2004 M 6.0 Parkfield earthquake. *Bull. Seism. Soc. Am.*, 96, 2006.

S. Crampin. A review of wave motion in anisotropic and cracked elasctic-media. *Wave Motion*, 1981.

S. Crampin. Geological and industrial implications of extensive-dilatancy anisotropy. *Nature*, 328, 1987.

S. Crampin and S. Peacock. A review of the current understanding of the seismic shear-wave splitting in the Earth's crust and common fallacies in interpretation. *Wave Motion*, 45, 2008.

S. Crampin, D. C. Boothe, R. Evens, S. Peacock, and J. B. Fletecher. Comment on "Quantitative measurements of s wave polarizations at the Anza seismic network, Southern California: implications for S wave splitting and earthquake prediction" by R. C. Aster, P. M. Shearer, J. Berger. *J. Geophys. Res.*, 96, 1991.

A. Derode, E. Larose, M. Campillo, and M. Fink. How to estimate the Green's function of a heterogeneous medium between two passive sensors? Application to acoustic waves. *Appl. Phys. Lett.*, 83(15), 2003a.

A. Derode, E. Larose, M. Tanter, J. de Rosny, A. Tourin, M. Campillo, and M. Fink. Recovering the Green's function from field-field correlations in an open scattering medium (L). *J. Acoust. Soc. Am.*, 113(6), 2003b.

C. Draeger and M. Fink. One-channel time reversal characteristic cavities: theoretical limits. *J. Acoust. Soc. Am.*, 105(2), 1999.

T. L. Duvall, S. M. Jefferies, J.W. Harvey, and M. A. Pomerantz. Time-distance helioseismology. *Nature*, 362, 1993.

T. Iidaka and Y. Hiramatsu. Shear-wave splitting analysis of the upper mantle at the Niigata-kobe tectonic zone with the data of the Joint Seismic Observations at NKTZ. *Earth Planets Space*, 61, 2009.

R. Ikuta and K. Yamaoka. Temporal variation in the shear wave anisotropy detected using the Accurately Controlled Routinely Operated Signal System (ACROSS). *J. Geophys. Res.*, 109, 2004.

H. Kern and H. R. Wenk. Fabric-related velocity anisotropy and shear wave splitting in rocks from Santa Rosa Mylonite Zone, California. *J. Geophys. Res.*, 95, 1990.

J. Langbein, R. Borcherdt, D. Dreger, J. Fletcher, J. L. Hardebeck, M. Hellweg, C. Ji, M. Johnston, J. R. Murray, R. Nadeau, M. J. Rymer, and J. A. Treiman. Preliminary report on the 28 September 2008, M 6.0 Parkfield, California earthquake. *Seism. Res. Lett.*, 76(1), 2005.

E. Larose, A. Derode, D. Clorennec, L. Margevin, and M. Campillo. Passive retrieval of Rayleigh waves in desordered elastic media. *Phys. Rev.*, 72(4), 2005.

E. Larose, L. Margevin, A. Derode, B. V. Tiggelen, M. Campillo, N. M. Shapiro, A. Paul, L. Stehly, and M. Tanter. Correlation of random wave fields: an interdisciplinary review. *Geophysics*, 71(4), 2006.

E. Larose, P. Roux, M. Campillo, and A. Derode. Fluctuations of the correlations and Green's function reconstruction: role of scattering. *J. Appl. Phys.*, 103(11), 2008.

P. C. Leary, S. Crampin, and T. McEvilly. Seismic fracture anisotropy in the eart's crust: an overview. *J. Geophys. Res.*, 95, 1990.

Y. Liu and T. L. Teng. Systematic analysis of shear wave splitting in aftershocks zone of the 1999 Chi-chi, Taiwan, earthquake: shallow crustal anisotropy and lack of precursory variations. *Bull. Seis. Soc. Am.*, 94, 2004.

Y. Liu, S. Crampin, and I. Main. Shear-wave anisotropy: spatial and temporal variations in time delays at Parkfield, Central California. *Geophys. J. Int.*, 130, 1997.

Y. Liu, H. Zhang, C. Thurber, and S. Roecker. Shear wave anisotropy in the crust around the San Andreas fault near Parkfield: spatial and temporal analysis. *Geophys. J. Int.*, 172, 2008.

T. V. McEvilly and L. R. Johnson. Stability of P and S velocities from central California quarry blast. *Bull. Seis. Soc. Am.*, 64, 1974.

M. C. Mueller. Prediction of lateral variability in fracture intensity using multicomponent shear wave seismic as precursor to horizontal drilling. *Geophys. J. Int.*, 107, 1991.

F. Niu, P. G. Silver, T. M. Daley, X. Cheng, and E. L. Majer. Preseismic velocity changes observed from active source monitoring at the Parkfield SAFOD drill site. *Nature*, 454(10), 2008.

A. Paul, M. Campillo, L. Margerin, and E. Larose. Emperical synthesis of time-asymmetrical Green function from the correlation of coda waves. *J. Geophys. Res.*, 110, 2005.

E. Roeloffs and J. Langbein. The Earthquake Prediction Experiment at Parkfield, California. *Rev. Geophys.*, 32(3), 1994.

P. Roux. Passive seismic imaging with directive ambient noise: application to the San Andreas Fault (SAF) in Parkfield. *Geophys. J. Int.*, 2009.

A. Saga, Y. Hiramitsu, J. Ooida, and K. Yamaoka. Spatial variations in the crustal anisotropy and its temporal variations associated with a moderate sized earthquake in the Tokari region, Central Japan. *Geophys. J. Int.*, 154(3), 2003.

A. N. Semenov. Variations in the travel time of transvers and longitudinal waves before violent earthquakes. *Izv. Akad. Naik. SSSR, Fiz. Zemli.*, 4, 1969.

N. M. Shapiro and M. Campillo. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise. *Geophys. Res. Lett.*, 31, 2004.

N. M. Shapiro, M. Campillo, L. Stehly, and M. H. Ritzwoller. High-resolution surface-wave tomography from ambient seismic noise. *Science*, 307, 2005.

L. Stehly, M. Campillo, and N. M. Shapiro. A study of the seismic noise from its long-range correlation properties. *J. Geophys. Res.*, 111, 2006.

L. Stehly, M. Campillo, and N. M. Shapiro. Travel time measurements from noise correlation: stability and detection of instrumental errors. *Geophys. J. Int.*, 171, 2007.

K. Tadokoro and M. Ando. Evidence for rapid fault healing derived from temporal changes in S wave splitting. *J. Res. Lett.*, 29(4), 2002.

K. Tadokoro, M. Ando, and Y. Umeda. S wave splitting in the aftershock region of the 1995 Hyogo-ken earthquake. *J. Geophys. Res*, 104(B1), 1999.

N. Teanby, J. M. Kendall, R. H. Jones, and O. Barkved. Stress-induced temporal variations in seismic anisotropy observed in microseismic data. *Geophys. J. Int.*, 156, 2004.

K. Wapenaar and J. Fokkema. Retrieving the Green's function in an open system by cross-crorelation: a comparison of approaches (L). *J. Acoust. Soc. Am.*, 118(5), 2005.

R. L. Weaver and O. L. Lobkis. Ultrasonic without a source: thermal fluctuation correlations at MHz frequencies. *Phys. Rev. Lett.*, 87(13), 2001.

H. Zhang, Y Liu, C. Thunber, and S. Roecker. Three-dimensional shear-wave splitting tomography in the Parkfield, California, region. *Geophys. Res. Lett.*, 34, 2007.