

The use of ambient seismic noise to investigate
internal changes in a tailings storage facility and
to image the subsurface geology in the Cradock
area of the Eastern Cape

By

Thulisile Kunjwa

A dissertation submitted to the Faculty of Science, Nelson Mandela
University in fulfilment of the requirements for the degree of

Master of Science

December 2021

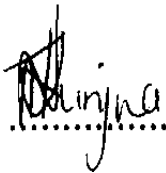
Supervisor: Prof. M. Doucouré



NELSON MANDELA
UNIVERSITY

DECLARATION

I, **THULISILE KUNJWA**, student number: **215122739**, hereby declare that the thesis: The use of ambient seismic noise to investigate internal changes in a tailings storage facility and to image the subsurface geology in the Cradock area of the Eastern Cape, for Masters in Geology is my own work and that it has not previously been submitted for assessment or completion of any postgraduate qualification to another University or for another qualification.



.....

THULISILE KUNJWA

DATE: 18 NOVEMBER 2021

In accordance with Rule (55.11.4) I hereby declare that the above-mentioned treatise/ dissertation/ thesis is my own work and that it has not previously been submitted for assessment to another University or for another qualification. However, material from publications by the student may be embodied in a treatise/dissertation/ thesis.

Abstract

Passive seismic interferometry is a process by which ambient noise data recorded at different seismic stations can be cross-correlated to estimate Green's functions. In the past, both surface waves and body waves have successfully been extracted by cross-correlation of ambient noise data on both regional and global scales. Recent advancements in ambient seismic noise techniques have the potential to provide new methods for subsurface imaging and monitoring. The ambient noise data processing procedure divides into four principal phases: (1) single station data preparation, (2) cross-correlation and temporal stacking, (3) measurement of dispersion curves and (4) inversion of dispersion curves to obtain 1-D shear wave profiles and computation of 2-D shear wave velocity cross-section. The purpose of this study is to investigate whether ambient seismic noise can be used to image and detect internal changes within a mine tailings dam wall and to image the subsurface geology of part of the Eastern Cape Karoo near Cradock.

In the first experiment, the investigation consisted of continuous ambient noise data recordings over a period of 3 days with 20 three-component short period geophones. The geophones were deployed over a survey wall of roughly 100 m in length at the Harmony Gold mine tailings dam in Welkom. In the second experiment, the investigation consisted of data recordings over a period of 35 days. The geophones were deployed in Cradock. The first phase of the data processing procedure included de-trending, de-meaning and band-pass filtering the data. This was done to ensure that any long period trends associated with instrument glitches are removed from the data. A spectrogram was then computed to view the spectrum of frequencies in the signal and to check if the filter that was designed was able to cut off the unwanted frequencies. The horizontal and vertical components of the ambient noise data were cross-correlated and picked between sensor pairs to create surface wave dispersion curves.

Subsequently, the dispersion curves were inverted to estimate the shear wave velocity of the dam wall and subsurface as a function of depth. The computed cross sections of shear wave velocity indicated a low-velocity zone between 2 and 10 m below the surface on the dam wall, this suggested that the phreatic surface is much closer to surface in this area. In the second experiment, the interpolated shear wave velocity profiles indicated that there is a layer of low velocity zone between depths 250 to 300 m below the surface. The cross-correlations were also used to compute group velocity maps from periods 1.5 seconds to 30 seconds. The group velocity maps showed various high and low velocity anomalies. The high velocity zones observed on the eastern section of the map were interpreted as evidence of dolerite

intrusions. The low velocity zones observed in the western and southern sections of the map interpreted as Karoo sediments that belong to the Adelaide Subgroup which is dominated by mudstones.

Acknowledgments

I would firstly like to thank God for giving me the strength to pursue this project, none of this would have been possible if it was not for Him. I would like to express my sincere gratitude to my supervisor Professor Moctar Doucouré for his scientific expertise, patience and guidance throughout my MSc career.

I am particularly grateful to Dr Nicholas Tonnelier who saw potential in me and encouraged me to pursue an MSc degree. I would like to thank Dr Lucian Bezuidenhout for teaching me how to install seismic stations in the field, assisting with MATLAB and for always having time for my questions in his very busy schedule.

I would also like to thank Gerrit Olivier from the Institute of mine seismology for giving me an opportunity to work on the tailings dam project and for outlining the different phases of the processing procedure which were crucial in meeting the aims of the study.

I would like to acknowledge the financial support from AEON (Africa Earth Observatory Network). I feel privileged to be part of such a great research group.

I would like to thank my friends and colleagues Sinazo Dlakavu, Sameera Musa, Abiel Kidane and Okuhle Poto for providing much needed breaks during my most stressful times and for always encouraging and motivating me to endure the hardships of my project and finish my degree. Thank you to my best friend Nelisiwe Nkutha for always being there for me. A very special thanks to my siblings Nomonde and Tendai for their love and support throughout my career.

Lastly, I would like to thank my mother for her love, support, encouragement and for the daily phone calls (yes daily) that kept me going when the road was tough. I am truly grateful and I owe everything I have and I am to you, ngiyabonga Ndlondlo enhle.

Table of Contents

1. Introduction	1
1.1 Background.....	3
1.2 Aims and objectives	6
1.3 Thesis Outline.....	6
2. Geological setting	8
2.1 The Karoo Basin.....	8
2.2 Study area 1: Welkom, Free State	9
2.3 Study area 2: Cradock, Eastern Cape	11
3. Previous Geophysical Studies of the Karoo	13
4. Tailings Dams.....	19
4.1 Methods of construction	19
4.2 Common monitoring and management techniques	21
5. Passive Seismic Interferometry	24
5.1 Historical development.....	24
5.2 Ambient seismic noise.....	26
5.3 Theoretical background of Cross correlations.....	27
5.5 Dispersion Curves	30
6. Data Acquisition	37
6.1 Study area 1	38
6.2 Study area 2	38
6.3 Data Processing	39
7. Data Processing.....	42
7.1 Study area 1: Harmony Dataset	42
7.2 Study area 2: Cradock Dataset.....	58
8. Discussion	77
8.1 Harmony Dataset	77
8.2 Cradock Dataset.....	79
9. Conclusion.....	82
9.1 Harmony Dataset	82
9.2 Cradock Dataset.....	83
9.3 Recommendations	84

10. References85
Appendix A.....93
Appendix B.....108

List of Figures

- Figure 1: Schematic diagram of a seismogram showing the arrival order of the waves (Bormann, et al., 2012)3
- Figure 2.1: Simplified map of the geology of the Karoo Basin. The study areas where the data was collected are indicated on the map by yellow circles.9
- Figure 2.2: Map of the geology that underlies Welkom, the study area. The red rectangle indicates the area where the tailings dam is situated.10
- Figure 2.3: Simplified geological Map of the second study area11
- Figure 3.1: Location map showing the occurrence of dolerites within deep SOEKER wells (triangles) within the Karoo Basin (thick black outline). Wells in the southwest and east intersect >150 m thickness of dolerite (red triangles), while those in the south-central Karoo Basin intersect (Sheiber-Enslin, et al., 201613
- Figure 3.2: Reduced to the Pole (RTP) magnetic residual anomaly map. The black arrows point to circular or ring structures which are some of the dolerite sills (Baiyegunhi, et al., 2019)....14
- Figure 3.3: Depth slices showing the appearance and disappearance of anomalous features with depth. The BMA and BSA represent the Beattie magnetic anomaly and the bean shape anomaly. (Baiyegunhi, et al., 2019)...15
- Figure 3.4: Map of southern Africa showing the study region, (right) zoom on part (18 km length) of the seismic line indicating the location of the borehole shot (star) and theseismic recorders (dots) (Ryberg, 2011).15
- Figure 3.5: Interpretation of the migrated section and subdivision of the continental crust: the upper crust region above the sub-horizontal reflectors show dipping strata (annotated in pink) of the Karoo Supergroup. The uppermost continuous reflectors (yellow and orange) represent 5 to 10 km thick Phanerozoic Cape Supergroup sediments with a shallow dip to the south. An unconformity separates the upper crust from north dipping mid-crust granitoid gneisses of the NNOB (green). The lower crystalline crustal slab (blue lines) displays an internal fabric that dips both to the north and south, but predominantly to the north. The boundary between the mid- and lower-crustal sections is not well defined. Reflectivity ~1 to 2 km above the undulating Moho depth suggests a possible additional layer. (Lindeque, et al., 2011).....16
- Figure 3.6: Rayleigh waves group velocity inversion maps using ambient seismic noise wave fields at periods: 3.0 s; 5.0 s; 6.0 s and 7.0 s. The black triangles show the location of the stations. The blue colour indicates the high velocity regions, and the red colour indicates the low velocity domains. (Bezuidenhout, et al., 2016).....17
- Figure 3.7: Array location indicating position with respect to the ocean. b) Normalized amplitude of the cross-correlation function for both positive and negative lag plotted with its corresponding network azimuth for Array C. c) Same as for Array A d) same for Array B.

(Bezuidenhout & Doucouré, 2020)	18
Figure 4.1: Diagram of a tailings dam constructed by upstream method (Cutifani, M. (2019)).....	20
Figure 4.2: Diagram of a tailings dam constructed by downstream method (Cutifani, M. (2019)).	21
Figure 4.3: Diagram of a tailings dam constructed by centerline method (Cutifani, M. (2019))....	21
Figure 5.1: 1-D diagram showing an example of direct-wave interferometry. (a) A plane wave traveling rightward along the x-axis, emitted by an impulsive source at $x=x_S$ and $t=0$, the response observed by a receiver at x_A . This is the Green's function $G(x_A, x_S, t)$ (c) As in (b) but for a receiver at x_B . (d) Cross correlation of the responses at x_A and x_B . (Wapenaar, et al., 2010).....	27
Figure 5.2: Illustration of direction of propagation of surface waves through a medium. Particles are represented by cubes in the models (a) In Love waves particle motion is horizontal and perpendicular to the direction of propagation (transverse) and (b) In Rayleigh waves, particle motion consists of elliptical motions (Boschi, 2014) in the vertical plane and parallel to the direction of propagation.....	30
Figure 5.3: Geometric dispersion of surface waves in vertically heterogeneous media. λ is the wavelength of the surface wave with phase velocity V and f is the. V_A and V_B indicate the generic shear wave velocity in the two layers affected by the surface wave propagation. (a) Sketch of amplitude decay of the fundamental mode at different wavelengths, (b) dispersion curve in the wavelength—phase velocity domain, (c) dispersion curve in the frequency—phase velocity domain. (Foti, et al., 2018)	31
Figure 5.4: An illustration of the overall procedure and main advantage of the MASW method. Complicated nature of seismic waves is carried over into the measurement (multichannel record). Then, dispersion nature of different types of waves is accurately imaged through a 2-D wave field transformation. Certain noise wave fields such as back- and side-scattered surface waves and several types of body waves are automatically filtered during this transformation. Dispersion curves are then extracted to be inverted for a 1-D VS profile, multiples of which can be prepared to make a 2-D VS map (Park, et al., 2007)	33
Figure 6.1: (a) Aerial view of the tailings dam where the seismic stations were installed. (b) Zoomed in picture of figure 6(a) showing the distribution of the seismic stations on the tailings dam and the location of the cone penetration tests. The area with seepage is also indicated in the picture	38
Figure 6.2: Google Earth image showing the distribution of the seismic stations in Cradock....	39
Figure 7.1: Data from station A9M and A9J, after the mean and linear trends are removed. The length of the time-series is 82800 seconds and the amplitude of the data is presented on the y-axis in digital counts which is an unscaled measure of displacement.	43
Figure 7.2: The raw signals (blue) from the two stations plotted with the filtered signals (Orange). Spikes and samples with high amplitudes have been removed and are not presented in the	

orange signal.....	44
Figure 7.3: Spectrogram showing the filtered and filtered signal. The spectrograms of the signal below show a frequency band between 7 and 40 Hz, all other frequency components that were not in the filtered range have been removed.....	45
Figure 7.4: Segment of filtered data that represents the first ten minutes that were recorded by station A9M.....	45
Figure 7.5: (a) Filtered data before temporal normalization using the running absolute mean method. (b) Filtered data after temporal normalization.....	46
Figure 7.6: (a) An example of the amplitude spectrum of station A9J before spectral whitening and (b) amplitude spectrum after spectral whitening is applied.....	47
Figure 7.7: (a) Cross-correlation function between stations A9M and A9J plotted between lags 20 and 20 seconds. The different colours represent signals travelling in opposite directions....	49
Figure 7.8: (a) The Love wave virtual source signals shown as a function of interstation distance....	50
Figure 7.9: Diagram of the computed dispersion curves (Phase velocity is plotted against frequency). The plotted lines represent the average dispersion curve for the different sensor pairs....	51
Figure 7.10: Dispersion curve showing the fundamental mode in blue, the 1st higher mode in purple, and the 2nd higher mode is plotted in yellow and the 3rd in red....	52
Figure 7.11: (a) 1-D inverted profile (b) Theoretical dispersion curve that corresponds to the velocity model.....	55
Figure 7.12: (a) An example of Love wave dispersion curve inverted with a neighborhood algorithm to find ground profile. The values in red indicate a low misfit.....	56
Figure 7.13: 2-D shear wave velocity profile over 85 m of the section. The low-velocity zone below the white line likely corresponds to the phreatic surface. The numbers at the bottom of the figure indicate geophone positions. The profile shows increasing distance from right to left and was plotted in this way to be consistent with the order in which the stations were planted. The stations were planted from left to right on the tailing dam wall.....	57
Figure 7.14: Data from station A1R and B5Y, after the mean and linear trends are removed. The length of the time-series is 86400 seconds and the amplitude of the data is presented on the y-axis in digital counts which is an unscaled measure of displacement.....	59
Figure 7.15: Data from stations A1R and B5Y, after band-pass filtering was applied.....	59
Figure 7.16: (Spectrogram computed from stations A1R and B5Y plotted against time in hours.....	60
Figure 7.17: Spectrogram after band-pass filter was applied.....	61
Figure 7.18: (a) Stacked cross-correlation functions (b) Stacked cross-correlation functions for 15 days.....	63
Figure 7.19: Dispersion curves computed using FTAN for the different stations. The black lines	

indicate the fundamental modes.....	66
Figure 7.20: (a) shows one of the inverted 1-D shear wave velocity profiles that were computed between stations AAD-B5Y. (b) Shows the inverted 1-D shear wave velocity profiles that were computed between A1R-B4X.....	68
Figure 7.21: 2-D shear wave velocity profile across the study area. There is a low-velocity zone between 5 and 15 km and at 220 m. High velocity zone is observed at a shallower depth between 15 and 20 km. The numbers at the top of the figure indicate the geophone positions.....	70
Figure 7.22: Density path maps plotted for periods 1.5 and 2.5 seconds.....	76
Figure 7.24: Ray coverage maps for periods 1.5 and 2.5 seconds.....	77
Figure 7.26: Group velocity maps plotted for periods 1.5 (a) and 2.5 (b) seconds. The yellow colours indicate high velocity zones and the purple indicates low velocity zones. Clear arrival fronts are visible at positive and negative lag times.....	79
Figure 7.27: Group velocity maps plotted for periods 5 and 10 seconds. The yellow colours indicate high velocity zones and the purple indicates low velocity zones.....	80
Figure 7.28: Group velocity maps plotted for periods 20 and 30 seconds. The yellow colours indicate high velocity zones and the purple indicates low velocity zones.....	81
Figure 8.1: (a) Schematic diagram showing the location of the CPTu tests and the geophones. The grey line indicates the position of the surface and the blue dotted line indicates the position of the phreatic surface. (b) Three panels; panel 1 is coneresistance, panel 2 is pore pressure, panel 3 is dissipation and panel 4 is S-wave velocity.....	84
Figure 8.2: Description of the SOEKOR deep borehole that was drilled in Cradock. (Linol,et al., 2016).....	86

List of Tables

Table 1: Typical Shear Wave Velocities of different types of soil (Chen & Ong, 2012).	44
---	----

List of Photographs

Photograph 1.....	38
Photograph 2.....	78

1. Introduction

The main goal in seismology is to study the subsurface of the Earth's interior and provide an image of what it looks like and get a better understanding of the processes that happen beneath, which are fundamental drivers in all geological processes. While serving a general scientific interest in the origin, history and development of our planet, understanding the subsurface geology plays a key role in mitigating a number of societal issues, ranging from natural hazard assessments (e.g. earthquakes, volcanoes, landslides) to natural resource exploration and production (e.g. water, minerals, ores, oil, gas, geothermal energy) and subsurface waste storage (e.g. nuclear waste, CO₂). In all these cases obtaining information about the subsurface is vital (Draganov, et al., 2006).

Seismic wave fields that are generated by earthquakes or artificial sources of energy provide a means to illuminate the Earth's interior and to create an image of the subsurface. Seismologists analyze seismic waves to examine the Earth's internal structures and processes.

The traditional methods used in seismology to image the Earth's interior require a recognizable pulse source of seismic energy (natural or artificial) to pass through the earth to make inferences about its structure and characteristics (Nicolson, et al., 2012). When the waves propagate within the subsurface, they are either reflected, refracted or diffracted when elastic contrasts occur at boundaries between layers or rock masses with different physical properties (Schuck & Lange, 2007). The recording of seismic waves returning from the subsurface to the surface allows inference on structures and lithological compositions of the subsurface.

Since traditional methods require an active source, their application is limited to seismologically active areas or to areas where controlled sources such as explosives can be used with minimal environmental impact. The other limiting factor with this method is the requirement for repetition of earthquake events and the need to generate a highly energetic wave of short duration that is repeatable and does not generate noise that interferes with the detection (Hirose, et al., 2017).

The demand for cost-effective exploration techniques that could characterize geological structures with minimum impact on the environment resulted in the development of a new technique known as passive seismic interferometry. The technique makes use of natural ambient seismic noise to infer information about the medium that through which waves propagate (Shapiro & Campillo, 2004).

Ambient seismic noise is a term encompassing the background, or “noise” seismic wave field (Sabra, et al., 2005). The amplitude of ambient noise is greater along coastal areas than at continental stations. The dominant peak frequency is about 0.15 Hz, which is twice the typical frequency of ocean swell of about 0.07 Hz. The ambient noise around this frequency can be described as a persistent seismic wave field that is excited by ocean swell activity and is a significant source of noise in observations of seismic waves from earthquakes (Nishida, 2017). Over time, multiple noise sources repeatedly enter the surface seismic waves, and the inhomogeneity of the earth's crust scatters these waves, further randomizing the wave field. Most studies have focused on surface waves in the microseismic frequency band (0.05 - 0.2 Hz), which contain strong signals generated by ocean waves (Bensen et al., 2007; Lin et al., 2007). High-frequency noise includes inputs from environmental and anthropogenic sources, such as wind and vehicle traffic.

When energy is released, body and surface waves radiate from the source at the same time. P waves are faster than S waves, they arrive at a seismogram first, followed by S waves and then by surface waves (Bormann, et al., 2012). Figure 1 depicts the arrival pattern of waves in a typical seismogram.

Theoretical and experimental studies have shown that the cross-correlation of the diffuse ambient noise wave field provides an estimate of the impulse response or the empirical Green's Function (EGF) of the subsurface between the two stations that are correlated (Shapiro, et al., 2005). Seismic interferometry is a concept derived from the Claerbout conjecture proposed in 1968. The conjecture states that cross-correlation of noise traces recorded at two different receiver positions gives the response observed at one of the receiver positions if there was a source at the other (Curtis, et al., 2006). This is called a virtual source and can be achieved if the receiver is in a heterogeneous three-dimensional medium with a diffuse noise wavefield. This cross-correlation response is also called Green's function, and "diffuse" means that the amplitude of the normal mode is not correlated but has the same expected energy in all directions (Wapenaar & Fokkema, 2006).

The Green's function is important because it contains information about how energy travels through the Earth between two locations which is based on changes in the relative velocity of body and surface waves travelling through the Earth.

Therefore, the use of ambient seismic noise in seismic imaging requires a better understanding of the ambient seismic noise wavefield and its source locations and mechanism.

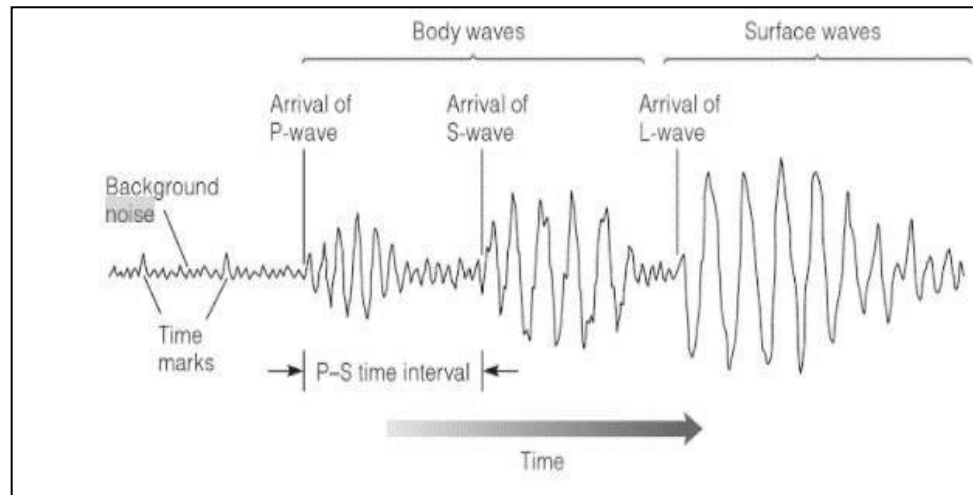


Figure 1: Schematic diagram of a seismogram showing the arrival order of the waves (Bormann, et al., 2012).

Seismic interferometry, developed in the last decade transformed the way in which seismologists think about both passive and active wave fields and how they use them to infer information about the subsurface. There has been a steady increase in the development of new interferometry techniques in the seismic community and these have been applied in broad disciplines.

(Das & Rai, 2016) showed that seismic interferometry can be used to calculate surface waves between receivers using recorded ambient noise data, and they proposed South India's first surface wave group velocity map based on ambient seismic noise that contained useful surface geological information.

This technique allows seismologists to provide high-resolution tomography studies from regional to continent scales. This method can also be applied in delicate areas where active sources cannot be used. The method is cost-effective; it is easy to measure the signal, which is available everywhere and anytime.

The dispersive nature of surface waves in a vertically heterogeneous medium provides important information regarding the stiffness properties of near-surface materials (Park, 2005). The basis of most surface wave analysis methods is to accurately determine the frequency-dependent phase velocities of fundamental mode Rayleigh waves and Love waves. In addition to being a function of frequency, the phase velocity of Rayleigh wave is also related to several sets of soil properties, the most important of which is the velocity of shear waves. Therefore, by inverting the experimental scattering curve, the shear wave velocity profile of the test field can be obtained.

Shear wave velocities play a vital role in the mitigation of seismic hazards because their low velocity values are the main reason for site amplification phenomena and can

significantly influence the susceptibility of an area to seismic-induced landslides (Paoletti, 2012). In general, the shear wave velocity profile of the uppermost 20–30 m is of most interest in engineering purposes (Lee & Trifunac, 2010). Shear wave velocity profiles can be used to evaluate the stiffness of the upper soil layers, and subsequently, e.g., their load bearing capacity. In addition, in earthquake design, the shear wave velocity is an important parameter for the assessment of liquefaction potential and soil amplification.

1.1 Background

In recent decades, the stability of tailings dams has attracted a lot of attention because many tailings dam failures have been recorded all over the world. According to reports, there are approximately 18,000 tailings dams in the world, and the failure rate in the past 100 years is estimated to be 1.2%. (Azam & Li, 2010). This is more than two orders of magnitude higher than the failure rate of traditional storage dams (Azam & Li, 2010). Generally, these storage facilities are prone to failure due to the following reasons: (1) Embankments constructed with mud, coarse waste and residual materials from mining operations; (2) As the height of the tailings dam increases, the amount of waste water increases; (3) Lack reasonable design standards and regulations; (4) The monitoring cost during the operation and closure of the tailings dam is high. (Lyu, et al., 2019).

The standard of public reporting varies considerably from country to country and from region to region. Many tailings dam failures were not reported or lacked basic information at the time of reporting. This seriously hindered the development of safety standards in these areas.

Despite advancement in technology, tailings dams still fail. It is impossible to determine the actual level of monitoring implemented by mining companies, but it seems that most companies operating large-scale projects still have active tailings dams and many of those are monitored through regular visual inspections. There are mining companies that still rely on physical field monitoring and visual monitoring by operators for analysis (O'Dwyer, 2019).

The monitoring of these activities and the need for personnel to carry out regular inspections of the dam walls on site are carried out on a regular basis with manual techniques. This method can range from manually measuring the position of the dam, installing downhole tools to measure pore pressure or even monitoring the structure of the fracture. Operators usually manually read most of the field instruments used.

It is not uncommon for readings from dam monitoring instruments to be manually inserted into a spreadsheet for analysis (Andrade, 2017). This approach creates many risks, including data entry errors, calculation errors, lack of audit trails, and the possibility of fraudulent data. Some tailings dams are in remote areas and this may hinder regular monitoring of the dams.

Some mining companies use geophysical methods to monitor tailings dams. Instruments in common use include piezometers, flow meters, and cone penetration tests. These instruments only provide point measurements, so it is not easy to detect the first signs of failure when the degraded area may be small and local (Olivier et al., 2018). In other cases, high resolution radars and or cameras are used to monitor small deformations in the damwall. This method measures surface disturbances and cannot detect internal changes in the dam wall. Another limitation of this technique is its poor performance during heavy rain. Due to the fluid on the wall, small deformations of the dam wall become very difficult to detect. Similarly, during heavy rains, it is difficult to interpret the seepage flow rate because rainwater flows down the embankments and contaminates the measurement results. It is well known that heavy rain is the main factor leading to the failure of the tailings dam wall (Azam & Li, 2010).

The failure of tailings dams is often difficult to predict. The results of these failures were great economic losses, environmental degradation and, often, human loss. Therefore, the development of a profitable method is urgently needed to monitor structural stability over time of these structures. The seismic interferometry technique has been applied in different ways. It was used as a tool in monitoring of temporal changes after phreatic eruption of Mt. Ruapehu in New Zealand. This technique revealed a statistically significant change in velocity because of the local volcanic process that occurred before eruptive activity (Mordret, et al., 2010).

Moreover, the technique has also been used to monitor groundwater storage at mesoscale (Lecocq, et al., 2017). The velocity changes identified in the study were interpreted as the effects of temperature diffusion and water storage changes. It was concluded that in addition to existing observation methods, ambient seismic noise recordings can become a valuable new tool for monitoring heterogeneous mesoscale groundwater systems. In addition, recent laboratory experiments have shown that the technology can detect internal erosion of earthen dams and levees (Planès, et al., 2016).

1.2 Aims and objectives

Aims of this thesis

- To investigate whether seismic interferometry can be applied and yield useful information that can be used in monitoring of tailings dam wall failures.
- To obtain empirical Green's functions by cross correlation of recorded ambient noise between stations
- To utilize the cross-correlation functions to compute dispersion curves to extract the Rayleigh and Love wave phase velocities as a function of frequency.
- To apply inversion to the computed dispersion curves to obtained ground profiles as a function of depth
- To compute 2D maps of shear velocity as a function of depth.

To meet the above-mentioned objectives, the ambient seismic noise experiments were conducted in two study areas. The first set of seismic stations were installed on a tailings dam in Welkom and the other in Cradock. The seismic networks recorded ambient noise (earth vibrations) at different periods in time. There were 20 sets of stations that were installed on a tailings dam in Welkom and were left to record data over a period of 3 days, from 27 August to the 30th of August 2017. The interstation distance for this set of data was 5 m. The second dataset consisted of a group of 24 stations that were installed in Cradock. The stations were left to record continuous data over a period of 35 days. The interstation distance for this set was 5 km. The data was sampled at 100 Hz. The main concept behind the technique is approximating empirical Green's function by calculating the cross-correlation of the ambient seismic noise between stations. The dispersion curves were then calculated from the cross-correlation functions. The dispersion curve measurements were inverted to infer information of the subsurface through 2-D shear wave velocity cross sections.

1.3 Thesis Outline

This section contains an overview and description of the chapters that follow. The thesis is divided into nine chapters. The current chapter, Chapter 1 is the Introduction and briefly describes the seismic interferometry technique and highlights the main objectives of this study.

Chapter 2 provides a review of the geological setting of the Karoo basin, the area where the ambient seismic noise experiments were conducted and includes geological maps of the study areas.

Chapter 3 briefly discusses results from previous geophysical studies performed in the Karoo basin.

Chapter 4 discusses tailings dams and provides a review of the different types of construction methods and provides a brief description of the common monitoring techniques that are applied by mining companies.

Chapter 5 discusses the theory behind seismic interferometry and briefly outlines the historical development of the technique. Furthermore, the chapter characterizes the ambient seismic noise. The chapter shows that by cross-correlating random diffuse wave fields between pairs of receivers; the empirical Green's function can be approximated. Finally, the chapter also provides a brief description of the analysis methods that are applied on surface waves to extract dispersion curves.

Chapter 6 covers data acquisition and describes the field installation steps. The distribution and deployment of the station for each location is also described. Lastly, the chapter outlines the processing workflow.

Chapter 7 presents the results from the ambient seismic noise data processing flow used to obtain the empirical Green's function from raw data. In addition, the chapter provides the spectral analysis, which shows the frequency range where the ambient noise is most energetic. The chapter also shows computed dispersion curves and 1-D shear wave velocity profiles obtained from inversion of the resulting dispersion curves. The chapter also shows the computed tomography maps that were obtained in the Cradock study area.

Chapter 8 shows the results from the inversion and the computed 2D shear wave cross sections. The results obtained from the inversions are discussed in correlation to the subsurface geology of the Cradock area and discusses the inversion results in relation to the interior of the tailings dam.

Chapter 9 provides a conclusion of the study and main findings as well as possible recommendations for future studies using the technique in the Karoo Basin.

2. Geological setting

The Karoo basin developed within the continental interior of Southwest Gondwana and provides a record of the evolution of paleo-environments and climate changes of the Earth's history, following the amalgamation of the Gondwana supercontinent during 750-550Ma until its break up during the period 750-130 Ma (Catuneanu, et al., 2005). The Basin covers nearly two thirds of South Africa (Figure 2.1) and accommodates >5-kilometer-thick sedimentary sequence that was deposited from the Late Carboniferous to Early Jurassic (Aarnes, et al., 2010). The extensive volcanic activity of the Early Jurassic (183 ± 1 Ma) ended the sedimentation of the basin. This chapter briefly describes the regional geology of the Karoo Basin and includes maps of the two study areas where the data was collected.

2.1 The Karoo Basin

The basin is comprised of thick sedimentary rock sequences that were deposited between mid-Paleozoic to Lower Jurassic. These sedimentary rock sequences are duplicated along the thrusts and folds in the southern margin of the basin. In the north, these sequences are intruded by dolerite sills and capped by basalts that are related to the eruption of Karoo Large Igneous province at approximately 182 Ma during the initial break-up between East and West Gondwana (Linol, et al., 2016).

The main Karoo basin is mainly underlain by a base of rocks that include the Kaapvaal Craton in the north and the Namaqua-Natal Metamorphic Belt in the south and confined along its southern margin by the Cape Fold Belt (Johnson, et al., 2006).

The deposition of the first-order sedimentary sequence of Karoo generally occurred in the late Carboniferous, around 300 Ma, following major inversion tectonics event along the southern margin of the supercontinent that led to the assemblage of Pangea (Catuneanu, et al., 2005). Deposition in the south-central section of the basin took place during the late Paleozoic to early Mesozoic.

Deposition and sedimentation continued until the disintegration of the supercontinent in the Middle Jurassic (Catuneanu, et al., 2005). The accumulated sedimentation was intruded by a large igneous province. In the post-Gondwana era, the upper part of the Karoo sequence was eroded, so the age of the younger Karoo deposits generally ranges from the Triassic to the Middle Jurassic. The sedimentary filling of the basin is accumulated under the control of two main allogenic bodies, namely tectonics and climate (Catuneanu et al., 2005). The first-level Karoo sedimentary sequence has a

superior lithostratigraphic level, including several groups defined according to the general characteristics of the sediments. These groups are called Dwyka, Ecca, Beaufort, Stormberg, and Drakensberg in order of strata. Sedimentary environments vary greatly, including glaciers (Dwyka Group), ocean and coastal plains (Ecca Group), non-marine rivers and aeolian (Beaufort and Stormberg Group).

The depositional environments largely range from glacial (Dwyka Group), to marine and coastal plain (Ecca Group), to non-marine fluvial and aeolian (Beaufort and Stormberg groups). Figure 2.1 shows the simplified geology of the Karoo basin and its extension across South Africa.

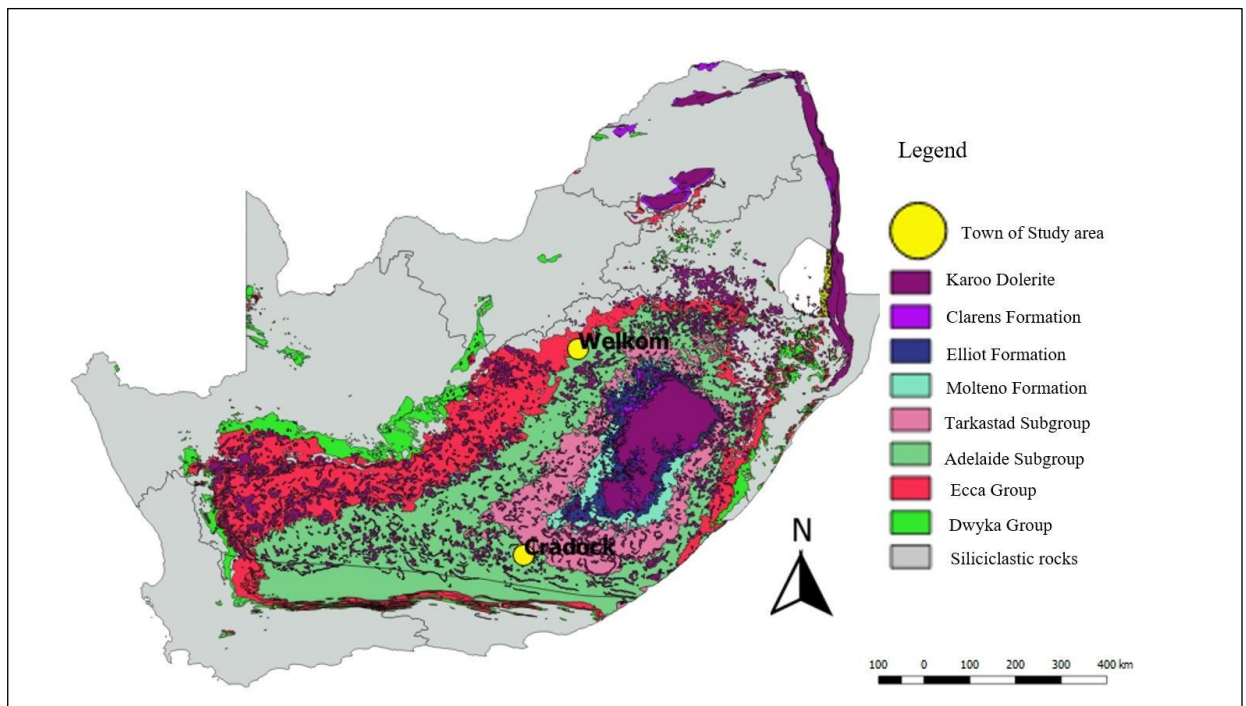


Figure 2.1: Simplified map of the geology of the Karoo Basin. The study areas where the data was collected are indicated on the map by yellow circles.

2.2 Study area 1: Welkom, Free State

The first study area is in Welkom, Free State. The ambient noise experiment was conducted on a tailings dam at the Harmony Gold processing plant. The Free State is situated in the core of the Karoo Sequence of rocks, containing shales, mudstones, sandstones and the Drakensberg Basalt which forms the youngest capping rocks. Mineral deposits are plentiful, with gold and diamonds being of importance and mostly found in the north and west of the province.

The geology of the province forms the basis for its mining activity. The physical landscape from the combination of local geology and both past and present geomorphological processes. The subterranean rocks in the Free State are dominantly mudstone, sandstone and shales, proving that the region formed part of the Great Karoo.

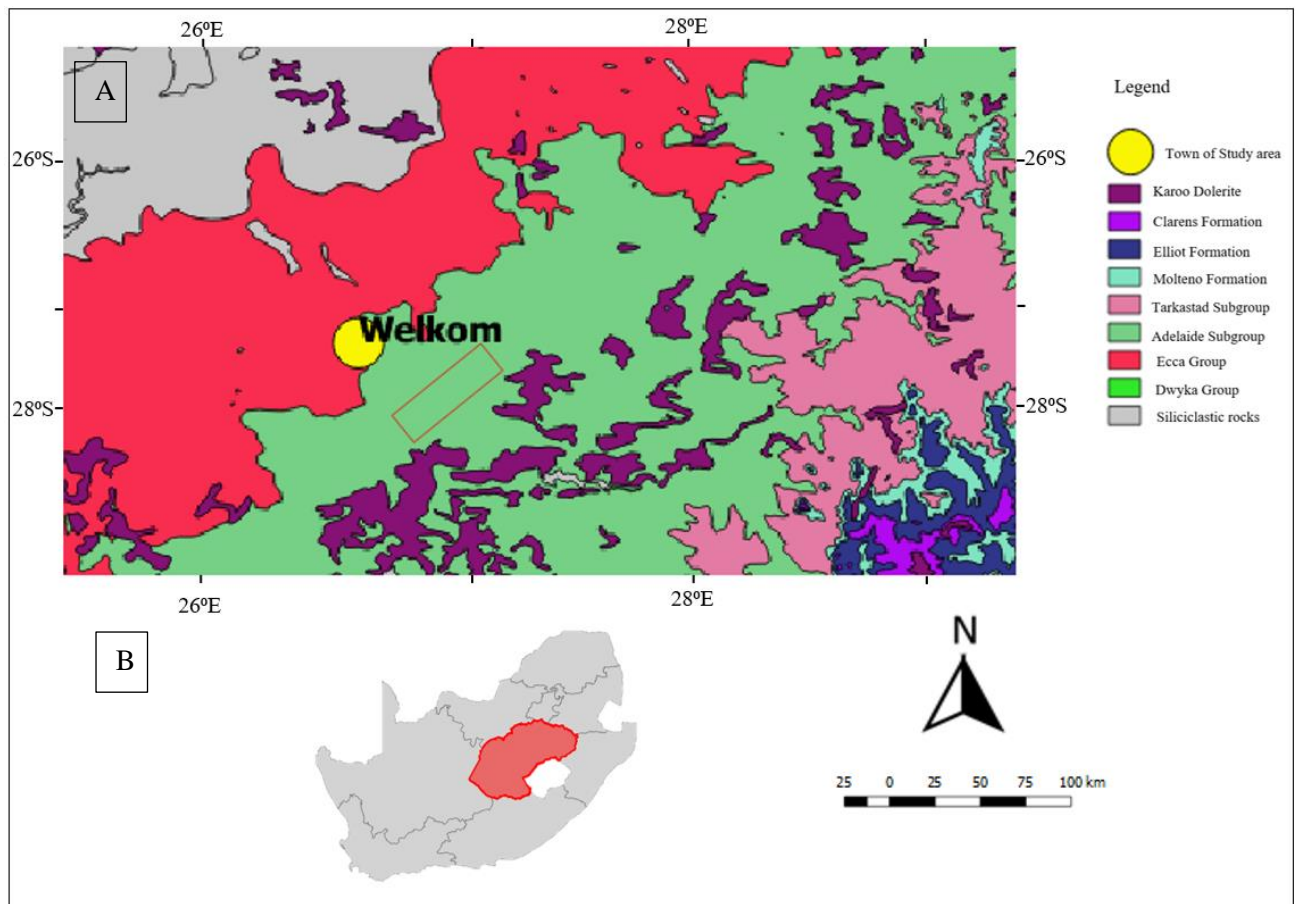


Figure 2.2: (a) Map of the geology that underlies Welkom, the study area. The red rectangle indicates the area where the tailings dam is situated. (b) Outline map of South Africa, the red area indicates the Free State province where Welkom is located.

The whole of the Free State province is covered by rocks of the Karoo Supergroup. The Precambrian geology is poorly represented and forms part of the basement rocks of the regional geology, the strata are represented in the northern Free State province at the Vredefort dome. The basement in this region is of Archean age and is composed of different types of granitoid rocks, there are several intrusive rocks of alkali granite that overlay the Witwatersrand series. (Reimold, 2006). The dominant Karoo geological formations underlying the Welkom study area are the Dwyka and Ecca Groups.

The Dwyka Group

The Dwyka group is the oldest group in the Karoo Supergroup. This group of rocks is from the Late Carboniferous to the Early Permian and is located on the pre-Cambrian bed of the glacial. The group also consist of various rock types that display features that reflect a glacial or glacial-related origin such as diamictite, conglomerate, fluvio-glacial pebbly sandstone and mud rock with drop stones (Johnson, et al., 1996). Rock layers from this group show distinct lithological differences over the basin.

The Ecça Group

The Ecça group overlies the Dwyka group and is estimated to be of Permian age. The group is found between the Dwyka and Beaufort and occupies most of the Permian. The rocks of the Ecça Group are basically a clastic sequence of mudstone, siltstone, sandstone, small conglomerate and coal (Catuneanu, et al., 2005). The group comprises of 16 formations that reflect the lateral facies changes that are used to characterize the group (Johnson, et al., 2006).

The regional geology of the study area also includes the Beaufort Group, as well as the overlying Molteno, Elliot and Clarens Formations in south eastern section of the map (Figure 2.2) and described in Lucas & Hancox (2001).

2.3 Study area 2: Cradock, Eastern Cape

The second study area is in Cradock situated in the Eastern Cape Province of South Africa. The area is a semi-arid, undulating hilly region within the eastern portion of the Great Karoo. The regional geology of Cradock is dominated by rocks that belong to the Beaufort Group. The group comprises of a lower Adelaide Subgroup and an upper Tarkastad Subgroup.

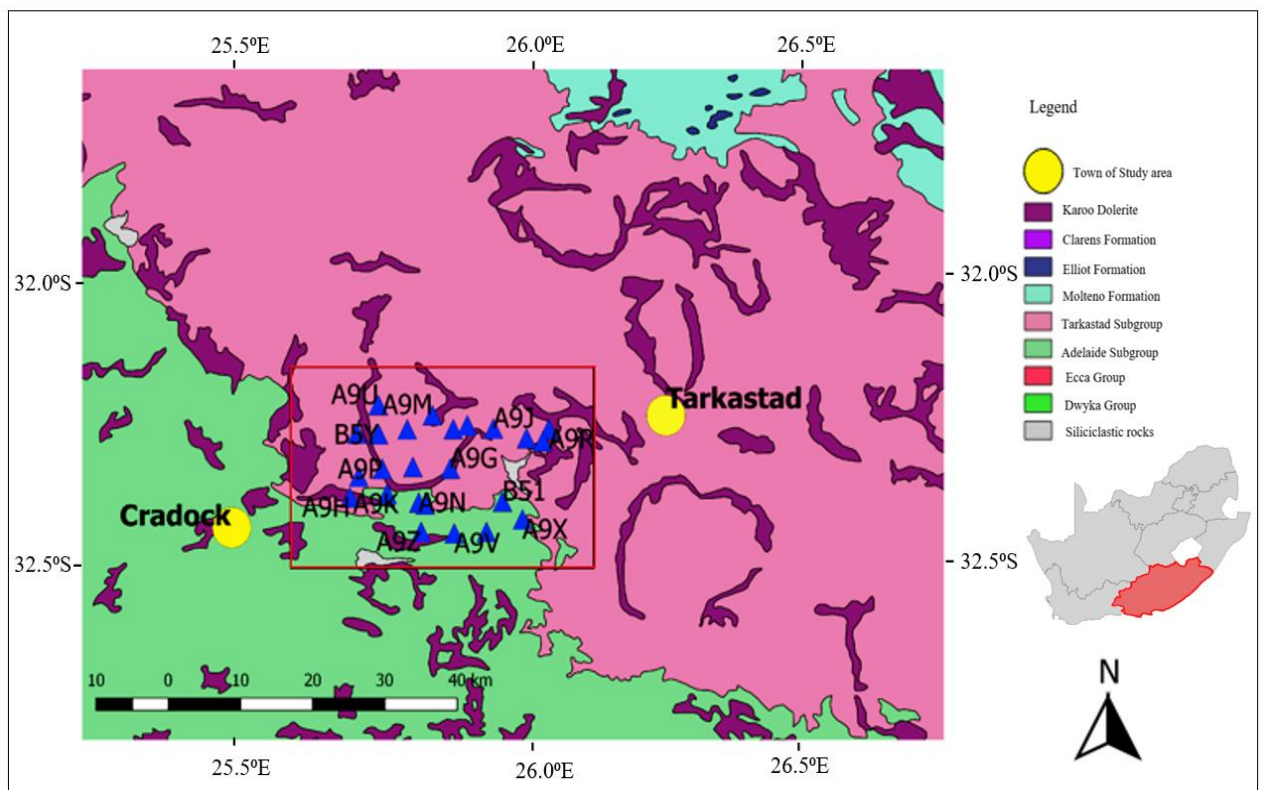


Figure 2.3: Simplified geological Map of the second study area in Cradock. The blue triangles represent

The Beaufort Group

The Beaufort group overlies the Ecca group, and its age ranges from the Middle Permian to the Middle Triassic (Catuneanu et al., 2005). This group is mainly composed of fossil-rich sedimentary rocks (Lucas and Hancox, 2001). This group is mainly composed of two main subgroups, the Adelaide subgroup and the Takastad subgroup. The Adelaide subgroup consists of 5 formations.

In the east of the Karoo Basin, the Adelaide subgroup includes the Koonap, Middleton, and Balfour formations, while the west is composed of the Abrahamskraal and Teekloof formations, and the north is the Volksrust and Normandien formations (Johnson et al., 2006). The Beaufort Group's stratum is mainly composed of alternating shale and siltstone, followed by lenticular sandstone and plate channels deposited by various river systems.

After the sedimentation of the basin stopped, a large amount of lava flowed out at around 182 Ma, and the magma invaded the basin at multiple levels and in the form of dams and foundations. The distribution of dolerites has changed throughout the basin. The northwest and east of the basin are the most abundant. The dolerite is concentrated in the middle and south of the basin, and the intrusions are mainly in the Beaufort Group.

3. Previous Geophysical Studies of the Karoo

Some of the earliest studies of geophysical research in the Karoo include the early exploration of seismic data by SOEKER in the 1970s. The digitized paper records that were acquired from the data indicated clear reflection targets (Webb, et al., 2016). This section discusses previous geophysical studies that were done in the Karoo and includes work that is compiled other authors.

The focus of most of the research conducted in Karoo is related to research on the impact of shale gas potential. Studies have shown that there are differences in the distribution of dolerites in the basin, and this difference is related to the location of the magma source that supplies the large igneous provinces on the west and southeast coasts of South Africa (Scheiber-Enslin, et al., 2016). The highest concentration of dolerites was found in the northwest and east of the basin, with a total dolerite thickness of > 150 m in each well (Figure 3.1). The volume of dolerite in the central and southern part of the basin has decreased significantly, and the thickness of the integrated dolerites is less than 150 m perwell, and the intrusion is limited to the Beaufort Group.

Understanding the geometry of the intrusions has been the core of most research because it has been found that the intrusions have intruded into the layers that were the current focus of shale gas exploration. The intrusions are suspected to have heated these shale gaslayers.

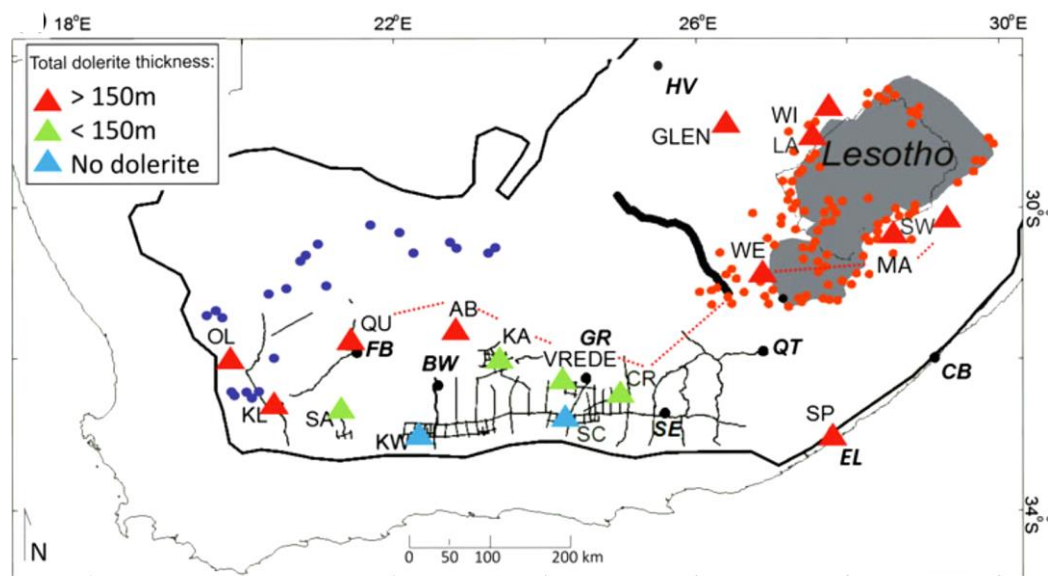


Figure 3.1: Location map showing the amount and presence of dolerites within deep SOEKOR wells (triangles) within the Karoo Basin (thick black outline). Wells in the southwest and east intersect >150 m thickness of dolerite (red triangles), while those in the south-central Karoo Basin intersect (Scheiber-Enslin, et al., 2016).

The geometry of these Karoo dolerites is still not well understood. Baiyegunhi et al. (2019) mapped and shown the dolerite intrusions that could have impacted shale reservoirs to understand their geometry within the main Karoo Basin. In this study, the aeromagnetic and gravity data that was acquired was combined to image the geometry of dolerites within the main basin (Figure 3.2).

The magnetic signatures from the dolerite intrusions are observed on the magnetic map and become broader up to the depth of 5400 m, they eventually disappear at a depth of about 15600 m. This possibly indicates that the Karoo intrusions were emplaced near the surface, existing at the targeted depth for shale gas exploration (Baiyegunhi, et al., 2019).

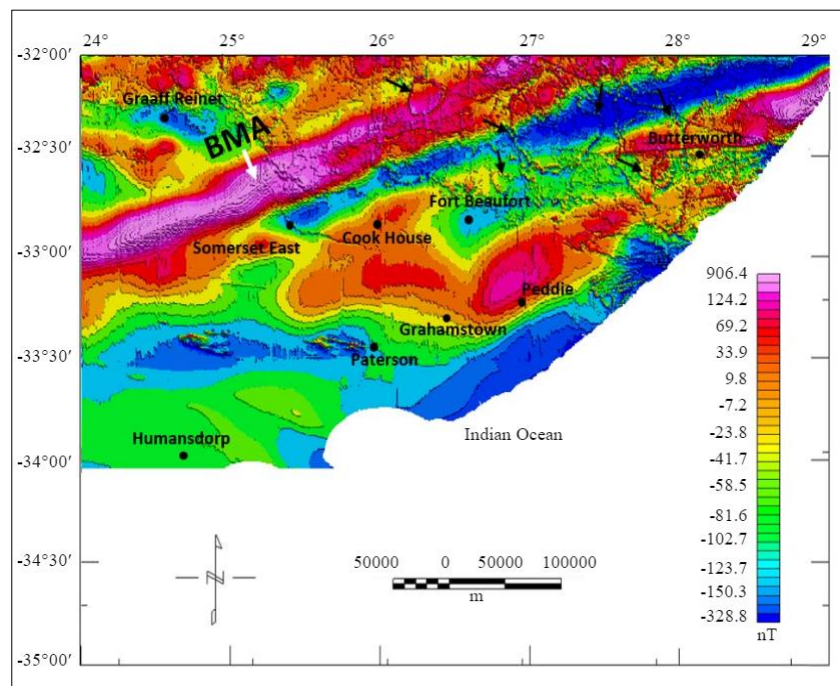


Figure 3.2: Reduced to the Pole (RTP) magnetic residual anomaly map. The black arrows point to circular or ring structures which are some of the dolerite sills (Baiyegunhi, et al., 2019).

The authors calculated the power spectrum to estimate the average depth to magnetic sources and analyzed the anomaly amplitude responses within depth slices (Figure 3.3). These depth slices will rather be called depth windows in the following as they represent differences in vertical height. Depth windows 1-3 revealed structural patterns that overlap with dolerites on the geological map of the area and also indicate that the dolerites have responses that are noticeable until a depth of up to 5400 m (Figure 3.3).

Depth window 4 revealed only minor anomalies that are linked to dolerite intrusions. A dolerite sill with a circular or ring structure was observed in the Northern part of the map. This ring structure becomes stronger and map visible as the depth increases to about 5400 m disappears on depth window 4 which indicate magnetic signatures at a depth of 15600 m (Baiyegunhi, et al., 2019).

Based on the gravity models and magnetic depth window that were obtained from the study, it was inferred that the mapped dolerites are interconnected at depth and extend from the surface down to a depth of approximately 5400 m.

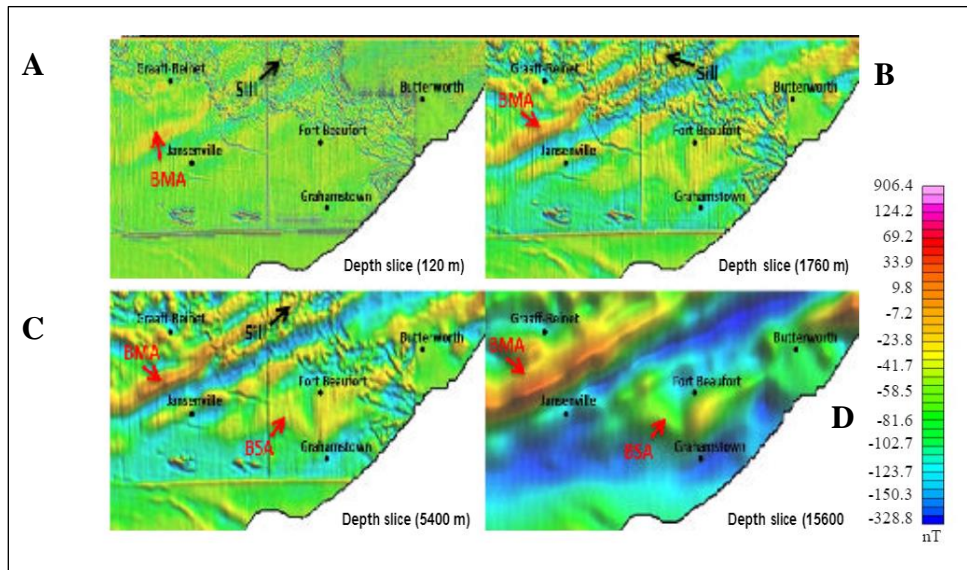


Figure 3.3: Depth windows showing the appearance and disappearance of anomalous features with depth. (a) Depth window 120m (b) Depth window 1760 m (c) Depth window 5400 m (d) Depth window 15600 m. The BMA and BSA represent the Beattie magnetic anomaly and the bean shape anomaly. (Baiyegunhi, et al., 2019).

There are several other seismic researches that have been conducted in the Karoo Basin that focused on understanding the reflectivity and velocity structure of the upper and lower crust. The reflection seismic imaging investigations produced by Lindeque et al. (2011) comprised of a seismic profile of 100 km length with 182 borehole shots. A continuous recording of the wave field of up to 60 hours was acquired as a "by-product" of the observational technique. As a main characteristic, they found that the upper crust of the Karoo basin is dominated by flat reflection zones, which are related to the Phanerozoic- Mesozoic angle and the sedimentary rocks of the Karoo supergroup (Figure 3.4).

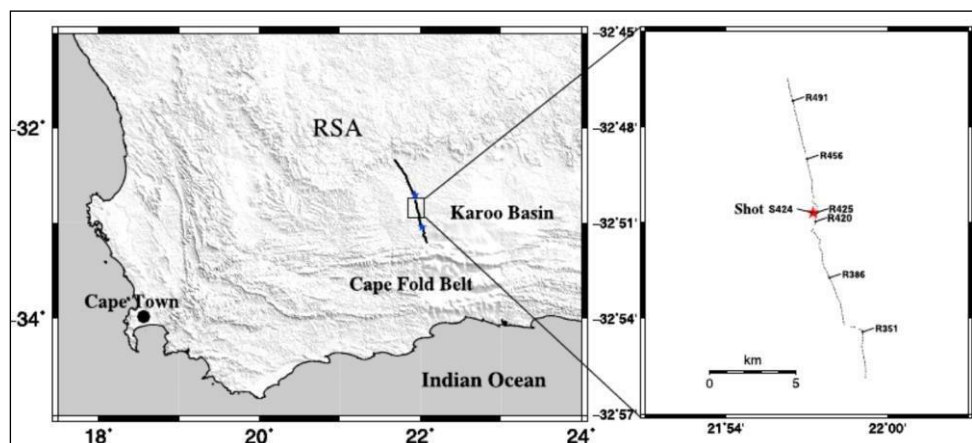


Figure 3.4(left) Map of Southern Africa showing the study area, (right) zoom on part (18 km length) of the seismic line indicating the location of the borehole shot (star) and the seismic recorders (dots) (Ryberg, 2011).

Lindeque (2011) also concluded that the area of the upper crust of the Mesozoic Cape and the sedimentary rocks of the Karoo Supergroup in the Phanerozoic (marked in pink, brown, yellow and red in Figure 3.5), increases in thickness increased by 5 km in the North. to the south 10 km. The inclined reflectors near the surface (marked in pink) are 5 km thick to the south and 2.5 km to the north and are interpreted as representative of the Karoo supergroup.

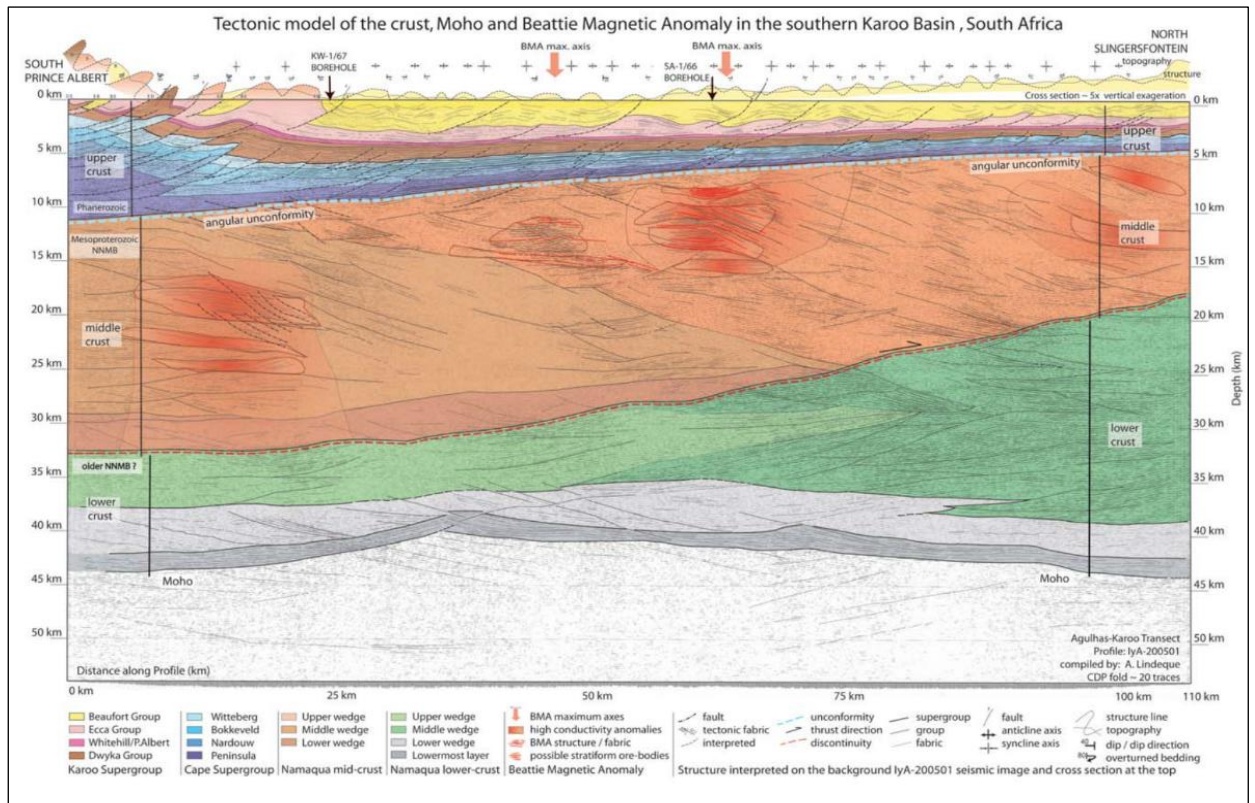


Figure 3.5: Interpretation of the migrated section and subdivision of the continental crust: the upper crust region above the sub-horizontal reflectors show dipping strata (annotated in pink) of the Karoo Supergroup. The uppermost continuous reflectors (yellow and orange) represent 5 to 10 km thick Phanerozoic Cape Supergroup sediments with a shallow dip to the south. An unconformity splits the upper crust from north dipping mid-crust granitoid gneisses of the NNOB (green). The lower crystalline crustal slab (blue lines) shows an internal fabric that dips both to the north and south, but mainly to the north. The boundary between the mid- and lower-crustal sections is not well defined. Reflectivity ~1 to 2 km above the undulating Moho depth suggests a possible additional layer. (Lindeque, et al., 2011).

Ryberg (2011) used the remaining recorded data to show that the analysis of processed ambient seismic noise can yield pseudo-shot gathers which have important body wave arrivals. He was able to detect that the high-frequency portion of the wave field contains direct (refracted) and reflected P-waves. These phases can be used as input for travel time tomography and were appropriate for imaging the shallow velocity structure. This study proved that the exploitation of ambient noise records for “source less” seismic reflection

imaging can have a huge potential and might be an efficient economical alternative to controlled source seismic imaging methods.

Surface waves are analyzed to compute group velocity and dispersion curves. The resulting group velocity curves used to compute tomography maps. Ambient Noise Tomography has been successfully used to map the subsurface of the Earth on a global and local scale.

Bezuidenhout et al. (2016) evaluated the usage of ambient seismic noise as a tool to map the subsurface of the Cape-Karoo transition in the Eastern Cape region. In this study, two passive seismic arrays were deployed in the Karoo. The recorded data that was acquired was processed and cross-correlated to obtain the empirical Green's function. The resulting cross-correlation functions were used to calculate group speeds which allowed for group velocity maps to be computed at depths between 3 and 7 km (periods between 3 to 7 seconds) to be obtained for the first time in the study area.

The maps obtained at different depths produced useful information about the geological structure of the subsurface of the Cape-Karoo region. The results showed two distinct velocity anomalies to the north and the south of the study area, which we can interpret as the thicker Karoo sediments and the younger less compacted sediments of the Algoa Basin, respectively.

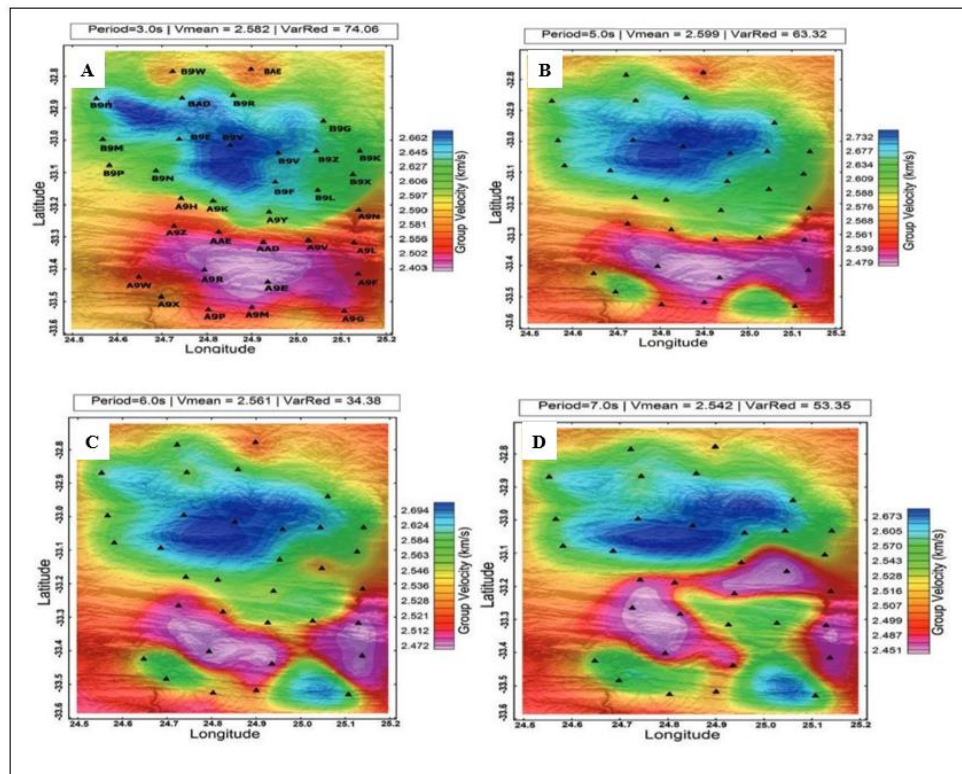


Figure 3.6: Rayleigh waves group velocity inversion maps using ambient seismic noise wave fields at periods: (a) 3.0 s; (b) 5.0 s; (c) 6.0 s and (d) 7.0 s. The black triangles show the location of the

stations. The blue colour indicates the high velocity regions, and the red colour indicates the low velocity domains. (Bezuidenhout, et al., 2016).

In a recent study, Bezuidenhout and Doucouré (2020) evaluated the cross-correlation functions that were computed between station pairs to determine the source of ambient noise source that was recorded in the Eastern Cape. The amplitudes were normalized, and station azimuths were determined from the positive and negative lags of the cross-correlation functions for periods between 1 and 10 s. The results that were obtained confirm that most of the ambient seismic noise recorded in the south eastern Karoo area originates from nonlinear interaction of ocean waves (secondary microseism) along the oceanic coast of South Africa (Figure 2.9).

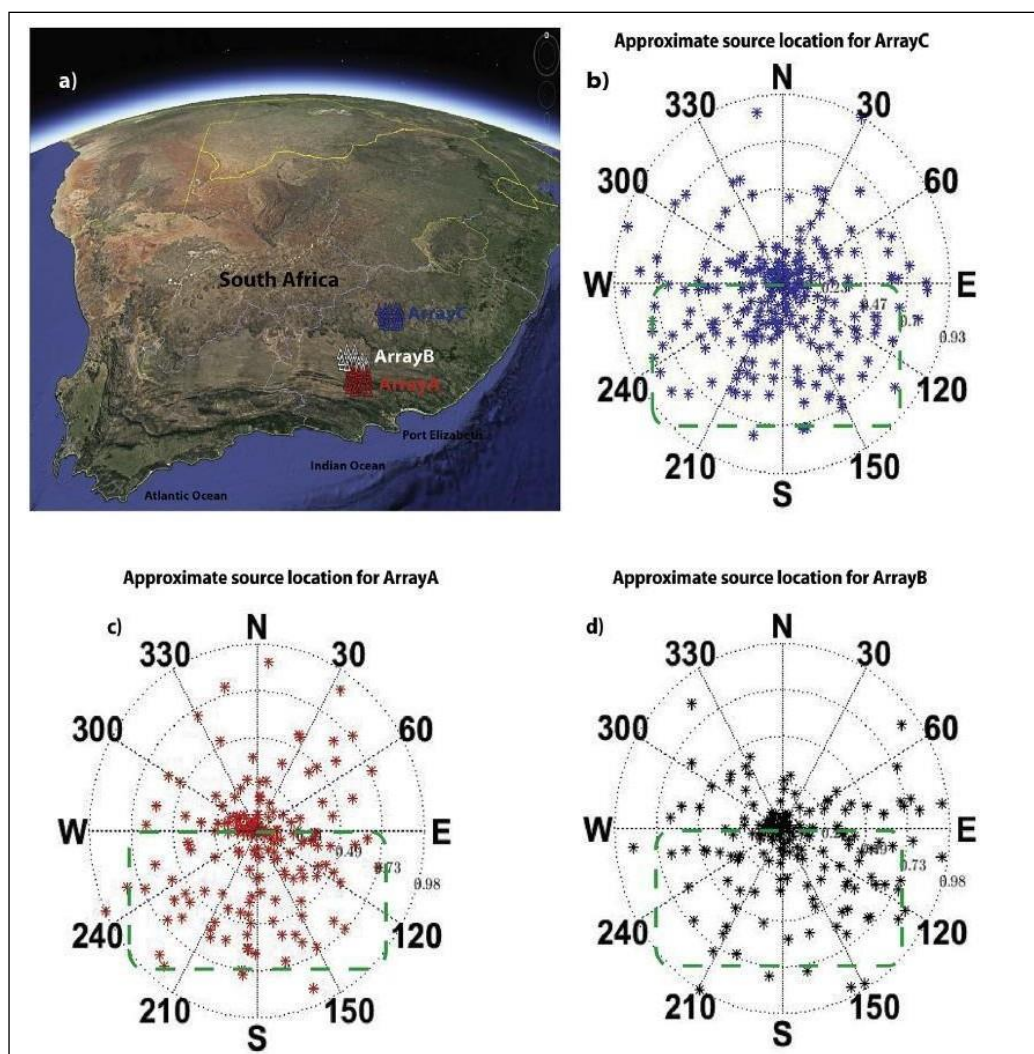


Figure 3.7: Array location representing position with respect to the ocean. b) Normalized amplitude of the cross-correlation function for both positive and negative lag plotted with its corresponding network azimuth for ArrayC. c) Same as for Array A d) same for Array B. (Bezuidenhout & Doucouré, 2020).

4. Tailings Dams

Tailings dams are earthen embankments that are used to store the by-products of mining operations after the gangue is separated from the ore (Azam & Li, 2010). Tailings are the materials that remain after the mineral separation process is complete. This chapter discusses tailings dams and provides a review of the different types of construction methods. Lastly, it provides a brief description of the common monitoring techniques that are applied by mining companies.

Tailings are different from waste rock, the latter being soil or rock that covers an orebody and is displaced during mining but is not processed. Tailings are a mixture of fine particles and water. They come from mineral processing, which involves crushing and grinding minerals into fine particles, which are then mixed with water and other reagents to facilitate the extraction of target minerals. After separation, water is removed from the tailings mixture and what remains is then stored in a tailings storage facility (Wallingford, et al., 2019).

The size and composition of the tailings depends on the mining method. The tailings that come from most hard rock mining companies like Gold mines is usually found in powder form or a very fine mud. Depending on the percentage of remaining water and solids, the tailings can be defined as slurry, paste or filtered tailings.

The water in the tailings mixture may contain certain chemicals related to the metal recovery process, such as cyanide in gold tailings and heavy metals in copper, lead, and zinc tailings (Bell, 2007). The tailings mixture may also contain sulphide minerals such as pyrite that can lead to acid mine drainage. Accordingly, contaminants transported in the tailings represent a source of pollution for both groundwater, surface water and soil.

4.1 Methods of construction

Many factors need to be considered when choosing the best site selection and construction method for tailings dams. Each dam is unique and is designed, constructed and operated according to the environmental and chemical characteristics of the tailings. Site conditions, such as topography, rainfall, seismic activity, and proximity to residential areas, determine the selection of a suitable tailings dam site.

The dams are usually constructed using readily available materials or dry tailings, rather than the concrete commonly used in water storage dams. The difference between the tailings dam and the impoundment dam is that the construction phase of the tailings dam will run through the entire operating cycle of the mine processing plant, sometimes for decades, and the parameters of the tailings dam will undergo many changes.

The operation and design period that occurred during this period. There are three main types of construction methods for tailings dams: upstream method, downstream method and centerline method.

(a) *The Upstream method*

The upstream method is the most commonly used method for constructing and erecting tailings dams. The main reason is that it is the most economical method and requires the least building material. This method starts with the construction of the starting dam. The tailings are separated naturally, the coarse material is deposited near the initial dam, and the liquid and fine material are deposited further away. With the increase of the material level, the dam crest is raised upstream with the support of the elevation of the front dam and the area of the tailings beach. The upstream embankments are not suitable for areas where the climate is arid, the amount of water that needs to be stored in the reservoir is very small, and it is impossible to quickly accumulate water (such as floods and upstreamwater flooding). The upstream embankment is not suitable for seismically active areas because the risk of liquefaction will increase due to potential dynamic seismic loads.

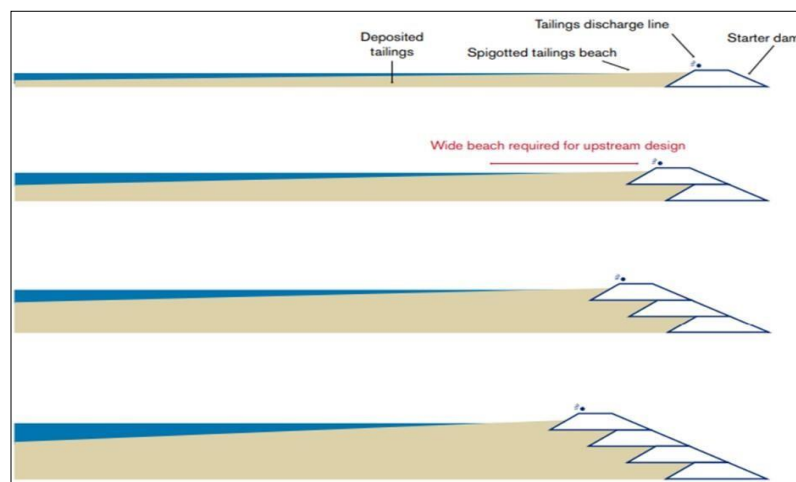


Figure 4.1: Diagram of a tailings dam constructed by upstream method (Cutifani, 2019).

(b) *The downstream method*

The construction of the downstream method also starts with the construction of the starting dam. The low permeability lining is installed in the starting dam to control and minimize water loss. The tailings are stored behind the dam, and the dam is raised by building a new wall on the upper and lower slopes. Therefore, the top of the dam moves downstream or away from the starting dam. This method requires more materials to construct than the upstream method, but its advantage is that the raised part can be designed to have a variable

porosity to solve any problems with the groundwater level of the dam. This method is more stable than the upstream method and is more suitable for areas with more seismic activity and rainfall.

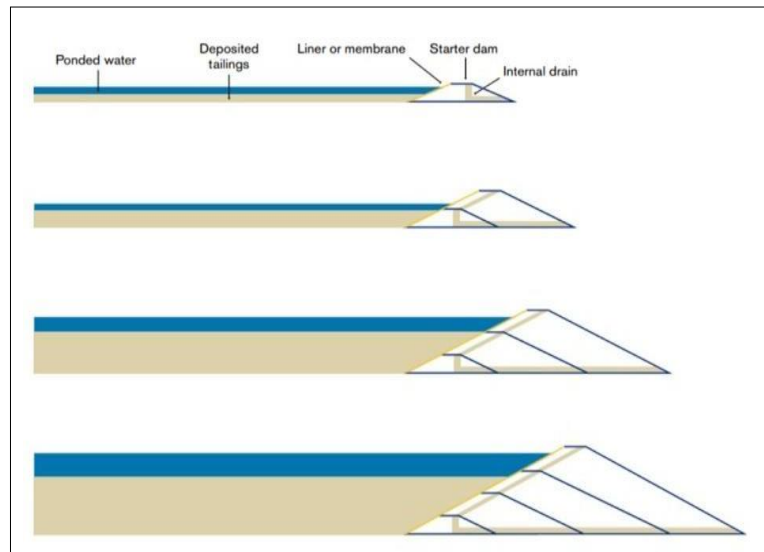


Figure 4.2: Diagram of a tailings dam constructed by downstream method (Cutifani, 2019).

(c) The centerline method

The centerline method is a compromise between both the upstream and downstream construction methods. Initially the starter dam is constructed and then the tailings are deposited on top of the dam to form a beach behind the dam wall. When subsequent raising is required, material is placed on both the tailings and the downstream slope of the existing embankment. The embankment crest is raised vertically on the centerline and does not move in relation to the upstream and downstream directions. This method is best suited in areas with moderate rainfall and moderate to high seismic risk.

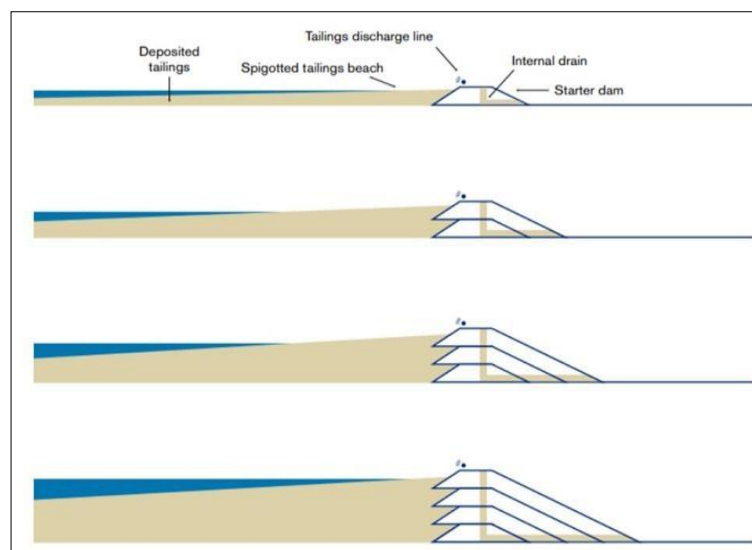


Figure 4.3: Diagram of a tailings dam constructed by centerline method (Cutifani, 2019).

4.2 Common monitoring and management techniques

There are several technologies that are utilized by mining companies to monitor and measure performances of tailings dams. The monitoring processes vary from manually measuring a dam wall position, to the installation of downhole instrumentation to measure pore pressure and crack structures. Typically, these activities are periodically undertaken with manual techniques that require personnel to be in the field, increasing the risk of fatigue or injury.

There are some companies that use Ground Penetrating Radar to monitor tailings dams. This method can monitor and detect movement on the embankment in real time and trigger an alarm. It works by sending radar pulses through the ground and timing the time it takes for the bounce to build an internal map of the ground. This method has proven to be most effective when used on solid structures such as cliffs or concrete retaining walls, so this technique is not ideal for tailings dams that will eventually move over time and potentially present new risks.

Terrestrial LiDAR

Unlike ground penetrating radar, LiDAR uses a laser scanner to measure the small movement of the dam wall, build a terrain model that can be compared to previous data, or programmed to continuously scan and trigger an alarm if movement is detected. LiDAR is a relatively new technology in the industry with few implementation examples. The adoption of this technology may be slowed down due to its high cost (Prevost, 2019).

It is also important to note that the effectiveness of this technique depends on the nature of the dam. For example, in the case of downstream tailings dams, continued expansion of the structure can cause false alarms or produce unusable erroneous data. LiDAR is more suitable for closed sites where the dam structure is completed. Due to the data, power and cost limitations of the LiDAR device, using it today requires very specific conditions to be successful.

Seismic sensors

These sensors are designed to detect earthquake-induced activities that can cause dam collapse but can also detect blasting operations and heavy plant movements. There are also sensors that use accelerometers and magnetometers to detect small movements in the walls of the dam, which can be configured to trigger alarms.

The seismic sensor also has the disadvantage that it sometimes produces false alarms, especially around active mines where blasting is common. These issues are manageable, but worth pointing out because they can complicate the data set. The sensors can be connected and designed to provide real-time data (Ganesh, 2006).

Total Station Automation

Automated site-wide surveys are something that many companies have or have considered. Typically, this form of monitoring is used in open pit mines to monitor slopes for wall movement. The principle of operation of the system is to set up multiple fixed measurement points at known locations on the wall or dam, and to configure one or more total stations in a hut or other protected area to measure these points at regular intervals; for example, Once every 30 minutes. The system then loads this data for surveyors to compare with old scans. This system has results similar to LiDAR or InSAR because it allows the user to see movement but has little information on the cause.

Real-time instrument monitoring

Real-time remote instrument monitoring describes connecting sensors such as piezometers, weather stations, ultrasonic altitude sensors, and phreatic board sensors to produce data at very close intervals. This method provides the most complete understanding of dam site conditions, helps to understand changes in site conditions overtime, and helps to receive alerts that indicate changing conditions. By adding powerful connectivity options (such as satellite), this method can be used anywhere in the world, all data is concentrated in a centralized system for site engineers, control centers, auditors and supervisors to help improve management tailings.

5. Passive Seismic Interferometry

The term interferometry refers to a type of measurement technique that uses the interference phenomena between pairs of signals or waves to obtain information from the differences between them (Curtis, et al., 2006). The technique makes use of the principle of superposition to combine waves in a way that will cause the result of their combination to have some meaningful property that is diagnostic of the original state of the waves. Passive seismic interferometry refers to the principle of producing new seismic responses of virtual sources by cross-correlating recorded ambient noise at different receiver locations (Rodrigues, et al., 2019). The technique is based on the extraction of the Green's function between two sensors.

The Green's function can be estimated by the cross correlation of long time series of ambient noise that is recorded between stations. A Green's function between two points is the signal recorded at one location due to an impulsive or instantaneous source of energy at the other. It contains information about how energy travels through the Earth between two locations (Wapenaar, et al., 2010).

It is possible to extract the Green's function continuously if there is an average random source of noise present and is persistent for long periods of time. This makes it particularly valuable for passive monitoring and provides an outline for studying possible temporal changes of the target of the survey.

In this chapter, the historical background of the ambient noise is briefly discussed and thereafter, the characterization of ambient seismic noise and the application of the technique are briefly introduced. In addition, the theoretical relationship between cross-correlation and Green's function and the ambient noise methodology is also discussed. Lastly, the dispersion measurement of surface waves is discussed in detail.

5.1 Historical development

The history of seismic interferometry dates back as far as 1968 when Jon Claerbout proved that it was possible to obtain the plane-wave reflection response of a horizontal layered medium from a plane wave transmission response of the same medium provided that the medium is one dimensional and bounded on top by free surface like the Earth's surface and below by a half-space or a homogeneous infinitely extensive Earth (Curtis, et al., 2006). Claerbout also concluded that by listening to noise at two receivers, it is possible to construct the signal that would have been observed if a source at one of the receiver locations was used. This method was applied to construct artificial seismic sources and was demonstrated and proven to be successful in helioseismology years later (Curtis, et al., 2006).

This meant that it was now possible to construct the Green's function from one point on the surface back to itself without the presence of a source emitting the waves.

Lobkis and Weaver (2001) established the feasibility of the technique experimentally on ultra sound waves and showed by calculations that the long-term average of random noise in an aluminium block produces the time domain impulse response between the two points. The development of the technique further verified that the Green's function can be extracted from the cross-correlation of surface waves between long time sequences of ambient noise observed at different stations. Claerbout's conjecture was finally confirmed mathematically by (Snieder, 2004) for acoustic media and (Wapenaar & Fokkema, 2006) for elastodynamic media.

Since then, research and development of the technique of seismic interferometry grew rapidly, each providing fundamental contribution to the understanding and knowledge of what we know the technique to be today. The first empirical seismological demonstrations of such an analysis was achieved by Campillo and Paul (2003) who showed that the cross-correlation computed from the recordings of a diffuse seismic noise wave field at two seismometers gives an approximate of the surface wave components of the Green's function between the two receivers as if one of the receivers had actually been a source. This was the first illustration of the geophysical reconstruction of the interferometric impulse response using noise and encouraged many geophysicists to consider the use of interferometric techniques in their fields.

Schuster et al. (2004) later showed that the cross-correlations of seismic responses that were resulted from man-made or natural sources at the surface could be used to form an image of the subsurface. Bakulin and Calvert (2004) produced the first practical application of seismic interferometry in an exploration setting confirming that it is possible to create a virtual source at a subsurface receiver location in practice.

As the years progressed, the technique grew increasingly popular amongst seismologists and the application in a variety of settings. The application of seismic interferometry to construct impulse responses for tomographic velocity inversions was shown by Shapiro et al. (2005) and by Sabra et al. (2005). By applying the virtual source method, they cross-correlated the noise recordings between multiple pairs of stations, thereby reconstructing the surface wave component of the impulse response between these stations. With tomographic velocity inversion, these surface waves are then used to generate a phase velocity map indicative of subsurface characteristics.

The history of interferometry shows that in relatively few articles, mainly in the last six years, seismic interferometry has changed from an unpopular and guess-based field to an extremely dynamic and active field. The field has great theoretical and practical progress being made almost monthly.

5.2 Ambient seismic noise

The ambient field is omnipresent, and its amplitude varies with position, time, and frequency. The surface of the Earth is affected by continuously acting forces caused by a variety of human activities, and by a wide range of natural phenomena. The ambient field generated by these forces constitutes the seismic background radiation of the Earth or seismic noise (Fichtner, et al., 2019). There are three dominant sources of ambient noise and can be grouped as short period, mid-period and long period (Kiwamu, 2017).

The most common source of short-period seismic noise is from human activity on or near the surface of the earth. This noise propagates mainly as short-period (0.1–1 s) surface waves that attenuate within several kilometers in distance and depth (Mcnamara & Boaz, 2019). Another common source of short-period seismic noise is observed from the interaction of wind with the surface of the earth. Wind can cause vibration of trees and structures as well as movement of debris, coupling energy into the solid earth.

Ambient seismic noise that occurs at mid period mainly results from microseisms observed in the primary (10-20 seconds) and secondary (5-10 seconds) bands (Yang & Ritzwoller, 2008). The seismic ambient noise spectrum of the earth is dominated by broad and dominant peaks near the periods of 7 and 14 seconds, principally due to seismic surface waves. This excitation, the microseism, is observed globally, even in deep continental interiors and is due to ocean swells that are caused by intense cyclonic low-pressure storm systems that have steep pressure gradients that cause strong surface winds which result in the transfer of atmospheric energy into the ocean subsequently becoming microseism energy (Bensen, et al., 2007). Primary microseisms are generated at the causative ocean swell period by direct pressure fluctuations at the ocean bottom in shallow water resulting from breaking and/or shoaling waves. The amplitude of the much higher power secondary microseism is proportional to the product of standing wave components of the ocean wavefield. The mid-period power also increases due to ocean wave activity created by the hurricane.

In general, the long-period bands (greater than 20 s) is of interest to the seismology community for earthquake surface waves and long-period resonances in the solid Earth. The sources of ambient noise are randomly and homogeneously distributed in space and

time (Wapenaar, et al., 2010). The conditions of the earth do not meet these conditions thus the best any data processing procedure can achieve is a plausible estimate of the Green's function (Fichtner, et al., 2019).

5.3 Theoretical background of Cross-correlations

Cross-correlations are an estimate of the Green's function. A wave field that is generated by a source can be represented by the convolution of a source wavelet and a Green's function (Wapenaar et al., 2010).

The basic idea of seismic interferometry is that the Green's function between two seismic stations can be estimated by cross-correlating long-time series of ambient noise recorded at the stations. A Green's function between two points may be thought of as the seismogram recorded at one location due to an impulsive or instantaneous source of energy at the other. The Green's function is important because it contains information about how energy travels through the Earth between the two locations. Figure 5.1 is an illustrative 1-D analysis of direct-wave interferometry. The figure shows a plane wave, radiated by an impulsive unit source at $x=x_S$ and $t=0$, propagating in the rightward direction along the x -axis.

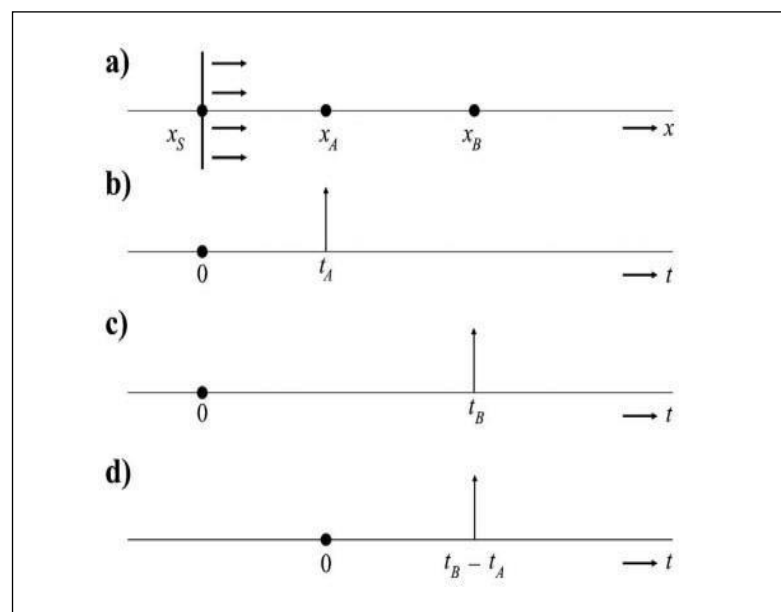


Figure 5.1: 1-D diagram showing an example of direct-wave interferometry. (a) A plane wave traveling rightward along the x -axis, emitted by an impulsive source at $x=x_S$ and $t=0$, the response observed by a receiver at x_A . This is the Green's function $G(x_A, x_S)$. (c) As in (b) but for a receiver at x_B . (d) Cross correlation of the responses at x_A and x_B . (Wapenaar, et al., 2010).

Figure 5.1 is a schematic diagram where a single wave propagates in the positive x direction. In this example, the propagation velocity c is assumed to be constant and the medium is lossless. There are two receivers along the x -axis at x_A and x_B . Figure 5.1 (a) is an illustration

of a simple cartoon model. A plane wave created by an impulsive source starting at time zero and position S, propagates from the left to the right in the horizontal direction, and is subsequently recorded by two receivers A and B. The recorded response is illustrated in Figure 5.1(b). The response observed by the first receiver at x_A and this response can be expressed as (Wapenaar, et al., 2010).

$$G(x_A, x_S, t) \quad 1$$

Where G represents the Green's function. The first two arguments in Equation 56 represent the receiver and source coordinates, x_A and x_S , whereas the last argument represents time t angular frequency ω . In this example, the Green's function consists of an impulse at;

$$t_A = (x_A - x_S) / c \quad 2$$

Therefore,

$$G(x_A, x_S, t) = \delta(t - t_A) \quad 3$$

Where $\delta(t)$ is the Dirac delta function. Similarly, the response at x_B is given by:

$$G(x_B, x_S, t) = \delta(t - t_B) \quad 4$$

The impulse shown in Figure 5.1(c) is then given by:

$$t_B = (x_B - x_S) / c \quad 5$$

The seismic interferometry technique involves the cross-correlation of responses at two receivers, in this case at x_A and x_B . In Figure 5.1(a), it seems that the ray paths associated with $G(x_A, x_S, t)$ and $G(x_B, x_S, t)$ have a common path from x_A to x_S . The travel time along this common path cancels in the cross-correlation process and leaves the travel time along the remaining path from x_A to x_B this then means

$$t_B - t_A = (x_B - x_A) / c \quad 6$$

Hence, the cross-correlation of the responses in Figure 5.1(b) and (c) is an impulse at $(t_B - t_A)$ as shown in Figure 5.1(d). 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 $G(x_B, x_S, t)$. the propagation velocity c and the position of the actual source x_S need not be known. The travel times along the common path from x_S to x_A compensate each other, independent of the propagation velocity and the length of this path.

The cross-correlation of the impulse responses at x_A and x_B is expressed as:

$$G(x_B, x_S, t) * G(x_A, x_S, -t) \quad 7$$

The asterisk represents temporal convolution, the time reversal of the second Green's function turns the convolution into a cross-correlation, defined as: (Wapenaar, et al., 2010)

$$G(x_B, x_S, t) * G(x_A, x_S, -t) = \int G(x_B, x_S, t + t') G(x_A, x_S, t') dt' \quad 8$$

Substituting the delta functions into the right-hand side gives the equation below:

$$\begin{aligned} \int \delta(t + t' - t_B) \delta(t' - t_A) dt' &= \delta(t - (t_B - t_A)) \\ &= \delta(t - (x_B - x_A) / c) \end{aligned} \quad 9$$

The Green's function propagating from x_A to x_B is expressed as $G(x_B, x_A, t)$. Since we started deriving the cross-correlation of the Green's function in this example, the 1D Green's function is represented as follows:

$$G(x_B, x_A, t) = G(x_B, x_S, t) * G(x_A, x_S, -t) \quad 10$$

There are many interesting phenomena may be deduced from Figure 5.1. The Green's function can be extracted from A to B without any knowledge about the original source location S, propagation velocity C and initial time of original source. This is because the propagating ray path associated with shares the same wave path from A to B. The travel times of signals spent along the common propagating path from S to A compensate each other, leaving the travel time along the remaining path from A to B. Equally, if there is a random initial source time rather than starting at time zero, the cross-correlation shifts the same amount of time from S to A, consequently, the absolute time when the source emits is cancelled in the cross-correlation, as well as the propagation velocity c and the paths from receiver A and B to source S.

The representation in equation 10 expresses the principle that the cross correlation of observations at two receivers x_A and x_B gives the response at one of those receivers x_B as if there were a source at the other receiver x_A . This shows why seismic interferometry is commonly referred to as the Green's function retrieval.

5.4 Application of Seismic Interferometry

The most broadly used application of passive seismic interferometry is the retrieval of surface waves between seismometers. The retrieved surface wave responses can be used to extract receiver-receiver phase and group velocities. The phase and group velocities frequently serve as input parameters for tomographic inverse problems and determination of the surface wave velocity distribution of the subsurface.

It has been established that in a layered medium, surface waves consist of several propagating modes in layered media, of which the fundamental mode is usually the strongest (Wapenaar, et al., 2010). If only the fundamental model is considered, surface waves be an approximate solution of a 2-D wave equation with a frequency dependent propagation velocity. The Green's function of the fundamental mode of the direct surface wave can thus be extracted by cross-correlating ambient noise recordings at two seismometers (Wapenaar, et al., 2010). When there are multiple seismometers available, the process can be repeated for any combination of two seismometers. In other words, each seismometer can be turned into a virtual source, the response of which is observed by all other seismometers. The responses obtained can then be used for tomographic inversion of the Rayleigh group and phase velocity of the crust (Breguier et al., 2007; Lin et al., 2009), and for measuring azimuthal anisotropy of the crust.

5.5 Dispersion Curves

Surface waves are generated in the presence of a free boundary, such as the surface of the Earth, and propagate parallel to this surface. There are two types of surface waves known as Love waves and Rayleigh waves after their discoverers and are classified with respect to the polarization of the ground motion during propagation (Foti, et al., 2018).

Love waves are the fastest type of surface waves and involve transverse motion moving the ground side-to-side (Figure 5.2a). The waves are confined to the surface of the crust and produce horizontal motion. Rayleigh waves involve elliptical motion (Figure 5.2b) and roll along the ground resulting in the ground moving up and down, and side-to-side in the same direction that the wave is moving (Rawlinson, 2016).

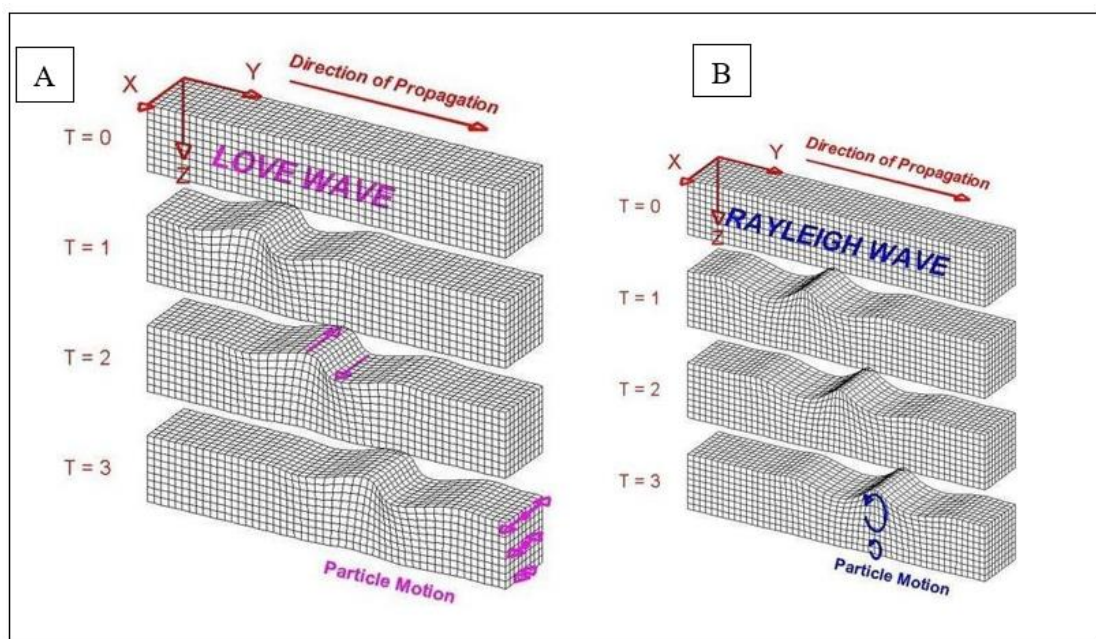


Figure 5.2: Illustration of direction of propagation of surface waves through a medium. Particles are represented by cubes in the models (a) In Love waves particle motion is horizontal and perpendicular to the direction of propagation (transverse) and (b) In Rayleigh waves, particle motion consists of elliptical motions (Boschi, 2014) in the vertical plane and parallel to the direction of propagation

For Rayleigh waves, the amplitude of the related motion decays exponentially with depth, becoming insignificant within about one wavelength (λ) from the surface in homogeneous media (Foti, et al., 2018). In vertically heterogeneous media, the decay of particle motion amplitude with depth cannot be predicted without knowledge of the subsurface structure. The velocity of Rayleigh waves depends on the elastic properties of the subsurface, mainly on the shear (S) wave velocity, and slightly on the compression (P) wave velocity and on the mass density (Boschi, 2014). Love waves do not exist in homogeneous media and in heterogeneous media Love wave velocity depends only on how VS and mass density vary with depth (Foti, et al., 2018).

In vertically heterogeneous media, the propagation of surface waves is governed by geometric dispersion, harmonic waves of different wavelengths propagate within different depth ranges (Figure 5.3(a)) and, hence, for each wavelength the phase velocity V depends on the elastic properties and density of the subsurface medium within the propagation depth range (Figure 5.3(b)). The distribution of phase velocities as a function of frequency or wavelength is called a dispersion curve (Figure 5.3(c)). In vertically heterogeneous media with increasing velocity with depth, the velocity of propagation of surface waves decreases for increasing frequency (Rawlinson, 2013).

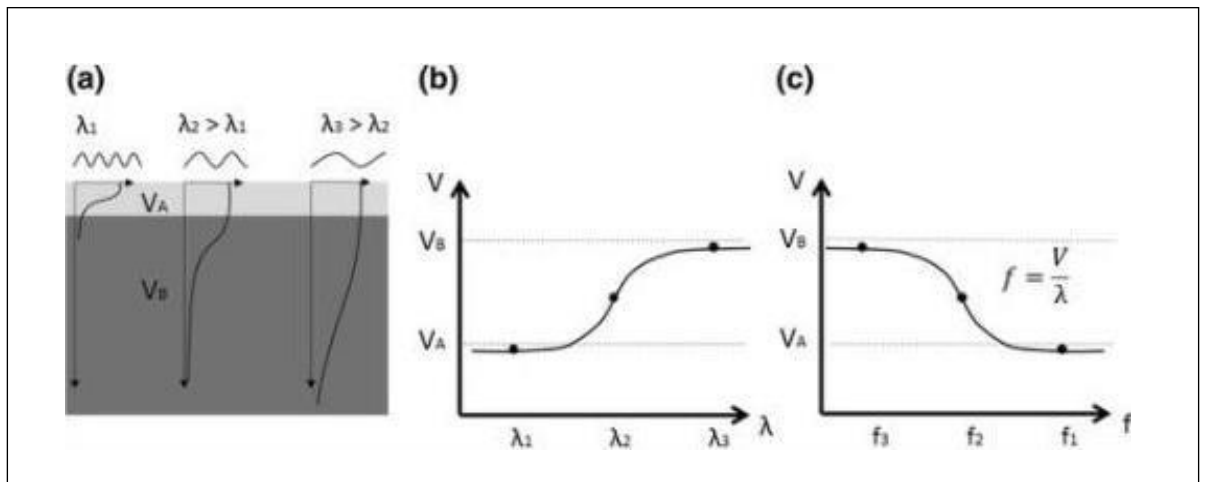


Figure 5.3: Geometric dispersion of surface waves in vertically heterogeneous media. λ is the wavelength of the surface wave with phase velocity V and f is the. V_A and V_B indicate the generic shear wave velocity in the two layers affected by the surface wave propagation. (a) Sketch of amplitude decay of the fundamental mode at different wavelengths, (b) dispersion curve in the wavelength—phase velocity domain, (c) dispersion curve in the frequency—phase velocity domain. (Foti, et al., 2018).

The basis of most surface wave analysis methods is to accurately determine the frequency dependent phase velocity of the fundamental mode Rayleigh wave (Park et al. 1998), i.e., the experimental fundamental-mode dispersion curve. There are different methods that are implemented to image and retrieve surface waves dispersion curves from ambient noise. In this study the multichannel analysis and frequency-time analysis methods were applied to retrieve dispersion curves.

The Multichannel Analysis of Surface Waves method

The Multichannel analysis of surface waves (MASW) is a non-destructive surface wave analysis method that is used to analyze and retrieve the dispersion properties of surface waves that propagate in a horizontal direction along the surface of measurement directly from impact points to receivers (hammer shot points to geophones, respectively), or from passive noise sources such as described in 5.2.

This method processes lower frequency surface waves (e.g., 1-30 Hz) recorded using a multi-channel recording system and a receiver array deployed at a distance of several hundred meters. The depth range of investigation is shallower (e.g., a few to a few tens of meters) (Park, et al., 2007). The maximum depth of investigation for ambient noise ranges with site, the configuration of the measurement profile and the natural frequency of the receivers. The main advantage of MASW is its ability to take into account the complex nature of passive seismic waves as well as fundamental-mode surface waves (Figure 5.4(a)). These waves may often influence each other in an unfavourable way during the analysis of their dispersion properties if they are not properly accounted for (Xia, et al., 2000).

The MASW approach images the dispersion properties of all types of waves (including body and surface waves) through a wave field-transformation method which converts the multichannel record directly into images in which specific dispersion pattern are identified in the transformed energy distribution (Figure 5.3(a)). Then extract necessary dispersion properties (like that of the fundamental mode) from the identified pattern. All other reflected/scattered waves are usually automatically removed during the conversion process. The procedure for MASW consists of three steps: (1) acquiring multichannel fieldrecords; (2) extracting dispersion curves; and (3) inverting these dispersion curves to obtain 1-D (depth) Vs profiles (Park, et al., 2007). Figure 4.4 illustrates the MASW process.

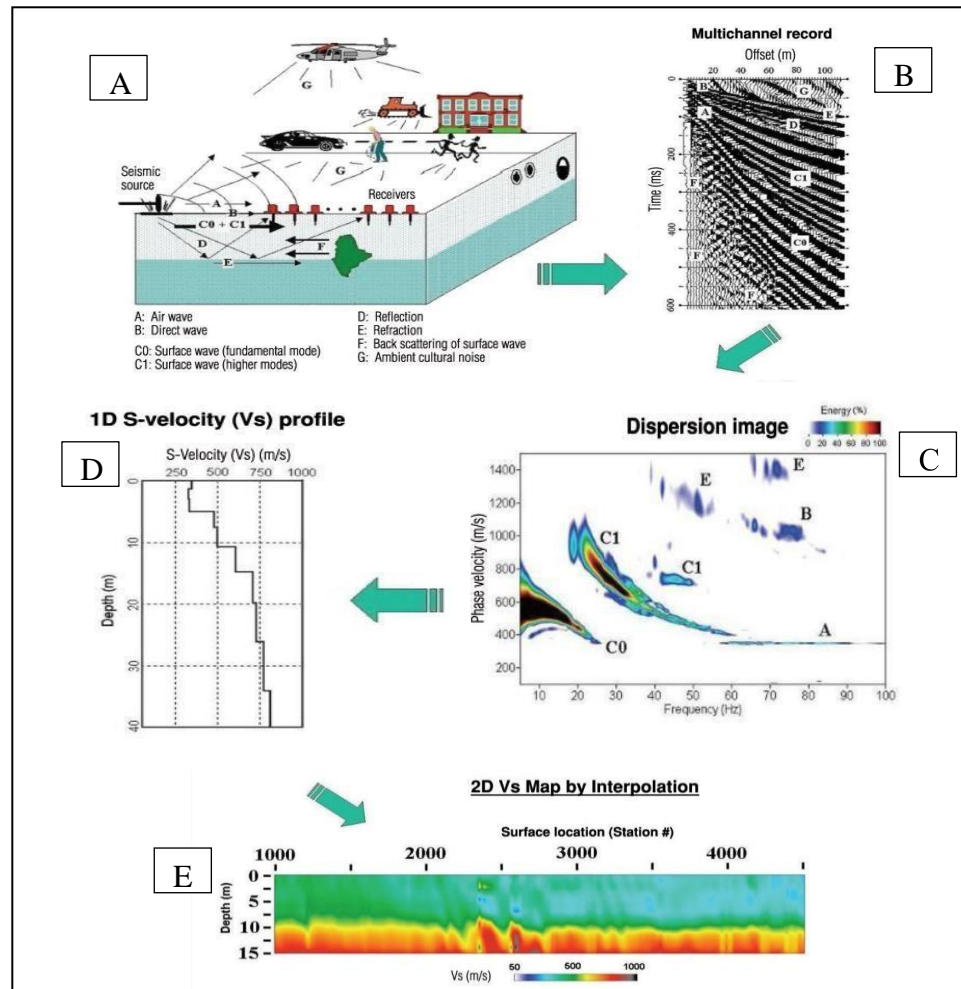


Figure 5.4: An illustration of the overall procedure and main advantage of the MASW method. (a) Complicated nature of seismic waves is carried over into the measurement (multichannel record). (b) Then, dispersion nature of different types of waves is accurately imaged through a 2-D wave field transformation. Certain noise wave fields such as back- and side-scattered surface waves and several types of body waves are automatically filtered during this transformation. (c) Dispersion curves are then extracted to be inverted for a 1-D VS profile (D) Computed 1-D profile from the dispersion curves, (D) Multiple 1-D profiles assembled and prepared to make a 2-D VS map. (Park, et al., 2007).

The frequency time analysis (FTAN) method

The frequency time analysis (FTAN) method is an efficient tool that can be used to study surface waves and useful in defining shear wave velocity profiles of shallow geological structures. The method was implemented by Levshin and colleagues and began to significantly improve the multiple filter analysis which Dziewonski and associates originally developed (Ndikum, et al., 2014). FTAN can be applied to single channel configuration, even if there is higher mode pollution, the group speed can be measured. When the source is known, the phase velocity can also be measured.

This method uses multiple narrow-band Gaussian filter systems, with a variable center frequency, does not introduce phase distortion and provides good resolution in the time-frequency domain. The waveform record is mapped in a two-dimensional domain; time (group velocity) and frequency (period). The group velocity measurement of Rayleigh waves and/or Love waves is carried out in the envelope of surface wave trains and can be robustly measured in a wide period band from a fraction of a second to hundreds of seconds.

For each filter frequency band, the square amplitude of the inverse Fourier transform of the filtered signal is the energy carried by the center frequency component of the original signal. Since the arrival time is inversely proportional to the group velocity, for a known distance, at a certain center frequency, energy is obtained as a function of the group velocity. Repeat the process for different center frequencies. The FTAN diagram is an image of a matrix, the columns of which are the energy values in a given period, and the rows are the energy values at a constant group velocity. A series of frequency filters and a time window are applied to the dispersion curve to facilitate the extraction of the fundamental mode. Combination of floating filter technology and phase equalization allows fundamental mode to be isolated from higher modes.

The computational structure of FTAN can be automated and the steps are defined as follows:

If a wave form is defined as $s(t)$. The Fourier transform of the waveform is defined with a positive exponent:

$$S(\omega) = \int_{-\infty}^{\infty} s(t) \exp(i\omega t) dt \quad 11$$

The dispersion measurements are computed from the analysis and properties of the analytical signal which is expressed in the frequency domain as:

$$S_a(\omega) = S(\omega)(1 + \text{sgn}(\omega)) \quad 12$$

The analytical signal is then transferred to the time domain by implementation of an inverse Fourier transform and is then expressed as:

$$S_a(t) = s(t) + iH(t) = |A(t)|\exp(i\phi(t)) \quad 13$$

Where $H(t)$ is the Hilbert transform of $s(t)$. The next step is to establish a frequency-time function. To do this, the analytical signal is put through a set of narrow band pass Gaussian filters with center frequencies ω_0 :

$$S_a(\omega, \omega_0) = S(\omega)(1 + \text{sgn}(\omega))G(\omega - \omega_0) \quad 14$$

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

Each band-passed function is transferred back to the time domain to produce a smooth 2-D envelope that is expressed as:

$$y(t) = |A(t, \omega_0)| \quad 16$$

The phase function is denoted as:

$$p = \phi(t, \omega_0) \quad 17$$

α in Equation 15 is a parameter that can be adjusted and describes the balancing resolutions in the frequency and time domains and is dependent on the distance. The group speed is measured using the function $|A(t, \omega_0)|$ and phase speed using $\phi(t, \omega_0)$. The group arrival time is denoted as $\tau(\omega_0)$ and is a function of the center frequency of the Gaussian filter and can be calculated from the peak of the envelope function so the group speed is then expressed as:

$$U(\omega_0) = r / \tau(\omega_0) \quad 18$$

Where r is the interstation distance and ω_0 is the instantaneous frequency. The instantaneous frequency is the time rate of change of the phase of the analytical signal at time r . Spectral leakage causes the spectrum of the input waveform not to be flat and this results in the central frequencies of the narrow-band filters to inaccurately represent the frequency content of the output filters. To correct this, the center frequency ω_0 of the narrow-band filter should be replaced with the instantaneous frequency which is defined as:

$$\omega = |d\phi(t, \omega_0) / dt|_{t=\tau(\omega_0)} \quad 19$$

Phase speed measurements

The group speed curve can be measured by analyzing the envelope function $|A(t, \omega)|$. Phase speeds cannot be measured directly from the group speed but group speed can be measured from the phase speed (Bensen, et al., 2007). To do this we define the group speed as:

$$U = \partial\omega / \partial \quad 20$$

The phase speed is denoted as:

$$c = \omega / k \quad 21$$

The group slowness is denoted as:

$$s_u = U^{-1} \quad 22$$

And the phase slowness as:

$$s_c = c^{-1} \quad 23$$

Where k is the wavenumber. Substituting U from Equation 20 in Equation 22, s_u can then be denoted as:

$$s_u = \partial k / \partial \omega = \partial(\omega S_c) / \partial \omega$$

The first-order differential equation relating to the group and phase slowness at frequency ω :

$$\frac{\partial S_c}{\partial \omega} + \omega^{-1} S_c = \omega^{-1} S_u \quad 25$$

If the phase speed curve $c(\omega)$ is known, the group speed curve $U(\omega)$ can be calculated directly from Equation 20. The differential equation must be solved to calculate the phase speed. This involves an integration constant that is generally unknown. The solution to the differential equation is as follows:

$$S_{c(\omega)} = \omega^{-1} \left(\int_{\omega_n}^{\omega} S_u(\omega) d\omega + \omega_n S_c^n \right) \quad 26$$

The constant of integration is written in terms of a boundary condition that the phase speed curve is known at some frequency:

$$\omega_n: S_c(\omega_n) = S_c^n \quad 27$$

To measure the phase speed curve, the information and envelope function on which the group speed has been measured must be known. This can be determined from the phase of the analytic signal which is composed of a propagation term, an initial source phase and a phase ambiguity term. The instantaneous frequency is denoted as:

$$\phi(t, \omega) = k\Delta - \omega t - \phi_s - \phi_a \quad 28$$

Where t is the travel time, Δ is the interstation distance, k is the wavenumber, ϕ_s is the source phase and ϕ_a is the phase ambiguity term. The phase at the observed group arrival time can be evaluated to find the phase slowness which is expressed as:

$$S_c = S_u + (\omega\Delta)^{-1}(\phi(t_u) + \phi_s + \phi_a) \quad 29$$

Where t_u is the group arrival time and is defined as $t_u = \Delta / U$ and $k = \omega S_c$

Equation 56 denotes the phase slowness and the phase speed curve. The use of this information depends on knowledge of the initial source phase and extra phase ambiguity term. Earthquake and ambient noise studies, the phase ambiguity term contains a part derived from the 2π ambiguity inherent to any phase spectrum:

$$\phi_a = 2\pi N \quad 30$$

Where $N = 0, \pm 1, \pm 2, ..$ The ambiguity term can be resolved by using a global 3-D model or phase velocity maps to predict phase speed at long periods. The value of N then is chosen to give the closest relation between these predictions and observations.

Snieder (2004) demonstrated that the cross-correlation phase between the displacement waveforms has a $\pi/4$ term, which rises in the lateral direction of the two stations from the fixed phase integration (effectively exceeding the source). in the direction transverse to the two stations. The sign of this term depends on the seismograph component. The vertical component of the Rayleigh wave is positive, and the radial component is negative. However, it is assumed that the sources are evenly distributed in azimuth. Non-uniform distribution will produce different phase changes, and since this distribution will vary with frequency, the variation may depend on the frequency Lin et al. (2007) shows that for the velocity waveform, the rate of the vertical component seems to be $-\pi / 4$ for the vertical component. Thus, the phase for vertical component ambient noise cross-correlations between velocity waveforms is expressed as (Lin, et al., 2007):

$$\phi_a = 2\pi N - \pi / 4 \quad 31$$

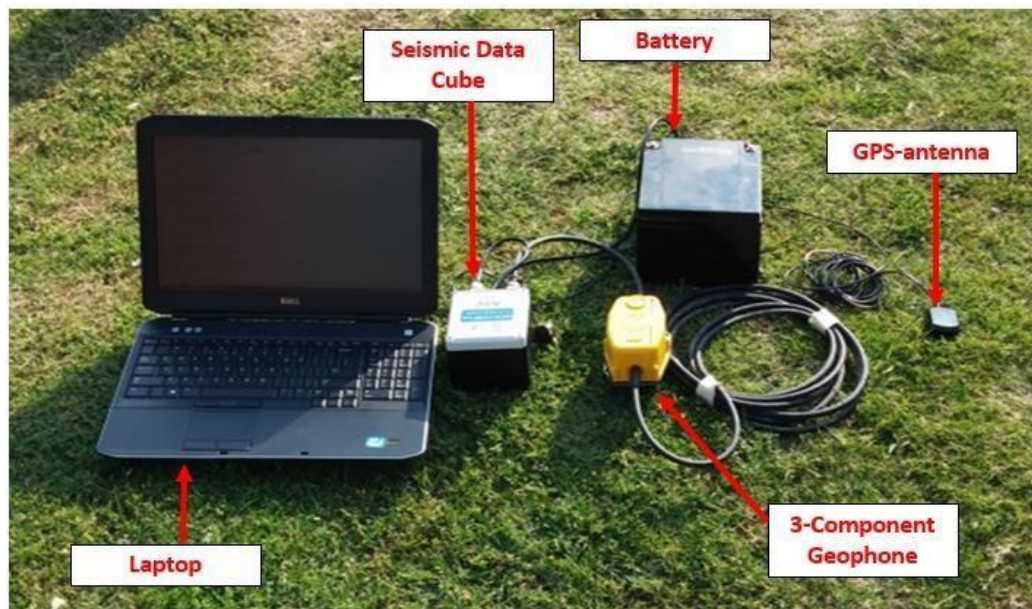
The value for phase-slowness derived from a vertical component ambient noise cross-correlation can be written as:

$$S_c = S_u + (\omega\Delta)^{-1}(\phi(t_u) + 2\pi N - \pi / 4) \quad 32$$

6. Data Acquisition

Equipment

The same set of instruments were used to acquire both datasets from the two locations. The data was acquired by installation of seismic stations in the ground. A station consists of a seismic data cube, a 3-component geophone, a battery and a GPS. The 14 Hz geophone measures ground motion in 3 directions (i.e. vertical and two horizontal directions). The seismic data-cube is a 3-channel, 24-bit seismic data recorder with an internal data storage capacity of up to 32GB. This cube stored the data that was recorded by the geophone. The battery provided power to the station and the GPS was used for time synchronization. This chapter covers data acquisition and describes the field installation steps. The distribution and deployment of the station for each location is also described. Lastly, the chapter outlines the processing workflow.



Photograph 1: Picture of the equipment that was used to collect the data.

Field Installation

The battery was connected to the external power output on the seismic data cube through a 4-way plug. The external GPS antenna was also connected to the seismic cube using the BNC socket and finally the geophone was connected to the seismic cube through socket in the upper corner for data recording. It is recommended to wait typically 3-5 minutes for reception of the GPS signal to test if the coordinates that were added in the configuration file match the location where the stations are to be installed. Adopting the correct GPS coordinates in the configuration file speeds up the GPS synchronization significantly. The seismic cube is ready and starts recording when the red and the blue LEDs are blinking at 1 second interval.

Once synchronization was achieved, a 40-50 cm hole was dug in the ground; the battery and seismic cube were placed in a 30 cm plastic box. The geophone was placed in the ground outside the plastic box and ensured that it made good contact with the surface. The box and the geophone were covered with soil and the GPS was left outside the hole.

6.1 Study area 1

A total of 20 seismic stations were deployed on a survey line of 100 m with an interstation distance of 5 m. The data was sampled at a frequency of 400 Hz. The seismic stations were installed on the tailings wall that was approximately 1.5 m high. The equipment was left to record continuous data over a period of 3 days. Figure 6.1 (a) below shows an aerial view of the tailings dam where the stations were installed and figure 6 (b) shows the distribution of the seismic stations. The stations were placed in between cone penetration tests that were installed on the tailings dam and above a seepage area which was seen at the foot of the tailings dam.



Figure 6.1: (a) Aerial view of the tailings dam where the seismic stations were installed. (b) Zoomed in picture of figure 6(a) showing the distribution of the seismic stations on the tailings dam and the location of the cone penetration tests. The area with seepage is also indicated in the picture.

6.2 Study area 2

A total of 24 seismic stations were deployed with an interstation distance of 5 km. The stations were left to record continuous seismic data over a period of 35 days. The data was sampled at a frequency of 100 Hz.

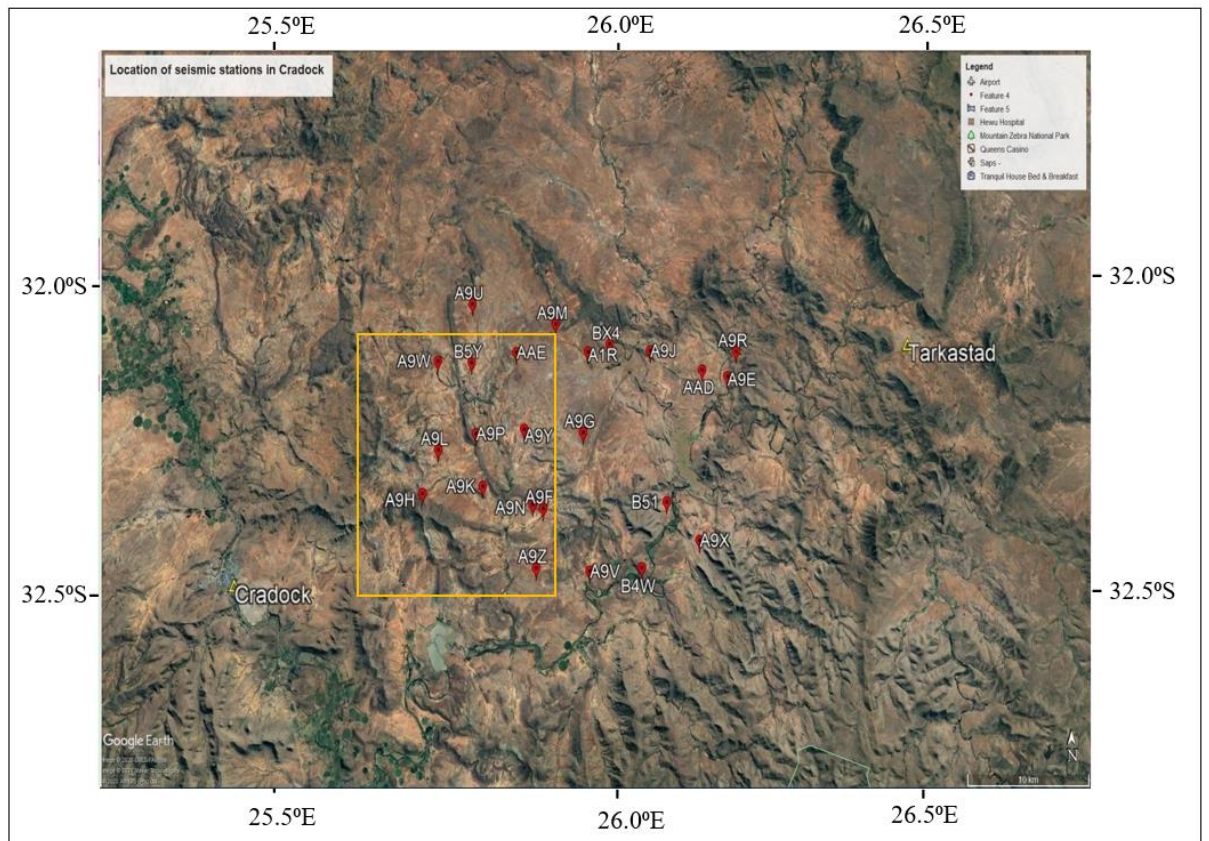


Figure 6.2: Google Earth image showing the distribution of the seismic stations in Cradock.

6.3 Data Processing

Prior to pre-processing and single station preparation, the continuous ambient seismic noise waveform data collected from all the components was converted from binary to Miniseed format. The conversion divided the data into daily 24-hour long segments. The data was then converted from Miniseed to ASCII and then lastly from Miniseed to SeisComP Data Structure (SDS) for further waveform analysis in MATLAB and MSNoise.

The single station data processing procedure that was applied was adopted from (Bensen, et al., 2007). The processing process is divided into three principal phases: (1) single station data preparation, (2) Calculating cross-correlation and temporal stacking, and (3) Computation of dispersion curves and inversion of ground profiles. In this project a MATLAB code was generated for the first two phases of the data processing procedure. The diagram below is a summary of the processing workflow.

Processing Workflow

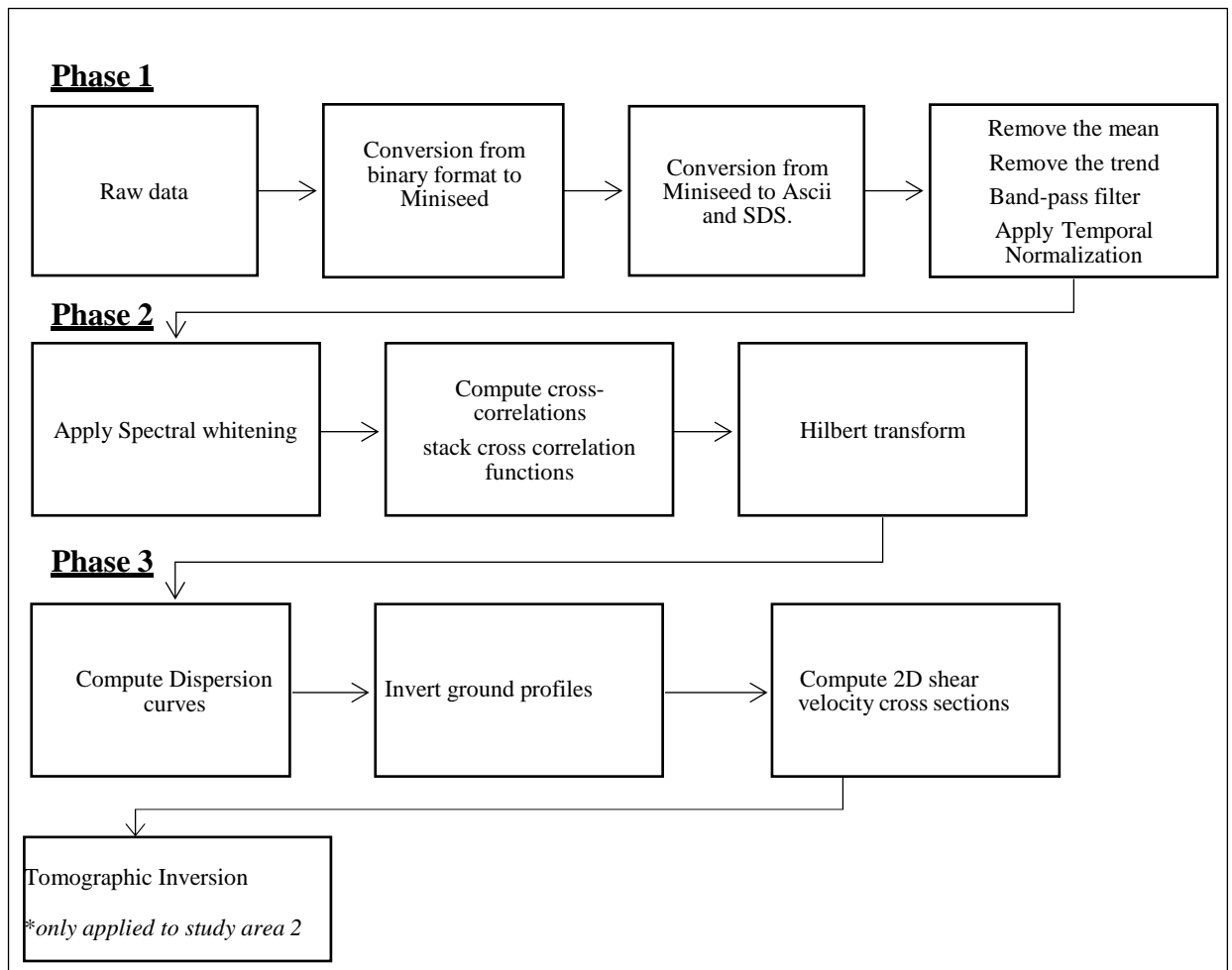


Figure 6.3: Schematic representation of the data processing scheme. Phase 1 shows the steps involved in preparing single-station data prior to cross-correlation. Phase 2 outlines the cross-correlation procedure and stacking. Phase 3 includes dispersion measurement and inversion process.

Data availability plot

The data that was acquired from the tailings dam comprised of the vertical component which is the Rayleigh wave and the two horizontal components are the Love waves. The two horizontal components were used in the processing. Once the data conversion process was completed, the data from all the stations was plotted to analyze the continuity of the recorded waveforms and to see which stations did not record data for the entire duration that they were expected to record the data for. Figure 6.4 is the data availability plot for the stations that were installed at the tailings dam. The plot shows that stations A9P and A9Y did not have data available for the first 16 hours of day 1 (28/08/2017) and were excluded in the processing. The data processing is discussed in the next chapter.



Figure 6.4: Data availability plot of seismic stations acquired from the tailings dam. Plots the data availability, as contained in the database. Every day which has a least some data will be coloured in red. Days with no data remain blank.

7. Data Processing

An understanding of the subsurface geology and soil profile of the area investigated is imperative when applying surface wave tomography. In earthen dams, it is important to delineate the soil saturation and the location of the phreatic surface. Raleigh waves are generated by the interference of the interactions of P and S_V waves with a free surface. It has been shown that P- and S-wave velocities have opposite responses to saturation, with P-wave velocities increasing and S-wave velocities decreasing with saturation (Carcione & Helle, 2002). This means that in areas where saturation is an important consideration, Raleigh waves can give ambiguous results when inverting the dispersion curve to determine the velocity profile of the shear wave. Love waves are only produced by the interaction of SH waves with free surfaces. Therefore, the velocity response of the Love wave is unique and decreases with saturation. The Love wave was chosen to process the data obtained from the tailings dam. Love wave processing needs to process two horizontal components at the same time.

This chapter presents the results from the ambient seismic noise data processing flow used to obtain the empirical Green's function from raw data. The processing workflow is discussed for both study areas. In addition, the chapter provides the spectral analysis, which shows the frequency range where the ambient noise is most energetic. The chapter also shows computed dispersion curves and 1-D shear wave velocity profiles obtained from inversion of the resulting dispersion curves. The chapter also shows the computed tomography maps that were obtained in the Cradock study area

7.1 Study area 1: Harmony Dataset

Phase 1(Pre-processing)

The first phase of data processing involves preparing waveform data from each station individually. The purpose of this phase is to accentuate the ambient noise signals by attempting to remove signals that resulted from earthquakes or instrument irregularities that tend to mask ambient noise. The mean and the trend of the daily component (E, N, and Z) were removed. Figure 6.1 shows raw wave forms from station A9M and A9J.

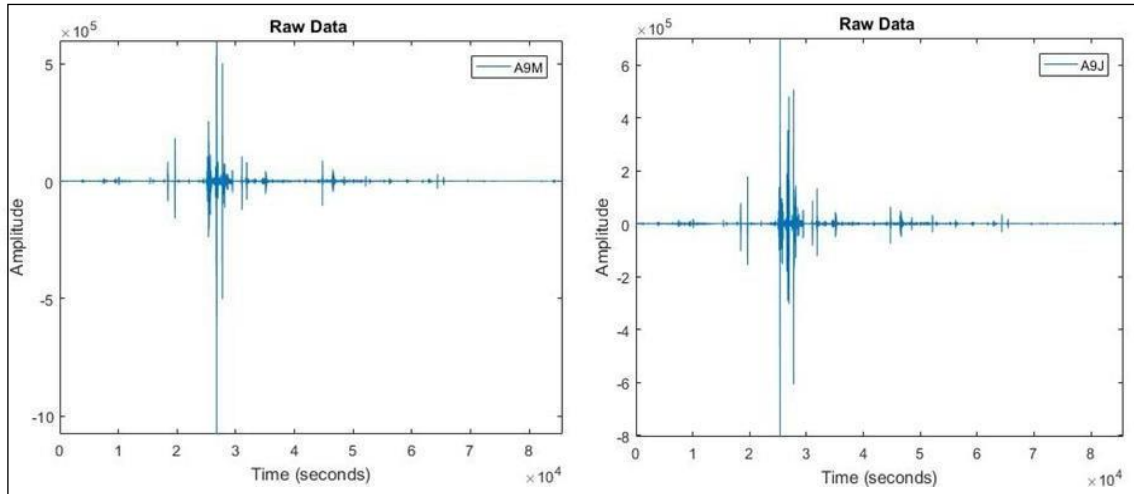


Figure 7.1: Data from station A9M and A9J, after the mean and linear trends are removed. The length of the time-series is 82800 seconds and the amplitude of the data is presented on the y-axis in digital counts which is an unscaled measure of displacement.

Band-pass filtering

It is important to determine what frequency range should be considered to design an appropriate filter. The objective was to image the top 20 m of the dam wall. This requires surface waves with a wavelength of at least 20 m. The wavelength is given by (Akademi, 2009):

$$\lambda = v/f \quad 33$$

The wave velocity is denoted by v and f is the frequency. The average shear wave velocity of sandy soil like that of the tailings dams is estimated to be approximately 200 m/s and the surface wave velocity 140 m/s (Chen & Ong, 2012). Using the wave formula, the lowest frequency that should then be considered is 7 Hz and the highest frequency selected for the filter is 40 Hz. The band pass filter limits were chosen to be 7 and 40 Hz. Ambient noise attenuated rapidly at frequencies higher than 40 Hz.

Soil classification	Shear wave velocity (m/sec)
Very soft soil	84–106
Soft soil	107–137
Medium soil	138–183
Stiff Soil	184–274
Very stiff soil	275–366
Soft rock/cemented soil	367–610
Rippability limit	670
Rock	>670
Concrete	2,286–2,438

Table 1: Typical Shear Wave Velocities of different types of soil (Chen & Ong, 2012).

Butterworth was chosen over the other filters because has the flattest possible pass-band magnitude response. The data was band-passed using a Butterworth filter between frequencies of 7 and 40 Hz. Band-passing the data ensured that any long period trends

associated with instrument glitches were removed. The presence of long period trends can create large distortions in the frequency spectrums of the data by invalidating the estimation of low frequency spectral quantities.

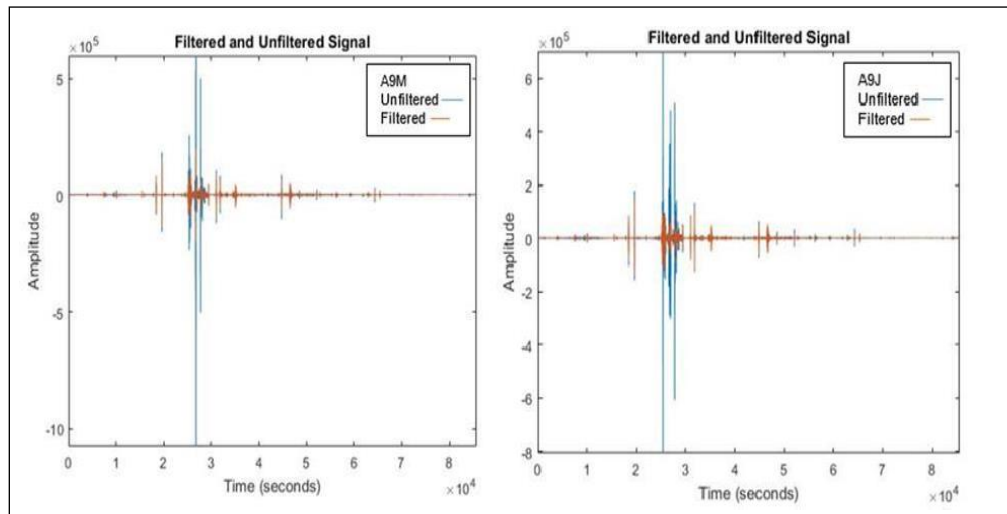


Figure 7.2: The raw signals (blue) from the two stations plotted with the filtered signals (Orange). Spikes and samples with high amplitudes have been removed and are not presented in the orange signal

The spectrograms show the temporal range of the frequency content over the duration of the recording period. There are various dominant peaks above 50 Hz (Figure 7.3 a-b) which is attributed to anthropogenic effects. The most common source of short-period seismic noise is human activities on or near the earth's surface. This is often referred to as "cultural noise", mainly due to the coupling of automobile traffic and mechanical energy to solid ground (Mcnamara & Boaz, 2019). Noise from cultural sources propagates mainly as short-period (0.1–1 s) surface waves that attenuate within several kilometers in distance and depth.

The spectrogram also shows that most of the seismic energy is concentrated below 50 Hz. The spectrograms were also computed to check if the filter that was designed was able to cut off the unwanted frequencies. Figure 7.3 (c-d) is spectrogram after the band pass filter was applied and shows that the designed filter was successful in cutting-off unwanted frequencies above 50 Hz.

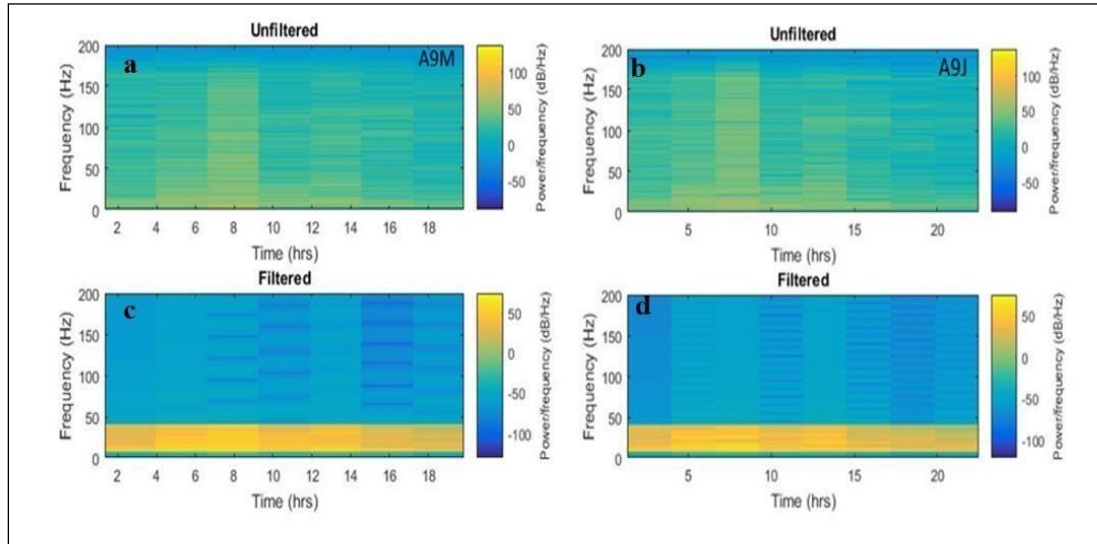


Figure 7.3: Spectrogram showing the filtered and filtered signal. The spectrograms of the signal below show a frequency band between 7 and 40 Hz, all other frequency components that were not in the filtered range have been removed.

The data was then parsed into ten-minute long segments. The reason for cutting the data is to provide a zoomed in picture of the signal. This is important for further analysis as it provides a clear picture of events and spikes that are not clearly visible in the signal. The sampling rate for each station is 400 samples per second, a ten minutes long data contains 240000 samples. Only 23.66 hours' worth of data was computed and prepared for the processing, there are 142 ten-minute intervals in 23.66 hours. For a single station's entire day of data, a total of 142 ten-minute long segments were computed in preparation for the next phase of data processing.

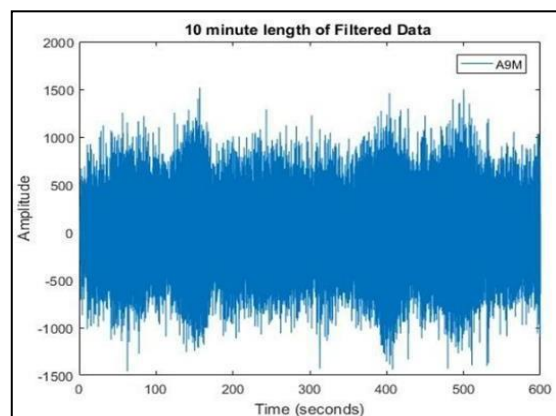


Figure 7.4: Segment of filtered data that represents the first ten minutes that were recorded by station A9M.

Time domain normalization is an important step in single station preparation. This procedure reduces the effect of cross-correlated earthquake signals, instrumental irregularities and non-stationary noise sources near to the stations. The method for temporal normalization that was used is the one promoted by (Bensen, et al., 2007) and is

called running-absolute-mean-normalization. This method calculates the running average of the absolute value of the waveform in a normalization time window of fixed length and weights the waveform at the center of the window by the inverse of this average. The normalization weight was applied to the filtered data using the formula below:

$$W_n = \frac{1}{2N + 1} \sum_{j=n-N}^{n+N} |dj| \quad 34$$

Where time is presented on a discrete grid. At a time grid point dj , the weight is the absolute mean of the signal in a time window of length N spanning the time point. The width of the normalization window $(2N + 1)$ determines how much amplitude information is retained. Typically, N is chosen to be half of the maximum period of the passband filter, this ensures that sufficient amplitude information of the signal is retained without rescaling the original signal. The filter limits were chosen to be 7-40 Hz, thus 0.0714 was chosen as the normalization window.

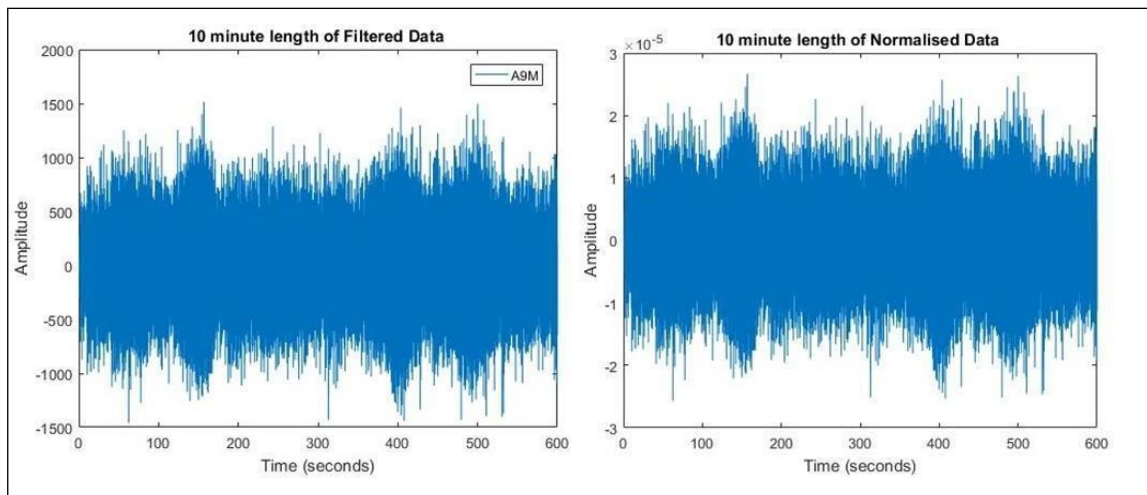


Figure 7.5: (a) Filtered data before temporal normalization using the running absolute mean method. (b) Filtered data after temporal normalization.

The last step in this phase of data processing is spectral whitening. The spectrum of ambient noise is not flat or uniform in the frequency domain. Whitening flattens the spectrum and acts to broaden the bandwidth of the ambient noise and reduces the impact of spatially isolated persistent noise sources. A function named Whitening was defined in MATLAB to perform spectral whitening. This function produces flat Fourier spectrum for a given signal (which is originally not white) either for the full range of 0 Hz to the Nyquist frequency or for a user defined frequency band. The process includes Fourier transforming the signal after applying Hann window, then normalizing its magnitude, and then inverse Fourier transforming it. As observed in Figure 7.6(a) The amplitude spectra of the waveform before whitening is not broad and uniform, there are several peaks that have different amplitudes that can be observed. As illustrated in Figure 7.6 (b) the result after application

of this step is a broader and uniform spectrum of the signal. Spectral whitening enhanced frequencies with low amplitude.

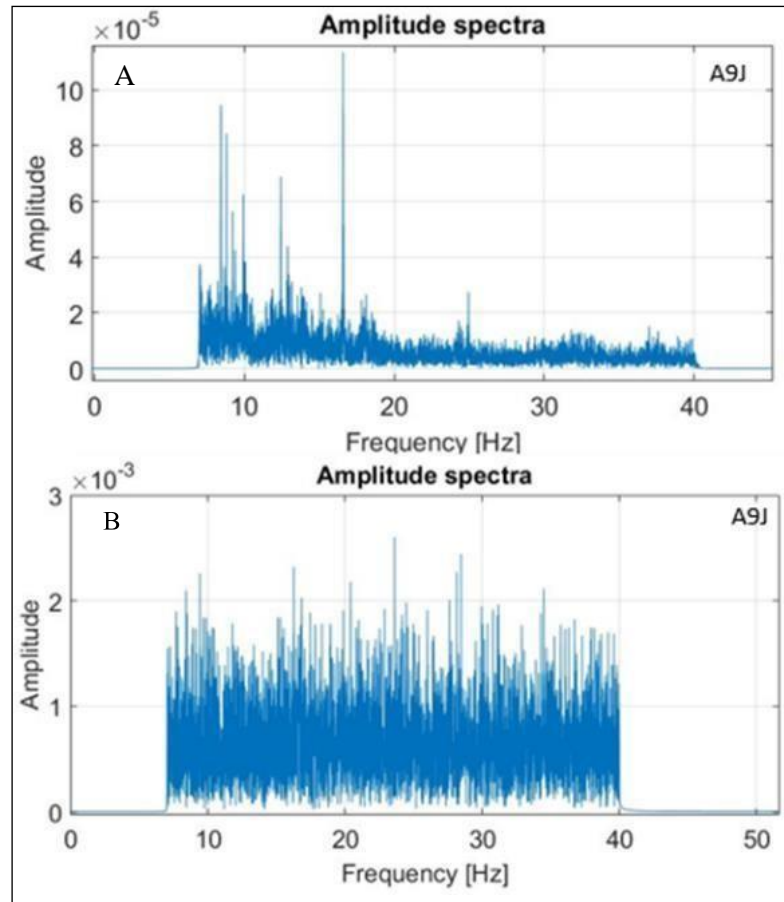


Figure 7.6: (a) An example of the amplitude spectrum of station A9J before spectral whitening and (b) amplitude spectrum after spectral whitening is applied

Phase 2

The first step of phase 2 of the data processing procedure is cross-correlation (obtaining empirical Green's function) between all station pairs. Cross-correlation (CC) is the process of measuring the similarity between two signals. The theoretical aspect of obtaining the empirical Green's function through the process of cross-correlation was discussed in section 5.3. The data that was prepared to be cross-correlated was from 17 stations. For the 17 stations used in the study, the maximum number of station pairs that were correlated were derived from the formula below (Lecocq, et al., 2014):

$$\frac{N(N-1)}{2} \quad 35$$

Where, N is the number of stations, $N = 17$ stations were used. Therefore, the total number of station pairs that were cross-correlated were 136. The 10-minute segments for each station pair were cross correlated using a function called `xcorr` in MATLAB. The cross-correlations

were performed on the daily pre-processed data in the frequency domain. Performing cross-correlations can be computationally intensive, so it is recommended that the procedure is done in the frequency domain to reduce computational time and speed (Bensen, et al., 2007).

The correlation is calculated for all components (north-north, northeast-east, east-east, and east-north) between each pair of sites recorded every 10 minutes. Because all operators are interchangeable with the rotation operators, the transverse—transverse, transverse—radial, radial—radial and radial—transverse cross-correlations between each station-pair can be determined by the linearity of those four components with coefficients related to the interstation azimuth θ and back azimuth ψ angles. These angles are defined by setting the first station as the event position and the second station as the receiver position so that the rotation can be expressed by the following equation (Lin, et al., 2008);

$$\begin{pmatrix} TT \\ RR \\ TR \\ RT \end{pmatrix} = \begin{pmatrix} -\cos\theta\cos\psi & \cos\theta\sin\psi & -\sin\theta\sin\psi & \sin\theta\cos\psi \\ -\sin\theta\sin\psi & -\sin\theta\cos\psi & -\cos\theta\cos\psi & -\cos\theta\sin\psi \\ -\cos\theta\sin\psi & -\cos\theta\cos\psi & \sin\theta\cos\psi & \sin\theta\sin\psi \\ -\sin\theta\cos\psi & \sin\theta\sin\psi & \cos\theta\sin\psi & -\cos\theta\cos\psi \end{pmatrix} \times \begin{pmatrix} EE \\ EN \\ NN \\ NE \end{pmatrix} \quad 36$$

The T–T cross-correlation displays the Love wave arrival and the TT equation was used to rotate the horizontal components. The cross-correlations are presented as a function of lag time and inter-stations distance. (Figure 7.7(a) shows cross-correlation between 2 stations, station A9M and A9J. The positive lag represents the ambient noise energy travelling from stations A9J to A9M, and the negative lag represents the ambient noise energy traveling in the opposite direction. The waveforms represent waves travelling in opposite directions between the stations and are represented by different colours. Figure 7.7(b) is a plot of the interferogram between stations A9M and A9J, the plot shows that most of the seismic energy from the calculated cross-correlations is concentrated between lag times -2 and 2 seconds.

After the cross-correlations are returned to the time domain they were stacked to correspond to longer time series. The stacked Love wave seismograms are shown in Figure 7.8 (a) as a function of inter-station spacing. The resulting cross-correlation functions are two-sided time functions with both positive and negative time coordinates. The negative and positive time lags represent waveforms travelling in opposite directions between the stations.

The cross-correlation functions for lag times between -2 and 2 seconds were stored. The

length of the time series needed was calculated from the wave equation, as discussed in the band-pass filtering subsection, the average surface wave velocity of sandy soil is 140 m/s and the longest interstation distance is 85 m. The required time-series needed for the cross-correlation functions is therefore 1.64 seconds. Each horizontal line denotes the stacked correlation functions for station pairs with the same inter-station distance.

These Love wave correlation functions show clear arrival fronts for positive and negative lag times. The arrival fronts correspond to the propagation velocity of the surface waves. The cross-correlation functions appear to be relatively symmetric around the zero-time lag; which would indicate the ambient seismic noise is evenly distributed through time and space.

There is substantial asymmetry in amplitude and spectral content that was observed in the cross-correlation functions. This observed asymmetry is observed indicative of significant difference in both source and distance to source in the directions radially away from the stations.

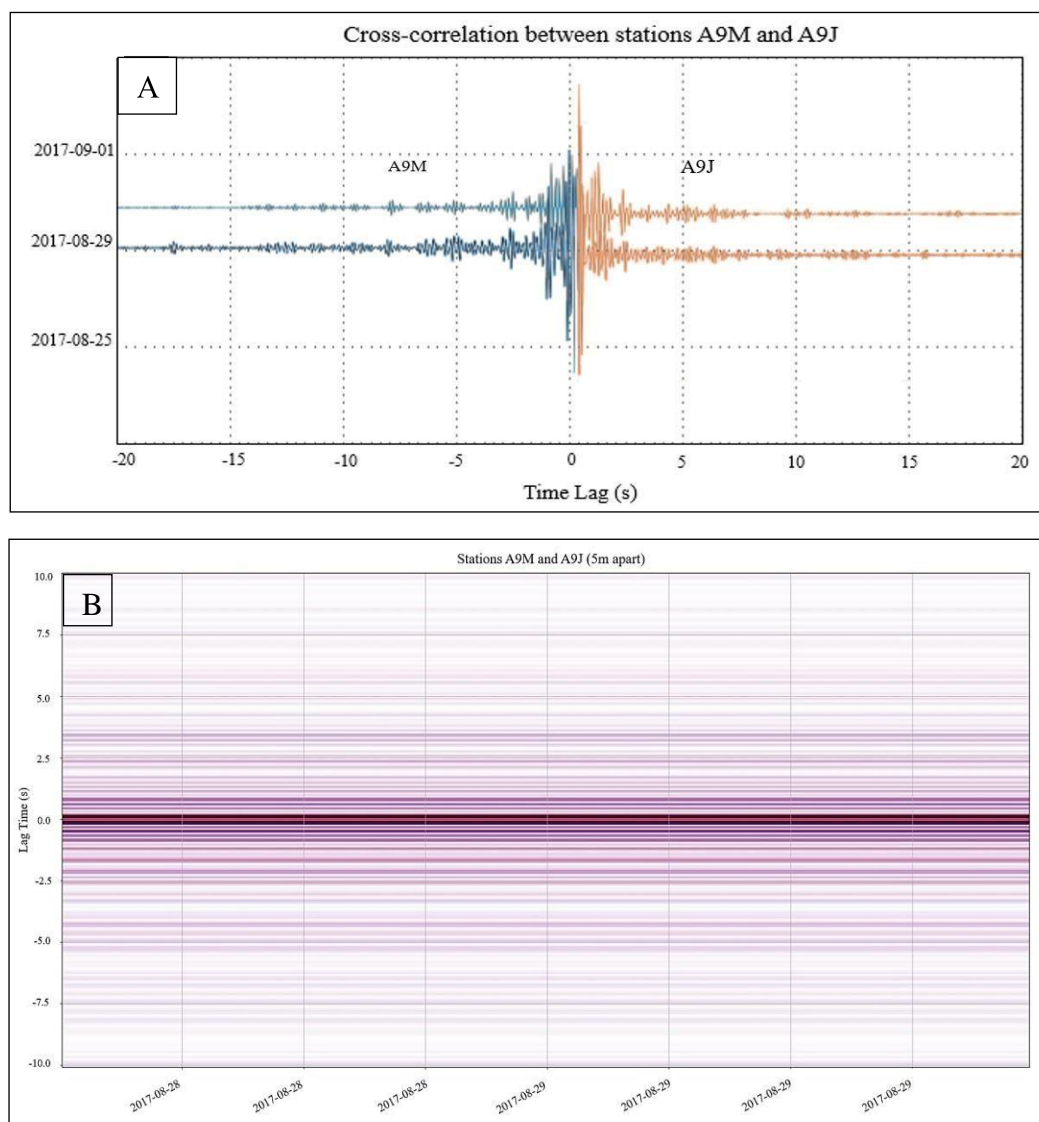


Figure 7.7: (a) Cross-correlation function between stations A9M and A9J plotted between lags 20

and 20 seconds. The different colours represent signals travelling in opposite directions. (b) Interferogram between stations A9M and A9J plotted between lags -10 and 10 seconds.

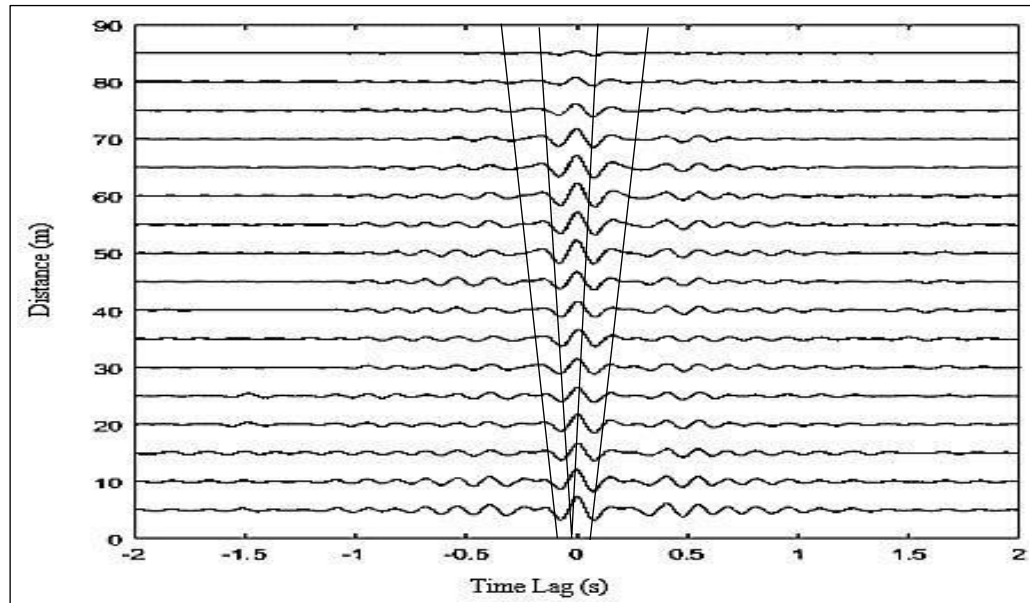


Figure 7.8: (a) The Love wave virtual source signals shown as a function of interstation distance. Clear arrival fronts are visible at positive and negative lag times

Phase 3: Computation of dispersion curves

Once the computation of daily cross correlations and the stacking was completed, phase speeds were measured as functions of the period on the resulting waveform data. The main goal of this step is to analyze how the Love wave speeds change as a function of frequency. The dispersion curves were computed using the multichannel analysis of surface waves method discussed in section 5.5. Each stacked correlation function was filtered in the narrow frequency bands between 7 and 14 Hz and the strongest arrival time was picked for each frequency. Since the distances between the seismic stations was known, the velocity was calculated by dividing the distance between the sensors by the picked arrival time. The curves were then stacked on top of each other for multiple sensor pairs to increase the signal-to-noise ratio. Figure 7.9 shows the calculated dispersion curves for sensor pairs.

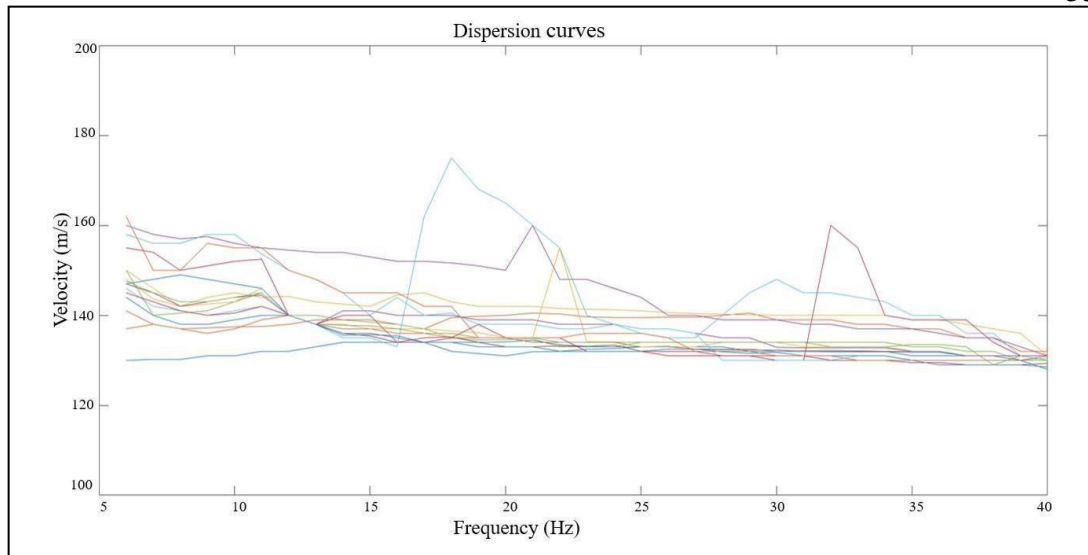


Figure 7.9: Diagram of the computed dispersion curves (Phase velocity is plotted against frequency). The plotted lines represent the average dispersion curve for the different sensorpairs.

A total of 30 dispersion curves were calculated. The average dispersion curve was calculated for each sensor pair. Consecutive sensor pairs were then grouped and the average dispersion curve for the group was calculated for each trial velocity at a given frequency. The general trend of all the curves is similar. Lower frequencies from 5 to 10 Hz have a high velocity range between 138 to 160 m/s. A few curves had relative high velocities between 10 and 15 Hz. The velocity starts to decrease at frequencies above 20 Hz.

Surface wave take about 70% of the total seismic energy and are dispersive in nature (Ivanov, et al., 2001). The different wavelengths have different penetration depths and therefore travel at different velocities. The propagation velocity of the waves depends on the density and the elasticity of the medium. The longer period waves are most sensitive at greater depths and penetrate deeper into the Earth's surface (Lin, et al., 2007). Surface waves propagate in several modes and this makes the dispersion property to be represented by several curves of different modes. The fundamental mode is the mode that is at the lowest velocity range and the ones that appear at higher velocity ranges than the previous are known as higher modes (Ivanov, et al., 2001).

In a horizontally layered medium, the surface wave propagation is a multimodal phenomenon: at each frequency, larger than a well-defined cut-off frequency, different modes of vibration exist. Each mode is characterized by its own propagation velocity, which always increases from the fundamental to the higher modes.

The higher modes of surface waves that exist in heterogeneous media result from the constructive interference phenomena that occurs amongst waves undergoing multiple reflections at the layer interfaces. The different overtones carry different energy which makes it difficult for them to be detected. The energy distribution is dependent on the frequency, a mode can be strongly dominating within a certain frequency band, while negligible in other frequency bands (Foti, et al., 2018).

The different modes are imaged through a wave field transformation process that directly converts the multi-channel record where the different dispersion curves are recognized in the transformed energy distribution. The cross-correlation functions were transformed by into the frequency domain by a Fast Fourier transform. A phase shift was then applied to the data and the relative energy for a particular frequency was calculated. The relative energy values were then added to obtain the slant-stacked amplitude to each set of virtual pairs. The dispersion curves were obtained by plotting the summed amplitude in the frequency domain.

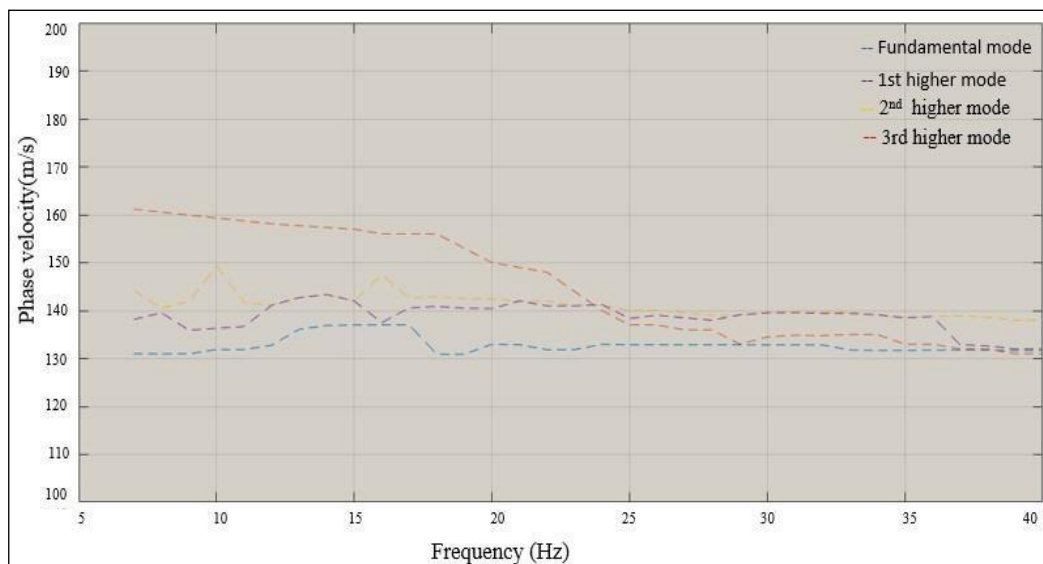


Figure 7.10: Dispersion curve showing the fundamental mode in blue, the 1st higher mode in purple, and the 2nd higher mode is plotted in yellow and the 3rd in red.

The propagation velocities are low, but reasonable for sandy soil. The fundamental mode velocity shows a constant increase in velocity between 12 and 16 Hz. The velocity in this range is 135m/s. The velocity then decreases at 17 Hz to 130m/s and continues to be constant at 131m/s till 40 Hz. The 1st and 2nd higher modes also show lower velocity in higher frequency than in low frequency.

Dispersion Curve Inversion

The next step in phase 3 of the data processing procedure is the inversion of the dispersion curves. The inversion of dispersion curves is a tool that produces P- and S-wave velocities that are needed for wave imaging. The phase velocity that is plotted on the dispersion curves is used for the estimation of shear-wave velocity (V_s) which is important for the evaluation of the stiffness properties of the near surface materials. Shear-wave velocity (V_s) increases with the increase of a materials shear strength or rigidity (Ivanov, et al., 2010).

Information about the shear-wave velocity along with compressional-wave velocity and bulk density can be used to estimate the rock mechanical properties. The shear wave velocity (V_s) and the modulus of each lithological unit are mainly influenced by the degree of fracturing and faulting, the porosity, the clay amount and the precipitation, with the latter two influencing the unit water content (Paoletti, 2012). The inversion process is aimed at searching the best subsurface model whose forward response fits well with the data and generates a 1-D shear wave velocity profile.

The calculated dispersion curves were inverted using the Geopsy software package called Dinver. Dinver operates by generating trial V_s profiles using an improved neighborhood algorithm (Wathelet, 2008). Dinver uses a modified neighborhood algorithm, first developed by Sambridge (1999) and modified by Wathelet (2008) The neighborhood algorithm makes use of Voronoi cells to model the misfit function across the parameter space. Previously generated samples guide the search for improved models through the direct search method to optimize the misfit.

The derivation of the 1-D shear-wave velocity profiles from surface wave dispersion is commonly solved using linearized methods (Wathelet, 2008).

In this study, the stochastic direct search method for finding models of acceptable data fit inside a multidimensional parameter space, using the neighborhood algorithm, is used (Wathelet, et al., 2004). The method is known to correctly map uncertainties of the problem

in the case of non-uniqueness and was first developed by Sambridge (1999) and later modified by Wathelet (2008).

The main parameters that are required for surface wave inversion in Dinver include the S-wave velocity, the P-wave velocity, density and thickness of the layer. The algorithm generates pseudo-random samples in the parameter space and the dispersion curves are computed for the models.

The comparison of the results from the computational dispersion curves with the measured dispersion curve provides a misfit value that indicates how far the generated model is from the true solution. Once the data misfit function is known at all previous samples, the neighborhood algorithm provides a way of interpolating an irregular distribution of points by making use of Voronoi geometry to find the most promising parts of the parameter space (Wathelet, et al., 2004). A high number of dispersion curves are computed to achieve a satisfactory investigation of the parameter space.

The final product from the inversion process is a subsurface variation of shear wave velocity in 1-D (depth). In theory, all surface wave methods are based on a layered earth model, the data analysis steps apply lateral averaging of subsurface conditions along the surface distance occupied by the receiver array.

The velocity model was defined with ten 2 m layers. The V_s value was varied between 50-200m/s for the soil layers and between 100-300 m/s for half-space. To get the best suited models, a total of 7 runs were generated and 168 active models were generated for each run, the misfit values ranged between 0.1 and 0.2. The values of P-wave velocity were linked to the values of V_s through the Poisson ratio, which was varied between 0.2 and 0.5. The density was kept constant and equal to 1800kg/m^3 . It should be noticed that the effect of the value of Poisson's ratio and density of the medium is not significant on the dispersion curves (Wathelet, et al., 2005).

Surface wave amplitudes decay exponentially with depth; therefore, it is conventional to use a third to a half of the wavelength to calculate a V_s profile. Using wavelength values beyond this tends to show higher variance in results. Figure 7.11(a) shows one of the

inverted 1-D shear wave velocity profiles that were computed while (b) shows the computed theoretical curve that corresponds to the alongside shear wave profile.

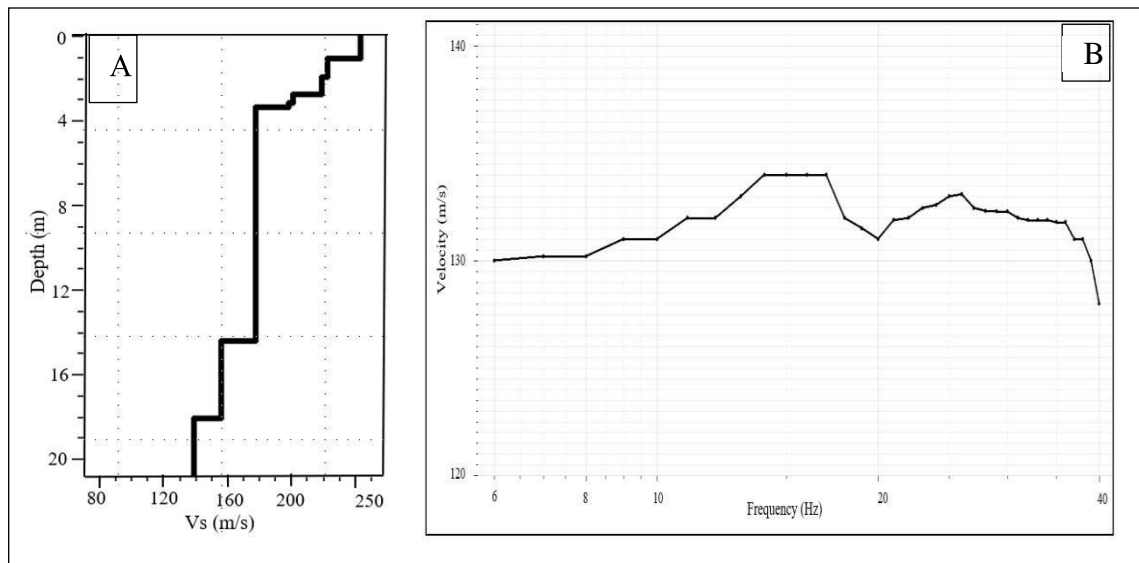


Figure 7.11: (a) 1-D inverted profile (b) Theoretical dispersion curve that corresponds to the velocity model.

At 0 to 4 m velocity varies between 240 to 164 m/s. The velocity is at a constant of 164 m/s from 4 m to 14 m. The velocity drops at 13 m to 158m/s. The velocity further drops below 18 m to 123m/s.

Figure 7.12 (a) Dispersion curve, where slowness plotted against frequency. The shear wave is at 0.0072 s/m. The wave has lower values between 10 and 15 Hz and the speed is at 0.007 s/m. The wave speed increases at higher frequencies between 20 to 40 Hz, the values increase from 0.0071 s/m to 0.008 s/m. Figure (b) is the computed shear wave profile. There is constant velocity of 140 m/s between 2 and 8m. The velocity drops at a constant value of 137 m/s between 8 and 17 m. Above 18 m, the velocity is 120 m/s.

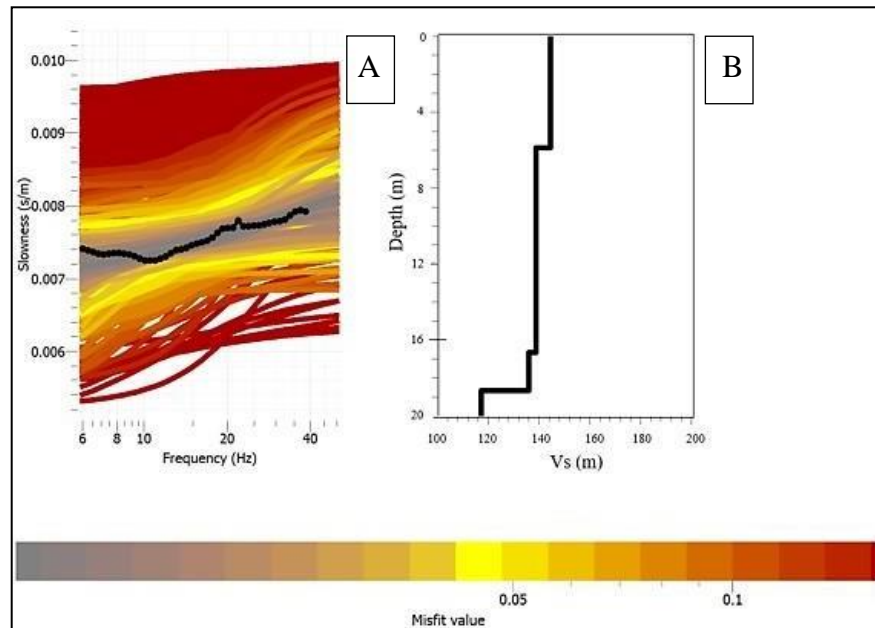


Figure 7.12: An example of Love wave dispersion curve inverted with a neighborhood algorithm to find ground profile. The values in red indicate a high misfit (Olivier, et al., 2018).

The overall shear wave velocities observed from all the profiles are low and do not exceed 250 m at shallow depth between 1 and 3m. In all the profiles, the shear wave velocity drops gradually above 4 m. The average shear wave velocity observed at 20 m is 120m/s. S-wave ground profile here shows higher preferential velocities in the shallow section (0–9 m). After the inversion of all 30 dispersion curves, the corresponding S-wave velocity profiles are allocated to the centroid of the sensor group used to create the 2-D shear wave velocity cross section.

2D shear wave velocity profile

The last step of the data processing procedure is to construct a 2-D shear wave velocity profile to produce a surface and depth shear-wave velocity (V_s) cross section. A total of 16 dispersion curves were inverted. To construct the 2-D shear wave velocity profile, a bilinear interpolation scheme was applied to generate a 2-D velocity grid-data set that in turn was used to construct a 2-D cross-section in which velocities were represented with different colour codes.

Bilinear interpolation is a 2-D interpolation method in which values are first interpolated in one direction and then in the orthogonal direction. Bilinear interpolation is an extension

of linear interpolation for interpolating functions of two variables on a rectilinear 2-D grid (Park, 2005).

The horizontal resolution of the map is predominantly influenced by two field parameters: the receiver spread length and the acquisition interval. The virtual source spread length sets the theoretical lower limit and any Vs structure with its lateral dimension smaller than this will not be properly resolved in the final Vs map. An acquisition interval smaller than the spread length will not improve this limitation as spatial smearing has already been introduced by the receiver spread. However, since all 1-D Vs profiles will always contain some error resulting from the imperfect analysis in dispersion-inversion processing, processing accuracy can also influence the resolution. To avoid processing inaccuracy, computer programs often interpolate stacking velocity in this way, first interpolating in time between picks at velocity analysis points and then spatially between velocity analysis locations. This may create discontinuities in the gradient.

The inverted dispersion curves were imported in MATLAB and a code was generated to compute the 2-D shear wave velocity profile. The interpolation for the 2-D gridded data was arranged in mesh grid format. The inverted profiles were extracted in the form of individual 1-D shear velocity depth profiles for each virtual source location along the survey. The profiles were then assembled according to the surface coordinate at the midpoint of the spread that was used to acquire the corresponding record. The result is shown in Figure 7.13.

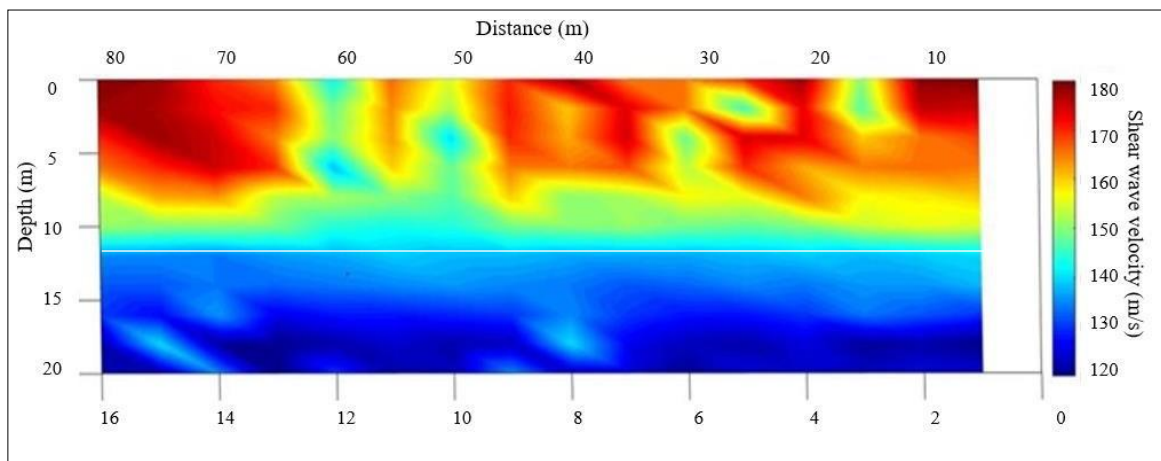


Figure 7.13: 2-D shear wave velocity profile over 85 m of the section. The low-velocity zone below the white line likely corresponds to the phreatic surface. The numbers at the bottom of the

figure indicate geophone positions. The profile shows increasing distance from right to left and was plotted in this way to be consistent with the order in which the stations were planted. The stations were planted from left to right on the tailing dam wall.

A lateral variation of the S-wave velocity can be observed in the upper layers between depths 0 and 5m. This variation possibly indicates the level of soil compaction where the high velocity could be due to the presence of dense sand while the low velocity could be due to the presence of loose sand or soft clay.

The near-surface low-velocity zone at 30 m, 50 m and 60 m along the profile probably indicates very loose sand or very soft clay in this region. The S-wave profile also shows a clear zone of low velocity around 10 m below the surface. Near the center of the array (between 50 and 65 m), there are two plumes below 3 m with lower velocities ranging from 133 to 140m/s possibly indicating that the phreatic surface is closer to the surface.

There is a high velocity zone that is observed between 70 and 80 m on the profile, this zone is interpreted as possibly indicating dense sand. The survey line was spatially well correlated with the area where the photograph that showed evidence of seepage was taken.

7.2 Study area 2: Cradock Dataset

Pre-processing

Rayleigh waves were chosen for the processing of the data acquired from Cradock. The processing of Rayleigh waves requires that vertical component of the data from all the stations be processed.

Pre-processing is done to accentuate the broadband ambient noise and consists of the mean and trend removal, temporal normalization and selected band pass filtering. These procedures remove the high amplitude ambient noise events that tend to obscure the desired ambient noise signal, prior to cross-correlation (Bensen, et al., 2007). Figure 7.14 depicts the raw waveforms (prior to pre-processing) recorded from the vertical component of two stations (A1R and B5Y).

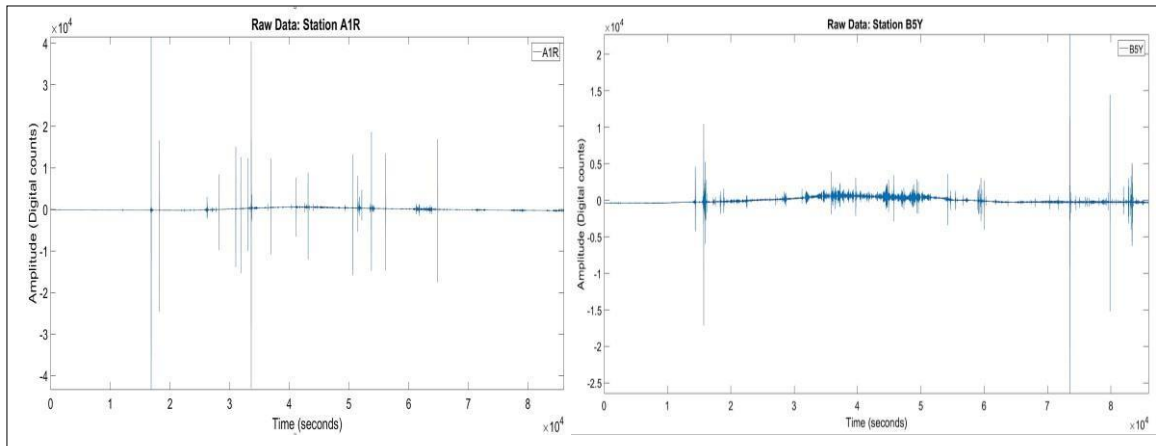


Figure 7.14: Data from station AIR and B5Y, after the mean and linear trends are removed. The length of the time-series is 86400 seconds and the amplitude of the data is presented on they-axis in digital counts which is an unscaled measure of displacement.

Band-pass filtering

The main purpose of a seismic filter is to remove unwanted effects from the seismic image. MATLAB's Butterworth filter was used to remove the unwanted noise effects from the seismic data, the daily traces were filtered between a low pass of 0.125 seconds (8 Hz) and a high pass of 100 seconds (0.01 Hz). This was selected to eliminate the high amplitude and long period signals produced by earthquakes. Earthquake signals generate high long period waves that tend to obscure ambient noise (Bensen, et al., 2007).

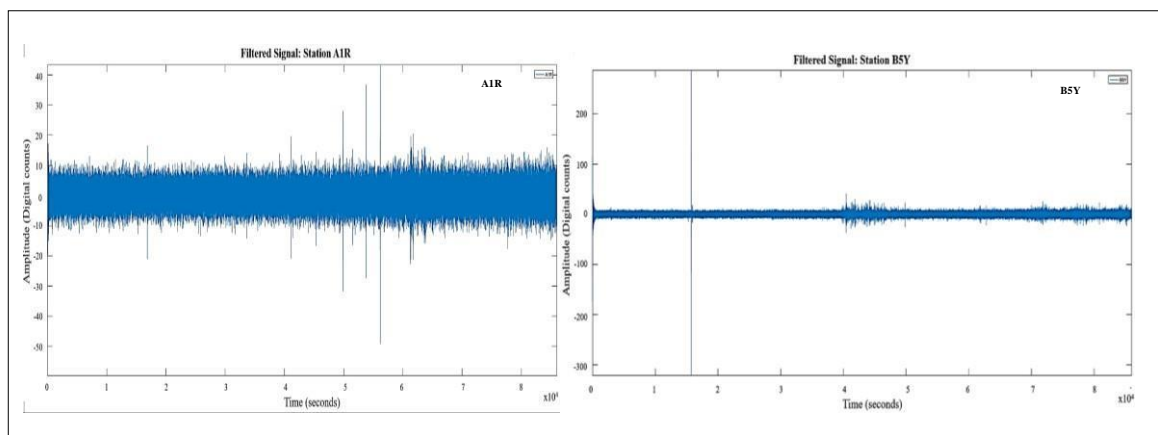


Figure 7.15: Data from stations AIR and B5Y, after band-pass filtering was applied.

Spectrogram

A spectrogram was performed in MATLAB to plot and analyze the range of frequency content over the duration of the recording period in the data. The ambient seismic energy for both the stations appears to be relatively distributed throughout the frequency range with a strong frequency peak occurring between approximately 0.1 and 1 Hz. The spectrogram also shows that most of the seismic energy is at a maximum below 0.5 Hz, which suggests that the seismic energy that was recorded is mainly due to storms and interactions of wind-driven ocean waves travelling in opposite directions in the shallow or deep water. This produced noise is also termed secondary microseisms (Lepore & Grad, 2018).

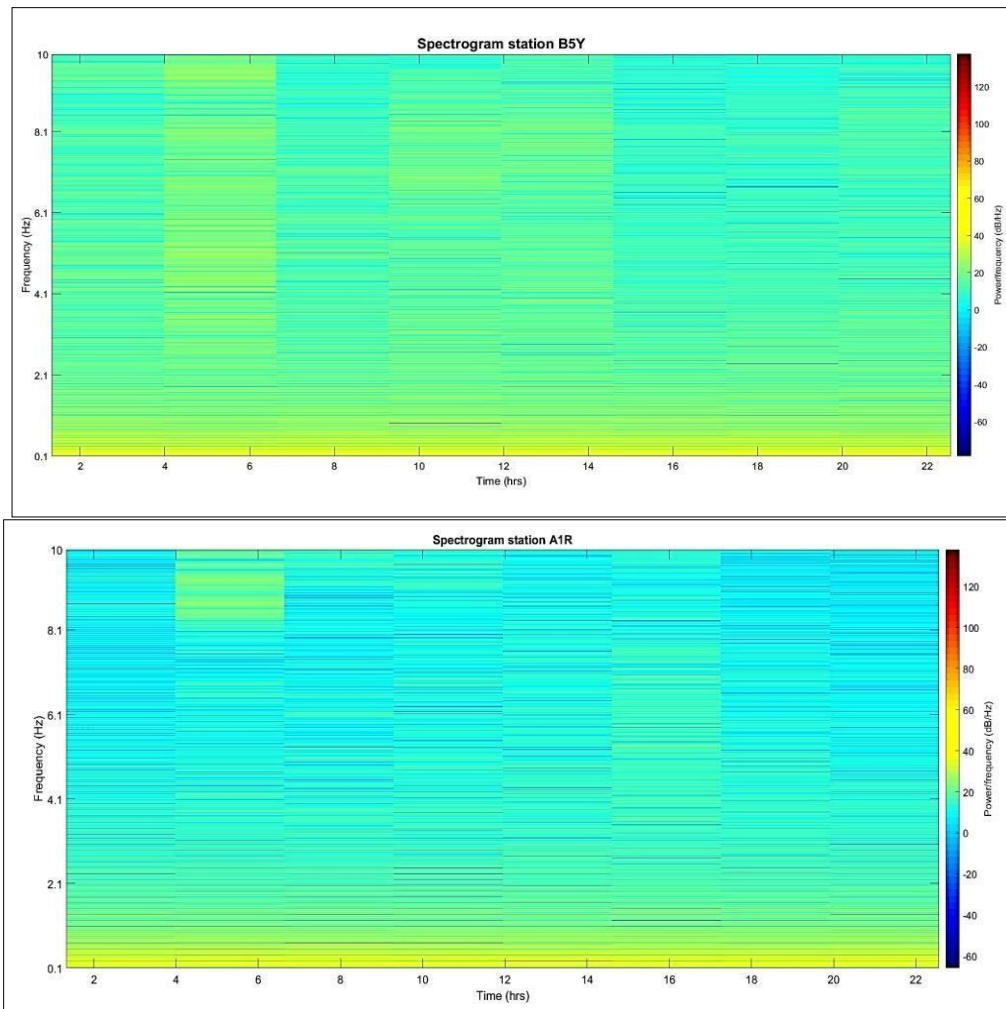


Figure 7.16: Spectrogram computed from stations A1R and B5Y plotted against time in hours.

The spectrogram was also computed to check if the filter that was designed was effective in cutting off the unwanted frequencies. Figure 7.17 shows the result of the spectrogram after the traces were filtered. The filter was effective in removing frequencies above 2 Hz.

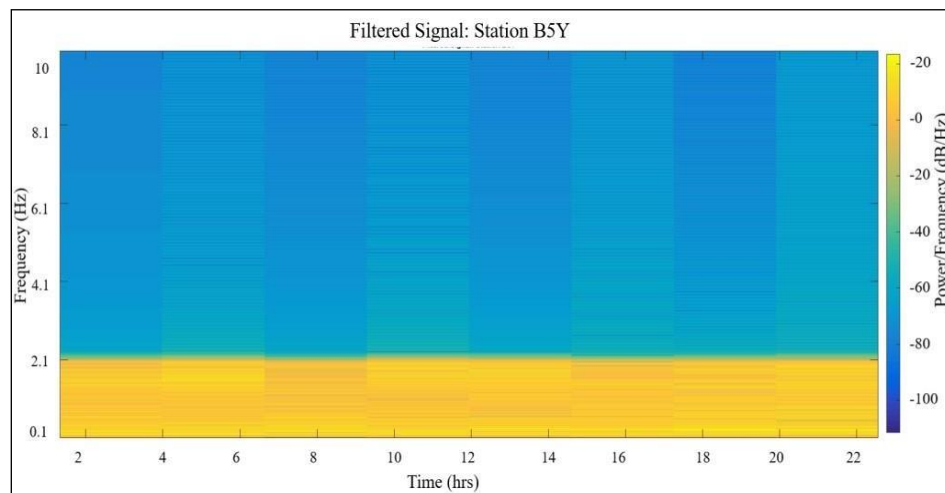
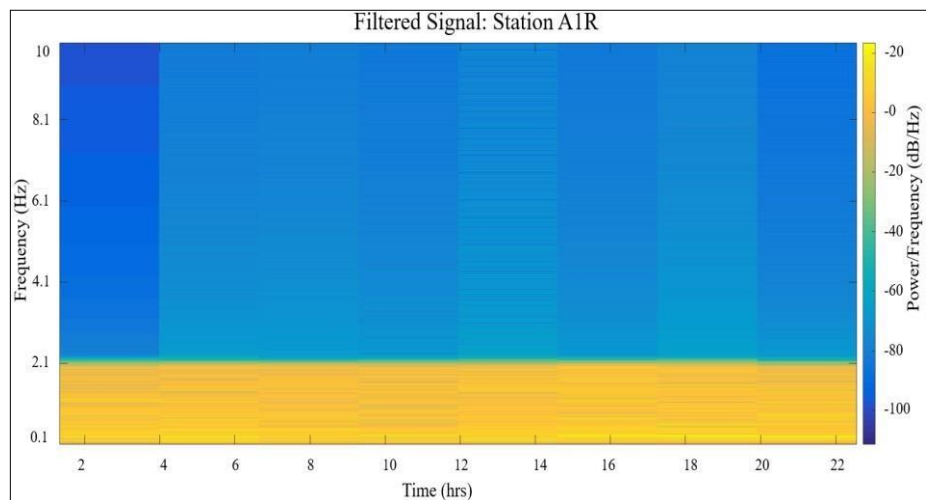


Figure 7.17: Spectrogram after band-pass filter was applied

The next phase of data processing was done in MSNoise. In preparation for computation of cross-correlations, the daily traces were decimated to reduce the sampling rate. This step was done to reduce the memory required to complete the processing. The daily traces were divided into cross-correlation windows and the duration for each window was set at 1800 seconds. Then, for each cross-correlation duration window in the signal, the traces were clipped to winsorizing times. A winsorizing of three times the root mean square (RMS) of the daily traces was applied. Winsorization is the process of minimizing the effect of outliers/parasitic signals from seismic events in the daily traces by modifying the outlier's

data point. The clipped traces were whitened between the frequency bounds of 0.1 and 1 Hz (1 to 10 sec) to broaden the band of the ambient noise signal in cross-correlation and prevent degradation that is caused by monochromatic sources (Bensen, et al., 2007).

Stacked cross correlations

The next step of the processing procedure is the calculation of cross-correlation functions between all station pairs. The cross-correlation calculations were performed on the daily pre-processed data in the frequency domain, to increase the speed of the processing. The cross-correlation functions for all station pairs were computed. 22 stations were prepared for cross-correlations, the maximum number of station pairs that were correlated were derived from Equation 35 (Lecocq, et al., 2014). Therefore, the total number of station pairs that were cross-correlated were 231.

The daily cross-correlation functions between all station pairs were then converted back into the time domain, by means of the inverse Fast Fourier Transform and stacked together. Each daily cross correlation function is divided into a correlation window of 30 minutes. The process of stacking the cross-correlation functions over an increasing duration of days is important because it increases the signal to noise ratio and improves the emergence of surface waves arrivals (Bensen, et al., 2007). Figure 7.18 (a) shows the stacked cross-correlation functions plotted as a function of interstation distance and lag time between - 60 and 60 seconds for the entire duration the data was recorded for. The lag times stored (-60 to 60 sec) is sufficient to obtain the Rayleigh surface waves arrivals, with the maximum interstation distance and the velocity of the surface waves considered.

Figure 7.18(b) shows the stacked cross-correlation functions plotted against the interstation distance for a period of 15 days. The arrival fronts and their corresponding frequencies are not as defined as in Figure 7.18 (a). The stacking of cross-correlation functions over longer periods increases the signal to noise ratio and improves the quality and resolution of the signal

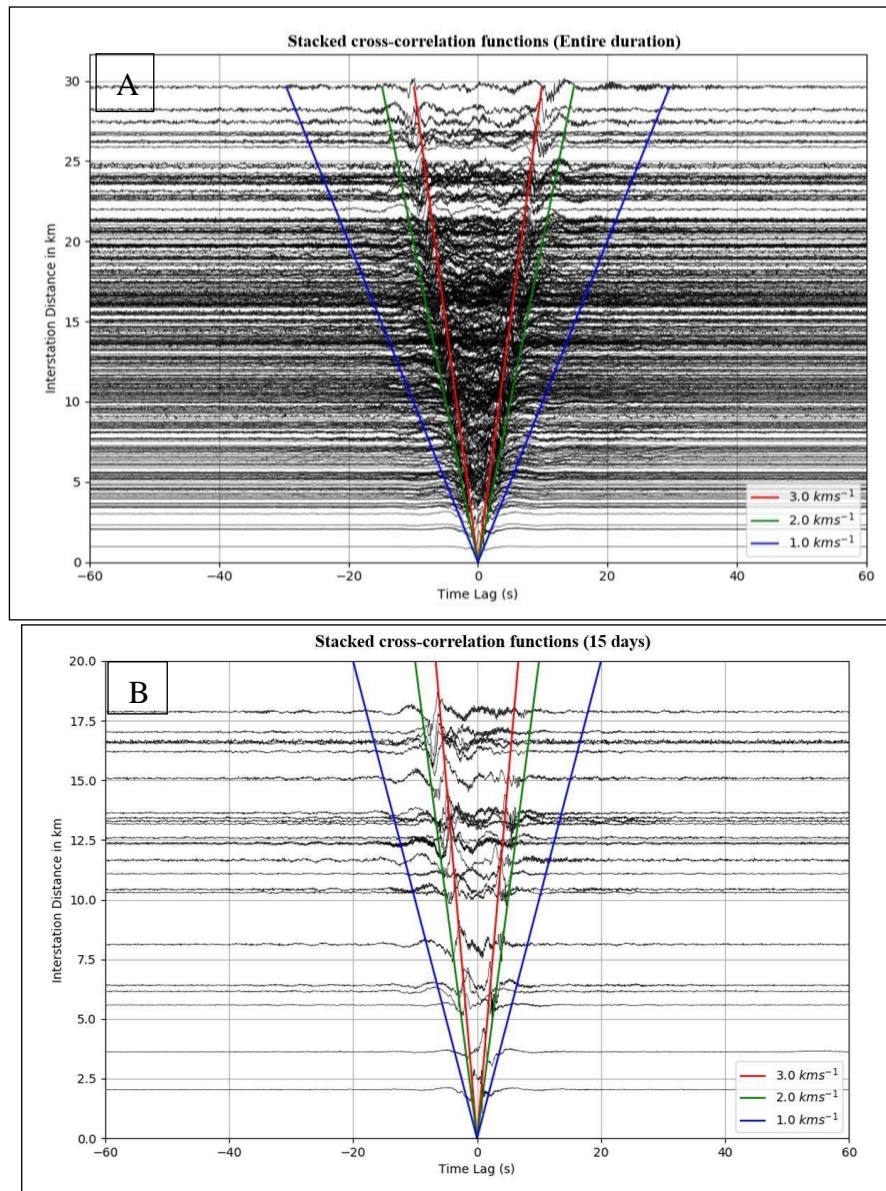


Figure 7.18: (a) Stacked cross-correlation functions (b) Stacked cross-correlation functions for 15 days

The coherency of the signal and the arrival times of the Rayleigh waves is evident in Figure 7.18 (a-b). The emergence of Rayleigh wave arrival can be observed on both the positive and negative lags between the station pairs. The waveforms represent waves travelling in opposite directions between the stations. Wave packets appear between the red and green solid lines on both the positive and negative time lags with a velocity between 2 and 3 km/s, which is presumed to be Rayleigh waves. The Rayleigh wave arrivals, travelling between stations, obtained from the vertical component waveforms display comparable trends to that observed.

There is considerable asymmetry in amplitude that can be observed, the causal and acausal signals are not identical, they differ in amplitude and this amplitude factor depends directly on the energy flux of the waves traveling from one station to the other. This suggests that there are differences in both the source process and the distance to the source in directions that are radial from the stations.

When ambient noise fields are homogeneously distributed in azimuth, the cross-correlation is expected to be approximately symmetric in amplitude and arrival time around the zero-time lag. A perfectly random distribution of the sources of ambient noise would result in symmetric cross correlations with energy arriving at both positive and negative correlation lag times. However, this is not always the case as considerable asymmetry of the cross-correlations is often observed, which results from stronger or closer ambient noise sources directed radially away from one station than the other (Stehly, et al., 2006).

An important consequence is that the asymmetry of the cross correlation computed between several pairs of stations of a network can be used to measure the main direction of the energy flux across the array.

Dispersion Curves

The next step of the data processing is the computation of the dispersion curves. The waveforms that were the product of the computed and stacked cross-correlation functions are an estimate of the Green's Function. The dispersion curves were computed in MSNoise using the TOMO plugin. The plugin calculates the dispersion curves using the frequency-time analysis method (FTAN). The calculations of phase velocity dispersion measurements were made by inputting the average of causal and acausal cross correlation functions from the station pairs into a frequency-time analysis (FTAN) scheme.

The study of surface wave dispersion allows for the extraction of seismic velocity variation with depth. The paths of the station pairs run through different geological regions, thus the dispersion curves were expected to vary for the different station pairing. Generally, periods below 20 seconds are sensitive to the crust and periods above 20 seconds are sensitive to the upper mantle. Short periods of the Rayleigh waves generally sample the shallower subsurface and longer periods the deeper subsurface.

A total of 276 dispersion curves were computed. The general trends of the dispersion curves are similar with lower velocities observed in higher periods than in lower periods. Phase velocity decrease with an increase of period. A maximum peak at period 1.5 seconds is observed with a velocity of 4-4.5 km/s. A second peak can be seen at 2.5 seconds with a velocity of 3 to 3.2 km/s. The last peak is seen at period 4 seconds with a velocity of 2.8 km/s.

The dispersions curves that were computed from the stations that were installed on the western part of the area (stations enclosed in orange box on the map in Figure 6.2) showed lower velocities than those installed in the north east section of the area. Two major peaks can be observed at periods 1.5 and 2.5 seconds at velocities 2.8 and 3 km/s.

Higher periods sample deeper into the crust and this suggests that the velocity at deeper depths is lower than that at shallow depths.

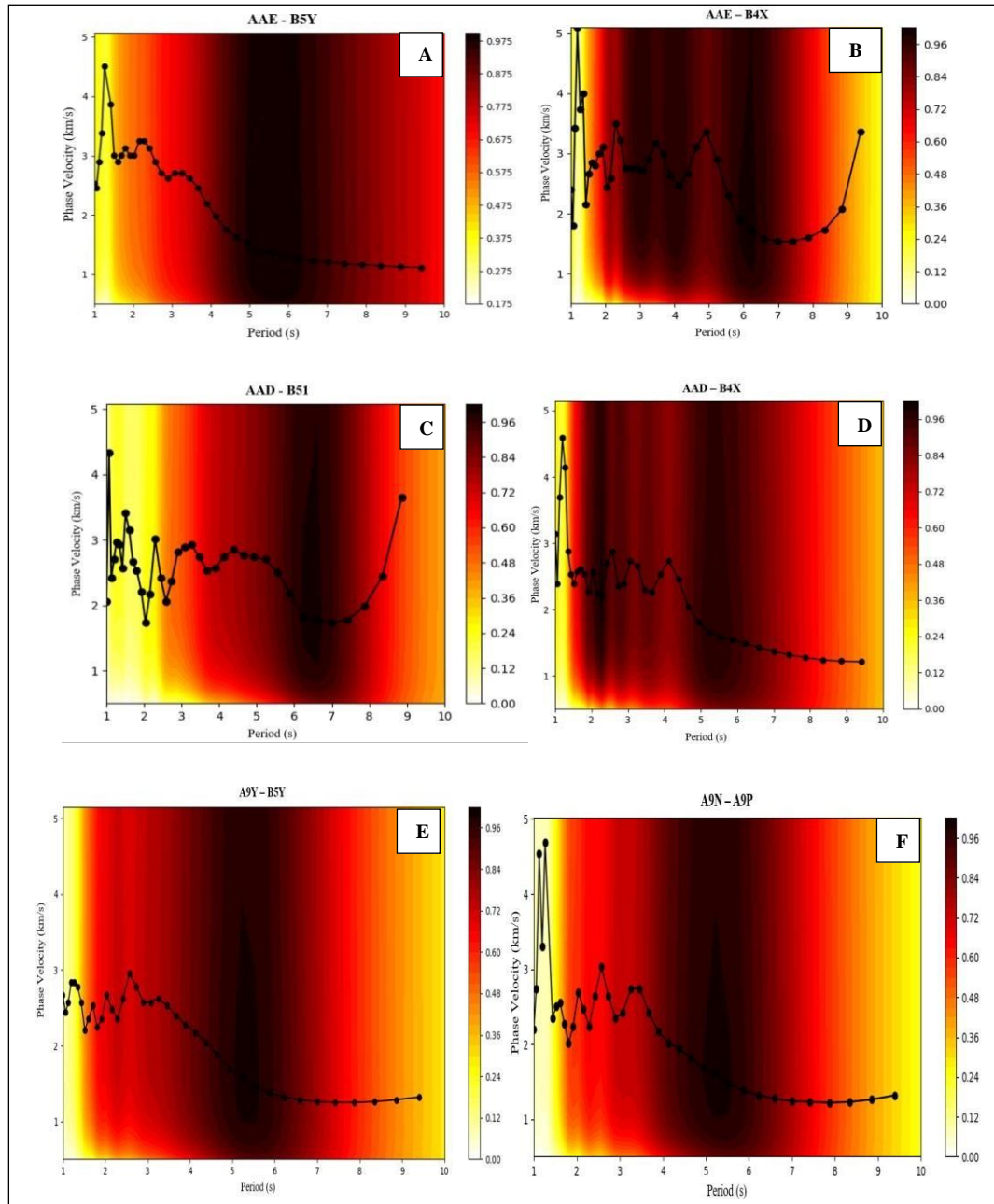


Figure 7.19: Dispersion curves computed using FTAN for the different stations. The black lines indicate the fundamental modes.

Dispersion Curve Inversion

To estimate the P- and S-wave velocity profiles, the dispersion curves were inverted using the Improved Neighborhood Algorithm implemented in the Dinver toolbox of the Geopsy package (Wathelet, 2008).

The largest wavelength in the dispersion curves for Rayleigh waves is about 7.5 km and corresponds to a wave velocity of about 5 km/s at a period of 1.5 seconds (Figure 7.19). A depth of approximately 1/3 of this wavelength can be resolved, that is about 2.5 km, using the phase velocity dispersion curve information alone.

The V_s of each layer was ranged from 50 to 3500 m/s and increased to 150-3500 m/s for the half space. The compressional velocity (V_p) was set to vary from 200 to 5000 m/s and the Poisson's Ratio remained uniform in the range 0.2-0.5. The density was set at a constant of 2000 kg/m³. The density has low influence on surface wave dispersion. The thickness limits are defined through the wavelengths (λ) derived from frequencies and phase-velocities of the actual dispersion curves.

The Neighborhood Algorithm Wathelet (2008) was used to invert the experimental curves. The tuning parameters were n_s , n_r , and n_i , corresponding to the number of samples randomly distributed or not over the parameter space, the number of best cells to consider and the number of times the process is repeated until an acceptable sampling of the solution is obtained. A misfit function is first computed for the initial set of n_i models. Within the n_r cells with the lowest misfit, a total of n_s new models is added (n_s/n_r samples generated per cell). The last two steps are repeated N times resulting in a total of $n_i + N*n_s$ models. The tuning parameters we used are $n_i = 100$, $n_s = 50$, $n_r = 50$, and $N = 4000$ resulting in a total of 200100 models that are enough to reach the plateau of the misfit trend.

Each set of parameterizations was repeated for 8 runs to test the robustness of the results. Inside each layer within the four parameterization groups the free parameters are V_s , the ratio V_s/V_p and the thickness in the limit range previously indicated. However, the most important parameter in surface wave inversion is V_s . In each layer, the V_p was linked to the V_s interface and this was done to have a large range of possible solutions without increasing excessively the number of degrees of freedom (Di Giulio, et al., 2010).

Figure 7.20(a) shows one of the inverted 1-D shear wave velocity profiles that were computed. (b) Is the theoretical curve that was computed that corresponds to the alongside shear wave profile.

In Figure 7.20 (a) at 0 to 180 m velocity varies between 200 and 700 m/s. The velocity is at a constant of 1100 m/s from 220 m to 580 m. The velocity increases at 600 m to 2500 m/s.

Figure 7.20 (b). Higher velocity is observed at a shallow depth of 100 m. There is sharp increase in velocity at 180 m to 1800 m/s. Above 380 m the velocity is 2500 m/s.

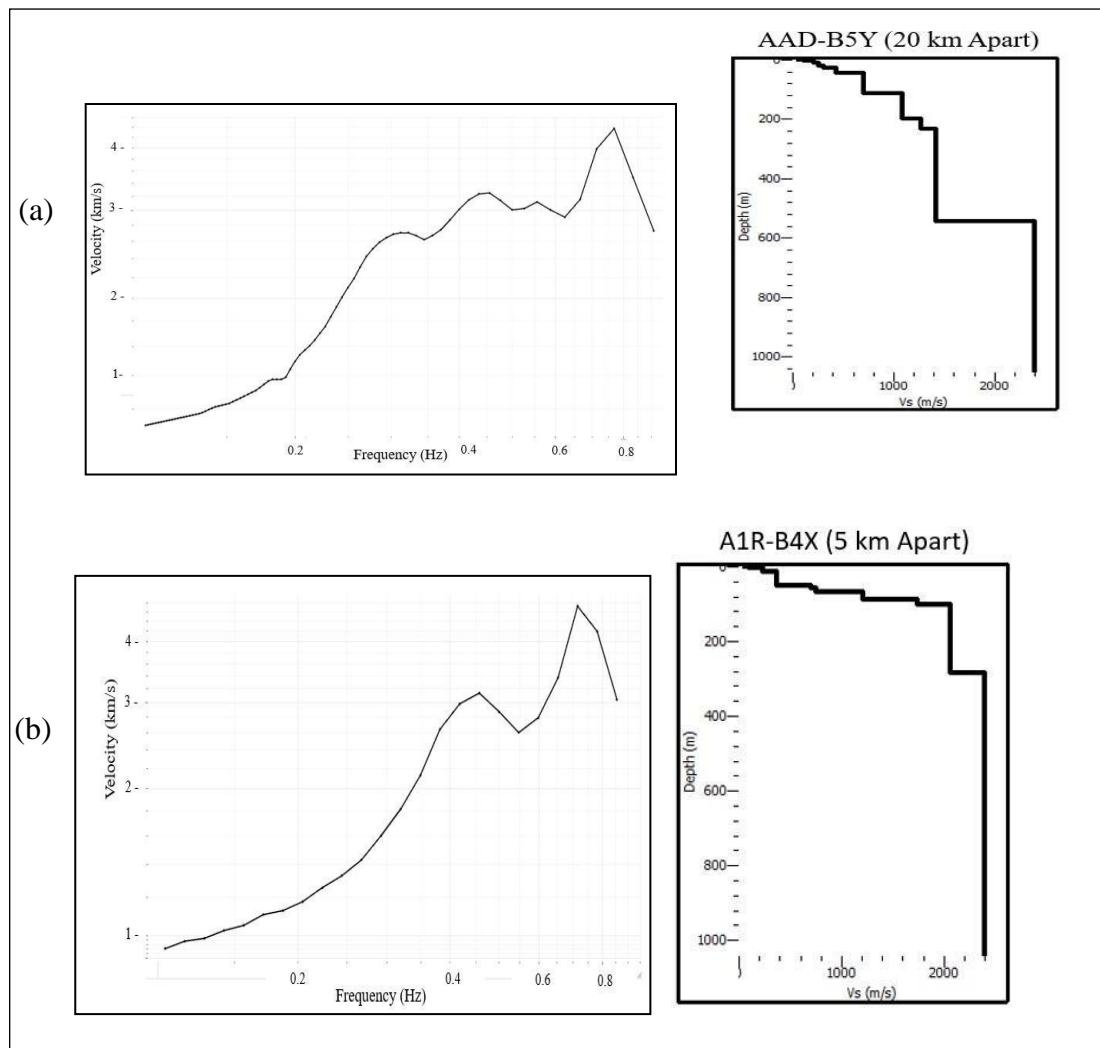


Figure 7.20: (a) shows one of the inverted 1-D shear wave velocity profiles that were computed between stations AAD-B5Y. (b) Shows the inverted 1-D shear wave velocity profiles that were computed between stations A1R—B4X.

2-D shear wave velocity profile

The next step in the processing of the ambient noise data was to compute a 2-D shear wave velocity profile that would show how shear-wave velocity changes with depth to provide an image of the subsurface. A shear wave cross section is important to interpret subsurface and is used to describe the characteristics of earth layers.

The dispersion curves that were computed from all the stations indicate that there is a change in velocity moving eastwards of the study area. The western region has lower velocities than the eastern region of the study area and to provide a subsurface image to further see how the velocities change with depth, a profile that runs across the study area was chosen for shear wave velocity cross-section. The profile was from station B51 to station AAD and is 25 km long (Figure 7.21). The resulting sets of 1-D plots of shear-wave velocity profiles were interpolated to produce a 2-D shear wave velocity cross-section at stations. The interpolation and plotting were done in MATLAB. A peak at 1.5 second is observed in most of the curves and this corresponds to upper crust at a depth of 1.5 km, the cross-section was chosen to be at a depth of 1 km.

The data was then assigned into a 2-D ($x-z$) grid. A bilinear interpolation and data smoothing technique was applied to the data. The shear wave velocities are assigned to the grid and this created ambiguity in the horizontal coordinate (x) because each V_s profile was obtained from a virtual source that spanned a distance too large to be considered as a single point. To resolve this, the center of the receiver spread was chosen to be the most appropriate point to be used because the analyzed V_s profile represented an average property within the spread length. Bilinear interpolation was used to interpolate each midpoint velocity and generate the 2-D V_s profile from the midpoint of the first spread line to the midpoint of last spread line.

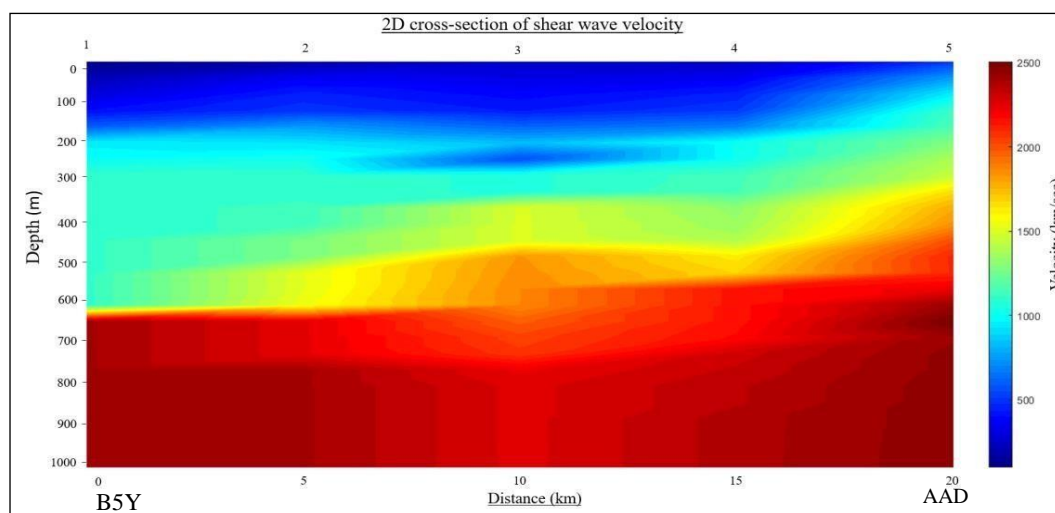


Figure 7.21: 2-D shear wave velocity profile across the study area. There is a low-velocity zone between 5 and 15 km and at 220 m. High velocity zone is observed at a shallower depth between 15 and 20 km. The numbers at the top of the figure indicate the geophone positions.

In general, the subsurface layers of the research area are composed of soft soil shown in dark blue, medium soil shown in light blue, soft rock shown in yellow and hard rock shown in red.

The cross-section indicates that shallow depth shear wave velocities vary up to 300 m/s, possibly associated with soil material consisting of loose and larger stones. The 2-D shear wave velocity profile down to about 100 m shows that the V_s distribution is irregular, and this could be due to layering and heterogeneity in the subsurface.

The velocity starts to increase from 200 m and the velocity varies from 1000 to 1500 m/s, the increase possibly indicates soft weathered rock or soft sediments, the distribution of this rock layer is not regular. There's low velocity zone between depths 250 to 300 m between 6 and 14 km, this could be due to artefacts from the processing of the data. There are higher velocities that are observed at shallower depths at a distance between 17 and 20 km and the velocity vary between 900 and 1300 m/s.

At a distance between 5 and 15 km and at a depth of 370 m a zone of high velocity can be seen and the velocity ranges between 1500 and 2000 m/s, this could indicate that there is a rock layer that is at this depth. For a depth from 600 m to 1000 m there is a hard rock layer, the shear wave velocity varies from 1500 m/s to 2500 m/s.

2-D Tomography maps

The next step in processing the ambient seismic noise is tomographic inversion.

Surface wave tomography provides an accurate velocity model for the middle and shallow crust. On the one hand, it also helps to use body wave tomography to better recover the deeper discontinuities and structures in the deep crust and superficial mantle. On the other hand, it serves as a reference model for the study of shallow structures (for example, more than 1 km deep). The sensitivity of deep surface waves depends on their frequency, the lower frequencies being sensitive at greater depths (Mordret, et al., 2013).

The 2-D tomographic method that was employed in this study is based on assumption that the seismic wave propagates along the great-circle path connecting the source and the

station. The theory that forms the basis of the tomographic method, together with expressions of the forward and inverse problem, is given in Appendix B.

Group velocity maps

The resolution of the group velocity maps depends mostly on the density of the ray paths and on the spatial and temporal distribution of seismicity and seismic stations (Stehly, et al., 2009). The path density was plotted and analyzed to assess the spatial averaging of the tomography inversion. Figure 7.22 shows the path density map at periods, 1.5, 2.5, 10 and 20 seconds. The path coverage reaches 96 paths per cell particularly for periods shorter than 5 s. For higher periods, the path coverage reaches 108 paths per cell. At the edges of the study region the resolution decreases rapidly due to fewer stations and less ray coverage.

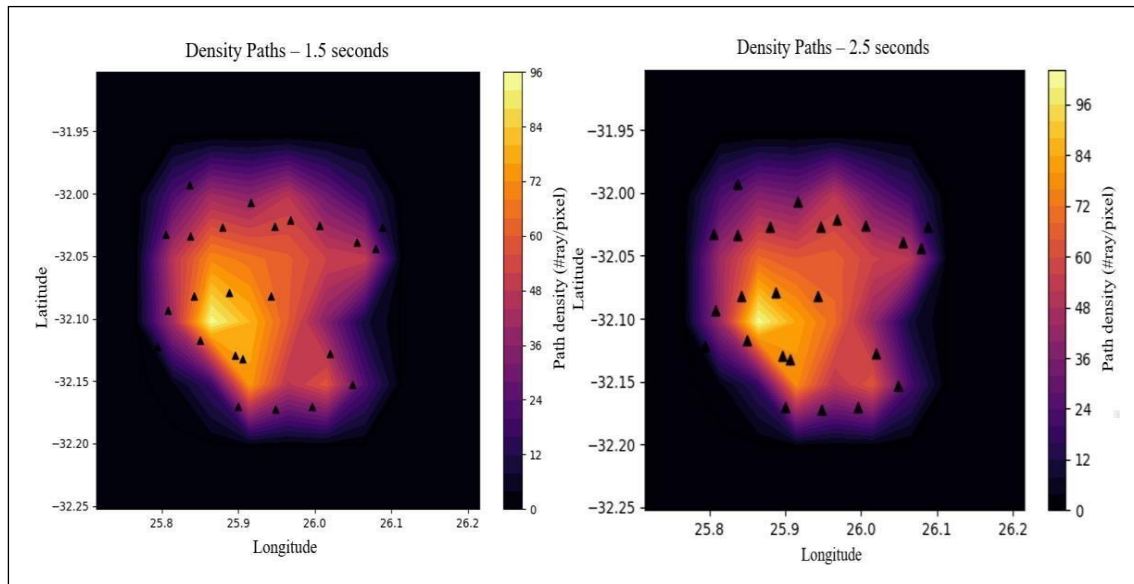


Figure 7.22: Density paths maps plotted for periods 1.5 and 2.5 seconds

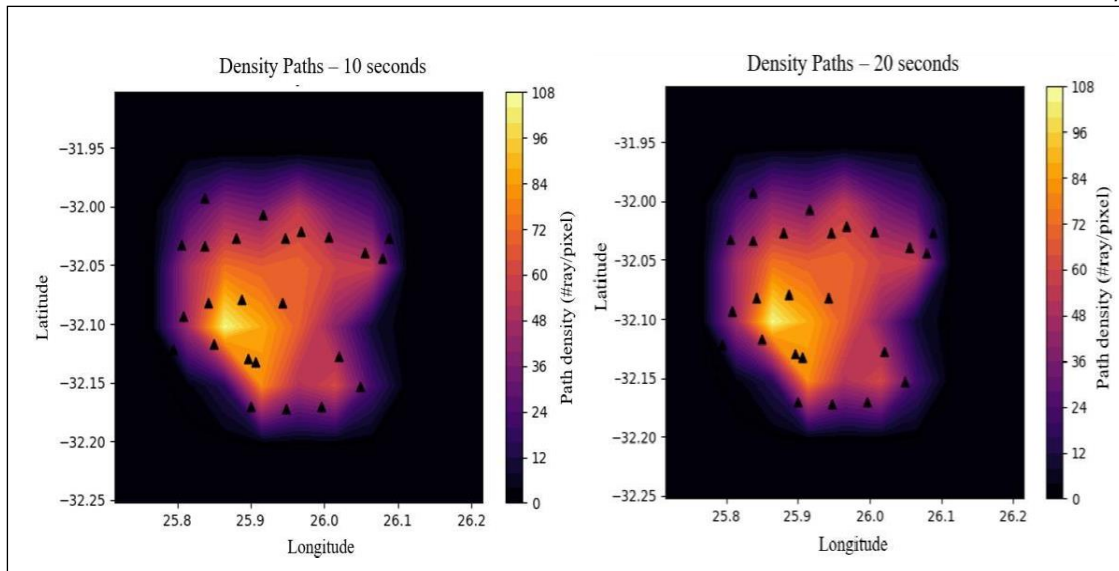


Figure 7.23: Density paths maps plotted for periods 10 and 20 seconds.

Ray coverage maps

The ray-path coverage (Figure 7.24) shows a quite dense coverage at periods of 1.5 and 2.5 while at 10 and 20 s the coverage is sparser. The ray coverage maps also indicate high and low velocity anomalies, high velocity indicated by yellow and low velocity by blue.

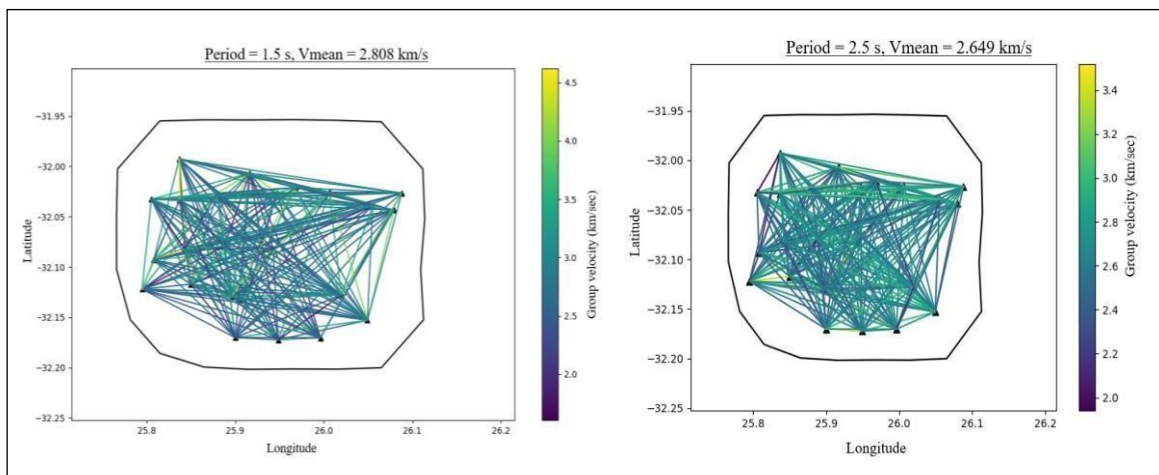


Figure 6.24: Ray coverage maps for periods 1.5 and 2.5 seconds.

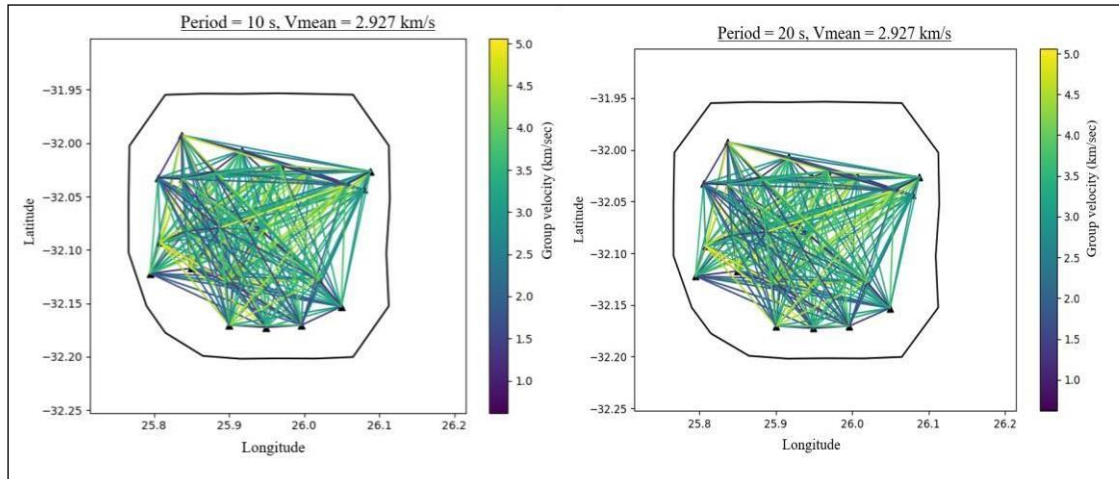


Figure 7.25: Ray coverage maps plotted for periods 10 and 20 seconds.

Rayleigh wave group velocity maps

Group velocities were obtained for six different periods: 1.5, 2.5, 10 and 20 seconds. High frequency or short period waves only penetrate the shallow subsurface, while long period waves penetrate deeper into the crust (Stankiewicz, et al., 2010).

1.5 and 2.5 second group velocities

The 1.5 second map provides an average velocity model of the uppermost 1-1.5 km. Short period maps are sensitive to mainly shear-wave velocities in the upper crust. The approximate depth sensitivity of Rayleigh wave group velocity is around one third of a wavelength, thus the 1.5 second period (Figure 7.26(a)) gives an average group velocity model of the shallow crustal structure (the uppermost 1-1.5 km) and the 2.5 second map (Figure 7.26(b)) is most sensitive to depths of up to 2-3 km. There is a zone of low velocity that is observed on the south-western section of the area Figure 7.26. This zone most likely reflects thick sedimentary deposits. The velocities change and start to increase when moving across from the west to the east (i.e. towards station AAD) and this could be due to a change in sedimentary thickness.

The main features of the 2.5 s group velocity map are broadly similar to those of the 1.5s map. Once again, low surface wave velocities are imaged west of the study area. Velocities in Figure 13 appear to increase abruptly eastwards. A band of higher group velocities ($>2.6 \text{ km s}^{-1}$) is observed on the north and eastern part of the area.

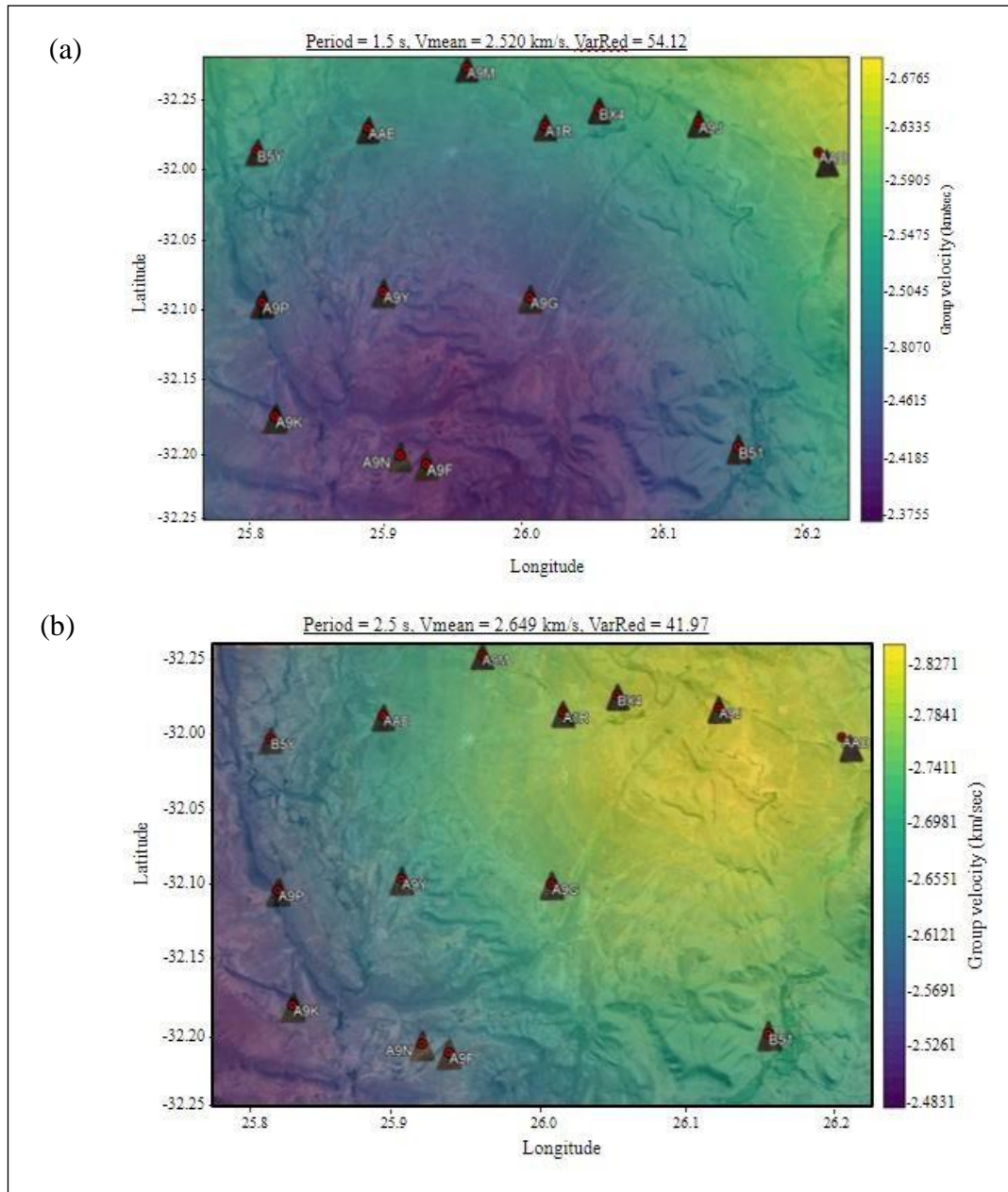


Figure 7.26: Group velocity maps plotted for periods 1.5 (a) and 2.5 (b) seconds. The yellow colours indicate high velocity zones and the purple indicates low velocity zones.

5 and 10 second group velocities

The 5 second group velocity map is similar to the 2.5 second group velocity map. Low velocity is observed on the western section of the region (towards Cradock) and there is a general increase in velocity eastwards. High velocity structures are prominent from the center to the North eastern side of the map and these structures are visible on the 2.5s and

5 s maps. They are indicated by the yellow colours and range in velocity between 2.4 km/s and 2.6 km/sec. The high velocity region is most sensitive to depths that range between 2.5 to 4 km.

The 10 second is slightly different from the 5 second map. lower velocity region is still seen on the western section of the map. Lower velocities can also be seen on the center of the map moving eastwards. The high velocity region that was prominent in the 5 second map shows a drop-in velocity and starts fade and become less prominent with increase in depth. The 10 second map shows structures that were imaged at a depth range 9 to 10 km.

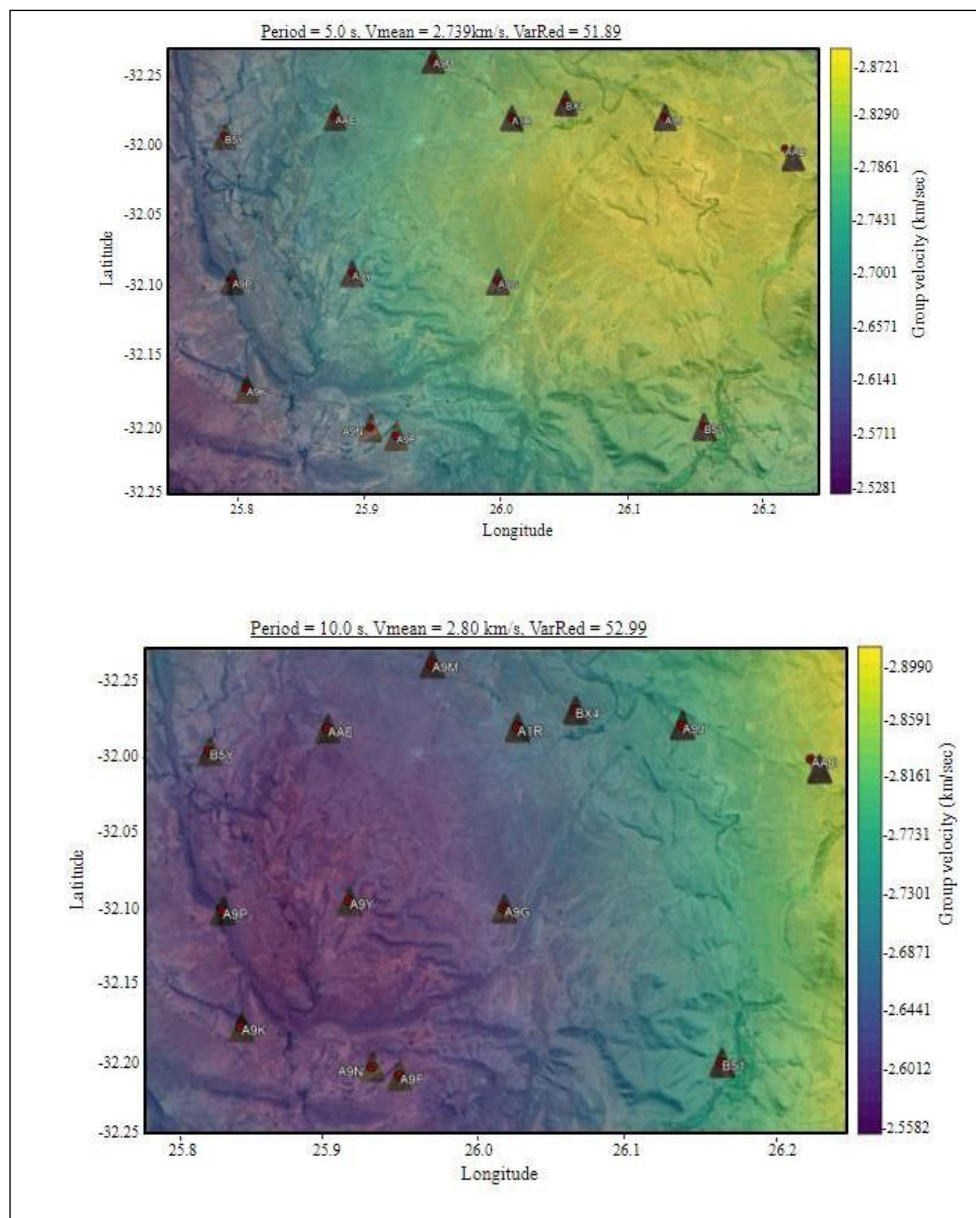


Figure 7.27: Group velocity maps plotted for periods 5 and 10 seconds. The yellow colours indicate high velocity zones and the purple indicates low velocity zones.

20 and 30 second group velocities

Compared to the shorter-period maps, the 20 and 30 seconds image structures at depths that range from 10 to 20 km. The high velocity region that was prominent in the 2.5 second map shows lower velocities. There's a zone of low velocity indicated by blue and purple colours from the western section of the map moving towards the center.

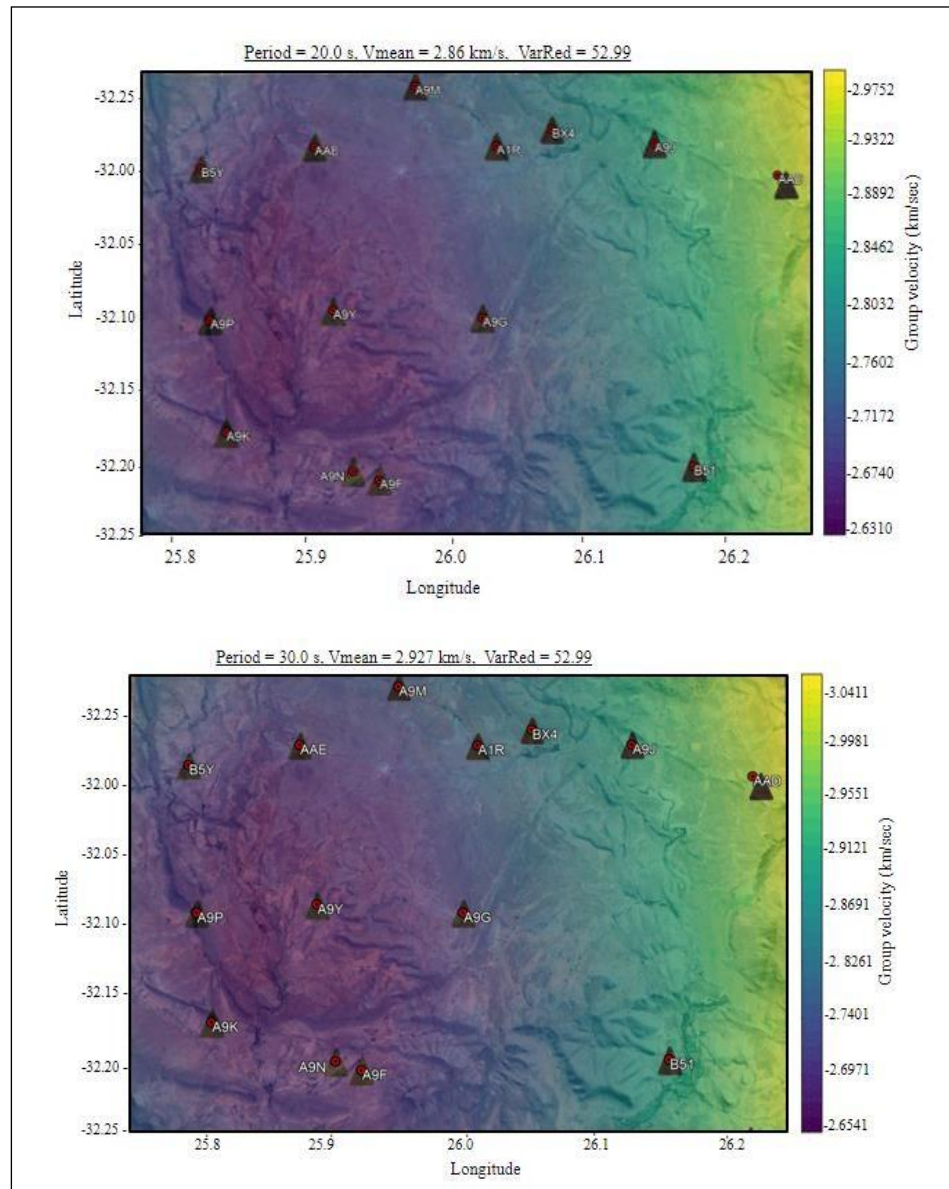


Figure 7.28: Group velocity maps plotted for periods 20 and 30 seconds. The yellow colours indicate high velocity zones and the purple indicates low velocity zones.

8. Discussion

8.1 Harmony Dataset

The FSSW8 tailings dam where the data was collected was constructed using the upstream method. An analysis of its stability is dependent on effective consolidation stresses, which in turn are controlled by pore pressure conditions. Pore pressure and saturation conditions govern whether the tailings of which the upstream tailings dams are often constructed will fail under drained or undrained conditions (Martin, 1999).



Photograph 2: Pictures that were taken at the foot of the FSSW8 tailings dam. The pictures indicate evidence of visible water on surface and increased vegetation growth at the foot of the dam.

The seismic stations were installed about 20 m below the cone penetration tests (CPTu). A CPTu is an in-situ test that is used to classify and identify subsurface conditions such as characterizing soil type, detecting layers and measuring strength, determining deformation characteristics and the permeability of foundation soils (Bard, 2019). The CPTu tests in the study area were conducted by Jones & Wagener Consultants (Olivier, et al., 2018).

The results from the CPTu tests and interpretations are shown in Figure 8.1(a). An example of the water pressure with depth as recorded by the piezocone probe during dissipation tests for hole P1B is given in Figure 8.2(b)

The first panel in Figure 8.2(b) indicates results from the cone resistance test. The resistance of the cone represents the ratio of the force measured at the tip of the cone to the projected area of the normal line of the tip of the cone.

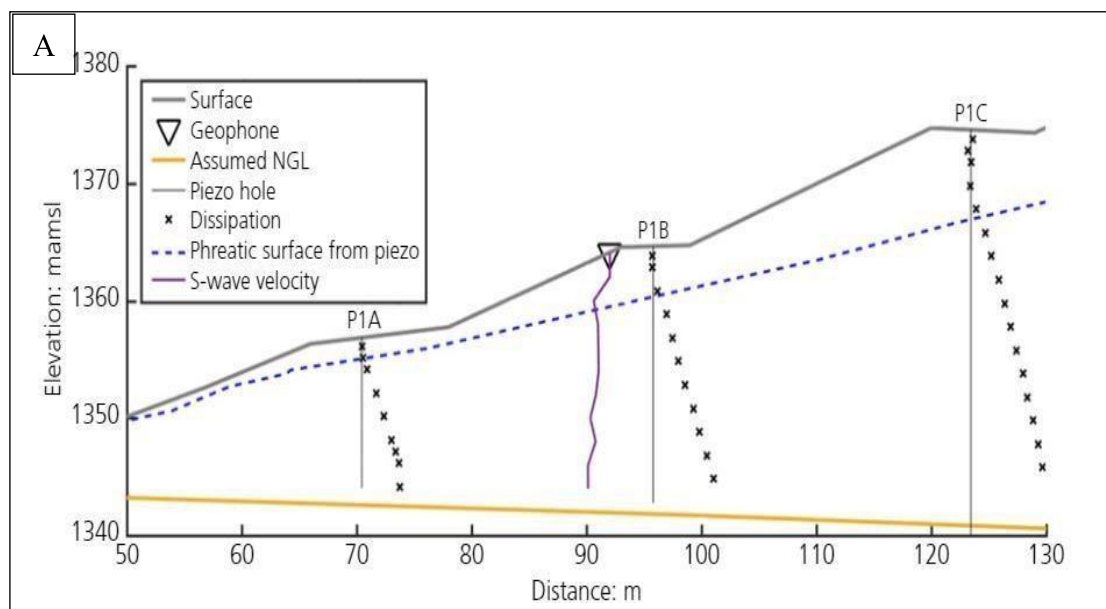
Cone resistance indicates the undrained (i.e., including in-situ moisture) shear strength of the soil (Morton, 2019). The cone resistance is higher between depths 0 and 4 m.

The data provided from pore pressure provides an estimate of the depth of the water table. This information is useful both for site characterization and for geo-environmental and remediation applications. The pore pressure is zero between 0 and 3m, this is the same depth that the S wave starts to increase, and this suggests that the water table is at this depth.

The third panel shows the results from the dissipation tests. The identification of the shape and location of the depression line is important in monitoring the stability of tailings dams. The line separates the unsaturated medium from the full saturated sediments. The dissipation line in panel 3 is at zero between 0 and 3 m depth, it starts to increase at a depth of 3 m.

The fourth panel indicates S- wave velocity. The S-wave velocity decreases between 2 and 4 m below the surface. This depth coincides with the depth at which pore pressure and dissipation increases. This suggests that the phreatic surface is most likely around 3 m.

This observation is consistent with the CPTu test, taking into account the position of the geophone line relative to the CPTu probe (Figure 8.1 (a)). Although the resolutions of the two methods are very different and difficult to compare directly, the obvious agreement between the results inspires people's confidence in the seismic survey results.



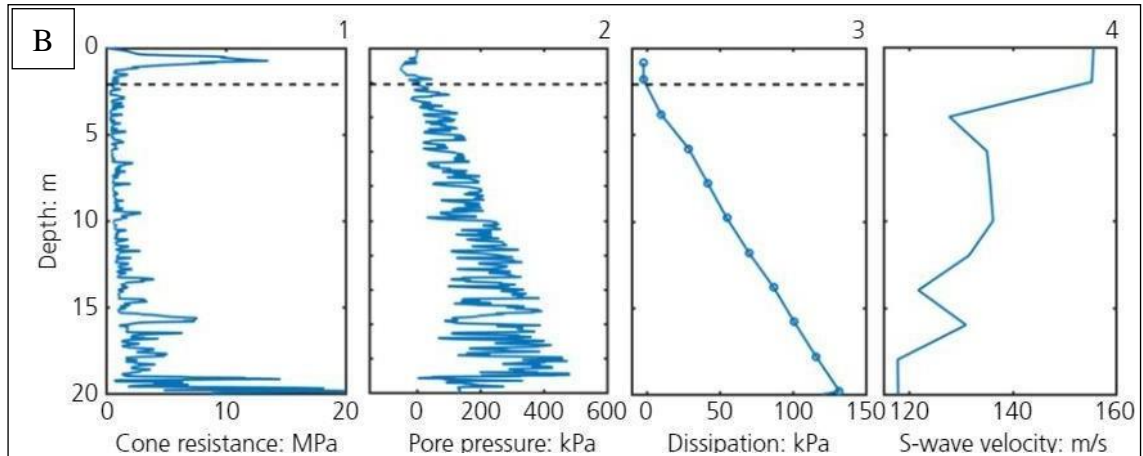


Figure 8.1: (a) Schematic diagram showing the location of the CPTu tests and the geophones. The grey line indicates the position of the surface and the blue dotted line indicates the position of the phreatic surface. (b) Three panels; panel 1 is cone resistance, panel 2 is pore pressure, panel 3 is dissipation and panel 4 is S-wave velocity (Olivier, et al., 2018).

8.2 Cradock Dataset

There are two distinct velocity zones that are observed on the western and eastern section of the tomography maps. The western section has low velocity structures and the eastern section shows high a velocity zone that is most prominent on the 2.5 second map which corresponds to a depth range between 2 and 3 km. Seismic velocity mainly depends on the temperature, pressure and composition of the rock. Generally, shear wave velocity increases with period and with depth. This suggests that the rocks that are at lower depth are denser than those that lie at shallower depths.

The tomography maps from the different periods do not correlate well with subsurface structures from the borehole data. The propagation of surface waves between the periods of 1.5 and 2.5 seconds are strongly influenced by the shallow crustal structures. The 1.5 second group velocity (Figure. 7.26), provide subsurface images that correspond to the shallow subsurface (1-1.5 km), the map indicates that sediments from the western and southern sections of the of the map have low velocities relative to other period maps, this is consistent with less dense and less compacted shallower sediments. The 2.5 second map shows that there is a general increase in the Rayleigh group velocity in the eastern part of the map. This suggests that there is a hard and denser rock layer on that lies at a shallower depth of 2 km. This high velocity structure fades with an increase in period and depth.

The Karoo Basin is characterized by a vast network of intrusive dolerite sills and dykes and there is a change in the distribution of dolerites throughout the basin. The largest concentration of dolerites is in the north-western and eastern parts of the basin with intrusions at all stratigraphic levels. Figure 8.3 is a stratigraphic column from SOEKOR borehole data and indicates that there are dolerites present at 386, 426 and 518 m. The 2-D shear wave velocity profile indicates that there is a high velocity layer at 350 m and there is a high velocity zone that is observed on the tomography maps on the eastern section of the map. The observed high velocity regions could be indicative of the dolerite sill intrusion, which is consistent with higher velocities often characterized with igneous and metamorphic rocks.

There are low velocity structures that are observed on all the maps (indicated by the blue and purple on Figures 7.26 –7.28) in the south (roughly between stations A9K, A9N, A9F and A9G) and north-west (roughly between stations A9P, A9Y, B5Y and AAE) of the study area. Station locations are shown in Figure 2.3. These low velocity structures, which extend to the north and south of the study area and lie on the transition from the Adelaide subgroup to the Tarkastad group. The Adelaide subgroup rocks are mudstone-dominated. The high velocity zone lies on the Tarkastad group of rocks which are dominated by fine to medium-grained sandstones. The difference in lithology of the overall study area could be the reason for the differences observed on the tomography maps.

The MSNoise software uses the Monte Carlo optimization method for the inversion and introduces variance reduction in the calculation to improve the accuracy of model fitting. Figure 7.26 to 7.28 shows the variance reduction of the remaining travel time and the average group velocity at different periods. The variance reduction obtained from the group velocity maps for periods is generally close to 50 % indicating that the modelled group velocity maps fits the data well.

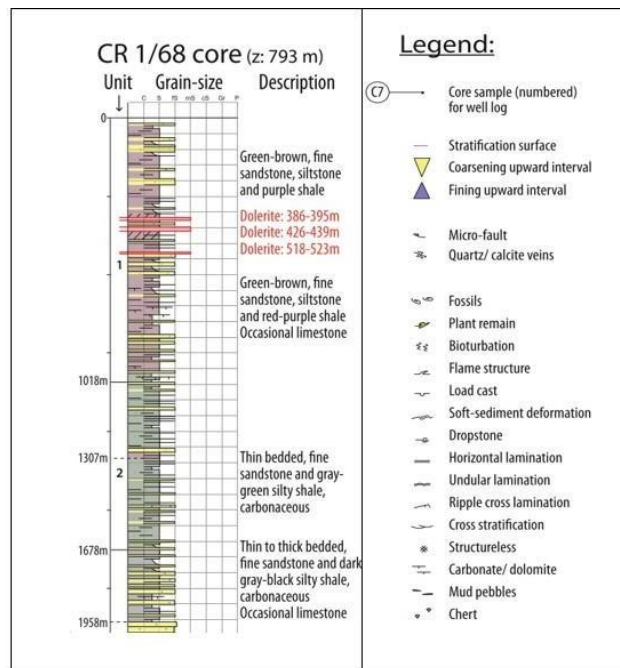


Figure 8.2: Description of the SOEKOR deep borehole that was drilled in Cradock. (Linol, et al., 2016).

9. Conclusion

9.1 Harmony Dataset

Tailings dams are constantly growing structures that require diligent monitoring and may be susceptible to static liquefaction and internal erosion. It is important for early internal signs of degradation, seepage or fluid saturation to be noted in time to prevent the dams from collapsing. Seismic methods are one of the few methods that can image the internal structure of these facilities continuously in space, as opposed to most geophysical methods that only provide point measurements. However, seismic methods have scarcely been used in this environment.

The main objective of this study was to investigate the use of ambient seismic noise as a tool to image the internal structure of a short section of a tailings dam that showed signs of increased seepage (Figure 8.1). To achieve this objective, estimates of Love wave virtual source signals between sensors were constructed by cross-correlating the seismic noise recorded in the horizontal transverse components of the geophones. These virtual source signals were used to create dispersion curves between different sensor pairs. The dispersion curves were inverted with an improved neighborhood algorithm to produce 1-D shear wave velocity profiles. The shear wave profiles were then interpolated to produce a 2-D shear wave cross-section across the survey line. The cross section indicated that there is a low-velocity zone around 12 m below surface. 15, 50 and 60 m of the array, plumes of low-velocity zones were found between 2 and 10 m below the surface indicating that the phreatic surface is much closer to surface in this area. This area was correlated to where seepage was identified on the surface of the dam wall. The results of the study were compared with the cone penetration tests (CPTu) carried out near the center of the geophone array and the results were comparable. The results obtained from this study indicated that ambient seismic noise interferometry can be used as a cost-effective and non-invasive method to image the internal structure of a tailings dam.

9.2 Cradock Dataset

The main objective of this study was to apply seismic interferometry to map the subsurface of Cradock. This objective was achieved by the deployment of seismic stations in the study area. The stations were left to record data for a period of 35 days. The Rayleigh waves extracted from the vertical component were used for the rest of the study. The waveforms were pre-processed to accentuate the broadband ambient noise and then they were cross-correlated and stacked to obtain an estimate of the Green's function.

The cross-correlation functions were used to calculate phase velocity and dispersion curves between different sensor pairs. The resulting dispersion curves were inverted with an improved neighborhood algorithm to produce 1-D shear wave velocity profiles. The shear wave profiles were then interpolated to produce a 2-D shear wave cross-section across the study area. The 2-D cross-section was compared with the SOEKER borehole and did not correlate very well. The cross section indicated that there is a layer of low velocity up to 200 m. This was interpreted as soil fill material that consists of loose and larger stones. The velocity starts to increase from 200 m and varies from 1000 to 1500m/s, the increase possibly indicates soft weathered rock or soft sediments, there is a low velocity zone between depths 250 to 300 m between 6 and 14 km. The borehole stratigraphic column indicates a presence of dolerite at 386, 426 and 518 m.

The cross-correlations were also used to compute group velocity maps from periods 1.5 seconds to 30 seconds. The group velocity maps showed various high and low velocity anomalies. The high velocity zones observed on the eastern section of the map were interpreted as evidence of dolerite intrusions. The low velocity zones observed in the western and southern sections of the map interpreted as Karoo sediments that belong to the Adelaide Subgroup which is dominated by mudstones.

This study demonstrates that ambient seismic noise interferometry can be used as an effective and non-invasive tool to image the subsurface and the interior of a tailings dams wall.

9.3 Recommendations

Possible future work that could increase the understanding of the subsurface in the Karoo region includes:

- Compute dispersion curves and inversions to obtain a 3D shear velocity model.
- Use a seismographic network with a much larger station density. Increasing station density would allow for the investigation of high frequency noise which attenuates rapidly and could not be investigated in my thesis.
- Investigate the radial and transverse (horizontal) components of ambient noise can be effective in identifying p-wave and s-wave phases. Then, these phases can be subject to polarization analysis (frequency-time analysis) which would compare the particle motion of different observed phases to that of a synthetic seismogram created at similar inter-station distance. This type of analysis was beyond the scope of this study but would be an effective way to confirm the observed phases in cross-correlations.

10. References

- Akademi, Ö., 2009. Introduction to Geophysics. *Slideshare*, [online]. Available at: <https://www.slideshare.net/oncel/seismic-reflection-4-acquisiton-processing-and-waveform-analysis-2755918> [Accessed 3 March 2021].
- Aarnes, I., Svensen, H., Polteau, S. & Planke, S., 2010. Contact metamorphic devolatilization of shales in the Karoo Basin, South Africa, and the effects of multiple sill intrusions. *Chemical Geology* 281, Volume 2011, pp. 181-194.
- Andrade, F., 2017. Safety and monitoring of tailings dams. *Visaya*, [online]. Available at: <https://visaya.solutions/en/article/safety-tailings-dam-monitoring> [Accessed 21 January 2020].
- Azam, S. & Li, Q., 2010. Tailings Dam Failures: A Review of the last one hundred years. *Geotechnical News*, Volume 28, pp. 50-54.
- Baiyegunhi, C., Gwavava, O., Liu, K. & Baiyegunhi, T., 2019. An Integrated Geophysical Approach to Mapping and Modelling the Karoo Dolerite Intrusions in the South-Eastern Karoo Basin of South Africa. *Journal of Engineering and Applied Sciences*, Volume 14, pp. 1885-1911.
- Bakulin, A. & Calvert, R., 2004. Virtual source: new method for imaging and 4D below complex overburden. *SEG technical program expanded abstracts*.
- Bard, E., 2019, *Seminar 1: Liquefaction Assessment of Deposited Thickened Tailings using CPTu Soundings Calibrated with Laboratory Testing: A Critical State Soil Mechanics Application*, lecture notes, GolderTv: tailings, delivered 11 July 2019.
- Bensen, G., Ritzwoller, M., Barmin, M., Levshin, A., Lin, F., Moschetti, M., Shapiro, M., Yang, Y., 2007. Processing seismic ambient noise data to obtain reliable broadband surface wave dispersion measurements. *Geophysical Journal International*, Volume 169, p. 1239–1260.
- Bezuidenhout, L., Doucouré, M., Wagener, V. & de Wit, M., 2016. Ambient noise tomography (Passive seismic) to image the Cape-Karoo Transition near Jansenville, Eastern Cape. In: M. de Wit & B. Linol, eds. *Origin and evolution of the Cape Mountains and Karoo Basin*. s.l.:Springer, pp. 27-32.
- Bezuidenhout, L. & Doucouré, M., 2020. Source azimuth determination of ambient seismic noise in the Eastern Cape. *Journal of African Earth Sciences*, Volume 164, pp. 1-9

Bormann, P., Engdahl, E. & Kind, R., 2012. Elastic modules and body waves. In: *New Manual of Seismological Observatory Practice 2*. Potsdam: GFZ German Research Centre for Geosciences, pp. 3-13.

Boschi, L., 2014. Seminar 2: Italian and Alpine crustal structure imaged by ambient-noise surface-wave dispersion, [online]. *American Geophysical Union*. Available at: <https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2015GC006176> [Accessed 16 December 2014].

Brenguier, F., Clarke, D., Aoki, Y., Shapiro, N., Campillo, M., Ferrazzini, V., 2011. Monitoring volcanoes using seismic noise correlations. *Geoscience*, Volume 343, pp. 633-638.

Campillo, M. & Paul, A., 2003. Long-Range Correlations in the Diffuse Seismic Coda. *Science*, Volume 299, pp. 547-549.

Carcione, M. J. & Helle, B. H., 2002. Rock physics of geopressure and prediction of abnormal pore fluid pressure using seismic data. *CSEG Recorder*, pp. 9-32.

Catuneanu, O., Wopfner, H., Eriksson, P., Cairncross, B., Rubidge, B., Smith, R., Hancox., 2005. The Karoo basins of south-central Africa. *Journal of African Earth Sciences*, Volume 43, pp. 211-253.

Chen, S. E. & Ong, C. K., 2012. Spectral Analysis of Surface Wave for Empirical Elastic Design of Anchored Foundations. *Advances in Civil Engineering*, Volume 2012, pp. 1-10.

Curtis, A., Gerstoft, P., Sato, H., Sneider, R., Wapenaar, K., 2006. Seismic Interferometry-turning noise into signal. *The leading edge*, Volume 25, pp. 1082-1092.

Cutifani, M., 2019. Tailings dam factsheet. *Anglo-american*, [online]. Available at: <https://www.angloamerican.com/~media/Files/A/Anglo-American-Group/PLC/sustainability/2019-tailings-factsheet.pdf> [Accessed 20 November 2019].

Das, R. & Rai, S., 2016. Seismic interferometry and ambient noise tomography: theoretical background and application in South India. *Journal Physics: Conference Series*, 759, pp. 1-7.

DiGiulio, G., Cornou, C., Ohrnberger, M., Wathelet, M., Rovelli, A., 2006. Deriving wavefield characteristics and shear-wave velocity profiles from two-dimensional small-aperture arrays analysis of ambient vibrations in a small-size alluvial basin, Colfiorito, Italy. *Bulletin of the Seismological Society of America*, 26(5), pp. 1915-1933.

Di Giulio, G., A. Savvaidi, N. Theodoulidi, M. Ohrnberger, B. Endrun., 2010. Inversion of surface wave dispersion at European strong motion sites using a multi-model

parameterization and an information-theoretic approach. *14th European Conference on Earthquake Engineering*.

Draganov, D., Wapenaar, K. & Thorbecke, J., 2006. Seismic interferometry: reconstructing the earth's reflection response. *Geophysics*, Volume 71, pp. SI61-SI70.

Fichtner, A., Gualtieri, L. & Nakata, N., 2019. Visualisation of the seismic ambient noise spectrum. In: N. Nakata, L. Gualtieri & A. Fichtner, eds. *Seismic ambient noise*. New York: Cambridge, pp. 1-6.

Foti, S., Hollender, F. & Garofalo, F., 2018. Guidelines for the good practice of surface wave analysis: a product of the Interpacific project. *Bull Earthquake Engineering*, Volume 16, p. 2367–2420.

Ganesh, M., 2006. *Monitoring of tailings dams with geophysical methods*, PhD thesis, Luleå University of Technology, Sweden.

Hirose, T., Nakahara, H. & Nishimura, T., 2017. Combined use of repeated active shots and ambient noise to detect temporal changes in seismic velocity: application to Sakurajima volcano, Japan. *Earth, Planets and Space*, Volume 69, pp. 1-12.

Ivanov, J., Miller, R. & Xia, J., 2010. Multi-mode inversion of multi-channel analysis of surface waves (MASW) dispersion curves and high-resolution linear radon transform (HRLRT). *SEG Denver 2010 Annual Meeting*, pp. 1902-1907.

Ivanov, J., Park, B. C., Miller, D. R. & Xia, J., 2001. Modal separation before dispersion curve extraction by MASW method. *Symposium on the Application of Geophysics to Engineering and Environmental Problems Proceedings*, pp. 1-11.

Johnson, M. R., Vuuren, V., Hegenberger, W., Key, R., Show, U., et al., 1996. Stratigraphy of the Karoo Supergroup in Southern Africa: a review. *Journal of African Earth Sciences*, Volume 23, pp. 3-15.

Johnson, M. R., Vuuren, V., Visser, J., Cole, D., Wickens, H., Christie, A., Robert, D., Brandl, G., et al., 2006. Sedimentary rocks of the Karoo. In: M. R. Johnson, C. R. Anhaeusser & R. J. Thomas, eds. *The Geology of South Africa*. Johannesburg: Council of Geoscience, pp. 461-495.

Kiwamu, N., 2017. Ambient seismic wave field. *Proceedings of the Japan Academy. Series B, Physical and Biological sciences*, Volume 93, pp. 423-448.

Lecocq, T., Longuevergne, L., Pederson, H., Brenguier, F., Stammler, K., 2017. Monitoring ground water storage at mesoscale using seismic noise: 30 years of continuous observation and thermo-elastic and hydrological modelling. *Scientific Reports*, Volume 7, pp. 1-13.

Lecocq, T., Caudron, C. & Brenguier, F., 2014. MSNoise, a Python Package for Monitoring Seismic Velocity Changes Using Ambient Seismic Noise. *Seismological Research Letter*, Volume 85, pp. 715-726.

Lepore, S. & Grad, M., 2018. Analysis of the primary and secondary microseisms in the wavefield of the ambient noise recorded in northern Poland. *Acta Geophysics*, Volume 66, pp. 915-929.

Lee, V. W. & Trifunac, M., 2010. Should average shear-wave velocity in the top 30 m of soil be used to describe seismic amplification?. *Soil Dynamics and Earthquake Engineering*, Volume 30, pp. 1250-1258.

Lin, F.-C., Ritzwoller, M., & Snieder, R. (2009). Eikonal tomography: surface wave tomography by phase front tracking across a regional broad-band seismic array. *Geophysical Journal International*, Volume 177(3), 1091-1110.

Lindeque, A., de Wit, M., Ryberg, T., Weber, M., Chevallier, L., 2011. Deep crustal profile across the southern Karoo basin and Beattie magnetic anomaly, South Africa: An integrated interpretation with tectonic implications. *South African Journal of Geology*, Volume 114, pp. 265-292.

Lin, F.-C., Ritzwoller, M., Townend, J., Bannister, S., Savage, M., 2007. Ambient Noise Rayleigh Wave Tomography of New Zealand. *Geophysical Journal International*, Volume 170, pp. 649-666.

Lin, F.-C., Moschetti, P. M. & Ritzwoller, H. M., 2008. Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps. *Geophysical Journal International*, Volume 173(1), pp. 281–298.

Linol, B., Chere, N., Muedi, T., Nengovhela, V., de Wit, M., 2016. Deep borehole lithostratigraphy and basin structure of the Southern Karoo basin re-visited. In: M. de Wit & B. Linol, eds. *Origin and evolution of the Cape Mountains and Karoo basin*, Springer, Switzerland, pp. 3-16.

Lobkis, O. & Weaver, R., 2001. On the emergence of the Green's function in the correlations of a diffuse field. *Journal of the acoustical society of America*, Volume 110, pp. 3011-3017.

Lucas, S. G. & Hancox, P. J., 2001. Tetrapod-based correlation of the non-marine Upper Triassic of Southern Africa. *Albertina*, Volume 25, pp. 5-9.

Lyu, Z., Chai, J., Xu, Z., Qin, Y., Cao, J., 2019. A Comprehensive Review on Reasons for Tailings Dam Failures Based on Case History. *Advances in Civil Engineering*, Volume 2019, pp. 1-18.

Martin, T., 1991. Characterization of pore pressure conditions in upstream tailings dams. *CiteSeerX*, [online]. Available at: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.489.13&rep=rep1&type=pdf> [Accessed 10 October 2020]

Mcnamara, D. E. & Boaz, R. I., 2019. Visualization of the Seismic Ambient Noise Spectrum. In: N. Nakata, L. Gualtieri & A. Fitchner, eds. *Seismic Ambient Noise*. Cambridge: Cambridge University Press, pp. 1-18.

Mordret, A., Jolly, A., Duputel, Z. & Fournier, N., 2010. Monitoring of phreatic eruptions using Interferometry on Retrieved Cross-Correlation Function from Ambient Seismic Noise: Results from Mt. Ruapehu, New Zealand. *Journal of Volcanology and Geothermal Research*, Volume 191, pp. 46-59.

Mordret, A., Landès, M., Shapiro, N., Singh, S., Roux, P., Barkved., 2013. Near-surface study at the Valhall oil field from ambient noise surface wave tomography. *Geophysical Journal International*, Volume 19, pp. 1627–1643.

Morton, K., 2019. Accurate pore pressure monitoring reduces tailings dam risk. *Mining Review Africa*, [online]. Available at: <https://www.miningreview.com/health-and-safety/accurate-pore-pressure-monitoring-reduces-tailings-dam-risk/> [Accessed 11 September 2020]

Ndikum, E., Tabod, C. & Tokam, A.-P., 2014. Frequency Time Analysis (FTAN) and Moment Tensor Inversion Solutions from Short Period Surface Waves in Cameroon (Central Africa). *Open Journal of Geology*, Volume 4, pp. 1-11.

Nishida, K., 2017. Ambient seismic wave field. *Proceedings of the Japan Academy, series B Physical and biological sciences*, Volume 93, pp. 423–448.

Nicolson, H., Curtis, A., Baptie, B. & Galetti, E., 2012. Seismic interferometry and ambient noise tomography in the British Isles. *Proceedings of the Geologists' Association*, Volume 123, pp. 74-86.

O'Dwyer, R., 2019. Innovating to avoid the next tailings dam failure. *Trends in data science*, [online]. Available at: <https://medium.com/trends-in-data-science/innovating-to-avoid-the-next-tailings-dam-failure-a86e29ab355a> [Accessed 15 January 2020].

Olivier, G., de Wit, T., Brenguier, F., Bezuidenhout, L., Kunjwa, T., 2018. Ambient noise Love wave tomography at a gold mine tailings storage facility. *Geotechnique Letters*, pp.1-6.

Paoletti, V., 2012. Remarks on factors influencing shear wave velocities and their role in evaluating susceptibilities to earthquake-triggered slope instability: Case study for the Campania area (Italy). *Natural Hazards and Earth System Science*, Volume 12, pp. 2147-2158.

Park, C., 2005. MASW: Horizontal Resolution in 2D Shear-Velocity (V_s) Mapping, *Kansas: Kansas Geological Survey, Kansas University*, [online].

Available at: http://www.kgs.ku.edu/Geophysics/OFR/2005/OFR05_04/OF-2005-4.pdf [Accessed 13 September 2020].

Park, C., Miller, R. & Xia, J., 1998. Imaging dispersion curves of surface waves on multichannel record. *SEG Technical program expanded abstracts*, pp. 1-4.

Park, C. B., Miller, R. D., Xia, J. & Ivanov, J., 2007. Multichannel analysis of surface waves (MASW)—active and passive methods. *Geophysics*, Volume 26, pp. 60-64.

Planès, T., Mooney, M., Rittgers, J., Parekh, M., Behm, M., Snieder, R., 2016. Time-lapse monitoring of internal erosion in earthen dams and levees using ambient seismic noise. *Géotechnique*, Volume 66, pp. 301-312.

Prevost, N., 2019. A guide to tailings dam monitoring. *Inmarsat*, [online]. Available at: <https://www.inmarsat.com/en/insights/enterprise/2019/a-guide-to-tailingsdam-monitoring.html> [Accessed 10 August 2020].

Rawlinson, N. & Spakman, W., 2016. On the use of sensitivity tests in seismic tomography. *Geophysical Journal International*, Volume 205(2), pp. 1221-1243.

Reimold, W. U., 2006. Impact structures in South Africa. In: M. R. Johnson, C. R. Anhaeusser & R. J. Thomas, eds. *The Geology of South Africa*. Pretoria: Council for Geoscience, pp. 629-645.

- Rodrigues, C., de Paula, A., Corrêa, T., Sebastião, C., Costa, O., Magalhães, G., Santana, L., 2019. Passive seismic interferometry's state-of-the-art - a literature review. In: J. Tournier, T. Bennett & J. Bibeau, eds. *Sustainable and safe dams around the world*. London: CRC Press.
- Ryberg, T., 2011. Body wave observations from cross-correlations of ambient seismic noise: A case study from the Karoo, RSA. *Geophysical Research letters*, 38(13), pp. 1-5.
- Sabra, K. et al., 2005. Extracting time-domain Green's function estimates from ambient seismic noise. *Geophysical research letters*, Volume 32, pp. 1-5.
- Sambridge, M. (1999). Geophysical inversion with a neighbourhood algorithm—I. Searching a parameter space. *Geophysical Journal International*, Volume 138 (2), pp. 479-494.
- Schuck, A. & Lange, G., 2007. Seismic Methods. In: *Environmental Geology*. s.l.:Springer, Berlin, Heidelberg, pp. 337-402.
- Schuster, G., Yu, J., Sheng, J. & Rickett, J., 2004. Interferometric/daylight imaging. *Geophysical Journal International*, Volume 157, pp. 838-852.
- Shapiro, N. M. & Campillo, M., 2004. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise. *Geophysical research letters*, Volume 31(7), pp. 1-4.
- Shapiro, N., Campillo, M., Stehly, L. & Ritzwoller, M., 2005. High-Resolution Surface Wave Tomography from Ambient Seismic Noise. *Science*, Volume 307, pp. 1615-1623.
- Scheiber-Enslin, S., Webb, S. & Manzi, S., 2016. Seismic imaging of Dolerite Sills in the Karoo Basin, with implications for shale gas potential. In: M. de Wit & B. Linol, eds. *Origin and evolution of the Cape Mountains and Karoo Basin*, Springer, pp. 17-26.
- Snieder, R., 2004. Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase. *Physical Review E*, Volume 69, pp. 046610-1-046610-8.

Stankiewicz, J., Trond, T., Haberland, C., Natawidjaja, F., Natawidjaja, D., 2010. Lake Toba volcano magma chamber imaged by ambient seismic noise tomography. *Geophysical research letters*, Volume 37, pp.1-5.

Stehly, L., Campillo, M. & Shapiro, M., 2006. A study of the seismic noise from its long-range properties. *Journal of geophysical research*, Volume 111, pp. 1-12.

Stehly, L., Campillo, M., Shapiro, N., Guilbert, J., Boschi, L., Giardini, D., 2009. Tomography of the Alpine region from observations of seismic ambient noise. *Geophysical Journal International*, Volume 178, pp. 338-350.

Wallingford, H. R., Roca, M., Murphy, A., Walker, L., Vallesi, S., 2019. A review of the risks posed by the failure of tailings dams, *Oxfordshire: HR Wallingford Ltd*, [online]. Available at: <https://damsat.org/wp-content/uploads/2019/01/BE-090-Tailings-dams-R1-Secured.pdf> [Accessed 18 July 2020].

Wapenaar, K., Draganov, D., Sneider, S., Campman, X., Verdel, A., 2010. Tutorial on seismic interferometry: Part1 - Basic principles and applications. *Geophysics*, Volume 75, pp. 75A195-75A209.

Wapenaar, K. & Fokkema, J., 2006. Green's function representations for seismic interferometry. *Geophysics*, Volume 71(4), pp. SI33-SI46.

Wathelet, M., 2008. An improved neighborhood algorithm: Parameter conditions and dynamic scaling. *Geophysical Research letters*, Volume 35(9), pp. 1-5.

Wathelet, M., Jongmans, D. & Ohrnberger, M., 2005. Direct Inversion of Spatial Autocorrelation Curves with the Neighborhood Algorithm. *Seismological Society of America*, Volume 95, pp. 1787-1800.

Wathelet, M., Jongmans, D. & Ohrnberger, M., 2004. Surface-wave inversion using a direct search algorithm and its application to ambient vibration measurements. *Near Surface Geophysics*, Volume 2, pp. 211-221.

Webb, S. J., Scheiber-Enslin, S. & Mnazi, M., 2016. Reprocessing Seismic Data - Using Wits Seismic Exploration Data to Image the Karoo Basin. *American Geophysical Union: Fall meeting abstracts*, pp. 1-5.

Xia, J., Miller, R., Park, C., Hunter, J., Harris, J., Ivanov, J., 2000. Comparing Shear-Wave Velocity Profiles from MASW with Borehole Measurements in Unconsolidated

Sediments, Fraser River Delta, B.C., Canada. *Journal of Environmental & Engineering Geophysics*, 5(3), pp. 1-13.

Yang, Y. & Ritzwoller, M., 2008. Characteristics of ambient seismic noise as a source for surface wave tomography. *Geochemistry, Geophysics, Geosystems*, Volume 9, pp. 1-18.

Appendix A

Pre-Processing

```

%%Loop for all data processing
% Nr 1 Importing the daily data for all stationstic clc clear all close
all

%loading all stations

    % filenames =
    {'A9E.mat', 'A9F.mat', 'A9G.mat', 'A9H.mat', 'A9J.mat', 'A9K.mat', 'A9L.mat',
    'A9M.mat', 'A9N.mat', 'A9P.mat'};    filenames
    =
    {'A9M.mat', 'A9J.mat', 'A9W.mat', 'A9H.mat', 'A9G.mat', 'A9U.mat', 'A9F.mat',
    'A9E.mat', 'A9R.mat', 'A9K.mat', 'A9X.mat', 'A9L.mat', 'A9Y.mat', 'A9V.mat',
    'A9P.mat', 'A9N.mat', 'A9Z', 'AAD.mat'};

for kk = 1:numel(filenames) load(filenames{kk}) end

%making the stations the same size
% x = [size(AEvert,1),size(A9Fvert,1),size(AGvert,1),size(A9Hvert,1),size(A9J
,1),size(A9Kvert,1),size(A9Lvert,1),size(AMvert,1),size(A9Nvert,1),size
(A9Pvert,1)]; x
=
[size(AMvert,1),size(A9J,1),size(A9Wvert,1),size(A9Hvert,1),size(AGvert
,1),size(A9Uvert,1),size(A9Fvert,1),size(AEvert,1),size(A9Rvert,1),size
(A9Kvert,1),size(A9Xvert,1),size(A9Lvert,1),size(A9Yvert,1),size(A9Vver
t,1),size(A9Pvert,1),size(A9Nvert,1),size(A9Zvert,1),size(AADvert,1)];
z=min(x);
%
A9Evert=AEvert(1:z);A9Fvert=A9Fvert(1:z);A9Gvert=AGvert(1:z);A9Hvert=A9
Hvert(1:z);A9Jvert=A9J(1:z);A9Kvert=A9Kvert(1:z);A9Lvert=A9Lvert(1:z);A
9Mvert=AMvert(1:z);A9Nvert=A9Nvert(1:z);A9Pvert=A9Pvert(1:z);
A9Mvert=AMvert(1:z);A9Jvert=A9J(1:z);A9Wvert=A9Wvert(1:z);A9Hvert=A9Hve
rt(1:z);A9Gvert=AGvert(1:z);A9Uvert=A9Uvert(1:z);A9Fvert=A9Fvert(1:z);A
9Evert=AEvert(1:z);A9Rvert=A9Rvert(1:z);A9Kvert=A9Kvert(1:z);A9Xvert=A9
Xvert(1:z);A9Lvert=A9Lvert(1:z);A9Yvert=A9Yvert(1:z);A9Vvert=A9Vvert(1:
z);A9Pvert=A9Pvert(1:z);A9Nvert=A9Nvert(1:z);A9Zvert=A9Zvert(1:z);AADve
rt=AADvert(1:z);

    disp('loaded and all thestations')

%%Divide stations into 10 minutes each nr_of_stations
= 18;
%
nr_of_ten_min_col = 142;%There are 142 ten minutes in 23 hours ten=240000;
nr_of_columns = nr_of_ten_min_col*nr_of_stations; count
= 0;
for i=1:nr_of_ten_min_col%nr_of_columnscount = count +1;
A9MNew(:,count) = A9Mvert((i-1)*ten)+1:ten*i);A9JNew(:,count) =
A9Jvert((i-1)*ten)+1:ten*i);A9WNew(:,count) = A9Wvert(((i-
1)*ten)+1:ten*i);A9HNew(:,count) = A9Hvert(((i-
1)*ten)+1:ten*i);A9GNew(:,count) = A9Gvert(((i-
1)*ten)+1:ten*i);A9UNew(:,count) = A9Uvert(((i-
1)*ten)+1:ten*i);A9FNew(:,count) = A9Fvert(((i-
1)*ten)+1:ten*i);A9ENew(:,count) = A9Evert(((i-
1)*tn)+1:ten*i);A9RNew(:,count) = A9Rvert(((i-
1)*ten)+1:ten*i);A9KNew(:,count) = A9Kvert(((i-
1)*ten)+1:ten*i);A9XNew(:,count) = A9Xvert(((i-
1)*ten)+1:ten*i);A9LNew(:,count) = A9Lvert(((i-
1)*ten)+1:ten*i);A9YNew(:,count) = A9Yvert(((i-

```

```

1)*ten)+1:ten*i);A9VNew(:,count) = A9Vvert((i-
1)*ten)+1:ten*i);A9PNew(:,count) = A9Pvert((i-
1)*ten)+1:ten*i);A9NNew(:,count) = A9Nvert((i-
1)*ten)+1:ten*i);A9ZNew(:,count) = A9Zvert((i-
1)*ten)+1:ten*i);AADNew(:,count) = AADvert((i-1)*ten)+1:ten*i); end
disp('cut all the data into 10 min segments')

%Combine all the stations
%
% alldata =
[A9ENew,A9FNew,A9GNew,A9HNew,A9JNew,A9KNew,A9LNew,A9MNew,A9NNew,A9PNew]
;
alldata =
[A9MNew,A9JNew,A9WNew,A9HNew,A9GNew,A9UNew,A9FNew,A9ENew,A9RNew,A9KNew,
A9XNew,A9LNew,A9YNew,A9VNew,A9PNew,A9NNew,A9ZNew,AADNew,];
disp('Combined all thestations')

%Determine the size of all data k
=size(alldata,2);
%
%Define matlab built-in bandpass filter digfilt =
designfilt('bandpassiir', 'StopbandFrequency1', 6,
'PassbandFrequency1', 7, 'PassbandFrequency2', 40,
'StopbandFrequency2', 41, 'StopbandAttenuation1', 60, 'PassbandRipple',
1, 'StopbandAttenuation2', 60, 'SampleRate', 400);
Fs=400;%Sampling frequencyfor j=1:k
%1. Remove the mean
Data_Demean(:,j) = alldata(:,j) - mean(alldata(:,j));%Removes mean
%2. Let's remove linear trend Data_Detrend(:,j) =
detrend(alldatanew(:,j));

%3. filtering data
Data_Filtered(:,j) = Data_Detrend(:, j); % Extract this one column
% fc = cutoff(k); % Extract this one cutoff value.
% Determine filter parameters.
% Do the filtering, and put result into column k of the output y
Data_Filtered(:,j) = filter(digfilt, Data_Detrend(:,j));
% disp('step3')

% 4. normalization
%The normalistaion method that is applied here is called the running
absolute mean normalisation(look Bensen et al)
%Time-domain normalization is a procedure for reducing the effect onthe
crosscorrelations of earthquakes, instrumental irregularities and
nonstationary noise sources near to stations.

```

```

%This method computes the running average of the absolute value of the
waveform in a normalization time window of fixed length and weights the
waveform at the centre of the window by the inverse of this average.
Norm_window=0.07; %%normilazation window
%The nrmalisation window should be half of the maximum frequency wichthe
data was filtered throuhg.
%The data was filtered between 7 and 40 Hz,the maximum is 0.142seconds,
half of that is 0.07
weightconstant = 1/((2*Norm_window)+1);
%This formula is from Bensen's paper and is used to determine the
normalisation weight.
allsums(j) = sum(abs(Data_Filtered(:,j)));
weight(j)=weightconstant*allsums(j);
newnormdata(:,j) = Data_Filtered(:,j)/weight(j);%newnormdata is the
normalised data

%5. Spectral whitening
whitendata(:,j) = whitening(newnormdata(:,j), Fs, 'freq', [7,40]);
end
disp('Completed preprocessing for all
the stations') clearvars -except
whitendata
% save whitendata

%%Calculate thw combination of recievers needed for the cross
%totalxcorrs = ((nr)*(nr-1))/2;

%%Perfoming Cross correlations (cc) using matlabs Xcorr function tenslice
= 142; %Defining the number of columns for each station nr_of_stations =
18; alltenslice = nr_of_stations*tenslice; count
= 0; mp=max(size(whitendata(:,1)));%Maximum where peaks appear for
i=1:alltenslice for j=i+tenslice:tenslice:alltenslice
count = count+1; cc_data(:,count) = xcorr(whitendata(:,i),
whitendata(:,j)); end end
toc
disp('completed the CC')

%%Pause program, Start definition of cc peaks%%

%Cutting data the data to where peaks appear
% % % %
kk =size(cc,2);%Get the size of cccount = 0; mp = 240000;
% mp=max(size(whitendata(:,1)));for col = 1:kk;
yy(:,col) = xx(mp-5*400:mp+5*400,col);
end
z=size(yy);
% % %%After the crosscorrelation take the hilbert transform of the signal
%
```

```

%%Take the hilbert transform to the signals.The bigger the factor the
smoother the signals become.
%Loop to test the value of the smoothing factor on the signals
smoothing_factor = 20; %More than 20 produces smoother envelope
for i =1:z(2);
    hilbert_cc(:,i)=envelope(yy(:,i), 20,'peak'); %envelope built in
matlab function to find the hilbert of the signal end % %
    %%%Averaging the cross correlations%% data = yyy; [n,m] =
size(data);
B = zeros(4001,153);
A = data; for i =1:153
B(:,i) = mean(A(:,(i-1)*142+1:i*142),2); end

%%Distance matrix created on seperate mfile distance =
nonzeros(flipud(tril(repmat((95:-5:5)', 1,95/5)))));

%%Plotting and stacking the cross correlationsfigure(1) hold on d =
distance; dnew = d/145; sr = 400; sz=size(B); n_pairs=min(sz);
l_xc=max(sz);
for i=1:n_pairs for j = 1:142:n_pairsplot((-
l_xc+1:2:l_xc)/(2*sr), (B(
:,i))+dnew(i), 'k')
%Possitve and negativeside3 end end d_max=max(d)*1.1;

```

Pre-processing workflow

```

%loading all stations
% filenames =
{'A9E.mat','A9F.mat','A9G.mat','A9H.mat','A9J.mat','A9K.mat','A9L.mat',
'A9M.mat','A9N.mat','A9P.mat'};
% filenames =

```



```

{'A9M.mat', 'A9J.mat', 'A9W.mat', 'A9H.mat', 'A9G.mat', 'A9U.mat', 'A9F.mat',
'A9E.mat', 'A9R.mat', 'A9K.mat', 'A9X.mat', 'A9L.mat', 'A9Y.mat', 'A9V.mat', '
A9P.mat', 'A9N.mat', 'A9Z', 'AAD.mat'};

filenames =
{'A9M.mat', 'A9J.mat', 'A9W.mat', 'A9H.mat', 'A9G.mat', 'A9U.mat'};

for kk = 1:numel(filenames) load(filenames{kk}) end

% %Remove zeros stations
A9M( ~any(A9M,2), : ) = []; %rows
A9J( ~any(A9J,2), : ) = []; %rows
A9W( ~any(A9W,2), : ) = []; %rows
A9H( ~any(A9H,2), : ) = []; %rows
A9G( ~any(A9G,2), : ) = []; %rows
A9U( ~any(A9U,2), : ) = []; %rows
%

%making the stations the same size
% x = [size(AEvert,1), size(A9Fvert,1), size(AGvert,1), size(A9Hvert,1), size(A9J
,1), size(A9Kvert,1), size(A9Lvert,1), size(AMvert,1), size(A9Nvert,1), size
(A9Pvert,1)]; x
=
[size(A9M,1), size(A9J,1), size(A9W,1), size(A9H,1), size(A9G,1), size(A9U,1
)];
z=min(x);
A9Mvert=A9M(1:z); A9Jvert=A9J(1:z); A9Wvert=A9W(1:z); A9Hvert=A9H(1:z); A9G
vert=A9G(1:z); A9Uvert=A9U(1:z);
%A9Gvert=AGvert(1:z);

%%Divide stations into 10 minutes each nr_of_stations
= 6;
%
nr_of_ten_min_col = 142;%There are 142 ten minutes in 23 hours ten=240000;
nr_of_columns = nr_of_ten_min_col*nr_of_stations; count
= 0;
for i=1:nr_of_ten_min_col%nr_of_columnscount = count +1;
A9MNew(:,count) = A9Mvert(((i-1)*ten)+1:ten*i); A9JNew(:,count) =
A9Jvert(((i-1)*ten)+1:ten*i); A9WNew(:,count) = A9Wvert(((i-
1)*ten)+1:ten*i); A9HNew(:,count) = A9Hvert(((i-
1)*ten)+1:ten*i); A9GNew(:,count) = A9Gvert(((i-
1)*ten)+1:ten*i); A9UNew(:,count) = A9Uvert(((i-1)*ten)+1:ten*i); end

%Combine all the stations
%
alldata =
[A9ENew, A9FNew, A9GNew, A9HNew, A9JNew, A9KNew, A9LNew, A9MNew, A9NNew, A9PNew]
; alldata = [A9MNew, A9JNew, A9WNew, A9HNew, A9GNew, A9UNew];

```

```

%Determine the size of all data k
=size(alldata,2);
%

%Define matlab built-in bandpass filter
digfilt = designfilt('bandpassiir', 'StopbandFrequency1', 6,
'PassbandFrequency1', 7, 'PassbandFrequency2', 40,
'StopbandFrequency2', 41, 'StopbandAttenuation1', 60, 'PassbandRipple',
1, 'StopbandAttenuation2', 60, 'SampleRate', 400);
Fs=400;%Sampling frequency
for j = 1:k demeaned(:,j) =rmean(alldata(:,j));

detrended(:,j) = rtrend(demeaned(:,j));

filtered_data(:,j) = bp_bu_co(detrended(:,j),5,50,Fs,2,1);
Norm_window=0.07; %%normilazation window
% %The nrmalisation window should be half of the maximum frequencywich the
data was filtered throuhg.
% %The data was filtered between 7 and 40 Hz,the maximum is 0.142seconds,
half of that is 0.07
weigthconstant = 1/((2*Norm_window)+1); %This formula is from Bensen's
paper and is used to determine the normalisation weight. %
allsums(j) = sum(abs(Data_Filtered(:,j))); allsums(j) =
sum(abs(filtered_data(:,j))); weight(j)=weigthconstant*allsums(j);
newnormdata(:,j) = filtered_data(:,j)/weight(j);%newnormdata is the
normalised data
% %5. Spectral whitening
whitendata(:,j) = whitening(newnormdata(:,j), Fs, 'freq', [5, 50]);end

% plot(detrended,'DisplayName','detrended')
% plot(filtered_data,'DisplayName','filtered_data')
% plot(newnormdata,'DisplayName','newnormdata') %
plot(whitendata,'DisplayName','whitendata') lagtime = 5000;
tenslice = 142; %Defining the number of columns for each station
nr_of_stations = 6;
alltenslice = nr_of_stations*tenslice; count
= 0;
mp=max(size(whitendata(:,1)));%Maximum where peaks appear for
i=1:alltenslice
for j=i+tenslice:tenslice:alltenslice count
= count+1;
cc_data(:,count) = cross_correlate(whitendata(:,i),
whitendata(:,j),Fs,Fs,lagtime,'norm'); end end

%%%Replace NaNs to zero cc_data(isnan(cc_data))=0;
% find(isnan(cc_data));%Find NaNs in matrix
% A_B(isnan(A_B))=0;%Code to replace NaNs with zero

```

```

%%NOTE:: NEED TO NORMALISE CROSS CORRELATIONS
% %Remove rows with zeros
% data = cc_data;
% % Remove zero rows
% data( all(~data,2), : ) = [];
% % Remove zero columns
% % data( :, all(~data,1) ) = [];

%Filter the cross correlation mm
=size(data,2);
  for r = 1:mm    fil_data(:,r) = bp_bu_co(data(:,r),5,50,Fs,2,1); end

% filtered_cc(:,j) = bp_bu_co(cc_data(:,j),5,50,Fs,2,1);

%The central value is the DC component of the data and does not contribute
%to the actual cross correlated signal
%Normalise the cross correlations

%The numerator is equal to the basic equation for cross correlation
%The denominator scales the result of the cross correlation by a factor
%that is related to the energy of each signal in cross correlation x and y
  data = fil_data;

% yy = cc_data;
% zz=size(yy,2);
% for i =1:zz
%     hilbert_cc(:,i)=envelope(yy(:,i), 2,'peak'); %envelope built in
matlab function to find the hilbert of the signal
% end
% data = hilbert_cc;%Perform hilbert transform on the crosscorrelations

%%Stacking nr_of_stations = 6; mm = size(data,2); tenslice =142; %Defining
the number of columns for each station
total_pairs = mm/tenslice;
%Split data according sensor pair
% A_B = data(:,1:142);
% A_C = data(:,143:284);
% A_D = data(:,285:426);
% B_C = data(:,427:568);
% B_D = data(:,569:710); %C_D = data(:,711:852);

%%For 6 stations

```

```

A_B = data(:,1:142); A_C = data(:,143:284); A_D = data(:,285:426); A_E =
data(:,427:568); A_F = data(:,569:710); B_C = data(:,711:852); B_D =
data(:,853:994); B_E = data(:,995:1136);
B_F = data(:,1137:1278); C_D = data(:,1279:1420); C_E = data(:,1421:1562);
C_F = data(:,1563:1704); D_E = data(:,1705:1846); D_F = data(:,1847:1988);
E_F = data(:,1988:2130);

```

```

%Stack cross correlations according sensor pairs

```

```

% pair1 = sum(A_B,2);
% pair2 = sum(A_C,2);
% pair3 = sum(A_D,2);
% pair4 = sum(B_C,2);
% pair5 = sum(B_D,2);
% pair6 = sum(C_D,2);

```

```

pair1 = sum(A_B,2);%5 pair2
= sum(A_C,2);%10 pair3 =sum(A_D,2);%15 pair4 = sum(A_E,2);%20 pair5 =
sum(A_F,2);%25 pair6 = sum(B_C,2);%5 pair7 = sum(B_D,2);%10 pair8 =
sum(B_E,2);%15 pair9 = sum(B_F,2);%20 pair10 = sum(C_D,2);%5 pair11 =
sum(C_E,2);%10 pair12 = sum(C_F,2);%15 pair13 = sum(D_E,2);%5 pair14 =
sum(D_F,2);%10 pair15 = sum(E_F,2);%5

```

```

% five = ([pair1,pair4,pair6]);% AT 5M
% ten = ([pair2,pair5]);%AT 10 M

```

```

five = ([pair1,pair6,pair10,pair13,pair15]);% AT 5M ten =
([pair2,pair7,pair11,pair14]);%AT 10 M fifteen = ([pair3,pair8,pair12]);
twenty = ([pair4,pair9]); twentifive =([pair5]);

```

```

% stack1 = sum(five,2);

```

```

% stack2 = sum(ten,2);

stack1 = sum(five,2); stack2 = sum(ten,2); stack3
= sum(fifteen,2); stack4 =sum(twenty,2); stack5 = sum(twentifive,2);

% BB = [stack1,stack2,pair3];

BB = [stack1,stack2,stack3,stack4,stack5];

% distance = d;d1(1:1197)
= 5; d2(1:1197)
= 10;
d3(1:1197) = 15;
% d4(1:397) = 5;
% d5(1:397) = 10;
% d6(1:397) = 5;

% d = [d1;d2;d3;d4;d5;d6]';d = [d1;d2;d3]';

%%
%Stack the cross correlations computed from 4 stations
%This is the code that works for stacking edited 20/11/2019 17:55 %figure(1)
% hold on
% d = distance/100;
% d = 5:5:15;
distance = d; sr =400; sz=size(BB); n_pairs=min(sz); l_xc=max(sz); count
= 0;

% source = [source_1,source_2,source_3,source_4];for i =1:3      plot((-
l_xc+1:2:l_xc)/(2*sr),BB(:,i)+distance(:,i))      holdon      for i
=2:3      plot((- l_xc+1:2:l_xc)/(2*sr),BB(:,i)+distance(:,i))      hold
on      for i =3:3      plot((-
l_xc+1:2:l_xc)/(2*sr),BB(:,i)+distance(:,i))      hold on %
      for i =4:6
%      plot((-l_xc+1:2:l_xc)/(2*sr),BB(:,i)+distance(:,i))

```

```

%           end         end         end         axis([-1.5
1.5 0 30])%x-axis begins at -1.5 end

%%
%This is the code that works for stacking edited 20/11/2019 17:55
%Stack the cross correlations computed from 6 stations
%Define the distance matrix d1(1:10001) = 5;
d2(1:10001) = 10;
d3(1:10001) = 15;
d4(1:10001) = 20; d5(1:10001) = 25; d = [d1;d2;d3;d4;d5]';
    distance = d; sr = 400; sz=size(BB); n_pairs=min(sz); l_xc=max(sz);
count
= 0; for
i =1:5
    plot((-l_xc+1:2:l_xc)/(2*sr),BB(:,i)+distance(:,i)) hold on
        for i =2:5
            plot((-
l_xc+1:2:l_xc)/(2*sr),BB(:,i)+distance(:,i)) hold on for i =3:5
                plot((- l_xc+1:2:l_xc)/(2*sr),BB(:,i)+distance(:,i)) hold on for i
=4:5
                    plot((- l_xc+1:2:l_xc)/(2*sr),BB(:,i)+distance(:,i))
end hold on for i =4:5
            plot((-l_xc+1:2:l_xc)/(2*sr),BB(:,i)+distance(:,i)) end end end
                axis([-1.5 1.5 0 30])%x-axis begins at -1.5 end

```

Pre-processing (Cradock dataset)

```

%%Preprocessing of stations and data from Cradock
%For the purpose of displaying the output from each preprocessing steps
%two stations are used: A1R and B5Y
%Vertical(pri.0) components are used to process Rayleigh waves
%Ascii files are arranged into days and each Day file consists of ,mat
files
%named IntensityArray. The .mat files contain data arranged into columns.
%Each column represents data from each station for each Day.

%Sampling frequency = 100 Hz;
%Sample points = 60x60x24x100 = 8640000 points
%Load IntensityArray files
% load('IntensityArrayDay1.mat')

```

```

%Extract column1(A1R) and column25(B5Y). NOTE: This is Day1's data.A1R =
IntensityArray(:,16);
B5Y = IntensityArray(:,28);
%Remove zeros
A1R( all(~A1R,2), : ) = [];
B5Y( all(~B5Y,2), : ) = [];

%Make the signals the same size

% clear vars -except A1R B5Y

%STEP 1%%
%Find the length of the signalx = B5Y; npts=length(x);
Fs=100;%samples per second = 400Hz,
%Calculate the time (x-axis) sampletime
= ((0:length(x)-1)/Fs);
%Transpose time from row vector to column vector sampletime
= sampletime';

% %1. Remove the mean meaningless = x - mean(x);
%
% %2. Remove the trend Data_Detrend = detrend(meanless);

% %3.apply a Tukey window to apply a 10% cosine Taper
% window=tukeywin(npts,0.1);
% tapered=trendless.*window;

%%Note Plotting to be done separatelyfigure
plot(sampletime,Data_Detrend) xlabel('Time (seconds)') ylabel('Amplitude
(Digital counts)')title('Raw Data: Station B5Y ') legend('B5Y') axis
tight

% legend('signal','filtered') figure % subplot(2,1,1)
spectrogram(x, [], [], [], Fs, 'yaxis');title('Unfiltered')
% subplot(2,1,2)
% spectrogram(x_filtered, [], [], [], Fs, 'yaxis');
% title('Filtered')

%Define matlab built-in bandpass filter
digfilt = designfilt('bandpassiir', 'StopbandFrequency1', 0.1,
'PassbandFrequency1', 0.2, 'PassbandFrequency2', 8,
'StopbandFrequency2', 9, 'StopbandAttenuation1', 60, 'PassbandRipple',

```

```

1, 'StopbandAttenuation2', 60, 'SampleRate', 400);
Fs=100;%Sampling frequency

x_filtered = filter(digfilt, Data_Detrend); figure plot(sampletime,
x_filtered);      xlabel('Time (seconds)')      ylabel('Amplitude')
title('Filtered Signal') legend('A1R') axis tight
spectrogram(x_filtered,[],[],[],Fs,'yaxis'); title('Filtered')

```

Computing Dispersion curves

%To create the phase velocity dispersion curves, you need to apply a phase-delay and stack method

You can easily write your own code if you understand the process well but it is easier to just apply narrow band-pass (or preferably Gauss) filters around each frequency and stack these ccfs for different test velocities

```

testVelocity = 100:300; freqVec=10:40;
sampling_rate =
??

```

% ccf is a NxM matrix with all your ccfs, with N the amount of pairs and M the amount of time samples

% dist is a Nx1 vector with the interstation distance for each of the N pairs

```

DispCurve=zeros(length(testVelocity), length(freqVec));

```

```

% Loop over frequencies
for fInd =
1:length(freqVec) freq=freqVec(fInd);

```

```

    [A,B]=butter(4,[freqVec-0.5, freqVec+0.5]./(sampling_rate/2)); % 1Hz
    windows, try others

```

```

    ccf_filtered=filtfilt(A,B,ccf); % Narrow band-pass filter applied to
    each ccf

```

```

% Loop over all velocities

```



```
for ind = 1:length(testVelocity)

    % find the sample in each ccf that corresponds to the testvelocity

    testVsamplePositiveLag=(M-1)/2 +round(dist./testVelocity(ind));

    testVsampleNegativeLag=(M-1)/2 -round(dist./testVelocity(ind));

    DispCurve(ind,fInd)=sum(ccf_filtered(:,testVsamplePositiveLag))
+ sum(ccf_filtered(:,testVsampleNegativeLag)); % sum positive andnegative
lag times

endend
```

Interpolation

```

%Values obtained from velocity profile

% Define x and y values
%
% x = [100:10:200];
%
% y = [0:10:20];
%
% % z = [180, 179, 180, 178, 177, 175, 170, 165, 160, 155, 150, 140,
132, 124, 122, 121, 120, 120, 120, 118, 119];
%
% [X,Y] = meshgrid(x,y);
%
% figure
% surf(X,Y)
%%
x = [1:16]'; %START FROM ZERO y
= [0:2:20]'; %START FROM ZERO
[X,Y] = meshgrid(x,y);

% X = X';
% Y = Y';
V = [c1, c2, c3, c4, c5, c6, c7, c8, c9, c10,c11, c12, c13, c14, c15,c16];
% figure
% surf(X,Y,Z)
% title('Original Sampling');
  figure imagesc(V)%WORKS BETTTER first plot option beforeinterpolation

%creating gridded data [Xq,Yq] = meshgrid(x,y);

% Interpolate at the query points.Vq = interp2(X,Y,V,Xq,Yq); figure
surf(Vq); figure
imagesc(Vq)%WORKS % 1st option still need to smooth. plot after
interpolation
  figure surf(Xq,Yq,Vq); %alternative option see if you canremove lines

%Smoothing the plot figuresurf(Xq,Yq,Vq); colormap summer % change to jet
shading interp

%%
%%Trial

```

```

%%Tested before solving interpolation method
% Vq = interp2(Z);%THIS WORKS
% VqQ = interp2(Z,5);
% figure
% pcolor(Z)
% shading interp;

% figure
% surf(Xq,Yq,Vq);

```

Appendix B

The measured group velocities were inverted using the method of Barmin *et al.* (2001) that is based on ray theory with a Gaussian-shaped lateral smoothing. The method also uses spherical geometry. The region of influence is defined by an arbitrary simple closed curve. Nodes are spaced by at approximately constant distances from one another, the interpolation is based on the three nearest neighbors. The surface waves are treated as rays that sample an infinite zone along the great circle linking source and receiver.

The resulting tomographic maps represent a local spatial average of the group velocity at each location on the map and summarize large volumes of surface wave dispersion information in a form that is both useful and easily transportable. It must be noted that the relationship between the observed seismic waveforms and the Earth model is not linear. This means the problem to use surface wave as a tool to image the structure of the crust is not linear. The inversion of surface waves is divided into two parts: (1) A step to estimate the 2-D dispersion maps and (2) a step where the dispersion maps are used to infer the Earth's structure. It is this second part that is referred to as surface-wave tomography (Barmin, *et al.*, 2001).

Surface wave tomography

Using ray theory, the forward problem for surface wave tomography consists of predicting a frequency-dependent travel time $t_{R/L}(\omega)$ for the Rayleigh waves (R) from a set of 2-D phase or group velocity maps, $c(r, \omega)$:

Where:

$$t_{R/L}(\omega) = \int_p c_{R/L}^{-1}(r, \omega) ds \quad 37$$

$r = [\theta, \phi]$ is the surface position vector, θ is the latitude and ϕ is the longitude, and p specifies the wave path. The dispersion maps are nonlinearly related to the seismic structure of the earth.

A group time t along a ray p can be computed as:

$$t = \int_p \frac{ds}{U(s)} \quad 38$$

Where s is the distance along the ray, p is the ray path and U is the group velocity. The problem is linear if the p is known. The wave path approximates that of a spherical symmetric model.

The forward problem

The travel time of surface waves is inversely related to velocity, a change in travel time δt relative to a reference velocity distribution U_0 is then expressed as:

$$\delta t = t - t_0 = \int_p \frac{ds}{U(s)} - \int_p \frac{ds}{U_0(s)} \quad 39$$

The Inverse problem

The main goal of this step is to estimate the vector function m using a set of observed travel-time residuals d relative to the reference model. The model is defined as:

$$m = \frac{U_0 - U}{U} \quad 40$$

Where, δt becomes a linear function of m .

Surface wave tomography is the set of observed travel times $T_{obs}(\omega)$ for many different paths p to infer group or phase velocity map, $c(r)$ at frequency ω . The observed travel time is a sum of the real travel time and an observational error:

$$T_{obs}(\omega) = t(\omega) + \varepsilon(\omega) \quad 41$$

Where ε is an observational error for a given path.

The datum can be expressed as:

$$d_i = \delta T_{obs(i)} = \delta T_{obs(i)} - t_0 \quad 42$$

Where t_0 is the travel time for a reference model U_0 . This results in the datum being defined as:

$$d_i = G_i(m) + \varepsilon_i \quad 43$$

The next step is to solve the inverse problem by estimating the model (m). This can be achieved by adjusting the model parameters that are subject to regularization constraints to fit the observed data. Regularization is also called damping and refers to the application of constraints on the estimated model during inversion. Damping is the focal point of the inversion method because tomographic maps that result from the inversion of Surface waves does not depend continuously on its parameters. The regularization involves a penalty function that consists of a spatial smoothing function and spatial constraint on the amplitude of the estimated model from a reference state (Barmin, et al., 2001)

The model m is estimated by minimizing the following penalty function (S):

$$S(m) = (G(m) - d)^T C_d^{-1} (G(m) - d) + \alpha \|F(m)\|^2 + \beta \|H(m)\|^2 \quad 44$$

Where m is the model, d is the data vector and G is the forward operator. C_d is the data covariance matrix which is assumed to be diagonal. The first term of the penalty function in Equation 56 represents the data misfit of the model prediction from the data, the second term is the spatial smoothing condition and the third term is the damping constraint that penalizes the weighted norm of the model. The process of minimizing the first term of the penalty function reveals a model that fits the data, and the last two terms penalize models which diverge too far away from the reference model. A norm of an arbitrary function $f(r)$ can be defined as:

$$\|f(r)\|^2 = \int_s f^2(r) dr \quad 45$$

The spatial smoothing term F involves a correlation length σ

$$F(m)(x) = m(x) - \int_s K(x, x') m(x') dx' \quad 46$$

Where, K is a smoothing kernel and is defined as:

$$K(x, x') = A \exp\left(-\frac{|x - x'|^2}{2\sigma^2}\right) \quad 47$$

The third term constraints the amplitude of the perturbations depending on local path density;

$$H(m)(x) = \exp(-\lambda\rho(x)) \quad 48$$

Where $\rho(x)$ is the path density around x and λ is constant. The parameters $\alpha, \beta, \sigma, \lambda$ are user-controlled parameters and are determined through systematic exploration of the misfit evolution with different realistic values. The velocity distribution is discretized with a Cartesian grid where each cell has a constant velocity. The next step of the tomographic inversion is discretization of the penalty function. The discretization process involves the formation of a discrete grid and the evaluation of the model on the grid. The goal is to generate a grid with nodes that are constantly spaced on a sphere so that the nearest neighbors can be identified.

The Inversion matrix

To construct the inversion matrix, the penalty function needs to be integrated such that it can be rewritten in matrix form:

Let N be the number of ray paths, $i = 1 \dots N$, let M be the number of cells in the model, $j, k = 1 \dots M$, then, $m(x) = m$ is an M -long vector containing the slowness for every cell and Equation 44 can be discretized and rewritten in matrix form:

$$S(m) = (Gm - d)^T C_d^{-1} (Gm - d) + m^T Qm \quad 49$$

Where, the smoothness and model norm constraints are included in the second term and where:

$$Q = \alpha F^T F + \beta H^T H \quad 50$$

The matrix G is an $N \times M$ matrix containing the length of every path in every cell of the model and is define as:

$$G_{ij} = \frac{l_{ij}}{U_{0j}}$$

Where, l_{ij} is the length of the i th path in the j th cell and U_{0j} is the initial group velocity in the j th cell. The matrices F and H are $M \times M$ matrices and their components are defined by:

$$F_{jk} = \delta_{jk} \exp - \frac{K_{jk}}{\sum_k K_{jk}} \quad 52$$

where δ_{jk} is the Kronecker function and $K_{jk} = A \exp(-d_{jk}^2/2\sigma^2)$ with d_{jk} being the distance between the j th and k th cells:

$$H_{jk} = (\delta_{jk} \exp(-\lambda\rho_j)) \quad 53$$

where ρ_j is the number of paths crossing the j th cell. With these definitions, the minimum of the function S is found at m_{min} with

$$m_{min} = (G^+ C_d^{-1} d) \quad 54$$

Where G^+ is the inversion operator and is defined as:

$$G^+ = (G^T C_d^{-1} G + Q)^{-1} G^T \quad 55$$

The resolution matrix can then be defined as:

$$Res = G^T C_d^{-1} G \quad 56$$