Crustal Structure Of The Western Himalaya

A Thesis

Submitted to

Indian Institute of Science Education and Research Pune in partial fulfillment of the requirements for the BS-MS Dual Degree Programme

By

Mohit Bokariya
20171128

Supervisor: Prof. Shyam S. Rai

Indian Institute of Science Education and Research Pune

Dr. Homi Bhabha Road,
Pashan, Pune 411008, INDIA.
April 2023

Certificate

This is to certify that this dissertation entitled “ Crustal Structure of the Western Himalaya ” towards the partial fulfillment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by “ Mohit Bokariya ” at “ IISER Pune ” under the supervision of “ Prof. Shyam S. Rai ”, Emeritus, Department of Earth and Climate Science” during the academic year 2022-2023

Prof. Shyam S. Rai

Declaration

I hereby declare that the research work presented in the report entitled “ Crustal Structure of the western Himalaya ” have been 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.

Mohit Bokariya

Acknowledgement

I would primarily like to thank my supervisor Prof. Shyam S. Rai for giving me the opportunity of MS project, his support and his belief in my abilities. I sincerely thank my expert Dr Arjun Datta for his valuable time, suggestions and honest review .

I am extremely thankful to my co-supervisor Dr Vivek Kumar for his continuous guidance , and clarifying several doubts, valuable comments and suggestions which strengthen my research work. I also want to thank my labmate Dipak Kumar Chaubey for his help in my MS project.

Table of Contents.

List of Figures.

1. Abstract.

Seismic waveforms recorded during quiet periods (i.e., no earthquakes) is known as ambient noise. It is generated by natural (e.g., oceanic gravity waves) or anthropogenic sources. I processed and modeled the ambient noise data from the 15 seismic networks in the western Himalayas to compute the seismic velocity structure of the underlying crust. The ambient noise processing data can be divided into 4 steps. (1) Processing of raw data , (2) cross-correlation and stacking, (3) measurement of group velocity using dispersion cruve , (4) inversion of dispersion curve. I used the Multiple Filter Technique (MFT) method for the surface wave dispersion of the fundamental mode Rayleigh wave. The inversion of the dispersion curves provides a crustal velocity model. The velocity model shows a velocity reduction between 10 and 30 km depth that indicates a Low-Velocity Layer (LVL) which may indicate the presence of partial melt or aqueous fluid or both.

2. Introduction and Theory

2.1. Introduction

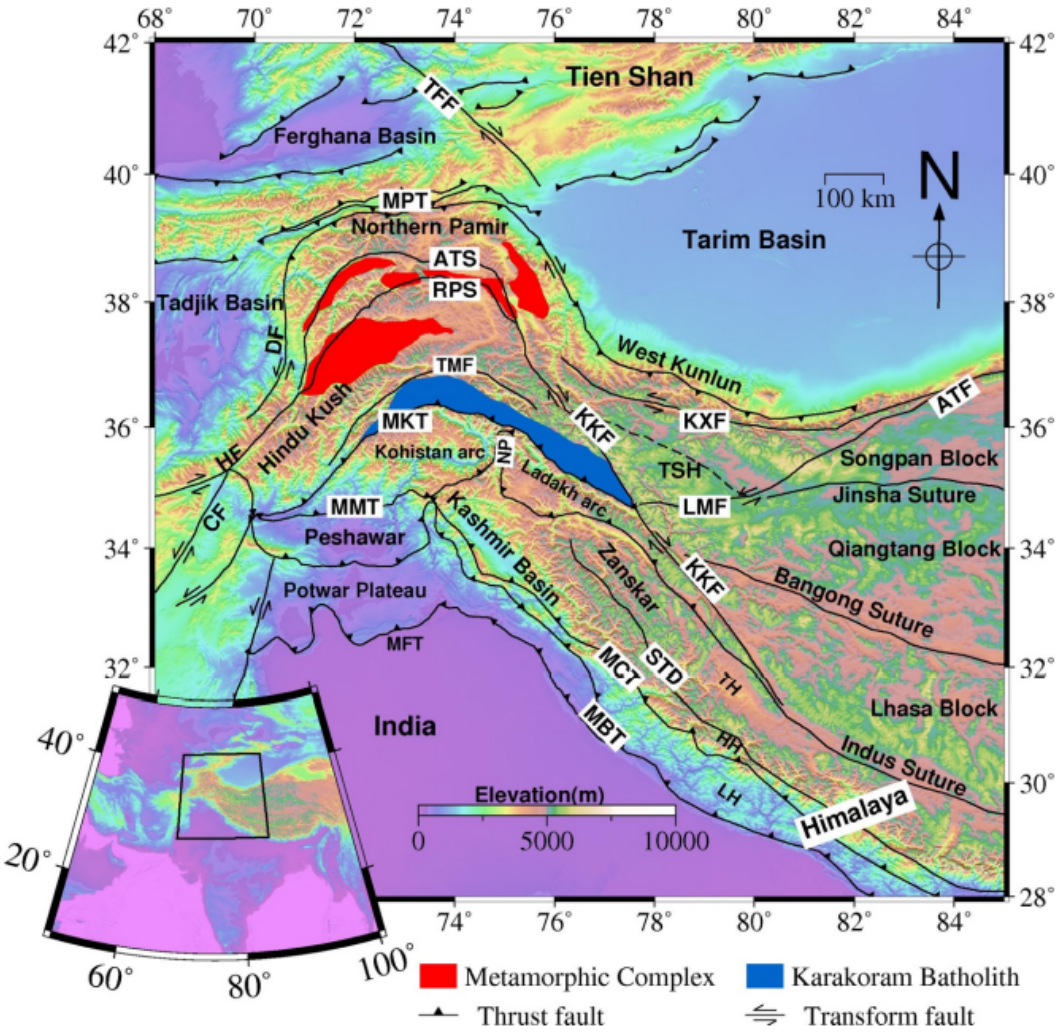
The formation and structure of Himalaya have become one of the most interesting subjects for geologists and geophysicists for more than a century. Himalaya formed due to the conversion of indian and asian plate and the collision of these two plates created several faults in Himalaya. Based on faults boundaries, Himalaya is divided into several regions that have different types of rocks. These regions in Himalaya have different crustal structure and my main objective is to get the crustal structure of western Himalaya. There are several studies on crustal structure of western Himalaya but still some questions and problems are not resolved and some of the studies are controversial, as discussed below. In this thesis, I use broadband stations in the western Himalaya to obtain crustal structure using the ambient noise analysis.

2.1.1.Formation Of Himalaya.

Around 300 Ma ago there existed a supercontinent of the Southern hemisphere called Gondwana and India was part of it. In early Permian(~300 Ma), a great rifting event (Panjal basalt explosion) occurred which split apart Gondwana and created a new oceanic basin. Before the collision of Indian and Asian plates, both were separated by the Tethys Ocean. There are two different motions of the Indian Plate. (1) Unusual fast convergence of India and Asia between 80 to 65 Ma. Mantle plume eruption at approximately 67 Ma and the formation of the Deccan traps have been used to explain the quick pace of the Indian plate. (2) Reduction in convergence rate after 65 Ma. Because of the collision, the convergence rate dropped from 160 mm/yr to 50 mm/yr after 70-50 Ma. Due to that convergence, the upper crust of India created the Himalayas and the middle and lower crust subducted beneath the Asian crust.

2.1.2.Thrust Faults in the Himalayas.

There are several faults in the Himalayas (see figure 1). From north to south, (i) Indus suture zone (ISZ), (ii) South Tibetan Detachment (STD), (iii) Main Central Thrust (MCT), (iv) Main Boundary Thrust (MBT), (v) Main Frontal Thrust (MFT). These thrust faults converge to a major detachment called the Main Himalayan Thrust (MHT), which separates the top of the underthrusting Indian plate from the overlying Himalayan wedge (Schelling & Arita, 1991). Based on these faults, the Himalayas is divided into three parts. (i)Tethyan Himalaya Sequence (THS): It is surrounded by ISZ in the north and STD in the south. Sedimentary rocks that are marine and fossiliferous are locally interbedded with volcanic flows (~650-40 ma). (ii) Greater Himalayan Sequence (GHS): Surrounded by STD in the north and MCT in the south. It Contains high-grade metamorphic rocks like gneisses migmatites sillimanite and kyanite (800-480 Ma). (iii) Lesser Himalaya Sequence (LHS): Surrounded by MCT in the north and MFT in the south, have low grade metamorphic rocks (1870-800 Ma).



Figure(1) : Topography map of the study region (Western Himalaya) (Kumar et al., 2022)

Classical divisions of Himalayas, as discussed above, become less applicable in western Himalayan west of ~ 77°E, where the MCT comes close to the MBT, and both extend as a parallel fault pair with a narrow but continuous (< 8 km wide) Lesser Himalayan sedimentary zone further west (DiPietro & Pogue, 2004; Yin, 2006). If I see the Chamba-Lahul and Kashmir Himalaya region,I notice that the high-grade rocks between the MCT and the THS are missing. If I go from east to west the continuity of the STD is broken in Lahul and further west I see the Zanskar crystallines region which is bounded by the Zanskar shear zone(ZSZ), which is considered equivalent to STD. Within the GHS of the western Himalaya, there are two windows (i) the

Kishtwar Window (KW) in the Zanskar region and (ii) the Kulu-Rampur Window (KRW) in Himachal Himalaya east of the Chamba region which expose the low-grade LHS rocks and represent the folding of the MCT zone. As a result of the motion along the MCT and STD, the high-grade rocks of the GHS in the central Himalaya are laterally exhumed as a slab.

2.1.3.Crustal Structure.

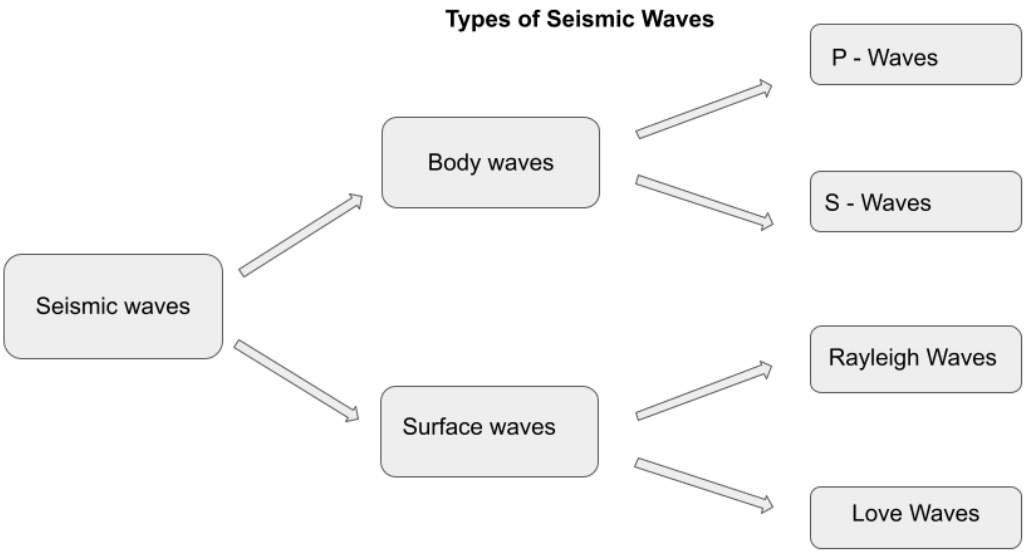
South of the Himalayan front (MFT), the Indian crust has a uniform internal structure with a crustal thickness of ~ 42 km and ~ 48 km near the southern and northern boundary of the Foreland basin, respectively (Gilligan & Priestley, 2018; Rai et al., 2006). Within the Himalayan range, the crustal thickness increases from ~ 50 km beneath the Lesser Himalaya to 60-65 km beneath the Tethys Himalaya (Nábělek et al., 2009; Rai et al., 2006). Beneath the Tethys Himalaya and ISZ, the Indian crust exhibits mantle velocities (Vs > 4.4 km/s, Vp ~ 8.4 km/s) commonly attributed to eclogite grade metamorphism (Monsalve et al., 2008; Sapin & Hirn, 1997; Nábělek et al., 2009). The surface wave dispersion analysis in the northwest Himalayas indicates continuous low shear wave velocities (with 7-17 % velocity reduction) at mid-crustal depth (~ 30 km) from Tethyan Himalaya to the Tibet plateau (Caldwell et al., 2009). It is thought that this ductile zone represents an active channel flow.

The crustal thickness of Ladakh is 60 to 70 km. The s-wave velocity in Tibet is observed anomalously low at depth of (20−40 km) and interpreted to indicate the presence of partial melts (Owens & Zandt, 1997, Makovsky & Klemperer, 1999, Yang et al., 2012). Low seismic velocities at mid-crustal depths (Caldwell et al., 2009) and high seismic attenuation (Rai et al., 2009) indicate a melt-dominated crust of thickness reaching up to ~ 75 km (Rai et al., 2006)

2.2.Seismic wave.

Seismic wave is an energy that is mostly generated by a sudden movement of tectonic plates/breaking of rocks within the earth. These waves are also generated by volcanic eruptions, landslides, rushing rivers, anthropogenic sources, etc. These waves can travel through the earth's interior and along the subsurface. Seismic waves are also called elastic waves because when these waves travel through rock/layer, they change their shape and position temporarily and that rock/layer comes back to its original shape and position after the wave passes through.

Seismic waves are divided into 2 types - (1.) Body waves. (2.) Surface waves (see figure 2).



Figure(2): Types of Seismic wave

2.2.1.Body waves.

Body waves propagate in three dimensions through the interior of the earth and have smaller amplitude and short wavelengths than surface waves. There are two types of body waves. (1). P-waves (2). S-waves

2.2.1.a.P-waves - P waves are the fastest kind of seismic waves, travel faster than s-waves, and appear first on seismograms that's why these waves are called primary waves. P waves can propagate through solids (such as rocks) and fluid (such as magma), so these waves can propagate through the earth's outer and inner core. While passing through the medium these waves compress and expand the medium in the direction of propagation of waves. P-waves are also called compressional waves.

2.2.1.b.S-waves- S-waves are 40 % slower than p-waves (for ideal Poisson solid), appearing after p-waves on seismograms that's why these waves are called secondary waves. These waves can travel through solids but can't travel through fluid that's why these waves can't penetrate earth's outer core. while passing through the solid medium, the horizontal component of the s-wave vibrates the earth up and down, and the vertical component of the s-wave vibrates the earth left and right. S-waves are transverse waves, that travel perpendicular to the direction of wave propagation

2.2.2. Surface waves.

Surface waves are formed from the interaction/combination of P-waves and S-waves (vertical and horizontal components). These waves propagate along the earth's surface. These waves are slower and have a wider range of frequencies than body waves. Surface waves are dispersive in nature which means the velocity of surface waves depends on frequency. Because of their high amplitude, these waves cause more

damage than body waves. These waves have a slower rate of decay so it can complete multiple rounds of the earth following a large earthquake. There are two basic kinds of surface waves. (1). Rayleigh waves. (2). Love waves

2.2.2.a.Rayleigh waves - Rayleigh waves are generated by the interaction between P-waves and SV-waves (vertical component of s-wave). Rayleigh waves show retrograde elliptical motion near the surface and the direction of this motion is opposite to the direction of propagation but as amplitude decreases with the depth it shows prograde motion in the deeper part of the earth. There is a 90° phase difference between the vertical and horizontal components of displacement and the component of the SV-wave decay slower than the component of the P-wave so the SV-component is dominant in the deeper part that's the reason why the Rayleigh wave shows prograde motion in deeper part. This wave moves particles side to side (because of p -waves) and up and down (because of SV-waves). Rayleigh waves are the only waves that can travel along the surface of a uniform medium.

2.2.2.b.Love waves - Love waves are produced by constructive interference between high-order SH surface multiples. These are transverse waves that vibrate the earth in the horizontal direction perpendicular to the direction of propagation which cause horizontal shearing of the earth during an earthquake because their side-to-side motion waves cause more damage and destruction during earthquakes. These waves are faster than Rayleigh waves and intrinsically dispersive in nature.

2.3.Ambient Noise.

Ambient noise is random fluctuations recorded on a seismograph during a seismically quiet period (i.e., no earthquakes). Ambient noise has a broad range of frequencies,

from 1 mHz to 100 Hz. Above 1 Hz is mainly generated by humans (Traffic, train, mining, oil drilling, sonar systems, etc) and below 1 Hz is mainly generated by natural activities (oceanic gravity waves are the dominant source of this field, wind waves, volcanic eruption, etc). Based on frequency range ambient noise is divided into 3 parts. Seismic hum (1 mHz - 20 mHz) - Produced by the Earth's continuous background free oscillations. Primary microseisms (20 mHz – 0.1 Hz) - Generated due to a direct interaction between ocean swells and shallow seafloor. Secondary microseisms (0.1–1 Hz) - Generated by a nonlinear interaction between two primary microseisms of the same frequency traveling in the opposite direction. Both primary microseisms and Secondary microseisms are called microseismic. The peak of seismic hum, primary microseismic, and secondary microseismic can be observed in the power spectrum of the seismic ambient noise wavefield near 0.01 Hz (100 s), 0.07 Hz(14 s), and 0.14 Hz(7 s) respectively. There are a few advantages to using ambient noise in place of earthquakes - (1) Ambient noise provides us with high-resolution surface wave tomography. (2) Earthquakes only occur in certain regions and directions. Ambient noise allows us to do subsurface imaging of that region where earthquakes occur rarely or not at all. (3) Earthquakes are unpredictable, to get subsurface imaging and other research on particular regions I have to wait for good earthquake data. Ambient noise provides us continuous seismic wave data.

2.4.Green's function.

The green function is an impulse response that is the output produced by an impulse function input, where Dirac delta function act as input . In Seismic Green's function is the cross-correlation of seismic time series between any two points (Seismic stations) of the Earth

We can get a green function between two stations using ambient noise data but there are two conditions. (1.) Ambient noise sources should be random. (2.) Ambient noise sources should be homogenous.

2.5.Surface Wave Dispersion.

Surface wave dispersion refers to the phenomenon where the speed of a surface wave changes with the frequency of the wave. This means that waves with different frequencies will travel at different speeds, causing the wave to disperse or spread out over time. The study of surface wave dispersion is important in geophysics and seismology, as it can provide information about the structure and composition of the Earth's subsurface. By analyzing the dispersion properties of surface waves, scientists can infer the physical properties of the Earth's crust and upper mantle, such as the thickness of the layers and the velocity structure of the underlying materials.

Let’s Consider two harmonic waves with slightly different wavenumbers and angular frequencies.

$$u_1(x,t) = Acos(\omega_1 t - k_1 x) \quad (1)$$

$$u_2(x,t) = Acos(\omega_2 t - k_2 x) \quad (2)$$

The sum of eq(1) and eq(2) -

$$u(x,t) = A[cos(\omega_1 t - k_1 x) + cos(\omega_2 t - k_2 x)] \quad (3)$$

If I assume that wave numbers (k_1 and k_2) and angular frequencies (ω_1 and ω_2) are -

$$k_1 = k + \delta k \qquad k_2 = k - \delta k \qquad k >> \delta k$$

$$\omega_1 = \omega + \delta\omega \qquad \omega_2 = \omega - \delta\omega \qquad \omega >> \delta\omega$$

After putting these wave numbers and angular frequencies in eq(3)Iget -

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

$$u(x,t) = 2A[cos(\omega t - kx)cos(\delta\omega t - \delta kx) \qquad (4)$$

The resultant wave equation is also a propagating harmonic wave equation. Since $k >> \delta k$ and $\omega >> \delta\omega$, so the first cosine term is faster than the second cosine term in time and space. So cosine term with ω and k will propagate as carrier wave and the term with $\delta\omega$ and δk will propagate as an envelope over a carrier wave.This envelope is superimposed over the carrier wave. The velocity with which the envelope propagate is called group velocity (U) and the velocity with which carrier wave propagate is called phase velocity (c). Phase velocity is dependent on wavelength

We can write phase velocity and group velocity respectively -

$$c = \frac{\omega}{k} \qquad (5)$$

$$U = \frac{\delta\omega}{\delta k} \tag{6}$$

After putting ω from eq(5) to eq(6) -

$$U = \frac{\delta(ck)}{k}$$

$$U = c + k\frac{\delta c}{\delta k}$$

$$U = c - \lambda\frac{\delta c}{\delta \lambda} \tag{7}$$

Normal dispersion - When phase velocity increases with increasing wavelength. In this case $\frac{\delta c}{\delta \lambda}$ is positive therefore group velocity U is slower than the phase velocity c.

2.6.Dispersion Measurement.

Let's say there is an event and station, the distance between them is r and azimuth angle is θ. $f(t)$ is post cross-correlated/stacked waveform signal. Where $f(t) = 0 \; for \; t < 0$. So Fourier transform of $f(t)$ will be -

$$f(\omega) = \int_0^\infty f(t)e^{-i\omega t}dt \tag{8}$$

We can write $f(\omega)$ in terms of amplitude and phase

$$f(\omega) = A(r, \theta, \omega)e^{i\phi(r,\theta,\omega)} \quad (9)$$

The amplitude term $A(r, \theta, \omega)$ has a dependency on some characteristics of the source (like event depth, fault plane solution), instrument response of recording the station and medium properties (Convolution). The Phase term is -

$$\phi(r, \theta, \omega) = k(\omega) \cdot r + \phi_{\circ}(\theta, \omega) + \phi_{i}(\theta, \omega) \quad (10)$$

Where $k(\omega)$ is the wavenumber, $\phi_{\circ}(\theta, \omega)$ is the source phase and $\phi_{i}(\theta, \omega)$ is the instrument phase.

Let's assume that the source and instrument phase shift is zero because source information cancels out during cross-correlation and I remove instrument response during raw data preparation. Now I take the inverse Fourier transform of eq(8) andIget -

$$f(t) = \int_{0}^{\infty} A(r, \theta, \omega)e^{i(\omega t - k(\omega)\cdot r)]}d\omega \quad (11)$$

If there is a monochromatic wave which has frequency $\omega_{\circ}$,I get

$$\omega_{\circ} t - k(\omega_{\circ})r = constant \quad (12)$$

We differentiate eq(12) with respect to r for all r (for k does not depend on r)

$$\omega_{\circ}(\frac{dt}{dr}) - k(\omega_{\circ}) = 0 \quad (13)$$

$$\frac{dr}{dt} = \frac{\omega_\circ}{k(\omega_\circ)} = c(\omega_\circ) \tag{14}$$

Where $c(\omega_\circ)$ is phase velocity.

To determine the group velocity I will use eq(11) and integrate it from $\omega_\circ - \varepsilon \ to \ \omega_\circ + \varepsilon$. These limits work as narrow bandpass filters.

$$f(t) = \int_{\omega_\circ - \varepsilon}^{\omega_\circ + \varepsilon} A(r, \theta, \omega) e^{i(\omega t - k(\omega)\cdot r)]} d\omega \tag{15}$$

The function $f(t)$ has its highest value when all waves in the frequency range are in phase since Group velocity is correlated with the maximum energy's propagation.

$$\frac{d}{d\omega}(\omega t - k(\omega) r)_{\omega=\omega_\circ} = 0 \tag{16}$$

$$t_{\omega_\circ} = \{\frac{dk(\omega)}{d\omega}\}_{\omega=\omega_\circ} \cdot r \tag{17}$$

$$u(\omega)_{\omega=\omega_\circ} = \{\frac{dk(\omega)}{d\omega}\}_{\omega=\omega_\circ} \tag{18}$$

Where $u(\omega)$ is the group velocity. If $u(\omega)$ or $c(\omega)$ value changes with changing ω value then our signal is dispersive. Using eq(14) and eq(18) I find the relation between group and phase velocity.

$$u(\omega) = \frac{dk(\omega)}{d\omega} = \frac{d[c(\omega)\cdot k(\omega)]}{dk(\omega)}$$

$$u(\omega) = c(\omega) + k(\omega)\frac{dc(\omega)}{dk(\omega)} \quad (19)$$

2.7.Least Square Inversion.

There is a relation between model parameters and data.I call this relationship a *quantitative model*.

$$d = Gm \quad (20)$$

Where d is the data set, m is model parameters, and G is a data kernel.I can represent d and m as a vector which have N and M elements respectively

$$\text{Data: } d = [d_1, d_2, d_3, d_4, \ldots\ldots d_N]^T \quad (21)$$

$$\text{Model parameters: } m = [m_1, m_2, m_3, m_4, \ldots\ldots m_M]^T \quad (22)$$

In the least square method I try to keep predicted data d^{pre} as close as possible to observed data d^{obs} So to do this I need to choose model parameters (slope and intercept) very carefully.

We take the difference between d^{obs} and d^{pre} for each observation which is called prediction error.

$$\text{Prediction error : } e_i = d_i^{obs} - d_i^{pre} \tag{23}$$

Which model parameters provide the smallest overall error E, called the best-fit line.

$$E = \sum_{i=1}^{N} e_i^2 = e^T e \tag{24}$$

This overall error E is also called squared Euclidean length of the vector e which is the sum of the square of the individual error e_i

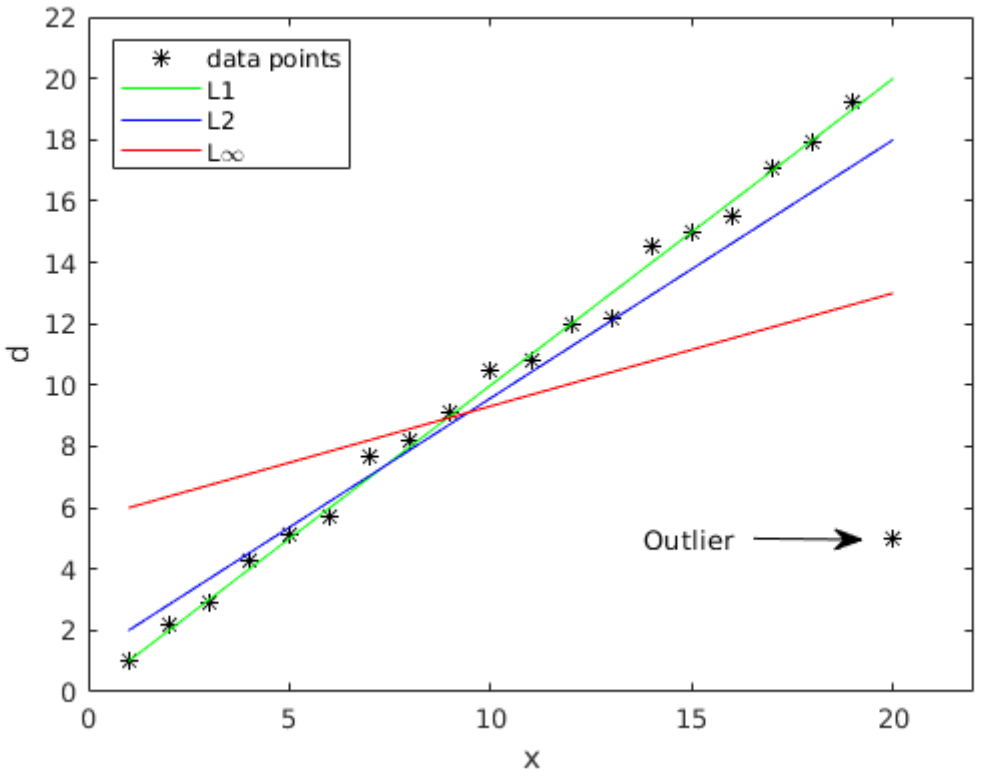
In this method there is a term *norm* which is the measure of length or size. In other words, it is the sum of n the power of the elements of the vector e represented as L_n

$$L_1 \text{ norm : } ||e||_1 = \left[\sum_i |e_i|^1\right] \tag{25}$$

$$L_2 \text{ norm : } ||e||_2 = \left[\sum_i |e_i|^2\right]^{1/2} \tag{26}$$

$$L_n \text{ norm : } ||e||_n = \left[\sum_i |e_i|^n\right]^{1/n} \tag{27}$$

Higher norm give larger weight to the largest element of prediction error (e). IfIhave very accurate data then a bigger difference between prediction and observed value is important. In this case,Iuse a high-order norm because it preferably gives weights to the larger errors. If I have data that is widely scattered about trend so for this case I use a low order norm because errors of different sizes are given more equal weight. Different types of data obey different types of statistics so based on the statistic I choose the norm.



Figure(3): This is a straight line fits using L_1, L_2 and L_∞ norm for measuring error. The L_1 gives least and L_∞ gives most weight to one outlier.

2.7.1.Solution of the Linear inverse problem using Least square method.

The primary objective of the least square method is to minimize the sum of the squared error. To get the general solution of the linear inverse problem using least squares I take the derivative of total error E (eq. 24) with respect to one of the model parameters m_p and equate it to zero. The total error is -

$$E = e^T e = (d - Gm)^T (d - Gm) \tag{28}$$

$$E = \sum_{i=1}^{N} \left[d_i - \sum_{j=1}^{M} G_{ij} m_j \right] \left[d_i - \sum_{k=1}^{M} G_{ik} m_k \right] \tag{29}$$

Size of G depend on the size of d and m . let's say if I have N numbers of elements in d and M numbers of elements in m then the size of G will be NxM.
Multiplying the term of (eq. 29) and reversing order give us -

$$E = \sum_{j=1}^{M} \sum_{k=1}^{M} m_j m_k \sum_{i=1}^{N} G_{ij} G_{ik} - 2 \sum_{j=1}^{M} m_j \sum_{i=1}^{N} G_{ij} d_i + \sum_{i=1}^{N} d_i d_i \tag{30}$$

NowIhave three terms in this equation soIwill calculate the derivative of each term with respect to m_p

$$\frac{\partial}{\partial m_p} \left[\sum_{j=1}^{M} \sum_{k=1}^{M} m_j m_k \sum_{i=1}^{N} G_{ij} G_{ik} \right] = \sum_{j=1}^{M} \sum_{k=1}^{M} \left[\delta_{jp} m_k + m_j \delta_{kp} \right] \sum_{i=1}^{N} G_{ij} G_{ik}$$

$$= 2 \sum_{k=1}^{M} m_k \sum_{i=1}^{N} G_{ip} G_{ik} \tag{31}$$

The derivatives of the form $\frac{\partial m_i}{\partial m_j}$ equals to 1 (unity) when $i = j$ or zero when $i \neq j$.

Therefore $\frac{\partial m_i}{\partial m_j}$ is Kronecker delta δ_{ij}

The derivative of a second term -

$$\frac{\partial}{\partial m_p}\left[-2\sum_{j=1}^{M} m_j \sum_{i=1}^{N} G_{ij} d_i\right] = -2\sum_{j=1}^{M} \delta_{jp} \sum_{i=1}^{N} G_{ij} d_i$$

$$= -2\sum_{i=1}^{N} G_{ip} d_i \qquad (32)$$

The third term doesn't have m term so it's derivative would be equal to zero

$$\frac{\partial}{\partial m_p}\left[\sum_{i=1}^{N} d_i d_i\right] = 0 \qquad (33)$$

NowIwill combine all of three terms

$$\frac{\partial E}{\partial m_p} = 2\sum_{k=1}^{M} m_k \sum_{i=1}^{N} G_{ip} G_{ik} - 2\sum_{i=1}^{N} G_{ip} d_i = 0 \qquad (34)$$

We can represent this equation in matrix form

$$G^T G m - G^T d = 0 \qquad (35)$$

$G^T G$ is a square matrix of MxM dimension and $G^T d$ is a vector of length M. This is a square matrix equation and using this equationIcan find unknown parameters. If $\left[G^T G\right]^{-1}$ exists then

$$m^{est} = \left[G^T G\right]^{-1} G^T d \tag{36}$$

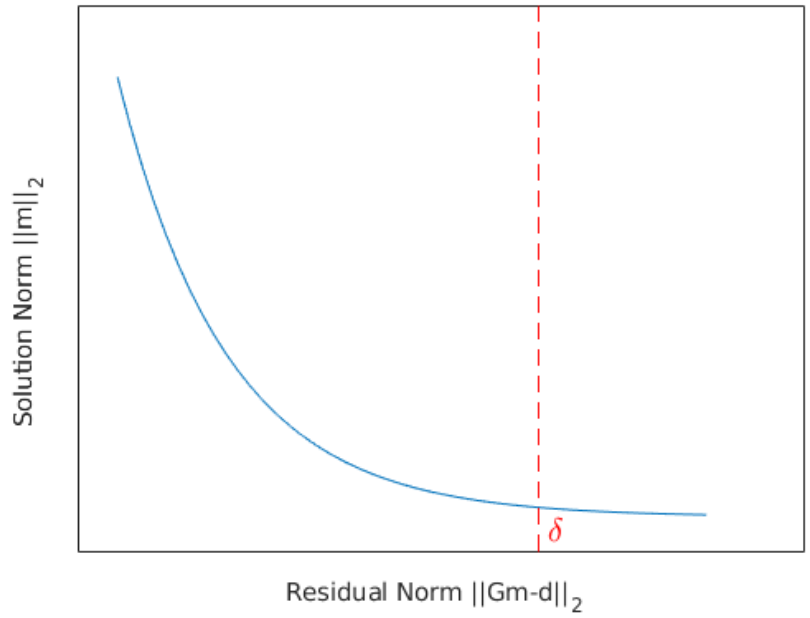
2.8.Regularization.

For the general linear least square problem I can have multiple solutions that can suitably fit the data so that $||Gm - d||_2$ is small enough.

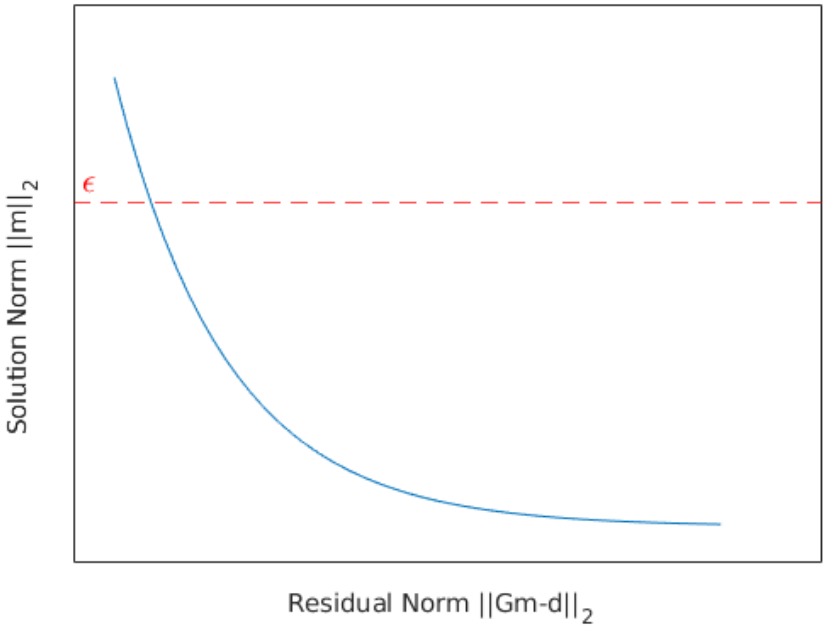
In zeroth-order Tikhonov regularization, in the first case, I try to minimize the norm of m and consider all possible solutions with$||Gm - d||_2 \leq \delta$. where δ is a particular misfit norm (see figure 4a)

$$\min||m||_2$$
$$||Gm - d||_2 \leq \delta. \tag{37}$$

The primary goal is to get minimum value of the norm of m is to find a solution with just the right amount of feature complexity, as determined by measuring a norm, to correctly fit the data. Note that with increment in δ , the number of possible models also increases, and the minimum value of norm m decreases. As a result, I can draw a curve showing norm m minimal values in relation to δ.



(a)



(b)

Figure(4): figure(a) The trade-off curve between $||m||_2$ and $||Gm - d||_2$ with varying δ figure(b) : The trade-off curve between $||m||_2$ and $||Gm - d||_2$ with varying ϵ [Aster et al., 2012]

In the Second case I consider all possible solutions of the norm $m \leq \epsilon$ and minimize the $||Gm - d||_2$. where ϵ is a particular model norm (see figure 4b)

$$\min||Gm - d||_2$$
$$||m||_2 \leq \epsilon \quad (38)$$

As ϵ decreases set of possible solution reduce and minimum value of $||Gm - d||_2$ increases. In order to get suitable values of $||m||_2$ and $||Gm - d||_2$Ineed to adjust ϵ. In both cases when I increase δ , the value of ϵ decreases and vice versa so it is hard to get suitable values of ϵ and δ together.

To get appropriate values of ϵ and δI consider the third case **damped least square** problem. After applying the Lagrange multipliers method on eq(37)Iget

$$\min||Gm - d||_2^2 + \alpha^2||m||_2^2 \quad (39)$$

Where α is a regularization parameter.I can get the solutions of eq.(37) and eq.(38) using eq.(39) by adjusting the α until the constraints are justified.I can find a suitable value of ϵ that will approximately minimize prediction error $E = (Gm - d)$ while approximately minimizing the Euclidean length of the solution$L = m^T m = \sum m^2$.

There is no simple or direct method to estimate a suitable value of ϵ, It is estimated by trial and error.

2.9.Continental Crust -

2.9.1.Seismic velocity and composition.

The outer layer of the earth is called crust there are 2 types of crust (i) Continental crust (ii) Oceanic crust . Continental crust thickness varies from 16 km to 72 km and the average thickness is 41 km. In three-dimension, the continental crust is very diverse. So the continental crust is divided into three-layer models. Upper, middle, and lower crust are the three layers of continental crust. The *Upper continental crust* is mainly composed of gneisses, tonalites, phyllite, mica quartz schist, granites, and granodiorites types of rock and have compressional wave velocities V_p range from 5.6 to 6.4 Km/s and its thickness varies from 10-25 km. The *middle crust* is usually composed of amphibolite facies and has compressional wave velocities Vp range from 6.4 to 6.8 km/s and its thickness varies from 5-15 km. The *lower crust* is composed of metamorphic rocks in granulite facies (mainly granulites, gabbros, diorites, and amphibolites) and has compressional wave velocities Vp range from 6.8 to 7.6 km/s (only Precambrian shield and platform areas have Vp from 7.2 to 7.6 km/s).
In general, the velocity increases with depth because of a gradual change from granite/granodiorite to gabbro. From the upper to middle crust the increase in velocity can be explained by a large volume of mafic mineral assemblages in the middle crust.

2.9.2.Moho discontinuity.

Moho is the boundary between the earth's crust and the mantle and here discontinuity means a surface at which seismic velocities change. The change in velocities is from 0.5 to 1.5 km/s.The P-wave velocity at the top of the mantle varies from 7.6 to 8.8 km/s and it’s referred to as P_n velocity. In some areas, the seismic velocities may gradually shift from crustal to mantle values and the Moho does not show a sharp seismic discontinuity. In some areas, the moho is not a simple boundary.

3. Methods.

3.1.Processing of Ambient Noise

To get the shear wave velocity model I processed the ambient noise and the process is divided into four principal phases (see Figure 5). I am going to explain these phases below.

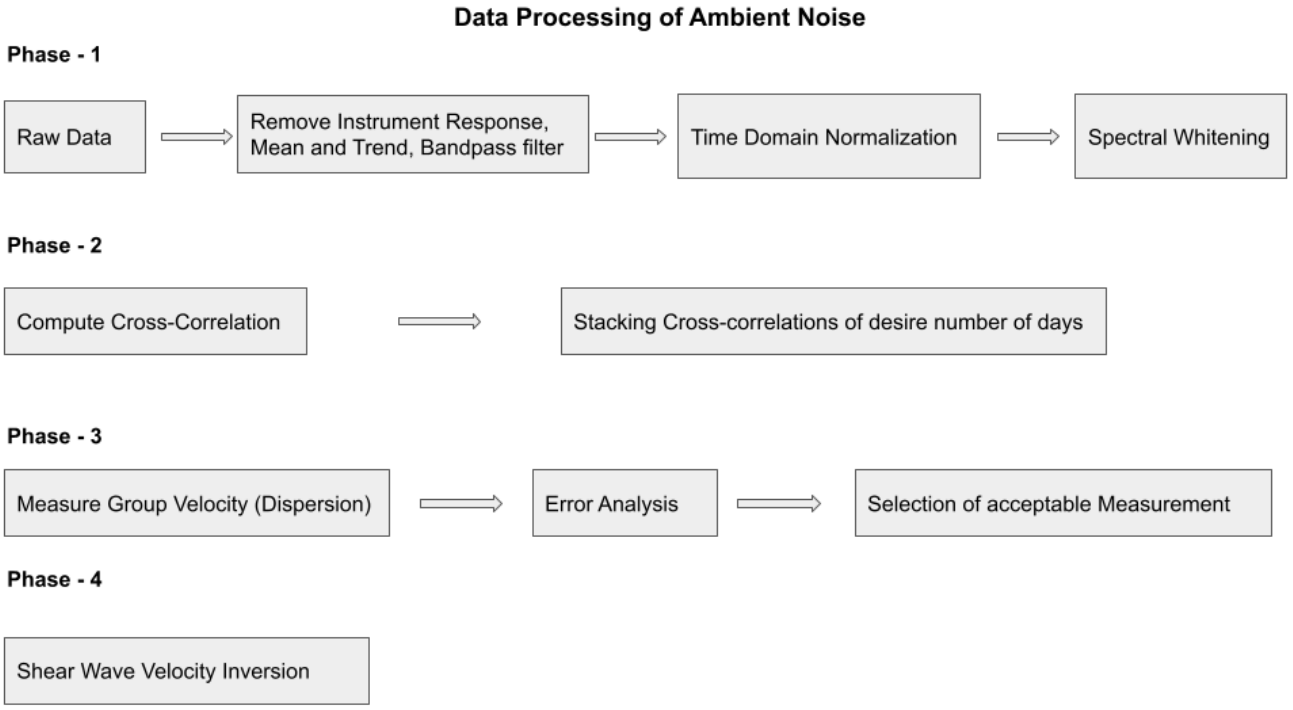


Figure (5): It is a schematic representation of Ambient noise data processing.I divided this process into 4 parts. Phase 1 is single-station data preparation removing noises, increasing the signal-to-noise ratio, and resolution enhancement. Phase 2 includes

Cross-Correlation between two stations and stacking cross-correlation of the desired number of days. Phase 3 is surface wave dispersion measurement (measuring group velocity and phase velocity) [Bensen et al., 2007]. Phase 4 is Shear wave velocity inversion

3.1.1. Phase -1 Initial Data Preparation.

In this Phase first, I download the Raw data. To remove instrument response I used Transfers in SAC (Seismic Analysis Code). The RESP and SAC_PZ file contains instrument response information. The transfer function uses one of the files to correct instrument response.

Removing trends helps to understand how data fluctuate about the trend. I also remove the mean from the raw data. Tapering the time series helps to eliminate the effect of the discontinuity between the beginning and end of the time series. To remove the high amplitude earthquake signals and non-stationary noise sources, I use time domain normalization. I used moving average time normalization. The microseism appears considerably more prominently on cross-correlations between distant stations because it propagates coherently over large distances. After This, I apply spectral whitening which helps to remove the effect of microseismic peaks on the cross-correlation of ambient noise. Spectral whitening helps to flatten the observed spectrum. In addition to reducing the influence of geographically separated persistent noise sources, whitening acts to increase the bandwidth of ambient noise.

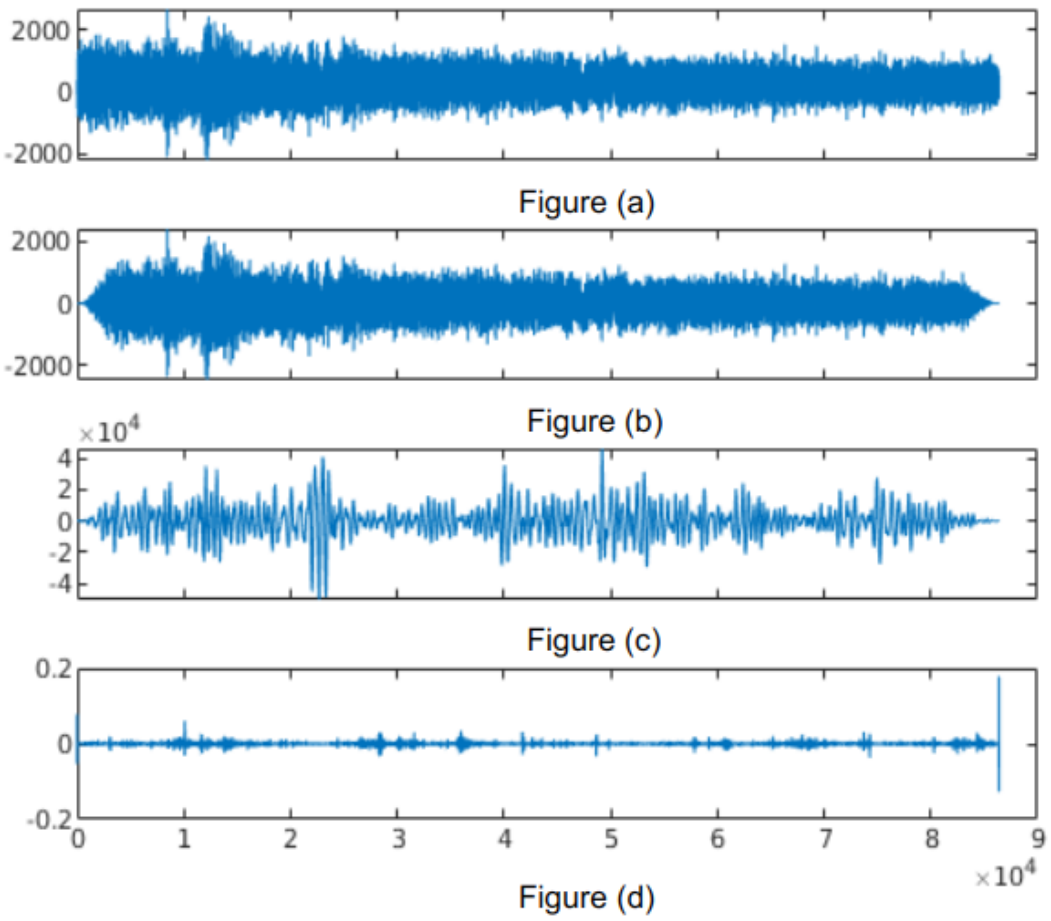


Figure (6): - Figure(a): It shows raw SAC data of 24 hours [Counts vs Time (sec)] recorded at station. **Figure(b): I** removed the mean, removed the trend, and the tapered figure(a). **Figure (c):** I removed the instrument response from Figure (b). **Figure (d):** I performed time normalization(moving average) and spectral whitening in Figure (c).

3.1.2.Phase-2 Cross-correlation and Stacking

After the data have been prepared for each day including time domain normalization and spectral whitening. Cross-correlation determines the similarity between two signals with respect to the time lag of one signal with respect to another. The higher

value of cross correlation implies higher similarity between two signals. I do cross-correlation over time between all possible station pairs of 24-hour data. Let's say I have n a number of stations so I have $n(n-1)/2$ possible station pairs. In cross-correlation the time series is divided into two parts positive lag and negative lag and the length of this time series depends on the group velocity and distance between the stations. If I have two stations A and B, the waveform $0 \rightarrow +ve$ will be $A \rightarrow B$ and the waveform $0 \rightarrow -ve$ will be $B \rightarrow A$. These waveforms are the waves that travel in opposite directions between two stations. If ambient noise sources are distributed homogeneously in azimuth then both waveforms will be identical However, it is frequently observed that there is a significant amount of amplitude and spectral asymmetry, which points to different source processes and distances from the source in directions that are radially out from the stations and non-uniform isotropic noise field distribution.

After computing cross-correlations of the desired number of days (let's say 2 years), now I have hundreds of cross-correlations of two stations then stack them together. The single-day cross-correlation does not give information about traveling time. That's why I stack a large number of cross-correlations to get the traveling time of waves from one station to another station.

Stacking is used to reduce random noise and amplify and boost the coherent signal. SNR ratio is generally improved by stacking over longer and longer time series. Stacking is a linear average of all cross-correlations.

Let's say I have n numbers of cross-correlations $C(t)$ then stacking $S(t)$ would be -

$$S(t) = \frac{1}{n}\sum_{i}^{n} C_i(t) \tag{40}$$

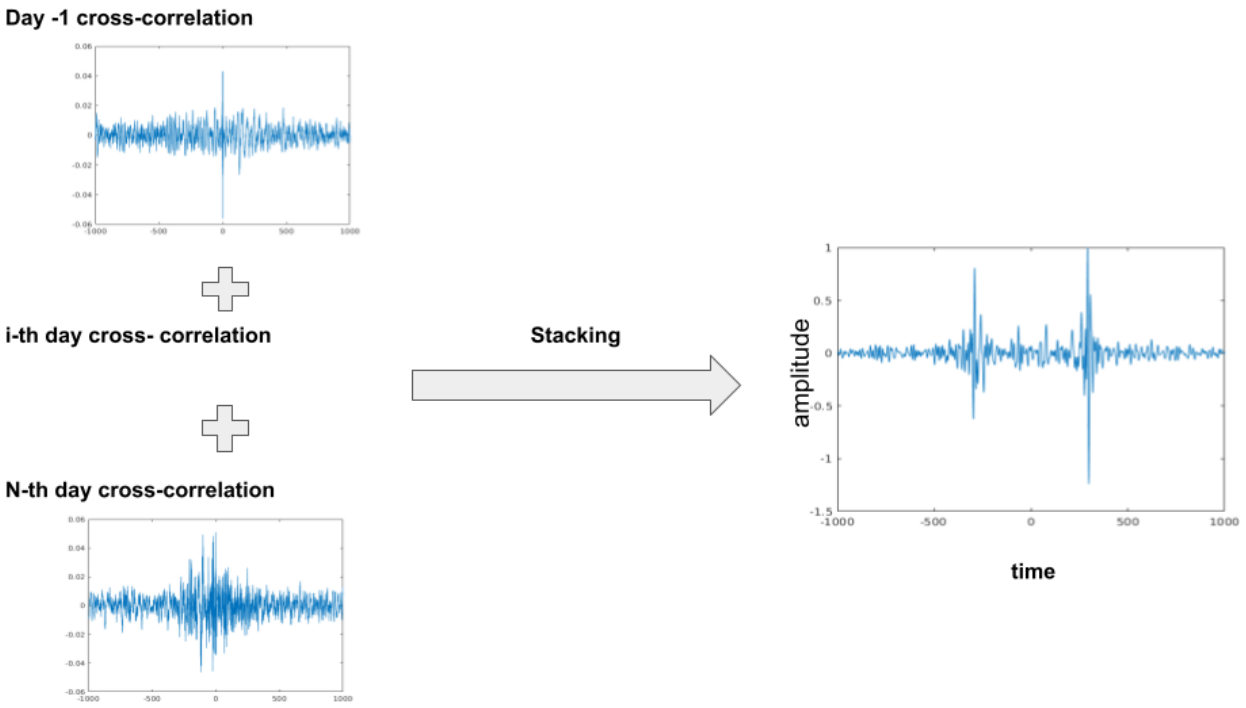
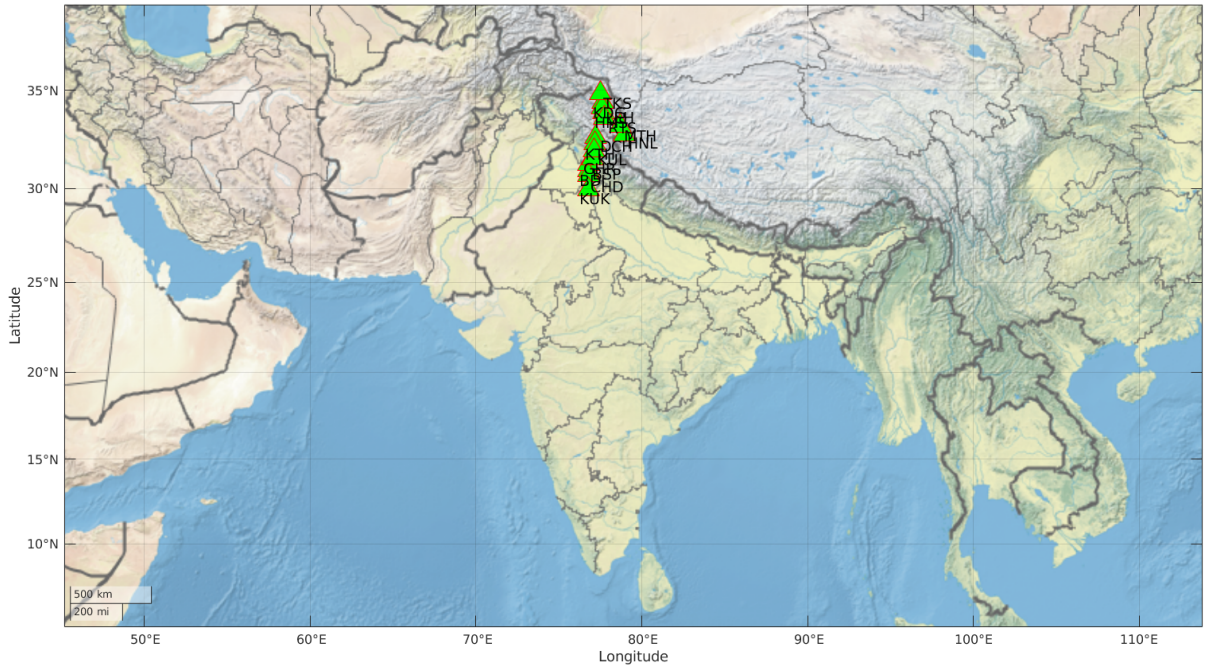
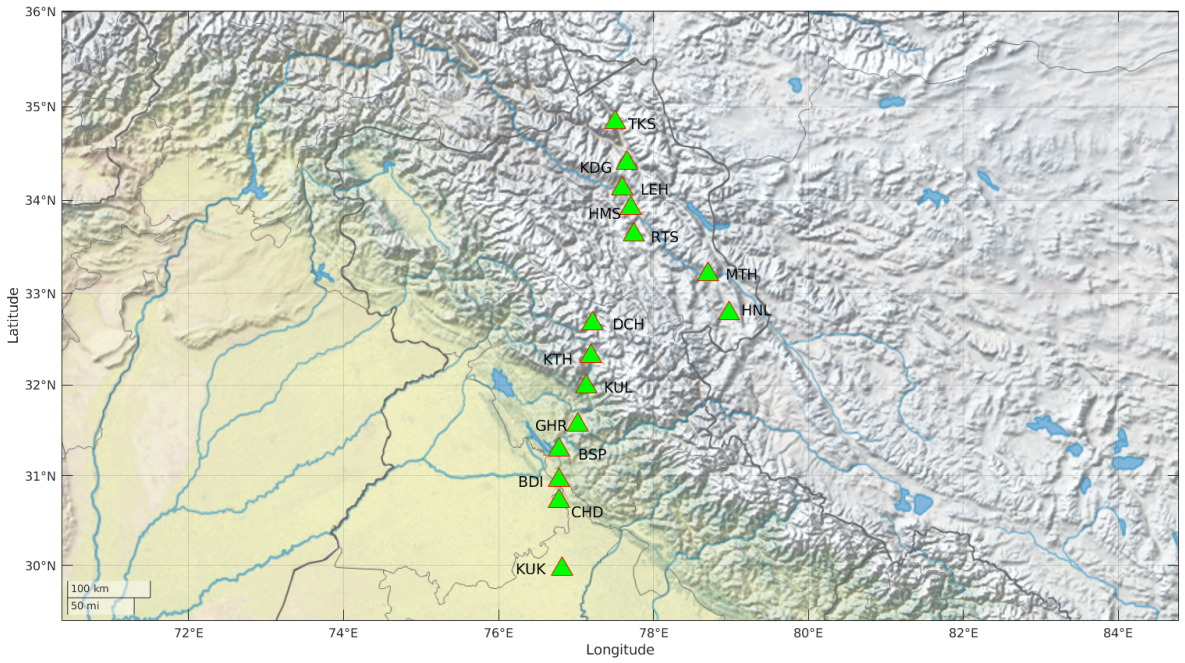


Figure (7): This shows single-day cross-correlation between two stations and stacking of cross-correlation of N number of days.



(a)



(b)

Figure (8): figure(a) &(b) shows 15 stations of Network **NW** located in north India. These stations recorded data from 2001 to 2003.

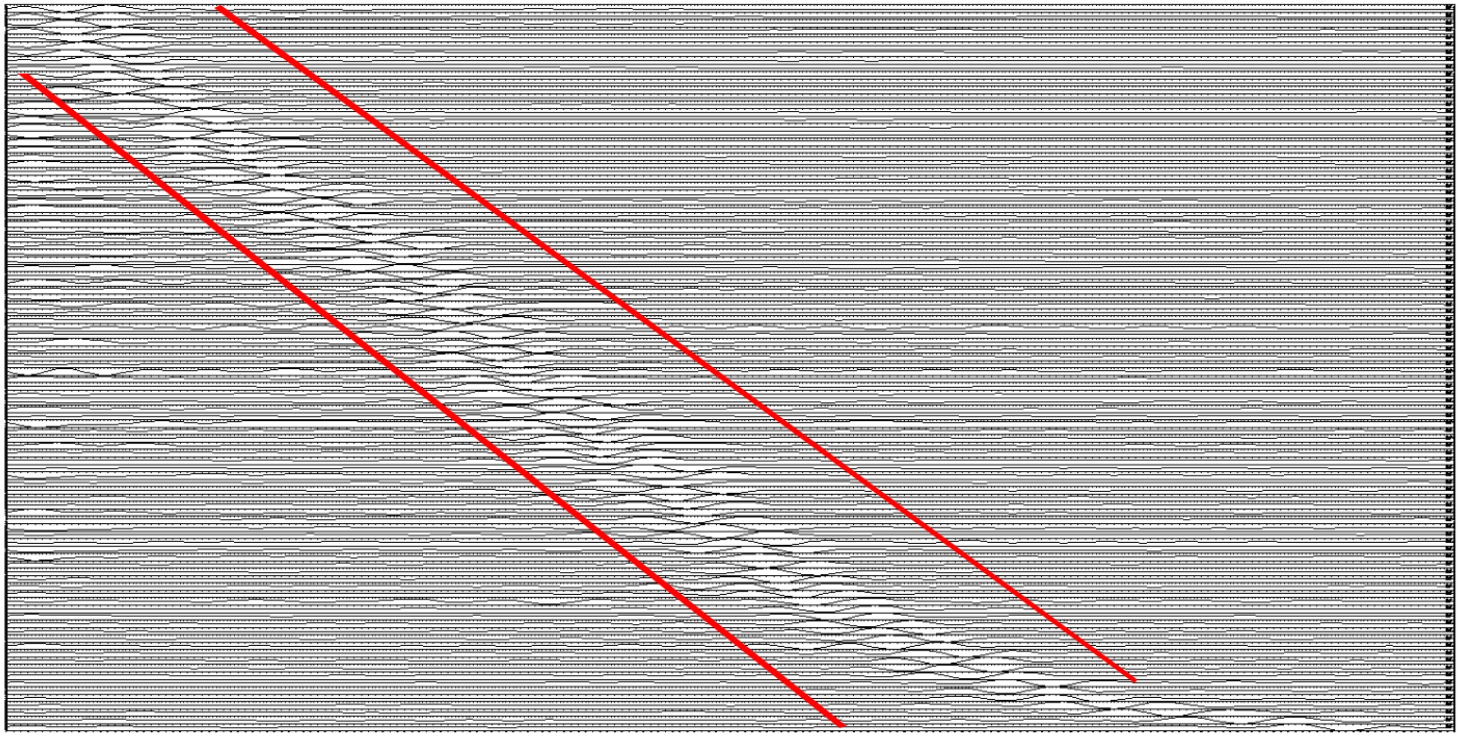


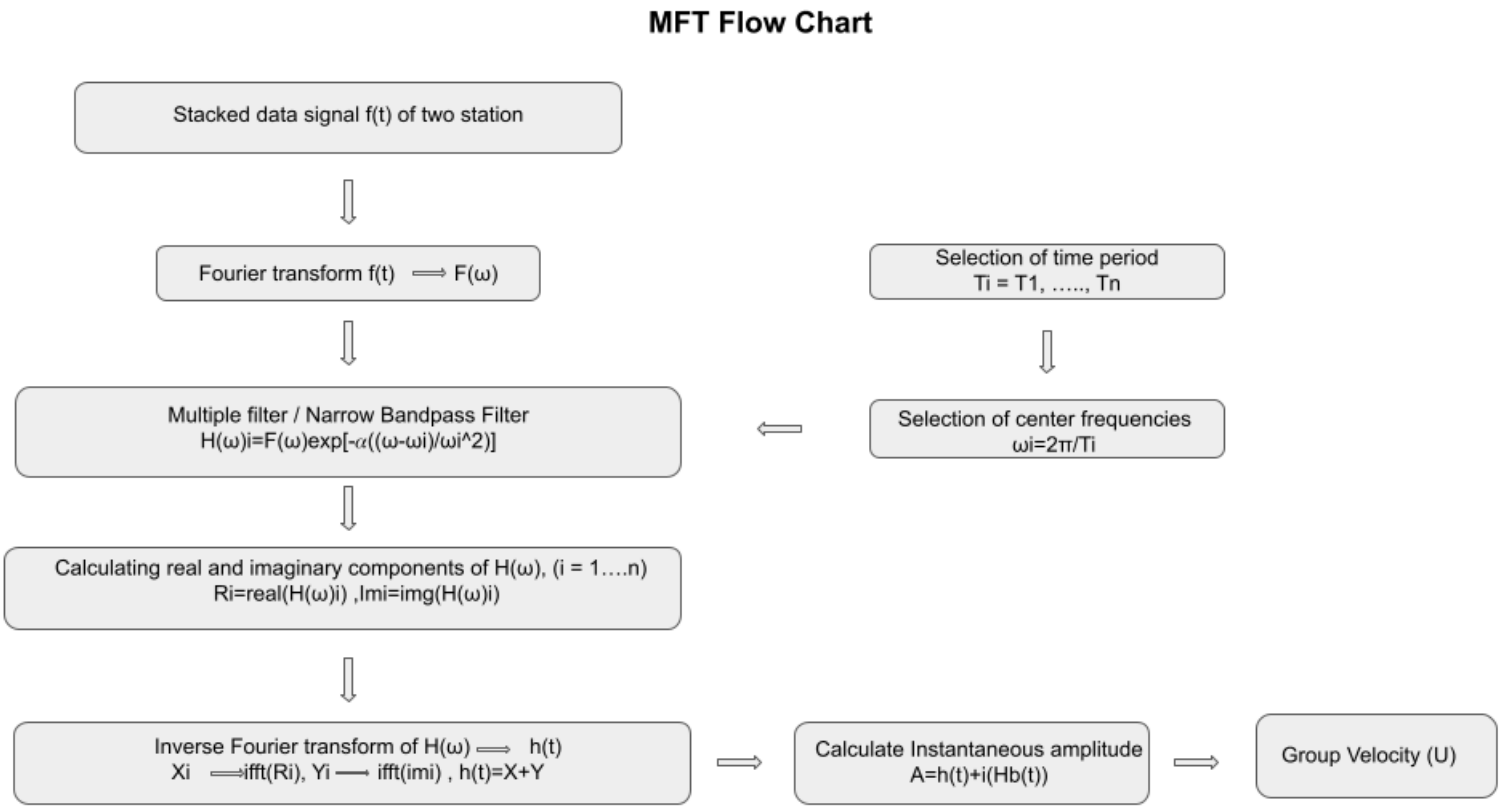
Figure (9): Shows the stacking of all possible cross-correlation between two stations in order of increasing distance. The area/amplitude between two red lines is green functions. This figure has 98 station pairs, the x-axis(time) and y-axis (amplitudes)

3.1.3. Group Velocity Measurement.

To measure the group velocity dispersion curve for Rayleigh waves use Multiple Filter Technique. In this method, I follow some steps (1)I Have stacked data $f(t)$ signal of two stations. (2)I take the Fourier transform of $f(t)$ and get $f(\omega)$. (3) I apply a narrow bandpass filter $H(\omega)$ on it. (4) In this filter I need to choose center frequencies of the desired time period. (5) Then I calculate the real and imaginary

components of $H(\omega)$. (6)I take the inverse Fourier transform of these components. (7). Calculate the instantaneous amplitude. (8)Then I plot this amplitude with the time period and velocity range and get a group velocity dispersion curve.

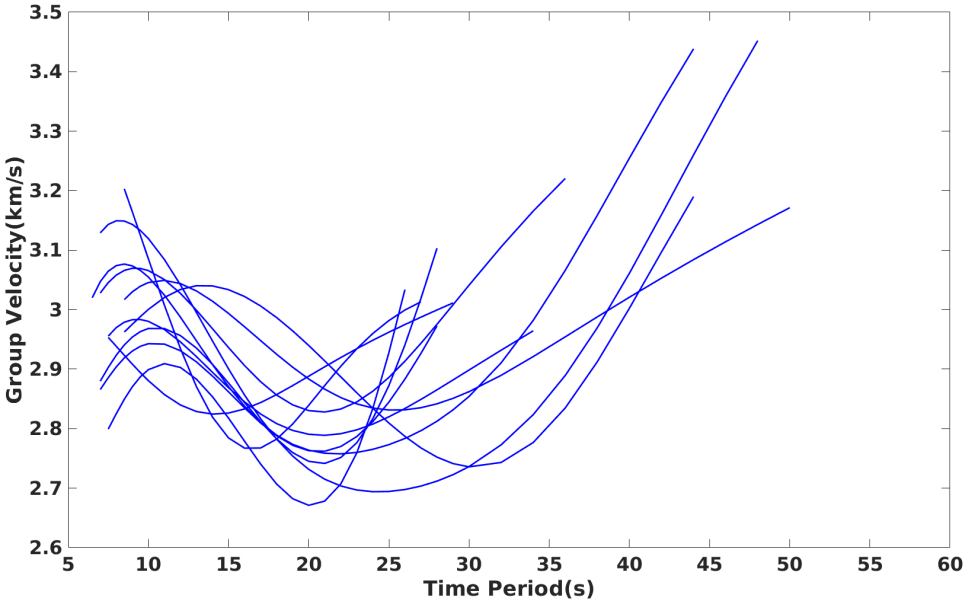
$$H_n(\omega) = F(\omega)exp[-\ \alpha((\omega - \omega_n)/\omega_n^2)] \tag{41}$$



Figure(10): This figure shows a Multiple Filter Technique flow chart from stacked data to get group velocity

3.1.4.Dispersion and Inversion

We measured the surface wave dispersion of the Rayleigh wave (fundamental mode) using the MFT method by (Herrmann and Ammon (2002). After stacking 15 stations with each other I got 98 possible station pairs out of 105. I performed the MFT method with 4 different filter parameter “alpha” values for all 98 pairs and manually picked the maximum energy(group velocity) between 6 and 50 sec. Out of 98 dispersion curves I selected 18 and discarded the rest due to unreliability and noisy data based on visual inspection.



Figure(11) : figure(a) shows the Group velocity dispersions of selected 11 station pairs

To get the inversion of the dispersion curve, I used the least square method for fitting and regularization (L-curve test) to keep the balance between error and maintain the

difference between test and predicted parameters for all 18 dispersion curves, and using the (Herrmann and Ammon (2002) program, I get 1-D shear-wave velocity profile. In inversion, the initial/earth model was prepared with layers thickness of 2 km, 5 km, and 10 km in the upper 30 km, 30 to 60 km, and 60 to 100 km depth respectively and velocity ranges 3 to 4.45 km/s (surface to100 km depth). With the increasing depth, I was having high error and low-resolution problems. Sensitivity kernels represent the relationship between dispersion velocity and earth structure. For longer the time period the surface wave becomes more sensitive to the deeper velocity structure. This implies that I can sample deeper structures with longer periods so I cut the inversion curve from surface to 40 km - 60 km depth because the minimum and maximum time period I got in dispersion was ~6 - 26 sec and ~ 6 - 55 sec respectively. In the inversion process, I used five different damping parameters and one starting model. With the combination of 18 station pairs and 5 damping parameters, I ran 90 inversions, and based on visual inspection(fitting and data quality, error) out of 18 pairs I selected 11 station pairs with 5 different damping parameters and discarded the rest 6 pairs due poor quality of data. In selected 11 station pair I performed L-curve test (see figure 12) for different damping parameters. Based on L-curve test I chosen 0.1 and 0.5 damping values.

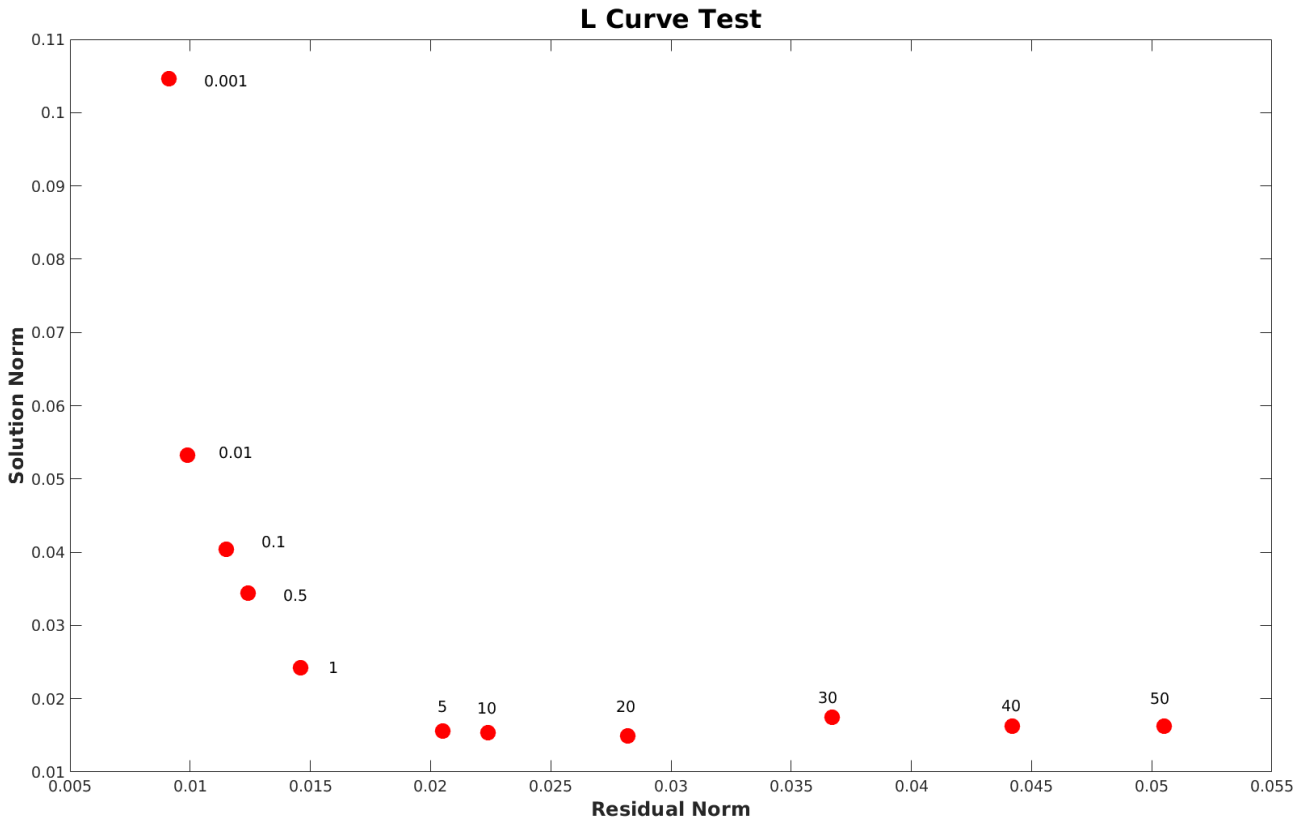


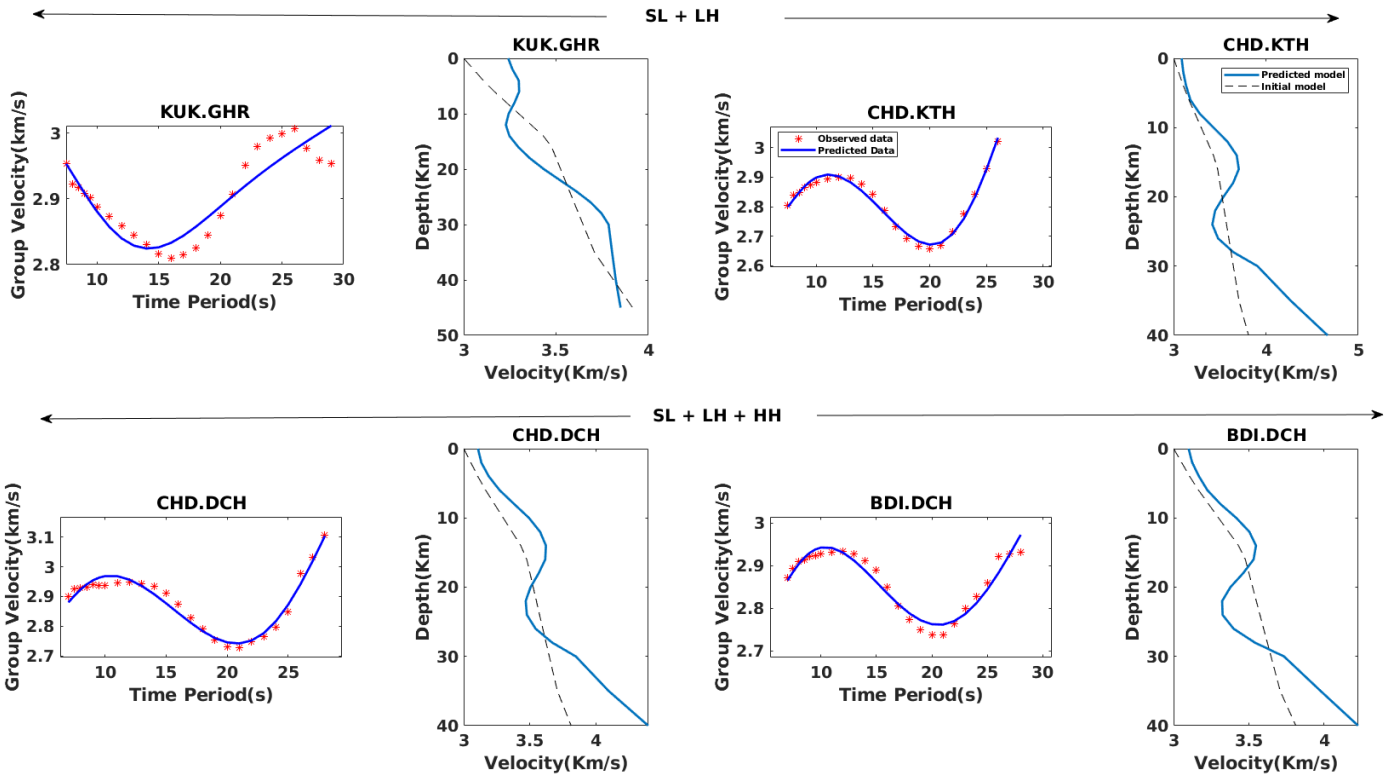
Figure (12) : L-curve test for inversion damping parameters

---

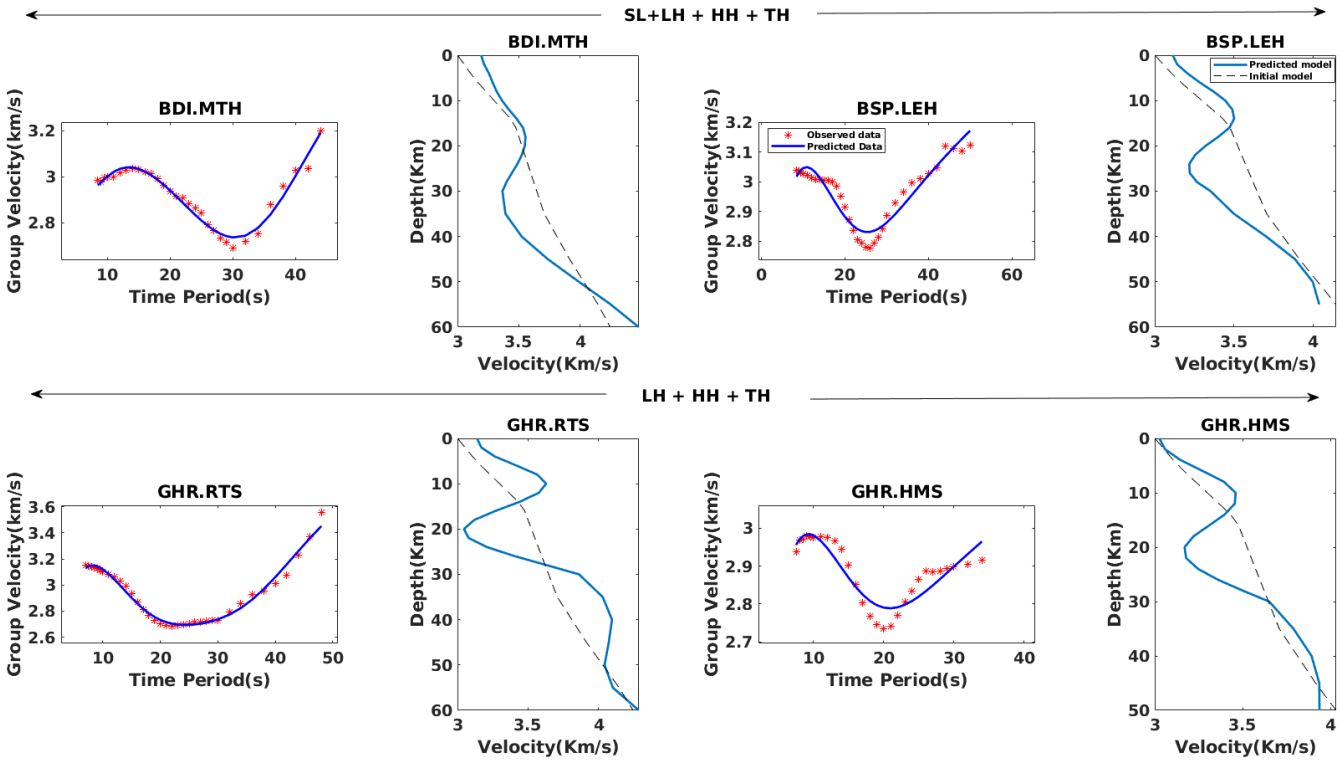
4. Results.

We calculated and plotted the one-dimensional s-wave velocity models between two stations for 13 station pairs (figure-3.1). In (figure-3.1) I have surface wave dispersion (left) and inversion of dispersion (right), Where SL, LH, HH, and TH represent Shivalik, Lesser Himalaya, High Himalaya, and Tethys Himalaya respectively. In all inversions I observed a pattern, first velocity increases with depth after a certain depth (5-15 km) velocity decreases till (20-30 km) depth and then increases again. In the SL+LH and SL+LH+HH region velocity reduction is upto ~3.4 km/s and ~3.3 but In both the LH+HH+TH and HH+TH region velocity reduction is upto ~3-3.2 km/s,

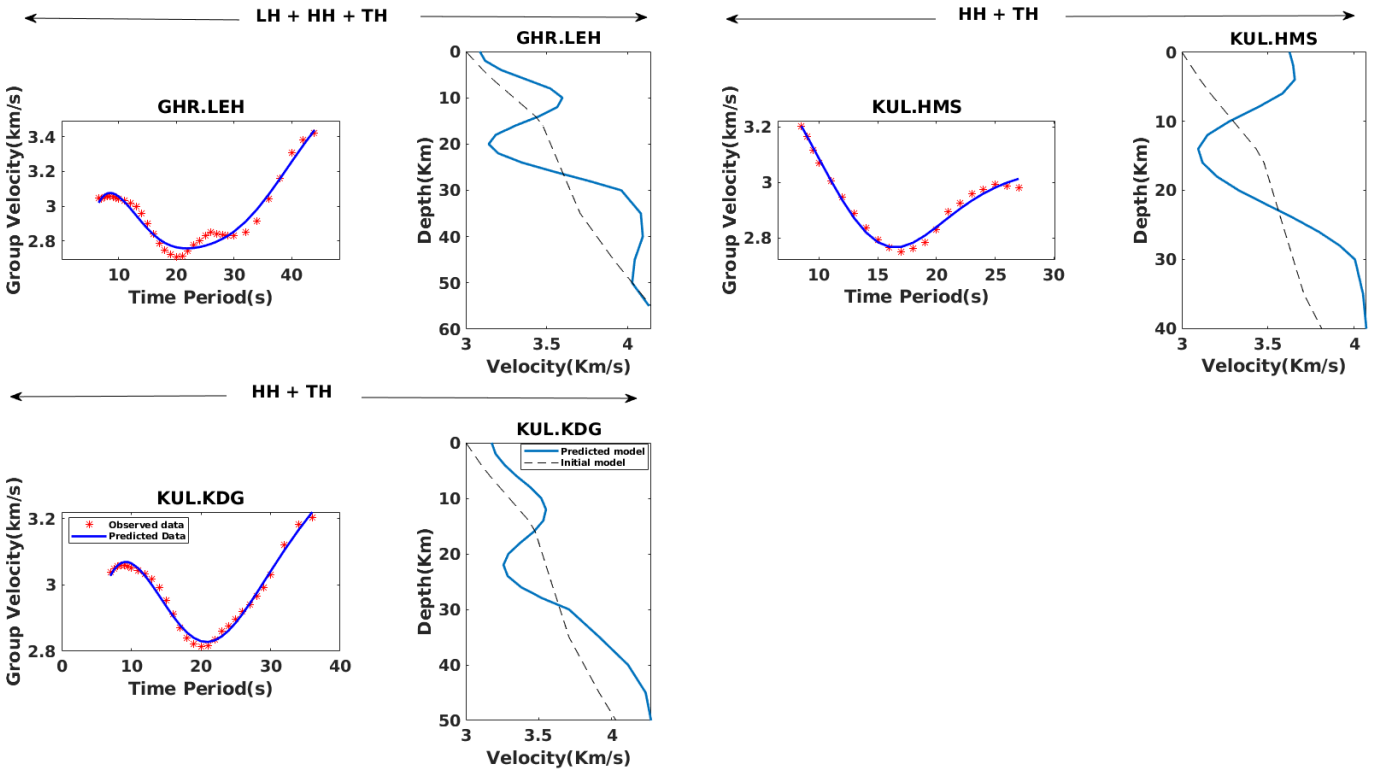
Since in our 1-D shear wave velocity models I could not covered individual regions (SL, LH, HH and TH) due less stations pair availability so it's hard to explain velocity variation individual region wise. So on the observation SL region does not have LVL and LH region is unclear but HH and TH both region show LVL so the low-velocity layer more dominating in western himalaya as we go from south (SL) to north (TH). Previous studies (i) (Kanna et al., 2021) also performed seismic crustal shear velocity structure using receiver function modeling with the same stations. They also found LVL in both regions LH and HH that is more significant and thicker in HH than LH. (ii) (Caldwell et al.,2009) obtained shear-wave velocity models for HH and TH using regional earthquakes with the same station's data and they observed LVL in both region. (iii) (Arora et al., 2007) found low resistivity in HH and TH that indicates low velocity in the mid-crust (centered at 30 km).



(a)


SL+LH + HH + TH
BDI.MTH
Group Velocity(km/s)
Time Period(s)
Depth(Km)
Velocity(Km/s)
BSP.LEH
Observed data
Predicted Data
Predicted model
Initial model
LH + HH + TH
GHR.RTS
GHR.HMS

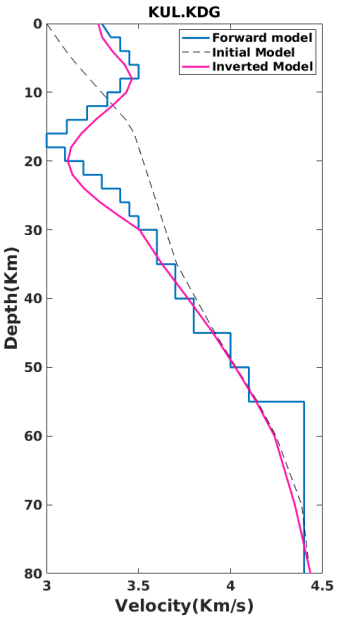
(b)


LH + HH + TH
GHR.LEH
Group Velocity(km/s)
Time Period(s)
Depth(Km)
Velocity(Km/s)
HH + TH
KUL.HMS
KUL.KDG
Observed data
Predicted Data
Predicted model
Initial model

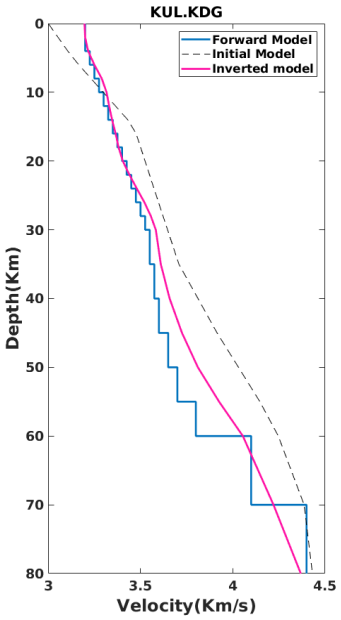
(c)

Figure(13): All these 3 figures show group velocity dispersion(left) and 1-D s-wave velocity model(right) for station pairs. Where SL(Shivalik), LH(Lesser Himalaya) and HH(High Himalaya) and TH(Tethys Himalaya)

4.1. Synthetic Test. Inversion models can be highly sensitive to measurement errors and other sources of noise, which can lead to inaccurate results**.** To qualify the accuracy of my Inversion results I performed 2 synthetic test (Forward modeling). (a) True model with LVL (b) True model with no LVL (see figure 14). In this synthetic test after performed forward modeling I added 5 % noise in dispersion data and then inverted the dispersion curve so in (a) true model with LVL shows LVL in inversion but in (b) true model with no LVL does not show any LVL. By comparing the inversion results with the forward modeling results, demonstrate the accuracy, validity, and robustness of our inversion method. This provides a strong justification for our inversion results that the LVL in inversion is not an artifact of inversion.



(a)



(b)

Figure(14) : Shows synthetic test (forward modeling) (a) true model with LVL (b) True model with no LVL.

5. Discussion.

4.1. Low-Velocity Layer (LVL).

Now I have questions about what are the factors which cause LVL. The low velocity of 3-3.2 km/s observed in crustal solid rock (melt-free) is not possible. Christensen et al. (1996) performed an experiment with different crustal rocks at the pressure of 1000 MPa (~lithostatic pressure at 30 km depth) and room temperature and measured the lowest shear wave velocities. The lowest shear wave velocity range in metamorphic rocks is 3.51-3.66 km/s and in plutonic rock (granite–granodiorite) the lowest velocity measured is 3.73 km/s. Using the summary of (Jamieson et al,.2004) I can conclude that the temperature of metamorphic rocks of the Greater Himalayan Sequence(GHS) in central Nepal lies between 500-800 ℃ but the temperature of the majority of rocks lies between 600-700 ℃. At the core of the orogenic wedge just above the MHT, at a depth of roughly 40 km near the Indus Tsangpo suture, coupled thermal-mechanical models predict present peak crustal temperatures of 650° to 900° that would decline southward (Beaumont et al., 2004). IfIdecrease shear wave velocity (V_s) to 0.2m/s/℃. and try to extrapolate V_s to mid crustal temperature, I don't get enough reduction in V_s as I observed. So around 25-30 km depth at mid-crustal temperature and pressure melt-free metamorphic rocks or plutonic rocks can’t be the reason for low velocity. So the most reasonable explanation of LVL is the presence of partial melt / aqueous fluid or both because both of them can easily produce a reduction in velocity that I observed.

Now another question raised is what type of fluids is causing LVL, partial melt / aqueous fluid, or a combination of both.

Lachenbruch and Sass (1977) also measured temperatures at a 30 km depth that is 650–850 °C. This result matches with Jamieson et al. (2004) and Beaumont et al. (2004) for GHS rocks discussed above. These temperatures are higher than the critical temperature for pelitic wet solidus. Thompson and Connolly (1995) indicate that if aqueous fluid is present then there is most probably partial melt also present because water reduces the melting temperature of rocks. At this temperature partial melt most likely does not exist because dry crustal rock starts melting at a temperature >900 ℃. The low-resistivity magnetotelluric anomalies found in the NW Indian Himalaya and in Tibet and partial melting offered the most possible explanation (Unsworth et al., 2005, Arora et al.,2007).

Rosenberg and Handy (2005) experimental results show that there is a drop in the magnitude of rock's strength that occurs between 0 and 7% melt. Based on MT measurements Unsworth et al. (2005) discovered 2–4% in NW India and 5–14% melt in southern Tibet. But our observation shows more velocity reduction in HH and TH and less in LH and SL that indicates if there is melting then NW India has more melt than southern Tibet.

---

6. Conclusions

We process ambient noise waveforms of 15 broadband stations in the northwest Himalayas and obtain surface wave dispersion between 5 and 55 sec. The inversion of the dispersion curve provides a one-dimensional crustal shear wave velocity model. I Observe a sequential reduction in mid-crustal velocities (Vs < 3.4 km/s) as I go from LH to TH, which shows that there is the existence of a LVL. The most probable reason to cause this LVL is either the aqueous fluid or partial melt or a combination of both.

7. References

- Arora, B.R., Unsworth, M.J., Rawat, G., 2007. Deep resistivity structure of the northwest Indian Himalaya and its tectonic implications. Geophysical Research Letters 34, L04307. doi:10.1029/2006GL029165.

- Aster, R. C., Borchers, B., & Thurber, C. H. (2018). Parameter estimation and inverse problems. Elsevier.

- Beaumont, C., Jamieson, R.A., Nguyen, M.H., Medvedev, S., 2004. Crustal channel flows: 1. Numerical models with applications to the tectonics of the Himalayan-Tibetan orogen. Journal of Geophysical Research 109, B06406. doi:10.1029/2003JB002809.

- Bensen, G. D., Ritzwoller, M. H., Barmin, M. P., Levshin, A. L., Lin, F., Moschetti, M. P., ... & Yang, Y. (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophysical journal international*, *169*(3), 1239-1260.

- Caldwell, W.B., Klemperer, S.L., Rai, S.S. and Lawrence, J.F., 2009. Partial melt in the upper-middle crust of the northwest Himalaya revealed by Rayleigh wave dispersion. Tectonophysics, 477(1-2), pp.58-65.

- Christensen, N.I., 1996. Poisson's ratio and crustal seismology. Journal of Geophysical Research 101 (B2), 3139–3156.

- DiPietro, J.A. and Pogue, K.R., 2004. Tectonostratigraphic subdivisions of the Himalaya: A view from the west. Tectonics, 23(5).

- Gilligan, A. and Priestley, K., 2018. Lateral variations in the crustal structure of the Indo–Eurasian collision zone. Geophysical Journal International, 214(2), pp.975-989.

- Herrmann, R.B., Ammon, C.J., 2002. Computer Programs in Seismology: Surface Waves. Receiver Functions and Crustal Structure, St. Louis University, St. Louis, MO.

- Jamieson, R.A., Beaumont, C., Medvedev, S., Nguyen, M.H., 2004. Crustal channel flows: 2. Numerical models with implications for metamorphism in the HimalayanTibetan orogen. Journal of Geophysical Research 109, B06407. doi:10.1029/ 2003JB002811.

- Kumar, V., Rai, S. S., Hawkins, R., & Bodin, T. (2022). Seismic Imaging of Crust Beneath the WEstern Tibet‐Pamir andIstern Himalaya Using Ambient Noise and Earthquake Data. Journal of Geophysical Research: Solid Earth, 127(6), e2021JB022574.

- Lachenbruch, A.H., Sass, J.H., 1977. Heat flow in the United States and the thermal regime of the crust. In: Heacock, J.G. (Ed.), The Earth's Crust: Its Nature and Physical Properties. American Geophysical Union Monograph, vol. 20, pp. 626–675. Washington, DC.

- Makovsky, Y. and Klemperer, S.L., 1999. Measuring the seismic properties of Tibetan bright spots: Evidence for free aqueous fluids in the Tibetan middle crust. Journal of Geophysical Research: Solid Earth, 104(B5), pp.10795-10825.

- Monsalve, G., Sheehan, A., Rowe, C. and Rajaure, S., 2008. Seismic structure of the crust and the upper mantle beneath the Himalayas: Evidence for eclogitization of lower crustal rocks in the Indian Plate. Journal of Geophysical Research: Solid Earth, 113(B8).

- Nábělek, J., Hetényi, G., Vergne, J., Sapkota, S., Kafle, B., Jiang, M., Su, H., Chen, J. and Huang, B.S., 2009. Underplating in the Himalaya-Tibet collision zone revealed by the Hi-CLIMB experiment. Science, 325(5946), pp.1371-1374.

- Nakata, N., Gualtieri, L., & Fichtner, A. (Eds.). (2019). *Seismic ambient noise*. Cambridge University Press.

- Nishida, K. (2017). Ambient seismic wave field. *Proceedings of the Japan Academy, Series B*, *93*(7), 423-448.

- Owens, T.J. and Zandt, G., 1997. Implications of crustal property variations for models of Tibetan plateau evolution. Nature, 387(6628), pp.37-43.

- Rai, S.S., Padhi, A. and Sarma, P.R., 2009. High crustal seismic attenuation in Ladakh–Karakoram. Bulletin of the Seismological Society of America, 99(1), pp.407-415.

- Rai, S.S., Priestley, K., Gaur, V.K., Mitra, S., Singh, M.P. and Searle, M., 2006. Configuration of the Indian Moho beneath the NW Himalaya and Ladakh. Geophysical Research Letters, 33(15).

- Rosenberg, C.L., Handy, M.R., 2005. Experimental deformation of partially melted granite revisited: Implications for the continental crust. Journal of Metamorphic Geology 23 (1), 19–28. doi:10.1111/j.1525-1314.2005.00555.x.

- Sapin, M. and Hirn, A., 1997. Seismic structure and evidence for eclogitization during the Himalayan convergence. Tectonophysics, 273(1-2), pp.1-16.

- Schelling, D. and Arita, K., 1991. Thrust tectonics, crustal shortening, and the structure of the far‐eastern Nepal Himalaya. Tectonics, 10(5), pp.851-862.

- Thompson, A.B., Connolly, J.A.D., 1995. Melting of the continental crust: some thermal and petrological constraints on anatexis in continental collision zones and other tectonic settings. Journal of Geophysical Research 100 (B8), 15,565–15,579.

- Unsworth, M.J., Jones, A.G.,Ii, W., Marquis, G., Gokarn, S.G., Spratt, J.E., 2005. Crustal rheology of the Himalaya and Southern Tibet inferred from magnetotelluric data. Nature 438 (7064), 78–81. doi:10.1038/nature04154.

- Yang, Y., Ritzwoller, M.H., Zheng, Y., Shen, W., Levshin, A.L. and Xie, Z., 2012. A synoptic view of the distribution and connectivity of the mid‐crustal low velocity zone beneath Tibet. Journal of Geophysical Research: Solid Earth, 117(B4).