Ambient Noise Seismic Tomography

A DISSERTATION REPORT

Submitted in partial fulfillment of the

requirement for the award of the degree

OF

INTEGRATED MASTER OF TECHNOLOGY

IN

GEOPHYSICAL TECHNOLOGY

BY

ANIMESH ANAND

Enrollment Number: 11411007

DEPARTMENT OF EARTH SCIENCES

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

ROORKEE – 247667 (INDIA)

MAY, 2016

CANDIDATE'S DECLARACTION

I hereby declare that the work which is presented in this dissertation report titled ***"Ambient Noise Seismic Tomography***" in partial fulfillment of the requirements for the award of the degree of ***"INTEGRATED MASTER OF TECHNOLOGY"*** in ***GEOPHYSICAL TECHNOLOGY***, submitted to the Department of Earth Sciences, Indian Institute of Technology Roorkee is an authentic record of my own work carried out during the period from June 2015 to April 2016 under the supervision of Dr. Kamal, Associate Professor, Department of Earth Sciences, Indian Institute of Technology Roorkee and Dr. S. C. Gupta, Scientific Officer, Department of Earthquake Engineering, IIT Roorkee.

I have not submitted the matter embodied in this dissertation report for the award of any other degree.

Date: **Animesh Anand** (11411007)

Place: IIT Roorkee Integrated M. Tech (Geophysical Technology)

CERTIFICATE

This is to certify that the above statement made by the candidate is correct to the best of my knowledge and belief.

I wish him a great success in life and bright future.

Date:

Place: IIT Roorkee

Signature of co-guide	**Signature of guide**
Dr. S. C. Gupta	Dr. Kamal
Scientific Officer	Associate Professor
Department of Earthquake Engineering	Department of Earth Sciences
IIT Roorkee, India	IIT Roorkee, India

CERTIFICATE

I …..**ANIMESH ANAND**.… hereby solemnly declare that the dissertation entitled ……**AMBIENT NOISE SEISMIC TOMOGRAPHY**…… being submitted by me towards the partial fulfillment of the requirements for the award of ....**INTEGRATED M.TECH. GEOPHYSICAL TECHNOLOGY**…. degree is a record of my own work and that I have not copied the work of any other person(s) including published literature and material from any web site.

Where ever the work of other person(s) has been used, it has been duly acknowledged and quoted with proper reference to the original work.

I fully understand the implications of plagiarism and that if at any stage the above statement made by me is found to be incorrect; I shall be fully responsible for my act(s).

(Signature ……………………………………………...)

Name …………….ANIMESH ANAND…………...

Place……ROORKEE……….

Date………………………….

ACKNOWLEDGEMENT

I wish to express my deep sense of gratitude and sincere indebtedness to my supervisors **Dr. Kamal**, Associate Professor, Department of Earth Sciences, Indian Institute of Technology Roorkee; **Dr S. C. Gupta**, Scientific Officer, Department of Earthquake Engineering, IIT Roorkee, for their kind cooperation and encouragement that they have provided me for developing new ideas and tackle the various situations which came across while doing the dissertation work. Their undying determination to get the best out of their students served as inspiration for completion of this report.

In addition to the **Department of Earth Sciences, IIT Roorkee** and **Department of Earthquake Engineering, IIT Roorkee**and the **administration** whose help and resources I have used from time to time, I thank the brains behind the scientific publications, earth science books and research papers whose guidance I have drawn on to make this report.

I am also thankful to all my friends for their continuous support and help during different stage of my dissertation work.

Lastly, I will be always grateful to my parents whose blessings are always with me and acts as a beacon in finding the right path.

Well, I must say at last that it was all God's grace that I gained so much of valuable knowledge and experience from completion of this dissertation work.

ANIMESH ANAND

CONTENTS

List of figures

ABSTRACT

Ambient noise can be used to generate tomographic images that give us an idea of various geological features prevalent in the subsurface. Ambient noise is used to generate shallow depth images. Such images can be used to infer about the presence of groundwater aquifers, or hydrocarbon traps, formed by faults and fractures. Faults hold lots of academic and economic significance. They can also act as structural traps for accumulation of hydrocarbons or can behave as migration pathways aiding the escape of hydrocarbons.

Ambient noise can be low frequency or high frequency. Low frequency noise are usually generated by ocean waves, and penetrate to depths of about 10 km. High frequency noise are generated by various cultural noise, etc. Here, I will be using high frequency noise that will be used as diffused wave fields. Such wave fields can be used to generate Greens function by cross correlating them with each other. The Green function can be used to give us an idea about the velocity structure of the region.

In order to achieve the above mentioned objective, various methods are prevalent. In my thesis, I will be working on the method given by *G. D. Bensen et al., 2007*. I used the concepts mentioned in the paper to generate MATLAB codes to generate cross correlated data and subsequent Green function. I have used noise data from five stations spread around Garhwal region. The data used is uploaded on an hourly basis, for a period of twenty-four hours. The date of the recording of the data is 30th December, 2007. The station names from where the data is recorded are: RAJGA, KHURM, AYARC, SRIKO, and CNTTE. After obtaining the Green function from these stations, I applied FTAN analysis using python package to obtain the dispersion curves.

CHAPTER 1

INTRODUCTION

1.1What is ambient noise seismic tomography?

In layman terms, tomography can be said to be the process by which one can know about the internal unseen characteristics of a body. The process to be applied in doing so should be to study the path of different rays, be it seismic rays or X rays, as it travels through the body. In case of seismic tomography, we use mathematical tools to study the path and the behaviour of P and S waves. We also calculate their travel times to know how their path has been affected by various undulations present within the earth. In case of ambient noise seismic tomography, the source that is used can be of varied origins. It can either be noise that is caused due to various cultural reasons, like people passing by, or vehicular noise, etc. It can also be noise generated by ocean waves that constitute low frequency noise. All these types of noise are recorded on seismograms after they travel through the earth. Using such a technology is similar to the one used by medical scientists while operating a CAT-scan. Computer assisted tomography (CAT) uses X-rays that travel through the body in many different directions. We then use mathematical tools to study different parts of the body. This study is based on different absorptive power of different organs of the body. The comparison between CAT-scans and seismic tomography differs because X-rays travel in straight paths, whereas the ray paths of sound waves bend with changes in the velocity structure of the medium.

1.2 What are the uses of ambient noise seismic tomography?

Ambient Noise Seismic tomography has several applications in exploration and global geophysics. Ambient Noise tomography produces images that are of shallow depth and are primarily used to identify groundwater aquifers, or shallow depth hydrocarbon traps, such as faults or fractures. It can also be used to characterize locate ore bodies. Ambient Noise Tomography opens up new possibilities in imaging the subsurface at many different scales and complements traditional passive source tomography.

1.3 What are the limitations of seismic tomography?

Ambient Noise data has certain limitations. Such limitations include prerequisite for small interstation distance. This is a requirement because a certain noise recorded at a certain station may not at all get recorded at some other station if the other station is far away. This will result in inaccurate data interpretation.

1.4 How does it work?

Ambient noise data processing comprises three major steps. The first of them involves preparing single station data. Secondly, we cross correlate the noise data to obtain green function. Then we stack the cross correlated data. Thirdly, we apply FTAN analysis to obtain the dispersion curves.

The cross correlation is done between a pair of stations and it returns the Green function corresponding the area between the two stations. For example, in the figure given below, the cross correlation between the two noisy waveform will give us the Green function corresponding to the region in between the two detectors. Similarly, if we take a number of stations, let’s say ‘n’, then we will have n(n-1)/2 number of pairs, and the same number of cross correlations

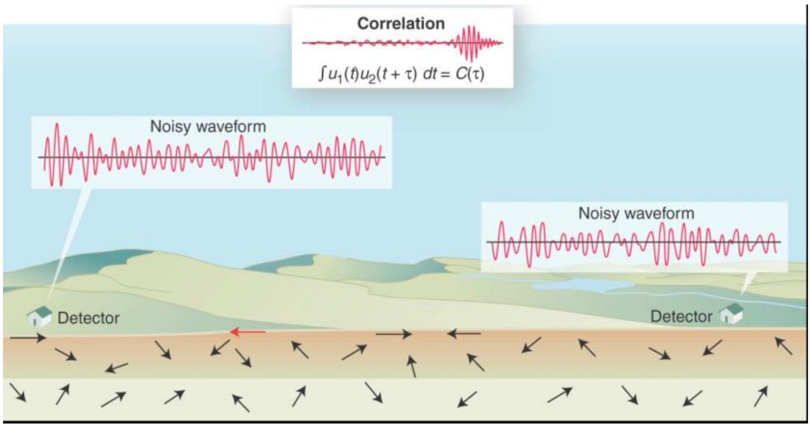


Figure 1.4: A schematic diagram showing how cross correlation between two noisy waveforms result in estimation of Greens function

The formula for cross correlation is as follows:

$$C(\tau) = \int u_1(t)u_2(t+\tau)dt$$

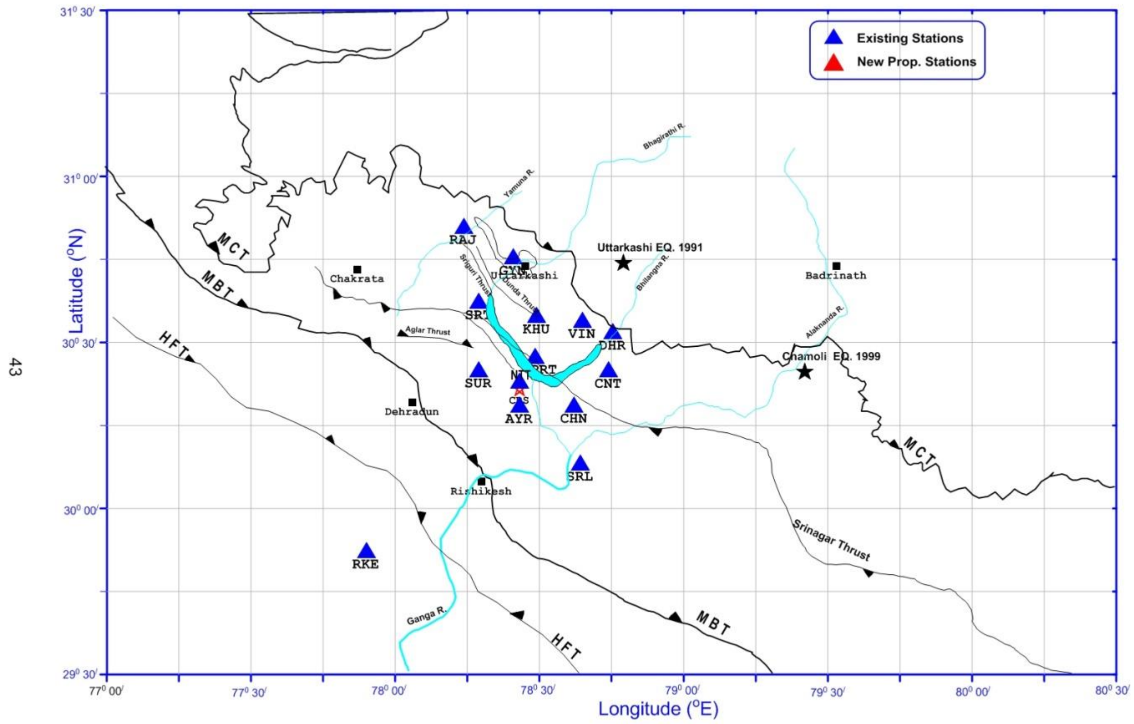
The above cross correlation is done in frequency domain. After we have obtained the cross correlated data, we apply inverse Fourier transform to obtain the data in time domain. Then, we apply stacking, in which the different time series are added together. After stacking, we apply FTAN, which is a modified form of a python package. FTAN returns dispersion curves.

CHAPTER 2

METHODOLOGY

The data that I have used here has been recorded in five seismic stations in the Garhwal region. The data is recorded on an hourly basis, with the sample rate of 100 samples per second. Thus, in an hour, there are about 360000 samples from one station. As I have taken data from five stations, so there are ten pairs of station that needs to be cross correlated.

Here is the map showing the locations of the stations



As discussed previously, the whole methodology comprises three major steps:

1) Preparation of data from single station
2) Cross correlation, and Stacking
3) Dispersion Curves using FTAN

2.1: Preparation of data from single station

Firstly, we will look into the first step. Preparation of data from single station involves the pre-processing steps. Such steps include de-meaning, de-trending, temporal whitening, and spectral whitening. The main aim of this step is to remove event data, i.e., earthquake data from the noise data. The earthquake data maligns the noise data and obscures desired result.

De-meaning and de-trending aligns the start of the plot with the zero line. Both these steps are done using MATLAB. Along with de-meaning and de-trending, instrument response is also removed, as it also obscures noise data, and needs to be removed to obtain accurate results.

After de-meaning and de-trending, we need to apply temporal normalization. Temporal normalization also helps in reducing the effect of earthquake in the cross correlated data as well as the effect of instrument irregularities. There are several methods to apply temporal normalization. These include:

1) One-bit normalization
2) A clipping threshold application
3) Automatic detection of events and their subsequent removal
4) Running absolute mean normalization

Out of the above mentioned methods, the one that is used here is one-bit normalization. It is a pretty aggressive method. In this method, all the positive amplitudes of the signal are converted to '1', and all the negative amplitude of the signal are converted to '-1'. So, basically it divides each sample in the signal with its absolute value.

After temporal normalization, comes spectral whitening. Spectral whitening is used to remove the peaks in the ambient noise data in the frequency domain. These peaks arise due to various reasons, primarily due to earth hums.

2.2: Cross correlation and Stacking

Now, after we are done with the pre-processing steps, we advance to the next step, i.e., cross correlation and stacking. As mentioned above, I have used data from five stations, which result in ten pairs. Thus there are ten cross correlations each hour for a period of twenty-four hours. The cross correlation is done in the frequency domain. So, Fourier transform is applied to the data to convert it into frequency domain. Now, for each hour, we apply cross correlation for ten pair of stations. So, in all, we get two hundred and forty cross correlated data sets.

After the cross correlation, the data is returned to temporal domain by applying inverse Fourier transform. Now, we apply stacking to the corresponding pairs of stations. For example, the cross correlated data for station 1 and station 2 are stacked for the twenty four hour period. Similarly, we work upon the rest of the nine pairs.

2.3: FTAN analysis

After we are done with cross correlation and stacking, what we are left with is the Greens function corresponding to the region. Now, we can use this Greens function to calculate group speed and phase speed. These two speeds are obtained as a function of time. This process is accomplished by applying frequency-time analysis (FTAN). Here, I have used the version of FTAN that is explained by *Levshin et al. (1989).* In FTAN analysis, first of all we obtain the analytic signal. Let's say, we have the original waveform as 'u (t)'. Now, the Fourier transform of this waveform will be:

$$U(\omega) = \int_{-\infty}^{\infty} u(t) \exp(i\omega t)\, dt.$$

Now, the analytic function, U_a is given by:

$$U_a(\omega) = U(\omega)(1 + sgn(\omega))$$

This analytic function is converted back to the time domain by applying inverse Fourier transform.

$$U_a(t) = u(t) + iH(t) = |A(t)| \exp(i\varphi(t))$$

Here, H(t) is the Hilbert transform of u(t). Now, the analytic signal is passed through a Gaussian filter. The mean frequency of this Gaussian filter is ω_0.

$$U_a(\omega, \omega_0) = U(\omega)\big(1 + sgn(\omega)\big)G(\omega - \omega_0)$$

$$G(\omega - \omega_0) = e^{-\alpha(\frac{\omega-\omega_0}{\omega_0})^2}$$

After obtaining the bandpassed function, we apply inverse Fourier transform to obtain an envelope function, $|A (t, \omega_0)|$, and a phase function, $\varphi (t, \omega_0)$. $|A (t, \omega_0)|$ is used to calculate group speed, whereas $\varphi (t, \omega_0)$ is used to calculate phase speed.

All this process of FTAN analysis can be done mathematically, taking each waveform separately. However, since in this case there is large amount of data, so the process needs to be automated. So, I have used a python package for FTAN analysis, that does all this work in an automated process, and returns a cleaned FTAN diagram.

As a recap, I would like to display a flowchart displaying the methods:

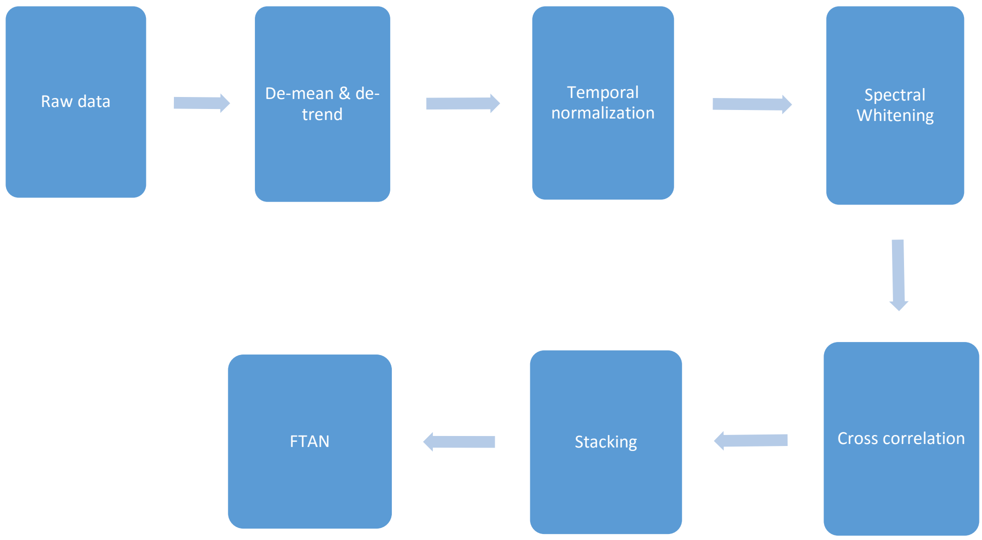


Figure 2.3: A schematic diagram showing the processes followed in this thesis

CHAPTER 3

RESULTS AND DISCUSSION

Before delving into the results, let us get through some of the raw data that I have used in my work. Since it is not possible to show all of them, so I will be showing data of a couple of stations that can be looked upon as a representative of all other data, since all of them are similar noise data.

3.1 Raw data

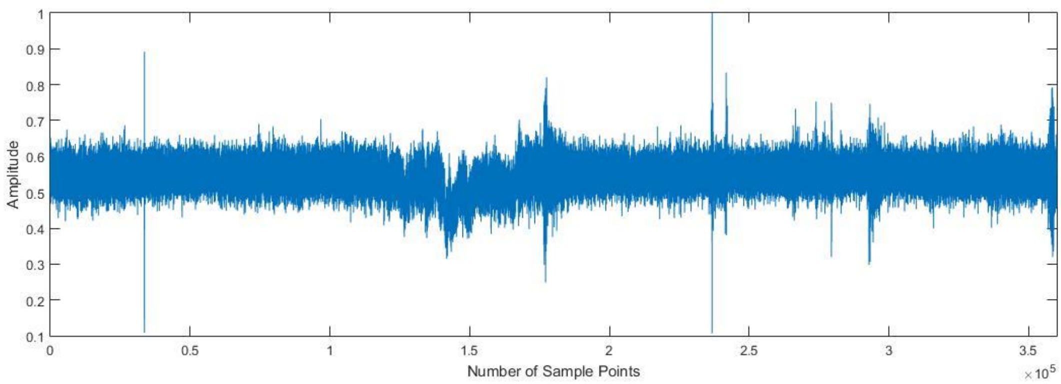


Figure 3.1(a): Raw data from the first station recorded for the first hour

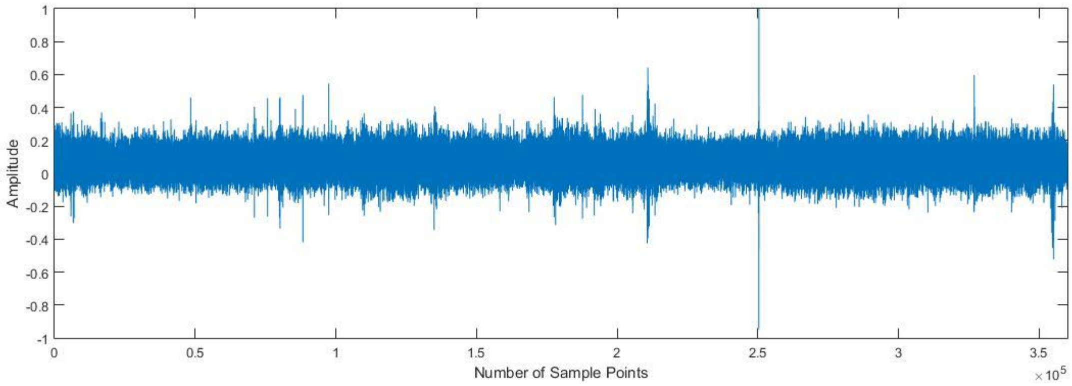


Figure 3.1(b): Raw data from the second station recorded for the first hour

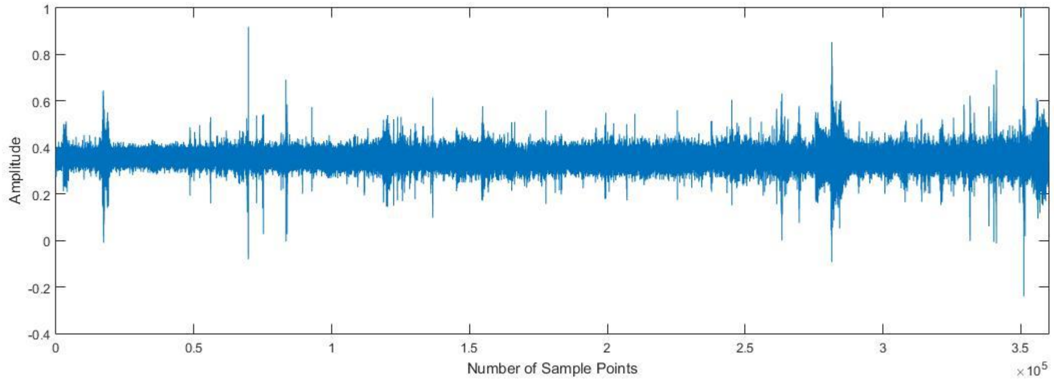


Figure 3.1(c): Raw data from the first station recorded for the second hour

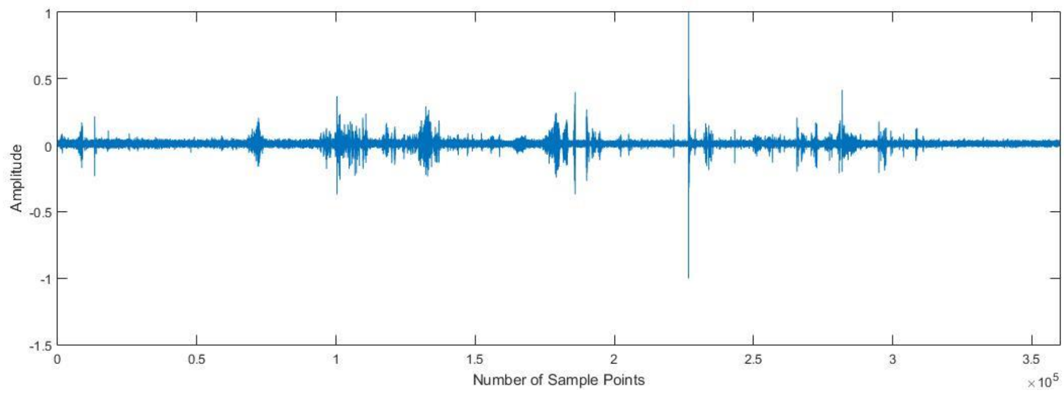


Figure 3.1(d): Raw data from the second station recorded for the second hour

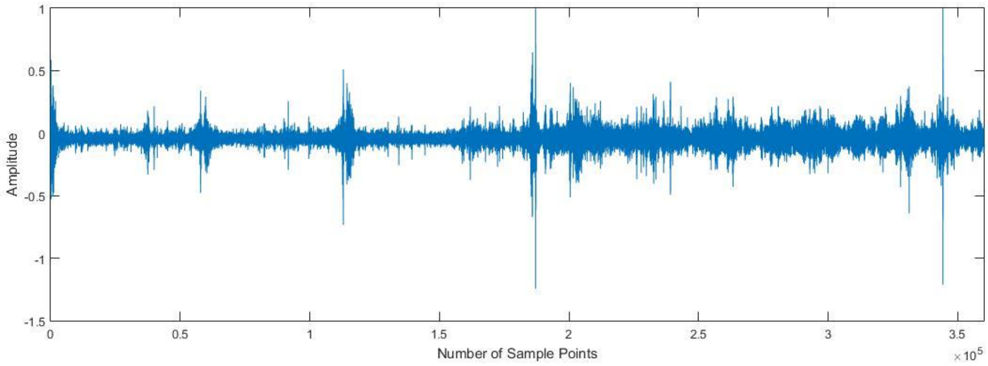


Figure 3.1(e): Raw data from the first station recorded for the third hour

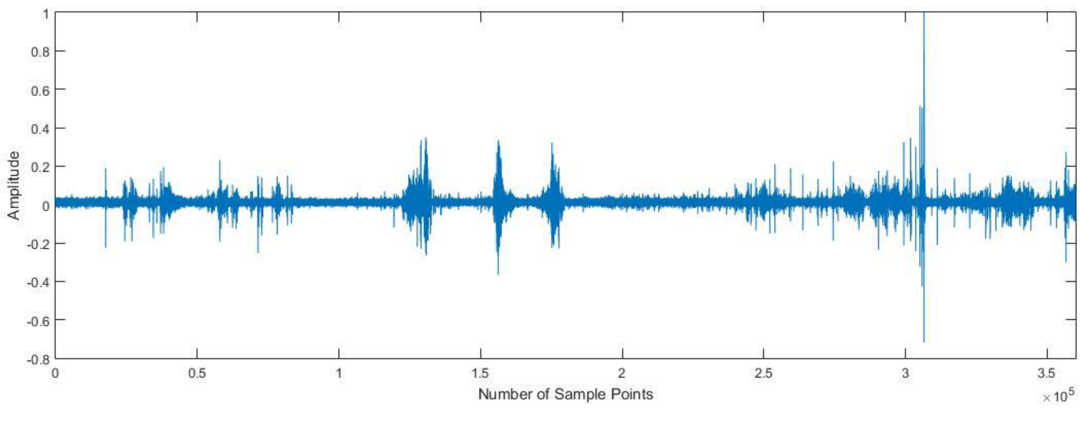


Figure 3.1(f): Raw data recorded from the second station for the third hour

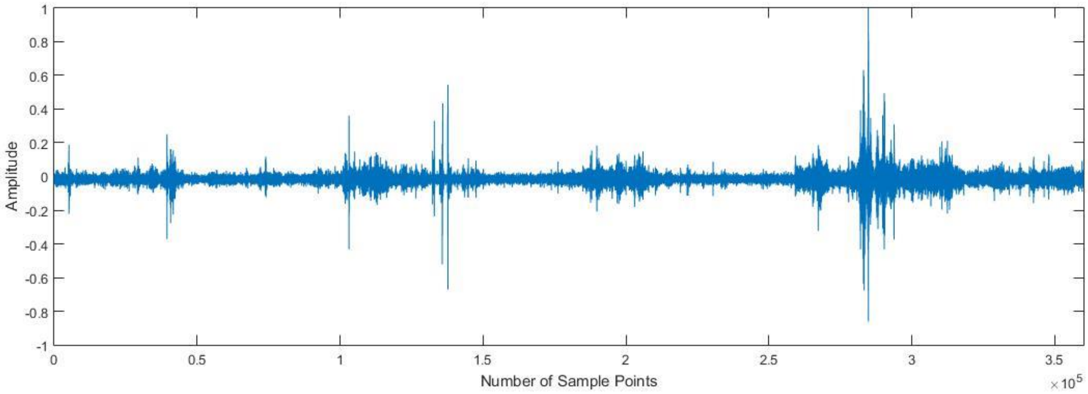


Figure 3.1(g): Raw data recorded from first station for fourth hour

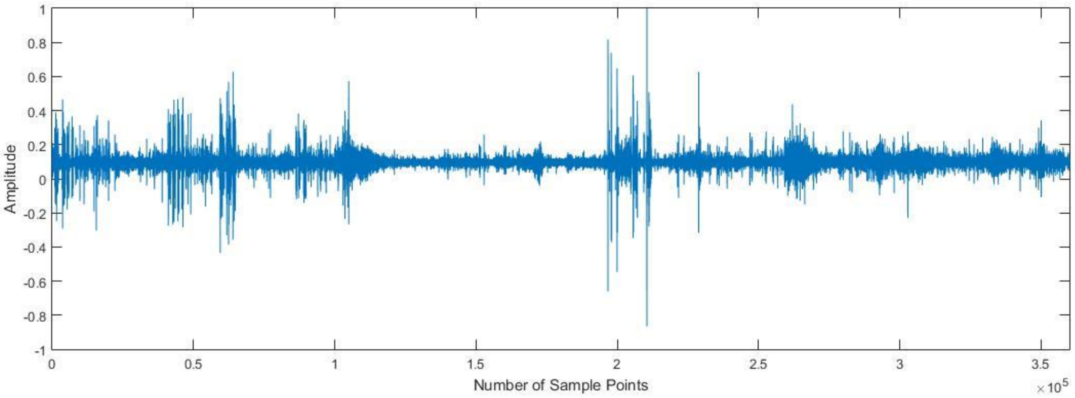


Figure 3.1(h): Raw data recorded from second station for fourth hour

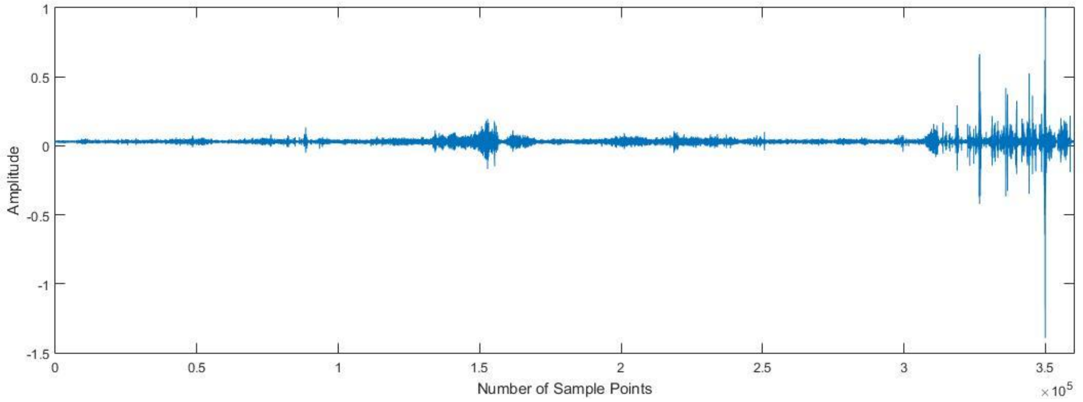


Figure 3.1(i): Raw data recorded from first station for fifth hour

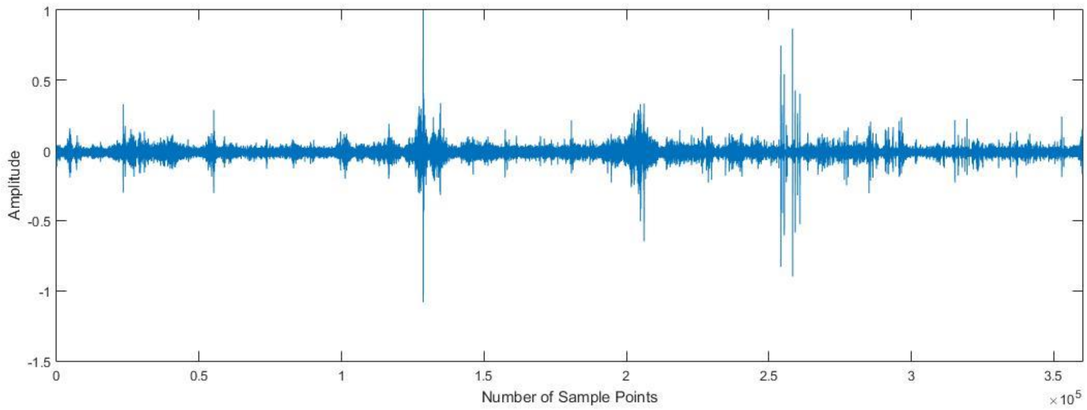


Figure 3.1(j): Raw data recorded from second station for fifth hour

3.2 Data after de-meaning and de-trending

The first step after collecting data is to de-mean and de-trend the data so as to align the start of the plot with the zero line. This is required so as to produce accurate results. Following plots are produced after the above shown raw data are made to go through the de-meaning and de-trending process.

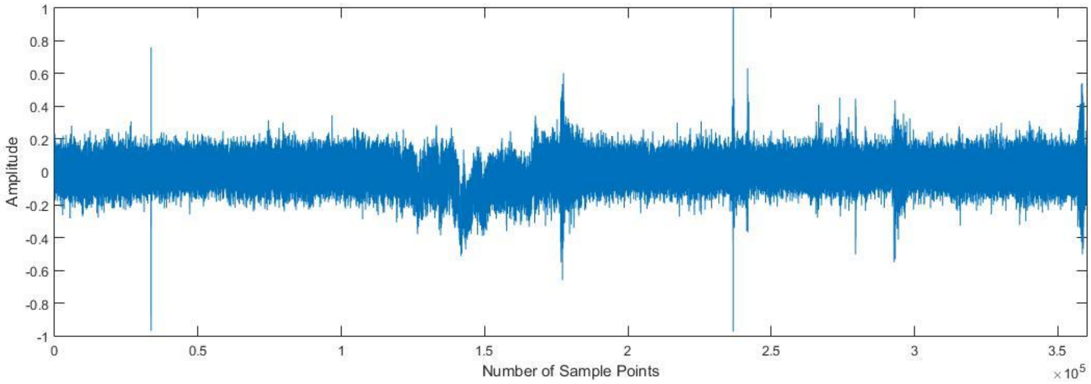


Figure 3.2(a): De-mean data from first station for first hour

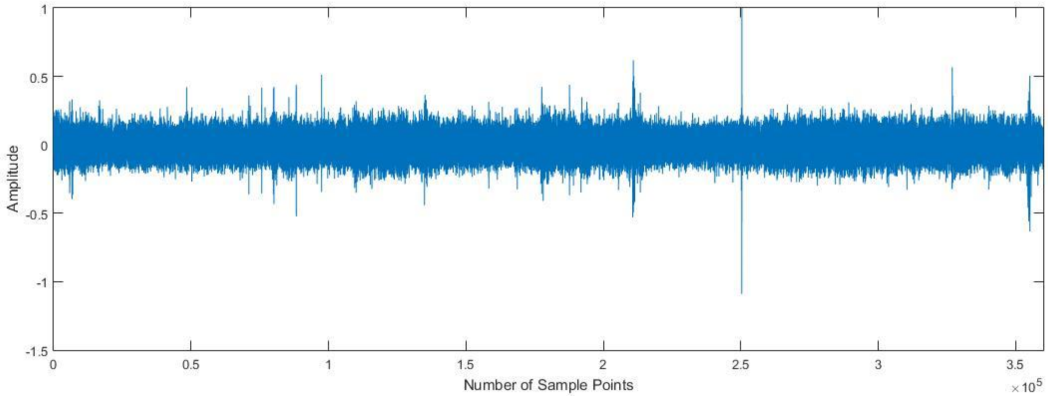


Figure 3.2(b): De-mean data from second station for first hour

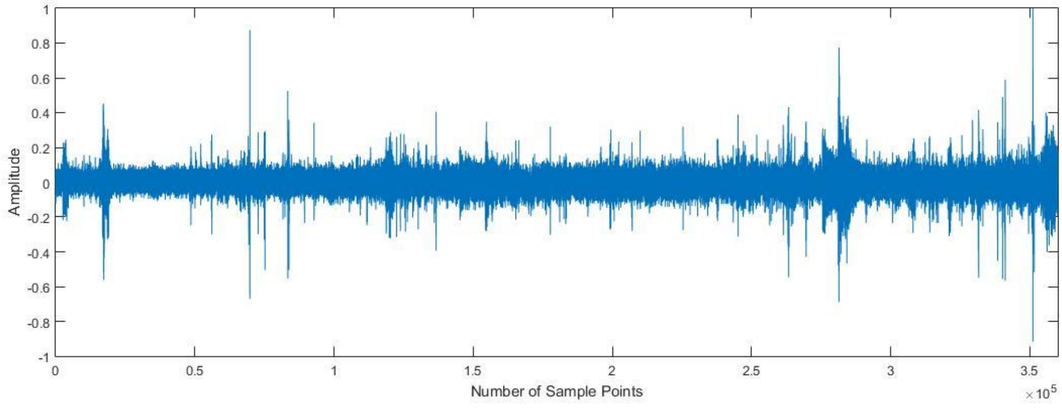


Figure 3.2(c): De-mean data from first station for second hour

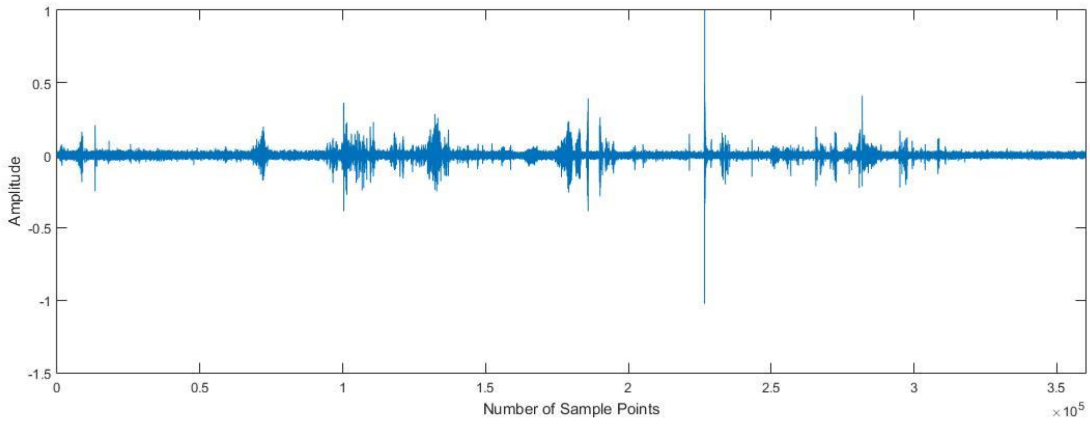


Figure 3.2(d): De-mean data from second station for second hour

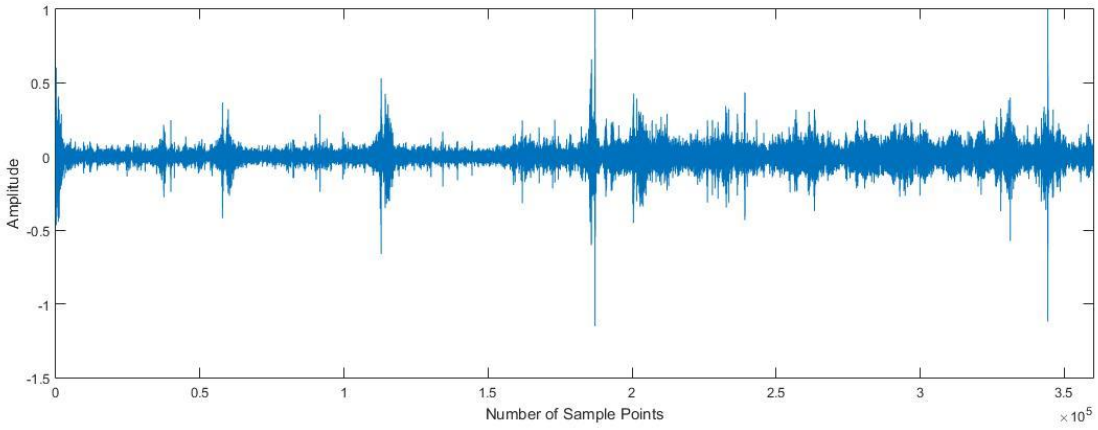


Figure 3.2(e): De-mean data from first station for third hour

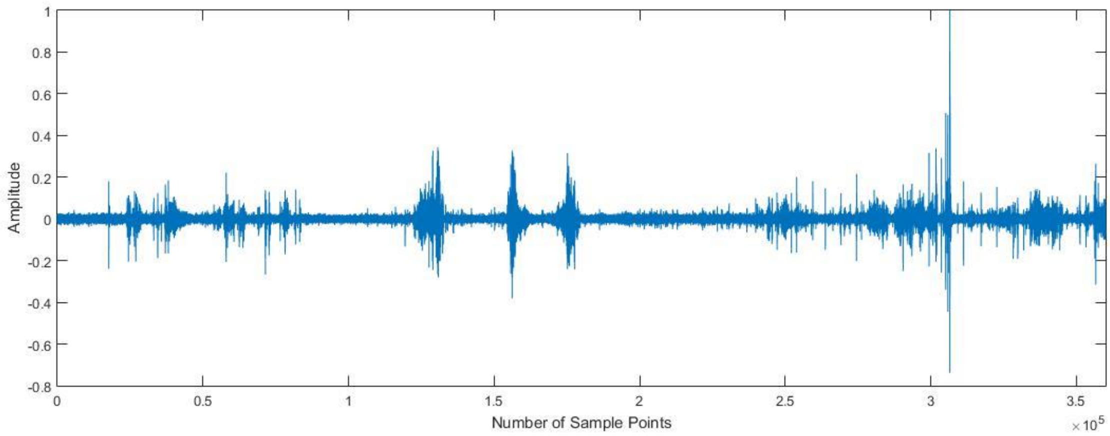


Figure 3.2(f): De-mean data from second station for third hour

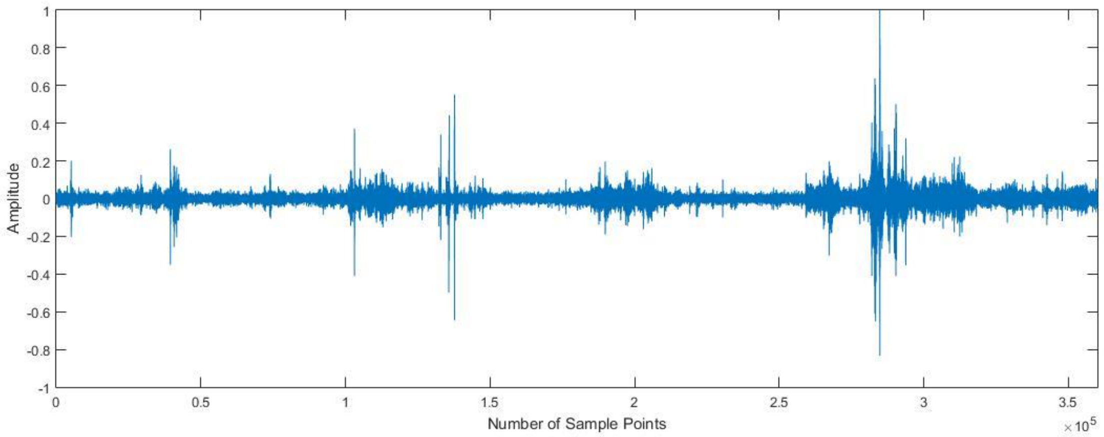


Figure 3.2(g): De-mean data from first station for fourth hour

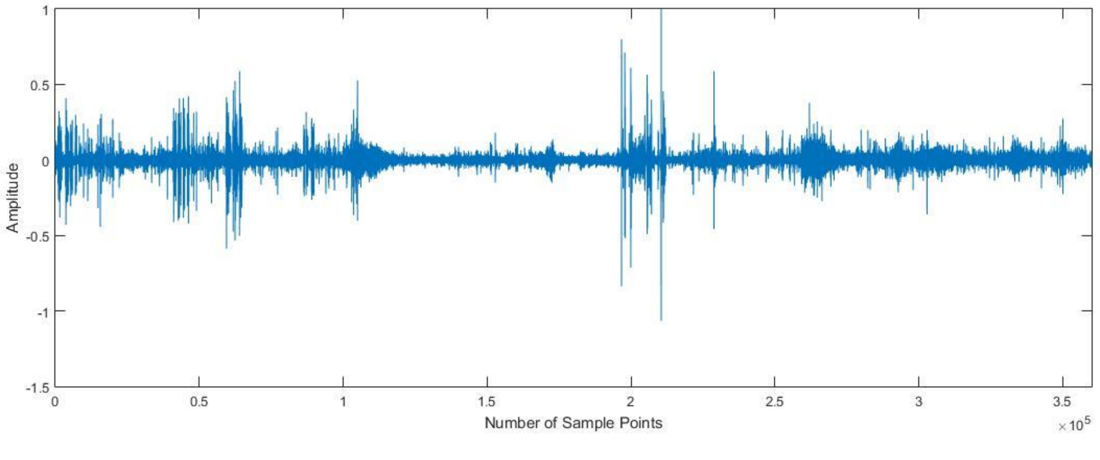


Figure 3.2(h): De-mean data from second station for fourth hour

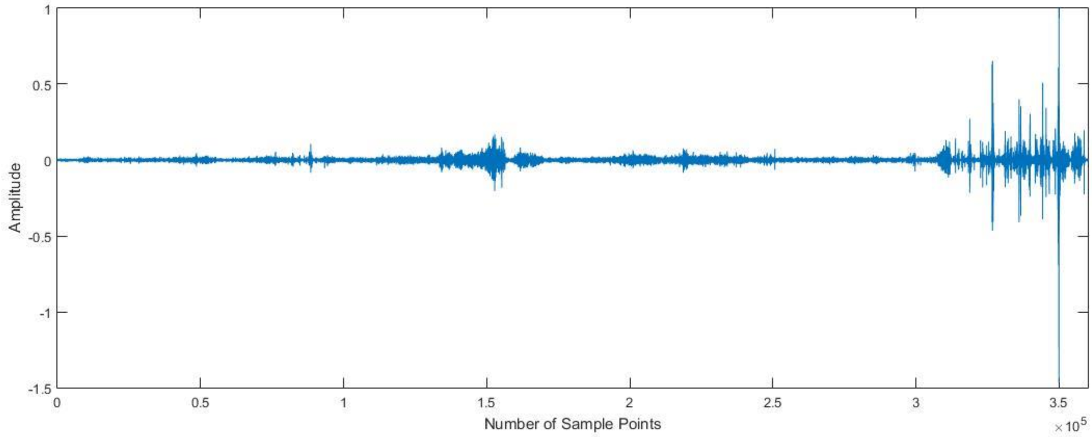


Figure 3.2(i): De-mean data from first station for fifth hour

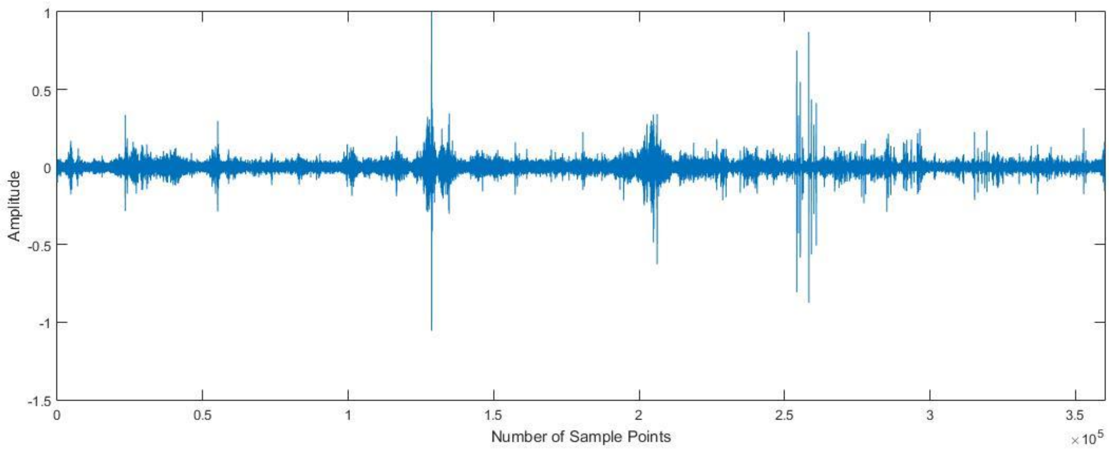


Figure 3.2(j): De-mean data from second station for fifth hour

3.3 Temporal normalization

After de-meaning and de-trending, it's time for temporal normalization. The method that I have used here, is called one-bit normalization. It converts all the positive values of the amplitude to '1', and all the negative values to '-1'.

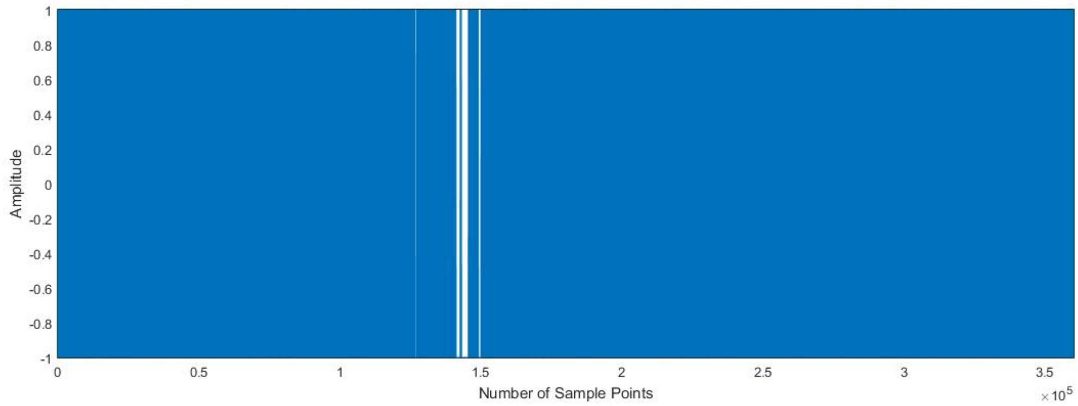


Figure 3.3: Normalized data from first station for first hour

Since all the plots after temporal normalization look more or less the same, so I will be showing just one plot for the temporal normalized data.

3.4 Cross Correlation

After temporal normalization and spectral whitening, the data is converted to frequency domain by applying Fourier transform. Now, we take stations on pair-basis and compute cross correlation for that pair. We do this process for each hour independently. Here, since I have data from five stations, so, there will be ten pairs of stations each hour. Taking into account twenty four hours of data that I am using, there will be a total of two hundred and forty sets of cross correlated data.

In the following plots, I will be showing cross correlation for stations one and two for different hours. Later on, we will be stacking them up in time domain.

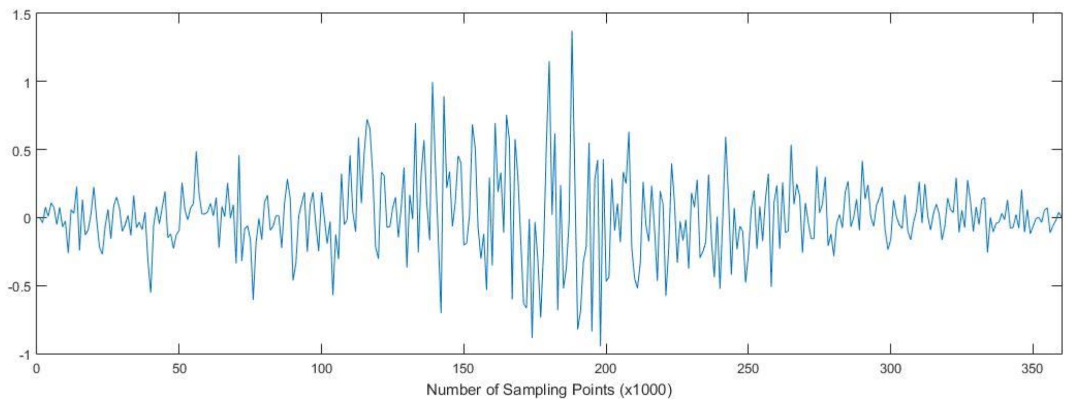


Figure 3.4(a): Cross correlated data between stations 1 and 2 for first hour

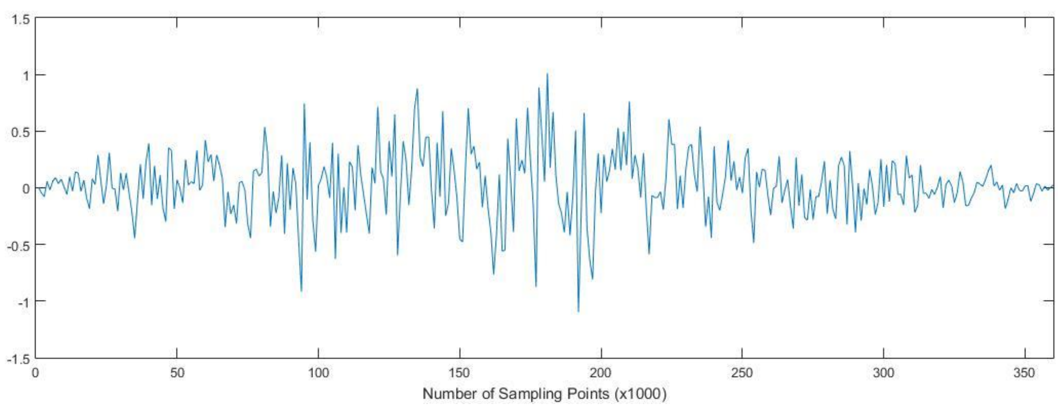


Figure 3.4(b): Cross correlated data between stations 1 and 2 for second hour

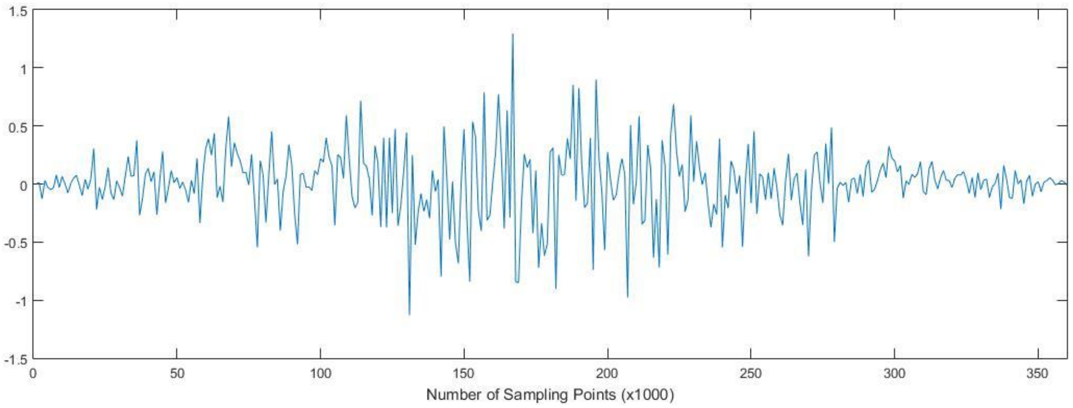


Figure 3.4(c): Cross correlated data between stations 1 and 2 for third hour

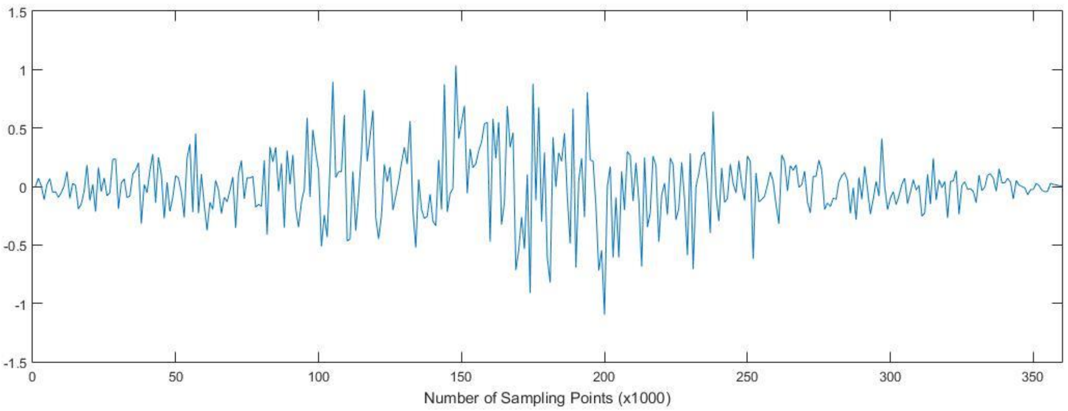


Figure 3.4(d): Cross correlated data between stations 1 and 2 for fourth hour

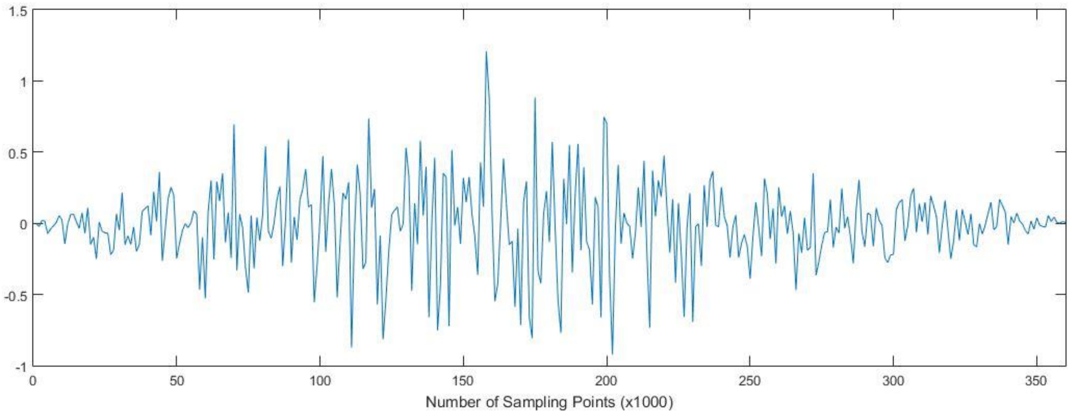


Figure 3.4(e): Cross correlated data between stations 1 and 2 for fifth hour

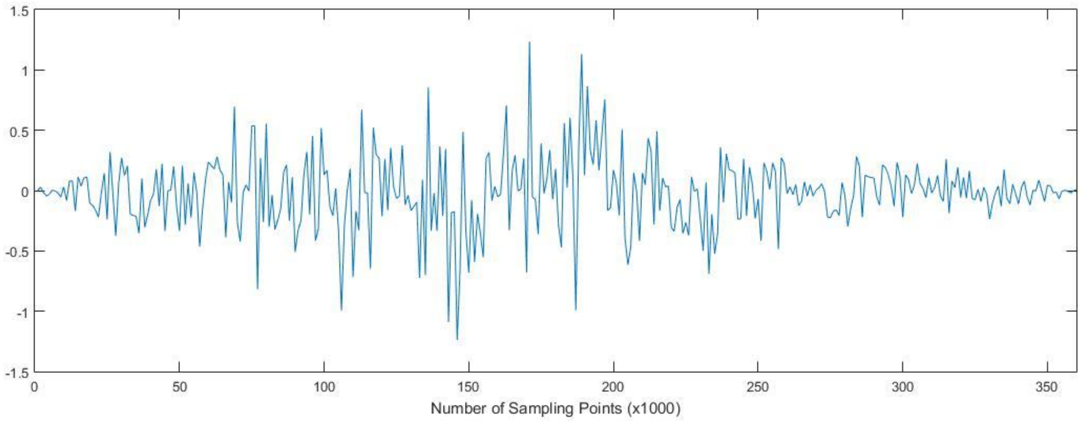


Figure 3.4(f): Cross correlated data between stations 1 and 2 for sixth hour

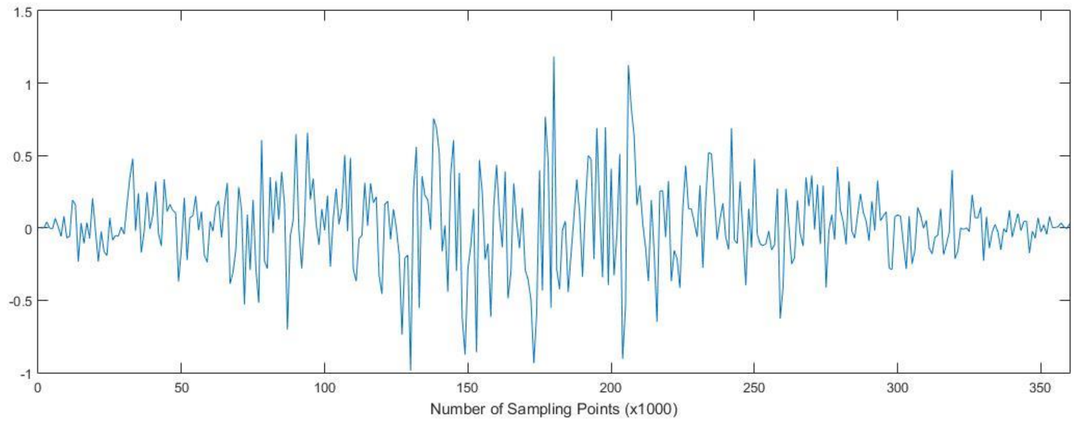


Figure 3.4(g): Cross correlated data between stations 1 and 2 for seventh hour

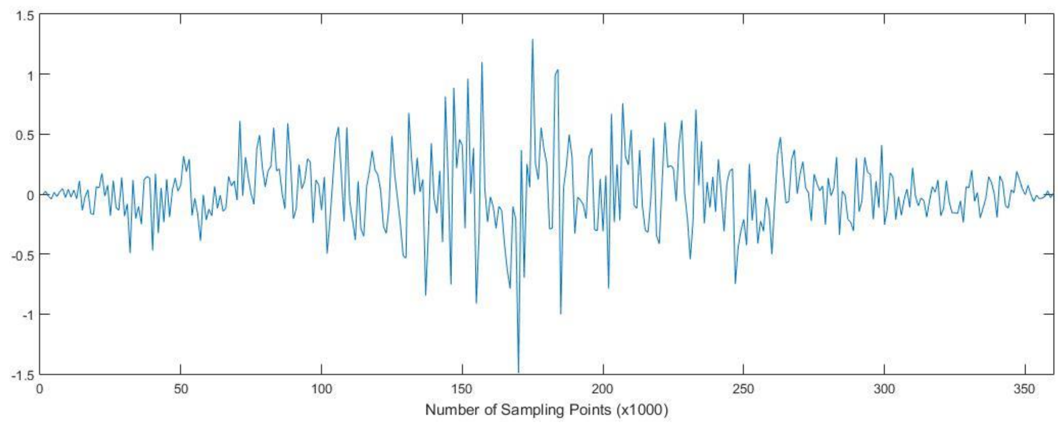


Figure 3.4(h): Cross correlated data between stations 1 and 2 for eighth hour

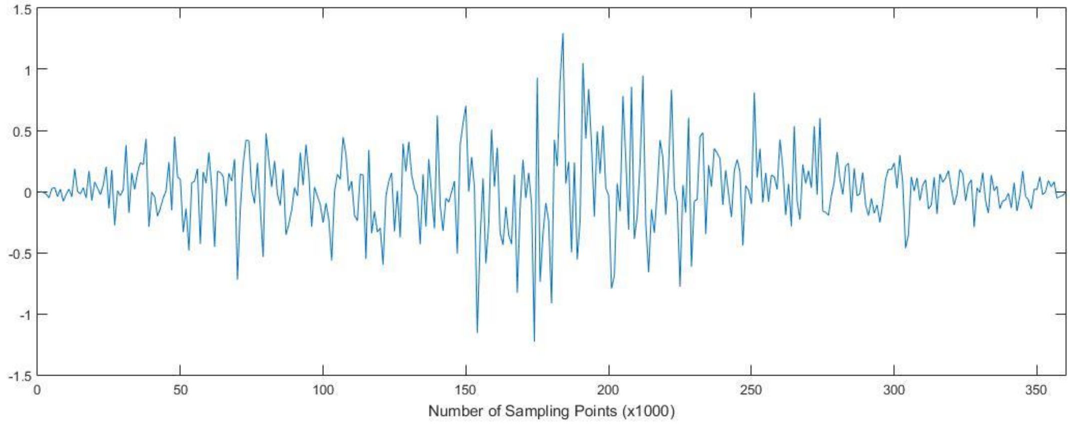


Figure 3.4(i): Cross correlated data between stations 1 and 2 for ninth hour

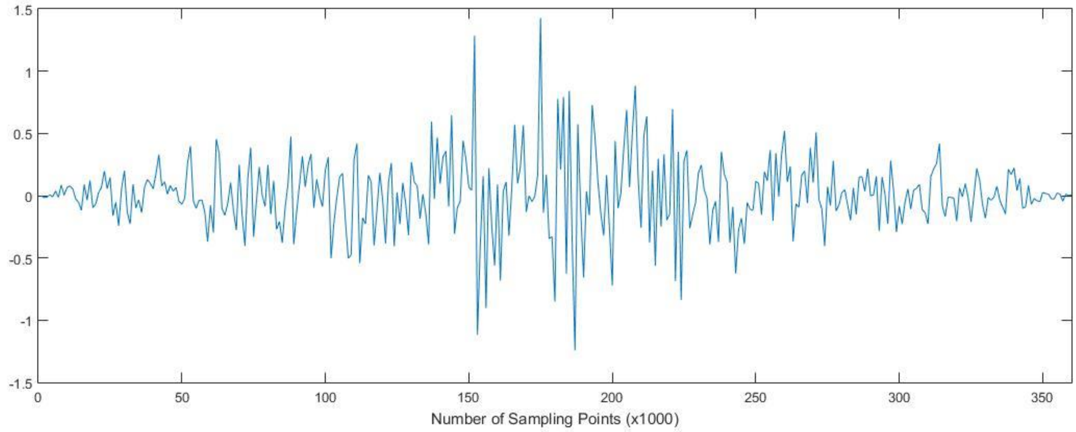


Figure 3.4(j): Cross correlated data between stations 1 and 2 for tenth hour

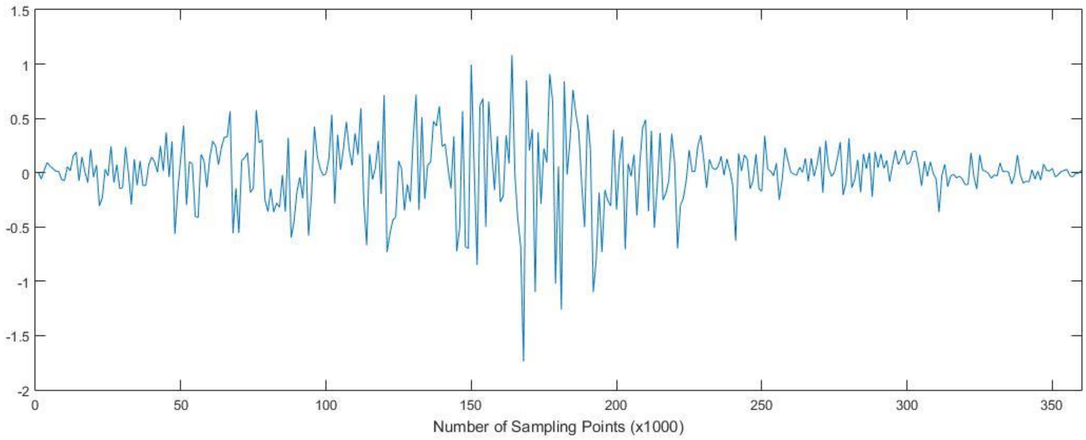


Figure 3.4(k): Cross correlated data between stations 1 and 2 for eleventh hour

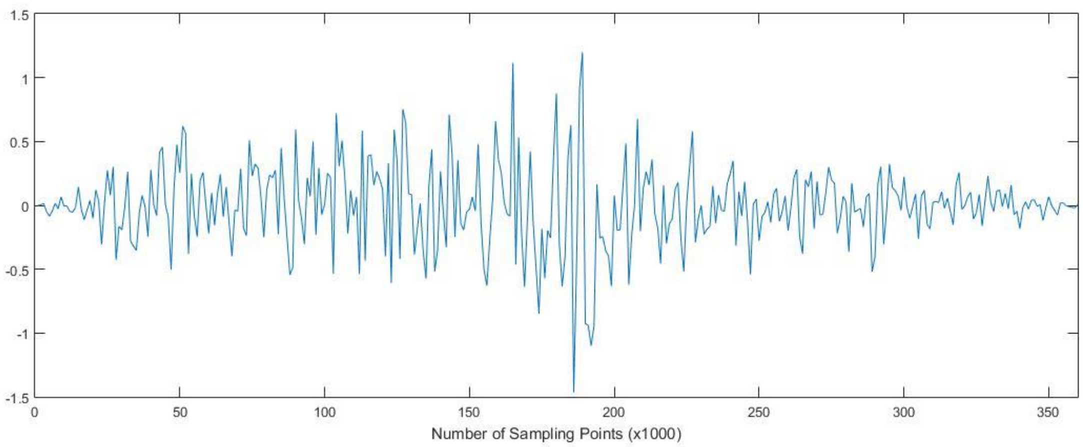


Figure 3.4(l): Cross correlated data between stations 1 and 2 for twelfth hour

Similarly, cross correlations are carried out for the remaining pairs for the full twenty four hours.

3.5 Staking

After hourly cross correlations have been done, data is returned to time domain by applying inverse Fourier transform. Here, I will show two stacks corresponding to the pairs of stations 1 and 2, and stations 1 and 3.

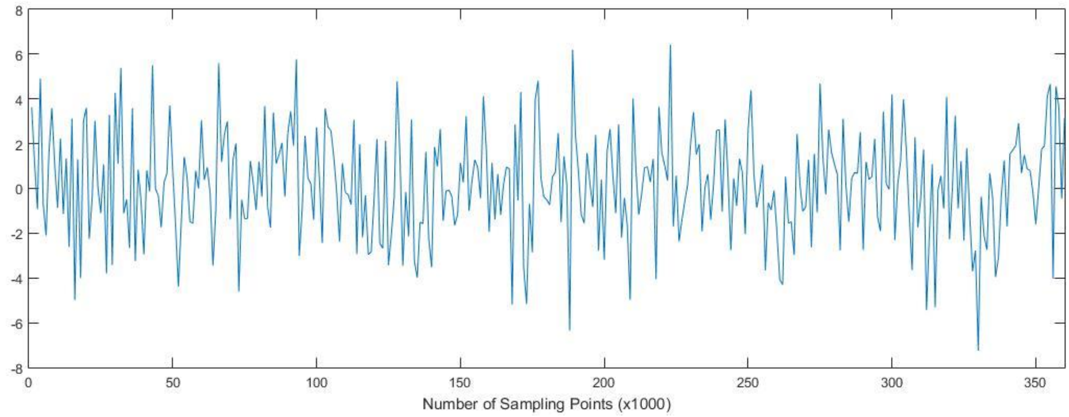


Figure 3.5(a): Stacked data corresponding to stations 1 and 2

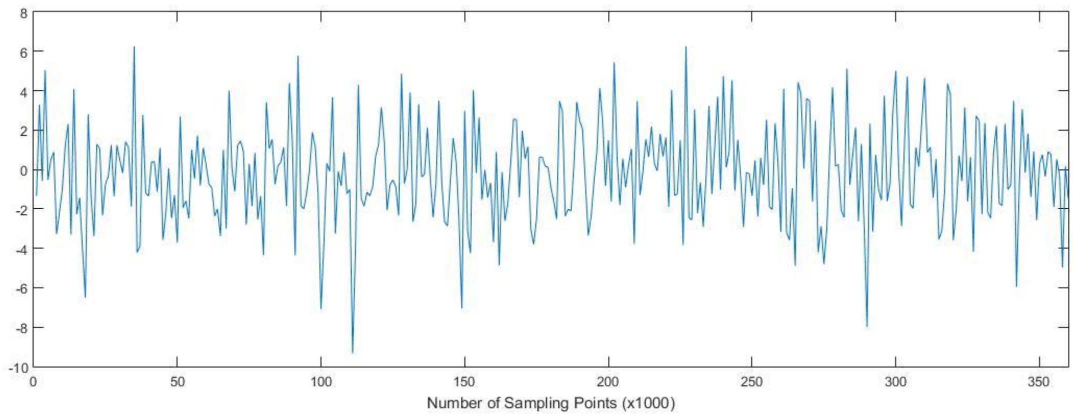
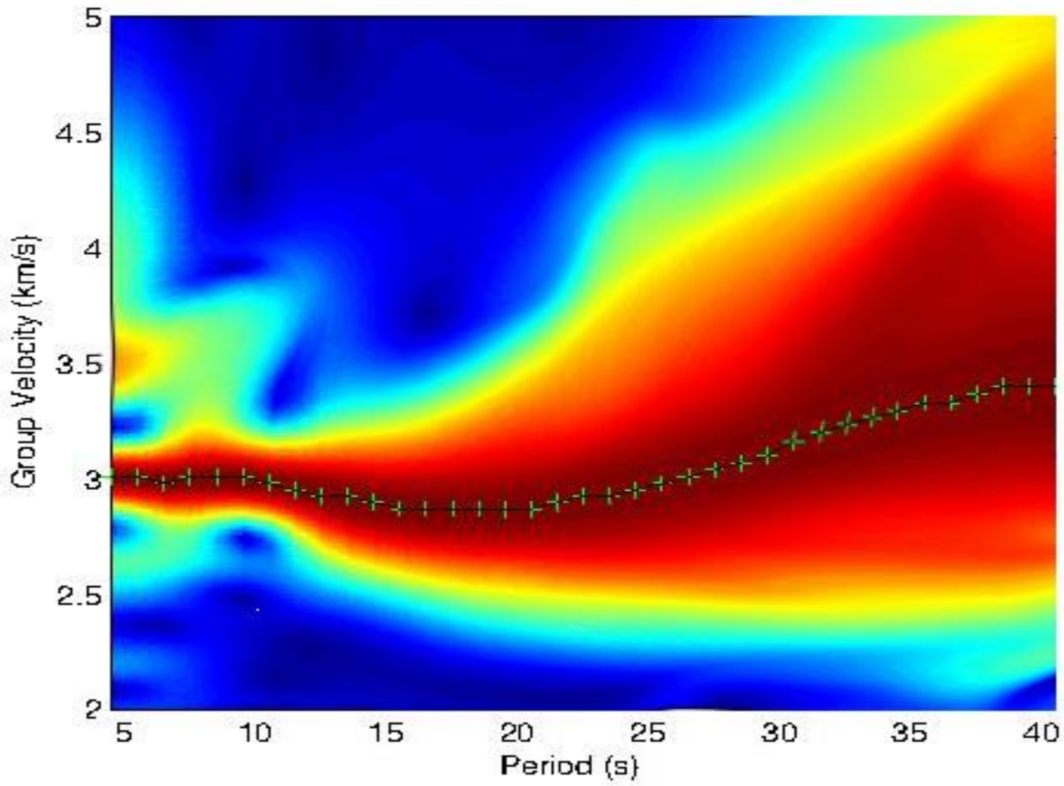


Figure 3.4(a): Stacked data between stations 1 and 3

3.6 FTAN Analysis

After obtaining the estimated Greens function, we can now get the group velocity as well as the phase velocity of the region. Here, I have used python package that takes in cross correlated data and outputs the dispersion curve of the region. The following plot is the dispersion curve I got from the package.



Here the black line indicates the group velocity curve. Further work involves inversion of the obtained dispersion curve, so as to obtain the geological profile of the region. Here, I have limited myself to the production of the dispersion curve, which has been produced by using python package for ambient noise seismic tomography. All other work has been done on MATLAB, the codes of which can be found in the annexure.

CHAPTER 4

Conclusion

The dispersion curve shows the variation of group velocity and phase velocity of seismic waves. From the dispersion curves, we can estimate the mechanical parameters of the medium. The advantages of using dispersion curves to obtain information about the geological features of a region include less complex inversion techniques as compared to dealing with inversion of waveforms.

Moreover, since in this technique, the dispersion curves are obtained from ambient noise, so there is no need of earthquake data, which is generally concentrated at few centres around the world. On the other hand, we can get ambient noise anywhere in the world, apply the above mentioned method, obtain dispersion curves of the region, and then perform inversion to get an insight of the geological profile of the region.

The ambient noise that has been used in this thesis is high frequency noise, produced by wind and various human activities. Such noise of low penetrating depth, and generally cover only the surface waves, i.e., Love or Rayleigh waves. However if use ambient noise generated from ocean waves, which are low frequency, then they can be used to generate dispersion curves corresponding to greater depths.

APPENDICES

Appendix A

MATLAB code for pre-processing, cross correlation, and stacking of data

```
%% Load raw data
clear all
clc
close all

load('a.mat')
delt = 1/100;
b=length(a(:,1));
t = delt*(1:b);
t = t';

%% De-mean and de-trend

for i=1:120

    me(:,i)=((a(:,i)-mean(a(:,i))));%./abs(a(:,i)-mean(a(:,i)));
end

for i=1:120
    nom(:,i)=(me(:,i))./abs(me(:,i));
end

for i=1:120
    a(:,i)=fft(nom(:,i));
end

%% Cross-correlation
p2=(a);
p1=p2(1:359999/2+1,:);
f = 100*(0:(359999/2))/359999;
p1(2:end-1,:) = 2*p1(2:end-1,:);
a=p1;
k=1;
h=4;
o=0;
for i=1:1:119
    if rem(i,5)==0
        h=4;
    end
    for j=i+1:1:120
        if rem(i,5)~=0
            d(:,k)=xcorr(a(:,i),a(:,j));
            k=k+1;
            o=o+1;
        end
        if o==h
            break
        end
```

```
    end
    if rem(i,5)~=0
        h=h-1;
    end
    o=0;
end

%% Stacking
d=ifft(d);

for i=1:1:10
    stacking(:,i)=sum(d(:,i:10:240),2);
end
```

References

- Bensen, G. D., RitzwollerM. H., BarminM. P., LevshinA. L., LinF.,MoschettiM. P., Shapiro N. M. and Yang Y., 2007, Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, Geophysics Journal International 169, 1239-1260
- Python package for getting dispersion curve has been obtained from https://github.com/bgoutorbe/seismic-noise-tomography
- SURFACE WAVE TOMOGRAPHY FROM AMBIENT SEISMIC NOISE, Center for Imaging the Earth's Interior, Department of Physics, University of Colorado at Boulder
- WeaverR. L.,2011,Seismic Noise Correlations, *U Illinois, Physics*
- YoungM. K., Rawlinson N.,Arroucau P., Reading A. M., and Tkalčić1H.,2011,High-frequency ambient noise tomography of southeast Australia: New constraints on Tasmania's tectonic past, GEOPHYSICAL RESEARCH LETTERS, VOL. 38, L13313