UNIVERSITEIT·STELLENBOSCH·UNIVERSITY
jou kennisvennoot • your knowledge partner

The Development and Parameter Optimisation of a Cetacean Localisation Algorithm

Thandi M. Colley
21564027

Report submitted in partial fulfilment of the requirements of the module Project (E) 448 for the degree Baccalaureus in Engineering in the Department of Electrical and Electronic Engineering at Stellenbosch University.

Supervisor: Prof D. J. J. Versfeld
Department of Electrical and Electronic Engineering

November 14, 2021

Acknowledgements

I would like to thank my supervisor, Professor Versfeld, for his support throughout this project. It has been a pleasure to work under his professional guidance and to have the opportunity to learn from and be inspired by his work.

I also extend thanks to my parents, my brother, my partner and the rest of my support structure for their love and encouragement, not only during this period, but throughout my academic career. I appreciate every celebration of success, consolation after disappointment, and distraction when necessary. I am ever grateful for them all.

UNIVERSITEIT·STELLENBOSCH·UNIVERSITY
jou kennisvennoot • your knowledge partner

Plagiaatverklaring / *Plagiarism Declaration*

1. Plagiaat is die oorneem en gebruik van die idees, materiaal en ander intellektuele eiendom van ander persone asof dit jou eie werk is.
   Plagiarism is the use of ideas, material and other intellectual property of another's work and to present is as my own.

2. Ek erken dat die pleeg van plagiaat 'n strafbare oortreding is aangesien dit 'n vorm van diefstal is.
   I agree that plagiarism is a punishable offence because it constitutes theft.

3. Ek verstaan ook dat direkte vertalings plagiaat is.
   I also understand that direct translations are plagiarism.

4. Dienooreenkomstig is alle aanhalings en bydraes vanuit enige bron (ingesluit die internet) volledig verwys (erken). Ek erken dat die woordelikse aanhaal van teks sonder aanhalingstekens (selfs al word die bron volledig erken) plagiaat is.
   Accordingly all quotations and contributions from any source whatsoever (including the internet) have been cited fully. I understand that the reproduction of text without quotation marks (even when the source is cited) is plagiarism

5. Ek verklaar dat die werk in hierdie skryfstuk vervat, behalwe waar anders aangedui, my eie oorspronklike werk is en dat ek dit nie vantevore in die geheel of gedeeltelik ingehandig het vir bepunting in hierdie module/werkstuk of 'n ander module/werkstuk nie.
   I declare that the work contained in this assignment, except where otherwise stated, is my original work and that I have not previously (in its entirety or in part) submitted it for grading in this module/assignment or another module/assignment.

21564027 Studentenommer / *Student number*	Handtekening / *Signature*
T.M. Colley Voorletters en van / *Initials and surname*	November 14, 2021 Datum / *Date*

Abstract

English

The occurrence of Bryde's whales along the coasts of South Africa has, to this day, not yet been accurately assessed. For effective management and conservation of this cetacean, the ability to track and monitor its movement is imperative. Passive acoustic monitoring (PAM) is an efficient method of recording the vocalisations of Bryde's whales; the value in PAM systems lies largely in the post-processing of such recordings.

This project details the development and optimisation of an algorithm that uses PAM recordings from an array of hydrophones to localise the source of significant acoustic events that occur within the recording. Cross-correlation and TDOA principles are applied to gather information from three separate recordings, with which hyperbolic trilateration is performed to produce the GPS coordinates of the source.

The performance of the algorithm is optimised through multiple sets of simulations. Upon optimisation, the algorithm accurately localises the signal source within 1 m. This accuracy is maintained for recordings with a signal-to-noise ratio as low as 0.5 dB.

Afrikaans

Die aanwesigheid van Bryde se walvisse langs die kus van Suid-Afrika is tot dusver nog nie akkuraat nagevors nie. Vir die effektiewe bewaring van hierdie walvis, is die vermoë om hulle op te spoor en hul bewegings te monitor, essensieel. Passiewe akoestiese monitering (PAM) is 'n doeltreffende metode om die vokalisasies van Bryde se walvisse op te neem; die waarde in PAM-stelsels lê grootliks in die na-prosessering van sulke opnames.

Hierdie projek beskryf die ontwikkeling en optimalisering van 'n algoritme wat PAM-opnames van 'n reeks hidrofone gebruik om die bron van beduidende akoestiese gebeure wat binne die opname voorkom, te lokaliseer. Kruiskorrelasie en verskil in aankomstydbeginsels word toegepas om inligting van drie afsonderlike opnames in te samel waarmee hiperboliese trilaterasie uitgevoer word om die GPS-koördinate van die bron te verskaf.

Die betroubaarheid van die algoritme word geoptimaliseer deur etlike stelle simulasies. Met optimalisering lokaliseer die algoritme die bron van 'n sein akkuraat binne 1 m. Hierdie akkuraatheid word gehandhaaf vir opnames met 'n sein-tot-geraas ratio so laag as 0.5 dB.

Contents

Acknowledgements **i**

Abstract **iii**

List of Figures **vi**

List of Tables **vii**

Nomenclature **viii**

1. Introduction **1**
1.1. Background 1
1.2. Problem Statement 2
1.3. Project Objectives 2
1.4. Work Done and Scope 2
1.5. Report Overview 3

2. Literature Review **4**
2.1. Related Work 4
2.1.1. Bryde's Whale Population Study In South Africa 4
2.1.2. Bottom-Mounted PAM Using Correlation Sum Estimation (CSE) . 5
2.1.3. Hydrophone Array PAM Using Teager-Kaiser (TK) Energy Detection 5
2.2. Relevant Concepts 6
2.2.1. Time Difference of Arrival (TDOA) Localisation 6
2.2.2. Hyperbola Theory 8
2.2.3. Measuring Error 10
2.3. Summary 11

3. Methodology **13**
3.1. Overview 13
3.2. Sensor and Whale Coordinates 14
3.2.1. Study Region 14
3.2.2. Actual Arrival Times 15
3.3. Sound Files 15
3.3.1. Vocalisation and Sea Noise Test Set 15

3.3.2. Fabricated Vocalisations . . . 17
3.4. Time Difference of Arrival (TDOA) . . . 17
3.4.1. Auto-Correlation (Energy Detection) . . . 18
3.4.2. Cross-Correlation . . . 19
3.5. Hyperbolic Trilateration . . . 20
3.6. Simulation Setup . . . 22
3.6.1. Simulation A: Sensor and Whale Coordinates . . . 23
3.6.2. Simulation B: Signal-to-Noise Ratio (SNR) . . . 23
3.6.3. Simulation C: Correlation Window Length . . . 24
3.7. Software Used . . . 24
3.7.1. Audacity . . . 24
3.7.2. Python . . . 24
3.8. Summary . . . 25

4. Simulation Results **26**
4.1. Simulation A: Sensor and Whale Coordinates . . . 26
4.2. Simulation B: Signal-to-Noise Ratio (SNR) . . . 27
4.3. Simulation C: Correlation Window Length . . . 28
4.4. Simulation D: Final Test . . . 30
4.5. Summary . . . 31

5. Conclusion **32**
5.1. Objectives . . . 32
5.2. Algorithm Design and Testing . . . 32
5.3. Future Work . . . 33
5.4. Conclusion . . . 33

Bibliography **34**

A. Project Planning Schedule **37**

B. ECSA Compliance **38**

C. GitHub Repository **40**

List of Figures

2.1. Geometric parameters of a hyperbola. 8
2.2. Geometric parameters of a hyperbola as applied to TDOA localisation. . . 9
2.3. Hyperbolic trilateration example. 10

3.1. Iterative design and testing process. 13
3.2. Study region. 14
3.3. Example sensor and source coordinates. 15
3.4. Amplitude visualisation of f_{voc}. 16
3.5. Test set source files. 16
3.6. Fabricated vocalisation files. 17
3.7. Amplitude visualisation of f_{0_j}. 18
3.8. Cross-correlation of f_1 and f_2 with f_{0_j}. 19
3.9. System visualisation with test set parameters. 21
3.10. Simulation A test region. 23

4.1. RMSE produced by Simulation A. 26
4.2. RMSE produced by Simulation B.1. 27
4.3. RMSE produced by Simulation B.2. 27
4.4. RMSE produced by Simulation C.1. 28
4.5. RMSE produced by Simulation C.2. 29
4.6. RMSE produced by Simulation D. 30

A.1. Project planning Gantt chart. 37

List of Tables

3.1. Intermediary parameter set for example data. 21
3.2. Base parameter set. 22

4.1. Optimal parameter ranges. 31

A.1. Project planning schedule. 37

B.1. ECSA exit level outcomes compliance. 38

Nomenclature

Acronyms and abbreviations

2D	Two-dimensional.
3D	Three-dimensional.
CSE	Correlation sum estimation.
GPS	Global positioning system.
HARP	High-frequency audio recording package.
MARU	Marine autonomous recording unit.
PAM	Passive acoustic monitoring.
RMSE	Root mean square error.
SNR	Signal-to-noise ratio.
SST	Sea surface temperature.
TDOA	Time-difference of arrival.
TK	Teager-Kaiser.
TOA	Time of arrival.
wav	Waveform audio.
WGN	White Gaussian noise.

Variables and functions

a, b, d	Hyperbolic parameters.
d_S	Distance between sensors.
d_W	Distance between sensor and source.
$\hat{d}_W$	Estimated distance between sensor and source.
$\Delta\hat{d}_W$	Estimated difference of distance between sensor and source.
e	Error.
f_{sea}	Sound file containing sea noise.
f_{voc}	Sound file containing vocalisation.
f	Fabricated vocalisation recording.
i	Correlation window.
k	Correlation lag.
$R_{f_n:f_n}$	Auto-correlation.
$R_{f_n:f_m}$	Cross-correlation.
m	Gradient.
q	Number of samples.
$\mathbf{S}$	A set of sensor coordinates.
SNR	Signal-to-noise ratio (SNR).
θ	Angle of rotation.
t	Arrival time.
$\hat{t}$	Estimated arrival time.
Δt	Time-difference of arrival (TDOA).
$\Delta\hat{t}$	Estimated TDOA.
t_{win}	Correlation window length.
v	Speed of sound in water.
$\mathbf{W}$	A set of source coordinates.
x	Latitude.
y	Longitude.

Chapter 1

Introduction

1.1 Background

The Bryde's whale (*Balaenoptera edeni*) is a critically understudied species of cetacean. Until as recently as 2008, it was classified as "data deficient" by the International Union for Conservation of Nature and Natural Resources (IUCN) [1]. To this day, its population and geographic range is still largely unresolved, with the IUCN calling for research on the Bryde's whale's taxonomy, population size, distribution and trends, threats, and population monitoring.

For effective wildlife management and conservation, it is fundamentally important to assess the occurrence, range, distribution, density and habitat selection of the species of interest [2]. Visual methods of cetacean monitoring, such as vessel survey, diving experiments and capture-related activity, are mostly inexpensive and traditionally common practice [3]. However, such visual methods can be largely inefficient and produce imperfect data. Visual sightings are limited by weather conditions, range of visibility, availability of daylight and the tendency of marine species to approach the surface. The Bryde's whale, in particular, exhibits unpredictable behaviour in that it breaks above the surface at irregular intervals and tends to change its direction of travel spontaneously [4].

Passive acoustic monitoring (PAM) is an indirect survey method which uses the vocalisations of animals to determine population densities. Since the Bryde's whale produces low-frequency pulsed calls [5], it is an ideal candidate for PAM survey. There are a variety of techniques that can be used to implement PAM. One example of such a technique is the implementation of multiple-hydrophone arrays that collect acoustic recordings over long periods of time for post-processing and examination. These arrays can be used as a standalone population survey method, or in conjunction with visual sightings to contextualise the data collected [6]. The efficacy of PAM depends largely on the post-processing performed on the raw acoustic data.

To assess the movement and behaviour of the Bryde's Whale along South African coasts, an array of PAM hydrophones can be deployed with the intention of recording the whales' vocalisations. Given the GPS coordinates of each hydrophone, time-difference of arrival (TDOA) techniques can be used to localise the whale at each vocalisation instance.

The development of such a TDOA localisation algorithm can contribute significantly to existing efforts to monitor the occurrence of Bryde's whales along the south coast of South Africa.

1.2 Problem Statement

The aim of this research project is to develop and optimise an algorithm which localises the source of an underwater acoustic event, particularly the vocalisation of an aquatic mammal, using TDOA principles. This algorithm will be used to track and monitor the movement of Bryde's whales.

1.3 Project Objectives

The algorithm developed in this project should return the GPS coordinates of a cetacean given synchronised vocalisation recordings from a distributed array of underwater sensors and their GPS coordinates.

The accuracy of algorithm should be optimised by investigating the effect of changing the values of all variable parameters. The variables of interest are:

- the coordinate values of each sensor and the vocalisation source;
- the signal-to-noise ratio (SNR), in dB, of the acoustic recording;
- the length, in seconds, of the correlation window used to detect the vocalisation within each recording.

The algorithm should accurately localise the source within 10 m. It should account for reasonable levels of noise within the recording and for drifting of the sensors.

1.4 Work Done and Scope

This project follows the development of a two-dimensional TDOA localisation algorithm for underwater acoustic events using hyperbolic trilateration. The algorithm is developed for an array of three floating hydrophones located within a 10 km^2 region, that produce separate synchronised audio recordings at a sample rate of 96 kHz. The algorithm could be scaled to include additional sensors, expand the range of monitoring or sample audio at a different rate, though these adaptations are not tested.

Test data is created with which to measure the performance of the algorithm. The test data uses one example vocalisation with which to fabricate sensor output. The test data files, to which noise is added, are 10 seconds long and contain a single vocalisation each. Signal attenuation is not accounted for in the creation of test data.

Piece-wise cross-correlation is used to detect energy peaks within recording snippets to determine the TDOA for each sensor. A sliding correlation window is used, at an overlap of half the window length. Algorithm performance related to correlation window overlap length is not investigated. Optimal length of the correlation window itself, however, is considered during testing.

Three sets of simulations are completed, using the test data, to analyse the effect of certain variables on the performance of the algorithm. The performance is measured by calculating the localisation RMSE, in meters, of each simulation. Computer processing performance is not measured during testing. After the simulations are complete, a range of values that will produce accurate results is defined for each variable.

Field testing of the localisation algorithm was not performed due to time constraints.

1.5 Report Overview

In Chapter 2, a selection of previous work relating to the project is detailed. This chapter also gives thorough descriptions of the theoretical concepts used throughout the report. Chapter 3 covers the creation of test data, the development of the localisation algorithm, and explains the context and setup for the test simulations. In Chapter 4, the results of each simulation are revealed, discussed and interpreted. Finally, Chapter 5 concludes the report and highlights recommendations for future work.

Chapter 2

Literature Review

2.1 Related Work

2.1.1 Bryde's Whale Population Study In South Africa

In response to the Bryde's whale's classification of "data deficient" by the International Union of Conservation of Nature (IUCN) [1], Penry et al. [7] conducted a study into the inshore population of Bryde's whales along the south-eastern coast of South Africa. This study sought to determine seasonal movement patterns within the region, define the environmental factors that affect population density and feeding activity, and investigate the whales' interactions with other species.

Along approximately 50 km of the coastline of Plettenberg Bay and Nature's Valley, data was collected almost daily between November 2005 and June 2008. Observers boarded whale-watching vessels and recorded visual sightings of Bryde's whales while completing a standard trip. Biopsy samples, photo-identification and feeding activity were routinely documented, in addition to digitally plotting the GPS coordinates of each sighting. Environmental variables such as sea surface temperature (SST), chlorophyll *a* concentration, and wind speed were also recorded for this period.

Throughout the duration of the study, 146 Bryde's whales were observed. There was a significant monthly sighting variation, with April having the highest rate of encounters. After modelling the encounter rate and the aforementioned environmental variables, a number of relations were discovered. Encounter rate showed a direct correlation to SST, an indirect correlation to wind speed, and poor correlation to chlorophyll *a* concentrations. Another noteworthy observation was the connection between social group size of Bryde's whales and an increase in feeding activity. In terms of multispecies interactions, it was reported that Bryde's whales have a tendency to associate with Cape gannets, Cape fur seals and common dolphins.

The chosen method of data collection for this study of using whale-watching vessels is commonplace and relatively simple to execute. Unfortunately, the results produced by such research can be incomplete and biased. This is because of the time and spatial limitations posed by vessel trips, and imprecise detection by observers.

2.1.2 Bottom-Mounted PAM Using Correlation Sum Estimation (CSE)

In 2009, Stanistreet, Risch and Van Parijs [8] performed a PAM study in the Gulf of Maine with the intention of tracking the movement of humpback whales (*Megaptera novaeangliae*) within a certain feeding ground (the Stellwagen Bank National Marine Sanctuary, or SBNMS).

The study used arrays of ten Marine Autonomous Recording Units (MARUs), developed by researchers at Cornell University, to capture continuous synchronised acoustic recordings. The MARU is a pressurised glass unit containing a hydrophone, computer, power source and hard drive, and was bottom-mounted (i.e. fixed on the seabed) upon use [9]. Each unit was programmed to produce recordings at a 2 kHz sampling rate with a resolution of 12 bits.

Localisation of individual humpback whales was done using correlation sum estimation (CSE) on selected vocalisation recordings by the MARU array. CSE is an iterative algorithm that determines the cross-correlation of each pair of recordings for a set of coordinate points and accumulates them. The coordinate point that produces the largest sum value represents the position of interest, i.e. the latitude and longitude of the whale at the moment of vocalisation [10]. CSE differs from standard hyperbolic multilateration methods of localisation in that it does not require known information on the TDOA of the peak cross-correlation point. The CSE method is considered to be less susceptible to background noise and clutter for this reason.

Various useful results in the field of marine mammal tracking were produced by this study. High-density presence of the humpback whale was observed in the SBNMS during the spring (March – June) and autumn (August – December) seasons, with peaks in mid-April and late-November. This confirmed that there is a consistent annual pattern of humpback whale presence in this area. The tracked whales showed numerous degrees of movement while vocalising: distances from 0.69 km to 20.13 km were travelled while vocalising. Furthermore, speed of movement ranged from 0.54 km/h to 6.05 km/h, with an average speed of 2.06 km/h.

The researchers concluded that PAM is an effective strategy for cetacean population study. However, this method is limited in its dependence on vocal activity of individuals. This restricts our holistic understanding of the social behaviour of whales, as female or non-vocal male individuals will not be accounted for.

2.1.3 Hydrophone Array PAM Using Teager-Kaiser (TK) Energy Detection

Gassmann et al. [11] developed and tested a method for 3D tracking of odontocete cetaceans that produce vocalisations at higher-than-average frequencies than baleen cetaceans. Such vocalisations, like those of a dolphin or beaked whale [12], attenuate at a higher rate over distance than those of larger whales, and so the need arose for improved detection methods.

The experimental setup used to assess the tracking method consisted of one array of large-aperture high-frequency audio recording packages (HARPs) and three arrays of small-aperture hydrophones. The arrays were calibrated by tracking the noise made by a vessel with known coordinates.

During the study period, the vocalisations of a group of five killer whales (*Orcinus orca*) were detected and monitored. Echolocation click vocalisations were automatically identified in the recordings using an energy measurement technique called Teager-Kaiser (TK) energy measurement. The TK energy operator is defined as follows:

$$\Psi(x[n]) = x^2[n] - x[n+1]x[n-1], \tag{2.1}$$

found in [13]. The TK energy operator allows for near-instant energy detection, since only three samples are required to determine the energy value at each instant.

Cross-correlation of the vocalisations recorded by each hydrophone was then done to determine the TDOA Δt of each individual vocalisation. These TDOA values were used to determine the arrival angle β of the sound wave at selected hydrophone pairs using Equation 2.2:

$$\beta = \cos^{-1}\left(\frac{v \times \Delta t}{d}\right), \tag{2.2}$$

where v is the speed of sound in water and d is the distance between hydrophones. The value for β, in conjunction with Δt and horizontal bearing adjustment, was used to localise each detected sound source. Sets of TDOAs (and by extension, localised positions), were assigned to individual killer whales by cross-correlating the recordings from different hydrophones.

The study proved successful in its endeavour to develop a 3D tracking method for echolocation clicks and high-frequency calls. A group of five killer whales was tracked for a total period of 8 minutes, as they traveled through the detection range of the hydrophone sensors. The tracking data was sufficient for calculating the speed of movement of each whale, and comparing the source level for different vocalisation types. The researchers concluded that the method developed could be used to study and track other small cetaceans. In particular, their recommendation was to used the tracking method alongside visual tracking to gain a better contextual understanding of the cetaceans' behaviour.

2.2 Relevant Concepts

2.2.1 Time Difference of Arrival (TDOA) Localisation

Signals detected by an array of sensors can be accurately localised using the time difference of arrival (TDOA) [14] method. This method comprises of two steps. Firstly, the TDOA

values Δt between individual sensors are determined using the linear velocity equation:

$$v = \frac{d}{t} = \frac{\Delta d}{\Delta t}, \tag{2.3}$$

where v is the speed of the signal in the propagation material and d is the distance traveled. Secondly, the TDOA values are substituted into a set of hyperbolic equations, which, when solved, return the estimated position of the source signal. This method assumes that the signal speed in the propagating material is known and constant, the effects of wind/current or other environmental movement are negligible, and that the sensor error is insignificant [15].

For a two-dimensional system with three sensors $S_{0:2}$ and a signal source W, the coordinates x_W and y_W of W can be found if the sensor coordinates $x_{S_{0:2}}$ and $y_{S_{0:2}}$ are known, as well as the TDOA values $\Delta t_{0:2}$. Equations 2.4a – 2.4c are used to do so:

$$d_{W_0} = v \times t_0 = \sqrt{(x_W - x_{S_0})^2 + (y_W - y_{S_0})^2}; \tag{2.4a}$$

$$d_{W_1} = v \times t_1 = \sqrt{(x_W - x_{S_1})^2 + (y_W - y_{S_1})^2}; \tag{2.4b}$$

$$d_{W_2} = v \times t_2 = \sqrt{(x_W - x_{S_2})^2 + (y_W - y_{S_2})^2}; \tag{2.4c}$$

where d_{W_n} is the distance between S_n and W, and t_n is the time taken for each sensor to receive the signal. This arrival time t_n can also be expressed as in Equations 2.5a – 2.5c:

$$t_0 = t_0 + \Delta t_0, \therefore \Delta t_0 = 0, \tag{2.5a}$$

$$t_1 = t_0 + \Delta t_1, \tag{2.5b}$$

$$t_2 = t_0 + \Delta t_2, \tag{2.5c}$$

where we use the arrival time of S_0 as a reference for the other sensors. In order to eliminate t_0 and solve for $\Delta t_{1:2}$, the above equations can be combined to form Equations 2.6a – 2.6c [16]:

$$d_{S_{0:1}} = d_{W_0} - d_{W_1} = v(t_0 + \Delta t_0) - v(t_0 + \Delta t_1) = v(\Delta t_0 - \Delta t_1), \tag{2.6a}$$

$$d_{S_{1:2}} = d_{W_1} - d_{W_2} = v(t_0 + \Delta t_1) - v(t_0 + \Delta t_2) = v(\Delta t_1 - \Delta t_2), \tag{2.6b}$$

$$d_{S_{2:0}} = d_{W_2} - d_{W_0} = v(t_0 + \Delta t_2) - v(t_0 + \Delta t_0) = v(\Delta t_2 - \Delta t_0). \tag{2.6c}$$

The definition of $d_{S_{n:n+1}}$ is shown in the following equation; it is essentially the equation of a hyperbola with the y-axis being its baseline [17]:

$$d_{S_{n:n+1}} = \sqrt{(x_W - x_{S_n})^2 + (y_W - y_{S_n})^2} - \sqrt{(x_W - x_{S_{n+1}})^2 + (y_W - y_{S_{n+1}})^2}. \tag{2.7}$$

We now have sufficient information to construct a hyperbola such that the location of W can be anywhere on its locus. By determining the formulae for all three hyperbole and identifying their intersection, the coordinates of W at the instance of signal emission can be found. Hyperbola theory explains why this is possible.

2.2.2 Hyperbola Theory

A hyperbola has the geometric property that every point (x, y) on its graph is at the same constant difference $|d_1 - d_2|$ from its two focus points $\pm d$ [18]. Equation 2.8 defines this property:

$$|d_1 - d_2| = 2a. \tag{2.8}$$

Equation 2.9 is the general formula of a hyperbola:

$$\frac{y^2}{a^2} - \frac{x^2}{b^2} = 1. \tag{2.9}$$

Figure 2.1 illustrates such a hyperbola and shows where the constant hyperbolic parameters a, b and d come from.

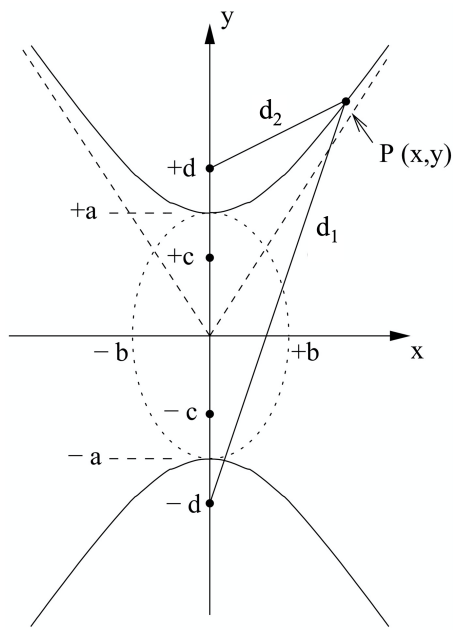


Figure 2.1: Geometric parameters of a hyperbola [17].

Equations 2.10 and 2.11 define the remaining relationships required to plot the hyperbola:

$$a^2 + b^2 = d^2, \tag{2.10}$$

$$y = \pm \left(\frac{a}{b}\right). \tag{2.11}$$

Hyperbola Theory Applied to Localisation

We can now approach TDOA localisation with an understanding of hyperbola theory to perform hyperbolic trilateration. The variables from Figure 2.1 are populated with the previously-determined TDOA parameters, as seen in Figure 2.2.

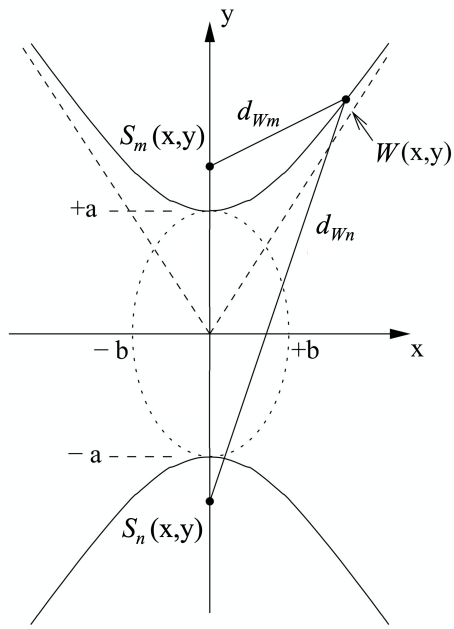


Figure 2.2: Geometric parameters of a hyperbola as applied to TDOA localisation.

First, the coordinates x, y_S of a pair of sensors $S_{n,m}$ are placed at the locations of $\pm d$, the foci of the hyperbola. This allows us to determine $\pm d$ as follows:

$$d = \frac{d_{S_{n:m}}}{2}. \tag{2.12}$$

Next, the sensor-to-source distances $d_{W_{n,m}}$ replace $d_{1,2}$ in Equation 2.8, which can be reworked to give us the following relationship:

$$a = \frac{|d_{W_n} - d_{W_m}|}{2} = \frac{|\Delta d_n - \Delta d_m|}{2}. \tag{2.13}$$

The final hyperbolic parameter b can now be determined in order to plot the hyperbola representing the potential source location W. If this is done for each pair of sensors

in a 2D array of three sensors, the hyperbole will intersect at two locations, giving us the solutions to our hyperbolic trilateration problem. An example of such a hyperbolic trilateration is shown in Figure 2.3.

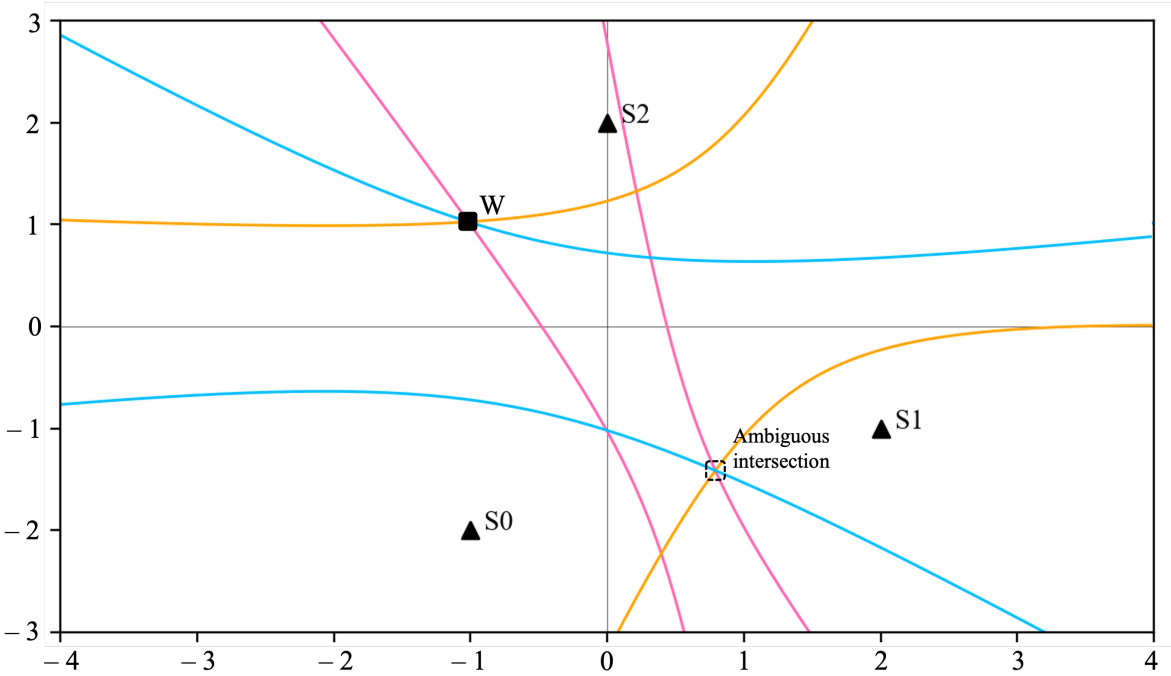


Figure 2.3: Hyperbolic trilateration example.

This method produces ambiguity in that there are two points where all three hyperbole intersect, not only once at the source location W. This ambiguity can be removed by adding a fourth sensor S_3 to the system [19], recording bearing measurements at one or more sensors, or by using the absolute sensor-to-source detection times to determine which sensor is closest to the source, and eliminating the furthest intersection as a source position [20].

2.2.3 Measuring Error

We can define error e_n as the difference between a true observation value z_n and the predicted observation value $\hat{z}_n$. Succinctly, $e_n = \hat{z}_n - z_n$. Below are some common metrics used to measure performance by calculating error values as explained by Vandeput in [21].

Mean Absolute Percentage Error (MAPE)

MAPE accumulates a percentage value for each observation by dividing the error by the true observation value, and then determines the average percentage by dividing the sum by the number of observations i:

$$\text{MAPE} = \frac{1}{i} \sum_{n=1}^{i} \frac{|e_n|}{z_n}. \tag{2.14}$$

MAPE is most useful for measuring the performance of data that contains few outliers and no zeros [22].

Mean Absolute Error (MAE)

MAE calculates the average absolute error for all observations:

$$\text{MAE} = \frac{1}{i}\sum_{n=1}^{i} |e_n|. \tag{2.15}$$

This method of measuring error is not very susceptible to outliers. MAE optimisation will target the median value of the true data [21].

Mean Squared Error (MSE)

MSE calculates the average squared error for all observations:

$$\text{MSE} = \frac{1}{i}\sum_{n=1}^{i} {e_n}^2. \tag{2.16}$$

MSE is more susceptible to outliers than MAE. It is not scaled to the data, which can make it challenging to interpret.

Root Mean Squared Error (RMSE)

RMSE is like MSE except that it takes the square root of the average squared error:

$$\text{RMSE} = \sqrt{\frac{1}{i}\sum_{n=1}^{i} {e_n}^2}. \tag{2.17}$$

RMSE leaves us with an error measurement that is scaled to the data, allowing for easy interpretation. It can help determine the line of best fit for a series of data [22]. Optimisation of RMSE will target the mean value of the true data [21], unlike the MAE which will target the median.

2.3 Summary

This chapter delivers insight into existing solutions for cetacean tracking, and provides an understanding of the theoretical concepts needed to proceed in the development of our own localisation algorithm. The figures and equations depicted in this chapter are referred to throughout Chapter 3.

The background knowledge gained in Section 2.1 allows us to make certain decisions regarding the methodology used in the remainder of the project. Firstly, our confidence in PAM as a device for accurate tracking is established. Cross-correlation techniques are

shown to be effective for use in tangent with PAM. For this reason, general cross-correlation techniques are used for call detection in our algorithm implementation.

Section 2.2 provides the mathematical basis of our TDOA localisation algorithm. The role played by hyperbola theory in this localisation method is also explained. The final subsection presents a selection of error measurement metrics with which to evaluate the performance of our algorithm. RMSE has been chosen to do so, as it is simple to interpret and is efficient in directing us towards the parameter values that produce the highest localisation accuracy.

Chapter 3

Methodology

3.1 Overview

This chapter discusses the preparation and experimental setup required to optimise our localisation algorithm. The aim is to arrive at a modular system which can receive 2D position data and acoustic recordings from an array of sensors, and return an accurate source location of any acoustic event detected within the set of recordings.

An example set of input data is used to test the localisation algorithm at each step of its design. This test set is manually created so as to most closely resemble realistic system input, and the process used to do so is detailed in this chapter.

As the localisation algorithm is developed, certain variable values are selected on an estimation basis, with the intention of revisiting once optimisation testing starts. For this reason, the design and testing process is iterative. Figure 3.1 illustrates this process.

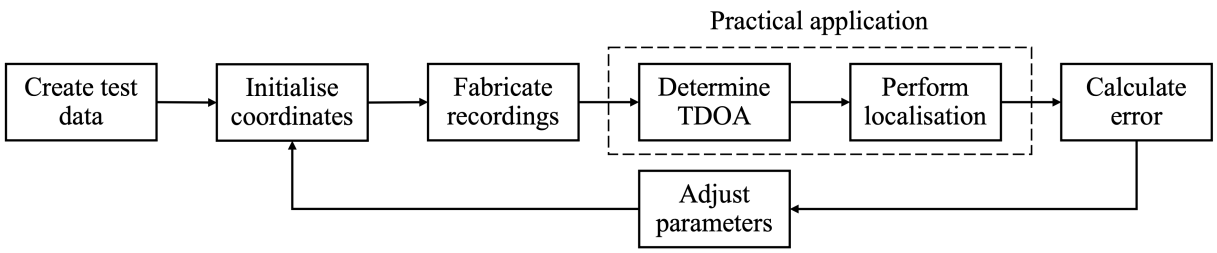


Figure 3.1: Iterative design and testing process.

As mentioned in Chapter 1, the variable values we investigate in detail are:

- the coordinate values of each sensor and the vocalisation source;
- the signal-to-noise ratio (SNR), in dB, of the acoustic recording;
- the length, in seconds, of the correlation window used to detect the vocalisation peak within each recording.

By varying these values and performing thorough testing, the effect each variable on the accuracy of the localisation algorithm can be better understood. Firstly, however, the experimental environment required for this testing needs to be created. The following sections offer detailed descriptions of how this is done.

3.2 Sensor and Whale Coordinates

3.2.1 Study Region

Since physical testing of the localisation algorithm is outside the scope of this project, the study region is practically inconsequential. It is, however, necessary to use geographical coordinates in simulation, and so a real location is chosen to represent the study region. The region of interest is False Bay, South Africa, seen in Figure 3.2. False Bay is the site of existing research into the presence of Bryde's whales along the coast of South Africa [23].

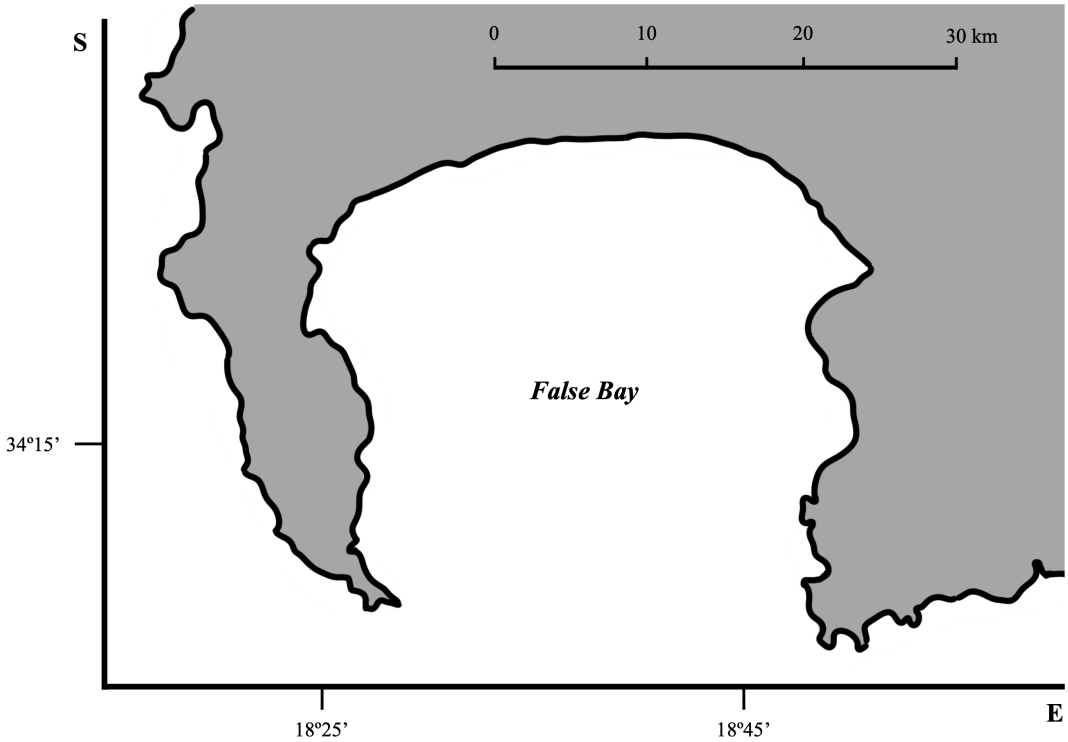


Figure 3.2: Study region.

The simulation positions chosen for the sensor units S_n and vocalisation source W are shown in Figure 3.3. These positions are situated within the same $10\,\text{km}^2$ region near Gordon's Bay, along the coast of False Bay. The corresponding coordinates are:

$$S_0 = [x_{S_0}, y_{S_0}] = [18.7542711777, -34.1458751443];$$
$$S_1 = [x_{S_1}, y_{S_1}] = [18.6936017061, -34.1843165453];$$
$$S_2 = [x_{S_2}, y_{S_2}] = [18.6819537842, -34.1011518069];$$
$$W = [x_W, y_W] = [18.7152089095, -34.1701228282].$$

These are the example coordinate values that are used to design and test the localisation algorithm.

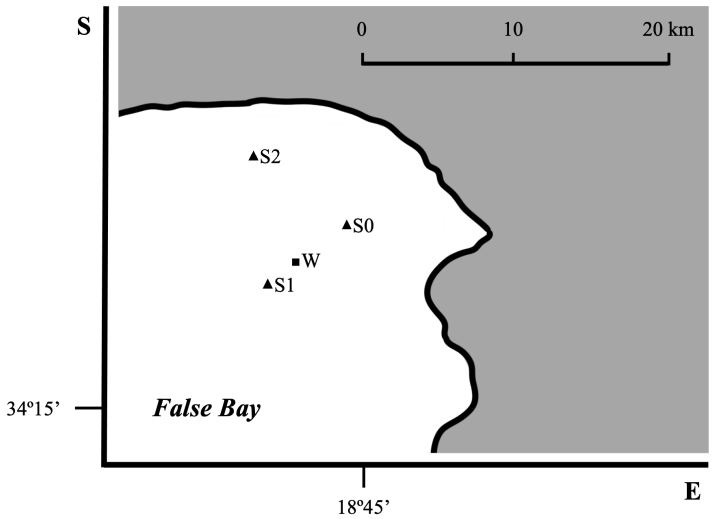


Figure 3.3: Example sensor and source coordinates.

3.2.2 Actual Arrival Times

From the sensor and source locations, x, y_{S_n} and x, y_W, the actual arrival times t_n of the vocalisation to each sensor can be calculated. First, the distance d_{W_n} between each sensor and the source are determined using Equation 2.7. These distances are:

$$d_{W_0} = 5.11232019 \text{ km}; \qquad d_{W_1} = 2.87462617 \text{ km}; \qquad d_{W_2} = 8.51415213 \text{ km}.$$

Given the speed of sound in water $v = 1.5\,\text{km/s}$, we can then obtain the arrival times using Equation 3.1, adapted from Equation 2.3:

$$t_n = \frac{d_{W_n}}{v}. \tag{3.1}$$

In this case, the arrival times t_n are found to be:

$$t_0 = 3.40821346 \text{ s}; \qquad t_1 = 1.91641745 \text{ s}; \qquad t_2 = 5.67610142 \text{ s}.$$

3.3 Sound Files

3.3.1 Vocalisation and Sea Noise Test Set

In order to test the accuracy of the TDOA and localisation algorithm, we require sound files that mimic the features that would be present in the system from Section 3.2. A single existing PAM recording containing Bryde's whale vocalisations is used to create a test set of sound files. The recording consists of an hour-long waveform audio (wav) file, which has a sample rate of 96 kHz and 24-bit resolution.

The aim is to fabricate three vocalisation files which can be used to test the localisation algorithm. For this, a single vocalisation instance is required, as well as three instances of

generic ocean noise. Manual inspection of the original recording is done in order to extract a single vocalisation f_{voc}, seen in Figure 3.4, and three 10-second instances of ocean noise f_{sea_n}, seen in Figures 3.5a, 3.5b and 3.5c. A recording length of 10 seconds is selected to be long enough to allow for a wide range of d_{W_n} distances, but also short enough to reduce strain on the processor during the performance testing stage.

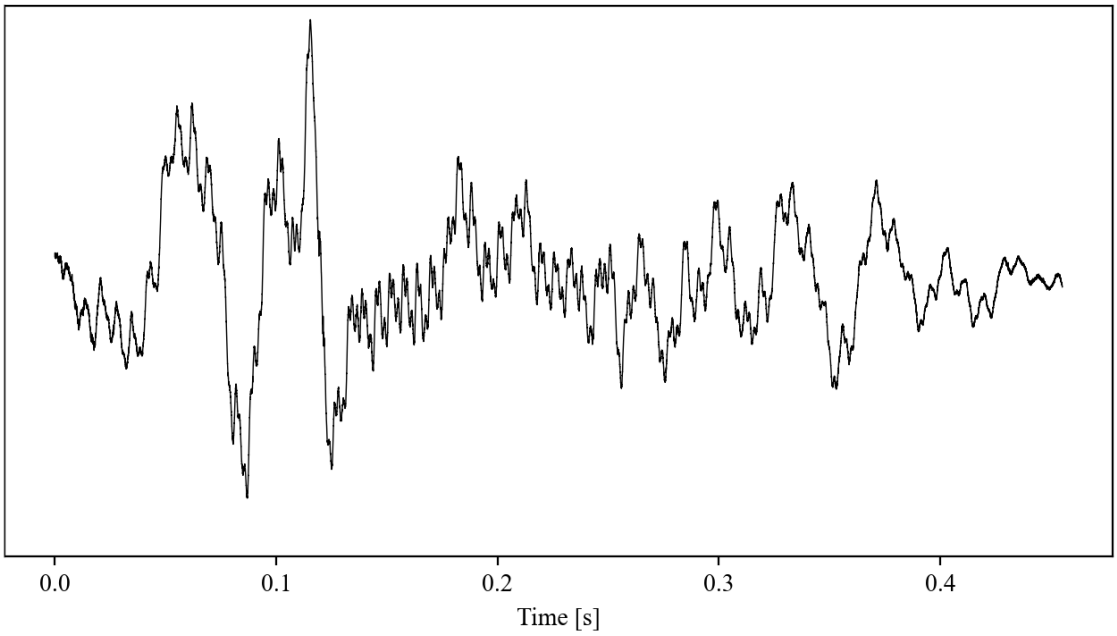
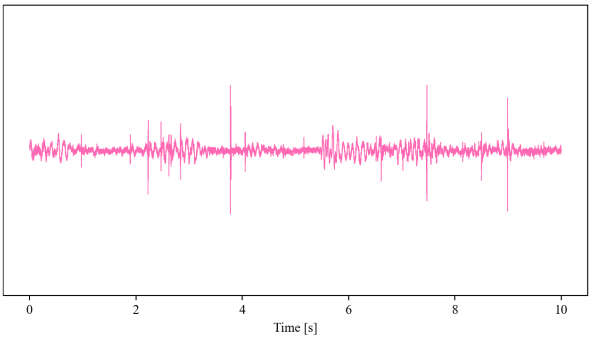
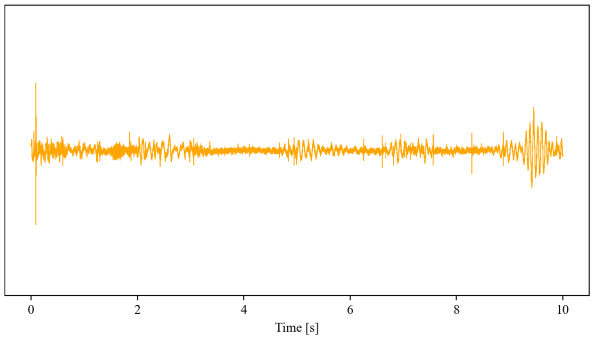


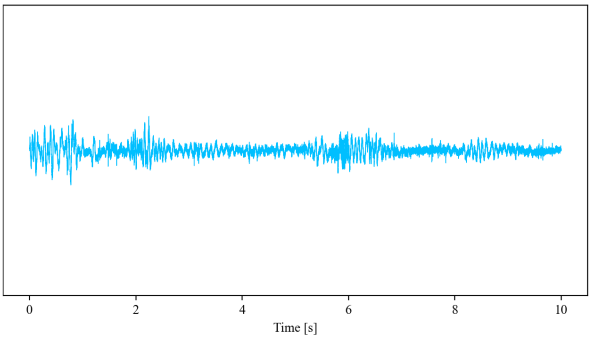
Figure 3.4: Amplitude visualisation of f_{voc}.



(a) Amplitude visualisation of f_{sea_0}.



(b) Amplitude visualisation of f_{sea_1}.



(c) Amplitude visualisation of f_{sea_2}.

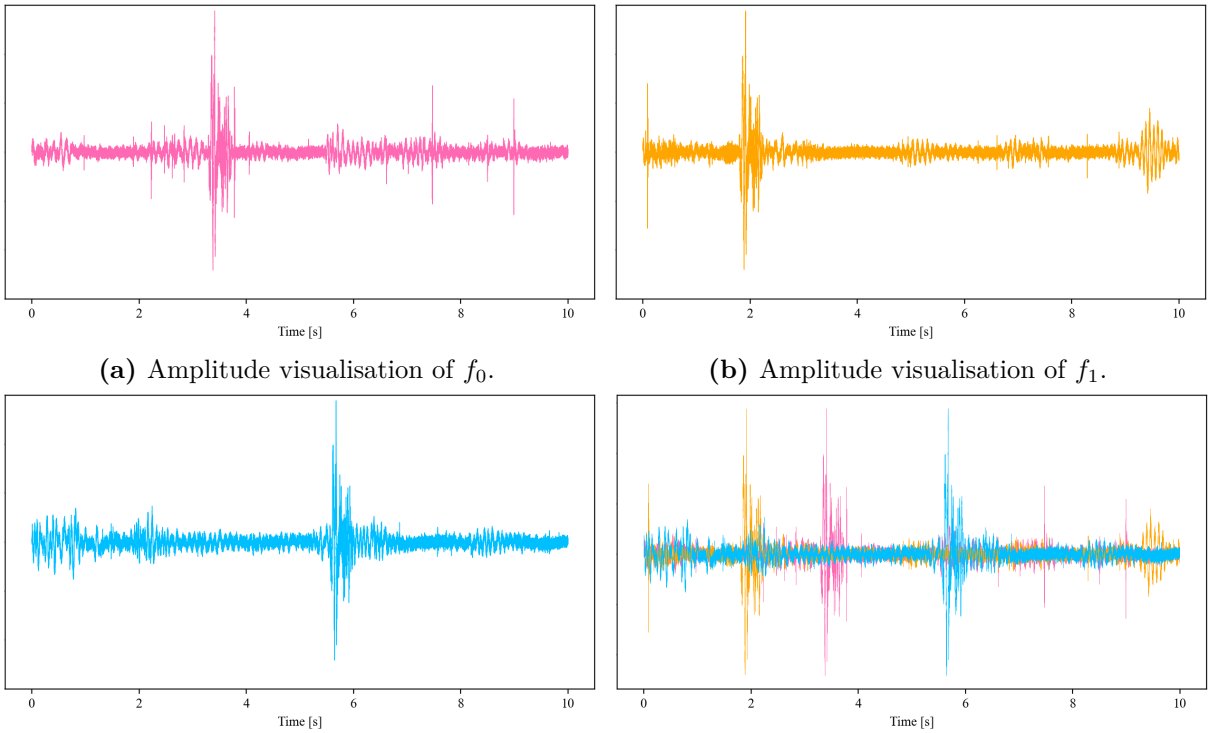
Figure 3.5: Test set source files.

3.3.2 Fabricated Vocalisations

The source files from Subsection 3.3.1 are used to fabricate the vocalisation files required to test the localisation algorithm.

Firstly, the vocalisation instance f_{voc} is inserted into each noise file f_{sea_n} at the appropriate arrival time t_n. White Gaussian noise (WGN) is then applied to the new file in order to simulate practical discrepancies in the detected vocalisation, and allow for more control over the signal-to-noise ratio (SNR). The SNR value is variable and its effect on the accuracy of the localisation algorithm is examined in Chapter 4.

Each fabricated file f_n acts as a recording snippet of a single vocalisation that may have been received by a hydrophone in the array. Figure 3.6 depicts f_0, f_1 and f_2 at $SNR = 100$.



(a) Amplitude visualisation of f_0.

(b) Amplitude visualisation of f_1.

(c) Amplitude visualisation of f_2.

(d) Amplitude visualisation of f_0 to f_2.

Figure 3.6: Fabricated vocalisation files.

3.4 Time Difference of Arrival (TDOA)

The estimated TDOA of the vocalisation to each sensor in the array $\Delta\hat{t}_n$ (not to be confused with the actual TDOA Δt_n) is determined by investigating the fabricated sound files f_n. By visual inspection of Figure 3.6d, it is clear that there is a time difference between the energy peaks of the respective vocalisation detections. The time difference values can be estimated to varying degrees of accuracy by performing piece-wise cross-correlation of the files and comparing the energy peak locations.

3.4.1 Auto-Correlation (Energy Detection)

Recording snippet f_0 is chosen as our benchmark vocalisation, allowing us to use t_0 as a reference from which to measure each estimated TDOA value $\Delta\hat{t}_n$. Piece-wise auto-correlation is performed on f_0, in order to find the energy peak of the fabricated vocalisation and record its time delay from the start of the file. This estimated delay is $\hat{t}_0$; if the auto-correlation is accurate, it should be very close to t_0. The accuracy of the energy detection has a direct relationship with the correlation window length t_{win} used to step through the sound file. The window length t_{win} is variable and its effect on the performance of the localisation algorithm is investigated in Chapter 4.

Equation 3.2, adapted from [24], shows how the auto-correlation $R_{f_0:f_0}$ of each window of f_0 is calculated.

$$R_{f_0:f_0}[i] = \sum f_0[i] f_0[i+k], \tag{3.2}$$

where i is the window and the lag k is a multiple of the window length [25]. Note that the algorithm has been developed such that consecutive windows overlap each other with half of the window length t_{win}.

Figure 3.7 depicts the snippet from f_0 containing the energy peak window $i = 6\,816$, at which $R_{f_0:f_0}$ is the highest, for a window length $t_{win} = 1.0 \times 10^{-3}$ s. This window, which is now referred to as f_{0_j}, represents the loudest point in the vocalisation recording. It is extracted from f_0 and used to determine the cross-correlation in the next step.

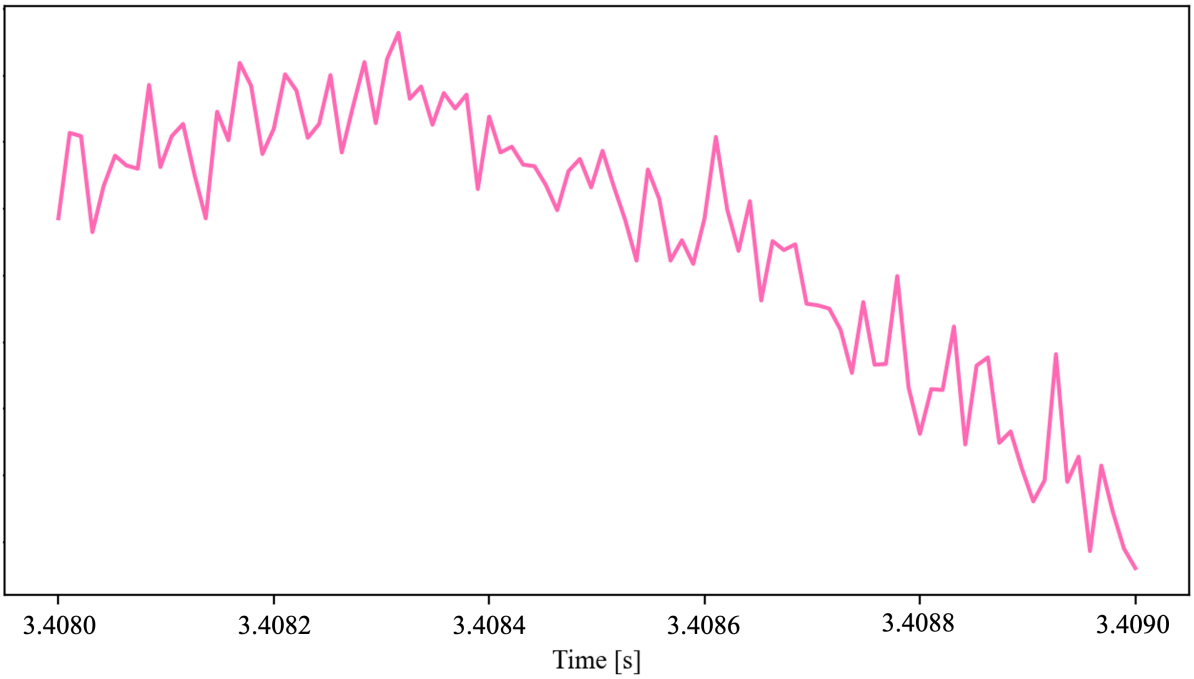


Figure 3.7: Amplitude visualisation of f_{0_j}.

3.4.2 Cross-Correlation

Since f_0 is the benchmark vocalisation, its TDOA value is set to $\Delta\hat{t}_0 = 0$ s. The estimated TDOA values $\Delta\hat{t}_n$ for f_1 and f_2 are determined by finding the estimated time delay $\hat{t}_n$ from the start of its associated sound file, and subtracting $\hat{t}_0$ from it, in accordance with Equation 2.5.

To find $\hat{t}_1$ and $\hat{t}_2$, piece-wise cross-correlation of their sound files with f_{0_j} is performed. The same window length t_{win} is used to do so. Equation 3.3, adapted from [24], shows how the cross-correlation $R_{f_{0_j}:f_n}$ of each window of f_1 and f_2 is calculated:

$$R_{f_{0_j}:f_n}[i] = \sum f_{0_j} f_n[i+k]. \tag{3.3}$$

Figure 3.8 depicts the piece-wise cross-correlation for f_1 and f_2. There are clear peak values indicating the time at which f_{0_j} had the highest correlation with a window in f_1 or f_2. The window number i is used to calculate the estimated time delay $\hat{t}_n$, from which the estimated TDOA $\Delta\hat{t}_n$ can be found.

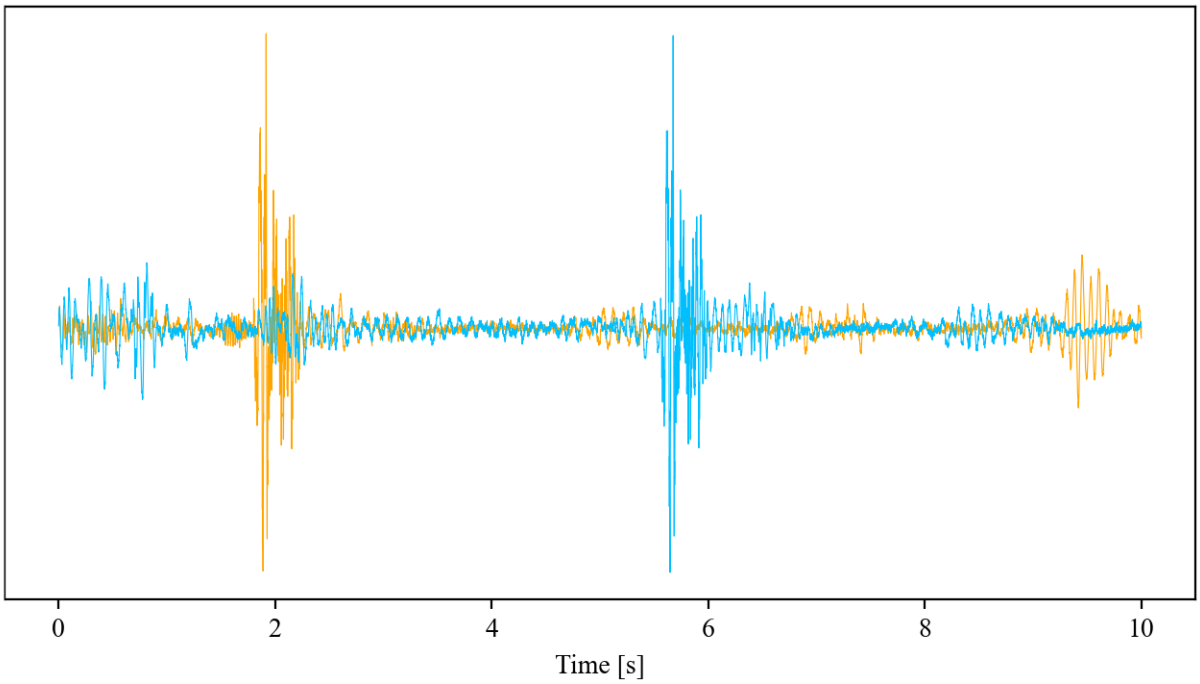


Figure 3.8: Cross-correlation of f_1 and f_2 with f_{0_j}.

Using an adapted linear velocity equation, the estimated difference of sensor-to-source distance $\Delta\hat{d}_{W_n}$ for each sensor is found. Equation 3.4 is this adapted equation:

$$\Delta\hat{d}_{W_n} = v \times \Delta\hat{t}_n. \tag{3.4}$$

Note that since $\Delta\hat{t}_0 = 0$ s, $\Delta\hat{d}_{W_0} = 0$ km.

3.5 Hyperbolic Trilateration

Once the estimated TDOAs $\Delta\hat{t}_n$ are obtained, it is possible to determine the values of the geometric parameters a_n, b_n and d_n required for hyperbolic trilateration. For more insight into the formulas and parameters used in this section, see Section 2.2.

We start with the sensor-to-sensor distances, which are found using the known sensor coordinates, and Equation 3.5, adapted from the Cartesian distance formula:

$$d_{S_{n:n+1}} = \sqrt{(x_{S_{n+1}} - x_{Sn})^2 + (y_{S_{n+1}} - y_{Sn})^2}. \tag{3.5}$$

The first hyperbolic parameter calculated is d_n. To find d_n, the sensor-to-sensor distance from the previous step is halved, as follows:

$$d_n = \frac{d_{S_{n:n+1}}}{2}. \tag{3.6}$$

The parameter a_n is found in Equation 3.7, adapted from Equation 2.8, using the estimated difference of sensor-to-source distance $\Delta\hat{d}_{W_n}$:

$$a_n = \frac{\left|\Delta\hat{d}_{W_n} - \Delta\hat{d}_{W_{n+1}}\right|}{2}. \tag{3.7}$$

Once a_n and d_n are determined, Equation 2.10 is reworked to obtain b_n:

$$b_n = \sqrt{d_n^2 - a_n^2}. \tag{3.8}$$

The hyperbolic formula for each pair of sensors are populated with the parameters a_n and b_n to plot the possible locations of W. Equation 3.9 is an adapted version of Equation 2.9, the original hyperbolic formula:

$$y = \sqrt{a_n^2 \times \left(1 + \frac{x^2}{b_n^2}\right)}. \tag{3.9}$$

Once the shape of each hyperbola is found, each graph needs to be rotated about the y axis in order to accurately produce the trilateration solutions. Equation 3.10 is adapted from the Cartesian gradient formula and returns the slope value m_n for the straight line between each pair of sensors:

$$m_n = \frac{y_{S_{n+1}} - y_{Sn}}{x_{S_{n+1}} - x_{Sn}}. \tag{3.10}$$

The slope value m_n is used in Equation 3.11 to calculate the angle of rotation θ_n of each set of hyperbole [26]:

$$\theta_n = \begin{cases} -\frac{\pi}{2} - 2\arctan(m_n), & \text{if } m_n \leq 0, \\ \frac{\pi}{2} - 2\arctan(m_n), & \text{otherwise.} \end{cases} \tag{3.11}$$

Finally, the matrix in Equation 3.12 is used to plot the hyperbole and localise W:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta_n & -\sin\theta_n \\ \sin\theta_n & \cos\theta_n \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \tag{3.12}$$

For the test data established earlier in this chapter, the intermediary parameter values in Table 3.1 are obtained. Note that $SNR = 100$ and $t_{win} = 5.0 \times 10^{-4}$ s.

Parameter	$n = 0$	$n = 1$	$n = 2$	**Unit**
$\Delta\hat{t}_n$	0	–1.492	2.2675	[s]
$\Delta\hat{d}_{W_n}$	0	–2.238	3.40125	[km]
$d_{S_{n:n+1}}$	7.9863	9.3378	9.4548	[km]
a_n	1.0063	2.5358	1.5294	$\times 10^{-2}$
b_n	3.4473	3.3467	3.9668	$\times 10^{-2}$
d_n	3.5911	4.1988	4.2515	$\times 10^{-2}$
m_n	0.6336	–7.1399	–0.6184	
θ_n	57.6408	–7.9729	–58.2661	[°]

Table 3.1: Intermediary parameter set for example data.

These intermediary parameters are used to plot the hyperbole in Figure 3.9.

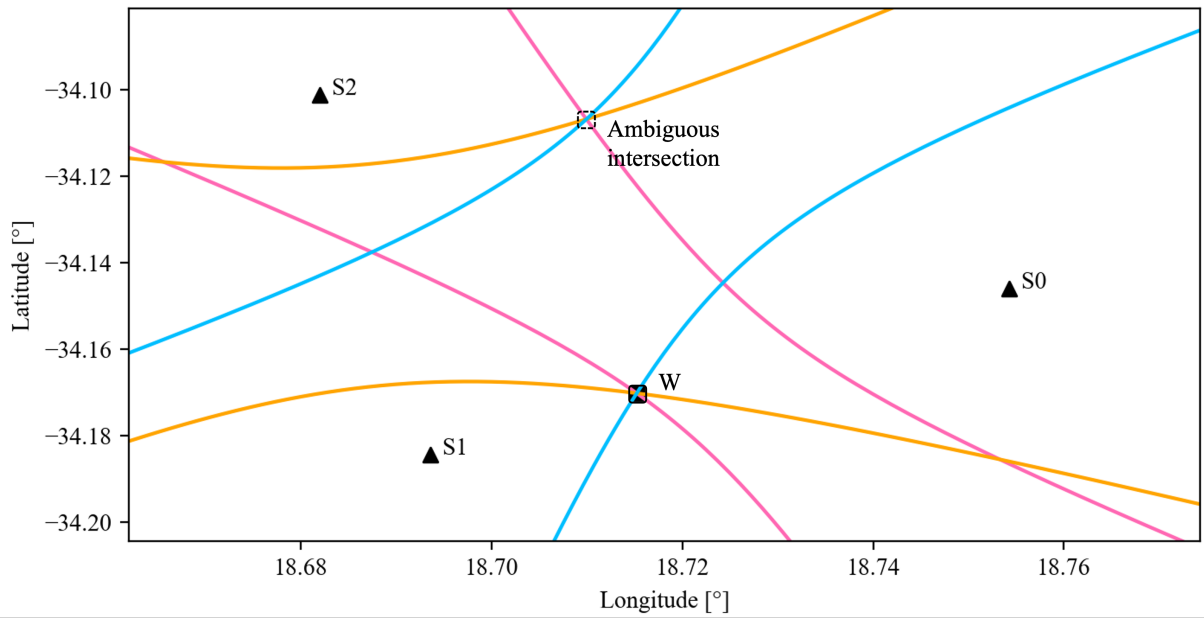


Figure 3.9: System visualisation with test set parameters.

Note the matching colours used to indicate the separate hyperbole associated with each

pair of sensors. The hyperbole produce two intersections: one at the location of W, and one ambiguous intersection. This ambiguity can be removed as described in Section 2.2.

3.6 Simulation Setup

A series of simulations are conducted with the aim of determining the optimal values (or ranges of values) for any variable parameters used in the design of the localisation algorithm. A set of base parameters are selected as input for the simulation, for their ability to produce relatively accurate results (discussed in Chapter 4). These base parameters, the values of which can be seen in Table 3.2, have been used in the example scenario throughout this chapter, and are mostly maintained for the test simulations.

Parameter category	Parameter	Value	Unit
Coordinates	S_0	[18.7542711777, –34.1458751443]	[°]
	S_1	[18.6936017061, –34.1843165453]	[°]
	S_2	[18.6819537842, –34.1011518069]	[°]
	W	[18.7152089095, –34.1701228282]	[°]
SNR	SNR	100	
Correlation window length	t_{win}	1.0×10^{-3}	[s]

Table 3.2: Base parameter set.

For each parameter category in Table 3.2, a set of simulations is performed. The parameters of interest are varied, and the effect of their variation on the performance of the localisation algorithm is measured by calculating an RMSE value and plotting it visually.

Equation 3.13, adapted from Section 2.2, defines the error value e_n for each observation as the difference between the actual sensor-to-source distance d_{W_n} and the estimated sensor-to-source distance $\hat{d}_{W_n}$:

$$e_n = \hat{d}_{W_n} - d_{W_n}. \tag{3.13}$$

This error is used to populate Equation 2.17 to calculate the RMSE value for all three sensors in the array at each iteration of the simulation. Manual inspection of the RMSE trend for each simulation assists in selecting an appropriate new base parameter to be used in following simulations. The aim is to reach a balance where error is minimised while processing speed is taken into account.

Once all parameters have been reselected to optimise the performance of the localisation algorithm, a final simulation is run to measure the error of the new and improved base parameter set.

The following subsections offer more detail on each simulation. All of these simulations are conducted from a single Python testing file. Variable parameters are manually entered and plots are printed to a Jupyter notebook for ease of use.

3.6.1 Simulation A: Sensor and Whale Coordinates

Simulation A measures the effect of changing the position of each sensor S_n as well as the vocalisation source W on the RMSE of the localisation algorithm. For this simulation, all base parameters are used except for the coordinates. Instead, each sensor coordinate x, y_{S_n} and source coordinate x, y_W is generated using a random, uniformly distributed process. The ranges within which each coordinate value falls are defined below:

$$x = [18.680°, 18.765°] = [18°40'48'' \text{ E}, 18°45'54'' \text{ E}],$$
$$y = [-34.185°, -34.1°] = [34°11'06'' \text{ S}, 34°06'00'' \text{ S}].$$

This is an approximately $10\,\text{km}^2$ region located within the False Bay study area, as seen in Figure 3.10.

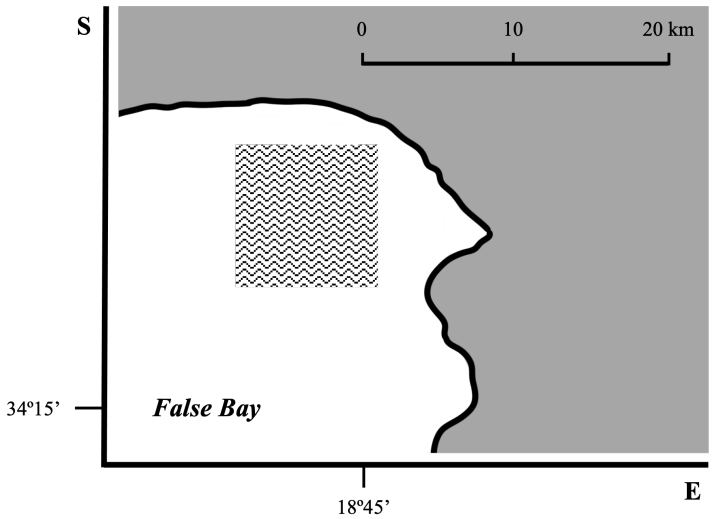


Figure 3.10: Simulation A test region.

3.6.2 Simulation B: Signal-to-Noise Ratio (SNR)

Simulation B is used to measure the effect of decreasing the SNR of the fabricated vocalisation sound files f_n on the RMSE of the localisation algorithm. There are two parts to this simulation.

First, a low range of SNR values is applied to f_n and used alongside the rest of the base parameter set to understand at which SNR the algorithm becomes completely inaccurate.

These ratios range from $SNR = 5$, down to $SNR = 0.1$.

The second run of Simulation B features a larger range of SNR values, ranging up to $SNR = 15$ dB, but only starting above the relative accuracy threshold identified in the previous part. Both parts of Simulation B plot the mean RMSE values for 1 000 iterations for each SNR value. This ensures that outliers are accounted for and results are reported on within context.

3.6.3 Simulation C: Correlation Window Length

Simulation C is used to measure the effect of using smaller and larger window lengths t_{win} when performing piece-wise auto- and cross-correlation on f_n. Once again, there are two parts to this simulation.

As in Simulation B, an initial relative accuracy threshold is determined by first iterating over a large range of t_{win} values. These t_{win} values are chosen to represent all whole factors of the number of samples q in f_n. In the case of our test set,

$$q = 96\,000\,\text{Hz} \times 10\,\text{s} = 960\,000\,\text{samples}.$$

Once the relative accuracy threshold is determined, the simulation is repeated with a smaller range of t_{win} values, in order to more closely inspect the RMSE trend for better-performing localisation. Both parts of Simulation C plot the mean RMSE values for 1 000 iterations for each t_{win}, to ensure that outliers are accounted for and results are reported on within context.

3.7 Software Used

3.7.1 Audacity ®

For the examination of the original audio file, and the extraction of f_{voc} and f_{sea_n}, an open-source recording and editing software called *Audacity* ® [27] has been used. This software is user-friendly while being platform-adaptable and sophisticated in utility.

3.7.2 Python

All further system design, signal processing, and simulation is done in Python. Separate modules containing functional classes have been written for efficiency of testing. Jupyter notebook is used to interface with the modules.

The *correlate* function in *NumPy* [28] calculates the auto- and cross-correlation of the fabricated vocalisations. The *SoundFile* Python library [29] is used to read and write waveform audio (wav) files. The only other noteworthy Python package is *ObsPy*, particularly its *geodetics* library [30]. *ObsPy* was developed for processing seismological data, and

its *geodetics* library allows for accurate conversion between geographical coordinates in degrees and kilometer distances.

3.8 Summary

This chapter describes the development of our localisation algorithm and the way in which its performance is tested.

Section 3.1 gives an overview of the iterative design process before selecting an example dataset with which to visualise the design in Section 3.2. In Section 3.3, the example dataset is used to fabricate acoustic recordings as though they had been produced by the sensors in our system. These fabricated recordings are used to determine the estimated TDOA values of the signal in Section 3.4, and localise the source of the signal in Section 3.5. Section 3.6 is where the simulation setup to test the performance of the localisation algorithm is explained. Finally, in Section 3.7, the software environment within which all of this design and testing takes place is outlined.

In the next chapter, the results from Simulations A – C are displayed and analysed.

Chapter 4

Simulation Results

4.1 Simulation A: Sensor and Whale Coordinates

In Figure 4.1, we observe the RMSE produced by the localisation algorithm when the sensor and source coordinates are randomised within a 10 km^2 region.

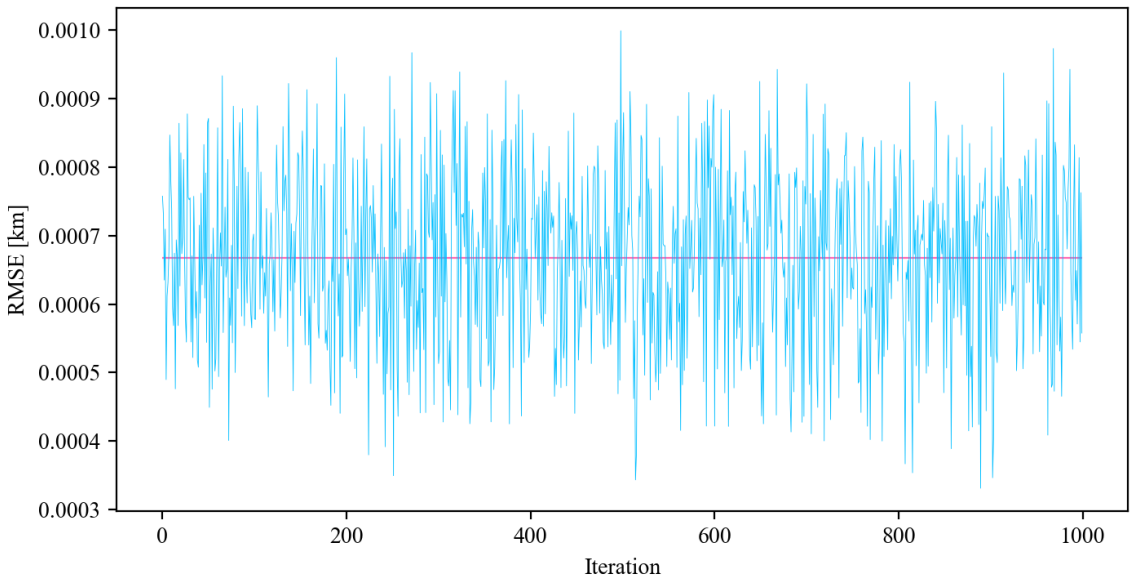


Figure 4.1: RMSE produced by Simulation A.

Over 1 000 iterations, the mean RMSE (indicated in pink) is found to be 0.66 meters, with a total error range of 0.33 m to 1.00 m. This is a relatively low variation, considering the size of the study region and, of course, of the vocalising cetacean.

The results of Simulation A indicate that in the ideal system dictated by our assumptions and using the base parameter set from Table 3.2, any combination of sensor placements can accurately localise a vocalisation source if it occurs within the 10 km^2 region that the sensors are located. This means that no specific geometric placement of the sensors is necessary, and that drifting of sensors within that same region will have an insignificant effect on the accuracy of the localisation algorithm, provided that the GPS coordinates of each sensor at the time of detection is known.

4.2 Simulation B: Signal-to-Noise Ratio (SNR)

Simulation B.1 – Figure 4.2 shows the RMSE for a low range ($SNR = [0.05, 5.00]$) of vocalisation signal SNR values after 1 000 iterations per ratio.

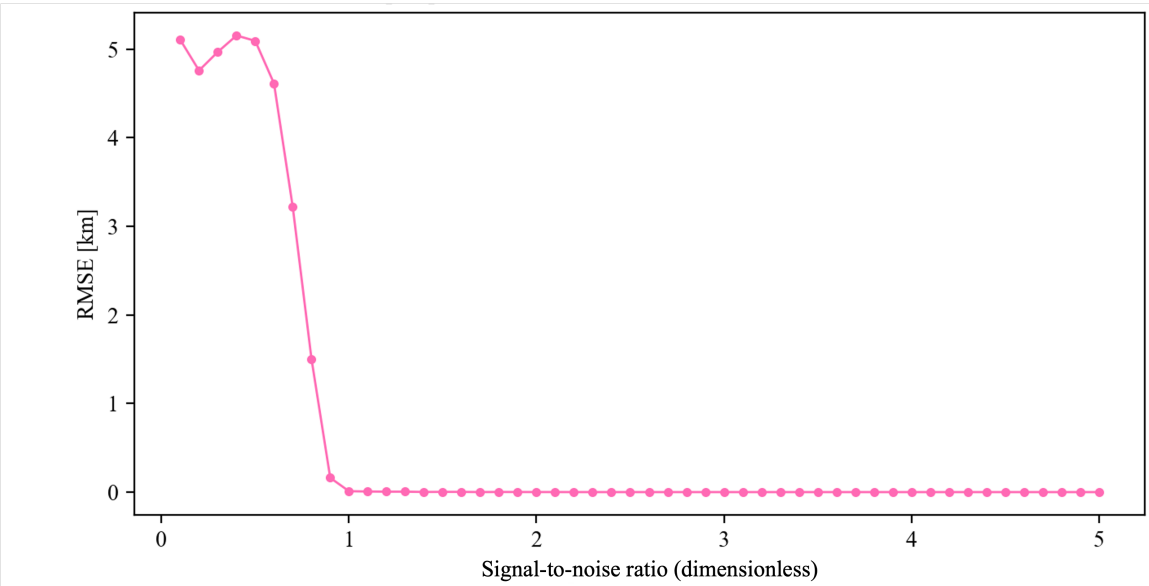


Figure 4.2: RMSE produced by Simulation B.1.

The RMSE is large, approximately 5 km, for $SNR < 0.5$. The error then sharply decreases as the SNR approaches 1. Ratios above 1:1 all produce an RMSE close to 0.

The results of Simulation B.1 indicate clearly that signal-to-noise ratios below 0 dB render the localisation algorithm unusable. This is expected, as $SNR < 0$ dB means that the noise level of a signal is higher than the signal level. Simulation B.2 further investigates the SNRs that produce more accurate localisation results.

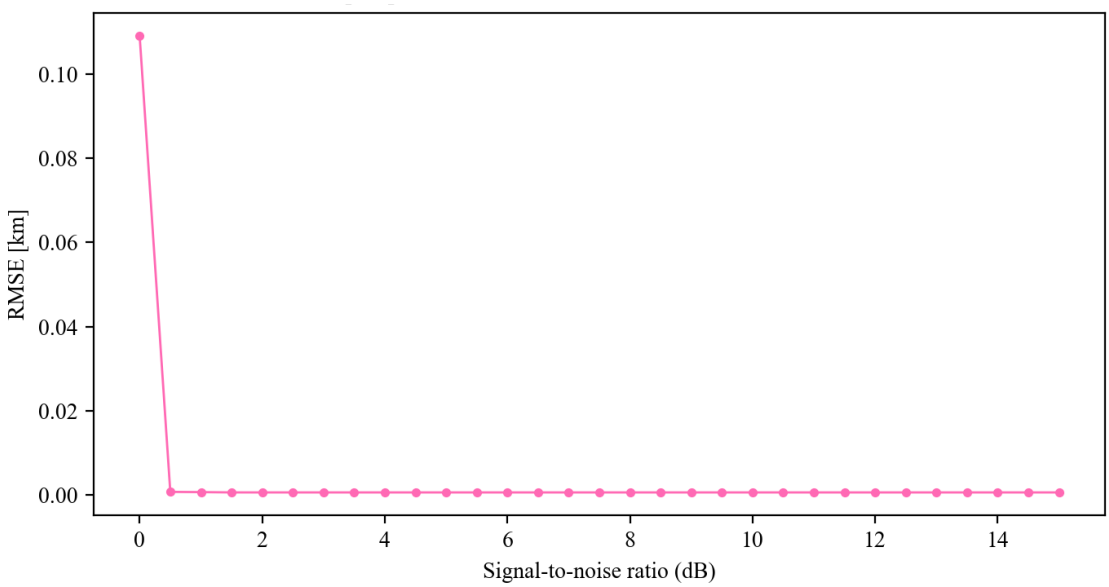


Figure 4.3: RMSE produced by Simulation B.2.

Simulation B.2 – In Figure 4.3 is the RMSE plot produced by 1 000 iterations of the localisation algorithm for signal-to-noise ratios from 0 dB to 15 dB. At 0 dB, the RMSE is approximately 110 m, which may be helpful for rough tracking, but is relatively inaccurate for detailed monitoring of movement. From 0.5 dB onwards, the RMSE graph reaches a plateau and maintains a value of 0.66 m.

The results of Simulation B.2 indicate that in the ideal system dictated by our assumptions and using the base parameter set from Table 3.2, the source will be accurately located, as long as the noise present in the vocalisation recording produces $SNR > 0.5$ dB.

4.3 Simulation C: Correlation Window Length

Simulation C.1 – Figure 4.4 illustrates the RMSE produced by 1 000 iterations of the localisation algorithm where the correlation window used to detect vocalisation peaks ranges from $t_{win} = 3.125 \times 10^{-4}$ s to 10 s.

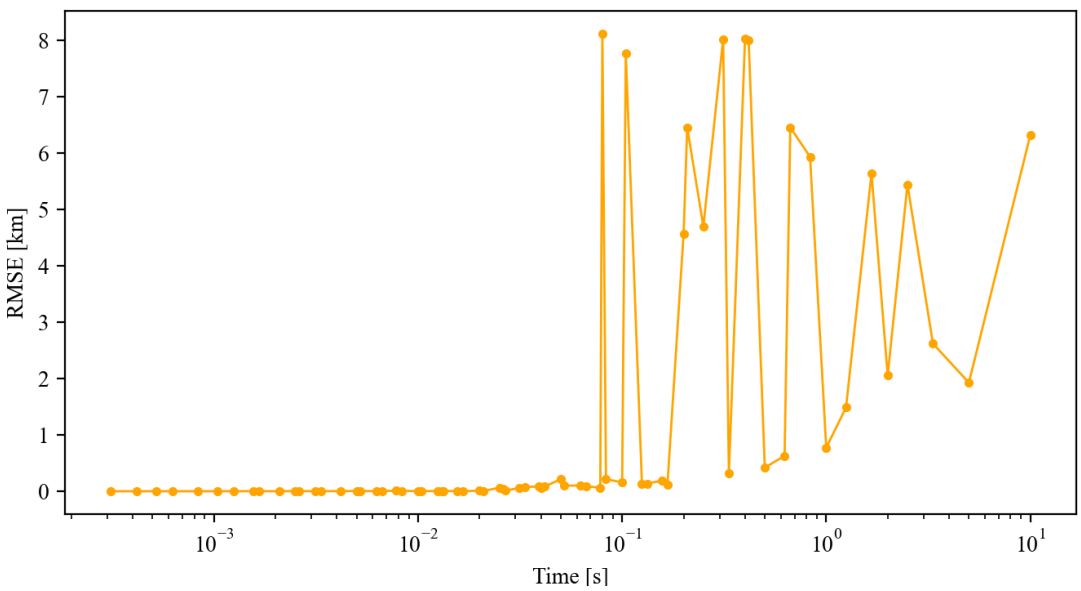


Figure 4.4: RMSE produced by Simulation C.1.

The longer windows, particularly those longer than 0.8 seconds, produce large and erratic error values. Correlation windows shorter than 0.8 seconds produce RMSE values close to 0.1 km, decreasing gradually with the value of t_{win}.

The results of Simulation C.1 indicate that for a Bryde's whale pulse vocalisation recorded at a sampling rate of 96 kHz, the correlation window used to detect the peak value of the vocalisation needs to be shorter than 0.8 seconds in order to accurately localise the source of the vocalisation. The accuracy of the localisation for $t_{win} < 0.8$ s is investigated in Simulation C.2.

Simulation C.2 – In Figure 4.5 we observe the RMSE for 1 000 iterations of the localisation algorithm for a smaller set of correlation window lengths ranging from

$t_{win} = 3.125 \times 10^{-4}$ s to 7.8125×10^{-2} s. This represents a range in the number of samples per window of 30 to 7500.

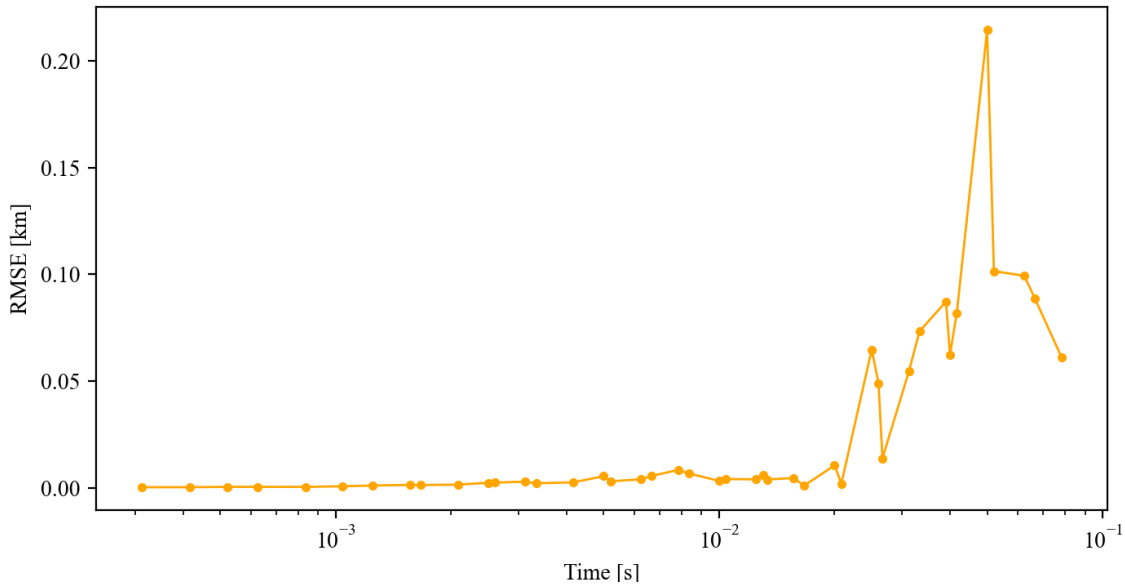


Figure 4.5: RMSE produced by Simulation C.2.

There is a downward trend in the error measurement as the correlation window becomes shorter, with RMSE = 61.25 m at $t_{win} = 7.8125 \times 10^{-2}$ s and RMSE = 0.204 m at $t_{win} = 3.125 \times 10^{-4}$ s. However, the peak RMSE value for this range is 214.69 m at $t_{win} = 5.0 \times 10^{-2}$ s.

Considering how many iterations have been completed of the localisation algorithm, the RMSE curve is unexpectedly irregular. This is likely because of cases where the peak value of a vocalisation occurs well within a correlation window, meaning that there is a pause between the start time of the window, selected as the estimated arrival time, and the actual time at which the peak occurs. The effect on accuracy of such cases will be more pronounced for longer correlation windows, hence the irregular appearance of the RMSE curve. So, while it is sufficient for rough tracking to have an error of a 200 m or so, as recorded in the right-hand-side of Figure 4.5, more predictability is observed for correlation windows shorter than 2×10^{-2} s.

The results of Simulation C.2 indicate that the accuracy of our localisation algorithm improves significantly as the correlation window used to detect vocalisations becomes shorter. In fact, t_{win} has the most consequential impact of all three parameters considered in this chapter. Unlike the RMSE trend of SNR in Figure 4.3, the plot in Figure 4.5 changes for every new value of t_{win}. This leads us to believe that changing the correlation window length would directly alter the RMSE "plateau" value of the SNR simulation. Therefore, the correlation window length used for this algorithm requires careful consideration before selection. For the best performance with recordings that have a sampling rate of 96 kHz, a t_{win} value of 2×10^{-2} s or less should be selected.

4.4 Simulation D: Final Test

After Simulations A – C are conducted, Simulation D is run with the intention of observing the consistency of the results obtained. The localisation algorithm is run for 1 000 iterations with a set of parameters chosen to reflect sufficient accuracy as determined by the previous simulations. No parameters are changed in between iterations, as the aim is to observe any variation in results that may be caused by the noise levels of a recording, or any other unexpected issues.

The parameter values chosen for Simulation D are as follows:

- **Coordinates:** The same sensor and source coordinates used in Chapter 3 are used here. The sensor-to-sensor values do not exceed $d_{S_{n:n+1}} = 9.45\,\text{km}$ and the sensor-to-source values do not exceed $d_{W_n} = 8.51\,\text{km}$. They all occur within the same $10\,\text{km}^2$ region, and therefore should produce accurate results as per Simulation A.

- **SNR:** A *SNR* value of 100 (or 2 dB) is selected. This is above the accuracy threshold of $SNR = 0.5\,\text{dB}$ as defined in Simulation B.

- **Correlation window length:** A window length $t_{win} = 1.0 \times 10^{-3}\,\text{s}$ is selected. This is sufficiently below the accuracy threshold of $2 \times 10^{-2}\,\text{s}$ defined in Simulation C to avoid any irregular results.

Figure 4.6 depicts the results of Simulation D. The RMSE is 0.66 m for all 1 000 iterations of the localisation algorithm with the parameters selected above. In relation to the study region and object, this error is inconsequential; it is certainly sufficient for the purpose of localising and tracking cetacean movement.

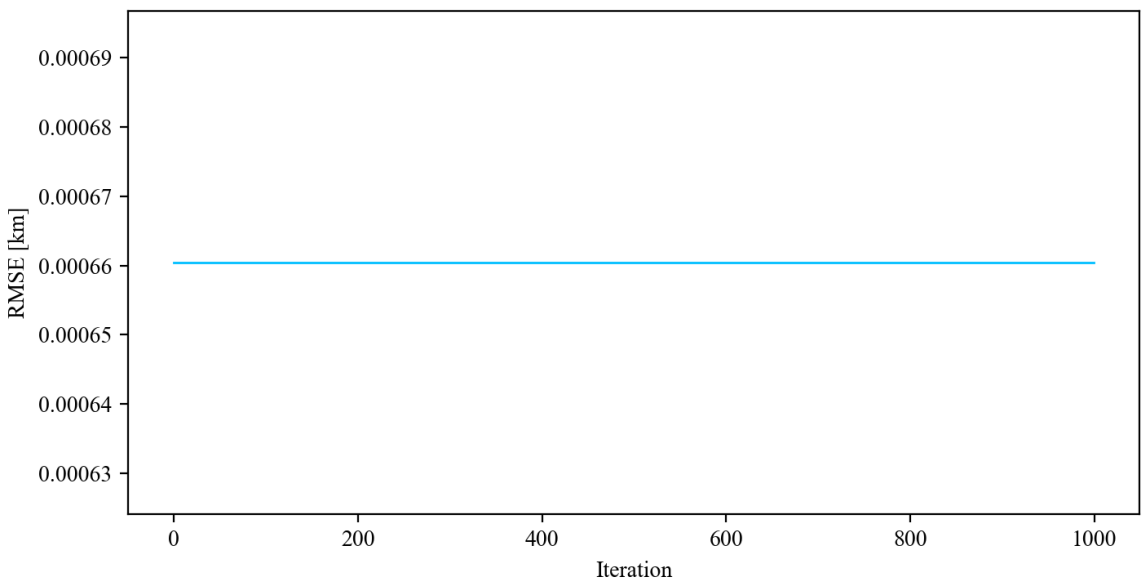


Figure 4.6: RMSE produced by Simulation D.

4.5 Summary

In this chapter, the results from Simulations A – D are displayed and analysed. Optimal parameter ranges are identified to produce localisation accurate to the nearest 1 m of the source.

In Section 4.1, we observe a low variation in localisation accuracy as the sensor and source coordinates are randomly selected within the same 10 km^2 region. These results demonstrate that the localisation algorithm allows for drifting of sensors and does not require them to be arranged in any particular manner in order to function.

Section 4.2 shows that for a $SNR = 0.5$ dB or higher, the localisation algorithm consistently produces accurate results.

The most significant results are observed in Section 4.3, where we note the erratic variation in localisation accuracy when the correlation window length of the algorithm is changed. Although an accuracy threshold of 2×10^{-2} s is determined, the results in this section emphasise the care that must be taken when selecting a value for t_{win}. This parameter has the largest impact on the localisation accuracy of our algorithm.

Section 4.4 reinforces our confidence in the optimal accuracy ranges identified in the previous sections. These optimal ranges can be seen in Table 4.1.

Parameter	**Tested**		**Optimal**		**Unit**
	Min.	Max.	Min.	Max.	
$\mathbf{d_{S_{n:m}}}$, $\mathbf{d_{W_n}}$	0	14.14	0	14.14	[km]
SNR	–13.0	15	0.5		[dB]
$\mathbf{t_{win}}$	3.125×10^{-4}	10		2×10^{-2}	[s]

Table 4.1: Optimal parameter ranges.

Note that for Simulation A, the tested coordinates produce accurate results across the entire range. For a 10 km^2 region, this translates to a maximum value for $d_{S_{n:m}}$ and d_{W_n} of 14.14 km. This is not necessarily the entire range within which accurate results would be shown, but since no other testing is done, this is the optimal range that is reported.

Chapter 5

Conclusion

5.1 Objectives

The project sought to develop an algorithm that localises the source of an underwater acoustic event in two dimensions to an accuracy of 1 m. This algorithm is to be integrated with a cetacean monitoring system which consists of three underwater sensors that produce audio recordings and can return their respective GPS coordinates at the time of recording. Reasonable levels of noise, as well as drifting of sensors, were to be accounted for in the design of the algorithm.

5.2 Algorithm Design and Testing

After consulting existing work in the field of acoustic localisation, the TDOA method has been implemented, using cross-correlation for the detection of vocalisations within a recording, and hyperbolic trilateration for returning the source position. Example data has been created to test the performance of the algorithm. This allows for easy adjustment of parameter values such as the locations of the sensors and source, the signal-to-noise ratio of the audio recording, and the cross-correlation window length used to detect vocalisations.

The algorithm has been tested using a series of simulations. The results of these simulations prove the following:

- the algorithm accurately locates the source of an audio event within 1 m;
- the algorithm works for any arrangement of sensor and source coordinates within a $10\,\mathrm{km}^2$ range;
- the algorithm works for SNR levels as low as 0.5 dB.

Upon achieving the above, all project objectives are met successfully.

The simulations have also identified parameter limits within which the algorithm performs most accurately. In particular, the maximum length of cross-correlation window for audio recordings sampled at 96 kHz should be 2×10^{-2} s. This variable has the greatest impact on the localisation accuracy produced by the algorithm.

5.3 Future Work

The performance of the current algorithm can be further optimised by investigating variables that were not considered in this report. One such variable is the example vocalisation snippet that was used to fabricate the recording files. Realistic levels of attenuation of the vocalisation signal could also be mimicked in the creation of these sound files. Attenuation is likely to have an effect on the SNR accuracy threshold and that of the sensor-to-sensor and sensor-to-source distances. Furthermore, the size of the window overlap for the piece-wise cross-correlation step has been maintained at half the length of the window, but could certainly be varied. Considering the impact of the window length itself on the localisation accuracy, a similar test for the overlap size might also produce significant results.

Adding a fourth sensor into the system would extend the range of signal detection, reduce localisation ambiguity in the 2D system, and make 3D localisation possible. The algorithm would need to be slightly adapted to accommodate the added sensor, but it would considerably improve the localisation capability of the system.

For this algorithm to be practically useful as it is now, it should be integrated with a hardware and pre-processing solution. The limits of the hardware used, such as the recording range of the hydrophone, will need to be considered in this integration, and the algorithm should be adjusted if necessary. Pre-processing of the sensor recordings should, at least, include the automatic detection of vocalisation instances. Filtration of noise at undesirable frequencies would also be useful.

5.4 Conclusion

All of the objectives of this project are met. Observations regarding the performance of TDOA localisation using cross-correlation and hyperbolic trilateration are documented and discussed, and future work recommendations for the implementation and improvement of the localisation algorithm are made. This algorithm is a valuable tool for the localisation of any underwater audio event, particularly for the tracking and monitoring of cetaceans.

Bibliography

[1] J. G. Cooke and R. L. Bronwell Jr. (2018) Balaenoptera Edeni, Bryde's Whale. International Union for Conservation of Nature and Natural Resources (IUCN). The IUCN Red List of Threatened Species. [Online]. Available: iucnredlist.org

[2] D. L. Mackenzie and J. A. Royle, "Designing Occupancy Studies: General Advice and Allocating Survey Effort," *Journal of Applied Ecology*, vol. 42, July 2005.

[3] S. M. Jarvis, R. P. Morissey, D. J. Moretti, N. A. DiMarzio, and J. A. Shaffer, "Marine Mammal Monitoring on Navy Ranges (M3R): A Toolset for Automated Detection, Localization, and Monitoring of Marine Mammals in Open Ocean Environments," *Marine Technology Society Journal*, vol. 48, no. 1, Feb 2014.

[4] G. Cooke. (2020, Jun) Whales of Hermanus. The Royal Portfolio. [Online]. Available: theroyalportfolio.com/whales-of-hermanus

[5] A. Sirovic, H. R. Basset, S. C. Johnson, S. M. Wiggins, and J. A. Hildebrand, "Bryde's Whale Calls Recorded in the Gulf of Mexico," *Marine Mammal Science*, vol. 30, no. 1, Jan 2014.

[6] T. L. Rogers, M. B. Ciaglia, H. Klinck, and C. Southwell, "Density Can Be Misleading for Low-Density Species: Benefits of Passive Acoustic Monitoring," *Public Library of Science (PLoS) ONE*, vol. 8, no. 1, Jan 2013.

[7] G. S. Penry, V. G. Cockroft, and P. S. Hammond, "Seasonal Fluctuations in Occurrence of Inshore Bryde's Whales In Plettenberg Bay, South Africa, with notes on Feeding and Multispecies Associations," *African Journal of Marine Science*, vol. 33, no. 3, Sep 2011.

[8] J. E. Stanistreet, D. Risch, and S. M. Van Parijs, "Passive Acoustic Tracking of Singing Humpback Whales (Megaptera Novaeangliae) on a Northwest Atlantic Feeding Ground," *Public Library of Science (PLoS) ONE*, vol. 8, no. 4, Apr 2013.

[9] G. Pavan. (2005, Aug) Underwater Bioacoustics. Centro Interdisciplinare di Bioacustica e Richere Ambientali (CIBRA). [Online]. Available: unipv.it/cibra/edu_underwaterbioacoustics_uk

[10] S. E. Parks, C. F. Hotchkin, and K. A. Cortopassi, "Characteristics of Gunshot Sound Displays by North Atlantic Right Whales in the Bay of Fundy," *The Journal of the Acoustical Society of America*, vol. 131, no. 3173, Jan 2012.

[11] M. Gassmann, E. E. Henderson, S. M. Wiggins, M. A. Roch, and J. A. Hildebrand, "Offshore Killer Whale Tracking Using Multiple Hydrophone Arrays," *The Journal of the Acoustical Society of America*, vol. 134, no. 3513, Sep 2013.

[12] S. M. Wiggins and J. A. Hildebrand, "High-Frequency Acoustic Recording Package (HARP) for Broad-Band, Long-Term Marine Mammal Monitoring," in *2007 Symposium on Underwater Technology and Workshop on Scientific Use of Submarine Cables and Related Technologies*. Tokyo, Japan: IEEE, Apr 2007.

[13] V. Kandia, Y. Stylianou, and T. Dutoit, "Improve the Accuracy of TDOA Measurement Using the Teager-Kaiser Energy Operator," in *2008 New Trends for Environmental Monitoring Using Passive Systems*. Hyeres, France: IEEE, Oct 2008.

[14] P. Wu, S. Su, Z. Zuo, X. Guo, B. Sun, and X. Wen, "Time Difference of Arrival (TDoA) Localization Combining Weighted Least Squares and Firefly Algorithm," *Sensors (Basel, Switzerland)*, vol. 19, Jun 2019.

[15] J. L. Spiesberger and K. M. Fristrup, "Passive Localization of Calling Animals and Sensing of their Acoustic Environment Using Acoustic Tomography," *The American Naturalist*, vol. 135, no. 1, Jan 1990.

[16] D. E. Chaitanya and G. S. Rao, "Unknown radio source localization based on a modified closed form solution using TDOA measurement technique," *Procedia Computer Science*, vol. 87, no. 184 – 189, May 2016, 4th International Conference on Recent Trends in Computer Science and Engineering.

[17] W. Rison, "Time of Arrival Location Technique," New Mexico Tech., 2008, EE 389 – Mathematical Engineering course notes.

[18] A. Mardjanian, S. Lashutka, R. Kamgang, A. Nasr, and R. Conti. Equations of Hyperbolas. OpenStax College. College Algebra: Online Course. Chapter 15: Conic Sections. [Online]. Available: courses.lumenlearning.com/waymakercollegealgebra/chapter/equations-of-hyperbolas

[19] B. Jin, X. Xu, and T. Zhang, "Robust Time-Difference-of-Arrival (TDOA) Localization Using Weighted Least Squares with Cone Tangent Plane Constraint," *Sensors (Basel, Switzerland)*, vol. 18, Mar 2018.

[20] S. Potluri, "Hyperbolic Position Location Estimator with TDOAs from Four Stations," Master's thesis, New Jersey Institute of Technology, 2002.

[21] N. Vandeput, *Data Science for Supply Chain Forecasting*, 2nd ed. De Gruyter, Mar 2021.

[22] S. Glen. (2021) Statistics Definitions in Plain English with Examples. Statistics How To: Elementary Statistics. [Online]. Available: statisticshowto.com/probability-and-statistics/statistics-definitions

[23] A. M. Usman, O. O. Ogundile, and D. J. J. Versfeld, "Review of Automatic Detection and Classification Techniques for Cetacean Vocalization," *IEEE Access*, vol. 8, Jun 2020.

[24] P. Z. Peebles Jr., *Probability, Random Variables and Random Signal Principles*, 2nd ed., ser. McGraw-Hill Series in Electrical Engineering, Communications and Information Theory. McGraw-Hill Inc., 1987, ch. 6: Random Processes.

[25] L. Simon and D. Young. Autocorrelation and Time Series Methods. Pennsylvania State University: Eberly College of Science. Statistics 462 Applied Regression Analysis Lesson 10. [Online]. Available: online.stat.psu.edu/stat462/node/188

[26] J. Abrahamson, V. Falduto, R. Gross, D. Lippman, M. Rasmussen, R. Norwood, N. Belloit, J.-M. Magnier, H. Whipple, and C. Fernandez. Rotation of Axes. OpenStax College. Precalculus II: Online Course. Chapter 10: Analytical Geometry. [Online]. Available: courses.lumenlearning.com/precalctwo/chapter/rotation-of-axes

[27] Audacity ®. Audacity Team. [Online]. Available: audacityteam.org

[28] T. E. Oliphant, *Guide to NumPy*, 1st ed., Dec 2006.

[29] Soundfile. PySoundFile. Python library documentation. [Online]. Available: pysoundfile.readthedocs.io

[30] ObsPy: A Python Framework for Seismology. ObsPy Development Team. [Online]. Available: docs.obspy.org

Appendix A

Project Planning Schedule

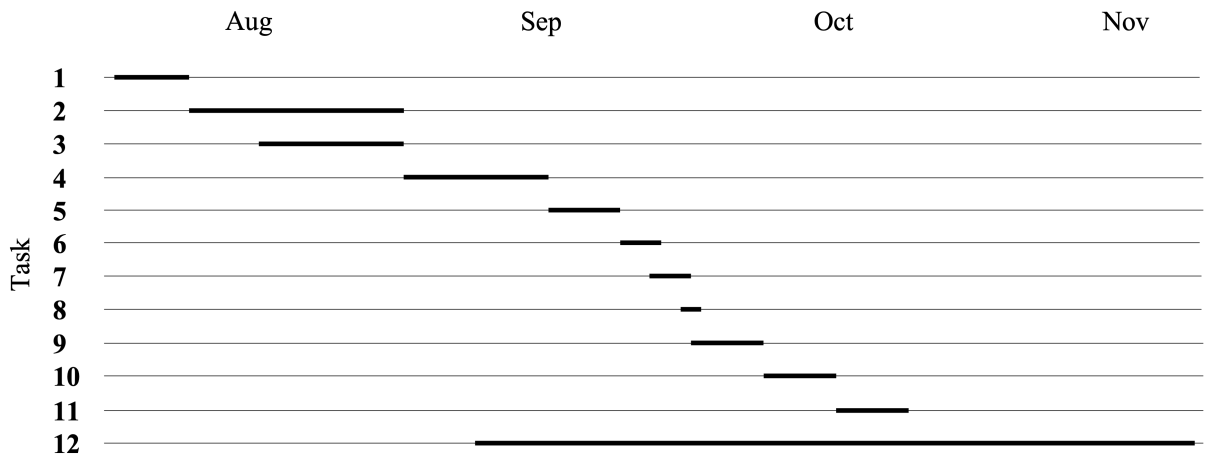


Figure A.1: Project planning Gantt chart.

Time Period	Task	
2021/08/02 – 2021/08/09	**1**	Plan; establish project aims.
2021/08/09 – 2021/08/29	**2**	Research related work and relevant concepts.
2021/08/16 – 2021/08/29	**3**	Write TDOA algorithm.
2021/08/30 – 2021/09/12	**4**	Write hyperbolic localisation algorithm.
2021/09/13 – 2021/09/19	**5**	Create test data.
2021/09/20 – 2021/09/23	**6**	Write correlation algorithm.
2021/09/23 – 2021/09/26	**7**	Integrate localisation and correlation algorithms.
2021/09/26 – 2021/09/27	**8**	Modularise code for efficient testing.
2021/09/27 – 2021/10/03	**9**	Test algorithm at different correlation window sizes.
2021/10/04 – 2021/10/10	**10**	Test algorithm at different SNRs.
2021/10/11 – 2021/10/17	**11**	Test algorithm at random coordinates.
2021/09/06 – 2021/11/14	**12**	Write report.

Table A.1: Project planning schedule.

Appendix B

ECSA Compliance

Table B.1: ECSA exit level outcomes compliance.

ECSA Outcome	Description of Outcome in Report	Section(s)
ELO 1: Problem solving	The scope of the localisation problem was ill-defined; the problem itself was identified in order to formulate the problem statement and objectives. The localisation algorithm was developed and optimised to best address and solve the objectives. Recommendations for ways in which to solve out-of-scope problems were also made.	1.2, 1.3, 1.4 3.2, 3.3, 3.4, 3.5, 3.6 5.3
ELO 2: Application of scientific and engineering knowledge	Engineering knowledge in the form of applied mathematics, signal processing and computer science was applied to the development and testing of the algorithm.	2.2 3.2, 3.3, 3.4, 3.5, 3.6, 3.7
ELO 3: Engineering design	An iterative process for the development and optimisation of the localisation algorithm was designed and performed.	3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7
ELO 4: Investigations, experiments and data analysis	A series of simulations were designed and run in order to assess the effect of altering certain variables on the performance of the localisation algorithm. The results of these simulations were analysed and reported on.	3.6 4.1, 4.2, 4.3, 4.4, 4.5

Continued on next page.

Table B.1 – *Continued from previous page.*

ECSA Outcome	Description of Outcome in Report	Section(s)
ELO 5: Engineering methods, skills and tools, including information technology	The entire project, including cross-correlation, localisation and testing, has been written in Python. The *SoundFile*, *ObsPy* and *NumPy* packages were utilised for their intended functionality.	3.2, 3.3, 3.4, 3.5, 3.6, 3,7 4.1, 4.2, 4.3, 4.4
ELO 6: Professional and technical communication	Correspondence with project supervisor was conducted online as a result of remote learning. This correspondence was done in a timely and professional manner to respect the time of Prof. Versfeld. The project report was written for an engineering audience in mind, and to meet the standards of technical academic writing.	All
ELO 8: Individual work	All work was done individually under the supervision of Prof. Versfeld.	All
ELO 9: Independent learning ability	Various new concepts were explored and implemented throughout this project. Such concepts include passive acoustic monitoring, time-difference of arrival localisation and hyperbolic multilateration.	2.1, 2.2

Appendix C

GitHub Repository

The repository containing all source code of the localisation algorithm as well as that of Simulations A – D is publicly available on GitHub at:

https://github.com/ThandiColley/PE448.