SEISMIC SIGNAL PROCESSING FOR STATIONS IN LSBB

Clara Castellanos López

Advisor: Stèphane Gaffet

July 2009

Contents

List of Figures

Chapter 1

Introduction: The LSBB laboratory

Initially designed as an underground missile launching control room, the LSBB laboratory (Laboratoire Souterraine a Bas Bruit, Low Noise Inter-Disciplinary Underground Science and Technology) is now used as site for multi disciplinary research and development. Located in the main seismic area in the South-East of France (the Vaucluse department), the laboratory is in a fractured massif above a broad aquifer. It allows access within the unsaturated zone of a carbonate platform in an environment with low human and a very low-noise because it is located far from large cities, industries or heavy traffic. LSBB has several rooms at different depths. The deepest part of the underground tunnel is 500m below the surface. A network of six three-dimensional seismological antennas have been installed, as shown in **Figure 1.1**. The seismometers record the amplitude variation of mechanical waves in three directions: two horizontal components (north-south (N), east-west (E)) and a vertical component (Z). 125 points per second for each direction are recorded in each station. All the stations: RAS, GAS, VES, GGB, MGS and EGS, are synchronized by a GPS system.

Data has been recorded, almost continuously, for the stations in the LSBB laboratory since 2006. The data is valuable because local geology leaves a unique footprint on the seismic motion recorded. The S-wave velocity of unconsolidated sediments, geometry of the bedrock, and impedance contrasts

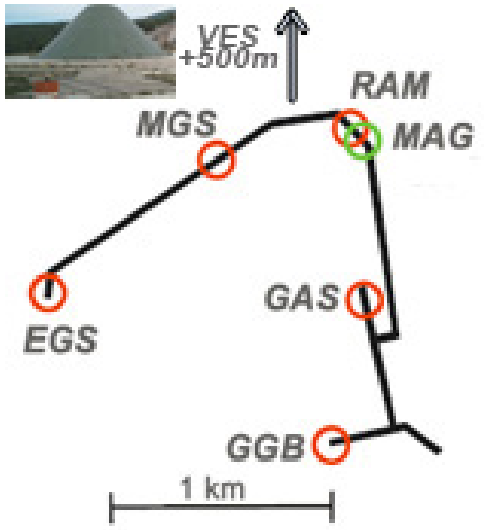


Figure 1.1: Array of antennas

between bedrocks are the main features influencing the behavior of the seismic noise signals. Trying to extract these parameters from the data recorded is an important step in attempting to describe and address the issues of each area. One way to find the detailed information desired are down-hole or cross-hole techniques. Nonetheless, drilling and logging is usually expensive and unpractical for urban areas. Fortunately, analyzing the seismic noise, which is presumably dominated by surface waves, provides information about the subsurface at less cost.

There are various interest in studying the data of the LSBB laboratory. The development of methods of observations, analysis and numerical modeling currently seeks to propose methods to quantify the risk. Another subject of study is the level of the water present in the aquifer that is located below the LSBB. The water stored in the aquifer is the supply for towns and villages in the surrounding area. Therefore, it is of interest to the community to have a precise estimation on the level of the aquifer. Finally, and more broadly, the data recorded by the stations is used to characterize and understand the geology of the area from its composition, to its geological structure. Everyday, experiments are being held which provide more information to the scientist interested in addressing these issues.

Chapter 2

Preliminary Concepts

2.1 Seismic Waves

There are two types of seismic waves, body waves and surface waves. Other modes of wave propagation exist than those described in this report, but they are of comparatively minor importance.

2.1.1 Body-Waves

Body waves travel through the interior of the Earth. They follow raypaths bent by the varying density and modulus (stiffness) of the Earth's interior. The density and modulus, in turn, vary according to temperature, composition, and phase. This effect is similar to the refraction of light waves.

2.1.1.1 P- waves

P waves (primary waves) are longitudinal or compressional waves, which means that the ground is alternately compressed and dilated in the direction of propagation. In solids, these waves generally travel almost twice as fast as S waves and can travel through any type of material. In air, these pressure waves take the form of sound waves, hence they travel at the speed of sound. Typical speeds are 330 m/s in air, 1450 m/s in water and about 5000 m/s in granite. When generated by an earthquake they are less destructive than the S waves and surface waves that follow them, due to their smaller amplitudes.

2.1.1.2 S - waves

S waves (secondary waves) are transverse or shear waves, which means that the ground is displaced perpendicularly to the direction of propagation. In the case of horizontally polarized S waves, the ground moves alternately to one side and then the other. S waves can travel only through solids, as fluids (liquids and gases) do not support shear stresses. Their speed is about 60% of that of P waves in a given material. S waves arrive second in a seismic station because of their slower speed. S waves are several times larger in amplitude than P waves for earthquake sources.

2.1.2 Surface Waves

Surface waves are analogous to water waves and travel just under the Earth's surface. They travel more slowly than body waves. Because of their low frequency, long duration, and large amplitude, they can be the most destructive type of seismic wave. There are two types of surface waves: Rayleigh waves and Love waves. Theoretically, surface waves can be understood as systems of interacting Primary and Secondary waves, which are also known as P waves and S waves.

2.1.2.1 Rayleigh Waves

Rayleigh waves, also called ground roll, are surface waves that travel as ripples with motions that are similar to those of waves on the surface of water (note, however, that the associated particle motion at shallow depths is retrograde, and that the restoring force in Rayleigh and in other seismic waves is elastic, not gravitational as for water waves). They are slower than body waves, roughly 90% of the velocity of S waves for typical homogeneous elastic media.

2.1.2.2 Love Waves

Love waves are surface waves that cause horizontal shearing of the ground. They usually travel slightly faster than Rayleigh waves.

2.1.3 Consolidated and Consolidated Soil

Consolidated sediments are materials that have been metamorphosed or cemented together, like limestone and sandstone.Ground water flows through fracture networks in these consolidated sediments. *Unconsolidated sediments* are sediments ranging from clay to sand to gravel, with connected pore spaces that allow ground water to be stored and transported.

2.1.4 Velocity of Seismic Waves

Often there is interest in describing the velocity in a region, and sometimes a velocity profile is searched because it reflects the geological footprints. The velocity depends on several variables which makes the velocity profile approximation non trivial, and which adds interest in describing it because it contains about information of other parameters such as the rock density or porosity.

$$\begin{aligned} v_P &= \sqrt{\frac{k + 4\mu/3}{\rho}} \\ v_S &= \sqrt{\mu/\rho} \end{aligned}$$

where

$$\begin{aligned} k &= \text{bulk modulus} \\ \mu &= \text{shear modulus} \\ \rho &= \text{density of the formation} \end{aligned}$$

The formation density is given by

$$\rho = (1 - \phi)\rho_{ma} + \rho_{fl}$$

where

$$\begin{aligned} \phi &= \text{porosity} \\ \rho_{ma} &= \text{matrix density} \\ \rho_{fl} &= \text{pore fluid density.} \end{aligned}$$

Throughout this report these relations are not used. However, since a target of this work is to obtain the veocity profile from the green function, it is important to keep in mind the complexity and the number parameters that are involved when describing the speed.

Chapter 3

Spectral Analysis: H/V ratio

3.1 H/V ratio

The H/V ratio (the ratio between the Fourier spectra of the *horizontal* and *vertical* components of microtremors) was introduced in the early seventies by Japanese scientists. They showed that it had a direct relationship with the elliptic trajectory of Rayleigh waves. They concluded that this ratio can be used to identify the fundamental frequency of soft soils, observing that the vertical component of Rayleigh wave motion almost systematically vanishes in the vicinity of the fundamental S wave resonant frequency.

Various sets of experimental data confirmed that these ratios are much more stable than the raw noise spectra. In addition, on soft soils sites, they exhibit a clear peak that is well correlated with the fundamental resonant frequency. Theoretical investigations show this occurs whenever the surface layer exhibits a sharp impedance contrast with the underlying stiffer formations. However, some studies also concluded that the amplitude of this peak is not well correlated with the S wave amplification at the site's resonant frequency.

Nakumara's technique, which is the one followed here, to do the H/V ratio is based on three main hypotheses. First, the ambient noises is generated by surface waves and multiply reflected and refracted shear waves trapped in the considered soils. Second, the sources of noise at the surface do not

influence the noise at the bottom of the soil. The last hypothesis is that any vertical amplification of the noise at the surface can be associated with propagation effects within the soil itself. Under these assumptions, the site response may be characterized by the spectral ratio of the vertical to horizontal components of the noise recorded at the surface.

Identifying the resonance frequencies of the H/V ratio, allows to infer information about the depth of the layer of soil bottom by the relation $f = \frac{\beta}{4H}$, where β is the mean shear velocity of the medium between the surface and the related bottom interface [1].

3.2 What we are interested in

Fluctuating peak frequencies in the spectral ratio may be attributed to changes in water saturation of superficial soils. Water saturation increase corresponds to a decrease of shear velocity and therefore a decrease of the related resonance frequency. Our target is to determine the fundamental resonance frequency of soft soils. Currently [2] only the first peak is used. Furthermore, it would be interesting to extract a water level time series by analyzing the evolution of the peaks of the H/V ratio. This would reveal information about the amount of groundwater storage through rainfall. There are usually two types of aquifers: unconsolidated, and bedrock aquifers. Bedrock aquifers, like the one below LSBB, are composed mainly of porous or fractured rock. Periodic fluctuations, or a periodic trend, in the peak frequencies are therefore expected, because of the seasonal behavior of the rainfall.

3.3 H/V analysis procedure

The data is filtered between 0.2 Hz and 40 Hz. The reason for applying this band pass filter is due to the fact that the spectral components with frequencies below 0.2 Hz are usually related to the natural period of wave tides. We are not interested in measuring these long period events, which resonance frequencies are not related to the soil itself. The upper cutoff is given by the nyquist frequency. For a sampling rate of $T_s = 0.008s$, the

nyquist frequency is $62.5Hz$. However, in practice, the nyquist frequency is approximated to $f = \frac{1}{3T_s} \approx 40Hz$. That is, variations higher that $40Hz$ can not be accurately measured. The data is then analyzed by one hour segments. Each hour segment is divided into windows of $T \approx 130s$ (2^{14} points), with 50 percent overlap. The H/V ratio is found for all the windows within an hour segment, and the median is finally chosen. Sometimes we will want to analyze the H/V ratio per days. In this case, the ratio is found for the 24 hours of the day, and then an average is taken.

The procedure to find the H/V ratio for each window segment, for any station, is as follows. Three components (z(t), n(t), e(t)) are initially detrended and the mean is removed. Following this, each window is apodized and the fast Fourier transform is taken for each component. The spectral ratio is then computed as:

$$r_i(t) = \frac{\sqrt{|FFT(n_i(t))|^2 + |FFT(e_i(t))|^2}}{|FFT(z_i(t))| + \epsilon} \tag{3.1}$$

The water level ϵ defined as $\epsilon = max(|FFT(z(t))|)/20$ is introduced in order to avoid a zero in the denominator, and physically to avoid over amplifications generated by a small energetic band. The water level is adapted empirically to preserve the amplitude of the ratios inside the frequency range under observation.

3.4 Results

In **Figures 3.1,3.2, 3.3, 3.4, 3.5** we can see the time evolution of the H/V ratio for five stations in 2007. If we look, for example, at station MGS in **Figure 3.1**, we can see a clear peak propagating in time in the low frequencies. The frequency range is 0.2Hz - 40Hz, because of the bandpass filter we performed earlier. However, in further graphs we will only plot until 30Hz because in the high frequency range there are no clear peaks evolving. As seen in **Figure 3.3**, station RAS has a similar behavior in the sense that it has a clear peak evolving in time. For stations GAS and GGB, **Figures 3.2, 3.4** more than one peak is seen. Also, both of these stations clearly record a high frequency event close to July 10 2007. The behavior of the station EGS,

Figure 3.5, is different from the rest. The HV ratio changes distinctly from the first to the second semester.

The frequencies for each station can be seen, alternately, on two dimensional graphs. All stations have clear peak frequencies propagating in time close to $5Hz$ and below. Let us start by looking at stations MGS and GAS, **Figures 3.6, 3.7**. The precise variations of the peak frequencies can be traced to obtain more detail about the changes it presents, as shown in **Figure 3.11**. The lowest resonant frequencies at MGS have values approximately of $0.8Hz \pm 0.6Hz$, $1.7Hz \pm 0.8Hz$ and $2.9Hz \pm 0.5Hz$. The lowest frequencies for GAS, **Figure 3.12**, is approximately $0.76Hz \pm 2.1Hz$. There is a low of precision in the detection of this frequency, due to a lack of data for some dates. The next resonance occurs at $5.6Hz \pm 1.5Hz$.

Stations RAS and GGB have resonant frequencies below 3Hz, as seen in **Figures 3.8 3.9**. The details of how the resonant peak frequencies vary can be seen in closer detail in **Figure 3.13**. The fundamental peak frequency of station RAS has a value of approximately $0.75Hz \pm 1.75Hz$. The next two frequencies are at $2.85 \pm 1.6Hz$ and at $3.2 \pm 1.8Hz$. The first clear peak frequency for station GGB, **Figure 3.14** is close to $0.85 \pm 1.5Hz$, and the next one is at $1.9Hz \pm 2Hz$. Station EGS, as seen in **Figure 3.10**, shows a clearly different behavior between the first and second semester.

Therefore, the fundamental resonance frequency of the soft soils has been determined for almost all stations. It would be desirable now, to extract a time series of the variation of the height of the water. Approximating the velocity of the p waves as $v_p = 5000m/s$, the velocity of the shear waves a $v_s = \frac{v_p}{\sqrt{3}} = 2887m/s$. The approximate depth of the layers for station RAS are, for example,

Station	v_p (km/s)	v_s(km/s)	f_0 (Hz)	h_0(m)	Δh	f_1 (Hz)	h_1(m)	Δh_1
RAS	5000	2887	0.74	962.2	412.4	2.85	253.2	451.05

Therefore, measured from the RAS station which is $518m$ below surface, the aquifer is expected to be between $32m$ and $856m$ below the RAS station.

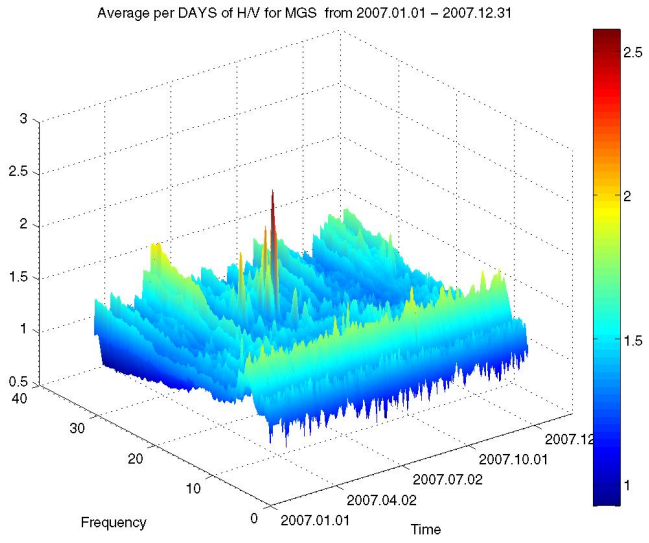


Figure 3.1: Time evolution of HV ratio. Station MGS, year 2007.

Doing a moving a time average, of thirty days, we were able to see in **Figure??** a time varying pattern of the lowest frequency which could be associated with the changing level of the water. However, a corrletaion between this, and the data registering the level of the water is needed.

3.4.1 Further Scope

There is yet more information to be extracted from the H/V ratio. The S-wave velocity depth profile, which obtains information about the geological structure, can be obtained by applying an H/V ratio inversion technique [2, 3]. Other applications of inversion techniques will be mentioned in the following chapter.

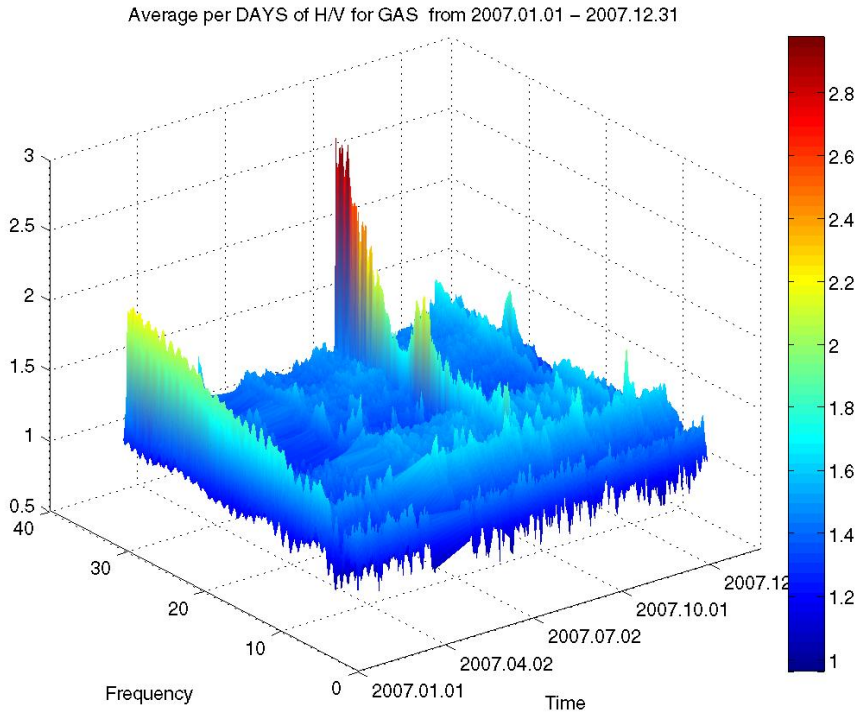


Figure 3.2: Time evolution of HV ratio. Station GAS, year 2007.

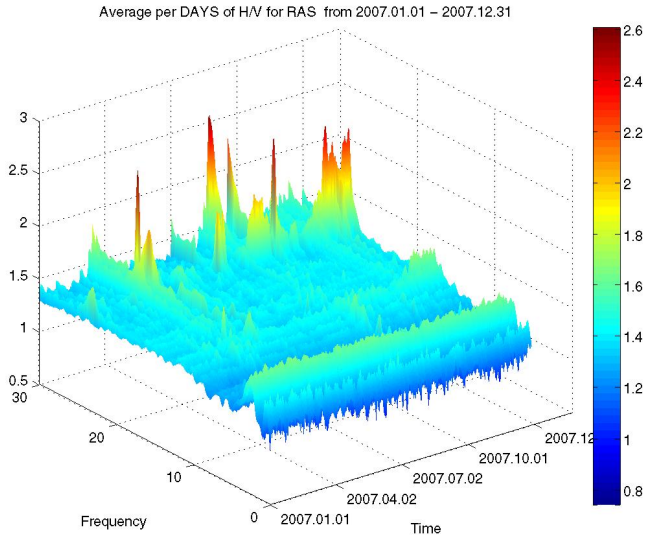


Figure 3.3: Time evolution of HV ratio. Station RAS, year 2007.

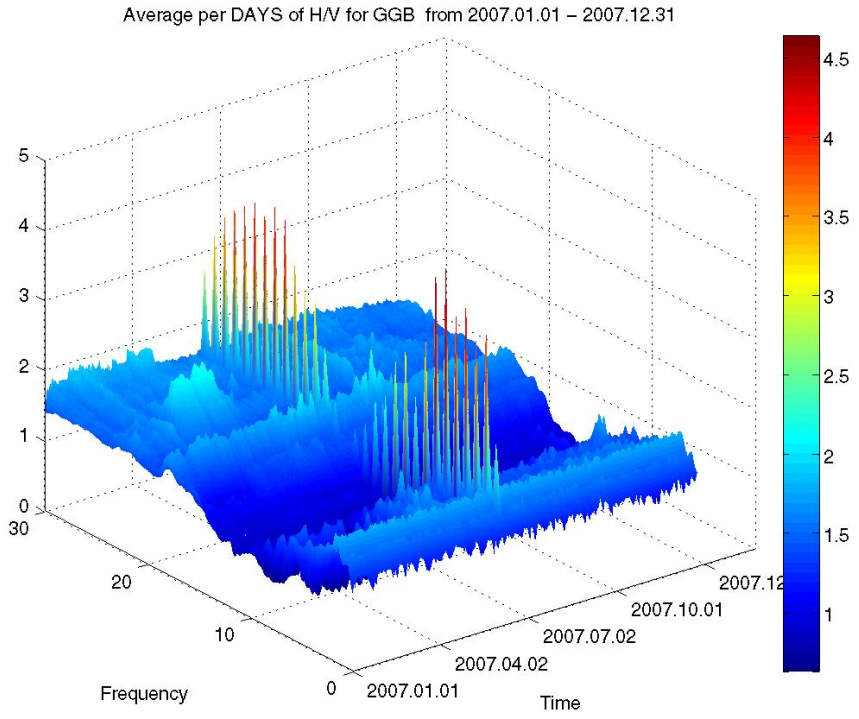


Figure 3.4: Time evolution of HV ratio. Station GGB, year 2007.

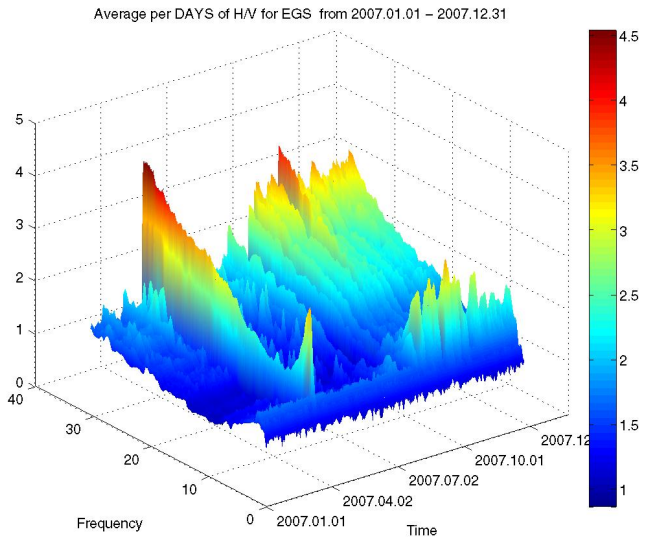


Figure 3.5: Evolution of HV ratio. Station EGS, year 2007.

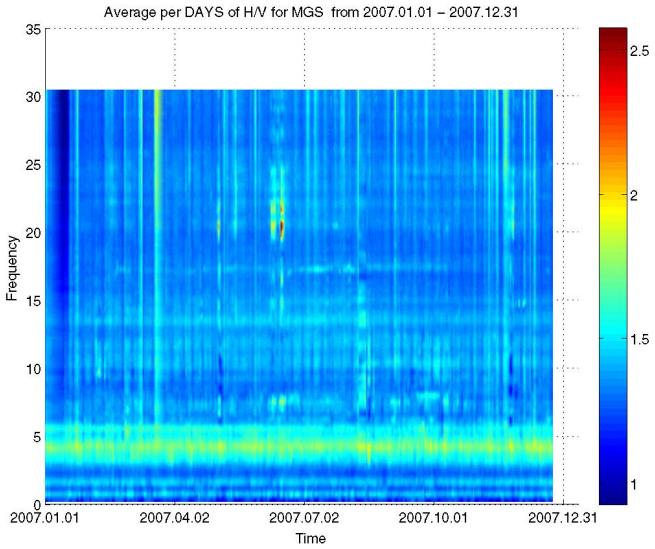


Figure 3.6: H/V ratio for station MGS . Year 2007.

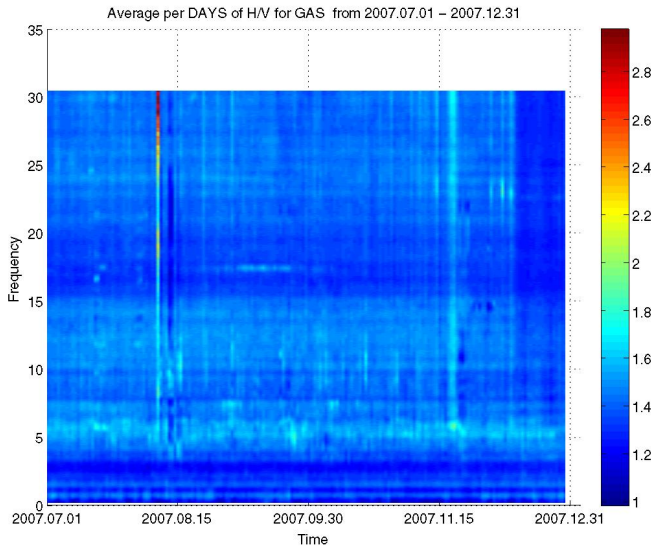


Figure 3.7: H/V ratio for station GAS. Year 2007.

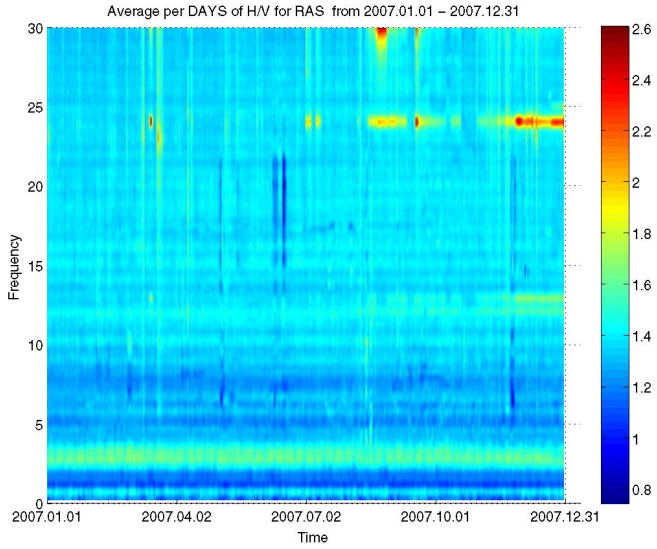


Figure 3.8: H/V ratio for station RAS. Year 2007

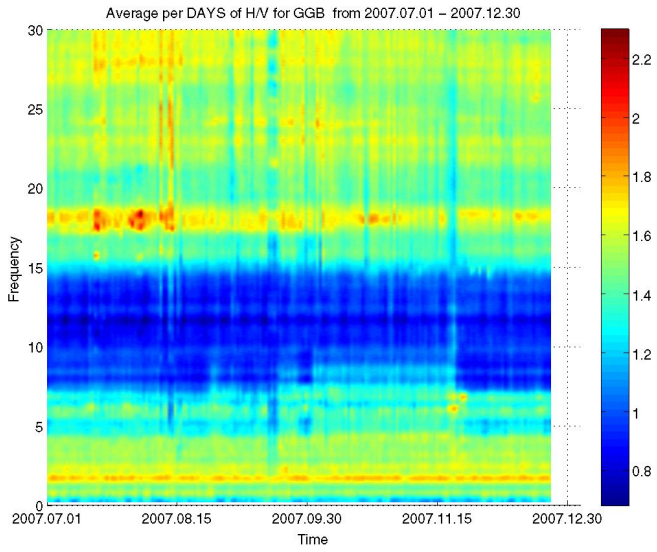


Figure 3.9: H/V ratio for station GGB. Year 2007.

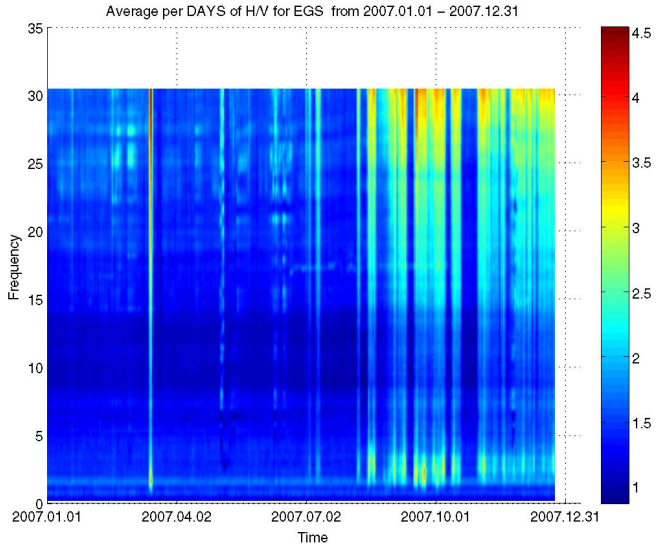


Figure 3.10: H/V ratio for station EGS. Year 2007

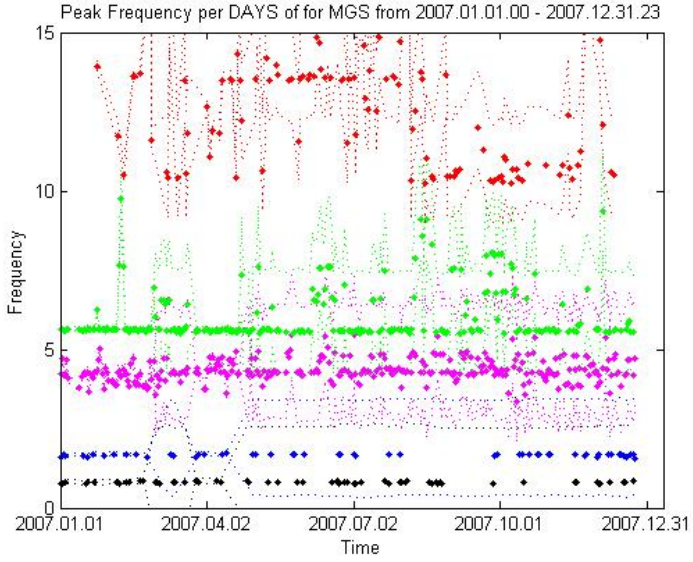


Figure 3.11: Trace of the main peak frequencies. The dashed lines represent the 70 percent confidence intervals.

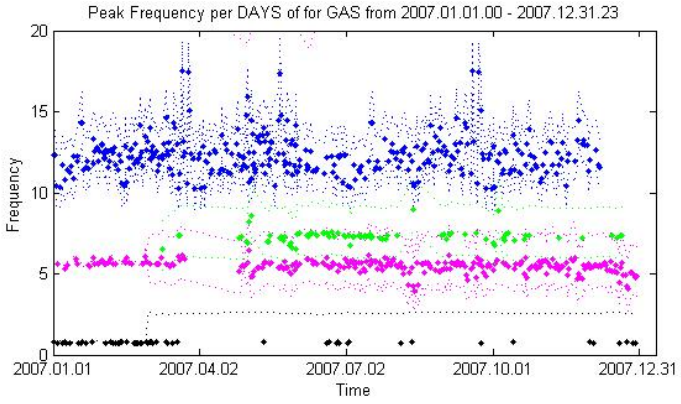


Figure 3.12: Trace of main peak frequencies. The dashed lines represent the 70 percent confidence intervals.

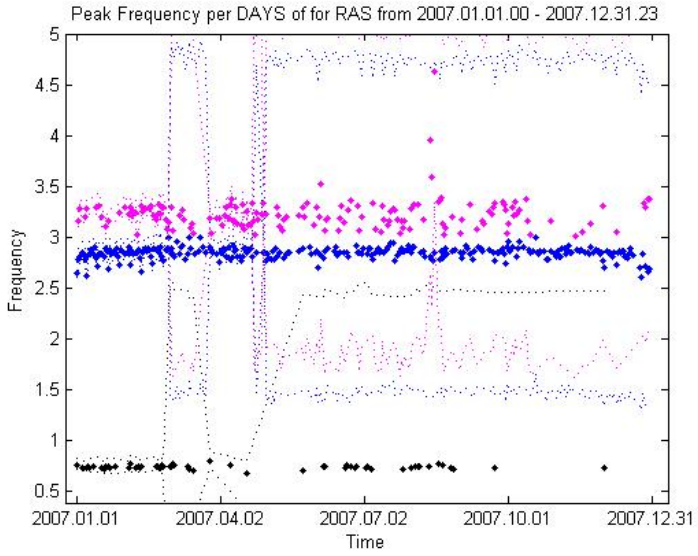


Figure 3.13: Trace of the main peak frequencies. The dashed lines represent the 70 percent confidence intervals.

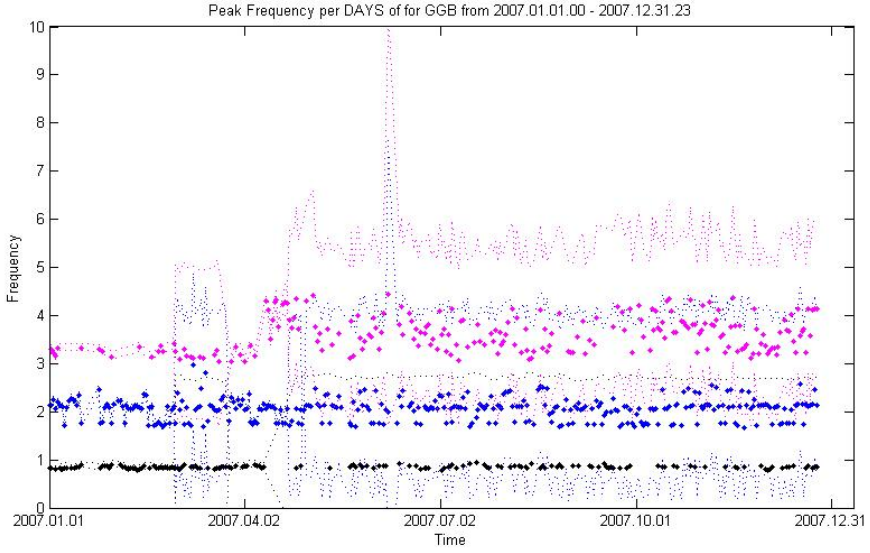
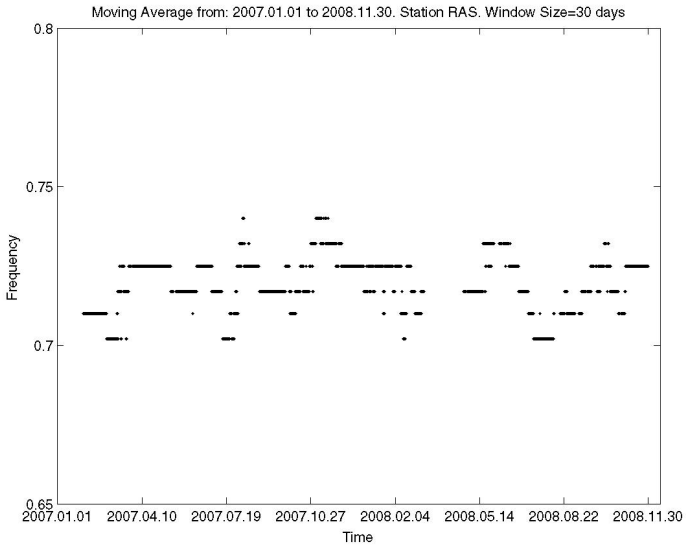


Figure 3.14: Trace of main peak frequencies. The dashed lines represent the 70 percent confidence intervals.


Moving Average from: 2007.01.01 to 2008.11.30. Station RAS. Window Size=30 days
0.8
0.75
0.7
0.65
Frequency
2007.01.01
2007.04.10
2007.07.19
2007.10.27
2008.02.04
2008.05.14
2008.08.22
2008.11.30
Time

Chapter 4

Cross Correlation of Seismic Noise

Traditionally, seismic imaging has been done with coherent seismic waves, waves emitted by explosions or earthquakes. These waves are used to measure travel times of the body waves and dispersion curves of the surface waves with the use of ray theory. Through these measurements, it is possible to gain information about the Earth's interior and structure. One setback of this technique, is that it requires energetic sources such as large explosions or earthquakes, in order to accurately locate the source. Therefore, this procedure has been used only in highly seismic areas.

Recently, a method to study low seismic regions has been proposed. Using fully diffuse wave fields with random amplitude and phases, and propagating in all directions, information about the earth has been extracted by computing the cross correlation between two stations. Assuming the modal amplitudes are uncorrelated random variables, it is possible to find the Green function between two stations [4].

At all times the stations are recording ambient seismic noise that is generated by sources such as ocean, rainfall or atmospheric perturbations. The distribution of the ambient seismic noise over long periods of times, is random. Multiple scattering tends to homogenize the phase space, the energy becomes uniform in the phase space when entering the diffusive regime. Therefore, one

can assume that the modal amplitudes are uncorrelated and, consequently, the deterministic Green function can be extracted from the ambient seismic noise. Since there is no need for an earthquake to occur, data can be registered for long periods of time.

It is natural to wonder why the Green's function is obtained by cross correlating the fields at two sites [5]. Consider two stations at points A and B, and a source S. $h_{AB}(t)$ is the wave field sensed at A when a Dirac $\delta(t)$ is sent by B. If $e(t)$ is the excitation function in S, then the wave fields received in A, ϕ_A, and in B, ϕ_B, will be the convolution $\phi_A = e(t) \otimes h_{AS}(t)$ and $\phi_B = e(t) \otimes h_{BS}(t)$. The cross correlation of the fields in A and B is

$$C_{AB}(t) = \int \phi_A(t+\tau)\phi_B(\tau)d\tau \tag{4.1}$$

$$= h_{AS}(-t) \otimes h_{BS}(t) \otimes e(t) \otimes e(-t) \tag{4.2}$$

Based on a time reversal (TR) symmetry argument, the Green's functions can be extracted from the correlation $C_{AB}(t)$. If we suppose the medium is not moving then $h_{AB}(t) = h_{BA}(t)$, so

$$C_{AB}(t) = h_{AS}(-t) \otimes h_{BS}(t) \otimes e(t) \otimes e(-t) \tag{4.3}$$

$$= h_{SA}(-t) \otimes h_{BS}(t) \otimes e(t) \otimes e(-t) \tag{4.4}$$

A typical imagined time reversal experiment would be like this [5]: A sends a pulse, S records $h_{SA}(t)$, performs a **time reversal** and sends it back. The resulting wave field observed in B would be $h_{SA}(t) \otimes h_{BS}(t)$, which is the cross correlation $C_{AB}(t)$ of the impulse responses received in A and B when S sends a pulse. So, C_{AB} is the same when the event is produced in S and we cross correlate A and B, and when the event is produced in A, time reversed in S, and observed in B.

If we assume the sources S were continuously distributed on a surface surrounding A and B, then no information would be lost during the TR operation. This means that when A sends a pulse that propagates everywhere in the medium, it is recorded on every point S. After the TR, the wave that goes backwards should hit B and then A. Which means that the field

received in B at times $t < 0$ is exactly $h_{AB}(-t)$, the time-reversed version of the Green's function. When the pulse reaches A, it continues to B, giving rise to h_{AB}. Therefore, the exact impulse response $h_{AB}(t)$ can be recovered from the casual ($t > 0$) or the ant casual part ($t < 0$) of the cross correlations C_{AB}.

This procedure of retrieving the Green's function has been tested numerically [6, 5, 7]. Experimental results show that even when the sources are not arranged as a perfect TR setup, the main features of the Green's function can be retrieved from the correlations because of the multiple scattering in the random sample. Through simulations it can also be seen that the backward and forward correlations are almost never identical in a real experiment, but are more similar when there is strong multiple scattering.

By doing the cross-correlation of the noise at two sites, the fundamental mode Rayleigh wave emerges [4]. Therefore, there is a possibility to invert the Rayleigh waves reconstructed from the cross correlation to find the group velocity dispersion curves.

4.1 Data Analysis

The seismic noise recorded by the stations is in the diffusive regime, and has a random distribution. However, to guarantee an adequate result for the cross correlation between two time series, the noise time series to be correlated have to be properly prepared. First, we synchronize the signals by taking segments of one hour, and guaranteeing they begin at the same time and have the same number of points. If this is not the case, we resample one of the series to have the same number of points as the other one. Following, we detrend the series and extract the mean. As mentioned previously, the possibility of finding the Green's function between two points is based on the fact that the amplitude of the noise has a random distribution. In order to guarantee this, we whiten the signal by making the amplitude of each frequency component equal. There has to be an additional previous step before calculating the cross correlation to find the Green function. The magnitude of the seismic noise varies by several orders of magnitude, depending on the distance the wave has traveled. Doing a cross correlation between two stations would overweight the part of the noise with greater amplitudes. However, to find the Green function there is only interest in the phase of the waves

being correlated, therefore the amplitude is disregarded by considering only one bit signals. Binarization gives equal weight to the longest paths, which have had more diffraction and scattering, helping to improve the conditions needed to retrieve successfully the Green's function.

Once the signal has been prepared in this way, the cross correlation between two stations is calculated. As with the H/V analysis, windows of $T \approx 130s$ are taken, with 25 percent overlap. The cross-correlations are calculated for all the windows within an hour, and the median and mean are taken to compare the results .

For an interval of 12 months, the correlations are then stacked. The stacking criteria determines which correlations are averaged, and which correlations are not. The criteria used for the results presented consists in rejecting the correlations whose maximum lies outside the $[-t^*\ t^*]$ interval. To determine the value t^*, we approximate the arrival of the S-wave. We can do this because for two given stations we know the distance between them, and we also know the approximate S-wave velocity. Therefore the expected time of arrival is $t = d_{AB}/V_s$. The range $[-t^*\ t^*]$ should be wide enough to ensure that the arrival lies within the interval despite some possible velocity variations.

When a the correlation function is stacked, the signal to noise ratio (SNR) is calculted. The signal is taken as the peak of the envelope of the cross correlation, and the noise is taken as the mean of the absolute value of signal, extracting the peak, as is done in [5]. By analyzing the SNR, it is possible to deduce weather the stacking procedure is contributing to the quality of the obtained Green function.

4.2 Inverse Problem

In the 1950s, a technique using seismic noise to study the S-wave velocity structure was proposed. By finding the correlation between two noise recordings, it was possible to derive dispersion curves. The S-wave velocity profiles

obtained through inversion methods where in good agreement with the measurements of down-hole techniques.

The inversion process consists in finding the velocity as a function of depth from the S-wave velocity profile. Over the past years numerous inversion techniques have been studied and tested, such as linearized inversion, genetic algorithms, and simplex methods, amongst others.

Phase velocities can be determined using ambient noise. The assumption is that noise represents the sum of waves propagating in a horizontal plane in different directions, but with the same phase velocity for any given frequency. Waves with different propagation directions are assumed to have different frequencies, and are statistically independent.

The dispersion curves of the Rayleigh waves have a non-linear relation with the density, with the S-wave velocity structure and with the P-wave velocity structure. Therefore, for most inversion methods, the final calculated parameters depend on the initial assumed model. When an accurate initial model may be proposed using a priori information, linearized inversions can find optimal solutions that are the global solutions.

4.3 Results

The stations EGS and GGB were chosen because, geographically, they are have similar depths. Using the procedure described previously, the cross-correlation of the vertical components was calculated, which correspond to a Rayleigh wave pulse. The result of the cross correlation is shown in **Figure 4.1**. A peak emerges at approximately $t = -0.616$, which corresponds to a S wave velocity of $V_s = 2464.29m/s$. This result is lower than we expected. In the same figure, we can see the behavior of the signal to noise ratio as the correlations are stacked. There is an interval where there is no enhancement of the signal. This indicates that the data being stacked in this interval is not similar to the rest of the data. The correlation function may be changing during periods of time due to the fact that the medium may be altered beacuse of weather factors such as rain. Therefore, we calculate again the green

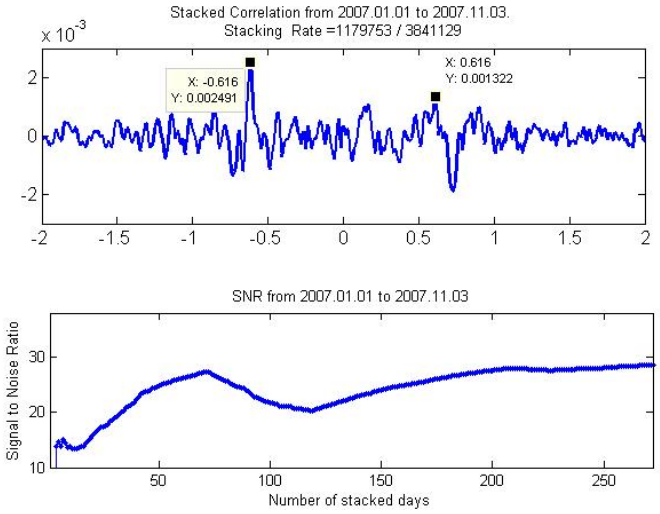


Figure 4.1: Above: Green Function. Below: Signal to noise ratio

function in the same period of time but this time stacking only those signals that increase the signal to noise ratio. The result is shown in **Figure 4.2**. The signal to noise ratio increases significantly, but the correlation function shows no improvement.

The variation of the SNR may be studied deeper, in order to determine if there is a relationship with the rainy season. In addition, the Green function can be estimated for different periods of time, and determine if the variations of the SNR can be detected in the correlation function. Once the correlation has been successfully obtained, it is possible to find the group velocity and phase velocity with the procedure based on the frequency time analysis described by *Campillo et. al* [8]. The dispersion curves and the inversion process still haven't been done, but the programs are ready to be tested.

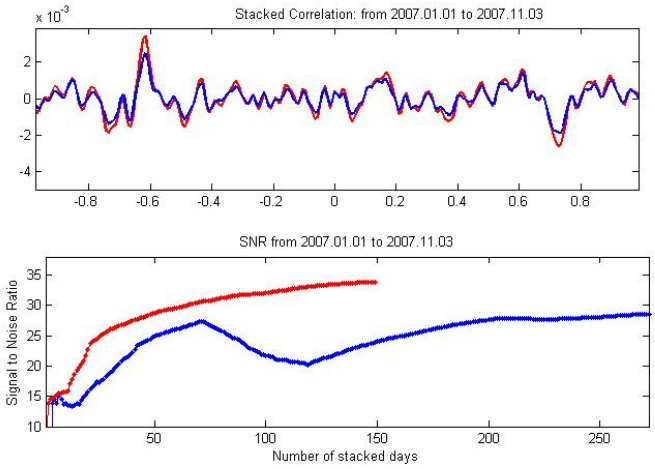


Figure 4.2: Above: Green Function. Below: Signal to noise ratio. Red line: stacking only occurs if the SNR increases.

Chapter 5

Programs

The programs are divided in two main parts which correspond to the two techniques developed throughout the research.

5.1 H/V analysis

The main program to find the H/V ratio for any period of time is ***Graphical_Analysis.m***. Inside this file you define the initial date, the final date and the stations we would like to analyze. This file calls several functions.

1. hsv_clapiere_evolution_station_d_dayint_Home2.m : This function divides the analysis in intervals of d-days. The function first calculates the interval to be analyzed, reads the TITAN files for this interval and converts them to SAC files, saves them in a folder /castella/SACfiles/. When the location of the TITAN files change, the script called test7.sh, which is called from this function, should be changed. The next function called is
   hsv_clapiere_evolution_onestationD.m, where the H/V ratio is calculated per intervals of one hour, and finds the average of all the hours in one day to find the H/V ratio per day. This function was almost entirely written by S.Gaffet, and I just made a few changes adapting it to what I wanted. The data of the ratio is saved in folders HrsTxt/ and DaysTxt/, under a subfolder with the station's name. This process

is the slowest, and the core of the H/V analysis. At the end of each interval, the SAC files are erased.

2. resonant_frequencies2.m: Calculates the 5 frequencies with the highest H/V ratio, for each hour and/or day. The variable $posfmax$ indicates until what position in the frequency vector will be analyzed. This variable is important because if one wants to concentrate on the low frequency regime, then a low value (for example 2500) should be given. Inside the program it is possible to see where the position should be taken according to the maximum frequency desired.

3. PeakFreq_Graph2.m: Plots the Peak frequencies calculated with the above function. The divisions should be introduced manually and empirically inside the function in the variables $d1, d2...d5$. The function resonant_frequencies2.m should be run first.

4. HV_hourMoving Average.m : In order to see a smoother variation of the peak frequencies, we can do a moving average time window. There are three additional parameters. winsize is the size of the window in hours. LowFreq and HiFreq determine the range where we think the peak frequency is located.

5. HV_Graphs_Avg3.m plots the H/V ratio as a function of frequency and time. It requires the first function to be run previously.

All of functions need the data generated by the first function. The folders in which this data is saved when it is generated should previously exist.

5.2 Green's function

The main file to find the Green function between two stations is called ***CorrelationHoursGraphicalAnalysis.m***. The initial and final date are defined in this file. Several functions are called here:

1. correlation_towstations_hoursFinal.m : This function divides the task of finding the correlation between two stations into intervals of d days. Initially, the function defines the interval to be analyzed, following it converts the TITAN files into SAC files. This time, however, the

drift files are included when converting to SAC format. This function calls *correlation_hours_drift.m* which is where the data to be correlated is read. Additionally, this function is in charge of synchronizing the two sets of data. The synchronization procedure consists in reading the start time and end time for each data set of by intervals of one hour. The two start times and end times are made as close as possible between the two time streams by removing points from the series that begins first or ends later. However, in this process it is assumed that there are 125 points per second, which may not be true at all times. When the synchronization is done as precise as possible, *correlationZZ_two_stations_hours_win.m* is invoked. Here the correlation between the Z components of the two stations is calculated for windows of two minutes within one hour. For each hour, the mean and the median of some *selected* correlations are saved in txt files. This function can be changed to calculate the correlation between any other components. This function receives as a parameter two matrices X and Y. Matrix X, for example, contains in the first column the Z component, the second column has the N component and the third column has the Z component of the first station. Therefore, when the ZZ correlation is desired, the first column of X and the first column of Y are chosen. If one wants the Z/E correlation, for example, then one must choose the first column of X and the third column of Y, which in the last mentioned function, correspond to $i = 1$ and $j = 3$. The way to select which data to average, or what is commonly referred to as the stacking criteria consists in searching the maximum in the correlation for each time window. If the maximum occurs in the interval $[-1, 1]$ seconds, then the correlation is stacked. Otherwise it's rejected. The time interval of rejection may vary depending on the stations chosen. For example, if we analyze stations GGB and EGS, the S wave is expected to arrive at $t = \frac{1518}{\frac{5000}{\sqrt{3}}}$ seconds, which is below one second.

2. stackZZ_corr_hours_win_perDays2.m: This function is simply used to stack all the files of the hours for one day, into one file.

3. group_velocityDays3.m : This function calculates the dispersion curve. Additionally, it plots the correlation function, the signal to noise ratio, and the dispersion curve, saving the plots at the end as .fig and .jpg files. The data is also written into text files.

Chapter 6

Conclusions

The H/V ratio technique was tested and used to find the resonant frequency of soft soils. Clear peaks were seen and those in the low frequency regime were used to find a range of depth where the aquifer might be located. Further analysis needs to be done to improve the dispersion in the results and to compare the peak's trend with the rainfall trend for each year. The resonant frequencies above 5Hz have not been analyzed. They correspond to other layers,and the thickness of these layers can be approximated. The analysis can be taken one step further. The H/V ratio can be inverted to find the velocity profile. This would be interesting to do, to compare it with the velocity profile obtained through the green's function method.

The cross correlation makes use of the noise signals to reconstruct Rayleigh wave. There is no need of an energetic signal so, this method has the advantage that it can be applied wherever seismograms are located, and that it is possible to have long time series. Finding Green's between two stations allows us to have information about the geological structure around these two points. To reconstruct the Rayleigh wave successfully , the cross correlations should be smartly stacked. Good criteria should be developed to determine which correlations should be stacked. Moreover, an analysis is needed to determine the minimum number of correlations needed in order to obtain a stable Green's function. The Rayleigh wave reconstructed here has an arrival time bigger than was expected. Shifts in time due to a lack of complete synchronization may be the cause. However, the results show that, with improvements yet to be done, the Rayleigh wave can be obtained.

Bibliography

[1] Claude Paumbrun Frederic Cappa Tony Monfret David Amitrano Stephane Gaffet, Yves Guglielmi. Monitoring of seismic and rainfall effects on rockslide rheology in the la clapiere mountainous enviroment, france. 2003.

[2] Fortunat Kind Donat Fah and Domenico Giardini. A theoretical investigation of average h/v ratios. *Geophysical J. Int, Volume 145, 535 - 549*, 2004.

[3] C. Milkereit S. Parolai, S.M Richwalski and D. Fah. S-wave velocity profiles for earthquake engineering purposes for the cologne area (germany). *Geophysical J. Int, Volume 145, 535 - 549*, 2005.

[4] N.M Shapiro and M.Campillo. Emergence of broadband rayleigh waves from correlations of the ambient seismic noise. *Geophysical Research Letter , Volume 31, L07614*, 2004.

[5] Michel Campillo Eric Larose, Arnaud Derode and Mathias Fink. Imaging from one-bit correlations of wideband diffuse wave fields. *Journal of Applied Physics, Volume 95, Number 12*, 2004.

[6] Michel Campillo and Anne Paul. Long range correlations in the diffuse seismic coda. *Science 299, 547*, 2003.

[7] Michel Campillo Alexandre Nercessian Florent Brenguier, Nikolai Shapiro and Valerie Ferrazzini. 3-d surface wave tomography of the piton de la fournaise volcano using seismic noise correlations. *Geophysical research letters, Volume 34, L02305*, 2007.

[8] A.Paul S.K. Singh D. Jomgmans N.M. Shapiro, M. Campillo and Sanchez Sesma. Surface-wave propagation across the mexican volcanic belt and the origin of the long-period seismic wave amplification in the valley of mexico. *Geophysical J. Int, Volume 128, 151-166*, 1997.