IDEA League

Master of Science in Applied Geophysics

Research Thesis

Quality control and clock error correction of passive seismic data from ocean bottom seismometers in the Red Sea

Guus Hoogewerf

August 5, 2022

Quality control and clock error correction of passive seismic data from ocean bottom seismometers in the Red Sea

Master of Science Thesis

for the degree of Master of Science in Applied Geophysics at
Delft University of Technology
ETH Zürich
RWTH Aachen University
by

Guus Hoogewerf

August 5, 2022

Department of Geoscience & Engineering · Delft University of Technology
Department of Earth Sciences · ETH Zürich
Faculty of Georesources and Material Engineering · RWTH Aachen University

Delft University of Technology

Printed in The Netherlands

IDEA LEAGUE
JOINT MASTER'S IN APPLIED GEOPHYSICS

Delft University of Technology, The Netherlands
ETH Zürich, Switzerland
RWTH Aachen, Germany

Dated: *August 5, 2022*

Supervisor(s):

C. Weemstra

L. Parisi

D.F. Naranjo

Committee Members:

C. Weemstra

L. Parisi

D.F. Naranjo

F. Wellmann

Abstract

Three-component seismometers deployed on the sea bed have been used for many years. These so-called ocean bottom seismometers (OBS) are used to register both compressional and shear waves. The recorded seismic information is used to enhance our understanding of areas of interest located offshore, like hydrocarbon reservoirs, mid oceanic ridges, and continental margins. A lot of progress has been made on OBS applications. However, there are still some distinctive problems mainly related to clock drift, orientation, and coupling. We use and test *OCloC*, a new method which seeks to improve the ability to detect and correct clock errors of OBSs. The method was successfully developed using an extensive dataset from a seismic network, which included many land- and ocean bottom-seismometers. One of the objectives is to understand under which conditions OCloC can be used when the dataset is relatively small. To do this, data from a network deployed in the Red Sea coastal area of the Kingdom of Saudi Arabia is used. This network includes two ocean bottom seismometers which are complemented by several land-stations. We collected all used data during five campaigns, both offshore and onshore. To analyse the data quality we designed and applied an extensive data quality control, where after the data is prepared for Ocloc through pre-processing. Both workflows are presented in a roadmap. This roadmap is made applicable for ocean bottom seismometers by including an orientation estimation, for which we used a technique based on Rayleigh wave polarization. Our quality control shows that the data from both ocean bottom seismometers show significant attenuation, primarily for short periods, which is possibly caused by a poor coupling to the loose sediments on the sea bed. The noise source illumination turns out to be highly anisotropic, with a high microseism noise source concentration in the Red Sea. This results in a poor signal to noise ratio for the interferometric surface wave responses. As a consequence, we find that the limited station deployment time and data quality interferes with the ability to obtain an accurate clock error. Nevertheless, the technique to estimate the OBS orientation turns out to be successful.

Acknowledgements

First of all I would like to express my deepest gratitude towards my supervisors, Kees Weemstra, Laura Parisi and David Naranjo, who guided me throughout the process and provided valuable insights into both the theoretical and practical aspects of my project.

I would like to thank the CES group at KAUST with PI Martin Mai for having me as a visiting student in Saudi Arabia, and for learning me a lot about all the different projects in the field of computational earthquake seismology.

I am also grateful to everyone that participated in the data collection fieldwork in Saudi Arabia, Adriano, Elisabeth, Laura, Marlin, Nicolas and Remi, without whom this project would not be possible.

I would like to thank the entire crew of the MV Handin Tide, both from Fugro and Tidewater, who made my first offshore campaign a truly amazing experience that I will never forget. There was so much to learn and everyone was very eager to explain me about their role on the ship. A special thanks to Anthony Gervasio, the party chief with whom I shared an office container on the ship, and who explained me everything about the ongoing operations. Also a special thanks to Tony Farn for making this experience possible.

I would like to thank all the amazing people I met during my stay in Saudi Arabia, all of you made my stay an unforgettable experience.

A special thanks to the entire Applied Geophysics 2020-2022 group, with whom I lived in Delft, Zurich and Aachen. It was great to have such a close group to study and enjoy our new living places together

Last but certainly not least I would like to thank my family for their unconditional support throughout my studies in 4 different countries.

Delft University of Technology
King Abdullah University of Science and Technology — Guus Hoogewerf

August 5, 2022

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The vast majority of our planet is covered by oceans. Exploration of the ocean's subsurface was challenging in the early 1900s, which resulted in much of the ocean floor being unexplored. In the early 1930s the first seismic refraction method was developed, which was soon tested offshore to obtain information on the ocean bottom. As part of these trials a first prototype of a so called ocean bottom seismometer (OBS) was made in 1938 [Schmidt-Aursch and Crawford, 2015]. This prototype used a rubber balloon filled with gasoline to make the device float above the seafloor. An automatic oscillograph was mounted below the balloon and connected to an external geophone, which was kept at the ocean bottom by a piece of iron [Ewing and Vine, 1938, Schmidt-Aursch and Crawford, 2015]. In these early days the use of an ocean bottom hydrophones was more established, this device was relatively simple and did not require coupling to the ocean bottom. The ocean bottom hydrophone is sufficient to monitor pressure waves, but shear waves information is unattainable [Bialas and Flueh, 1999]. This shear wave information became increasingly desired in the late 1950s, when seismic monitoring of nuclear explosions became suddenly important [Sutton, 1986]. Shear wave information turned out to be a great way to distinguish man made explosions from earthquakes, as earthquakes tend to produce a higher fraction of shear waves [Kornei, 2019]. Since then the use of OBS has become more and more common, as many areas of interest are located offshore, including gas and oil reservoirs, mid-oceanic ridges and continental margins [Schmidt-Aursch and Crawford, 2015].

Although a lot of progress has been made on OBS technology, challenges still exist. One of these challenges is the possibility of a timing error in the internal clock of the OBS. Normally a seismic station is fitted with a GPS module to synchronise the internal clock with a reference time. Since an OBS is placed on the ocean bottom it is often not possible to facilitate a satellite connection through the GPS module. This satellite connection is crucial to synchronize the internal clock with the exact time given by the satellite's internal atomic clock. Without this synchronization the OBS clock might drift at unknown rates [Hannemann et al., 2014]. Usually a linear correction is applied by using the known synchronized times before and after deployment of the OBS (e.g. [Geissler et al., 2010]). However, this correction is not possible if the battery dies before the OBS is recovered. Additionally, a timing error can occur during deployment caused by rapid changes in the environment around the OBS, mainly related to temperature and pressure changes [Weemstra et al., 2021]. To deal with these shortcomings other methods have been introduced with the aim to use the data directly to account for potential clock errors. Many of these methods utilize the temporal stability of interferometric responses. These responses are obtained by cross-correlating the data from two different seismic stations over a longer time period [Sens-Schönfelder, 2008, Loviknes et al., 2020, Hannemann et al., 2014, Weemstra et al., 2021].

A new method aiming to solve all these shortcomings is OCloC (OBS Clock Correction), a python package for the detection and correction of clock errors in OBS data [Naranjo et al., 2021]. The OCloC package is an implementation and extension of the method introduced by [Weemstra et al., 2021]. In the OCloC package two approaches are combined for clock error detection; temporal stability of cross-correlations and time symmetry of the retrieved interferometric responses. The OCloC package was tested and proven successful with data from an extensive seismic network deployed in southwest Iceland, which included 30 land stations and 24 OBS [Naranjo et al., 2021, Weemstra et al., 2021]. For many OBS systems that require a clock correction, such an extensive seismic network is not available. In this research we seek to detect and correct the timing and orientation errors for a less extensive dataset and thereby test if the OCloC package is suitable for a non-ideal case with significantly less data. The seismic network we used only includes 8 land-stations and 2 OBSs with a limited deployment time. Due to technical difficulties both OBSs were not deployed simultaneously. Because of this we could not test the package with only OBS data.

One of the inputs for the OCloC package are interferometric responses between seismic stations. To obtain these the seismic data has to be pre-processed. In this research we present a roadmap for seismic station pre-processing, with the aim to retrieve interferometric responses. Additionally, this roadmap is made applicable to OBSs by adding extra steps for orientation and position correction. The roadmap we present also includes a comprehensive quality control of the raw data. An important remark to make is that many steps in pre-processing can be tailored to the datasets specifications, this has also been done with this specific case. This means that the roadmap provides a clear direction, but will not always provide the best result for each dataset.

Modern seismic stations usually include a three-component seismometer, which typically resolve three axes of ground motion into one vertical and two orthogonal horizontal directions [Stachnik et al., 2012]. During installation one of the horizontal components has to be aligned with the true north to effectively use the horizontal components. This alignment can be very challenging in the case of OBSs [Stachnik et al., 2012]. Often the orientation has to be determined after the deployment. To do this airgun shots have been used in the past [Duennebier et al., 1987], however this is not an error- and cost free procedure [Stachnik et al., 2012]. In this research we obtain the orientation of the OBSs by using a method based on the seismic data. This method inspired by [Stachnik et al., 2012] utilizes the elliptical particle motion of Rayleigh waves in combination with teleseismic events.

The seismic network is located in and around the Red Sea. The Kingdom of Saudi Arabia's Red Sea coastline stretches about 1760 kilometers. This coastline is undergoing rapid developments to accommodate the growing demand for luxurious tourism destinations within the Kingdom. Saudi Arabia was closed to most tourist until 2019. Now it is opening up and wants to increase its tourist sector to 10% of the GDP. To achieve this the kingdom announced several billion dollar projects, including development of parts of their Red Sea coast into luxurious tourist resorts. Most famous are the Red Sea Project and NEOM, located in the Tubuk province.

The Red Sea is an elongated depression with a length of roughly 2000 kilometer. In the north the sea has a width of only 180 kilometers. This widens southward to 360 kilometers, before it narrows back down to only 28 kilometers in the south (strait of Bab el-Mandeb). It is thought to be the result of the Arabian Peninsula drifting away from the African continent during the early Pliocene, the outline of western Arabia and northeast Africa reinforce this theory, however not everything is known about the early formation. [Coleman, 1974]. Currently the Red Sea rift is spreading with an average rate of 13 millimeters per year [Chu and Gordon, 1998]. Since the rifting zone is relatively new it is studied to improve our understanding of continental breakup.

The bathimetry of the Red Sea as mapped by [Laughton, 1970] shows a clear trough in the middle of the red sea extending northward from the Zebayir islands to the southern tip of the Sinai Peninsula in Egypt. The main through is cut by an axial trough, containing high-temperature brine pools and hydro thermal sediments [Coleman, 1974]. The bottom of the red sea is relatively shallow and flat, including volcanic flows covering parts of the plains. Landward from this coastal plain the coast of Saudi Arabia can be best described in segments. South of Jeddah magnificent erosional escarpment can be found, following an extremely irregular erosional pattern [Coleman, 1974]. These scarps extend inland as far as 120 kilometers and can reach altitudes over 2000 meters. North of Jeddah no escarpment takes place and the coastline has a more gradual slope.

Strong positive gravity anomalies and high seismic velocities within the axial trough show that the trough is related to volcanic activity. Present-day seismic activity is mainly located along the same axial trough and strike-slip faults in the northeast. From surveys it is known that new oceanic crust is forming at this rifting structure, it is not exactly known when this started. In the southern section of the Red sea large areas of this new oceanic crust are exposed, in the north most of it is covered by salt which makes it difficult to obtain samples [Augustin et al., 2021]. The difference is mainly due to different extensional styles. The south is characterized by a localized extensional style [Almalki et al., 2015] and the north by a diffuse extensional style [Roobol and Stewart, 2019]. The northern diffuse extension extends beyond the Red Sea itself, and is seen in the form of dikes, grabens and volcanism outside the rift zone (Figure 1-1)[Aldaajani and Furlong, 2022].

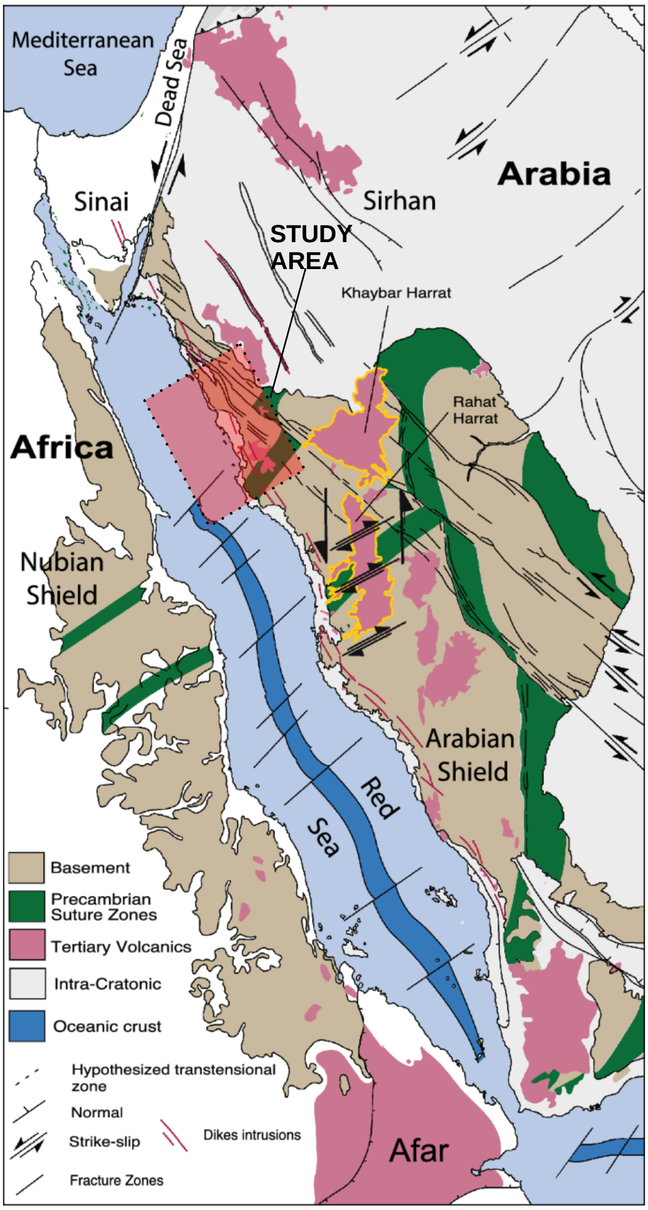


Figure 1-1: Current tectonic setting of the Arabian margin, the map shows the difference between extensional style between the North and the South. The study area is market by the red square in the northern part of the Red Sea. Adapted from [Aldaajani and Furlong, 2022]

Chapter 2

Seismic network

The seismic network used for this study is located in the Red Sea coastal area of the Kingdom of Saudi Arabia. It consists of 2 ocean bottom seismometers (OBSs) and 8 landstations. Their specifications can be seen in Table 2-1. It is important to include the sensor and data logger type, as they are used to retrieve the instrument response. Due to the strong difference in deployment environment the stations installed on land are split between land-stations and island-stations.

Network	Instrument	Data logger	Sample freq.	Stations	Needs correction
RS-OBS	Trillium Compact, 120s, AT1	Pegasus OBS PGSM-S	100 Hz	001, 002	YES
ZF-ISLAND	Trillium Compact, 120s, 754	V/m/s-Centaur, 40 vpp	125 Hz	BREEM, QUMAN	NO
ZF-LAND	Trillium Compact, 120s, 754	V/m/s-Centaur, 40 vpp	125 Hz	LAVA, KHUF, DEEP, SOUTH, WEST, EAST	NO

Table 2-1: Overview of stations used for this research with type of instrument and data logger, as well as the used sample frequency and if the station requires a clock error correction. Due to the difference in deployment environment the stations installed on land are split between land-stations and island-stations.

The used seismic stations were installed with different purposes. The two OBSs 001 and 002 are installed as part of the Zafran project together with 10 other OBSs. They are placed to study locations and mechanisms of seismic events in the so called Zabargad fracture zone, one of the largest rift offsets in the northern Red Sea [Marshak et al., 1992]. The exact location of OBS 001 is at the edge of the Mabahiss mons, a volcano with an average height of 450 meters. Island-stations BREEM and QUMAN, and land-stations LAVA and KHUF are also installed as part of the Zafran project to study the Zabargad fracture zone by augmenting the offshore OBS network. The stations DEEP, SOUTH, WEST and EAST are placed on Harrat Lunayyir, a volcanic field in the west of Saudi Arabia. This location saw earthquake swarms in the past, mainly caused by dike intrusions. The stations are placed to map any occurring seismic activity [Trippanera et al., 2019].

The three-component sensors were installed in holes, drilled in the bedrock formation. This was done to facilitate shielding from wind and temperature and create a good coupling to the bedrock which enhances the signal transfer. The power source needs to be constant and reliable. Since at all land station locations sun is abundant, solar panels in combination with small battery packs are used to power the seismic stations. A photo of a land-station is shown in Figure 2-1 (a), the grey box contains the battery and data logger. The data logger is connected to the sensor and the GPS antenna. The stations BREEM and QUMAN are installed on two small uninhabited islands also called Breem and Quman. These islands are part of the Red Sea development project (See Introduction). Because the islands are uninhabited there is little interference from human activity and a lower risk of the stations being stolen or damaged. A downside is the increased difficulty to reach the locations and permission for a visit is required from the Red Sea development project.

(a) **(b)**

Figure 2-1: (a): Land seismic station (the sensor cannot be seen since it is buried underground). **(b)**: Ocean Bottom Lander with several instruments, OBS is circled in red

The two OBSs are placed on an Ocean Bottom Lander designed by Fugro, which includes many sensors like flow and salinity meters. A photo of an ocean bottom lander can be seen in Figure 2-1 (b), the three-component sensor is circled in red. The lander has a long lasting battery, providing power to all sensors for up to one year. After one year the lander has to be picked up from the seafloor to retrieve the data. Since the seafloor is mainly mud, the OBS can not always be installed on the bedrock, which can lead to an attenuated signal [Bialas and Flueh, 1999].

The deployment time of each station is shown by the 'real situation' in Figure 2-2. To remove the clock error from the OBSs we require cross-correlations between land-stations and the OBSs (see section 4-3). For this we can only use the period when both stations were running, which limits the length of the potential cross-correlation period significantly. Additionally, the OBSs did not have any overlap in deployment time, which means that both OBSs have a different set of simultaneously deployed stations. This significantly restricts the cross-correlation options for each OBS and makes it impossible to test the code with OBS data only. The obtained data is only a part of the planned data as Figure 2-2 shows. There are several reasons why some data was not retrieved, some stations were not installed on the planned time, other stations could not be recovered in time to be used for this research. Some seismic stations also broke down during deployment, the reason for this is unknown, but it is thought to be a result of the intense heat the seismic sensors are exposed to.

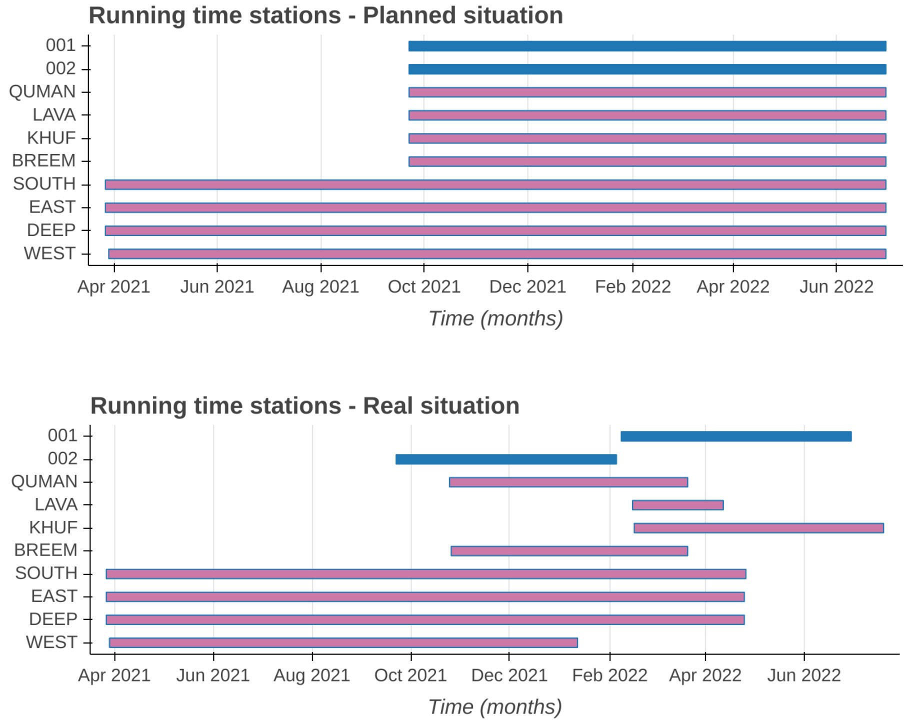


Figure 2-2: Running time of seismic stations, the top figure shows the planned situation. The bottom figure shows the actual retrieved data.

2-1 Data retrieval campaigns

We retrieved the seismic data during several campaigns, both on land and offshore, an overview of all campaigns can be seen in table 2-2. This table only includes the retrieval of data, not the installation of seismic stations. Due to long distances in between stations these campaigns sometimes covered multiple days. To reach the land-stations a proper off-road vehicle was a necessity, since it required driving long off-road sections through desert areas. The boat needed for reaching the Island stations was provided by the Red Sea project and brought us as close to the islands as possible, the last part was done on foot. Since the time for this project was very limited some data had to be retrieved during the summer months, when temperatures in Saudi Arabia often surpass 40 degrees celsius.

We retrieved the OBS data with a ship from Fugro: the MV Handin Tide. Since this ship is primarily used for R&D it is well equipped. Part of the equipment is an ROV (remotely operated vehicle) which can operate up to 4500 meters depth. The landers were slowly lowered and lifted using a heavy-lift line under guidance of the ROV, despite the lander design allowing them to be dropped off the ship. During the second offshore campaign a cable was placed in between the two landers and a new buoy system was coupled to the north lander. In the future this buoy should surface automatically once every 24 hours to send the data via a satellite connection. In this way (almost) live data from the landers becomes possible. This automatic buoy system can theoretically also be used to synchronise the internal clock of the OBS.

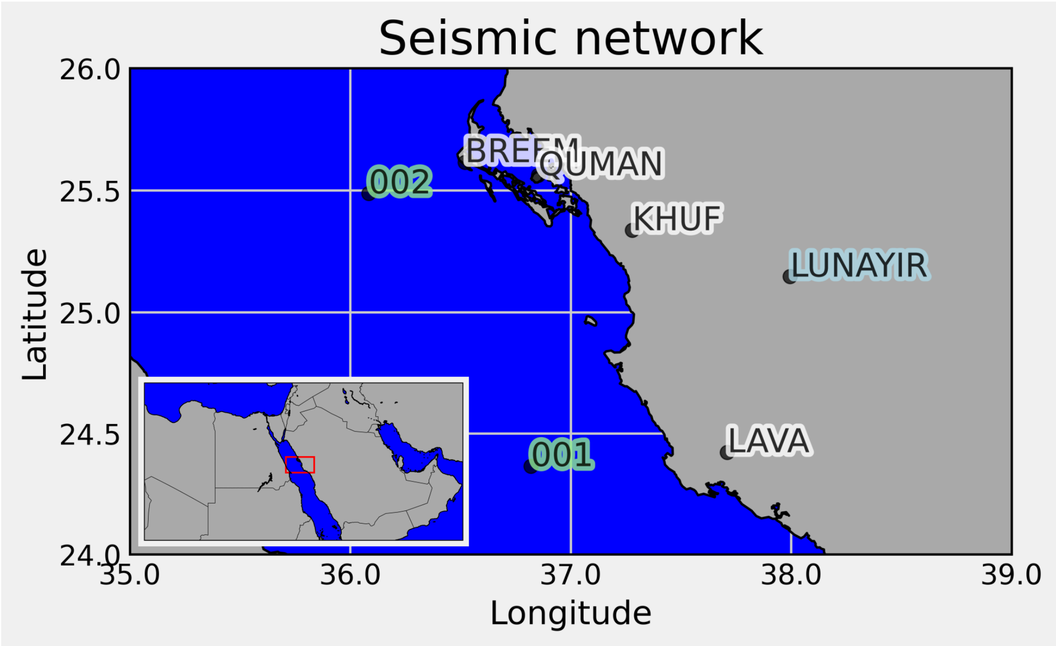


Figure 2-3: Map of seismic network, OBS 001 and 002 are shown in green. the LUNAYYIR group is shown in blue and includes stations [SOUTH, DEEP, EAST, WEST]

	Date	Retrieved data	Requirements
1	1/02/2022 14/02/2022	001, 002	Fugro MV Handin Tide
2	20/03/2022 21/03/2022	BREEM, QUMAN, LAVA, KHUF	Off-road vehicle Small boat
3	25/04/2022 26/04/2022	DEEP, SOUTH, WEST, EAST	2x Off-road vehicle
4	26/06/2022 11/07/2022	001, 002	Fugro MV Handin Tide
5	15/07/2022	LAVA, KHUF	Off-road vehicle

Table 2-2: Fieldwork campaigns to retrieve data used for this research. Only data retrieval is included, no station installation

Chapter 3

Theoretical background

3-1 Seismology

The OCloC package uses continuous recordings of ambient seismic noise to detect and correct a potential clock drift in ocean bottom seismometers (OBSs). To do this the package utilizes so called seismic interferometry to obtain an interferometric response (section 3-2). We focus on the surface wave part of this response. Surface waves are a type of seismic wave which propagates along the interface between differing media. To provide a general introduction to the different types of seismic waves the concept of seismology is briefly explained in this chapter. Seismology is the study of seismic waves that propagate through a medium. Seismic waves can be caused by both natural and artificial processes. Natural phenomena causing seismic waves can be volcanic, tectonic, glacial, fluvial, oceanic or atmospheric. Examples of artificial processes causing seismic waves are explosives or traffic. The different seismic wave types are observed to learn more about the medium they propagated through, or reflected upon. Earthquakes are a sudden release of energy in the Earth's lithosphere creating seismic waves, in this research different types of earthquakes are used, primarily for the data quality control.

3-1-1 Wave propagation

Seismic wave propagation is a ground motion governed by the seismic wave equation, which links force and displacement in three dimensions. The equation can be described in many forms, a simplified form for homogeneous media is shown here:

$$\rho \ddot{\mathbf{u}} = (\lambda + 2\mu)\nabla\nabla \cdot \mathbf{u} - \mu \nabla x \nabla x \mu \tag{3-1}$$

Where ρ is density, $\mathbf{u}$ is the displacement vector and λ, μ are the Lamé parameters, two material dependent quantities that arise in strain-stress relationships [Shearer, 2019]. This equation is an approximation where gravity and velocity gradient terms are neglected and the Earth is assumed to be isotropic. The general wave equation can be specified for either P-waves or S-waves by taking the divergence and curl, respectively. The divergence of equation 3-1 leads to the P-wave solution:

$$\nabla^2(\nabla \cdot \mathbf{u}) - \frac{1}{\alpha^2}\frac{\delta^2(\nabla \cdot \mathbf{u})}{\delta t^2} = 0 \tag{3-2}$$

With a P-wave velocity described as α by:

$$\alpha^2 = \frac{\lambda + 2\mu}{\rho} \tag{3-3}$$

The curl of equation 3-1 leads to the S-wave solution:

$$\nabla^2(\nabla \times \mathbf{u}) - \frac{1}{\beta^2}\frac{\delta^2(\nabla X \mathbf{u})}{\delta t^2} = 0 \tag{3-4}$$

With a S-wave velocity describe as β by:

$$\beta^2 = \frac{\mu}{\rho} \tag{3-5}$$

For a more elaborated explanation of the wave equation and it's different forms we refer to [Shearer, 2019].

P-waves induce compressive and tensile strains on the medium in a radial direction away from the source. S-waves induce shear strains on the medium perpendicular to the radial direction. When these so called body waves interact with a free surface, surface waves are generated. P-waves arriving at the surface induce Rayleigh waves with both a radial and lateral movement in the vertical plane. S-waves induce Love waves, characterized by movement lateral movement in the horizontal plane as shown in Figure 3-1 [D.G. Honegger, 2013].

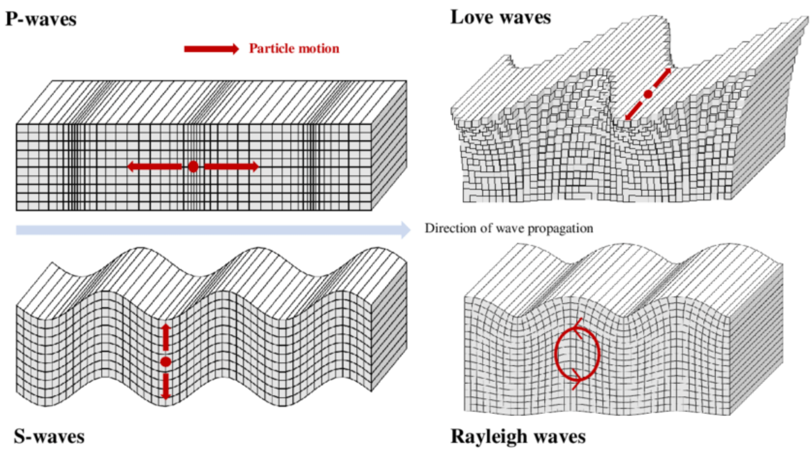


Figure 3-1: Body and surface waves, the P- and S-waves are body waves, Love and Rayleigh are surface waves (adopted from [Hohensinn, 2019])

The group-velocity of seismic waves depend primarily on the density and elasticity of the medium, as well as the type of wave. The velocity can be calculated for different locations with a tomographic inversion as done by [Mokhtar et al., 2001]. P-waves (or primary waves) travel faster than any other wave, which makes them to arrive first at the seismic station. S-waves (or secondary waves) travel considerably slower, and arrive after the faster-moving P-waves. Velocities of S-waves are typically around 60% of that of P-waves, depending on conditions. S-waves cannot propagate through any liquid medium, in contrast to P-waves. Surface waves are slightly slower than body waves, with the Love wave traveling at roughly 90% of the velocity of the S-wave, and the Rayleigh wave traveling slightly slower than the Love wave.

3-1-2 Earthquakes

In this research both local and teleseismic events are used to analyse data from seismic stations and to determine the orientation of the OBSs. For this earthquake we use the Rayleigh wave polarization, as explained in section 4-1-3. To do this we use both 'local' and 'teleseismic' events. An earthquake is classified as local when the epicenter is within 600 kilometers, it is classified as teleseismic when the epicenter distance is greater than 1000 kilometers. By increasing the distance to the earthquake the difference in velocity between the different wave types can be utilized to analyse the wave types separately. An earthquake is a sudden release of energy by movement of a part of the Earth's crust. The caused shaking can result in different types of seismic waves. The magnitude of an earthquake is a measure of released energy based on the amplitude of the recorded seismic waves. Two major features in the lithosphere are responsible for the generation of earthquakes: (I) fracturing of the rocks and (II) movements of distinctive volumes driven by tectonic forces [Keilis-Borok, 2002]. The majority of all earthquakes occur in fault networks, which are located at the border of separate rock volumes, or 'blocks'. Tectonic energy is stored in the whole volume of the 'block', but released through the relatively thin fault network.

The data used for this research is obtained in the Saudi Arabian Red Sea coastal area. The Western region of Saudi Arabia is considered to be a moderately active seismic zone [El-Quliti et al., 2016]. Many earthquakes of the swarm type are a result of the volcanic activity in the area, but earthquakes also occur along the Red Sea rifting system or at regional land faults [El-Isa and Shanti, 1989, El-Quliti et al., 2016]. The transition area from oceanic to continental crust on the bottom of the Red Sea is reported to have a high level of earthquakes with a relatively low magnitude [Brillinger, 1993]. One of the most active seismic regions in the area is the Gulf of Aqaba in the Northwest of Saudi Arabia, the activity is thought to have a relation to pull-apart tectonics [El-Isa and Shanti, 1989]. Also the Harrat Lunayyir volcanic field, where four stations used for this project are installed (DEEP, WEST, EAST and SOUTH) has experienced several occurrences of seismic swarms, mainly related to volcanic intrusions [Trippanera et al., 2019].

3-2 Seismic Interferometry

Seismic interferometry is a technique that has been around for several decades. The first forms were pioneered by ([Aki, 1957, Claerbout, 1968]. The OCloC package used in this research is built around this technique. Through years of development, seismic interferometry has been used in a wide variety of applications like subsurface characterization [Draganov et al., 2007] and reservoir monitoring [Sánchez-Pastor et al., 2019]. Seismic interferometry is a methodology that allows the retrieval of seismic responses by cross-correlating the waveforms recorded at two or more receiver locations [Wapenaar and Fokkema, 2006]. In the context of passive seismic interferometry, the retrieved interferometric responses might correspond to noise sources (ambient or anthropogenic) or earthquake activity. These interferometric responses are proportional to the Green's function of the medium between the receivers if the following conditions are fulfilled: (I) there exists a uniform noise source illumination pattern, (II) the noise sources are uncorrelated, (III) the medium is lossless and (IV) the noise sources have coinciding amplitude spectra [Wapenaar and Fokkema, 2006]. One critical feature of the recovered interferometric responses, under the aforementioned conditions, is that the responses are symmetrical signals, i.e. they are even signals. For correcting timing errors, based on OCloC's package and methodology, the time symmetry of the cross-correlations is exploited to identify potential clock errors related to the receivers [Naranjo et al., 2021]. For this technique, no prior knowledge of the subsurface medium parameters or of the sources' positions is required [Wapenaar et al., 2010, Wapenaar and Fokkema, 2006, Claerbout, 1968]. In this chapter, the concept of seismic interferometry is briefly explained, with a focus on direct-wave interferometry. The ambient noise that is used to obtain the interferometric responses is described and grouped. Finally, noise source illumination patterns, and their influence on the interferometric responses, are explained.

3-2-1 Direct wave interferometry

Time averaging of a cross-correlation between two ambient seismic noise fields from different receivers often provides a new seismic response [Weemstra et al., 2021]. The cross-correlation essentially turns one receiver into a virtual source, whose response is retrieved at the other receiver. This is mainly the result of the common raypath from any noise sources to one of the receivers cancelling out due to the applied cross-correlation, because of this the actual source location is of no importance [Wapenaar and Fokkema, 2006]. In case the used ambient seismic noise fields provide an isotropic illumination (section 3-2-3) of the receiver pair, two virtual source responses will be retrieved. One will be at positive and one at negative time, both should occur at time lags of equal magnitude but opposite sign [Hable et al., 2018]. Under special conditions that we mentioned before these virtual source responses can be related to the Green's function of the medium [Wapenaar and Fokkema, 2006]. The observed ambient seismic noise fields include both surface- and body-wave energy. This means that both waves are represented by the obtained Green's function. In this research we focus on the surface-wave part of the Green's function. We can do this by designing a filter with a pass band in the so called 'microseism' period band (section 3-2-2) In this period band the amplitude of the body-waves is in general considerably lower than the amplitude of the surface-waves [Weemstra et al., 2021, Wapenaar et al., 2010]. This is due to the cylindrical spreading of surface-waves, whereas body-waves lose their energy faster by spreading in three dimensions.

Most of the microseism noise is generated by interaction in the ocean, which is often relatively far away in terms of wavelenghts [Weemstra et al., 2021].

3-2-2 Ambient noise

In passive seismology we use ambient seismic noise instead of conventional active seismic sources to retrieve the virtual surface waves. Ambient noise can be defined as any diffuse wave generated by any natural earth vibrations (e.g. microseisms), cultural sources, instrumental glitches, or a combination of these [Peterson, 1993]. Because the nature of these events is often random, ambient noise is also mostly random. We can group ambient seismic noise based on their period. It is useful to make an assessment of the temporal and spatial distribution of possible ambient noise sources contributing to the wave field when a seismometer is installed [Yang and Ritzwoller, 2008]. An example can be a nearby highway, when periods in the cultural band are studied this can significantly influence the recordings. The temporal noise source distribution can also be described based on the noise spectrum with a PPSD and/or spectrogram. These methods were used to analyse the seismic noise in this research, they are described in more detail in part PPSD and Spectrogram.

Several publications have been made on Earth noise models (e.g. [Brune and Oliver, 1959, Frantti et al., 1962, Peterson, 1993]). The latest model by Peterson, which is nowadays considered the standard, is obtained using a worldwide network of broadband stations. With these the earthquakes could be filtered out, and a low-noise model (NLNM) and high noise model (NHNM) was constructed [Peterson, 1993]. The constructed noise model provides a lot of insights in the major components of the ambient noise spectrum. Short-term (0.1-1s) ambient noise power levels are mainly dominated by human-generated seismic energy; radiation from the electrical grid, cars, trains, and industry within a few kilometers of the recording station. Microseisms, which can be many orders of magnitude more powerful than the rest of the seismic spectrum, dominate the intermediate periods (1-30 s), they are mainly caused by pressure fluctuations on the ocean. Long-period (30-500 s) signals are typically formed by ocean waves caused by storm-forced, shoreward-directed winds, also known as "Earth Hum" [nak, 2019]. These low frequency waves induce an excitation of the Earth's normal modes to a nearly constant level [Webb, 2008].

In this research we focus on the waves in the microseisms interval (1-30 s). These periods are concentrated around two broad and dominant peaks called the primary peak (around 14 s) and the secondary peak (around 7 s) 3-2. Although observed globally, this period range experiences the strongest noise on data obtained from coastal and island stations [nak, 2019]. It is constructed out of a superposition of surface waves generated by strong surface winds due to intense cyclonic low-pressure storms. This movement of atmospheric pressure turns into ocean swell, which becomes the microseism energy observed in seismic data. Since the severity of these ocean storm are seasonal, a significant variation in microseism energy level can be observed based on differing season and climate factors [Aster et al., 2008]. The primary microseism peak results from direct pressure fluctuations at shallow parts of the ocean bottom, generated by breaking waves. The secondary microseism peak contains much more energy, and originates in standing waves created by the interaction between incoming swells and coastal reflections [Bromirski, 2009].

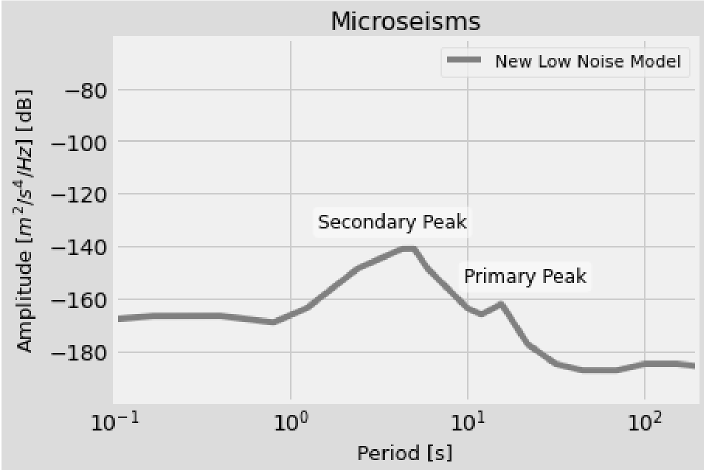


Figure 3-2: New Low Noise Model (by [Peterson, 1993]) clearly shows the primary microseisms peak around 14s and the secondary microseisms peak around 7s

3-2-3 Noise source illumination pattern

The lack of an uniform, or isotropic noise source illumination pattern leads to deviations of the retrieved surface wave responses from the actual, correct surface wave responses [Weemstra et al., 2021]. An isotropic illumination pattern is defined as an equal noise source distribution in all directions, or in other words; a net power flux of the illuminating wave field that is close to zero [Wapenaar et al., 2010]. Figure 3-3 shows different source illumination patterns. An anisotropic source illumination pattern can lead to a non-symmetric response, with differences between causal and acausal surface wave arrivals. Mainly sources in the Fresnel zones, indicated by the parabolic blue dotted lines in Figure 3-3, contribute to the interferometric response [Wapenaar et al., 2010]. Sources outside of this zone interfere destructively in case of similar magnitude, which makes that they give no coherent contribution. If there are no sources located within the Fresnel zone which emit ambient noise within the pass band of the used filter, no interferomteric response will be found [Wapenaar et al., 2010]. The situation of the noise source illumination pattern has to be assessed based on the period of the pass band of the filter that is used to compute the interferometric response. This is important since many noise sources mainly emit noise within a certain period band [Yang and Ritzwoller, 2008]. Also the possibility of spatial and/or temporal changes in the noise source illumination pattern has to be considered, examples for this can be seasonality or diurnal variations caused by cultural noise sources (section 3-2-2). These changes can be wrongly attributed to clock drift as explained in section OBS clock drift.

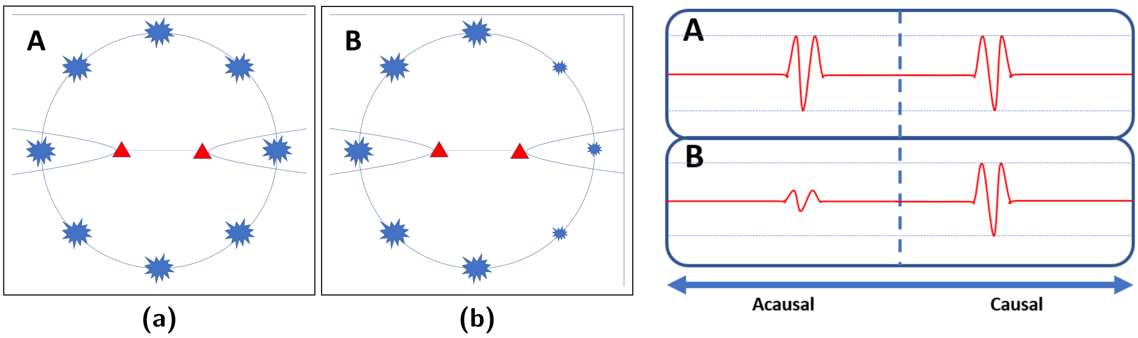


Figure 3-3: Map view: Red triangles indicate receiver locations; blue icons indicate source locations **(a)**: Receiver pair with uniform source illumination pattern; **(b)**: Receiver pair with non-uniform source illumination pattern; **(c)**: Interferometric surface wave response of **(a)** and **(b)**

3-3 OBS clock drift

A variety of methods have been developed to correct a potential time error in ocean bottom seismometers (OBS), e.g. [Loviknes et al., 2020, Hable et al., 2018, Weemstra et al., 2021, Naranjo et al., 2021]. The OCloC package used in this research aims to improve this process by building on the method presented by [Weemstra et al., 2021]. This chapter gives a general introduction on why a timing error can occur in an OBS, and what methods for correction already exist. Only a selection of methods are presented based on their relevance relative to OCloC.

3-3-1 Clock drift

The internal clock of an OBS can experience a drift over time, normally this drift can be corrected by using a reference time, given by the atomic clock on-board a satellite. On the ocean bottom this satellite connection is often not possible. The rate at which this drift happens is defined as the clock drift. Usually a linear correction is applied for OBSs by using the known synchronized times before and after deployment. The linearity of drift largely depends on the stability of sea water temperature and pressure [Shariat-Panahi et al., 2009]. Since both factors are mostly stable on the ocean bottom a linear correction is almost always considered sufficient. There are cases where it is impossible to apply a linear correction based on two points before and after deployment, for instance when the battery of the OBS runs out before recovery. This results in a missing synchronized end time to carry out the linear correction. For this reason other techniques using ambient noise cross-correlation functions or teleseismic P-wave arrivals are proposed to correct for the clock error [Loviknes et al., 2020, Hable et al., 2018, Weemstra et al., 2021, Naranjo et al., 2021, Sukhovich et al., 2021].

3-3-2 Methods based on noise cross-correlations

Several methods have proven to be successful in detecting and correcting the clock drift in OBSs by utilizing noise cross-correlations. Often these methods require an isotropic noise source illumination pattern, with sufficient energy located on both sides of the network. OCloC also makes use of these so called interferometric responses obtained by cross-correlating ambient noise from two receivers. The methods that utilize noise cross-correlations have a lot of similarities, three of them are explained briefly in the next section; a method based on temporal stability [Hable et al., 2018, Loviknes et al., 2020], a method based on time symmetry [Sens-Schönfelder, 2008, Weemstra et al., 2021], and OCloC [Naranjo et al., 2021] which is extensively explained in the Methods, chapter 4-3.

Temporal stability

To determine the clock drift the temporal stability of time-averaged cross-correlations can be used. To do this only the causal or acausal part of the interferometric response is required. For this method to work one of the stations used for the cross-correlation needs to be devoid of timing errors.

Noise cross-correlations that are time-averaged over a short period (e.g. 1 day) are computed and selected based on their SNR, this is called the short period cross-correlation function (CCF). A reference noise cross-correlation is constructed by stacking all short period CCFs from early in the deployment with a sufficient SNR [Hable et al., 2018], this is called a reference CCF. Determining the length of the reference CCF is prone to ambiguity, different methods are proposed by [Hable et al., 2018] and [Loviknes et al., 2020]. The short period CCFs are cross-correlated with the reference CCF, the clock error can then be determined by maximizing the Pearson correlation coefficient, which indicates the maximum correlation shift.

Temporal changes in noise source locations can also cause independent shape changes of the CCFs causal and acausal parts. This implies that the mentioned method based on temporal stability is only reliable in areas with a constant noise source illumination pattern. The accuracy of this method can be improved by averaging over a large amount of station pairs if they are available. Another method to improve the accuracy is disregarding cross-correlations with a normalised result below a certain threshold (e.g. 85 percent [Hable et al., 2018]). This method has proven to be able to correct for clock changes with a standard deviation of only 20 ms [Hable et al., 2018].

Methods based on temporal stability ignore the possibility of timing errors induced during the deployment of the OBS. These changes can occur due to a sudden change in temperature and pressure while the OBS is sinking to the ocean floor [Naranjo et al., 2021]. However, they do work with short inter-receiver distances, where the interferometric response is largely symmetric with the highest peak around 0 seconds [Loviknes et al., 2020].

Time symmetry

Other methods utilize the time symmetry of ambient seismic noise CCFs, one of these methods is described by [Weemstra et al., 2021]. The apparent arrival times of the interferometric surface waves are used to recover the clock error. This is possible by using the time symmetry usually inherent in time averaged cross-correlations as a result of the law of reciprocity. Contrary to the methods based on temporal stability both the causal and acausal surface wave responses are required. This can be problematic with an anisotropic noise source illumination pattern or when receivers are close to each other. When one or more of the analyzed receivers is devoid of instrumental timing errors, the timing errors of all other receivers can be determined. By using the time-averaged cross-correlation between two receivers the apparent arrival times of the direct surface-wave at positive time and the arrival of the surface wave at negative time is quantified [Weemstra et al., 2021]. The procedure which is used for this is defined in [Weemstra et al., 2021] and described in more detail in section 4-3.

The apparent arrival times are governed by a set of equations (given in section 4-3; equation 4-7), which can be inverted with a least-squares inversion. This is only possible if we assume that the spurious energy and non-uniform illumination patterns are cancelled out. Since especially the uniform illumination pattern is not necessarily true weights are added, based on the distance between the receivers. Longer distances decrease the adverse effect of inhomogeneities in the noise illumination pattern [Weemstra et al., 2021]. This is included in a weighted least-squares inversion. Inversion results are optimized by selecting CCFs with a sufficient signal-to-noise ratio (SNR) which reduces the risk of misidentifying the surface wave arrival. Additionally a receiver distance threshold is set, only receiver pairs with a certain distance are used. This is done to prevent including signals where the causal and acausal signal arrive at almost the same time [Weemstra et al., 2021]. This method does not directly calculate possible sudden time shifts that occurred during deployment due to pressure and temperature changes. Although such an error is not plausible it would be good to rule out [Naranjo et al., 2021].

Correction with OCloC

Although other methods have been proven to be successful, there is still room for improvement. Especially the possibility of time errors resulting from a sudden change in temperature and pressure during deployment would be good to rule out. The proposed correction method, called OCloC, is described by [Naranjo et al., 2021]. It allows for the estimation of an initial timing error, induced by the rapid temperature and pressure changes during deployment. The method combines two available techniques for clock error detection: temporal stability of CCFs and time symmetry of the retrieved interferometric responses. The method is an extension of the method described by [Weemstra et al., 2021] and is made available as a community code. Additionally the package aims to reduce the time a researched needs to spend in correcting the clock errors significantly, by being universally applicable and user friendly.

OCloC is the method used to correct OBS clock errors in this research, we explain it in more detail in the methods chapter OCloC.

3-3-3 Method based on P-wave arrivals

The methods we mentioned before are based on ambient noise cross-correlation functions. For completeness we mention one method which does not rely on these functions. The method described in this section uses P-waves to determine the clock error. We can determine the clock-drift from arrival times of teleseismic P-wave arrivals with the use of ray-tracing code as described by [Sukhovich et al., 2021, Hannemann et al., 2014]. This method assumes clock synchronization at the time of deployment. The clock drift can be determined by calculating the P-wave arrival of a teleseismic event using a global seismic-velocity model, and subtract this from the real P-wave arrival, as observed in the data. The uncertainty can be reduced by improving the event hypocenter location, origin times and the seismic velocity model [Sukhovich et al., 2021]. For this method to be accurate the distances between the OBS need to be significant, and the teleseismic events need to be above a certain severity [Loviknes et al., 2020].

A single P-wave arrival is not enough to reliably estimate the time-shift. Ideally a large set of teleseismic events occurred over the entire deployment of the OBS. The time axis of the OBS signal can be adjusted with the least-squares method to agree with the predicted arrival times of the P-waves. It should be noted that this method has a lot of uncertainties. Some worth noticing are: uncertainty in observed arrival time, uncertainty of hypocenter and uncertainty of seismic-velocity model [Sukhovich et al., 2021]. However, since this method does not require interferometric responses, it can be used for stations where these are difficult to obtain, for example due to an anisotropic noise source illumination pattern.

Chapter 4

Methods

4-1 Seismic noise analysis

The seismic data we use for this research originates from different seismic stations located in the Red Sea coastal area of the Kingdom of Saudi Arabia. The data is new and has not been used before, therefore it is crucial to conduct an extensive quality analysis to identify possible quality issues prior to using the data. To do this in a orderly way a roadmap was put together, containing the different steps for both land-stations and OBSs separately. This roadmap can be seen in Figure 4-1. The first step (checking filenames/filesizes) is useful to check if the data is harvested correctly and if the files are complete. The other steps are elaborated in the following sections.

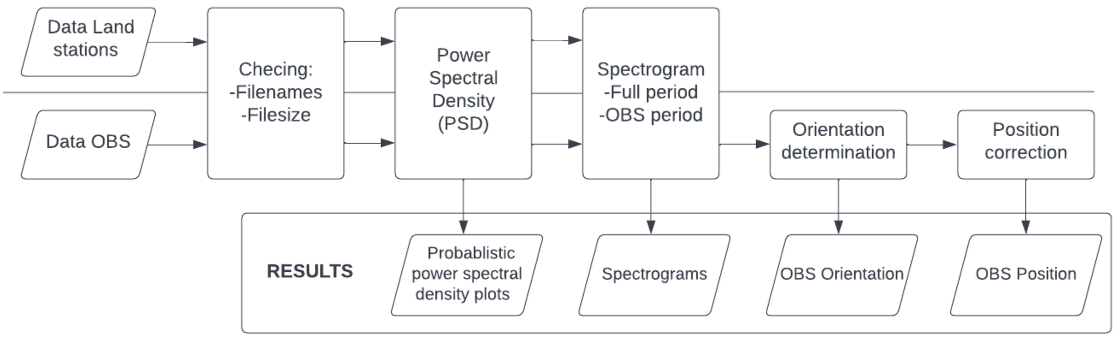


Figure 4-1: Roadmap showing steps taken for the data quality control, orientation determination and position correction are specifically for OBSs

4-1-1 PPSD

To qualitatively check the unprocessed data we compute the probabilistic power spectral density (PPSD) as suggested by [McNamara and Buland, 2004]. This means that the data can include instrumental glitches such as data gaps, spikes and clippings. These occurrences do not contaminate the seismic noise when used to obtain the PPSD, since they occur generally unrelated to each other, and with a low frequency. Still, it is important to check the data if these instrumental glitches occurred. A convenient way to do this is the spectrogram, which we explain in the next section 4-1-2.

To calculate the PPSD we first calculate the power spectral density (PSD) for each time segment of one hour. We do this individually for each component of each station. Calculation of the PSD is considered the standard method for quantifying seismic background noise. There are different ways to calculate the PSD for seismic noise data, in this research we use 'Cooley-Tukey method', which is considered the most common method [Cooley and Tukey, 1965, McNamara and Buland, 2004]. We calculate the PSD in the frequency domain by using a finite-range fast Fourier transform (FFT). First we cut the continuous time series in segments of 1 hour with an overlap of 50% to reduce variance in the estimated power spectral density. Each continuous segment of one hour is processed individually. The one hour segment is cut again into 13 segments with an overlap of 75%. To significantly increase the FFT computation speed we truncate the number of samples to the next lowest power of 2 [Bellanger and Daguet, 1974]. We remove long linear trends by demeaning with the average slope method. Finally, we use a 10% cosine taper on the time series to reduce side lobe leakage as a result of the carried out FFT [Schwarz et al., 1990].

The total power of the spectrum of each segment is defined as the square of the amplitude with a normalization factor of $2\Delta t/N$, where $N = T_r/\Delta t$, T_r is the length of the time segment and Δt is the sample interval [McNamara and Buland, 2004]. Since our data is discrete, we use the discrete Fourier transform as shown in Equation 4-1

$$Y_k = \frac{Y(f_k, T_r)}{\Delta t} \tag{4-1}$$

With the discrete frequency being $f_k = k/N\Delta t$ when $k = 1, 2, ..., N-1$. With the use of this discrete Fourier transform, the PSD estimate (P_k) is defined as:

$$P_k = \frac{2\Delta t}{N}|Y_k|^2 \tag{4-2}$$

We compute the power spectral density estimation for each segment, and remove the instrument response. We calculate the final PSD estimate for one hour as the average of the 13 segments converted to decibel (dB) [McNamara and Buland, 2004]. The power spectral density per hour is converted to a probability density function to construct the PPSD. The frequencies are sampled in 1/8 octave intervals, which reduces the frequencies by a factor of 169, within this interval the power is averaged and stored under the center frequency. This process can be seen in Figure 4-2 (a). After this is done for every 1 hour segment we accumulate the obtained powers in 1 dB bins to construct a histogram for each period, stacking the amount of occurrences which can be seen in Figure 4-2 (b).

Finally, we can compute the probability of occurrence of a given power at a particular period. we can plot this to obtain the final PPSD plot, which can be seen in Figure 4-2, In the results the PPSD plots for the data we used in this research are discussed and interpreted. As reference the new low noise model (NLNM) and new high noise model (NHNM) constructed by [Peterson, 1993] are added to the plot. Both models are constructed from seismic background noise obtained from a worldwide network of seismograph stations. They indicate the minimum and maximum background noise present without local disturbances and regiaonal variations [Peterson, 1993].

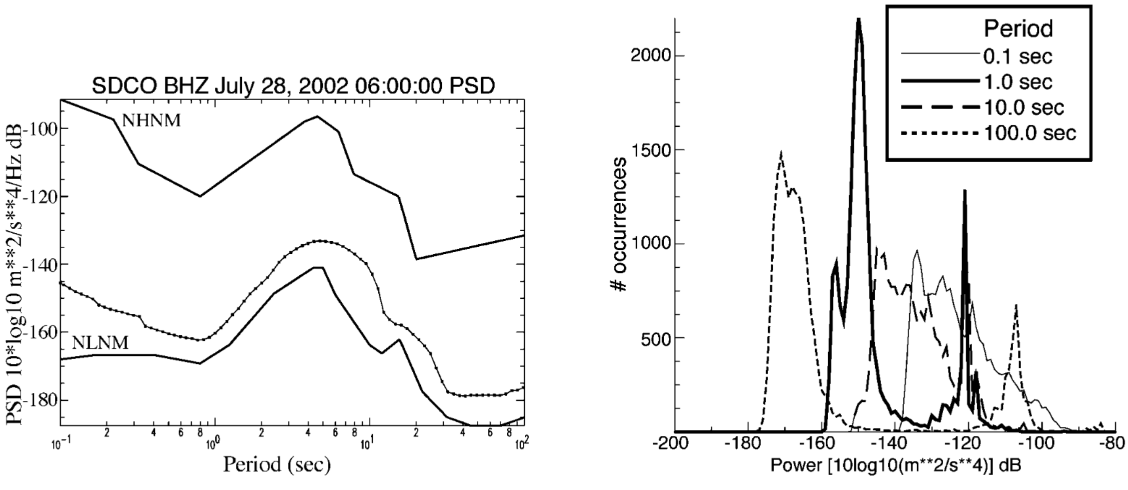


Figure 4-2: **(a)**: Frequencies of one hour segment are sampled in 1/8 octave intervals, averaged and stored under the center frequency. **(b)**: Powers of one hour segments are stacked and a histogram is constructed based on number of occurrences. adopted from [McNamara and Buland, 2004]

4-1-2 Spectrogram

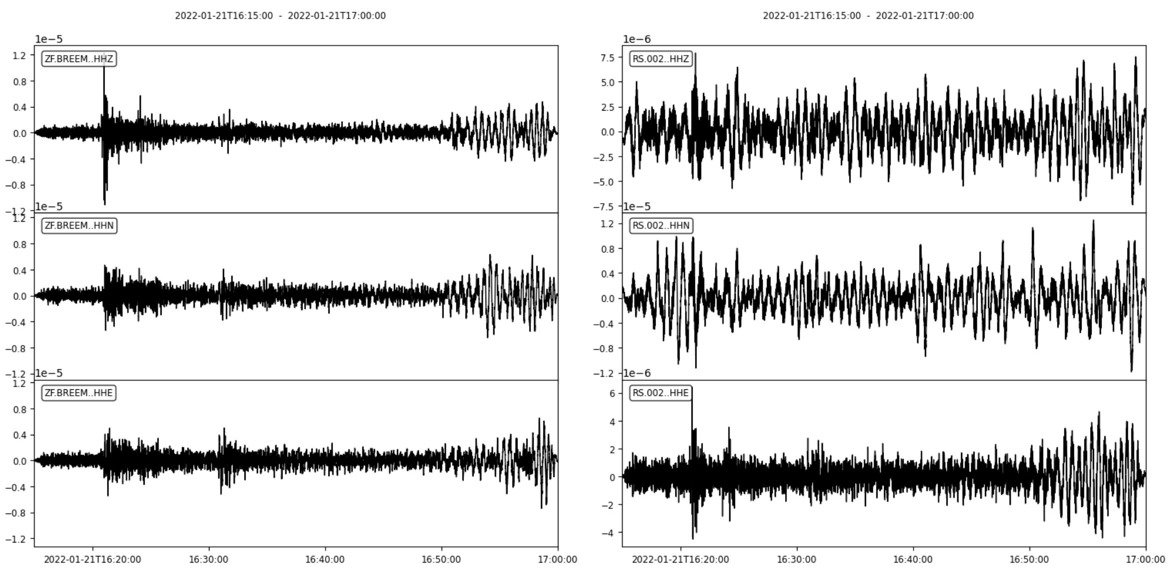
A spectrogram can be constructed from the temporal evolution of the power spectral density. The result can be used to improve understanding of the temporal evolution of the seismic noise and its potential periodicity. It is also a useful tool to observe the data completeness and to check how successful transient events have been removed, since periodic transient occurrences like traffic or weather are easily identified on the spectrogram.

In this research we construct the spectrogram from the hourly PSD bins (the PSD bin construction is explained in the previous section). Every hourly bin is plotted against the amplitude of a specific noise period. All hourly sections together generate a spectrogram, which we can use to see the temporal changes in the noise for a specific station. With this method it is possible to observe diurnal or seasonal variations and estimate its origin based on the period of the noise band. So can diurnal variations on the period band of 0.01 - 1 second (cultural band) often be assigned to difference in human activity between day and night.

4-1-3 OBS orientation

The OBSs require some additional analysis. Both stations feature a three-component seismic sensor. As mentioned in in the Chapter 2, the orientation of the device was not considered during installation on the bottom of the Red Sea. Which means that they can be in any possible orientation. The OBSs feature an auto-leveling sequence, which ensures that one component is exactly vertical en two components are horizontal [Nanometrics, 2019]. We do know that the OBS was placed on it's side to make it lay more stable on the sea bed. This would suggest that the vertical 'Z' component and one of the horizontal components are switched. A visual check of the waveforms from both OBSs clearly shows long period noise is dominating the vertical 'Z' and horizontal 'N' component (Figure 4-3), this is later confirmed by the Power Spectral Density 5-4. A strong long period noise and unclear signal is typical for horizontal components of OBSs, since the OBS can experience slight tilting caused by ocean currents [An et al., 2021, Lindner et al., 2017].

To check if this assumption holds we compare the data to another station that has a correct orientation. For this comparison we select a teleseismic event with a significant magnitude (6.3). The event took place in Shikoku, Japan, on the 21st of January, 2022. The stations selected for the comparison are island-station BREEM and OBS 002. Both stations were recording at the time of the event. We slice the data from both stations to fit the arrival of the teleseismic event. We remove the instrument response from both datasets to be able to compare the data. The comparison is shown in Figure 4-3, we clearly see the teleseismic event on all three components for island-station BREEM. First we see the direct P-wave arrival, followed by the P-wave reflected from the outer core. Approximately 10 minutes after the direct P-wave arrival, we see the direct S-wave arrival. On the data from OBS 002 the teleseismic event is clear on the 'E' component, the other components are dominated by long period noise. Since this long period noise is not visible in the other horizontal 'E' component we come to the conclusion that the vertical 'Z' component and the horizontal 'E' components are switched. This is normal for this type of OBS and does not lead to problems, since it can be oriented in any of the three positions [Nanometrics, 2019].



(a) **(b)**

Figure 4-3: Comparison land-station BREEM (left) and OBS 002 (right) for a teleseismic event in Japan with a magnitude of 6.3. We clearly see that components Z and N of OBS 002 are very noisy, this is a known problem for OBS horizontal components

After we defined the vertical component, we analyse the orientation of the horizontal components. Part of installing a seismic station is normally aligning the horizontal 'N' component with the true North [Stachnik et al., 2012]. For OBSs this alignment is often not possible since they are normally deployed from the side of a ship. Also the OBS stations we used in this research were not aligned with the true North. This means that the horizontal components can be in any possible orientation on the horizontal plane. Although this does not affect the vertical component, which we use in this research, it is still desirable to determine the actual orientation of the horizontal components. To do this we tested several methods with the aim to try and determine the orientation in a data driven way [Stachnik et al., 2012, Hofman et al., 2017].

The first method is inspired by [Hofman et al., 2017], where the orientation of borehole geophones was retrieved based on nearby accelerometers with known orientation. The distance between geophone and accelerometer was in the vertical direction and always less than 200 meters. Since this is not the case in our situation some changes had to be made, which mainly concerned the local earthquake direction. For the correction we used the OBS data and a nearby land-station (approx. 40 km) of which the horizontal 'N' component is aligned with the true North. Local earthquakes with a similar distance to both stations were selected. These local events had a relatively small magnitude. The station with aligned orientation is rotated in a way that the 'N' component is in radial alignment with the local earthquake. The OBS is iteratively rotated with steps of 1 degree, after which the signal is cross-correlated with the signal of the aligned land-station. The maximum correlation should theoretically occur when the OBS is rotated in a way that the 'N' component is in radial alignment with the local earthquake [Hofman et al., 2017]. This method has several disadvantages, mainly concerned to the different locations of the OBS and the aligned land-station, which results in a very different arrival signal from the local earthquakes. Another problem can be the use of local events. Since these events have a low magnitude, and our OBS systems seem to show an attenuation of the measured seismic signals (more about this in the Discussion), the events might not be measured sufficiently. The result did indicate a clear orientation range, this was however not yet precise enough. Because of that a second method is tested which should not suffer from the problems regarding small size earthquakes and different station locations.

The second method only requires data from the OBS with unknown orientation. It exploits the elliptical particle motion of Rayleigh waves (Chapter 3-1), which should theoretically only be observed in the vertical and radial directions. Instead of local events we make use of teleseismic events (section 3-1-2). The disadvantages of the first method do not exist, since only one station is used in combination with strong teleseismic events. The polarization analysis is performed by cross-correlating the vertical component with the Hilbert-transformed radial component [Stachnik et al., 2012]. A Hilbert transform is a linear operator which applies a phase shift of $\pi/2$ radians, or 90 degrees, on every (frequency-) component of a function. Theoretically this leads to a linear relationship between the two signals [Stachnik et al., 2012].

Teleseismic events that occurred during the OBS deployment are extracted from an earthquake catalog provided by IRIS (Incorporated Research Institutions for Seismology). Only events with a magnitude greater than 6.0 are used [Stachnik et al., 2012]. A time window is constructed based on the assumed Rayleigh wave arrival. This window runs from 20 seconds before a 4.0 km/s phase arrival to 600 seconds after it. The window is tapered with a 10% cosine taper and bandpass filtered for low frequencies from 0.02 to 0.04 Hz. These low frequencies are selected since the events occurs at a large distance and low frequencies tend to travel longer distances. The horizontal component is rotated iteratively with steps of 1 degree, for every new rotation the cross correlation is calculated given by:

$$S_{z\overline{r}} = \sum_{\tau=1}^{N} x_z(\tau) x_{\overline{r}}(\tau) \tag{4-3}$$

Where $x_z(\tau)$ is the vertical component and $x_{\overline{r}}(\tau)$ is the Hilbert-transformed radial component for that specific rotation. The cross-correlation is computed for every rotation and normalized to obtain the correlation coefficient. This step is repeated for every teleseismic event that occurred during the deployment of the OBS. The maximum correlation coefficient occurs when the radial component is rotated in such way that it points in direction of the teleseismic event. This angle can be used to compute the true orientation of the OBS with some simple geometry [Stachnik et al., 2012].

4-1-4 OBS position

Another known problem for OBSs can be the exact location. Normally an OBS is dropped of a ship, after which it slowly sinks down to the ocean bottom. It can take a long time for the OBS to reach the bottom, the OBS can drift laterally during this process. This can lead to an offset between the location the OBS was dropped and the actual location the OBS lands. However, this offset shows to be only problematic when the OBS is used for active seismic data, a small offset in location for passive seismic data does not lead to difficulties [Oshida et al., 2008], which is why no specific method is presented in this research. For the OBS data used in this research ultra short baseline acoustic positioning (USBL) was used to localise the OBS exact location on the bottom of the Red Sea. This was done with a transceiver mounted below the ship, the signal emitted is reflected by a transponder attached to the OBS. Software is used to calculate the location relative to the ship. This method is often not available during the deployment of OBSs since it requires specific equipment.

4-2 Pre-processing

After the quality of the data is analysed it can be pre-processed in preparation for the OCloC package. This requires several steps which we present in a second roadmap. The steps and parameters are again optimized for this specific case, but can be tailored to fit any other seismic network, similar to the quality analysis. First we look at the data of each individual station and apply single-station processing steps. Then the data is combined, cross-correlated and stacked to obtain the interferometric responses based on the method explained by [Bensen et al., 2007]. The pre-processing is very similar for the land-stations and OBSs. The steps used for pre-processing are shown in the roadmap in Figure 4-4 and are discussed in more detail in the following sections.

4-2-1 Instrument response

Every instrument has a distinctive instrument response. This response is removed individually from every station. Neglecting this instrument response removal can lead to significant amplitude and phase errors [Wilson et al., 2013]. Since we stack the cross-correlated data we are only interested in the phase response, as the amplitude does not provide any value after stacking. The phase response within the pass band of an instrument is typically presumed to be flat. If this is the case for all instruments used, the instrument response does not need to be removed. However, the majority of instruments do not have a truly flat response as

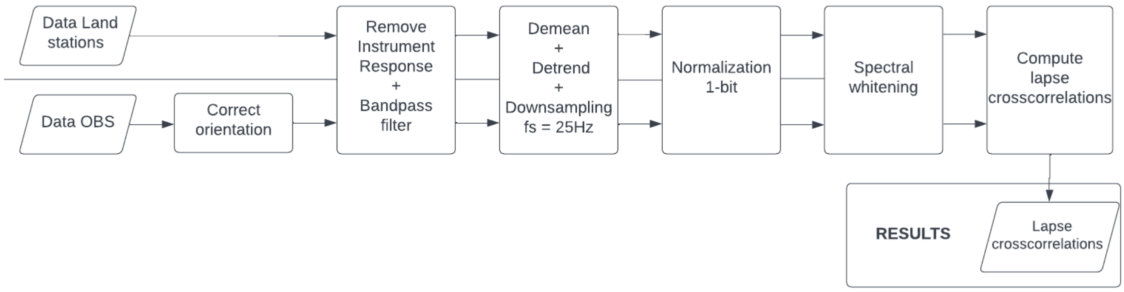


Figure 4-4: Roadmap showing steps taken to pre-process the data. Most steps are applied to 1-day time windows. Only for the computation of lapse correlations the data is cut to 1-hour time windows.

observed by [Wilson et al., 2013]. If in this case the instrument response correction is not carried out, it can result in timing errors that can affect the data significantly.

The instrument response can be defined as a relationship between the recorded data and the actual ground motion. The gain of the instrument is the ratio between recorded data and the actual ground motion, this information is obtained by calibration and is embedded in the seismic sensor as the complete instrument response function [Trebbin and Wassermann, 2010]. After retrieving this function it can be removed with the ObsPy package. The correction is carried out by deconvolving the instrument response function from the seismic data. To prevent over-amplification while deconvolving the instrument response a pre-filter is applied in the frequency domain. The used filter is a cosine taper. The corner frequencies are chosen as wide as possible, but narrow enough to already filter out the nyquist frequency, which results from the downsampling done in the next steps. In this way the pre-filter has a double function as anti-aliasing filter.

4-2-2 Downsampling

Two different sampling frequencies are used in this research. The land stations sample at a frequency of 125 Hz. The OBSs sample at a frequency of 100 Hz. To be able to compare the data the same frequency should be used. Additionally, the noise frequencies of interest are very low compared to the sampling frequency. This suggests that the data can be significantly downsampled, while retaining the desired noise information. In this case downsampling is desirable, as it considerably decreases the computing power required for pre-processing.

Downsampling to a unlike denominator (e.g.: 50 Hz) requires interpolation, which uses more computing power and is prone to errors. To prevent the need of interpolation, the decision is made to decimate both sampling frequencies, which means downsampling to a like, or common, denominator. The new sampling frequency: fs becomes 25 Hz, as this is the largest common denominator of both sampling frequencies. This means that for the land-stations 80% and for the OBS 75% of the data can be omitted.

4-2-3 Normalization

To look at ambient noise it is important to exclude strong events like earthquakes. If this is not done correctly the strong amplitude of these events will dominate the noise signal. Computing the different arrival times of earthquakes for each station based on an earthquake catalogue is time consuming and prone to errors, it also does not include strong amplitude events that are induced by instrument errors or local events. The preferred method to correct for strong signals is in a data-adaptive manner, as suggested by [Bensen et al., 2007]. There are different methods to correct the data, examples are one-bit, (iterative) clipping and running absolute mean normalisation (RAMN). With the one-bit normalisation only the sign of the signal is kept, all amplitude information is disregarded. This method is known to significantly increase the signal-to-noise ratio. One major downside is the lack of flexibility; i.e. it is not possible to change the method based on local circumstances [Bensen et al., 2007]. The clipping method scales down the signal above a given amplitude, this results in high amplitude events being cut of, and getting weaker relative to the low amplitude noise. This can be done once or iteratively until the entirety of the signal is below a certain clipping/water level. Selecting the clipping factor and downscale factor can be arbitrary and result in errors [Bensen et al., 2007]. The running absolute mean normalisation computes the running average of the absolute value of the waveform in a normalization time window and weights the waveform at the centre of the window by the inverse of its average. Since the seismic signal is a discrete time series, we compute the normalization weight for time point n as:

$$w_n = \frac{1}{2N+1} \sum_{j=n-N}^{n+N} |d_j| \tag{4-4}$$

The signal is then normalised by dividing the signal at the middle of the normalization time window by a weighing factor that can be defined based on the local circumstances [Bensen et al., 2007]. The selection of a suitable normalization method should be based on the situation, there is not one specific method that works best in each situation. When high amplitude contaminants pass the normalization they tend to appear on cross-correlations as spurious precursory arrivals. This can be prevented by adding temporal weights, down scaling regional seismic activity and fine tuning them [Bensen et al., 2007].

The right normalization method was selected by testing different methods and parameters on a teleseismic event for different stations. An earthquake in southern Iran was selected with a magnitude of 6.0. 3 hours around the earthquake arrival are shown to include all arrivals regarding to the earthquake, as well as ambient noise. The results for land-station 'DEEP' can be seen in Figure 4-5, where A is the raw signal, and B, C, D, E and F are different methods of normalization. Figure 4-5 A clearly shows how the high amplitude of the earthquake dominates the signal, the ambient noise occurring before the earthquake is not visible at all, this shows the importance of a correct normalization to prevent the domination of strong signals. Figure 4-5 B shows the one-bit method. We see that the signal's amplitude is transformed to either 1 or -1, while the phase information is retained, this method seems to be most successful in correcting the high amplitude events and keeping the ambient noise. Figure 4-5 C and D show the running mean average normalization method (RAMN) with different parameters. C uses a normalization window of 10 seconds, as suggested by [Bensen et al., 2007]. We see that the ambient noise becomes visible, but the earthquake still dominates the signal. D uses a normalization window of 10 traces (10/125 seconds). We see a clear noise signal, but the overall power of the earthquake signal still is very strong, and would dominate after cross-correlation. Also other normalization parameters were tried, but they did not result in a better normalization capacity. Figure 4-5 E and F show the iterative clipping method, both again with different parameters. E uses a clipping factor of 6 and a down scale weight of 10 as suggested by [Bensen et al., 2007]. This significantly increases the ambient noise power, but does not give the desired result, the same can be said for F, where a clipping factor of 3 is used in combination with a down scale weight of 10.

After testing the different methods we can conclude that only the one-bit method is successful at normalizing the signal sufficiently, according to [Bensen et al., 2007] this method also provides a decent signal to noise ratio. This justifies the one-bit method being used for normalization in this study.

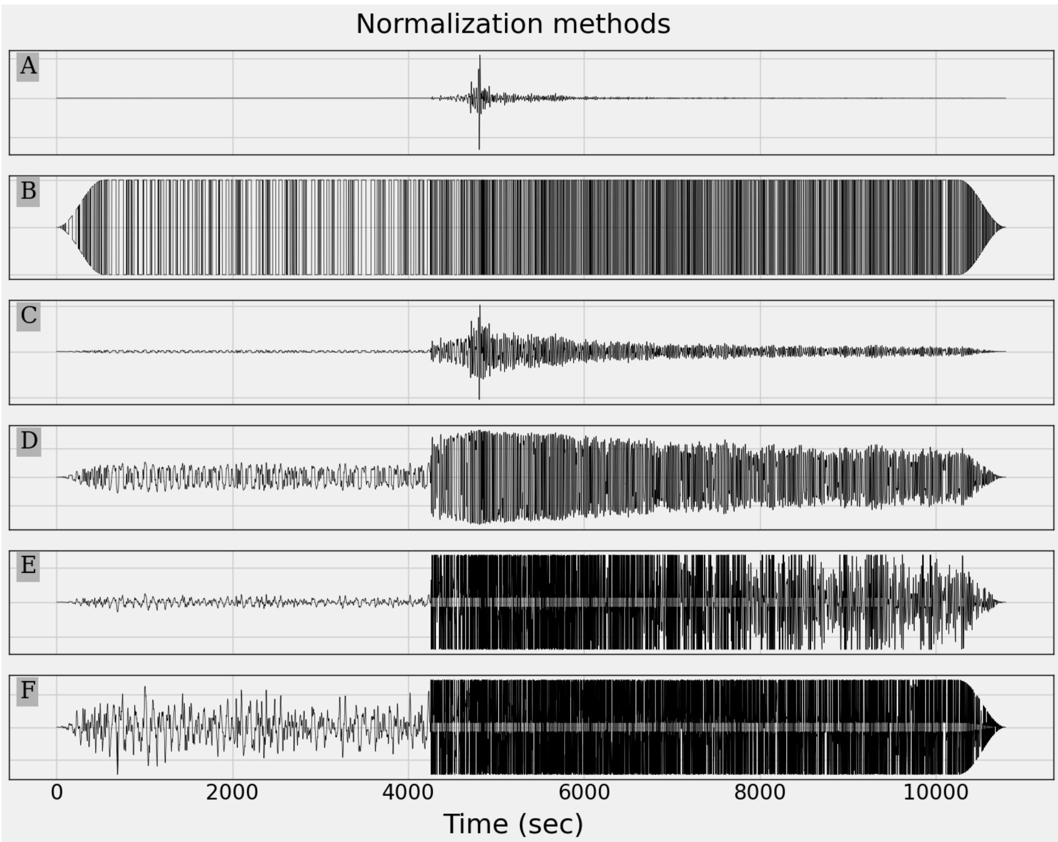


Figure 4-5: Different methods of normalization for land-station 'DEEP', performed on 3 hour signal around an earthquake of magnitude 6.0 in Southern Iran. A) raw signal, without any normalization. B) one-bit normalization. C) running mean average normalization (RAMN) with a normalization window of 10 seconds. D) RAMN with a normalization window of 10 traces (10/125 seconds). E) iterative clipping normalization with a clipping factor of 6 and a down scale weight of 10. F) iterative clipping normalization with a clipping factor of 3 and a down scale weight of 10.

4-2-4 Whitening

Although ambient noise is necessary to obtain an interferometric response, isolated and nearly monochromatic noise sources can corrupt the cross-correlation when no whitening is applied. Whitening, or spectral normalisation, seeks to smoothen the amplitude spectrum of the time-averaged cross-correlation functions and decreases the degradation caused by isolated, nearly monochromatic noise sources [Bensen et al., 2007]. Whitening can be performed in both the time- or frequency-domain. Time-variant spectral whitening is performed by applying a series of narrow bandpass filters on the signal. A gain function is computed for each frequency band, creating a series of gain functions. The inverses of these gain functions are applied to each frequency band, and the results are summed. The method used in this research is in the frequency domain. It uses an apodization function to smooth discontinuities in the frequency domain, which has the form of a cosine tapered boxcar bounded by a minimum and maximum frequency.

4-2-5 Cross-correlation and stacking

The cross-correlation is performed in the frequency domain on data slices of one hour. The translation is done with the fast Fourier transform (FFT) algorithm, which is known for its fast computation speed [Brigham and Morrow, 1967]. After translation back to the time domain the result is a so called cross-correlation function (CCF), which can be 'stacked' by simple addition to obtain the time series for the whole period [Bensen et al., 2007]. Splitting the cross-correlation function into sections and creating a stack is possible because of the linearity of the operation, this guarantees that the result is the same as cross-correlating a longer time-series. The final CCF consists out of a causal (positive) and acausal (negative) signal. These signals represent waves traveling in opposite direction between the selected stations. The signal can be compressed into a one-sided signal by averaging the causal and acausal parts. This is called the 'symmetric signal' [Bensen et al., 2007].

Stacking over a longer time-series usually increases the signal-to-noise ratio (SNR). Since the SNR is described as a function of the square root of the acquisition time it should theoretically improve proportional to the square root of the increase in time used for the cross-correlation. This would mean an increase from 2 to 8 months would double the SNR under perfect, constant conditions. The amount of data used to stack should be a trade-off between computational cost and the improvement of the signal-to-noise ratio. When the cross-correlations are used as lapse cross-correlations the mid time offset should also be considered.

The resulting stack is filtered using a narrow band pass filter. The aim is to maximise the SNR of the retrieved surface waves, to do this different period pass bands are tested since the surface wave does not necessarily exist on each period due to differences in noise illumination. When filtering for surface waves the wavelength and station distance should be monitored. This is calculated using the surface wave velocity and period range. A rule of thumb is to have at least 2 full wavelengths in between the stations.

4-2-6 Lapse cross-correlations

OCloC uses time-lapse cross-correlations to estimate possible clock errors in OBSs. This time-lapse is put together from different 'lapse cross-correlations', all with an equal time window length, but shifted in time. These time-averaged lapse cross-correlations are constructed from data within this given time window (e.g. 30 days). They are marked by the average time of all data that contributed to the lapse cross-correlation. That means that it is not necessarily the mid-time of the time window, gaps in the data can skew the average time towards the beginning or the end of this time window [Naranjo et al., 2021]. To implement this every minute of data that contributed to the lapse cross-correlation is stored, this is later used to compute the average time. A wide variety of lapse cross-correlations are constructed by changing their time window and overlap. An example can be: lapse cross-correlations with a time window of 30 days and an overlap of 0 days. If the total data length is 100 days 3 different lapse cross-correlations can be obtained.

4-3 OCloC

In this research the time-dependent clock error of OBS is detected and corrected with the OCloC (OBS Clock Correction) package. This is done by utilizing the temporal stability of the computed cross-correlations in combination with the time symmetry of retrieved interferometric responses [Naranjo et al., 2021].

The time-dependent clock error of the OBS is defined in [Naranjo et al., 2021] as:

$$\delta t_i^{(ins)}(t^{(lps)}) = a_i t^{(lps)} + b_i \tag{4-5}$$

Where $\delta t_i^{(ins)}$ indicates the clock error of station i at time $t^{(lps)}$, a_i is the clock drift rate of station i and b_i is the initial timing error of station i at zero time. Our goal is to retrieve both a_i and b_i for each different station i. This section explains the theory behind OCloC. First we discuss what criteria are used to select eligible lapse cross-correlations. Then we take a closer look at the theory behind the determination of the surface wave arrival times and finally we look at the solution of the inverse problem.

4-3-1 Selecting lapse cross-correlations

The different lapse cross-correlations are given by: $C_{i,j}(t, t^{(lps)})$, in which $t^{(lps)}$ is the average time of the lapse cross-correlation, and i and j indicate the different seismic stations. The lapse cross-correlations are a result from the pre-proccesing and are imported in the package. It is useful to plot the lapse cross-correlations in a single graph to get a first indication if clock drift exists in the OBS data, when this is the case the interferometric responses should show a temporal shift in between lapse cross-correlations. Some lapse cross-correlations should be filtered out, to select eligible lapse cross-correlations both the signal-to-noise ratio (SNR) and station distance are important. The ability of the algorithm to identify the surface wave arrival times strongly depends on the SNR. A low SNR will result in difficulties finding the surface wave arrival times and can be subject to cycle skipping. Cycle skipping in this context can be explained as the surface wave arrival time being off a value of approximately one center period of the filter's pass band (in this case 10-20 s) [Naranjo et al., 2021, Weemstra et al., 2021], this would be similar to what is referred to as cycle skipping in full waveform inversion [Weemstra et al., 2021, Warner and Guasch, 2014].

Station separation has to be sufficiently large to prevent the causal and acausal surface wave response from overlapping, if this happens the algorithm will be unable to differentiate between the causal and acausal waves. To prevent this, a minimum station separation is determined based on a threshold expressed in wavelengths [Naranjo et al., 2021].

4-3-2 Determining surface wave arrival times

A property of the noise cross-correlation is the time symmetry of the interferometric surface waves. Under certain assumptions (see chapter 3-3) the surface wave arrivals are symmetric. A deviation from this symmetry could indicate the presence of a clock error [Naranjo et al., 2021]. To measure this, the apparent arrival time of the causal and acausal direct surface wave in each lapse cross-correlation $C_{i,j,k}(t, t^{(lps)})$ is measured, where k indicates the mid-time of that lapse cross-correlation between station i and j. The measured, or apparent, arrival times of the direct surface wave are then defined as $t_{i,j,k}^{(+,app)}$ for the causal situation and as $t_{i,j,k}^{(-,app)}$ for the acausal situation. The theoretical time symmetry due to the law of reciprocity, which states that the same seismic signal should be recorded if the locations of the source and receiver are exchanged [Knopoff and Gangi, 1959], means that by definition $t_{i,j,k}^{(+,app)} = t_{i,j,k}^{(-,app)}$ [Naranjo et al., 2021]. When we use this to define an expression for the apparent arrival time we obtain:

$$t_{i,j,k}^{(+,app)} + t_{i,j,k}^{(-,app)} = 2a_i t_k^{(lps)} + 2b_i - 2a_j t_k^{(lps)} - 2b_j \tag{4-6}$$

For an ideal case with (I) an isotropic noise source illumination pattern, (II) cancellation of spurious energy due to the time-averaging process, and (III) recordings that are not subject to clock errors, the right hand side of Equation 4-6 evaluates to zero [Naranjo et al., 2021]. However, in the case clock drift occurred the signals are shifted, making the interferometric response asymmetric. As explained in chapter OBS clock drift these shift can also happen due to a change in noise-source distribution or a change in medium properties, however then the shifts occur in the opposite direction, which retains the symmetry [Loviknes et al., 2020]. Timing issues cause a shift in the cross-correlation which yields an asymmetrical signal.

The calculation of the apparent arrival times $t_{i,j,k}^{(+,app)}$ and $t_{i,j,k}^{(-,app)}$ is based on a procedure mentioned in [Weemstra et al., 2021]. It is carried out on the bandpass filtered time-averaged cross-correlations, of which the central frequency of the pass band is denoted as f_c. The procedure first disregards the time-averaged cross-correlations of station pairs with an insufficient distance in between them, this is determined based on the number of wavelenghts that fit in between the stations and the given distance threshold, which is also measured in wavelenghts. Then the expected surface wave arrival windows are calculated, based on station offset and a given reference surface wave velocity, this is defined as the a priori estimate. The expected surface wave arrival windows are used to determine the signal to noise ratio (SNR), time averaged cross-correlations with a SNR below a set threshold are disregarded.

For the set of time averaged cross-correlations that remain the peak and trough envelope is computed, this is done to detect the peaks and troughs of the signal. The time at which the difference between both envelopes is maximum is determined, this is only done for signal within the expected surface wave arrival windows (also known as the a priori estimate). A time window of $1/f_c$ around this maximum offset time is interpolated, this is required to ensure that the temporal resolution of the apparent surface wave arrival is sufficiently high [Weemstra et al., 2021]. The signal in the acausal interpolated time window is time reversed and thus forms an overlay with the causal interpolated time window. The start of both windows is set to $t = 0$, i.e. the start of the time reversed acausal time window and the causal time window. The signals within the interpolated windows are cross-correlated, the time lag for which this cross-correlation attains its maximum is used to derive the apparent arrival times $t_{i,j,k}^{(+,app)}$ and $t_{i,j,k}^{(-,app)}$ [Weemstra et al., 2021].

The workflow used by OCloC is described in Figure 4-6. We start with the lapse cross-correlations of different station pairs. A bandpass filter is used to retrieve the interferometric responses. The periods used for this bandpass filter should maximize the signal-to-noise ratios of the surface waves present in these interferometric responses. Then the lapse cross-correlations are selected, based on the just retrieved SNR and station offset, which is explained in the previous section. With this selection of lapse cross-correlations the measurement of the time symmetry shift is carried out. First an estimate of the drift is determined by cross-correlating the earliest lapse cross-correlation with the latest lapse cross-correlation for each station pair, with the assumption that the drift is linear [Naranjo et al., 2021]. This drift is refined to obtain the apparent arrival times $\mathbf{t}^{(app)}$ by using the procedure mentioned before, that detects the direct surface wave arrival for each lapse cross-correlation.

Processing step	Methods	Input
1. Retrieval of interferometric responses	- Stacked cross-correlations - Pre-processing + Bandpass filtering	- Seismic data - Filter ranges
2. Data selection	- Calculate signal-to-noise ratio - Calculate station offset thresholds	- Lapse cross-correlations - Station locations
3. Measuring the time symmetry shift	- Calculate Apriori estimate of time symmetry shift - Calculate time symmetry shift using approach of Weemstra et. Al., (2020)	- First and last lapse cross-correlations - Surface wave velocity - Station locations
4. Linear system of equations and inverse calculations	- Building the linear system of equations - Weighted least-squares inversion	- Lapse cross-correlations - Apriori estimates - Station locations
5. Removing wrong measurements	- Identify and remove outliers in the measurements	- Computed time symmetry shift - Max. error
6. Results refinement	- Re-using the obtained results for several iterations	- First iteration result - Number of iterations

Figure 4-6: Overview OCloC package with processing steps, used methods and required input, adapted from [Naranjo et al., 2021]

4-3-3 Solving inverse problem

The apparent arrival times are combined in a system of equations given by:

$$\mathbf{A}\mathbf{t}^{(ins)} + \mathbf{n}^{(src)} + \mathbf{n}^{(spur)} = \mathbf{t}^{(app)} \tag{4-7}$$

in which the vector $\mathbf{t}^{(ins)}$ contains the clock drift rates a and initial timing error b. Matrix $\mathbf{A}$ contains the different lapse times for each station pair. The vector $\mathbf{t}^{(app)}$ contains the measured, or apparent, arrival time of direct surface waves. More details on the exact formulation can be found in [Naranjo et al., 2021]. A difference in amplitude between the causal and acausal arrival can introduce an additional time shift, which can either be caused by the presence of spurious energy or an unequal source distribution. In this method the n terms regarding spurious energy and non-uniform illumination patterns are cancelled out. To do that we make the assumption that the ensemble-averaged cross-correlations cancel out the spurious energy, since a sufficient time span is used. Additionally we assume to have a uniform noise source illumination pattern [Naranjo et al., 2021]. These assumptions reduce Equation 4-7 to:

$$\mathbf{t}^{(app)} = \mathbf{A}\mathbf{t}^{(ins)} \tag{4-8}$$

This system of equations can be solved using a least-squares estimator with the Numpy Linear Algebra Package [Harris et al., 2020], which provides the a and b values for each station. Since the assumption of non-uniform illumination pattern is not necessarily true, the option of computing a weighted least-squares inversion based on station separation is provided [Naranjo et al., 2021].

To improve the model, the measured arrival time of direct surface waves is compared with the expected arrival time, which is computed based on the a and b from the first inversion iteration. This is done to prevent cycle skipping, a deviation from the true arrival time by approximately one center period of the filter's pass band (in this case 10-20 s), leading to incorrect a and b values. [Weemstra et al., 2021]. The comparison shows a linear trend, of which outliers due to cycle skipping can be excluded with the use of a pre-defined threshold. After the outliers are defined and removed the remaining values are used for the next inversion. This process is repeated until both the a and b stop showing significant changes with each iteration, which can be determined by plotting the evolution of a and b per iteration [Naranjo et al., 2021]. The resulting a and b can be used for the clock drift correction and checked by looking if the corrected interferometric responses align properly. The clock drift correction was successful if the response align.

Chapter 5

Results

5-1 Noise analysis

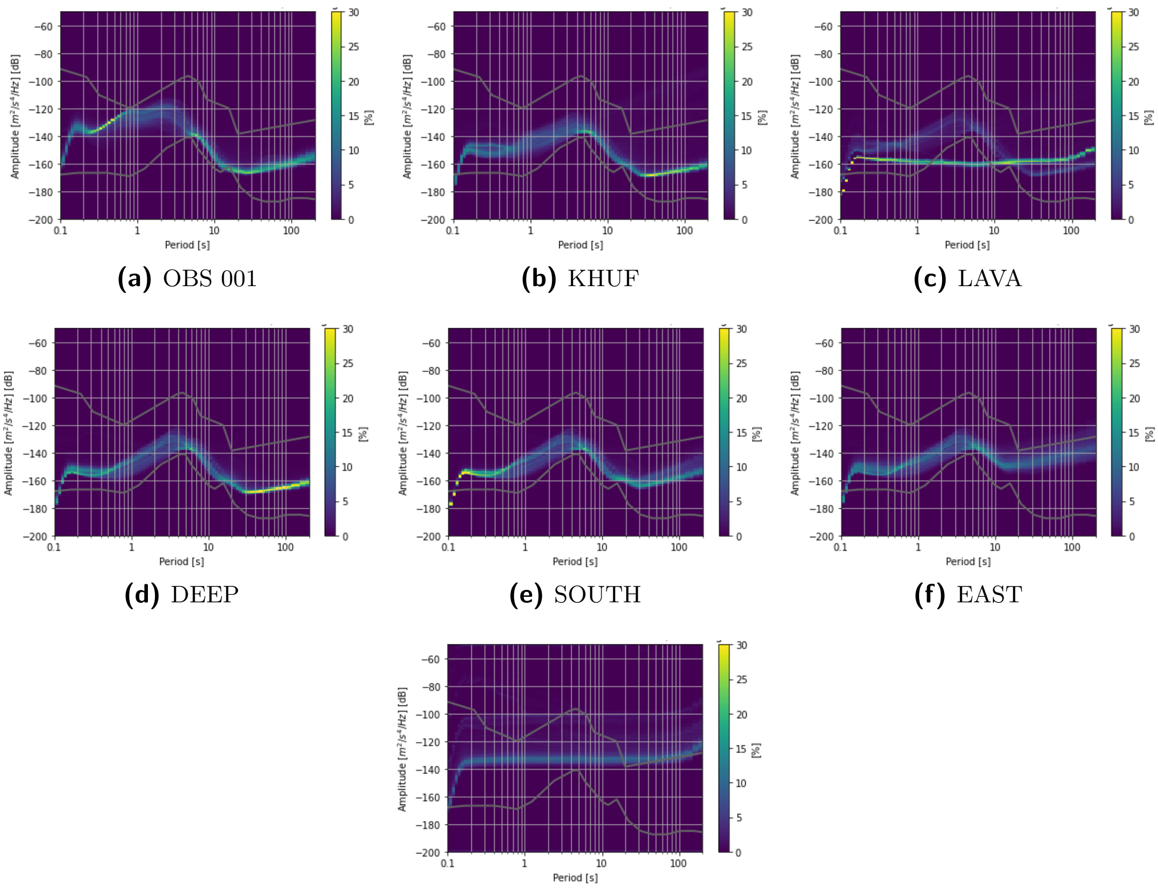
In this chapter we give the results obtained from the quality control and pre-processing. We obtained these results by using methods given in two roadmaps, which we explained in section 4-1. The roadmaps can be seen in Figure 4-1 and Figure 4-4. The relevant results from the quality control are analysed, starting with the power spectral density (method explained in section 4-1-1), then the spectrogram (method explained in section 4-1-2) and OBS orientation (method explained in section 4-1-3). Then we show the final result (i.e. cross-correlations) from the pre-processing (method explained in section 4-2-5). By doing the quality control and pre-processing we can analyse if the data is complete, the quality is sufficient and if the interferometric responses show a distinct causal and acausal wave.

5-1-1 Power spectral density

The power spectral density of each station is computed and visualized using a probabilistic approach, as described in theory section 4-1-1 and by [McNamara and Buland, 2004]. We group the seismic data based on a common deployment time with either OBS 001 or 002. The deployment time of the OBS is setting the boundaries for the other stations, i.e. data outside the OBS deployment time is disregarded.

First we take a look at the seismic data obtained during the deployment of OBS 001. The results can be seen in Figure 5-1. The probablistic power spectral density from OBS 001 and land-station KHUF (Figure 7-2 (a) and (b) respectively) show an expected result; the power spectral density stays within the low and high noise model (NLNM and NHNM by [Peterson, 1993]). We see a strong presence of short period noise in the OBS station compared to the land-station, this is according to the expectations mentioned in section 4-1-3.

When we take a look at the result from land-station LAVA in Figure 5-1 (c) we see abnormal behaviour. Two separate probability bands can be seen, with the majority concentrated in the lower band. The lower band is located roughly at the same amplitude for every period and includes amplitudes lower than the new low noise model between 2 and 10 seconds. This in non-physical behaviour and indicates a problem with the station. The fact that we see some probability in the upper band, following a more expected trajectory, indicates that the problems with the sensor started soon after deployment [McNamara and Buland, 2004]. A similar pattern is visible for land-station WEST in Figure 5-1 (g), although in this case



(a) OBS 001 **(b)** KHUF **(c)** LAVA

(d) DEEP **(e)** SOUTH **(f)** EAST

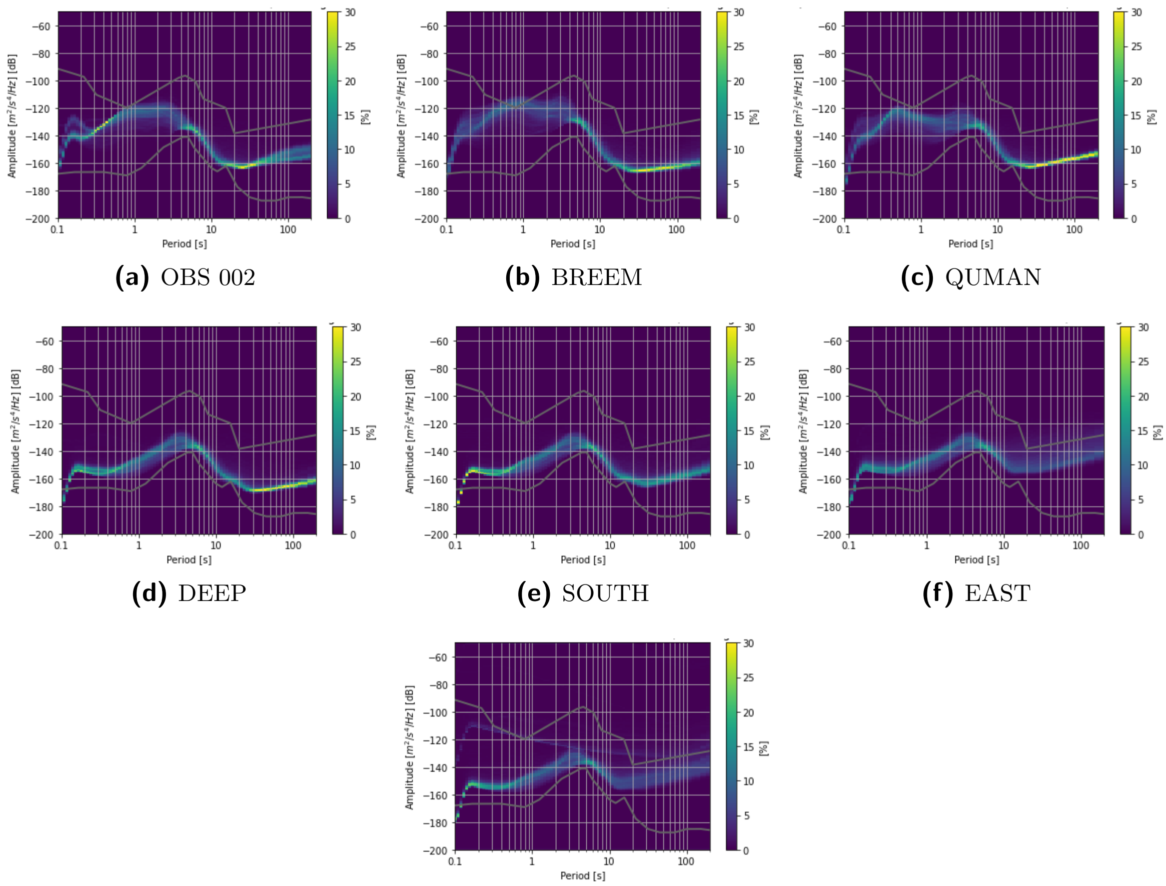
(g) WEST

Figure 5-1: PPSD of stations, same running time as OBS 001. **(a)**: OBS 001. **(b)**: Land-station KHUF. **(c)**: Land-station LAVA. **(d)**: Land-station DEEP. **(e)**: Land-station SOUTH. **(f)**: Land-station EAST. **(g)**: Land-station WEST.

only one faulty probability band is visible, indicating the problems started before the used timespan.

Next we take a look at the stations with an overlap of running time with OBS 002 As before we only include data obtained during the deployment of OBS 002. All stations show the expected result, with amplitudes above the new low noise model. The difference in noise between land-stations placed on Saudi Arabian soil, and stations close to, or in, the Red Sea becomes very clear. Both the OBS (Figure 5-2(a)) and island-stations (Figures 5-2(b)(c)) have stronger short period [0.2-5 s] noise levels compared to the land-stations (Figure 5-2(d)(e)(f)(g)), which are very similar to each other.

The difference between OBS, island- and land-stations is more obvious when we plot the mean PPSD of each station in one figure, as can be seen in Figure 5-3. We clearly see the short period noise generated by the Red Sea dissipating land inwards. As expected, long period waves keep more of their energy over a longer distance. A significantly higher amplitude can be seen at short period noise, e.g. [0.1-2 s] for the OBS and island stations compared to the coastal and inland stations.



(a) OBS 002 **(b)** BREEM **(c)** QUMAN

(d) DEEP **(e)** SOUTH **(f)** EAST

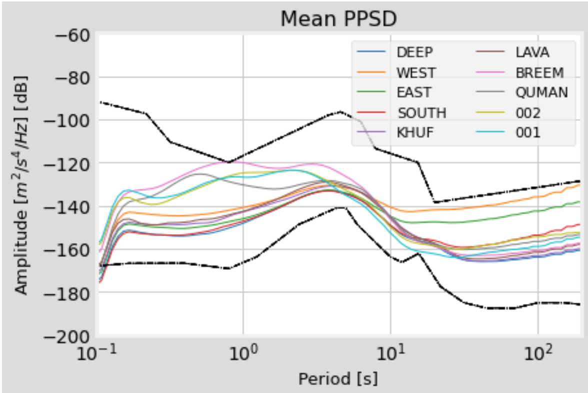
(g) WEST

Figure 5-2: PPSD of stations, same running time as OBS 002. **(a)**: OBS 002. **(b)**: Island-station BREEM. **(c)**: Island-station QUMAN. **(d)**: Land-station DEEP. **(e)**: Land-station SOUTH. **(f)**: Land-station EAST. **(g)**: Land-station WEST

Figure 5-4 shows a PPSD plot for each component of OBS 001 and 002, which were deployed on the bottom of the Red Sea. We clearly see the excessive long period noise present on component 'Z' and 'N'. Overall, long period noise levels are higher in horizontal components, as compared to the vertical component for OBSs [Uthaman et al., 2022]. Since the used OBS device can be oriented with any component functioning as vertical component (it does not have a preferred orientation), the observation is made that component 'E' is the vertical component. This is further explained in section 4-1-3

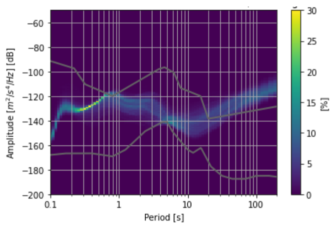
5-1-2 Spectrogram

The spectrograms are analysed according to a similar approach as the PPSD in the previous section. The aim of this method is to obtain information on the temporal evolution of the ambient noise and to check if the stations functioned properly, it is extensively discussed in section 4-1-2. The OBS deployment time is too short to clearly identify seasonal changes in the frequency bands. First we take a look at the spectrograms with the same running time

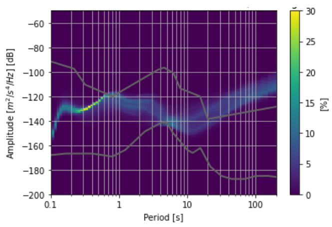


(a) Mean PPSD per station

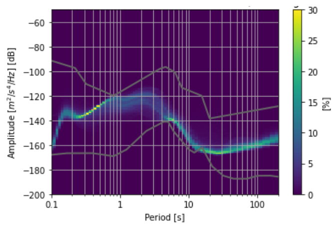
Figure 5-3: Mean of Probablistic Power Spectral Density (PPSD) per station. The lower and upper black lines show the NLNM (new low noise model) and NHNM (new high noise model) respectively



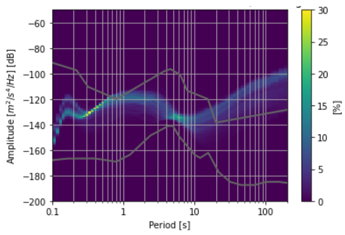
(a) OBS 001 Z



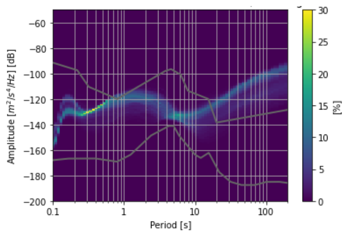
(b) OBS 001 N



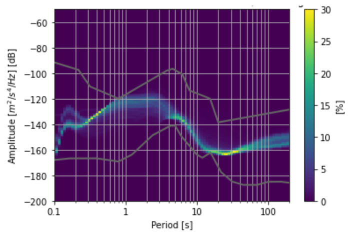
(c) OBS 001 E



(d) OBS 002 Z



(e) OBS 002 N



(f) OBS 002 E

Figure 5-4: PPSD of OBS 001 and 002, component Z, N, E

as OBS 001, which is shown in Figure 5-5. We see a stronger presence of short period noise in the OBS station compared to the land stations, which agrees to our findings in the PPSD shown before. The spectrogram of land-station LAVA Figure 5-5 (c) confirms our hypothesis originating from the PPSD data visible in Figure 5-1 (c). The station malfunctioned around halfway April, and afterwards did not provide reliable data. The data from land-station LAVA after half April can therefore not be used. The same is true for land-station WEST, which can be seen in Figure 5-1 (g). During the deployment of OBS 001 this station did not provide any usable data, and is therefore disregarded.

Similarly we take a look at the spectrograms with the same running time as OBS 002, the result can be seen in Figure 5-6. There is a clear difference between the stations close to the Red Sea coast (5-6 (a) (b) (c)) and the stations more inland (5-6 (d) (e) (f) (g)) visible

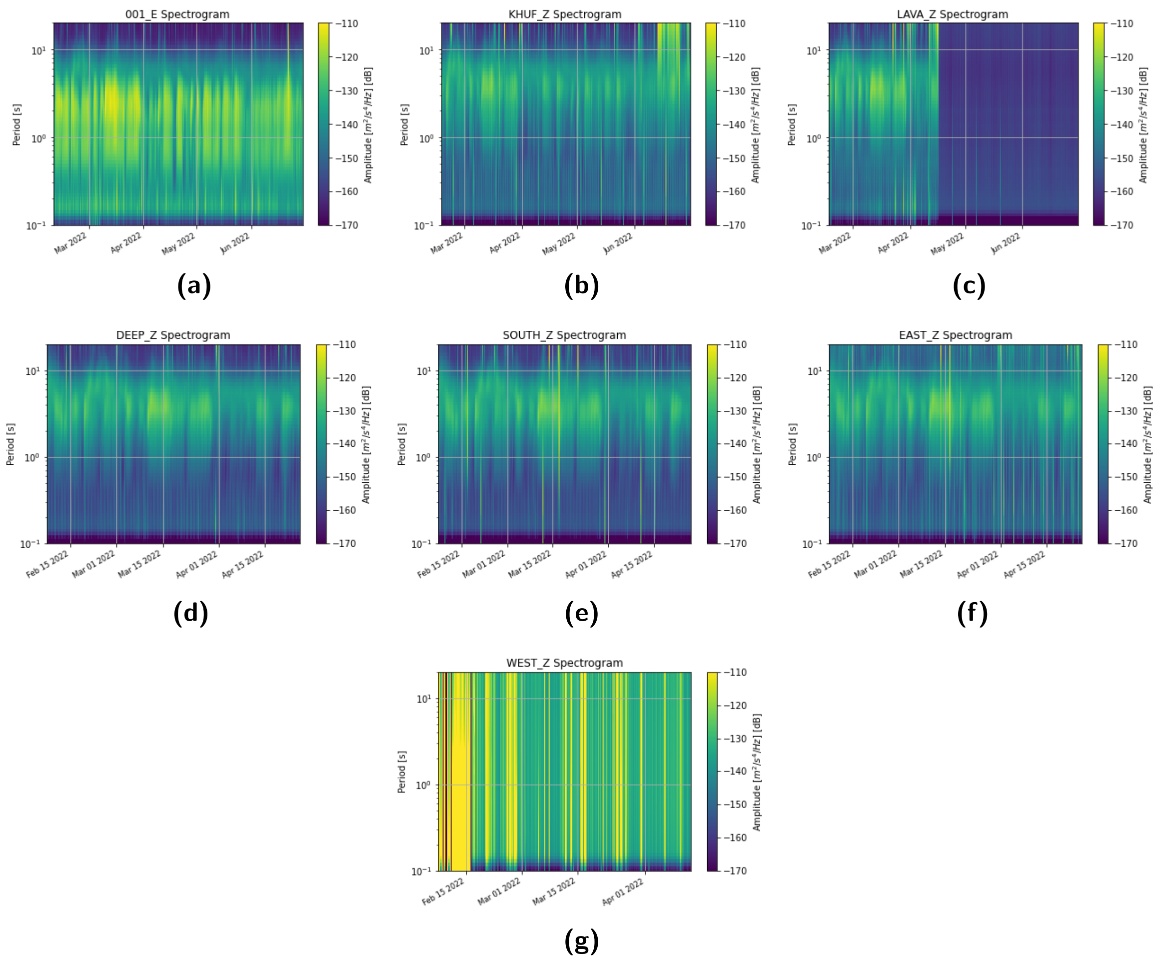


Figure 5-5: Spectrogram of stations, same running time as OBS 001. **(a)**: OBS 001. **(b)**: Land-station KHUF. **(c)**: Land-station LAVA. **(d)**: Land-station DEEP. **(e)**: Land-station SOUTH. **(f)**: Land-station EAST. **(g)**: Land-station WEST.

in the lower periods. An extra noise band can be seen around 0.5 seconds. Most likely a result of high-frequency microseisms generated by the Red Sea. The relatively low period of this noise band makes that the signal is dissipated once it arrives at the stations further inland. An irregularity can be seen at land-station WEST, Figure 5-6 (g). This has to do with an unknown instrumental failure which kept existing throughout the deployment time of OBS 001 .Afterwards the seismometer has been replaced to measure reliable data again in the future, The faulty data from this station is excluded from further computations.

5-1-3 OBS rotation

Both ocean bottom seismometers are installed on board of an ocean bottom lander, which contains a variety of sensors. This is discussed more extensively in chapter 2. The landers are slowly lowered from a ship in the Red Sea using a heavy lift system. Usually a three-component seismic station is installed with the horizontal 'N' component aligned with the

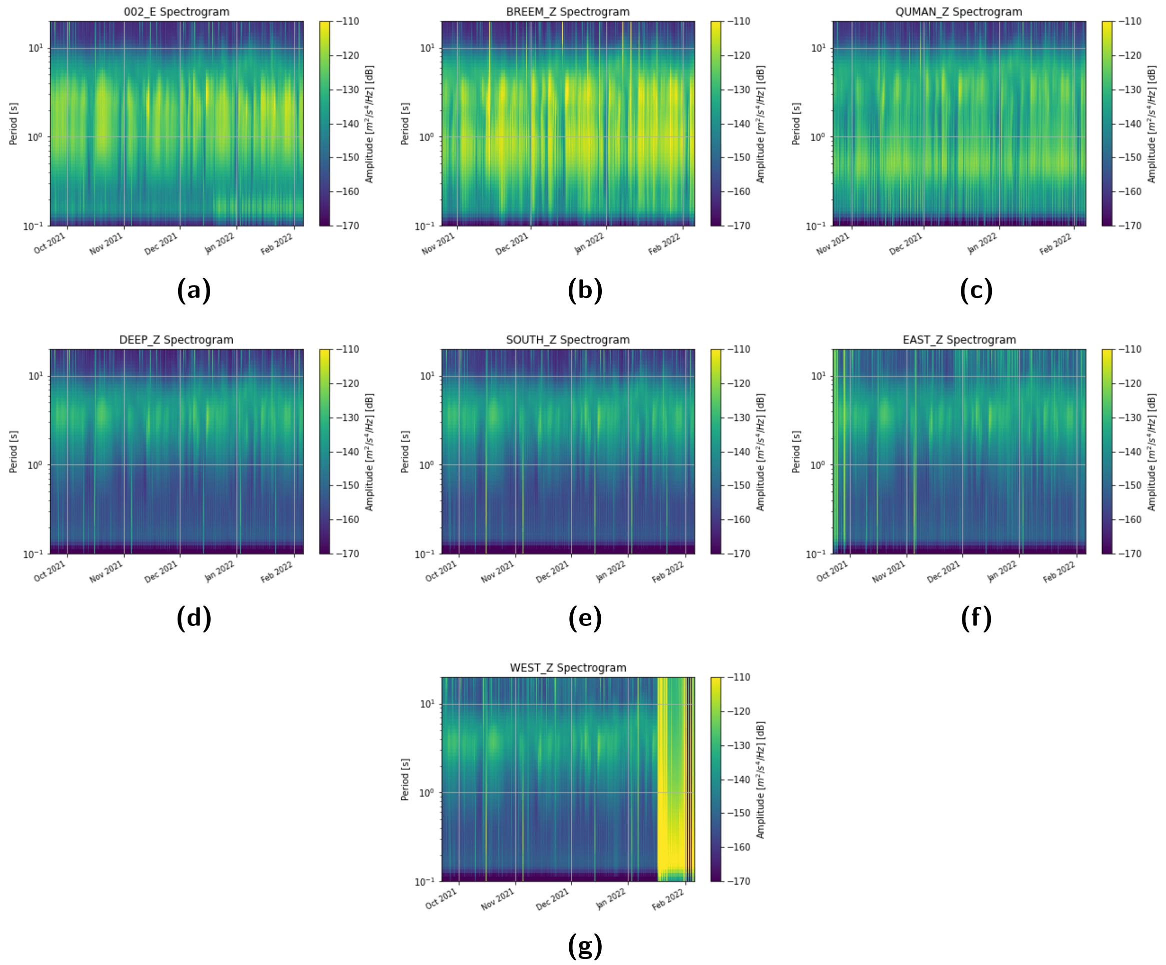


Figure 5-6: Spectrogram of stations, same running time as OBS 002. **(a)**: OBS 002. **(b)**: Island-station BREEM. **(c)**: Island-station QUMAN. **(d)**: Land-station DEEP. **(e)**: Land-station SOUTH. **(f)**: Land-station EAST. **(g)**: Land-station WEST

true North. In this case the OBSs were installed at roughly 1000 meters depth by an ROV (remotely operated underwater vehicle). During installation the alignment of the horizontal 'N' component with the true North was not considered. To use the horizontal components in a meaningful way this orientation is crucial to know [Stachnik et al., 2012]. Therefore, several techniques were used to determine the true orientation, the most promising is shown in this section for both OBS 001 and 002 and is explained in more detail in theory section 4-1-3.

This method utilizes the vertical component and iteratively rotated Hilbert transformed horizontal component 'N' of the OBS. Both signals are cross-correlated for a selection of teleseismic events (magnitude > 6.0) that occurred during the deployment time of the OBS. The maximum correlation corresponds to the alignment of the horizontal component 'N' with the radial direction (i.e. pointing towards the epicenter of the teleseismic event). The rotation of horizontal component 'N', for which maximum correlation occurred for each teleseismic event is shown in Figure 5-7(a) for OBS '001'. The rotation of horizontal component 'N' is corrected for the earthquake location, i.e. every blue dot indicates the amount of degrees the horizontal 'N' component is rotated from the true North according to that specific teleseismic event. The red line shows the median of all events combined. The direction shown on the polar plot shows the orientation of the used horizontal component 'N based on the median.

The result shows a clear pattern, with some outliers. These outliers are reduced by selecting only teleseismic events with an event depth of less than 100 kilometer and excluding normalized correlations lower than 0.4 as suggested by [Stachnik et al., 2012]. After this selection the observations within a 95% confidence interval of the circular mean are retained. The result after narrowing down is shown in Figure 5-7(b), again with the median indicated by the red line and the direction shown on the polar plot. This direction is used as the orientation of the 'N' component of OBS 001.

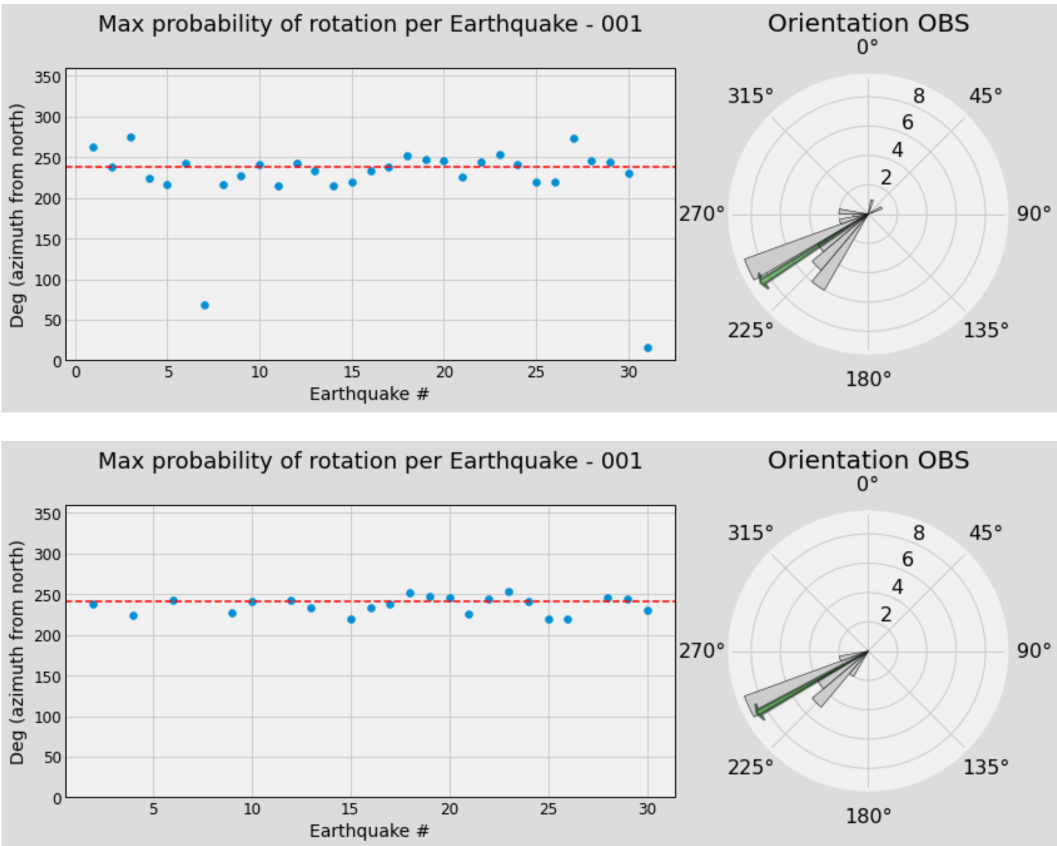


Figure 5-7: Orientation OBS 001 - Every blue dot indicates the heading of horizontal component 'N' according to that earthquake, The red line indicates the mean result from all earthquakes. The polar histogram binned the individual earthquake results and indicates the mean result from all earthquakes by the green arrow. Top shows the result for all earthquakes. Bottom shows the result after filtering the earthquakes.

The same procedure is applied to OBS 002. Since this station has a different deployment period, different teleseismic events had to be selected. We see that during the deployment of this OBS significantly more teleseismic events occurred, this should strengthen the final result obtained with this method. The results are shown in Figure 5-8, this is again corrected for the location of the teleseismic events. For this OBS we can clearly identify a concentration of events around 290 deg. However, significantly more outliers are present for this OBS. The exact reason for this is unknown, but the most obvious explanation would be the different selection of teleseismic events being used. Another possiblity is a movement of the OBS during deployment, which is plausible since the used OBS is very light in weight. To narrow down, earthquakes are excluded in the same way described for OBS 001. The final result can be seen in Figure 5-2 (b).

To test this method the same procedure is applied to different land-stations in the network. During the installation of these land-stations it is made sure that the horizontal 'N' component is properly aligned with the true North. We can test the method by assuming the horizontal 'N' component orientation is unknown. Theoretically this should result in a median orientation of 0 deg after correction for the earthquake locations. Both subsets of earthquakes were tested by selecting stations with different deployment times. For the stations tested the median was in all cases around 0 deg, confirming this method is a robust way to determine the orientation of the horizontal OBS components.

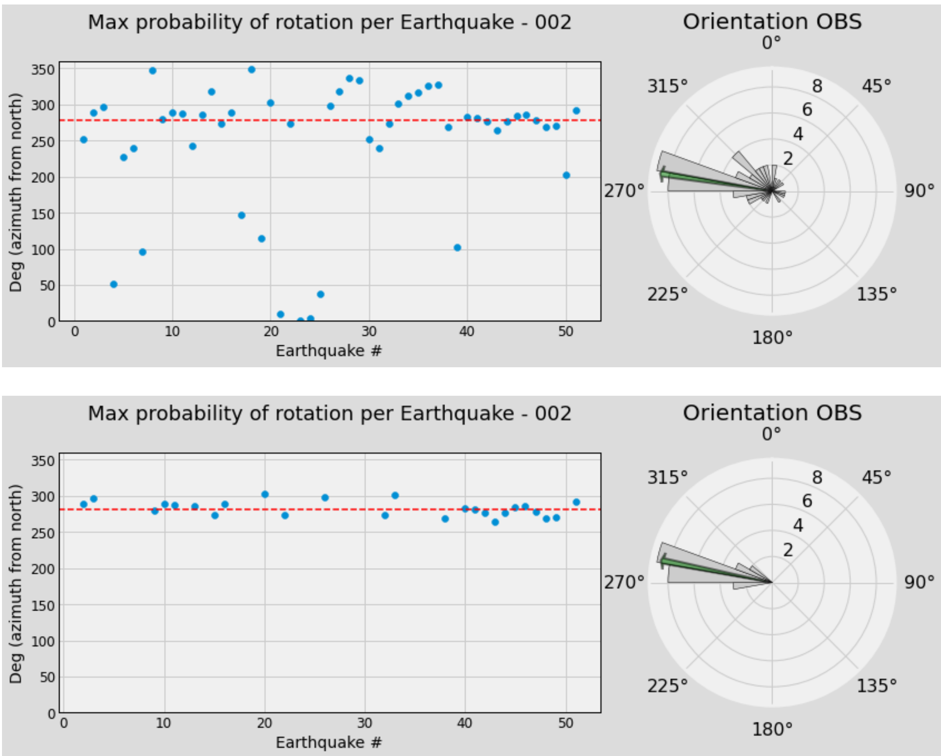


Figure 5-8: Orientation OBS 002 - Every blue dot indicates the heading of component 'N' according to that earthquake, The red line indicates the mean result from all earthquakes. The polar histogram binned the individual earthquake results and indicates the mean result from all earthquakes by the green arrow. Top shows the result for all earthquakes. Bottom shows the result after filtering the earthquakes.

5-1-4 Cross-correlations

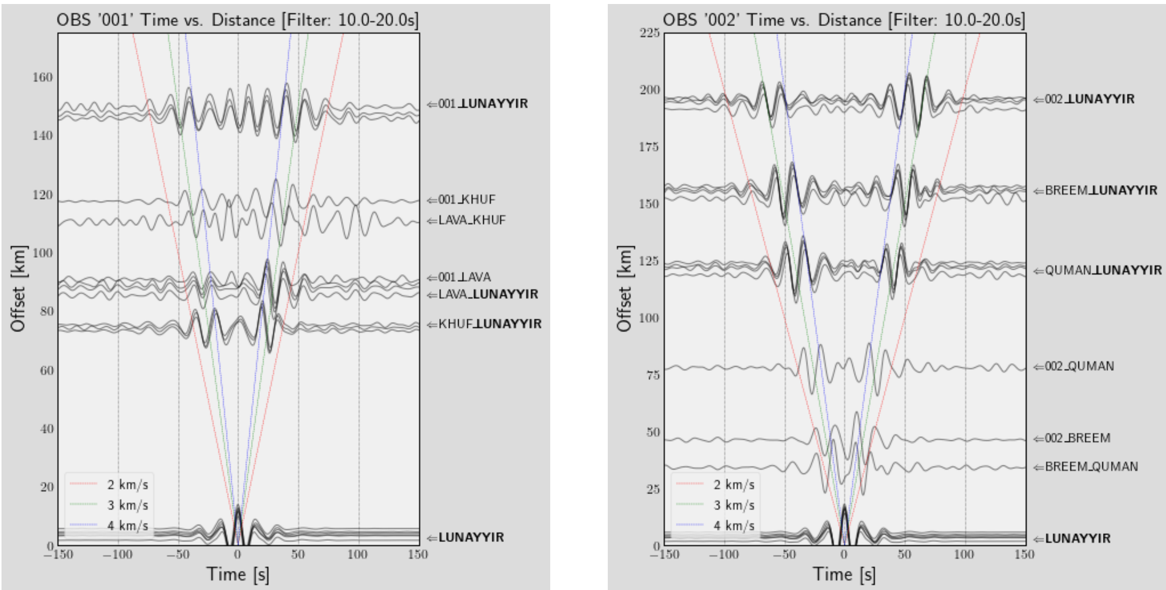
To detect and correct the timing errors of both OBSs the temporal stability and time symmetry of the retrieved interferometric responses is utilized, more about this is explained in chapter 4-3 and the paper [Naranjo et al., 2021]. For the code to work interferometric surface wave responses with a sufficient signal-to-noise ratio are required. To achieve this we cross-correlate the different signals based on the deployment time of both OBSs. Since there is no overlap in deployment time of the OBSs this is done in two group, each including stations with a deployment time corresponding to the OBS. The cross-correlated result is filtered with an adequate bandpass filter, providing the highest SNR for the interferometric responses. For this seismic network the options in filters were very limited, only cross-correlations filtered between 10 and 20 seconds provided a sufficient interferometric response. Since this filter retains relatively long periods difficulties arise concerning the distance between stations. A minimum of 2 wavelengths in between two stations is desired to maintain data quality and prevent the causal and acausal surface waves from merging [Bensen et al., 2007, Naranjo et al., 2021]. With these long wavelengths we have to exclude nearby stations according to this rule of thumb. Many attempts have been made to filter for shorter periods, however this does not give any satisfactory interferometric responses. Ultimatelym we made the decision to work with the relatively long periods between 10 and 20 seconds, despite the downsides. More about why this could be the case is explained in the Discussion.

The receiver pairs shown for each cross-correlation are ordered in a way that the receiver of the acausal (left) signal is put first. The receiver of the causal signal is put second. e.g. 001-KHUF indicates that OBS 001 received the acausal signal and land-station KHUF received the causal signal. Additionally the stations are grouped in a way that the station located closest to the Red Sea is shown first and the station land inward is shown second. In all shown examples the causal signal is significantly stronger than the acausal signal. This means that more seismic noise is travelling from the Red Sea towards land than in the opposite direction. This is according to expectation when the origin of seismic noise in the microseism band we use is considered to be the sea, more about this in section 3-2-2.

The cross-correlations for both station groups can be seen in Figure 5-9. The cross-correlations are sorted along the y-axis based on receiver offset, all cross-correlations are bandpass-filtered between 10 and 20 seconds. Reference lines have been added to indicate the velocity for 2, 3 and 4 kilometers/s. Figure 5-9 (a) shows the cross-correlations for stations with a similar deployment time as OBS 001. The result is poor, although some surface waves can be distinguished, the SNR remains low. Mainly the cross-correlations between land-stations (KHUF, LAVA and the 'LUNAYYIR' group) show surface waves, albeit with very low SNR. However, the OCloC algorithm requires interferometric surface wave responses for receiver pairs which include the OBS.

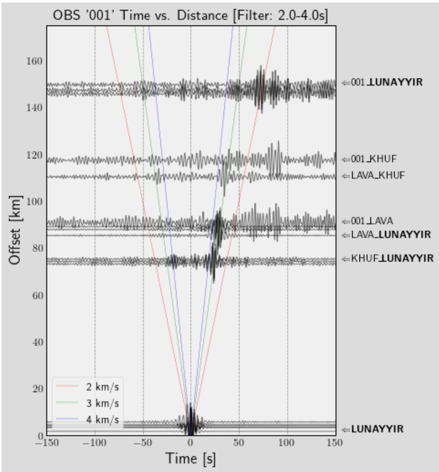
The cross-correlations for stations with a similar deployment time as OBS 002 can be seen in Figure 5-9 (b). The cross-correlations are again bandpass-filtered between 10 and 20 seconds. This time we can identify clear interferometric responses for all receiver pairs, including the pairs with the OBS. For the most proximate receiver pairs the causal and acausal responses are merging, which is to be expected [Bensen et al., 2007]. Overall the interferometric responses seem to be much more clear compared to the result in 5-9 (a)

Some examples of cross-correlation functions filtered for lower periods [2-4s] are shown in Figure 5-10. In the top row ((a) and (b)) we see the cross-correlations of all stations grouped with the same deployment time as OBS 001 and 002 respectively. In the bottom row ((c) and (d)) we see the same groups, but now only the cross-correlation functions of land-stations are shown, i.e. the OBSs are omitted. Overall the results are very noisy and most receiver pairs do not show a clear surface wave response. However if we look at the results for stations with an equal timespan to OBS 001 in Figure 5-10 (c), we clearly see surface waves at the causal sides. This indicates that enough energy in this specific period band travelled land inwards, originating from the direction of the Red Sea. There can be several reasons why receiver pairs with the OBS do not show these clear surface waves. One could be the longer distance, since low period noise dissipates faster than high period noise [McNamara and Buland, 2004]. Another reason can be the coupling of the OBS, since the Red Sea floor at location of deployment was very loose and soft, more about this is explained in the Discussion.

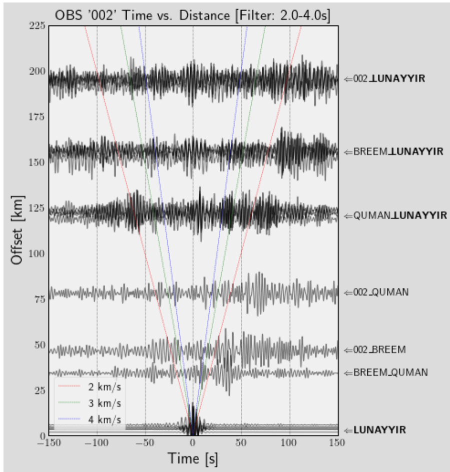


(a) **(b)**

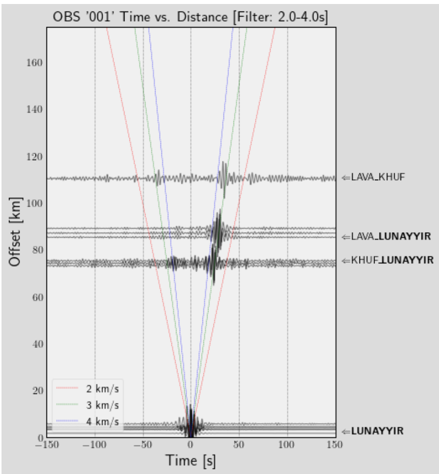
Figure 5-9: Cross-correlations of station pairs sorted along the y-axis based on Offset (km). **(a)** shows station pairs with the same deployment time as OBS 001. **(b)** shows station pairs with the same deployment time as OBS 002. **LUNAYYIR** indicates the stations installed in close vicinity of each other at 'hyatt lunayyir', these include [WEST, SOUTH, EAST, DEEP]



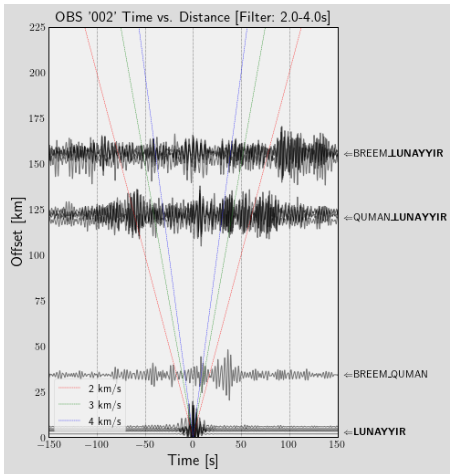
(a)



(b)



(c)



(d)

Figure 5-10: Lower period cross-correlation results. (a) and (b) show the results of all stations groups of OBS 001 and 002 respectively. (c) and (d) only show the land-stations i.e. the pairs including the OBSs are omitted.

5-2 Time shift correction

To correct the clock error that might be present in the OBS data, we use the OCloC (OBS Clock Correction) package. In this chapter we show the results of this process. First we take a look at how eligible lapse cross-correlations were selected. We do this based on the signal-to-noise ratio and distance between stations. We look at the SNR changes caused by selecting different timespans and overlaps for the lapse cross-correlations. The final selection of lapse cross-correlation is inverted and the results are analysed and discussed. A more extensive explanation on how the OCloC package works can be found in chapter 4-3.

5-2-1 Lapse cross-correlation selection

The OCloC package uses the temporal stability of cross-correlations and time symmetry of the retrieved interferometric responses as described in Chapter 4-3. To achieve this, cross-correlations between the OBSs that require a correction, and land-stations with synchronized timing, are computed. Each lapse cross-correlation needs to cover a set time window, where an overlap is possible if desired.

This means that there are two parameters that need to be determined to construct the most suitable lapse cross-correlations for the OCloC code: timespan and overlap. Each lapse cross-correlation still needs to have a distinctive causal and acausal response with a sufficient signal-to-noise ratio (SNR). Different timespans are tested, for each timespan the SNR of the causal and acausal interferometric signal is determined. The results are shown as a box plot in Figure 5-11. The box plots visualizes the SNR individually for the causal and acausal response, as well as the combination of both. It does this for both OBSs; 001 and 002. Since OBS 001 does not have many stations with a sufficiently overlapping deployment time, most SNR contributions are from the pair KHUF-001, which makes the spread relatively small compared to the SNR for OBS 002.

As expected, we clearly see an increase in SNR when more days are used for the lapse cross-correlation. This suggests that a maximum timespan is desired to retain a high SNR. However, as Figure 5-12 shows, an increase in timespan often leads to a decrease in maximum mid-time offset. Since clock drift happens at a rate in the order of milliseconds per day [Hable et al., 2018], a long enough mid-time offset is desirable to measure a clock error big enough to be out of the error margin. The causal responses show a significantly higher SNR for both OBSs, which can be explained by the high concentration of microseism noise sources at the Red Sea side of the network (section 3-2-2).

Figure 5-12 shows some examples of lapse correlation possibilities for the station pair QUMAN-002. It illustrates the importance of selecting the right timespan and overlap, since it can heavily impact the SNR, amount of data being used and maximum mid-time offset. One way to maximize SNR while retaining a sufficient maximum mid-time offset is including an overlap in the data. This figure suggests that maximizing overlap should improve the results, since a longer timespan can be used to achieve a higher mid-time offset. However, small timeshifts in between lapse correlations makes it challenging for the algorithm to detect the timeshift in between lapse correlations. So a sufficient increase in time in between lapse correlations should be pursued.

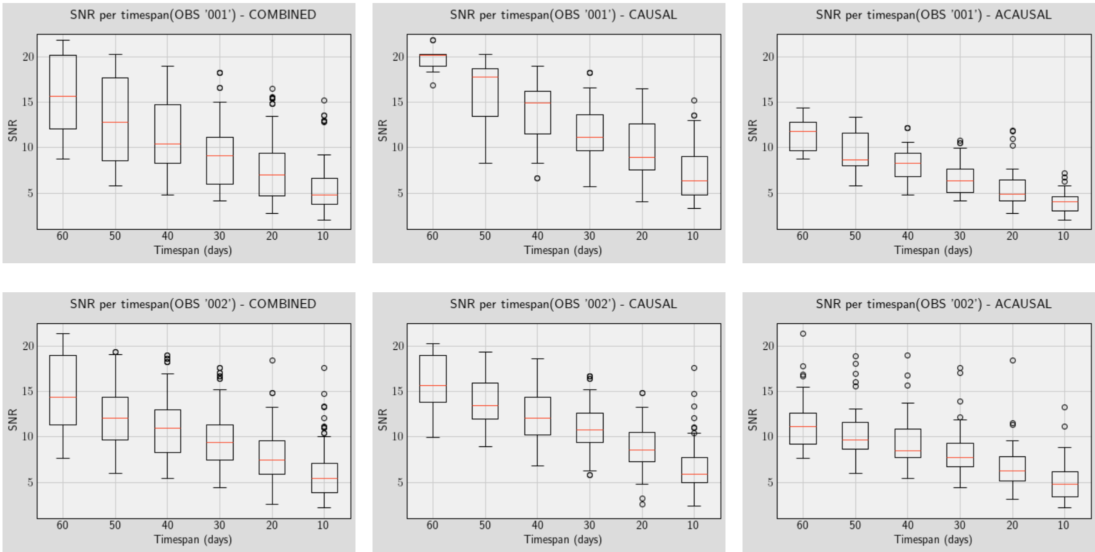


Figure 5-11: Signal-to-noise ratio for different timespans and overlaps for OBS 001 and OBS 002, the red line indicates the median, the box is drawn in between the 25th and 75th percentiles, the whiskers extend to 150% to both sides. Everything outside is indicated as outlier/flier

5-2-2 Clock drift estimation

To see if a clock drift is present the data is visually checked. This is done by zooming in on the lapse cross-correlations of a specific station pair, if a clock drift is present a shift in should be visible. A selection of four different station pairs can be seen in Figure 5-13, for each station pair a timespan of 40 days with an overlap of 20 days is selected. Bear in mind that the data obtained for OBS 001 does not seem to provide clear surface waves. For these pairs the clock drift in both situations seems to be in the opposite direction, which is a clear indication something is wrong. However, for the station pair SOUTH_001 we only have two lapse cross-correlations, which does not provide a solid backing. For the station pairs with OBS 002 we clearly see a shift in the same direction. This indicates that a clock drift is present for OBS 002.

The decision is made to focus on the lapse cross-correlations with a lapse time of 30 and 40 days. This is based on the deployment time of the OBSs in combination with the SNR visualized in Figure 5-11. A higher lapse time would most likely increase the SNR, however for some station pairs the mutual deployment time would be not sufficient to construct multiple lapse cross-correlations. An attempt is made to compute the clock drift of both OBSs by inverting the data. This means that we also did an attempt to compute the clock drift for OBS 001, despite the poor interferometric responses found in section 5-2-1 and contradicting clock drifts found in figure 5-13. Since these crucial confirmations are missing it is important to realize that the obtained clock drifts for OBS 001 are not reliable. OBS 002 did show clear interferometric responses and a matching clock drift, this makes the results much more reliable. The results for the clock drift (a) and incurred timing error (b) can be found in Table 5-1. We see that the values for OBS 002 are fairly consistent, with some outliers (30_10 and 40_10). This is easier to observe by looking at the scatter plot in Figure 5-14, where the

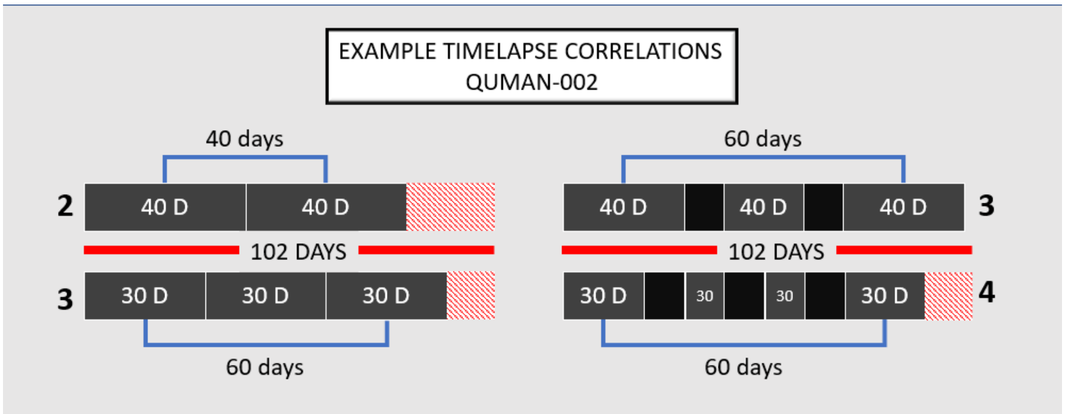


Figure 5-12: Lapse time for QUMAN-002, the red line shows the overlap in deployment time of OBS 002 and land-station QUMAN (102 days), the grey blocks indicate possible lapse cross-correlations, the black blocks indicate possible overlap. Left shows no offset, right shows an offset of 10 days. We see that the choice of time span and overlap strongly influences the maximum time offset between lapse cross-correlations.

results for OBS 002 are shown in red. The results for OBS 001 are more variable, which was to be expected considering the poor quality of the interferometric responses that was used for the OCloC package. This is also visualized in Figure 5-14 in the color blue.

Timespan [days]	Overlap [days]	Sensor	Lapse-correlations	a: Clock drift [s/day]	b: Incurred timing error at t=0 [s]	Timing error at recovery [s]
30	0		8	-0,003441	0,056482	-0,432117
30	10		8	x	x	x
30	20		18	-0,005887	0,157753	-0,678221
40	0	001	6	-0,007841	0,121603	-0,991828
40	10		7	-0,007264	0,087961	-0,943481
40	20		11	-0,00917	0,310187	-0,991969
40	30		19	-0,010853	0,445831	-1,095301
30	0		15	-0,004307	0,319788	-0,266016
30	10		23	-0,000345	0,000647	-0,046306
30	20		42	-0,003981	0,281936	-0,259481
40	0	002	7	x	x	x
40	10		15	-0,006157	0,379092	-0,458308
40	20		20	-0,004316	0,291563	-0,295355
40	30		41	-0,002828	0,213107	-0,171525

Table 5-1: Results from OCloC package for OBS 001 and 002

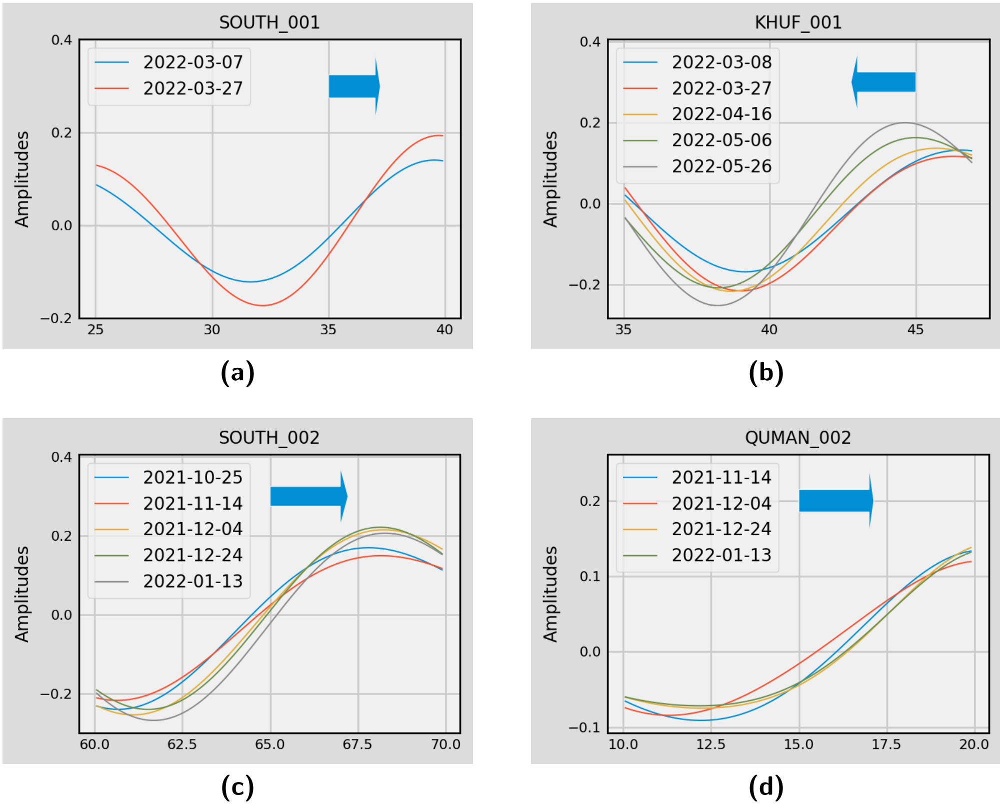


Figure 5-13: A visual check if a clock drift is present in between lapse cross-correlations. (a) and (b) show different station pairs for OBS '001', (c) and (d) show different station pairs for OBS '002'

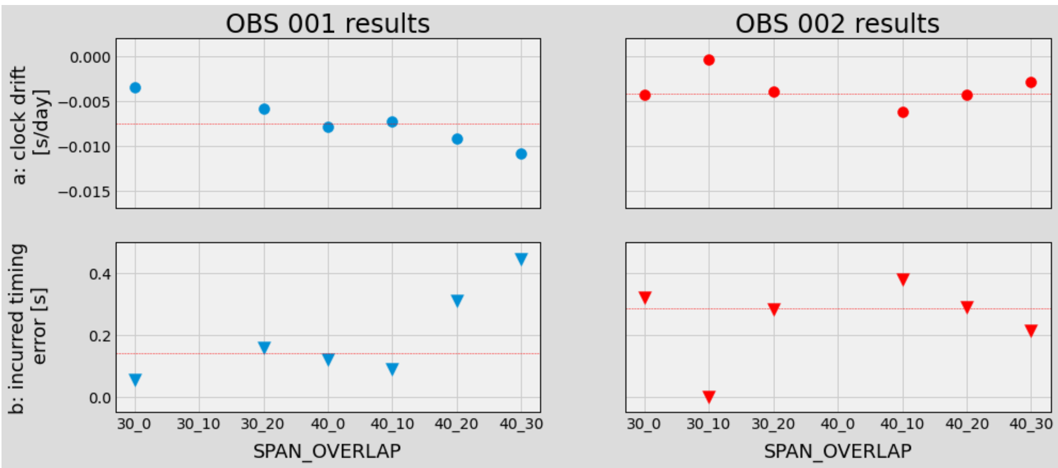


Figure 5-14: Results from OCloC package for OBS 001 (in BLUE) on the left, and 002 (in RED) on the right. The red line indicates the median value

Chapter 6

Discussion

In the discussion we will mainly reflect on our ability to detect and correct clock errors with the OCloC package. For completeness we also discuss the quality of the data used, and what could be reasons for observed quality issues.

6-1 OBS coupling

We installed two ocean bottom seismometers (OBSs) on the bottom of the Red Sea, at a depth of roughly 1000 meters. Both OBSs are part of an ocean bottom lander, which has many different sensors attached. We installed the OBSs with a remotely operated vehicle (ROV), the robotic arms are used to place the sensor in the vicinity of the lander. Since a flat seabed is required for the lander to be stable, the surface in both situations consisted of very loose sediments, in contrast to bedrock which is the preferred installation medium.

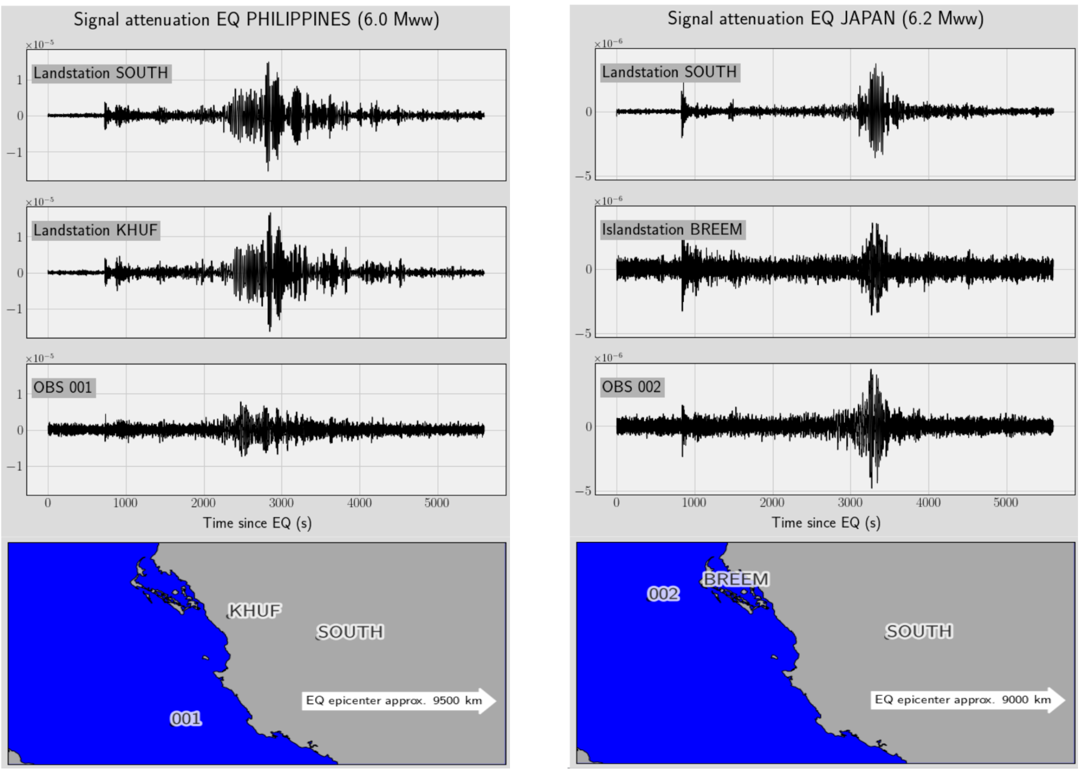
According to studies done with ocean bottom seismometers, the placement of the station on loose saturated sediments could lead to coupling problems with signal distortion and high levels of noise [Mangano et al., 2011]. Resonant amplification and attenuation of short periods also occur [Sutton et al., 1981, Bialas and Flueh, 1999]. The horizontal components of both OBSs showed strong noise levels, especially for long periods. This led to a significant difference in quality between horizontal and vertical components. According to studies on OBS data quality, a quality difference between vertical and horizontal components is common and mainly caused by the interaction of water currents, inducing a slight tilt on the OBS [Lindner et al., 2017].

The OBS type we used in this research is very lightweight and compact to make it easily fit on the ocean bottom lander. This makes the OBS easier to handle and orientate on the seafloor, but also increases its vulnerability to movement by ocean currents or even wildlife. Furthermore is the overall response strongly dependent on the mass of the OBS [Sutton et al., 1981]. The data used to test OCloC included 83 seismic stations, of which 23 OBSs, installed in Iceland for the IMAGE (Integrated Methods for Advanced Geothermal Exploration) project [Weemstra et al., 2021]. These OBSs were standalone, installed with the sole purpose to collect seismic data. This means that the whole station included the battery, data logger and sensor installed on a frame with flotation devices, making it bulkier and increasing its mass. This is beneficiary for the stability and coupling of the sensor on the seabed. As shown by [Sutton et al., 1981] a lower mass factor leads to a lower coupling frequency, which leads to the attenuation of higher frequencies. A lower sediment density and rigidity modulus also lowers the coupling frequency as can be seen in Equation 6-1 (adopted from [Sutton et al., 1981]). Additionally low frequencies can be amplified due to coupling resonances.

$$\omega_c^2 = (a_s r + b_s r^2)/(M_1 + c_1 {r'}^3) \tag{6-1}$$

where:

ω_c = natural angular frequency of the OBS coupling
r^2 = bearing area
${r'}^3$ = volume of sediment and water entrained with the OBS
c_1 = depends on OBS configuration
M_1 = mass of OBS
a_s = proportional to sediment rigidity modulus
b_1 = depends upon sediment density



(a) (b)

Figure 6-1: Attenuation of short period noise based on an earthquake example for OBS 001 and OBS 002 compared to land-stations (Lowpass filter 10s)

The low mass factor in combination with a coupling to loose sediments would explain why OBS-landstation pairs only showed interferometric surface waves for a relatively long period pass band (10-20 s) (Figure 5-9, since the lower periods would be highly attenuated. Another reason could be that no low period noise sources exist within the Fresnel zones of the station pair (chapter 3-2), however, when we look at the landstation-landstation pairs with a short period pass band (2-4 s) we clearly see a interferometric surface wave, albeit only on the causal side (Figure 5-10). We test this by looking at teleseismic events (Figure 6-1), we clearly see that the signal is highly attenuated at both OBS. One could argue that this is because the OBS is located further from the earthquake, however since the epicenter of both events is very far (9500 and 7000 km respectively) the relatively short distance in between the stations should not matter. Another example of the overall attenuation can be seen in the Appendix, this example is filtered below 10 seconds: Figure A-1.

Additionally, we found that the island stations 'BREEM' and 'QUMAN' show relatively high noise levels. Part of this can be explained due to the vicinity of the Red Sea, these high noise levels are found in the microseism band and mainly between 1 and 4 seconds as shown in section 5-1-1. However also for shorter periods below 1 second we see a high noise level for both island stations. The reason for this is found in the poor material below the station which might not be connected to the bedrock. The material contains many fractures and pores which causes the signal to scatter [Dainty, 1996].

6-2 Data availability

The quantity and quality of the data drastically influences the ability to detect and correct a clock drift for an OBS. The aim of this research was to stretch the ability of the OCloC package and see if it works with a less comprehensive dataset. However, the data obtained was less than anticipated as shown in Figure 2-2. Reasons for this vary from instrument failure to difficulty in reaching the location to obtain the data. The available data only covered 133 days for OBS 001 and 137 days for OBS 002. The OBSs do not have any overlap in deployment time, and only part of their deployment time has overlap with other stations. This available time has to be split up in smaller lapse cross-correlations to make the time-lapse method possible. An important trade-off between SNR and maximum lapse time shift has to be made: When longer lapse cross-correlations are used, the SNR will improve (as is shown in Figure 5-11). However, the maximum lapse time shift decreases, making the clock error smaller and harder to detect [Naranjo et al., 2021]. OCloC was designed for extensive networks with many OBSs. We see that OCloC struggles to find a reliable clock error with the small network we used in this research, where we only have two OBSs with a limited amount of complementing land-stations.

Clocks used for OBS are designed to be as accurate as possible. This means that the clock drift will only be in the order of milliseconds per day. A sufficient recording interval is required to reliably estimate the clock drift rates, taking into account the standard deviation [Hable et al., 2018]. More possible station pairs with overlap in deployment time decreases the standard deviation. Since the possibility of pairs for both OBS stations are limited the standard deviation in this research is significant. The overall standard deviation for a station is given by the formula [Hable et al., 2018]:

$$\sigma = \frac{\sigma_i}{\sqrt{n}} \tag{6-2}$$

Where σ is the standard deviation for the averaged clock error. σ_i is the standard deviation for the clock error resulting from one station pair and n is the number of possible station pairs. This equation tells us that the standard deviation is reduced by the square root of the possible station pairs.

6-3 Interferometric responses

The interferometric responses are essential for determining and correcting the clock drift using the OCloC package. The data we used for this project only showed a satisfactory interferometric response for long periods (10-20 seconds). This causes some issues concerning the distance between stations. Such long period waveforms means we are looking at wavelengths of around 40 kilometers, depending on the wavetype and assuming a velocity of 4km/s. These long waves require a large distance in between stations to make sure at least 2 wavelengths fit. This means that stations close to the OBS cannot be used and are excluded, which is a big loss regarding the available stations are already limited.

Different reasons can be dedicated to the interferometric response only showing a satisfactory result for long periods (10-20 s). One is the coupling of the OBS, which is described earlier on in the discussion. A unfavorable coupling on loose saturated soil can cause short period noise to attenuate faster. Another possible reason can be found in the noise sources available around the Arabian peninsula. When we look at the PPSD and spectrogram results for this method, long period noise seems to be dominant (section 4-1-1 and 5-1-2). It is possible that no noise sources emitting shorter period noise exist in the fresnel zones of the station pairs.

Other studies like [de Ridder and Biondi, 2010, Civilini et al., 2019] were able to find a surface wave on the Arabian Peninsula for shorter periods (0.14-1 s and 5-12s respectively). However, to achieve this only land stations were used. As mentioned by [Carrière and Gerstoft, 2013], lower frequencies tend to work better for interferometry when OBS are used due to the nature of source distributions. When we look at the cross-correlations for only land-stations, filtered for shorter periods in Figure 5-10, we can clearly see an interferometric response, albeit only at the causal side.

Another issue was the relatively low SNR of the obtained surface waves. in [Naranjo et al., 2021] the OCloC package is tested with a comprehensive dataset. The SNR is sufficient to have a SNR threshold of 30 (i.e. all interferometric responses with a SNR below 30 are omitted). If the same SNR threshold would be used on this dataset no interferometric responses would remain. Because of that we had to drop the SNR threshold significantly to 5. The quality of the clock drift measurement strongly depends on SNR. If the SNR is too low the algorithm can struggle to determine the arrival times of the surface waves [Naranjo et al., 2021]. The acausal SNR is outstandingly lower than the causal SNR. This can be assigned to the noise source illumination pattern, with the Red Sea on one side providing a high concentration of microseism sources.

6-4 Noise source illumination pattern

The seismic network used for this research is located on the Red Sea coast of Saudi Arabia. The periods used to to construct the interferometric responses $[10 - 20s]$ are within the microseism band (1-30 s) [nak, 2019]. The main source of ambient noise within this band are pressure fluctuations in the oceans, as described in Section Ambient noise. This means that a major noise source concentration is located at one side of the seismic network, which results in an uneven noise source illumination pattern. This uneven distribution shows in the interferometric responses as a difference in amplitude between the causal and acausal side.

6-5 OCloC

Working with OCloC has been very convenient, the package is clearly designed with user friendliness in mind. However, this can also lead to carelessness when using the package. Some steps involve visual checks, for instance the ability of the software to detect the surface waves. If not carried out with care the package might still provide an answer within the reasonable range of clock errors, which can be misleading. This occurred for OBS 001 in this research. It was clear that the interferometric responses would not be sufficient to reliably determine the clock error. Additionally the visual check for clock drifts showed contradicting results. Despite this we fed the data into OCloC to see analyse the result. Surprisingly the package provided a satisfactory answer. If the background of the data was less known this might be assumed as the true clock error.

A suggestion to tackle this might be using different overlaps, while retaining the same timespan of lapse cross-correlations. Theoretically these should provide similar results; the equal length of the lapse correlations should provide a similar SNR, and the maximum timeshift between lapse cross-correlations should also roughly correspond. Only the amount of available lapse cross-correlations might differ, this should not influence the final result when enough data is available.

We used different overlaps in this research. For OBS 001 they generated a wide variety of results, which could be an indication something is wrong with the data fed into the OCloC package. For OBS 002 the results were significantly closer, albeit still not within the limit that can be expected when the method is successful.

Chapter 7

Conclusions

The primary objective of this research was to determine the clock error for two OBS stations deployed in the Red Sea of Saudi Arabia. To do this we make use of the new OCloC package. Additionally we wanted to obtain valuable insights on how the OCloC package functions when a less extensive dataset is used. As secondary objectives we did an extensive quality control and data pre-processing with the aim to prepare eligible lapse cross-correlations for the OCloC package. To do this in an orderly way we present a roadmap, this roadmap is made applicable to OBS data by adding some extra steps to tackle typical OBS challenges like the orientation.

The obtained results for OBS 001 and 002 are not equally reliable, the used lapse cross-correlations for OBS 001 did not show clear surface wave responses, which is essential for OCloC to work properly. As a result we did not see constant clock drift results for different lapse cross-correlations with a differing timespan and/or overlap, which is a strong indication the result is not reliable. The lapse cross-correlations we used for OBS 002 did show clear surface wave responses, albeit with a relatively low SNR compared to the dataset used in [Naranjo et al., 2021]. The clock drift results for OBS 002 show significantly more consistency when we used different lapse cross-correlations with a differing timespan and/or overlap.

We present several possible reasons for these varying results. First and foremost are the low signal-to-noise ratios of the interferometric surface wave responses, which are used by the OCloC package. The low SNR has several reasons, first we have an anisotropic noise source illumination pattern, i.e. the Red Sea provides a high microseism source concentration on one side of the seismic network. Second, we see that a poor OBS coupling to the loose sediments on the Red Sea bottom is present. This results in a strong attenuation of short period noise and a possible amplification of long period noise. This was supported by the fact that interferometric surface wave responses for short periods could only be retrieved between two land-stations. Attempts were made for station pairs that included an OBS, but they were not successful.

The period range for which we obtained a decent SNR for the interferometric surface wave responses is between 10 and 20 seconds. The reason for this is most likely the coupling issues of the OBS discussed before. Because a long period leads to a long wavelength, and we want to keep 2 wavelengths in between a station pair at all times, stations with a low inter station distance had to be omitted.

Usually OBSs that are deployed with the aim to measure ambient seismic noise record for a longer time. This would most certainly increase the accuracy of OCloC, since a larger clock error would be easier to measure. For the small deployment time we tested in this research the clock error can be too small to be recovered with a low uncertainty.

For future studies it would be interesting to see how OCloC functions when more land-stations are deployed that cover the full deployment time of the OBSs. In this research this was often not the case which reduced the amount of possible lapse cross-correlation for each OBS. This seemed to be one of the main issues regarding the clock error estimation. Something else that could be considered is using a different type of OBS, the OBSs used in this research were very compact and lightweight, this made handling them convenient, but also increased vulnerability to movements and a bad coupling to the sea bed, which could cause damping of short period noise. It would be interesting to see a comparison between different OBS types with different weights on a similar sea bed.

In summary, our main findings are:

- The clock drift in OBS 001 could not be determined because of the combination between low quality data, short deployment and lack of simultaneously deployed land-stations.
- Conversely, although the deployment of the OBS 002 was also short, the clock drift could be retrieved because of better quality of the dataset and more simultaneously deployed land-stations.
- The noise source illumination pattern of the used period pass band is strongly anisotropic, with the Red Sea providing a high (microseism) source concentration on one side of the network. This complicated the clock error estimation.
- The poor coupling of the OBSs to the loose ocean bottom sediment caused a strong attenuation of short period noise. This was primarily caused by the low mass of the used OBSs.
- Although working with OCloC is very convenient, a clear check is missing (besides suggested visual checks). The use of different time spans and/or overlaps could provide this.

Overall we can conclude that the new OCloC package is a promising addition to obtain OBS clock error corrections. However, there are still some challenges that require further study to see if OCloC is able to work around them, these are primarily related to anisotropic noise source patterns and short deployment times.

Bibliography

[nak, 2019] (2019). *Visualization of the Seismic Ambient Noise Spectrum*, page 1–29. Cambridge University Press.

[Aki, 1957] Aki, K. (1957). Space and time spectra of stationary stochastic waves, with special reference to microtremors. *Bulletin of the Earthquake Research Institute*, 35:415–456.

[Aldaajani and Furlong, 2022] Aldaajani, T. and Furlong, K. P. (2022). The makkah–madinah transform zone: a relic rift-to-rift continental transform formed during early arabia–nubia plate separation. *Geoscience Letters*, 9(1):1–12.

[Almalki et al., 2015] Almalki, K. A., Betts, P. G., and Ailleres, L. (2015). The red sea–50 years of geological and geophysical research. *Earth-Science Reviews*, 147:109–140.

[An et al., 2021] An, C., Cai, C., Zhou, L., and Yang, T. (2021). Characteristics of Low-Frequency Horizontal Noise of Ocean-Bottom Seismic Data. *Seismological Research Letters*, 93(1):257–267.

[Aster et al., 2008] Aster, R. C., McNamara, D. E., and Bromirski, P. D. (2008). Multidecadal climate-induced variability in microseisms. *Seismological Research Letters*, 79(2):194–202.

[Augustin et al., 2021] Augustin, N., Van der Zwan, F. M., Devey, C. W., and Brandsdóttir, B. (2021). 13 million years of seafloor spreading throughout the red sea basin. *Nature communications*, 12(1):1–10.

[Bellanger and Daguet, 1974] Bellanger, M. and Daguet, J. (1974). Tdm-fdm transmultiplexer: Digital polyphase and fft. *IEEE Transactions on Communications*, 22(9):1199–1205.

[Bensen et al., 2007] Bensen, G., Ritzwoller, M., Barmin, M., Levshin, A. L., Lin, F., Moschetti, M., Shapiro, N., and Yang, Y. (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophysical journal international*, 169(3):1239–1260.

[Bialas and Flueh, 1999] Bialas, J. and Flueh, E. R. (1999). Ocean bottom seismometers. *Sea Technol*, 40(4):41–46.

[Brigham and Morrow, 1967] Brigham, E. O. and Morrow, R. (1967). The fast fourier transform. *IEEE spectrum*, 4(12):63–70.

[Brillinger, 1993] Brillinger, D. R. (1993). Earthquake risk and insurance. *Environmetrics*, 4(1):1–21.

[Bromirski, 2009] Bromirski, P. D. (2009). Earth vibrations. *Science*, 324(5930):1026–1027.

[Brune and Oliver, 1959] Brune, J. N. and Oliver, J. (1959). The seismic noise of the earth's surface. *Bulletin of the Seismological Society of America*, 49(4):349–353.

[Carrière and Gerstoft, 2013] Carrière, O. and Gerstoft, P. (2013). Deep-water subsurface imaging using obs interferometry. *Geophysics*, 78(2):Q15–Q24.

[Chu and Gordon, 1998] Chu, D. and Gordon, R. G. (1998). Current plate motions across the Red Sea. *Geophysical Journal International*, 135(2):313–328.

[Civilini et al., 2019] Civilini, F., Mooney, W., Savage, M., Townend, J., and Zahran, H. (2019). Crustal imaging of northern harrat rahat, saudi arabia, from ambient noise tomography. *Geophysical Journal International*, 219(3):1532–1549.

[Claerbout, 1968] Claerbout, J. F. (1968). Synthesis of a layered medium from its acoustic transmission response. *Geophysics*, 33(2):264–269.

[Coleman, 1974] Coleman, R. G. (1974). Geologic background of the red sea. In *The geology of continental margins*, pages 743–751. Springer.

[Cooley and Tukey, 1965] Cooley, J. and Tukey, J. (1965). An algorithm for machine calculation of complex fourier series. mathematics of computing, reprinted 1972. *Digital signal processing. IEEE Press, New York, NY*, pages 223–227.

[Dainty, 1996] Dainty, A. M. (1996). The influence of seismic scattering on monitoring. In *Monitoring a Comprehensive Test Ban Treaty*, pages 663–688. Springer.

[de Ridder and Biondi, 2010] de Ridder, S. and Biondi, B. (2010). Low frequency passive seismic interferometry for land data. In *2010 SEG Annual Meeting*. OnePetro.

[D.G. Honegger, 2013] D.G. Honegger, D. W. (2013). Handbook of seismic risk analaysis and management of civil infrastructure systems.

[Draganov et al., 2007] Draganov, D., Wapenaar, K., Mulder, W., Singer, J., and Verdel, A. (2007). Retrieval of reflections from seismic background-noise measurements. *Geophysical Research Letters*, 34(4).

[Duennebier et al., 1987] Duennebier, F., Anderson, P., and Fryer, G. (1987). Azimuth determination of and from horizontal ocean bottom seismic sensors. *Journal of Geophysical Research: Solid Earth*, 92(B5):3567–3572.

[El-Isa and Shanti, 1989] El-Isa, Z. and Shanti, A. A. (1989). Seismicity and tectonics of the red sea and western arabia. *Geophysical Journal International*, 97(3):449–457.

[El-Quliti et al., 2016] El-Quliti, S. A., Al-Harbi, T. B. S., Al-Yami, M. B. S., Al-Ghamdi, A. B. M., Al-Shammari, M. B. M., et al. (2016). Assessment of main parameters of extreme earthquakes in red sea, west coast of saudi arabia. *Open Journal of Earthquake Research*, 5(02):122.

[Ewing and Vine, 1938] Ewing, M. and Vine, A. (1938). Deep-sea measurements without wires or cables. *Eos, Transactions American Geophysical Union*, 19(1):248–251.

[Frantti et al., 1962] Frantti, G., Willis, D., and Wilson, J. T. (1962). The spectrum of seismic noise. *Bulletin of the Seismological Society of America*, 52(1):113–121.

[Geissler et al., 2010] Geissler, W., Matias, L., Stich, D., Carrilho, F., Jokat, W., Monna, S., IbenBrahim, A., Mancilla, F., Gutscher, M.-A., Sallarès, V., et al. (2010). Focal mechanisms for sub-crustal earthquakes in the gulf of cadiz from a dense obs deployment. *Geophysical Research Letters*, 37(18).

[Hable et al., 2018] Hable, S., Sigloch, K., Barruol, G., Stähler, S. C., and Hadziioannou, C. (2018). Clock errors in land and ocean bottom seismograms: high-accuracy estimates from multiple-component noise cross-correlations. *Geophysical Journal International*, 214(3):2014–2034.

[Hannemann et al., 2014] Hannemann, K., Krüger, F., and Dahm, T. (2014). Measuring of clock drift rates and static time offsets of ocean bottom stations by means of ambient noise. *Geophysical Journal International*, 196(2):1034–1042.

[Harris et al., 2020] Harris, C. R., Millman, K. J., Van Der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., et al. (2020). Array programming with numpy. *Nature*, 585(7825):357–362.

[Hofman et al., 2017] Hofman, L. J., Ruigrok, E., Dost, B., and Paulssen, H. (2017). A shallow seismic velocity model for the groningen area in the netherlands. *Journal of Geophysical Research: Solid Earth*, 122(10):8035–8050.

[Hohensinn, 2019] Hohensinn, R. (2019). *Detection of Hazardous Ground Movements with Instantaneous Velocity Estimates by GNSS*. PhD thesis.

[Keilis-Borok, 2002] Keilis-Borok, V. (2002). Earthquake prediction. *Annu. Rev. Earth Planet. Sci*, 30:1–33.

[Knopoff and Gangi, 1959] Knopoff, L. and Gangi, A. F. (1959). Seismic reciprocity. *Geophysics*, 24(4):681–691.

[Kornei, 2019] Kornei, K. (2019). Nuclear bomb or earthquake? explosions reveal the differences. *Eos*, 100.

[Laughton, 1970] Laughton, A. S. (1970). A new bathymetric chart of the red sea. *Philosophical Transactions for the Royal Society of London. Series A, Mathematical and Physical Sciences*, pages 21–22.

[Lindner et al., 2017] Lindner, F., Wassermann, J., Schmidt-Aursch, M. C., Schreiber, K. U., and Igel, H. (2017). Seafloor ground rotation observations: Potential for improving signal-to-noise ratio on horizontal obs components. *Seismological Research Letters*, 88(1):32–38.

[Loviknes et al., 2020] Loviknes, K., Jeddi, Z., Ottemöller, L., and Barreyre, T. (2020). When clocks are not working: Obs time correction. *Seismological Research Letters*, 91(4):2247–2258.

[Mangano et al., 2011] Mangano, G., D'Alessandro, A., and D'Anna, G. (2011). Long term underwater monitoring of seismic areas: Design of an ocean bottom seismometer with hydrophone and its performance evaluation. In *OCEANS 2011 IEEE - Spain*, pages 1–9.

[Marshak et al., 1992] Marshak, S., Bonatti, E., Brueckner, H., and Paulsen, T. (1992). Fracture-zone tectonics at zabargad island, red sea (egypt). *Tectonophysics*, 216(3-4):379–385.

[McNamara and Buland, 2004] McNamara, D. E. and Buland, R. P. (2004). Ambient noise levels in the continental united states. *Bulletin of the seismological society of America*, 94(4):1517–1527.

[Mokhtar et al., 2001] Mokhtar, T., Ammon, C., Herrmann, R., and Ghalib, H. (2001). Surface wave velocities across arabia. In *Monitoring the Comprehensive Nuclear-Test-Ban Treaty: Surface Waves*, pages 1425–1444. Springer.

[Nanometrics, 2019] Nanometrics (2019). Trillium compact obs manual.

[Naranjo et al., 2021] Naranjo, D., Parisi, L., Jousset, P., Weemstra, C., and Jónsson, S. (2021). Determining obs clock drift using ambient seismic noise. In *EGU General Assembly 2021*.

[Oshida et al., 2008] Oshida, A., Kubota, R., Nishiyama, E., Ando, J., Kasahara, J., Nishizawa, A., and Kaneda, K. (2008). A new method for determining obs positions for crustal structure studies, using airgun shots and precise bathymetric data. *Exploration Geophysics*, 39(1):15–25.

[Peterson, 1993] Peterson, J. R. (1993). Observations and modeling of seismic background noise. Technical report, US Geological Survey.

[Roobol and Stewart, 2019] Roobol, M. J. and Stewart, I. C. (2019). Cenozoic faults and seismicity in northwest saudi arabia and the gulf of aqaba region. In *Geological Setting, Palaeoenvironment and Archaeology of the Red Sea*, pages 275–305. Springer.

[Sánchez-Pastor et al., 2019] Sánchez-Pastor, P., Obermann, A., Schimmel, M., Weemstra, C., Verdel, A., and Jousset, P. (2019). Short-and long-term variations in the reykjanes geothermal reservoir from seismic noise interferometry. *Geophysical Research Letters*, 46(11):5788–5798.

[Schmidt-Aursch and Crawford, 2015] Schmidt-Aursch, M. C. and Crawford, W. C. (2015). *Ocean-Bottom Seismometer*, pages 1735–1748. Springer Berlin Heidelberg, Berlin, Heidelberg.

[Schwarz et al., 1990] Schwarz, K. P., Sideris, M. G., and Forsberg, R. (1990). The use of FFT techniques in physical geodesy. *Geophysical Journal International*, 100(3):485–514.

[Sens-Schönfelder, 2008] Sens-Schönfelder, C. (2008). Synchronizing seismic networks with ambient noise. *Geophysical Journal International*, 174(3):966–970.

[Shariat-Panahi et al., 2009] Shariat-Panahi, S., Alegria, F. C., Làzaro, A. M., and Del Rio, J. (2009). Time drift of ocean bottom seismometers (obs). In *Proceedings of the 19th IMEKO World Congress on Fundamental and Applied Metrology*, pages 2548–2553. Citeseer.

[Shearer, 2019] Shearer, P. M. (2019). *Introduction to seismology.* Cambridge university press.

[Stachnik et al., 2012] Stachnik, J., Sheehan, A. F., Zietlow, D., Yang, Z., Collins, J., and Ferris, A. (2012). Determination of new zealand ocean bottom seismometer orientation via rayleigh-wave polarization. *Seismological Research Letters*, 83(4):704–713.

[Sukhovich et al., 2021] Sukhovich, A., Perrot, J., and Royer, J.-Y. (2021). Using teleseismic p-wave arrivals to calibrate the clock drift of autonomous underwater hydrophones. *Bulletin of the Seismological Society of America*, 111(1):21–35.

[Sutton et al., 1981] Sutton, G., Duennebier, F., and Iwatake, B. (1981). Coupling of ocean bottom seismometers to soft bottom. *Marine Geophysical Researches*, 5(1):35–51.

[Sutton, 1986] Sutton, G. H. (1986). *Ocean Bottom Seismology: History and Current Status*, pages 821–840. Springer US, Boston, MA.

[Trebbin and Wassermann, 2010] Trebbin, N. and Wassermann, J. (2010). Seismic data analysis with obspy. *BSc, Ludwig-Maximilians-University Munich.*

[Trippanera et al., 2019] Trippanera, D., Ruch, J., Passone, L., and Jónsson, S. (2019). Structural mapping of dike-induced faulting in harrat lunayyir (saudi arabia) by using high resolution drone imagery. *Frontiers in Earth Science*, 7:168.

[Uthaman et al., 2022] Uthaman, M., Singh, C., Singh, A., Jana, N., Dubey, A. K., Sarkar, S., and Tiwari, A. K. (2022). Spatial and temporal variation of the ambient noise environment of the sikkim himalaya. *Scientific reports*, 12(1):1–13.

[Wapenaar et al., 2010] Wapenaar, K., Draganov, D., Snieder, R., Campman, X., and Verdel, A. (2010). Tutorial on seismic interferometry: Part 1—basic principles and applications. *Geophysics*, 75(5):75A195–75A209.

[Wapenaar and Fokkema, 2006] Wapenaar, K. and Fokkema, J. (2006). Green's function representations for seismic interferometry. *Geophysics*, 71(4):SI33–SI46.

[Warner and Guasch, 2014] Warner, M. and Guasch, L. (2014). Adaptive waveform inversion-fwi without cycle skipping-theory. In *76th EAGE Conference and Exhibition 2014*, volume 2014, pages 1–5. European Association of Geoscientists & Engineers.

[Webb, 2008] Webb, S. C. (2008). The earth's hum: the excitation of earth normal modes by ocean waves. *Geophysical Journal International*, 174(2):542–566.

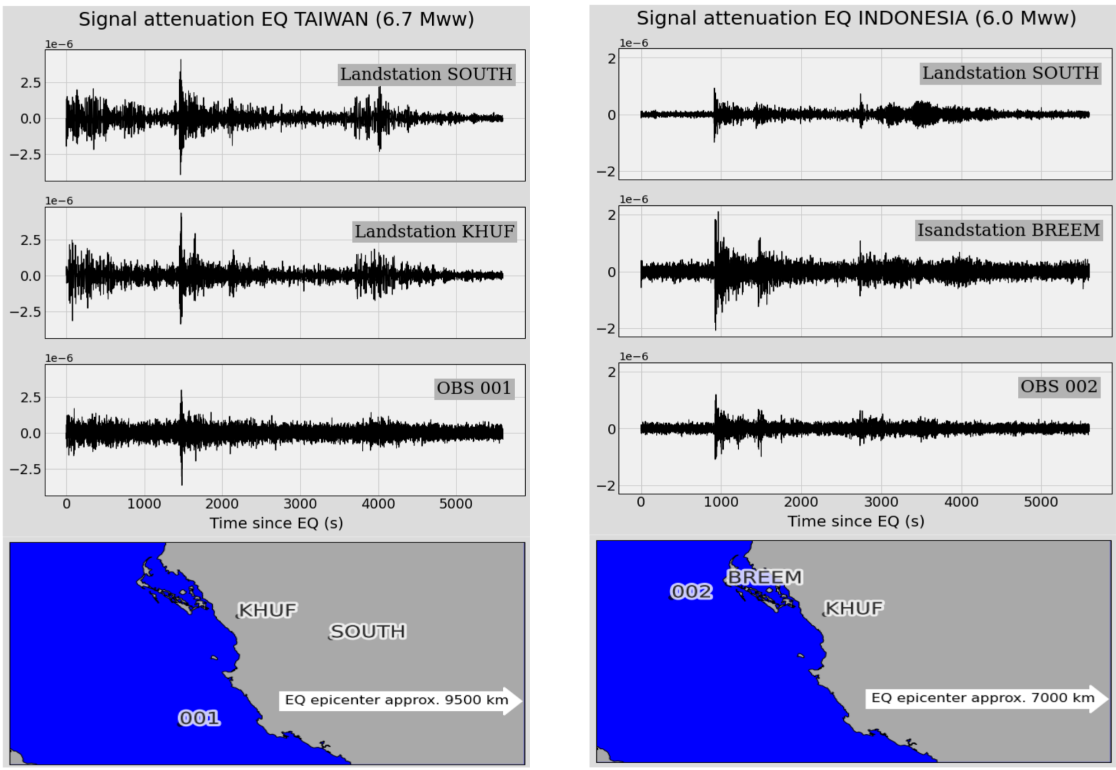
[Weemstra et al., 2021] Weemstra, C., De Laat, J. I., Verdel, A., and Smets, P. (2021). Systematic recovery of instrumental timing and phase errors using interferometric surface-waves retrieved from large-n seismic arrays. *Geophysical Journal International*, 224(2):1028–1055.

[Wilson et al., 2013] Wilson, D., Ringler, A., Storm, T., Hutt, C., and Gee, L. (2013). The importance of removing seismic sensor instrument response. In *AGU Fall Meeting Abstracts*, volume 2013, pages S42A–04.

[Yang and Ritzwoller, 2008] Yang, Y. and Ritzwoller, M. H. (2008). Characteristics of ambient seismic noise as a source for surface wave tomography. *Geochemistry, Geophysics, Geosystems*, 9(2).

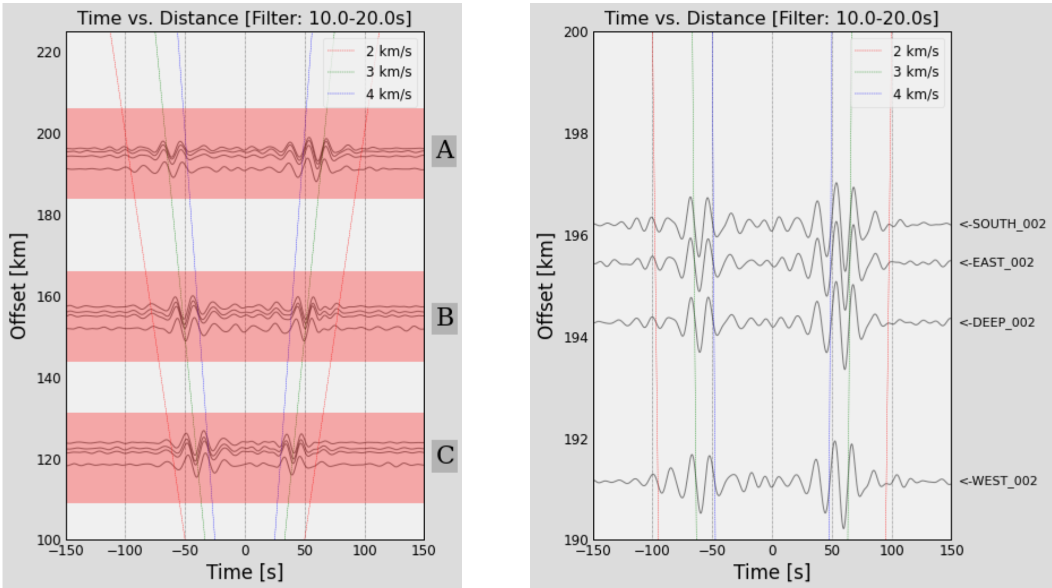
Appendix A

Appendix A

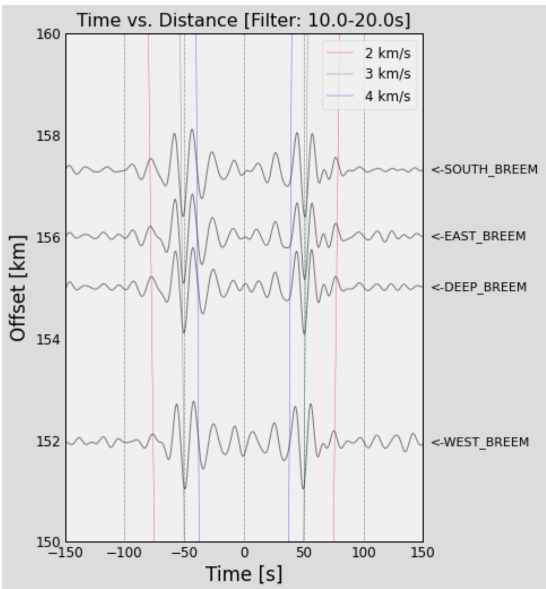


(a) **(b)**

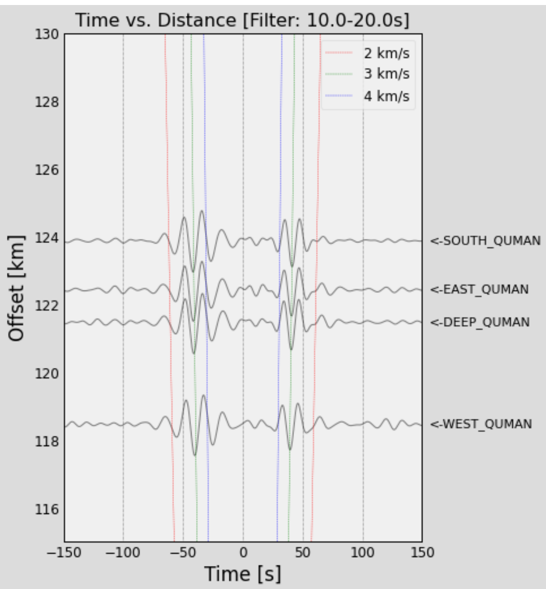
Figure A-1: Attenuation of short period noise based on an earthquake example for OBS 001 and OBS 002 compared to land-stations (Lowpass filter 10s)



(a)



(b)



(c)

Figure A-2: Zoomed in cross-corrleations from Harrat Lunayyir