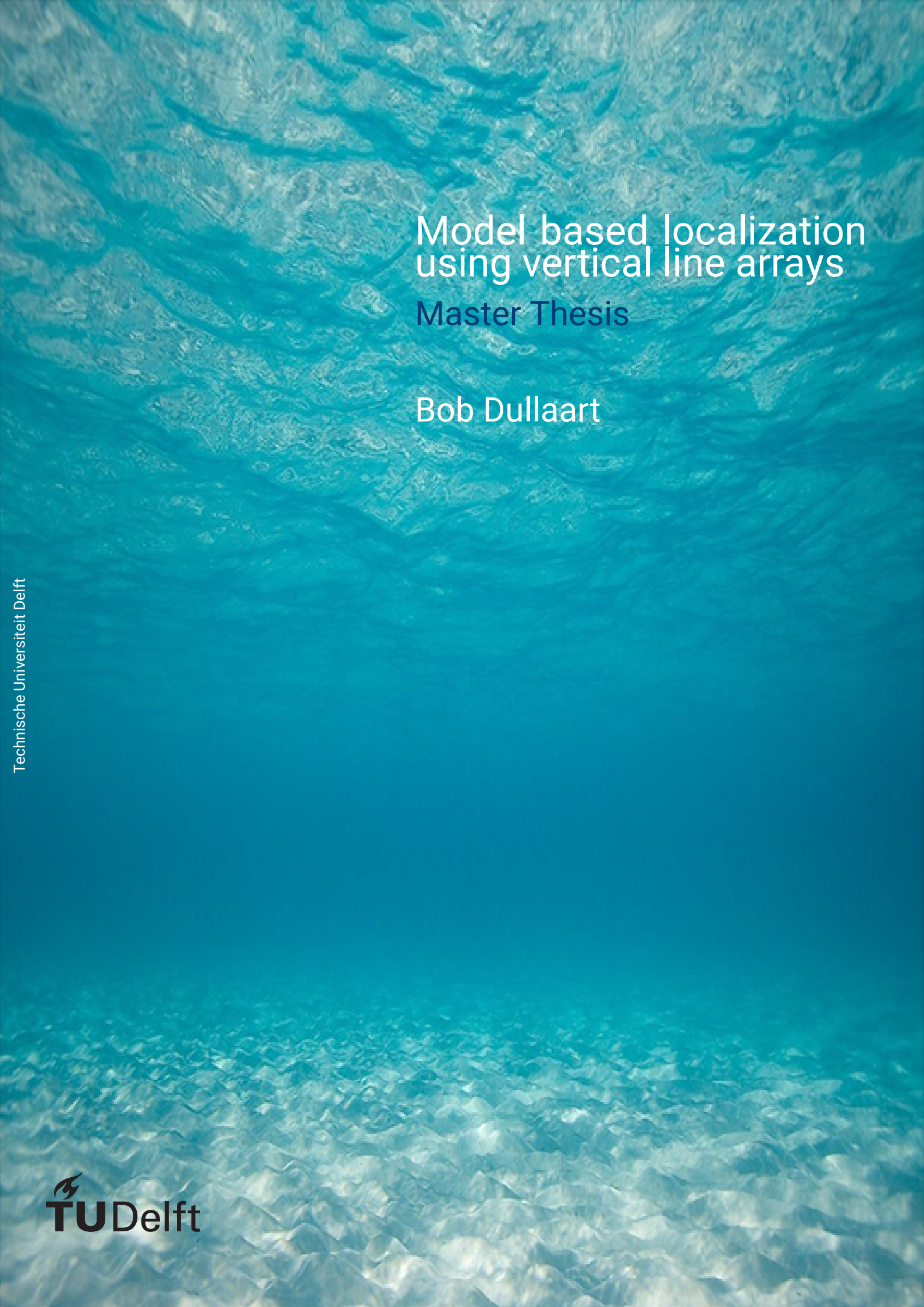


Model based localization using vertical line arrays

Master Thesis

Bob Dullaart

Technische Universiteit Delft

TU Delft

Model based localization using vertical line arrays

Master Thesis

by

Bob Dullaart

Abstract

On land, localization or ranging is typically performed by using electromagnetic waves. In an underwater environment, this becomes difficult, as electromagnetic waves dampen out fast. As a solution often acoustic waves are used to perform localization. Because of the complexity of underwater acoustic propagation, an acoustic propagation model is used of which the predicted output for different candidate locations is compared to the measurements. Methods to localize low-frequency sound sources with widely spaced receivers have been thoroughly studied in literature. More recently, also for high-frequency sources and receivers with a line-array structure successful localization has been demonstrated. In this thesis, the typical methods used to achieve this are gathered and typical challenges of these methods are identified. Based on these challenges, a new method is proposed, in an attempt to reduce the effects caused by these challenges and thus improve the localization performance. Also, Gibbs sampling is used to perform Bayesian inversion on the localization problem and simultaneously estimate environmental parameters. The results show that a high-frequency acoustic source at a range of 1500-2000m in an ocean environment with a depth of ~ 200 meters can be localized by using the proposed method, including an accurately estimated tilt of the receiver array.

Student number:	4290631	
Project duration:	March, 2020 – December, 2020	
Thesis committee:	Prof. dr. ir. G.J.T. Leus,	TU Delft, chair
	Dr. A.M. von Benda-Beckmann,	TNO, supervisor
	Dr. F. Fioranelli,	TU Delft

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TU Delft

Contents

1 Introduction

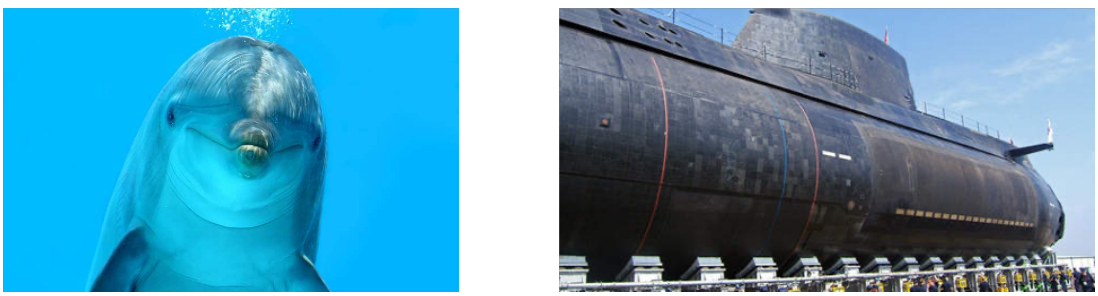
In localization problems, measuring the propagation of electromagnetic (EM) waves often offers a good solution. This is for example used in the field of radar technology or the common known GPS system. In some situations however, it will not be possible to use EM waves. One of these situations is when considering an underwater scenario. Due to the high electric permittivity, EM waves dampen out fast in water. An alternative is found in acoustic waves, which have more suitable propagation properties for an underwater environment.

Acoustic waves are also used for localization in situations where they travel through air. In that case, it is usually sufficient to assume linear propagation between a source and one or more receivers. By measuring the time between transmitting and receiving, the distance between a source and receiver can be determined. By combining multiple of these distances, the location of the source can be derived.

In an ocean environment, the acoustic localization process becomes more complex. Acoustic waves do not longer only travel directly from the source to a receiver, but are also reflected by the surface and bottom. Whilst this provides a more complex problem, the additional arriving reflections can also be used as an extra source of information when determining the source location.

The sound speed in water is determined by temperature and pressure. This causes the sound speed to differ through the environment, introducing another challenge. When a sound wave travels through an area of changing sound speed, the travelling direction changes. This makes it not possible anymore to assume linear propagation and makes other localization approaches required.

The field of underwater acoustic localization is very broad. Where the first incentives seem to come from the detection of submarines in a warfare setting, the field has expanded itself to many other domains, including experimental studies [20, 33], the tracking of mammals [2, 6, 17, 25–27], localization of automated underwater vehicles [5, 9] and the positioning of sensors in a sensor network [15, 21, 42].



(a) A dolphin is an example of a high-frequency sound source.

(b) Modern submarines are equipped with a hydrophone array.

Figure 1.1: Example of a high-frequency source and an application of a hydrophone array.

The typical way to solve an acoustic localization problem is by using a model-based approach. In a model-based approach, for different candidate locations the data from an acoustic propagation model is compared to the acoustic data that was measured. Which aspects are compared exactly and how they are compared differ per method. In the more classical approach, sound pressure levels are directly compared to determine which source location matches best [3, 8, 11]. These methods are known as matched field processing (MFP) and are known to be most useful for localizing a low-frequency acoustic source using widely spaced hydrophones [35].

More recently it has been shown that also for acoustic sources with higher frequencies model-based localization techniques can be used to locate or track a source [20]. This offers opportunities to for example track marine mammals that produce high-pitched sounds or for tracking UAVs, for which a high spatial resolution is required. Due to the higher frequencies of the source signal, phase shifts occur when reflected by the surface or bottom, which makes MFP not a suitable solution anymore. Instead of that, the general approach is to measure the arrival times of different reflections and compare these for different modeled scenarios [2, 6, 17, 20, 25, 26, 30, 31]. Even though this general principle is clear, work in this field remains. Multiple methods are considered through literature, but these methods still face several challenges when robust and reliable localization is required.

Another consideration when performing model-based localization is the hydrophone configuration. In more classic approaches the hydrophones are typically bottom mounted and widely spaced, as for example in [25]. The advantage of this approach is that the hydrophones will easily remain fixed and, since the distance between receiver positions is fairly large, their measurements are expected to clearly differ. It has been shown that also when using closely-spaced hydrophones in a vertical line-array configuration acoustic localization can successfully be performed [14, 20].Since such a structure is a lot more compact, it becomes easier to place or move than the bottom-mounted receivers. The compact structure also allows an array to be mounted onto a vessel, which is for example done on modern submarines, of which an example is depicted in figure 1.1b.

In this thesis, the case of high-frequency source localization using a model-based approach and a vertical line-array in a shallow ocean environment will be considered. For this case the currently available methods will be studied and based on challenges in these methods improvements will be proposed.

1.1. Research Questions

The goal of this thesis is to develop a robust localization algorithm to localize high-frequency sources using compact vertical line-arrays.

To achieve this goal, the following research questions will be addressed:

1. What model-based localization methods are currently used?
2. What are typical challenges that these methods face?
3. How can the effect of these challenges be reduced?
4. How can the accuracy of an estimated source location be determined?
5. What is the influence of the size and configuration of an array on detection performance?

1.2. Report Outline

In chapter 2, background information will be provided on underwater acoustic propagation, the ocean environment that is considered in this thesis and the acoustic propagation model that is used. In chapter 3, the existing methods for acoustic localization are introduced and a summary of their challenges is given. Based on these methods and their challenges, in chapter 4 a new method is proposed that attempts to reduce the effect of these challenges. To indicate the accuracy of a localization and incorporate environmental parameters in the estimation, methods to perform Bayesian inference are introduced in chapter 5. In chapters 6 and 7, the results of the discussed methods are shown. In chapter 7 also some considerations on the size of the receiver array are discussed.

2
Background

2.1. Underwater Acoustic Propagation

In localization problems typically electromagnetic (EM) waves are used to estimate a location, for example in radar technology or GPS. In an underwater environment these EM waves are dampened out much faster than in air, making it difficult to transmit over a larger distance. This is due to the fact that water has a much higher relative permittivity than air: 1 F/m for air against 72-80 F/m for seawater [41]. This makes EM waves an unsuitable solution for underwater localization.

An alternative is found in acoustic waves. Where in approaches where acoustic waves propagate through air often the assumption of linear propagation suffices, this is not the case when working with acoustic waves in an underwater environment. Two phenomena make underwater acoustic propagation in general complex: reflections through the bottom or surface and variable sound speed.

2.1.1. Reflections

Assume a sound source uniformly emitting acoustic waves in an ocean environment, where the sound speed is constant throughout the ocean. In some directions, the ones that are (close to) horizontal, the acoustic wave will find a clear and open path as far as the ocean spreads. However, in the other directions, the propagation is at some point interrupted by either the ocean's surface or the ocean's bottom.

When this happens, two things occur. A part of the acoustic wave is reflected back into the ocean, whilst another part continues its path into the air or the bottom. There are complete fields of research to study these interactions with for example the ocean bottom [38]. The exact details are out of the scope of this project, but it is important to know that acoustic waves interact with the bottom and surface, where they are partially reflected having an attenuation as result. An impression of the different paths a sound wave can propagate along can be seen in figure 2.1. It can be clearly seen that a relatively shallow ocean basically acts as an acoustic waveguide.

2.1.2. Sound Speed

Another property of the ocean that makes the propagation of sound complex is the varying sound speed. The sound speed in water is influenced by different parameters such as temperature, pressure and salinity, where salinity plays a relatively small role compared to pressure and temperature [1]. In literature attempts have been made to approximate the sound speed in an ocean environment as a function of temperature and depth (pressure), for example in [16].

This difference in sound speed causes acoustic waves to bend. In figure 2.1 ten different paths from a sound source are calculated using the BELLHOP model, which will be further introduced in section 2.2. It can be clearly seen that some of the paths do not propagate linear at all. When this kind of model is used to match to real measured data, as will be done later on in this report, the sound speed has to be taken into account to get realistic modeled data. It has been shown that even small errors in the sound speed can make model based localization difficult or even impossible [40].

To get accurate modeled data, the sound speed is usually measured for different depths at the site where an experimental measurement takes place. Or the temperature and pressure are measured, from which the sound speed is then derived. This results in a sound speed profile, of which one is depicted in figure 2.2. Using this sound speed profile improves the accuracy of the modeled data, but the sound speed is sensitive to change over time. This is due to different factors, for example that the surface layer of water can warm up or cool down due to the weather, or in different layers of the ocean colder/warmer water can flow from other areas. This might cause the measured sound speeds to differ from the sound speeds at the time of the actual measurement, which in turn might cause errors in the data predicted by the model.

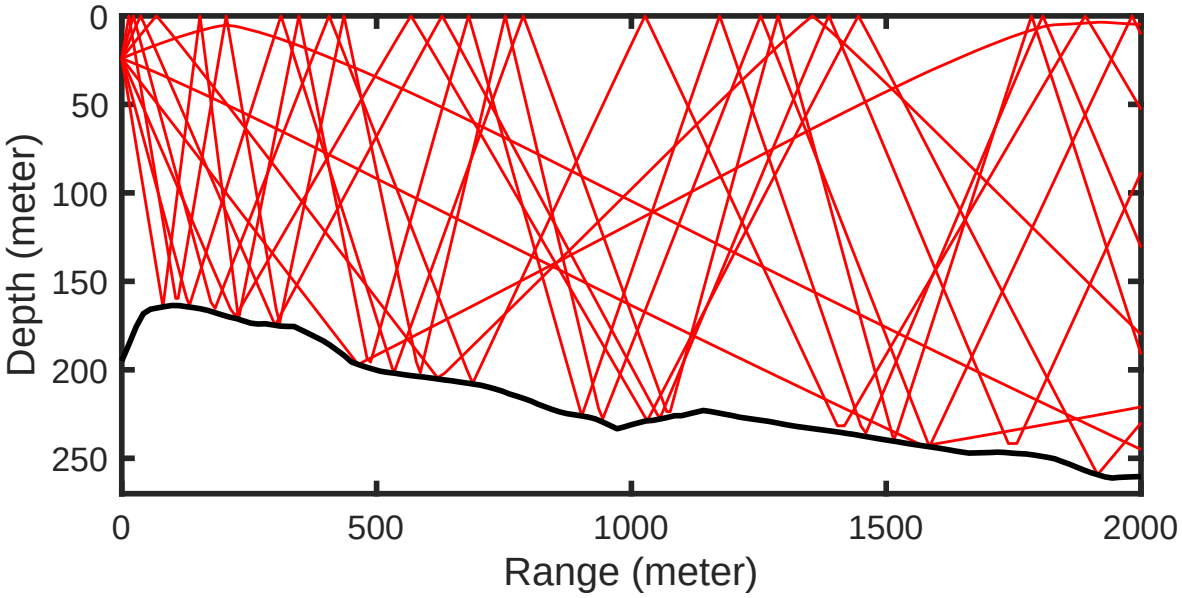


Figure 2.1: Acoustic propagation through an ocean environment using BELLHOP modeling software.

2.2. The BELLHOP Model

For the modeled data in this thesis, the BELLHOP ray-tracing software is used [28]. BELLHOP models the propagation of a number of rays (acoustic propagation paths) through an ocean environment. An example of this is given in figure 2.1. In this figure, 10 rays are emitted from a source at a depth of 60 meters. The paths of these rays through the ocean are then modeled according to Snell's law.

For a given source location, BELLHOP emits typically multiple thousands of rays under uniformly distributed angles. For a given receiver location, or multiple receiver locations, BELLHOP then determines which of these rays reaches this receiver location. The rays that do are referred to as *eigenrays* of a receiver. For each of the eigenrays, the travel time from the source and the attenuation caused by the propagation and reflections are computed. The output given by BELLHOP consists of the the travel time of each of the eigenrays, but also their amplitudes and direction of arrival. An overview of the inputs and outputs of BELLHOP that are considered in this project can be found in figure 2.3.

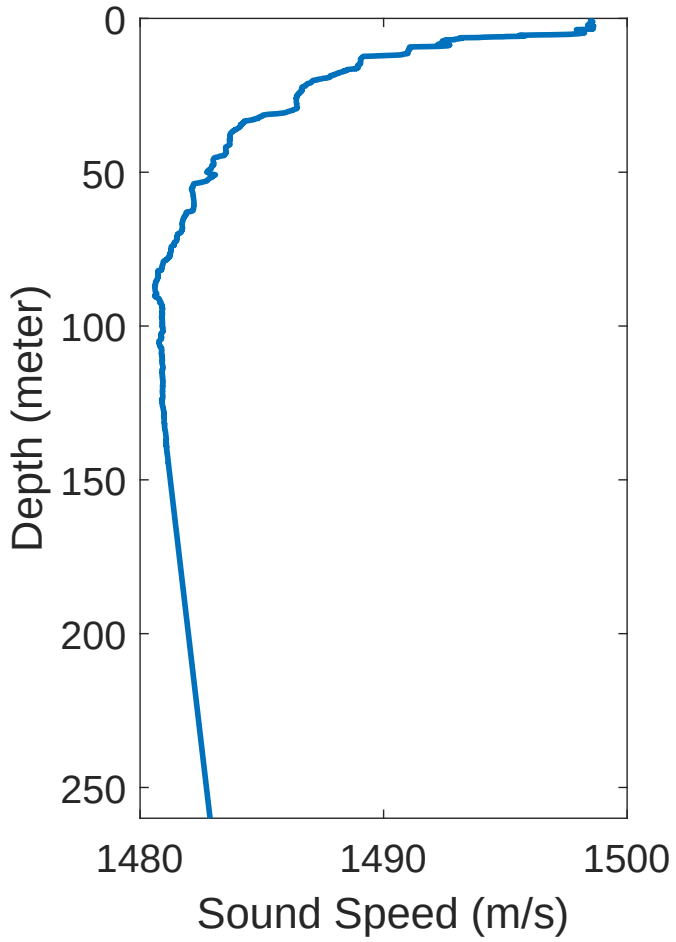


Figure 2.2: The sound speed profile as measured on the site of the UCAC measurements.

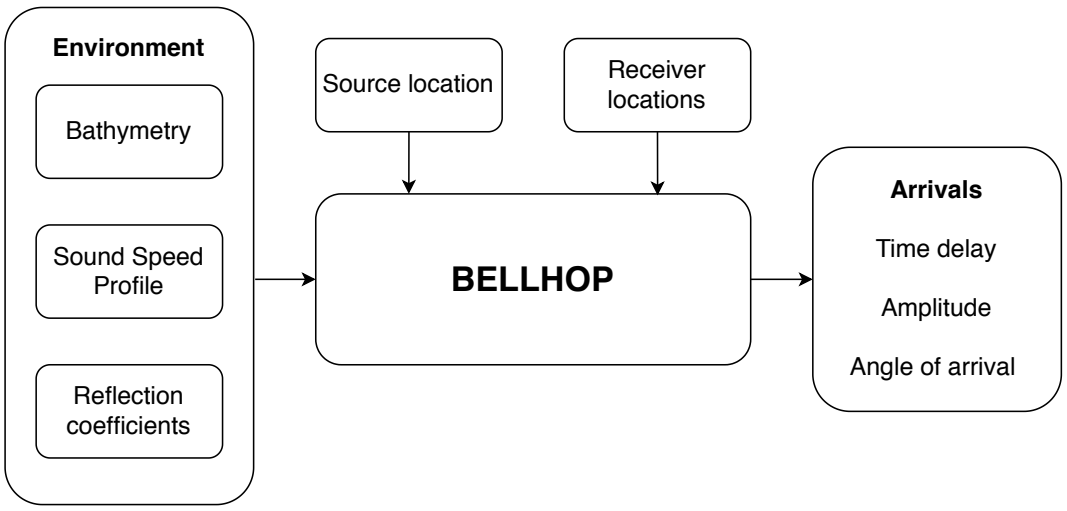


Figure 2.3: An overview of the in- and outputs of the BELLHOP ray-tracing software.

2.3. The UCAC Data set

In this thesis, measurement data from a past project named UCAC will be used to test the methods with a real-world scenario. In the UCAC project, acoustic measurements where made by placing a receiver array at a fixed location and towing an acoustic source away from this location. The result is multiple data sets, containing the acoustic measurements, but also the measured bathymetry at the site and a measured sound speed profile.

In the UCAC project, an array consisting of 128 elements has been used. For this thesis, only the middle 62 elements with a constant spacing of 15 cm will be considered. The measured sound speed profile is shown in figure 2.2 and the bathymetry considered for the UCAC scenario can be found in figure 2.5.

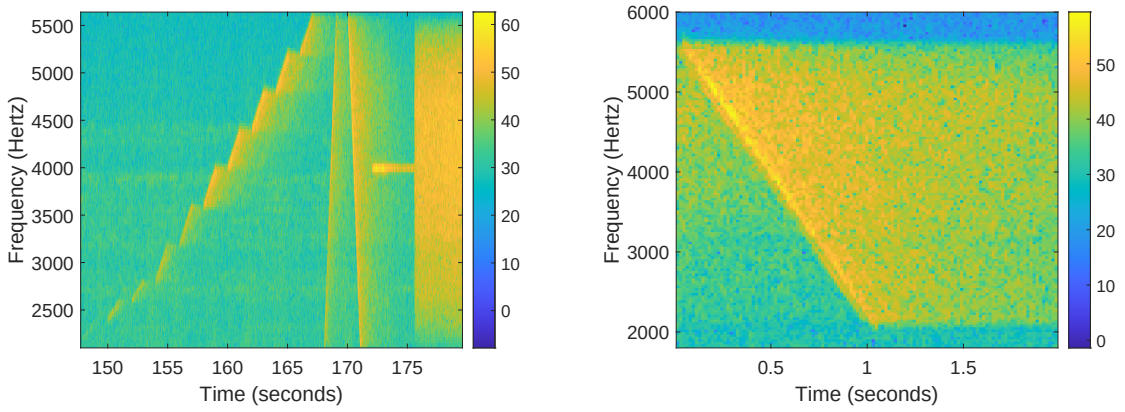
For different ranges, an identical series of signals was emitted. The first 11 signals are LFM chirps, which are sinusoidal signals of which the frequency increases or decreases over time. The complete specifications of the different LFM chirps can be found in table 2.1 and are visualised in figure 2.4a. For this thesis, the localization process will be performed considering the LFM chirp that ranges from 5.6 to 2.1 kHz, since it is one of the signals with the largest bandwidth. The spectrogram of just this chirp is shown in figure 2.4b.

2.4. Considered Scenario

In the major part of this thesis, a set scenario will be considered. The overview of this scenario is shown in figure 2.5. In this figure, the red marker indicates the source and the black marker indicates the position of the receiver array. The blue line indicates the bathymetry. The receiver array consists of the middle 62 elements of the UCAC array, of which the upper element is located on 24.32 meter and the spacing between the elements is 0.15 meter. When nothing else is mentioned, this scenario is used.

Starting frequency (Hz)	Target frequency (Hz)	Duration (s)	Bandwidth (Hz)
2400	2600	1	200
2600	2800	1	200
2800	3200	1	200
3200	3600	1	400
3600	4000	1	400
4000	4400	1	400
4400	4800	1	400
4800	5200	1	400
5200	5600	1	400
2100	5600	1	3500
5600	2100	1	3500

Table 2.1: Different LFM sweeps transmitted by the UCAC source.



(a) Complete UCAC transmission.

(b) LFM Chirp.

Figure 2.4: Spectrogram of a complete UCAC transmission and a selection of the LFM chirp used considered in this report. The color scale indicates amplitude on a linear scale.

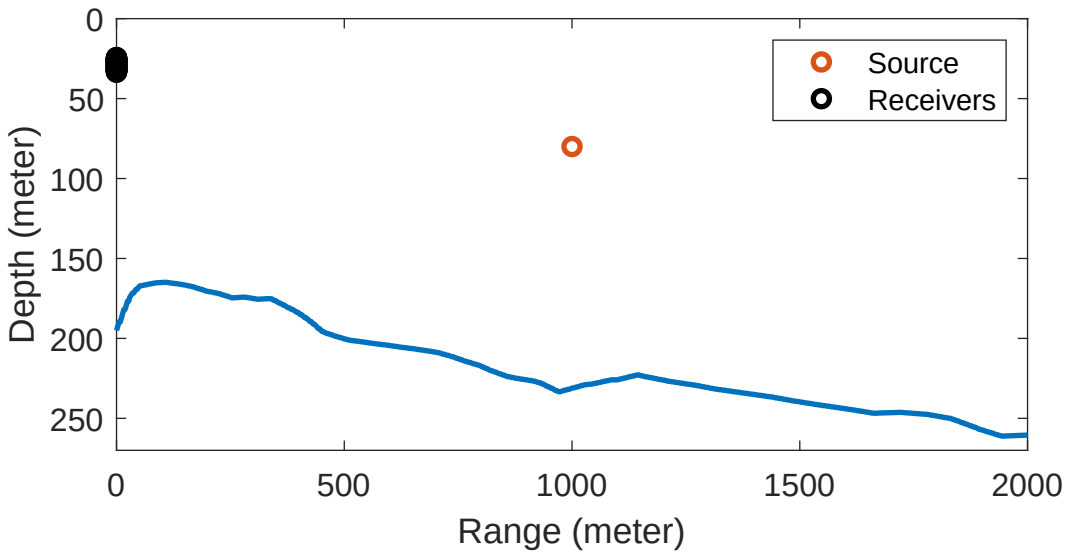


Figure 2.5: The scenario considered in the major part of this thesis. The red marker indicates a possible source location, the black markers the receiver array.

3

Localization Methods

In the past multiple studies have been conducted in the field of underwater acoustic localization. The different methods mostly differ in which quantity of the measured data is compared to a modeled prediction. An overview of the general structure of model-based localization methods is shown in figure 3.1. In section 3.1, first a signal model for the received signals is introduced. Using this signal model, in section 3.2 is described how information as the impulse response and Time Differences of Arrival (TDOAs) can be extracted from the received signals. In sections 3.3 to 3.8 is described how the extracted quantities can be used to perform localization. To conclude, in section 3.10 an overview of the typical challenges that occur when using the presented methods is given.

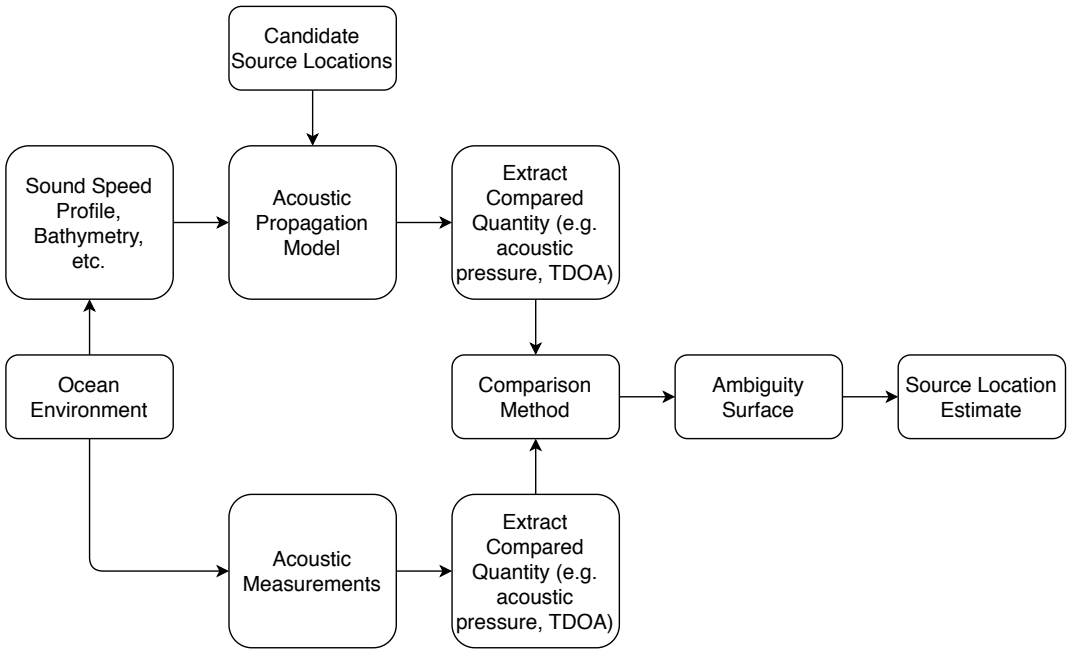


Figure 3.1: An overview of the general steps of model-based localization methods.

3.1. Signal Model

As was described in section 2.1, multiple reflections of the source signal are expected to arrive at the receivers. Consider a discrete source signal $s[n]$ is given, with sample index n. To match with the UCAC data set, the sample rate is chosen to be 15.625 kHz. If then I different reflections arrive at a receiver, with indices i, propagation delay m_i (in samples) and amplitude a_i, the received signal can be described as follows:

$$r[n] = \sum_{i=0}^{I} a_i s[n - m_i] + w[n] \tag{3.1}$$

Where $w[n]$ describes a noise component. If multiple receivers are considered, they will be described with an index k, and K indicating the total amount of receivers:

$$r_k[n] = \sum_{i=0}^{I} a_{k,i} s[n - m_{k,i}] + w_k[n] \tag{3.2}$$

3.2. Finding the impulse response and TDOA's

Multiple methods that will be described later in this chapter require the estimation of the Times of Arrival (TOAs), the TDOAs or the impulse responses from a received signal. In this section, methods to determine these quantities are introduced. Two cases will be discussed: the case where the source signal is known and the case where the source signal is unknown. A situation in which the source signal is known is for example an experimental setting or tracking an UAV. A situation in which the source signal is not (exactly) known is the tracking of marine mammals or when a source still has to be identified.

3.2.1. Known source signal

In case the source signal is known, it can be used to determine the impulse response of a received signal $r[n]$. In an ideal case, this impulse response is as follows:

$$h[n] = \begin{cases} a_i & \text{if } n = m_i \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

If this impulse response where to be known, the arrival times could be easily extracted by checking for nonzero elements. Note that by using the impulse response, the received signal can also be written as:

$$r[n] = h[n] \circledast s[n] + w[n] \tag{3.4}$$

Where $\circledast$ denotes the convolution operator. In a realistic case, the impulse response can not be extracted exactly from a received signal, due to the limited bandwidth of the source signal. However, using a matched filter an approximation of the impulse response can be made. The matched filter is known to be the optimal detection statistic for a signal in white noise [1, 22], since it optimizes the signal-to-noise ratio (SNR). The matched filter $m[n]$ is given by:

$$m[n] = \begin{cases} s[N-1-n] & \text{if } n = 0, 1, .., N-1 \\ 0 & \text{otherwise} \end{cases} \tag{3.5}$$

Applying this matched filter to the receiver signal is identical to using a replica correlator [22], which cross-correlates the receiver signal with a copy of the source signal that has been mirrored in time:

$$y[n] = \sum_{t=-\infty}^{\infty} r[t] s[t - n] \tag{3.6}$$

Which using the property of the cross-correlation that the cross-correlation between $r[n]$ and $s[n]$ given by $R_{rs}[n]$ is equivalent to $R_{sr}[-n]$ can be written as:

$$y[n] = \sum_{t=-\infty}^{\infty} s[t] r[t + n] \tag{3.7}$$

Combining equations 3.1 and 3.7 gives:

$$y[n] = \sum_{t=-\infty}^{\infty} s[t] \left(\sum_{i=0}^{I} a_i s[t+n-m_i] + w[t+n] \right) \tag{3.8}$$

Which after multiplication becomes:

$$y[n] = \sum_{t=-\infty}^{\infty} \sum_{i=0}^{I} a_i s[t] s[t+n-m_i] + s[t] w[t+n] \tag{3.9}$$

Now, by introducing the autocorrelation of the source signal, which is the cross-correlation of the source signal with itself:

$$u[n] = \sum_{t=-\infty}^{\infty} s[t] s[t-n] \tag{3.10}$$

Which due to symmetry is equivalent to:

$$u[n] = \sum_{t=-\infty}^{\infty} s[t] s[t+n] \tag{3.11}$$

And by introducing a new noise component containing the noise in the replica-correlator output $y[n]$ as:

$$w_y[n] = \sum_{t=-\infty}^{\infty} s[t] w[t+n] \tag{3.12}$$

The replica-correlator output can also be written as:

$$y[n] = \sum_{i=0}^{I} a_i * u[n-m_i] + w_c[n] \tag{3.13}$$

This result shows that the output of the replica-correlator consists of a noise component and copies of the autocorrelation $u[n]$, delayed with the same time delays as the different source signal copies in $r[n]$. If a source signal is used that has an autocorrelation with a sharp peak the output of the replica correlator approximates the impulse response:

$$\hat{h}[n] = y[n] \tag{3.14}$$

The LFM chirp that is used as source signal in this thesis has an autocorrelation containing such a sharp peak, as is shown in figure 3.4. To illustrate the procedure as described, a matched filter is applied to a receiver signal from the UCAC data set. The original source signal ($r[n]$) is shown in figure 3.2 and the matched filter output ($y[n]$) in figure 3.3. Note that where no individual reflections can be distinguished in the original signal, multiple peaks can be distinguished in the matched filtered signal.

From the impulse response approximation $\hat{h}[n]$, the arrival times can be estimated. Where in the ideal case just the nonzero samples could be detected, now also a noise component $w_y[n]$ and the side-lobes of the auto-correlations $u[n]$ are present. For this case, the arrival times are estimated by finding the local maxima that exceed a detection threshold. Let $\hat{\boldsymbol{M}}$ denote the set of all arrival time estimates $\hat{m}_i$ and $\boldsymbol{N}$ the set of all sample indices n, then:

$$\hat{\boldsymbol{M}} \in \{\hat{m}_i \mid \hat{m}_i \in \boldsymbol{N},\ y[\hat{m}_i - 1] \leq y[\hat{m}_i]\ ,\ y[\hat{m}_i] \geq y[\hat{m}_i + 1] \text{ and } y[\hat{m}_i] \geq \lambda\} \tag{3.15}$$

Where λ is a constant threshold. Now that the arrival times are estimated, the TDOAs can be computed. Two cases are distinguished: either the TDOAs can be computed with respect to a specific arrival time (e.g. the first or the strongest peak):

$$TDOA_i = \hat{m}_i - \min_i \hat{m}_i \quad \forall \hat{m}_i \in \hat{\boldsymbol{M}} \tag{3.16}$$

$$TDOA_i = \hat{m}_i - \arg\max_{\hat{m}_i} y[\hat{m}_i] \quad \forall \hat{m}_i \in \hat{\boldsymbol{M}} \tag{3.17}$$

Or the TDOAs can be computed by taking the difference between each possible combination of arrival times:

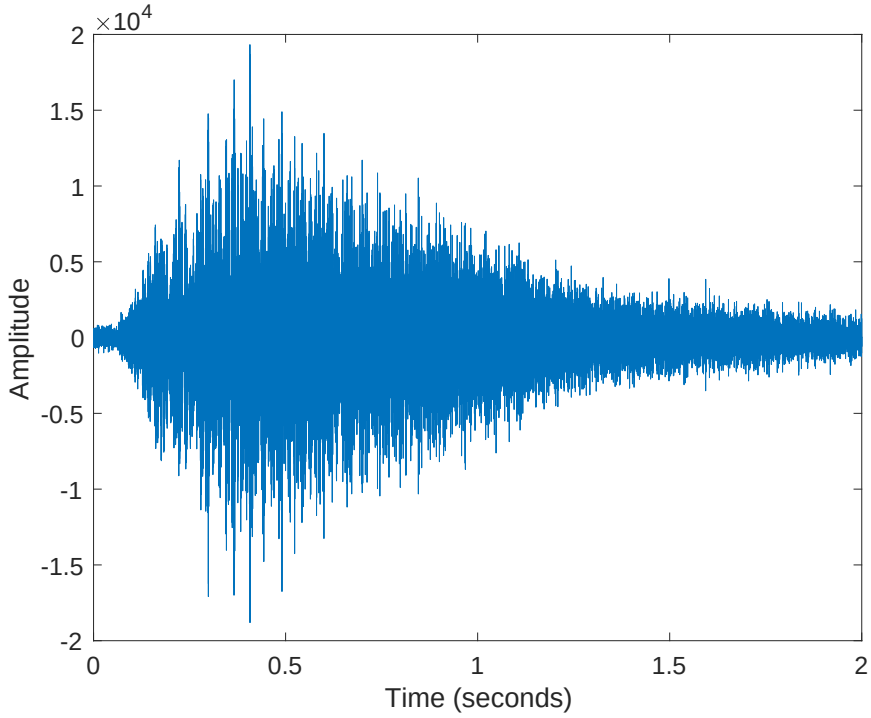


Figure 3.2: Hydrophone recording from the UCAC data set, consisting of multiple arriving LFM chirps and background noise.

$$TDOA_{i,j} = \hat{m}_i - \hat{m}_j \quad \forall \hat{m}_i, \hat{m}_j \in \hat{\boldsymbol{M}} \tag{3.18}$$

The drawback of computing TDOAs as in equations 3.16 and 3.17 is that all TDOAs depend on the correctness of one arrival time. If the first arrival time or the arrival time with the largest amplitude are estimated wrong, immediately all TDOAs will be wrong. This makes computing TDOAs as in equation 3.18 a safer option, as then only a subset of the TDOAs will be wrong. This does however result in $|\hat{\boldsymbol{M}}|^2$ instead of $|\hat{\boldsymbol{M}}|$ TDOAs, which might be more difficult to use in further processing steps.

TDOAs can also be computed between a pair of receivers. Let the estimated arrival times at receiver k be denoted as $\hat{m}_{k,i}$ with index i and at receiver p as $\hat{m}_{p,j}$ with index j. Then the TDOAs between receiver k and p are defined as:

$$TDOA_{i,j} = \hat{m}_{k,i} - \hat{m}_{p,j} \quad \forall i, j \tag{3.19}$$

Which of these TDOA definitions is most suitable to be used, depends on the method and the environment in which localization is performed.

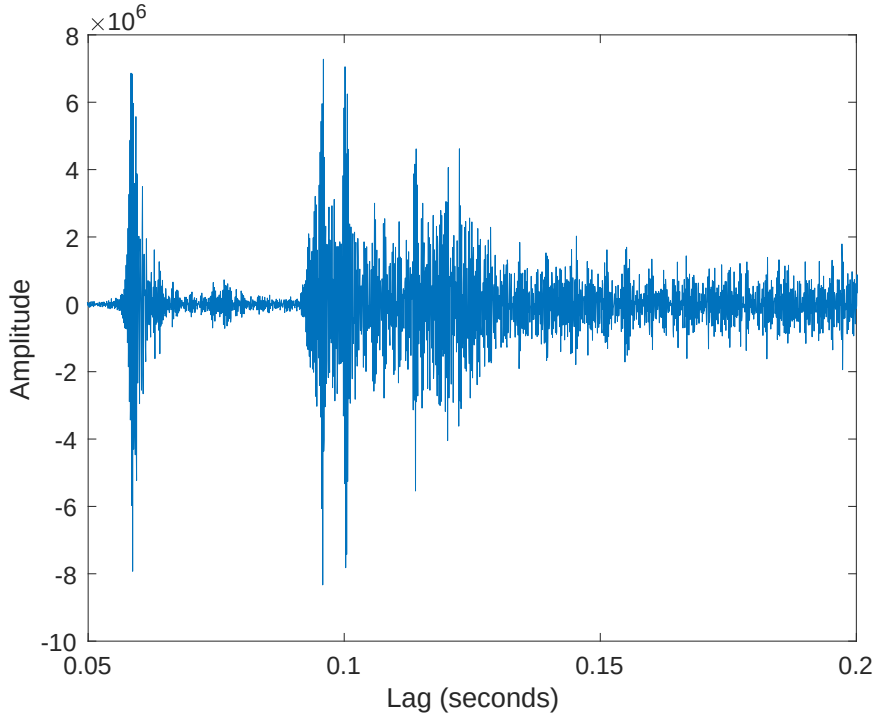


Figure 3.3: The data from figure 3.2 after applying a matched filter with the known LFM source signal.

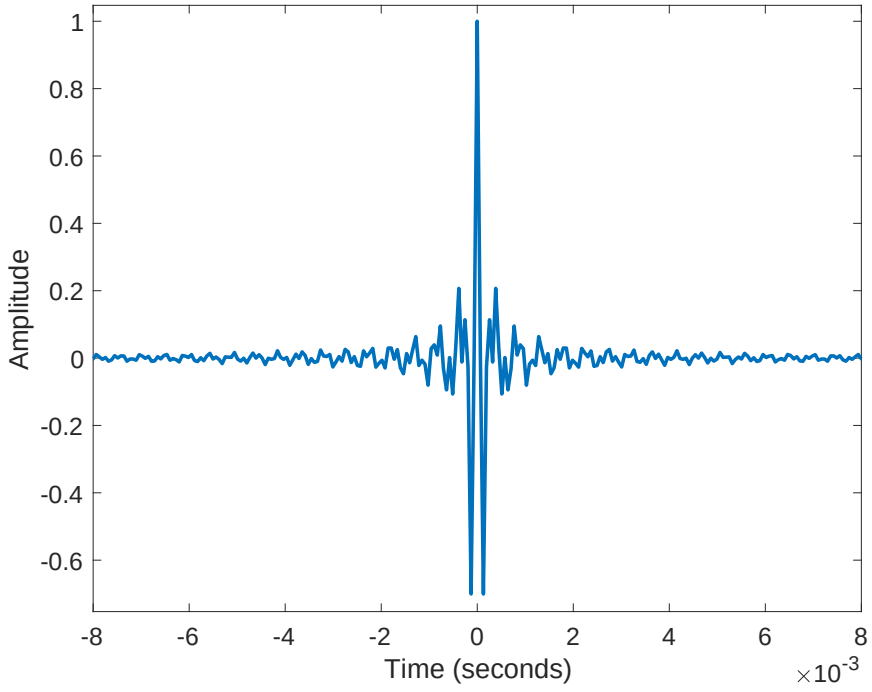


Figure 3.4: Autocorrelation from the LFM chirp as described in section 2.3.

3.2.2. Unknown source signal

In practical applications the source signal that is required for matched filtering is often not known. If that is the case, the matched filter as described in the previous section can not be used. However, with other techniques the impulse response and TDOAs can still be extracted, as will be shown in this section.

Let $H(\omega)$ denote the channel transfer function as in equation 3.3, but now in the frequency domain, and let $S(\omega)$ be the frequency spectrum of the source signal $s[n]$. The frequency spectrum of the received signal $R(\omega)$ is then given as:

$$R(\omega) = H(\omega)S(\omega) \tag{3.20}$$

Then, if the autocorrelation $a[n]$ is computed from the received signal:

$$a[n] = \sum_{t=-\infty}^{\infty} r[t]r[t-n] \tag{3.21}$$

This gives in the frequency spectrum:

$$A(\omega) = |H(\omega)|^2|S(\omega)|^2 \tag{3.22}$$

Where, $|H(\omega)|^2$ can be seen as the frequency spectrum of the autocorrelation of the impulse response and $|S(\omega)|^2$ can be seen as the frequency spectrum of the autocorrelation of the source signal. By using a method as the SCOT filter [4, 30], which is a pre-whitening filter, the effects of the source signal can be filtered, leaving the autocorrelation of the impulse response. As acoustic propagation models can predict this impulse response, and thus also the autocorrelation of the impulse response, this information suffices to compare them for different candidate locations, as is demonstrated in [30].

To extract the TDOAs, conversion to the frequency domain is not necessary. If the TDOAs between the reflections arriving at two receivers are desired, the cross-correlation between the receiver signals can be used. For receiver signals $r_k[n]$ and $r_p[n]$ the cross-correlation is given by:

$$c[n] = \sum_{t=-\infty}^{\infty} r_k[t]r_p[t-n] \tag{3.23}$$

This procedure is similar to what happens in the matched filter, except now copies of the sources autocorrelation occur at the arrival time differences. If the time delays of the different reflections are given by $m_{k,i}$ and $m_{p,j}$, in a similar way as for the matched filter it can be shown that the cross-correlation can be written as:

$$c[n] = \sum_{i=1}^{I}\sum_{j=1}^{J} u[n - m_{k,i} - m_{p,j}] + w_a[n] \tag{3.24}$$

Where $w_a[n]$ describes a noise component. Note that as in the matched filter a noiseless copy of the source signal was used for correlation, the SNR resulting from this cross-correlation will be worse. More noise will be present, since now not only the noise from the receiver correlates with the source signal, but the noise in both of the receiver signals correlates with the different source signal copies in the other receiver signal and the noise will also correlate with itself.

The TDOAs can be directly extracted from signal $c[n]$ by detecting the local maxima, in the same way as was described for the detection of arrival times in equation 3.15. Note that the resulting TDOAs align with the TDOAs as specified for the known source signal in equation 3.19. To illustrate the cross-correlation procedure, two receiver signals from the UCAC data set and their cross-correlation are shown in figure 3.5. Note that clear peaks are visible, but for some peaks it is difficult to say if they indicate TDOAs or are part of the noise.

If the TDOAs between the different reflections at a single receiver are desired, the same procedure as in equations 3.23 and 3.24 can be used, but now by correlating the receiver signal with itself. Note that for this case the resulting TDOAs align with the TDOAs as specified for the known source signal in equation 3.18.

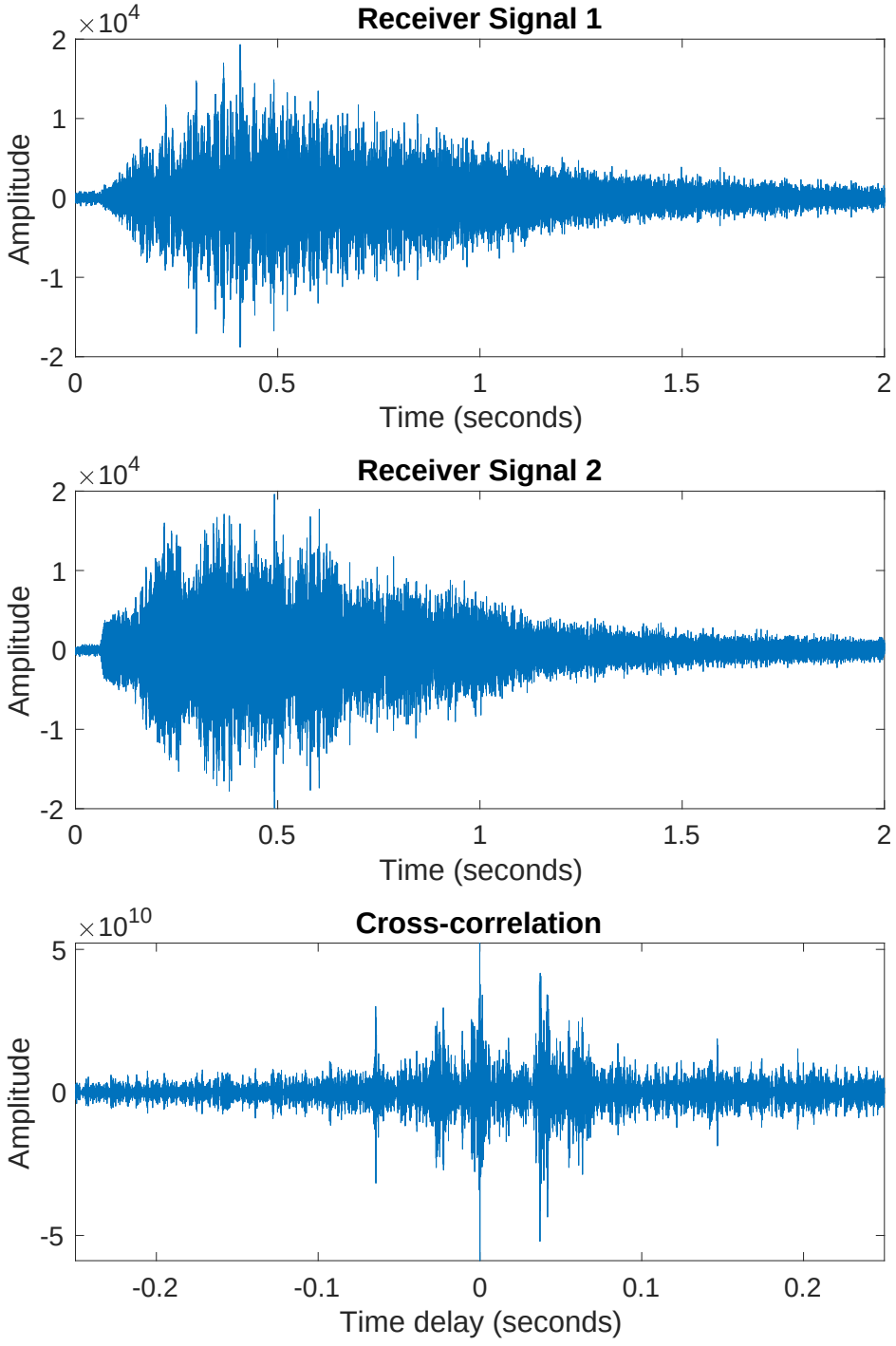


Figure 3.5: Cross-correlating the measurement of two of the receivers from the UCAC data set to find their TDOAs.

Besides the lower SNR than when a matched filter is applied, another drawback of these methods is the possible presence of interfering sources. While the different arriving copies of the source signal cause a peak in the correlation, also interfering sources might create such a peak which might lead to the extraction of a wrong TDOA.

3.3. Matched Field Processing

Matched Field Processing (MFP) is without doubt the most used and studied acoustic localization method. It is for example used in [11], [8] and [10], summarized by [3] and even an event to compare the different approaches has taken place [29].

In MFP processing, the measured receiver data is directly compared to a receiver signal predicted by an acoustic propagation model. This happens coherently, which means that the phase of the received signal is also measured and modeled.

Because the measured samples for these method are complex values, the previously defined receiver signal model does not suffice. Let a complex vector $\boldsymbol{r}_{measured}$ denote the complex measurements stacked for multiple hydrophones and let a complex vector $\boldsymbol{r}_{modeled}(r,d)$ denote the complex measurements as predicted by a model for a given range r and depth d, both at the same time instance. A score can then for example be computed by one of the following functions:

- The Bartlett (classical) Processor

$$P(r,d) = |\boldsymbol{r}^*_{measured}\boldsymbol{r}_{modeled}(r,d)|^2 \tag{3.25}$$

- The Capon Processor (minimum variance / maximum likelihood processor)

$$P(r,d) = \frac{1}{\boldsymbol{r}^*_{modeled}(r,d)\left(\boldsymbol{r}_{measured}\boldsymbol{r}^*_{measured}\right)^{-1}\boldsymbol{r}_{modeled}(r,d)} \tag{3.26}$$

Where these scores can be computed for multiple time instances and can then for example be averaged or summed.

When a high-frequency source is used, coherent methods like MFP are not suitable. In contrast to low-frequency sources, phase shifts will occur when interaction with the ocean surface or bottom takes place, due to the shorter wavelength. For this reason, MFP will not be further considered throughout this report.

3.4. Direct path TDOA

In section 3.2, it has been discussed how TDOAs can be extracted from measurement data. Using for example the BELLHOP model, these TDOAs can also be predicted for a series of candidate locations. In the direct-path TDOA approach, only the arrival time of the source signal that arrives via the direct path, thus without interacting with the surface or the bottom, is taken into account. The TDOAs are then computed between one or more pairs of receivers, as in equation 3.19 with only a single $\hat{m}_{k,i}$ and $\hat{m}_{p,j}$.

In some cases, the direct-path arrivals can be clearly distinguished by taking the first arrival time or the one that arrives with the largest amplitude. In many cases it is more challenging to distinguish this direct path, as due to the complex acoustic propagation that was discussed in section 2.1, it is also possible that the direct path is longer or has a higher propagation loss than paths that interact with the surface or bottom. When this is the case, it will be challenging to effectively implement this method. In literature, methods are sought that resolve this problem, as for example in [12].

Different approaches can be taken to match the direct-path TDOAs from the measurements with the TDOAs from the model. For example, a match can be scored as 1 if the difference between the real and modeled TDOA is within a certain range and 0 when outside this range. A more delicate approach, that assumes the errors in the TDOAs to be Gaussian distributed, is given by [26]. The score function used in this paper is given by:

$$\varphi_{TDOA}(r,d) = \prod_{k\in\boldsymbol{R},p\in\boldsymbol{R}} \exp\left[\frac{-1}{2\sigma^2}\left(TDOA_{meas,k,p} - TDOA_{mod,k,p}(r,d)\right)^2\right] \tag{3.27}$$

Where φ_{TDOA} indicates the resulting match, r and d are the range and depth of the candidate location, σ indicates the standard deviation, $TDOA_{meas,k,p}$ is the measured TDOA between receivers k and p and $TDOA_{mod,k,p}(r,d)$ is the modeled TDOA between these receivers. The set $\boldsymbol{R}$ contains all considered receivers.

In [2], this method is used to track sperm wales, using measurements from 7 widely spaced (several km) hydrophones. In [6], the direct path TDOA method is used to track bowhead whales, using 5 widely spaced (roughly 1 km) hydrophones. In [17], also whales are tracked using 5 widely spaced (around 3 km) hydrophones.

That this method is typically used for widely-spaced and bottom-mounted hydrophones is no coincidence. In this case, the direct path can usually be distinguished as the first arriving signal. Also, when two hydrophones are spaced closely together, the direct path arrival times will typically be also close together, leading to small TDOAs which makes them more sensitive to errors and noise.

3.5. Multipath TDOA

In the previous section, only the direct-path arrivals where considered. By doing this, the information captured in the other arrivals remains unused. In the multipath TDOA method, all detected arrivals are considered when determining how well a measurement matches with a candidate location. In literature, it has been shown that the combination of the different arrivals on a single hydrophone can, under specific circumstances, already be enough information to estimate a source location [31].

In this method there are different approaches that can be considered. The TDOAs of the arrivals at each individual receiver can be used, as in equation 3.18, or the TDOAs can be determined between one or more pairs of receivers, as in equation 3.19. A combination of the two is also possible.

When multiple paths are considered, since multiple TDOAs are detected for each receiver or a receiver pair, it might not directly be clear which measured TDOA belongs to which modeled TDOA in the matching process. This issue might introduce extra ambiguity, since there might be a candidate source location other than the real source location that has similar TDOAs.

Several approaches can be considered to assign a score to a candidate location. For example the following equation can be used:

$$\varphi_{TDOA}(r,d) = \prod_{i,j} \max_{q,v} \exp\left[\frac{-1}{2\sigma^2}\left(TDOA_{meas,i,j} - TDOA_{mod,q,v}(r,d)\right)^2\right] \tag{3.28}$$

Where each measured TDOA (between arrivals i and j) is matched with the modeled TDOA (between arrivals q and v) that has the smallest difference.

A problem that might follow from this approach, is when multiple similar measured TDOAs are matched with the same modeled TDOA, or the other way around. This might lead to an unintentional good scores, since multiple matches are made when there actually was a single one. This effect can be prevented by allowing each modeled TDOA to be matched only once, however this might in turn also lead to errors as if a false match is performed, the modeled TDOA will not be available anymore for the actual match.

Since the success of this method relies on multipath arrivals, it is applicable in ocean environments where multiple paths from the source to the receiver occur. This can either be caused by reflections in a shallow ocean, as in the environment considered in this thesis, or by the bending of acoustic waves due to the differing sound speed, in a deeper ocean environment.

This method has already known different successful implementations. For example, using a single hydrophone in [31]. Also, Nosal has used this method to track marine mammals [25, 26]. In [30], the auto-correlation between receivers is used to track an acoustic source.

3.6. Matching Impulse Responses

The previous two methods, where based on matching modeled and measured TDOAs. Several reasons where named why this matching might be challenging. Another method that avoids this individual matching, is to match the whole impulse response at once. This can be done by cross-correlating the measured impulse responses with modeled ones for different candidate locations.

Given this measured impulse response, generating a signal with the impulse response predicted by the model and convolving this with the autocorrelation of the source signal:

$$\hat{h}_{mod,r,d}[n] = h[n] \circledast u[n] \tag{3.29}$$

Gives a modeled version of the approximated impulse response for a source range r and depth d. From that, a score can be calculated for example by:

$$\varphi(r,d) = \max \text{xcorr}\left(\hat{h}_{mod,r,d}[n], \hat{h}[n]\right) \tag{3.30}$$

For a single receiver. For multiple receivers this score could be extended to:

$$\varphi(r,d) = \prod_{k=1}^{K} \max \text{xcorr}\left(\hat{h}_{mod,k,r,d}[n], \hat{h}_k[n]\right) \tag{3.31}$$

An advantage of this method is that no separate TDOA detection or association has to take place. A disadvantage is that this method does not allow for mismatch between modeled and measured arrival times. If some measured arrivals have a slight negative bias in time and others a slight positive bias, the overall impulse response will not correlate well.

This method has been used to track a high-frequency source using a compact hydrophone array by [20].

3.7. Bayesian Approach

In the approaches mentioned in this chapter, the parameters used as the inputs for the model are considered to be deterministic. In many cases it is more realistic to assume they are random variables, which is exactly what is done in the Bayesian approach. For example, the sound speed is measured at some point, but changes over time. It is expected that it has slightly changed at the moment the acoustic measurement actually takes place. But uncertainty is also present in many other variables, for example the bathymetry, reflection coefficients, the receiver positions, all of which contribute to the results of the model. Some of these parameters might even be unknown for a certain measurement set.

In literature [7, 13], promising results have been found by optimizing the match between measured and modeled data by considering uncertainty in the input parameters. One could say the match between model and data is not only optimized over the candidate source locations, but also over the different environmental parameters. Typically, an *a posteriori* probability density function of the range and depth are generated, indicating the probability of different ranges and depths based on the observed data and prior knowledge. This also gives an indication of the accuracy of a source localization.

This approach can also be used to estimate these unknown or uncertain environmental parameters. An example of this can be found in figures 3.6 and 3.7 (taken from [13]), where several environmental parameters are estimated besides the source range and depth.

A more extensive discussion of this process and related methods can be found in chapter 5 on Bayesian inference.

3.8. Other Methods

In the previous sections, the most common methods to perform model-based acoustic localization are discussed. However, it should be mentioned that this does not span the whole field of model-based localization. For example in [18], localization is performed by formulating an optimization problem, in [24], machine

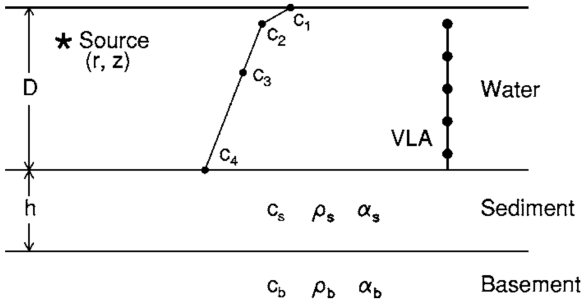


Figure 3.6: Situation sketch, indicating the parameters estimated in figure 3.7. Figure taken from [13].

learning is applied and in [19] an approach originating from information theory is used to find the source location.

3.9. Overview

Several methods have been discussed in the preceding part of this chapter. For the scenario considered in this thesis, the multipath TDOA approach seems the most promising. In matched field processing, coherent detection takes place, which is not suitable when considering high-frequency sources [35]. The direct-path TDOA is, as was mentioned, mostly appropriate for widely spaced and bottom mounted receivers. For the impulse response method, all the arrival time information is contained in a the estimated impulse response. This does not allow to allow for any error in arrival times, as can be allowed with the TDOA based methods.

To conclude this part, in table 3.1 a summary of the previously mentioned methods is given.

Method	**Coherent**	**Source signal required**	**Relevant literature**
Matched Field Processing	Yes	Yes	[3, 8, 11, 29, 35, 40]
Direct path TDOA	No	Yes	[2, 6, 17, 25, 26]
Multipath TDOA	No	Yes	[25, 26, 31]
TDOA using cross- and autocorrelation	No	No	[30]
Matching impulse response	No	Both possible	[20]
Bayesian approach	Both	Both possible	[7, 13]

Table 3.1: Overview of model-based localization methods

3.10. Challenges

In section 3.2 approaches to determine TDOAs and/or impulse responses from measurement data where introduced. On some occasions, these methods will sufficiently perform, however this will not always be the case. When multiple arrivals are present in a short amount of time, it becomes difficult to determine which peaks in the correlated signal belong to actual arrivals. In a realistic scenario, this issue becomes even worse, due to noise being present. The noise will also correlate and cause the peak detection to become even more challenging. One or more incorrect TDOAs might easily be detected, which can lead to significant errors in the localization process.

Another challenge, which was already mentioned with the multipath TDOA method, is the association of detected TDOAs. When matching TDOAs, it might not be completely clear which by the model predicted TDOA corresponds with the TDOA extracted from the measured data. This is especially the case when TDOAs are close. When in the matching process two TDOAs are mis-associated, this might in an equation as 3.28, lead to an unintentional bad score. The other way around can also happen, when a candidate location has TDOAs that accidentally match with a measured TDOA, which in its turn might lead to an unintentional good score.

A known issue in model-based localization is when the environment considered in the model does not completely align with the environment at the moment the measurements where performed. This phenomena is

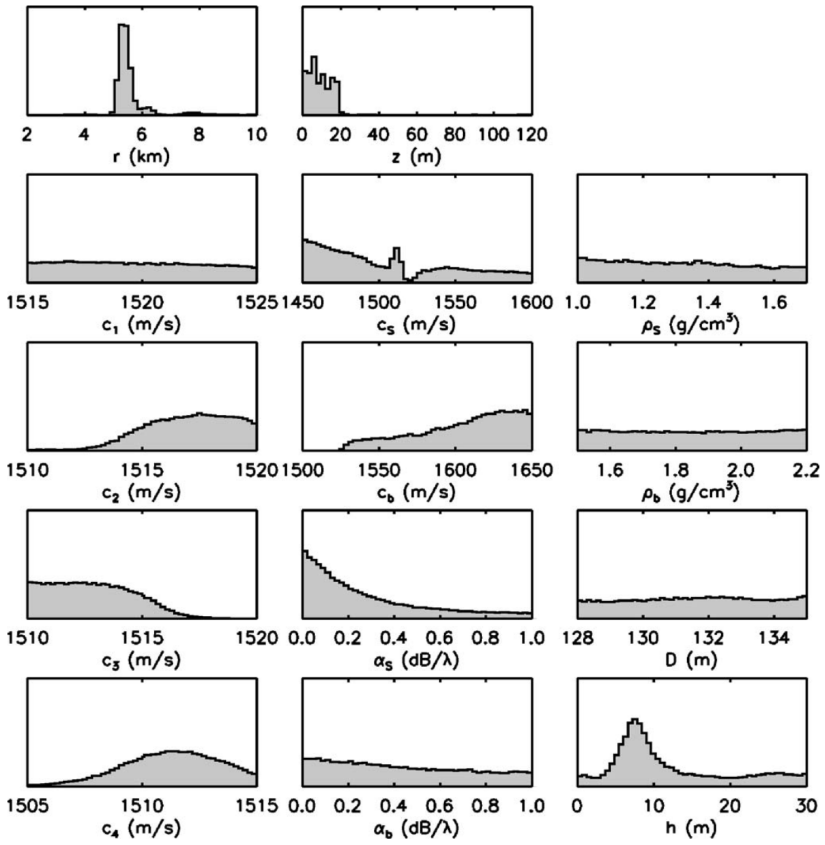


Figure 3.7: Example of marginal probability density functions that result from using a Bayesian approach. Figure taken from [13].

also called environmental mismatch. The sound-speed profile might have slightly changed between the moment it was measured and the acoustic measurement took place, there could be an error or uncertainty in the bathymetry measurements or anything similar. This environmental mismatch can lead to severe estimation errors [40].

When acoustic localization is performed, it is in general difficult to say something about the reliability of an estimated location. Consider performing maximum-likelihood estimation with the likelihood function in equation 3.27 or 3.28. The outcome will then be the range and depth with the highest likelihood score. All other information, for example if there are a lot of other relatively high scores, is lost.

Summarizing, the following challenges are found:

- Accurate TDOA estimation by peak detection from a matched filter, cross- or autocorrelation.
- Accurate TDOA association preventing unintended good or bad scores.
- Faulty location estimates due to the influence of environmental mismatch.
- Uncertainty about the reliability of / confidence in a location estimate.

In chapter 4, a method will be introduced that reduces the effects of the first two mentioned challenges. In chapter 5 an attempt will be made to overcome the latter two challenges.

3.11. Additional Considerations

Whilst not all methods considered in this chapter can always determine an accurate source location from a single measurement, even methods with an uncertain or multiple solutions might be useful. Two examples

will be given here. Both techniques are inspired by [20], however other sources use similar approaches.

In all mentioned methods, in some way the receiver measurements are correlated to either find the arrival times, the impulse response or TDOAs. In each of these approaches, also noise will correlate, which might result in errors. In the case considered by [20], not a single LFM chirp, but 40 are transmitted in 1 second. The different measurements are then matched filtered and averaged. Due to this averaging, the noise is significantly reduced. A similar approach is obviously possible in other configurations.

Another aspect to consider when working with a real-world implementation is that almost no realistic scenario requires a localization at a single time instance. For example, when locating a marine mammal, it is expected to be detectable for multiple successive moments. This, in combination with probably knowing the expected speed of the mammal that is located, makes it possible to track the animal over time. An example where this is done can be seen in figure 3.8. For a longer time an experimental acoustic source is towed away from a receiver array. In the figure, a range estimation can be found for a fixed depth and over time. It can be clearly seen that at some time instances, a range could not be clearly determined due to ambiguity. However, looking at the bigger picture over time, a clear trace can be easily distinguished. This shows that even when a certain method shows ambiguity at other source locations, it can still give useful results, for example when tracking a source. A lot of work is done in the field of tracking one or multiple source, for example by [34].

In [25], another example of such a tracking method is given for the case of tracking multiple marine mammals. In this method, for a given time instance a grid-search is performed over range and depth to find the positions with a high score, where these scores are computed based on measured TDOAs. Over multiple time instances, these found positions are then clustered based on identifying sets of locations that could form a realistic path over time. If a path can be distinguished for a given number of successive localizations, it is assumed that the estimated location peaks in this cluster are correct. If the peaks do not fill this criterion, they are discarded.

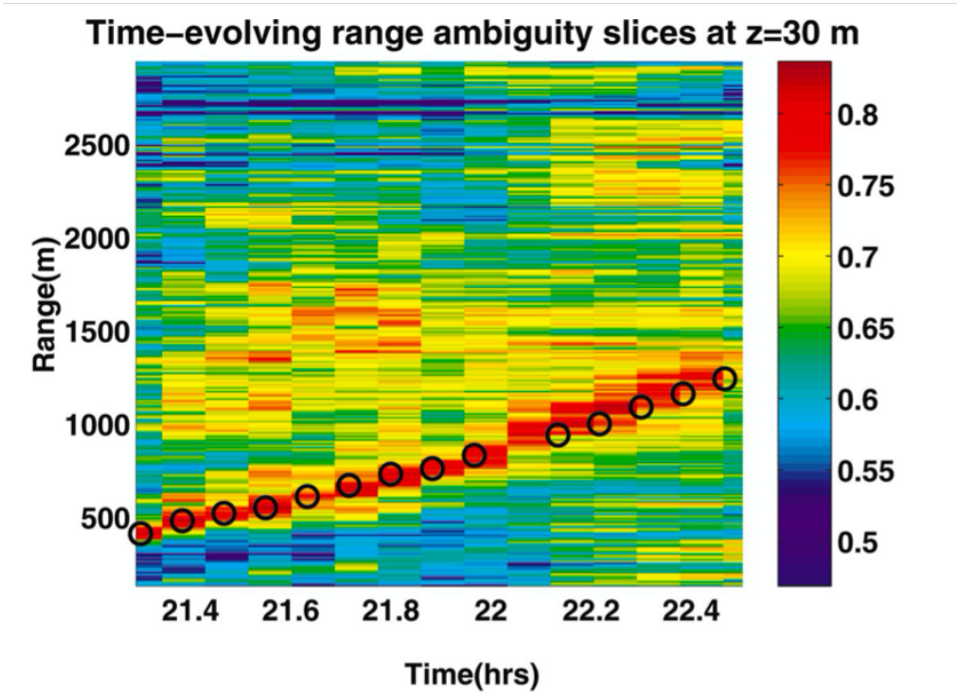


Figure 3.8: Using location estimates over time to track an acoustic source. Figure taken from [20].

4

Proposed Method for Localization Using Vertical Line Arrays

4.1. Overview

As was presented chapter 3, most existing localization methods rely on matching of the arrival times from the measured and modeled data sets, for example via TDOAs or the impulse response. The functioning of these methods is limited by (1) wrong detection of arrival times due to close arrivals and/or noise and (2) incorrectly relating a modeled TDOA and measured TDOA. With the method presented in this section, the effects of these limitations are reduced.

In the proposed method, the line-array structure of the receivers considered in this thesis is exploited by beamforming the received data. Based on the beamformed data, the Directions-Of-Arrival (DOAs) of the different reflections are determined. Using the beamformed data at the detected angles, the TDOAs between the different reflections can be computed.

The advantage of this method is that the different reflections can not only be associated based on their TDOA, but also on their DOA. Adding this other dimension decreases the chance of misassociation. Another advantage of beamforming is that the beamformed signal at a detected DOA approximately only contains the signal of this specific reflection in noise. This means that when a cross-correlation is performed, the chance of missing peaks because of their overlap or detecting a noise peak is reduced.

In the remainder of this chapter, the exact steps will be discussed and some intermediate results of this method will be shown.

4.2. Finding measured DOAs and TDOAs

As a first step, the DOAs of the measured arrivals are estimated. To achieve this, a delay-and-sum beamformer is applied to the measured data. Let $\mathbf{\Theta}$ denote the discrete set of angles for which the beamformer is applied:

$$\mathbf{\Theta} = \{-90, -89, ..., 89, 90\} \tag{4.1}$$

And let the beamformer output $b[n,\theta]$ be computed as:

$$b[n,\theta] = \sum_{k=1}^{K} r_k \left[n - \left\lfloor \frac{d_k * sin(\theta) * \text{sample rate}}{c} \right\rceil \right] \qquad \forall\, \theta \in \mathbf{\Theta} \tag{4.2}$$

Where d_k is the distance between receiver r_k and receiver r_1 and c is the average sound speed. $\lfloor x \rceil$ is used to indicate that x is rounded tot the nearest integer.

To illustrate the effect of applying this beamformer, receiver signals $r_k[n]$ where generated using equation 3.2 in combination with amplitudes and propagation delays computed by BELLHOP. As source signal, the

LFM chirp described in section 2.3 was used. This was done for a total of 62 receivers with intermediate distances of 15cm, to match with the configuration of the UCAC array size. These receiver signals where then run through the beamformer in equation 4.2. The result can be seen in figure 4.1.

To detect the directions θ from which a reflection arrives, the total energy for all directions is computed:

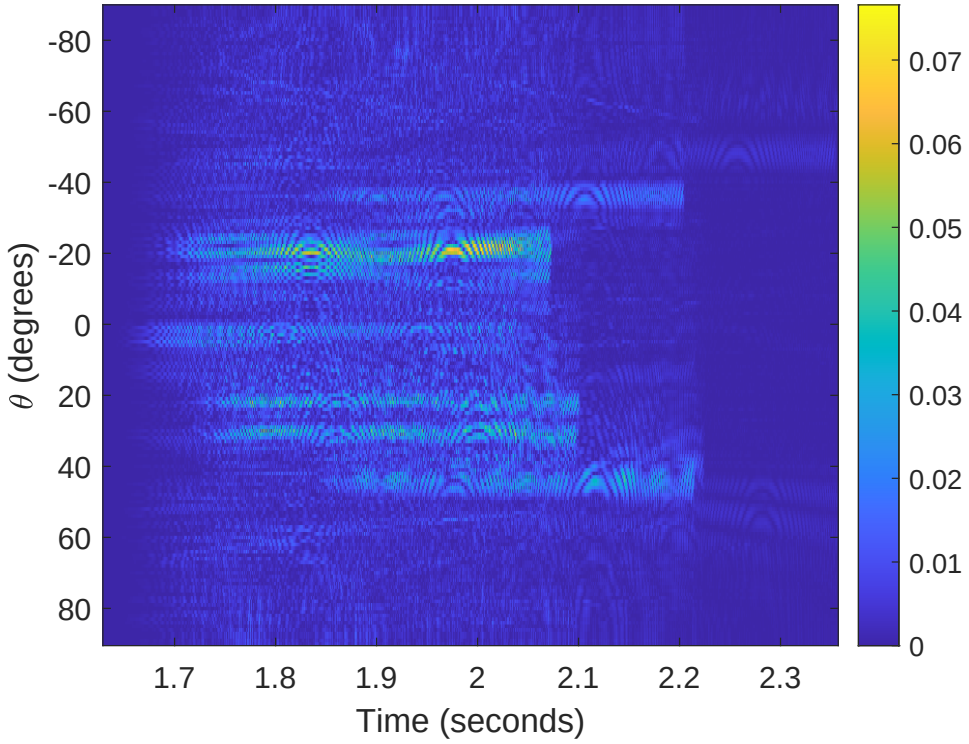


Figure 4.1: Example of the resulting signals after beamforming. This figure was generated using modeled data.

$$E[\theta] = \sum_{n} b[n,\theta]^2 \qquad \forall\, \theta \in \mathbf{\Theta} \tag{4.3}$$

From the resulting energy per θ, the local maxima that exceed a predefined threshold ϕ are detected as measured DOAs. Consider the set of estimated DOAs $\mathbf{\Theta}_{meas}$:

$$\mathbf{\Theta}_{meas} \in \{\hat{\theta}_{meas,i} \mid \hat{\theta}_{meas,i} \in \mathbf{\Theta} \wedge E[\hat{\theta}_{meas,i}-1] \geq E[\hat{\theta}_{meas,i}] \wedge E[\hat{\theta}_{meas,i}+1] \geq E[\hat{\theta}_{meas,i}] \wedge E[\hat{\theta}_{meas,i}] \geq \phi\} \tag{4.4}$$

To illustrate this process the energy $E[\theta]$ is plotted against θ, for receiver data generated from BELLHOP as described before and receiver data from a UCAC measurement in figures 4.2 and 4.3 respectively. It can be clearly seen that more noise is present in the case with UCAC data and peaks are clearer distinguishable for the receiver data generated using BELLHOP.

At this point, it is assumed that each estimated arrival angle $\theta_{meas,i}$ is related to a reflection measured with the receiver array, denoted by index i. The effect of the beamformer in equation 4.2 is that for a direction $\theta_{meas,i}$, for all receivers the arriving signals belonging to that $\theta_{meas,i}$ are aligned, whilst for all other reflections the signals are spread and thus suppressed. Due to this effect the beamformer output for a given $\theta_{meas,i}$ can be described as:

$$b[n,\hat{\theta}_{meas,i}] \approx A_i * s[n-m_i] + w_b[n] \tag{4.5}$$

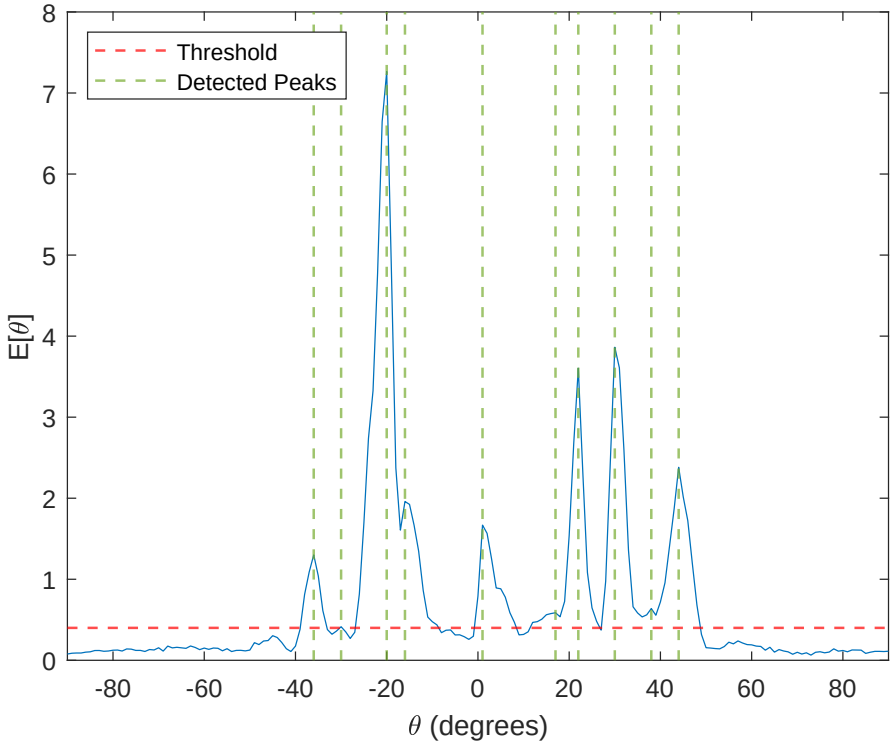


Figure 4.2: Example of energy per DOA after applying a beamformer, with both the threshold and the the detected local maxima indicated. This figure was generated using modeled data.

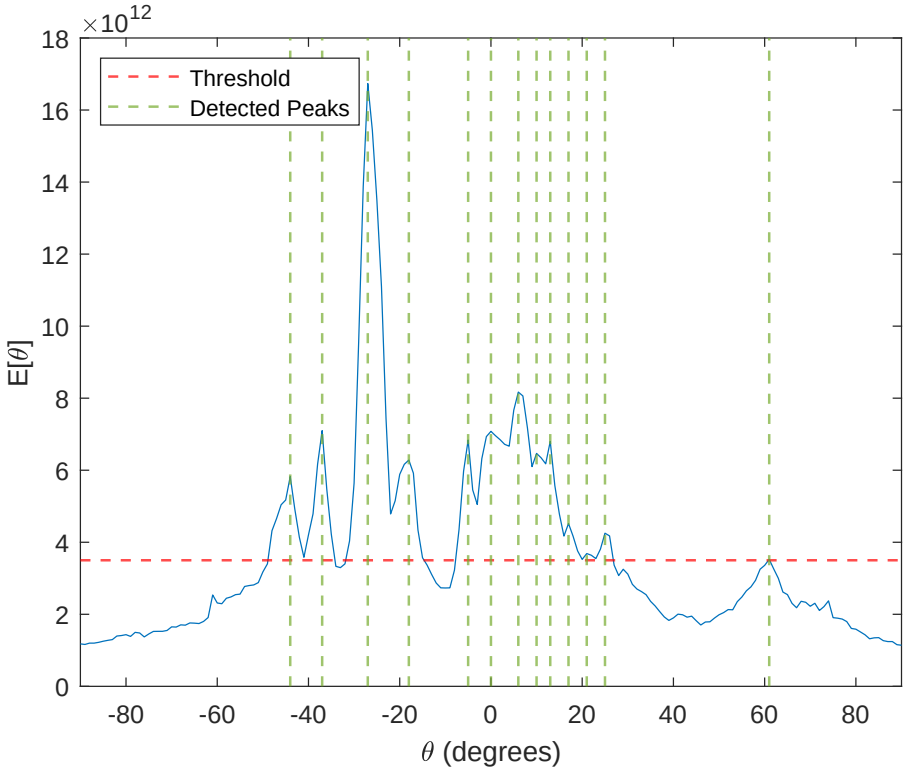


Figure 4.3: Example of energy per DOA after applying a beamformer, with both the threshold and the the detected local maxima indicated. This figure was generated using UCAC data.

Where A denotes an amplitude factor, m_i denotes the time delay of the arriving signal corresponding to $\theta_{meas,i}$ and w_b denotes a noise component that now also includes the suppressed signals of other reflections. Note that as all receiver signals are delayed in the beamformer with respect to the first receiver, these arrival angles and times are related to the position of this first receiver.

With this knowledge, the TDOA between a reflection i and a reflection j can be computed by cross-correlating $b[n,\theta_{meas,i}]$ and $b[n,\theta_{meas,j}]$:

$$TDOA_{samples,i,j} = \arg\max_{n} \sum_{t=-\infty}^{\infty} b[t,\hat{\theta}_{meas,i}]b[t-n,\hat{\theta}_{meas,j}] \tag{4.6}$$

Note that this TDOA is determined in a number of samples, but for convenience will be converted to seconds:

$$TDOA_{i,j} = \frac{TDOA_{samples,i,j}}{\text{sample rate}} \tag{4.7}$$

In the next section, pairs of DOAs and corresponding TDOAs will be compared between the measurement and the prediction of BELLHOP for a certain candidate location. To be able to do this, it is required to define how these TDOAs and DOAs can be computed in a way that they will eventually match, or at least at the real source location.

To achieve this, a reference reflection has to be chosen such that the TDOAs can be computed with respect to this reference reflection. Since all other TDOAs will depend on this reference reflection, it is important to choose this reference as reliable as possible. It is also desired that this reflection is both present in the modeled and the measured set of reflections.

To make sure that both these conditions are met, first all pairs of measured DOAs $\hat{\theta}_{meas,i}$ and DOAs predicted by the model $\theta_{mod,h}$ are selected that differ less than a certain value $\Delta\theta$. From these pairs, the most reliable arrivals are tried to be selected by taking the remaining $\hat{\theta}_{meas,i}$ and $\theta_{mod,h}$ such that the amplitude of both corresponding reflections is as high as possible. To achieve this, the product of the energy $E[\hat{\theta}_{meas,i}]$ and the modeled amplitude corresponding to $\theta_{mod,h}$ is maximized.

Let for a certain candidate location a total of $\mathscr{H}$ reflections be computed by BELLHOP, with each a propagation delay of $m_{mod,h}$ and an amplitude of $a_{mod,h}$. And let:

$$\mathscr{E}[\theta_{mod,h}] = a_{mod,h} \tag{4.8}$$

Then for both the modeled and the measured data a reference arrival is chosen as:

$$(\theta_{mod,ref}\,,\,\hat{\theta}_{meas,ref}) = \underset{\theta_{mod,h}\,,\hat{\theta}_{meas,i}}{\arg\max} \quad E[\hat{\theta}_{meas,i}] * \mathscr{E}[\theta_{mod,h}] \qquad \forall|\theta_{mod,h} - \hat{\theta}_{meas,i}| < \Delta\theta \tag{4.9}$$

With the reference DOAs known, the TDOAs can be computed with respect to these references:

$$TDOA_{meas,samples,i} = \arg\max_{n} \sum_{t=-\infty}^{\infty} b[t,\hat{\theta}_{meas,i}]b[t-n,\hat{\theta}_{meas,ref}] \qquad \forall\hat{\theta}_{meas,i} \in \boldsymbol{\Theta}_{meas} \tag{4.10}$$

$$TDOA_{meas,i} = \frac{TDOA_{meas,samples,i}}{\text{sample rate}} \tag{4.11}$$

and:

$$TDOA_{mod,h} = \frac{m_{mod,h} - m_{mod,ref}}{\text{sample rate}} \tag{4.12}$$

Where $m_{mod,ref}$ is the delay time in samples corresponding to the same modeled reflection as $\theta_{mod,ref}$.

Note that as a result of this procedure, a DOA-TDOA pair ($\hat{\theta}_{meas,i}$, $TDOA_{meas,i}$) is estimated for all measured reflections and similar, a DOA-TDOA pair ($\theta_{mod,h}$, $TDOA_{mod,h}$) is computed for all reflections that are predicted by BELLHOP for a given candidate location.

4.3. Matching DOAs and TDOAs

In the previous section, all DOA-TDOA pairs have been determined from both the measurement data and the BELLHOP model. Note that from the measured data a single set of DOA-TDOA pairs is derived, whilst from the model a set of DOA-TDOA pairs is derived for every candidate location. Recall that all θ's and $TDOA$s from the measured set where indexed by a subscript i and all modeled θ's and and $TDOA$s for a given candidate location where indexed by a subscript h. For each possible combination of DOA-TDOA pairs a score ℓ is than computed as:

$$\ell(i,h) = \exp\left[-\frac{1}{2\sigma_{TDOA}^2}(TDOA_{mod,h} - TDOA_{meas,i})^2\right] * \exp\left[-\frac{1}{2\sigma_{DOA}^2}(\theta_{mod,h} - \hat{\theta}_{meas,i})^2\right] \tag{4.13}$$

Where this score indicates how well a measured DOA-TDOA pair i matches with a modeled DOA-TDOA pair h. From these scores, it is desired to associate the different DOA-TDOA pairs that belong to the most similar modeled and measured arrivals.

To achieve this, first the DOA-TDOA pairs h and i are associated that have the highest score $\ell(i,h)$. Then, the second pair is associated based on the highest score that remains when the already associate pairs are not taken into account anymore. This process continues until either all measured or all modeled pairs (whichever has the least elements) have been associated. Since it turns out that typically there are some false DOA estimates in the measured set and also sometimes there are predicted DOAs that do not show in the measurement, the overall score of a candidate location is computed by taking the highest $\mathcal{N}$ scores resulting from the association process.

Let t be an index for the following iterative process, with a total of T iterations, where:

$$T = \min(I, H) \tag{4.14}$$

In each iteration t, a DOA-TDOA pair from the measured set will be associated with a modeled DOA-TDOA pair. Let the scores of the associated pairs be denoted by d, then:

$$d_t = \max \ell(i,h) \quad \forall i \in \{i | i \in \boldsymbol{I} \wedge i \notin \boldsymbol{I}_{associated,t}\}\ \forall h \in \{h | h \in \boldsymbol{H} \wedge h \notin \boldsymbol{H}_{associated,t}\} \tag{4.15}$$

Where $\boldsymbol{I}_{associated,t}$ and $\boldsymbol{H}_{associated,t}$ are the sets containing the i's and h's that have already been associated in a previous iteration.

The overall score of how well this set of DOA-TDOA pairs matches is than computed as:

$$\mathscr{L} = \prod_{t=1}^{\mathcal{N}} d_t \tag{4.16}$$

4.4. Choosing the threshold, σ_{DOA} and σ_{TDOA}

The detection threshold ϕ, as is used in equation 4.4, is selected manually in all realisations in this report. It is chosen such that all clear peaks reach above the threshold, but the noise floor and low-energy peaks remain below. This approach shows to be sufficient to demonstrate the concepts in this thesis, but in a realistic application this process would have to be automated. One possible solution for this might be to filter out the peaks and set the threshold at the maximum of the remaining signal.

In equation 4.13, the errors between the measured and modeled DOAs and TDOAs are assumed to be Gaussian distributed. The score is then computed based on this assumption, by using two standard-deviations σ_{DOA} and σ_{TDOA}. For these scores to be computed accurately, these standard-deviations should be chosen such that they resemble the actual inaccuracy of the considered scenario. By taking different causes of uncertainty into account, a standard deviation can be derived. For example in [25], the standard deviation in an arrival time is based on summing the expected uncertainties due to the measurement, receiver position and sound speed profile.

5

Bayesian Inference

5.1. Introduction

In probability theory, the question often is what the probability of a certain outcome of an experiment is. For example, if we consider the experiment of throwing two dice, the question could be what the probability is that the sum of the two values that are thrown is equal to a certain number. In Bayesian inference, this principle is turned backwards. The outcome of the experiment is known, also called the observation, but the input parameters are not. In mentioned example, the sum of the two dice will be known, whilst the probabilities of individual throw values are unknown and desired.

This desired probability is also known as the *a posteriori* probability, or in short posterior probability. This posterior distribution can be derived when two things are known: the likelihood of the observations to have occurred and the prior probability of the parameter values. The relationship between these three distributions is given by Bayes' rule:

$$P(\theta|\phi) = \frac{P(\phi|\theta)P(\theta)}{P(\phi)} \tag{5.1}$$

With θ as the parameter of interest, ϕ contains the observation, $P(\theta|\phi)$ is the posterior distribution and $P(\theta)$ is the prior distribution. $P(\phi|\theta)$ is called the likelihood function, which specifies the probability of a certain observation ϕ occurring for the different values of θ, or in other words, how likely it is for ϕ to occur for a given θ. $P(\phi)$ in the denominator is the probability of a certain observation ϕ occurring.

Figure 5.1 (a and b) gives a visual interpretation of what happens during Bayesian inference. The blue curve, the prior distribution, contains the knowledge of θ before any measurement or observation has taken place. If θ for example represents the number of visitors of a zoo, the prior could indicate the distribution of the amount of visitors over the past year. The red curve, the likelihood, contains knowledge based on the observation ϕ. For the case of the zoo, it could for example be that the visitors of a certain a part of the zoo are counted, scaled up for the whole zoo and a normal distributed error in the result was assumed. The black curve indicates the desired posterior distribution. This distribution combines the information in the prior and likelihood.

As can be seen from figure 5.1a, if both the prior and likelihood contain a similar level of information, the posterior will be somewhere in the middle of them. However, if one of them contains relatively few information, the other takes the overhand in the posterior, as can be seen in figure 5.1b.

In the cases where there is both a limited amount of parameters that have to be estimated and a simple relation between the observations and these desired parameters, the derivation of the posterior distribution can be done analytically. When the complexity of this relation and the parameter space grows, it becomes difficult to do this. If the posterior is desired in such a case, Markov-Chain Monte Carlo (MCMC) methods offer a solution.

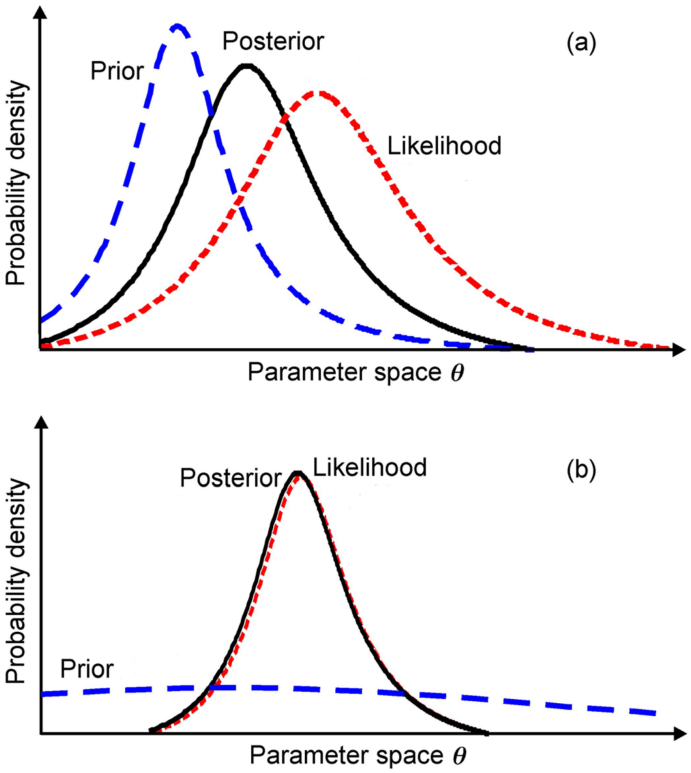


Figure 5.1: Visualisation of how a posterior density is derived from a prior density and likelihood function, for (a) a prior and likelihood both containing significant information and (b) a prior containing few information, whilst the likelihood contains relatively much information. Figure from [43].

With MCMC methods, samples are drawn from the target distribution, which in the case of Bayesian inference is the posterior. However, this distribution is in general not known. If the prior and likelihood function are known, and the relation in:

$$P(\theta|\phi) \propto P(\phi|\theta)P(\theta) \tag{5.2}$$

Can be used to approximate the desired posterior. To do this, $P(\phi)$ is omitted. This can be done without influencing the shape of the posterior, since it functions as a normalizing factor, note for a fact that it does not depend on θ [23]. The shape, mean and variances of θ will thus be identical for both equations 5.1 and 5.2.

5.2. Markov Chains

To understand MCMC methods, it is important to first understand Markov chains. A Markov chain is a system containing a number of different states, and a set of probabilities indicating the chance of moving from one state to another. An example can be seen in figure 5.2, where there are a total of three states: A,B and C. The arrows moving from one to another state indicate that this move is possible with the probability given besides the arrow.

Based on a Markov chain, a transition matrix T can be derived, containing all the probabilities to move from one state to another. For the Markov chain in this example, the transition matrix is given by:

$$T = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0.1 & 0.9 \\ 0.6 & 0.4 & 0 \end{bmatrix} \tag{5.3}$$

If a certain distribution of starting positions is considered, one can compute the probability of being in a certain state after a certain steps through the chain. If the probability for step N is denoted by P_N, and the starting probability is chosen as $P_0 = [0.2\ 0.6\ 0.2]$, matrix multiplication as in:

$$P_{n+1} = P_n * T \tag{5.4}$$

Shows that $P_1 = [0.12, 0.34, 0.54]$. After running about 30 steps through the Markov chain, this probability converges to $P_{30+} \approx [0.22, 0.41, 0.37]$.

The power of a Markov chain is that, under certain conditions, no matter what starting point P_0 is chosen, the eventual probability will converge to the same value, indicating the probabilities of each state to occur. For this to be guaranteed, each state must have a positive probability to move to another state and it must not be possible for the Markov chain to get trapped in a circle. If these conditions are met, a Markov chain becomes a powerful tool to determine the probabilities of being in a certain state.

Instead of determining the probability of each state by using the transition matrix, one can also approximate this distribution empirically. To illustrate this, 10^4 samples are drawn according to the Markov chain in figure 5.2, starting at state A. The eventual distribution of the drawn samples is shown in figure 5.3. As expected, the probability of each state occurring corresponds with the probabilities found using the transition matrix.

5.3. Markov-Chain Monte Carlo Methods

The general concept of MCMC methods is to sample from a desired distribution in such a way that the eventual samples represent this distribution. To do this, the parameter space of interest is considered as a Markov chain, where each possible set of parameters forms a state. Then, from some initial state, samples are drawn by moving through the Markov chain. Eventually, the distribution of these samples will indicate the desired distribution. The exact rules of which state is moved to with what probability, differs for the different MCMC methods.

Two common methods that will be discussed are Gibbs sampling and the Metropolis algorithm. The Metropolis algorithm is in general more complex, however instead of Gibbs sampling it still gives the desired result

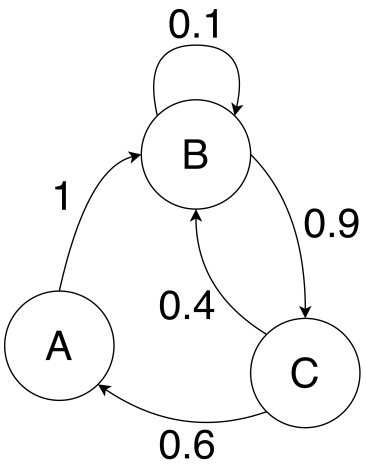


Figure 5.2: An example of a Markov chain.

when two of the considered parameters are highly correlated [23]. More complex methods are available, such as Hamiltonian Monte-Carlo (HMC) [**?**]. HMC converges faster and another big advantage is that easy to use software has been developed to work with it, such as STAN [37]. For simplicity, in this thesis only Gibbs sampling and the Metropolis algorithm are considered.

5.4. Gibbs Sampling

A relatively easy to understand MCMC method is Gibbs sampling. Gibbs sampling is, in contrast to the Metropolis algorithm that will be introduced in the next section, only applicable for models consisting of 2 or more parameters. Consider the goal is to draw samples from the parameter space $\boldsymbol{\theta} = [\theta_1, \theta_2, \theta_3]$, and there is a certain observation $\boldsymbol{\phi}$. To perform Gibbs sampling, the conditional probability densities, as in equations 5.5 need to be known.

$$\begin{aligned} &P(\theta_1|\theta_2, \theta_3, \boldsymbol{\phi}) \\ &P(\theta_2|\theta_1, \theta_3, \boldsymbol{\phi}) \\ &P(\theta_3|\theta_1, \theta_2, \boldsymbol{\phi}) \end{aligned} \tag{5.5}$$

To get started, first a starting point $[\theta_1^0, \theta_2^0, \theta_3^0]$ needs to be chosen. When possible a combination of parameters that gives a high probability is chosen here, since it will lead to a faster convergence. However, if knowledge of a high probability combination is not available, any point will do, it might only take more time. Once the starting point is set, the drawing of samples can start. Let each sample n of a total of N samples be denoted by θ^n. Then the samples are found by iteratively running through the steps in equations 5.6.

$$\begin{aligned} \theta_1^n &\sim P(\theta_1|\theta_2^{n-1}, \theta_3^{n-1}, \boldsymbol{\phi}) \\ \theta_2^n &\sim P(\theta_2|\theta_1^n, \theta_3^{n-1}, \boldsymbol{\phi}) \\ \theta_3^n &\sim P(\theta_3|\theta_1^n, \theta_2^n, \boldsymbol{\phi}) \end{aligned} \tag{5.6}$$

Note that for drawing θ_2^n and θ_3^n, the just drawn values of the other parameters are used. When all N samples have been drawn, the set of samples will represent the distribution of $\boldsymbol{\theta}$.

In figure 5.4 a visual version of the Gibbs sampling process is given. Let two parameters be considered: θ_1 and θ_2. Let them be distributed as:

$$\boldsymbol{\theta} = [\theta_1, \theta_2] \qquad \boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\sigma}^2) \qquad \boldsymbol{\mu} = [0, 0] \qquad \boldsymbol{\sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{5.7}$$

The same distribution is also displayed in figure 5.4a and 5.4b. In the latter also the starting values $[\theta_1^0, \theta_2^0] =$

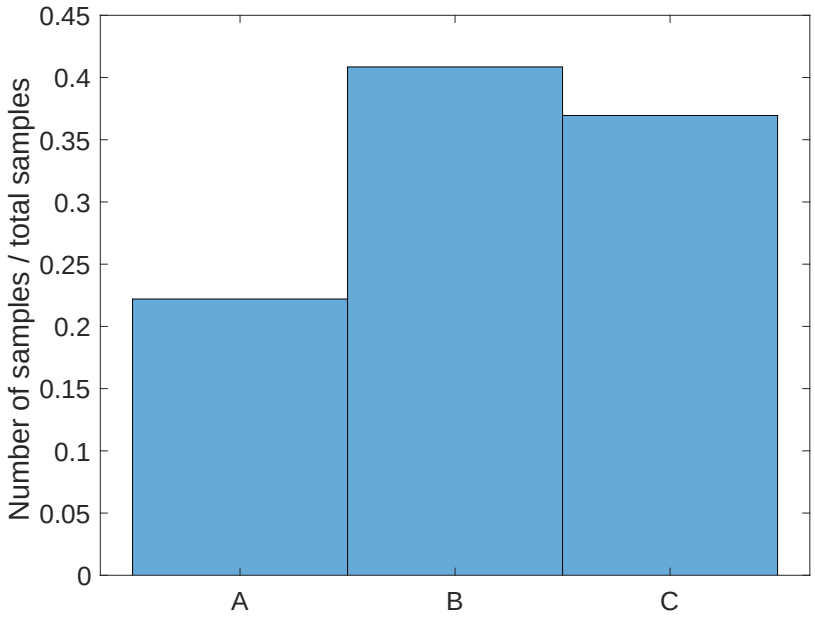


Figure 5.3: Resulting distribution of drawing 10^4 samples from the Markov chain in figure 5.2

$[1,1]$ are indicated by a marker. Suppose the next value to draw is θ_1^1. Then the conditional probability $P(\theta_1|\theta_2^0)$ is derived. In this case, this comes down to taking a vertical slice of the probability in figure 5.4b, of which the result is shown in figure 5.4c. From this distribution a value θ_2^1 is drawn. This new sample is indicated in figure 5.4d.

This process continues, as in figures 5.4e and 5.4f for the next two steps, untill all samples are drawn. In figure 3.8 the trace of the sampling is shown after 200 samples. Note that it can already be clearly seen that most samples are drawn from the region with high probability. Also note that once the trace moves into a region with less probability, it typically returns quickly to the more probable region. In figure 5.4h, a histogram of 10^4 samples is depicted. It can be clearly seen that the distribution of the samples approaches the distribution in figure 5.4a.

For these type of methods it might be challenging to determine after how many samples their distribution has converged to the desired distribution, since this distribution is not known. The general approach to check for convergence, is to run multiple sampling processes, but each with different starting points. Special measures are then available, for example as described in [23], to judge if the samples in each process are distributed similarly. When that is the case, the distribution is decided to be converged.

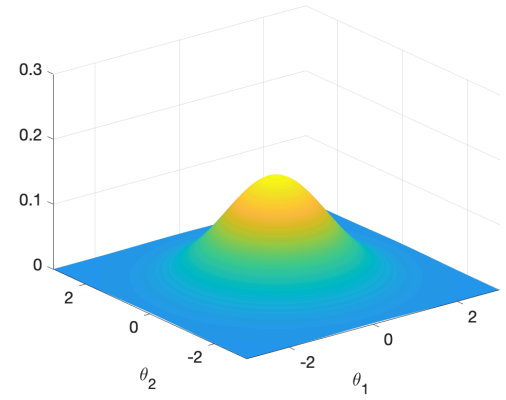
5.5. The Metropolis algorithm

The Metropolis algorithm, in the version discussed here also called the random walk Metropolis algorithm, is a bit more complex than Gibbs sampling. The biggest difference is that not all drawn samples are immediately added to the eventual set of samples, but they first have to pass a test.

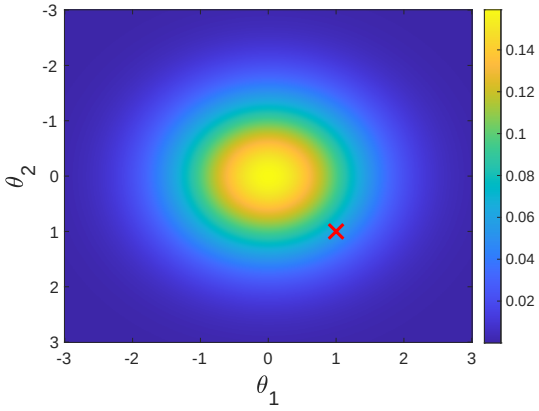
Consider the goal is to sample from the distribution of a parameter θ given an observation $\boldsymbol{\phi}$: $P(\theta|\boldsymbol{\phi})$. Just like with Gibbs sampling, first an initial point θ_0 has to be determined, preferably in a region with high probability. To get started, a random step s (hence the 'random walk' in the name) is drawn from some symmetric distribution. Here, a normal distribution will be used:

$$s \sim \mathcal{N}(0, \sigma_s^2) \tag{5.8}$$

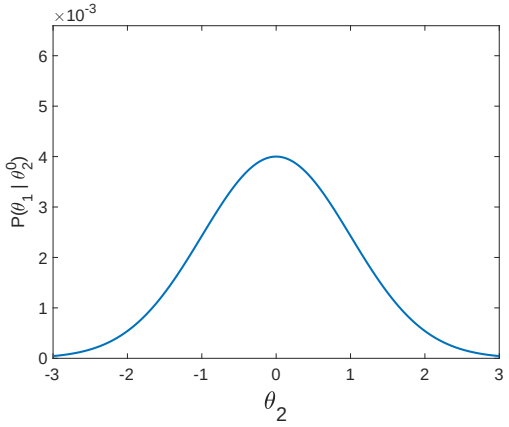
The candidate location, that might become the next samples θ_n, is then chosen as:



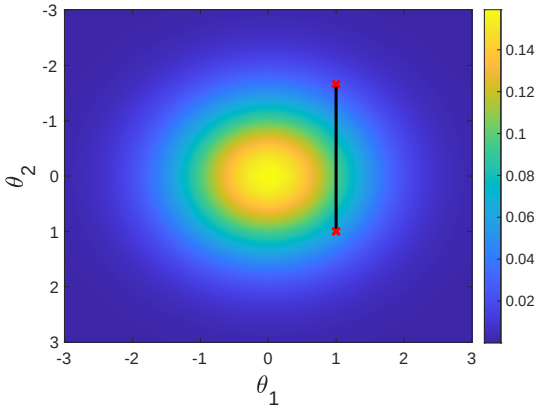
(a) The distribution of the considered parameters θ_1 and θ_2 in 3D.



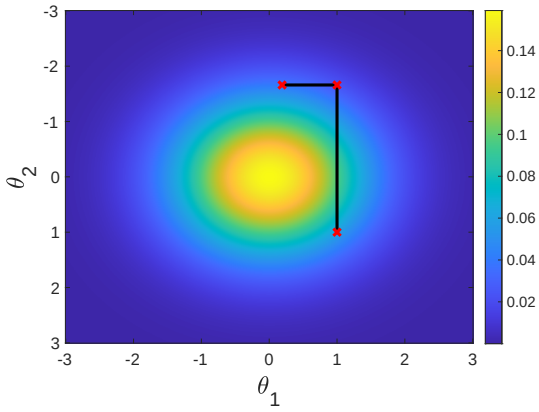
(b) The distribution from (a) in 2D. The maker indicates the initial values.



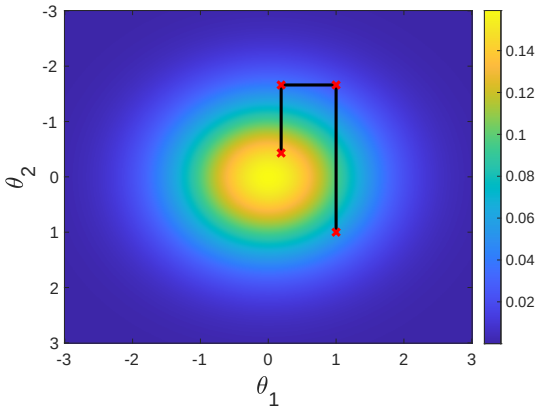
(c) The conditional probability used to draw a sample for θ_2.



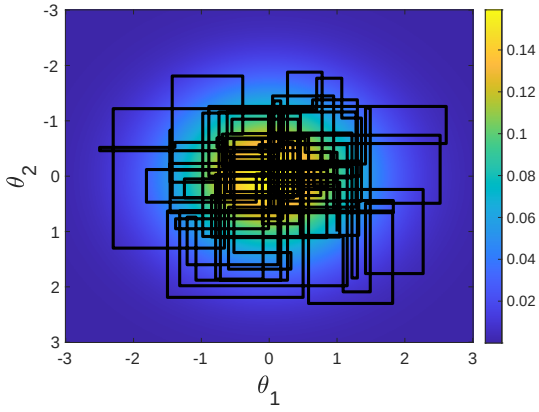
(d) Drawing the second sample.



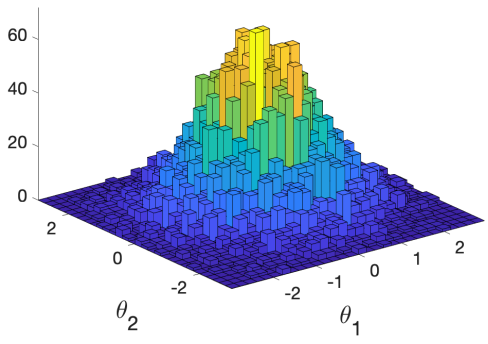
(e) Drawing the third sample.



(f) Drawing the fourth sample.



(g) Trace of the first 200 samples.



(h) Histogram plot after drawing 10^4 samples.

Figure 5.4: Visualisation of the Gibbs sampling process.

$$\theta_{candidate} = \theta^{n-1} + s \tag{5.9}$$

Then, the probability after making this step is computed: $P(\theta_{candidate}|\boldsymbol{\phi})$. Now if this probability is higher than the probability at the previous sample, this candidate sample is immediately accepted. However, if that is not the case, the candidate sample is only accepted with a probability equal to the ratio of the probability of θ_{n-1} and the candidate. This process is summarized in:

$$r = \begin{cases} 1, & \text{if } P(\theta_{candidate}|\boldsymbol{\phi}) \geq P(\theta^{n-1}|\boldsymbol{\phi}) \\ \frac{P(\theta_{candidate}|\boldsymbol{\phi})}{P(\theta^{n-1}|\boldsymbol{\phi})}, & \text{if } P(\theta_{candidate}|\boldsymbol{\phi}) < P(\theta^{n-1}|\boldsymbol{\phi}) \end{cases} \tag{5.10}$$

Where r indicates the probability that $\theta_{candidate}$ is accepted as θ^n.

When $\theta_{candidate}$ is not accepted as a new sample, the previous sample θ^{n-1} will be added to the list again and the process will start over from equation 5.8. This process is repeated until the desired amount of samples is gathered.

In figure 5.5, a visualisation of the steps of the Metropolis algorithm is provided. Considering sampling for a parameter θ, of which the distribution can be seen in figure 5.5a. In this same figure also the starting position θ^0 is indicated. To determine the first sample, θ^1, a candidate is selected according to equation 5.9. This candidate is marked in magenta in figure 5.5b. Now since the probability of this $\theta_{candidate}$ is rather small, following equation 5.10 the chance of it being accepted is also small. By random draw it is decided that this candidate indeed will not be accepted and $\theta^1 = \theta^0$.

As a next step, another $\theta_{candidate}$ is drawn, as pictured in figure 5.5c. Since the probability of this candidate is higher than of the last sample, it will be immediately accepted, as indicated in figure 5.5d. After that, another candidate is drawn, as indicated in figure 5.5e. The probability of this $\theta_{candidate}$ is about half of that of the previous sample, which means the chance of it being accepted is also about 50%. By random draw this sample is accepted, and the process continues like this as indicated in figures 5.5f and 5.5g. The resulting distribution after drawing 10^3 samples is shown in figure 5.5h, note the similarity with the distribution of θ.

The advantage of the metropolis algorithm is that due to the random step it also samples in places that have low probability and does not get caught in zones of high probability. A disadvantage of the metropolis algorithm is that it converges slower, due to not accepting all samples. This makes it take longer, before the whole parameter space is sampled representative. Another disadvantage is that some kind of step scale (in the case discussed before, σ_s) has to be chosen. If this step size is to big, some areas in the sample space might be skipped, whilst if the step size is to small, it takes longer to converge and/or some parts of the distribution might not be reached.

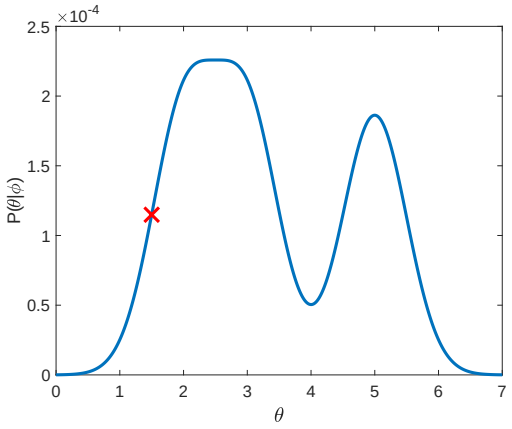
The metropolis algorithm can be simply expanded to sampling from a larger parameter space, by drawing from a multivariate (normal) distribution in equation 5.8.

5.6. Application in acoustic localization

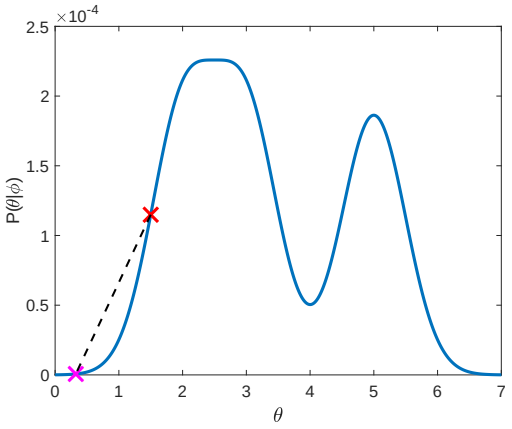
In the previous section it has been introduced how samples can be drawn according to a certain distribution, to eventually approximate this distribution. To apply this to Bayesian inference, samples are drawn from the distribution on the right-hand side of equation 5.2.

In the descriptions of Gibbs sampling and the Metropolis algorithm, samples have been drawn from the posterior distribution $P(\theta|\boldsymbol{\phi})$ or the conditional posterior distributions as $P(\theta^1|\theta^2, \boldsymbol{\phi})$. It is assumed that these functions can be approximated from the likelihood and the prior function. Now what is important to notice is that it might seem that if $P(\theta|\boldsymbol{\phi})$ and $P(\theta^1|\theta^2, \boldsymbol{\phi})$ are already available, the posterior distribution is already known. However, the values of these functions can only be evaluated in a single point in the parameter space, whilst the posterior distribution of an individual parameter is desired.

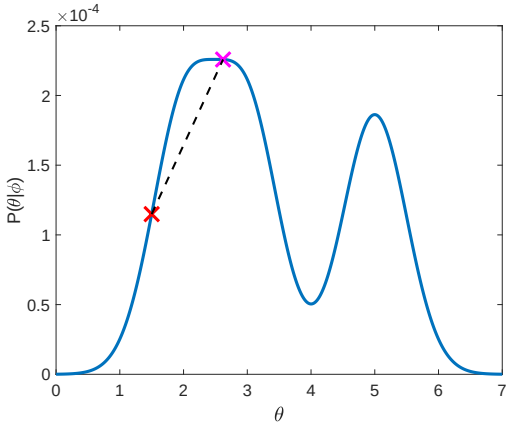
Let this be clarified with an example. Assume that Gibbs sampling is performed for three parameters. Then



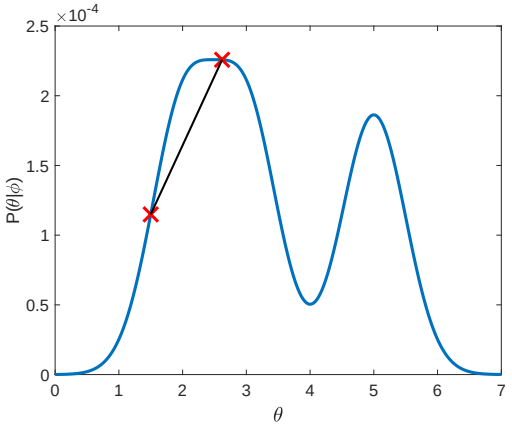
(a) The distribution of θ with the starting point marked.



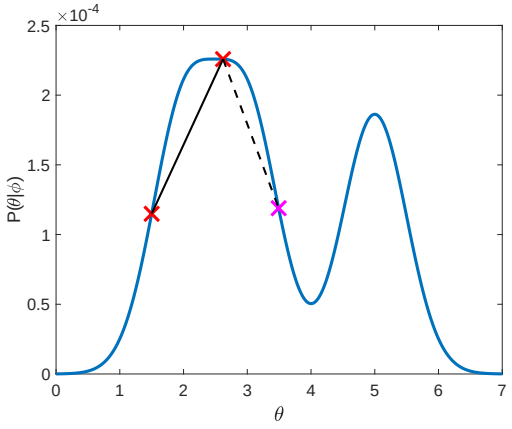
(b) The first candidate location, which will not be accepted.



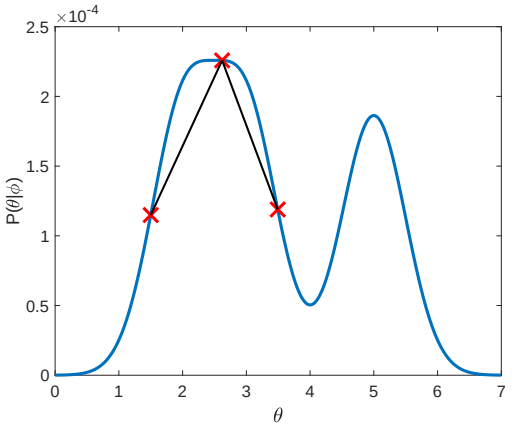
(c) The second candidate location, which will immediately be accepted.



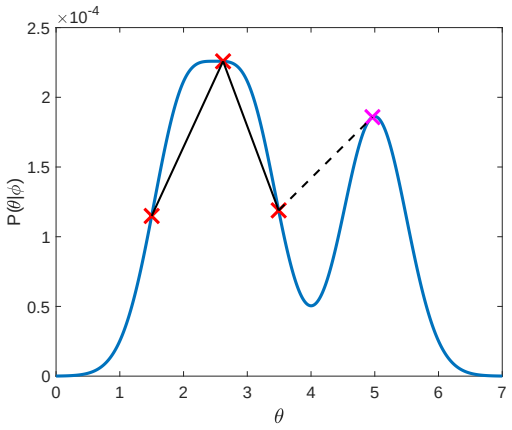
(d) The accepted second candidate sample.



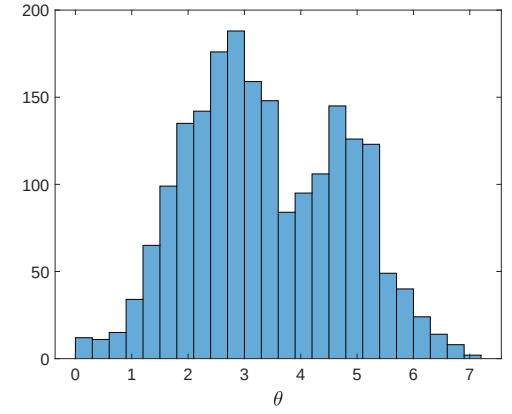
(e) The third candidate sample, which roughly has a 50/50 chance of being accepted.



(f) The third candidate sample is accepted.



(g) The fourth candidate location, which will immediately be accepted.



(h) Histogram plot after drawing 10^3 samples. Notice the similarity with figure (a).

Figure 5.5: Visualisation of sampling with the Metropolis algorithm.

by using the likelihood and prior the distribution $P(\theta^1|\theta^2,\theta^3,\boldsymbol{\phi})$ can be approximated. What in the end is the desired is the distribution $P(\theta^1|\boldsymbol{\phi})$. But if this has to be derived from $P(\theta^1|\theta^2,\theta^3,\boldsymbol{\phi})$, the distributions $P(\theta^2|\boldsymbol{\phi})$ and $P(\theta^3|\boldsymbol{\phi})$ would be required. But to derive these two distributions, $P(\theta^1|\boldsymbol{\phi})$ is required. This is a vicious circle, which can be broken up by the Gibbs sampling process.

The same holds for the metropolis algorithm, where $P(\theta|\boldsymbol{\phi})$ is assumed to be known. Now indeed, if the parameter space is limited to only one parameter θ, just evaluating this function over θ will give it's posterior. But when multiple parameters are considered, for example if:

$$\boldsymbol{\theta} = [\theta_1, \theta_2, \theta_3] \tag{5.11}$$

$P(\boldsymbol{\theta}|\boldsymbol{\phi})$ can only be evaluated by using a value for all three parameters, and just as in the previous example, to be able to determine $P(\theta^1|\boldsymbol{\phi})$, the distributions of $P(\theta^2|\boldsymbol{\phi})$ and $P(\theta^3|\boldsymbol{\phi})$ are required, which leads to the same vicious circle.

Breaking this circle is exactly where the power of Gibbs sampling and the Metropolis algorithm lies. By drawing samples of θ_1 from the conditional or multivariate posterior densities, whilst using sample values of the other variables that are representative for their individual posterior distributions, the eventual samples of θ_1 will represent $P(\theta_1|\boldsymbol{\phi})$.

In acoustic source localization, the main interest when applying Bayesian inference is to determine the posterior distribution of the range and depth estimates. To achieve this, a likelihood function based on measurement data is formulated. Examples of a likelihood function are equations 3.27 and 3.28 or the score as described for the method described in chapter 4. Using this likelihood function, the likelihood for a given data set can be computed for all considered candidate ranges and depths. In general, no prior information is available, which translates to a uniform prior along the candidate ranges and depths. However in some cases, such as tracking, a prior assumption may be made.

Considering the scenario where only range (r) and depth (d) are estimated, and the observed data set is $\boldsymbol{\phi}$, the equations for a Gibbs sampling iteration in equation 5.6 can be adapted as:

$$\begin{aligned} r_n &\sim P(r|d_{n-1},\boldsymbol{\phi}) * P(r) \\ d_n &\sim P(d|r_n,\boldsymbol{\phi}) * P(d) \end{aligned} \tag{5.12}$$

For the case of the metropolis algorithm, consider that $\boldsymbol{\theta} = [r, d]$. Since now two parameters are to be estimated, the random step has to be drawn from a two-dimensional normal distribution as:

$$\boldsymbol{s} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\sigma_s}^2) \tag{5.13}$$

After which the procedure can continue as in equations 5.9 and 5.10, but then with $\theta_{candidate}$ and θ_{n-1} replaced by their vector version $\boldsymbol{\theta}$. Results of the procedure using Gibbs sampling will be presented in sections 6.1 and 6.2.

At this point, the posterior distributions of the range and depth can be approximated. Bayesian inference can also be used to include environmental parameter uncertainty into the estimation process or to estimate these environmental parameters in the case they are unknown. This concept is demonstrated several times in literature, for example in [32] or [39].

Consider for example the case where the bottom and surface reflection coefficients are unknown and the goal is to estimate them. These parameters can then be added to the parameter space and a uniform prior can be assumed. Instead of only computing the likelihoods for all the ranges and depths, now the likelihoods for all combinations of range, depth and the two reflection coefficients are computed. Note that another likelihood function should be used as mentioned before, since these do not rely on the amplitudes of the arrivals. If there is a strong relation between the real reflection coefficients and the likelihood score, the posterior for that reflection coefficient will give a peak around the real value.

Another situation could be where there is a measurement of the reflection coefficients, but there is some uncertainty in it, perhaps following from a measurement uncertainty. The same procedure as before can be applied, but now with a prior distribution that represents this uncertainty, for example a normal distribution centered around the measured value. During the MCMC sampling procedure that follows, the values of the reflection coefficients that are drawn follow a combination of the prior and their likelihood function. If these values align (the measurement was correct), this means that most samples will be drawn from the measured value. However, if the likelihood function indicates that another value seems to be more accurate, the mean of the sample values will shift towards this value, as is indicated in figure 5.1.

In section 6.2 an implementation and results of this procedure will be given for an uncertain array tilt. In the remainder of this report, Gibbs sampling will be used for Bayesian inference, since it is simpler, converges faster and is not dependent on a step size [23].

6

Results using modeled input

In chapter 4, a new method has been proposed to perform high-frequency acoustic source localisation with a vertical hydrophone array. In chapter 5, methods have been introduced to estimate the posterior distributions for a sources range and depth. To get an indication of how well these methods work, they have been implemented and the results will be shown in this and the next chapter. In this chapter, the results will be given when input generated by BELLHOP is used. To be more precise, the different receiver signals are computed as in equation 3.2, where the delay times and amplitudes are predicted with BELLHOP. In chapter 7, results will be shown when using the measurement data from the UCAC data set.

The environment that is used here has been described in section 2.4. As receivers an array of 62 elements with intermediate spacing of 15cm is used, of which the upper element is located ~ 24 meters under the surface. A visualisation of this environment and the receiver positions can be found in figures 2.5 and 2.2.

For the implementation of the methods, a grid of candidate source locations is introduced. The range in this grid runs from 0 to 2km and has an intermediate step-size of 4 meter. The depth runs from 0 to 150 meter and has a step-size of 1 meter. For each of these candidate locations, BELLHOP is used to predict the arrival times and angles. Using the method as described in chapter 4 a score is then computed for each candidate location. For 4 different source locations the energy-per-DOA and the resulting scores in the form an ambiguity surface are shown in figures 6.1 and 6.2.

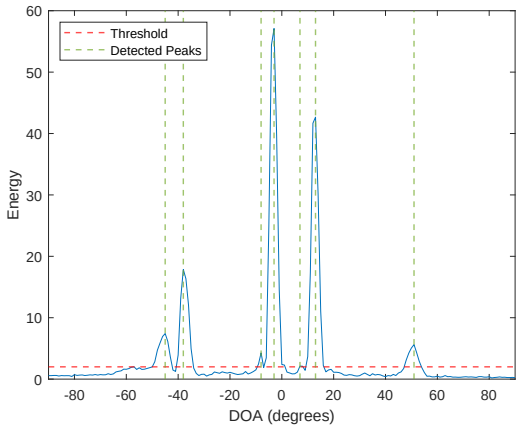
To overcome an error that might be caused by the beamformer only considering integer DOA's, σ_{DOA} is chosen to be 1 degree. To allow for small errors in arrival times, for example due to rounding in the conversion from time to samples and back, σ_{TDOA} is chosen to be 5 milliseconds. $\Delta\theta$, that is used to choose a reference reflection, is set to 2. The number of arrivals that is taken into account $\mathcal{N}$ is for a total of I detected arrivals chosen as:

$$\mathcal{N} = \min(6, I-1) \tag{6.1}$$

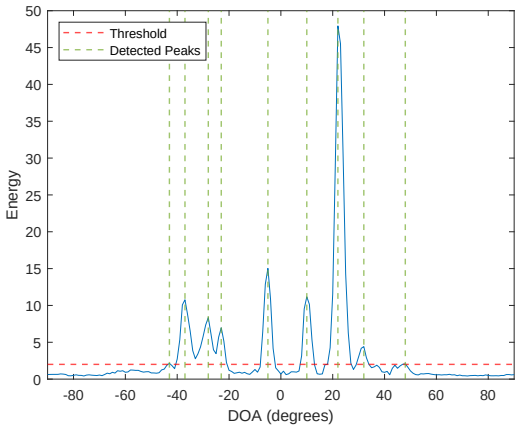
To illustrate the process of Gibbs sampling, the scores for different candidate locations are assumed to indicate the likelihood of these locations and are combined with uniform priors to estimate the posterior distributions of the range and depth. These results are also included in figures 6.1 and 6.2, where the red lines indicate the real parameter values. The typical accuracy for range in these figures is ~ 100 meters and for depth ~ 25 meters.

In section 6.2, the principle of incorporating environmental parameters in the estimation process is demonstrated. By artificially shifting the measured DOAs, an array tilt is induced in the data. The scores are computed for the same grid as described before but now also for all array tilts ranging from -30 to 30 degrees with an increment of 1 degree. The resulting posterior distributions of the range, depth and array-tilt are shown in figures 6.3 to 6.6.

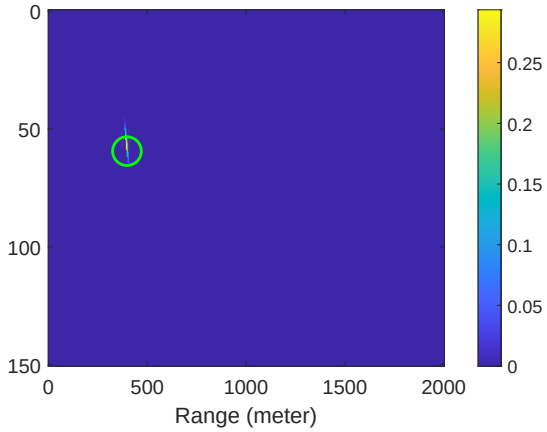
6.1. Proposed Method with Gibbs Sampling



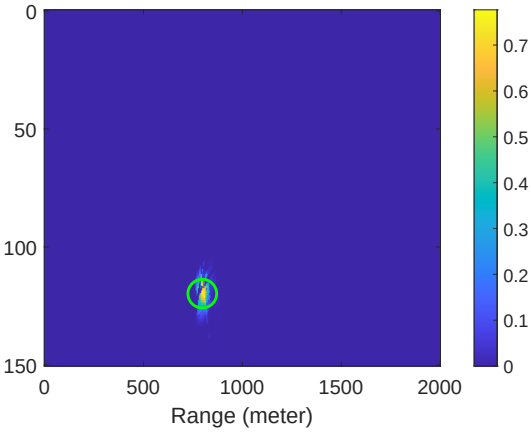
(a) Energy per DOA.



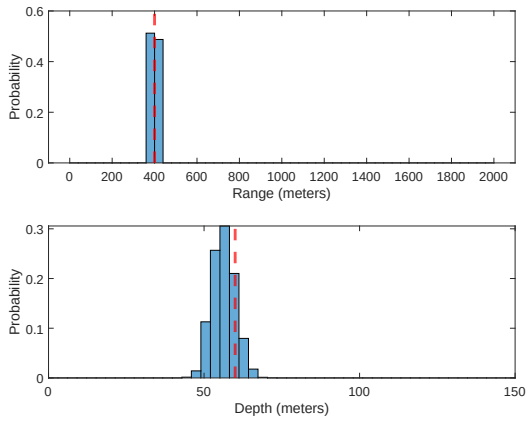
(b) Energy per DOA.



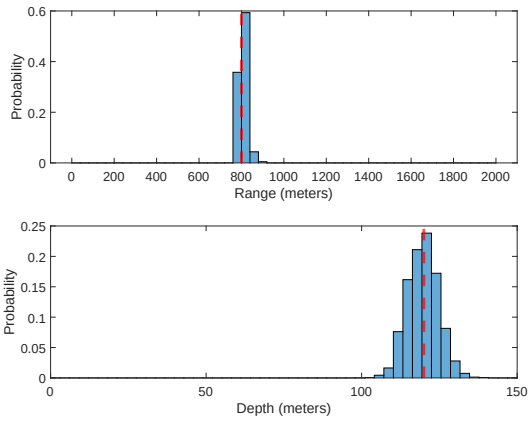
(c) Ambiguity surface.



(d) Ambiguity surface.

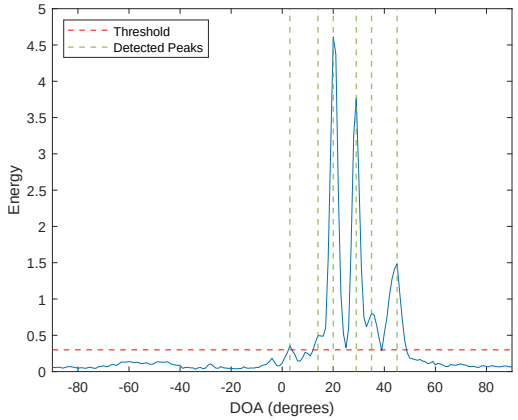


(e) Posterior probability densities.

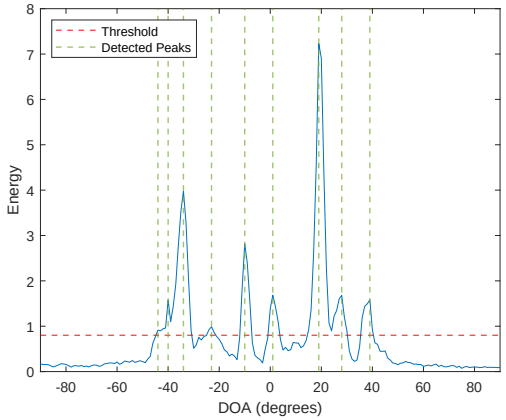


(f) Posterior probability densities.

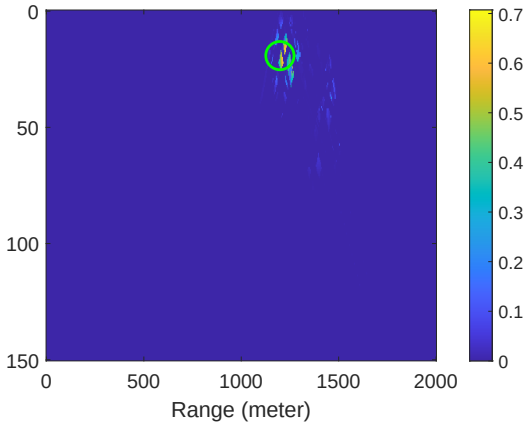
Figure 6.1: Results using modeled data as input with a source at a range of 400m and a depth of 60 meter (a, c, e) and a source at a range of 800m and a depth of 120m (b, d, f). The green marker in figures c and d and the red markers in figured e and f indicate the real source and depth.



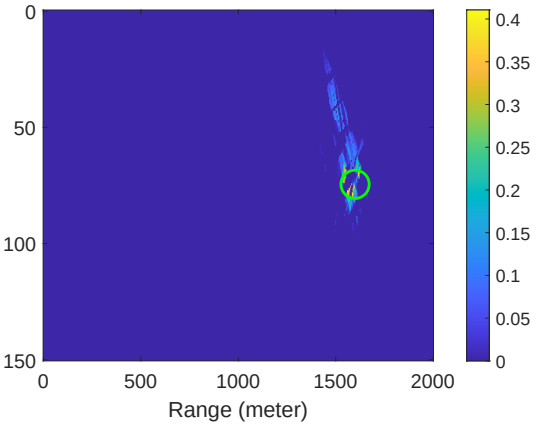
(a) Energy per DOA.



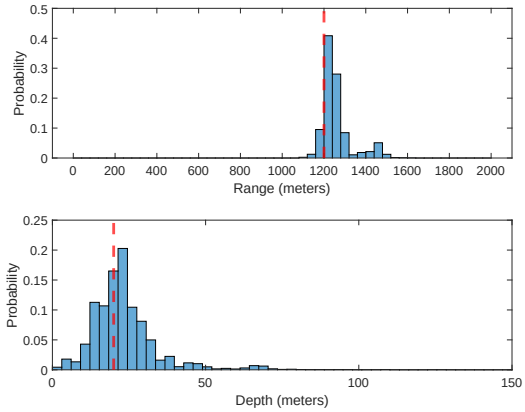
(b) Energy per DOA.



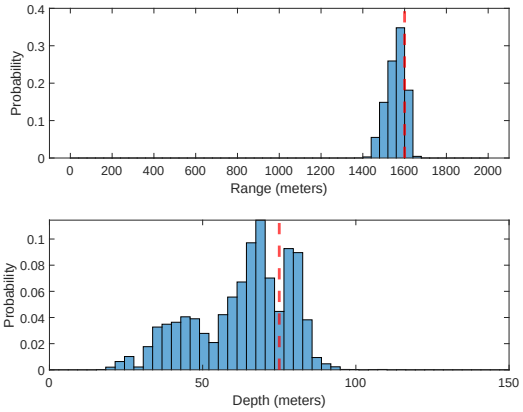
(c) Ambiguity surface.



(d) Ambiguity surface.



(e) Posterior probability densities.



(f) Posterior probability densities.

Figure 6.2: Results using modeled data as input with a source at a range of 1200m and a depth of 20 meter (a, c, e) and a source at a range of 1400m and a depth of 75m (b, d, f). The green marker in figures c and d and the red markers in figured e and f indicate the real source and depth.

6.2. Estimating Environmental Parameters: Array Tilt

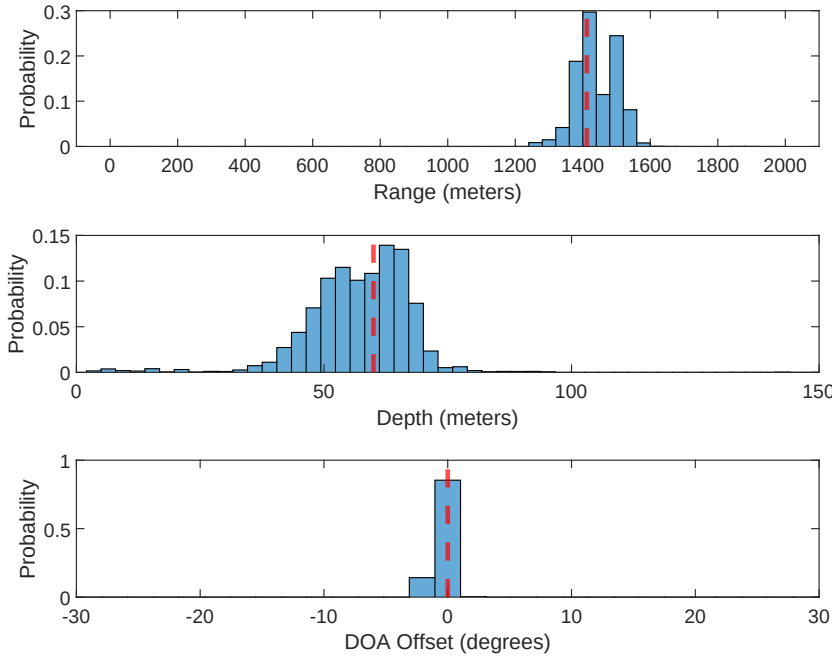


Figure 6.3: Posterior distributions with an array tilt of 0 degrees.

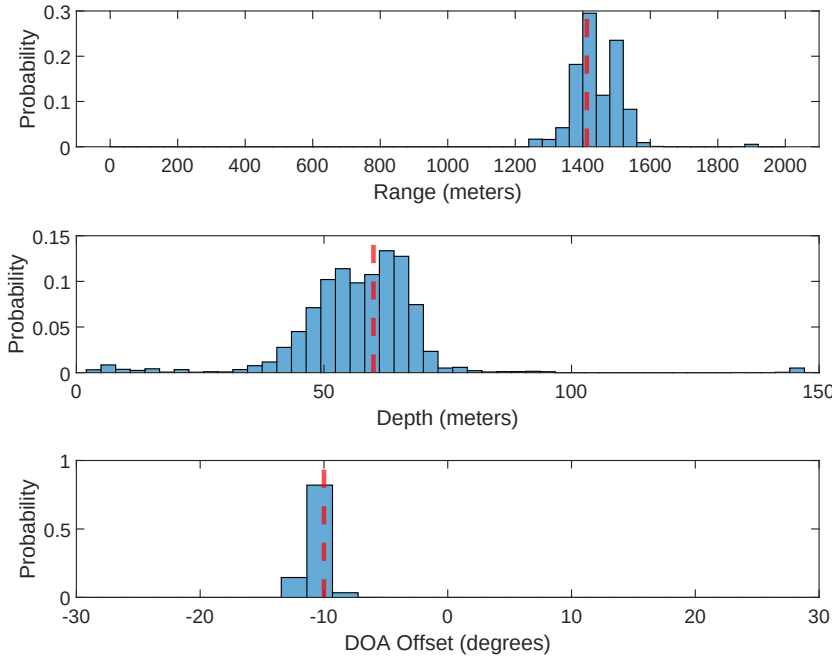


Figure 6.4: Posterior distributions with an array tilt of -10 degrees.

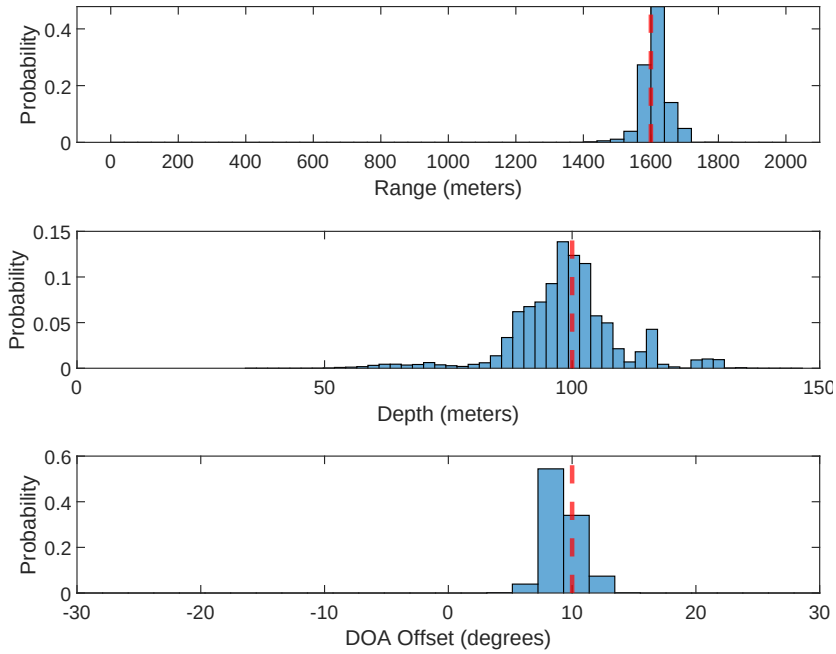


Figure 6.5: Posterior distributions with an array tilt of 10 degrees.

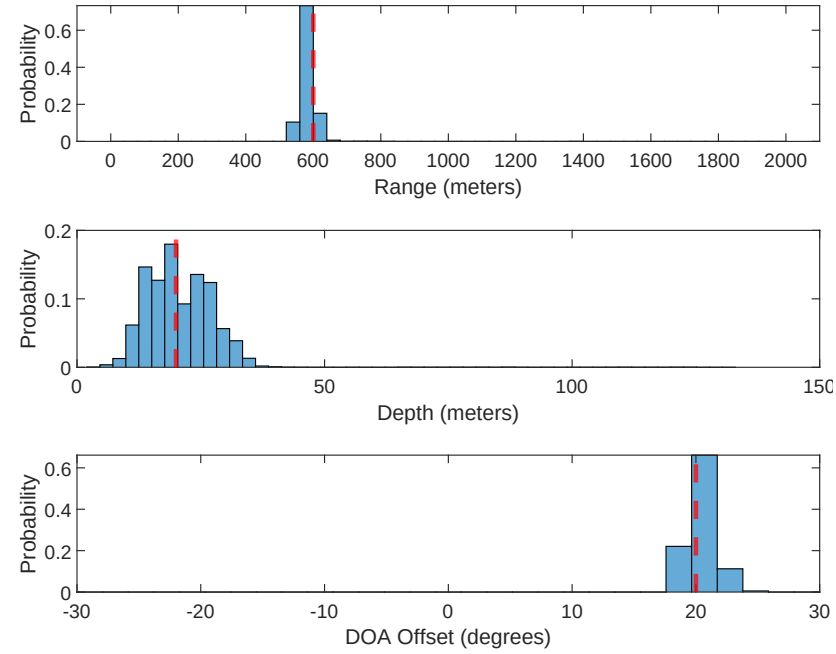


Figure 6.6: Posterior distributions with an array tilt of 20 degrees.

7

Results using UCAC data

In the previous chapter several results have been shown using input as modeled by BELLHOP. The eventual goal of the methods presented in this report is to be able to perform localisation with a real measurement data set, including noise and environmental uncertainty. In this chapter, the results are shown when the proposed method is used in combination with Gibbs sampling including the array tilt, but now using measurement data from the UCAC data set as input. The approach is, besides the different input data, the same as was described in the previous chapter. The results for a UCAC data set measured at a range of ~ 1410 meter and ~ 1760 meter are shown in figures 7.1 and 7.2. In section 7.2, considerations on the size of the used array are made. To compensate for additional mismatch in the environment, σ_{TDOA} is increased to 10 milliseconds.

7.1. Estimating Environmental Parameters: Array Tilt

In figure 7.1 it can be seen that an accurate and precise estimate is made for the range, where the accuracy is comparable with the results using modeled data in the previous chapter. The center of the main area of probability for the depth is off with about 40 meters. The array tilt, which is actually unknown, gives high probability on and around ~ -8 degrees. As this is also the case in the other UCAC measurement presented here, this is assumed to be correct.

In figure 7.2, also an accurate estimate of the range is displayed but for this range it is less precise. For the depth the posterior clearly indicates that there is much uncertainty. Note that in the modeled results, a source at a location close to the source in this case also gives relatively much ambiguity in the depth. This might indicate that in this area just similar DOA-TDOA patterns are present.

7.2. Considerations on Array Configuration and Size

So far, for all simulations and measurements the middle part of the UCAC array has been used, consisting of 62 elements separated by 15cm, thus resulting in a total array of 9.3 meters. The source signal used is an LFM chirp which runs from 2.1 to 5.6 kHz in 1 second.

In the method presented in chapter 4, a beamformer is introduced to determine the DOAs of different arrivals. Although methods are being developed that overcome this, a challenge that occurs when working with a beamformer and high-frequency source signals is that spatial aliasing might occur. To prevent spatial aliasing, there should at least be two receivers in a wavelength, or:

$$\Delta \leq \frac{\lambda}{2} \tag{7.1}$$

Where Δ indicates the receiver spacing and λ the wavelength. When this principle is translated to source frequency, this results in the following equation:

$$\Delta \leq \frac{c}{2f_{max}} \tag{7.2}$$

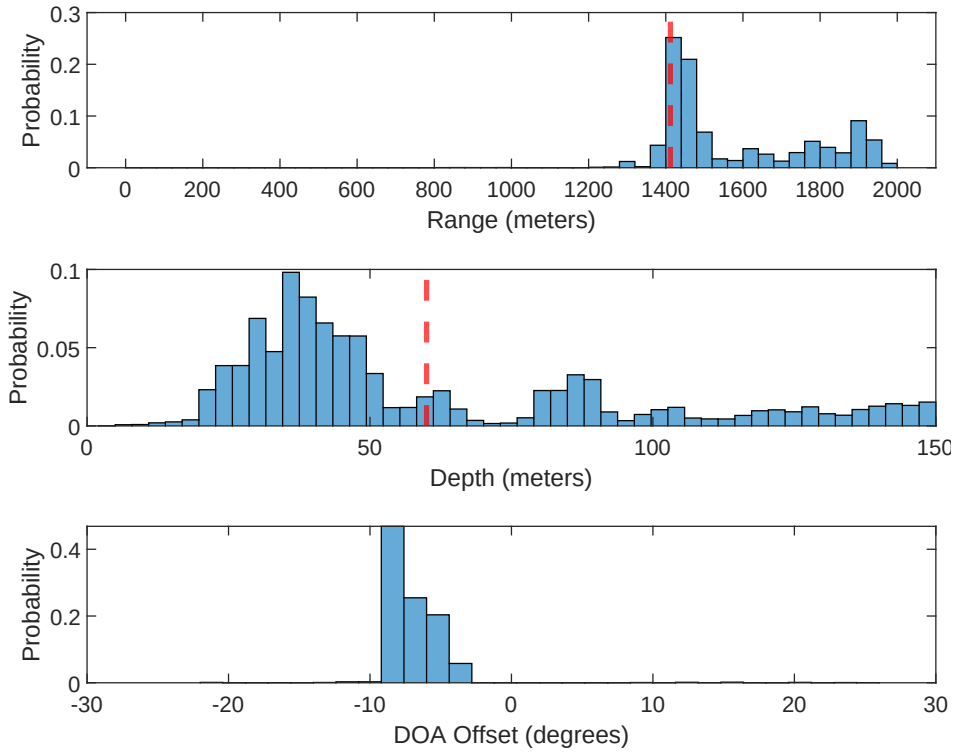


Figure 7.1: Posterior distributions when using data from the UCAC measurement recorded with a range of ~ 1410 meter.

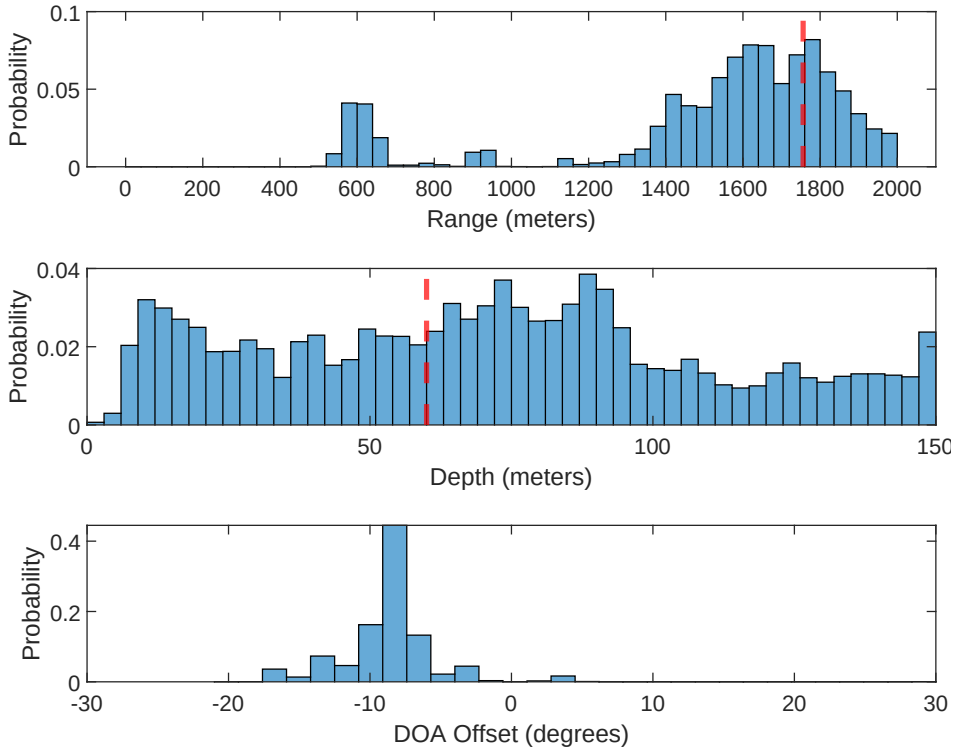


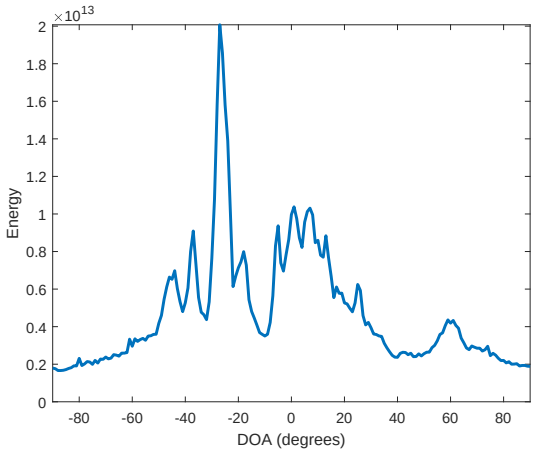
Figure 7.2: Posterior distributions when using data from the UCAC measurement recorded with a range of ~ 1760 meter.

Where c indicates the sound speed and f_{max} the maximum frequency of the source signal. Considering that f_{max} is 5.6 kHz for the source signal used in this thesis and $c \approx 1500$ m/s, Δ should be at most around 13 cm, which is not the case for the UCAC array. Given the spacing of 15 cm, a maximum frequency is allowed of at most $\sim$ 5 kHz. There are methods available in literature that provide anti-alias beamforming, for example in [36], but for simplicity these are not used in this work. To prevent aliasing and to filter noise a filter with a pass band from 2 to 4.5 kHz has been used. This does however filter part of the signal, reducing the SNR which in turn might reduce the localization performance. For this reason, when designing an array for localizing high-frequency acoustic sources the relation in equation 7.2 should be taken into account.

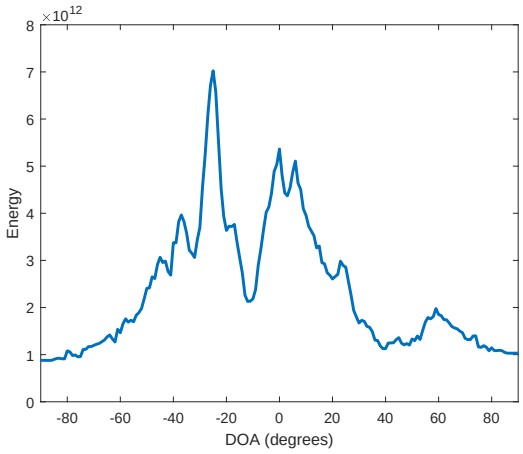
Another consideration is the size of the array. Up until now 62 elements ranging over more than 9 meters are considered, but in many practical applications this will not be a feasible option. For example when the array needs to be mounted to a vessel or simply because the budget is limited.

To give an idea what an acceptable size is to still be able to detect sufficient (>3-4) DOAs, the DOA energies for a changing number of array elements in the UCAC 1400m scenario are shown in figure 7.3. For the full 62-element array, multiple DOA peaks that seem to be reliable DOAs are distinguishable. For both the 31-element and 22-element array, still at least 3-4 of these peaks seem to be distinguishable, even though the resolution clearly decreases and some of the peaks disappear. For the 15-element array, the amount of clearly detectable peaks seems to be reduced to at most 3, which is expected to not contain sufficient information anymore.

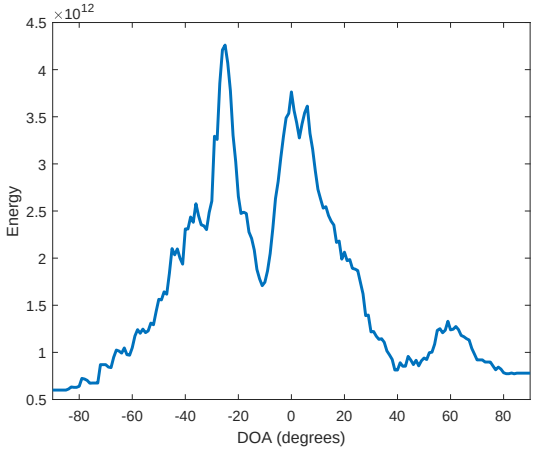
To further investigate this, the localization algorithm was applied to data from just the upper 31 receivers of the previously used 62-element array. Using this data, the algorithm was not able to find a combination of parameters with a good score. Since enough peaks where still distinguishable (as can be seen in figure 7.3b) and also where detected, it is assumed that the reduced SNR at the beamformers output caused faulty TDOA estimation(s), which in turn prevented a good match from occurring.



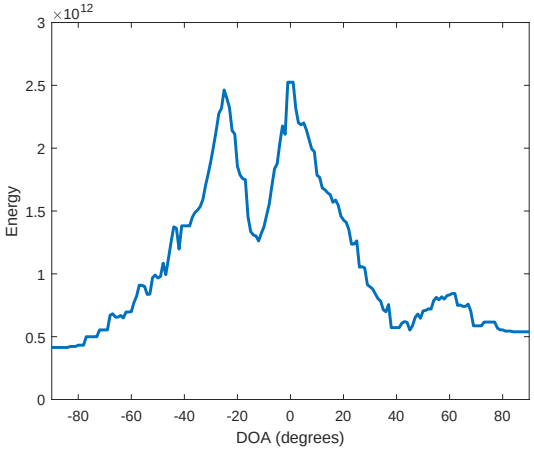
(a) 62 array elements.



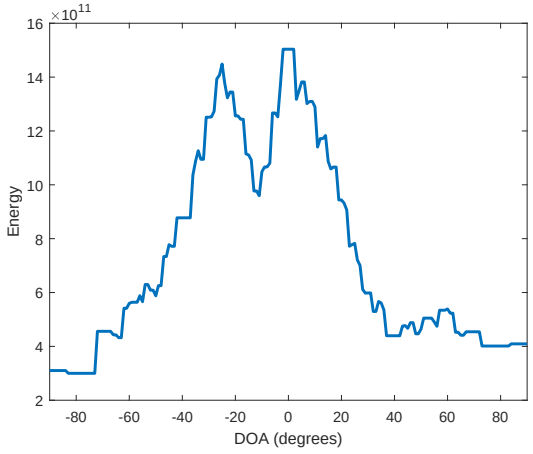
(b) 31 array elements.



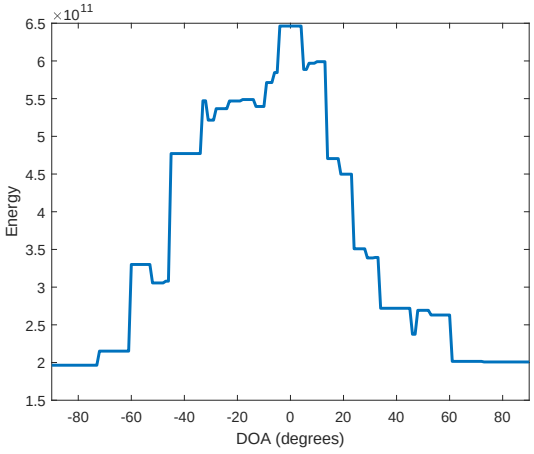
(c) 22 array elements.



(d) 15 array elements.



(e) 10 array elements.



(f) 5 array elements.

Figure 7.3: DOA energy in the UCAC 1413m scenario for a changing number of array elements.

8

Conclusions

8.1. Conclusions

In this thesis, it has been demonstrated that by introducing a new method for source localization a high-frequency acoustic source at a range of 1500-2000m in an ocean environment with a depth of ~ 200 meters can be localized with the use of a vertical line-array of hydrophones.

The main challenges that had to be overcome to achieve this included: (1) accurate TDOA estimation by peak detection from a matched filter, cross- or autocorrelation, (2) accurate TDOA association preventing unintended good or bad scores, (3) uncertainty about the reliability of or confidence in a location estimate and (4) faulty location estimates due to the influence of environmental mismatch.

Based on the first two challenges and the limitations of a high-frequency source and vertical line-array configuration, a new method was proposed in chapter 4. In this method, a delay-and-sum beamformer was used to both separate the different arrivals based on their DOA and retrieve the time series signals from the beamformer containing approximately only the time series of the reflection arriving from that DOA. This DOA was then used to add another dimension to the arrival in the association process, preventing wrong associations. The time series where used to determine TDOAs with less chance of detecting the wrong peak.

In section 6.1, the results are shown of implementing this method using receiver signals generated by BELLHOP as input, thus without the effects of noise and environmental uncertainty. It can be clearly seen that the actual source location is indicated by a peak in the ambiguity surfaces, indicating that the proposed method works as expected and indeed prevents unintended good- or bad scores.

As a solution for the third challenge, the posterior distributions of the range and depth where estimated by Bayesian inference through the use of Gibbs sampling, as was introduced in chapter 5. Also in section 6.1, the resulting posterior distributions are shown. The posterior distributions clearly show that most probability is centered around the real range and depth and also give an indication how certain these estimates are, as was desired.

Gibbs sampling also provides opportunities for the fourth challenge, since it facilitates to incorporate uncertainty in environmental parameters into the location estimation. In this thesis, this has been demonstrated by taking the tilt of the receiver array into account. In section 6.2, the posterior distributions of the range, depth and array tilt estimates are shown, whilst an artificial array tilt is created in the data. It can be clearly seen that the array tilt is estimated very precise in all of the demonstrated cases. Note that when there is an array tilt and it is not considered in the localization process, this most probably would cause faulty estimates.

Up until this point, only receiver data as produced by the BELLHOP model was considered. In section 7.1, the proposed method including Gibbs sampling and estimating the array tilt are applied to two sets of receiver signals from the UCAC data set. It can be clearly seen that both of the real ranges are within the main area of probability. For the range of 1400m the range is estimated quite accurately. The depth estimates show a lot

more uncertainty. An array tilt of ~ -8 degrees is estimated from both the measurements. Information of the actual array tilt is unavailable, but since the same tilt is indicated for both measurements it is assumed that this was the actual array tilt at the moment of the measurements.

In section 7.2, the influence of the size of the receiver array was briefly studied. The results indicate that up to ~ 22 receivers enough DOAs could be still estimated, but reduced SNR seems to cause the further localization to fail.

8.2. Discussion

Throughout this thesis, several assumptions have been made. The validity of these assumptions and thus of the results and conclusions as presented in this work require discussion, which will be given in this section.

To start, throughout this whole thesis only a single ocean environment is considered. This environment is chosen since this is where the UCAC data set was recorded, which allows to also test methods using this real measurement data. As performance is now only shown and tested for this specific environment, to be able to draw more general conclusions on the applicability of these methods, a study should be conducted also including other typical ocean environments.

In this work only a single source signal has been considered, being an LFM chirp as was described in section 2.3. This signal is chosen since it is used in the UCAC transmissions and since LFM chirps in general are good detectable due to their autocorrelation containing a peak, which gets sharper for an increasing bandwidth. In some applications, as tracking an UAV or experimental settings, this is a realistic and usable source signal. In some other scenarios, for example tracking marine mammals, one is limited to detecting the vocalisations of these mammals, which are expected to be less good detectable than the LFM chirp. To draw any conclusions on this, the methods should also be tested with other source signals.

Another point of attention is that in many realistic situations not a single acoustic source is present, but multiple sources. This can for example be the case when tracking marine mammals. If the other sources also require detection, an adapted algorithm is required. Several examples of such an adapted algorithm are presented throughout literature [2, 14, 39] and it is expected that the methods as presented here can be extended to multiple-source localisation in a similar manner, which would be an interesting option for future work.

In TDOA-based methods interfering signals (received signals that do not originate from the source(s) of interest) are a known problem, since they also correlate when TDOAs are estimated in a cross- or autocorrelation. In the methods as proposed here, it is expected that since the arriving signals are approximately separated by the beamformer, the effect of this problem is reduced. This should however also be verified.

In the method as presented in chapter 4, a manual set threshold is used in the peak detection. This is sufficient to demonstrate the concept and performance as in this report, but in a realistic situation it would be a significant burden. It is possible to automatically set a threshold, for example by filtering out the peaks and setting the threshold at the maximum of the remaining signal. Implementation of this is however also left for future work.

Also in the proposed method, a resolution of 1 degree is used in the beamformer. This chosen resolution proves to be sufficient for the methods to work in the remainder of this thesis, but it is expected that the DOAs can be estimated more accurate when either a smaller resolution is used or when the energy-per-DOA data is interpolated to more accurately estimate the local maxima. Also this would be a relevant subject to give more attention.

Results have shown that when the size of the receiver line-array is halved, localisation in a realistic situation is not possible anymore. The expected cause is the reduced SNR in the beamformer output. To be able to decrease the size of the array, this SNR should be increased in another way. This can for example be achieved by placing more hydrophones over half the length or by averaging out multiple copies of the source signal as was discussed in section 3.11.

8.3. Recommendations

In the previous section, already some aspects of this thesis that require further work to be validated are mentioned. In this section, some recommendations for further improvement of the methods will be added to this list.

Whilst developing the proposed method as in chapter 4, the amplitudes of the arrivals was intentionally not used in the comparison of modeled and measured data. This choice was made based on the fact that BELLHOP could not accurately predict these amplitudes, as the bottom- and surface reflection coefficients are not known. However, it has been illustrated in section 6.2 that exactly environmental parameters as these coefficients can be estimated alongside the range and depth. It is expected that if this is also done with the reflection coefficients, the additional amplitude data of the reflections can be used to further reduce uncertainty in location estimates.

Another environmental parameter that is expected to play a relevant role when improving accuracy is desired, is the sound speed profile. Since the sound-speed profile changes over time, errors might be introduced in the DOA due to different bending of the paths and in the TDOAs due to faster or slower traveling through the paths. In for example [13], the SSP is linearized on three segments and a shift of these linear segments is incorporated as environmental parameters, which improves the localizations accuracy. It would be interesting to see if this same results could be achieved with the situation studied in this thesis.

In chapter 5, Gibbs sampling is chosen over the Metropolis algorithm, since the parameter space that is considered in this work is limited enough to store and work with data for all possible parameter combinations. If other environmental parameters will be taken into account, this will not be possible anymore and a method as Metropolis needs to be used. It would be an interesting addition to also make an implementation using the Metropolis algorithm and compare on for example accuracy and computation time.

Bibliography

[1] D.A. Abrahams. *Underwater Acoustic Signal Processing - Modeling, Detection, and Estimation*. Springer, 2019.

[2] P. Baggenstoss. An algorithm for the localization of multiple interfering sperm whales using multi-sensor time difference of arrival. *The Journal of the Acoustical Society of America*, 130:102–12, 07 2011.

[3] A. B. Baggeroer, W. A. Kuperman, and P. N. Mikhalevsky. An overview of matched field methods in ocean acoustics. *IEEE Journal of Oceanic Engineering*, 18(4):401–424, 1993.

[4] C. Carter, A.H. Nuttal, and P.G. Cable. The smoothed coherence transform. 1973.

[5] J. Choi and H. Choi. Underwater vehicle localization using angular measurements of underwater acoustic sources. In *2015 12th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, pages 235–238, 2015.

[6] C. Clark, R. Charif, S. Mitchell, and J. Colby. Distribution and behavior of the bowhead whale, balaena mysticetus, based on analysis of acoustic data collected during the 1993 spring migration off point barrow, alaska. *Report for the International Whaling Commission*, 46, 01 1996.

[7] M.D. Collins and W. A. Kuperman. Focalization: Environmental focusing and source localization. *The Journal of the Acoustical Society of America*, 90(3):1410–1422, 1991.

[8] M.D. Collins, .T. Fialkowski, W. A. Kuperman, and J.S. Perkins. The multivalued bartlett processor and source tracking. *The Journal of the Acoustical Society of America*, 97(1):235–241, 1995.

[9] P. Corke, C. Detweiler, M. Dunbabin, M. Hamilton, D. Rus, and I. Vasilescu. Experiments with underwater robot localization and tracking. pages 4556–4561, 2007.

[10] C. Debever and W. A. Kuperman. Robust matched-field processing using a coherent broadband white noise constraint processor. *The Journal of the Acoustical Society of America*, 122(4):1979–1986, 2007.

[11] C. Debever and W.A. Kuperman. Robust matched-field processing using a coherent broadband white noise constraint processor. *The Journal of the Acoustical Society of America*, 122(4):1979–1986, 2007.

[12] R. Diamant, H. Tan, and L. Lampe. Los and nlos classification for underwater acoustic localization. *IEEE Transactions on Mobile Computing*, 13(2):311–323, 2014.

[13] S.E. Dosso and M.J. Wilmut. Bayesian focalization: Quantifying source localization with environmental uncertainty. *The Journal of the Acoustical Society of America*, 121(5):2567–2574, 2007.

[14] S.E. Dosso and M.J. Wilmut. Bayesian multiple-source localization in an uncertain ocean environment. *The Journal of the Acoustical Society of America*, 129(6):3577–3589, 2011.

[15] M. Erol-Kantarci, H.T. Mouftah, and S. Oktug. Localization techniques for underwater acoustic sensor networks. *IEEE Communications Magazine*, 48(12):152–158, 2010.

[16] H. W. Frye and J. D. Pugh. A new equation for the speed of sound in seawater. *The Journal of the Acoustical Society of America*, 50(1B):384–386, 1971.

[17] P. Giraudet and H. Glotin. Real-time 3d tracking of whales by echo-robust precise tdoa estimates with a widely-spaced hydrophone array. *Applied Acoustics*, 67:1106–1117, 11 2006.

[18] B.F. Harrison. An l-norm estimator for environmentally robust, shallow-water source localization. *The Journal of the Acoustical Society of America*, 105(1):252–259, 1999.

[19] T.J. Hayward. Information-theoretic analysis of iterated bayesian acoustic source localization in a static ocean waveguide. *The Journal of the Acoustical Society of America*, 137(5):2758–2772, 2015.

[20] P. Hursky, M.B. Porter, M. Siderius, and V.K. McDonald. High-frequency (8–16 khz) model-based source localization. *The Journal of the Acoustical Society of America*, 115(6):3021–3032, 2004.

[21] Y. Jiang and M. R. Azimi-Sadjadi. A robust source localization algorithm applied to acoustic sensor network. In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*, volume 3, pages III–1233–III–1236, 2007.

[22] S.M. Kay. *Fundamentals of Statistical Signal Processing - Detection Theory*. Prentice Hall PTR, 1998.

[23] B. Lambert. *A Student's Guide to Bayesian Statistics*. Sage, 2018.

[24] H. Niu, E. Reeves, and P. Gerstoft. Source localization in an ocean waveguide using supervised machine learning. *The Journal of the Acoustical Society of America*, 142(3):1176–1188, 2017.

[25] E-M. Nosal. Methods for tracking multiple marine mammals with wide-baseline passive acoustic arrays. *The Journal of the Acoustical Society of America*, 134(3):2383–2392, 2013.

[26] E-M. Nosal. *Chapter 8: Model-based marine mammal localization methods. In: Eds. O Adam and F Samaran, Detection Classification and Localization of Marine Mammal using Passive Acoustics – 10 years of progress*. Dirac NGO, 2013.

[27] E-M. Nosal and L. Frazer. Pair-wise spectrogram processing used to track a sperm whale. *Canadian Acoustics - Acoustique Canadienne*, 36:132–138, 03 2008.

[28] M.B. Porter. *The bellhop manual and user's guide: Preliminary draft*. 2011.

[29] M.B. Porter and A. Tolstoy. The matched field processing benchmark problems. *Journal of Computational Acoustics*, 02(03):161–185, 1994.

[30] M.B. Porter, P. Hursky, C.O. Tiemann, and M. Stevenson. Model-based tracking for autonomous arrays. In *MTS/IEEE Oceans 2001. An Ocean Odyssey. Conference Proceedings (IEEE Cat. No.01CH37295)*, volume 2, pages 786–792 vol.2, 2001.

[31] B.P. Rideout and E-M Nosal. Single-hydrophone automated passive acoustic ranging of fin whales at station aloha. *The Journal of the Acoustical Society of America*, 144(3):1958–1958, 2018.

[32] B.P. Rideout, S.E. Dosso, and D.E. Hannay. Underwater passive acoustic localization of pacific walruses in the northeastern chukchi sea. *The Journal of the Acoustical Society of America*, 134(3):2534–2545, 2013.

[33] H. Schmidt, A. B. Baggeroer, W. A. Kuperman, and E. K. Scheer. Environmentally tolerant beamforming for high-resolution matched field processing: Deterministic mismatch. *The Journal of the Acoustical Society of America*, 88(4):1851–1862, 1990.

[34] Lawrence D. Stone, Thomas L. Corwin, and Carl A. Barlow. *Bayesian Multiple Target Tracking*. Artech House, Inc., USA, 1st edition, 1999. ISBN 1580530249.

[35] D. R. Sweet. Statistical modelling, matched-field processing and matched-field inversion at high frequencies. In *1999 Information, Decision and Control. Data and Information Fusion Symposium, Signal Processing and Communications Symposium and Decision and Control Symposium. Proceedings (Cat. No.99EX251)*, pages 193–199, 1999.

[36] Z. Tang, G. Blacquiere, and G. Leus. Aliasing-free wideband beamforming using sparse signal representation. *Signal Processing, IEEE Transactions on*, 59:3464 – 3469, 08 2011. doi: 10.1109/TSP.2011.2140108.

[37] The Stan Development Team. *Stan Modeling Language Users Guide and Reference Manual, version 2.25*. 2019.

[38] E. I. Thorsos, K. L. Williams, N. P. Chotiros, J. T. Christoff, K. W. Commander, C. F. Greenlaw, D. V. Holliday, D. R. Jackson, J. L. Lopes, D. E. McGehee, J. E. Piper, M. D. Richardson, and Dajun Tang. An overview of sax99: acoustic measurements. *IEEE Journal of Oceanic Engineering*, 26(1):4–25, 2001.

[39] D. Tollefsen and S.E. Dosso. Three-dimensional multiple-source focalization in an uncertain ocean environment. *The Journal of the Acoustical Society of America*, 134(5):EL426–EL431, 2013.

[40] A. Tolstoy. Sensitivity of matched field processing to sound-speed profile mismatch for vertical arrays in a deep water pacific environment. *The Journal of the Acoustical Society of America*, 85(6):2394–2404, 1989.

[41] F.T. Ulaby and U. Ravaioli. *Fundamentals of Applied Electromagnetics*. Pearson, 2015.

[42] Z. Wu and X. Li. An improved underwater acoustic network localization algorithm. *China Communications*, 12(3):77–83, 2015.

[43] N. Xiang. Model-based bayesian analysis in acoustics—a tutorial. *The Journal of the Acoustical Society of America*, 148(2):1101–1120, 2020.