Rayleigh wave Tomography of NW Himalaya using Ambient Noise and Earthquake data

Thesis submitted in partial fulfillment of the requirements for the

BS-MS Dual Degree Programme

IISER PUNE

Indian Institute of Science Education and Research, Pune

By

Abhishikth Vaddineni

20111021

Under the guidance of

Prof. Shyam S. Rai

Chair, Department of Earth and Climate Science,

IISER Pune.

CERTIFICATE

This is to certify that this thesis entitled "Rayleigh wave Tomography of NW Himalaya using Ambient Noise and Earthquake data" submitted towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents original research carried out by Abhishikth Vaddineni at IISER, Pune under the supervision of Prof. Shyam S. Rai during the academic year 2015-2016.

Prof. Shyam S. Rai,

Chair, Department of Earth and Climate Science,

IISER Pune.

DECLARATION

I hereby declare that the matter embodied in the report entitled "Rayleigh wave Tomography of NW Himalaya using Ambient Noise and Earthquake data" are the results of the investigations carried out by me at the Department of Earth and Climate Science, IISER Pune, under the supervision of Prof. Shyam S. Rai and the same has not been submitted elsewhere for any other degree.

Abhishikth Vaddineni,
B.S – M.S Dual Degree Student,
IISER Pune.

ABSTRACT

Seismic tomography is an approach to reconstruct 2-D or 3-D velocity variation in a region of interest. This is performed using body or surface wave data from earthquake sources. Application of body wave data includes using travel time from the source to receiver, while surface waves use group or phase velocity between source and receiver at different time periods. Given the poor distribution of earthquakes (source) and seismic stations (receivers), it is difficult to obtain a well illuminated tomographic map of the region. To improve this, a new tool called seismic interferometry which uses ambient seismic noise to extract Green's functions between pairs of stations is recently being applied to perform tomography imaging. The approach is widely referred as Ambient noise Tomography (ANT). Seismic noise is mostly concentrated near the Earth's surface and contains information related to Earth's shallow structure which is generally absent in the earthquake records due to attenuation and scattering. In the recent years, ANT has been widely used to image the Earth's crust and upper-mantle on both regional and continental scales. In this study, we combine both Earthquake and Ambient noise data recorded at 15 broadband stations in North West (NW) Himalaya to obtain a higher resolution tomographic image of the region using seismic surface wave at time periods ranging from 8-50 seconds. It is also shown that the use of ambient noise correlations in this area together with the surface wave dispersion measurements of earthquakes improves both the lateral and vertical resolution.

Table of Contents

List of figures and Tables

ACKNOWLEDGEMENTS

Chapter 1

Introduction

1.1 Motive behind the thesis

Conventional deep earth exploration techniques use earthquake records which use arrival time or waveform data recorded over a network of seismic stations. In this case the spatial resolution depends on the geometry of the earthquake locations and recording stations. Given the anisotropic distribution of sources, we cannot expect to extract reliable 3D imaging data of the shallow earth completely. Also, due to the relatively large distance between the source and the receiver, the recorded waveform will not have the high frequencies as they will attenuate and scatter over the distance. This prohibits the traditional earthquake tomography to provide a detailed mapping of the shallow crustal structure as it requires high frequency information for shallow earth imaging.

To overcome these shortcomings, we apply a recently developed methodology of seismic interferometry (Ritzwoller, 2008) in conjunction with traditional earthquake tomography to better resolve the shallow earth crust in the Northwestern Himalaya region. The NW Himalaya and Ladakh form the western extremity of the Himalaya-Tibet orogenic system, formed due to collision of the Indian plate with the central Asian continental plate during mid to late Eocene.

Seismic interferometry is a technique used to obtain information about the earth's subsurface using the ambient noise wavefield caused by wind, ocean waves, rock fracturing and anthropogenic activity. This noise which is a permanent vibration due to the above sources and travels along the Earth and hence contains information about its subsurface structure. Therefore, naturally occurring ambient noise is a very economical source of energy that is usually regarded as of not much use and is removed from seismic data for analysis because of its non-impulsive nature.

The main idea in interferometry is to consider seismic noise as a wave field produced by randomly and homogeneously distributed sources when averaged over

long time series. Recent developments made it clear that long term cross correlations of ambient noise can be used as a virtual seismic source where, the cross correlation between the noise recordings at two stations produces the Green's function between these two stations (Wapenaar, 2004, 2006; Gouedard et al., 2008). If the noise sources are uniformly distributed, then the cross-correlation of noise records converges to the complete Green's function of the medium, which includes various modes like reflection, scattering and propagation modes. But, in the case of the Earth, most of the ambient seismic noise is generated by atmospheric and oceanic disturbances at the surface and hence surface waves have predominantly higher contribution in the Earth response between two stations at the surface. This shows that the surface wave part of the Green's function can be easily obtained from the noise cross-correlations. Shear wave velocity distribution can then be obtained from Green's functions using conventional imaging methods.

The idea of extracting coherent signal by such cross-correlation of noise was first applied to seismic waves in helioseismology (Duvall et al., 1993). The relationship between the cross-correlation and the Green's function of the wave propagating between pair of stations is well established (Weaver and Lobkis, 2001; Snieder, 2004). Extracting Green's function from the background noise has been applied in many fields such as helioseismology, ultrasonics, exploration seismology and marine acoustics. It has many advantages over the conventional approach as it does not require any kind of artificial source, provides uniform illumination of the study region, can be used in the region without earthquakes and is highly economical.

Shapiro and Campillo (2004) applied this technique for the first time to reconstruct the surface wave part of the Earth response by using data from stations separated by large distances (100 to 1000 km) and then measured the dispersion curves (from periods of 5-150 seconds). The first application of ambient noise tomography (also called as passive seismic imaging) using ambient noise in California (Shapiro et al., 2005; Sabra et al., 2005) provided a much greater spatial accuracy than the conventional techniques. From these results we see that the use of seismic noise

has a great potential in the study of the structure of the Earth, mainly at the crustal and upper mantle level.

Many studies have shown that, the noise sources become sufficiently well distributed when considered over long times, and also that the dispersion curves of fundamental mode surface waves can be obtained from the Green's function between the station pairs. These dispersion curves are used to calculate travel times which are then inverted to get a 3D variation of shear wave velocities in the crust and the uppermost mantle. This procedure has been used with many regional seismological networks (Yao *et al.*, 2006; Lin *et al.*, 2007; Yang *et al.*, 2008a).

Various new techniques have been developed and integrated to enhance tomographic imaging which consists mainly a forward calculation and solving the inverse problem. In this study we calculate traveltimes between stations (and earthquake epicenters) using a Fast marching algorithm (FMM) developed initially as a grid based eikonal equation solver. Once we have calculated traveltimes, we invert them using an inversion technique known as subspace inversion which is based on gradient methods. In this way we obtain tomographic maps of the NW Himalaya region using Fast Marching Surface Tomography (FMST) code written by N.Rawlinson (2004).

Here, we use the fact that ambient noise tomography can be easily combined with conventional earthquake-based tomography to extend the resolution to much shallower depths (e.g., Yang *et al.*, 2008b) and apply it to the NW Himalaya region. We get the velocity variations i.e., tomographic maps at different periods in this region using the combined noise and earthquake data.
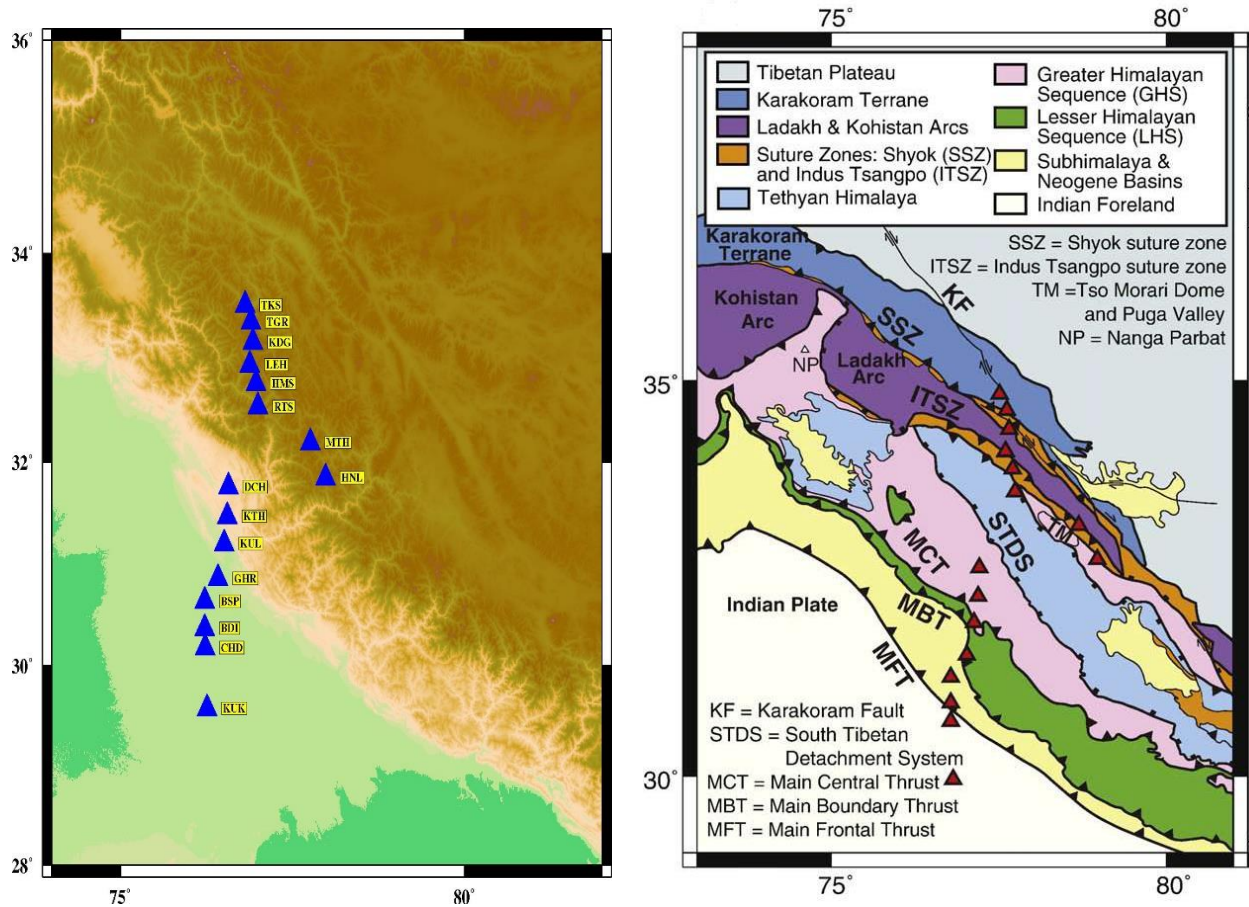
Figure 1.1: a) Geographical area (NW Himalaya) under study with locations of stations.
b) Geology map (Goscombe et al. 2006)

1.2 Geographical area under study

The area of this study is shown in the Figure 1.1a with the names of stations used to record the data. It lies in the North West Himalaya region, where the southernmost terrain is the Ladakh arc complex which is separated from the crust on the Indian plate by the Indus Tsangpo suture zone (ITSZ). It is one of the most seismologically active regions of the world. As shown in the Figure 1.2a, various geological features can be noticed such as the Tethyan Himalaya, which is on the south of ITSZ, Greater Himalayan Sequence (GHS), South Tibetan Detachment (STD), lesser Himalayan Sequence (LHS) and the Main Central Thrust (MCT). (Caldwell et al., 2009). This is also the most actively studied regions of the earth.

Chapter 2

Theory

2.1 Cross-correlation and Green's Function

Definitions:

Cross-correlation is a measure of the similarity of two series as a function of the lag of one series relative to the other series. Green's function is defined as the impulse response of an inhomogeneous differential equation defined on a domain, with specified initial conditions or boundary conditions.

The principle that is used here is that the Green's function between two stations can be obtained by cross-correlations of random wavefields recorded by these stations. In order to prove that the Green's functions can be estimated from the stack of cross-correlations of noise records, different mathematical approaches were developed (Weaver and Lobkis, 2001; Snieder, 2004) and various assumptions were made about noise characteristics and the properties of the medium (Yao *et al.*, 2009)

It was commonly believed that the diffuse wave fields reveal no information about the medium in which they propagate. But in ultrasonics, it was shown that noise correlation function gives the waveform that would be obtained in a direct pulse/echo measurement (Weaver and Lobkis, 2001). After this, Campillo and Paul (2003) applied this technique in seismology with real data and thus started a new branch of ambient noise tomography.

Here we follow the derivation proposed by Gouedard *et al.*, (2008). The main idea is that when averaged over long time series, seismic noise can be approximated as a seismic noise field.
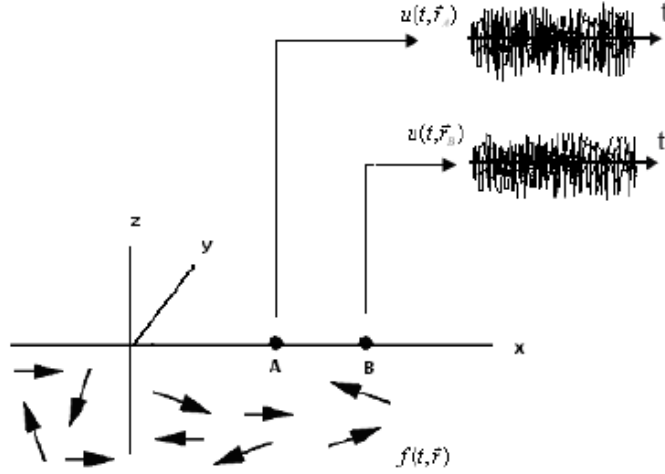
Figure 2.1: Noise wavefield and recorded waveforms at receivers A and B

Consider the displacement fields $u(t, r_A)$ and $u(t, r_B)$ recorded at two receivers at locations A and B in a medium with a random noise field $f(t, r)$, (which is assumed to be a white noise distributed everywhere in the medium) as shown in Figure 2.1. The time domain cross-correlation between the two receiver locations can be defined as:

$$C(\tau, r_A, r_B) = \lim_{T \to \infty} \frac{1}{T} \int_0^T u(t, r_A) \, \overline{u(t + \tau, r_B)} dt \qquad (2.1)$$

In equation 2.1, the bar denotes the conjugate. Using the Green's function and the source function $f(t, r)$ we can write the displacement field as

$$u(t, r) = \int_0^\infty dt' \int_x G_a(t', r, r_s) \, f(t - t', r_s) \, dr_s \qquad (2.2)$$

Then the cross-correlation becomes

$$C(\tau, r_A, r_B) = \lim_{T \to \infty} \frac{1}{T} \int_0^T dt \int_0^\infty ds \int_x dr_s \, G_a(s, r_A, r_s) \, f(t - s, r_s)$$

$$* \int_0^\infty ds' \int_x dr_s' \, \overline{G_a(s', r_B, r_s') \, f(t + \tau - s', r_s')}$$

In the frequency domain, white noise contains all frequencies with a random phase and in the time domain, this is a random wavefield such that the position and firing time of

each source is uncorrelated (Gouedard et al., 2008) and in this case, the limit T can be replaced by an ensemble average:

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T f(t-s,r_s)\, f(t+\tau-s',r_s')dt = E[f(t-s,r_s)\, f(t+\tau-s',r_s')]$$

$$= \sigma^2\, \delta(\tau+s-s')\delta(r_s-r_s')$$

where σ is variance of white noise

Now we have

$$C(\tau,r_A,r_B) = \sigma^2 \int_0^\infty ds \int_x dr_s\, G_a(s,r_A,r_s)\overline{G_a(s+\tau,r_B,r_s)} \tag{2.3}$$

By using the expressions of the Green's function in an attenuating medium, Green's functions of the positive and negative lags can be obtained as:

$$\frac{d}{d\tau}\, C(\tau,r_A,r_B) = \frac{-\sigma^2}{4a}\, (G_a(\tau,r_A,r_s) - G_a(-\tau,r_A,r_s)) \tag{2.4}$$

The Equation 2.4 implies that the time derivative of the cross-correlation computed between the wavefields recorded at A and B gives the Green's function of the medium (Gouedard *et al.*, 2008)

In the context of seismology the Green's function of a medium between two points A and B represents the recorded signal at A if an impulsive source is applied at B. In a completely random wavefield, the cross-correlation of signals recorded between two points converges to the complete Green's function of the medium, including all reflection, scattering and propagation modes (Weaver & Lobkis, 2001).

2.2 Seismic Interferometry

The term "Seismic Interferometry" refers to the principle of generating new seismic responses of virtual sources by cross-correlating seismic observations at different receiver locations.
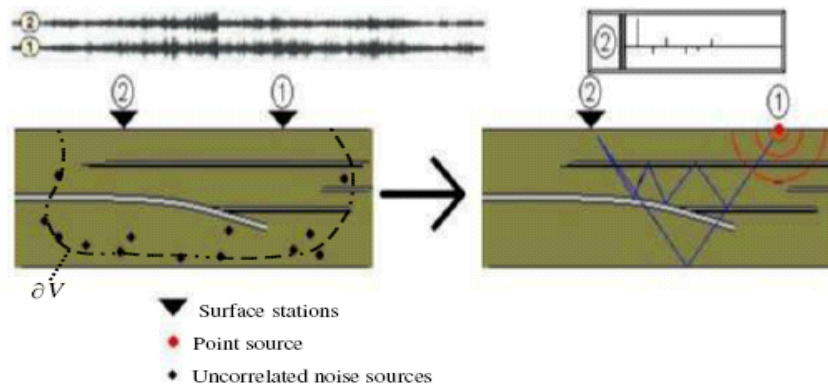
Figure 2.2: Schematic diagram showing the cross-correlation of noise at two surface stations

The idea behind seismic interferometry (Green's Function retrieval) started when Claerbout (1967), postulated that in 1-D, the autocorrelation of noise generated at depth at a surface station yields the reflection profile of the earth under that station. This implies that using two different surface stations yields the impulse response function between these stations as shown in the Figure 2.2. Recent works by Wapenaar & Fokkema (2005), and others have verified both in theory and in practice, the 3D elastodynamic generalization of the 1-D postulate by Claerbout (1967).

The basic theory behind seismic interferometry is that Green's function between two seismic stations can be estimated by cross correlating long time series of ambient noise recoded at those stations. This Green's function may be thought of as the seismogram recorded at one location due to an impulsive or instantaneous source of energy at the other.
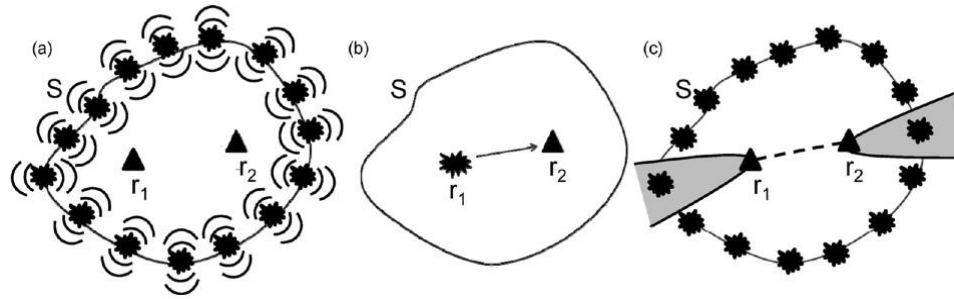
Figure 2.3: (a) Two receivers (triangles) are surrounded by a boundary S of sources each of which sends a wavefield into the interior and exterior of S. (b) The seismic interferometry method turns one of the receivers (r_1) into virtual source from which a real seismogram is obtained. (c) Sources located within the grey regions contribute the most to the Green's function computation. (Nicolson et al., 2012).

Consider two receivers at positions r_1 and r_2 which are surrounded by energy sources located on an arbitrary surroundings with a boundary S as shown in the Figure 2.3a. The wavefield emanating from each source propagates into medium in the interior of S and is recorded at both receivers. The signals recorded at the two receivers are then cross-correlated. When the cross-correlation of all the sources are added together, the energy that travels along the path will add constructively and the energy that doesn't travel along the path will add destructively. Thus the Green's function obtained between r_1 and r_2 will be as if one of the receivers had actually been a source (Wapenaar, 2003, 2004).

For the case of random noise, a surface S exists such that it joins up all the noise sources and since noise sources may all fire at the same or at overlapping time, their recorded signals at two receivers are already summed together which actually takes place naturally. Snieder (2004) showed that the seismic sources located around the extensions of the inter-receiver path contribute most to the interferometric Green's function construction and thus whole boundary of source is not necessary in order to approximate the inter-receiver Green's function. Since strong sources of seismic noise are in general restricted to locations within the Earth's crust and surface wave travel along interfaces between different layers over crust and upper mantle, the surface wave

part of inter-receiver Green's function appear clear in the seismogram constructed from seismic interferometry.
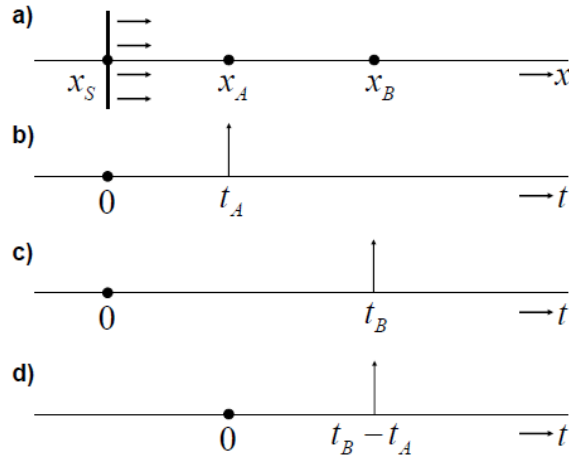
Figure 2.4: 1D example of direct wave interferometry

To describe the phenomenon of seismic interferometry in detail, we use the demonstration as given in Wapenaar et al., (2004) for one dimensional direct wave interferometry.

Seismic interferometry involves the cross-correlation of responses at two receivers at x_A and x_B when a plane wave (an impulse) is incident on them from a source x_s (Figure 2.4a). The response recorded by the first receiver is $G(x_A, x_S\ t)$, and the response by the second receiver is $G(x_B, x_S\ t)$ where G is the Green's function. Here, $G(x_A, x_S\ t) = \delta(t - t_A)$ and $G(x_B, x_S\ t) = \delta(t - t_B)$. Where t_A and t_B are the times at which the signal is received at the respective receivers.

In the cross-correlation, the traveltime along the common path between the receivers cancels leaving the traveltime along the remaining path which is shown as an impulse at $t_B - t_A$ in the Figure 2.4d. This impulse can be interpreted as the response of a source at x_A, observed by a receiver at x_B, i.e., the Green's function $G(x_B, x_A\ t)$. By using the definition of cross correlation, we have:

$$G(x_B, x_A\ t) = G(x_B, x_S\ t) * G(x_A, x_S, -t) \qquad (2.5)$$

If we replace the source (here an impulse) by a noise wavelet $s(t)$, then the responses at x_A and x_B are given by $u(x_A, x_S, t) = G(x_A, x_S, t) * s(t)$ and $u(x_B, x_S, t) = G(x_B, x_S, t) * s(t)$. Now, the cross-correlation of $u(x_A, x_S, t)$ and $u(x_B, x_S, t)$ gives:

$$G(x_B, x_A, t) * S_s(t) = u(x_B, x_S, t) * u(x_A, x_S, -t) \qquad (2.6)$$

Where $S_s(t)$ is the autocorrelation of the wavelet $s(t)$, i.e., $S_s(t) = s(t) * s(-t)$.
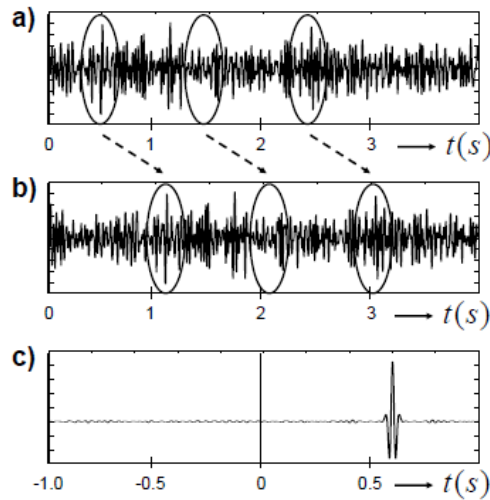
Figure 2.5: Signal recorded at A and B for a Noise source $N(t)$ at x_S and cross-correlation

Figures 2.5a and 2.5b show the responses at x_A and x_B, of a noise source $N(t)$ at x_S. If the receivers are separated by a distance of 2400m and the propagation velocity is 4000 m/s, the traveltime between these receivers is 0.6 s. Therefore, the signal at x_B is delayed by 0.6s relative to the signal at x_A.

In the figure 2.5c, the cross-correlation contains an impulse at t = 0.6 s which is the traveltime from x_A to x_B. We get the velocity between the receivers by the traveltime estimated from the Green's function (0.6 s) as $\frac{2400}{0.6} = 4000\ m/s$. This shows that seismic interferometry can be used for tomographic inversion.

This explanation can be extended to 2D and 3D wave interferometry where there are noise sources all around (Snieder, 2004). Here we examine the 2D configuration and note some important points.
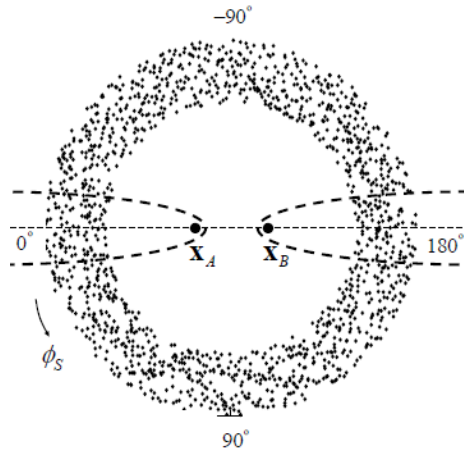
Figure 2.6: Distribution of point sources illuminating the receivers x_A and x_B. The thick dashed lines indicate the Fresnel zones (Wapenaar, 2003)

In the 2D configuration as shown in the Figure 2.6, there are many point sources denoted by black dots, located as in the geometry of a "pineapple slice". If we consider polar coordinates, the positions of the sources are denoted (r_s, ϕ_s). Here we notice that not only the sources exactly at $\phi_s = 0°$ and $\phi_s = 180°$ contribute to these events, but also the sources in Fresnel zones around these angles, denoted by the dashed lines in Figure 2.6. The events in all traces outside the Fresnel zones interfere destructively and hence give no coherent contribution (Snieder, 2004).

In a layered medium, we see that surface waves can be approximated as the solutions of a wave equation with a frequency dependent propagation velocity (will be discussed in the next section). We can use the same argument as above for noise sources as well and extract the Green's function by cross-correlating the noise recordings at two seismometers. This process can be augmented by the use of many seismometers as the number of cross correlation pairs increases as a power of two i.e., if there are n seismometers, the number of cross correlation pairs is given by

$$n(n-1)/2$$

Hence, each seismometer can be used as a virtual source, whose response is recorded by all other seismometers which is very useful when there are many seismometers in the array.

2.3 Surface waves and their dispersive nature

2.3.1 Wave equation

Once we have the Empirical Green's function, it can be regarded as a virtual seismogram. The equations governing the wave motion in seismology are

$$\nabla^2 \phi(x,t) = \frac{1}{\alpha^2}\frac{\partial^2 \phi(x,t)}{\partial t^2} \quad \text{For P waves with a scalar potential } \phi(x,t) \qquad (2.7)$$

$$\nabla^2 \Gamma(x,t) = \frac{1}{\beta^2}\frac{\partial^2 \Gamma(x,t)}{\partial t^2} \quad \text{For S waves with a vector potential } \Gamma(x,t) \qquad (2.8)$$

To understand the displacement caused by the two types of waves, consider a plane wave propagating in the z direction. The scalar potential for a harmonic plane P wave satisfying Equation 2.7 is $\phi(z,t) = A \exp(i(\omega t - kz))$. So the resulting displacement is the gradient $u(z,t) = \nabla\phi(z,t) = (0,0,-ik)A \exp(i(\omega t - kz))$, which has non-zero component only along the propagation direction z.

For S wave, the vector potential satisfying Equation 2.8 is given by $\Gamma(z,t) = (A_x, A_y, A_z)\exp(i(\omega t - kz))$ and the resulting displacement field is given by the curl $u(z,t) = \nabla \times \Gamma(z,t) = (ikA_y, -ikA_z, 0) \exp(i(\omega t - kz))$, whose component along the propagation direction is zero. Thus the only displacement associated with a propagating shear wave is perpendicular to the direction of propagation.

Let us define z axis as the vertical direction and orient the $x - z$ plane along the great circle connecting a seismic source and receiver. Plane waves travelling on a direct path between the source and the receiver thus propagate in the $x - z$ plane. The shear wave polarization directions are SV (Displacement in the vertical plane) and SH (Displacement in the horizontal y-direction). The SH and SV wave displacements are perpendicular to the propagation direction as shown in the Figure 2.7
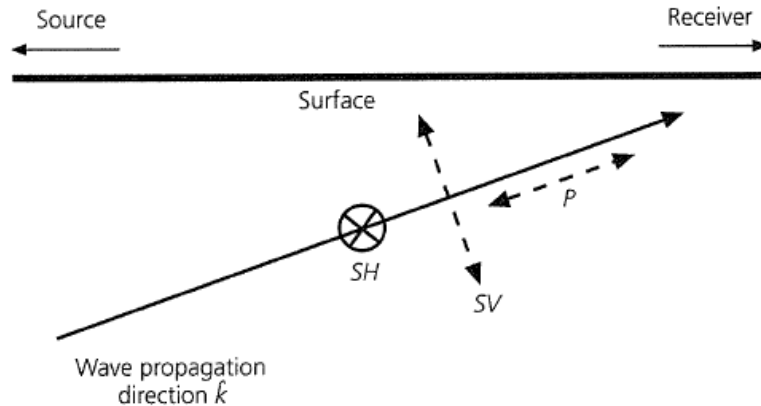
Figure 2.7: Displacement fields for plane P and S waves in x-z plane showing SV and SH polarizations

With this framework, in a layered medium we find that P and SV waves are coupled to each other since the displacement produced by their respective potentials are in $x - z$ plane. Whereas SH wave is not coupled to any wave in the sense that the displacement produced by the SH wave potential is only in the y-direction.

2.3.2 Surface Waves

A typical seismogram consists of longer period waves with large amplitudes that arrive after P and S waves. These waves are called surface waves because their energy is concentrated near the earth's surface. Due to geometric spreading, their energy spreads two-dimensionally and decays with distance r from the source as r^{-1}, whereas energy of the body waves spreads three dimensionally and decays as r^{-2}. Thus at large distances from the source, surface waves are more prominent in the seismograms.

At a free surface (a traction-free or stress-free surface), constructive interference of incident P and S waves generate surface waves which propagate parallel to the surface. There are two types of surface waves known as Love waves and Rayleigh waves named after their discoverers. Love waves (named after A.H.E. Love, British mathematician, 1911) are the result of SH waves trapped near the surface and

Rayleigh waves (named after Lord Rayleigh, 1885) are the result of a combination of P and SV motions. In this study, we consider only Rayleigh wave dispersion to perform ambient noise tomography.

For laterally homogeneous models, Rayleigh waves are radially polarized and exist at any free surface, whereas Love waves are transversely polarized and require some velocity increase with depth. Due to their dispersive nature these waves provide useful information on the structure of crust and the upper mantle in the Earth. Rayleigh waves exist in a half space but do not show dispersion in this scenario. But when we consider a layered model of earth then Rayleigh waves become dispersive in nature, a phenomena which is central to our study. First let us consider Rayleigh waves in a half space to see why there is no dispersion in this case.

2.3.3 Rayleigh waves in a half space

Rayleigh waves are a combination of P and SV motions. Consider a wave travelling in the x-z plane with propagation along x direction. Hence, the potentials giving rise to those motions are given by

$$\phi = A\, exp(-ikrz + ik(x - ct)) \quad \text{For P waves} \qquad (2.9)$$

$$\psi = B\, exp(-iksz + ik(x - ct)) \quad \text{For SV waves} \qquad (2.10)$$

Here A and B are the amplitudes, k is the wavenumber

$$r = \left(\frac{c^2}{\alpha^2} - 1\right)^{\frac{1}{2}} \quad ; \qquad s = \left(\frac{c^2}{\beta^2} - 1\right)^{\frac{1}{2}}$$

Also, c is the apparent velocity of the waves along x direction, α and β are the velocities of P and S waves respectively.

The following boundary conditions are to be met for the surface waves to exist:

1. The energy does not propagate away from the surface

For the energy to be trapped near the surface, the exponentials $\exp(-ikrz)$ and $\exp(-iksz)$ must have negative real exponents, so that the displacement will decay as $z \to \infty$. Because $r = \left(\frac{c^2}{\alpha^2} - 1\right)^{\frac{1}{2}}$ and $s = \left(\frac{c^2}{\beta^2} - 1\right)^{\frac{1}{2}}$ this condition requires that $c < \beta < \alpha$, so that both square roots become imaginary, with a choice of sign such that

$r = -i\left(\frac{c^2}{\alpha^2} - 1\right)^{\frac{1}{2}}$ and $s = -i\left(\frac{c^2}{\beta^2} - 1\right)^{\frac{1}{2}}$. Thus c, the apparent velocity must be less than the shear velocity.

2. Free surface boundary conditions are to be met

At the free surface, i.e., $z = 0$, the traction vector and hence the stress components $\sigma_{xz}, \sigma_{yz}, \sigma_{zz}$ must be zero for all x and t. σ_{yz} is automatically zero for P-SV waves in this geometry. Expressing stress components in terms of potentials,

$$\sigma_{xz} = 2\mu e_{xz} = \mu\left(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x}\right) = \mu\left(2\frac{\partial^2 \phi}{\partial x \, \partial z} + \frac{\partial^2 \psi}{\partial x^2} - \frac{\partial^2 \psi}{\partial z^2}\right) = 0$$

$$\sigma_{zz} = \lambda\theta + 2\mu e_{zz} = \lambda\left(\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial z^2}\right) + 2\mu\left(\frac{\partial^2 \phi}{\partial z^2} + \frac{\partial^2 \psi}{\partial x \partial z}\right) = 0$$

By substituting the potentials in the equations, we get,

$$2rA - (1 - s^2)B = 0 \tag{2.11}$$

$$[\alpha^2(r^2 + 1) - 2\beta^2]A - 2\beta^2 sB = 0 \tag{2.12}$$

This is a homogeneous system of equations; apart from the trivial solution, we get the other solution by equating the determinant to be zero:

$$\begin{vmatrix} 2r & -(1 - s^2) \\ \alpha^2(r^2 + 1) - 2\beta^2 & 2\beta^2 s \end{vmatrix} = 0$$

Solving the determinant gives

$$[\alpha^2(r^2 + 1) - 2\beta^2](1 - s^2) - 4rs\beta^2 = 0$$

23

On replacing the values of r and s from above and taking in to account that they are imaginary, we have:

$$(2 - \frac{c^2}{\beta^2})^2 = 4(1 - \frac{c^2}{\alpha^2})^{1/2}(1 - \frac{c^2}{\beta^2})^{1/2} \tag{2.13}$$

The above Equation 2.13 is known as Rayleigh's equation in honor of Lord Rayleigh, who solved this problem for the first time in 1887. This equation shows that there is no dispersion in this case because the apparent velocity is a constant and not a function of wavenumber or frequency $c(k)\ or\ c(\omega)$. From the above relation it can be shown that the Rayleigh wave velocity is related to the body wave velocities by Poisson's ratio. Since body wave velocities are constant with depth, the Rayleigh wave velocity in a homogeneous half-space is independent of frequency.

2.3.4 Rayleigh waves in an elastic layer over a half-space
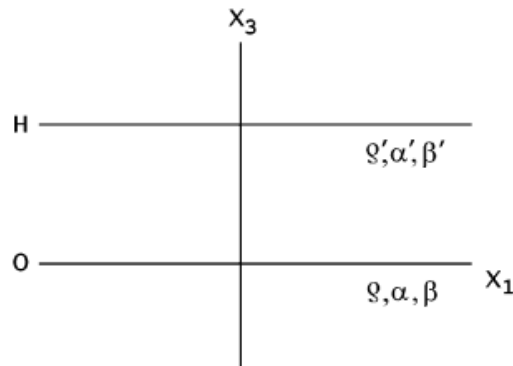
Figure 2.8: An elastic layer of thickness H over an elastic Half-space here $x_1 = x, x_3 = z$ axes respectively.

In this case we consider P and SV motions that propagate in the x-direction with a velocity c in a medium consisting of a layer of thickness H, density ρ' and velocities α', β' over a half space of density ρ and velocities α and β (Figure2.8). Here, the waves travel in both the positive and negative directions of z-axis inside the layer but only in the positive direction in the half-space. Taking in to account that amplitudes must decrease with depth in the half-space, the potentials are given by:

$$\phi' = A' \exp\big(ikr'z - ik(x - ct)\big) + B' \exp\big(-ikr'z - ik(x - ct)\big) \qquad (2.14)$$

$$\psi' = C' \exp\big(iks'z - ik(x - ct)\big) + D' \exp\big(-iks'z - ik(x - ct)\big) \qquad (2.15)$$

$$\phi = A \exp(-ikrz - ik(x - ct)) \qquad (2.16)$$

$$\psi = B \exp(-ikrz - ik(x - ct)) \qquad (2.17)$$

Where r and s are given by $r = -i\left(\frac{c^2}{\alpha^2} - 1\right)^{\frac{1}{2}}, s = -i\left(\frac{c^2}{\beta^2} - 1\right)^{\frac{1}{2}}$

and similarly r' and s' (on replacing α by α' and β by β'). Since amplitudes must decrease with depth $(-z)$, r and s must be imaginary and positive and consequently we have $c < \beta < \alpha$.

The boundary conditions for Rayleigh waves to exist are:

1. $z = H$, for the free surface, the components of stress are equal to zero.

2. At the contact surface $z = 0$. Here the stress and displacement components should be continuous.

As functions of displacements, the boundary conditions result in the following equations:

At $z = H$

$$\tau'_{31} = 0 \quad \Longrightarrow \quad \frac{\partial u'_z}{\partial x} + \frac{\partial u'_x}{\partial z} = 0$$

$$\tau'_{33} = 0 \quad \Longrightarrow \quad \lambda' \frac{\partial u'_x}{\partial x} + (\lambda' + 2\mu')\frac{\partial u'_z}{\partial z} = 0$$

At $z = 0$,

$$u'_x = u_x$$

$$u'_z = u_z$$

$$\mu'\left(\frac{\partial u'_z}{\partial x} + \frac{\partial u'_x}{\partial z}\right) = \mu\left(\frac{\partial u_z}{\partial x} + \frac{\partial u_x}{\partial z}\right)$$

$$\lambda' \frac{\partial u'_x}{\partial x} + (\lambda' + 2\mu') \frac{\partial u'_z}{\partial z} = \lambda \frac{\partial u_x}{\partial x} + (\lambda + 2\mu) \frac{\partial u_z}{\partial z}$$

We know that the displacements are given by:

$$u_x = \frac{\partial \phi}{\partial x} - \frac{\partial \psi}{\partial z}$$

$$u_z = \frac{\partial \phi}{\partial z} + \frac{\partial \psi}{\partial x}$$

On substituting the displacements in terms of the potentials according to the equations 2.14, 2.15, 2.16, 2.17, we get a system of six equations for six unknowns $A', B', C', D', A \ and \ B$ i.e., the amplitudes of potentials in the layer and half space. The condition for a solution is that the determinant of the system is null. Making the determinant equal to zero, we find an equation for c, the velocity of Rayleigh waves. Since this equation implies that the velocity $c(k)$ is a function of the frequency, Rayleigh waves in a layer over a half-space are dispersed. The phenomenon of dispersion is explained in the next section.
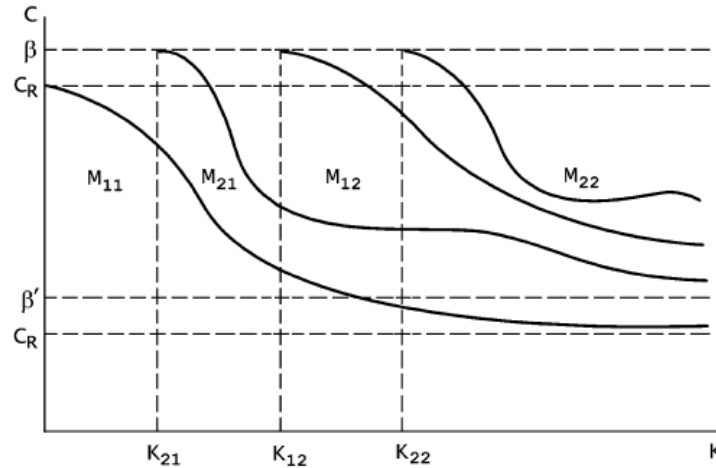
Figure 2.9: Dispersion curves for Rayleigh waves in an elastic layer over a half space

Solutions to the dispersion equation give the velocity of the Rayleigh waves as a function of frequency $c(k)$ as shown in the Figure 2.9. There are an infinite

number of solutions corresponding to different modes. There are Symmetric modes for which vertical displacements at free surface and contact surface have opposite signs and Anti symmetric modes in which they have same sign. In this study we consider only M_{11} i.e., fundamental mode Rayleigh waves to obtain the dispersion curves.

2.4 Dispersion

In the last section we saw that the Rayleigh waves are dispersive in a layer over a half-space because the apparent velocity along the surface varied with frequency. To explain the phenomenon of dispersion let us consider the net effect of two harmonic waves with slightly different frequencies and wavenumbers.

$$u(x,t) = A[cos(\omega_1 t - k_1 x) + cos(\omega_2 t - k_2 x)] \tag{2.18}$$

The angular frequencies and wavenumbers can be written in terms of the differences from their average values frequency ω and wave number k as

$$\omega_1 = \omega - \delta\omega, \qquad k_1 = k - \delta k$$

$$\omega_2 = \omega + \delta\omega, \qquad k_2 = k + \delta k$$

Therefore

$$u(x,t) = A[\cos(\omega t - \delta\omega t - kx + \delta kx) + \cos(\omega t + \delta\omega t - kx - \delta kx)]$$
$$= [2A\cos(\delta kx - \delta\omega t)]\cos(\omega t - kx)$$

Thus the sum of two harmonic waves is a product of two cosine functions which correspond to two propagating harmonic waves. Because $\delta\omega$ and δk are less than ω and k , the first term varies more slowly with time and also in space. Thus we have a carrier wave with angular frequency ω and wave number k on which a slower varying envelope with angular frequency $\delta\omega$ and wave number δk is superimposed. They travel with a velocity ω/k and $\delta\omega/\delta k$ respectively. The former is the phase velocity c and the latter is the group velocity U.

2.4.1 Group velocity

The property of a dispersive signal is that different frequency components arrive at different times. If we apply an array of narrow filters to the complex seismic signal, the filters may resolve transient signals composed of several dominant periods that arrive at the recording station almost simultaneously. Using the filtered amplitudes, we can calculate the group velocity. Since Rayleigh waves contain waves of different periods with different velocities, we can find group velocity from the recorded waveform at a seismic station in the following way. We use Frequency Time Analysis (FTAN) to separate different frequencies.

Let $S(t)$ be the waveform obtained as the Green's function from cross correlating noise records at two stations. The Fourier transform of $S(t)$ with positive exponent is

$$S(\omega) = \int_{-\infty}^{\infty} S(t)\, e^{i\omega t} dt \qquad (2.19)$$

The dispersion measurements are obtained by considering only the analytic signal which is defined in the frequency domain as

$$S_a(\omega) = S(\omega)\big(1 + sgn(\omega)\big)$$

The analytic signal is subjected to a set of narrow band pass Gaussian filters with center frequencies ω_o to separate different frequencies.

$$S_a(\omega, \omega_o) = S(\omega)\big(1 + sgn(\omega)\big)G(\omega - \omega_o) \qquad (2.20)$$

$$\text{Where } G(\omega - \omega_o) = e^{-\alpha\left(\frac{\omega - \omega_o}{\omega_o}\right)^2}$$

Applying inverse Fourier transform to each band passed function of equation (2.20) i.e. for different ω_o brings us back to the time domain which yields a smooth 2-D envelope function $|A(t)|$ and $\phi(t, \omega_o)$. Group velocity is measured using $|A(t)|$ and phase velocity is measured using $\phi(t, \omega_o)$.

The group arrival time $\tau(\omega_\circ)$ as a function of the centre frequency of the Gaussian filter is calculated by using the peak of the envelope function. The group velocity is given by

$$U(\omega_\circ) = \frac{r}{\tau(\omega_\circ)} \qquad\qquad (2.21)$$

Where r is the distance between the source and the receiver. Dziewonski *et al* (1969) developed the multiple filter technique (do_MFT) to determine the group velocities of a dispersive wave. The energy in the wave propagates as the envelope of the wave packet at a speed known as group velocity. The packet of energy that propagates as a surface wave contains a spectrum of periods. Each period can be measured from the time between successive peaks or troughs. The wave with longest period travels fastest and appears first on the seismogram. The group velocity is calculated by dividing the interstation distance by travel time of the wave group as in the equation. For e.g. if distance between two stations is 500 km, then wave group with a period 20 sec and travel time of 156 sec will have a group velocity of $\frac{500\ km}{156\ s} = 3.2\ km/s$. The later arriving group with period 15 sec and travel time of about 168 sec has a group velocity $\frac{500\ km}{168\ s} = 2.97\ km/s$. In this study we use only Group velocity dispersion curves obtained with the above procedure for performing traveltime tomography.

2.5 Tomography

2.5.1 Introduction

Once we have obtained dispersion curves from the empirical Green's functions, we can calculate travel times of Rayleigh waves at a given frequency. A 2D tomographic inversion can be performed on the travel times to estimate the variations in group velocity over the region under consideration at the specified frequency. We can obtain different 2D models for each frequency and then use the relationship between Rayleigh wave frequency and sampling depth to infer the velocity variations with depth.

Tomography can be defined as the reconstruction of an internal property of a medium from line integrals through the medium. If we represent some elastic property of the subsurface, (say velocity) by a set of model parameters **m**, then we can predict a set of data **d** (here traveltimes) for a given source receiver array by the use of a physical model **g**. The relation between data and model parameters is given as

$$\boldsymbol{d} = \boldsymbol{g}(m) \tag{2.22}$$

For an observed data set $\boldsymbol{d}_{obs}$ and an initial model m_0, the difference $\boldsymbol{d}_{obs} - \boldsymbol{g}(m_0)$ gives the error in the estimated data set. The inverse problem is to manipulate m in order to minimize the difference between observed and estimated data subject to any regularization like constraints on model parameters etc.

There are four major steps to produce a tomographic image from seismic data. They are:

1. Model parameterization: The seismic structure of the region being mapped is defined in terms of a set of unknown model parameters. Tomographic methods generally require an initial estimate of model parameter values to be specified.

2. Forward calculation: Calculation of model data (e.g. traveltimes) for a set of given model parameters.

3. Inversion: Updating the model parameter values to better match the model data to the observed data subject to regularization conditions.

4. Analysis of solution robustness: Based on estimates of covariance and resolution from linear theory or on the reconstruction of test models using synthetic datasets.

In seismic traveltime tomography the model data are traveltimes and the model parameters define velocity variations. The traveltime of a ray in a continuous velocity medium $v(x)$ is:

$$t = \int_{L(v)} \frac{1}{v(x)} dl \tag{2.23}$$

Where $L(v)$ is the ray path and $v(x)$ is the velocity field. The equation is non-linear since the integration path depends on the velocity.

The linearization assumption commonly adopted in traveltime tomography is reasonable provided it can be shown that the source-receiver path is not significantly perturbed by the adjustments made to the model parameter values in the inverse step.

We find that for a perturbation $\delta v(x)$ in $v_0(x)$, so that $v(x) = v_0(x) + \delta v(x)$, the new path becomes $L(v) = L_0 + \delta L$ and $t = t_0 + \delta t$. The travel time is then

$$t = \int_{L_0+\delta L} \frac{1}{v_0 + \delta v} \, dl \tag{2.24}$$

By expanding the integrand using geometric series and ignoring higher order terms, we get the perturbation in travel time as:

$$\delta t = -\int_{L_0} \frac{\delta v}{v_0^2} \, dl + O(\delta v^2) \tag{2.25}$$

The above Equation 2.25 implies that if velocity along the path is perturbed, then the corresponding traveltime perturbation calculated along the original path will be accurate to first order.

In terms of slowness, we have $s(x) = \frac{1}{v(x)}$ and the traveltime in terms of slowness is given by:

$$\delta t = -\int_{L_0} \delta s \, dl + O(\delta s^2) \tag{2.26}$$

Here, Traveltime perturbation δt is linearly dependent on δs. This is the linearization assumption.

Now, the four important steps to produce tomographic images are dealt in detail.

2.5.2 Model Parameterization

The traveltime of a seismic wave between source and receiver is solely dependent on the velocity structure of the medium through which the wave propagates. Therefore, subsurface structure in a seismic traveltime inversion is represented by variations in *P* or *S* wave velocity (or slowness). These velocity variations may be defined by a set of interfaces whose geometry is varied to satisfy the data, a set of constant velocity blocks or nodes with a specified interpolation function, or a combination of velocity and interface parameters. The most appropriate choice will depend on the *a priori* information (e.g. known faults or other interfaces), whether or not the data indicates the presence of interfaces (e.g. reflections, mode conversions),

whether data coverage is adequate to resolve the trade-off between interface position and velocity, and finally, the capabilities of the inversion routine.

Various ways of Velocity Parameterization are

1. Constant velocity blocks, which are simple to define and result in linear ray paths within each block but are not natural choice for representing smooth variations in subsurface structure due to the velocity discontinuities that exist between adjacent blocks.

2. To define velocities at the vertices of a rectangular grid together with a specified interpolation function. Higher order interpolation functions must be used if the velocity field is to have continuous first and second derivatives which are required for some ray tracing methods (Thomson & Gubbins, 1982).

In this study we use grid based velocity parameterization for solving the Eikonal equation as will be shown in the next section.

2.5.3 Forward calculation – Traveltime determination using FMM

The calculation of ray traveltimes between known end points through a given velocity structure is called the forward problem. In traveltime tomography, there are many ways of determining source-receiver traveltimes such as ray tracing (shooting and bending), wavefront tracking schemes (such as finite difference solutions of the eikonal equation). There are certain drawbacks in the method of ray tracing such that robustness, speed and uniqueness of the derived ray path. To overcome these problems the method of Fast Marching (a variant of wavefront tracking) has been used by Rawlinson and Sambridge (2003, 2004a, b) rather than using ray tracing and then the inversion is carried out. The method (FMM) was originally developed by Sethian (1996) for tracking advancing interfaces.

The fast marching method (FMM) is a grid based numerical scheme for tracking the evolution of monotonically advancing interfaces via finite-difference solution of the eikonal equation. The eikonal equation implies that the magnitude of the travel

time gradient at any point along a wave front is equal to the inverse of the velocity at that point as given in the Equation 2.27:

$$|\nabla_x T| = \frac{1}{v(x)} = s(x) \tag{2.27}$$

Where ∇_x is gradient operator, T is travel time and $s(x)$ is slowness. The eikonal equation cannot be easily solved in the presence of gradient discontinuities because the equation itself requires $\nabla_x T$ to be defined. Thus we need to find a weak solution to the eikonal equation which results in a continuous $T(x)$, but not necessarily continuous $\nabla_x T$. A weak solution to a differential equation is an entropy satisfying approximate solution that is not differentiable everywhere but satisfies an integral formulation of the equation. One way to solve this is to solve the viscous version of the eikonal equation

$$|\nabla_x T| = s(x) + \in \nabla_x^2 T \tag{2.28}$$

As $\in \to 0$ (the viscous limit), where the parameter $\in$ controls the smoothness imposed on the solution. The limit of smooth solutions is a weak solution that corresponds to the first-arriving wave front. The FMM of Sethian and Popovici (1999) for solving the Eikonal equation on a 3-D grid uses this approach. In its simplest form, the FMM uses the first-order upwind difference scheme as below:

$$[\max(D_{ijk}^{-x}T, -D_{ijk}^{+x}T, 0)^2 + \max(D_{ijk}^{-y}T, -D_{ijk}^{+y}T, 0)^2 + \max(D_{ijk}^{-z}T, -D_{ijk}^{+z}T, 0)^2]^{1/2} = S_{ijk}$$

Where S_{ijk} is the slowness at the grid point (i, j, k). Forward and backward finite difference operator notation used is:

$$D_{ijk}^{+x} T = \frac{T_{(i+1)jk} - T_{ijk}}{h} \tag{2.29}$$

$$D_{ijk}^{-x} T = \frac{T_{ijk} - T_{(i-1)jk}}{h}$$

where T_{ijk} is the value of T at grid point (i, j, k) with grid spacing h. Similar expressions can be written for y and z. The upwind scheme given by eq. (2.29) describes how to calculate new traveltimes using known traveltimes from adjacent grid points. The updating of nodes should be in the direction of flow and to achieve this, FMM systematically constructs traveltimes T in a downwind fashion from known values

upwind by employing a narrow-band approach. The narrow band represents the propagating wavefront and grid points are tagged as *alive*: the points where the travel times are already correctly assigned, *close*: the points have some assigned trial values and *far*: the points without the computed travel times (Figure 2.10).
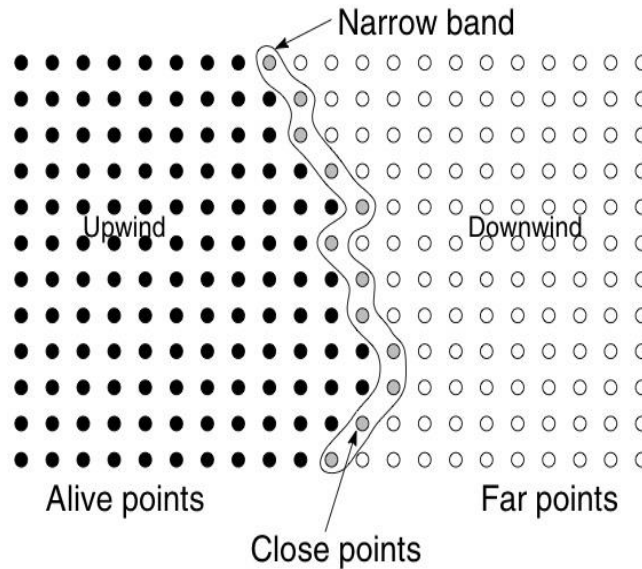
Figure 2.10: Principle of the narrow band method. Alive points lie upwind of the narrow band while far points lay downwind (Rawlinson and Sambridge, 2003).

The FMM begins from a source point (wave front) and travel time at neighboring grid points are calculated using Equation 2.29 to form the first stage of the narrow band (Figure 2.10). The point with minimum travel time is then accepted as *alive*, and all neighboring points to this *alive* point are updated (if *close*) or calculated for the first time (if *far*), in which case they become *close*. The calculation scheme will be repeated till all of the *far* points become *alive* points hence the propagation of the wavefront is tracked completely.
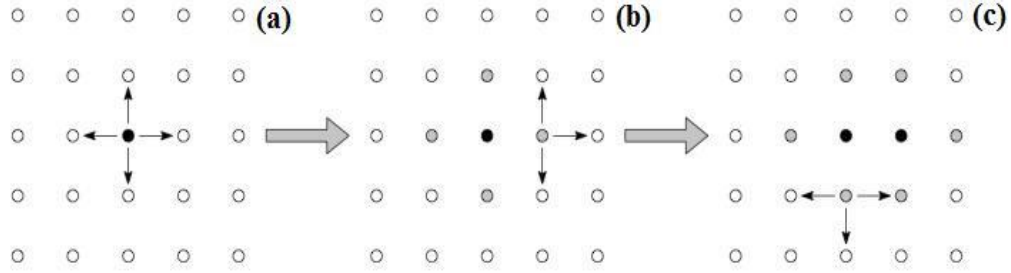
Figure 2.11: Principle of FMM in 2-D. (a) Starting from the source point (black dot) in the center of a grid, travel times to the four neighboring grid points are determined using equation 2.29. (b) The smallest of these four values (gray dots) must be correct, so all *close* neighbors to this point that are not *alive* (white dots) have their values computed, and added to the narrow band defined by the gray dots. (c) The smallest of these six close points again must be correct, and all neighboring points have their values computed (or recomputed) (Rawlinson and Sambridge, 2003).

Figure 2.11 illustrates several evolution steps of the narrow-band method from a source point. A heap sort algorithm is used to rapidly locate the grid point in the narrow band with minimum traveltime. Once the travel time between all station pairs are calculated, we invert them using subspace inversion.

2.5.4 Inversion of traveltime data

The inversion step, which involves the correction of the model parameters **m** to better satisfy the observed data d_{obs} through the known relationship $d = g(m)$, can be performed in a number of ways such as back projection, gradient methods and global optimization. Here we use Subspace inversion technique based on gradient methods to solve the inverse step. To set the inverse problem, the underlying medium is divided into 3-D blocks where slowness can be reconstructed. If the traveltime integral is discretized for a single ray in the jth block, then

$$\delta t = -\int_{L_0} \delta s \, dl + O(\delta s^2)$$

35

$$\delta t = \sum_j l_j \, \Delta s_j \tag{2.30}$$

Where Δs_j is slowness in block j and l_j is the path length in the corresponding block. For i number of multiple rays the relation becomes

$$\delta t_i = \sum_j l_{ij} \, s_j \tag{2.31}$$

In general it can be written as $d = Gm$, where d is observed data, G is data kernel and m is model parameter matrix. Solution of the above equation using generalized inverse method provides estimate of the model parameter, m_{est} and can be written as:

$$m_{est} = [G^T G]^{-1} G^T G d \tag{2.32}$$

Then misfit becomes $\boldsymbol{d}_{obs} - \boldsymbol{g}(m_{est})$. Now we update m iteratively again by minimizing the new misfit until sufficient convergence is achieved.

But the inverse problem in seismic tomography can be formulated as minimizing an objective function $S(m)$ consisting of a data residual term and one or more regularization terms. The misfit can be quantified by constructing an objective function $S(m)$ consisting of a weighted sum of data misfit and regularization terms.

If part of the problem is underdetermined, the solution is not unique. To overcome this, an *a priori* information related to model parameters can be introduced in the objective function as a regularization term:

$$S(m) = (d - Gm)^T (d - Gm) + \varepsilon (m - m_0)^T (m - m_0) \tag{2.33}$$

Where ε is the damping parameter to control the underdetermined part on the solution and m_0 is the priori information about the model as reference model.

The first term in the R.H.S of equation (2.33) measures the difference between the observed and predicted data which can be a L_1 norm or L_2 norm etc. But, if it is assumed that the error in the relationship is Gaussian, then a least square or L_2 measure is considered. If we have uncertainty estimates for the observed data, then we can give greater weight to more accurate data by introducing a data weighting term called data weighting matrix C_d^{-1}. The second term in the R.H.S of equation (2.33) provides an additional constraint on the model parameters to reduce the non-

uniqueness of the solution. Also some part of the apriori information (model parameters) may have greater significance than the others and hence the relative weight can be given to the parameters by an inverse model parameter covariance matrix C_m^{-1}.

A smoothness/flatness matrix D which attempts to find a trade-off between satisfying the data and finding a model with minimum amount of structural variation is also included to the formulation to control the outcome of the model parameter variations (Sambridge, 1990). These constraints are incorporated in to the objective function as:

$$S(m) = (d - Gm)^T C_d^{-1}(d - Gm) + \varepsilon(m - m_0)^T C_m^{-1}(m - m_0) + \eta m^T D^T D m \qquad (2.34)$$

Rewriting $S(m)$ as the sum of three functions:

$$S(m) = \Psi(m) + \epsilon\Phi(m) + \eta\Omega(m)$$
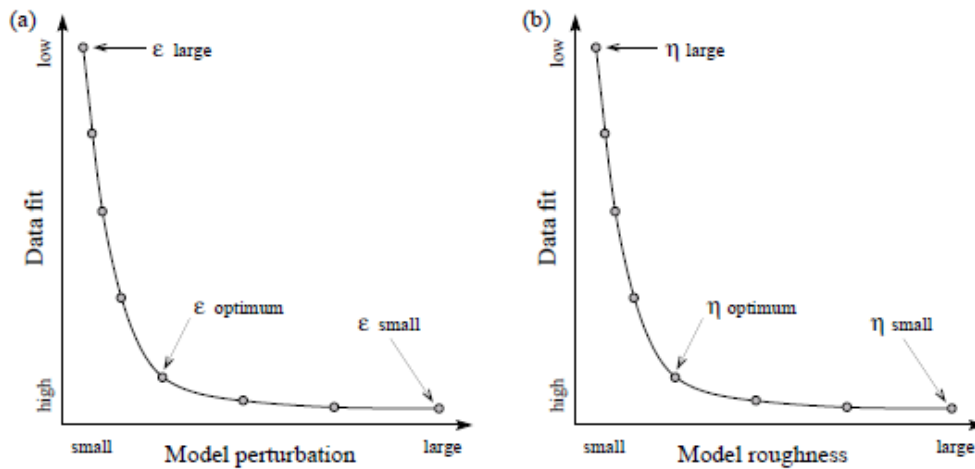
Figure 2.12: Trade-off curves to choose appropriate damping and smoothing parameters.

Here, ε and η govern the trade-off between how well the solution m_{est} will satisfy the data, how closely m_{est} is to m_{obs} and the smoothness of m_{est}. By solving equation slowness at each block can be obtained and thus the image of the region of interest can be constructed.

To minimize $S(m)$, Gradient based inversion methods use the derivatives of $S(m)$ at a specified point in model space. A basic assumption is that $S(m)$ can be expanded to allow a local quadratic approximation about a current model as:

$$S(m + \delta m) \approx S(m) + \hat{\gamma}\,\delta m + \tfrac{1}{2}\,\delta m^T \hat{H} \delta m \qquad (2.35)$$

Where δm is the perturbation to current model and $\hat{\gamma} = \frac{\partial S}{\partial m}$ and $\hat{H} = \frac{\partial^2 S}{\partial m^2}$ are the derivatives of $S(m)$ w.r.t m

Evaluating these partial derivatives using Equation 2.34 gives:

$$\hat{\gamma} = G^T C_d^{-1}[g(m) - d_{obs}] + \varepsilon C_m^{-1}(m - m_0) + \eta D^T D m \qquad (2.36)$$

$$\hat{H} = G^T C_d^{-1} G + \nabla_m G^T C_d^{-1}[g(m) - d_{obs}] + \varepsilon C_m^{-1} + \eta D^T D \qquad (2.37)$$

Where $G = \frac{\partial g}{\partial m}$ is the Frechet matrix of partial derivatives.

The minimization of the objective function $S(m)$ in equation is to be performed using an iterative approach to overcome the non-linearity of $g(m)$ by updating m in each iteration:

$$m_{n+1} = m_n + \delta m \qquad (2.38)$$

where m_0 is the initial model. The objective function is minimized for the current ray path estimate at each step to produce m_{n+1}, after which new ray paths are computed for the next iteration. We stop the iterations either when the observed traveltimes are satisfied or when the change in $S(m)$ with each iteration gets sufficiently small.

There are many methods like Gauss-Newton method and damped least squares which locates the updated point m_{n+1} by finding the minimum of the tangent paraboloid to $S(m)$ at m_n. This requires an $M \times M$ Matrix equation to be solved and is computationally expensive, if number of model parameters is large. Other methods are steepest descent which is based on the idea that $S(m)$ may be minimized by successive searches along local directions of steepest descent and conjugate gradient method which uses a new search direction at each iteration that is conjugate to all previous directions. Here large number of equations need not be solved compared to Newton's method but, $S(m)$ is minimized in only one dimension at each iteration and is prone to converge slowly.

In this study, δm is found iteratively using the subspace method. Both steepest descent and conjugate gradient method are 1D subspace methods in that they perform line minimization in each iteration. However, Subspace methods can be constructed in which the minimization is carried out simultaneously along several search directions that together span a subspace of the model parameter space. At each iteration, the subspace method restricts the minimization of the quadratic approximation of $S(m)$ to a *p*-dimensional subspace of model space, so that the perturbation δm occurs in the space spanned by a set of p *M*-dimensional basis vectors $\{a^j\}$:

$$\delta m = \sum_{j=1}^{p} \mu_j a^j = A\mu \qquad (2.39)$$

Where $A = [a^j]$ is the $M \times p$ projection matrix. The component μ_j determines the length of the corresponding vector a^j that minimizes the quadratic form of $S(m)$ in the space spanned by a^j. Using this expression for δm in Equation 2.35 for $S(m)$, we get:

$$S(m + \delta m) = S(m) + \sum_{j=1}^{p} \mu_j \, \hat{\gamma}^T a^j + \frac{1}{2} \sum_{j=1}^{p} \sum_{k=1}^{p} \mu_j \mu_k [a^k]^T \hat{H} [a^j] \qquad (2.40)$$

Now locating the minimum of $S(m)$ with respect to μ we get

$$\frac{\partial S(m)}{\partial \mu_q} = \hat{\gamma}^T a^q + \sum_{k=1}^{p} \mu_k [a^k]^T \hat{H} [a^j] = 0 \qquad (2.41)$$

For $q = 1,2 \dots p$. Solving for μ gives:

$$\mu = -[A^T \hat{H} A]^{-1} A^T \hat{\gamma} \qquad (2.42)$$

We need δm which is given by $\delta m = A\mu = -A[A^T \hat{H} A]^{-1} A^T \hat{\gamma}$ and from Equation 2.37,

$$\delta m = -A[A^T (G^T C_d^{-1} G + \epsilon C_m^{-1} + \eta D^T D)A]^{-1} A^T \hat{\gamma} \qquad (2.43)$$

Which can be used iteratively as mentioned above. The quantities, $A, \hat{\gamma}$ and G in Equation 2.43 are re-evaluated between successive iterations. The detailed discussion

of the subspace method is given in Kennett *et al.* (1998) and Rawlinson and Sambridge (2003).
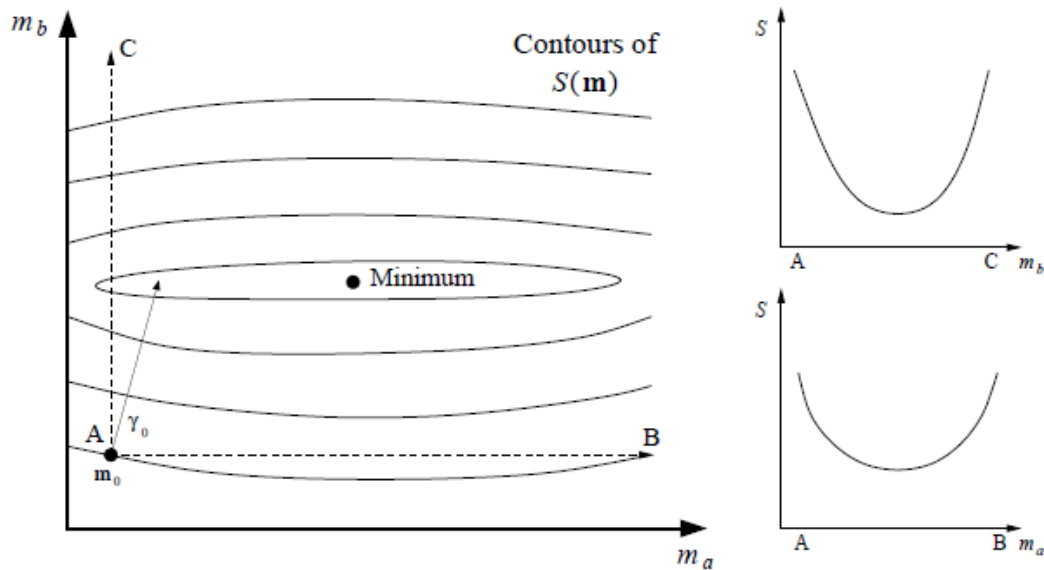
Figure 2.13: A contour plot of $S(m)$ which is a function of two parameters of different physical dimensions. Here $S(m)$ is more sensitive to m_b than m_a. Searching in directions specified by basis vectors (dotted lines) that only lie in space defined by a single parameter class is shown adjacent. $S(m)$ is a function of only one parameter class in each of these directions.

In the case of Figure 2.13, a gradient method like steepest descent will converge slowly if there are many model parameters. But searching in the directions of basis vectors as done in subspace inversion in the Figure 2.13, quickly gets to the local minimum in the direction of each basis vector.

2.5.5 Analysis of solution quality

The inverse problem demands that along with updating model parameters using any of the procedures, an analysis of that solution robustness and the procedure must be made. Only calculating a solution that minimizes the objective function (i.e. which best satisfies data and a priori information) without the knowledge of resolution or non-uniqueness is inadequate. Here, we perform a checkerboard test by reconstructing a synthetic model using the same source-receiver geometry as in the real data.

Chapter 3

Methodology

3.1 Data Analysis

The seismic data used in this study comprise of both Ambient Noise and earthquake data. Noise data is a continuous recording of earth vibration for more than a year whereas earthquake data is an event data in which only the duration of earthquake is picked for analysis. As a consequence, noise data is very large compared to the earthquake data. They have different schemes of data preparation and analysis until the calculation of Dispersion curves. Green's function is extracted from the cross-correlation of noise data which we take as a virtual seismogram. For earthquake data, we directly calculate the dispersion curves.
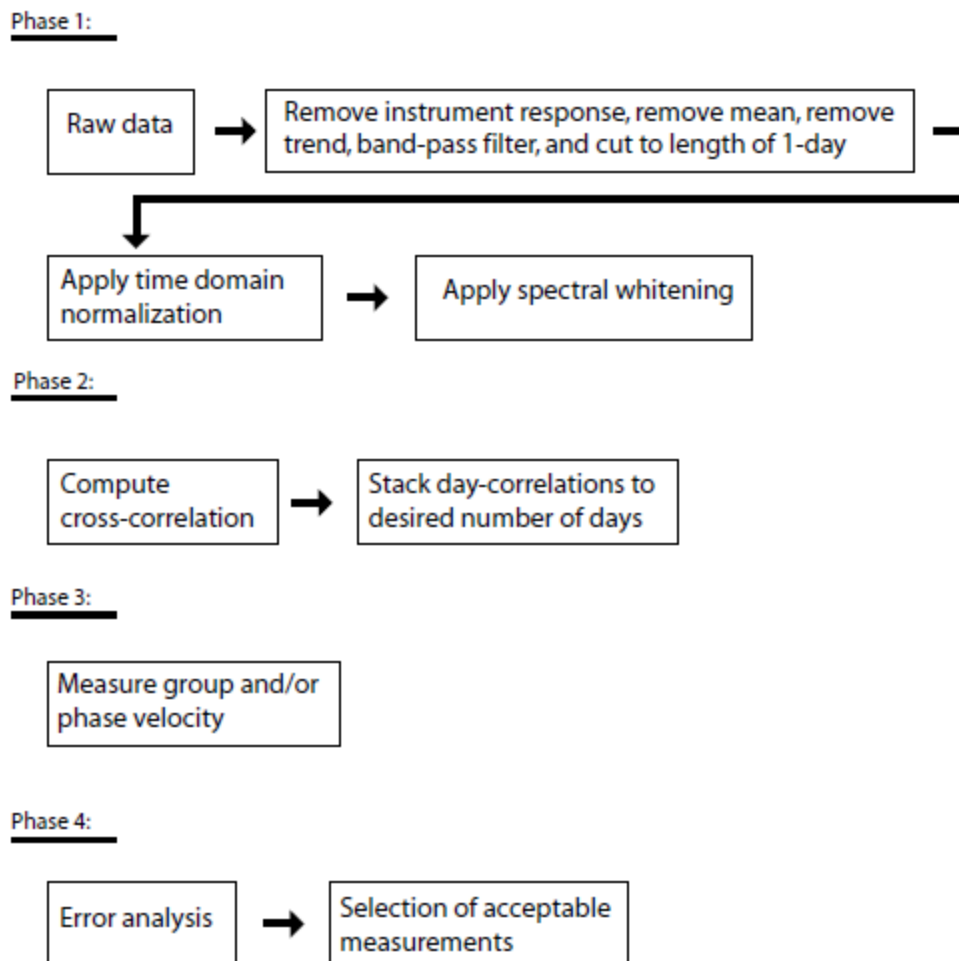
In the whole process, we extensively used Linux Shell scripting to organize, manipulate and edit data. It is very useful for the co-ordination of various programs that are necessary for the analysis by integrating them in to shell scripts. To analyze seismic data we used Seismic Analysis Code (SAC). It is an all-in-one tool for manipulating, editing, formatting seismic data. All geographical figures were generated using Generic Mapping Tools version 4.5.0. Different scripts and programs were used to do cross-correlation, stacking, removing impulse response, measuring dispersion and tomography, all written by various authors as will be discussed in the next section.

3.2 Ambient noise data analysis

Seismograms recorded at 15 Broadband sites have been analyzed for the study. The data used is approximately two years data (from 2002 to 2003) and consisted of 15 broadband seismometers managed by the CSIR-National Geophysical Research Institute (NGRI) along a 500km profile. Since all the stations are not operated at the same time, it is not possible to have a complete combination of all station pairs. The ambient noise processing procedures applied here are basically the same as the one described by Bensen et al. (2007). We use only the vertical component of ambient

noise data which means that the cross-correlations provide only the Green's functions which are dominated by Rayleigh wave signals.

Four stages of data processing procedure were applied to the data as shown in the Figure 3.1. First stage shows the steps involved in preparing single-station data before the cross-correlation. Second stage outlines the cross-correlation procedure and stacking. Third stage includes dispersion measurement and finally error analysis and data selection process.

Phase 1:

| Raw data | → | Remove instrument response, remove mean, remove trend, band-pass filter, and cut to length of 1-day |

| Apply time domain normalization | → | Apply spectral whitening |

Phase 2:

| Compute cross-correlation | → | Stack day-correlations to desired number of days |

Phase 3:

| Measure group and/or phase velocity |

Phase 4:

| Error analysis | → | Selection of acceptable measurements |

Figure 3.1: Schematic representation of Data processing scheme, (Bensen *et al*., 2007)

3.2.1. Single station data preparation

The first phase of data processing consists of preparing waveform data from each station individually. The raw data recorded at the station is in .ref format. We convert this to .sac format for analysis using a program called ref2sac. 24 hour long data segments are created for each station from the continuous data recordings of broadband seismic stations to amplify broad-band ambient noise by removing earthquake signals and instrumental irregularities that tend to obscure ambient noise. For each station the data has been resampled to 20 samples per second and split into one day segments, followed by the removal of mean, trend and instrument response. If we provide the data of poles and zeroes of each seismmometer to SAC, it will remove the instrument response from each one day data file. The resulting waveform is then tapered, bandpass filtered between periods 1-60s followed by time domain normalization and spectral whitening.

In time domain normalization, an event detection and signal removal technique is used where the data was first divided into one hour blocks and the blocks containing signals above a threshold are discarded. Setting the threshold very high provides no normalization and setting it very low removes most of the data. Hence, the threshold usually is chosen to be around ten times the absolute mean. Spectral whitening is used after normalization to reduce the natural imbalances in ambient noise spectrum and to broaden the ambient noise signal in cross-correlation.

3.2.2. Cross-correlation and stacking

After preparation of the daily time-series, the next step in the data processing scheme is cross-correlation and stacking. The cross-correlation between each of the resulting day segment was computed which are two sided functions, where causal and acausal components represent energy travelling in opposite directions between the pair of stations. The cross-correlation functions are twice as long as the input files, i.e. two days, and it was sufficient to consider only the sections around 400 seconds of either side of the origin depending upon the inter station distance. The

cross-correlation functions (CCFs) were calculated for the daily waveforms of each pair of station and stacked day by day to increase the signal to noise ratio (SNR).

3.2.3 Dispersion measurements and data selection

Once the daily cross-correlations are computed and stacked, the resulting waveform is an estimated Green function. Using this Green's function, the group velocities as a function of period are be calculated by using traditional frequency-time analysis (Levshin and Ritzwoller, 2011). Dziewonski et al (1969) developed the multiple filter technique (do_MFT) using which we determine the group velocities of the dispersive wave. We manually selected the dispersion curves by considering the maximum amplitude at each period and checking the quality of the curve.

3.3 Earthquake data analysis

Here, we select the event data i.e., 300 seconds before the origin time of the earthquake and 1200 seconds after the origin time. A total of 42 earthquakes that occurred in the year 2002 and 44 earthquakes in the year 2003 were considered for this study. The data processing sequence remains the same for earthquake data as for the noise except that we need not perform cross-correlations here. We remove the mean, trend and instrument response for each event and calculate group speed dispersion curves directly using the Multiple Filter Analysis (do_mft) technique.

3.4 Performing Tomography

After computing the group velocity dispersion measurement a tomographic inversion is performed for periods between 8 and 50 s using the non-linear 2-D tomographic inversion technique developed by Rawlinson and Sambridge (2003). Once we have obtained dispersion curves from the empirical Green's functions, we can calculate travel times of Rayleigh waves at a given frequency. A 2D tomographic inversion can be performed on the travel times to estimate the variations in group velocity over the region under consideration at the particular frequency. We can obtain

different 2D models for each frequency and then use the relationship between Rayleigh wave frequency and sampling depth to infer the velocity variations with depth.

The method combines the Fast Marching Method (FMM) (Rawlinson and Sambridge, 2004a, b) for calculation of forward problem. To perform the inverse step, a gradient method based on subspace technique (Kennett et al., 1988) is applied, where the minimization is carried out simultaneously along several search directions that together span a subspace of the model space. In this study we use the Fast Marching Surface Tomography (FMST) code developed by N. Rawlinson.

To verify the resolving capabilities of the station array, checkerboard tests (Iyer and Hirahara, 1993) were applied. These tests involve the calculation of synthetic travel times for available station pairs through an artificial velocity model of alternating positive and negative anomalies. These times are then inverted by using a uniform starting model. When the original anomalies can be recovered, it implies that a real anomaly of similar size, location and amplitude can be resolved.

Chapter 4

Results and Discussion

The data used for this study is collected using 15 broadband seismometers in the NW Himalaya region. A list of stations with their locations is given in the Appendix A.
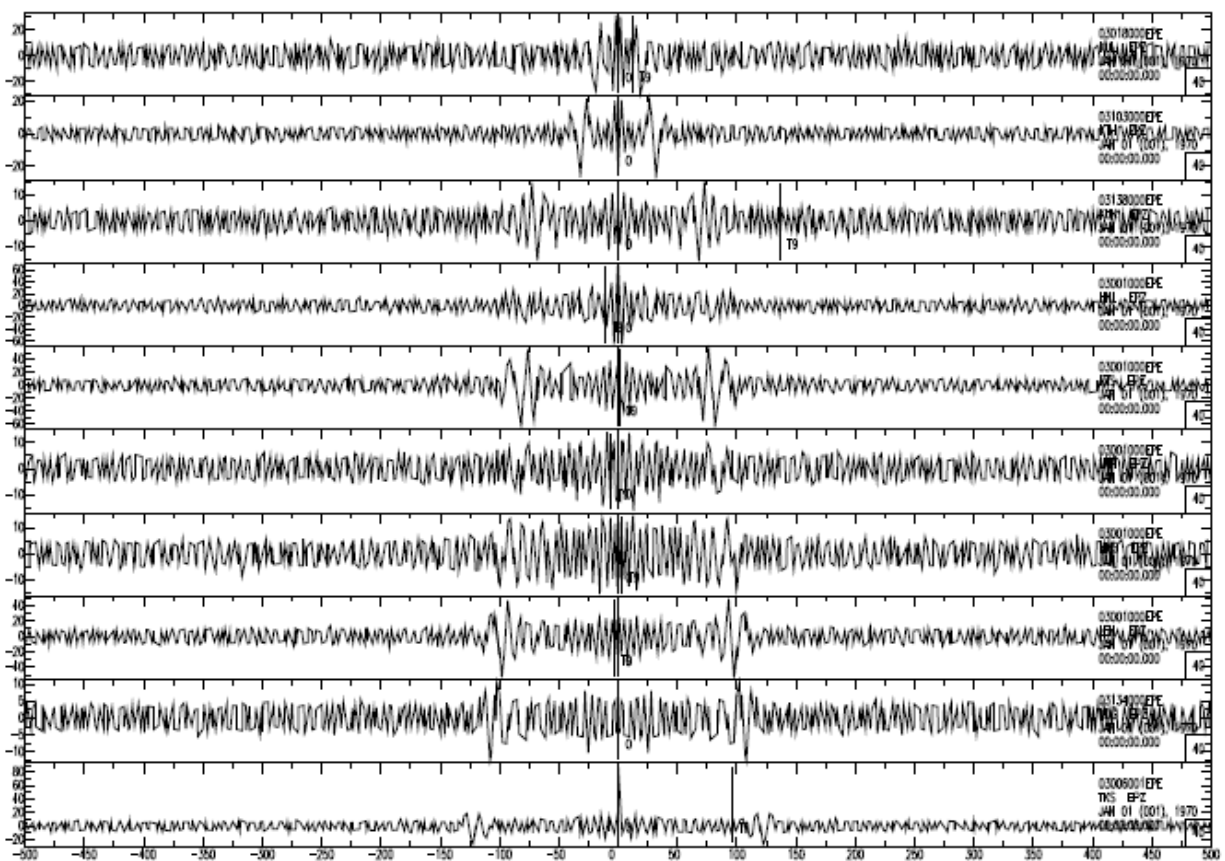
Noise Data:
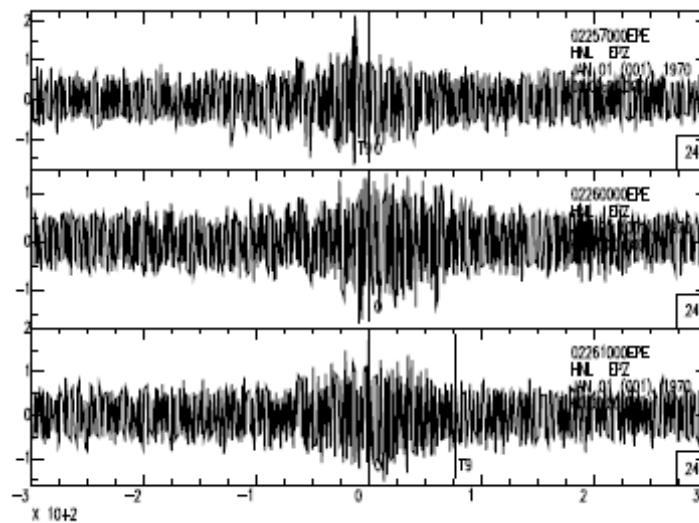
Figure 4.1: Cross-correlation of the stations HMS and HNL for Julian days 257, 260 and 261 of the year 2002

The daily cross-correlation functions were stacked for each station pair to increase the Signal to Noise ratio.
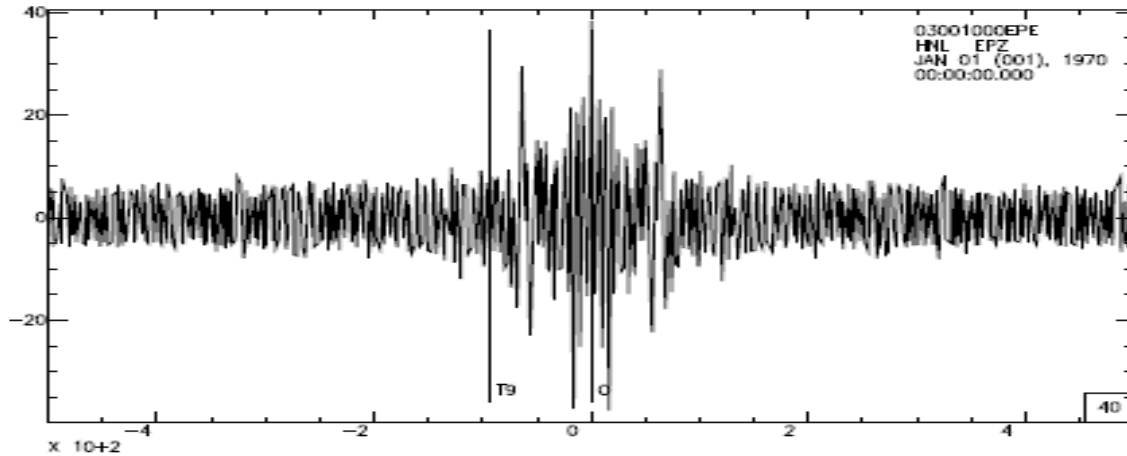
Figure 4.2: One year stacking of Cross correlations of the stations HMS and HNL of the year 2003

Figure 4.3: Emergence of Rayleigh waveforms in the stacks of daily cross-correlations of the signal recorded by station GHR with all other stations, plotted against distance between stations.
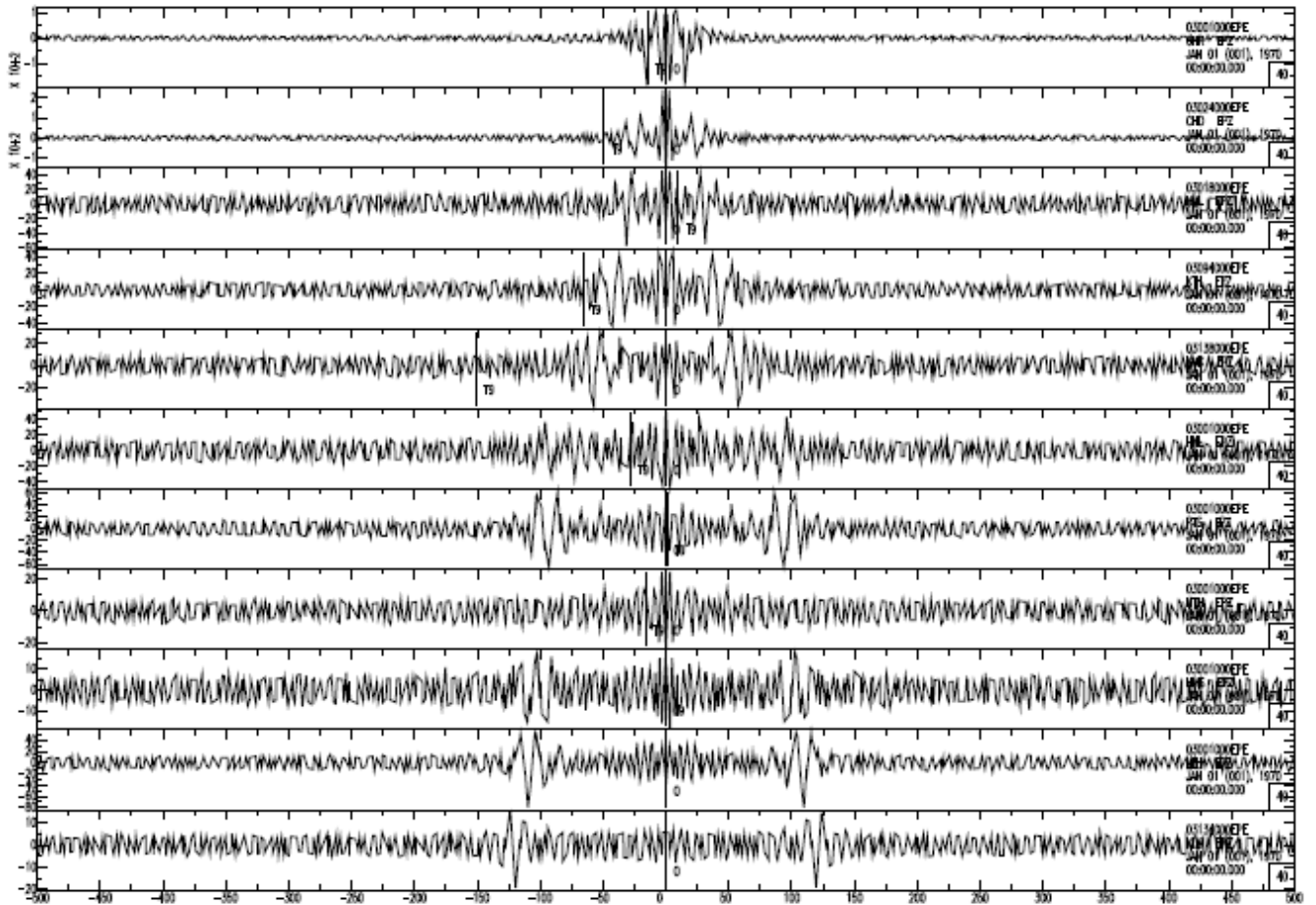
Figure 4.4: Emergence of Rayleigh waveforms in the stacks of daily cross-correlations of the signal recorded by station BSP with all other stations, plotted against the distance which increases downwards between the stations.

In the above figures 4.3 and 4.4 we can see that as the distance increases between the stations, the arrival of the waveform is clearly delayed because it takes more time for the energy to travel the increasing inter station distance

After stacking the daily records, the frequency dependent group velocities were calculated using Multiple Filter Technique (do_MFT program) and group velocity vs period were plotted as shown. The points with maximum amplitude at each period were chosen manually to get the dispersion curve.
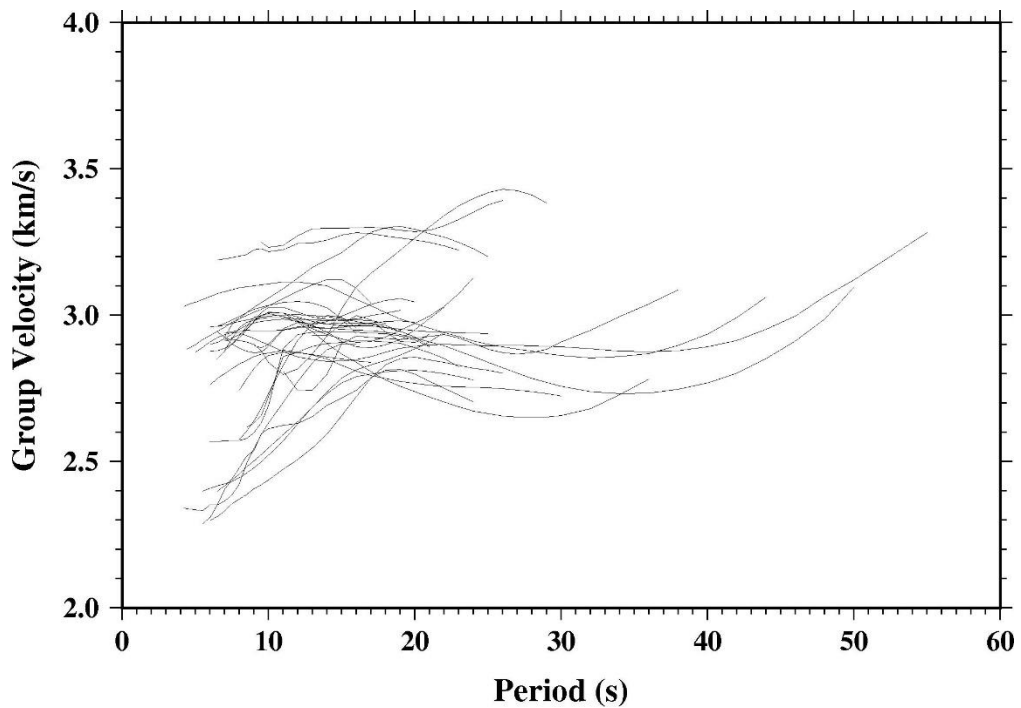
Dispersion curves for Noise Data:
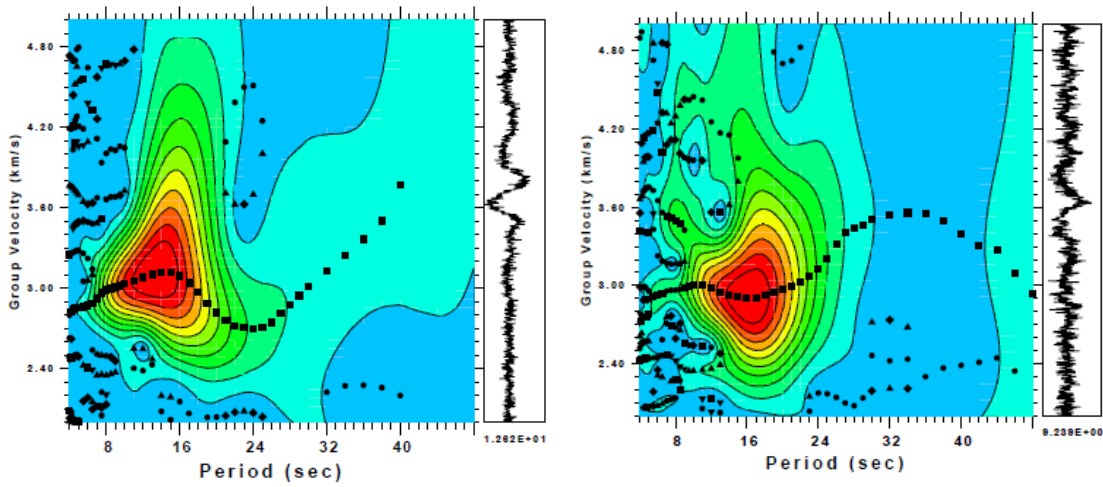
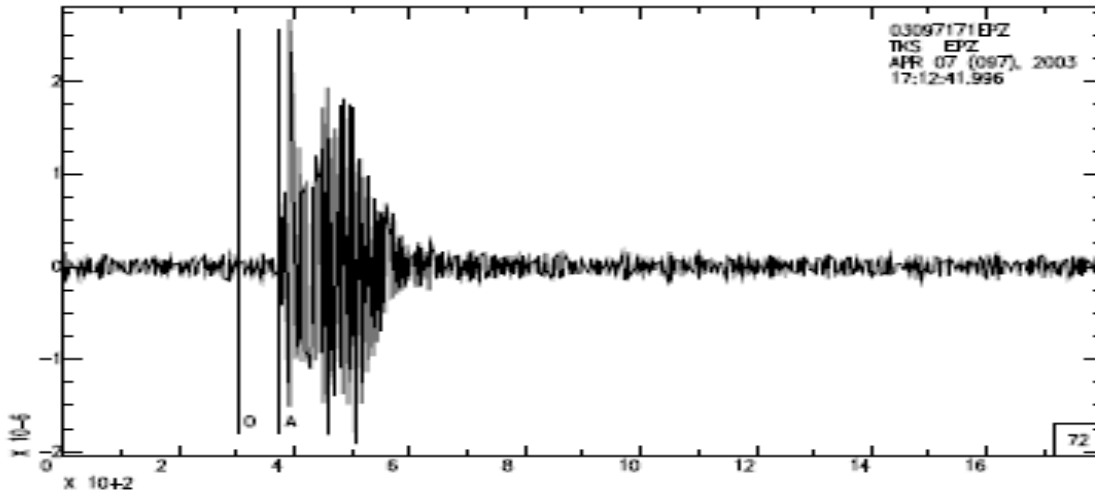**Figure 4.5:** Dispersion curves between stations KDG and CHD; HMS and GHR
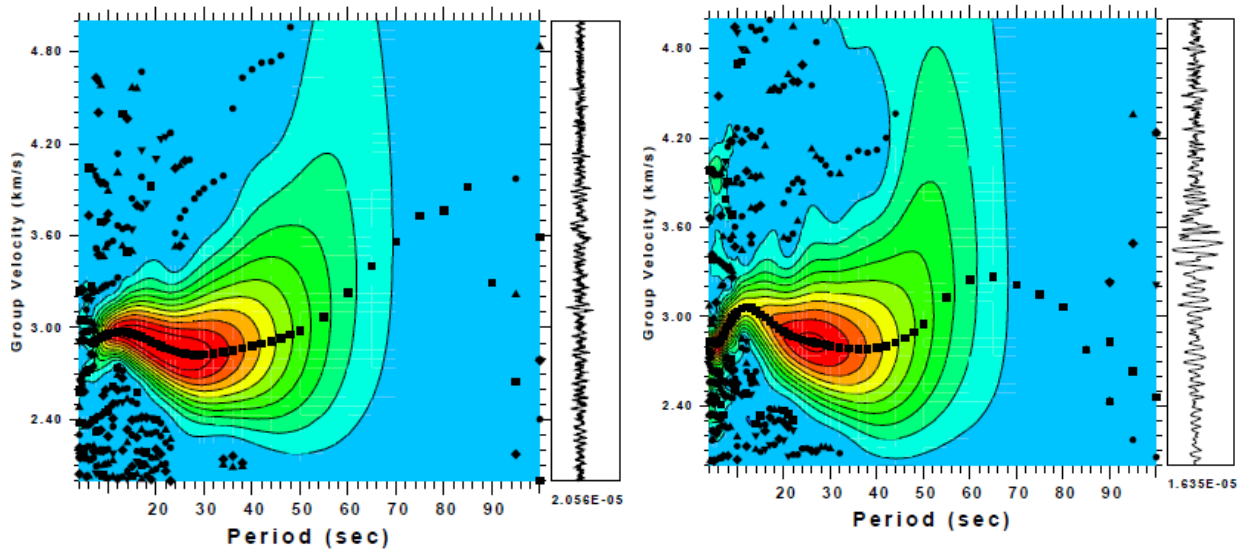
All dispersion curves for Noise Data:

Figure 4.6: All the dispersion curves between the stations were plotted together showing the similarity due to their N-S orientation

Earthquake data

The earthquake data is picked during the period and one of the event is given in the figure:
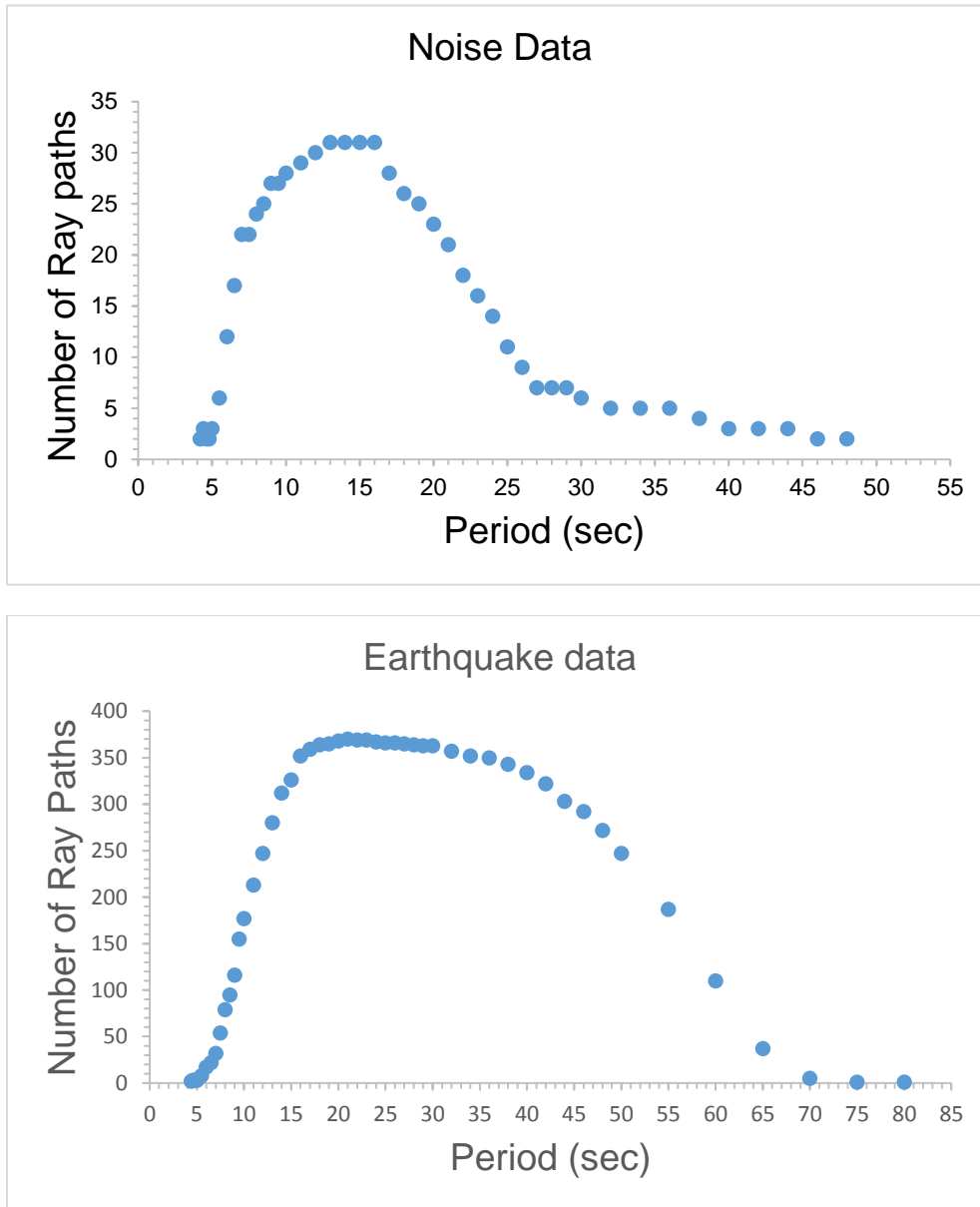

Figure 4.7: An earthquake recorded at the station TKS on April 7th, 2003.

Figure 4.8: Dispersion curves for station the KUK during the event on Julian day 149 of 2003; Dispersion curves for station the BDI during the event on Julian day 149 of 2003

Next we plot the total number of ray paths vs period for both Noise data and earthquake data:

Figure 4.9: Plots representing the total number of selected ray paths as a function of period for Noise and earthquake data.

These plots show that earthquake data has relatively less measurements for lower periods like 5-10 seconds compared to the noise data. Whereas earthquake data has measurements till 40-50 second, noise data has data only till 20 seconds.

Therefore, we can expect to derive greater resolution in shallow depths using Noise data combined with the earthquake data. The following figures show the ray paths for different periods obtained for Noise and Earthquake data:
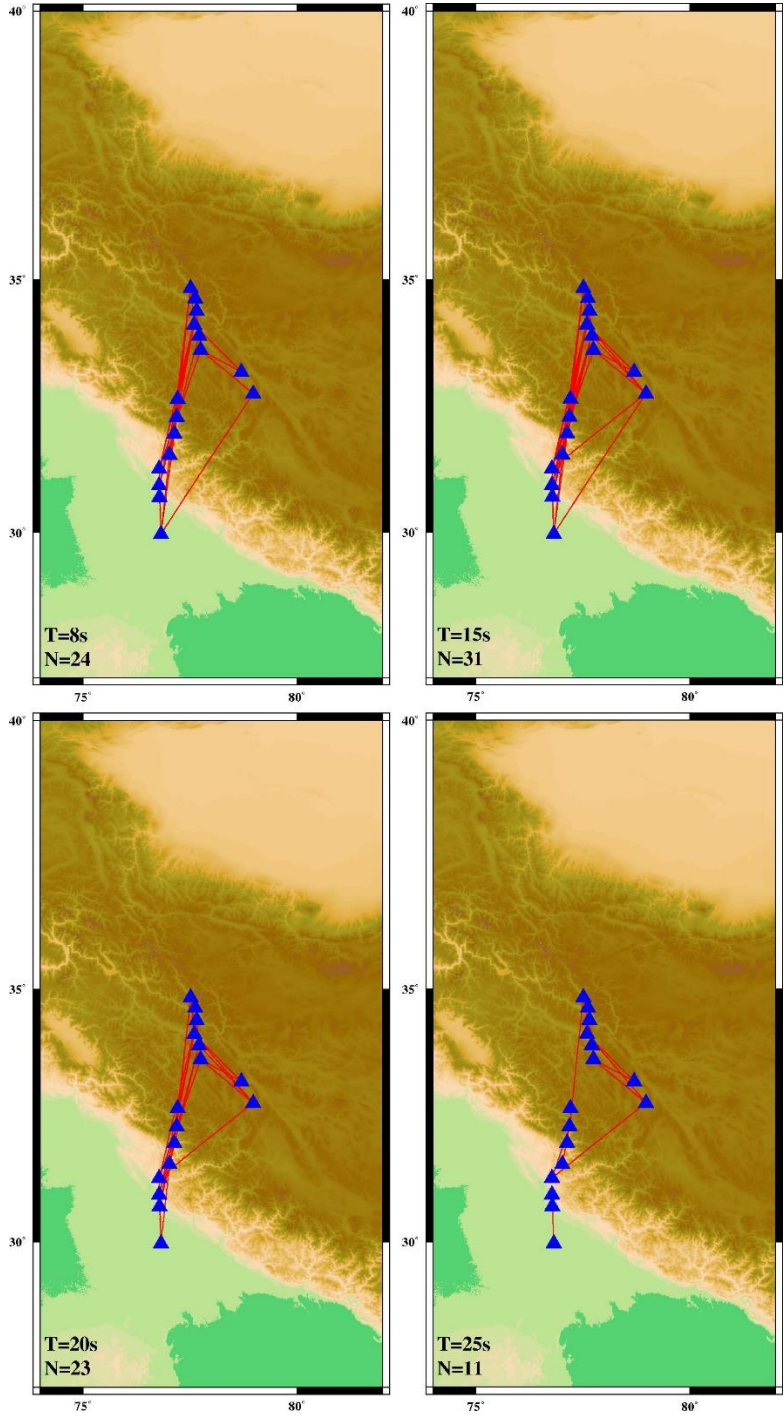
Figure 4.10: Diagram showing the selected ray paths for Noise data at different periods. The period and no of ray paths are given at bottom left corner.
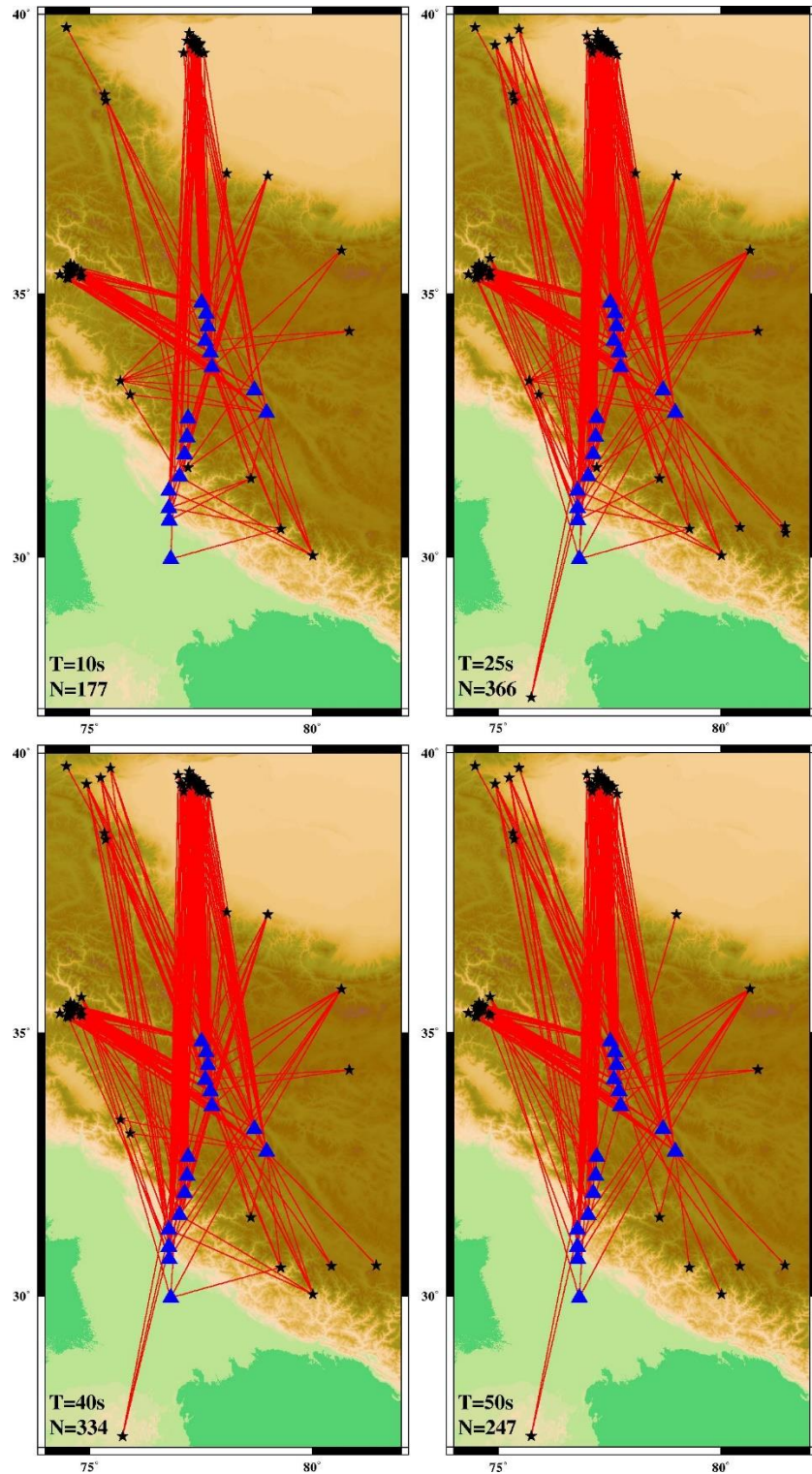
**Figure 4.11:** Diagram showing the selected ray paths for Earthquake data at different periods. The period and number of ray paths are given at bottom left corner. The events are marked with a star and the stations are marked with a triangle

Having obtained traveltimes of Rayleigh waves at a given frequency, a 2D tomographic inversion is performed to get the variations in group velocity at a given frequency.

To investigate the resolving capability of the station array, checkerboard tests were performed to test how well the geometry of stations and virtual sources might resolve the subsurface structure. The cell size 1°×1° provides us the maximum resolution. Figure 4.12 shows the resolution test for selected time periods.
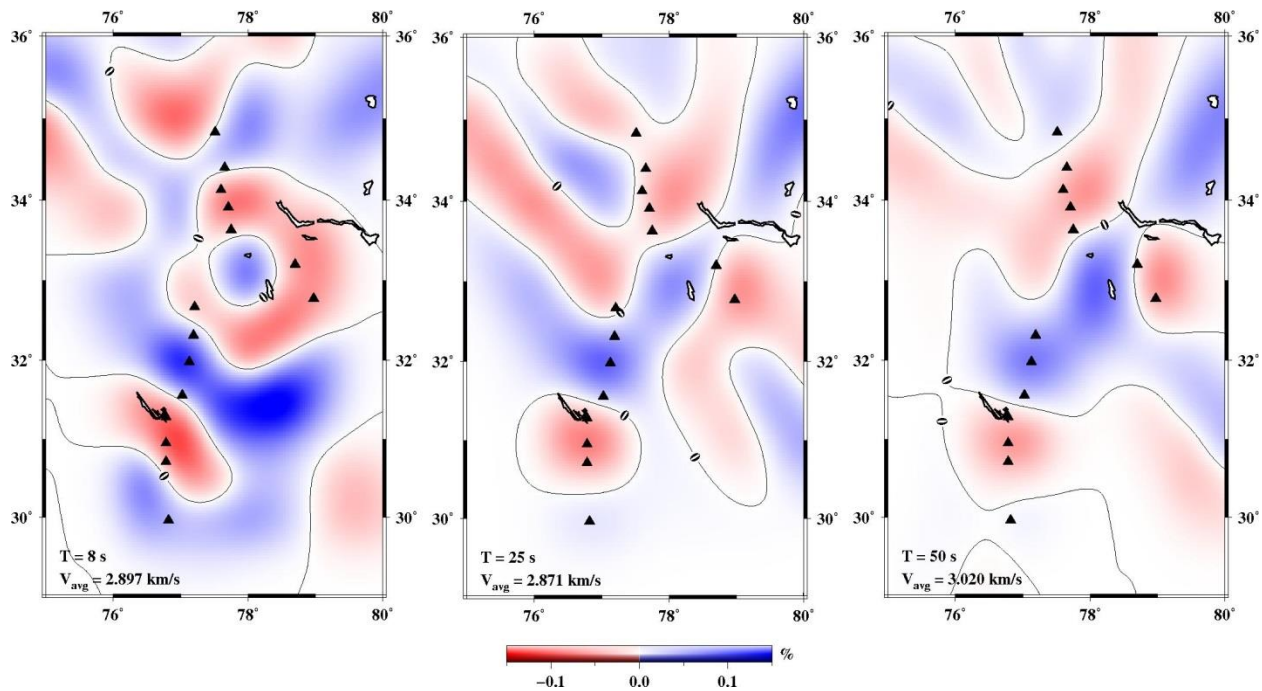
Fig 4.12: Results of the checkerboard tests used to estimate resolution.

From the checkerboard tests in the figure 4.12, we find that the region is well resolved in the area around the stations compared to the edges of the plot where we notice blur and poor resolution. Hence we can perform tomography to obtain sufficiently well resolved maps for the selected periods.

54

The Rayleigh wave group velocity models calculated for different periods are shown below:

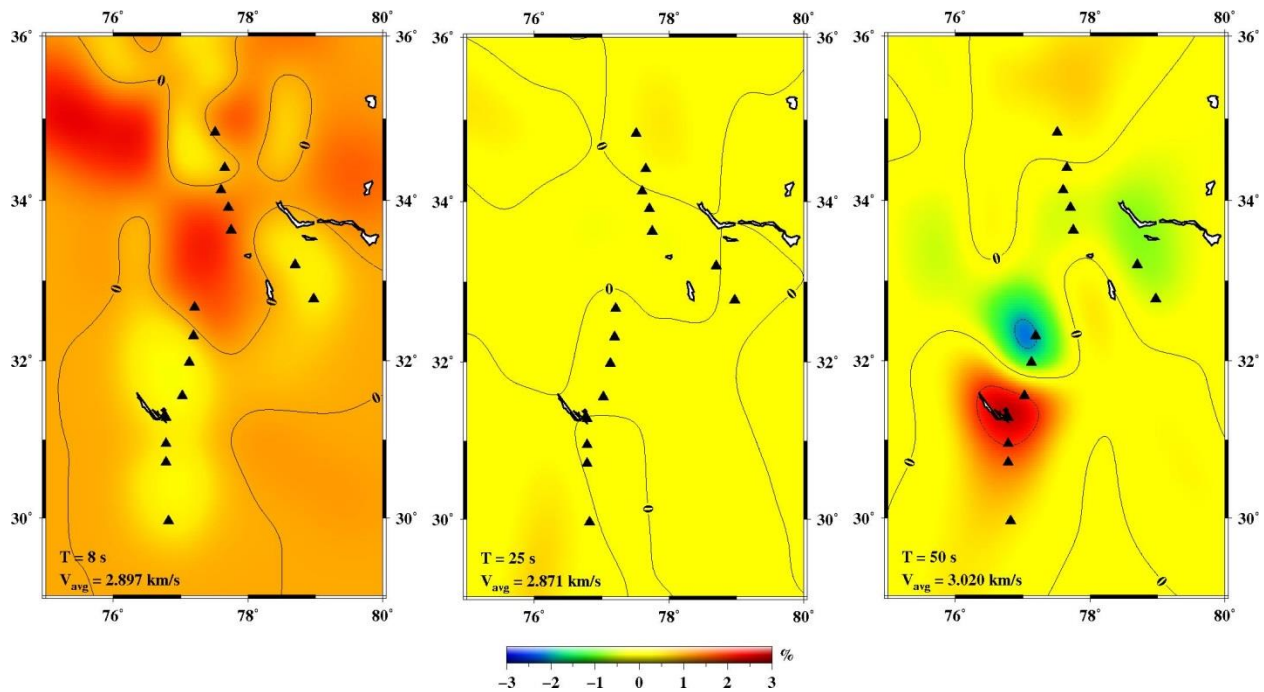Figure 4.13: Surface wave Group velocity perturbation maps computed for different periods, and thus corresponding to different depths. Velocity perturbations are relative to the average group velocity plotted at the bottom left corner of each plot.

Here we get the tomographic map of the region at different periods as show in the Figure 4.13. We notice velocity anomalies in 8 sec period and 50 sec period as shown.

Discussion and Conclusion

We have obtained the Rayleigh wave group velocity map of NW-Himalaya using both ambient noise and earthquake data. There have been quite a few studies done on the study region. Rai et al (2006), using the same stations data jointly inverted 15-60 sec Rayleigh wave group velocities with Receiver functions from teleseismic arrivals. The study mainly focused on the Moho depth variation in the area. Rai et al (2009), also using the same stations, modeled seismic attenuation using L_g waves and found high attenuation in Ladakh and low attenuation in Tethyan Himalaya and Himalayan thrust belt. Oreshin et al (2008) also using the same dataset and modeled V_p and V_s for the crust and mantle using teleseismic body waves. They found a mid-crustal low velocity in the Indus Tsangpo suture zone (ITSZ). Also they found a lower crustal low velocity at Ladakh and correlated the result with channel flow. Caldwell et al (2009) using the same dataset found a intra-crustal low velocity starting at depth of 10-15 km and continuing up to 28-35 km depth. They interpret it as the presence of 3-7% partial melt at present day beneath and north of South Tibetan Detachment in the mid crust of Himalaya and Tibetan plateau of NW India.

Our results shows a continuous low Rayleigh wave group velocity after 30 s period in the north of MCT (Main Central Thrust). This may be accredited to the presence of partial melts or aqueous fluids. Further study will be conducted to convert the group velocities to shear velocities (using 1D Shear wave velocity inversion) for better explanation of the structure and geodynamics of the region.

Appendix A: List of stations

Station Name	Latitude	Longitude	Elevation(m)
TKS	34.835	77.511	3210
TGR	34.639	77.616	3148
KDG	34.399	77.654	3967
LEH	34.128	77.6	3314
HMS	33.913	77.708	3626
RTS	33.633	77.749	4128
MTH	33.205	78.7	4195
HNL	32.778	78.973	4284
KTH	32.317	77.192	2513
KUL	31.983	77.13	1251
GHR	31.561	77.023	1293
BSP	31.284	76.783	690
BDI	30.952	76.782	367
CHD	30.714	76.785	289
KUK	29.961	76.82	214

References

1. Bensen, G. D., M. H. Ritzwoller, M. P. Barmin, A. L. Levshin, F. Lin, M. P. Moschetti, N. M. Shapiro, and Y. Yang (2007), Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, Geophys. J. Int., *169*(3), 1239–1260.

2. Caldwell, W. B., S. L. Klemperer, S. S. Rai, and J. F. Lawrence (2009). Partial melt in the upper-middle crust of the northwest Himalaya revealed by Rayleigh wave dispersion. Tectonophysics, *477*, 58-65.

3. Campillo, M., Paul, A. (2003). Long-range correlations in the diffuse seismic coda. Science *299*, 547–549.

4. Claerbout, J. F. (1968). Synthesis of a layered medium from its acoustic transmission response, Geophysics *33*, 264.

5. Duvall, Jr., T.L.; Jefferies, S.M.; Harvey, J.W., Pomerantz, M.A. (1993). Time-distance helioseismology. Nature *362,* 430–432.

6. Dziewonski, A., S. Bloch, and M. Landisman (1969). A technique for the analysis of transient seismic signals, Bull. Seismol. Soc. Am., *59*(1), 427–444.

7. Gouedard, P., L. Stehly, F. Brenguier, M. Campillo, Y. Colin de Verdiere, E. Larose, L. Margerin, P.Roux, F. Sanchez-Sesma, N. Shapiro, and R.Weaver. (2008). Cross-correlation of random fields: mathematical approach and applications, Geophysical prospecting, *56*(3), 375-393.

8. Goscombe, B., Gray, D., Hand, M., (2006). Crustal architecture of the Himalayan metamorphic front in eastern Nepal. Gondwana Research *10* (3–4), 232–255.

9. Iyer, H.M. and K. Hirahara (1993). Seismic Tomography Theory and Practice, Chapman & Hall, New York.

10. Kennett, B. L. N., M. S. Sambridge, and P. R. Williamson (1988), Subspace methods for large inverse problems with multiple parameter classes, Geophys. J., *94*(2), 237–247.

11. Levshin, A. L., and M. H. Ritzwoller (2001). Automated Detection, Extraction, and Measurement of Regional Surface Waves, Pure Appl. Geophys., *158*(8), 1531–1545.

12. Lin, F.-C., M. H. Ritzwoller, J. Townend, S. Bannister, and M. K. Savage (2007). Ambient noise Rayleigh wave tomography of New Zealand, Geophys. J. Int., *170*(2), 649–666.

13. Oreshin, S., Kiselev, S., Vinnik, L., Surya Prakasam, K., Rai, S.S., Makeyeva, L., Savvin, Y., (2008). Crust and mantle beneath western Himalaya, Ladakh and western Tibet from integrated seismic data. Earth and Planetary Science Letters 271 (1–4), 75–87.

14. Rai, S. S., K. Priestley, V. K. Gaur, S. Mitra, M. P. Singh, and M. Searle (2006). Configuration of the Indian Moho beneath the NW Himalaya and Ladakh, Geophys. Res. Lett., 33, L15308.

15. Rai, S.S., Ashish, A., Padhi, A., Sarma, P.R., (2009). High crustal seismic attenuation in Ladakh-Karakoram. Bulletin of the Seismological Society of America 99 (1).

16. Rawlinson, N., and M. Sambridge (2003). Seismic traveltime tomography of the crust and lithosphere, Advances in Geophysics, Elsevier pp. 81–198.

17. Rawlinson, N., and M. Sambridge (2004a). Multiple reflection and transmission phases in complex layered media using a multistage fast marching method, Geophysics, *69*(5), 1338.

18. Rawlinson, N., and M. Sambridge (2004b). Wave front evolution in strongly heterogeneous layered media using the fast marching method, Geophys. J. Int., *156*(3), 631–647.

19. Ren, Y., B. Grecu, G. Stuart, G. Houseman, and E. Hegedus (2013). Crustal structure of the Carpathian-Pannonian region from ambient noise tomography, Geophys. J. Int., 195(2), 1351–1369.

20. Ritzwoller, M.H., (2008). Ambient noise seismic imaging, McGraw Hill 2008 Yearbook of Science & Technology, 1-13.

21. Sabra, K. G., P. Gerstoft, P. Roux, W. A. Kuperman, and M. C. Fehler (2005a). Extracting time-domain Greens function estimates from ambient seismic noise, Geophys. Res. Lett., *32*, L03310.

22. Sabra, K. G., P. Gerstoft, P. Roux, W. A. Kuperman, and M. C. Fehler (2005b). Surface wave tomography from microseism in Southern California, Geophys. Res. Lett., *32*, L14311.

23. Sambridge, M. S. (1990). Non-linear arrival time inversion: Constraining velocity anomalies by seeking smooth models in 3-D, Geophys. J. Int., *102*, 653–677.

24. Saygin, E., and B. L. N. Kennett (2012). Crustal structure of Australia from ambient seismic noise tomography, J. Geophys. Res. Solid Earth, *117*(1).

25. Sethian, J. A., and A. M. Popovici (1999). 3-D traveltime computation using the fast marching method, Geophysics, *64*(2), 516.

26. Shapiro, N., Campillo, M., (2004). Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise. Geophysical Research Letters *31* (7), 1615–1619.

27. Shapiro, N. M., M. Campillo, L. Stehly, and M. H. Ritzwoller (2005). High-resolution surface-wave tomography from ambient seismic noise. Science, *307*(5715), 1615–1618.

28. Snieder, R. (2004). Extracting the Green's function from the correlation of coda waves: a derivation based on stationary phase. Phys. Rev. E, *69*(4 Pt 2), 046610.

29. Thomson, C.J & Gubbins D. (1982) Three Dimensional lithospheric modelling at NORSAR: linearity of the method and amplitude variations from the anomalies. Geophysics Journal, *71*, 1-36.

30. Wapenaar, K. & Fokkema J. (2005) Seismic interferometry, time-reversal and reciprocity, 67th EAGE Conference & Exhibition.

31. Wapenaar, K. (2003). Synthesis of an inhomogeneous medium from its acoustic Transmission response. Geophysics *68* (5), 1756–1759.

32. Wapenaar, K. (2004), Retrieving the Elastodynamic Green's Function of an Arbitrary Inhomogeneous Medium by Cross Correlation, Phys. Rev. Lett., *93*(25), 254301.

33. Weinberg, R.F., Dunlap, W.J., 2000. Growth and deformation of the Ladakh Batholith, Northwest Himalayas: implications for timing of continental collision and origin of calc-alkaline batholiths. The Journal of Geology *108* (3), 303–320.

34. Weaver, R. L., and O. I. Lobkis (2001), Ultrasonics without a source: Thermal fluctuation correlations at MHz frequencies, Phys. Rev. Lett., *87*, 134301.

35. Weaver, R. L., and O. I. Lobkis (2005), Fluctuations in diffuse field-field correlations and the emergence of the Green's function in open systems. J. Acoust. Soc. Am., *117*(6), 3432–3439.

36. Yang, Y., M. H. Ritzwoller, A. L. Levshin, and N. M. Shapiro (2007). Ambient noise Rayleigh wave tomography across Europe, Geophysics J. Int., *168*(1), 259–274.

37. Yao, H., R. D. Van Der Hilst, and M. V de Hoop (2006). Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis - I. Phase velocity maps, Geophysics Journal-Oxford, *166*(2), 732–744.

38. Young, M. K., N. Rawlinson, P. Arroucau, A. M. Reading, and H. Tkalčič (2011). High-frequency ambient noise tomography of southeast Australia: New constraints on Tasmania's tectonic past, Geophysics Res. Lett., *38*(13).