Department of Precision and Microsystems Engineering

Cepstral Ranging Performance in Coastal Waters

Jan Lavoo

Report no : 2024.078
Coach : ir. P. Wessels, Dr. ir. P. Steeneken
Professor : Dr. ir. P. Steeneken
Specialisation : Dynamics of Micro- and Nanosystems
Type of report : Master Thesis
Date : 12 September 2024

Delft University of Technology

Master Thesis
High Tech Engineering - Dynamics of Micro- and Nanosystems
ME56035

Cepstral Ranging Performance in Coastal Waters

Jan Lavoo - 4971450

Supervisors:
Prof. Dr. Ir. Peter Steeneken - TU Delft
Ir. Peter Wessels - TNO

September 12, 2024

Abstract Multiple recent events have shown the alarming vulnerability of infrastructure on sea to sabotage. This infrastructure can be monitored and protected with passive sonar systems, which provide a number of advantages over other surveillance methods such as RADAR or active sonar. An overview of different localisation methods suitable for the Dutch littoral zone that perform bearing estimation, ranging or complete localisation on the surface plane is made from methods available in literature. From this overview a proposal is made to select cepstral ranging for further research, as there is a limited amount of literature available on this subject. Experimental data from the Dutch coastal area has been gathered with a RHIB and stationary sensor to verify important parameters for both computation and the ranging environment. An effective cepstral ranging method based on experimental data has been developed, although ranging with tidal effects and SNR mismatch remains somewhat problematic. To compensate for this a novel ray tracing method based on open-source measurements has been developed that offers a significant increase in accuracy while effectively compensating for variations in bathymetry and tide.

Acknowledgments
I would like to thank my supervisor Peter Wessels for his kind support during my research at TNO. Even with a filled schedule and after I left employment at TNO it was always possible to ask Peter for advice, he made me feel at home at TNO during my time there and he saved me a lot of time and headaches with elegant solutions to coding problems.

I also want to express my gratitude to Peter Steeneken in his role as supervisor from the TU Delft. His guidance helped me structure this thesis and he suggested some interesting literature and solutions. His patience at a time where I was stuck in research made me feel at ease.

My collegues from TNO, Bas, Elisabeth and Erwin also helped me writing this thesis by offering support and insight. They created opportunities to get practical experience in the hydro-acoustic domain by giving me the chance to help with experiments, gave advise on topics of interest and provided me with the experimental data that was essential for this thesis. For this they have my gratitude.

My parents Marc & Britta and brother Jasper also have my gratitude as they supported me through this period of research, even when it took longer than expected. Although most of this research might go beyond their understanding they still expressed a lot of interest in my work and kept me going.

I also want to thank my loving girlfriend Hansje for her unwavering support. Not only was writing our theses together fun, it also helped me concentrate. Her support and understanding made the thesis I am presenting here possible, and even improved the result as her spell checks are incorporated.

Finally I want to thank my friend the heron (figure 1), who made sure to evaluate the progress of this thesis (in exchange for some cheese) during his daily visits, and kept me company while writing.

Figure 1: The heron

Contents

List of Figures

List of Tables

Figures that are presented in this thesis without citation are of own work.

1 Introduction

1.1 Background

Recent world events have brought to light the vulnerabilities of underwater infrastructure like subsea cables, pipelines and wind turbines and their power cables on sea. Notably the Nord-stream 1 and 2 pipelines were presumably sabotaged on 26 September 2022 by a yet unknown party [3], and on 20 October 2022 both subsea cables connecting Shetland to the rest of the world were cut [4]. In these examples damage to infrastructure was first noticed by failure of the connected system and afterwards it is hard, if not impossible, to positively identify the party responsible for the damage. The reason for this is that these subsea cables and pipelines are completely unprotected and while most modern subsea cables have build-in acoustic surveillance capabilities, older cables do not [5].

Not only cables and pipelines are at risk, but also the numerous wind farms that provide a significant amount of energy for the Netherlands. In 2022 a Russian vessel was found gathering information on Dutch wind farms which was suspected to be a preparation for sabotage [6]. If these wind farms would be disabled this might have profound effects on Dutch energy security with disastrous results from both an economic and social perspective.

Surveillance of these areas of interest can provide deterrence against acts of sabotage. Passive sonar systems are an excellent tool for this. They are able to detect small targets that might be undetected by radar or even submerged targets, as long as the targets emit an acoustic signal themselves. As passive sonar does not use a self-emitted signal it does not damage or hinder sea-life and targets do not know that they are being tracked. In the case of a stationary system there is also no requirement for permanent crewing of a vessel for surveillance as monitoring can be performed from land-based installations or even automatically.

It is important to note that at the start of this thesis the topic of research itself was not clearly defined yet. The starting objective was to perform a literature survey on methods suited for stationary and passive surveillance in coastal areas as mentioned above. From this literature survey it was concluded that cepstral ranging seems like a promising and recently emerging ranging technique, while this specific method had not yet been tested or used by TNO. Ranging itself, no matter the technique used, is one of the most important and hardest parts of sonar based localisation. In the future ranging techniques like this might be used to detect if a vessel comes too close to a prohibited part of Dutch coastal areas, i.e. wind farms or other infrastructure. Therefore the decision was made to expand on this topic in this thesis.

Cepstral ranging itself is based on a specific algorithm that is able to discern the Time Difference Of Arrival or TDOA of different travel paths of sound waves underwater. This is the so-called multi-path environment of a shallow water waveguide. The physical mechanism behind this is explained in greater detail in section 2.1.3. From this TDOA it is possible to determine the distance to a target by a geometric relation if certain parameters are known. This thesis researches the possibilities and limitations of cepstral ranging and determines which factors and parameters influence its ranging performance.

1.2 Problem statement

Defense of critical underwater infrastructure has high priority at this time, especially for countries that possess a significant amount of sub-sea and on-sea infrastructure that is important for their society and economy. The Netherlands possesses a lot of wind farms in its coastal areas and receives a significant amount of internet data through undersea cables. For these wind farms specifically the surrounding waters are restricted to normal sea traffic to prevent accidents, but possibly also sabotage. It is however difficult to actively patrol these areas as they are numerous and large. Radar surveillance is mostly used for surveillance at sea, but

smaller vessels might be lost in noise. Passive sonar might offer better performance in this case, as any vessel that uses an engine or any other source of sound can be heard and located by this method when in range of this system. Additionally passive sonar does not pose a risk for marine animals as it does not emit sound itself.

During the cold war a vast amount of resources has been invested in sonar R&D, although arguably mostly aimed at Anti Submarine Warfare or ASW. Using sonar detection and localisation methods for small vessels with the purpose of coastal defense seems to be a less mature field. Considering the risks to underwater infrastructure as described in section 1.1 there is now a great interest in expanding on this topic. Bearing estimation methods for this purpose have been researched to some degree as will be shown in section 2.2.1 of the literature survey, but methods to find the range to a source or target are less developed as shown in section 2.2.2.

When comparing the methods described in this section, the 'cepstral analysis' method from section 2.2.2.2 seems most promising because of its relative ease of implementation and straightforward method of operation. While some studies have been done on this topic data was found to be relatively sparse and factors contributing to an increase of effectiveness of cepstral ranging have not been thoroughly researched. Cepstral ranging has been shown to function well in certain settings that are fitting for Dutch coastal surveillance, e.g. a shallow water environment, but which factors have influence on this, and how well this technique would be suited for actual use in coastal defence is not yet clear. Especially in shallow waters it can be expected that differences in water depth caused by tidal effects, storms etc. will have a significant influence on the exact multi-path structure and thus the TDOA. An effect like this is readily visible in figure 2, which shows the water depth varying over the course of multiple hours for measurements in the same location. In this figure it is also apparent that there is a significant variation in bathymetry for different locations. Both will result in inaccurate or skewed ranging values, as cepstral ranging experiments in literature assume a spatially and temporally uniform water depth [7] [8] [9] [10] [11]. Another issue is the matching algorithm, which matches measurements to the correct range. These and other issues should be investigated and resolved first before cepstral ranging is feasible for practical use.

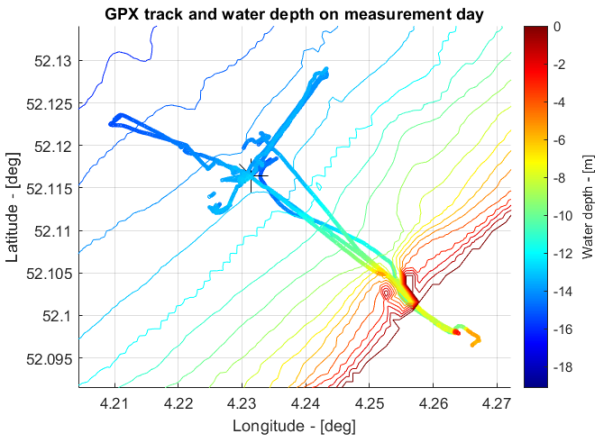


Figure 2: Measured and charted water depth during the acoustic measurements used in this thesis

1.3 Objectives

As described in the previous sections 1.1 and 1.2, the aim of this thesis is to provide an overview on the preparation of data for cepstral ranging, methods to perform cepstral ranging and to give substantiated advise to improve the performance of cepstral ranging. This thesis should provide the reader with a solid understanding of cepstral ranging, the advantages and disadvantages of this ranging method and suitable algorithms to use for this purpose. To accomplish this the following points will be thoroughly addressed, with the overarching research question mentioned first:

1. How accurate can the range of sound sources be estimated in a real-world situation using cepstral based ranging methods and a single hydrophone in relation to the true range between receiver and source? This main research question should indicate if cepstral ranging is indeed suitable for deployment in coastal surveillance and should offer an indication of performance compared to other ranging methods.

To answer this main research question first a few sub-questions should be answered. These are formulated to solve challenges that have to be overcome before cepstral ranging could be considered usable in a deployable and well-functioning system. These sub-questions including context are shown below:

2. Which cepstrum settings and methods of data preprocessing are suitable for cepstral ranging? A cepstrum is computed with parameters comparable to an FFT computation, i.e. a window size and overlap. It is also possible to adjust the sample rate to increase computation speed, which might impact the ranging performance. When a cepstrum is generated there will certainly be noise present. Removing this noise as much as possible with various methods will presumably increase the ranging performance.

3. Which methods of building a ranging library are suitable for cepstral ranging? Cepstral ranging involves comparing a cepstrum generated from a measurement to a so-called library which contains cepstra of which the accompanying range is known. Various methods to generate this library are available, either based on theoretical models or earlier measurements. A few of these methods of library building should be researched and compared.

4. Which matching algorithms are available and what is their performance? When one has decided on a type of library as mentioned above, an algorithm should be selected to automatically compare the measurement cepstrum to the library. The accuracy of this algorithm is expected to have a significant influence on the accuracy of cepstral ranging, as well as on the maximum range for which a source can be ranged.

5. What is the influence of (tidal) changes in water depth on the performance on cepstral ranging and how to negate resulting inaccuracies in ranging? As mentioned in section 1.2 it is expected that a change in water depth will have a significant impact on the ranging values generated by cepstral ranging. The Dutch coastal areas have a constantly changing water depth because of the tidal movement of water, which can result in a depth varying by several meters on a single day. This effect must be accounted for to prepare cepstral ranging for practical applications.

It is anticipated that answering these questions will result in an overview of cepstral ranging that is sufficient to prove whether cepstral ranging is applicable to real-world deployment and that the results can be used as a guideline for both the development of a cepstral ranging system and to compare it to other ranging systems.

1.4 Scope and limitations

Cepstral analysis and ranging is a broad topic which still requires a significant amount of research to enable it to be fielded reliably and accurate. It is not realistic to cover and solve all issues related with cepstral ranging in this thesis. Therefore a certain amount of compromises and limitations have to be made in the research for this thesis. These limitations will be outlined in the following section.

- Already before making the literature survey it was decided that the focus of this thesis should be on *stationary* systems. A moving receiver could perform range estimation by taking multiple bearing measurements while manoeuvring the receiver to infer a range. A moving system however introduces a plethora of other issues concerning reliability, maintenance and general feasibility, especially when the system has to be unmanned.
- No modelling or simulations to create synthetic data will be carried out. These models are often computationally intensive and creating a time sequence would take considerable resources. Next to that, creating accurate synthetic data is a complex task and it would be difficult to verify the accuracy of this data. As this thesis tries to verify the usefulness of cepstral ranging in a real environment, experimental data will suffice.
- While the dataset used in this thesis was recorded on a vector hydrophone, which is able to estimate the bearing of a source, only the included hydrophone *sound* data will be used in this report. Potentially a vector hydrophone could be able to estimate both a bearing and range using the methods discussed in this thesis, which could be very effective for coastal surveillance as it should result in a complete 2D localisation on the horizontal plane. To be able to focus completely on the more challenging ranging part, the vector data will not be used.
- Because of the relatively shallow depth of Dutch coastal areas and to simplify the research for this thesis considerably, all sound sources that are to be ranged are assumed to be on or very close to the surface. The multi-path structure of sound travelling in the waveguide of a body of water is essential for cepstral ranging, and this structure changes considerably if the depth of the source changes. This will result in different TDOA values for the same range in the case of a varying depth and thus inaccurate ranging values. This thesis focuses on surface sources like most vessels, and not on submersed sources like Unmanned Underwater Vehicles (UUV). While in theory it should be possible to estimate both the range *and* depth from cepstral analysis, this is deemed out of scope for this thesis.
- The focus of this thesis will be on measurements in open waters, i.e. in areas with no sound wave reflections by underwater objects or walls. Since cepstral ranging relies on the TDOA of multi-path reflections, reflections from objects other than the water surface or bottom might interfere with the performance of cepstral ranging. This could limit the usefulness of cepstral ranging in for instance a port or harbour, or any other area that contains significant underwater walls or objects that might reflect sound waves from the source to the receiver. While this apparent issue has to be overcome before cepstral ranging can be used in the aforementioned environments, priority should be given to create a reliable system for cepstral ranging without the addition of these challenges. Therefore any environment other than open water without significant underwater structures are deemed out of scope for this thesis.

Even without research done in the areas mentioned above, this thesis has the potential to result in a ranging method that is suitable for coastal defence in shallow water after overcoming some engineering challenges with regards to the deployment of such a system. Addressing these limitations would only offer an *improvement* to the ranging or localisation performance.

1.5 Overview of content

This thesis presents an exhaustive literature survey on topics of significance for coastal surveillance, which include the basics of hydroacoustics, methods to find both bearing and range or to perform a 2D localisation and an introduction to sound signatures.

This literature survey identifies cepstral ranging as an emerging field of interest with the potential to be of significance for coastal surveillance. This method elegantly identifies the Time Difference of Arrival or TDOA of various echoes, even with a stationary signal. These echoes can then be used to find the range to a source by geometric relations. As not a lot of information about this topic is yet available, a project is proposed to identify which parameters and factors are of importance during cepstral ranging and to quantify the practical accuracy of this method.

A methodology is presented which expands on the experimental data that has been used to generate these results. This data consists of both GPS and audio measurements taken a few kilometers from the coast near Scheveningen with a RHIB. The receiver was placed on the bottom after which the RHIB performed a series of runs at various distances of the receiver, while constantly recording its position with a GPS receiver. Next to the detailing of data it is explained how the measurement cepstra and libraries which contain cepstra with a known range are generated, how the step between library and measurement can be made and parameters of potential significance are identified. To compensate for tidal effects, other environmental effects and bathymetry, a novel ray tracing algorithm has been developed that incorporates open source data to compute sound speed and expected reflection path lengths.

The methodology is then converted to results where the errors between distance known by GPS and estimated distance are compared for various parameters of interest. These results offer insight on which parameters for cepstrogram generation offer the most accurate results and other implications. Issues with variations in tide and SNR are identified, researched and alleviated. A novel ray tracing method however offers impressive performance while not yet fully operational, as it is able to effectively compensate for bathymetry, tide and sound speed during TDOA estimation. These results are used to give recommendations to further the development of the ray tracing algorithm and to expand on the mechanism behind ranging issues with tidal variations and variations in SNR.

2 Literature survey

2.1 Principles of shallow water sound propagation

Water is an excellent medium for sound propagation. A body of water will act as a natural waveguide for sound waves, constrained by the sea floor and surface. Nevertheless, acoustic recognition and localisation is hampered by sound attenuation or so called transmission loss. Because sound absorption in water is very small for low frequencies, it is mostly dominated by geometrical spreading, volume attenuation, bottom reflection loss and boundary and volume scattering loss [12]. Another limiting factor is the refraction of sound waves caused by density differences in different water layers. Especially important in shallow water are sound reflections, bouncing between water surface, bottom and underwater obstacles. As these phenomena have a profound influence on detection and localisation underwater, the mechanisms behind them and their effect on shallow water sound propagation are discussed in this section.

2.1.1 Transmission loss

Transmission loss is defined as the ratio in decibels between the sound intensity $I(r, z)$ at a field point and the intensity I_0 at a 1m distance from the source. It plays an important role in the Sonar equation 1. While relatively simple for passive Sonar systems, it's a useful tool to determine the Signal to Noise Ratio or SNR, which has some connection to the detectability of targets by passive Sonar [1].

$$SNR = SL - TL - N + DI \tag{1}$$

Here SL stands for Source Level, TL for Transmission Loss, N the noise at a single hydrophone and DI for the Directivity Index, which is applicable when an array that has some noise rejection is used.

Geometrical spreading accounts for the largest part of transmission loss. When neglecting attenuation the radiated energy at first will spread over the surface of a sphere with radius r which is equal to the distance to the point source. This implies that the energy at distance r will be inversely proportional to the sphere surface area, which results in equation 2 [12].

$$TL = 20 \log r \qquad \text{(dB re 1m)} \tag{2}$$

As the sound energy is captured in the water waveguide, spreading in the vertical direction will cease once the sound travel distance r becomes larger than the water depth D. As a result the spreading loss gradually becomes inversely proportional to the surface area of a cylinder, which results in equation 3 [12].

$$TL = 10 \log r \qquad \text{(dB re 1m)} \tag{3}$$

Equation 2 holds for the cases where $r < D$ and equation 3 for the farfield when $r >> D$.

2.1.1.1 Volume attenuation and scattering losses

When sound travels through the ocean it will continuously transfer energy to the water which will turn into heat. Sound will also scatter while propagating through water, for a significant part by biological organisms or by surface and bottom roughness during reflections[12]. An approximation of these losses is given in equation 4 [13].

$$\alpha' \simeq 3.3 * 10^{-3} + \frac{0.11 f^2}{1 + f^2} + \frac{44 f^2}{4100 + f^2} + 3.0 * 10^{-4} f^2 \qquad \text{(dB/km)} \tag{4}$$

While equation 4 does not account for pressure, temperature and salinity, which all have an effect on attenuation, it is sufficient as an approximation and to demonstrate the effect of frequency on attenuation.

2.1.1.2 Boundary reflection loss

In shallow water waveguides, sound waves will inevitably interact with both the bottom and the surface of this waveguide. This constrasts with sound propagation in an ocean, where refraction as explained in section 2.1.2 can capture sound waves in a layer of low sound speed [14]. Depending on the incident angle in regard to the boundary plane and the difference in sound speed between both media, a certain fraction of the incident energy is reflected and another fraction leaves the waveguide.

Surface interaction also results in attenuation [12], a doppler shift and amplitude modulation by wave movement on the surface and multiple scattering mechanisms [15]. One of these mechanisms unique for the surface boundary is the mixing of air bubbles into the top layer with increased wind speeds. Sound energy reflected by the boundaries will also result in interference or so-called 'multipath propagation' which will be discussed in section 2.1.3.

Bottom interaction has much of the same mechanisms as surface interaction, but on the other hand the reflections are more complex as the sea bottom often is made up of multiple layers of materials with different geoacoustics. This layering can also have a large variation in lateral directions, making extrapolating from known data difficult [16]. Depending on the media these layers are made of, they can be modelled as fluids which only support compressional waves, or if the media are closer to a solid they have to be modelled as elastic. In that case they support both compressional and shear waves. In reality these media are more or less viscoelastic and thus lossy [12].

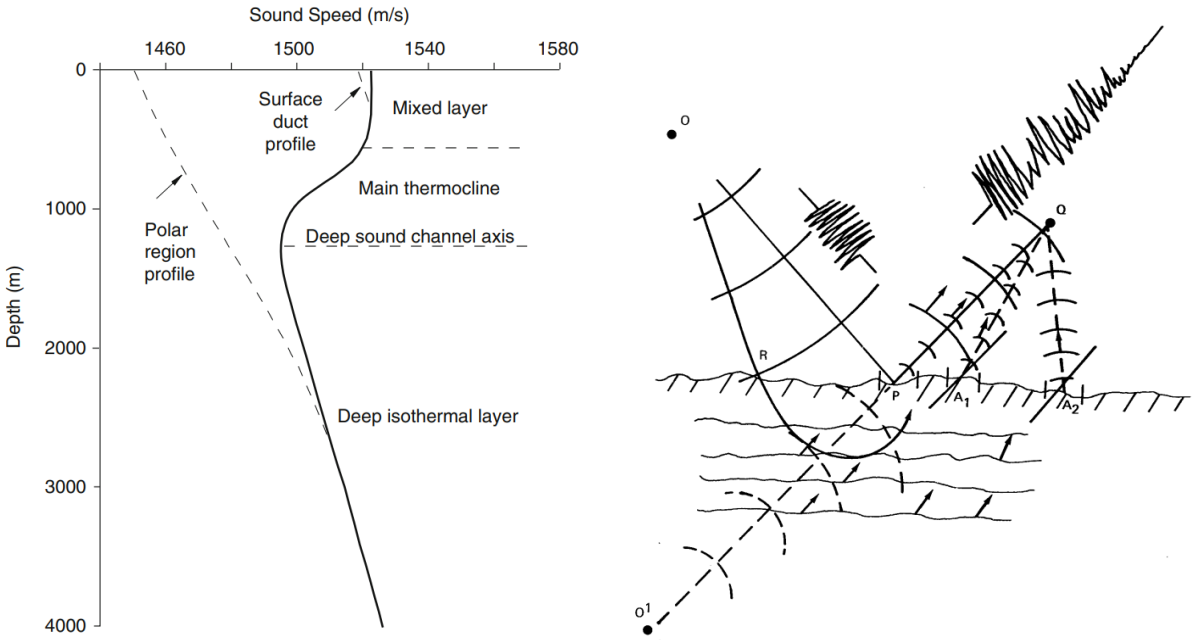
For bottom reflections there can be a critical incident angle for which reflections would be theoretically lossless because no incident energy will propagate into the boundary layer. This can only occur when the sound speed in the boundary layer is higher than the sound speed of water. The equation to calculate this critical angle is given in equation 5 [12].

$$\theta_c = \arccos\left(\frac{c_1}{c_2}\right) \tag{5}$$

In equation 5 c_1 is the water sound speed and c_2 the sound speed in the boundary layer. In reality there will be a phase shift depending on the incident angle and a loss of energy during the reflection.

2.1.2 Sound wave refraction

The speed of sound in water is in reality often a function of depth, especially in deep water. Water at the surface will (especially during the summer and daytime) rise in temperature and therefore see an increase in sound speed. At a certain depth there can be a large temperature gradient where the temperature quickly lowers with increasing depth, the so-called thermocline. With the lower temperature the sound speed decreases as well. As depth increases pressure increases too so sound speed gradually increases with an increase in depth. An approximate graph of the sound speed in an ocean as a function of depth is shown in figure 3a.



(a) Sound speed vs. depth in an ocean [12]

(b) Refraction and reflection in layered media [16]

Figure 3: Sound speed and refraction in an ocean

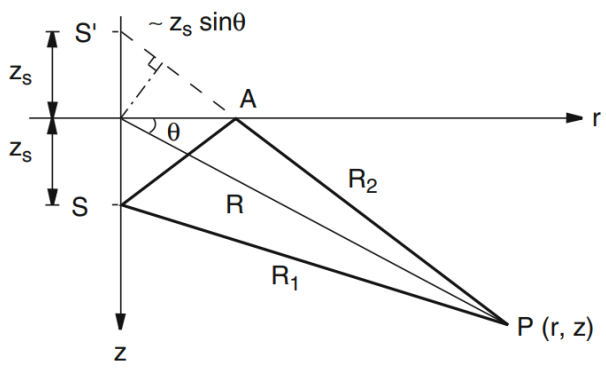
According to Snell's law, these changes in density will refract sound waves during their travel, potentially trapping sound waves in water layers with low sound speed [12]. For very shallow waters like ports or Dutch wind turbine farms with an average water depth of 20m these effects are practically nonexistent. While refraction by water layers might occur it will not prevent sound waves from reflecting of the surface or bottom [17].

Refraction can play a role in the layered media, where sound waves that have penetrated into a layer encounter a sound speed gradient and bounce or curve back to the water on different positions. This is shown in figure 3b and results in a distorted signal for an omnidirectional receiver Q.

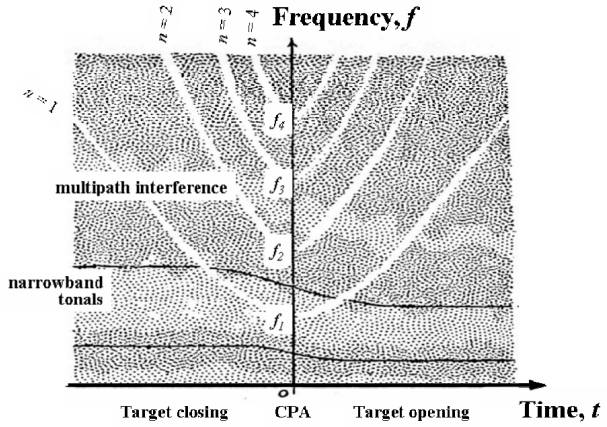
2.1.3 Interference in the multipath environment

Both refraction and reflection create many different paths that sound waves can travel from a transmitter to a receiver in a shallow water environment. As these paths have different travel times and transmission loss they will produce a distorted signal for the receiver [16].

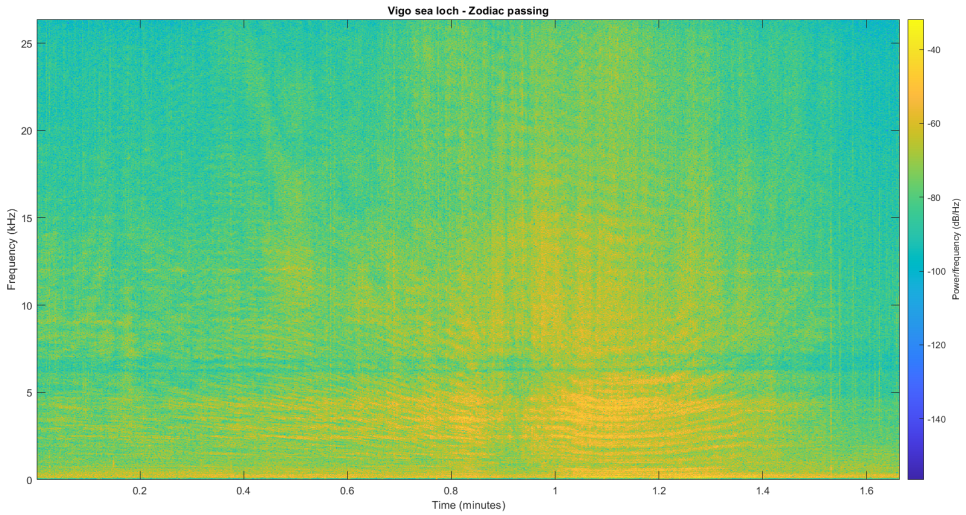
A classical example of this is the Lloyd mirror pattern. This effect is created by the summation of the direct path and one or more reflected paths when they arrive from a broadband source S at the receiver P as shown in figure 4a [18]. The time delay between arrival of these different paths and phase reversal of reflections will create both constructive- and destructive interference for specific frequencies, depending on source depth and range.



(a) Lloyd mirror geometry [12]



(b) Diagram of a typical Lloyd mirror pattern on a spectrogram [18]



(c) Spectrogram of real data displaying a Lloyd mirror pattern

Figure 4: Visualisation of the Lloyd mirror pattern

The parabolic shape of the interference striations in a spectrogram is caused by movement of the source relative to the receiver. As distance (or depth) decreases or increases the striations move to different frequencies, creating the characteristic 'bathtub' shape as seen in figure 4b and 4c [18].

While multipath interference will distort the received signal it can also be a valuable tool for ranging a target, as the interference pattern is a function of both depth and range. If the bathymetry is known it is possible to calculate the range to a target with use of the interference pattern. This is further explained in section 2.2.2.

2.2 Detection and localisation algorithms

To effectively survey an area the location of targets in or near this area has to be known. This section addresses common or promising algorithms that accomplish this goal. These algorithms can be broadly divided into three different categories on which location parameters of the target they can acquire.

The most common algorithms will result in an estimation of the target bearing relative to the receiver. This can be done in its simplest form by cross correlation of the signal of two or more hydrophones or beamforming algorithms that can use hundreds or hydrophones. These algorithms are explained in section 2.2.1. Algorithms for ranging are much less common. If the source strength and bathymetry of the surrounding water are unknown it is not possible to effectively find a target range with a stationary receiver, except when using triangulation methods which are addressed in section 2.2.4. There are some methods based on multipath interference that can work with a smaller hardware setup, these are addressed in section 2.2.2. Finally there are algorithms that can find a target location in 2D or 3D space. They often require extensive knowledge of the bathymetry or use a large amount of sensors and are addressed in section 2.2.3.

2.2.1 Bearing estimation methods

2.2.1.1 Cross correlation

In literature there is a large selection of bearing estimation methods to be found. For shorter ranges and very shallow water like harbour environments the (generalized) cross correlation method is frequently used [19] [20] [21] [22]. This method is based on the principle that sound waves will have different arrival times for two hydrophones with some spatial distance between them. The difference in arrival times can be calculated using equation 6 [19].

$$\Delta T = \frac{L \sin \alpha}{c} \quad \text{or} \quad \alpha = \arcsin\left(\frac{\Delta T * c}{L}\right) \tag{6}$$

Here L means the distance between the hydrophone pair, α the angle between the source and the normal of a line connecting the two hydrophones, and c means the local speed of sound.

The cross correlation algorithm compares two data sets and shifts them per sample point relative to each other. It returns a peak for a certain shift of samples that aligns the two data sets correctly. The amount of sample shifts that produced a peak in the cross correlation multiplied by the sample rate will give the time delay ΔT. The cross correlation is defined as shown in equation 7 [19].

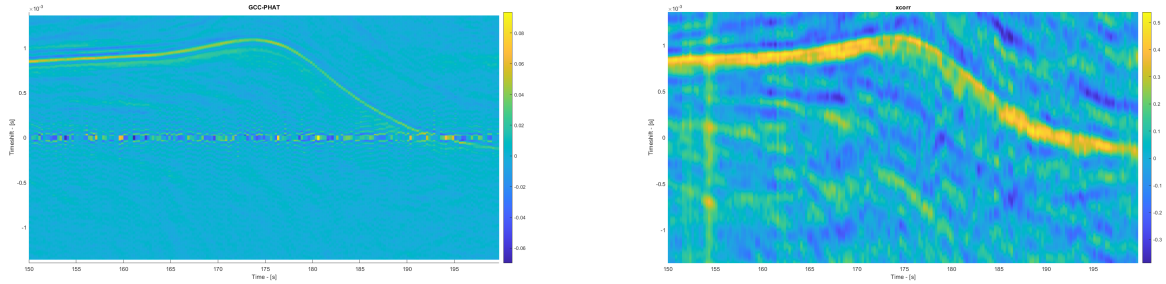
$$R_{12}(\tau) = \int_{-\infty}^{\infty} h_1(t') h_2(\tau - t') dt' \tag{7}$$

With h_1 and h_2 being the signals of hydrophone 1 and 2 respectively.

A variant of the cross correlation called the GCC-PHAT algorithm uses a frequency weighing function that normalizes the signal spectral density by the spectrum magnitude. As this whitens the input signals it sharpens the peak of the cross correlation function and thus increases resolution [23]. The cross correlation function 7 is multiplied by the weighing function. This specific algorithm has been used to detect divers [20].

While it is technically possible to place the sensors far apart this can compromise its performance for bearing estimation. As sound waves are made up from a multitude of sinusoidal signals there is also a multitude of phase shifts or time delays that will align the sinusoidal signals, depending on the specific frequencies of these sound waves. If the distance L between the two sensors is kept within half the wavelength of the smallest wavelength that is to be measured it is ensured that only the closest phase shift that matches is actually correct. If this is not done the cross correlogram will show grating lobes on other time delays that also match

the signal frequency as shown in figure 5a as tracks parallel to the main correlated track. In practice this means that it will be harder or even impossible to exactly pinpoint the correct bearing if the signal does not contain unique features in the time domain.



(a) GCC-PHAT correlogram

(b) Standard correlogram

Figure 5: Correlogram comparison of a RHIB passing in a harbour channel

2.2.1.2 Beamforming

Beamforming is probably the most popular family of algorithms for hydroacoustic detection. The SOSUS system, which was built and used by the USA to detect and track Soviet submarines throughout the Atlantic, employed large arrays and beamforming techniques for this purpose. This technique is also extensively used in RADAR and wireless communication [24]. The most straightforward beamforming method is the principle of delaying and summing the received signals of all elements of a hydrophone array in the time domain to look in a certain direction using the same principle of time delay as in equation 6 [25]. Summation of the hydrophone datasets after delaying them the appropriate amount has the advantage of creating noise rejection of signals coming from other directions. If it is assumed that noise and other sources are uncorrelated in the looking direction these sources will get cancelled somewhat by summing, while amplifying the signal in the looking direction [1]. The bearing of a source can be found by looking for peaks in the power output of the beamformer. Other relevant beamforming methods are generally applied in the frequency domain after a Fourier transformation of the received signals, like the conventional (Bartlett) beamforming algorithm. The delay and sum algorithm in the time domain is also called conventional beamforming or vice versa in some literature [26], but [1] uses this term for the frequency domain version. The Bartlett beamformer is similar to the delay and sum beamformer but instead of a timeshift a phase shift is applied that is equal to a timeshift on each frequency. By increasing the amount of sensors in an array the noise filtering effect becomes larger and the beamwidth becomes smaller [25]. As the array still is finite there will be sidelobes, angles other than the looking direction that have some gain and thus will let directional noise through. One method to suppress these is by using a weighed beamformer, where each sensor is assigned a specific weight. This does come at the price of increasing the mainlobe width [1].

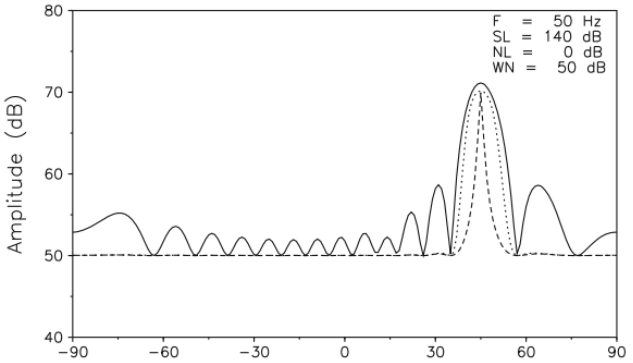


Figure 6: Comparison of beamforming algorithms for a single source in 100m water, the solid line represents CBF, the dotted line MCM and the dashed line MVDR [1]

While conventional and weighed beamforming are capable of handling wideband signals there are also other beamforming methods that will only handle narrowband signals with a reduction in sidelobes and mainlobe width [16]. The Minimum Variance Distortionless Responce processor (or MVDR) shows excellent source separation and resolution for sources close to each other as shown in figure 6 (note the sidelobes for CBF), but is sensitive to phase errors on sensors, in that case the Fourier Integral Method (FIM) might be more robust[27] [28].

2.2.1.3 MUSIC Algorithm

The MUltiple SIgnal Classification algorithm can be used as a direction finding algorithm for signals. It offers high resolution and is based on the eigenvalue decomposition of the sensor covariance subspace of an array. In its original form it only works with narrowband signals but modifications that support broadband signals are available at some computational cost[29]. In the case of diver detection with passive sonar MUSIC offers a comparable performance to MDVR [30]. The algorithm has as shortcomings that the amount of sources needs to be known in advance and that correlated sources are difficult to separate but it offers better performance than other bearing estimation methods [31].

2.2.2 Range estimation methods

Range estimation is one of the challenges of passive SONAR. While with active SONAR it is possible to infer the range to a target by recording the travel time between emission of a signal and reception of its reflection this is not an option with passive SONAR. If the velocity of a source is unknown it is also impossible to infer the range to a source from the bearing rate $\dot{\theta}$ and as moving receivers are outside of the scope of this report inferring the range by receiver manoeuvring is also not feasible. It is however possible to perform ranging based on the acoustic energy distribution over frequencies measured at one or more stationary receivers, based on multipath interference. This section discusses and explains these methods.

2.2.2.1 The waveguide invariant

The striation pattern that can be observed in figure 4 is caused by interference of reflected paths as explained in section 2.1.3. From normal mode propagation theory it can be shown that similar striations occur for ranges where the Lloyd mirror pattern is not valid anymore [32]. In this far field the striations are created by interference between multiple normal modes. According to [33] the received pressure spectrum $I(r, z_r, z_s, \omega)$ of a source is a function of range r, receiver depth z_r, source depth z_s and angular frequency ω can be written as equation 8, which is the sum of all relevant modes. The amount of relevant modes N is a function of frequency [32].

$$I(r, z_r, z_s\omega) = P(\omega)(\sum_n B_n^2 + 2 \sum_{n,m;n\neq m} B_n B_m \cos(\Delta k_{mn}(\omega) r)) \tag{8}$$

where

$$B_n(z_s, z_r) = (\frac{2\pi}{k_n r})^{\frac{1}{2}} \Psi_n(z_s)\Psi_n(z_r) \quad \text{and} \quad \Delta k_{mn}(\omega) = k_m(\omega) - k_n(\omega) \tag{9}$$

with $P(\omega)$ the source specrum, k_n the mode eigenvalue and horizontal wavenumber and $\Psi_n(z)$ the accompanying eigenfunction. The cosine term in equation 8 is created by the interference of two modes and it is this part of the equation that creates the striation pattern mentioned above. It also becomes clear from equation 8 that the contours of equal sound intensity (the striations in this case) are a function of range and frequency. The slope of these contours can be calculated using equation 10 [33].

$$\frac{\delta\omega}{\delta r} = \beta_{ml}\frac{\omega}{r} \tag{10}$$

Here m and l are wavemode numbers, ω the frequency in radians per second and r the target range. β is an interesting parameter determining if the contour is linear or not and is shown to be ≈ 1 for ideal waveguides. In ideal waveguides it is also shown to be mostly independent of m and l, and importantly not strongly dependent on the details of sound speed profile (SSP) or seafloor properties in shallow water, hence the name waveguide *invariant* [33]. If these assumptions are made then equation 10 can be rewritten to equation 11.

$$r = \beta * f * \frac{\delta r}{\delta f} \tag{11}$$

This potentially makes it a powerful tool in passive ranging in shallow water as it offers a robust option where errors in SSP or bottom properties would only have a weak effect on ranging errors.

2.2.2.2 Cepstral analysis

The term 'cepstrum' was originally coined in [34] by reversing the first syllable of 'spectrum' as it was then explained as being the spectrum of a spectrum. Other relevant terms like 'frequency', 'harmonic' or 'filter' were changed to 'quefrency', 'rhamonic' and 'lifter' respectively in the same fashion [35]. These terms will be used for the remainder of this report.

The current definition of the power cepstrum is shown in equation 12.

$$C_p = \mathcal{F}^{-1}(\log(F_{xx}(f))) \tag{12}$$

Here F_{xx} denotes the power spectrum and $\mathcal{F}^{-1}$ the inverse Fourier transform. Essentially this operation returns peaks on the periodicity of peaks in the spectrum. The log operation performs linear separation or deconvolution on the multipath signal, separating a signal containing an echo in two additive parts, the original signal and the echo which contains a periodic component.

If a signal containing one single echo is considered it can be expressed as 13. Here $s(t)$ is the signal and $\alpha s(t-\tau)$ an echo with strength α and a TDOA of τ.

$$x(t) = s(t) + \alpha s(t-\tau) \tag{13}$$

Taking the Fourier spectral density of this signal would result in 14, where the echo and the original signal are convoluted.

$$|F_{xx}(f)|^2 = |S(f)|^2(1 + \alpha^2 + 2\alpha\cos(2\pi f\tau)) \tag{14}$$

By taking the logarithm of the power spectral density deconvolution is performed on the signal and its echo, as shown in 15 [36].

$$C(f) = \log|F_{xx}(f)|^2 = \log|S(f)|^2 + \log(1 + \alpha^2 + 2\alpha\cos(2\pi f\tau)) \tag{15}$$

In this sense it is indeed a spectrum of a spectrum, as a spectrum will result in a peak on the periodicity of a signal. Because phase information is lost by using the power spectrum it is not possible to return to the time domain, hence the term 'quefrency' denoting the temporal unit of the power cepstrum [34].

The power cepstrum is a very useful tool for echo detection, as an echo will result in both constructive and destructive interference on specific frequencies depending on phase. This is an exact analogy to the Lloyd mirror pattern of section 2.1.3. Therefore the power cepstrum can be used to determine the timedelay of multipath arrivals or Time Difference Of

Arrival (TDOA). It was shown that for shallow water SONAR applications it outperforms the autocorrelation function [7]. While other cepstra with different properties exist, like the complex cepstrum which is reversible to the time domain because it retains phase information, it was shown that for TDOA estimation the power cepstrum performs best [37]. Using cepstral analysis for range estimation in shallow water SONAR seems to be a relatively new concept. Model based predictions of TDOA seem to introduce a large error from ranges further than 200m - 300m, which can be solved by using trained convolutional neural networks for the replica fields [8]. The error in the model based predictions was explained in this study as being caused by unpredictable scattering of the reflected signal on smaller angles of attack. By using a more advanced reflection model which accounted for a rough seafloor the error was significantly reduced for intermediate ranges. This method can be combined with good results with the GCC method from section 2.2.1.1 for bearing estimation in a small array to get both bearing and range estimation [9]. A comparison for a number of different methods to create replica fields is given in [10].

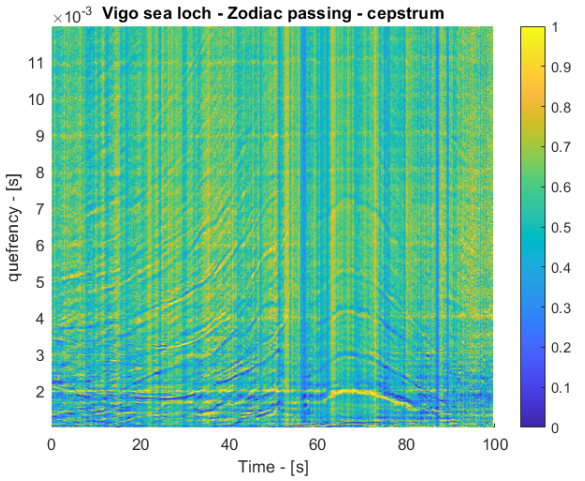


Figure 7: Cepstrogram of passing Zodiac, using the same data as in figure 4c

Recently [11] showed an impressive performance of cepstral range detection using a single hydrophone with accurate range estimates for a distance of up to 1km in 25m deep water. In this experiment normalized, replica cepstra were derived from a passing RHIB with a single known track and cepstra from other hydrophones in the same array were matched to this library for multiple tracks. This increase in performance was explained by the increase of the maximum quefrency which should therefore include other paths with longer travel times. These longer paths still have a significant TDOA even on longer ranges. Data was also downsampled to 2kHz.

2.2.3 Matched Field Processing

One solution to find range, depth and possibly bearing estimations is the family of algorithms called Matched Field Processing (MFP). This method can be broadly seen as the three dimensional generalization of lower dimensional, plane-wave beamformers. In general for MFP the sound field is recorded at a (often vertical) linear array and matched to a known set of sound field replica's and their spatial location. The unique spatial structure of the sound field then permits localisation in range, depth and bearing, depending on the array geometry and the complexity of the ocean environment [1]. These replica fields are mostly computed using sound propagation models, but it is also possible to use sources of opportunity with a known location to estimate the replica fields for these locations [38] [39].

MFP uses the depth dependent Green's function as a basis to represent the used signals. The ocean is modelled as a waveguide and the signal from the source can be considered to be the point source solution to the wave equation. The point source solution is composed of modal components which propagate long distances at shallow grazing angles and vertical components which decay rapidly because of strong interaction with the bottom. The modal components have discrete values of the wavenumber spectrum for the Green's function while the vertical components occupy a continuum in the spectrum at lower wavenumbers or higher phase speeds, as shown in figure 8. For sources at horizontal ranges that are greater than a few water depths from the receiving array, the discrete portion of the field is the major contributor. Surface noise

can have both discrete and continuous contributions and can thus have some overlap with the source signal at the array [40].

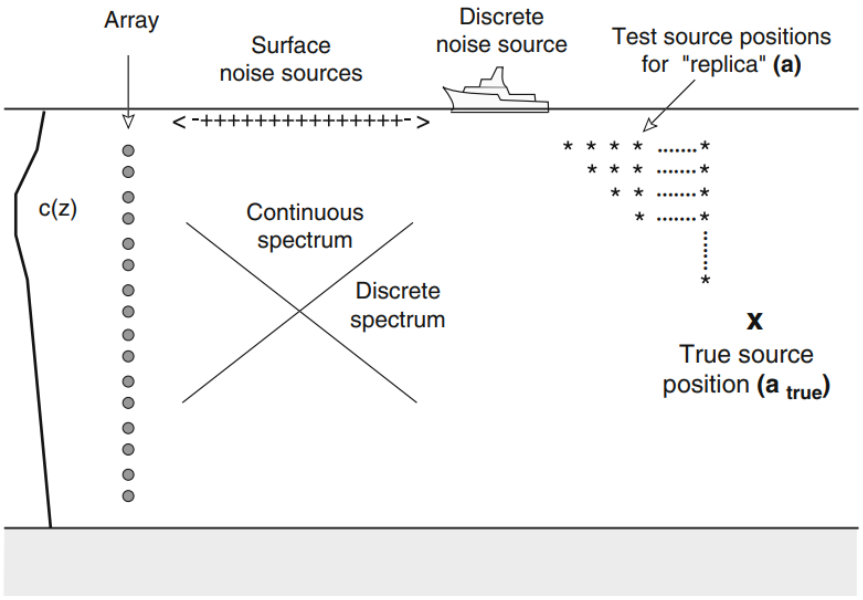


Figure 8: MFP schematic showing the vertical array, replica field positions, sound speed profile and the spectrum distribution [1]

The MFP algorithms generally use narrowband signals, although some broadband methods exist [40]. $\mathbf{G}(\mathbf{r}, \mathbf{z})$ is originally used to denote the velocity potential but as the field is a function of frequency it can also be conveniently expressed as $\mathbf{G}(f, \mathbf{A})$ where A denotes parametric dependencies on source location and environmental parameters. The spatial structure of the signal at each array element N can then be written as 16

$$\mathbf{G}(\mathbf{f}, \mathbf{A}) = \begin{bmatrix} G(\mathbf{r}_1, z_1, \mathbf{A}) \\ G(\mathbf{r}_2, z_2, \mathbf{A}) \\ \vdots \\ G(\mathbf{r}_N, z_N, \mathbf{A}) \end{bmatrix} \tag{16}$$

with $r_1, z_1 ... r_N, z_N$ denoting the ranges and depths of each sensor respectively. The covariance of the signal is given by equation 17,

$$\mathbf{K_s} = \sigma_b^2 \mathbf{G}(\mathbf{f}, \mathbf{A}) \mathbf{G^H}(\mathbf{f}, \mathbf{A}) \tag{17}$$

where σ_b^2 is the variance of a complex Gaussian random variable that is equal to the source strength. The ensemble covariance matrix $\mathbf{K}$ is then made up of the signal covariance matrix $\mathbf{K_s}$ and the noise covariance matrix $\mathbf{K_n}$. The noise covariance matrix contains mostly sensor (self)noise, surface noise and discrete noise sources that are in fact point sources just like the signal of interest [40]. A method to derive the covariance matrix directly by signal measurements is to take snapshots of length l of the data of each sensor, taking the Fourier transform and use them to create the sample covariance matrix $\hat{\mathbf{K}}(f)$ as shown in equation 18.

$$\hat{\mathbf{K}}(f) = \frac{1}{L} \sum_{l=1}^{L} \mathbf{R}^l(f) \mathbf{R}^l(f)^H \tag{18}$$

where L is the number of snapshots [40] and $\mathbf{R}$ denotes the datavector containing the snapshot on all sensors. When conventional MFP is used, normalized weight vectors $\mathbf{w_c}(A)$ are

constructed for every spatial point $\mathbf{A}$ according to equation 19. From hereon the frequency dependence is removed from the equations for simplicity.

$$\mathbf{w_c}(\mathbf{A}) = \frac{\mathbf{G}(\mathbf{A})}{|\mathbf{G}(\mathbf{A})|} \tag{19}$$

Then the conventional MFP output $B_c(\mathbf{A})$ for every point $\mathbf{A}$ in reference to the true source location $\mathbf{A_{true}}$ is given by equation 20.

$$B_c(\mathbf{A}) = \mathbf{w_c}(\mathbf{A})^{\mathbf{H}}\hat{\mathbf{K}}(\mathbf{A_{true}})\mathbf{w_c}(\mathbf{A}) \tag{20}$$

Using equation 20 a grid of values of B_c can be formed over every parameter variation of $\mathbf{A}$ to search for the best match of $\mathbf{A}$ to $\mathbf{A_{true}}$, which is indicated by the global maximum of $B_c(\mathbf{A})$.

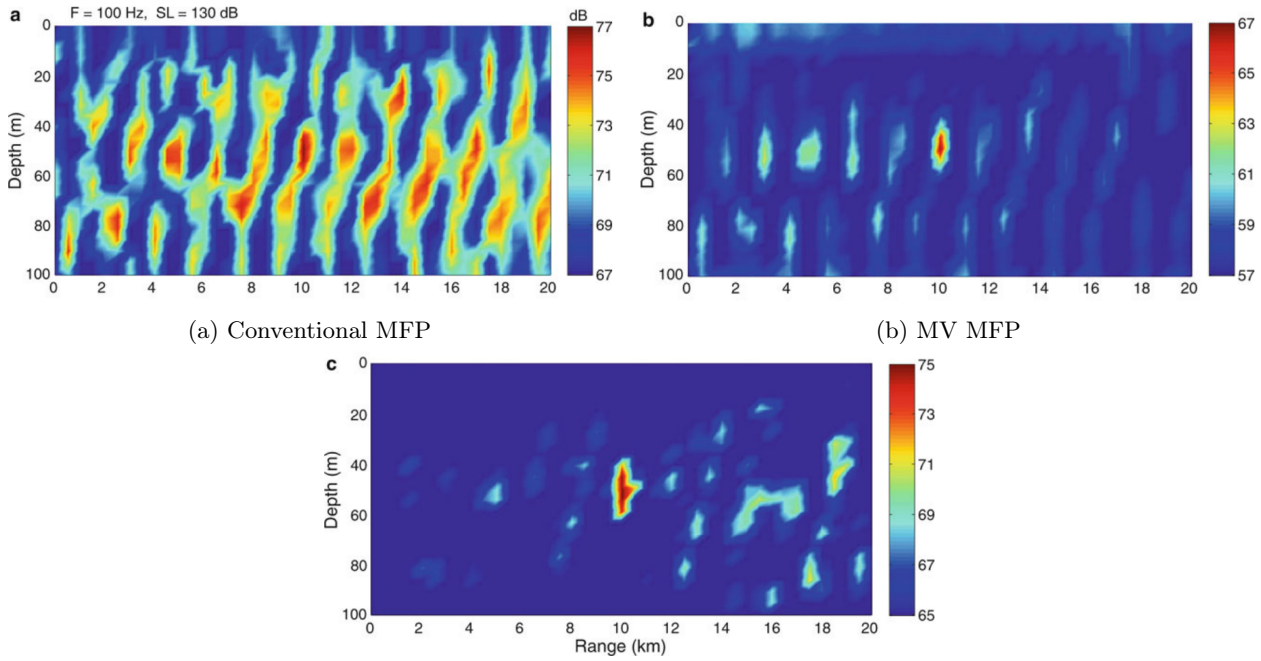
While it is possible to include every spatial parameter in $\mathbf{A}$ to perform localisation in 3D one has to account for the spatial uniqueness of the signal. If the surroundings of the array are not exactly unique in criteria that affect the weight vectors $\mathbf{w_c}(\mathbf{A})$ then there will be ambiguity with one or more variations of $\mathbf{A}$ that are not the location of $\mathbf{A_{true}}$. In that case other MFP algorithms might remove some ambiguity [40].

The MVDR beamformer from section 2.2.1.2 when converted to MFP provides higher resolution and is able to handle ambiguity better than conventional MFP. The MFP algorithm uses equation 21 as the algorithm and equation 22 for the weight vectors.

$$B_{MV}(\mathbf{A}) = [\mathbf{w_{MV}}(\mathbf{A})^{\mathbf{H}}\hat{\mathbf{K}}(\mathbf{A_{true}})^{-\mathbf{1}}\mathbf{w_{MV}}(\mathbf{A})]^{-1} \tag{21}$$

$$\mathbf{w_{MV}}(\mathbf{A}) = \frac{\mathbf{K^{-1}}\mathbf{w_{MV}}(\mathbf{A})}{\mathbf{w_{MV}}(\mathbf{A})^{\mathbf{H}}\mathbf{K^{-1}}\mathbf{w_{MV}}(\mathbf{A})} \tag{22}$$

The increased resolution of this adaptive MFP algorithm however has also an increased resolution in it's parameters, and therefore is more prone to mismatch than the conventional MFP algorithm [40]. So when for instance the sound speed profile in the water column or the bottom profile has some error this will result in a more serious mismatch with the MV algorithm than with the conventional algorithm. For this reason more tolerant algorithms exist, like the Multiple Constraints Matching algorithm [1]. A comparison of the three algorithms with a slight mismatch in a shallow water example is given in figure 9. Promising results are also shown for the Matched Mode Processing algorithm, which has the advantage that replica's can be formed by using only certain modes that are easier to correctly model [40] [41].



(a) Conventional MFP

(b) MV MFP

(c) MCM MFP

Figure 9: Comparison of three MFP algorithms with a mismatch of 2m/s at the bottom of the water column [1]

For this research submerged targets are deemed out of scope, so it might be advantageous to assume all sources of interest are surface targets. This would mean one less parameter to search and match against, which might improve the robustness of MFP. While Dutch coastal areas have some slight depth variation because of submerged sand banks this will probably not influence the sound propagation models enough to enable 2D localisation on the sea surface plane. In that case MFP might still be used for range estimation if it has advantages over other ranging methods.

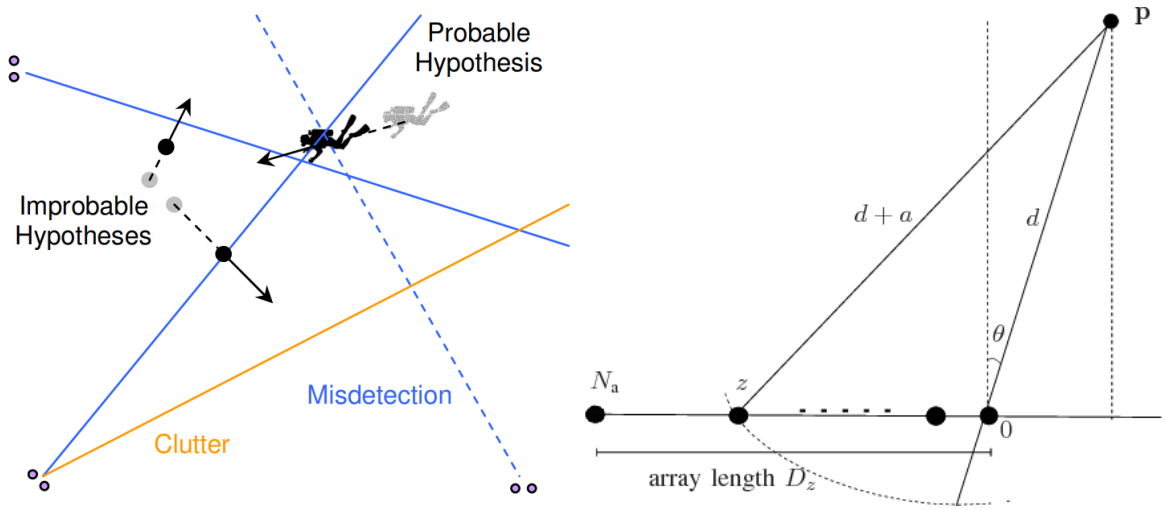
2.2.4 Triangulation

Another probably more obvious method of acoustic localisation would be triangulation. In literature this can be broadly divided into bearing triangulation [42] [43] and TDOA triangulation using wavefront curvature [44] [45] [43]. Triangulation using moving receivers [46] is another method but out of scope for this survey. In the case of bearing triangulation in 2D at least two separate arrays capable of finding the bearing of a target are needed. With this setup the localisation will be ambiguous if the surrounding terrain does not make one of the found locations impossible, but this ambiguity can be removed by using another array. The target's location will be on the intersection of the bearings found by each array as indicated in figure 10a. For the bearing estimations any bearing finding method as described in section 2.2.1 can be used. [43] Proposes the use of two arrays and beamforming to get an ambiguous solution, while [47] uses cross correlation to find the bearing of a target.

Wavefront curvature localisation is only feasible in the nearfield of the array. This nearfield is limited by what is often called the Fresnel range and can be computed using equation 23 [43].

$$R_f = L^2/\lambda \tag{23}$$

Here R_f denotes the Fresnel range, L the baseline of the array which is called D_z in figure 10b and wavelength λ. When considering the optimal propagation frequency of 800Hz, i.e. a wavelength of 1.875m in a 25m deep waveguide [17] with a sound speed of 1500m/s the array



(a) Schematic of bearing triangulation [42]

(b) Schematic of wavefront curvature localisation [48]

Figure 10: Triangulation methods

needs an aperture of more than 43m to be able to perform ranging up to one kilometre. On lower frequencies the array's aperture will need to be even larger. One possible solution is using wide aperture (sparse) arrays for wavefront curvature localisation, i.e. arrays with an inter-sensor spacing larger than half the maximum wavelength [47] [49]. This makes arrays with a large baseline feasible with a minimal amount of hardware, as in theory only three spatially separated measurement points are needed to find the curvature of the wavefront [50]. Studies regarding wavefront curvature used transient signals in experiments, like dolphin clicks [50] or by hammering on a metal pole [47], although the latter also was able to correlate the signal received of a running jet engine which emits a more stationary signal. Surface targets do usually not produce much transients so for wavefront curvature ranging to be applicable to surface targets it should be able to work reliably with continuous signals. In this regard it is possibly limited in the same way as cross correlation based bearing estimation from section 2.2.1.1, as the cross correlation function is mostly used for correlating the signals and finding the time delay [49]. Multipath arrivals might also result in ambiguous solutions to the time delay estimation [43].

2.3 Signature recognition

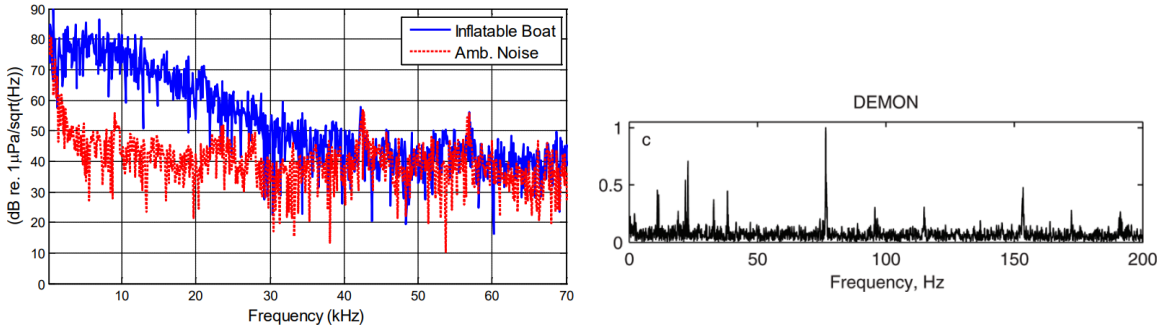
Not unlike human voices machinery also has distinctive parameters in their sound emissions that make it possible to separate them from each other. This is certainly the case for vessels. Being able to discern different vessels from each other is fundamental in a military setting, especially for for instance submarines which largely rely on sound for sensing. Because of this the signatures and their recognition methods are often classified. Still some information on this is available in literature, which will be addressed here.

[19] Cites two widely used methods for signature recognition. The first method is based on the sound spectrum emitted by the target. Vibrations of the engine and other machinery on board will carry over to the water and have higher energy levels on certain frequencies (harmonics) than others. The specific distribution of these energy levels per frequency are characteristic of a certain vessel. An example of this compared to ambient noise is given in figure 11a. Large vessels will have higher energy levels on lower frequencies than small vessels [51]. It was also shown that the frequencies of these harmonics vary with (engine) speed [52] as expected, but no method for predicting how this shift will exactly occur was found.

The other method of collecting a signature is by demodulating the cavitation noise caused by the vessel's propeller and extracting information on the parameters of this propeller [19]. Cavitation occurs on almost all vessel propellers and produces non uniform noise with a maximum frequency inversely proportional to the cavitation bubbles size [53]. Cavitation is stronger in the upwards position where the wake of the hull creates a slower inflow velocity, which means a blade pass rate can be found using demodulation techniques. Often one blade of the propeller will cavitate more than the others. In that case it is also possible to discern a propeller rotational rate [52].

Demodulation itself is performed by the DEMON algorithm, which stands for Detection of Envelope MOdulation on Noise [19]. This algorithm works by windowing the signal, applying a Hilbert transform to the window, computing the root mean square of this transformed window and then performing a fast Fourier transform to extract the frequency data [21]. An example of this is shown in figure 11b. While normally the signal of one hydrophone is used for DEMON it is also possible to use the cross correlation of two or more hydrophones [19] [21], in cases where the cross correlation is already used for bearing estimation. This can also be reversed, where narrowband components of the DEMON signal of two hydrophones are selected and the phase difference is computed between the two components to find a direction [52].

The added benefit of being able to recognise certain vessels can be that possibly hostile vessels that might be used for espionage or sabotage can be recognised if a signature is available. When such a vessel approaches a monitored area a different reaction by law enforcement might be applicable than when a known fishing vessel approaches the same area. In the case of demodulation there might be added benefits in direction finding as the demodulated signals have much increased wavelengths compared to the modulated signals, possibly making a larger array aperture feasible while still maintaining the spatial Nyquist criterion as addressed in section 2.2.1.1.



(a) Example of the PSD of a small inflatable boat [51]

(b) Example of DEMON output [19]

Figure 11: Examples of signature outputs

An interesting observation made while writing this report is that cepstral analysis might aid in the extraction of signatures. The quefrency of horizontal lines in the cepstrum of the measurement shown in figure 5 were found to exactly match the demodulated frequency with the highest level found using the DEMON algorithm. For the measurement shown in figure 7 the horizontal lines visible there did not match the demodulated frequency with the highest level but one of the secondary lines. A possible explanation for this is that cepstral analysis will pick up sidebands and harmonics created by the shaft rotation. As these sidebands and harmonics will have a periodicity in frequency these will show up in a cepstrum. Further research into this observation might be interesting.

2.4 Project proposal

After conducting the literature survey the project proposal from section 2.4.1 was drafted. Soon it was discovered however that this plan was too ambitious to complete within the time allocated for this thesis, as coding issues with wavefront curvature ranging were found. Therefore it was decided to simplify the research and to focus solely on cepstral ranging. This algorithm was chosen as ranging is a much more complex task than bearing estimation, and cepstral ranging showed potential in retrieved literature.

The original project proposal has been kept in this thesis for the sake of completeness. The revised project proposal of the project that was actually carried out can be found in section 2.4.2.

2.4.1 Original project proposal

This report provided an overview of techniques found in literature for localisation and recognition of surface targets from a stationary, passive sonar system in shallow water. Almost all retrieved literature focused on one specific algorithm, or at most provided a comparison on different algorithms in the same family. If a surveillance system for the aforementioned case would have to be designed first a thorough investigation and comparison of the techniques mentioned in this report should be made to be able to select the optimal localisation techniques for this system.

As the comparison mentioned above is not available it is proposed that one is made. This research project should include all relevant algorithms and methods as mentioned in this report and make an objective and preferably quantitative overview and comparison of these methods to be used as a guideline for the design of such a surveillance system. Different combinations of methods can be proposed to achieve improved localisation performance. The setting of this research should be centered around the Dutch littoral zone, i.e. the North sea, where most infrastructure with Dutch interests are found. An excellent case in point would be a typical wind farm in the North sea as it is expected that wind farms would be one of the use cases of such a surveillance system.

Unfortunately acoustic simulations are known to give incorrect results with even the smallest error in bottom- or sound speed profile. Other errors are introduced compared to real world performance by the small temporal deviations in SSP, scattering and signal that are impossible to simulate accurately. Consequently it is proposed to base the comparison as much as possible on experimental measurements. This way it is ensured that the obtained results are accurate and additionally it provides some insight on the real world performance of algorithms that are only tested in simulations. TNO has some relevant data from previous experiments available and additional measurements will be taken in TNO's anechoic basin and the port of Scheveningen. If these are not sufficient additional measurements can be taken on other locations. Acoustic simulations might still be useful in a supporting role to conduct sensitivity studies for instance.

When sufficient data has been gathered and analyzed the performance of the different algorithms can be compared. Parameters like the maximum functional distance, resolution, size of systematic errors, accuracy, resistance to disturbances, reliability on low SNR and ability to discern between multiple sources can be used to quantify the results. It is doubtful one algorithm can be picked as the 'best' algorithm as the use cases might differ significantly, but the goal of this study will be to show the strengths and weaknesses of these algorithms in different relevant settings.

2.4.2 Revised project proposal

This report provided an overview of techniques found in literature for localisation and recognition of surface targets from a stationary, passive sonar system in shallow water. Of these techniques the cepstral ranging method was identified as an algorithm of interest in regard to surveillance of the Dutch coastal areas. Ranging is the more challenging part of hydroacoustic localisation and cepstral ranging seems to offer accurate results with a comprehensive visualisation in TDOA.

While some literature is available on cepstral ranging which sometimes shows impressive performance, the circumstances of these results and parameters of interest for cepstral ranging are often unclear. It is therefore proposed to conduct an in-depth analysis of the performance of cepstral ranging in a real-world situation. To accomplish this data has to be gathered, preferably recorded on a location of interest like the Dutch littoral zone. Algorithms applicable to cepstral ranging should be developed, their key parameters should be identified and potential problems with cepstral ranging should be addressed.

Unfortunately acoustic simulations are known to give incorrect results with even the smallest error in bottom- or sound speed profile. Other errors are introduced compared to real world performance by the small temporal deviations in SSP, scattering and signal that are impossible to simulate accurately. Consequently it is proposed to base the comparison as much as possible on experimental measurements. This way it is ensured that the obtained results are accurate and additionally it provides some insight on the real world performance of algorithms that are only tested in simulations. TNO has some relevant data from previous experiments available and additional measurements will be taken in TNO's anechoic basin and the port of Scheveningen. If these are not sufficient additional measurements can be taken on other locations. Acoustic simulations might still be useful in a supporting role to conduct sensitivity studies for instance.

Using the developed algorithms and measurement data recorded in an area of interest the effectiveness of cepstral ranging should be researched. This should preferably result in quantative data expressed in parameters like the expected error or maximum range so that during future work a comparison can be made between cepstral ranging and other ranging methods. The goal of this study will be to identify the strenghs and weaknesses of cepstral ranging and identify problems that should be addressed before cepstral ranging can be fielded in a practical setting.

2.5 Conclusion of literature survey

There is a vast amount of literature covering passive sonar techniques as during the cold war significant efforts were made to perfect the art of ASW. However most efforts went into techniques most suitable for this tactical scenario, i.e. beamforming algorithms that work well with the large towed arrays or the SOSUS system that increase the detection range with their array gain. Now with the large increase of infrastructure at sea and global tensions rising the stationary surveillance of said infrastructure becomes more and more relevant.

This report focused on promising techniques to detect and localize surface targets with stationary, passive sonar systems, including the challenges that need to be overcome. A number of bearing estimation, ranging, direct localisation and signature recognition algorithms were addressed which can be added in varying combinations to arrive at a functional system to provide surveillance over critical underwater infrastructure in shallow water.

The multipath environment created by the shallow water waveguide brings both advantages and disadvantages. While it might increase the amount available methods to perform ranging like cepstral analysis there is also a risk of interference of the echoes when using cross correlation algorithms, especially when performing wavefront curvature localisation. If the interference is too severe for cross correlation to find a direction reliably beamforming might have an advantage at the cost of an increased amount of sensors. The MUSIC algorithm shows excellent DOA resolution but has as a distinct disadvantage that the amount of sources need to be known in advance. While it might be possible to first identify the amount of sources with another algorithm and then compute accurate bearings using MUSIC the presence of grating lobes must be eliminated in that case, as these grating lobes might present themselves as false positives or hide an actual source.

For ranging a number of suitable methods are presented. Cepstral analysis is relatively new in sonar applications and shows varying results in accuracy and range. The maximum range of cepstral analysis seems to increase when more reflections are included. Best results seem to be achieved when using replica fields derived from actual sources instead of propagation models. The waveguide invariant method uses similar analysis methods based on a different physical principle. While cepstral analysis is based on multipath interference the waveguide invariant method is based on the interference of multiple propagation modes. It might be possible to combine these two methods to perform ranging on short and long ranges. It might be that the long range cepstral ranging [11] already used the waveguide invariant to achieve its longer range compared to other studies, as not analytical models but replica fields from actual sources were used.

Matched Field Processing has in theory the ability to locate sources directly using a single array. It is however doubtful if this can be achieved in the shallow North sea where the lack of spatially unique features might result in a significant amount of ambiguity. In that case MFP might still be used for ranging, but the amount of sensors needed is probably a disadvantage compared to other ranging methods. MFP is also prone to mismatch and needs accurate information on the waveguide, which can vary significantly from both a spatial and temporal point of view.

Triangulation methods are potentially reliable and easy to implement. They however need a large aperture or sensor spacing to achieve a sufficient range and accuracy, and wavefront curvature might result in some ambiguity if there are significant reflections.

Finally a short overview of simple signature recognition methods is given. The recognition of vessels might aid in protecting an area of interest and demodulation can provide extra tools to perform localisation as it provides new signals to be used with a much increased wavelength.

3 Methodology

The research outlined in the sections 1 and 2.4 will be carried out using the data, methods, algorithms and equations as described in this section. While some parts or solutions were extracted from literature, most had to be built specifically for this thesis. Fortunately TNO possessed an experimental dataset which included almost all data necessary for this thesis, and further data could be acquired through open-source means, e.g. Rijkswaterstaat.

First the experimental data is described, which includes audio, GPS, AIS and depth data. Subsequently the algorithms needed to build and enhance the cepstra are discussed, followed by methods to determine the TDOA. Next the algorithms needed for cepstral ranging are investigated, while not an exhaustive list because of time constraints, preliminary research showed that these algorithms offered the best chance at successful cepstral ranging. Finally the method by which the accuracy of the described cepstral ranging methods will be measured is discussed.

3.1 Data collection and experimental setup

Cepstral ranging was tested on an already existing dataset made available by TNO. This dataset was recorded on 17-05-2022 using a vector sensor, which uses three sensors to measure particle motion for direction finding and a regular pressure hydrophone for a total of four recorded channels. This device was placed at a depth of 14.6m at the time of placement in the North Sea near Scheveningen, with a latitude of N52.1164 and longitude of E4.2314. A RHIB with a GPS receiver of which the location was constantly recorded performed a series of experiments and runs around the sensor. Especially these runs as shown in figure 12, which are performed under full power and thus emit a maximum amount of sound energy are useful for this thesis.

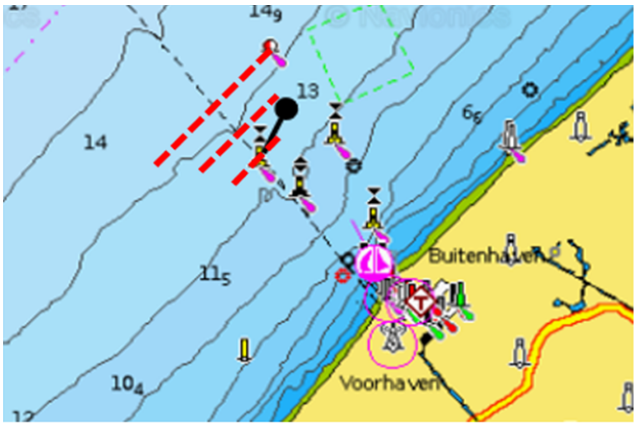


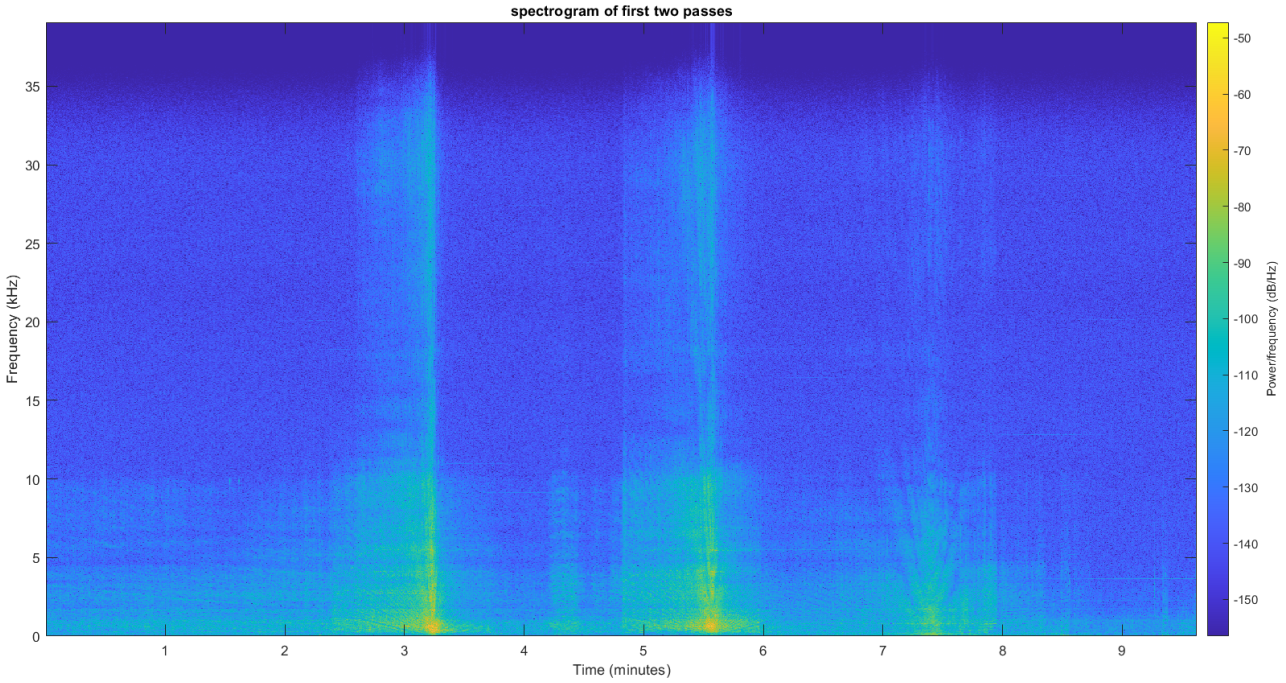
Figure 12: Multiple high-speed run tracks marked in red, sensor location marked with the black 'balloon'

3.1.1 Audio data

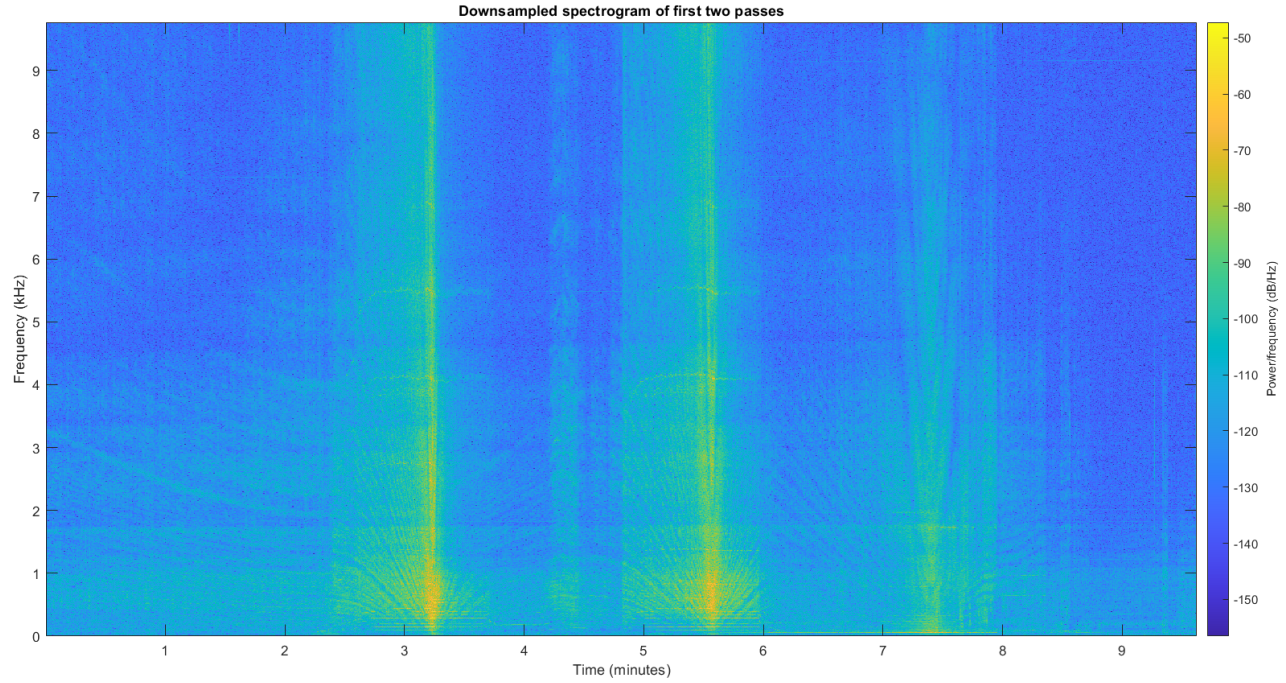
For this thesis only the data from the hydrophone was used, which was recorded with a sample rate of 78125Hz. While the vector data might produce a bearing to the target it was deemed out of scope as explained in section 1.4. As the pressure sensor was part of the vector sensor setup, which are inherently dependent on wavelength for their operation, there was a filterbank included that unfortunately also filtered the hydrophone sensor data above 10kHz. The consequences are clearly visible in the spectrogram of figure 13a, where most sound energy disappears at approximately 10kHz. Therefore it can be expected that downsampling with a factor of approximately 4 (shown in figure 13b) will not result in a ranging performance degradation, while reducing the amount of data that needs to be processed and thus decreasing the computational load. Important to note is that the third Lloyd mirror pattern in figure 13 between the 7 and 8 minute mark is not caused by the RHIB, but by another unknown sound source (presumed to be a vessel). The presence of another vessel gives the opportunity to check if the cepstrograms of different vessels will vary.

The vector- and sound data was stored in a sequence of .WAV files of each 264kB, with a recording time of 4 minutes and 49 seconds. The starting time of the recording has been included in each title. Both the time of the .WAV files and the .GPX file, which is detailed in section 3.1.2, were recorded in the UTC time standard. Therefore there is a time difference of -2 hours compared to the local time at the moment of recording which had to be corrected to enable comparisons with written notes of the experiment. Additionally, it was found that there was a time shift of 3 seconds between the .GPX data and the .WAV files after comparing the Closest point of Approach or CPA of multiple passes between these two files. Finding the CPA for the GPS data is trivial, and the CPA for the audio data was identified by either finding the minima in the Lloyd mirror patterns and maximum sound energy of generated spectrograms or by finding the maxima of quefrency values in cepstrograms.

These .WAV files were generated for the entire three days that the sensor was left on the seafloor. RHIB measurements and experiments were only carried out on the first day of recording and this thesis only uses the data of this first day.



(a) Full spectrogram



(b) Downsampled (4x) spectrogram

Figure 13: Comparison between full and downsampled spectrogram of the first run back and forth, directly over the sensor setup

3.1.2 GPS and depth data

Next to the sound files a .GPX file was included that constantly recorded the exact location of the RHIB in a single track segment with a temporal resolution of 1 second, spanning the entire day of measurements. Additionally, a number of waypoints was included, which among other things contained the location of the hydrophone, the starting and stopping locations of the high speed runs and the location of airgun measurements. These points were also described in a short measurement report, which additionally detailed the depth of the receiver as 14.6m below sealevel.

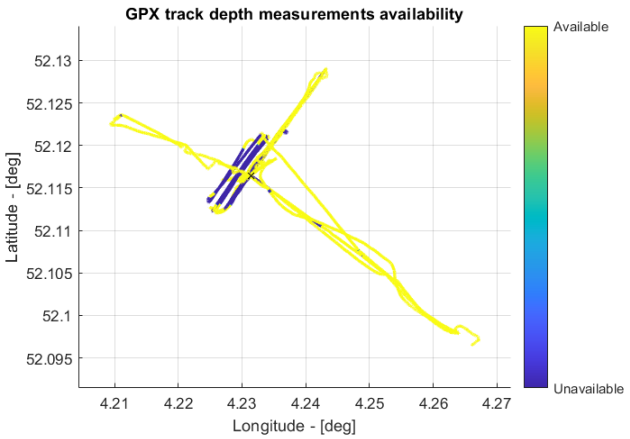


Figure 14: The availability of depth data in the GPS track recording

While analysing this .GPX file it was found that the file also contained a segment of depth measurements. These depth recordings were available for most gps coordinates, except for the high speed track segments. The depth sensor presumably could not perform measurements on high speeds. An overview of the entire track and the availability of depth data is shown in figure 14. These depth measurements can later be used for comparison against available depth charts and measuments from third parties to enhance the accuracy of the multipath prediction for ranging.

3.1.3 AIS data

Finally a set of .CSV files containing recorded AIS messages sent and received in the vicinity of the experiment for the entire three days was included in the dataset. These messages were recorded with a temporal resolution of approximately 1 minute and contained all message types. Of these message types only the type 1, 2, 3 and 18 are relevant for this thesis as these messages contain position reports. The .CSV files were imported into MATLAB and filtered for relevant message types, data and proximity to the sensor setup with a custom-made function which is presented in appendix A.1. The data distilled from the AIS messages is described below:

- **Time of message** This section contains the timestamp of when the message was received. These timestamps are, just like the GPS data, presented in the UTC format and thus have to be calibrated to local time by shifting them 2 hours ahead.
- **Message type** This section contains the message type, which shows what kind of information is contained in the message. Only the message types 1, 2, 3, and 18 contain position reports, where type 1 to 3 are sent by class A AIS devices. These transmitters are placed on larger vessels with a length of >20m where AIS is mandatory. Type 18 messages are sent by class B AIS devices where the transmission of AIS data is voluntary.
- **Latitude** The latitude of the vessel at the time of sending.
- **Longitude** The longitude of the vessel at the time of sending.
- **MMSI** The Maritime Mobile Service Identity number of the sending vessel. This is the unique ID of the AIS transmitter and vessel, which is used to identify the sender of the message.
- **Speed** The speed of the vessel in knots at the time of sending.

As the locations of recorded AIS messages were often at a great distance to the sensor setup and therefore no signals of these sources could possibly be recognised on the sensor, only messages sent in a grid of a longitude between 4.1943° and 4.2822°, and a latitude between 52.0815° and 52.1440° were used.

3.2 Cepstrogram generation

In this section the processing of the audio data from section 3.1.1 into a cepstrogram is detailed. The principle behind cepstral analysis for ranging purposes has already briefly been addressed in section 2.2.2.2 of the literature survey. In summary, a cepstrum is generated by application of a Fourier transformation on a signal. The generated signal in the frequency domain is then enhanced and the echo and main signal are deconvoluted by computing the logarithm of the spectral amplitude, after which the signal is brought back by an *inverse* Fourier transform. This is reiterated in equation 12 shown below [35].

$$C_p = \mathcal{F}^{-1}(\log(F_{xx}(f))) \qquad \text{(12 revisited)}$$

While different types of cepstrum exist, e.g. the complex cepstrum which retains phase information, real cepstrum or power cepstrum, only the *power* cepstrum C_p is used in this thesis. The literature survey showed that using the power cepstrum for ranging applications is the norm, and personal testing indicated that the power cepstrum is more computationally efficient than the complex cepstrum and offers better contrast for echo detection than both the complex and real cepstrum. Because phase information is lost and the signal is deconvoluted during the computation of the power cepstrum, the inverse Fourier transform does not result in a return of the signal to the original time domain. However, the cepstrum is expressed on a timescale called *quefrency*, usually in seconds.

$$x(t) = s(t) + \alpha s(t - \tau) \qquad \text{(13 revisited)}$$

$$|F_{xx}(f)|^2 = |S(f)|^2(1 + \alpha^2 + 2\alpha\cos(2\pi f\tau)) \qquad \text{(14 revisited)}$$

$$C(f) = \log|F_{xx}(f)|^2 = \log|S(f)|^2 + \log(1 + \alpha^2 + 2\alpha\cos(2\pi f\tau)) \qquad \text{(15 revisited)}$$

The revisited equations 13, 14 and 15 displayed above show how the cepstrum computation results in the deconvolution of a signal $s(t)$ and an echo with amplitude α and time delay τ. Equation 15 implies that an echo will result in a periodic set of peaks on the quefrency value equal to τ and its rahmonics (the name for harmonics in a cepstrum). This ability to detect echoes, even in stationary signals, is the basis for cepstral ranging as it makes it possible to use the multipath structure of a shallow water waveguide for ranging.

3.2.1 Cepstrum parameters and the implications of discrete data

In this specific case the cepstra have to be computed from a large discrete dataset as discussed in section 3.1.1. This means that a cepstrum generated from this data will also be discrete. As cepstra are returned to a time unit without removing sample points the quefrency spacing will be equal to the spacing between the original data points as shown in equation 24.

$$\Delta_\tau = \frac{1}{f_s} \qquad (24)$$

It is important to note that therefore reducing the sample rate f_s will result in a lower quefrency resolution. The implications of this trade-off between computational load and cepstral resolution will be researched later in this report.

The maximum quefrency that can be detected in a cepstrum also depends on both the amount of data fed into the algorithm that computes the cepstrum and the sample rate of the

data. This is comparable to how the maximum frequency visible in a spectrum is equal to the Nyquist frequency. Just like a discrete spectrum generated by a Fast Fourier Transform or FFT the power cepstrum algorithm will produce a cepstrum that is symmetrical. While the first value is unique, the second value will be equal to the last value, the third value will be equal to the second-last value and so on. The total length of the data vector produced by the power cepstrum algorithm will be equal to the length of the data vector fed to it. However, the amount of unique data is less and can be computed using equation 25

$$K_{\text{quefrencies}} = \frac{N_{samplepoints}}{2} + 1 \tag{25}$$

Of these unique quefrencies and accompanying data the first entry is equal to a quefrency of zero (0). Combined with equation 24 the maximum quefrency value τ_K, expressed in seconds, can be computed using equation 26.

$$\tau_K = \frac{N_{samplepoints}}{2f_s} \tag{26}$$

This relationship can be used to compute the minimum block-size necessary to see all relevant quefrencies. The time-delay between direct path and the reflected paths increases as the source gets closer to the receiver, but approached a limit. When the source is directly overhead from the receivers point of view the time-delay is at its maximum. As shown later in this report the Bottom-Surface or BS path was found to be dominant. As it might be obvious from this title, this path leaves the source (on the surface) and reflects after reaching the bottom to the surface, from which it reflects to the bottom mounted receiver. Consequently, the maximum quefrency value of interest is equal to approximately twice the time sound needs to travel from the surface to the bottom in the area of interest. For example, when using the measured depth as explained in section 3.1.2 and a sound speed of 1500 m/s the maximum quefrency value of interest is equal to approximately 19.5ms.

3.2.1.1 Expected cepstral shape

The dominance of the BS path can be anticipated when computing expected time-delay values for various propagation paths. The following equations were used to accomplish this, based on equations published in [2] and an expansion of these by own work for the third order reflections of the SBS- and BSB path. First the path length of the direct path (equation 27) and the various reflection paths (equations 28 to 33) are computed, after which the time-delays can be derived using these results and equation 34.

$$R_D = \sqrt{X^2 + (H_r - H_t)^2} \tag{27}$$

$$R_S = \sqrt{X^2 + (H_t + H_r)^2} \tag{28}$$

$$R_B = \sqrt{X^2 + (2H_s - H_t - H_r)^2} \tag{29}$$

$$R_{SB} = \sqrt{X^2 + (2H_s + H_t - H_r)^2} \tag{30}$$

$$R_{BS} = \sqrt{X^2 + (2H_s - H_t + H_r)^2} \tag{31}$$

$$R_{SBS} = \sqrt{X^2 + (2H_s + H_t + H_r)^2} \tag{32}$$

$$R_{BSB} = \sqrt{X^2 + (4H_s - H_t - H_r)^2} \tag{33}$$

$$\tau = \frac{R_{\text{reflection}} - R_D}{c} \tag{34}$$

In these equations X refers to the horizontal distance between source and target, in other words the all-important parameter for this thesis, H_s to the vertical distance between bottom and surface, H_t to the vertical distance between source and surface and H_r to the vertical distance between receiver and surface. A graphical representation of these parameters is given in figure 15.

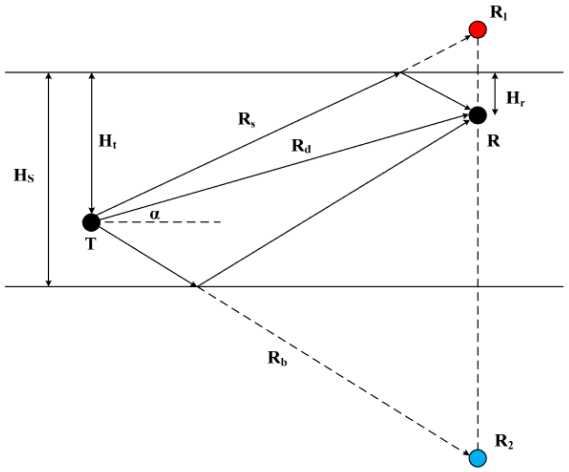


Figure 15: Representation of time-delay computation parameters [2]

These equations only hold for a mirror smooth bottom profile, something which certainly is not representative of the seafloor near Dutch coastal areas. They do however provide some insight on expected time-delay magnitudes and how their values will change for various ranges. Another issue pointed out by these equations is the inclusion of source depth as a parameter. While this thesis solely targets surface vessels it is assumed that even this class of vessels will exhibit variations of effective source depth. The primary sound source of surface vessels can be assumed to be propeller cavitation, and the propeller of a large cargo ship will be located several meters deeper than the propeller of a RHIB with an outboard engine. In theory equation 28 to 33 provide a system of equations that can be solved for both the range X and source depth H_t, although it is unsure if all these different propagation paths will be clearly recognisable in the cepstra, especially when considering that these will inevitably have clutter from rahmonics.

The different approximate time-delay values have been plotted using the equations above in figure 16. As parameters a depth of $H_s = 14.6m$, a receiver depth of $H_r = 14.3m$, a source depth $H_s = 0.4m$ and a sound speed $c = 1500m/s$ were used.

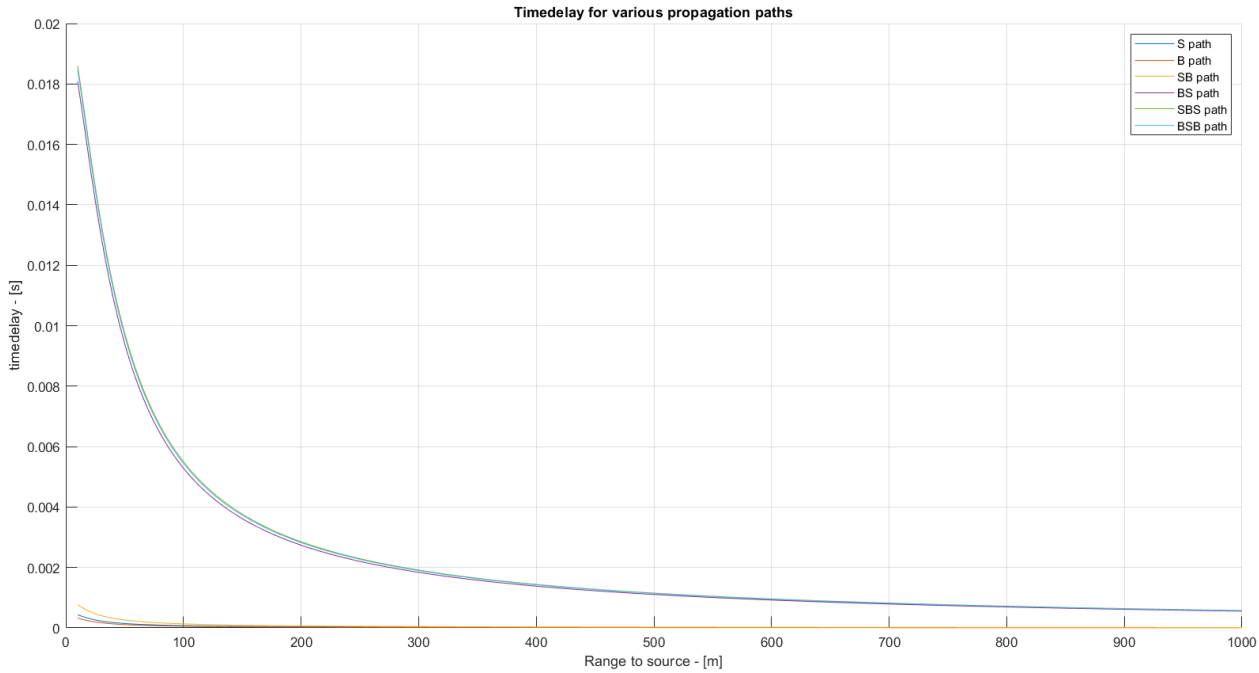


Figure 16: Expected time-delay values for various propagation paths at range

This figure shows that the reflections for the SB, SBS and BSB paths are tightly clustered together for virtually all ranges. The shorter paths are severely dependent on the distance between either source and surface or receiver and bottom for a significant time-delay over the direct path. Since these distances are very short in this case these shorter paths result in a time-delay that is only recognisable when the source is extremely close to the receiver. Therefore it can be expected that the longer paths will dominate in the analysis of the measurements described in section 3.1. Figure 16 also shows a shape similar to an exponential function, where the time-delay sensitivity swiftly decreases at range. This might contribute to larger ranging errors at range, together with the loss of SNR.

As the power cepstrum algorithm makes use of both the FFT- and iFFT algorithms it has the same limitations when it comes to block-sizes. The FFT algorithm is much faster when the block-size is equal to a power of two (2). Therefore it is advisable to choose a block-size of a power of two, close to the preferred size as computed. It is also possible and probably preferable to make use of a window function like the Hann or Hamming windows to reduce spectral leakage during the FFT computation.

3.2.2 From cepstrum to cepstrogram

As mentioned earlier the cepstra have to be computed from a large discrete dataset as discussed in section 3.1.1. Of course, computing one single cesptrum from the recording of a dynamic scenario, i.e. a vessel moving in the vicinity of the sensor, would not provide adequate results as the multipath structure will constantly change. Therefore multiple cepstra are generated to create a K by M sized *cepstrogram*, which is analogue to how a spectrogram is generated from multiple spectra.

To accomplish this the dataset is divided into multiple, smaller sets of a certain size. These smaller sets or windows make it possible to evaluate the available data at multiple points in time, which in this scenario means that ranging is performed at certain intervals. Usually these windows or blocks are created with an overlap with regards to the previous dataset, expressed in n datapoints or a percentage of the window size N. For example, a window size of $N = 100$ datapoints and an overlap of 50% means that the first 50 datapoints of the new window are equal to the last 50 datapoints of the previous window. Using an overlap makes it possible to increase the temporal resolution while still keeping an adequate quefrency range. A typical cepstrogram is shown in figure 17a.

An addition to the algorithm to remove some of the noise in the cepstra is proposed by Kam W. Lo in his paper on cepstral ranging [11]. Here a method is presented that splits the selected block for cepstrum computation in five (5) blocks of half the original block size, with 75% overlap. The cepstra of all sub-blocks are then computed after which the mean of the five cepstra are used as the final cepstrum. The idea behind this technique is that the echoes will stay relatively stationary during this small time-frame and thus will not shift in quefrency, while noise will be random and vary over each cepstrum computation. By taking the mean of the results this noise is possibly removed or reduced. Of course it is imperative that the time span of the blocks is small enough that the echo time-delay will not shift significantly during the sub-block computation, or else the cepstrum peaks might be reduced or washed out by computing the mean as well. It is also important to note that when using this technique the value of τ_K as discussed in section 3.2.1 will depend on the sub-block size, and not the overall block size. The effectiveness of this technique will be tested later in this report.

3.3 Cepstrum processing

The cepstrogram was generated using an algorithm built around equation 12. From this algorithm the cepsta have a very small spacing between minima and maxima as seen in figure 17a which can be solved by normalization like in figure 17b using equation 35, where τ_k denotes the discrete quefrency values τ_1, τ_2, ..., τ_K [11].

$$\hat{C}(\tau_k) = \frac{C(\tau_k) - \min[C(\tau_k)]}{\max[C(\tau_k) - \min[C(\tau_k)]]} \tag{35}$$

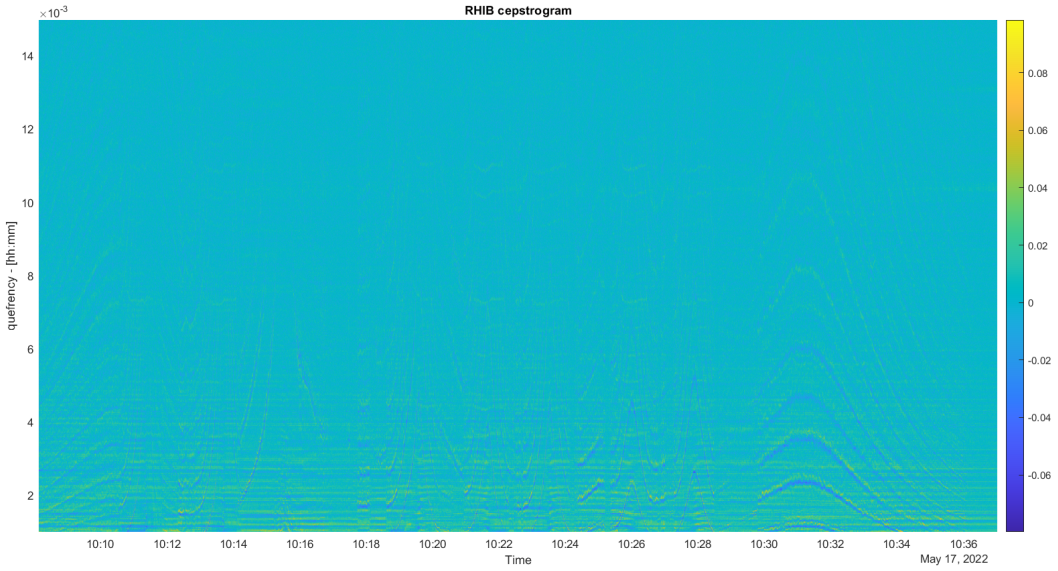
The normalization result $\hat{C}(\tau_k)$ however results in a larger mean value of cepstra that were recorded at a higher overall sound level, i.e. with the sound source nearby, which is clearly visible in figure 17b. This can be solved by subtracting the mean of every cepstrum from itself as shown in equation 36, which results in figure 17c. This subtraction removed the difference in contrast between individual cepstra and is denoted by $\tilde{C}(\tau_k)$.

$$\tilde{C}(\tau_k) = \hat{C}(\tau_k) - \frac{1}{K}(\sum_{k=1}^{K} \hat{C}(\tau_k)) \tag{36}$$

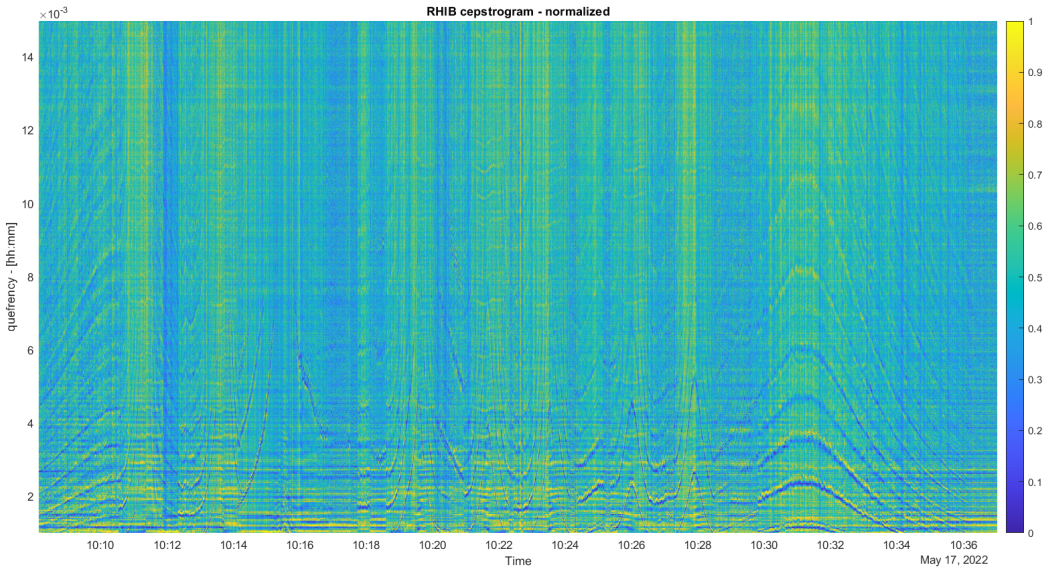
Close observation reveals that there are horizontal lines at specific quefrencies visible. These are presumably caused by irrelevant, stationary processes or noise. By taking the total mean of the cepstrogram, formed by M cepstra, for each quefrency bin and subtracting this from each respective quefrency bin some of these stationary signals can be removed. One should note that this cannot be done to individual cepstra as one needs a significant amount of cepstra, i.e. a cepstrogram, to accomplish this. It should however be possible to accomplish this by either always using the same total mean in regards to time, or by for instance saving the mean of the last few minutes of recording for this operation. For a deployed, functional system it would be possible to record when no vessels are near and use an 'empty' cepstrum generated from this data. This is shown in equation 37, which was acquired from [8], and provides us with the final result of figure 17d.

$$\bar{C}_m(\tau_k) = \tilde{C}_m(\tau_k) - \frac{1}{M}(\sum_{m=1}^{M} \tilde{C}_m(\tau_k)) \tag{37}$$

In addition it is necessary to remove the first few quefrency bins from all cepstra, up to a quefrency of approximately 0.5ms. This should be done before normalization. Unfortunately a cepstrum contains peaks of high amplitude around a quefrency of zero (0) which drown out the smaller peaks generated by echo arrivals.

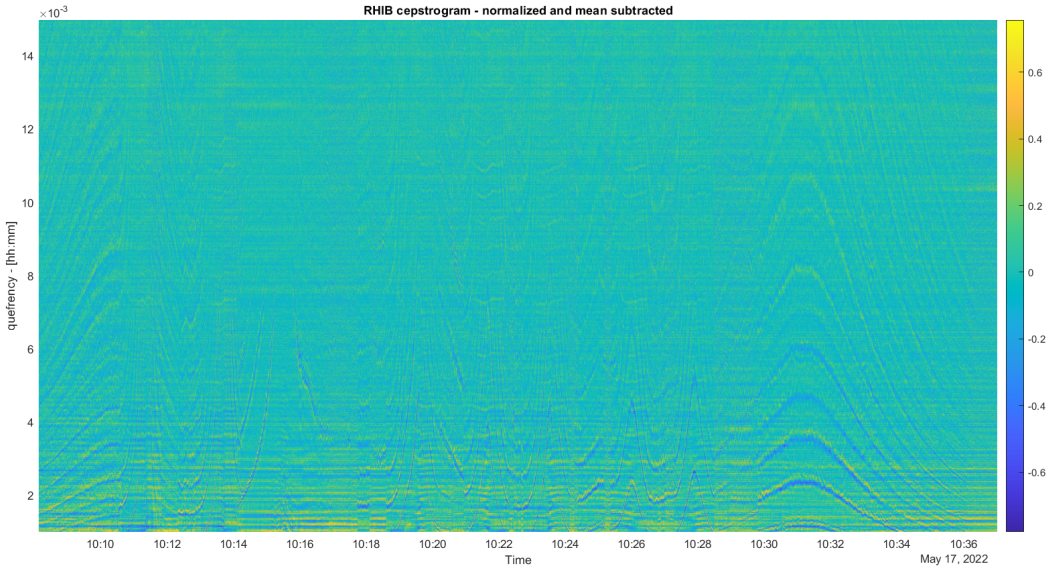


(a) Standard cepstrogram

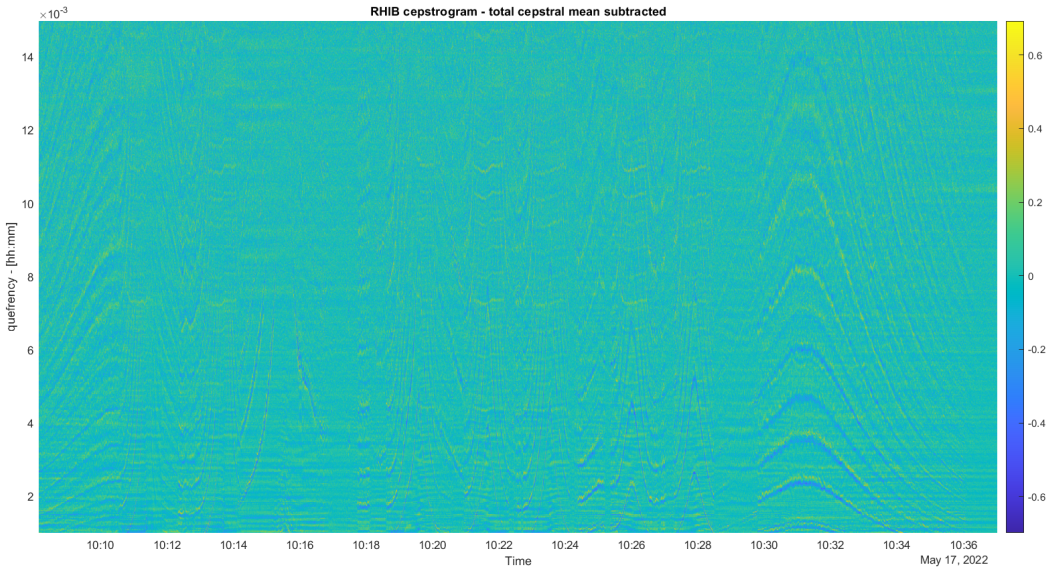


(b) Normalized cepstrogram

Figure 17: Different stages of cepstrogram processing



(c) Normalized cepstrogram with cepstral mean subtracted



(d) Normalized cepstrogram with total and cepstral mean mean subtracted

Figure 17: Different stages of cepstrogram processing

After this final operation that resulted in figure 17d the cepstrum or cepstrogram is ready to be used for ranging. For the remainder of this report a cepstrum normalized and processed using the methods as described in this section will be denoted using $\bar{C}(\tau_k)$ where τ_k denotes the discrete values of the quefrency domain.

3.4 Raytracing for TDOA estimation

Next to methods to build a cepstrum library from experimental data it should also be possible to build one from simulations. In the case of a constant bathymetry the analytical equations as given in section 3.2.1 would suffice. However, this is usually not the case and certainly not in the area where the measurements used in this thesis were taken. This area is sloped because of its proximity to the coast and might contain sandbanks that create some additional spatial variation in bathymetry. An overview of the bathymetry from both the RHIB measurements and open source data in the NAP (Normaal Amsterdams Peil, approximately the average water height in the North sea) standard is visible in figure 18.

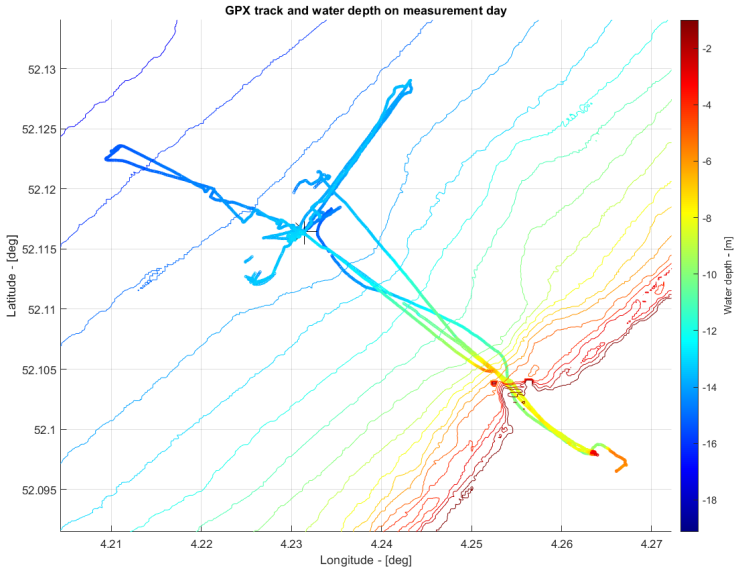


Figure 18: Measured and retrieved bathymetry for the testing area

Therefore it can be assumed that more accurate results can be obtained by estimating TDOA values while taking in account the varying bathymetry. It was decided to build a simple raytracing algorithm to accomplish this task.

3.4.1 Sound speed computation

First, open source data was acquired from Rijkswaterstaat concerning the salinity and temperature of the water near the measurement location for the measurement day. Unfortunately the nearest measurement station that collected this data for two or more depths was located at 'Haringvliet 10' near Rotterdam, about 38km from the measurement location at Scheveningen. This station provided both temperature and salinity readings at 10 minute intervals, and at a depth of both 250cm and 900cm. Both the salinity and the temperature measurements varied either not at all or an insignificant amount between those two depths. The nearby measurement station 'lichteiland Goeree', which provides only readings at 600cm depth but is placed in deeper water reported very similar measurements. Therefore a raytracing method that does not account for refraction by a change in sound speed in the water column was deemed accurate enough, while greatly simplifying the coding process.

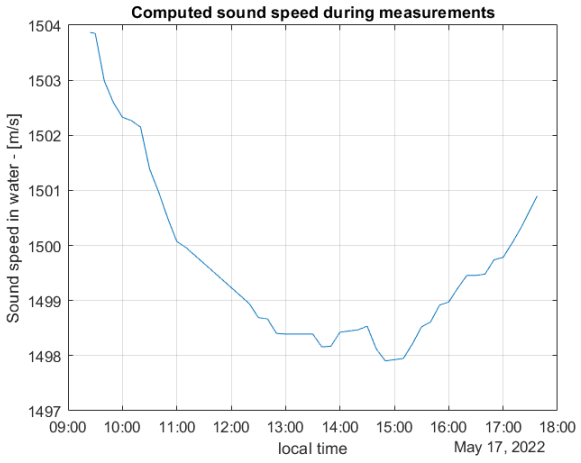


Figure 19: Computed sound speed during measurements

In section 3.2.1 it was explained that the BS path was expected to be dominant in the cepstra, and this was confirmed by preliminary cepstra generation. Because of this the raytracing algorithm was written to solely compute this specific propagation path.

To generate accurate TDOA estimations accurate data on sound speed and bathymetry are essential. Sound speed measurements were not available so equation 38 was used to compute the sound speed from the salinity and temperature measurements which were acquired from Rijkswaterstaat.

$$c = 1449.2 + 4.6T - 0.055T^2 + 0.00029T^3 + (1.34 - 0.01T)(S - 35) + 0.016z \qquad (38)$$

Here c denotes sound speed in m/s, T denotes temperature in degrees Celsius, S denotes salinity in parts per thousand and z denotes depth in meters. According to the source this equation should give reasonably accurate results [12]. Using the data from the measurement day a plot of the sound speed can be generated with a resolution of 10 minutes. As the GPS transmitter on the RHIB provides a position update on every second the sound speed data is linearly interpolated to match this resolution.

The resulting estimated sound speed is visible in figure 19. A spread of approximately 6 m/s can be observed during the measurements.

3.4.2 Bathymetry integration

The most important part for this solution is the bathymetry integration to account for range dependent environments while raytracing. This system is based on open source depth charts available in the .XYZ format, where X contains the WGS84 longitude, Y the WGS84 latitude and Z the water depth in meters at that point. First depth charts from the EMODNET database from 2022 were used with a spatial resolution of $\frac{1}{16}$ arc minutes, which translates to roughly 115 meters. This resolution was deemed unsatisfactory and a different chart was requested from Rijkswaterstaat. This chart has a resolution of 20m but first had to be converted from a GEOTIFF file to an .XYZ file using QGIS software (open source geographic information system). A GEOTIFF file is much more efficient when it comes to storage space but it was

found that an .XYZ file is easier to implement. To compensate for the increase in storage the GEOTIFF file was cropped to the area of measurement before conversion to .XYZ. As the Rijkswaterstaat bathymetry dataset used the 'RD NEW' geodetic datum a conversion to WGS84 had to be carried out as well, which was also possible using QGIS. A resolution of 20m was the highest available at this time but an even higher resolution might offer better results. Once the correct bathymetry has been loaded the complete dataset has to be calibrated for tidal changes in the water column. It would be possible to compute the theoretical tidal height as a forecast, but fortunately a measurement station was available at Scheveningen which measured the water height compared to NAP every 10 minutes. As visible in figure 20, highlighted in orange, a tidal difference of approximately 2 meters occurred during the measurements.

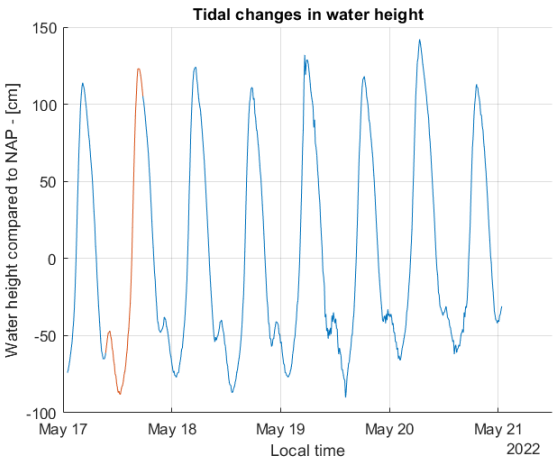


Figure 20: Measured tidal height relative to NAP at Scheveningen, orange denotes when the RHIB was active

The EMODNET file used a datum different than NAP and therefore required some calibration. Fortunately the Rijkswaterstaat file is also relative to NAP, just like the tidal information. This makes it possible to add the tidal height directly to the bathymetry dataset. To accomplish this the tidal dataset was first linearly interpolated over the same timescale as the GPS recording time, just like the salinity and temperature for the sound speed computation. After this the correct water height for each GPS location and its corresponding point in time can be added to the bathymetry dataset.

3.4.3 Raytracing algorithm

With the correct bathymetry, calibrated for tidal effects, and the computed sound speed it becomes possible to estimate the TDOA values for each location using raytracing. The concept of this technique is based on making a vertical 'slice' of the waveguide between the source and the receiver. The probability of this slice containing bathymetry datapoints is very low, so a bathymetry profile is linearly interpolated in the vertical plane between source and receiver, using the calibrated bathymetry dataset. This profile contains a fixed number of interpolation points between source and receiver for uniform computational load and to scale resolution with distance. Some additional datapoints are added to the set *behind* the receiver location to allow some overshoot of rays. An example of the result for the RHIB at 09:35 local time is shown in figure 21.

This figure clearly shows the range dependent bathymetry and implies that accurate cepstral ranging can only be accomplished when accounting for bathymetry.

The interpolated bathymetry is then used to compute the first ray. An angle θ_1 is selected and a uniformly spaced vector x is created, spanning the entire range of interpolated bathymetry for x_1, x_2, ... x_N, with a selected resolution r being the spatial spacing between adjacent values of x. The bathymetry is again (linearly) interpolated over the vector x to have the bathymetry depths y_{b_n}. The corresponding y values for the first ray y_{1_n} are given by equation 39 with y_s being the source depth, i.e. the propeller depth.

$$y_{1_n} = \tan(\theta_1)x_n + y_s \tag{39}$$

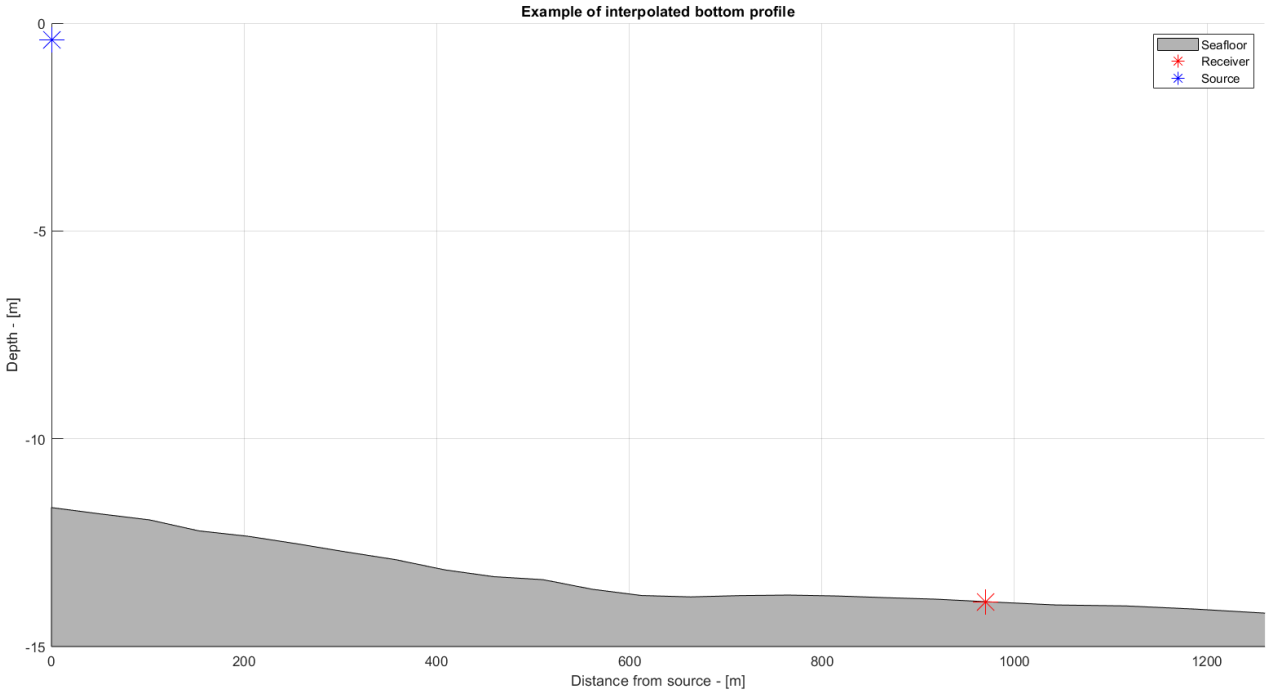


Figure 21: Interpolated bottom profile with receiver and source (RHIB) locations during departure from Scheveningen

The algorithm then computes the absolute difference between the first ray and bathymetry using equation 40 and finds the indices i_1, i_2 of the two smallest results as these belong to the closest points to the intersection between the bottom and first ray. These points s_{i_1} and s_{i_2} are then used to compute the exact intersection using equation 41 using geometric relationships. The result of 41 can be plugged into equation 39 to compute the corresponding y value.

$$s_{i_1,i_2} = \min_{1,2} |y_{1_n} - y_{b_n}| \tag{40}$$

$$xr_1 = x_{i_1} + \frac{s_{i_1} * r}{\sum s} \tag{41}$$

With the exact intersection point known the segment of bathymetry that is hit by the ray is selected and its local angle θ_b used to compute the reflection angle of the second ray. The angle of the second ray θ_2 can be computed using equation 42. It follows that the second ray y_{2_n} can be computed using equation 43.

$$\theta_2 = 2\theta_b - \theta_1 \tag{42}$$

$$y_{2_n} = \tan(\theta_2)(x_n - yr_1) + xr_1 \tag{43}$$

As the sea surface is assumed to be horizontal, smooth and located on $y = 0$ the second reflection point can be computed straightforward using the previous results using equation 44 with the y coordinate yr_2 being the trivial solution.

$$xr_2 = xr_1 - \frac{yr_1}{\tan(\theta_2)} \tag{44}$$

The procedure for the third and final ray is analogue to the procedure for the first ray as they both follow a downward path and end when they intersect the interpolated bathymetry. With the corresponding y values y_{3_n} given by equation 45 the equations 40 and 41 can be used to compute where the final ray will touch the bottom, i.e. (hopefully) the receiver again.

$$y_{3_n} = \tan(\theta_2)(xr_2 - x_n) \tag{45}$$

With the endpoint (x_{3_n}, y_{3_n}) known the entire expected propagation path for the ray can be plotted as shown in figure 22.

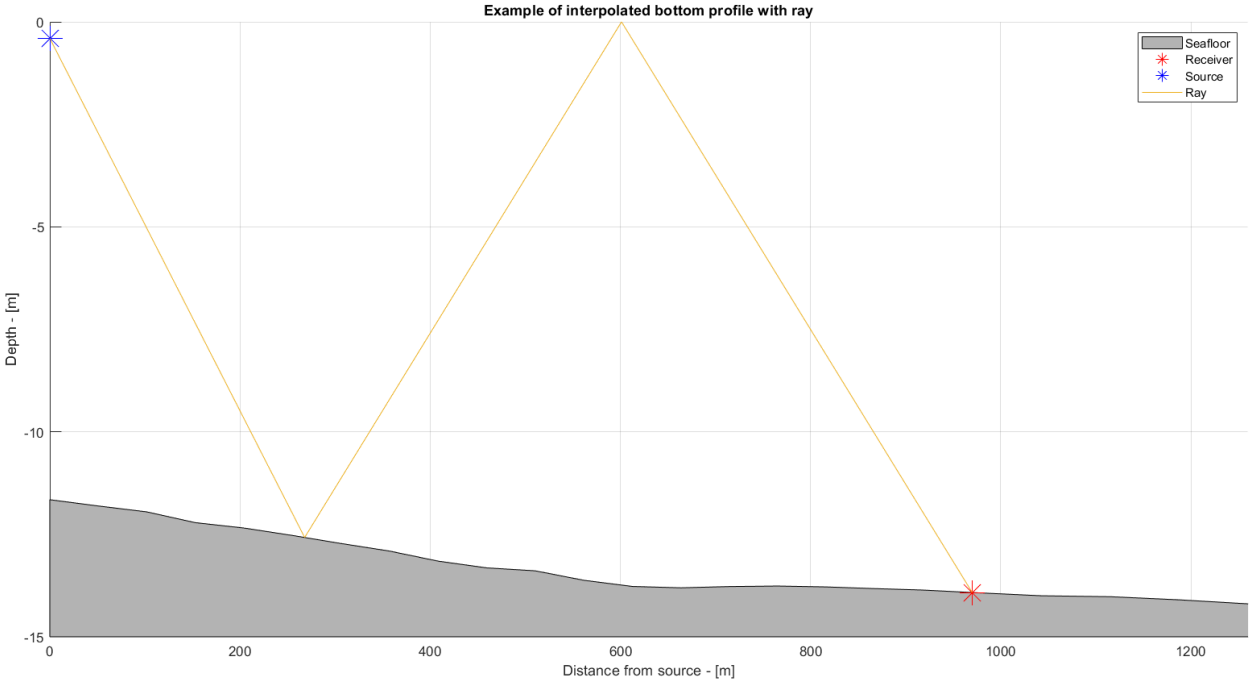


Figure 22: The example from figure 21 with the computed ray included

These equations can be used to compute the total trajectory of the BS-path for a certain starting angle θ_1. Of course some simplifications are made in this model. Refraction in the water is not accounted for and neither is bottom penetration and refraction. Including these might offer better estimation results, but at the cost of requiring a more refined and complicated algorithm with an increased computational load. The accuracy of the estimation process described above will be measured in section 4.

To arrive at a quantity usable for cepstral ranging the expected TDOA has to be computed. This is simply the difference between the total path-length of the computed rays and the length of the direct path, divided by the sound speed at that time as computed in section 3.4.1. The code for the algorithm is shown in appendix A.3.

3.4.4 Iterative solver

The raytracing method from section 3.4.3 only works for a certain starting angle θ_1 and computes the expected path, regardless of whether this ends at the receiver or not. Because of the range dependent bathymetry it is impossible to find an exact solution beforehand. Therefore an iterative process or solver is needed to find the starting angle that *will* result in the ray ending with the receiver.

A relatively simple solving algorithm was developed. While not completely robust to find the global minimum if was found to function without error for this specific case when boundary conditions were satisfied.

First the algorithm from section 3.4.3 is executed for a broad range of angles which should positively include the exact solution, but with a relatively course resolution to decrease computational load. For each computed ray the distance between the endpoint of the ray and the receiver, i.e. error ε, is measured by computing the difference between their respective x coordinates. From these ε values, which can be either negative or positive, the *largest* minimum value and the *smallest* positive value are selected, i.e. the two values closest to zero with opposite

signs. It is reasoned that the exact solution should lie between these two error values and their respective starting angles $\theta_{\max}$ and $\theta_{\min}$.

Consequently a new range of starting angles is formulated, spanning the range between the two angles $\theta_{\max}$ and $\theta_{\min}$ as found above. This enhances the resolution in the area where the solution is to be found. Again the raytracing algorithm is executed for these new values of θ_1 and again two new angles $\theta_{\max}$ and $\theta_{\min}$ are selected. This process continues until the absolute error $|\varepsilon|$ has been reduced to a certain threshold.

Because the first evaluation computes the errors for a broad range of angles which should include the exact solution, the risk of finding a local minimum is reduced. To remove the risk of the solver getting stuck on a local minimum which is not able to satisfy the maximum error condition the solver terminates after 100 iterations, shows an error and selects the minimum value of $|\varepsilon|$ present at that moment. This did not happen for this specific use-case, except when the RHIB was located in the port of Scheveningen and no direct path to the receiver was present. For a more challenging bathymetry that includes steep ridges this solver might not be suitable, although one could argue if this method of ranging itself would be suitable for such a scenario.

The code for the solver can be found in appendix A.4.

3.5 Building a cepstrum library

It is possible to perform the ranging by cepstral analysis automatically by comparing a cepstrum generated from a measurement to a set of cepstra of which their respective distances are known. This is however complicated by varying bathymetry. When using polar coordinates with angle θ and range r, and the bathymetry varies with respect to both θ and r, a bearing finding method should be included as well, and the library should contain cepstra for all bearings and ranges of interest. An additional complicating factor is found when transitioning to a 3D coordinate system, e.g. cylindrical coordinates, to include tidal variations in water depth. To compensate for this the tidal depth should be included as a variable as well.

The sensor setup used in the measurements for this thesis contained a vector sensor and would therefore be able to determine a bearing to a source as well. However, insufficient data was available to build a library that uses the bearing θ as a variable since the amount of passes from the source and the variation in their location was limited. The raytracing method from section 3.4 can be used to create synthetic data for the library, but no satisfying results were yet gained. Therefore it was decided to exclude θ as a variable during this analysis.

Regardless of whether θ is included in the library as a variable, multiple methods of library building are possible. These methods include both experimental and synthetic options.

3.5.1 Experimental methods

On first glance it seems relatively straightforward to generate a library from experimental measurements, as no simulations have to take place. After all, measurements from a real scenario remove any errors that might be present in synthetic models. However, this severely complicates the ability to vary different variables when it comes to bathymetry. In other words, one has to perform a measurement for any possible bearing, range, tide and source depth if depth is assumed to be a variable.

Nevertheless, using experimental data provides a method to quickly generate expected cepstra for ranging purposes. The accuracy of these methods are evaluated using the data from section 3.1.1. As the RHIB that was used as a sound source had its position measured constantly using a GPS receiver, determining its distance to the receiver is straightforward. Using this information it is possible to directly appoint a certain cepstrum to a certain distance, discarding bearing and tidal height as variables.

Using this technique, two different methods are identified:

Single-pass library

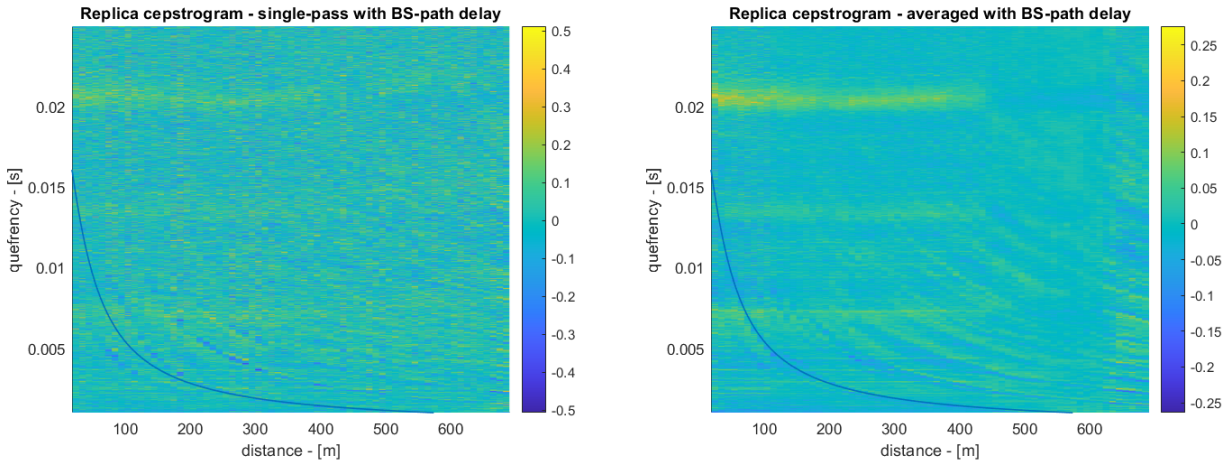
For this library a collection of cepstra is selected that together span a single run of the RHIB closing in on the receiver, passing overhead and moving away again. This section should include all relevant ranges, i.e. the RHIB should start at maximum distance, pass directly overhead for minimum distance and move away to maximum distance again. These cepstra are then sorted on distance and have their respective distance appointed to them. A matching algorithm as explained in section 3.6 can then be used to find the library cepstrum that is most similar to a cepstrum of unknown distance, i.e. a cepstrum generated during surveillance. The appointed range of the library cepstrum is then assumed to be equal to the range of the unknown cepstrum. An example of this library is shown in figure 23a.

Averaged library

This library is similar to the single-pass library, but instead of a single run it used multiple runs on multiple trajectories or locations. The cepstra of these runs are then grouped together on M range intervals, e.g. all cepstra recorded at 200m to 210m are grouped together when a 10m interval is used. The N cepstra of each group or interval are then averaged over each quefrency as shown in equation 46 to generate M library cepstra, one for each interval, denoted by L.

$$L_m(\tau_k) = \frac{1}{N} \sum_{i=1}^{N} \bar{C}_n(\tau_k) \tag{46}$$

This technique has a few advantages. Primarily, false cepstrum peaks caused by another sound source or mechanism are expected to not be located on the same quefrency for multiple measurements and thus will be filtered by taking the mean of multiple measurements. Peaks in the cepstra that are inherent to the specific range will be located on the same quefrency for every measurement and will therefore not be filtered. Secondly, small differences of real cepstrum peak locations in respect to quefrency, caused by small changes in bathymetry when the source location differs in respect to θ, will be averaged. When the bathymetry of a single-pass differs from other passes this is expected to reduce the selective bias created by that specific bathymetry. However, if the quefrencies of real cepstrum peaks differ too much between runs, the peak amplitude might be reduced and result in smearing. An example of this is shown in figure 23b.



(a) Example of single-pass library

(b) Example of averaged library

Figure 23: Examples of experimental libraries, including the expected TDOA of the BS-path for the receiver depth

Figure 23 clearly shows the differences between these two different libraries. While the single-pass library contains more noise, the amplitude of the cepstra has a greater magnitude and the peaks are sharper. Meanwhile the averaged library contains significantly less noise and less defined peaks in the cepstra. The peaks of the single-pass library however do not align accurately with the expected TDOA for the receiver depth, while the averaged library aligns quite well.

3.5.2 Synthetic methods

Another option would be to use the raytracing model from section 3.4 to compute the expected quefrencies of cepstrum peaks. This method also significantly reduces the difficulty of including bearing and tidal height as a variable. After all, only changes in parameters for the algorithm need to be made instead of performing experiments for all the possible combinations of variables. Difficulties arise however during the conversion of the computed TDOA values to cepstra that can be used to match measurement cepstra. A peakfinding algorithm that can detect the TDOA in cepstra seems like the perfect solution to this issue, however most cepstra contain so much noise that no reliable peakfinding algorithm could be found. Attempts to generate replica cepstra from computed TDOA values were unsuccessful as well. Therefore comparison of computed TDOA values will be done by hand in section 4, and automating this process will be recommended to future researchers of this subject.

3.6 Library matching algorithms

When a library containing cepstra with known range has been built, it is advantageous to use an algorithm that can automatically detect the library cepstrum that best fits the cepstrum, generated from a measurement with unknown distance. This way the range of a source can automatically be detected. Unfortunately no satisfying methods were found for TDOA values generated by the raytracing algorithm from section 3.4 so the following matching methods are solely meant for libraries built from experimental data which contain cepstra.

3.6.1 Minimized sum

The most straightforward method to compare two sets of data for similarity would be the method of minimized sum. This method was also used in [11] and subtracts every cepstrum set in the library from the measured cepstrum set for each quefrency value. Subsequently it looks for the minimal sum of the absolute value of the results, as shown in equation 47 [11].

$$\hat{n} = \min_{1 \leq n \leq N} \sum_{k=1}^{K} |\bar{C}(\tau_k) - \bar{C}_n^o(\tau_k)| \tag{47}$$

If two sets of data would be equal for every quefrency value the result of equation 47 should be zero. The selected index $\hat{n}$ indicates the estimated range $\hat{r} = r_{\hat{n}}$ for the source. This algorithm is expected to be computationally efficient as very few operations are required to compute $\hat{n}$.

3.6.2 Pearson Correlation Coefficient

The Pearson Correlation Coefficient or PCC is a well known statistical tool which is used to measure the linear correlation between two sets of data. This PCC can best be described as the ratio between the covariants of two variables and the product of their standard deviations. In the case of discrete datasets like the cepstral library this can be expressed as equation 48.

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{48}$$

Here r_{xy} is the Discrete Pearson Correlation Coefficient or DPCC, n the number of datapoints, x_i and y_i are the individual datapoints of the two sets and $\bar{x}$ and $\bar{y}$ are their respective means. Ranging is performed by finding the largest r_{xy} value after r_{xy} has been computed for the measurement cepstrum and all cepstra in the library.

While the PCC is meant to be used for linear datasets it was found that it is also able to correlate the highly nonlinear datasets of the cepstral library and incoming measurements.

3.7 Accuracy estimation

The results in ranging performance should be compared, preferably quantitative. In the case of comparison of accuracy both the Mean Absolute Error (MAE) and Symmetric Mean Absolute Percentage Error (SMAPE) as shown in equation 49 and 50 respectively offer a convenient single value for an entire measurement sequence. Here the MAE value is expressed in the unit of measurement, in this case meters, while the SMAPE value shows a percentage error. In these equations $\hat{y}_i$ represents the i-th *predicted* value of the observation, y_i represents the i-th *true* value of the observation and n represents the number of measurements.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{49}$$

$$\text{SMAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|\hat{y}_i - y_i|}{(|\hat{y}_i| + |y_i|)/2} * 100\% \tag{50}$$

A comparison relative to true distance might offer insight as well. In this case one can remove the sum and division by n from equation 49 and 50 and calculate these values for every distance. However, an error comparison relative to sound energy or intensity might be more beneficial. Vessels or other sound sources are expected to produce different sound intensity levels. In other words, some vessels will certainly be quieter than others, and it is reasonable to expect a louder vessel resulting in a further maximum range in which it can be detected and ranged, or that it might result in more accurate ranging values.

3.8 From methodology to results

The previously discussed methodological subjects are meant to answer the research questions found in section 1.3. The following list details how the previous sections will be used to answer these questions. Some of these questions have already been partly answered during the methodology research, but their implications will be discussed in the coming sections.

2. Which cepstrum settings and methods of data preprocessing are suitable for cepstral ranging? While this question has been partly answered in section 3.2 and 3.3, the ideal parameters have not been determined yet. The performance of a variation of parameters like block-size and sample rate will be measured in section 4.

3. Which methods of building a ranging library are suitable for cepstral ranging? These methods have been discussed in section 3.5. Their performance will be measured in section 4.

4. Which matching algorithms are available and what is their performance? These algorithms have been discussed in section 3.6. Their performance for different types of libraries will be measured in section 4.

5. What is the influence of (tidal) changes in water depth on the performance on cepstral ranging and how to negate resulting inaccuracies in ranging? The data gathered and algorithm designed in section 3.4 is meant to negate the influence of tidal changes in water depth while also incorporating variations of bathymetry. The accuracy of this method when it comes to determining the TDOA for varying parameters will be measured in section 4.

Finally these research questions will be used to answer the main research question: **1. How accurate can the range of sound sources be estimated in a real-world situation using cepstral based ranging methods and a single hydrophone in relation to the true range between receiver and source?**

Other details on data, e.g. section 3.1.3, will be used to gather additional information on parameters that might influence the ranging potential of cepstral analysis.

4 Results

The methods and algorithms as discussed and explained in section 3 are put in practice in this section, to generate the results necessary to answer the research questions as given in section 1.3. Some unexpected issues were encountered while generating these results, therefore a few subsections also contain an explanation on how these issues were researched and solved. The result subsections contain references to their respective discussions in section 5.

4.1 Cepstrogram parameters

First some research was carried out to find the settings for cepstrum and cepstrogram generation that provide the best results. As made evident in section 3.2 there are a multitude of different parameters available with either continuous or integer possible values. No realistic means of completely optimizing this optimisation problem are available, so only the most significant parameters were varied over a set of realistic parameters.

The block-size and sub block-size were kept fixed on a power of two to keep the FFT algorithm efficient, and it's exact size was determined by the selected sample rate while keeping τ_K above the value of 19.5ms as computed in section 3.2.1, using an edited form of equation 26 which is shown in equation 51, and equation 52 where the $\log 2$ term is rounded to the nearest upper integer.

$$N_{samplepoints} = 19.5^{-3} * 2 * f_s \tag{51}$$

$$BS = 2^{\log 2(N)} \tag{52}$$

It is important to remember here that if a sub block structure is used as described in section 3.2.2, the sub block size determines the value of τ_K and the governing block size has to be increased by a factor of 2.

Considering this block size contains only about 40ms of data, the location of the source is updated extremely often, even with no overlap in cepstrogram generation. To speed op computation a trial was done where only one in five data blocks was used to compute a cepstrum. The resulting accuracy is shown in table 3 and 4.

4.1.1 Effects of sample rate

Determining the ideal, or rather lowest sample rate that does not result in a breakdown of accuracy is essential for increasing the efficacy of cepstral ranging. A lower sample rate means less data to process and thus an increase in computational speed. However, decreasing the sample rate will not only decrease the maximum frequency that is used for cepstrum generation as it lowers the Nyquist frequency, but also the quefrency resolution as shown in equation 24.

Unfortunately the dataset used was filtered above 10 kHz as evident in figure 13a. This makes it impossible to verify a change in ranging performance when higher frequencies are included. Considering that sound attenuation increases with an increase in frequency it can be assumed that using ultrasound for ranging on distance will be ineffective. Still, with this dataset the impact of an increase in Δ_τ, and decreasing the Nyquist frequency under 10kHz can be observed.

Down-sampling was performed by removing a certain amount of equally spaced sampling points from the original audio data. This was tried with and without a filter to reduce aliasing, which was included in the `decimate` function in Matlab. It was found that using this filter had a positive effect on ranging performance, more noticeably as the amount of down-sampling was increased. Therefore all down-sampling was performed using this Chebyshev Type I IIR filter.

Using the `decimate` function down-sampling was compared with a factor of 2, 4, 8, 16 and finally 32. Considering the original sample rate of 78125Hz this results in a sample rate of 39602Hz, 19531Hz, 9766Hz, 4884Hz and 2441Hz. Computational speeds for these sample rates

were compared as well. The results for these sample rates and their computational speeds with their accuracy are expressed in MAE and SMAPE, as explained in section 3.7, and are shown in table 1 and table 2 respectively.

Sample rate performance in MAE					
Down-sample factor	Pearson [m]	Minimisation [m]	Averaged library (Pearson) [m]	Cross correlation [m]	Computation time [s]
1x	156	192	25	158	466
2x	147	151	34	150	722
4x	141	120	39	147	405
8x	128	107	46	151	243
16x	120	110	49	154	157
32x	112	93	49	209	112

Table 1: Sample rate performance in Mean Absolute Error

Sample rate performance in SMAPE					
Down-sample factor	Pearson [%]	Minimisation [%]	Averaged library (Pearson) [%]	Cross correlation [%]	Computation time [s]
1x	46	62	7	47	466
2x	44	47	9	44	722
4x	42	36	11	44	405
8x	38	30	12	45	243
16x	35	30	13	46	157
32x	32	25	13	66	112

Table 2: Sample rate performance in Symmetric Mean Absolute Percentage Error

Sample rate performance in MAE					
Down-sample factor	Pearson [m]	Minimisation [m]	Averaged library (Pearson) [m]	Cross correlation [m]	Computation time [s]
1x	153	170	4	153	149
2x	144	171	9	147	176
4x	139	155	19	149	119
8x	125	130	33	144	94
16x	122	112	44	153	80
32x	116	101	46	198	72

Table 3: Sample rate performance in Mean Absolute Error, negative overlap

Sample rate performance in SMAPE					
Down-sample factor	Pearson [%]	Minimisation [%]	Averaged library (Pearson) [%]	Cross correlation [%]	Computation time [s]
1x	45	45	1	47	149
2x	43	46	2	45	176
4x	41	42	5	45	119
8x	36	35	8	44	94
16x	34	30	11	45	80
32x	32	27	12	61	72

Table 4: Sample rate performance in Symmetric Mean Absolute Percentage Error, negative overlap

These results were generated using the methods described in sections 3.5 and 3.6. As input data the runs close to the receiver at low tide, as detailed in section 3.1, were used. The distance during this entire time-span is shown in figure 24. All matching algorithms except the 'averaged library' used only the data in the orange segment to build a library, this segment was chosen because it includes all ranges of all other segments. The averaged library used the entire range visible in figure 24 to generate a library for the specific results of the tables shown above.

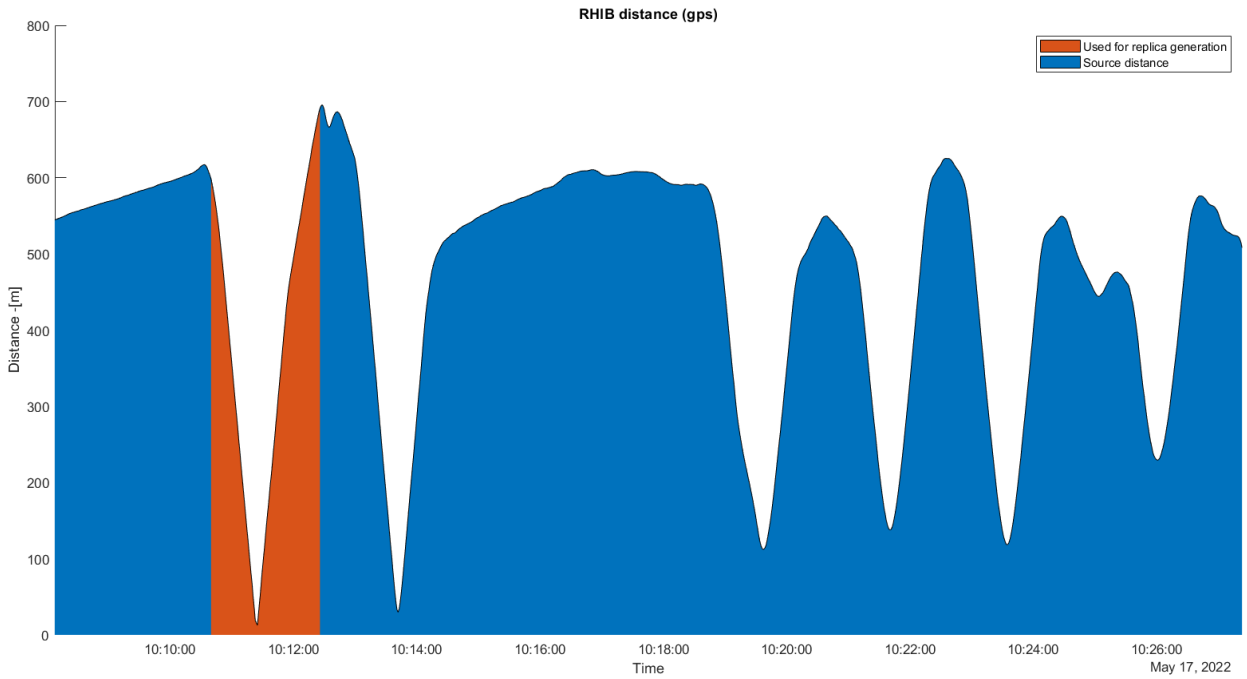


Figure 24: Range to RHIB from receiver based on GPS measurements

The error values of the various tables above are computed over the entire range of measurements shown in figure 24. This includes moments where the source is relatively stationary. It can be assumed that the engine produces less noise in these moments, resulting in a lower SNR. It was chosen to nevertheless include these segments when computing the MAE and SMAPE values to be able to judge the overall effectiveness of ranging. An in depth comparison of error values for different ranges and sound levels follows later. It can be observed that, with the exception of the cross-correlation matching algorithm, the accuracy actually monotonically increases with an increase in down-sample factor. Meanwhile the averaged library, here only shown using the Pearson matching algorithm, shows an error increase while increasing the down-sample factor. The errors shown in table 3 and 4 are comparable to the errors of table 1 and 2, while a significant decrease of computation time was gained by using a negative overlap. These results are discussed in section 5.1.

4.1.2 Usage of sub blocks

In literature a structure was proposed where a data block was split into five blocks of half the original block size with a 75% overlap. This might result in a reduction of noise in the generated cepstrum. To verify if this method provides any reduction in error the results of table 1 and 2, which were generated using this technique, were also produced *without* this technique. These are shown in tables 5 and 6.

Sample rate performance in MAE					
Down-sample factor	Pearson [m]	Minimisation [m]	Averaged library (Pearson) [m]	Cross correlation [m]	Computation time [s]
1x	196	200	32	211	952
2x	187	177	39	205	1317
4x	180	176	52	201	769
8x	169	156	61	205	462
16x	160	150	66	206	293
32x	145	120	68	190	220

Table 5: Sample rate performance in Mean Absolute Error, no sub blocks

Sample rate performance in SMAPE					
Down-sample factor	Pearson [%]	Minimisation [%]	Averaged library (Pearson) [%]	Cross correlation [%]	Computation time [s]
1x	57	56	8	63	952
2x	55	48	11	61	1317
4x	53	49	15	59	769
8x	50	44	16	60	462
16x	47	41	17	61	293
32x	44	33	19	59	220

Table 6: Sample rate performance in Symmetric Mean Absolute Percentage Error, no sub blocks

Because of the halved block size to keep τ_K close to 19.5ms a large increase in computation time can be observed. This halved block size also doubles the times that ranging is being performed on this dataset, i.e. a block now contains approximately 20ms of data instead of 40ms of data. Therefore this method without sub blocks was also compared to the same block size as used when averaging sub blocks. To compensate for the doubling of τ_k the quefrencies above 20ms were simply deleted. It's results are shown in table 7 and 8.

Sample rate performance in MAE					
Down-sample factor	Pearson [m]	Minimisation [m]	Averaged library (Pearson) [m]	Cross correlation [m]	Computation time [s]
1x	175	263	21	172	449
2x	161	187	33	162	656
4x	150	148	43	153	387
8x	138	132	51	148	240
16x	126	127	55	157	158
32x	115	100	59	221	121

Table 7: Sample rate performance in Mean Absolute Error, no sub blocks, standard block size

Sample rate performance in SMAPE					
Down-sample factor	Pearson [%]	Minimisation [%]	Averaged library (Pearson) [%]	Cross correlation [%]	Computation time [s]
1x	52	87	5	51	449
2x	48	58	9	48	656
4x	45	45	12	45	387
8x	41	38	15	45	240
16x	38	37	16	48	158
32x	35	29	16	72	121

Table 8: Sample rate performance in Symmetric Mean Absolute Percentage Error, no sub blocks, standard block size

All tables in section 4.1.2 show a worse performance than their counterparts in section 4.1.1, which use sub-blocks. This effect can be observed across all sample-rates.

4.1.3 Selection of parameters for further results

Virtually all results previously in this section have shown that, for a single-pass library, the error is actually *reduced* by down-sampling, with a minimum error at a down-sample factor of 32. The exception is the cross-correlation algorithm, which performs better at smaller down-sample factors. Meanwhile the averaged library seems to perform better with a higher sample rate. Therefore for the coming sections a down-sample factor of 32 will be used in conjunction with the single-pass library, while a higher sample rate will be used for the averaged library.

Section 4.1.2 shows that the usage of sub-blocks reduces the error significantly for both block-sizes. Therefore this technique will be used for the remainder of the results. When it comes to using a negative overlap in cepstrum generation, e.g. 'skipping' blocks, no significant increase in error can be seen while computation time is significantly reduced. The one in five ratio used in tables 3 and 4 will be used for the remainder of this report unless a different ratio is specified.

4.2 Library and matching algorithm effectiveness

Two different types of experimental libraries were discussed in section 3.5. This includes the single-pass library where only a small segment of measurements is used to create a library as shown in figure 24, and the averaged library where much more data is included and averaged for certain range increments. The results of the tables above mostly include the single-pass library, while the averaged library is represented paired with the Pearson algorithm. In this section these results will be expanded, and the averaged library will be tested in conjunction with the other algorithms as well.

4.2.1 Single-pass library

The single-pass library only uses the very first run of the RHIB to generate the library, as indicated in figure 24. This range is also indicated in the different sub-figures of figure 26. These sub-figures show the range as measured by GPS in the top position and the estimated range by using the single-pass library and the corresponding algorithm is shown in the middle position. The bottom graph shows the error between the GPS measurements and the estimated distance. The error spread is mostly centered around zero for the mean algorithm, while the Pearson algorithm shows a slight bias to overestimating the range. The cross correlation algorithm however shows a strong bias to underestimating the range.

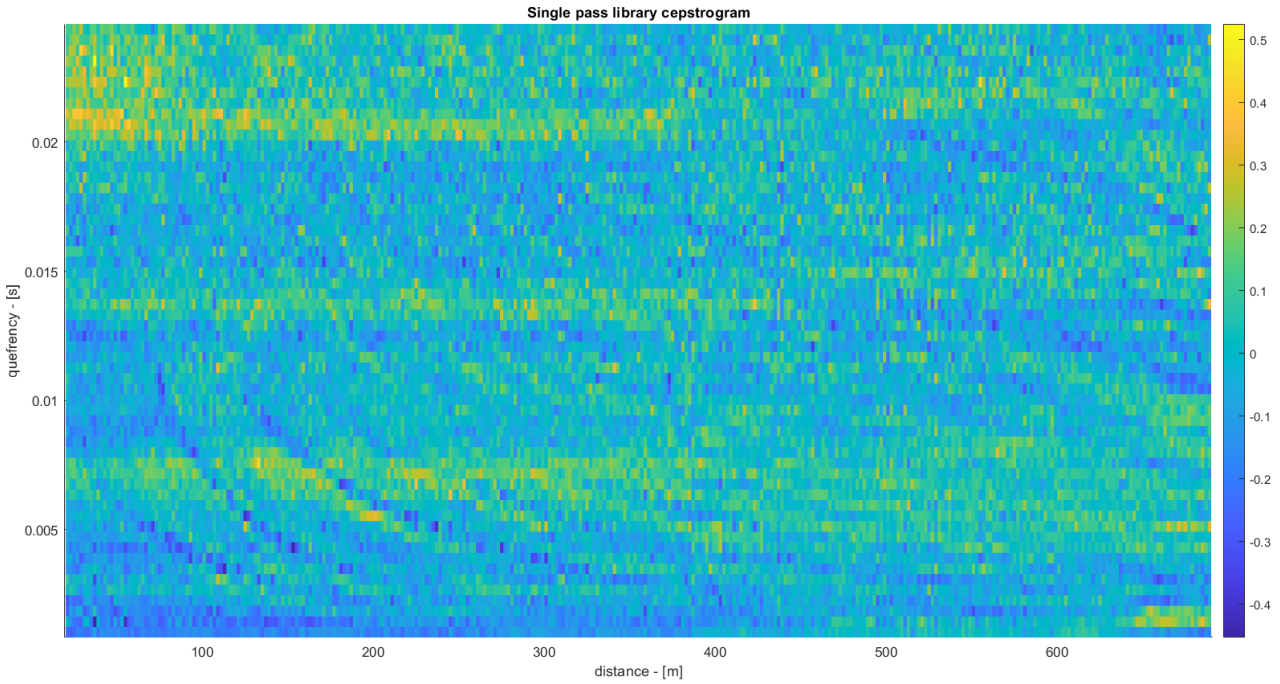
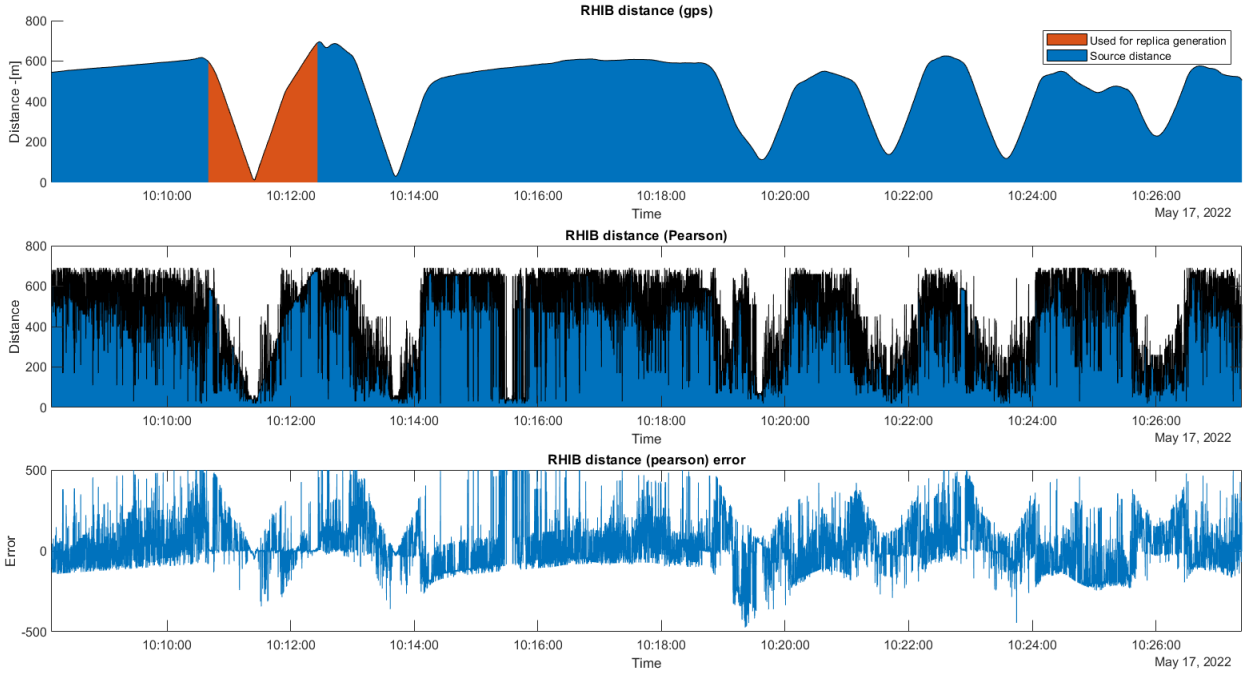
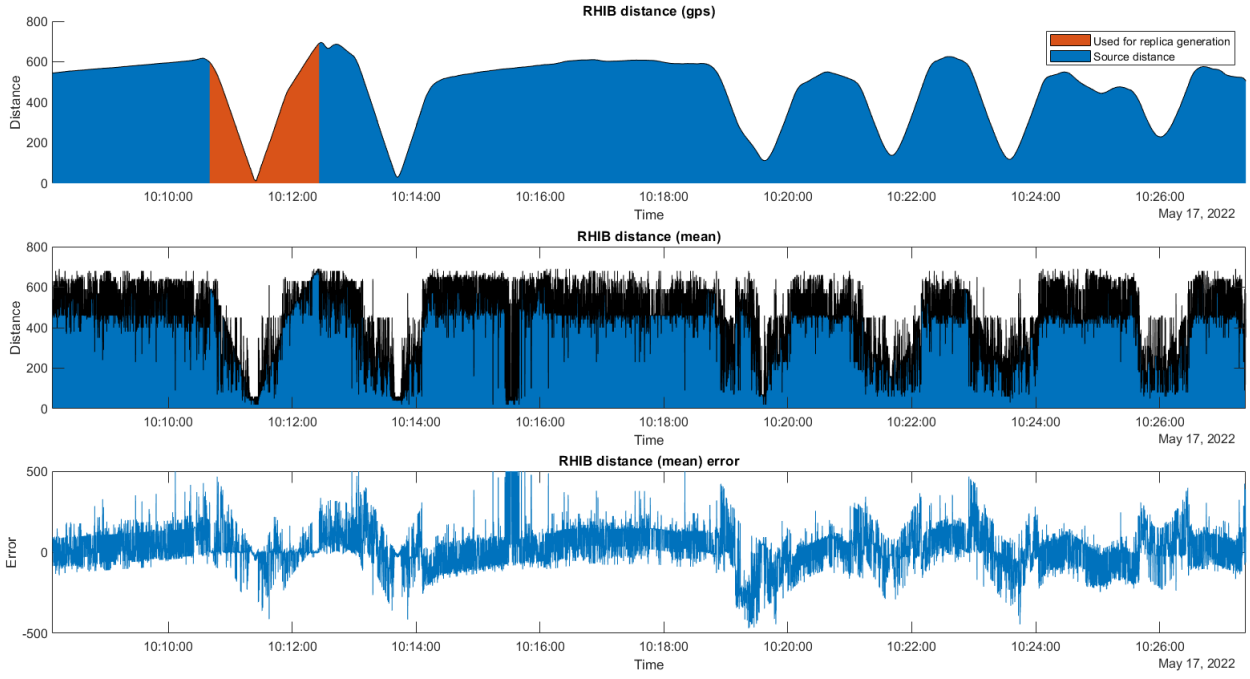


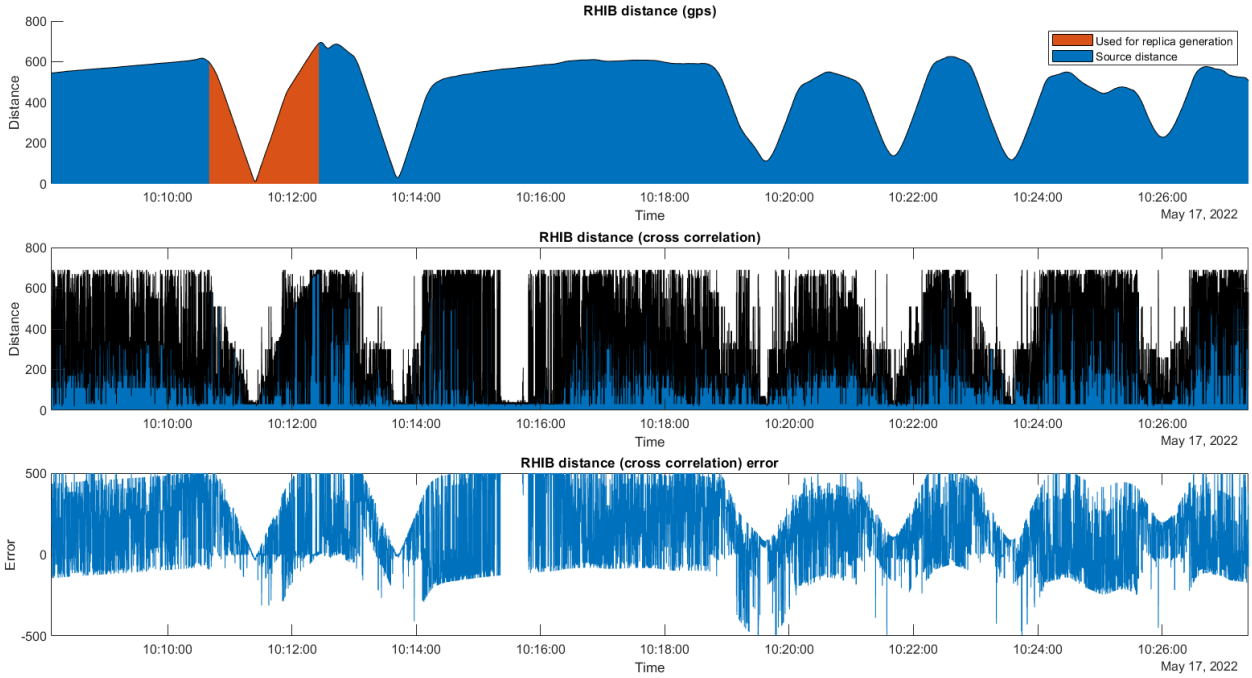
Figure 25: Single-pass library - 32x down-sampling



(a) Pearson ranging - single-pass



(b) Minimisation ranging - single-pass



(c) Cross correlation ranging - single-pass

Figure 26: Single pass library results

4.2.2 Averaged library

In section 4.1.1 the averaged library combined with the Pearson matching algorithm impressed with lower MAE and SMAPE magnitudes compared to averaged library algorithms. However, since the library shown in figure 27 contains all cepstra on which ranging is performed, even when they are averaged, these results might be subjective. Therefore the algorithms were tested on an averaged library that excluded the data between 10:13:00 and 10:18:45. This way the second run at around 10:14:00 can be used to verify the effectiveness, of which the results are shown in figure 28. The ranging performance of a complete averaged library is also shown for comparison in figure 28d. These results were generated without down-sampling since section 4.1.1 showed that the averaged library performs better without down-sampling.

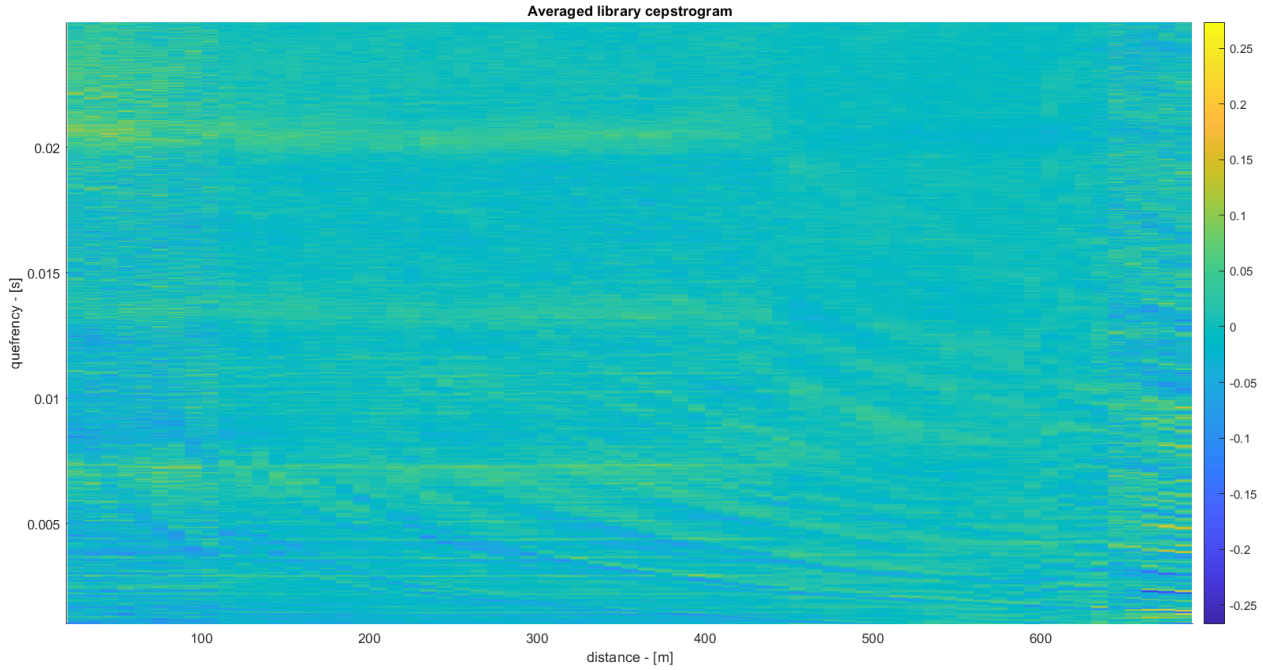
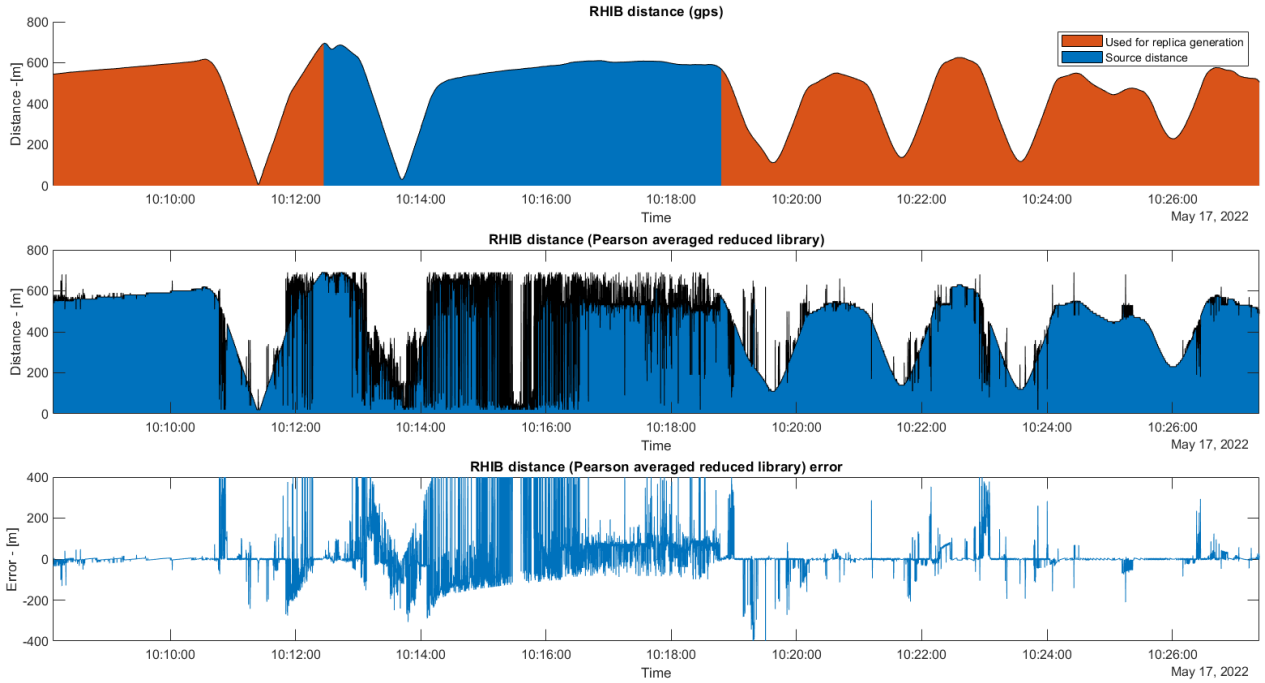
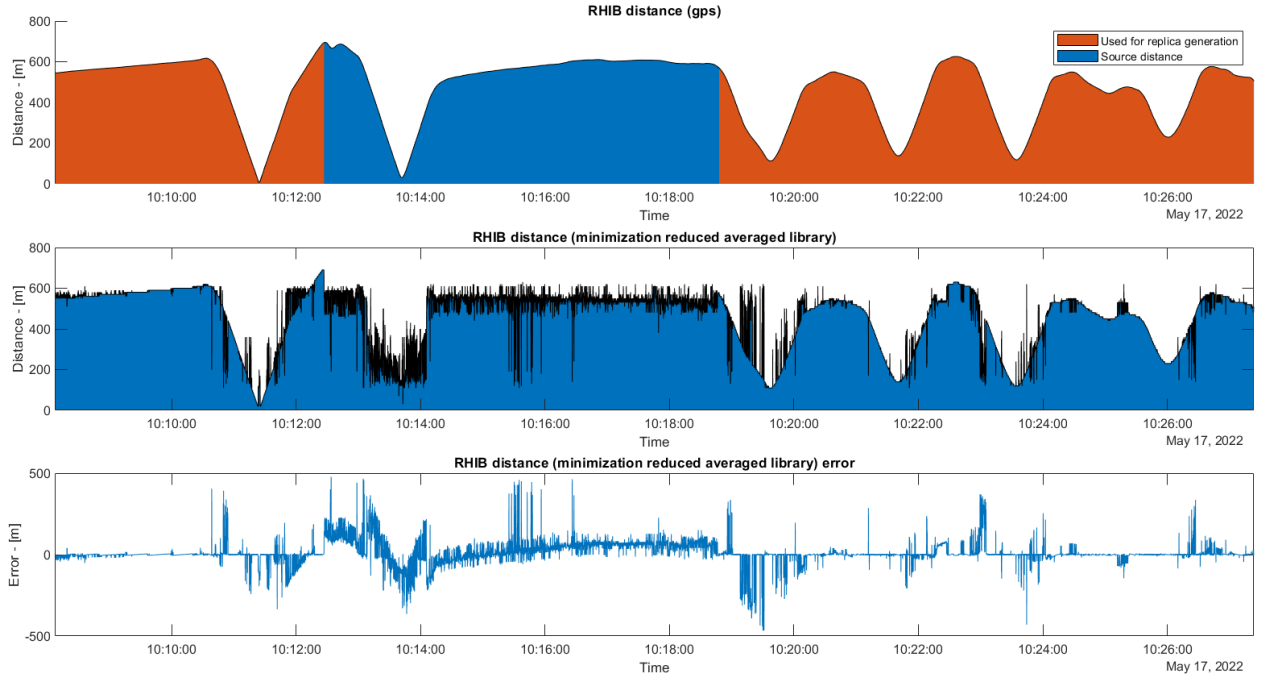


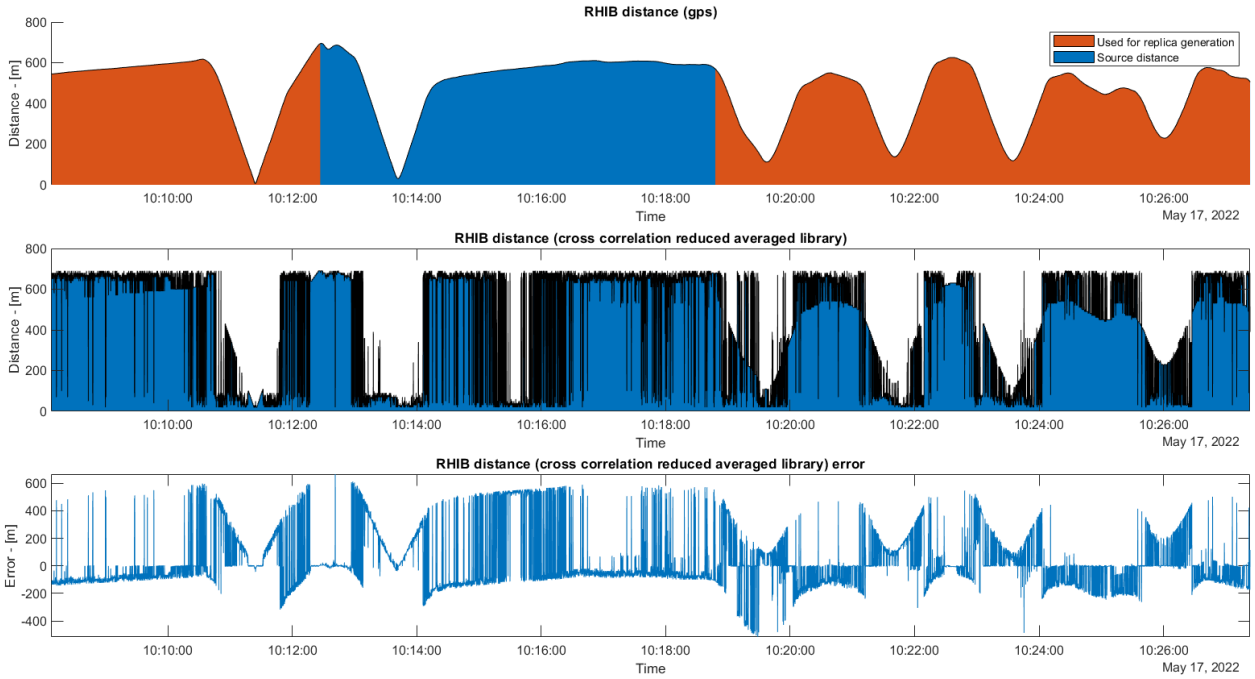
Figure 27: Averaged library - no down-sampling



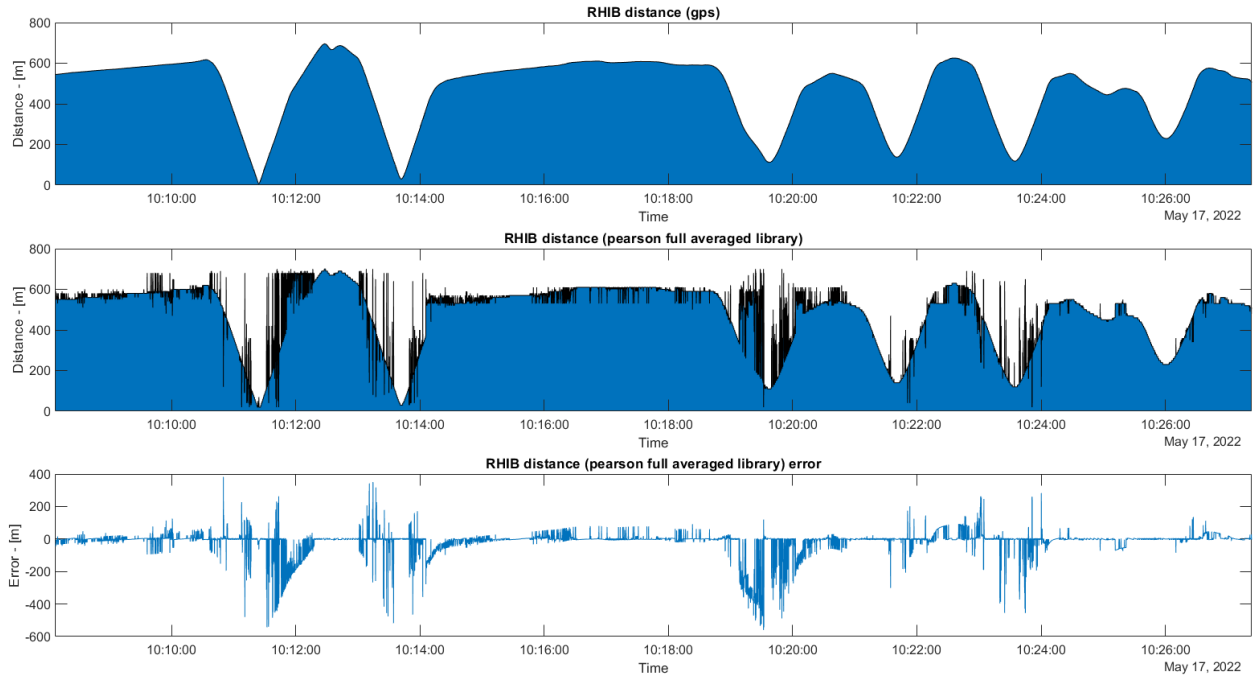
(a) Pearson ranging - averaged library



(b) Minimisation ranging - averaged library



(c) cross correlation ranging - averaged library



(d) Pearson ranging - full averaged library

Figure 28: Averaged library results at high tide

4.2.3 Tidal influence

The previous results have all been generated using experimental libraries that were recorded at virtually the same time as the measurements themselves. To see how these libraries and matching algorithms perform at a change in water height by tidal effects a measurement from ca. 6 hours later was used.

Here the same RHIB passes the sensor, although at a significantly lower speed at first. Some acceleration occurs after the closest point of passage, although only half the speed of the runs at low tide is reached. The tides created a difference with respect to NAP of approximately -50cm at the time of library creation, which is close to low tide. During the measurements used in this section however the tide created a difference of 105cm with respect to NAP, creating a change in water height of 155 cm compared to the moment of library creation.

4.2.3.1 Single-pass library

The single pass library is not able to perform ranging with any of the algorithms with the new measurements. No trend following the GPS measured range of the RHIB can be observed in figure 30 and the resulting range of all algorithms seems to be mostly noise.

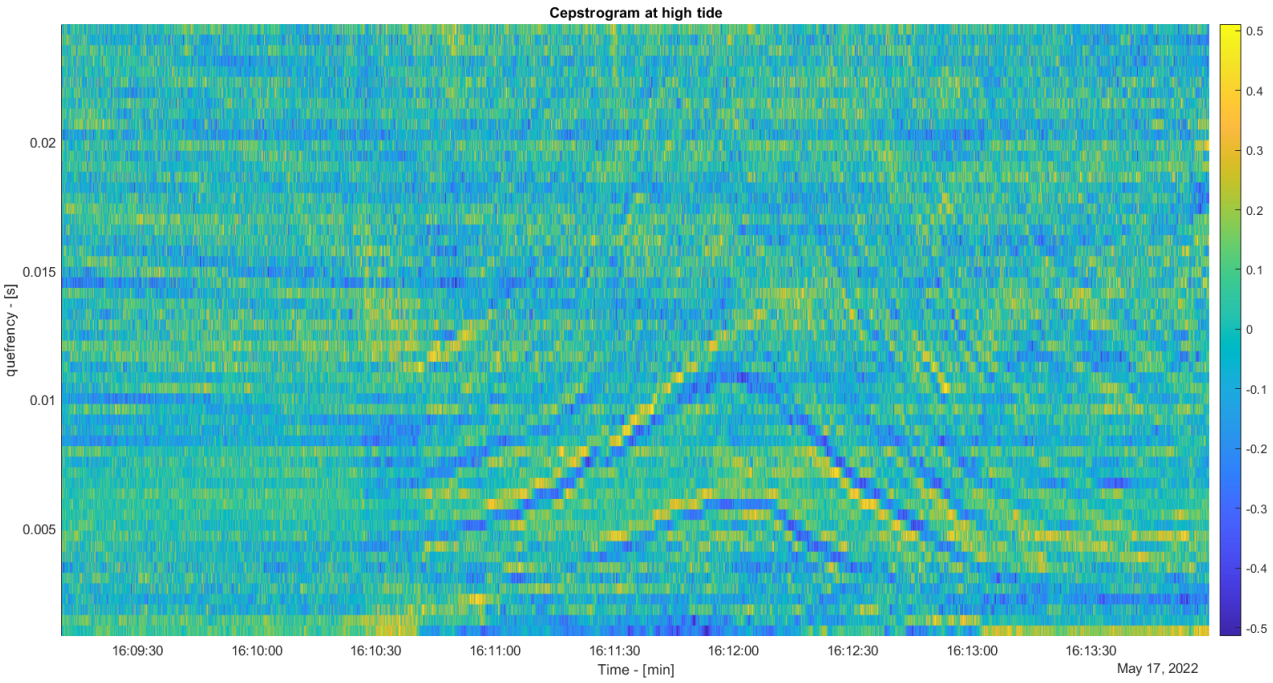
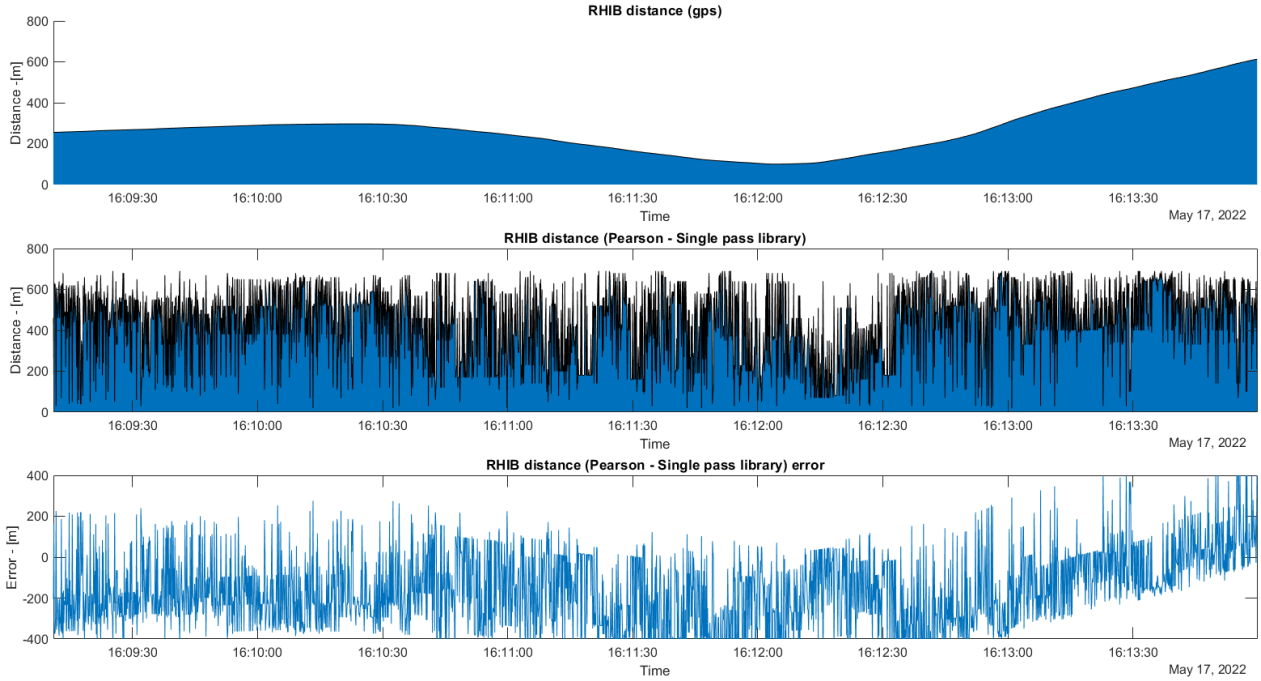
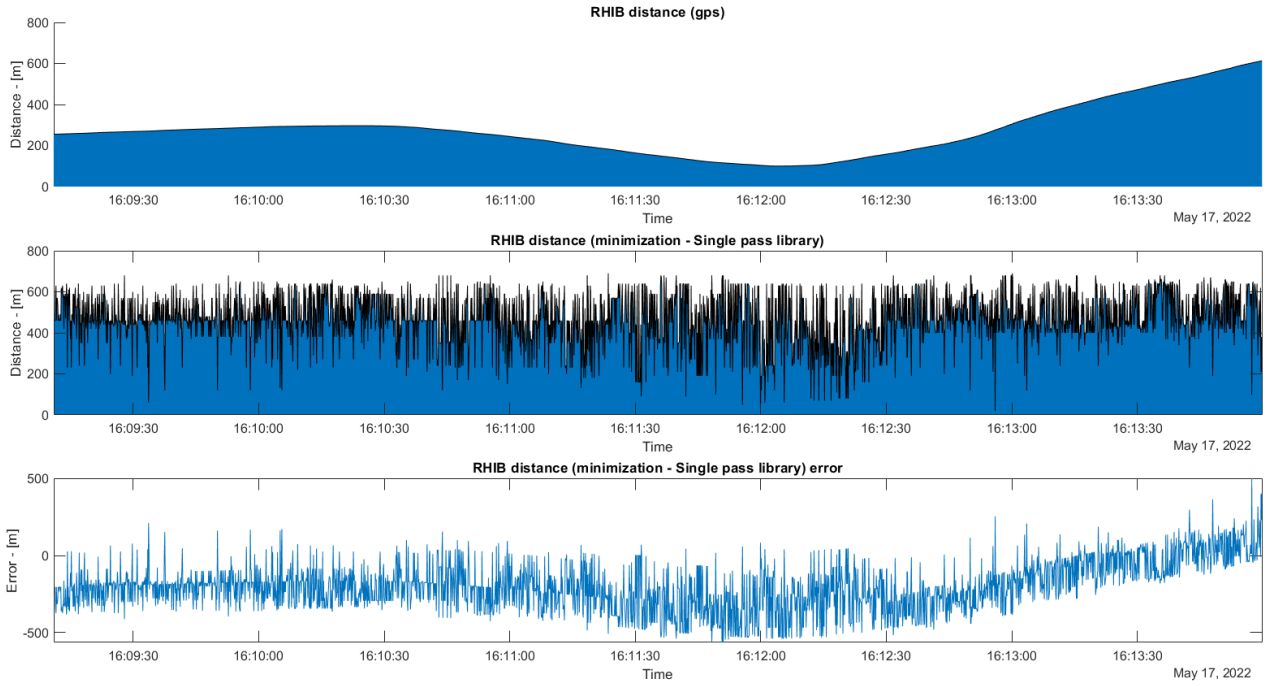


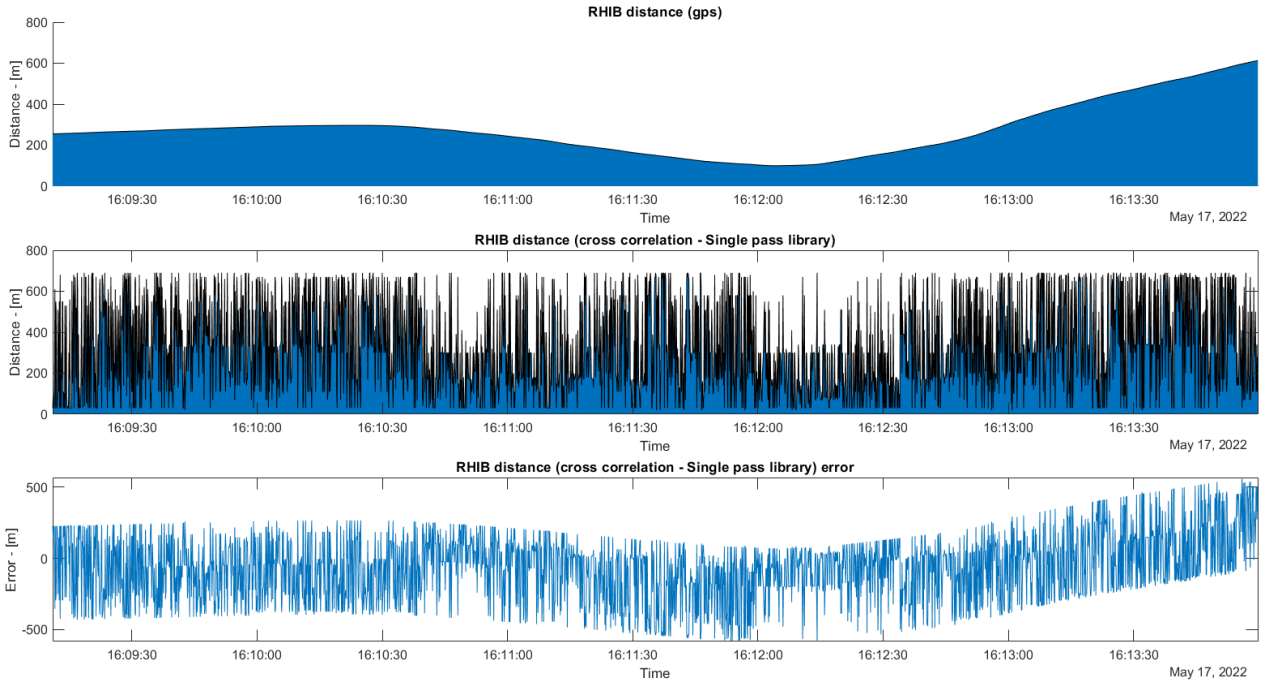
Figure 29: Cepstrogram of RHIB at high tide - 32x down-sampling



(a) Pearson ranging at high tide- single-pass



(b) Minimisation ranging - single-pass



(c) Cross correlation ranging - single-pass

Figure 30: Single-pass library results at high tide

4.2.3.2 Averaged library

The averaged library is not able to perform ranging on the new measurements with any algorithm, just like the single-pass library. The Pearson algorithm shows some tendency towards lower range values at the closest passing point in figure 32a, but both the minimization algorithm in figure 32b and the cross correlation algorithm in figure 32c result in either maximum or minimum ranging values.

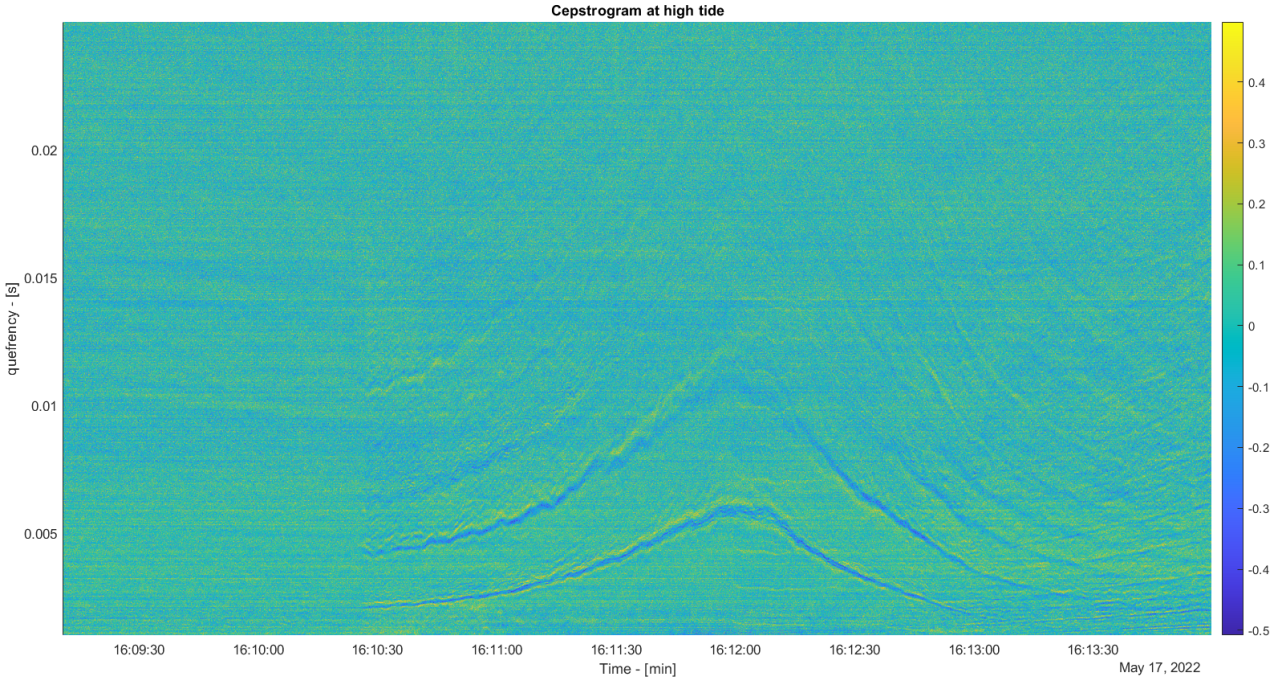
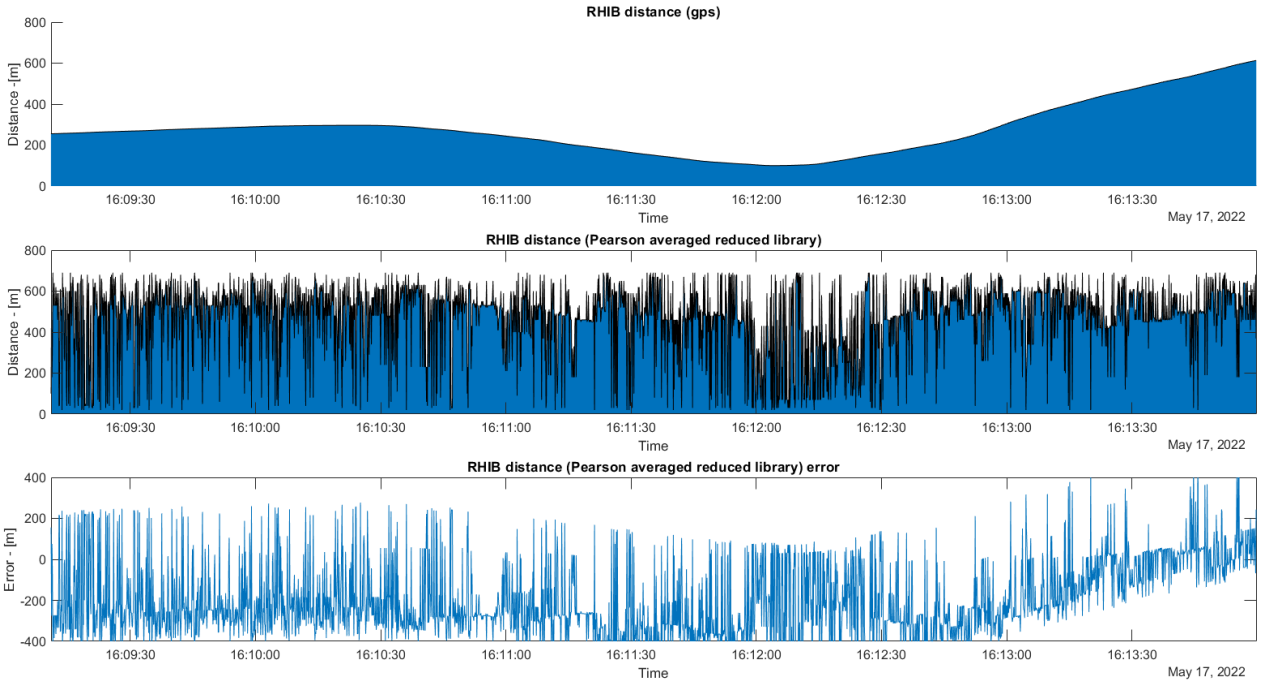
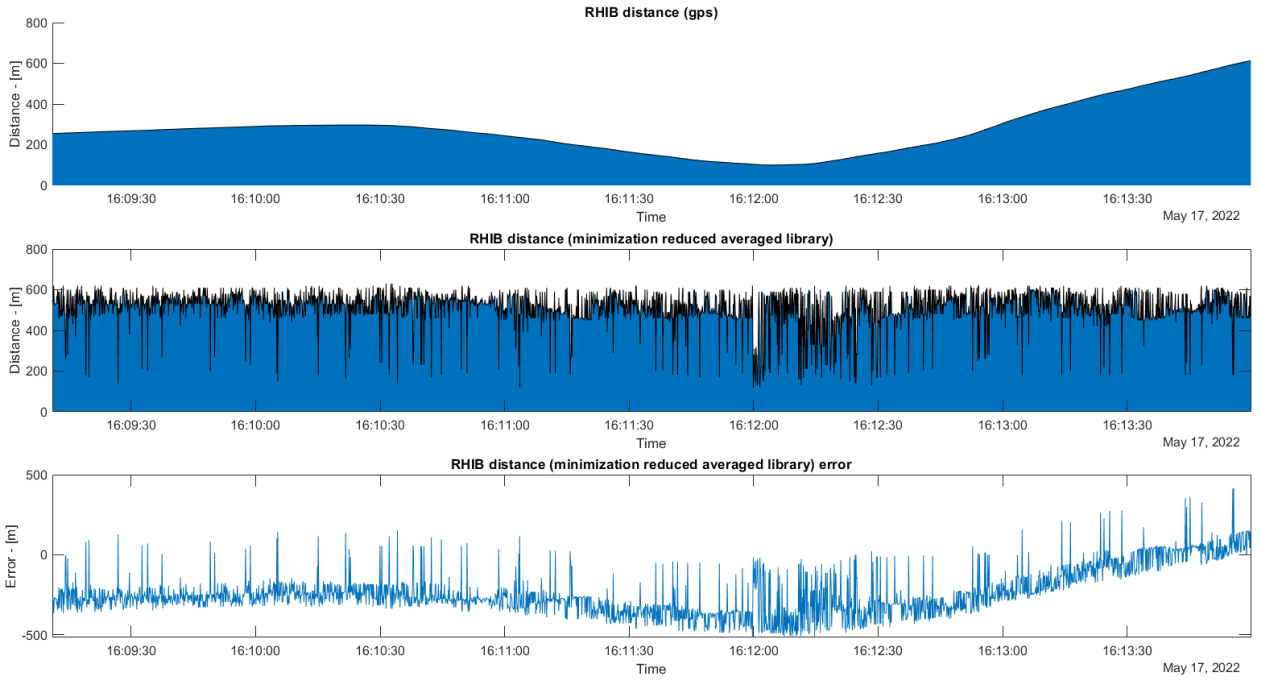


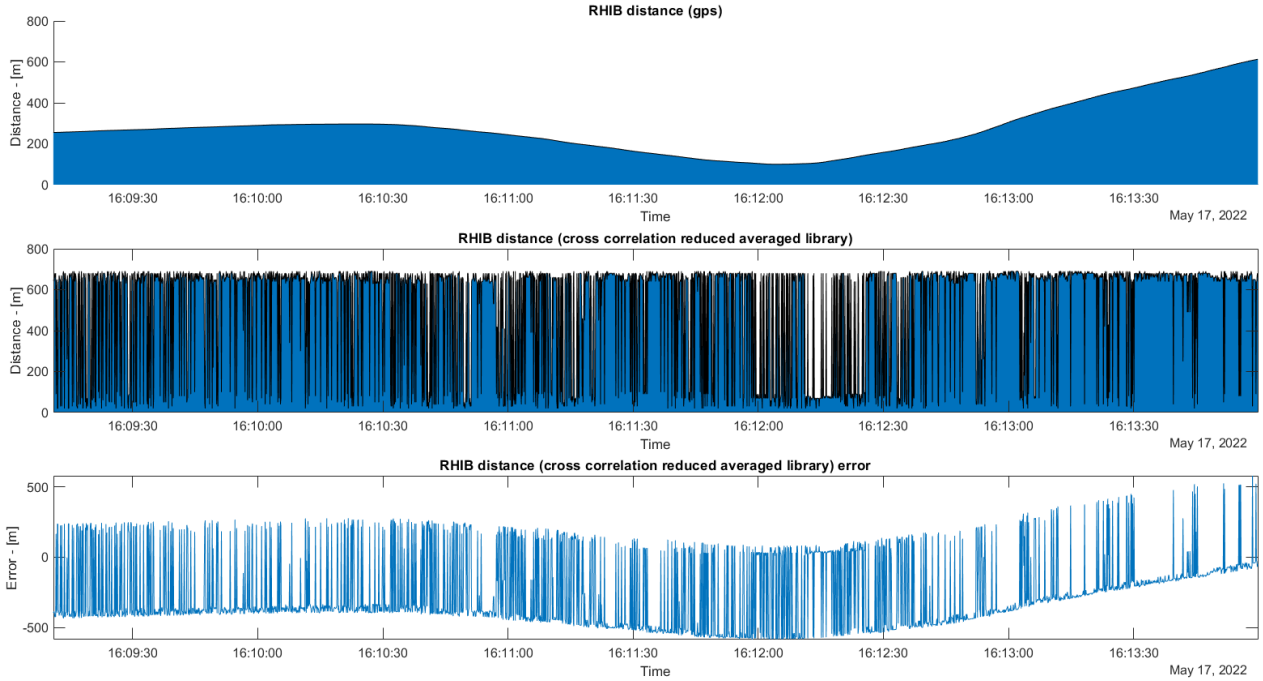
Figure 31: Cepstrogram of RHIB at high tide



(a) Pearson ranging at high tide- averaged library



(b) Minimisation ranging at high tide - averaged library



(c) cross correlation ranging at high tide - averaged library

Figure 32: Averaged library results at high tide

4.2.3.3 Identifying problems with tidal ranging

To come to an explanation for the abysmal ranging performance with a change in water height a comparison was made between two cepstra. The first cepstrum was measured at high tide at 16:12:35 with a distance of 160m, which is visible in figure 29, while the second cepstrum was extracted from the single-pass library for a distance of 175m generated at low tide as visible in figure 25. These cepstra were both generated from data that had been downsampled 8 times as a tradeoff between noise and resolution. The result is visible in figure 33.

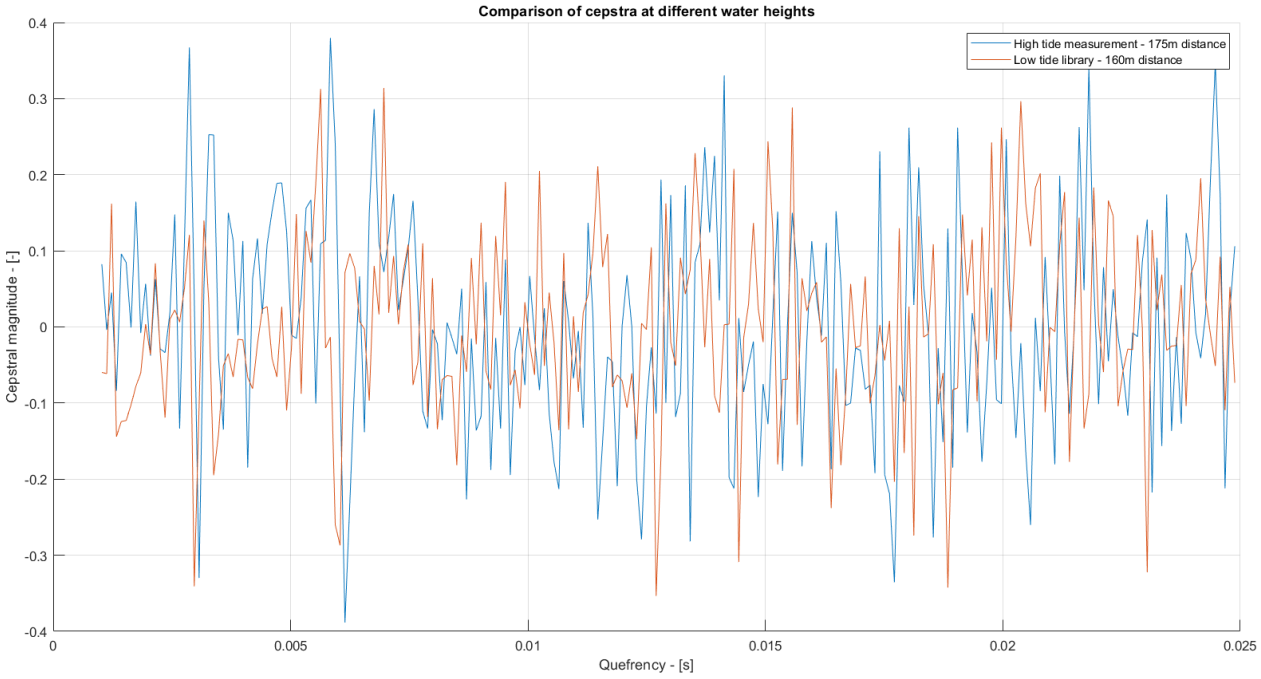
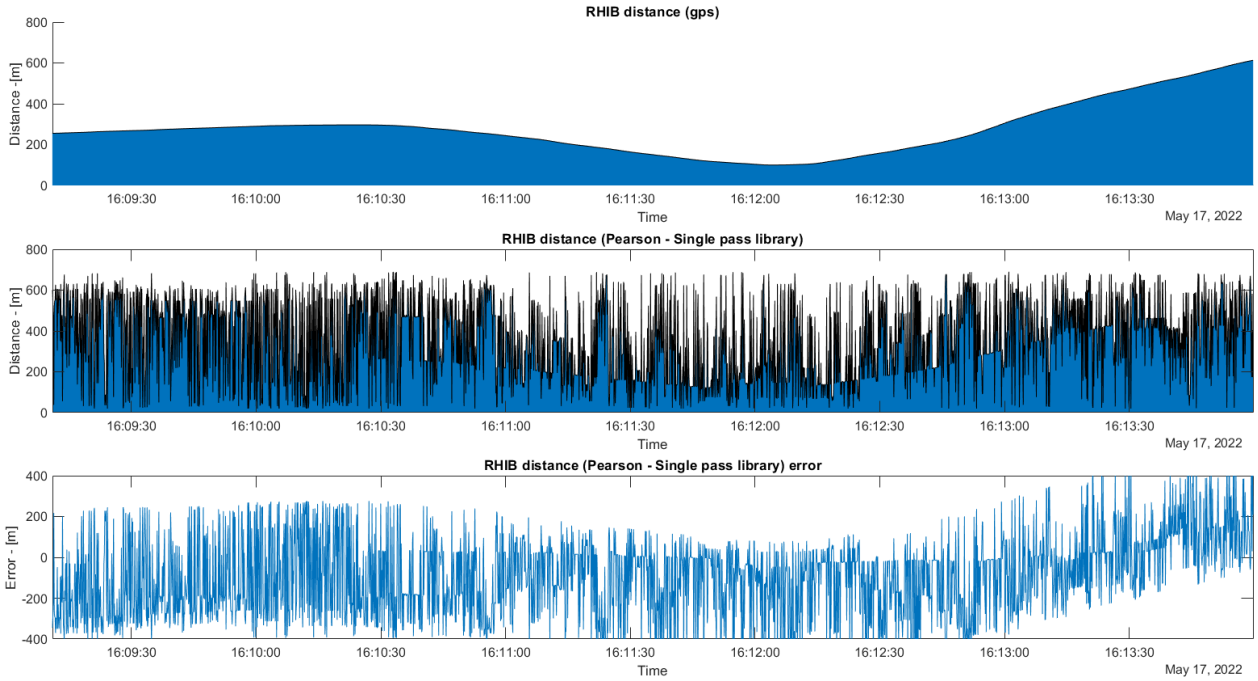


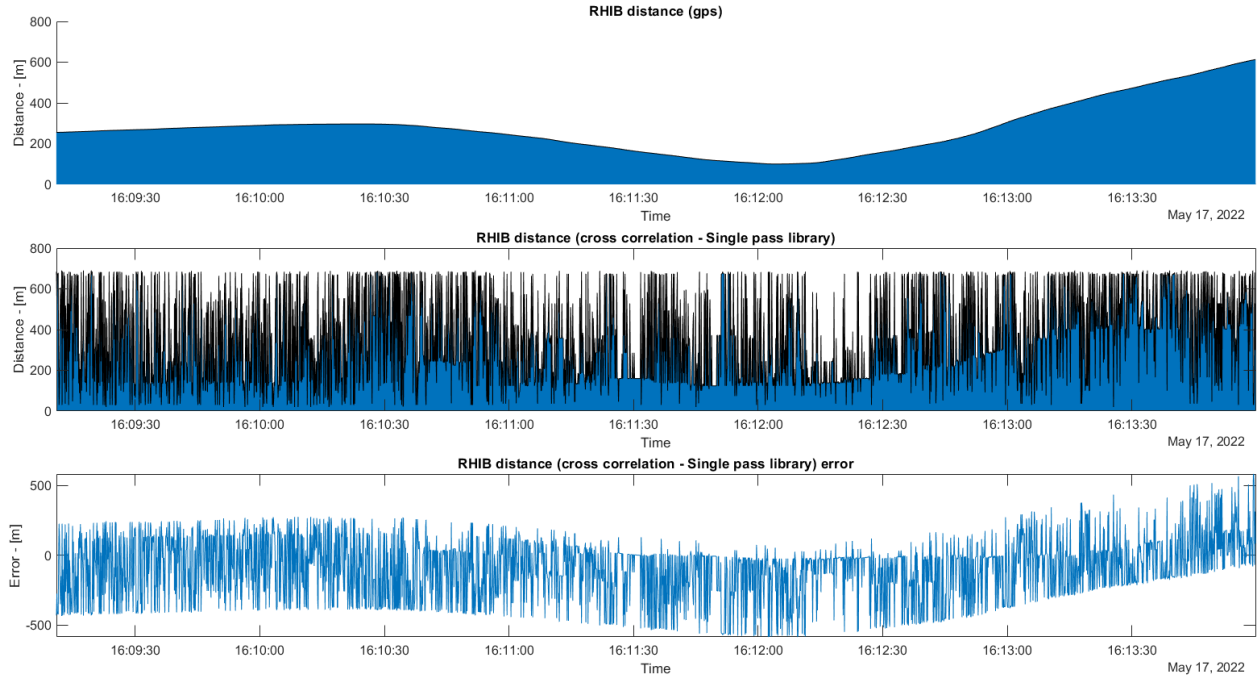
Figure 33: Comparison of cepstra at high- and low tide

Only the first two negative peaks at a quefrency of 3ms and 6ms respectively are of significance. These are created by the arrival from the bottom-surface reflection path and it's first rahmonic. While these peaks seem to closely align between the two cepstra, no correlation can be seen at higher quefrencies.

These results point to an issue with ranging created in the higher quefrency range. Therefore ranging was attempted again, but with the maximum quefrency limited to 12ms, which corresponds approximately to the quefrency of the first rahmonic at the closest point of passage during high tide. Only the result of the Pearson algorithm and cross correlation algorithm are displayed in figure 34 as the minimisation matching algorithm only generated noise. These results were generated with data downsampled with a factor of 8 as this gave the most accurate ranging performance in this specific case. Decreasing the down-sample factor further did not result in an increase in ranging performance.



(a) Pearson ranging at high tide - maximum quefrency of 12ms



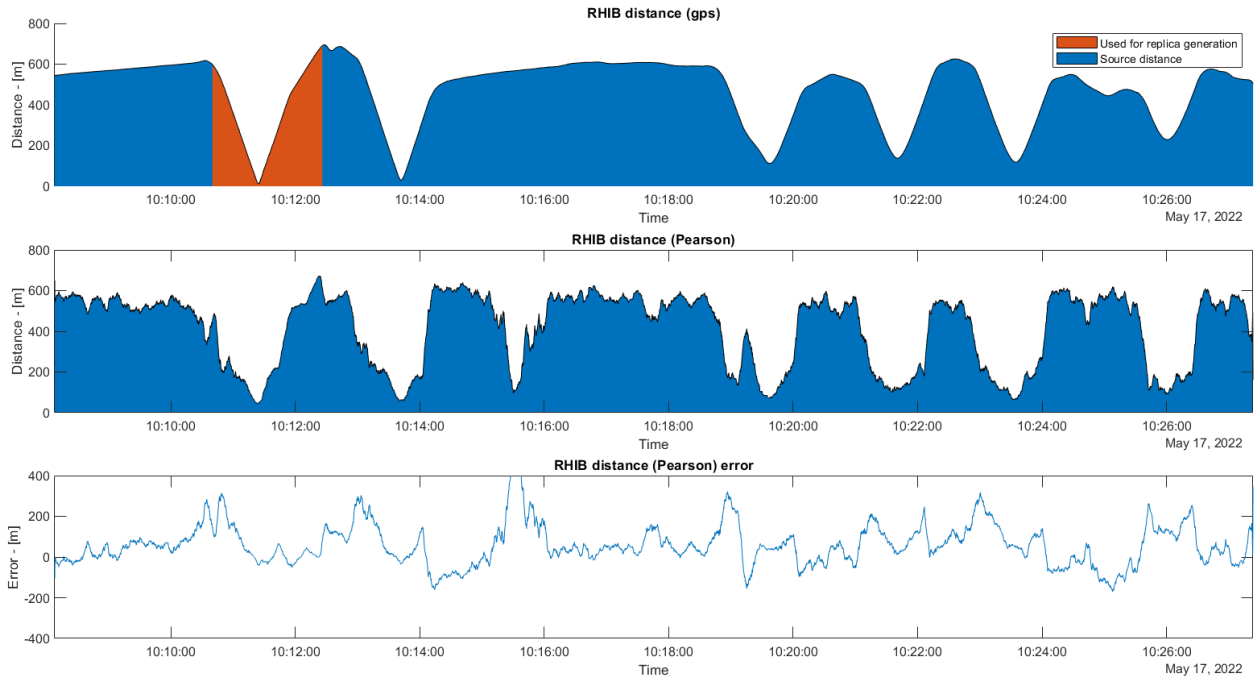
(b) Cross correlation ranging at high tide - maximum quefrency of 12ms

Figure 34: Single-pass library results at high tide - maximum quefrency of 12ms

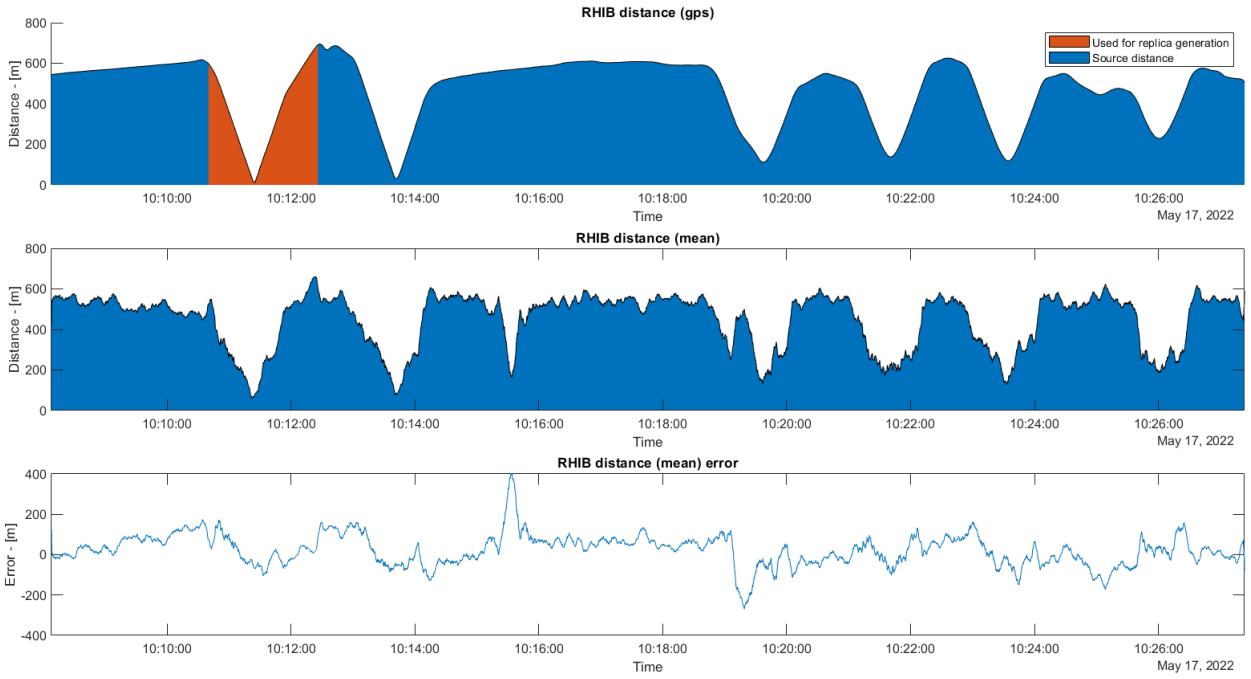
Interestingly decreasing the quefrency improved the ranging performance considerably with a tide mismatch. Although a significant amount of noise is still present in the range estimations, a trend following the actual range where the error magnitudes are close to zero can be seen in this data, especially after the closest point of passage. These results are further discussed in section 5.3.

4.2.4 Smoothing of ranging results

As evident in figures 26 and 28 there is significant noise contained in the range estimations, regardless of matching algorithm. The averaged library seems to be less affected. This noise seems to fluctuate around an error value of zero on most ranges. Therefore a simple moving average window filter was added to the ranging results of the single-pass library in conjunction with the Pearson and minimized sum matching algorithms. The window span was set on 49, which contains approximately 5 seconds of data. While this may seem as a large window for a source at high speed it was found that this size gives the best results for this specific series of measurements.



(a) Pearson ranging - single-pass - moving average filter



(b) Minimisation ranging - single-pass - moving average filter

Figure 35: Single-pass library results with moving average filter

The readability of ranging results is increased significantly by adding a moving average filter. Some 'rounding' of ranging results can be observed where there is a sharp change in range, i.e. where the RHIB suddenly accelerates or decelerates. As a result the MAE was reduced to 85m and 62m for the Pearson and minimisation matching algorithms respectively, and the SMAPE to 25% and 16% respectively. A sudden decrease in range estimate can be observed around 10:16, especially with the Pearson algorithm in figure 35a, while the RHIB stays at an approximately constant distance. Analysis of the AIS data showed that this is the passage of another vessel. This vessel was identified by its MMSI as the MARKENJE, a small dredging vessel. The TDOA lines created by this vessel can be seen in figure 36 around the 10:16 mark. The implications of this result will be discussed in section 5.4.

4.3 Raytracing accuracy for TDOA estimation

While a raytracing algorithm has been developed for this thesis that can incorporate open-source depth measurements as well as data regarding salinity, temperature and tidal effects, no results were gained in converting this to an automatic ranging algorithm. Therefore the results of this algorithm will only be presented visually, plotted over generated cepstrograms. The data has not been down-sampled to sharpen the cepstrograms. As the TDOA in a cepstrum will invariably result in rahmonics the TDOA computed by raytracing has been plotted including two rahmonics, visible as black lines. These rahmonics are simply computed by multiplying the computed TDOA with a factor of two and three respectively.

As visible in figure 36 and more notably figure 38a, some passes of the RHIB seem to have a time-shift between the GPS and audio data. This has been corrected by shifting the GPS time by two seconds in figure 38b. The results of the ray tracing algorithm are discussed in section 5.5.

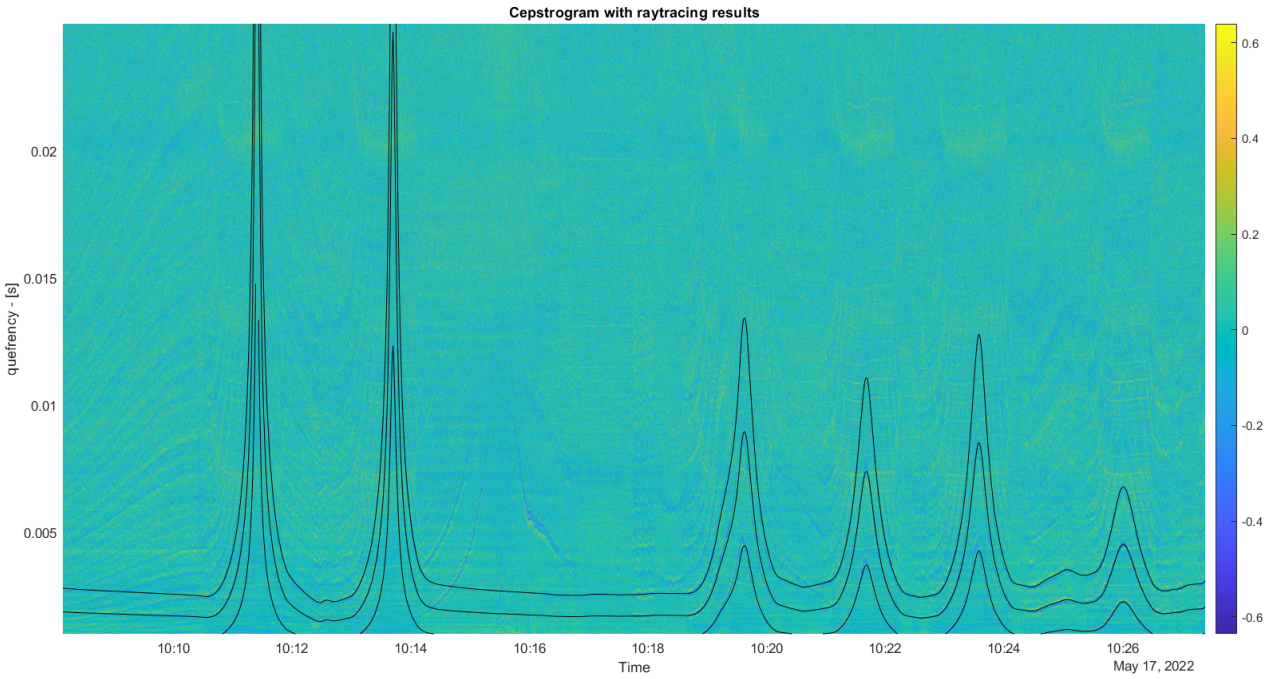
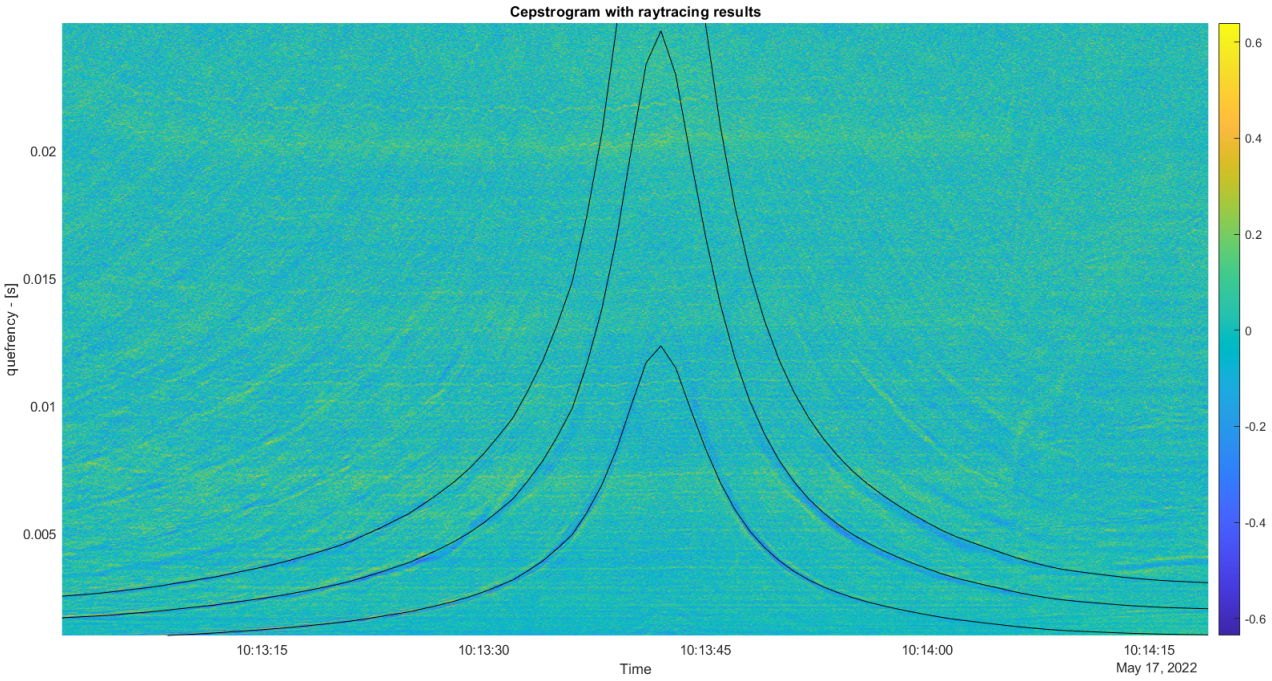
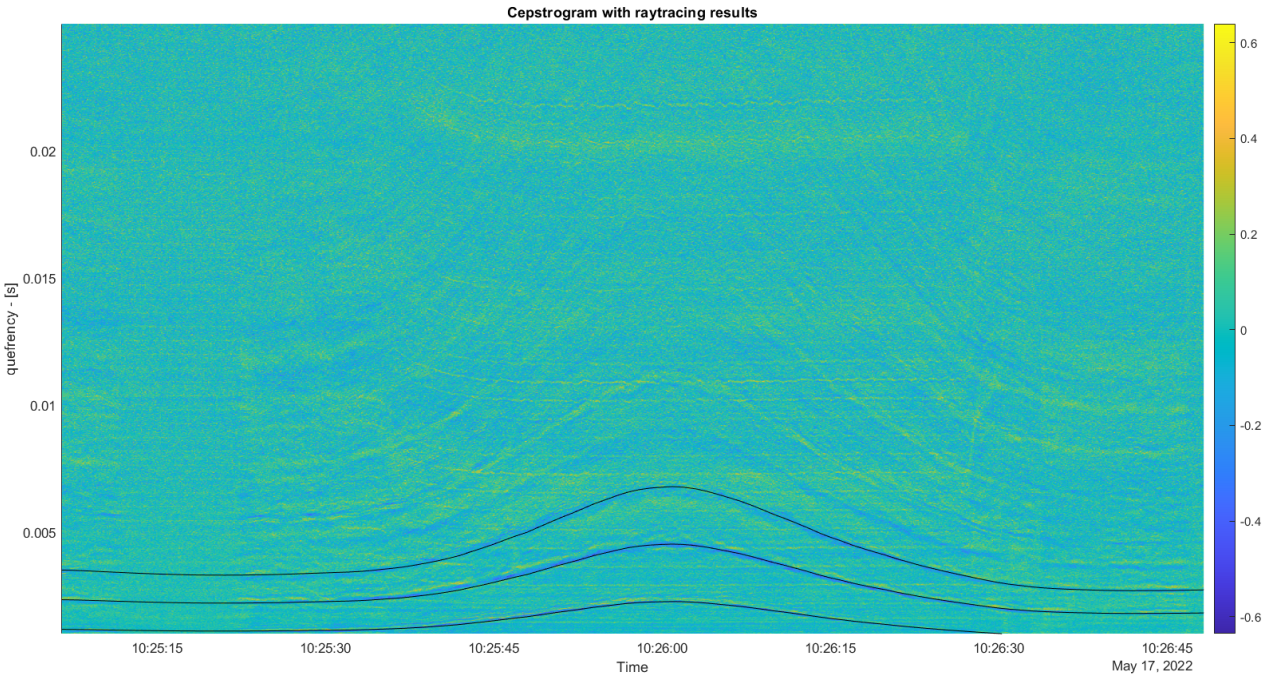


Figure 36: Cepstrogram including raytracing result



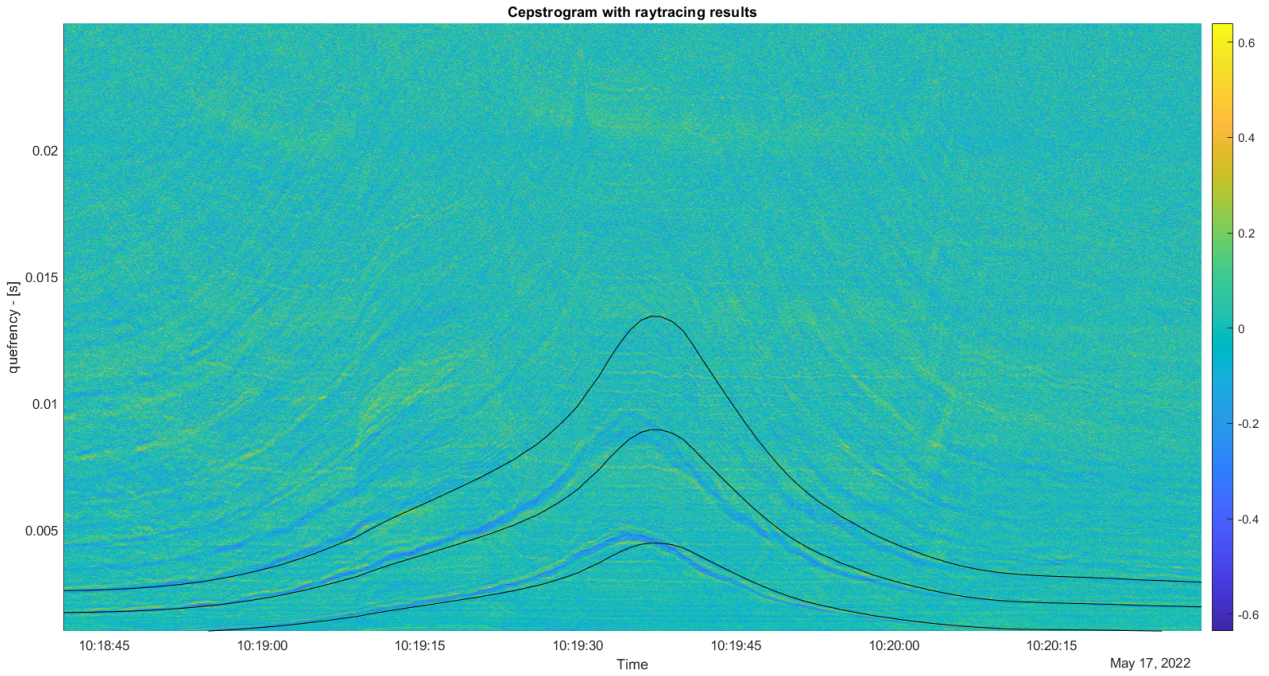
(a) Second pass of RHIB - raytracing result



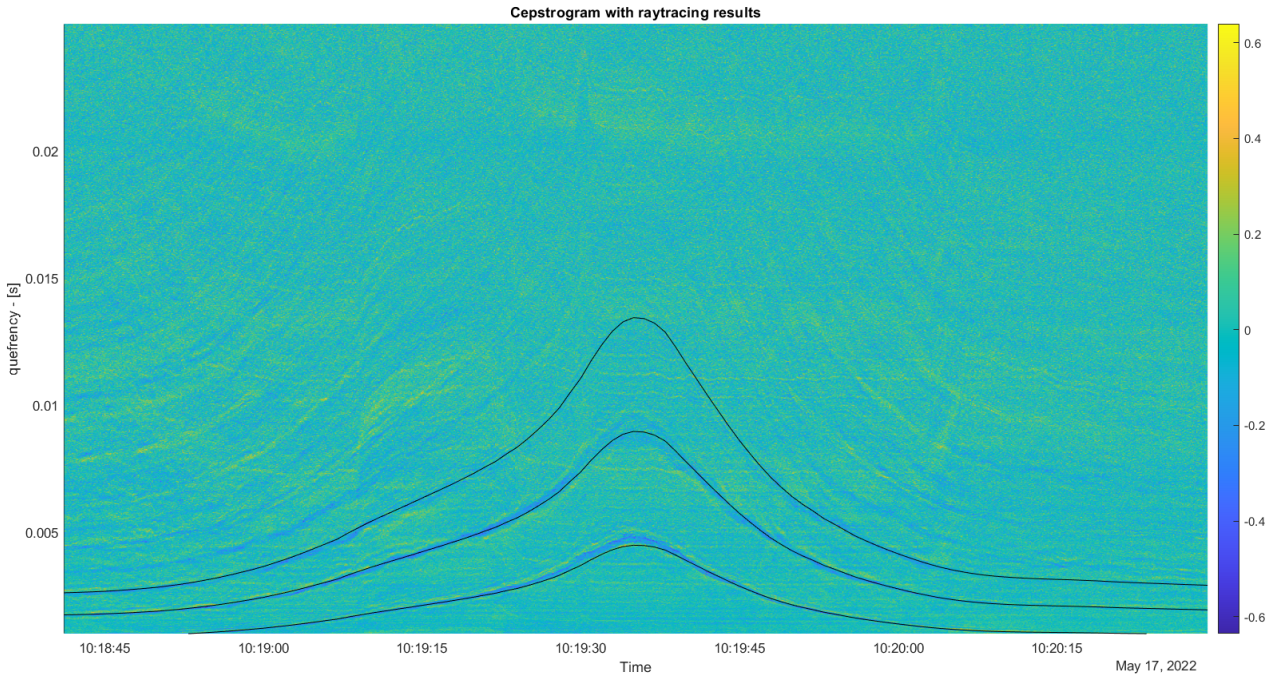
(b) Sixth pass of RHIB - raytracing result

Figure 37: Two examples of correct raytracing results

Figure 37 shows an exact match between the TDOA line and its rahmonics in the cepstrograms and the computed TDOA values.



(a) Third pass of RHIB - ray tracing result



(b) Sixth pass of RHIB - ray tracing result shifted 2 seconds

Figure 38: Example of time mismatch and correction

By shifting the ray tracing results 2 seconds, a very close match between the cesptrogram and computed TDOA is attained. There seems to be a slight mismatch at the closest point of passage, although barely significant.

4.3.1 Raytracing effectiveness of tidal compensation

The TDOA has also been computed for the measurements from section 4.2.3 to visualise if and how well the raytracing algorithm is able to compensate for tidal effects the TDOA has also been computed for the measurements from section 4.2.3. The result can be seen in figure 39. Again a time-shift is visible in this plot. This is corrected in figure 40 by shifting the GPS time by four seconds. This figure compares the (correct) result for high tide in black with the result for the same track at the tidal height in red of the earlier measurements as in figure 36.

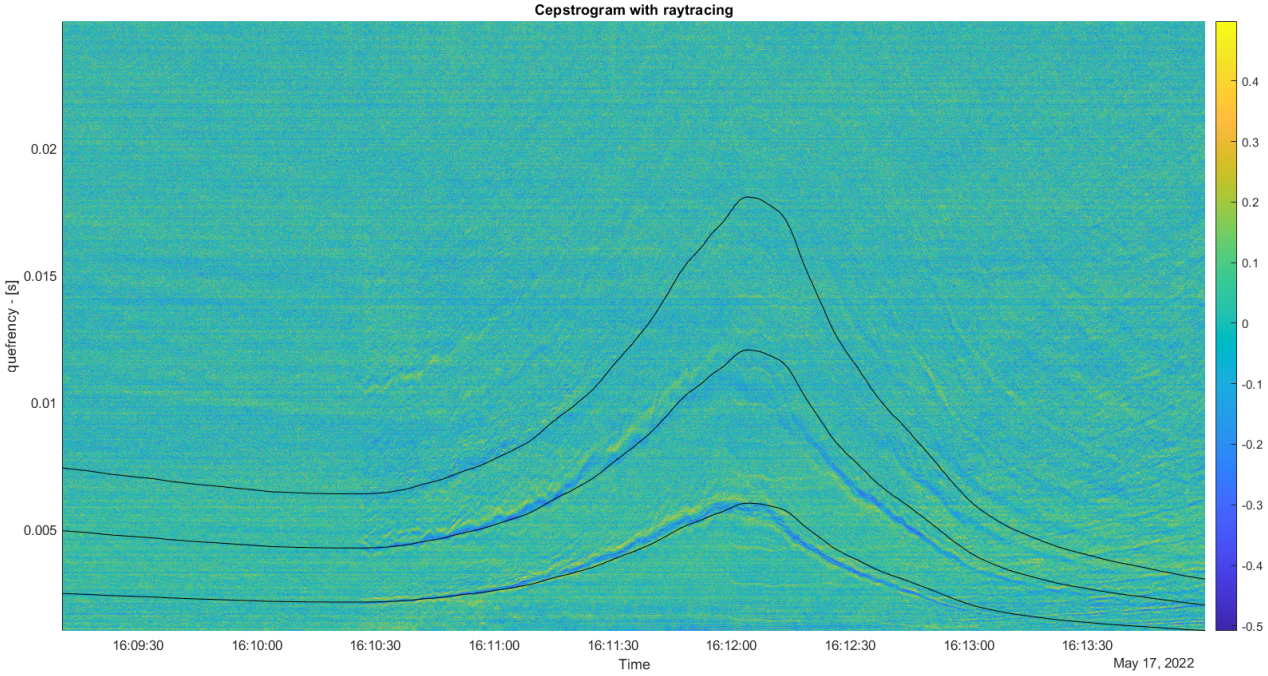


Figure 39: Cepstrogram including raytracing result - high tide

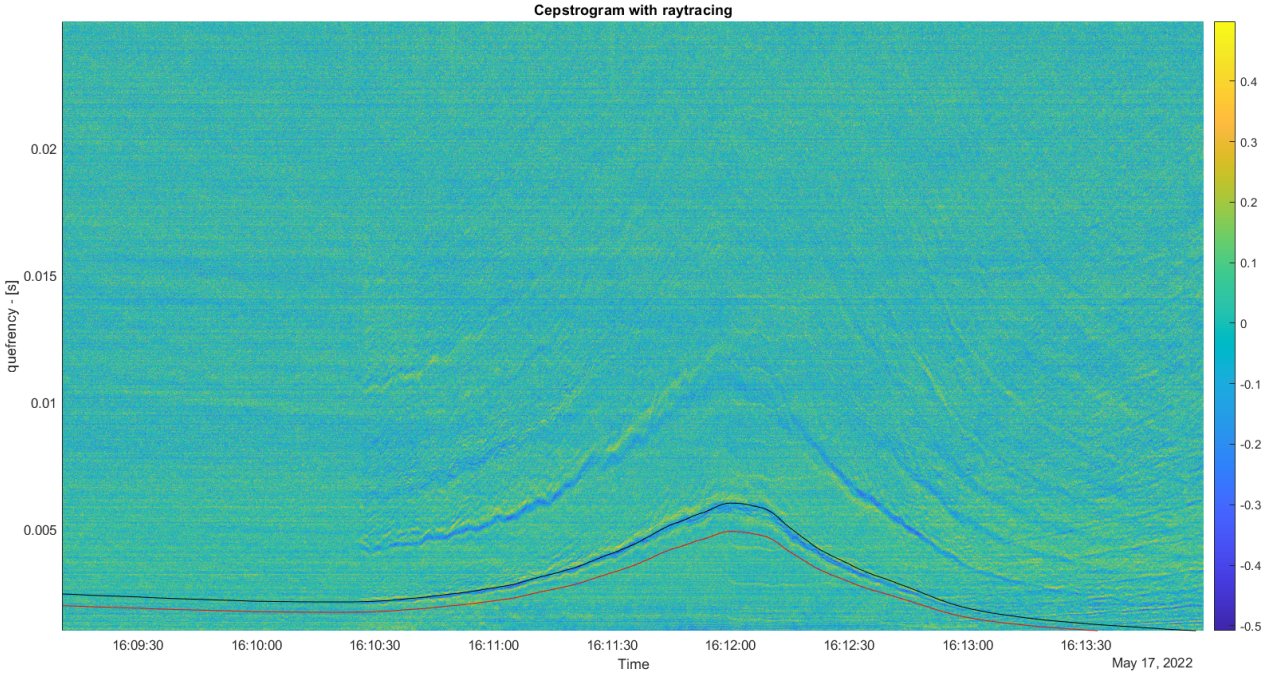


Figure 40: Cepstrogram including raytracing result - high tide

5 Discussion

In this section the results as presented in section 4 are analyzed and where possible explained. The previous section showed the effects of down-sampling, usage of sub-blocks and skipping data, expressed in error compared to GPS data and computation time.

Following these results and taking them into account, the effectiveness of the single-pass library and averaged library were explored in conjunction with the Pearson, minimization and cross correlation algorithms. These methods were tested on measurements with varying tidal height and visualised with error graphs.

Finally the accuracy of the ray tracing algorithm was visualised for varying tidal heights.

5.1 Cepstrogram parameters and sample rate

Section 4.1 provided insight on the effect that parameters have when it comes to cepstrogram generation and the resulting accuracy.

Interestingly for the single-pass library errors are reduced significantly by down-sampling, with a monotone decrease of error compared to an increase in down-sampling factor. This can be explained by the fact that the amount of quefrencies K is reduced and their spacing Δ_τ is increased by this down-sampling operation. Fewer discrete quefrencies result in less noise relative to the peaks of interest in the cepstra. Next to that the peaks of interest, which are usually a negative peak flanked by two positive peaks, are represented by only three discrete quefrencies when down-sampling with a factor of 32. This reduces the risk of mismatch caused by a TDOA peak that has a slightly different shape in the higher resolution representation created without down-sampling.

For the averaged library however there is a monotone *increase* of error visible when the amount of down-sampling is increased. It is assumed that this is created by the same mechanism that decreases the error for the single-pass library, namely the increase in Δ_τ. As visible in figure 27 the TDOA peaks are washed out and often multiple traces are visible in close proximity. This is almost definitely caused by differences in bathymetry for different positions of the RHIB, which are all used to create this library. With larger magnitudes of Δ_τ these TDOA traces end up merging, resulting in a TDOA peak that spans more quefrencies.

Apparently the Lloyd mirror pattern is also strong enough on the lower frequency range generated by down-sampling as no definite reduction in accuracy by down-sampling could be measured.

Sub-blocks have been proven as a simple yet effective method to reduce the ranging error. In both table 6 and 8 a significant decrease of error can be seen when compared to table 2, while computation time was not significantly affected. Apparently this method of averaging while the signal is relatively stationary due to the short time-span is able to reduce noise in the cepstra.

As expected does the skipping of data to generate cepstra not have significant effects on the ranging error. Since the data blocks for cepstrum generation span such a short time there is no need to compute them for all available data. This result can be used to reduce the computation time significantly.

Finally an interesting observation can be made when comparing the computation time in the various tables regarding the down-sampling results. Here an increase of computation time is actually visible when down-sampling with a factor of two. This is created by the `decimate` algorithm and the included low-pass filter which takes more time to execute than the time needed to perform the computation on the original sample rate. Only when down-sampling with larger factors is the computation time actually reduced. When recording on a lower sample rate this would obviously not be an issue.

5.2 Library and matching algorithm effectiveness

In section 4.2 the effectiveness of both library types in combination with three algorithms were explored, with and without tidal influence.

When using the single-pass library both the Pearson algorithm and the minimisation algorithm show a comparable performance. The minimisation algorithm however seems to show a small decrease in spread between adjacent ranging values. As this algorithm has the lowest computation speed of all matching algorithms these results suggest that the minimisation algorithm is the best choice in combination with a single-pass library. This can be interpreted as showing that the covariance of the library and measurement is a less suited indicator for matching than comparing magnitude as with the minimisation algorithm. Unfortunately the cross-correlation algorithm has a comparably worse performance and tends to underestimate the range to the RHIB.

An increase in error magnitude when on maximum range is not visible in the data and it is expected that cepstral ranging can be used for ranges further than the maximum of approximately 700m found in this dataset. A definite maximum can not be found from this dataset and is expected to also depend on the sound level of the source. Unfortunately no quantitative sound levels of the RHIB from these measurements can be attained.

The averaged library continues these trends as this library also shows less spread for the minimisation algorithm. This is especially noticeable during the stationary section in figure 28b. The shape of the second pass is however better recreated by the Pearson algorithm. Again the cross correlation algorithm shows a comparably worse performance, resulting often in incorrect minimum or maximum ranges. While the averaged library shows a superior performance compared to the single-pass library at first glance, this is certainly created by having all measurements which are to be ranged in the library. It was hoped that this would not be the case by averaging multiple ranges but apparently these exact cepstra are still available in original form in the library. When comparing the second pass of the averaged library, which was not included in the library, to the second pass in the single-pass library it becomes evident that the single-pass library actually offers better ranging performance. This might be explained by the reduced library cepstra magnitudes as visible in figure 27. These reduced magnitudes are probably caused by averaging, as an inclusion of a few negative values for a certain discrete quefrency and range will reduce the overall magnitude of that quefrency and range in the library.

5.3 Tidal influence on ranging

When it comes to ranging with a change in water height are both libraries and their algorithms virtually unable to perform ranging with the default parameters as selected previously. One would assume that an increase in water height would only result in an underestimation of the range to the source. Figure 33 validates this assumption, as the cepstrum peaks of the bottom-surface path and it's first rahmonic align with the same peaks of the library with an approximate 10% reduction in range for a 10% increase of water height. However, there is no correlation between both cepstra after these two peaks. Consequently the ranging performance with tidal mismatch can be improved in this case by limiting the maximum generated quefrency to the quefrency of the first rahmonic. This limit improved the ranging performance considerably as visible in figure 34.

The most probable explanation for this result is the lower velocity of the vessel and consequently a lower sound level. It was theorized that possibly another reflection path, different than the bottom-surface path as expected according to section 3.2.1, was present in the data. This could possible result in a nonlinear scaling between quefrency peaks with a change in bathymetry and distance, and therefore hinder matching. However no proof for this could

be found, as all cepstrum peaks for both the high- and low tide data were found to be on quefrencies scalable with an integer factor from the first peak, i.e. the original peak and it's rahmonics.

Considering that the ranging performance is increased in the later half of the run, after the RHIB has accelerated, it can be assumed that this is the primary factor that decreases ranging performance during the high tide measurements. The SNR is certainly increased by the higher engine load, which is assumed to increase the amount or magnitude of rahmonics present at higher quefrencies. Limiting the quefrency is therefore an obvious solution, however this limits the minimum range at which cepstral ranging can be performed. Therefore limiting of the maximum quefrency should only be done when it is certain that the source is at a further range than the minimum range imposed by the quefrency limit.

Another factor that improved the ranging performance with tidal mismatch was decreasing the down-sample factor to 8, while decreasing it further did not result in an decrease of ranging error. This result can possibly be explained by the increase in resolution. The lower speed and change in water height possibly introduces small perturbations in the cepstra which influence the TDOA peak and rahmonics on this lower resolution, while the higher resolution is able to keep the perturbations separate.

5.4 Responsiveness to other vessels

The measurements used in this thesis all involve the same RHIB as a sound source. However, at 10:16 a different vessel passes the sensor, which was identified as the MARKENJE, a small dredging vessel. Judging by the slope of the cepstral lines in figure 36 around 10:16 this vessel had a significantly lower speed than the RHIB. This lower speed and resulting lower engine sound level possibly explains why less rahmonics are visible in the cepstrogram. Unfortunately less rahmonics in a cepstrum also mean that matching against a library with cepstra that contain more rahmonics is less reliable. This, together with the sound levels of greater magnitude caused by the RHIB, explains why the ranging algorithm only picked up this vessel when it got within reasonable distance of the sensor. This explanation matches the explanation for tidal influence as given in section 5.3.

Unfortunately the AIS location was only recorded once every minute, which means that the ranging accuracy for this vessel can not be measured. This result does however show that an experimentally generated cepstrum library shows some sensitivity to specific vessels or engine speed, even if their characteristics are quite similar. The MARKENJE has an inboard diesel engine with Z-drive propulsion and while the size of the vessel and propeller depth, and thus *source* depth, are roughly comparable to the RHIB, the generated cepstrum still shows small differences.

5.5 Ray tracing for TDOA estimation

As this method is not automated there is no accurate method to compare the results of ray tracing to the results of the experimental methods. The visualised results do however look very promising. The ray tracing method is able to match the TDOA lines in the cepstra almost exactly while correctly predicting the influence of tidal changes. If automated in similar fashion as the experimental methods, a SMAPE of less than 5% can be expected. This accuracy also brings an issue with the GPS or audio data to light. As evident in figure 38 there seems to be a mismatch between the time of the GPS data and the time of the audio data. That is the only logical explanation as sound travel times in water for these ranges are significantly smaller than the mismatch. Determining which one of these two sets is responsible for this mismatch was found to be impossible. As this mismatch was not constant it was impossible to accurately compensate for this issue. This means that this issue must also be present in the previous

results, although not directly evident because of the lower accuracy.

The assumption that, based on salinity and temperature measurements, there will be no significant refraction of sound waves in the measurement waveguide was found to be correct. If not larger errors in TDOA estimation from the ray tracing method should have been encountered. As the day of measurements was a sunny day with temperatures of 23 °C at sea, some density differences between layers in the waveguide could have been expected, but fortunately this effect was found to be insignificant. To completely rule out refraction some measurements should be taken at the end of summer, when the coastal areas might have been warmed by longer periods of warm weather. However, for practical applications during most of the year, refraction is not expected to count a significant shift in TDOA estimation based on the data in this thesis.

5.6 Recommendations for future research

The aim of this thesis was to lay the foundation for a practical cepstral ranging method that can be used to survey coastal areas at a time where sub-sea infrastructure is at an elevated risk. Unfortunately some issues were encountered that first have to be solved before any practical deployment can be considered. Therefore the recommendations for future research will mostly be aimed at the mitigation of these problems.

The primary recommendation is to automate the ray tracing algorithm as presented in section 3.4.3. This might be done by finding an effective method to create synthetic cepstra from the ray tracing TDOA results which should be suitable for matching from a library analog to the experimental methods in this thesis. Another suggestion would be to create an effective peak-finding algorithm that is able to reliably find the first two or three TDOA rahmonics in a cepstrum to match them to a ray tracing TDOA value. While this peak-finding algorithm might impose a serious challenge it would offer a very elegant solution as it removes the need for comparison of complete cepstra. If a reliable solution has been found that can couple ray tracing results to measured cepstra the result would most probably be very effective and accurate. Next to that does ranging based on only the cepstrum peak with smallest quefrency and not the following rahmonics increase the robustness of ranging. As shown in section 5.4 cepstra can vary somewhat between vessels of comparable characteristics as rahmonics can vary in strength, but the cepstrum peak with the smallest quefrency is always expected to be present. Using only this peak can increase the reliability between different vessels. Of course the source depth needs to be comparable between vessels to keep the multi-path propagation in the waveguide constant.

The second recommendation would be to further investigate the influence of source sound level and consequently SNR on the matching performance. The most probable explanation for the difficulties of performing ranging on a different vessel or the same vessel at high tide with a lower engine speed was found to be differences in sound level compared to the library. The data available however was insufficient to say this with certainty. Different measurements with controlled tidal height and engine speed to find a correlation between these factors and ranging performance are expected to offer further insight into the exact cause. When the cause is identified it can be expected that methods to circumvent the degradation of ranging performance can be found.

Finally it might prove effective to combine the methods in this thesis with machine learning or artificial intelligence. This relatively new field has a huge potential when it comes to pattern recognition and might therefore be effective at matching a measurement to a library.

5.6.1 Possible practical applications

First the current issues with cepstral ranging as documented in this report need to be solved. Once this method has matured sufficiently it offers a robust ranging method that only requires a single, readily available sensor and can perform in shallow water.

When paired with a method to perform bearing estimation it offers the possibility of 2D localisation on the horizontal plane for an area with a radius of at least 700 meters. One possible configuration could be three hydrophones, placed within short distances of each other but not on the same line in the horizontal plane. These three hydrophones can be split in three possible pairings, which together with a cross correlation algorithm can find the unambiguous bearing to a target as shown in section 2.2.1.1. Cepstral ranging can then be performed on any of the hydrophone signals, possibly including beamforming for a small array gain, to find the accompanying range and thus the polar coordinates of the source.

Another possible configuration would be to use a vector sensor as used for the measurements in this report. The internal hydrophone signal can be used for cepstral ranging while the signals of the three movement sensors can be used to find the unambiguous bearing, resulting in a similar localisation method as described above.

Both these suggested configurations require only a small amount of readily available hardware and can perform localisation from a single point. Using a vector sensor results in an even smaller sensor setup, the sensor package used to generate data for this thesis for example takes up less than one square meter. However, the first suggested configuration of three hydrophones uses less expensive and possibly more robust hardware. It's size will mostly depend on the maximum wavelength that needs to be detected for bearing estimation.

6 Conclusion

While cepstral ranging has had some articles dedicated to it in literature it was found that research on this topic was mostly incomplete. No data could be found on the effects of important parameters for cepstrogram generation nor on the algorithms suitable to convert a cepstrum to an estimate of range and the implications of environmental factors. The aim of this thesis was therefore to generate this data and to give an overview on important factors to reckon with when selecting a cepstral ranging method for practical usage, and to show what can be expected of cesptral ranging when it comes to performance.

6.1 Cepstral parameters

It was found that a lower sample rate of approximately 2.5 kHz actually increases the accuracy when paired with the single-pass library as long as the library cepstra match the measurement cepstra well. With mismatch created by a change in bathymetry or SNR, a slightly higher sample rate of 10 kHz gave better results. Another interesting conclusion is that not all data has to be analyzed for maximum accuracy as the data-blocks relatively small when limited to the maximum quefrency of interest. The method of using sub-blocks to reduce noise was found to be effective.

6.2 Library types

Two (experimental) library types were conceived in this thesis. This includes the single-pass library, where only a small part of experimental data is used to generate a library, and the averaged library where a larger part of experimental data is selected and averaged in groupings of comparable ranges to generate a library. It was found that the single-pass library is superior to the averaged library when measurements are taken in an environment with varying bathymetry. This variation produces smearing in the averaged library which hinders ranging as it decreases the amplitude of cepstrum peaks of interest. In an environment of virtually range independent bathymetry, or where data is used with the exact same environmental parameters and location, the averaged library might offer some advantage in noise reduction.

Next to these experimental methods a ray tracing algorithm was developed with the aim to create a synthetic library. While efforts to create this synthetic library to automate ranging failed, this algorithm still proved to be impressively accurate in forecasting TDOA values. No TDOA estimation errors by refraction should be expected when using this method in the environment of measurements. Once a method has been found to create a library from this algorithm, or any other solution to compare measurement cepstra to the ray tracing results, the ray tracing method is expected to offer significantly superior performance over the experimental libraries. Next to an increase in ranging performance it offers the possibility of creating the libraries from open source data, without the need for time consuming and expensive measurements.

6.3 Matching algorithms

Three different matching algorithms were investigated in this thesis. The Pearson algorithm which compares covariance of two sets, the minimisation algorithm which compares the magnitude for each discrete quefrency and the cross correlation algorithm to correlate two sets of data. While the Pearson and minimisation algorithms showed comparable performance the spread of ranging values was deemed to be lower for the minimisation algorithm. It is recommended to use this algorithm in combination with a low sample rate as it also is computationally more efficient than the Pearson algorithm. The cross correlation algorithm showed a significantly worse performance, unless there was a mismatch between library cepstra and measurement cepstra. In that case the cross correlation algorithm performed comparable to the Pearson algorithm, while the minimisation algorithm suffered a more significant reduction from this mismatch.

6.4 Tidal influence and mismatch

As expected the change in water height by tidal movement has a significant effect on ranging performance, although this effect is more profound than anticipated. A systematic error depending on range can be expected from this change in water height, however with the cepstrum computation settings as selected in section 4.1 no ranging can be performed and the range estimation mostly consists of noise. The main cause was identified to be in a variation between the library cepstra and the measurement cepstra at high tide at the higher quefrency range. No other reflection paths that show nonlinear behaviour with the tidal mismatch, and thus hinder ranging, could be identified. It is therefore assumed that this decreased ranging performance is caused by the significantly lower engine speed of the RHIB during the high tide measurements. By reducing the maximum generated quefrency and increasing the sample rate the ranging performance could be increased significantly, although not to the levels without mismatch.

A second vessel that passed the sensor at lower speed during the low tide measurements where the libraries were created was also ranged less accurate. This reinforces the conclusion that the decrease in ranging performance is the result of a lower engine speed and SNR.

6.5 Overall accuracy

The maximum realistic accuracy attained in this thesis has been a SMAPE of 25% using the single-pass library and the minimisation matching algorithm, which can be improved to 16% when a moving average filter is added. The used library however has to be created at exactly the correct water height for these results. When using a directional sensor setup and a library that also accounts for the bearing of the source and therefore bathymetry this value might be improved. The ray tracing method, while not yet automated, shows potential to have a SMAPE of approximately 5%.

When it comes to range, no specific dependency of error and range can be discovered. Cepstral ranging performs well up to the maximum range of the RHIB in this dataset, which is approximately 700m. This result is impressive considering the shallow depth of the waveguide and demonstrates the effectiveness of cepstral ranging in shallow water. In practice this decreases the amount of sensors needed to survey a specific area.

6.6 Contributions to the field of cepstral ranging

The emerging field of cepstral ranging shows impressive ranging results in literature. However, the data on which these findings were based is often generated in highly controlled experiments with little research presented on possible detrimental factors. The aim of this thesis was to investigate the potential of cepstral ranging in a less controlled environment and therefore to validate it's performance in practical, real world situations.

Therefore a dataset was selected that was not necessarily recorded for this thesis, but still contained relevant data to evaluate cepstral ranging in a real world setting. Using this data the performance of cepstral analysis was found to be indeed spectacular with an experimentally generated library, but severely dependent on an exact match between library and measurements, which was not apparant from literature. Literature on cepstral ranging mostly use the same data to generate both the library and the measurements that need to be ranged.

Multiple matching algorithms, two of which were not yet used for cepstral ranging in literature, were evaluated. Ranging performance was found to be dependent on engine speed or SNR, especially in the higher quefrency range, a conclusion which was not found in any existing literature. This thesis shows that the negative effect of SNR on ranging performance can be reduced by limiting the maximum quefrency. It was also concluded that a low sample rate in the 2Khz range offers the best performance without mismatch between measurements and library, while a higher sample rate in the 10kHz range offers better performance with a mismatch between data.

The foundation was laid of a ray tracing method that can accurately predict TDOA values with varying environmental parameters like tidal height or sound speed. This algorithm only uses open-source data to generate TDOA estimations which were found to be extremely accurate for all measurements available in the TNO dataset. While the algorithm is not yet able to perform automated ranging, it provides a step to a system that is not dependent on experimentally generated data for TDOA estimation, nor prone to mismatch by environmental factors.

In conclusion this thesis provides an advancement of cepstral ranging to convert it from a theoretical method to practical application. Encountered problems with cepstral ranging, not yet presented in literature, were identified and at least partially solved, or suggestions for further research to solve these problems was given.

Bibliography

[1] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics*, 2nd ed. New York: Springer-Verlag, 2011, ch. 10.

[2] L. AN and L. CHEN, "Underwater acoustic passive localization base on multipath arrival structure," 11 2014.

[3] BBC, "Nord stream blast 'blew away 50 metres of pipe," 2022, last accessed 21 November 2022. [Online]. Available: https://www.bbc.com/news/world-europe-63297085

[4] ——, "Damaged cable leaves shetland cut off from mainland," 2022, last accessed 21 November 2022. [Online]. Available: https://www.bbc.com/news/uk-scotland-north-east-orkney-shetland-63326102

[5] C. Bueger, T. Liebetrau, and J. Franken, "Security threats to undersea communications cables and infrastructure – consequences for the eu," European Union, Report, 2022.

[6] NOS, "Mivd: Russen willen nederlandse energievoorziening verstoren," 2023, last accessed 25 March 2023. [Online]. Available: https://nos.nl/artikel/2464616-mivd-russen-willen-nederlandse-energievoorziening-verstoren

[7] Y. Gao, M. Clark, and P. Cooper, "Time delay estimate using cepstrum analysis in a shallow littoral environment," in *Conference: Undersea Defence Technology, Europe*, 06 2008.

[8] E. L. Ferguson, S. B. Williams, and C. T. Jin, "Convolutional neural network for single-sensor acoustic localization of a transiting broadband source in very shallow water," *The Journal of the Acoustical Society of America*, vol. 146, no. 6, pp. 4687–4698, 2019.

[9] E. Ferguson, S. Williams, and C. Jin, "Sound source localization in a multipath environment using convolutional neural networks," 04 2018, pp. 2386–2390.

[10] E. L. Ferguson, "Multitask convolutional neural network for acoustic localization of a transiting broadband source using a hydrophone array," *The Journal of the Acoustical Society of America*, vol. 150, no. 1, pp. 248–256, 2021.

[11] K. W. Lo, "A matched-field processing approach to ranging surface vessels using a single hydrophone and measured replica fields," *The Journal of the Acoustical Society of America*, vol. 149, no. 3, pp. 1466–1474, 2021.

[12] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics*, 2nd ed. New York: Springer-Verlag, 2011, ch. 1.

[13] R. Urick, "Sound propagation in the sea," DARPA, Washington, D.C. 20402, Tech. Rep., 1979, chapter 4.

[14] ——, "Sound propagation in the sea," DARPA, Washington, D.C. 20402, Tech. Rep., 1979, chapter 6.

[15] ——, "Sound propagation in the sea," DARPA, Washington, D.C. 20402, Tech. Rep., 1979, chapter 9.

[16] ——, "Sound propagation in the sea," DARPA, Washington, D.C. 20402, Tech. Rep., 1979, chapter 10.

[17] F. Jensen and W. Kuperman, "Optimum frequency of propagation in shallow water environments," *Acoustical Society of America*.

[18] J. K. Wilson, "Maritime surveillance using a wideband hydrophone," Master's thesis, Naval Postgraduate School, Monterey, California, sep 2007.

[19] L. Fillinger, A. Sutin, and A. Sedunov, "Acoustic ship signature measurements by cross-correlation method," *The Journal of the Acoustical Society of America*, vol. 129, pp. 774–778, 2011.

[20] L. Fillinger, A. J. Hunter, M. Zampolli, M. C. Clarijs, and K. Verolme, "Improving protection against intruders using passive sonar," in *Proceedings 4th International Conference and Exhibition on Underwater Acoustic Measurements - Technologies & Results*, ser. UAM, 2011.

[21] K. W. Chung, A. Sutin, A. Sedunov, and M. Bruno, "Demon acoustic ship signature measurements in an urban harbor," *Advances in Acoustics and Vibration*, 2011.

[22] J. Gebbie and M. Siderius, "A two-hydrophone range and bearing localization algorithm with performance analysis," *The Journal of the Acoustical Society of America*, vol. 137, p. 1586–1597, 2015.

[23] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, no. 4, pp. 320–327, 1976.

[24] M. Fakharzadeh, S. Jamali, P. Mousavi, and S. Safavi-Naeini, "Fast beamforming for mobile satellite receiver phased arrays: Theory and experiment," *Antennas and Propagation, IEEE Transactions on*, vol. 57, pp. 1645 – 1654, 07 2009.

[25] M. E. G. D. COLIN, "Underwater detection, classification and localisation: Improving the capabilities of towed sonar arrays," Ph.D. dissertation, TU Delft, Delft, Nederland, 04 2011.

[26] J. Grythe, "Beamforming algorithms - beamformers," Norsonic AS, Tech. Rep., 2016.

[27] S. Ijsselmuide and P. Beerens, "Adaptive beamforming algorithms for passive sonar arrays," 06 2001.

[28] I. Solomon, A. Knight, and M. Greening, "Sonar array signal processing for sparse linear arrays," in *ISSPA '99. Proceedings of the Fifth International Symposium on Signal Processing and its Applications (IEEE Cat. No.99EX359)*, vol. 2, 1999, pp. 527–530 vol.2.

[29] W. Coventry, C. Clemente, and J. Soraghan, "Polynomial root-music algorithm for efficient broadband direction of arrival estimation," in *2017 Sensor Signal Processing for Defence Conference (SSPD)*, 2017, pp. 1–5.

[30] L. Fillinger, A. J. Hunter, M. Zampolli, and M. C. Clarijs, "Passive acoustic detection of closed-circuit underwater breathing apparatus in an operational port environment," *The Journal of the Acoustical Society of America*, vol. 132, no. 4, pp. EL310–EL316, 2012.

[31] H. Zhang and H. Zhang, "Research on doa estimation method of sonar radar target based on music algorithm," *Journal of Physics: Conference Series*, vol. 1176, p. 032001, 03 2019.

[32] G. L. D'Spain and W. A. Kuperman, "Application of waveguide invariants to analysis of spectrograms from shallow water environments that vary in range and azimuth," *The Journal of the Acoustical Society of America*, vol. 106, no. 5, pp. 2454–2468, 1999.

[33] K. L. Cockrell, "Understanding and utilizing waveguide invariant range-frequency striations in ocean acoustic waveguides," Ph.D. dissertation, Massachusetts Institute of Technology, 2011.

[34] B. Bogert, M. Healy, and J. Tukey, "The quefrency alanysis of time series for echoes: Cepstrum, pseudo-autocovariance, cross-cepstrum, and saphe cracking," in *Proc. of the Symp. on Time Series Analysis*, 1963, pp. 209–243.

[35] R. Randall, "A history of cepstrum analysis and its application to mechanical problems," *Mechanical Systems and Signal Processing*, vol. 97, 12 2016.

[36] A. Oppenheim and R. Schafer, "From frequency to quefrency: a history of the cepstrum," *IEEE Signal Processing Magazine*, vol. 21, no. 5, pp. 95–106, 2004.

[37] S. Tavakkoli, "Application of the cepstrum technique to location of acoustic sources in the presence of a reflective surface," Master's thesis, Virginia Polytechnic Institute and State University, 1986.

[38] N. C. Durofchalk, J. Jin, H. J. Vazquez, K. L. Gemba, J. Romberg, and K. G. Sabra, "Data driven source localization using a library of nearby shipping sources of opportunity," *JASA Express Letters*, vol. 1, no. 12, p. 124802, 2021.

[39] C. M. A. Verlinden, J. Sarkar, W. S. Hodgkiss, W. A. Kuperman, and K. G. Sabra, "Passive acoustic source localization using sources of opportunity," *The Journal of the Acoustical Society of America*, vol. 138, no. 1, pp. EL54–EL59, 2015.

[40] A. Baggeroer, W. Kuperman, and P. Mikhalevsky, "An overview of matched field methods in ocean acoustics," *IEEE Journal of Oceanic Engineering*, vol. 18, no. 4, pp. 401–424, 1993.

[41] B. M. Worthmann, H. C. Song, and D. R. Dowling, "Adaptive frequency-difference matched field processing for high frequency source localization in a noisy shallow ocean," *The Journal of the Acoustical Society of America*, vol. 141, no. 1, pp. 543–556, 2017.

[42] A. Hunter, L. Fillinger, M. Zampolli, and M. Clarijs, "Data fusion from multiple passive sonar nodes for target localisation and false alarm reduction," 2012, 11th European Conference on Underwater Acoustics ; Conference date: 01-07-2012 Through 06-07-2012.

[43] P. Beerens, S. Ijsselmuide, and A. Koersel, 06 2003.

[44] S. von Benda-Beckmann, P. Beerens, and S. Ijsselmuide, "Effect of towed array stability on instantaneous localization of marine mammals," *The Journal of the Acoustical Society of America*, vol. 134, pp. 2409–17, 09 2013.

[45] A. Tesei, F. Meyer, and R. Been, "Tracking of multiple surface vessels based on passive acoustic underwater arrays," *The Journal of the Acoustical Society of America*, vol. 147, no. 2, pp. EL87–EL92, 2020.

[46] C. Zhu, H. Garcia, A. Kaplan, M. Schinault, N. O. Handegard, O. R. Godø, W. Huang, and P. Ratilal, "Detection, localization and classification of multiple mechanized ocean vessels over continental-shelf scale regions with passive ocean acoustic waveguide remote sensing," *Remote Sensing*, vol. 10, no. 11, 2018. [Online]. Available: https://www.mdpi.com/2072-4292/10/11/1699

[47] B. Ferguson and R. Wyber, "Wavefront curvature passive ranging in a temporally varying sound propagation medium," in *MTS/IEEE Oceans 2001. An Ocean Odyssey. Conference Proceedings (IEEE Cat. No.01CH37295)*, vol. 4, 2001, pp. 2359–2365 vol.4.

[48] D. Dardari and F. Guidi, "Direct position estimation from wavefront curvature with single antenna array," in *2018 8th International Conference on Localization and GNSS (ICL-GNSS)*, 2018, pp. 1–5.

[49] B. G. Ferguson, "Wide aperture arrays for locating impulsive sound sources in air and underwater," in *Unattended Ground, Sea, and Air Sensor Technologies and Applications VIII*, E. M. Carapezza, Ed., vol. 6231, International Society for Optics and Photonics. SPIE, 2006, p. 62310X.

[50] E. L. Ferguson and B. G. Ferguson, "High-precision acoustic localization of dolphin sonar click transmissions using a modified method of passive ranging by wavefront curvature," *The Journal of the Acoustical Society of America*, vol. 146, no. 6, pp. 4790–4801, 2019.

[51] A. Tesei, S. Fioravanti, V. Grandi, P.Guerrini, and A.Maguer, "Acoustic surveillance of small surface vessels in confined areas," in *PROCEEDINGS OF SYMPOL-2011*, 2011.

[52] A. Pollara, A. Sutin, and H. Salloum, "Improvement of the detection of envelope modulation on noise (demon) and its application to small boats," in *OCEANS 2016 MTS/IEEE Monterey*, 2016, pp. 1–10.

[53] Y. Wen and M. Henry, "Time Frequency Characteristics of the Vibroacoustic Signal of Hydrodynamic Cavitation ," *Journal of Vibration and Acoustics*, vol. 124, no. 4, pp. 469–475, 09 2002.

Appendices

A Matlab code

A.1 AIS data import

```
function AISDATA = import_AIS(latlim2,lonlim2)
%% Import AIS data

B1 = readtable("2022-05-17_decoded_type1.csv");
B2 = readtable("2022-05-17_decoded_rest.csv");

B2 = B2((B2.msg_type < 4 | B2.msg_type == 18),:);

B11 = B1(:,["timestamp","msg_type","lat","lon","mmsi","speed"]);
B22 = B2(:,["timestamp","msg_type","lat","lon","mmsi","speed"]);

B33 = [B11;B22];

B33 = sortrows(B33,"mmsi");

B33 = B33(~(B33.lat < latlim2(1) | B33.lat > latlim2(2)),:);
B33 = B33(~(B33.lon < lonlim2(1) | B33.lon > lonlim2(2)),:);

B33.timestamp = datetime(B33.timestamp*1e6,'ConvertFrom','epochtime','TicksPerSecond',1e6,'
    Format','dd-MMM-yyyy HH:mm:ss.SSSSSS') + hours(2);

AISDATA = B33;

end
```

A.2 Cepstrogram generation

```
function [rsb,tb,tt] = pceps(y,fs,bs,sbs,bov,sbov,limits)
% [rsb,tb,tt] = pceps(y,fs,bs,sbs,bov,sbov,limits) calculates the power
% cepstrum of dataset y which is sampled with samplerate fs.
%
% The dataset is split into blocks of size bs with an overlap of bov.
% The blocks can also be divided into 5 subblocks with size sbs and
% overlap sbov. If this is not required, fill in [] for sbs and/or sbov.
%
% Maximum time delay can be set with the vector 'limits' containing the
% lower and upper limit. If not required, fill in [] for limits which
% results in a maximum timedelay of sbs/2/fs or bs/2/fs.

i = 1;
j = 0;
k = 1;

if or(isempty(sbs),isempty(sbov)) == true

    sbs = [];
    sbov = [];
    perform_averaging = false;

else

    perform_averaging = true;

end

if perform_averaging == true

    steps = (bs-sbs)/((1-sbov)*sbs);

end

while i < length(y)

```

```
    if i+bs-1 > length(y)

        break

    end

        yb = y(i:i+bs-1);

        if perform_averaging == true

            while j <= steps

                ysb = yb((1:sbs)+(j*(1-sbov)*sbs));

%                 fsb(:,j+1) = abs(fft(hanning(length(ysb)).*ysb)).^2;
                fsb(:,j+1) = abs(fft(ysb)).^2;
                j = j+1;

            end

            fsb_mean(:,k) = mean(fsb,2);
            rsb(:,k) = ifft(log10(fsb_mean(:,k)));

        else

%             fsb = abs(fft(hanning(length(yb)).*yb)).^2;
            fsb = abs(fft(yb)).^2;
            rsb(:,k) = ifft(log10(fsb));


        end

        tb(k) = i/fs;


    i = i + (bov*bs);
    j = 0;
    k = k+1;

end

rsb = rsb(1:round(size(rsb,1)/2),:);

%     i = i-(bov*bs)+bs-1;

% Constrain time lag and perform normalisation

if isempty(limits) == false

    if round(limits(2)*fs) > size(rsb,1)

        limit_upper = size(rsb,1);

    else

        limit_upper = round(limits(2)*fs);

    end

    limit_lower = round(limits(1)*fs);

    rsb(limit_upper+1:end,:) = [];
    rsb(1:limit_lower,:) = [];
    tt = (limit_lower:limit_upper-1)/fs;

else

    tt = (0:size(rsb,1)-1)/fs;

end


end
```

A.3 Raytracing algorithm

```
function [pathlength,directlength,delta_t,x_ref,y_ref,err] = raytracing_function_num3(xb,yb,xs
    ,ys,target,theta,c_0,res)

xbb = 0:res:max(xb);
ybb = interp1(xb,yb,xbb);

for ik = 1:length(theta)

%% First ray

% First ray setup
p0 = tand(theta(ik));

ray1 = p0*xbb + ys;

% first reflection computation
delta_1 = abs(ray1 - ybb);
[~,sol11_I] = mink(delta_1,2);
delta_points11 = delta_1(sort(sol11_I));
sol11_I = sort(sol11_I);

sol1x(ik) = xbb(sol11_I(1)) + (delta_points11(1)/(sum(delta_points11))) * res;
sol1y(ik) = p0 * sol1x(ik) + ys;

% angle of reflecting surface
p1 = (ybb(sol11_I(2)) - ybb(sol11_I(1))) / res;

% angle of reflected ray
theta_p2 = -theta(ik) + 2*atand(p1);

%% second ray

% second ray setup
p2 = tand(theta_p2);

ray2 = p2*xbb + sol1y(ik) - (p2 * sol1x(ik));

sol2x(ik) = sol1x(ik) + -sol1y(ik)/p2;
sol2y(ik) = 0;


%% Third ray

ray3 = -p2*xbb - (-p2 * sol2x(ik));

delta_3 = abs(ray3 - ybb);
[test3,sol33_I] = mink(delta_3,2);
delta_points33 = delta_3(sort(sol33_I));
sol33_I = sort(sol33_I);

if test3(1) > 2 * res

    sol1x(ik) = [];
    sol1y(ik) = [];
    sol2x(ik) = [];
    sol2y(ik) = [];
    break

end

sol3x(ik) = xbb(sol33_I(1)) + (delta_points33(1)/(sum(delta_points33))) * res;
sol3y(ik) = -p2*sol3x(ik) - (-p2 * sol2x(ik));

err(ik) = sol3x(ik) - target;
x_ref_temp = [xs sol1x(ik) sol2x(ik) sol3x(ik)]';
y_ref_temp = [ys sol1y(ik) sol2y(ik) sol3y(ik)]';

d = diff([x_ref_temp y_ref_temp]);

tl = sum(sqrt(sum(d.*d,2)));

```

```
pathlength(ik) = tl(end);

directlength(ik) = sqrt(sol3x(ik)^2 + (sol3y(ik) - ys)^2);
delta_t(ik) = (pathlength(ik)-directlength(ik))/c_0;

x_ref(ik,:) = x_ref_temp';
y_ref(ik,:) = y_ref_temp';

end

if exist('sol3x',"var") == 0

    pathlength = NaN;
    err = NaN;
    x_ref = [NaN NaN NaN NaN];
    y_ref = [NaN NaN NaN NaN];
    directlength = NaN;
    delta_t = NaN;

else

end

end
```

A.4 Raytracing solver

```

function [pathlength,directlength,delta_t,x_ref,y_ref,err] = iterative_raysolver(xb,yb,xs,ys,
    target,margin,theta,c_0,res)

%% Iterative raytracing

[pathlength,directlength,delta_t,x_ref,y_ref,err] = raytracing_function_num3(xb,yb,xs,ys,
    target,theta,c_0,res);

iq = 1;

while min(abs(err)) > margin

    err_neg = err;
    err_pos = err;
    err_neg(err>0) = nan;
    err_pos(err<0) = nan;
    [~,I_neg] = max(err_neg);
    [~,I_pos] = min(err_pos);

    theta = linspace(theta(I_neg),theta(I_pos),10);

    [pathlength,directlength,delta_t,x_ref,y_ref,err] = raytracing_function_num3(xb,yb,xs,ys,
    target,theta,c_0,res);

    if iq > 100

        disp(theta)
        disp(err)
        disp('stuck in iteration')
        break

    end

    iq = iq+1;

end

[err,I_search] = min(err,[],"ComparisonMethod","abs");
pathlength = pathlength(I_search);
directlength = directlength(I_search);
delta_t = delta_t(I_search);
x_ref = x_ref(I_search,:);
y_ref = y_ref(I_search,:);

```

```
43 end
```