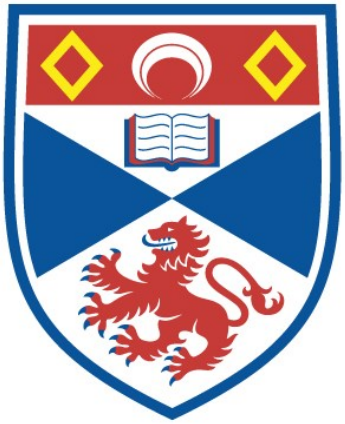
Advancements in methods for estimating the abundance of marine megafauna using novel sampling techniques

Felix Thomas Petersma

A thesis submitted for the degree of PhD
at the
University of St Andrews



2024

Identifier to use to cite or link to this thesis:
DOI: https://doi.org/10.17630/sta/889

Declarations

Candidate's declaration

I, Felix T. Petersma, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 30,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree. I confirm that any appendices included in my thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

I was admitted as a research student at the University of St Andrews in September 2019.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date: 22-11-2023 Signature of candidate:

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree. I confirm that any appendices included in the thesis contain only material permitted by the 'Assessment of Postgraduate Research Students' policy.

Date: 22-11-2023 Signature of supervisor:

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Felix T. Petersma, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

- No embargo on print copy.

Electronic copy

- No embargo on electronic copy.

Date: 22-11-2023 Signature of candidate:

Date: 22-11-2023 Signature of supervisor:

Underpinning research data or digital outputs

Candidate's declaration

I, Felix T. Petersma, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

Date: 22-11-2023 Signature of candidate:

'Getting a PhD is basically just constantly worrying you're wrong whilst convincing others you're not.'

Acknowledgements

General acknowledgements

I would first like to thank the Centre for Research into Ecological and Environmental Modelling (CREEM) for funding this PhD and making this journey possible. Being able to work for four years on a topic that I love has been a privilege.

Special thanks to my collaborators, without whom much of this research would not have been possible. For the extension of acoustic spatial capture-recapture, I express my gratitude to Aaron Thode, who kept answering my emails and helped me prepare the data so that I could complete this work; Tiago Marques for his contributions in the earlier stages of the project; Gisela Cheoo for her work on the case study that formed the foundation on which my research was built; and Katherine Kim and the rest of Greeneridge Sciences for collecting such an outstanding and unique data set. For the close-kin mark-recapture simulation study, I thank Darcy Bradley and Yannis Papastamatiou for sharing their knowledge on sharks and allowing me to work on their samples; Tanya Sneddon for letting me use her lab space to cut small pieces of shark into even smaller pieces; and of course Mark Bravington, Mr CKMR himself, whose enthusiasm always inspired me during of our talks over the years. For my work on estimating absolute population size from baited cameras, I thank Marcus Rowcliffe for his help in the early stages, and Tom Letessier for his contributions and sharing his data from baited remote underwater video systems.

I am not sure when this journey towards a PhD in statistical ecology started. Maybe when I did my first scuba dive in 2009 and discovered the magic of the underwater world. My parents might say that it started when I watched Free Willy for the first time, or one of the many times thereafter. Even though both surely hold some truth, I think that it really started during my MSc programme at the University of St Andrews. I was able to tailor my masters to focus on statistical ecology and marine science, by including modules on advanced bioacoustics and population dynamics modelling. On top of that, CREEM was, and still is, a place where anyone can always walk into someone's office with a question, even if you do not know them. I want to thank

everyone at CREEM for creating such an welcoming environment. Also, the cake was great. It quickly started to feel like home, as did St Andrews as a town.

Over the past four years, numerous individuals have shaped the experience of this PhD in one way or another. Here, I would like to mention some who have had a particular influence on me: Savannah and Taylor, who created a home not just for themselves, but for everyone close to them, and with whom I share so many great memories; Janneke, one of the first people I became friends with in St Andrews, and who remains a close friend to this day; Marena, whom I explored much of the Scottish lands with, and who will always have a special place in my heart; Cher, one of the best cooks that I know, and whose sourdough bread always turned out better than mine; Maeva, a fellow student from my masters, organiser of the infamous movie nights with Indian food, and who over the years became a dear friend; and Cal and Abinand, two of my closest friends during the PhD. Cal and I started our research at the same time under Len's guidance, and Abi joined the CREEM crew a year later. We were there through good times and through bad times, through love and through heartbreak, through biryani and through macaroni cheese, and I truly hope we never lose touch. Of course, there are many others without whom these years would not have been the same. I thank you all.

Throughout this PhD, I was lucky enough to have been supervised by two amazing people. Danielle, I am so glad that you were able to actively join as a supervisor after first putting a tiny human on this planet. Your support and expertise have been invaluable, and it was great having someone guide me who had gone through this process not too long ago herself. I am really looking forward to DCLDE 2024 this summer, and I hope that this is not the last time we work together.
Len, I often say that you were at least half the reason I started this PhD to begin with. Every conversation we had, whether it was during the independent study module or any of the PhD meetings after, motivated me and got me more excited about whichever topic we had discussed. Your knowledge and experience keep amazing me, and I was lucky to have you guide me through these sometimes very challenging times. You were fair and had to be a bit strict by times, but always supportive of my decisions. This was by far the hardest thing I ever had to do, and I am sure I was not always the easiest person to work with, but you were always (or at least, almost always, as you were occasionally out of service on aeroplanes) there with the support and guidance.

I extend my gratitude to my examiners, Chris Sutherland (internal) and Paul Conn (external) for their contributions. I was more nervous for the viva than I had expected, but it ended being a great chat with snacks and excellent comments.

I also want to mention my friends back home. Living in Scotland made it very clear how important and special these friendships are. Most of us have known each other longer than we have not, with some friendships going back almost 20 years. Living abroad and seeing how rare this is has made me appreciate it so much more. My visits were often packed because I wanted to see as many of them as I could. I would then sit on the plane back to Scotland more tired than when I left, but completely happy and satisfied.

Last, I thank my family. I always found it particularly difficult being away from them. Even though Scotland was less than 1.5 hours by plane, it felt farther than that. This was partly due to Brexit, but the travel regulations in response to the COVID-19 pandemic did not help either. Maybe more than anything, it all felt so far away because my life in Scotland was so different from life back home. Scotland is such a wonderfully beautiful place, in some ways as different from the Netherlands as you can find in such close proximity. Whenever I returned to Amsterdam it felt like coming home, and family played a major part in that. Mum, dad, your support during this journey has been without limits. I really struggled at times, and you were always there, either in person or on the phone, to support in any way that you could. I dedicate this PhD to you. I love you, and the rest of my family, dearly.

Funding

This work was supported by the University of St Andrews (Centre for Research into Ecological and Environmental Modelling).

Abstract

Population size ('abundance') is a key parameter for wildlife conservation. As technology advances, new sampling techniques become available for which established estimation methods are extended and new methods developed. We explore, extend, and test new methods for estimating abundance of large marine megafauna that utilise novel sampling techniques: passive acoustic monitoring (PAM), modern genetic analyses, and baited video systems.

The PAM study is motivated by acoustic recordings of bowhead whales (*Balaena mysticetus*) from multiple sensors in the Beaufort Sea. A high proportion of detections on only one sensor ('singletons') are false positives, resulting from the automated detection algorithm. We extend acoustic spatial capture recapture methods to enable singletons to be discarded without biasing the abundance estimate. In our model, detection probability is dependent on received sound level, but we also propose an extension where detection probability is a function of signal-to-noise ratio, enabling the use of detections at lower received levels.

In the genetic study, we test the recently-developed close-kin mark-recapture (CKMR) method in the situation where the ages of sampled animals are estimated with error. Our work is motivated by a case study of grey reef sharks (*Carcharhinus amblyrhynchos*) at Palmyra Atoll, Pacific Ocean. CKMR relies on knowing animal age, and for many species this is inferred from measured length. We use simulation to test the effects of misspecifying the age-length relationship ('growth curve') and the amount of length measurement error on abundance estimates. A misspecified growth curve produces substantial bias while misspecifying measurement error does not.

Lastly, we explore the potential of extending random encounter modelling (REM), which is used to estimate population size from remote wildlife cameras, for the situation where baited cameras are used as is common in shark studies. We believe that potential exists for REM with baited cameras if bait plume intensity is considered.

To my parents

Contents

Chapter 1

Introduction

1.1 Abundance estimation

Sometimes, the easiest questions are the hardest to answer. A good example of this is the question: 'how many animals are there?'. We cannot feasibly count all animals and thus we would first have to restrict the question. By narrowing the question it becomes clearer what a correct answer could be. For example, 'how many roe deer (*Capreolus capreolus*) live in the Hoge Veluwe in the Netherlands on January 1st 2023?' would be more narrowly-defined and easier to answer question. It is species-specific and concerns a relatively small area. Still, counting all roe deer constitutes a major task, even in a restricted area of 55km^2. Such challenges arise in many instances where the abundance of an animal species is of interest, e.g., the area is too large to cover (e.g., bluefin tuna in the Atlantic Ocean; *Thunnus thynnus*), animals are rarely seen (e.g., deep-diving, elusive beaked whales; *Ziphiidae*), or habitats are hard to access (e.g., the snow leopards in the Himalayas; *Panthera uncia*). The point is that counting all animals, which is called a 'census', is rarely feasible, and that we instead have to estimate the abundance. But how do we estimate population size, and why is it important that we do it at all?

The size of an animal population at a certain time point is a key metric to determine the conservation status of a species (IUCN, 2012; Nichols and Williams, 2006). This can apply to, among other examples, species that are essential to proper functioning of ecosystems, such as bees and whales (Buchmann and Ascher, 2005; Roman et al.,

2014), and species that we harvest and thus livelihoods depend on, such as commercial fish species (Jacinto et al., 2015). When trying to count animals, there are often two main problems: i) we cannot feasibly cover the entire area of interest, and ii) we cannot properly observe the animals. The first problem can be resolved by (systematically) randomly sampling a part of the study area, and then extrapolating over the entire area (Buckland et al., 2015, Chapter 2). The second problem, which is that animals are often hard to see or 'detect', can be resolved by estimating a detection probability and correcting the estimated abundance accordingly (Buckland et al., 2015, Chapter 1). This is the probability that an animal is observed when it is available to be observed at a certain time and place. An example of these concepts is the use of aerial surveys to detect humpback whales (*Megaptera novaeangliae*). Even when the whales are present, one could easily imagine that one will be unable to spot a humpback whale when there is a substantial distance between the observer and the whale. Weather might also play a role, and so does the skill of the observer. Furthermore, whales are not always available for detection, i.e., they only spend some of the time in the detectable water column; when they are too deep underwater once could never see them, even when flying right over them. It is therefore necessary to adjust the total number of whales estimated from a survey for the ones that were there but that were not available to be detected. The main lesson from this example is that estimating this detection probability can be challenging. Several methods have been developed to estimate this detection probability, and technological advances have allowed for improved or new ways of estimating the abundance of animals (Buckland et al., 2023). In this thesis we test, develop, and discuss such novel methods and extensions. First, however, we will present some background on some of the most common abundance estimation methods, some of which are essential to the ones we further cover in Chapter 2–5.

1.2 Capture-recapture methods

1.2.1 Closed populations

Using capture-recapture (CR) to estimate the abundance of closed animal populations dates to the first half of the 20^{th} century (Otis et al., 1978). Closed populations are populations without any births and deaths, and no migration, during some time period. Conventional CR relies on perfect identification of individuals, either through

natural markings or human markings such as tags (Otis et al., 1978). The basic concept of closed-population CR is as follows (Seber, 1982). At the first sampling occasion, a large enough number of animals out of a population are sampled. The total number of individuals sampled at the first occasion is defined as n_1. If they are not individually identifiable, for example, through natural markings such as fur patterns, these individuals are marked or tagged; if the animals were physically captured, they are released again. At a second occasion, animals are sampled again. The total number of animals sampled on the second occasion is denoted n_2, and the number of animals sighted at the second occasion that were seen before (which we know due to their markings) are denoted m_2. In this simple, two-occasion study, where we assume all animals to have the same constant detection probability, our estimate for the total number of animals in this population N is

$$\widehat{N} = \frac{n_1 n_2}{m_2}. \tag{1.1}$$

In other words, the ratio of the number of initial captures to the total abundance is the same as the ratio of recaptures to the number of captures on the second occasion. This estimator is known as the Lincoln-Petersen estimator (Seber, 1982). Besides the closed population and equal detectability assumptions mentioned before, this estimator relies only on the assumption that animals are never misidentified, for example, through lost tags or through mistakes by the observer (Otis et al., 1978). Studies can also involve more than two sampling occasions. In such situations a simple, closed-form estimator such as Equation (1.1) rarely exists (Otis et al., 1978). Otis et al. (1978) provided an unified overview of closed-population CR for more than two trapping occasions up to that point in time, and extended it to estimate density from grid trapping studies. They identified four main models: a base model M_0 using the same assumptions as the Lincoln-Petersen estimator, ; model M_t, where capture probability varies with time or occasion, for example due to weather; model M_b, where capture probability varies due to behavioural responses, for example when an animal becomes trap-shy after getting trapped for the first time; and model M_h, where capture probability varies between individuals (i.e., there is capture heterogeneity among animals), for example when bigger individuals are more likely to be seen, or when a certain sex is more easily observed. Moreover, Otis et al. (1978) summarised all combinations of these models, which results in the additional four models M_{tb}, M_{th}, M_{bh}, and M_{tbh}.

1.2.2 Open populations

Jolly (1965), Seber (1965), and Cormack (1964) all developed some of the first versions of open-population capture-recapture. The version by Cormack only allowed for apparent survival (the complement of deaths and emigration) but not for recruitment (births and immigration), and thus those models are generally referred to as the Cormack-Jolly-Seber (CJS) models (Schwarz, 2001); models that allow for both survival and recruitment are generally referred to as Jolly-Seber (JS) models (Schwarz, 2001). Unlike CJS models, the JS models also estimate abundance by assuming that all animals have equal capture and survival probabilities at each sampling occasion (Pledger et al., 2010). Schwarz and Arnason (1996) achieved this by adding two parameters to the model, namely the size of the super population from which animals could be recruited to the main population, and the probability of recruitment. Pollock (1982) developed a version of the JS model that was robust to variable capture probabilities by having sets of closely-spaced sampling occasions ('secondary periods') during which the population was assumed closed, separated by longer intervals during which the population was assumed open ('primary periods'). Closed-population models could be used for the secondary periods, whereas open populations models could be used for the primary periods (Pollock, 1982).

1.2.3 Spatial capture-recapture

Spatial capture-recapture was simultaneously developed by Borchers and Efford (2008) and Royle and Young (2008), works that both built on previous work by Efford (2004). They took a big step by associating abundance with a particular spatial location and area, thereby estimating a density (surface); total abundance can subsequently be derived by integrating the density over the entire study area. This solved a gap in capture-recapture, where often it was hard or impossible to estimate density from the abundance estimate, as the area that the studied population was associated with was often undefined (Borchers and Efford, 2008). Spatial capture-recapture (SCR; or spatially explicate capture-recapture) resolves this by combining two models: a state model, which generally involved a two-dimensional density surface, through which density was explicitly modelled as a function of spatial covariates, and an observation model, which modelled the detection process for detectors/sensors placed in the study

area (Borchers and Efford, 2008). SCR has gained in popularity since its introduction, resulting in at least 364 published studies by the 3rd of March 2020 that used SCR to answer ecological questions (Tourani, 2022). The main idea behind SCR is that every animal has an activity centre with a (circular) home range or activity area, where the animal spends most of its time at the centre and gradually less as the distance from the centre is increased (Borchers and Efford, 2008). The shorter the distance between an activity centre and a sensor, the higher the probability that an individual is detected within a sampling interval. The detection probability is therefore a combination of two probabilities: the probability that the animal is close enough to the sensor for it to be detected in the first place, and the probability that the sensor detects the animal given that it is within its detectable range. These two probabilities are often not explicitly separated in SCR studies, but it is worth noting that the detection probability is actually a combination of the two processes.

Let n denote the number of unique animals that were detected at least once. The recorded information for each of these n animals generally consists of a 1 if detected at a sensor and a 0 if not, for every sensor in the study area, and for every sampling occasion or interval. The true location of the activity centre is unknown, and thus the method integrates out all possible locations. Variations exists, for example, when a trap can only capture one animal every interval, and thus can no longer 'observe' other animals once it has captured, i.e., observed, the first one (Borchers and Efford, 2008). These variations are mainly determined by the sampling technique used, as SCR can be applied to a variety of sampling methods. Among these sampling methods exist direct sampling methods like camera traps or continuous video (e.g., Karanth and Nichols, 1998; Jiménez et al., 2023), or (multi-catch) live traps (e.g., Flowerdew et al., 2004); or indirect sampling, by sampling hairs left behind or collecting faeces (Mowat and Paetkau, 2002; Poutanen et al., 2019). Although these sampling techniques have worked well in terrestrial contexts, marine-based contexts are not always as suited.

1.3 Challenges of the marine environment

Studying marine wildlife poses additional challenges relative to terrestrial wildlife. This means that several common and well-tested surveying methods are no longer feasible when studying aquatic life. This is mostly due to reduced visibility associated

with the marine environment. To illustrate, imagine that a researcher is planning to observe a relatively rare, large terrestrial species such as a giraffe (*Giraffa*). One excellent option would be a line transect survey from an aeroplane. This way, a large area can be covered relatively quickly, and giraffes are easily observed even from great distance, especially using high-resolution cameras. Imagine using a similar reasoning for the marine environment: someone is planning to study whale sharks (*Rhincodon typus*). These large animals should be distinguishable from great distance, except that, whenever someone flies over a whale shark, they would not see them unless they are close to the surface. This is a challenge when visually assessing the ocean in general: an ocean crowded with animals and one absent of them can seem very similar from an aeroplane or the shore. One must generally go underwater to properly assess what is going on beneath the surface, and this introduces difficulties (Marques et al., 2013). For example, visibility underwater is greatly reduced compared to air, which makes visual observations infeasible for most of the ocean volume. Even in the shallow waters, it is hard to see anything farther than maybe a few dozen meters away when waters are clear, and only a few meters (or even less) when waters are murky. Moreover, animal densities can be very low for marine species, especially when considering megafauna.

1.4 Other abundance estimation methods

Besides the methods covered so far and in the other chapters of this thesis, there are several that we believe should be mentioned here for completeness. First, a census involves a count of all the animals in an area. This method is generally only feasible in small areas when all animals can be counted with certainty. Whenever an area is too large to completely cover given the available resources, one could alternatively use a method called plot sampling (Borchers et al., 2002). In plot sampling, the researchers (randomly) select a number of plots in the study area and count all the animals present in that plot. The total abundance is then estimated by extrapolating the abundance in the sampled plots across the entire study area. Similar to a census, one of the main assumptions is still that animals are detected with certainty. In many contexts animals are not detected with certainty even when they are in observable range of a researchers. In such scenarios it is required to estimate a detection probability. In many contexts, the detection probability of an animal decreases as the distance between the animal and an observer increases. This intuitive concept was formalised in distance sampling

(e.g., Buckland et al., 2015). In distance sampling, researchers sample (a part of) the study area using either point or line transects. Whenever an animal is sighted, researchers recorded the distance from the point to the animal (for point transects) or the perpendicular distance from the transect line (for line transects). This collection of observed distances is then used to fit a detection function, which here is a function that models detectability as a function of distance (Buckland et al., 2015). The detection probability at zero distance is assumed to be perfect, i.e., equal to one, and decreases gradually as distance increases; eventually, it should level at approximately zero. When this assumption fails, the detection probability at zero distance can be estimated using mark-recapture distance sampling, for example by having two observers performing the survey simultaneously (Burt et al., 2014). The most commonly-used detection function in DS is called the half-normal detection function, but others exist such as the hazard-rate and uniform detection functions. With the fitted detection function one then estimates the effective sampled area, which is the sampled area corrected for the probability of detecting the animal. From this, together with the number of animals actually sighted, one can then finally estimate total absolute abundance. Finally, the N-mixture model can be used to estimate animal abundance from count data (Royle, 2004). However, this method has been criticised as it is sensitive to deviations from the assumed distribution of counts (Link et al., 2018).

1.5 Novel sampling techniques

The abundance estimation methods covered were all initially developed for human observers. However, human observation is not always feasible, especially when visibility is limited or continuous surveying is required (Marques et al., 2013). In such scenarios other sampling techniques are preferable, which has resulted in adaptations of old models and the development of new ones. For example, SCR was extended to work with detection data from acoustic recorders (Efford et al., 2009; Dawson and Efford, 2009); genetic sampling led to the development of methods that utilised the information on genetic relatedness (Skaug, 2001; Bravington et al., 2016b); and distance sampling was extended to work with data from remote cameras (Howe et al., 2017). In the following section we further elaborate on these sampling techniques.

1.5.1 Acoustics

The field of acoustics covers the way sound is created and behaves in our world. Sounds are pressure waves travelling through a medium. In general, sound travels quicker and with less energy loss in media that are denser. Hence, a sound that would only be detectable over several hundreds of meters when propagating through air is detectable over dozens of kilometres underwater. (Zimmer, 2011). The dominant unit used in underwater acoustics is the decibel unit (dB), which is the ratio of sound intensities on the base-10 logarithmic scale ($\log_{10}$; Jensen et al., 2011). Expressing an absolute intensity can be done by using a (common) reference intensity. The currently accepted reference intensity is one micropascal (μPa), which means that intensities are often expressed in dB re 1 μPa. For more details on sound intensities see Jensen et al. (2011, Section 1.3.2)). When a sound is produced in a medium and spreads, the initial energy in that sound is spread out over an increasing surface. The way a sound loses energy as it travels through the medium is called *propagation* or *transmission loss* (Zimmer, 2011). Transmission loss (TL) is derived as the change in sound intensity relative to the sound intensity at 1 meter from the source (dB re 1 m):

$$\mathrm{TL} = \mathrm{SL} - \beta_r \times \log_{10}(x) \qquad (\text{dB re 1 m}), \tag{1.2}$$

where SL is the source level measured at one meter from the source, x is the distance travelled by the sound wave from the source, and β_r is the transmission loss parameter. At small scales the spread is often approximately spherical, which is captured by $\beta_r = 20$, whereas at larger scales the spread is approximately cylindrical, which is captured by $\beta_r = 10$ (Zimmer, 2011, Section 3.3.2). One could imagine that the true energy loss is often, if not always, somewhere between these two 'perfect' scenarios. However, this is not the case. There are a variety of environmental conditions that affect sound attenuation in various ways, which can lead to higher energy loss, i.e., $\beta_r > 20$ (Jensen et al., 2011). Especially in air the transmission loss per meter travelled is generally high, as air is a gas and thus absorbs much of the energy of the sound wave. Water is more dense and absorbs much less energy (Jensen et al., 2011). Many underwater animals therefore communicate and navigate using sounds, by producing clicks, moans, songs, and other vocalisations. We can sample/detect these sounds using underwater microphones, or 'hydrophones'. These can, for example, be placed on fixed arrays, towed behind boats, or be used manually by researchers. The use

of hydrophones to passively record acoustics signals, so-called 'passive acoustics', has become increasingly popular as it offers four main advantages over other sampling techniques (Marques et al., 2013). Firstly, animals that produce loud and/or frequent sounds can be detectable at much greater distances than for other sampling techniques. Second, passive acoustics work under any light conditions. A third reason is that passive acoustics are amenable to the use of automated data processing tools, allowing for the processing of vast amounts of data in little time. Lastly, this automated data collection means that data can be collected in places that are too hostile for human observers.

1.5.2 Genetics

Genetics is a field that focuses on the study of genes and heredity, i.e., how certain traits are passed on genetically from parent to offspring. The field has grown greatly over the past 20 years due to technological advancements (Nature, 2010) and the completion of the Human Genome Project (Durmaz et al., 2015). These advancements have resulted in a cost reduction of more than 99.9% to sequence ('read in') a single human genome (Wetterstrand, 2023), making the use of large-scale genetic data much more feasible. Genetic information is generally stored in deoxyribonucleic acid (DNA; Hartl, 2020). DNA is a double-stranded molecule that is composed of a long strain of four bases/nucleotides/constituents which are abbreviated to A (for adenine), G (for guanine), T (for thymine), and C (for cytosine). DNA is stored in chromosomes, of which humans generally have 23 pairs stored in our cells. A specific subset of the DNA that is transmitted from parent to offspring is called a gene, and the total DNA in a cell is called the genome. The position of a gene along the chromosome is called the 'locus' of that gene. The set of genes present in an individual is called the genotype, and the physical expression of this genotype is called the phenotype (Hartl, 2020).

Genes exist in different variants or forms called 'alleles'. Diploid organisms (such as humans) carry two alleles at every locus, one inherited from the mother and one from the father (Hartl, 2020). Polymorphisms are genetic differences that are common among individuals of the same species; genetic divergence constitutes the genetic diversity that accumulates between species. A single-nucleotide polymorphism (SNP) is one in which individuals within a population can differ in which nucleotide pair is present at a locus. An SNP thus defines two alleles, and the analysis of these SNPs can be

used to make inferences about the relationship/kinship among individuals within a population.

Until recently, genetics has mainly been used as a way to uniquely identify individuals for capture-recapture studies (e.g., Woods et al., 1999; Mills et al., 2000; Lukacs and Burnham, 2005; Burgar et al., 2018), for example as an alternative to visual or acoustic surveys. However, the genetic profile of an individual holds more information than simply a mark—it also holds information about the genetic relatedness to others in the population (see above). In 2001, Skaug published the first version of a method that could use the frequency of allele occurrence to estimate the abundance of animals in a capture-recapture framework. This method generally relied on the sequencing of many genetic samples, which was too costly at the time. In the decade that followed, however, the reduction in sequencing costs stimulated further development of the method, which was formalised in 2016 as close-kin mark-recapture (CKMR; Bravington et al., 2016b). The main idea behind this method is that probability of certain types of relatedness/kinship among individuals (e.g., parent-offspring, half-sibling) can be expressed as a function of the expected relative reproductive output (ERRO) of an adult, which is strongly related to the total adult population size (Bravington et al., 2016b). CKMR therefore, can only be used to estimate the size of the reproductive population; other methods and/or assumptions are required to estimate the total abundance. More details on this method are presented in Chapter 4.

1.5.3 Photo and video

Over the past several decades, researchers have increasingly used remote photo and video to study wildlife (Rowcliffe and Carbone, 2008; Gilbert et al., 2021). One type of use involves camera trapping, which means that cameras take at least one photo or a short video whenever they are triggered by a motion sensor. Camera trapping has been used to study various elements of wildlife ecology, such as behaviour, abundance, distribution and community structure (Rovero et al., 2013). Most of the studies that estimate absolute abundance from camera traps did so only for marked species (Burton et al., 2015; Gilbert et al., 2021). Nearly all studies on unmarked populations only estimated relative abundance or focused on occupancy modelling (MacKenzie et al., 2002). Another method that estimates abundance without the need for (natural) marks is random encounter modelling (REM; Rowcliffe et al., 2008). REM estimates

density from encounter rates on camera systems by assuming that animals' movement can be approximated by the way gas molecules move around in space. In Chapter 5 we present more detail on REM. Moreover, we present a study that explores potential to using REM to estimate density from baited camera traps, which is challenging as REM relies on the assumption that animals move independently with respect to the cameras.

1.6 Statistical notation

Throughout this work, we do not distinguish between probability mass functions (pmfs) or probability density functions (pdfs)—both are denoted $f(x)$, where x denotes the random variable. We try to keep notation as clear and consistent as possible, whilst also respecting the existing literature on the various topics. This inevitably leads to some contradicting (re-)use of variables between chapters. These changes are always made explicit; in fact, variables are clearly defined when first introduced in each chapter, even if they have been declared previously in this work. In general, we do not distinguish between random variables and observed values. This distinction should be clear from context, or is made explicitly where the context does not make it clear.

1.7 Summary of thesis

This thesis is structured following three case studies: acoustic recordings of a migrating population of bowhead whales north of Alaska and Canada (Chapters 2 and 3), genetic samples of a grey reef shark population from the tropical Palmyra Atoll (Chapter 4), and baited video recordings of sharks that live around Chagos Archipelago (Chapter 5). Chapter 6 consists of a closing discussion and remarks on future research. Below we present summaries of the three case studies.

1.7.1 Acoustic spatial-capture recapture for migrating bowhead whales in the Beaufort Sea

Passive acoustic monitoring is a promising method for surveying wildlife populations that are easier to detect acoustically than visually. When animal vocalisations can be

uniquely identified on an array of sensors, the potential exists to estimate population density through acoustic spatial capture-recapture (ASCR). However, sound classification is imperfect, and in some situations a high proportion of sounds detected on just a single sensor ('singletons') are not from the target species. We present a case study of bowhead whale calls (*Balaena mysticetus*) collected in the Beaufort Sea in 2010 containing such false positives. In Chapter 2, we propose a novel extension of ASCR that is robust to false positives by truncating singletons and conditioning on calls being detected by at least two sensors. We allow for individual-level detection heterogeneity through modelling a variable sound source level, model inhomogeneous call spatial density, and include bearings with varying measurement error. We show via simulation that the method produces near-unbiased estimates when correctly specified. In Chapter 3, we propose a further extension of the model that involves a more sophisticated detection function using the signal-to-noise ratio. This involves more methodological complexity and additional information on (ocean) noise.

1.7.2 Close-kin mark-recapture for grey reef sharks on Palmyra Atoll

Close-kin mark-recapture (CKMR) relies on the observed frequency and type of kinship among individuals sampled from the population to estimate population size. Knowledge of the age of the individuals, or a surrogate thereof, is essential for inference with acceptable precision. One common approach, particularly in fish studies, is to measure animal length and infer age using an assumed age-length relationship (a 'growth curve').

In Chapter 4, we present a simulation study to test the effect of misspecifying the length measurement error and the growth curve on population size estimation. Simulated populations represented two fictional shark species, one with a relatively simple life history and the other with a more complex life history based on the grey reef shark (*Carcharhinus amblyrhynchos*). We estimated sex-specific adult abundance, which we assumed to be constant in time. We observed small median biases in these estimates ranging from 1.35% to 2.79% when specifying the correct measurement error standard deviation and growth curve. CI coverage was adequate whenever the growth curve was correctly specified. Introducing error via misspecified growth curves resulted in changes in the magnitude of the estimated adult population, where underestimating age negatively biased the abundance estimates. Over- and underestimating the standard

deviation of length measurement error did not introduce a bias and had negligible effect on the variance in the estimates.

1.7.3 bREM: adapting random encounter modelling for data from baited remote underwater video systems

Abundance estimation from camera footage has mostly been undertaken using capture-recapture methods for marked populations. Random encounter modelling (REM) can estimate density and absolute abundance for unmarked populations from the encounter rate on camera traps. One of the main assumptions is that animals move around randomly with respect to the cameras. In Chapter 5, we explore the potential of using REM for baited cameras. The bait is placed in front of the camera to lure animals in, resulting in an inflated encounter rate relative to unbaited cameras. When unaccounted for, this leads to positively biased estimates of density and absolute abundance. This study was motivated by footage from baited remote underwater video (BRUV) systems of several shark species on and around the Chagos Archipelago, a remote island system in the Indian Ocean.

One challenge introduced by the bait is that the bait plume (i.e., the area over which scent from the bait can be detected) covers a much larger area than the field-of-view (FOV) of the cameras. Accurately estimating the bait plume area therefore is essential to properly estimate the sampled area, which in turn is required to estimate density. It also means that the encounter moment is no longer the moment the shark enters the FOV of the camera, but rather the moment it encounters the plume. We estimate this moment, we find an analytical solution to the expected distance from the camera for a given bait plume. Moreover, we perform a simulation study to evaluate how a bait plume spreads under various current conditions.

1.8 Data and code availability

The code and data used for model fitting and simulation in Chapter 2 are publicly available at `https://www.doi.org/10.5281/zenodo.10848178`.

The code used for simulation in Chapter 4 is publicly available at `https://www.doi.org/10.5281/zenodo.10727210`.

Chapter 2

Accommodating false positives within acoustic spatial capture-recapture, with variable source levels, noisy bearings and an inhomogeneous spatial density

2.1 Introduction

In recent decades, passive acoustic monitoring (PAM) has increasingly been used to study both terrestrial (Sugai et al., 2019) and marine animals, particularly cetaceans (Zimmer, 2011). Compared with more traditional visual survey methods, acoustic monitoring works day and night, is robust to variation in environmental conditions such as weather, and in some habitats has the potential to detect animals at greater distances, hence increasing the area surveyed (Marques et al., 2011). It has enabled studies of rare and elusive species such as the vaquita (Thomas et al., 2017) and several species of beaked whale (Hildebrand et al., 2015; Yack et al., 2013) which, despite being visually cryptic, produce frequent sounds that can be detected by PAM systems.

One important application of PAM is to estimate population density or abundance (Marques et al., 2013). In the situation where multiple acoustic sensors are deployed simultaneously with a spatial separation such that some vocalisations can be detected on multiple sensors then an appropriate statistical framework for estimating density is spatial capture-recapture (SCR; also known as spatially explicit capture-recapture) (e.g., Borchers et al., 2015; Stevenson et al., 2015). SCR is an extension of long-established capture-recapture (otherwise known as mark-recapture or mark-resight) methods where data on the detection ('capture') of individually-identified animals is supplemented by data on the spatial location of the survey effort and the detections (Borchers and Efford, 2008; Royle et al., 2009; Kidney et al., 2016). Recording not just whether but also where each animal was detected increases the accuracy of abundance and density estimates and potentially allows estimation of a spatially inhomogeneous animal density surface.

Acoustic spatial capture-recapture (ASCR) is a special case of SCR where the detections are of individual animal vocalisations or 'cues'. Standard SCR relies on animals moving between detection locations and hence typically requires multiple capture occasions, while in ASCR the sound travels almost instantaneously from its source and hence can be detected on multiple sensors in a single occasion. Estimation is of cue spatial density; to convert to animal density an estimate of average cue production rate is required (Marques et al., 2013; Stevenson et al., 2021). As well as the location of detection, additional information is often available about the location of the vocalisation, e.g., the bearing, received sound level or time of arrival on multiple sensors. Borchers et al. (2015) showed that using this additional information further improved estimation accuracy.

PAM data can also present challenges that need to be accounted for in order to avoid bias in ASCR analysis. First, sound classification is imperfect, leading to false positive detections of sounds not originating from the target species. Second, vocalisation source level can vary considerably, causing heterogeneity in detectability. Third, there can be considerable measurement error in the additional information, particularly the bearings. Fourth (in common with other SCR studies), spatial density of source locations can vary substantially.

The methods presented here are motivated by a case study that demonstrates all four

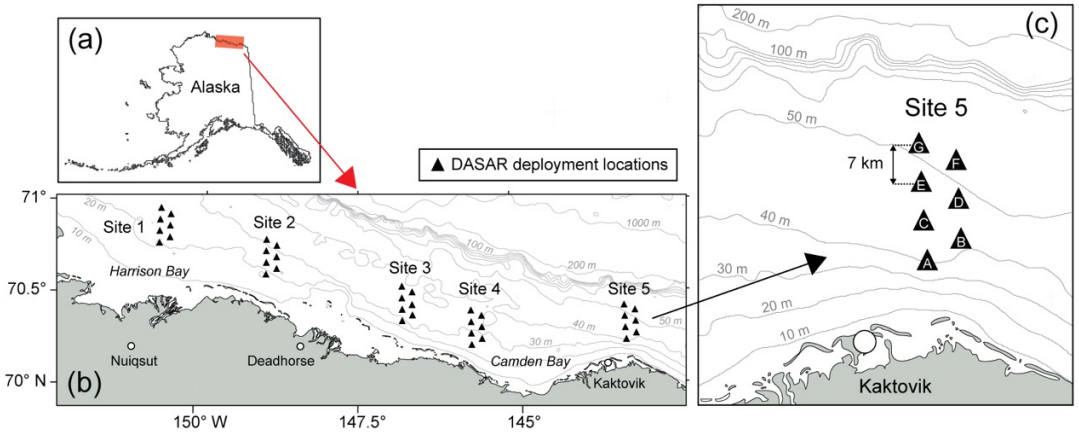


FIGURE 2.1: Map of the DASAR deployment locations from 2008-2014, with a detailed view of site 5. In 2010, detector F at site 5 malfunctioned and its data was omitted from the analysis. This image was originally presented in Thode et al. (2020).

of the above issues: estimation of call density of bowhead whales (*Balaena mysticetus*) migrating through the Beaufort Sea. Multiple arrays of acoustic sensors were deployed in the Beaufort Sea during the migration season and recorded millions of bowhead whale calls (see Figure 2.1). Automated detection and classification methods were therefore used to process the data, yielding call detections, received sound levels and bearings, and linking calls across detectors. However, a high proportion of the detections made on only one sensor ('singletons') were thought to be false positives. Naively including these singletons in the analysis would lead to a positive bias in the estimated abundance of unknown but likely substantial magnitude. To avoid this, we excluded all singletons from the analysis and conditioned the ASCR likelihood to only include calls detected on multiple sensors ('multiply-detected calls'). Truncating the data in this manner, rather than attempting to model the proportion of false positives, is a good strategy when data are plentiful (see Discussion). Conn et al. (2011) proposed a similar procedure in a mark-resight study as a way to differentiate between resident and transient bottlenose dolphin populations, assuming that transients were never detected more than once. To our knowledge, this is the first time this approach has been used in SCR.

Our case study features some additional complications that may commonly appear in real-world data but are not typically all dealt with. Call source level was thought to vary substantially and so we include source level as a random effect in our model. Exploratory analysis showed that, while most estimated bearings were accurate, some

appeared to be very inaccurate. We therefore developed a two-part discrete mixture model for bearing error, extending the bearing error methods of Borchers et al. (2015). Finally, we allowed for an inhomogeneous density model to accommodate the spatial preference of the migrating whales (and thus their calls).

In Section 2.2 we describe the case study in more detail. We then present the extended ASCR model in Section 2.3. In Section 2.4 we evaluate the model performance via simulation, and show how ignoring some of the real-world issues can result in substantial bias. Section 2.5 gives results from a proof-of-concept application of the model to a single day of data. Finally, in Section 2.6, we discuss results, limitations, and alternatives.

2.2 Case study

Every year from August to October, the Bering–Chukchi–Beaufort population of bowhead whales (*Balaena mysticetus*) migrates westwards through the Beaufort Sea to their wintering areas in the Bering Sea (Blackwell et al., 2007). They travel mainly over the continental shelf in waters less than 25 meters deep, approximately 10–75 km offshore (Greene et al., 2004). During this migration, bowhead whales are known to produce a wide variety of calls (Ljungblad et al., 1982). The purpose of these calls remains largely unknown, although they may be used for long-range communication (Thode et al., 2020) or to navigate through ice (George et al., 1989).

Between 2007 and 2014, up to 35 Directional Autonomous Seafloor Acoustic Recorders (DASARs; Greene et al., 2004) were deployed at several sites off the north coast of Alaska to monitor the calling behaviours of the migrating whales during seismic surveys. An automated detection and classification procedure was developed to handle the more than one million detected calls over the period 2007–2014 (Thode et al., 2012, 2020). This procedure could identify discrete sounds as bowhead whale calls, and subsequently link them with detections from other DASARs within the array if they were the same call (Thode et al., 2012).

Even though it was not the original purpose of the monitoring, the detection and linking of calls created detection histories for every detected call, making these data suitable for ASCR. ASCR theory assumes capture histories without errors, i.e., calls can be

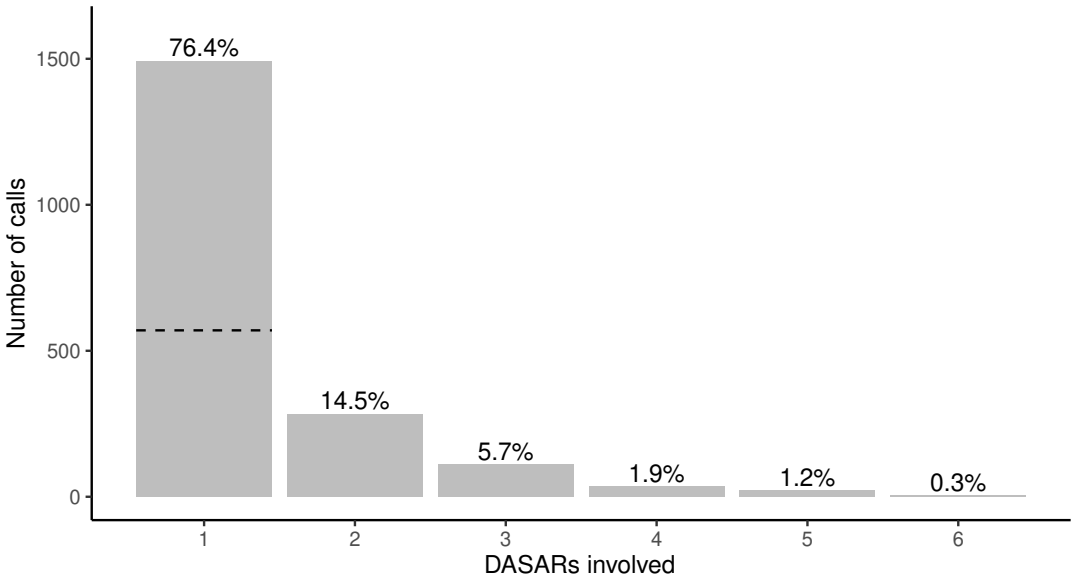


FIGURE 2.2: Counts of detected calls by number of DASARs (sensors) involved per call at site 5 on 31 August 2010. Detections were included if the received level was at least 96 dB. The high proportion of singletons (detections on a single DASAR; 1417) raised concerns about the validity of those calls. The dashed line shows the number of singletons (570) that were estimated to be valid using the methods developed in this paper.

missed but not incorrectly positively identified or matched. Several data cleaning procedures were required to meet this assumption as far as possible. Sometimes, calls would be wrongly identified as other discrete sounds. Distant airgun signals were occasionally misidentified as bowhead whale calls; bearded seals (*Erignathus barbatus*) and walruses (*Odobensu rosmarus divergens*) could appear similar as well, but these were generally rare and much quieter (Thode et al., 2012). Moreover, if bowhead whale calls overlapped in time, they could be incorrectly matched as being the same call. Cheoo (2019) showed that the number of detection histories that involved just one sensor ('singletons') in the automated data was not in line with expectations based on sound propagation theory (see Figure 2.2). It was therefore hypothesised that these singletons contain a high-degree of incorrectly classified calls ('false positives'), mostly consisting of random fluctuations in the noise field or reverberations of whale calls, which were unlikely to be associated among multiple detectors. The solution we present here is to exclude singleton detection histories all together and modify the ASCR likelihood to be conditional on the capture histories involving at least two sensors. To ensure that the multiply-detected calls would contain no false positives, we cleaned

the remaining data using a procedure described in Appendix A in Supplementary Materials.

In this study, we focus on the data from one specific day, 31 August 2010, from site 5, the most eastward site. Calls accumulated somewhat evenly across the day, resulting in a low expected rate of overlapping calls. The array consisted of six functioning sensors spaced at 7 kilometres from each other (see Figure 2.1). For every call the following data were recorded: a detection history with the recorded bearings and received sound levels, as well as noise levels for every involved DASAR. The source level and origin of the call are unobserved, and hence treated as latent variables. Throughout this chapter, all sound measurements are denoted on the decibel scale (RMS; dB re 1 μPa). More details on the background, availability, and pre-processing of the data are presented in Appendix A in Supplementary Materials.

2.3 Model

In this section, we first introduce the full likelihood. Following that, we derive each element of the likelihood individually. Consider an acoustic survey of an array of K sensors operating within a survey region $\mathcal{A} \subset \mathbb{R}^2$ over a period of time T. A total of n unique animal calls are detected by at least two sensors. For any multiply-detected call $i \in 1, ..., n$, let ω_{ij} be 1 if the call was detected at sensor $j \in 1, ..., K$, and 0 otherwise. We define the matrix containing all detection histories as $\boldsymbol{\Omega} = (\boldsymbol{\omega}_1, ..., \boldsymbol{\omega}_n)^\top$, where the detection history for call i is denoted $\boldsymbol{\omega}_i = (\omega_{i1}, ..., \omega_{iK})^\top$ and $\top$ denotes the transpose. For every call i that was detected at sensor j, we also observe the bearing y_{ij}, measured in degrees clockwise relative to true north, and the received sound level r_{ij}. These data are contained in $\boldsymbol{Y} = (\boldsymbol{y}_1, ..., \boldsymbol{y}_n)^\top$ and $\boldsymbol{R} = (\boldsymbol{r}_1, ..., \boldsymbol{r}_n)^\top$, respectively. Latent variables are the spatial origins of calls, denoted by $\boldsymbol{X} = (\boldsymbol{x}_1, ..., \boldsymbol{x}_n)^\top$ where $\boldsymbol{x}_i$ is a location in the Cartesian plane, and the source levels, denoted by $\boldsymbol{s} = (s_1, ..., s_n)^\top$. The support for s is denoted $\mathcal{S}$. Throughout this chapter, we do not distinguish explicitly in notation between random variables and specific observations/realisations – this should be clear from the context.

The parameter vectors used in the model are (following previous literature as closely as possible): $\boldsymbol{\phi}$ for parameters associated with the spatial density of emitted calls, $\boldsymbol{\nu}$

for those associated with the source levels of calls, $\boldsymbol{\eta}$ for those associated with source sound propagation and received levels, $\boldsymbol{\theta}$ for those associated with detectability of calls, and $\boldsymbol{\gamma}$ for those associated with measurement error of the bearings. For notational convenience we define the joint parameter vector $\boldsymbol{\xi} = (\boldsymbol{\phi}, \boldsymbol{\nu}, \boldsymbol{\eta}, \boldsymbol{\theta}, \boldsymbol{\gamma})$. A list of model assumptions is presented in Section 2.3.10.

TABLE 2.1: Summary of notation.

Symbol	Description	Type
	General	
n	Number of multiply-detected calls	Observed
p	Detection probability of an emitted call	Function
t_r	Detection threshold for calls	Known
w	Detection indicator	Observed
r	Received level	Observed
y	Bearing	Observed
$\boldsymbol{x}$	Spatial origin	Latent
s	Source level	Latent
$\mathcal{S}$	Support for source level	Support
f	Probability density/mass function	Function
$\mathcal{L}_f, \mathcal{L}_c$	Full likelihood, conditional likelihood	Function
	Quantities related to the survey	
K	Total number of sensors	Observed
$\mathcal{A}$	Study area	Support
T	Study period	Observed
	Estimable parameters	
$\boldsymbol{\phi}$	Spatial density of emitted calls	Parameter vector
$\boldsymbol{\eta}$	Source sound propagation	Parameter vector
$\boldsymbol{\nu}$	Distribution of source levels	Parameter vector
$\boldsymbol{\theta}$	Detectability of calls	Parameter vector
$\boldsymbol{\gamma}$	Measurement error of bearings	Parameter vector
$\boldsymbol{\xi}$	All estimable parameters	Parameter vector
	Subscript and sets	
i	Call index	
j	Sensor index	
$\boldsymbol{R}$	Set of all received levels	Observed
$\boldsymbol{\Omega}$	Set of all detection histories	Observed
$\boldsymbol{Y}$	Set of all bearings	Observed
$\boldsymbol{X}$	Set of all spatial origins	Latent
$\boldsymbol{s}$	Vector of all source levels	Latent

2.3.1 Likelihood specification

The likelihood is formed from the joint distribution of all observed and latent variables introduced above. We denote this likelihood $\mathcal{L}_f(\boldsymbol{\xi})$ where the subscript f stands for 'full' to distinguish it from the conditional likelihood $\mathcal{L}_c(\boldsymbol{\xi})$ which is conditional on the observed number of detections. Denoting both probability mass and density functions as $f(.)$, we define the full likelihood as

$$\begin{aligned}\mathcal{L}_f(\boldsymbol{\xi}) &= f(n, \boldsymbol{\Omega}, \boldsymbol{Y}, \boldsymbol{R}, \boldsymbol{X}, \boldsymbol{s}; \boldsymbol{\xi}) \\ &= f(n; \boldsymbol{\phi}, \boldsymbol{\theta}, \boldsymbol{\nu}) \times f(\boldsymbol{\Omega}, \boldsymbol{Y}, \boldsymbol{R}, \boldsymbol{X}, \boldsymbol{s} | n; \boldsymbol{\xi}) \\ &= f(n; \boldsymbol{\phi}, \boldsymbol{\theta}, \boldsymbol{\nu}) \times \mathcal{L}_c(\boldsymbol{\xi}).\end{aligned} \tag{2.1}$$

As we do not observe the call locations and source levels, we marginalise over the unobserved $\boldsymbol{X}$ and $\boldsymbol{s}$ to obtain

$$\mathcal{L}_c(\boldsymbol{\xi}) = \int_{\mathcal{S}} \int_{\mathcal{A}} f(\boldsymbol{\Omega}, \boldsymbol{Y}, \boldsymbol{R}, \boldsymbol{X}, \boldsymbol{s} | n; \boldsymbol{\xi}) d\boldsymbol{X} d\boldsymbol{s}. \tag{2.2}$$

The double integral in Equation (2.2) is of dimension $3n$, making this likelihood intractable. We follow Stevenson et al. (2015) in assuming calls to be independent of each other in all respects, allowing us to specify Equation (2.2) as the product of n 3-dimensional integrals:

$$\begin{aligned}\mathcal{L}_c(\boldsymbol{\xi}) &\equiv \prod_{i \in \{1:n\}} \int_{\mathcal{S}} \int_{\mathcal{A}} f(\boldsymbol{\omega}_i, \boldsymbol{y}_i, \boldsymbol{r}_i, \boldsymbol{x}_i, s_i | \omega_i^* \geq 2; \boldsymbol{\xi}) d\boldsymbol{x}_i ds_i \\ &= \prod_{i \in \{1:n\}} \int_{\mathcal{S}} \int_{\mathcal{A}} f(\boldsymbol{\omega}_i | \boldsymbol{x}_i, \boldsymbol{s}_i, \omega_i^* \geq 2; \boldsymbol{\theta}) \\ &\qquad \times f(\boldsymbol{y}_i | \boldsymbol{\omega}_i, \boldsymbol{x}_i; \boldsymbol{\gamma}) \times f(\boldsymbol{r}_i | \boldsymbol{\omega}_i, \boldsymbol{x}_i, s_i; \boldsymbol{\eta}) \\ &\qquad \times f(\boldsymbol{x}_i | s_i, \omega_i^* \geq 2; \boldsymbol{\phi}, \boldsymbol{\theta}) \\ &\qquad \times f(s_i | \omega_i^* \geq 2; \boldsymbol{\phi}, \boldsymbol{\theta}, \boldsymbol{\nu}) d\boldsymbol{x}_i ds_i\end{aligned} \tag{2.3}$$

where we assume independence between $\boldsymbol{y}_i$ and $\boldsymbol{r}_i$ given $\boldsymbol{x}_i$, and $\omega_i^* := \sum_{j \in 1:k} w_{ij}$ denotes the total number of sensors involved in the detection of call i. The separation of the joint distribution inside the integrals in Equation (2.3) follows from repeatedly applying Bayes' formula. Note that conditioning the joint distribution on n in Equation (2.2) is equivalent to conditioning every marginal distribution on involving at least two

sensors in Equation (2.3); for the second and third element this conditioning is implicit in conditioning on detection history $\boldsymbol{\omega}_i$. If we were to assume a constant spatial density of calls, it would be sufficient to simply maximise $\mathcal{L}_c$, as the parameter estimates from the conditional MLE are identical to those obtained by maximising the full likelihood – a Horvitz Thompson-like estimator could then be used to derive the mean density (Borchers and Efford, 2008). In our case study, however, the spatial density of calls is known to be inhomogeneous, so the full likelihood is required.

In the following sections, we specify in further detail $f(n; \boldsymbol{\phi}, \boldsymbol{\theta}, \boldsymbol{\nu})$ and the five components in Equation (2.3). For readability we will hereforth omit the indexing of parameters in the probability functions.

2.3.2 Detection probability, $p(\boldsymbol{x}_i, s_i)$

A fundamental part of ASCR is the concept of a detection probability, which is the probability that an emitted call is detected by a sensor—this is the way ASCR accommodates for missed calls, i.e., false negatives. The function that relates this probability to covariates is called the detection function, denoted $p(.)$. For ASCR it was proposed to be most appropriate to use a detection function based on the received (sound) level, also known as signal strength, which is primarily a function of source level s_i and range (i.e., distance to the origin of the sound, $d(\boldsymbol{x})$) (Efford et al., 2009; Stevenson et al., 2015). We propose a detection function for sensor j where $g_0 \in (0, 1)$ denotes the detection probability when the true received level of a call surpasses threshold t_r, and zero probability else, as follows

$$p_j(\boldsymbol{x}_i, s_i) = g_0 \times \left(1 - \Phi\left(\frac{t_r - \mathbb{E}[r_{ij}|\boldsymbol{x}_i, s_i])}{\sigma_r}\right)\right), \tag{2.4}$$

where Φ denotes the standard normal cumulative distribution function (cdf) and σ_r denotes the measurement and propagation error on received level (see Section 2.3.4). The threshold t_r should be set by the analyst, and in this study is roughly equal to the maximum level of ocean background noise. As there is error on the received levels, calls with an expected received level close to t_r can have a detection probability between 0 and g_0. For example, we get a detection probability of $0.5 \times g_0$ whenever the expected received level equals the threshold.

2.3.3 Detected calls, $f(n)$

To construct a probability function for the number of detected calls, we start with the distribution of emitted calls, which are assumed to occur independently of one another in space and time. Thus, let the number of emitted calls at point $\boldsymbol{x}$ in period T be a realisation of a spatially-inhomogeneous Poisson point process with intensity $D(\boldsymbol{x})$. As we only observe multiply-detected calls, we take the product of this intensity and the probability of a detection involving multiple sensors, denoted by

$$\begin{aligned} p.(\boldsymbol{x}_i, s_i) &= \mathbb{P}(\omega^* \geq 2 | \boldsymbol{x}_i, s_i) \\ &= 1 - \mathbb{P}(\omega^* = 0 | \boldsymbol{x}_i, s_i) - \mathbb{P}(\omega^* = 1 | \boldsymbol{x}_i, s_i), \end{aligned} \tag{2.5}$$

where $\omega^* := \sum_{j \in 1:k} \omega_j$. Note that Equation (2.5) is the part of the method that deviates from conventional ASCR and allows us to exclude all singletons. We rewrite Equation (2.5) and marginalise over source level to get the filtered Poisson point process with spatially varying rate parameter $\int_{\mathcal{S}} D(\boldsymbol{x}) p.(\boldsymbol{x}, s) f(s) ds$. Lastly, we marginalise over $\boldsymbol{x}$ to get the distribution of the total number of multiply-detected calls over time T:

$$n \sim \text{Poisson}\left(\int_{\mathcal{A}} \int_{\mathcal{S}} D(\boldsymbol{x}) p.(\boldsymbol{x}, s) f(s) ds d\boldsymbol{x} \right). \tag{2.6}$$

The total number of (emitted) calls is then estimated by integrating the estimated density over the study area, such that $\widehat{N} = \int_{\mathcal{A}} \widehat{D}(\boldsymbol{x}) d\boldsymbol{x}$..

2.3.4 Received level, $f(\boldsymbol{r}_i | \boldsymbol{\omega}_i, \boldsymbol{x}_i, s_i; \boldsymbol{\eta})$

Sound waves propagating through water lose strength through various mechanisms (Jensen et al., 2011), a process known as 'transmission loss'. We approximate this process by allowing a single parameter β_r to determine the acoustic transmission loss, such that

$$\mathbb{E}[r_{ij} | \boldsymbol{x}_i, s_i] = s_i - \beta_r \log_{10}(d_j(\boldsymbol{x}_i)), \tag{2.7}$$

where $d_j(\boldsymbol{x}_i)$ returns the distance from sensor j to location $\boldsymbol{x}_i$. Here, $\beta_r = 20$ would reflect purely spherical spreading loss while $\beta_r = 10$ would reflect cylindrical spreading loss; in reality the dominant process will be range- and depth-dependent, with other factors also playing a role (Jensen et al., 2011). To capture potential error in the propagation model and the measurement of received level, we follow Stevenson et al.

(2015) and assume Gaussian error on the received levels, giving

$$r_{ij}|\boldsymbol{x}_i, s_i \sim \mathcal{N}\left(\mathbb{E}[r_{ij}|\boldsymbol{x}_i, s_i], \sigma_r^2\right) \tag{2.8}$$

where $\mathcal{N}\left(\mu, \sigma^2\right)$ denotes a normal distribution with mean μ and variance σ^2. Unlike Stevenson et al. (2015), we do not assume all calls above threshold t_r to be detected with certainty. Instead, we allow for a single detection probability for calls with received levels above the threshold, since the signal processing used in our case study meant that other factors also determined detectability (see Section 2.3.2). As r_{ij} is only recorded if sensor j detected the call, we condition the indexing on the j^{th} DASAR detecting the call and thus only evaluate cases where $\omega_{ij} = 1$. Assuming independence between the sensors, the third component of Equation (2.3) becomes

$$\begin{aligned} f(\boldsymbol{r}_i|\boldsymbol{\omega}_i, \boldsymbol{x}_i, s_i; \boldsymbol{\eta}) &= \prod_{j \in \{1:k|\omega_{ij}=1\}} f(r_{ij}|\omega_{ij} = 1, \boldsymbol{x}_i, s_i) \\ &= \prod_{j \in \{1:K|\omega_{ij}=1\}} \frac{1}{\sigma_r} \frac{\phi\left((r_{ij} - \mathbb{E}[r_{ij}|\boldsymbol{x}_i, s_i])/\sigma_r\right)}{1 - \Phi((t_r - \mathbb{E}[r_{ij}|\boldsymbol{x}_i, s_i])/\sigma_r)}, \end{aligned} \tag{2.9}$$

where ϕ denotes the standard normal probability density function (pdf). This is in effect a normal distribution truncated at t_r. An alternative formulation would be to take the product over all K sensors but set $f(r_{ij}|\omega_{ij}, \boldsymbol{x}_i, s_i)$ equal to 1 whenever $w_{ij} = 0$ (see Stevenson et al., 2015). The reason that no detection probability appears in Equation (2.9) is that we condition on a positive detection.

2.3.5 Bearings, $f(\boldsymbol{y}_i|\boldsymbol{\omega}_i, \boldsymbol{x}_i; \boldsymbol{\gamma})$

DASARs were designed to record the direction to discrete sounds; these recorded bearings contain measurement errors. Following Stevenson et al. (2015) and Borchers et al. (2015), we capture this error by assuming a Von Mises distribution with concentration parameter κ on the bearings. Analogous to the received levels, we only record bearings at DASARs that detected a call. Assuming that the sensors are independent of each other, the second component of Equation (2.3) would therefore be

$$\begin{aligned} f(\boldsymbol{y}_i|\boldsymbol{\omega}_i, \boldsymbol{x}_i; \boldsymbol{\gamma}) = \\ \sum_{j \in \{1:K|w_{ij}=1\}} \frac{\exp\{\kappa \cos(y_{ij} - \mathbb{E}[y_{ij}|\boldsymbol{x}_i])\}}{2\pi I_0(\kappa)}, \end{aligned} \tag{2.10}$$

with I_0 denoting the modified Bessel function of degree 0. Exploratory research found that, while most bearings appeared relatively accurate, a small proportion seemed highly inaccurate, which could have been the result of i) only a fraction of the call getting captured by the measurement window of the sensor, or ii) a low signal-to-noise ratio at the sensor (the latter should be rare as we truncated the data at a sound level equal to the highest noise observed). Moreover, some of the mismatched calls could still have been present in the data, thus leading to disagreement among the bearings. To accommodate these inaccurate bearings we apply a 2-part discrete mixture model on the bearings. In effect, we fit a Von Mises distribution with a lower accuracy (dispersion is κ) for some proportion of the bearings, ψ_κ, and a Von Mises distribution with higher accuracy (dispersion is $\kappa + \delta_\kappa$, where increment δ_κ is non-negative) to the remaining proportion of the bearings, $1-\psi_\kappa$. The second component of Equation (2.3) thus becomes

$$\begin{aligned} f(\boldsymbol{y}_i|\boldsymbol{\omega}_i, \boldsymbol{x}_i; \boldsymbol{\gamma}) = & \\ & \sum_{j\in\{1:K|w_{ij}=1\}} \psi_\kappa \frac{\exp\{\kappa\cos(y_{ij} - \mathbb{E}[y_{ij}|\boldsymbol{x}_i])\}}{2\pi I_0(\kappa)} \\ & \times (1-\psi_\kappa)\frac{\exp\{(\kappa+\delta_\kappa)\cos(y_{ij} - \mathbb{E}[y_{ij}|\boldsymbol{x}_i])\}}{2\pi I_0(\kappa+\delta_\kappa)}. \end{aligned} \tag{2.11}$$

2.3.6 Call location given s and at least two detections, $f(\boldsymbol{x}_i|s_i, \omega_i^* \geq 2; \boldsymbol{\phi}, \boldsymbol{\theta})$

To evaluate the pdf of call locations, we assume independence between call locations and source levels, and use Bayes' formula to obtain

$$\begin{aligned} f(\boldsymbol{x}_i|s_i, \omega_i^* \geq 2; \boldsymbol{\phi}, \boldsymbol{\theta}) = & \frac{f(\omega_i^* \geq 2|\boldsymbol{x}_i, s_i) f(\boldsymbol{x}_i|s_i)}{\int_{\mathcal{A}} f(\omega_i^* \geq 2|\boldsymbol{x}, s_i) f(\boldsymbol{x}|s_i) d\boldsymbol{x}} \\ = & \frac{p.(\boldsymbol{x}_i, s_i) f(\boldsymbol{x}_i)}{\int_{\mathcal{A}} p.(\boldsymbol{x}, s_i) f(\boldsymbol{x}) d\boldsymbol{x}}. \end{aligned} \tag{2.12}$$

Although $f(\boldsymbol{x}_i)$ is unknown, we do know that it is proportional to the emitted call density, such that $f(\boldsymbol{x}_i) = D(\boldsymbol{x}_i)/\int_{\mathcal{A}} D(\boldsymbol{x}) d\boldsymbol{x}$. Thus we can simplify Equation (2.12)

as

$$\begin{aligned} f(\boldsymbol{x}_i|s_i, \omega_i^* \geq 2; \boldsymbol{\phi}, \boldsymbol{\theta}) =& \frac{p.(\boldsymbol{x}_i, s_i)D(\boldsymbol{x}_i)/\int_{\mathcal{A}} D(\boldsymbol{x})d\boldsymbol{x}}{\int_{\mathcal{A}} p.(\boldsymbol{x}, s_i)D(\boldsymbol{x})/\int_{\mathcal{A}} D(\boldsymbol{q})d\boldsymbol{q}d\boldsymbol{x}} \\ =& \frac{p.(\boldsymbol{x}_i, s_i)D(\boldsymbol{x}_i)}{\int_{\mathcal{A}} p.(\boldsymbol{x}, s_i)D(\boldsymbol{x})d\boldsymbol{x}}. \end{aligned} \tag{2.13}$$

2.3.7 Source level given at least two detections, $f(s_i|\omega_i^* \geq 2; \boldsymbol{\phi}, \boldsymbol{\theta}, \boldsymbol{\nu})$

Analogous to above, we assume independence between and among call locations and source levels, to derive

$$\begin{aligned} f(s_i|\omega_i^* \geq 2; \boldsymbol{\phi}, \boldsymbol{\theta}, \boldsymbol{\nu}) =& \frac{f(\omega_i^* \geq 2|s_i)f(s_i)}{f(\omega_i^* \geq 2)} \\ =& \frac{\int_{\mathcal{A}} p.(\boldsymbol{x}, s_i)D(\boldsymbol{x})d\boldsymbol{x} \times f(s_i)}{\int_{\mathcal{A}}\int_{\mathcal{S}} p.(\boldsymbol{x}, s)f(s)D(\boldsymbol{x})dsd\boldsymbol{x}}. \end{aligned} \tag{2.14}$$

Note that a part of the numerator in Equation (2.14) cancels out against the denominator in Equation (2.13), and that the denominator of Equation (2.14) denotes the effective sampled area. Thode et al. (2020) estimated source levels using the estimated origin of the call, the received level on the sensor nearest the origin, and a transmission loss parameter β_r of 15. Based on the observed distribution of these estimated source levels, we assume a normal distribution on s_i truncated at 0, such that

$$s \sim \mathcal{N}_0^{\infty}(\mu_s, \sigma_s^2). \tag{2.15}$$

2.3.8 Detection history, $f(\boldsymbol{\omega}_i|\boldsymbol{x}_i, s_i, \omega_i^* \geq 2; \boldsymbol{\theta})$

If we assume sensors to be independent, we can view the detection history of a call as a realisation of a binomial process with size K and non-constant probability p_j with $j = 1, .., K$, where the order is relevant and hence the binomial coefficient is absent. This gives

$$f(\boldsymbol{\omega}_i|\boldsymbol{x}_i, s_i, \omega_i^* \geq 2; \boldsymbol{\theta}) = \frac{\prod_{j=1}^{K} p_j(\boldsymbol{x}_i, s_i)^{\omega_{ij}}(1 - p_j(\boldsymbol{x}_i, s_i))^{1-\omega_{ij}}}{p.(\boldsymbol{x}_i, s_i)}, \tag{2.16}$$

where $p.(\boldsymbol{x}_i, s_i)$ appears in the denominator to account for the conditioning on at least two sensors in every call detection history (see Equation (2.5)).

2.3.9 Variance estimation

We do not use the Hessian matrix to extract standard errors, as these are only asymptotically normal. Instead, we estimate uncertainty using a non-parametric bootstrap, where we re-sample the calls with replacement (Borchers et al., 2002) and fit the model every time. Following that, we estimate the standard error by taking the standard deviation of all bootstrapped parameter estimates, and we take the 2.5% and 97.5% percentiles to estimate their 95% confidence intervals.

2.3.10 Assumptions

The method presented in this chapter relies on several assumptions, as follows. Note that these assumptions apply to all emitted calls irrespective of whether they were detected and how many sensors were involved in a detection.

1. **Call origins are a realisation of a Poisson point process, thus calls are spatially and temporally independent.** Because of this, we assume that the number of calls is a Poisson random variable. The assumption is likely to be violated due to the spatial dependence of a whale (when it calls several times), the theorised communicative nature of the calls, and the possibility that bowhead whales swim in groups (Shelden and Rugh, 1995). Violation likely mainly affects uncertainty estimates, not point estimates. Thinning the data could ensure that this assumption is closer to being met. Alternatively, one could bootstrap by time blocks to ensure more robust variance estimates.

2. **Calls are omnidirectional and equally detectable given only the received level (i.e., no unmodelled heterogeneity).** It is very likely that other factors impact the detectability of these calls, such as the frequency and duration. The direction and severity of a potential bias due to an incorrectly specified detection function is not known.

3. **Sensors are identical in performance and operate independently.** This means that we can model a single detection probability.

4. **Calls are matched without error and identified correctly, but can be missed (i.e., no false positives, but false negatives are allowed).** This is

an essential assumption to the model, as we only allow for false negatives through the concept of a detection probability, but not for false positives. Wrongly identified calls will lead to an overestimate of abundance. Incorrectly matched calls will tend to lead to an underestimate of abundance.

5. **The transmission loss model is correctly specified.** Our propagation model is relatively simple and assumes a single transmission loss parameter, and is surely an oversimplification. A more complex (and realistic) propagation model could be used, potentially calculating propagation loss in the target frequency band of calls between all grid points and all sensors, and then using this as a look-up table in a non-Euclidean SCR framework (Royle, 2018; Phillips, 2016). For the current case study, we believe this would make relatively little difference to the results since bathymetry is the main driver of variation in propagation loss in this area and almost all calls are believed to have originated from the same shallow shelf area as the sensors. However, this a topic worthy of further investigation.

6. **Uncertainty on bearings and received levels are independent.** One way this assumption could be violated is if there are certain characteristics of the water column that affect both the precision of the bearing measurements and the received level measurements. This will likely not affect the point estimates.

7. **Source level is independent of space and time.** This assumption could be violated for a variety of reasons. Firstly, if calls are communicative in nature, then potentially this could make the loudness depend on space and time. Some whales could also be louder than others, which means that the source level of a call is dependent on space and time through the individuals who produced it. Moreover, calls could be louder when ocean background noise is louder, an effect that is also known as the 'Lombard effect' (Thode et al., 2020). When ocean noise is then spatio-temporally independent, so is source level. We believe that a violation of this assumption would mainly affect uncertainty estimates and not point estimates.

8. **Sounds that are incorrectly identified as bowhead whale calls (i.e., false positives) are never detected on more than sensor.** This means that by truncating the singletons we remove all potential false positives from the data.

2.3.11 Practical implementation

We fitted the model using maximum likelihood estimation (MLE) in `R 4.1.0` (R Core Team, 2023) with some components written in `C++` and linked to `R` through the `Rcpp` library (Eddelbuettel, 2013). We standardised the covariates in the density model to improve the convergence. The estimates were found through optimisation with the function `nlminb()` (R Core Team, 2023). We used a spatial mesh with non-uniform grid spacing to reduce run-time, with increased grid spacing farther from the sensor array (see Appendix B in Supplementary Materials for details).

2.4 Simulation study

We used simulations to evaluate the performance of the model under variable source levels (scenario 1) and fixed source levels (scenario 2). Both scenarios featured measurement error on bearings simulated from a two-part mixture model, and inhomogeneous call density specified as follows:

$$\log(D) = \beta_0 + \beta_1 d + \beta_2 d^2, \tag{2.17}$$

where d denotes the (scaled) shortest distance to the coast. The parameter values used were chosen to match those from the case study data analysis and are given in Table 2.2.

For each scenario, we simulated 100 data sets and analysed each data set with 5 models: (a) the true model, i.e., that corresponding to the simulated scenario; (b) a model with incorrect assumption about source level, i.e., for scenario 1 the model assumed fixed source level and for scenario 2 the model assumed variable source level; (c) a simpler bearing model assuming a von Mises distribution on bearing error but no two-part mixture; (d) a model that omitted the bearing information altogether; and (e) a model assuming a homogeneous spatial density. We did not include a simulation scenario where we naively fit a model to false positives, as the effect on the estimates is already known: it will induce a positive bias that will increase with increasing false positive rate.

Parameters	Variable SL	Constant SL
g_0	0.6	0.6
β_r	18.0	14.5
σ_r	2.7	4.5
μ_s	163.0	155.0
σ_s	5.0	-
κ	0.3	0.3
δ_κ	36.7	34.7
ψ_κ	0.1	0.1
β_0	-12.0	-16.0
β_1	45.0	57.0
β_2	-53.0	-68.5

TABLE 2.2: The parameter values used for the simulation study. These were based on initial fits to the real data, in order to keep the simulations as realistic as possible.

For each scenario and model combination (1a-e and 2a-e) we evaluated performance by calculating the coefficient of variation (CV), relative error and relative bias in estimated total number of calls. Further details are given in Appendix A.

2.4.1 Results

Both correctly specified models (1a and 2a) gave near-unbiased results (Figure 2.3). Fitting a single source level model in the variable source level scenario (1b) introduced a strong negative bias with low variance; fitting a variable source level in the fixed source level scenario (2b) introduced a positive bias with high variance. Using an incorrect, simple bearing model (1c and 2c) did not induce bias but did increase variance. Ignoring the bearing information induced a small positive bias and greater variance in the variable source level scenario (1d) but had little effect on the fixed source level scenario (2d). Lastly, using an incorrectly specified constant spatial density model (1e and 2e) resulted in severe overestimation of abundance ($> 300\%$ and $> 200\%$ bias for variable and fixed source level scenarios, respectively).

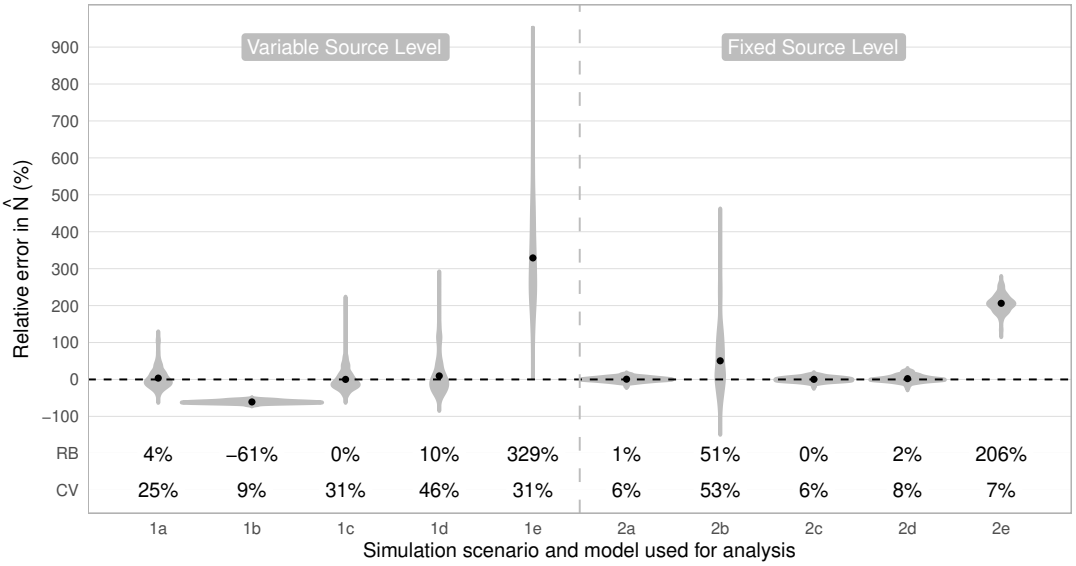


FIGURE 2.3: Relative error, relative bias and coefficient of variation for $\hat{N}$ from 100 simulations. The black dots correspond to the relative bias (RB; the mean relative error) and the grey shading shows the spread of relative error per scenario. The RB and coefficient of variation (CV) for every scenario and model combination are shown just above the x-axis. Simulation scenario 1 is variable sound source level and 2 is fixed sound source level. Analysis model (a) is the true model, (b) the incorrect source level model, (c) an incorrect simple bearing model, (d) a model with bearing data, and (e) an incorrect homogeneous spatial density model.

2.5 Bowhead whale analysis

We used the proposed method to estimate bowhead whale call density and abundance at site 5 on 31 August 2010. We used data from a single day, and as we present our estimate as call density per day, we did not need to adjust for our study time ($T = 1$). We set t_r at 96 dB, as the observed background noise never surpassed this level. The true call density surface is unknown and therefore we fitted several candidate models, which we compared using Akaike's information criterion (AIC; Akaike (1998)). As the bowhead whales are thought to exhibit a spatial preference based on their distance from the coast and ocean depth during fall migration (Greene et al., 2004), we fitted a candidate set of 35 models containing combinations of distance and depth as linear, quadratic or smooth covariates—see Appendix A for full details. Measures of uncertainty in abundance (standard error, CV, and quantile coefficient of dispersion (QCD)) were derived by bootstrapping the calls 999 times, refitting the AIC-best model each time. The QCD is a relative measure and insensitive to outliers, and is defined as

$(Q_3 - Q_1)/(Q_3 + Q_1)$, where Q_1 and Q_3 are the first and third quartile, respectively. To further illustrate the potential benefits of modelling source level heterogeneity, we included point estimates of the best model equivalent with a fixed source level.

2.5.1 Results

The lowest AIC model included density modelled as a smooth function of distance to coast and a quadratic relation with depth. We observe an increase in density, followed by a decrease, as we increase the distance from the coast (Figure 2.4, left panel). Moreover, density is higher just east of the sensor array, potentially due to some effect of ocean depth. Even though we observe increases in uncertainty in the southern regions and slightly in the north, overall QCD estimates are low (Figure 2.4, right panel). Total abundance of bowhead whale calls was estimated at 5741 in the study area, and the CVs of the parameter estimates varied considerably, ranging from 0.95% for $\hat{\mu}_s$ to 58.22% for $\widehat{\beta_{\text{s(depth).3}}}$ (Table 2.3).

2.6 Discussion

We have developed and tested a novel extension of acoustic spatial capture-recapture by conditioning on a minimum of two detections per individual call. Removing single detections may be necessary when the validity of these calls is challenged, and our simulation study shows that the method gives unbiased results in both variable and single source level scenarios. Even though this model applies to (marine) acoustic data, the extension proposed can be applied to all forms of SCR. Fitting a single source level model to variable source level data, thus ignoring heterogeneity in the detectability of the calls, resulted in severe underestimation of abundance. This represents a cautionary tale for other applications of SCR, both terrestrial and aquatic, when it is hypothesised that there is some random variable affecting individual detectability—it does not have to be something as tangible as a source level. Moreover, we confirmed results from Stevenson et al. (2015) on adding bearing information to a SCR model and found a further increase in precision when allowing for a mixture of 'good' and 'bad' bearings. Finally, we also confirm that incorrectly assuming a constant call density surface can lead to severe overestimates of the abundance. The notion that density of bowhead

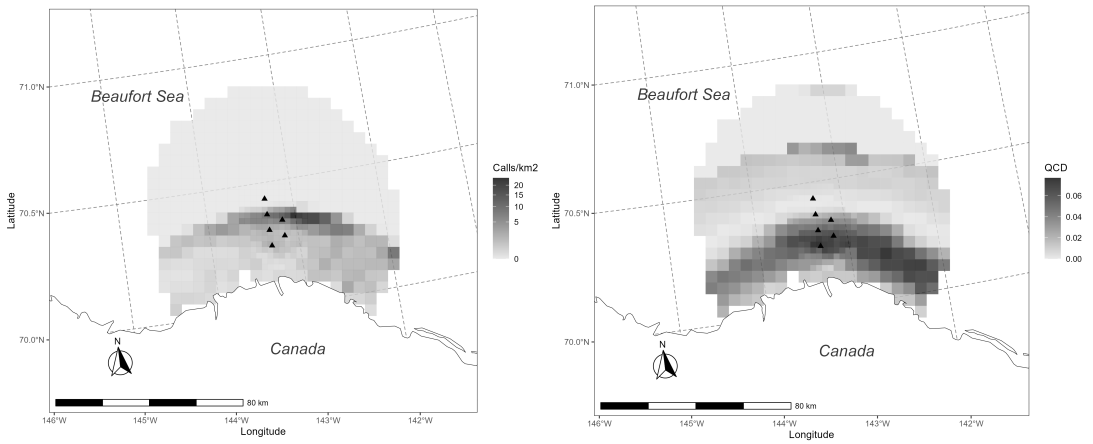


FIGURE 2.4: A map of the study site 5 from 31 August 2010 with the estimated density surface (left) and the associated empirical quartile coefficient of dispersion (QCD; right). Locations of the 6 DASARs are indicated with the triangles. The square grid cells within 10 km of the array each have an area of 6.25 km^2, and and area of 25 km^2 elsewhere. The density map corresponds to the model with the lowest AIC, which is model 33 in Appendix A. The QCD is derived from 999 Monte Carlo simulations based on the same model, where the calls are re-sampled.

whale(s) (calls) is highest 10–75 km from the coast (Greene et al., 2004) seems to be confirmed by the best model, which finds a density that is highest a bit away from the coast, but not too far. The high call spatial density in the northeastern part of the study area (Figure 2.4) was unexpected, but could have been a consequence of using only data from one day. This likely strongly violated the Poisson assumption about call spatial location, since whale location is spatially auto-correlated, especially on small time scales (Wursig et al., 1985). Alternatively, it has been hypothesised that the higher estimated density of calls in the east could be due to the directionality of the calls or the effect of water depth on sound propagation (Blackwell et al., 2012). In this study, the authors observed 2.1 times as many calls east of the array as west. Future research could explore whether a more gradual increasing/decreasing density slope is found if more days were included in the analysis, resulting in reduced autocorrelation among the call origins. Another extension could be to use a longer time-series and consider a two-dimensional spatial density surface, e.g., through splines on latitude and longitude.

Quantity	**Estimate**	**SE**	**CV (%)**	**Single SL**
N	5741.39	576.01	10.55	1916.06
g_0	0.55	0.02	2.73	0.67
β_r	17.05	0.33	1.91	13.86
σ_r	2.75	0.11	4.06	4.64
μ_s	158.46	1.50	0.95	151.04
σ_s	5.47	0.26	4.90	-
κ	0.77	0.17	21.95	0.62
δ_κ	45.49	3.02	6.54	40.23
ψ_κ	0.12	0.01	10.75	0.09
$\beta_{\text{intercept}}$	-265.65	7.92	2.99	-146.81
$\beta_{\text{s(depth).1}}$	42.43	18.07	45.55	80.65
$\beta_{\text{s(depth).2}}$	302.26	6.72	2.23	8.25
$\beta_{\text{s(depth).3}}$	41.14	22.32	58.22	88.65
$\beta_{\text{s(depth).4}}$	-141.57	6.88	4.89	-8.78
$\beta_{\text{s(depth).5}}$	154.60	6.19	4.01	24.23
$\beta_{\text{s(d2c).1}}$	159.48	4.47	2.82	59.19
$\beta_{\text{s(d2c).2}}$	-449.58	15.55	3.47	-185.40
$\beta_{\text{s(d2c).3}}$	101.06	4.29	4.26	33.88
$\beta_{\text{s(d2c).4}}$	-109.30	5.10	4.69	5.14
$\beta_{\text{s(d2c).5}}$	-224.66	11.13	4.96	-142.24

TABLE 2.3: The point estimates and associated uncertainties. SE is standard error, CV stands for coefficient of variation, and Single SL are point estimates from the equivalent model without source level heterogeneity. N is a derived quantity; g_0 through ψ_κ are the observation parameters presented on the real scale; and the remainder are the density parameters presented on their log link scale.

Another extension could be to use a longer time-series and consider a two-dimensional spatial density surface, e.g., through splines on latitude and longitude. However, the biological pattern of migration along the coast combined with limited sensor spacing in the east-west direction means this may not produce improved inferences.

The model-based expected number of singletons was 570 whereas there were 1417 observed singletons, which is 149% more than expected and confirmed our suspicions regarding false positives (see Figure 2.2).

An alternative to truncating singletons would be to retain all data and try to model the occurrence of false positive detections. For example, one could assume that an unknown proportion of detections are false positives and include a parameter for this

proportion in an analogous way to the $M_{t,\alpha}$ model presented by Yoshizaki (2007). Detection at any detector j could have three outcomes: no detection with probability $1-p_j$, false detection with probability αp_j, and correct detection with probability p_j. Moreover, we would have to implement some modification of the M_b model (Otis et al., 1978) to account for a change in α for the multiply-detected calls, which would likely involve an additional parameter. A benefit of this $M_{b,\alpha}$ model would be that it could allow for false positives in multiply-detected calls, however, it increases complexity and introduces hard-to-test assumptions on the false positive rate. Conn et al. (2014) estimated a false positives rate for automated detections of seals from aerial surveys by using a double sampling procedure. The model performed reasonably, but they also found that the increased model complexity could cause issues. Given that a large amount of acoustic data are potentially available (we used data from just one day), including poor-quality data seemed undesirable given the additional complexity and run time. Hence we consider the truncation approach developed here to be best for our application of ASCR.

The main model in our study conditions on at least two positive detections for every detection history, but can readily be extended to condition on a minimum of three or more involved sensors. This requires generalising Equation 2.5 to

$$\begin{aligned} p_\cdot(\boldsymbol{x}, s) =& \mathbb{P}(\omega^* \geq \omega^*_{\min} | \boldsymbol{x}, s) \\ =& 1 - \mathbb{P}(\omega^* = 0 | \boldsymbol{x}, s) - \mathbb{P}(\omega^* = 1 | \boldsymbol{x}, s) \\ & - \ldots - \mathbb{P}(\omega^* = \omega^*_{\min} - 1 | \boldsymbol{x}, s), \end{aligned} \tag{2.18}$$

where $\omega^*_{\min}$ denotes the minimum number of sensors involved. Preliminary simulations showed no apparent bias, although depending on the situation, conditioning this way can rapidly reduce the amount of data available and hence decrease estimator precision. This is illustrated by the relative frequencies in Figure 2.2. Moreover, excluding calls with fewer associated sensors results in the remaining calls generally being louder calls. If there is heterogeneity in source level between the bowhead whales (which violates assumption 7), this could increasingly eliminate quieter whales.

We estimated variance empirically by bootstrapping the calls. This assumes independence between the calls, which we know is violated as the calls are produced by whales and whales are spatially auto-correlated. Moreover, calls can potentially trigger responses from nearby whales, leading to temporal correlation between them (Thode

et al., 2020). Some of the spatial and temporal dependence was eliminated by removing all detections with a received level below 96 dB, as this thinned the data, assuming independence between call characteristics and ocean noise. To create more accurate variance estimates, one could stratify the data by time and bootstrap these time chunks, thereby removing some of the autocorrelation. Even better, one could sample short time frames from several days and combine those to obtain the desired amount of data to analyse, in which case the analysed observations themselves could be assumed independent and likelihood-based variance estimates are acceptable (given a large enough data set). The length and frequency of these time chunks will be case specific, as this will depend on the calling behaviour of the studied species.

We estimated model parameters using MLE as opposed to in a Bayesian inferential framework, with MLE having coding simplicity and reduced run time as the main benefits. We did find that convergence was sometimes slowed due to flat likelihood surfaces as a consequence of the many sources of variation in our model. Evaluating the likelihood for the best model took around 19 seconds on a single core. Consequently, it took over two weeks to fit the 999 models for the bootstrap procedure, where model fitting was performed in parallel on 32 threads of an Intel® Xeon® Gold 6152 CPU. A benefit of using a Bayesian framework would be the possibility to include informative prior distributions based on previous research, especially on the latent variable source level, which would likely improve convergence and may improve precision of density and abundance estimates.

Our method derives a call abundance, which can be converted into an animal abundance by correcting this estimate for the call rate, similarly to Marques et al. (2013). Ideally, this call rate is estimated for the same population and alongside the collection of the data, but this is not always possible. Blackwell et al. (2021) estimated a call rate for the Bering–Chukchi–Beaufort population over a longer period of time. Naively dividing our estimate of total calls by their median call rate estimate of 31.2 calls/whale/day (interquartile range of 12–129.6 calls/whale/day) gives us $\widehat{N}_{\text{whales}} = 5741.39/31.2 \approx 184$ individuals. We present this number for illustration alone; for correct inference, one should properly account for the variance around call rate and call abundance estimates. Alternatively, there are methods that directly estimate animal density based on cues (for more information, see e.g., Fewster et al. (2016) and Stevenson et al. (2019)).

When background noise is highly variable, selecting a truncation level that ensures that noise (almost) never surpasses the received level might result in most data being discarded. When data are scarce, it may then be beneficial to use a detection function based on the signal-to-noise ratio (SNR). We develop such a method in Chapter 3.

A potential weakness of the model is the fact that we assume the same propagation loss model for the entire study area. For most of the study site this assumption is reasonable, as it concerns a shallow plateau with little variation in ocean depth. However, the ocean floor drops in the northern part of the study site, resulting in altered propagation conditions. It is assumed that the bowhead whales migrate mainly over the shallow plateau and hence depth-induced inhomogeneous propagation is not a practical issue in our case study. In general, however, it is something that should be considered when modelling sound propagation in variable (ocean) landscapes. Phillips (2016) and Royle (2018) presented ways in which it is possible explicitly to model variable sound attenuation. Such a model requires accurate and sometimes high resolution information on environmental variables that affect sound attenuation, such as ocean depth.

The model presented here extends the scope of SCR to provide reliable inference on spatial density and abundance from passive acoustic data. Removing single detections will also be of use in other SCR applications where single detections are unreliable – for example when association between detections is not perfect such that repeat detections of the same individual are sometimes incorrectly classified as single detections of a new individual. This can happen in situations where natural animal characteristics are used for identification, such as photo- and genetic-ID

Chapter 3

A signal-to-noise ratio detection function for acoustic spatial capture-recapture

3.1 Introduction

In Chapter 2, we truncated the data at a sound level that would avoid any situation in which the background noise would be louder than the received sound level of a call. To achieve this, we set the truncation value equal to at least the maximum noise level. In acoustic spatial capture-recapture (ASCR), it is common practice to set a truncation value and assuming a single detection probability of calls that surpass this value(Dawson and Efford, 2009; Efford et al., 2009; Stevenson et al., 2015). When noise is (highly) variable, however, truncating the data using the highest noise level can result in removing a substantial amount of data, which can be a problem when data are scarce. A model that uses a detection function that relies on the signal-to-noise ratio (SNR), which is the ratio of the loudness of the signal to the loudness of the noise, would not require the truncation of data. Occasional periods of high background noise, when calls are harder to detect, no longer affect the periods with low background noise, when calls are better more easily detected. In this chapter, we provide details of such a model. We first derive the full likelihood, for which we use several sources of recorded information: the detection histories, the bearings, the received levels and the noise

levels. We assume that calls are independent of one other in all their characteristics, similar to the previous chapter. In Section 3.2, we first derive the likelihood for a single call. From there, we expand it to include all calls (Section 3.3). Finally, in Section 3.4 we provide some concluding remarks on the method and suggestions for future steps. In this chapter, we only develop the methodology; we do not apply it to any real or simulated data. However, it is anchored on the same case study as Chapter 2 (see Section 2.2).

3.2 Probability density for a single call

We start by presenting this method for a single call, and then later expand to n calls. Consider an acoustic survey using an array of K sensors operating within a survey region $\mathcal{A} \subset \mathbb{R}^2$ over a period of time T. The detection history is denoted by the vector $\boldsymbol{\omega}$, where element $\omega_j = 1$ if the call was detected at sensor j, and 0 otherwise. Whether a call gets detected at a sensor is a Bernoulli trial with some detection probability. We define $\omega^* = \sum_{j=1}^{K} \omega_j$, where K is the total number of sensors in the array. Bearings are denoted by the vector $\boldsymbol{y}$, where element y_j is the bearing from sensor j to the call in radians clockwise relative to north if the call was detected, and NA (denoting not available) otherwise. The measured received levels are the root-mean-square (RMS) pressure over the frequency band of the call in dB re 1 μPa. They are denoted by vector $\boldsymbol{r}$, where element r_j is the received level of the call at sensor j if the call was detected, and NA otherwise. Noise, which refers to all acoustics other than the signal of interest that could interfere with the detection, is denoted by vector $\boldsymbol{c}$, where element c_j is the RMS in dB re 1 μPa of the noise at sensor j, measured over the same frequency band as the call. Putting all of this together, we denote the joint probability density function (pdf) for a single call as

$$f(\boldsymbol{\omega}, \boldsymbol{y}, \boldsymbol{r} | \omega^* \geq 2, \boldsymbol{c}). \tag{3.1}$$

There are some call characteristics that are relevant to this density distribution, namely the source level s and spatial origin of the call $\boldsymbol{x}$, which are latent but we still include in the model. We deal with these latent variables by marginalising over these variables, as described in Section 3.2.2.

3.2.1 Detection probability for sensor j, $g_j(\boldsymbol{x}, s, c_j)$

Before we derive the joint density for the random variables, we need to define probability of detection. Unlike the detection function in Chapter 2, we do not propose a step function with a single detection probability, but rather a gradual detection function that is a function of SNR (Küsel et al., 2011; Thode et al., 2020). The simplicity of a step function might be preferable when data are scarce and/or many parameters are to be estimated, but a gradual function is likely to be closer to the truth, as empirically derived SNR detection functions have been found to in general be gradually increasing (Küsel et al., 2011; Thode et al., 2020). In line with previous literature, we denote the detection function for sensor j as

$$g_j(\boldsymbol{x}, s, c_j) = \int_{\mathcal{R}} p(r, c_j) f(r|\boldsymbol{x}, s, c_j) dr, \tag{3.2}$$

where $\mathcal{R} \in [0, \infty)$ denotes the support for the received level, $p(r, c_j)$ denotes the detection characteristic function, and $f(r|\boldsymbol{x}, s, c_j)$ is the pdf of r given $\boldsymbol{x}, s$ and c_j (see Equation (3.18)). The detection characteristic function defines the relationship between SNR and detectability of the call at a sensor and can be any sensible function. As an example, we will suggest one for the bowhead whale case study in Chapter 2. Exploratory analysis showed that the detection probability of bowhead whale calls could be well below 1 even for calls with high SNR. We therefore should allow for an upper asymptote smaller than one, as we want to allow for non-perfect detection even at (very) high values of SNR. For example, one could use the Janoschek function (Janoschek, 1957), which is denoted as follows:

$$p(r, c) = \theta_U - (\theta_U - \theta_L) \exp\left\{-\theta_R (r - c)^{\theta_I}\right\}, \tag{3.3}$$

where θ_L and θ_U are the respective lower and upper asymptote, $\theta_R > 0$ is the rate of increase, $\theta_I > 1$ controls the inflection point, and $r - c$ is the SNR, which is a subtraction as r and c are denoted on the decibel scale. We set the lower limit θ_L to 0 to avoid negative probabilities.

3.2.2 Joint distribution of the detection histories, bearings and received levels, given two or more detections and noise, $f(\boldsymbol{\omega}, \boldsymbol{y}, \boldsymbol{r}|\omega^* \geq 2, \boldsymbol{c})$

We start by specifying the joint distribution of the observed random variables. We know this distribution does not only depend on detection, but also on spatial origin and the source level of the detected call. Since these variables are latent, we integrate them out, which gives the following formula for the joint distribution

$$f(\boldsymbol{\omega}, \boldsymbol{y}, \boldsymbol{r}|\omega^* \geq 2, \boldsymbol{c}) = \int_{\mathcal{S}} \int_{\mathcal{A}} f(\boldsymbol{\omega}, \boldsymbol{y}, \boldsymbol{r}, \boldsymbol{x}, s|\omega^* \geq 2, \boldsymbol{c}) d\boldsymbol{x} ds, \tag{3.4}$$

where $\mathcal{S}$ is the support for s, which in theory is $(0, \infty)$ but in practice it suffices to use a smaller support as long as it sufficiently covers the probability density of s. Using standard conditional probability theory we can rewrite this distribution as

$$\begin{aligned} &\int_{\mathcal{S}} \int_{\mathcal{A}} f(\boldsymbol{\omega}, \boldsymbol{y}, \boldsymbol{r}, \boldsymbol{x}, s|\omega^* \geq 2, \boldsymbol{c}) d\boldsymbol{x} ds = \\ &\qquad \int_{\mathcal{S}} \int_{\mathcal{A}} f(\boldsymbol{\omega}, \boldsymbol{y}, \boldsymbol{r}|\boldsymbol{x}, s, \omega^* \geq 2, \boldsymbol{c}) f(\boldsymbol{x}|s, \omega^* \geq 2, \boldsymbol{c}) f(s|\omega^* \geq 2, \boldsymbol{c}) d\boldsymbol{x} ds. \end{aligned} \tag{3.5}$$

We will now focus on each component in Equation (3.5) separately, starting with $f(s|\omega^* \geq 2, \boldsymbol{c})$.

3.2.3 Distribution of the source level of the call, given at least two detections and noise, $f(s|\omega^* \geq 2, \boldsymbol{c})$

Using Bayes' formula and the addition of latent variables similar to before, we can rewrite the density $f(s|\omega^* \geq 2, \boldsymbol{c})$ as

$$\begin{aligned} f(s|\omega^* \geq 2, \boldsymbol{c}) &= \frac{f(\omega^* \geq 2|s, \boldsymbol{c}) f(s)}{f(\omega^* \geq 2|\boldsymbol{c})} \\ &= \frac{\int_{\mathcal{A}} f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c}) f(\boldsymbol{x}) d\boldsymbol{x} \times f(s)}{\int_{\mathcal{S}} \int_{\mathcal{A}} f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c}) f(\boldsymbol{x}) f(s) d\boldsymbol{x} ds}. \end{aligned} \tag{3.6}$$

Note that some probability density distributions, such as the one for source level, are not denoted as conditional on noise. This is because the variables s and $\boldsymbol{x}$ are assumed independent of noise. Thode et al. (2020) showed that under some circumstances the

source level of bowhead whale vocalisations can be related to noise and other discrete sounds; here we assume the two variables to be independent.

3.2.4 Distribution of the spatial origin and source level of the call, given at least two detections, noise, and source level, $f(\boldsymbol{x}|s, \omega^* \geq 2, \boldsymbol{c})$

In this section, we will start by emphasising the distinction between an *emitted* and a *detected* call; we need to define the former before defining the latter. The probability density of the spatial origin of an emitted call $\boldsymbol{x}$ is proportional to the call density at $\boldsymbol{x}$, $D(\boldsymbol{x})$, that is $f(\boldsymbol{x}) \propto D(\boldsymbol{x})$. The normalising constant is therefore the integral over the study area, $\int_{\mathcal{A}} D(\boldsymbol{x})d\boldsymbol{x}$, and this gives

$$f(\boldsymbol{x}) = \frac{D(\boldsymbol{x})}{\int_{\mathcal{A}} D(\boldsymbol{x})d\boldsymbol{x}}. \tag{3.7}$$

We have not defined a spatial density model $D(\boldsymbol{x})$ for the calls yet. It is common practice to try several density models and then choose to best one using some criterion such as Akaike's information criterion (AIC; Akaike, 1998). The density $f(\boldsymbol{x}|s, \omega^* \geq 2, \boldsymbol{c})$ is

$$\begin{aligned} f(\boldsymbol{x}|s, \omega^* \geq 2, \boldsymbol{c}) &= \frac{f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c})f(\boldsymbol{x}|s, \boldsymbol{c})}{f(\omega^* \geq 2|s, \boldsymbol{c})} \\ &= \frac{f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c})f(\boldsymbol{x})}{\int_{\mathcal{A}} f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c})f(\boldsymbol{x})d\boldsymbol{x}}. \end{aligned} \tag{3.8}$$

In both definitions we assume the distribution of *emitted* calls to be unrelated to s and $\boldsymbol{c}$, thus $f(\boldsymbol{x}|s, \boldsymbol{c}) \equiv f(\boldsymbol{x})$ (note that this is not true for the distribution of *detected* calls, which changes when s and $\boldsymbol{c}$ changes). As with Chapter 2, we assume $s \sim \mathcal{N}(\mu_s, \sigma_s^2)$.

Finally, we take the product of Equations (3.6) and (3.8), which gives

$$\begin{aligned} &f(\boldsymbol{x}|s, \omega^* \geq 2, \boldsymbol{c})f(s|\omega^* \geq 2, \boldsymbol{c}) \\ &= \frac{f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c})f(\boldsymbol{x})}{\int_{\mathcal{A}} f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c})f(\boldsymbol{x})d\boldsymbol{x}} \times \frac{\int_{\mathcal{A}} f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c})f(\boldsymbol{x})d\boldsymbol{x}f(s)}{\int_{\mathcal{S}}\int_{\mathcal{A}} f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c})f(\boldsymbol{x})f(s)d\boldsymbol{x}ds} \\ &= \frac{f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c})f(s)}{\int_{\mathcal{S}}\int_{\mathcal{A}} f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c})\frac{D(\boldsymbol{x})}{\int_{\mathcal{A}} D(\boldsymbol{x})d\boldsymbol{x}}f(s)d\boldsymbol{x}ds} \times \frac{D(\boldsymbol{x})}{\int_{\mathcal{A}} D(\boldsymbol{x})d\boldsymbol{x}} \\ &= \frac{f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c})D(\boldsymbol{x})f(s)}{a}, \end{aligned} \tag{3.9}$$

where in the last step the constants $\int_{\mathcal{A}} D(\boldsymbol{x})d\boldsymbol{x}$ cancel out and $a = \int_{\mathcal{S}}\int_{\mathcal{A}} f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c})D(\boldsymbol{x})f(s)d\boldsymbol{x}ds$.

Finally, the distribution of $f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c})$ is a probability mass function (pmf), where $\omega^* \geq 2$ can either attain the value 0 when a call is detected at most once or 1 when detected twice or more. This is represented by the following:

$$\begin{aligned} f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c}) &= 1 - \mathbb{P}(\omega^* = 0|\boldsymbol{x}, s, \boldsymbol{c}) - \mathbb{P}(\omega^* = 1|\boldsymbol{x}, s, \boldsymbol{c}) \\ &= 1 - \prod_{j=1}^{K}(1 - g_j(\boldsymbol{x}, s, c_j)) \\ &\quad - \sum_{j=1}^{K}\left(g_j(\boldsymbol{x}, s, c_j)\prod_{k\in\{1:K|k\neq j\}}(1 - g_k(\boldsymbol{x}, s, c_k))\right). \end{aligned} \tag{3.10}$$

where detection function $g_j(\boldsymbol{x}, s, c_j)$ is specified in Equation (3.2). Equation (3.10) can readily be extended to set the minimum number of detections per call at 3, 4, etc. (although the formula and computational complexity significantly increases, as was discussed in Chapter 2).

3.2.5 Joint distribution of the capture history, bearings and received level for a call, given its spatial origin, source level, at least two detections and noise, $f(\boldsymbol{\omega}, \boldsymbol{y}, \boldsymbol{r}|\boldsymbol{x}, s, \omega^* \geq 2, \boldsymbol{c})$

First, we observe that

$$\begin{aligned} f(\boldsymbol{\omega}, \boldsymbol{y}, \boldsymbol{r}|\boldsymbol{x}, s, \omega^* \geq 2, \boldsymbol{c}) = f(\boldsymbol{\omega}|\boldsymbol{x}, s, \omega^* \geq 2, \boldsymbol{c}) \times f(\boldsymbol{y}|\boldsymbol{\omega}, \boldsymbol{x}, s, \omega^* \geq 2, \boldsymbol{c}) \\ \times f(\boldsymbol{r}|\boldsymbol{y}, \boldsymbol{\omega}, \boldsymbol{x}, s, \omega^* \geq 2, \boldsymbol{c}). \end{aligned} \tag{3.11}$$

Since $\boldsymbol{y}$ and s are unrelated, and so are $\boldsymbol{r}$ and $\boldsymbol{y}$, we can rewrite Equation (3.11) as

$$\begin{aligned} f(\boldsymbol{\omega}, \boldsymbol{y}, \boldsymbol{r}|\boldsymbol{x}, s, \omega^* \geq 2, \boldsymbol{c}) = f(\boldsymbol{\omega}|\boldsymbol{x}, s, \omega^* \geq 2, \boldsymbol{c}) \times f(\boldsymbol{y}|\boldsymbol{\omega}, \boldsymbol{x}, \omega^* \geq 2, \boldsymbol{c}) \\ \times f(\boldsymbol{r}|\boldsymbol{\omega}, \boldsymbol{x}, s, \omega^* \geq 2, \boldsymbol{c}). \end{aligned} \tag{3.12}$$

We now evaluate this joint density from left to right, starting with the density of $f(\boldsymbol{\omega}|\boldsymbol{x}, s, \omega^* \geq 2, \boldsymbol{c})$. First, we rewrite $f(\boldsymbol{\omega}|\boldsymbol{x}, s, \omega^* \geq 2, \boldsymbol{c})$ as

$$\begin{aligned} f(\boldsymbol{\omega}|\boldsymbol{x}, s, \omega^* \geq 2, \boldsymbol{c}) &= \frac{f(\omega^* \geq 2|\boldsymbol{\omega}, \boldsymbol{x}, s, \boldsymbol{c}) f(\boldsymbol{\omega}|\boldsymbol{x}, s, \boldsymbol{c})}{f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c})} \\ &= \frac{1 \times f(\boldsymbol{\omega}|\boldsymbol{x}, s, \boldsymbol{c})}{f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c})} = \frac{f(\boldsymbol{\omega}|\boldsymbol{x}, s, \boldsymbol{c})}{f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c})}. \end{aligned} \tag{3.13}$$

Conditioning on the detection history gives a probability of two or more detections, since the probability of detection at least twice is implied by the existence of the detection history.

We can view the detection history of a call as a set of Bernoulli trials of size K and non-constant probability $g_j(\boldsymbol{x}, s, \boldsymbol{c})$ with $j = 1, .., K$. This gives

$$f(\boldsymbol{\omega}|\boldsymbol{x}, s, \omega^* \geq 2, \boldsymbol{c}) = \frac{\prod_{j=1}^{K} g_j(\boldsymbol{x}, s, c_j)^{\omega_j} (1 - g_j(\boldsymbol{x}, s, c_j))^{1-\omega_j}}{f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c})}. \tag{3.14}$$

The second term in Equation (3.11) is $f(\boldsymbol{y}|\boldsymbol{\omega}, \boldsymbol{x}, \omega^* \geq 2, \boldsymbol{c})$, which we can rewrite as

$$\begin{aligned} f(\boldsymbol{y}|\boldsymbol{\omega}, \boldsymbol{x}, \omega^* \geq 2, \boldsymbol{c}) &= \frac{f(\omega^* \geq 2|\boldsymbol{y}, \boldsymbol{\omega}, \boldsymbol{x}, \boldsymbol{c}) f(\boldsymbol{y}|\boldsymbol{\omega}, \boldsymbol{x})}{f(\omega^* \geq 2|\boldsymbol{\omega}, \boldsymbol{x}, \boldsymbol{c})} \\ &= \frac{1 \times f(\boldsymbol{y}|\boldsymbol{\omega}, \boldsymbol{x})}{1} = f(\boldsymbol{y}|\boldsymbol{\omega}, \boldsymbol{x}), \end{aligned} \tag{3.15}$$

where we assume independence between bearings and noise.

To keep this section as simple as possible, we assume that all bearings follow a Von Mises distribution with a single concentration parameter κ (unlike Chapter 2, where we assumed a 2-part mixture), which gives

$$f(\boldsymbol{y}|\boldsymbol{\omega}, \boldsymbol{x}) = \prod_{j \in \{1:K|\omega_j=1\}} \left(\frac{\exp\{\kappa \cos(y_j - \mathbb{E}[y_j|\boldsymbol{x}])\}}{2\pi I_0(\kappa)} \right), \tag{3.16}$$

where I_0 is the modified Bessel function of order 0 and $\mathbb{E}[y_j|\boldsymbol{x}]$ is the expected bearing at sensor j for spatial call origin $\boldsymbol{x}$.

Finally, we rewrite the density $f(\boldsymbol{r}|\boldsymbol{\omega}, \boldsymbol{x}, s, \omega^* \geq 2, \boldsymbol{c})$ as

$$\begin{aligned} f(\boldsymbol{r}|\boldsymbol{\omega}, \boldsymbol{x}, s, \omega^* \geq 2, \boldsymbol{c}) &= \frac{f(\omega^* \geq 2|\boldsymbol{r}, \boldsymbol{\omega}, \boldsymbol{x}, s, \boldsymbol{c}) f(\boldsymbol{r}|\boldsymbol{\omega}, \boldsymbol{x}, s, \boldsymbol{c})}{f(\omega^* \geq 2|\boldsymbol{\omega}, \boldsymbol{x}, s, \boldsymbol{c})} \\ &= \frac{1 \times f(\boldsymbol{r}|\boldsymbol{\omega}, \boldsymbol{x}, s)}{1} = f(\boldsymbol{r}|\boldsymbol{\omega}, \boldsymbol{x}, s, \boldsymbol{c}). \end{aligned} \tag{3.17}$$

Following the model described in the previous chapter, we assume a Gaussian error on received levels to capture measurement error and error in the propagation model, such that

$$r_j|\boldsymbol{x}, s \sim \mathcal{N}(\mathbb{E}[r_j|\boldsymbol{x}, s], \sigma_r^2). \tag{3.18}$$

Received level itself is a function of i) source level, denoted by s; ii) the Euclidean distance in meters, denoted by $d_j(\boldsymbol{x})$, between the spatial origin of the call $\boldsymbol{x}$ and sensor j; and iii) the transmission loss per distance travelled by the sound, β_r. We approximate the transmission loss by using sound propagation characteristics of a lossless medium (Jensen et al., 2011), which gives the following formula for the expectation for r_j:

$$\mathbb{E}[r_j|s, \boldsymbol{x}] = s - \beta_r \log_{10}\left(d_j(\boldsymbol{x})\right), \tag{3.19}$$

where β_r is a parameter indexing the rate at which received level decreases with distance. More complex propagation models can be used, as was shown by Phillips (2016) and Royle (2018). As r_j and c_j are both denoted on the dB scale which is logarithmic, the expected signal-to-noise ratio at sensor j is simply given by

$$\mathbb{E}[\mathrm{SNR}_j|s, \boldsymbol{x}, c_j] = \mathbb{E}[r_j|s, \boldsymbol{x}] - c_j = s - \beta_r \log_{10}\left(d_j(\boldsymbol{x})\right) - c_j. \tag{3.20}$$

Signal strength is only recorded if the call is detected, and thus the probability density can only be derived when $\omega_j = 1$. Assuming independence between the detectors, we can write

$$f(\boldsymbol{r}|\boldsymbol{\omega}, \boldsymbol{x}, s, \boldsymbol{c}) = \prod_{j \in \{1:K|\omega_j = 1\}} f(r_j|\omega_j = 1, \boldsymbol{x}, s, c_j), \tag{3.21}$$

where Bayes' formula gives

$$f(r_j|\omega_j = 1, \boldsymbol{x}, s, c_j) = \frac{f(\omega_j = 1|r_j, \boldsymbol{x}, s, c_j) \times f(r_j|\boldsymbol{x}, s, c_j)}{f(\omega_j = 1|\boldsymbol{x}, s, c_j)} \tag{3.22}$$

The denominator in Equation (3.22) denotes the probability of detection given spatial

origin, source level, and noise, and is thus the same as the detection function (see Equation (3.2). The first element in the numerator in Equation (3.22) denotes the probability of detection given received level and noise, and is the same as the detection characteristic function. This means that $\boldsymbol{x}$ and s are redundant, thus giving

$$f(\omega_j = 1|r_j, \boldsymbol{x}, s, c_j) = f(\omega_j = 1|r_j, c_j) = p(r_j, c_j) \tag{3.23}$$

The second element of the numerator in Equation (3.22) is the probability density function of received level given the spatial origin and the source level, which is specified in Equation (3.18). As received level and noise are assumed independent, we get

$$f(r_j|\boldsymbol{x}, s, c_j) = f(r_j|\boldsymbol{x}, s) = \frac{1}{\sigma_r}\varphi\left(\frac{r_j - \mathbb{E}(r_j|\boldsymbol{x}, s)}{\sigma_r}\right), \tag{3.24}$$

where φ denotes the standard normal probability density function. We combine the three elements to rewrite Equation (3.21) as

$$f(\boldsymbol{r}|\boldsymbol{\omega}, \boldsymbol{x}, s, \boldsymbol{c}) = \prod_{j \in \{1:K|\omega_j=1\}} \frac{1}{\sigma_r} \times \frac{p(r_j, c_j) \times \varphi\left(\frac{r_j - \mathbb{E}(r_j|\boldsymbol{x}, s)}{\sigma_r}\right)}{g_j(\boldsymbol{x}, s, c_j)}. \tag{3.25}$$

Putting all of this and Equation (3.9) together gives the following joint density function

$$
\begin{aligned}
&f(\boldsymbol{\omega}, \boldsymbol{y}, \boldsymbol{r}, |\omega^* \geq 2, \boldsymbol{c}) \\
&\quad = \int_{\mathcal{S}} \int_{\mathcal{A}} \frac{f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c}) D(\boldsymbol{x}) f(s)}{a} \\
&\qquad \times \frac{\prod_{j=1}^{K} g_j(\boldsymbol{x}, s, c_j)^{\omega_j} (1 - g_j(\boldsymbol{x}, s, c_j))^{1-\omega_j}}{f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c})} \\
&\qquad \times \prod_{j \in \{1:K|\omega_j=1\}} \frac{\exp\{\kappa \cos(y_j - \mathbb{E}[y_j|\boldsymbol{x}])\}}{2\pi I_0(\kappa)} \\
&\qquad \times \prod_{j \in \{1:K|\omega_j=1\}} \frac{p(r_j, c_j) \times \varphi\left(\frac{r_j - \mathbb{E}(r_j|\boldsymbol{x}, s)}{\sigma_r}\right)}{\int_{\mathcal{R}} p(r, c_j) \times \varphi\left(\frac{r - \mathbb{E}(r|\boldsymbol{x}, s)}{\sigma_r}\right) dr} d\boldsymbol{x} ds \qquad (3.26) \\
&\quad = \frac{1}{a} \int_{\mathcal{S}} \int_{\mathcal{A}} D(\boldsymbol{x}) f(s) \prod_{j=1}^{K} g_j(\boldsymbol{x}, s, c_j)^{\omega_j} (1 - g_j(\boldsymbol{x}, s, c_j))^{1-\omega_j} \\
&\qquad \times \prod_{j \in \{1:K|\omega_j=1\}} \frac{\exp\{\kappa \cos(y_j - \mathbb{E}[y_j|\boldsymbol{x}])\}}{2\pi I_0(\kappa)} \\
&\qquad\qquad \times \frac{p(r_j, c_j) \times \varphi\left(\frac{r_j - \mathbb{E}(r_j|\boldsymbol{x}, s)}{\sigma_r}\right)}{\int_{\mathcal{R}} p(r, c_j) \times \varphi\left(\frac{r - \mathbb{E}(r|\boldsymbol{x}, s)}{\sigma_r}\right) dr} d\boldsymbol{x} ds.
\end{aligned}
$$

3.3 Generalising to n calls

We have so far derived the density of the detection history, bearing and received level for a single call. We now generalise to n calls that were detected on at least two sensors. We assume that all calls are independent of one another. We define the $(n \times 2)$-matrix of all spatial origins $\boldsymbol{X} = (\boldsymbol{x}_1, ..., \boldsymbol{x}_n)^\top$, the $(n \times 1)$-vector of all source levels $\boldsymbol{s} = (s_1, ..., s_n)^\top$, the $(n \times K)$-matrix of all detection histories $\boldsymbol{\Omega} = (\boldsymbol{\omega}_1, ..., \boldsymbol{\omega}_n)^\top$, the $(n \times K)$-matrix of all bearings $\boldsymbol{Y} = (\boldsymbol{y}_1, ..., \boldsymbol{y}_n)^\top$, the $(n \times K)$-matrix of all received levels as $\boldsymbol{R} = (\boldsymbol{r}_1, ..., \boldsymbol{r}_n)^\top$ and the $(n \times K)$-matrix of all the noise as $\boldsymbol{C} = (\boldsymbol{c}_1, ..., \boldsymbol{c}_n)^\top$. Using this notation, we can create the joint distribution $f(n, \boldsymbol{\Omega}, \boldsymbol{Y}, \boldsymbol{R}|\boldsymbol{C})$ as

$$
\begin{aligned}
f(n, \boldsymbol{\Omega}, \boldsymbol{Y}, \boldsymbol{R}|\boldsymbol{C}) &= f(n|\boldsymbol{C}) f(\boldsymbol{\Omega}, \boldsymbol{Y}, \boldsymbol{R}|n, \boldsymbol{C}) \\
&= f(n) \int_{\mathcal{S}} \int_{\mathcal{A}} f(\boldsymbol{\Omega}, \boldsymbol{Y}, \boldsymbol{R}, \boldsymbol{X}, \boldsymbol{s}|n, \boldsymbol{C}) d\boldsymbol{X} d\boldsymbol{s}
\end{aligned} \qquad (3.27)
$$

We show below why $f(n|\boldsymbol{C})$ is equivalent to $f(n)$.

3.3.1 Number of detected animals, n

In Chapter 2 we assumed that the emitted calls occur in space according to an inhomogeneous Poisson point process with intensity $D(\boldsymbol{x})$. As not all calls are detected, the *detected* calls occur in space according to a filtered inhomogeneous Poisson point process with the rate parameter being a function of $\boldsymbol{x}$, i.e., $D(\boldsymbol{x})$, multiplied by the probability that this call was detected at least twice, $\int_{\mathcal{S}} f(\omega^* \geq 2|\boldsymbol{x}, s)f(s)ds$. The total number of detected calls then follows a Poisson distribution with rate parameter equal to the integral over $\mathcal{A}$, giving $\lambda = \int_{\mathcal{A}} \int_{\mathcal{S}} D(\boldsymbol{x})f(s)f(\omega^* \geq 2|\boldsymbol{x}, s)dsd\boldsymbol{x}$.

So far, our spatial process was *space-inhomogeneous*, but *time-homogeneous*. In fact, this is one of the assumptions of a Poisson process; their intensity or rate parameter is constant throughout time. The information on noise, however, varies over time, which means that our rate parameter λ varies with it. When a Poisson count process is time-inhomogeneous, the rate parameter of the Poisson process over the entire study period can be derived by integrating over time. As our time-related variable is noise, this means we have to integrate over noise. Instead of assuming a distribution on noise and weighing by that, we propose using a (systematic) random sample of b noise recordings for all K sensors. We denote the matrix of these noise recordings by $\mathcal{C}$. The rate parameter for the underlying Poisson point process thus is approximated by

$$\lambda \approx \frac{\sum_{\boldsymbol{c}\in\mathcal{C}} \int_{\mathcal{S}} D(\boldsymbol{x})f(s)f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c})ds}{b},$$

involving a Monte-Carlo integration over noise. The distribution of n follows as

$$f(n) \sim \text{Poisson}\left(\frac{\int_{\mathcal{A}} \sum_{\boldsymbol{c}\in\mathcal{C}} \int_{\mathcal{S}} D(\boldsymbol{x})f(s)f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c})dsd\boldsymbol{x}}{b}\right). \tag{3.28}$$

3.3.2 Joint distribution of all detection histories, bearings, received levels, spatial origins and source levels, given the detected calls, $\boldsymbol{\Omega}, \boldsymbol{Y}, \boldsymbol{R}, \boldsymbol{X}, \boldsymbol{s}|n, \boldsymbol{C}$

As we assume calls to be independent of one another, the part inside the integrals becomes the product of the density for the individual calls. Moreover, conditioning on n is equivalent to conditioning on every individual call being detected at least twice

(Borchers and Efford, 2008). This gives:

$$
\begin{aligned}
&f(\boldsymbol{\Omega}, \boldsymbol{Y}, \boldsymbol{R}, \boldsymbol{X}, \boldsymbol{s} | n, \boldsymbol{C}) \\
&\quad = f(\boldsymbol{\Omega} | n, \boldsymbol{C}) f(\boldsymbol{Y}, \boldsymbol{R}, \boldsymbol{X}, \boldsymbol{s} | \boldsymbol{\Omega}, n, \boldsymbol{C}) \\
&\quad = \binom{n}{n_1, ..., n_{\mathcal{U}}} \prod_{i=1}^{n} f(\boldsymbol{\omega}_i | \omega^* \geq 2_i, \boldsymbol{c}_i) f(\boldsymbol{y}_i, \boldsymbol{r}_i, \boldsymbol{x}_i, s_i | \boldsymbol{\omega}_i, \omega^* \geq 2_i, \boldsymbol{c}_i) \\
&\quad = \binom{n}{n_1, ..., n_{\mathcal{U}}} \prod_{i=1}^{n} f(\boldsymbol{\omega}_i, \boldsymbol{y}_i, \boldsymbol{r}_i, \boldsymbol{x}_i, s_i | \omega^* \geq 2_i, \boldsymbol{c}_i).
\end{aligned} \tag{3.29}
$$

with $\binom{n}{n_1,...,n_{\mathcal{U}}}$ being the multinomial coefficient where $n_1, ..., n_{\mathcal{U}}$ are the frequencies for $\mathcal{U}$ unique detection histories. This coefficient appears as the order of the calls does not matter. Integrating out the latent variables, and using the independence assumption between calls, we derive the following

$$
\begin{aligned}
&f(n) \times \int_{\mathcal{S}} \int_{\mathcal{A}} f(\boldsymbol{\Omega}, \boldsymbol{Y}, \boldsymbol{R}, \boldsymbol{X}, \boldsymbol{s} | n, \boldsymbol{C}) d\boldsymbol{X} d\boldsymbol{s} \\
&\quad = f(n) \binom{n}{n_1, ..., n_{\mathcal{U}}} \\
&\qquad \times \int_{\mathcal{S}} \cdots \int_{\mathcal{S}} \int_{\mathcal{A}} \cdots \int_{\mathcal{A}} \prod_{i=1}^{n} f(\boldsymbol{\omega}_i, \boldsymbol{y}_i, \boldsymbol{r}_i, \boldsymbol{x}_i, s_i | \omega^* \geq 2_i, \boldsymbol{c}_i) d\boldsymbol{x}_1 \cdots d\boldsymbol{x}_n ds_1 \cdots ds_n \\
&\quad = f(n) \binom{n}{n_1, ..., n_{\mathcal{U}}} \\
&\qquad \times \int_{\mathcal{S}} \int_{\mathcal{A}} \cdots \int_{\mathcal{S}} \int_{\mathcal{A}} \prod_{i=1}^{n} f(\boldsymbol{\omega}_i, \boldsymbol{y}_i, \boldsymbol{r}_i, \boldsymbol{x}_i, s_i | \omega^* \geq 2_i, \boldsymbol{c}_i) d\boldsymbol{x}_1 ds_1 \cdots d\boldsymbol{x}_n ds_n \\
&\quad = f(n) \binom{n}{n_1, ..., n_{\mathcal{U}}} \prod_{i=1}^{n} \int_{\mathcal{S}} \int_{\mathcal{A}} f(\boldsymbol{\omega}_i, \boldsymbol{y}_i, \boldsymbol{r}_i, \boldsymbol{x}_i, s_i | \omega^* \geq 2_i, \boldsymbol{c}_i) d\boldsymbol{x}_i ds_i,
\end{aligned} \tag{3.30}
$$

where

$$
f(n) = \frac{\lambda^n \exp\{-\lambda\}}{n!}, \lambda = \frac{\int_{\mathcal{A}} \sum_{\boldsymbol{c} \in \mathcal{C}} \int_{\mathcal{S}} D(\boldsymbol{x}) f(s) f(\omega^* \geq 2 | \boldsymbol{x}, s, \boldsymbol{c}) ds d\boldsymbol{x}}{b}. \tag{3.31}
$$

3.3.3 Full likelihood

We now use the result from the previous section to obtain the density of the data conditional on noise levels, $f(n, \boldsymbol{\Omega}, \boldsymbol{Y}, \boldsymbol{R}|\boldsymbol{C})$:

$$
\begin{aligned}
&f(n, \boldsymbol{\Omega}, \boldsymbol{Y}, \boldsymbol{R}|\boldsymbol{C}) \\
&\quad = \int_{\mathcal{S}} \int_{\mathcal{A}} f(n, \boldsymbol{\Omega}, \boldsymbol{Y}, \boldsymbol{R}, \boldsymbol{X}, \boldsymbol{s}|\boldsymbol{C}) d\boldsymbol{X} d\boldsymbol{s} \\
&\quad = f(n) \int_{\mathcal{S}} \int_{\mathcal{A}} f(\boldsymbol{\Omega}, \boldsymbol{Y}, \boldsymbol{R}, \boldsymbol{X}, \boldsymbol{s}|n, \boldsymbol{C}) d\boldsymbol{X} d\boldsymbol{s} \\
&\quad = \frac{\lambda^n \exp\{-\lambda\}}{n!} \binom{n}{n_1, ..., n_{\mathcal{U}}} \\
&\qquad \times \prod_{i=1}^{n} \frac{1}{a_i} \int_{\mathcal{S}} \int_{\mathcal{A}} D(\boldsymbol{x}_i) f(s_i) \prod_{j=1}^{K} g_j(\boldsymbol{x}_i, s_i, c_{ij})^{\omega_{ij}} (1 - g_j(\boldsymbol{x}_i, s_i, c_{ij}))^{1-\omega_{ij}} \\
&\qquad \times \prod_{j \in \{1:K|\omega_j=1\}} \left(\frac{\exp\{\kappa \cos(y_{ij} - \mathbb{E}[y_{ij}|\boldsymbol{x}_i])\}}{2\pi I_0(\kappa)} \right. \\
&\qquad\qquad \left. \times \frac{\varphi\left(\frac{r_{ij} - \mathbb{E}[r_{ij}|\boldsymbol{x}_i, s_i, c_{ij}]}{\sigma_r}\right)}{\sigma_r \left(1 - \Phi\left(\frac{c_{ij} - \mathbb{E}[r_{ij}|\boldsymbol{x}_i, s_i, c_{ij}]}{\sigma_r}\right)\right)} \right) d\boldsymbol{x}_i ds_i,
\end{aligned}
\tag{3.32}
$$

where $a_i = \int_{\mathcal{S}} \int_{\mathcal{A}} f(\omega^* \geq 2|\boldsymbol{x}, s, \boldsymbol{c}_i) D(\boldsymbol{x}) f(s) d\boldsymbol{x} ds$.

Given our data, we obtain from this the following full likelihood

$$
\mathcal{L}_f(\boldsymbol{\phi}, \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\nu}, \kappa|n, \boldsymbol{\Omega}, \boldsymbol{Y}, \boldsymbol{R}, \boldsymbol{C}), \tag{3.33}
$$

where $\boldsymbol{\phi}$ contains the parameters associated with the specified density model for $\boldsymbol{x}_i$, $\boldsymbol{\theta}$ contains the parameters associated with the Janoschek detection function, $\boldsymbol{\eta}$ contains the parameters related to received level, $\boldsymbol{\nu}$ contains the parameters associated with the distribution of source levels, and κ is the concentration parameter associated with the bearings. Note that $n, \boldsymbol{\Omega}, \boldsymbol{Y}$ and $\boldsymbol{R}$ now refer to the observed data and no longer to the random variables. Parameter estimation could take place via maximising the likelihood function, or rather minimising the negative log-likelihood.

3.4 Concluding remarks

When acoustic detections are scarce and background noise is variable, the model presented in Chapter 2 can result in the removal of most of the available data, leaving too little to properly fit a model to. In this chapter we presented a possible solution for this problem, at the cost of complexity. We have not performed a simulation study to test the model, as this was computationally not possible within the time constraints of this PhD. If anyone considers applying this model, we believe a simulation study should be an initial step before implementing the method, which almost surely would involve improving the computational tractability of the method.

We assumed independence between several variables in this model. The consequences of violations of these independence assumptions, and potential solutions, are discussed in Chapter 2. The fitting of this model will likely involve long computation time due to the nested integrals. This was already a major challenge for the model we used in Chapter 2, where the model fitting for the simulation study took several weeks in parallel on 30+ cores, and the model presented in this chapter included an additional nested integral. To minimise the computation time, we would recommend looking for other solutions than only parallel processing, for example working with a software that uses automatic differentiation, such as `TMB` (Kristensen et al., 2016). Further, the model could be made more simple by using a step detection function similar to Chapter 2, instead of the Janoschek detection function presented here.

Chapter 4

Age is not just a number: how incorrect ageing impacts close-kin mark-recapture estimates of adult population size

4.1 Introduction

Close-kin mark-recapture (CKMR) is a method for estimating population size and other key parameters such as fecundity (and population growth and survival rates) using data on the relatedness of individuals sampled from the population (Skaug, 2001; Bravington et al., 2016b). The key rationale is that small populations will tend to contain a higher proportion of closely-related individuals than large populations.

One of the main advantages of CKMR over capture-recapture (Otis et al., 1978) and its extensions such as spatial capture-recapture (Borchers and Efford, 2008), is that it can be applied in cases when sampling is necessarily lethal, such as fisheries, and when physical recaptures are rare or impossible, where alternative metrics are often relative (e.g., catch-per-unit-effort) and potentially unreliable (Bravington et al., 2016a; Casas and Saborido-Rey, 2023). This is because CKMR does not require the recapturing of individuals, but rather their genetic markers. Offspring share genetic information

with their parents (hence 'kin'), thus they 'mark' their parents when born; through modern genetics we can compare sampled individuals with one another to see if these marks are 'recaptured'. So far, CKMR has been developed for parent-offspring pairs (POPs; e.g., Bravington et al., 2016a; Ruzzante et al., 2019; Trenkel et al., 2022), half-sibling pairs (HSPs; e.g., Hillary et al., 2018; Bravington et al., 2019; Patterson et al., 2022), and the combination of both (e.g., Bradford et al., 2018). The rise in popularity of the method has become clear from an increase in published studies involving CKMR, although the total number of applications is still small (Delaval et al., 2023). Most of the applications up to this point involved marine or aquatic species. Several salmonids have been studied (Wacker et al., 2021; Ruzzante et al., 2019; Prystupa et al., 2021), as well as large pelagic species such as southern bluefin tuna (*Thunnus maccoyii*; Bravington et al., 2016a) and the pelagic bluefin tuna (*Thunnus orientalis*; Tsukahara et al., 2023), and a variety of elasmobranchs such as white sharks (*Carcharodon carcharias*; Hillary et al., 2018), lemon sharks (*Negaprion brevirostris*; Swenson et al., 2024), thornback rays (*Raja clavata*; Trenkel et al., 2022), blue skates (*Dipturus batis*; Delaval et al., 2023), and grey nurse sharks (*Carcharias taurus*; Bradford et al., 2018). The Christmas Island flying-fox (*Pteropus natalis*; Lloyd-Jones et al., 2023) and the yellow fever mosquito (*Aedes aegypti*; Sharma et al., 2022) were the only terrestrial species that we could identify in published CKMR studies to date.

CKMR with POPs can estimate the size of the entire adult population, whereas with HSPs only the *breeding* adult population is estimated, e.g., post-reproductive adults are 'invisible' for the method (Bradford et al., 2018). Here, we focus only on POPs. For any comparison between two individuals, the probability that a potential offspring truly is the offspring of the parent is inversely related to the number of mature animals alive in the birth year of the offspring. Probabilities of finding a kin pair are expressed as a function of the expected relative reproductive output (ERRO) of the parent in the year that the offspring was conceived. This approach is parent-centric, as it starts from the point that the parent is sampled and then formulates a probability for a PO relationship (an alternative, offspring-centric formulation was proposed by Skaug (2017)).

In the simplest scenario, the probability of any adult being the parent of a juvenile reduces to two over the number of potential parents, assuming a 50:50 sex ratio; in reality, this probability is often more complicated, e.g., when reproductive output is

related to age, or when there is stock structure or population trend. To use relatedness to estimate adult population abundance with acceptable precision, it is therefore essential to accurately age the studied animals because birth year is derived from their age. Accurate ageing can be challenging: for example, epigenetic ageing requires calibration using individuals of known age (Polanowski et al., 2014; De Paoli-Iseppi et al., 2017), which is not always possible; ageing via otoliths, which are calcium carbonate structures in the inner ear, can be relatively accurate (Campana, 2001) but requires lethal sampling and is only possible for animals that have otoliths (and sharks are not among those); and ageing by counting the dental or cementum growth layer groups in teeth is not necessarily lethal and commonly used for (marine) mammals (Hohn, 2009, Chapter 9), but cannot be applied to fish species. Sharks can be aged from their vertebrae, but this is a lethal procedure and can be biased in various ways or even unusable depending on the species (Burke et al., 2020). Alternatively, length can be used to infer age through growth curves, which seems appealing as length is often recorded during sampling. Accurate estimates for growth curves of the studied species are not always available, however, and age as a function of length (age-at-length) can vary substantially between populations of the same species (e.g., Bradley et al., 2017a). Moreover, length measurements often involve measurement error. Swenson et al. (2024) studied the effects of ageing error from incorrect length measurement through simulation and found that incorrect ageing can induce substantial bias in CKMR parameter estimates. Various degrees of error were added to the true lengths of individuals, after which these were converted to ages using a von Bertalanffy growth curve. These 'incorrect' ages were then used as inputs for the CKMR model without explicitly modelling the length measurement error.

Simulation is an important tool to assess the robustness of statistical methods to violations of model assumptions (DiRenzo et al., 2023) and their performance more generally (Morris et al., 2019). Through simulation, Conn et al. (2020) studied the effects of unmodelled spatial heterogeneity on CKMR estimation and found that this can induce a negative bias in the abundance estimates; Sévêque et al. (2024) found that fitting overly simplistic CKMR models (that do not account for complex life-history traits or selective sampling) can cause biases in survival and estimates in non-trivial directions; and Waples and Feutry (2022) showed, among other things, that age-specific vital rates can bias abundance estimates from CKMR. We follow an agent-based simulation approach similar to Swenson et al. (2024) to explore the effects of

incorrect ageing on the CKMR adult abundance estimator. Unlike Swenson et al. (2024), our model does not assume length (and thus age) to be perfectly known but rather we explicitly account for the measurement error on lengths.

It is often important for demographic modelling to account for the uncertainty in the age estimates, especially when sampling probabilities depend on the age of individuals (i.e., when there is 'selectivity'), which is fundamental to fishing (Vasilakopoulos et al., 2020). Correctly accounting for ageing error is therefore still an active part of fisheries research (e.g., Hulson and Williams, 2024). Fournier and Archibald (1982) showed how ageing error in catch-at-age data can be accounted for as long as the ageing error is known. Later, Richards et al. (1992) developed statistical methodology to account for ageing error when the error is unknown, using multiple readings of fish. We are unaware of any CKMR studies in which ageing error is directly modelled and estimated. Bravington et al. (2019) accounted for the uncertainty in the ageing by first fitting a known-age CKMR model to the data and then refitting the model ten times, resampling ages from the age-at-length curve each time. In our simulation, ageing error is introduced in two ways: i) through misspecified growth curves, and ii) through incorrect length measurements, i.e., measurement error. In reality, error could also be (and almost surely is) introduced within a population through natural variation in length-at-age, e.g., as a function of genetic and environmental factors. We assume that all individuals follow the growth curve perfectly; however, one could readily interpret the length measurement error as the joint error of length measurement and length-at-age variation, or even solely as length-at-age variation if that is more appropriate for a particular case study. We assume ageing error from incorrect length measurements to be known and explicitly account for it in our model (Bravington et al., 2016b, Section 3.1.4).

The research presented in this chapter is centred around two fictional shark species that are based on a grey reef shark population (*Carcharhinus amblyrhynchos*) at Palmyra Atoll, in the central Pacific Ocean (Bradley et al., 2017b,a; Papastamatiou et al., 2018). This motivating case study consists of genetic samples that were collected from this population in 2013 and 2014. One fictional species is a simplification of the real species (hereafter referred to as the 'simple species'), and was included to test the basic performance of the model. The other fictional species has more realistic life history traits (hereafter referred to as the 'complex species'), and was included to more closely

match a real empirical study. We also compare the results for both species. It is paramount to first explore the feasibility of CKMR, for example through simulation, before committing the resources and time required for the correct collection and genetic analysis of the samples. Moreover, the findings will be be relevant to other CKMR studies when age is uncertain.

4.2 Materials and Methods

We first present our setup of the simulations for the two fictional shark species. Simulated time series are 100 years long, with sampling occurring in the final two years (mimicking the two years of sampling in the Palmyra Atoll case study). Following that, we present the POP-based CKMR models for our two species using these two years of data, followed by our estimation method and performance diagnostics. We assume that kinship relationships are known with certainty; in real life situations, one often needs to account for uncertainty in this process (Bravington et al., 2016b). All variables and quantities used in this study are summarised in Table 4.1. Code for the simulation and fitting of models was written in `R 4.3.2` and `C++14`, where the latter was linked to `R` through `Rcpp 1.0.12` (R Core Team, 2023; Eddelbuettel, 2013).

4.2.1 Simulation

We used stochastic individual-based ('agent-based') simulation. Two different 'species' were simulated separately, one with simple life history characteristics, and one with a more complex life history. For each simulation, sampling in the last two years was random, and mating occurred at random as well, i.e., mothers and fathers were matched at random, where all non-gestating mothers mated and mature males could father multiple litters in the same mating cycle. Females of the simple species always produced two offspring, whereas the litter size for the complex species ranged from 3 to 6, with equal probability. Females of the simple species reproduced every year as gestation was negligible; females of the complex species gestated for a year and therefore reproduced every other year. Newborns had age zero and sex was assigned at random with an expected 50:50 sex ratio. The survival process was Bernoulli where the annual survival probability ϕ was the same for all ages and sexes, but different

TABLE 4.1: Summary of notation.

Symbol	Description	Type
	General	
n	Number of sampled individuals (individuals can be sampled more than once)	Observed
p	Detection probability of an individual	Function
f	Probability density/mass function	Function
K	Kinship category	Latent/observed
N	Adult abundance	Parameter
	Quantities related to a captured individual	
y	Birth year	Latent
c	Capture/sampling year	Observed
ℓ	Length (when captured)	Observed
σ_ℓ	Standard deviation of the length measurement error	Parameter
a	Age (when captured)	Latent
s	Sex	Observed
$\boldsymbol{z}$	Vector of observed covariates at time of capture/sampling	Observed
	Population dynamics and demography	
ϕ	Survival probability from one year to the next	Parameter
r	Growth rate parameter from one year to the next	Parameter
α	Age of maturity	Parameter
	Subscript	
i, j	Individual i and j	
$\female, \male$	Sex, either female or male	
t	Year	
m	Simulation index	

between the two species and empirically set at a level that resulted in the yearly population growth rate equaling approximately one, i.e., no growth. Natural mortality was the only source of mortality we considered, and all individuals that reached the maximum age perished at the next survival event, i.e., animals could go through at most $a_{\max} + 1$ yearly cycles. The maximum age for sharks of the simple species was 19 years and 63 years for the complex species, where the latter matches the results from Bradley et al. (2017a). For a given species, all individuals of the same sex matured at the same age: males and females in the simple species matured at 10 years old, whereas in the complex species males matured at 17 years and females matured at 19

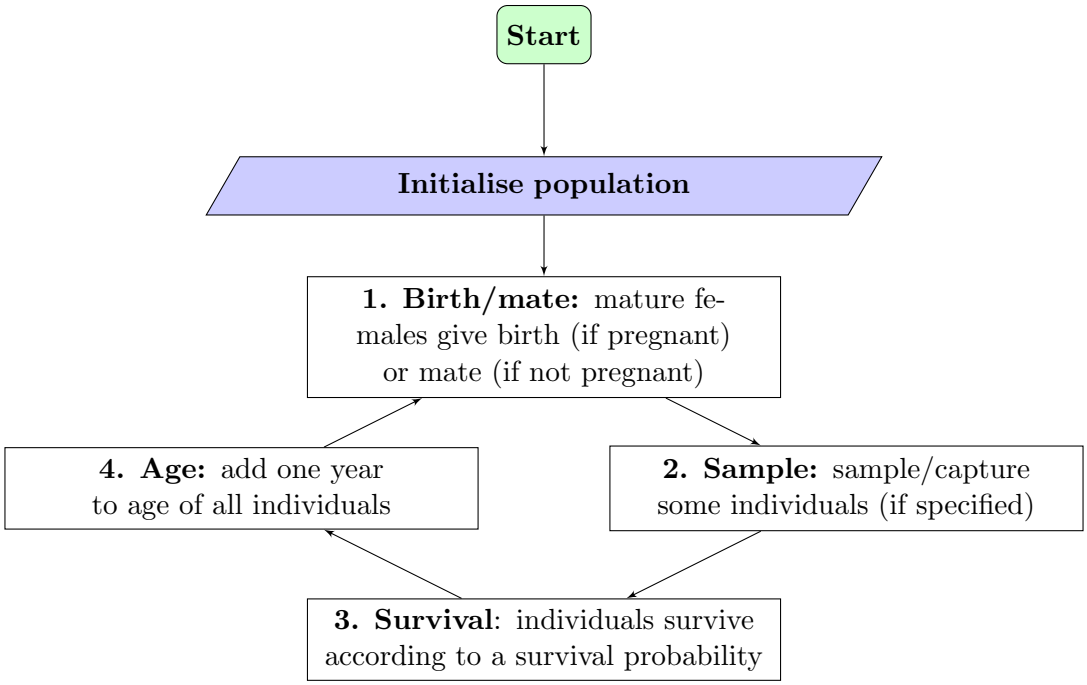


FIGURE 4.1: Flowchart representing the different stages of the life-cycle for the simulation. A population is initialised at the start of a simulation. Following that, it loops through stage 1–4 every year the simulation runs.

years of age. The length of an animal was the same for all individuals of a certain age, irrespective of sex and species. After the initialisation of a population in year zero, the simulation looped through four distinct events: a birthing/mating event, a sampling event (only in the final two years of the simulation), a survival event, and an ageing event (Figure 4.1).

For both species, we ran the simulations for one hundred years, to ensure that all animals of the initial populations would have died off. Every simulation started with 8500 individuals to stay close to the population size estimate of 8,433 by Bradley et al. (2017b), with an expected 50:50 sex ratio. At each sampling event, 375 individuals were randomly and non-lethally sampled, where re-captures were possible between sampling events. This resulted in at most 750 unique sampled individuals across the two years of sampling, which is of a similar scale as the number of genetic samples available in the motivating case study. All 750 samples were retained for analysis as there was no particular reason exclude recaptures, unlike, for example, Hillary et al. (2018), where duplicate samples were excluded from the analysis to avoid them aliasing as half-sibling pairs. For every sampled individual, the age, year of capture, and sex

were recorded; true length was derived through a von Bertalanffy growth function (VBGF; von Bertalanffy, 1938; Francis, 1988), that was specified as

$$l(a) = l_\infty \times (1 - e^{-k(a-a_0)}), \tag{4.1}$$

where $\ell_\infty = 163$ cm is the asymptotic length, $a_0 = -8.27$ the theoretical age at length zero, and $k = 0.0554$ denotes the growth coefficient. These values match the estimates of the best model in Table 2 of Bradley et al. (2017a). Gaussian noise was added to reflect (symmetric) length measurement error with variance $\sigma_\ell^2 = 2.89^2$, with over and underestimates being equally likely, after which this 'observed' length was rounded to the nearest integer. Based on these parameters, we generated 1000 different realisations of a 100-year-long population history for each species, using functions based on those from the `fishSim`-package (Baylis, 2022).

4.2.2 POP-based estimator

We developed estimators for both populations based only on POPs. Any other possible genetic relationship (such as half-sibling or self-capture) was categorised as 'not a POP'. CKMR models are generally fitted through a likelihood (function), which is constructed from the joint distribution of all pairwise comparisons between the samples, i.e., the product of approximately $n(n-1)/2$ Bernoulli trials for a POP, where n is the number of samples. We only consider pairwise comparisons and treat these as independent, whereas they clearly are not: an offspring can only have one parent of each sex. Because we ignore these higher dependencies, our likelihood is not a *true* likelihood but rather a *pseudo*-likelihood. Working with such a likelihood should not affect the point estimates but could affect other properties of likelihood-based estimation, such as variance estimation, although this effect is likely minor or even negligible provided that a small proportion of the total population is sampled, i.e., $n \ll N$ (Skaug, 2001; Bravington et al., 2016b). Because length is measured with error and age is inferred from length, age is uncertain and hence we cannot assume directionality in the comparison, i.e., who is the parent and who is the offspring. Therefore, for any comparison for individual i and j, we test both directions (parent-offspring and offspring-parent), denoted PO/OP. In practice, we tend to optimise the logarithm of the pseudo-likelihood, the so-called 'pseudo-log-likelihood', as this is generally easier

to work with and numerically more stable. Our pseudo-log-likelihood is given by

$$\begin{aligned}\log \mathcal{L}_P(\boldsymbol{\theta}|\boldsymbol{x}) =& \ell_P(\boldsymbol{\theta}|\boldsymbol{x}) = \\ &\sum_i \sum_j \log \Big\{ \Pr(K_{ij} = \text{PO/OP}|\boldsymbol{z}_i, \boldsymbol{z}_j)^{\omega_{ij}} (1 - \Pr(K_{ij} = \text{PO/OP}|\boldsymbol{z}_i, \boldsymbol{z}_j))^{1-\omega_{ij}} \Big\},\end{aligned} \tag{4.2}$$

where $\boldsymbol{\theta}$ is the parameter vector, $\boldsymbol{x}$ denotes the observed data, K_{ij} is the kinship between i and j, Pr() is the probability function, ω_{ij} is an indicator that is 1 if the kinship between i and j is observed to be PO/OP and 0 otherwise, and $\boldsymbol{z}$ denotes the information recorded about a captured individual, such as length. Age is required to calculate the probability of observing kinship, and therefore we sum over all potential ages for i, j and multiply by the probability density of that age given the measured length, $f(a|l^*)$:

$$\begin{aligned}\ell_P(\boldsymbol{\theta}|\boldsymbol{x}) = \sum_i \sum_j \log \Big\{ &\sum_{a_i} \sum_{a_j} \Pr(K_{ij} = \text{PO/OP}|\boldsymbol{z}_i, \boldsymbol{z}_j, a_i, a_j)^{\omega_{ij}} \times \\ &(1 - \Pr(K_{ij} = \text{PO/OP}|\boldsymbol{z}_i, \boldsymbol{z}_j, a_i, a_j))^{1-\omega_{ij}} \times f(a_i|l_i^*) f(a_j|l_j^*) \Big\}.\end{aligned} \tag{4.3}$$

We will now specify the two main elements of Equation (4.3), namely the probability of observing the PO/OP kinship, and the probability density of age given length.

4.2.2.1 Probability of kinship

We modelled the female and male adult abundance separately; thus, for every PO/OP comparison between two individuals we had to consider, conditional on the sexes, both combinations of which individual in older and thus the potential parent. We will first present the formulae for the simple species, followed by those for the complex species. The probability of any comparison between i and j being PO/OP is the same as the sum of testing for PO and OP separately, thus we only present the PO probabilities. For the simple species this became

$$\begin{aligned}&\Pr(K_{ij} = \text{MO}|\boldsymbol{z}_i, \boldsymbol{z}_j, a_i, a_j) = \\ &\qquad \mathbb{I}(y_i + \alpha_{\text{♀}} \leq y_j) \times \left(N_{\text{♀},y_j}\right)^{-1} \times \begin{Bmatrix} 1; & \text{if } c_i \geq y_j \\ \phi_i(c_i, y_j); & \text{if } c_i < y_j \end{Bmatrix}\end{aligned} \tag{4.4}$$

for the females, and

$$\Pr(K_{ij} = \text{FO}|\boldsymbol{z}_i, \boldsymbol{z}_j, a_i, a_j) = \\ \mathbb{I}(y_i + \alpha_{\male} \leq y_j) \times \left(N_{\male, y_j}\right)^{-1} \times \begin{Bmatrix} 1; & \text{if } c_i \geq y_j \\ \phi_i(c_i, y_j); & \text{if } c_i < y_j \end{Bmatrix} \tag{4.5}$$

for the males. Here, $\mathbb{I}()$ is an indicator function that returns 1 if its argument is true and 0 otherwise, MO and FO refer to mother-offspring and father-offspring, respectively, y denotes the birth year, α the age of maturity, $N_{s,t}$ the total adult abundance of sex s in year t, c the year of capture, and $\phi_i(t_1, t_2)$ the survival function for individual i from t_1 to t_2. As survival was assumed constant, $\phi_i(t_1, t_2)$ was defined as $\phi^{t_2 - t_1}$. Even though females could only have one litter whereas males could father multiple litters, their ERROs were formulated similarly, i.e., the reciprocal of the total mature abundance of their respective sexes. For the complex species, the probability of an MO pair thus became

$$\Pr(K_{ij} = \text{MO}|\boldsymbol{z}_i, \boldsymbol{z}_j, a_i, a_j) = \\ \mathbb{I}(y_i + \alpha_{\female} \leq y_j - 1) \times \left(N_{\female, y_j - 1} \times \phi\right)^{-1} \times \begin{Bmatrix} 1; & \text{if } c_i \geq y_j \\ \phi_i(c_i, y_j); & \text{if } c_i < y_j \end{Bmatrix} \tag{4.6}$$

and the probability of an FO pair became

$$\Pr(K_{ij} = \text{FO}|\boldsymbol{z}_i, \boldsymbol{z}_j, a_i, a_j) = \\ \mathbb{I}(y_i + \alpha_{\male} \leq y_j - 1) \times \left(N_{\male, y_j - 1}\right)^{-1} \times \begin{Bmatrix} 1; & \text{if } c_i \geq y_j - 1 \\ \phi_i(c_i, y_j - 1); & \text{if } c_i < y_j - 1 \end{Bmatrix}. \tag{4.7}$$

The two key differences between the complex species relative to the simple one were that 1) a potential father only needed to have been alive the year before the birth of the offspring, whereas a potential mother needed to have survived until birthing, and 2) the potential parents needed to have matured at least one year before the birth year. To illustrate this, imagine that we are comparing two individuals from the complex species, where the parent is female, and we know the individuals' ages. The offspring was caught in year 50 at age 3, and thus born in year 47. The potential parent was female, and caught in year 45 and would have needed to survive for at least two years in order to be a potential parent; she was 36 years old at the time of capture, and thus

born in year 9. The ERRO for this parent in the year of mating, i.e., the year before the birth year of j, is the reciprocal of the number of females alive in that year who also survived one year of gestation, which is ϕ. Therefore, the probability that i is the mother of j would be:

$$\begin{aligned}\Pr(K_{ij} = \text{MO}|\boldsymbol{z}_i, \boldsymbol{z}_j, a_i, a_j) =& \mathbb{I}(9 + 19 \leq 47 - 1) \times \left(N_{♀,y_j-1} \times \phi\right)^{-1} \times \phi^{47-45} \\ =& 1 \times \left(N_{♀,y_j-1} \times \phi\right)^{-1} \times \phi^2 \\ =& \phi \left(N_{♀,y_j-1}\right)^{-1}\end{aligned} \tag{4.8}$$

Every comparison, given a_i and a_j, contains a signal about the adult population in a specific year. We assumed a constant population size, and thus $N_{s,t} = N_s$. We also developed and tested a model that included sex-specific growth parameters. This model was internally inconsistent and therefore not included in the main body of this thesis for any formal inference. However, we did include the derivations and some results in Appendix B.3.

4.2.2.2 Probability density of age given length

We had an assumed true length-at-age curve $l(a)$ (Equation (4.1)) and we knew that there was measurement error on lengths. Denoting the *measured* length as l^*, we derived the probability density $f(a|l^*)$ using Bayes' rule as follows:

$$f(a|l^*) = f(l^*|a)f(a)f(l^*)^{-1}. \tag{4.9}$$

Measured length given age $l^*|a$ was assumed to follow a discretised Normal distribution, as lengths were rounded to the nearest centimetre. We followed Roy (2003) in defining this distribution as

$$f(l^*|a) = \Phi\left(\frac{l^* - \mu - 0.5}{\sigma_l}\right) - \Phi\left(\frac{l^* - \mu + 0.5}{\sigma_l}\right), \tag{4.10}$$

where Φ denotes the standard normal cumulative distribution function, the expectation μ is given by Equation (4.1), and σ_l captures the standard deviation of length measurement error. As the sampling probability in the simulation was unrelated to age, the age distribution of sampled individuals was the same as the age distribution

in the whole population, and we did not need to distinguish between the two. We assumed that the population had a stable age distribution with no growth, which meant that the distribution of ages, had we not imposed a maximum age, would have been geometric with shape parameter being equal to the mortality rate, which is $1 - \phi$. Acknowledging that there was a maximum age, $a_{\max}$, we needed to condition on the age being at most this age, and thus

$$f(a) = \begin{cases} \frac{(\phi)^a(1-\phi)}{1-(\phi)^{a_{\max}+1}}; & \text{if } 0 \leq a \leq a_{\max} \\ 0; & \text{otherwise} \end{cases}, \tag{4.11}$$

where the numerator and denominator were the geometric probability mass and cumulative distribution functions, respectively. Note here that we used the definition of a geometrically distributed variable being the number of failures (survival) until a success (death) occurs. Finally, the probability density function on measured length became

$$f(l^*) = \sum_{a=0}^{a_{\max}} f(l^*|a)f(a). \tag{4.12}$$

4.2.3 Fitting

The parameters in the CKMR model were estimated from the samples collected in the last to years of the simulation by maximising the pseudo-log-likelihood, which can involve prohibitively long computation time. To resolve this, we restricted the number of pairwise comparisons. Many pairwise comparisons resulted in identical probabilistic statements, and thus in practice only needed to be derived once. As we considered adult abundance for both sexes separately, we estimated two parameters: $N_{\female}$ and $N_{\male}$. All other parameters, such as ϕ, were assumed known and fixed. To each of the 2000 population realisations (1000 for each species) we fitted the appropriate POP-model with varying degrees of length measurement error and growth curves, which was achieved by altering some of the fixed parameters. Specifically, we assumed five different standard deviations for length measurement error: the correct one, a 33% and 67% underestimate, and a 33% and 67% overestimate. We also considered five different growth curves: the correct one, two that were shifted upwards by 5% and 10%, and two that were shifted downwards by 5% and 10%. These growth curve shifts were aimed to represent real variation in growth curves between populations of the

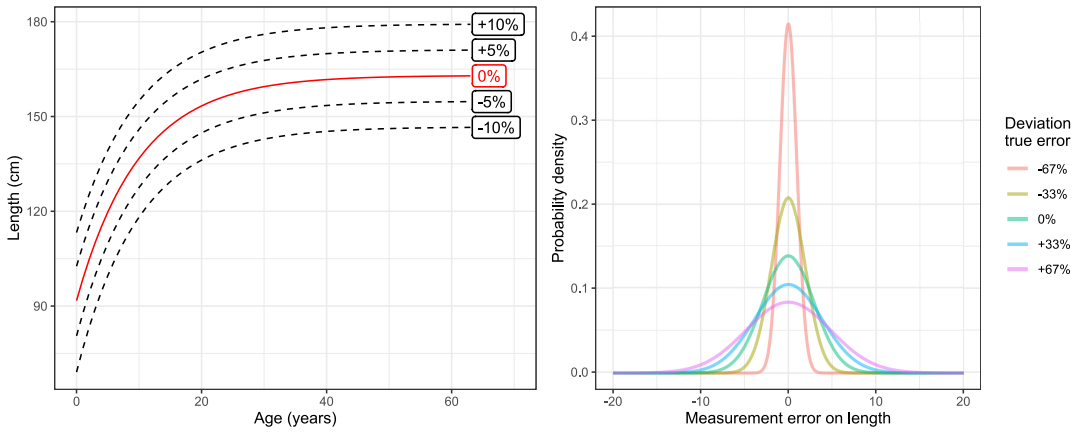


FIGURE 4.2: The left panel shows the five growth curves that were used in the scenarios tested in this study. The true growth curve is indicated in red; the black dotted-lines show the incorrect ones, which were constructed by shifting the growth curve up and down in steps of 5 percent. These shifts were aimed to represent real variation in growth curves between populations of the same shark species (Bradley et al., 2017a). The right panel shows five measurement errors used in this study. The true simulated error was 2.89 cm, and the other measurements errors were chosen by deviating from this error in both directions.

same shark species (Bradley et al., 2017a). This resulted in a total of 25 combinations or 'scenarios'. We labelled these scenarios using the format 'ME±XX:GC±YY', where ME refers to the measurement error, XX denotes the percentage over- or underestimate, GC stands for growth curve, and YY denotes the percentage of up- or downwards shifting; for example, the scenario with a 33% overestimated standard deviation of length measurement error and a 5% downshifted growth curve had the label ME+33:GC-5. These measurement errors and growth curves are visualised in more detail in Figure 4.2. Considering 25 scenarios for every simulation resulted in the fitting of 50,000 models in total. To keep computation time to a minimum, we implemented most of the fitting process in `C++`.

4.2.4 Variance and performance

To evaluate the performance of the estimator, we present the following metrics: i) mean error and mean relative error to evaluate a potential bias; ii) median error and median relative error to evaluate the median bias, which uses the median instead of the mean, as the median is often more appropriate when distributions are skewed. In addition, the mean absolute error (MAE) and root mean square error (RMSE)

are presented in supplemental tables. The definitions of the six metrics are given in Appendix B.1.1. Furthermore, we derived the 95% log-normal confidence interval (CI) coverage to evaluate the performance of these CIs in the correct growth curve scenarios. Variance was estimated from the Hessian matrix produced by the maximum likelihood estimation, and averaged over these 1000 estimated standard errors. We can treat the pseudo-likelihood as a true likelihood as long as sampling was sparse (see Section 4.2.2). It is unclear if this criterion was met in our study, as we took 750 samples from a population with roughly 8,500 individuals. If sampling is not sparse, the estimated variance could be negatively biased as the pairwise comparisons are not approximately independent. To explore the extent of this potential bias, we evaluated how well the average estimated standard error estimated the empirical standard deviation of population estimate errors across the 1000 simulations for each species. We include definitions of these in Appendix B.2.1.

4.3 Results

The mean number of POPs for all sampling realisations was 48.6 (range: 25–76) for the simple species and 55.6 (range: 31–90) for the complex species. Mean simulated adult abundances in the final year of the simulation were 794 and 793 (range: 630–992 and 600–1019; ♀ and ♂) for the simple species and 514 and 650 (range: 400–683 and 516–824; ♀ and ♂) for the complex species. A small number of recaptures between the two years of sampling, i.e., that some individuals were sampled at more than one sampling event, occurred in every simulation. This ranged from 4 to 25 recaptured individuals for the simple species, and 5 to 30 individuals for the complex species. The simulated mean annual growth rate was 0.999 for both sexes of the simple species, and 1.001 and 0.998 for the males and females of the complex species, respectively; the mean annual growth for any simulation was always within 0.3 percent point from the mean across all simulations. The fitting algorithm did not always converge when the measurement error standard deviation and/or the growth curve was (very) negatively biased. Whenever this happened, it happened for most of the simulations in that scenario. Therefore, we excluded the scenarios where this happened from the analysis, which led to the exclusion of scenarios ME-67:GC-10, ME-67:GC-5, ME-67:GC+0, ME-33:GC-10, ME-33:GC-5, and ME+0:GC-10. In the other scenarios, all models converged successfully.

For the simple species, median errors for $\widehat{N}_s$ when using correct measurement error and growth curve specification (ME+0:GC+0) were 20.83 and 22.52 (relative: 2.57% and 2.79%; ♀ and ♂) individuals (Figure 4.3; Tables B.1 and B.2). For the complex species, median errors for $\widehat{N}_s$ when using correct measurement error and growth curve specification were 6.48 and 10.28 (relative: 1.35% and 1.59%; ♀ and ♂) individuals (Figure 4.4; Tables B.3 and B.4). For the simple species, median relative errors in abundance estimates were positive but close to zero for all deviations from the true standard deviation of length measurement error provided that the growth curve was correctly specified, although they were slightly larger for the females (Figure 4.3, also Tables B.1–4). For any given measurement error standard deviation, we observed a trend from a positive median error to a negative median error as we shifted the growth curve upwards (Figures 4.3 and 4.4). When growth curves were shifted down 5%, this resulted in median relative errors of around 30% for the simple species, and between 30 and 60% for the complex species. Shifting growth curves up by 5% resulted in median relative errors of -30% for the simple species, and between -30 and -40% for the complex species.

When the growth curve was correctly specified, the 95% CI coverage (rounded to one decimal place) for the simple species adult abundance estimates ranged from 96.1% to 96.4%, and ranged from 94.4% to 95.9% for the complex species estimates (Table B.7). For a given growth curve, no relation between the measurement error standard deviation and CI coverage became apparent. Incorrectly specified growth curves severely lowered the CI coverage for all measurement errors for both species (Table B.7). When the growth curve was correctly specified, the empirical standard errors ranged from 180.97 to 182.56 for the male simple species, and 190.85 to 192.56 for the female simple species (Table B.5); for the complex species, these errors ranged from 145.77 to 148.42 for the males, and from 107.51 to 109.27 for the females (Table B.6). Given a growth curve, increasing the measurement error standard deviation seemed slightly decrease the empirical standard errors, for all species and sexes. Given an assumed standard deviation of measurement error, the empirical standard errors decreased as the growth curve was shifted upwards. When the growth curve was correctly specified, the average estimated standard errors ranged from 181.04 to 183.04 for the male simple species, and 183.54 to 185.51 for the female simple species (Table B.5); for the complex species, these errors ranged from 139.79 to 142.41 for the males, and from 102.99 to 105.71 for

the females (Table B.6). The changes in average estimated standard errors between the scenarios follow a similar pattern to the empirical standard errors.

Whenever the growth curve was correctly specified, average model estimated standard errors were always (slightly) lower than empirical standard errors all scenarios and both species, except for the male simple species (Tables B.5 and B.6). The underestimation of the empirical standard error by the average estimated standard error was always $< 5\%$. Deviations from the correct growth curve increased underestimation in all cases (Tables B.5 and B.6).

4.4 Discussion

In this study, we explored the effects of incorrect age inference from length measurements on CKMR estimates of adult abundance through misspecifying the length measurement error and the growth curve in various ways. The number of POPs discovered in our simulation was in the vicinity of the 50–100 kin pairs recommended for a CKMR application (Bravington et al., 2016b), albeit on the lower end. Overall, an incorrect assumed standard deviation of measurement error mostly impacted the convergence likelihood of the fitting algorithm, whenever this standard deviation was assumed to be smaller in the fitting than was true for the simulation. Whenever the measurement error standard deviation was high enough to allow for convergence, it made little difference whether it was the true value or if a much higher standard deviation was assumed. This would suggest that, if researches are ever unsure about whether their assumed degree of spread in length measurement error is correct, it is safer to overestimate it. A misspecified growth curve, on the other hand, had drastic effects on the estimation of all parameters: a 5% shift away from the true growth curve resulted in biases ranging from -60% to +40%; estimated and empirical standard errors seemed to scale with the abundance estimates, and shifting away from the true growth curve resulted in an increased underestimation of empirical standard errors.

The model performed well under correct specification (scenario ME+0:GC+0), although the positive median relative error in adult abundance estimates suggests a positive median bias. This error was more extreme for the simple than for the complex species. A bias in the estimates is not uncommon for maximum likelihood methods

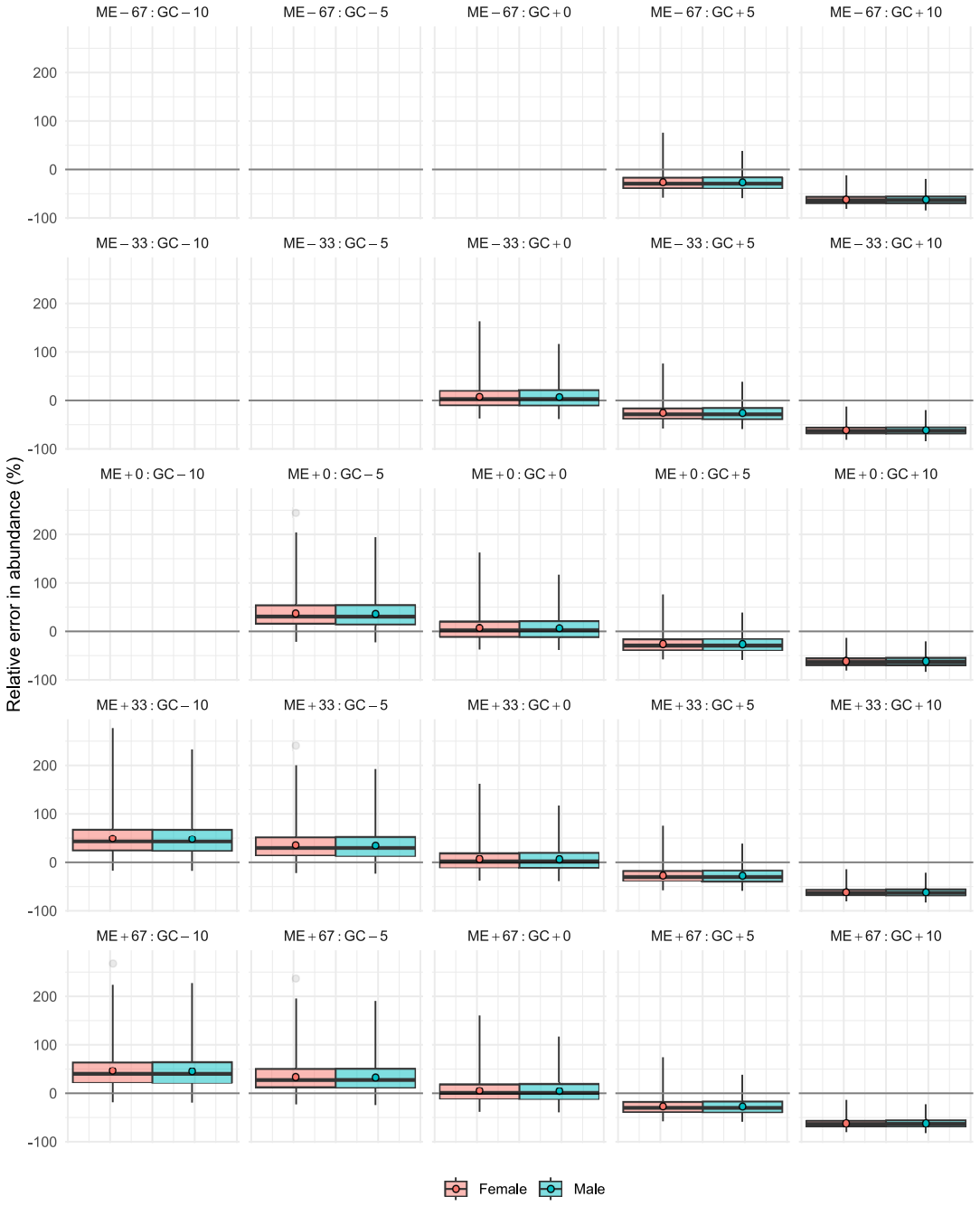


FIGURE 4.3: Box plots for the error in estimated sex-specific adult abundance relative to the true abundance for the simple species. We only present the results for scenarios in which the optimiser consistently converged; this meant that some scenarios were left blank. Box plots show the interquartile range (IQR) and the median; the mean is indicated by the darker filled circle; the vertical lines cover five times the IQR; and all values outside of that are indicated as outliers. The scenarios were labelled using the format 'ME±XX:GC±YY', where ME refers to the measurement error, XX denotes the percentage over- or underestimate, GC stands for growth curve, and YY denotes the percentage of up- or downwards shifting; for example, the scenario with a 33% overestimated standard deviation of length measurement error and a 5% downshifted growth curve had label ME+33:GC-5.

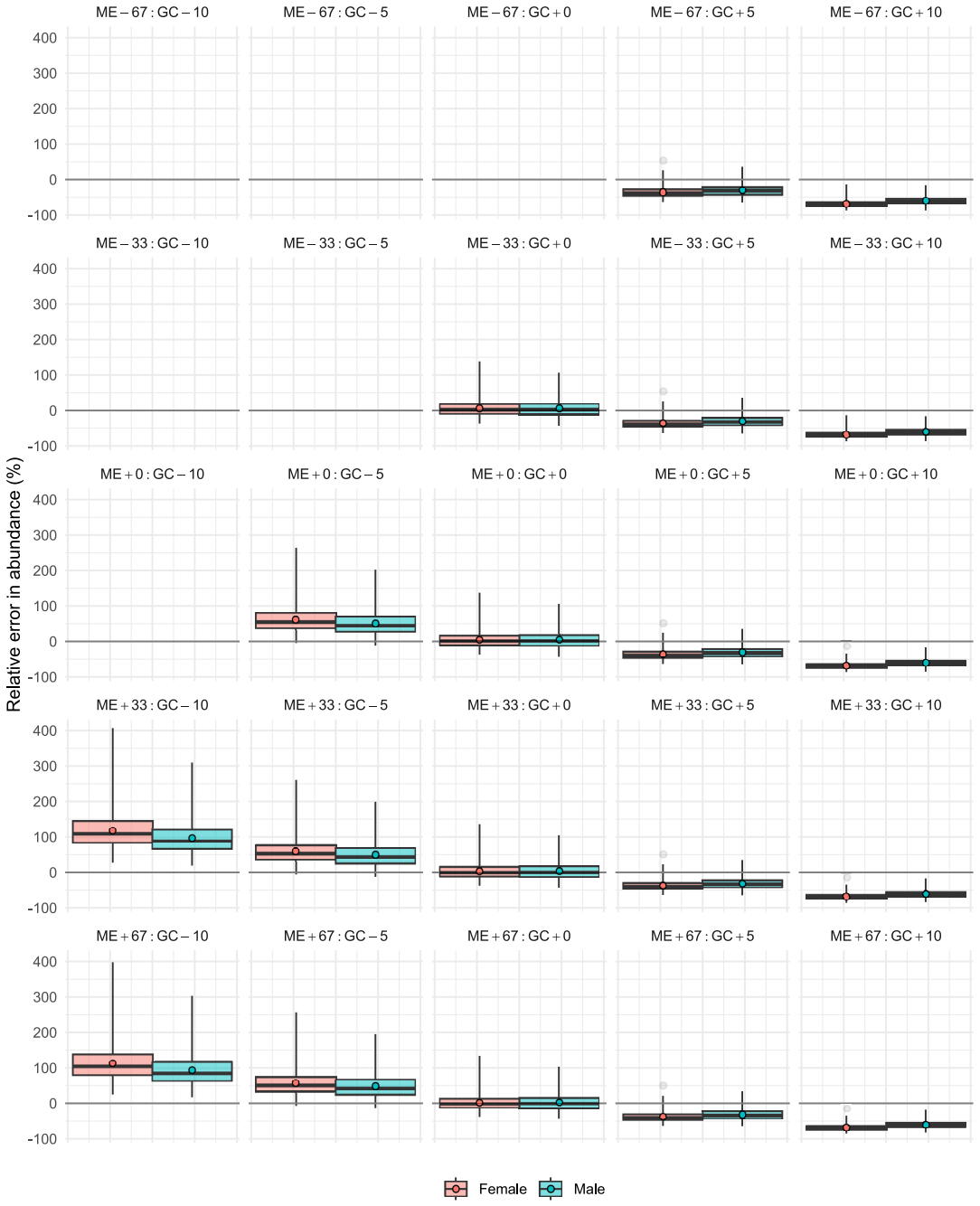


FIGURE 4.4: Box plots for the error in estimated sex-specific adult abundance relative to the true abundance for the complex species. We only present the results for scenarios in which the optimiser consistently converged; this meant that some scenarios were left blank. Box plots show the interquartile range (IQR) and the median; the mean is indicated by the darker filled circle; the vertical lines cover five times the IQR; and all values outside of that are indicated as outliers. The scenarios were labelled using the format 'ME±XX:GC±YY', where ME refers to the measurement error, XX denotes the percentage over- or underestimate, GC stands for growth curve, and YY denotes the percentage of up- or downwards shifting; for example, the scenario with a 33% overestimated standard deviation of length measurement error and a 5% downshifted growth curve had label ME+33:GC-5.

when the sample size is small, which could be true in our study as the number of sampled POPs never surpassed 76 for the simple species and 90 for the complex species. However, it could also be that this shows a slight positive bias in the method itself, especially as a previous CKMR simulation study by Conn et al. (2020) found small positive biases in the abundance estimates, too. We can express the empirical standard errors for the correct scenarios as percentages of the associated mean simulated abundances in the final year. This gives standard errors relative to the mean (also known as coefficients of variation) of 23.0% and 24.3% for the male and female estimates of the simple species, respectively, and 22.7% and 21.2% for the male and female estimates of the complex species, respectively. These are high but not uncommon for real-life population studies. The 95% log-normal confidence intervals (CIs) seemed to accurately represent the uncertainty around the estimates whenever the correct growth curve specification was used, as the coverage ranged from 94.4% to 96.4%. Nonetheless, the coverage always exceeded 95% when the entire model was correctly specified, which could indicate that the 95% log-normal CIs were slightly conservative.

In this study, we assumed that all individual sharks followed the specified growth curve perfectly, and any variation in lengths for a given age resulted from measurement error. This is a simplification of reality, and future research could focus on ways to accommodate natural variation in length at a given age, which could be a function of age in itself. As an incorrect standard deviation of length measurement error seemed to have little effect on point estimates, we believe that, when in doubt, it is preferable to assume a higher standard deviation as this improves how likely it is that the fitting algorithm converges.

The effects of deviating from the true growth curve on the adult abundance estimates were substantial. When growth curves were shifted by 5% we often observed median relative errors of over 30%. This strongly highlights the sensitivity of the method to correct age estimation. Empirical standard errors were also increasingly underestimated when growth curves were shifted away from the the truth. This effect was stronger when the growth curve was shifted upwards, i.e., when ages were being underestimated. An underestimation of uncertainty could be a consequence of the comparisons not being truly independent, i.e., a violation of the sparse sampling assumption. It is important here to note that we did not evaluate the standard deviation of abundance estimates but rather of the error in abundance estimates (see Appendix

B.2.1). The true abundance was different for every simulation, so we could not use the standard deviation of the abundance estimates itself, since this would partly capture the stochasticity of the simulation process. To overcome this, we used the standard deviation of the error in abundance estimates, which should be a more robust measure of the true standard error. This should not be a problem as long as estimation is unbiased; however, our results indicate a slight positive bias, which could have impacted the accurateness of the empirical standard error in being a measure of true standard error. CI coverage was most severely impacted by incorrect growth curves; however, this was likely mostly due to the bias in the estimates in those scenarios. In real-world applications, researchers could potentially check the correctness of their assumed growth curve by assessing the distribution of lengths/ages among the sampled individuals. If many of the observed lengths are either associated with very low ages or are close to asymptote, or in some other way exhibit an unexpected sampled age distribution given the sampling scheme, then this could be an indication that the assumed growth curve is incorrect (or that sampling assumptions are violated).

Even though recaptures should be rare (as long as sampling is sparse), they did occur in our simulations between sampling years. These recaptures did not pose any problems within the analysis, e.g., getting mistaken for a different genetic relationship, and thus we retained the recaptures in our data. Alternatively, duplicate samples can be excluded from the analysis when there is reason to do so. We hypothesise that excluding recaptures would likely increase estimates of uncertainty, as fewer observations are used for the analysis. We are unaware of any study that investigated the extent to which including recaptures could potentially affect precision or even bias in CKMR estimation, and we believe that this could be a great topic for future research. Whenever it is known that multiple samples belong to a single individual, there exists the potential for extending CKMR by incorporating some form of capture-recapture into the method (Otis et al., 1978; Bravington et al., 2016b). Additionally, it could also allow us to fit the growth curve jointly with the CKMR model, instead of assuming it to be known by extrapolating from other studies (Bravington et al., 2019). This could create a situation where one collects new samples every year to update the model, thereby continuously improving the estimates not only of the abundance and trend, but also of the growth curve: in a Bayesian framework, one could use the initial growth curve as prior information, and then update the posterior every year as more information is collected.

In our model, we did not allow for any growth or decline in the population size over time. Our simulated populations exhibited no systematic growth, but the stochastic nature of the process did lead to some random growth/decline. One could consider estimating a growth rate, or assume a small range of growth rates (see Hillary et al. (2018) for an example with white sharks (*Carcharodon carcharias*)). The main challenge would be to understand how including a growth rate parameter affects the assumed age distribution $f(a)$. We can imagine three general population growth scenarios. If a population is stable but growing or in decline, the assumed age distribution will be geometric and depend on a combination of survival and growth rate (Caswell, 2006, Section 4.5.2.1). The second scenario is when a population exhibits changing growth or decline, in which case there is no stable age structure. We believe that this scenario is intractable, and it would make a good subject for a robustness study to see how much it affects estimation. The third scenario would be where there is no expected population growth or decline but there is demographic stochasticity, which in practice could result in deviations from the stable age structure. For this scenario, an option could be to use the method described by Hillary et al. (2018), where the measured lengths were binned and a multinomial distribution was fit to these binned data to estimate the distribution of sampled ages. Still, this could be a topic for future research to see what other methods exist to find the distribution of (sampled) ages.

CKMR involves many pairwise comparisons, which often involves many identical probabilistic statements. To limit computation time, we evaluated unique probabilistic statements only once. If further computational improvements are required, it is possible to reduce the number of pairwise comparisons that are evaluated by excluding a subset of comparisons from the analysis. For example, the length-age relationships are often much clearer for younger animals, and therefore one could choose to only consider animals up to a certain size as potential offspring (Trenkel et al., 2022).

In our simulation and model, we assumed some life history traits to be fixed and known, but this is not always required for CKMR. We estimated sex-specific adult abundance only in our model and assumed quantities such as survival to be known and fixed. In order to relax the assumption of a fixed and known survival parameter ϕ, one could estimate it by including half-sibling pairs alongside parent-offspring pairs (Bravington et al., 2016b). Parent-offspring pairs can be used to model fecundity as long as the parameter appears explicitly in the model, which could be the case when

fecundity varies with the size or age of animals (Bravington et al., 2016b, Section 3.1.4) We are unaware of any attempts to estimate time-varying fecundity or survival, and we believe this to be a potential direction for future research. Moreover, we assumed maturity to be knife-edge as it slightly reduces complexity of the model. However, if maturity occurs more gradually, then this can be accommodated by adding a fecundity curve to the model (e.g., a logistic curve; Conn et al., 2020). We also imposed a fixed and known maximum age in the simulation, mostly to reduce computation time. In reality, animals do not always have a maximum age; in such cases, one could set the maximum age equal to an age that the animal has practically zero probability of reaching. Further, we made the assumption that sampling was random with respect to age, i.e., that there is no selectivity, which does not necessarily need to be true in reality. When accounting for ageing error when there is selectivity, it will be necessary to include some function relating true age to observed age, which would depend on the probability of being sampled at a given true age. Finally, we have not considered fishing-induced mortality, as our case study concerned an area protected from fishing. This and other anthropomorphic sources of mortality should be accounted for whenever they are present, analogously to Bravington et al. (2016a).

When a promising method like CKMR is first presented, one can see the appeal to start studying populations as quickly as possible. Benchmark comparisons could be useful (e.g., Ruzzante et al., 2019) to compare a new method to some 'truth'. However, these comparisons can be ambiguous when it is unclear how accurate the benchmark truly is. Simulation studies, such as this one (and see Conn et al. (2020) for the effects of unmodelled spatial heterogeneity on CKMR), are a key part of understanding when the CKMR method works well and when it does not. We believe the CKMR method has great potential and, in some cases, is an improvement over other methods, but our study confirms that care that needs to be taken when ageing is biased. In such cases epigenetic ageing could be preferable, even though epigenetic ageing can still involve substantial uncertainty (e.g., Larison et al., 2021; Prado et al., 2021) and relies strongly on the quality of the training data (Mayne et al., 2023).

Chapter 5

Baited random encounter modelling for sharks on the Chagos Archipelago

5.1 Introduction

The methods presented in this thesis so far have relied on some unique identifier, that can be *captured* and *recaptured* to estimate the abundance of individuals or cues. A common way of introducing such a marker is by capturing animals and physically marking them, either using tags or permanently marking some part of their bodies. However, these methods have been criticised over the years due to their invasive nature (e.g., White et al., 2012; Jepsen et al., 2005). Alternatively, when animals are genetically unique, their marker can be of a genetic nature (Yoshizaki, 2007). This only works for animals that are genetically unique, which is not always the case (e.g., certain armadillo species; Bravington et al., 2016b). Genetic markers can also be challenging to obtain without catching the animal, especially in a marine environment where sources of genetic material such as faeces quickly dissipate (Creel et al., 2003). Some animals can be uniquely identified by their physical markings. Examples of these in a terrestrial context are snow leopards or tigers (Jackson et al., 2006; Karanth and Nichols, 1998); examples of marine animals are whale sharks, sea turtles, and cetaceans (McCoy et al., 2018; Schofield et al., 2008; Constantine et al., 2012). This allows for

much less invasive sampling techniques as physical captures are no longer necessary. Furthermore, it allows for the usage of camera traps or even continuous video, which means that researchers no longer have to make observations *in situ*, but rather can analyse the footage later. This is especially favourable for elusive animals, as you need to sample over a large period and/or area to obtain enough detections for any reliable inference, and for animals that live in places that are hostile for human observers, such as marine environments. However, using (remote) cameras is not suited for all animals, as some do not have natural markings or do not have markings that reliably allow for unique identification, such as certain species of rodents (Caravaggi et al., 2016). The study presented in this chapter was motivated by a data set consisting of shark footage from baited cameras placed on the Chagos Archipelago. Such a camera system is generally referred to as a baited remote underwater video (BRUV) system. These systems generally consist of a rig with one or two cameras (so-called 'stereo-BRUV') mounted in the centre and a container in front of the camera to put the bait in (Mallet and Pelletier, 2014) This container has holes in it so that the water interacts with the bait and the scent of the bait (the 'bait plume') can spread through the water, attracting animals from beyond the field-of-view (FOV) of the camera(s) (Langlois et al., 2020). BRUV systems can be placed on the sea bed (e.g., Harvey et al., 2012; Brooks et al., 2011), or drift in mid-water (e.g., Bouchet and Meeuwig, 2015; Letessier et al., 2019). As sharks are generally elusive, using remote underwater systems seems appealing; however, sharks can have little or no natural markings to uniquely identify the individuals.

The lack of natural markings makes the application of capture-recapture methods challenging in such scenarios. An alternative method that does not rely on individual identification is random encounter modelling (REM; Rowcliffe et al., 2008). Another possible framework would be camera trap distance sampling (e.g., Howe et al., 2017) if distances from the animals to the camera could be obtained; here, we focus on REM. REM utilises the ideal gas theory, i.e., the theory describing how gas molecules move around in a volume. If one assumes that animals move around an area randomly, then this method can be used to produce unbiased estimates of animal density by relating the number of animals that enter the FOV of a camera to the total number of animals in an area. Over a given time frame, the number of animals entering the FOV of the video is proportional to the total number of animals. REM produces unbiased estimates even when animal movement is non-random, as long as their non-random

movement is unrelated to the presence of cameras (Rowcliffe et al., 2013). More detail on REM is presented in Section 5.2.

REM overcomes the issue introduced by the lack of individual identification, but the elusive nature of sharks still poses a challenge. Visual observation is inefficient underwater (see Chapter 1) which reduces the area that is sampled by the cameras. The solution to this by researchers is to add bait to lure animals towards the cameras from beyond the FOV (Mallet and Pelletier, 2014). This greatly increases the area sampled and hence the encounter rate. There are, however, at least five challenges when using these data for absolute abundance estimation. First, the movement of animals is no longer random with respect to the cameras, which violates one of the main assumptions of REM. Second, the sampled area is now much harder to estimate, as it is no longer defined by FOV of the cameras, but by the area covered by the bait plume. The bait plume spreads after the bait is placed, and so the sampled area becomes a function of time. REM estimates density rather than abundance alone, so knowledge of the sampled area is required. Third, as a shark is only visible once it enters the FOV of the camera, it is challenging to determine exactly when it encountered the bait plume. The distance travelled could be informative about the density of the sharks, but strong assumptions have to be made in order to create such an estimator (Priede et al., 1990, 1994; Priede and Merrett, 1996; Priede and Bagley, 2000). Fourth, individuals might not always respond to the bait plume. The bait plume dilutes as it spreads, which means that the effective sampled area is potentially much smaller than the absolute sampled area, and one needs to account for this to avoid underestimating density. There could also be other reasons for a shark not to respond to the bait, for example because it is satiated. Last, as sharks might circle the bait and the camera for some time after the initial encounter, it could be challenging to distinguish what really is a novel encounter and what is not.

In this chapter, we will first present a summary of previous work on using BRUV data for absolute abundance estimation. Next, we provide a summary of REM and some extensions of the method since its introduction. Third, we discuss the ways in which REM can be extended to work with data from baited cameras, for which we derived an analytical solution for the expected distance from the animal to the camera and used simulation to better understand the area covered by the bait plume. Finally, we present our findings and make recommendations for future directions.

5.1.1 Previous work

The challenges presented in the previous section have resulted in data from BRUV systems generally having been used for relative abundance measures, often Max_N (Sherman et al., 2018; Schobernd et al., 2014). Max_N is the maximum number of individuals in one frame over all recorded video frames, and hence produces an absolute underestimate of abundance, assuming that no individuals are counted twice. The link between Max_N and absolute abundance has been studied (e.g., Dunlop et al., 2015), but the link remains highly context-dependent due to differences in animal behaviour. Max_N is an order statistic, which limits its usability for population size inference. Another limitation is that Max_N does not have a linear relationship with abundance or density, as activity at the detector might attract some animals and repel others. Animals might also leave the bait after varying amounts of time. Another metric that is commonly extracted from BRUV systems is time-to-first-arrival (TFA; Priede et al., 1990). The time it takes until the first individual of the focal species arrives at the bait should be inversely related to the area covered by the bait plume, the species' density, and the movement speed of the animals. Priede and Merrett (1996) attempted to quantify this relationship for several abyssal demersal fish species. They viewed the ocean floor as a two-dimensional area and made the following assumptions: first, all fish are evenly distributed along the sea floor and occupy a regular hexagon with common sides length r and thus area $3\sqrt{3}r^2/2$; second, all fish are static until they encounter the bait plume; third, when a fish encounters the bait plume, it always responds by swimming to the bait in a straight line; fourth, the odour of the bait will spread down-current with the same speed as the water velocity; and finally, the odour can never pass between the two nearest fish that are down-current. Using these five assumptions, they derived the following relationship between the hexagonal-shaped area occupied by an individual, A_indiv and TFA:

$$A_\text{indiv} = \frac{3\sqrt{3}r^2}{2}, \tag{5.1}$$

where $r = t_\text{arr}/(V_f^{-1} + V_w^{-1})$, t_arr denotes the TFA in seconds, V_f the approach speed of fish and V_w the water velocity, both in meters per second. Total abundance N is then estimated by dividing the total study area by A_indiv.

The assumptions introduce a challenge, as they strongly restrict the set of animals that this method could be applied to. If animals are evenly distributed and if they move

only when they detect the bait, then this method might work; this could be the case for the type of abyssal species that motivated the development of this method (Priede et al., 1990). However, sharks (see Section 5.1.2) are not static and some species are known to congregate (e.g., Speed et al., 2011), which would mean that the assumptions underpinning this method are violated. Moreover, this method assumed a process in two-dimensional space. Even though this might work for species that tend to live at the bottom such as rays, the modelling of movement for more pelagic (shark) species could require a model that uses three-dimensional space.

5.1.2 Chagos Archipelago

The Chagos Archipelago is positioned in the centre of the British Indian Ocean Territory (BIOT; Figure 5.1). The BIOT consists of seven atolls and contains a joint coral reef surface of roughly 15,000 km^2 (Ferretti et al., 2018). The shark communities living around the Chagos Archipelago were still considered pristine only 50 years ago, but the number of sharks sighted decreased drastically from 4.2 per scientific dive in the 1970s to 0.4 in 2006 (Graham et al., 2010). This decline is most likely due to poaching by large distant fishing fleets, indicating that even these remote marine areas are vulnerable (Graham et al., 2010). In 2010, the BIOT became a no-take zone of 640,000 km^2 and thereby one of the largest marine protected areas in the world (Wilhelm et al., 2014). The decline of shark populations on Chagos was confirmed by Ferretti et al. (2018). They estimated 571,130 (standard deviation: 124,915) grey reef sharks (*Carcharhinus amblyrhynchos*) and 31,693 (standard deviation: 18,513) silvertip sharks (*Carcharhinus albimarginatus*) inhabited the archipelago in 2012, about 79 and 7% of their respective baselines. These estimates were derived through state-space Bayesian surplus production models using historical catch data from illegal fishing and monitored industrial fishing, historical abundance indices from scuba and longline surveys, and ecological theory.

Our attempt to estimate absolute abundance from BRUV data was motivated by a case study on these same species. The data available consisted of footage from mid-water and seafloor BRUV systems from 2012–2016, and additional information extracted from that footage such as TFA and Max_N. Even though the main species of interest were grey reef sharks and silvertip sharks as these were most often encountered, data

were also available for other species such as silky sharks (*Carcharhinus falciformis*) and whitetip reef sharks (*Triaenodon obesus*).

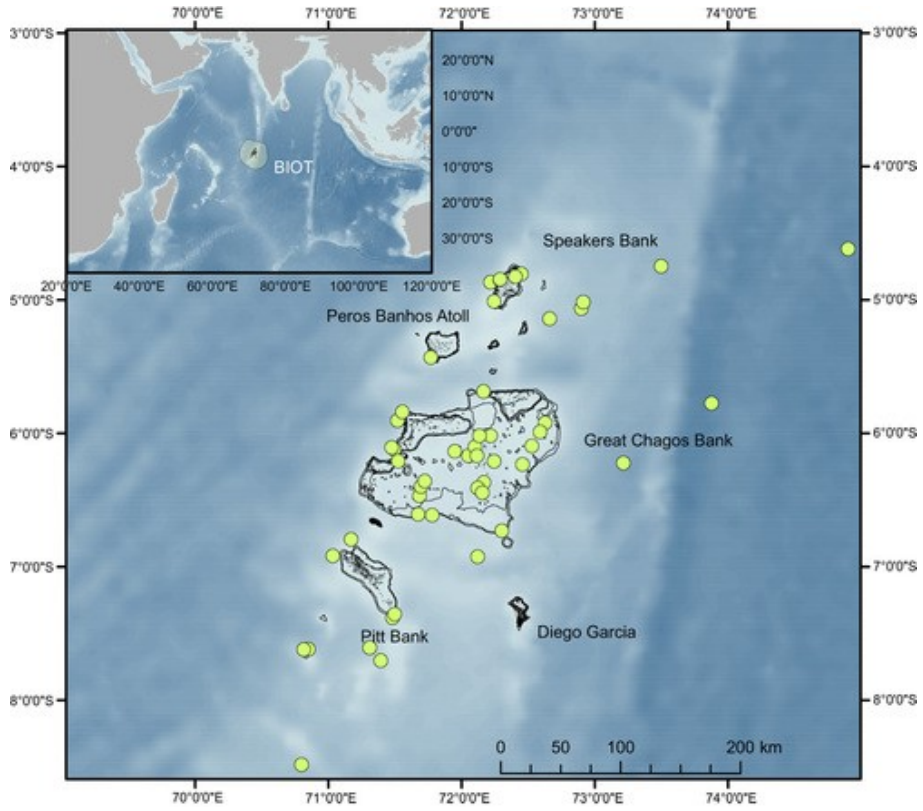


FIGURE 5.1: Map of the British Indian Ocean Territory (BIOT) and the Chagos Archipelago. The dots indicate the locations where illegal fishing was intercepted by a BIOT patrolling vessel. Image extracted from Ferretti et al. (2018).

5.2 Random encounter modelling

We seek to overcome the challenges introduced in Section 5.1 by extending REM to allow for baited cameras, which we henceforth refer to as baited random encounter modelling (bREM). REM was developed at the same time as spatial capture-recapture and there have recently been several new developments, namely generalised REM (gREM; Lucas et al., 2015), the addition of staying time (the so-called REST model; Nakashima et al., 2018) and research into the sources of variance in REM (Jourdain et al., 2020). The core assumption of REM is that cameras are placed independently of the movement of the animals; the animals do not need to move independently with respect to

each other. This is the main challenge when using bait to lure animals to the camera: not only does REM have the same challenge as SCR—namely that the estimation of the sampled area is more complex—-but also that the animals no longer move independently with respect to the cameras.

5.2.1 Summary of REM

In this section we summarise REM as presented by Rowcliffe et al. (2008). REM was initially developed for data from camera traps but is increasingly being used for video. Camera trapping uses fixed cameras that are triggered by infra-red sensors. The number of photographs taken per unit of time contains information about the density of animal population. Using this as an index of abundance has been controversial, primarily because it does not estimate the probability of detection (Rowcliffe et al., 2008) (however, recent work by Howe et al. (2017) showed that this is possible with camera trap data). Rowcliffe et al. (2008) modelled the underlying observation process, thereby developing a way to derive density estimates from camera trapping rates.

This observation process was based on mechanistic models to describe the interaction between gas molecules, i.e., the ideal gas theory. This model assumes a circular zone around objects at which contact occurs. A circular particle moving through two-dimensional space covers an area that is the product of its width, $2r$, and the total distance moved in a time frame, vt, which gives

$$a = 2rvt \tag{5.2}$$

where a denotes covered area, v denotes velocity, t denotes time, and r the radius. The total area covered by n particles is simply Equation (5.2) multiplied by n, which is equal to the product of density D and total area A. Therefore, the total covered area a_{t} by n particles of width $2r$ is given by

$$a_{\mathrm{t}} = 2rvtDA. \tag{5.3}$$

If we place a single sensor in area A with negligible radius, then the expected total number of encounters between the sensor and the particles is equal to the total covered

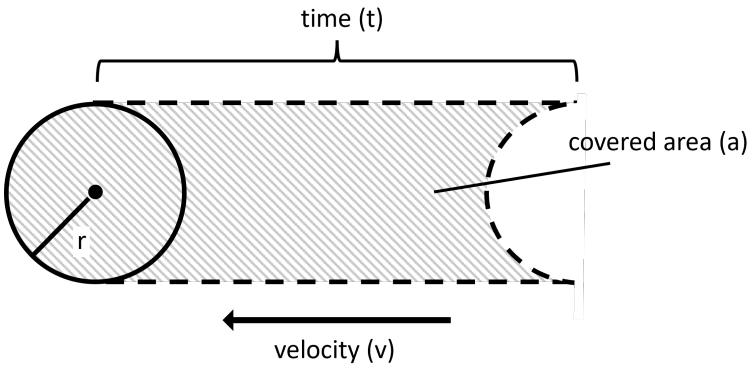


FIGURE 5.2: Area covered by particle moving through two-dimensional space with detection range r.

area divided by the area A, such that

$$\mathbb{E}[y] = \frac{a_{\mathrm{t}}}{A} = 2rvtD, \tag{5.4}$$

where an encounter y is defined as the moment a the edge of the particle interacts with the sensor. Equation (5.4) still holds if we treat the particles as points with negligible radius and the sensor as having a circular area or *detection zone* with radius r. If these particles now represent animals, then we can rewrite (5.4) to get an estimate of animal density:

$$\widehat{D} = \frac{y}{2rvt}, \tag{5.5}$$

where y is the number of encounters, where an encounter is now defined as the moment an animal enters the detection zone of the sensor. For cameras the detection area is often not circular, but rather a sector with radius r an angle ψ. The width of an animal's covered path, p, is therefore no longer $2r$, but depends on the angle of approach relative to the direction of the camera. We describe the six limiting cases for π possible approach angles, north to south (counterclockwise), in Figure 5.3. As the angle of approach is generally not known, we integrate over all possible angles $\gamma \in [0, 2\pi]$ to get the expected profile width. This involves integrating over $[0, \pi]$, dividing it by π to get the average over a half circle, and then multiplying by two to account for the full circle. Instead of the average profile width being equal to $2r$, it

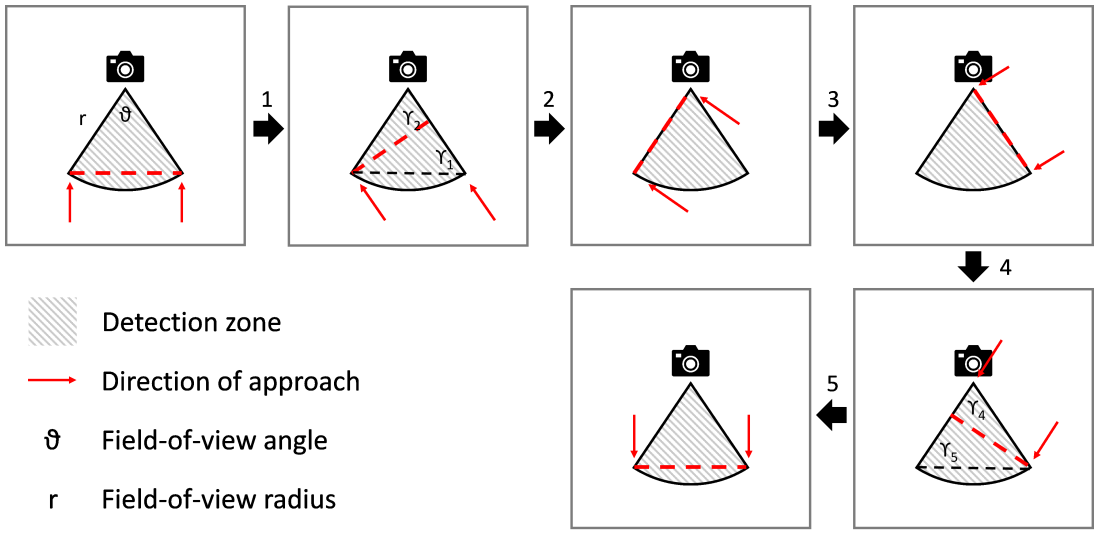


FIGURE 5.3: Six limiting cases for π approach angles (north to south, counterclockwise; red arrows). The numbers 1–5 refer to the 5 transitions between the cases as the approach angle changes. The width of the profile (p; dashed red) changes as the approach angle γ changes, which is the angle opposite the profile when drawing a right-angled triangle. The subscripts in γ refer to which transition phase the angle is related to. The profile width is $2r\sin(\theta/2)\sin(\gamma)$ for $(\pi-\theta)/2 \leq \gamma \leq \pi/2$ (transition 1 and 5); the profile width is $r\sin(\gamma)$ for $\theta \leq \gamma \leq \phi/2$ (transitions 2 and 4); and the profile width is r for approach angle θ (transition 3). This figure is redrawn from Fig. 1 from Rowcliffe et al. (2008).

becomes

$$\begin{aligned}
\bar{p} &= \frac{2}{2\pi}\left(2\int_{\frac{(\pi-\theta)}{2}}^{\frac{2}{\pi}} 2r\sin\left(\frac{\pi}{2}\right)\sin(\gamma)d\gamma + 2\int_{\theta}^{\frac{\pi}{2}} r\sin(\gamma)d\gamma + r\theta\right) \\
&= \frac{r}{\pi}\left(4\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\pi-\theta}{2}\right) + 2\cos(\theta) + \theta\right) \\
&= \frac{r}{\pi}\left(4\times\frac{1}{2}\left(\sin\left(\frac{\theta+\pi-\theta}{2}\right) + \sin\left(\frac{\theta-\pi+\theta}{2}\right)\right) + 2\cos(\theta) + \theta\right) \\
&= \frac{r}{\pi}\left(2\left(\sin\left(\frac{\pi}{2}\right) - \sin\left(\frac{\pi}{2}-\theta\right)\right) + 2\cos(\theta) + \theta\right) \\
&= \frac{r}{\pi}\left(2\left(1-\cos(\theta)\right) + 2\cos(\theta) + \theta\right) \\
&= \frac{r(2+\theta)}{\pi}
\end{aligned} \tag{5.6}$$

Incorporating the above in Equation (5.4) gives an updated expected number of contacts

$$\mathbb{E}[y] = \frac{2+\theta}{\pi} rvtD. \tag{5.7}$$

Rearranging this gives an estimator for density as a function of trapping rate y/t, animal movement v, and the camera dimensions θ and r:

$$\widehat{D} = \frac{y}{t}\frac{\pi}{vr(2+\theta)} \tag{5.8}$$

When animals move in groups and y denotes the number of encounters of groups, then Equation (5.8) must be multiplied by an unbiased estimate of group size g to get an estimate of animal density.

Variance is estimated using a non-parametric bootstrap, whereby the camera locations are re-sampled with replacement and the model is refitted to the bootstrapped data (Rowcliffe et al., 2008). The variance of $\widehat{D}$ is then estimated by the empirical variance of a large number of these bootstrapped density estimates. Variance of the other, independently estimated parameters (which are v, g, r, and θ), can be incorporated using the delta method (Seber, 1982).

5.2.2 Extensions and critiques

Some critiques of the method came from Foster and Harmsen (2012), who claimed that the assumption of random movement was certainly violated, and also that random placement of cameras was not feasible. Rowcliffe et al. (2013) responded to this by stating that REM is insensitive to non-random and non-independent movement of animals, as long as they do move randomly with respect to the cameras. Moreover, they stated that cameras can be placed at non-random locations, as long as the parts of the landscape that either attract or repel the animals are sampled proportionally to their coverage of the total sample area. We note that we believe that this is correct regarding the unbiasedness of the point estimator; however, when sampling is proportional to coverage of preferential environments in the study area, then variance estimated from a non-parametric bootstrap is no longer unbiased (Borchers et al., 2002, Chapter 4).

A challenge with REM is the inflation of estimated density when animals' movement speed is underestimated (Rowcliffe et al., 2012). Rowcliffe et al. (2016) propose a possible solution by estimating animal movement from camera footage, however, this requires video and could be challenging to adapt to conventional camera trapping. Two additional challenges of REM as presented by Rowcliffe et al. (2008) were that

i) it allowed only for sensor widths of up to $\pi/2$ (thereby excluding the use of the popular wide-angle cameras), and ii) the method was developed for camera-trap data and assumed that signals were omnidirectional, which excluded applying the method to acoustic data when acoustic signals were directional (Lucas et al., 2015). These challenges were resolved with a generalised version of REM (gREM) that was presented by Lucas et al. (2015). In this study, the authors developed several models and limiting scenarios by extending the range of sensor width θ from $[0, \pi/2]$ to $[0, 2\pi]$ and allowing for the signal to be directional by modelling a signal width $\alpha \in [0, 2\pi]$. In total, 8 models were developed to derive the expected profile width $\bar{p}$, depending on the values of α and θ. Similar to Rowcliffe et al. (2008), an encounter was defined as the moment the signal of an animal entered the detection zone. A simulation study was carried out to test the performance of gREM. Here, the accuracy was tested comparing the true and estimated densities from the correct model for various combinations of α, θ, total number of captures, and for range of animal movement patterns. All the gREM submodels showed low bias, but precision varied strongly with sensor width and signal width. The authors also tested for model sensitivity to error in the parameter estimates, which they found to be reasonable as percentage error in the density estimate was never more than 5% larger than the error in the parameter estimate (Appendix S6; Lucas et al., 2015). For future field studies, the authors concluded that i) sensor width should be maximised, ii) the number of captures has a strong effect on precision, and thus researchers should focus on collecting a sufficient amount of data, and iii) capture-recapture studies should be included alongside gREM to test performance under field conditions.

Finally, REM is a deterministic model and relies on parameters estimated from other studies. Jourdain et al. (2020) derived the REM estimator as a maximum likelihood estimator by assuming that the number of encounters between animals and sensors follows a Poisson distribution. This allowed for the incorporation of covariates, such as habitat type. They further proposed an extension of this maximum likelihood REM estimator called the integrated random encounter model (iREM). This extension jointly models the encounters of animals and sensors and the movement speed of the animals, instead of assuming a fixed movement speed. They show that REM and iREM can both be used for density in heterogeneous habitats, but that iREM is preferred whenever one wants to incorporate camera random effects or account for variability in animal

speed. However, it is noted that the fitting of iREM can be substantially more complex and computationally intensive.

5.2.3 REM for sharks on Chagos

The main challenge for using REM to estimate shark abundance for our case study is that the cameras were baited, which violates the standard assumption of independence between the movement of the sharks and the cameras. In addition, the effective sampled area is primarily determined by the area and intensity of the bait plume, and the response behaviour of sharks to the bait, and less so by the FOV of the camera.

We believe that the sampled area can be estimated using an extension of the method proposed by Dunlop et al. (2015). In this study, the bait plume was modelled using the current speed, diffusional velocity, and elapsed time. However, the conditions affecting the bait plume spread will need to be estimated or recorded separately for every BRUV and *in situ*, as these conditions can vary locally (Heagney et al., 2007). For example, when estimating the density of stone crabs (*Lithodidae*) using BRUV, the camera systems carried a current meter to estimate individual bait plume spread (Collins et al., 2002). A primary challenge is how to model the process of the shark responding to the bait. Sharks respond to bait differently depending on species, and the type and amount of bait used (Kilfoil et al., 2021). In a previous study researchers assumed a response probability of 1 and that the animal would swim to the bait in a straight line upon detection (Dunlop et al., 2015), but we are unaware of any studies that support the validity of such an assumption. Moreover, we should consider that the presence of a shark at or near the bait might affect the behaviour of other sharks and hence the arrival of these sharks into the FOV. Another challenge is to account for disruption in movement caused by sensing the bait. REM assumes movement to be random and unaffected by the detector; however, here it is likely that animals will spend some time near the camera exploring the bait. This disruption in movement can generally be categorised in two ways:

- the disruption in the movement is temporary, which means that the shark is attracted, swims to the camera, remains close to the camera for some time, and then continues the initial movement pattern, albeit from a different location, or

- the shark interacts with the bait and camera in a similar way, but now changes its swimming behaviour after encountering the bait.

The second category of disruption would be more challenging to model as one would have to know how encountering the bait affects the general movement of the shark, whereas the former could potentially be modelled more easily.

A further question that requires attention is the metric that will be used to quantify an encounter. In Rowcliffe et al. (2008), the moment an animal enters the FOV of a sensor was used to calculate encounter rate. If an animal leaves the area and re-enters, then this is counted as a new encounter. This does not bias the estimates, as long as the movement of the animals in independent of the sensor. As this is no longer so when using baited cameras, a new definition of an encounter could be required. Others have suggested using association, which is defined as the number of ongoing encounters within an area in a give time frame (Hutchinson and Waser, 2007; Campos-Candela et al., 2018). The absolute animal density can then be estimated by the averaged counts across a number of independent frames divided by the sampled area (Hutchinson and Waser, 2007). This estimator was later extended for data from animals displaying home range behaviour (Campos-Candela et al., 2018). Potentially, it is also possible to incorporate time-to-first-arrival (TFA) as additional information in a REM framework, similarly to the way Nakashima et al. (2018) incorporated staying time to develop the random encounter and staying time model. There could even be information in the distribution of all arrival times. Of course, it will often be impossible to say for sure whether a new appearance of a shark is actually a new shark, or just an old one that re-entered the FOV. Even though we assumed the sharks are not uniquely identifiable, it might occasionally be possible to tell some apart, for example when one shark is clearly of a different size than the other or a (nearly) unique marker is found, such as a scar. With this one would be able to create a lower bound number of unique individuals encountered on a BRUV system. Moreover, there could exist the potential to model these arrival patterns and obtain some signal about density. This information could then be incorporated through an additional element in the likelihood if the likelihood-based method is used. Some examples of adaptations to partially identifiable populations are Chandler and Royle (2013) and McClintock et al. (2009). We do not explore using only TFA to estimate abundance, as this would

reduce every BRUV deployment to a single data point, greatly reducing the available information.

5.3 Baited random encounter modelling: exploratory studies

We performed two studies to test and explore the potential of resolving some of the questions and challenges presented in the previous section. We explored whether REM could be extended to baited cameras (bREM) under some strong assumptions. One of the metrics that is important here is the Euclidean distance travelled by the shark after it encounters the bait plume. We first derive an analytical solution to the expected distance the shark must swim when the point of contact with the plume is unknown. Second, we explored through deterministic numerical simulation the ways in which bait plumes spread under various current scenarios, with constant and variable plume intensity.

5.3.1 Distance from camera to plume edge

It is not generally known how a shark approaches bait upon encountering a bait plume. Even though sharks might not approach bait in a straight-line after encountering the scent, modelling it this way can still work if the general direction of the shark is towards the bait. The distance from the edge of the bait plume to the camera/bait thus remains an important quantity for determining the time that the shark takes to get to the bait and hence be detected and time-stamped. In this section, we find an analytical solution to this expected distance, $\mathbb{E}(d)$.

We assumed that the water was moving in one direction with a constant velocity, and that the bait plume spread at angle θ. We approximated the shape of the bait plume by a circular sector with angle θ and radius r, similarly to Dunlop et al. (2015). The approach angle γ of the shark was set at 0 when it approaches directly against the current. As we did not know the angle at which the shark approaches the bait plume, we considered all possible angles from 0 to π and averaged over them, similarly to Rowcliffe et al. (2008). We only needed to consider half of the full circle as π to 2π

is identical. The radius of the plume is denoted by r. This resulted in three limiting cases for distance d: $d = r$, if $0 \leq \gamma \leq \frac{\theta}{2}$; $d = r\left(1 - \frac{1}{2}\frac{\sin(\gamma-\theta/2)}{\sin(\gamma+\theta/2)}\right)$, if $\frac{\theta}{2} \leq \gamma \leq \frac{\pi}{2}$; and $d = \frac{1}{2}r$, if $\frac{\pi}{2} \leq \gamma < \theta$. The expected distance was therefore the integral of d for $\gamma \in [0, \pi]$, where the equation for d depended on the value of γ (see above), divided by π, such that

$$\begin{aligned}
\mathbb{E}(d) =& \frac{1}{\pi}\left(\int_0^{\frac{\theta}{2}} r d\gamma + \int_{\frac{\theta}{2}}^{\frac{\pi}{2}} r\left(1 - \frac{1}{2}\frac{\sin(\gamma-\theta/2)}{\sin(\gamma+\theta/2)}\right) d\gamma + \int_{\frac{\pi}{2}}^{\pi} \frac{1}{2} r d\gamma\right) \\
=& \frac{r}{\pi}\left(\frac{\theta}{2} + \frac{\pi}{2} - \frac{\theta}{2} - \frac{1}{2}\int_{\frac{\theta}{2}}^{\frac{\pi}{2}} \frac{\sin(\gamma-\theta/2)}{\sin(\gamma+\theta/2)} d\gamma + \frac{\pi}{4}\right) \\
=& \frac{3r}{4} - \frac{r}{2\pi}\left(\frac{(\pi-\theta)\cos(\theta)}{2} - \log\left(\frac{\cos(\theta/2)}{\sin(\theta)}\right)\sin\left(\frac{\theta}{4}\right)\right)
\end{aligned} \tag{5.9}$$

Figure 5.4 shows the expected distance plotted against $\theta \in (0, \pi)$ for a range of values for r. What was promising is that there seemed to be little variation in the expected distance as θ increases.

This is useful for a scenario with the following assumptions:

1. odour intensity is constant throughout the bait plume;
2. the camera detection angle is equal or less to the bait plume angle;
3. sharks swim in a straight line towards the bait upon encountering the bait plume; and
4. the circle sector representation of the bait plume sufficiently holds.

The second assumption is mostly important when the FOV is substantial relative to the area covered by the bait plume; when it is negligible, this assumption can be relaxed. Even if the shark does not approach in a straight line, one could still work with the straight line distance but use a corrected swimming speed. These are strong assumptions, still, and should be tested before our derivations can be used, especially assumption 1 and 4. For this reason, we further explore the way the plume spreads when currents are less predictable than we assumed.

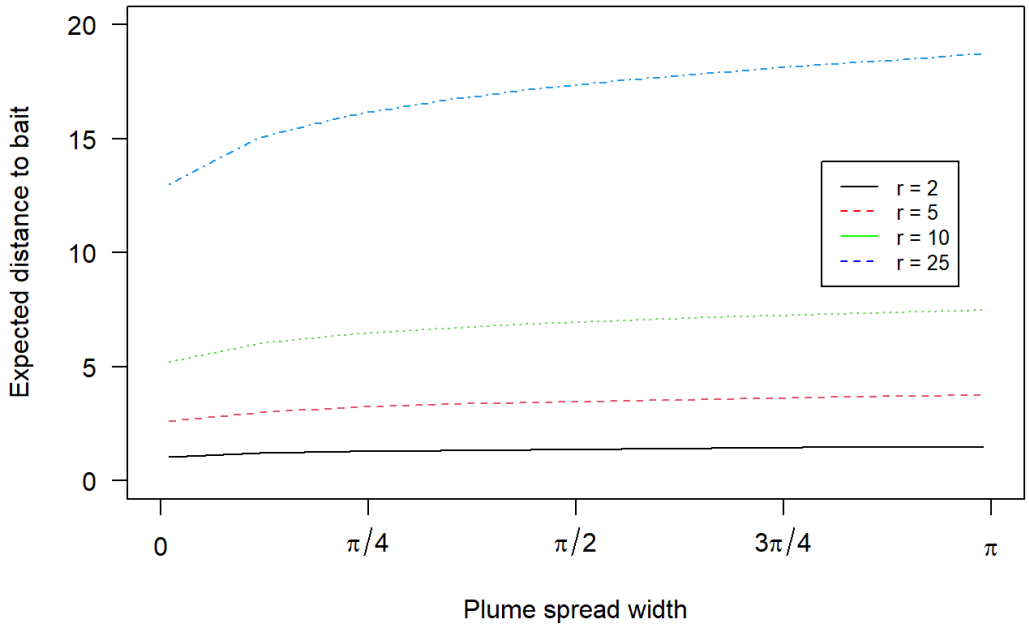


FIGURE 5.4: Expected distance from the edge of the bait plume to the bait/camera, as a function of plume spread angle, θ. We considered four different values for plume length, r.

5.3.2 Bait plume spread

In our second study, we considered bait plume spread under various current conditions, and the ways in which they affect the total (effective) sampled area. Many different plume spread models exist, but many are complex and require much calibration depending on specific bait type and water conditions (Westerberg and Westerberg, 2011). Therefore, we started with a simplified model proposed by Dunlop et al. (2015). The spread of the bait plume was modelled using the following three equations:

$$L_{\text{pl}} = V_w T, \tag{5.10}$$

where L_{pl} denotes the bait plume length, V_w the mean current speed, and T the time in seconds;

$$\theta_{\text{pl}} = 2 \tan^{-1} \left(\frac{B_y}{V_w} \right), \tag{5.11}$$

where θ_{pl} denotes the plume spread angle in radians, and B_y the diffusional velocity; and

$$B_a = \left(\frac{\theta_{\text{pl}}}{2}\right) L_{\text{pl}}^2, \tag{5.12}$$

where B_a denotes the area covered. As currents can vary substantially locally (Heagney et al., 2007) we now consider the following questions: 1) is B_a affected by changes in the direction of the current speed, and 2) what role does the distribution of current speeds play (i.e., can we rely solely on an average current speed or is it important to account for variation in current speeds)?

5.3.3 Methods

The goal was to predict the total sampled area. Here we made the distinction between *absolute* and *effective* area. The absolute area is the total area that is covered by the bait plume, irrespective of the intensity of the plume. On the other hand, the effective area is the total area corrected for the intensity of the plume. When a bait plume spreads, the plume dilutes and thus the intensity decreases. We included this because we believed the intensity could affect the probability that a shark responds the plume when it interacts with it. For our simulation we discretised the two-dimensional space into equally-sized squares, where every square had an area of size 1. We assumed that the current could only move in four directions: west, north, east, and south. The bait was placed at the point (0,0). We also discretised time: at every turn, each cell that was already covered by the bait plume 'baited' the next cell down-current and the cells left and right of it, relative to the current—the original cell was no longer covered by the plume, as long as it was not baited again by one of the neighbouring cells. The one cell that was always covered by the plume was the bait itself, the 'source'. This approximated a spread model with plume angle $\pi/2$ radians.

We considered five scenarios of bait plume spread, where we varied the current direction and the number of steps between changes in direction. The first scenario consisted of a stable westward current; the second scenario consisted of a predominantly westward current, that occasionally moved north- or southwards and rarely eastwards; the third scenario consisted of unstable current that was randomly moving either westwards or eastwards; the fourth and fifth scenario consisted of currents that were stable for some time going one direction and then flip to the opposite direction—this would happen

once or several times, depending on the scenario. The five scenarios are described in more detail in Table 5.1. Every simulation consisted of 100 time steps. At every time step, the cell directly down-current from a cell covered by the plume were baited with 50% of its plume intensity, and 25% of its intensity baited the cells left and right of the cell directly down-current. To illustrate: a cell with a bait plume intensity of 0.8 would bait the cells directly down-current with intensity $0.5 \times 0.8 = 0.4$ and the cells left and right of it with intensity $0.25 \times 0.8 = 0.2$. The source of the bait plume, the bait itself, always had plume intensity 1. The plume intensity at all cells was capped at 1 to reflect the concept that at some point an increase in plume intensity no longer affects the responsive behaviour of the shark. The total plume area is denoted A_T and is total number of cells covered by the plume; the effective plume area is denoted A_E and equals the sum of the plume intensities over the whole area.

5.3.4 Results

The results are presented in Table 5.1 and visualised in Figure 5.5. Whenever currents were stable, i.e., rarely changed direction, then the effective plume area was very similar (A_E is 101 in scenario 1 and 96 in scenario 4). As current changes became more frequent, and thus the same cells got repeatedly infected, the effective plume area dropped (A_E is 82 in scenario 2 and 87 in scenario 5). When the current constantly moved back and forth, the effective plume area was smaller still (A_E is 56 in scenario 3). Overall, the largest effective plume area (101; scenario 1) was almost twice as large as the smallest effective plume area (56; scenario 3). There was also a gap between total plume area and effective plume area, with the total plume area always being much larger. Moreover, the total plume area was more sensitive to changes in the current scenario (A_T is 10201 in scenario 1 and 1741 in scenario 3).

5.4 Concluding remarks

In this chapter we explored the potential of deriving absolute abundance estimates from BRUV data using an extension of the random encounter model. The main challenges that were evaluated were i) that shark behaviour would no longer be independent of the

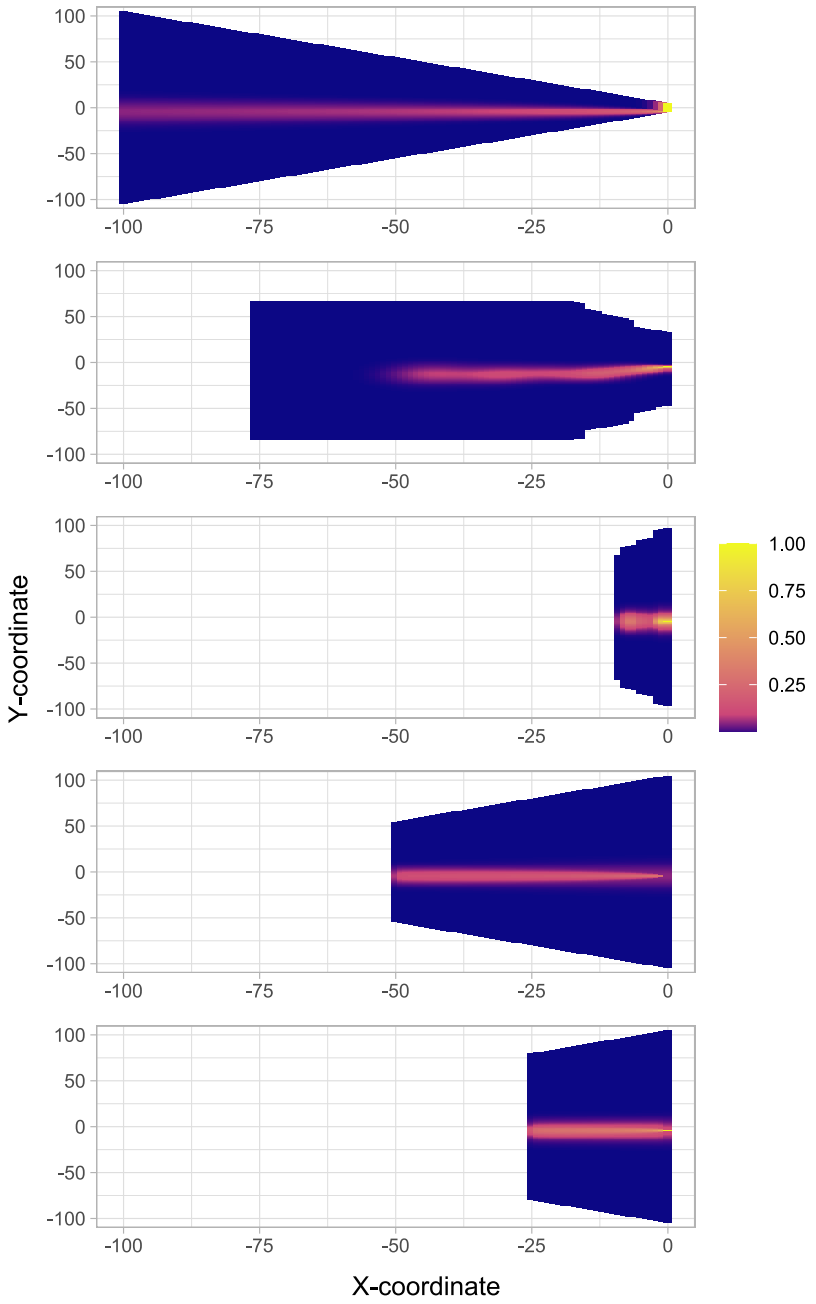


FIGURE 5.5: Plots of bait plumes for five scenarios. In every scenario, the current was simulated for 100 time steps. At every time step, the direction of the current either followed a fixed pattern or was randomly drawn from four options: west (W), east (E), north (N), and south (S). The five scenarios are (top-to-bottom): 1) constant west current; 2) mostly west current (direction probabilities W: 55%, E: 9%, N: 18%, S: 18%); 3) rocking water (direction probabilities are W: 5%, E: 50%, N: 10%, S: 0%); 4) rarely changing current (50 times W, then 50 times E); and 5) regularly changing current (25 times W, the 25 times E, and then repeat).

Scenario	Description	A_T	A_E
1	The simulation was run for 100 time steps with current going westwards at every step.	10201	101
2	The simulation was run for 100 time steps where the direction of the current at every step was sampled with probabilities north (18%), west (55%), south (18%), and east (9%), i.e., predominantly westwards.	10008	82
3	The simulation was run for 100 time steps where the direction of the current at every step was sampled with probabilities north (0%), west (50%), south (0%), and east (50%), trying to mimic a body of water that moves back and forth without a true dominant current.	1741	56
4	The simulation was run for 100 time steps where the direction of the current was exclusively westwards for the first 50 steps, followed by 50 steps eastwards, trying to imitate a current that flips direction infrequently, or potentially a rising and falling tide.	7701	96
5	The simulation was run for 100 time steps where the direction of the current was exclusively westwards for the first 25 steps, followed by 25 steps eastwards, 25 steps westwards, and finally 25 steps eastwards, trying to imitate a current that flips more frequently.	4576	87

TABLE 5.1: Overview of the plume spread simulation scenarios and simulated bait plume areas. A_T denotes the simulated total plume area, i.e., the total number of cells that were covered by the bait plume, and A_E denotes the simulated effective plume area, i.e., the sum of all bait plume intensities. The areas correspond to the plots presented in Figure 5.5.

presence of the camera systems and ii) that the effective sampled area would depend on the dynamics of the water column.

We observed large reductions in effective plume area when currents changed, whereby a constantly changing back-and-forth current resulted in the lowest area. This was a result of the plume intensity being capped at 1, i.e., which was the intensity at the source; would the plume intensity have not been capped, then the effective plume area would have been the same for every scenario. This cap was introduced to reflect the fact that at some point more intensity does not necessarily result in a stronger response from a shark. However, this was a simplifying assumption, and other plume intensity structures could be tried. Moreover, we only tried a simplistic, deterministic model

of bait plume spread, and the benefits of using more sophisticated (e.g., stochastic) models, such as Gaussian plume models, should be tested (Westerberg and Westerberg, 2011). Finally, the substantial gap between total plume area and effective plume area highlights that one should consider adding bait concentration as a part of the simulation.

Previous studies concluded that methods using baited cameras work in theory, but that observations from BRUV data remain sensitive to effects that will be hard to model or accurately estimate, such as local current conditions or shark behaviour (Harvey et al., 2007; Kilfoil et al., 2021). The explorations undertaken here were based on strong assumptions, too, and these require further testing. The most important assumption one would be the behavioural effect of encountering the bait on the movement of the shark. One could get an approximation of when the shark encountered the plume by assuming a straight-line approach with an altered swimming speed, but we are not aware of information about the staying time and changes to movement once the shark leaves the bait. Furthermore, the fact that a shark could be circling the bait for some point before leaving, during which time it could enter the FOV of the camera many times, is something that will require attention. A potential way to handle this is to omit data from the camera from the start of the encounter as long as there is good reason to be believe that a shark is still present at the bait, and only continue using that data once the shark has left. This would mean that potential new encounters are not counted, but the issue of a repelling effect of a shark at the bait on other sharks is potentially reduced and that the data is of higher quality. However, this would require some level of individual identification, which can be challenging for some shark species. Moreover, if individual identification is possible, then one should probably consider using a capture-recapture method.

Nonetheless, there still appears to be potential in using BRUV data for estimating absolute abundance of sharks through a random encounter framework. There are several field considerations that will enhance this potential. First, we believe that it is important to record as much information *in situ* about the environmental conditions, such as current direction and speed. Second, one should consider whether the placement of the BRUV systems allows for well-estimated bait plume spread. For example, mid-water placement of BRUV systems might allow for more predictable and consistent currents directions and speed but would require modelling in three-dimensional

space as opposed to two-dimensional. On the other hand, bottom placed BRUV systems could allow for a two-dimensional model, but current speeds and directions at sea floors could be less predictable, especially if the environment is more turbulent (e.g., due to coral reef or other structures).

Last, we note that the use of baited traps or sensors is not limited to the marine context. Adding a scent to traps to increase the probability of observing an animal has been used for small mammals (Parmenter et al., 2003), large predatory mammals (Braczkowski et al., 2016), for insects such as beetles and moths (Dahlsten et al., 2004; McNeil, 1991), and for birds (Reid et al., 1999; Klavitter et al., 2003). This means that the insights from this study could be relevant also to researchers seeking to apply REM with baited traps in a terrestrial context.

5.4.1 Future research

We believe that one of the most important future steps is to further explore the behavioural response of sharks to varying intensities and types of bait. Novel techniques such as drones could allow for aerial observations of shark behaviour when sharks are clearly visible from such footage, for example when present in shallow, coastal waters. The bait could be dyed (as long as this does not affect the behaviour of the sharks), so that the spread of the plume is easily observable from a drone. A potential challenge is that it remains unknown to what degree such results could be extrapolated to other species and environmental contexts.

Moreover, a clear next step would be to develop an actual estimator for density. Once such an estimator is derived, one should perform further simulation studies to explore robustness and bias under different scenarios. This is because such an estimator will likely rely on strong and hard-to-verify assumptions about the bait plume and the behaviour of the shark. We believe that agent-based simulations studies would be vital to understanding where the sensitivities to applying REM in these scenarios lie, similar to Kilfoil et al. (2021) (see also Chapter 4), as these allow for the incorporation of individual movement of animals. For example, it could be that effects on the swimming behaviour of sharks after interaction with the bait are not a problem given that the expected behaviour remains the same. Finally, another area of future work would be to test out continuous time and space differential equations of bait intensity (Garvey

and Fowler, 2023), as these could be used to infer on how the bait intensity spreads through time, how long it takes to reach a steady state, and what that steady state is.

Chapter 6

Discussion

6.1 Summary

As technology advances, new sampling techniques become available to wildlife researchers. We increasingly appreciate the importance of wildlife conservation, and thus continue to extend existing abundance estimation methods and develop new ones. In this thesis we studied various methods that use novel sampling techniques to estimate animal abundance. All research was motivated by case studies involving marine megafauna from the Arctic to the tropics.

In Chapter 2 and 3 we developed an extension of acoustic spatial capture-recapture (ASCR) motivated by recordings of bowhead whale calls. This extension allowed us to remove all calls that were detected on only one sensor, which were assumed to contain a high proportion of false positives. This extension is readily applicable to any ASCR study that faces a similar issue. It can also be applied to regular spatial capture-recapture (SCR), for scenarios in which there is reason to believe that animals only recorded once are not part of the population for which inference is desired (e.g., when a transient population mixes with the population of interest; Conn et al., 2011), or when repeat captures are mistakenly thought to be captures of new individuals (e.g., for snow leopard (*Panthera uncia*) detections from camera traps; Kodi et al., 2024). We showed that the method is nearly unbiased with a received level step detection function. Using the step detection function meant that we had to discard all calls below a certain truncation level, which was set so that noise never surpassed

the received levels; calls with a received level above this truncation level were assumed to have the same detection probability. As this can lead to discarding a sizeable proportion of the data, which can be undesirable, we also presented mathematical derivations of a model with a detection function based on the signal-to-noise ratio. This method requires additional information on noise, which is often readily obtained from acoustic sensors.

In Chapter 4, we assessed the extent to which incorrect ageing affects close-kin mark-recapture (CKMR) estimates. This research was motivated by a case study of genetic samples from grey reef sharks. The exact age is often unknown and thus often estimated by inferring the age from the length of the animal. In previous studies, researchers found that growth curves can vary between different populations of the same species. We undertook a simulation study using a range of length measurement errors and where the length-age relationships were misspecified. These simulations involved two fictional species to see whether the complexity of the population dynamics played any key role. We found that the assumed length measurement error had negligible effect on the estimated parameters, but that a bias in the growth curve induced strong bias in the estimates.

In Chapter 5, we explored the potential to use an adapted form of random encounter modelling (REM) to estimate abundance from baited sensors. Using bait violated a key assumption of REM, which is that animals move independently of the sensors. Moreover, the sampled area was no longer equal to the field-of-view of the camera, but rather the area of the bait plume. Conventional REM estimates density from the encounter rate, which is the number of encounters of animals with the edge of the area sampled by a sensor over some defined time period. We derived an analytical solution for the expected distance from the camera to the edge of the bait plume, which can be used to estimate the time the animal first encountered the bait plume based on when the animal reached the camera. Further, we ran a simulation study to explore how the area covered by the bait plume changed under various current conditions. We showed that this can vary substantially and that it is important to properly understand the current conditions when modelling the bait plume. Also, we believe that one should account for the intensity of the plume within the covered area, as the intensity of the plume can have an effect on the behavioural response of the animal.

6.2 Future directions

One of the main conclusions we drew from working with these novel abundance estimation techniques is that they contain inherent trade-offs. Established and rigorously validated approaches such as distance sampling and spatial capture-recapture work well when the context is suited for their application; however, not all scenarios are. The methodologies elaborated upon in this thesis were devised for such scenarios through the incorporation of novel sources of information. The signal-to-noise ratio detection function gives us the opportunity to include additional information relative to a conventional ASCR analysis but is more complex than the step detection formulation from Chapter 2. CKMR uses the richness of genetic information, but to do so this method requires expertise in 1) the biology of the animals to define the population dynamics model correctly; 2) the genetics to accurately determine kinship and incorporate any uncertainty associated with this relationship; and 3) statistics to implement the method in a computationally feasible way and correctly interpret the results. Combining these three fields allowed for improved abundance estimation of populations such as the northern Atlantic bluefin tuna (*Thunnus thynnus*) stock (Bravington et al., 2016a); however, it means that CKMR studies require extensive collaboration. Finally, using data from baited sensors for unmarked populations means that we either must make strong and hard-to-verify assumptions about the behaviour of the animals, or that we need to perform separate studies on this behaviour and include them in our model. For all these methods, we believe that a priority should be to understand the sensitivity of the models to violations of the assumptions and error in the various sources of information; in the absence of such knowledge, it will be hard to judge whether an application of the method is suited for studying a particular species.

A potential future step for the ASCR method is the actual implementation in code of the signal-to-noise extension. As part of that process, it would be essential to implement a simulation study to explore the overall performance of this method, i.e., to discover any bias and optimise it until the computation time is acceptable, but also to make sure that the implementation is correct. Another avenue to explore would be finding a way to incorporate the directionality of the calls, i.e., calls that are louder in front of the whale than behind it. These extensions already exist in the current ASCR literature, and we believe that incorporating such extensions for this case study could offer a significant improvement in the estimated call density surface.

To our knowledge, we performed the first simulation-based study of the effects of incorrect ageing through misspecified growth curves on CKMR estimates. We found that even a slightly shifted (i.e., biased) growth curve induced severe bias in the estimates. Moreover, we noted a difference in these effects between the two fictional species that we studied. We believe that this difference induced by changes in population life history traits would be an excellent direction for future research. For example, one could explore how incorrect ageing affects species of various life spans and various ages of maturity. Furthermore, it would be valuable to improve our understanding of how these violations affect parameter estimates when we allow for growth in the population. We included some mathematical derivations of such a model in Appendix B; however, we did not fully succeed in this as the distribution of age given sampled lengths was inconsistent with a population that exhibited either growth or decline. Nonetheless, there were clear indications that a misspecified growth curve could not only affect abundance estimates on average, but also the direction of the trend (e.g., growth or decline). As the trend is one of the key metrics for wildlife conservation, understanding the relationship between estimated trend and incorrect ageing in a CKMR framework would be a natural step for future research. In addition to this, one could also further explore the effects of individual variation in the length-at-age relationship, potentially in combination with measurement error. Finally, we believe that a future direction for research with great potential is the joint fitting of a CKMR and capture-recapture model whenever recaptures occur.

We provided some groundwork for baited random encounter modelling; however, no actual estimator was developed. For this reason, we believe that this could be an excellent step for future research. To do this, we believe that the most feasible part to tackle is the sampled area. If the current conditions allow for proper modelling of the bait plume, one could estimate an effective sampled area by making some assumptions about the behavioural response of a shark. For example, one could assume that sharks respond with certainty whenever the plume intensity surpasses some threshold value, and not whenever the intensity is below the threshold. This is a simplification, and further behavioural studies are required to determine whether such a modelling assumption is acceptable, and perhaps a sensitivity study to variation in bait plume response. The more challenging part would be how to handle the disruption in the shark's movement pattern. One way to handle this is to 'pause' the data from the camera whilst the shark is still exploring the bait and only continue using it for new

encounters once the shark has moved on. If one then assumes that the shark resumes its behaviour from before the bait plume encounter, then this could work within the REM framework. Simple assumptions could be made regarding the attracting or repelling effect of one shark at the bait on other sharks when developing the initial estimator, which can then later be extended as more behavioural data become available.

As we continue to push the boundaries of wildlife research, there is an increasing need to integrate various scientific disciplines, including biology, ecology, physics, genetics, computer science, and statistics. Exploring novel approaches for studying wildlife brings forth new challenges in data analysis, with statistics playing a crucial role in ensuring precise inference from these data. Advancements in computing power have enabled the development and assessment of methods that were previously infeasible. We developed novel methodology to utilise the potential of acoustic data and overcome some of its challenges. Additionally, we validated a recently developed method for genetic data by evaluating its strengths and weaknesses through simulation. While standalone validation research remains essential, we advocate for a proactive approach. Researchers should not delay the adoption of new methods for estimating abundance. Instead, we propose incorporating simulation into the application of novel methods, preferably prior to allocating significant resources to the primary study.

Appendix A

Supplementary materials for Chapter 2

A.1 Data background, availability & cleaning

A.1.1 Data availability

To handle the vast number of calls detected by the DASARs, Thode et al. (2012) developed a method to detect and match calls using a neural network trained using manually processed data from 2008 and 2009. From these data a collection of detection histories could be derived, where for every call and every DASAR it contains a 1 if detected, and a 0 otherwise. Sometimes, calls would be wrongly identified and/or matched, as other discrete sounds produced by, among others, airguns, bearded seals (*Erignathus barbatus*) and walruses (*Odobensu rosmarus divergens*) can appear similar to bowhead whale calls (Thode et al., 2012). It was thought that single detection calls ('singletons') contain a high proportion of falsely identified calls (false positives), thus we removed these detections altogether.

In this study, we focus on the data from one specific day, 31 August 2010, from site 5, the most eastwards site, as this was of one the more intensely studied subsets of the data. Moreover, calls accumulated somewhat evenly across the day, resulting in a low expected rate of overlapping calls. For this day, the following observations were extracted:

- A matrix containing detection histories for all calls and all sensors, with 1 indicating detections and 0 indicating no detection.
- A matrix containing received levels for the positive detections, and NA values otherwise.
- A matrix containing noise measurements for all sensors and all calls. Noise at DASARs that detected a call was derived using a 6 second window, consisting of 3 seconds before and 3 seconds after the call, using the same frequency band as the call. Noise recordings for DASARs that did *not* detect a call were derived using the widest observed frequency window for the DASARs that were involved in the detection of the call, using a similar time window.
- A matrix containing noise measurements derived over 6-second windows, taken at a random time points and frequency bands within the study period. These were subsequently checked to not contain any calls, as that would not qualify as 'noise' - other discrete sounds, such as walrus calls or airguns, do qualify as noise, as they potentially interfere with the identification of bowhead whale calls.

The last two matrices containing noise information are only used for the detection function based on SNR (see Chapter 3).

Two variables in our likelihood are latent: the origin of the call and the source level of the call. To limit the run time, we only integrate over the smallest subset of the domain that includes at least 99.9% of the associated probability density/surface, at the lowest resolution that ensures good estimates. Thus, the remainder of our data includes:

- A matrix of at least 300 evenly spaced spatial points covering an area such that calls that originated at the bounds have a probability of being detected on at least two sensors of < 0.1. For every grid point, the ocean depth and distance to coast was also known; these were used as potential covariates in the density model.
- A vector of source levels, ranging from 100 to 220 dB, with increments of 3 dB, which were the largest increments that still resulted in good parameter estimates.

A.1.2 Data cleaning

ASCR theory requires perfect identification, and to meet this assumption as closely as possible, several cleaning procedures were required. The false positives among calls that involved two or more detections did not pose a problem for the original studies (Thode et al., 2012). Here, the main goal was to triangulate the bearings of detected calls to estimate the spatial origin of the call. Every involved DASAR would record a bearing, and a localisation procedure would add varying weights to these bearings until it resulted in a successful localisation, or not (Thode et al., 2012). For example, if the bearings of a call with two detections would never meet, e.g., because one of them belonged to a wrongly identified sound, a localisation would fail. If, however, a call that involved three detections had two bearings that matched and one that did not (the 'outlier'), it would give the latter a weight close to 0, and the others a weight closer to 1, thus resulting in a successful localisation. As we only wanted the 'valid' detections in our detection histories, we used these weights to remove wrongly identified and matched calls. This was done using the following procedure:

1. We start by only including call events that led to a successful localisation. This first step got rid of misidentified airgun signals, as those originated too far from to array to be localisable.

2. If the call event contained at least one bearing with a weight < 0.02, we removed the detection with the lowest weight.

3. The call was localised again (now excluding the previously removed bearing), and the newly assigned weights were evaluated.

4. If the call was successfully localised and still contained at least one weight < 0.02, we removed that detection and went back to step 2. Otherwise, the call was now considered valid.

The weight of 0.02 was found to most accurately remove wrongly identified and matched calls; through visual checks we found that lower values still allowed for some incorrect bearings to be included, and that increase the weight did not make a difference in how many incorrect bearings were removed. Finally, one call had a failed noise recording and was thus excluded from the analysis.

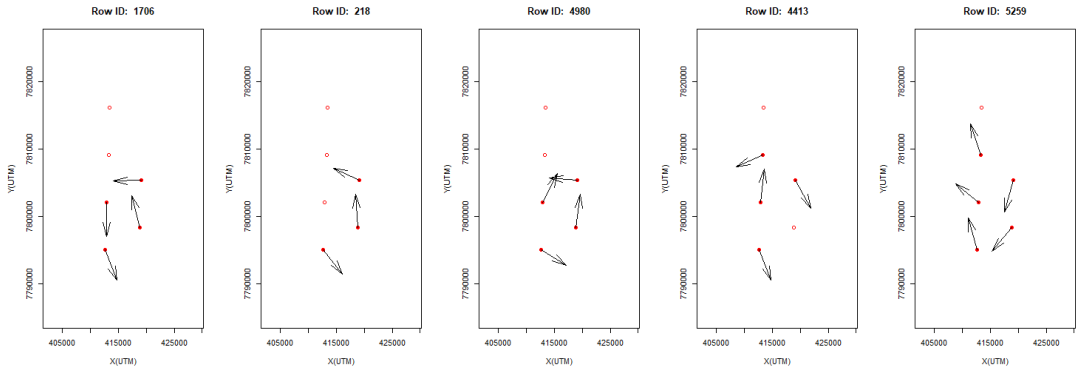


FIGURE A.1: Five call events with strongly misaligned bearings from the data from Site 5 on 31 August 2020. The red dots are the sensors in the array, and the black arrows indicate the bearings for sensors that were involved in the call event. Generally, bearings were fairly accurate, but these five plots show that occasionally bearings could be far off.

This left us with a total of 5793 'correct' calls detected on at least two sensors at site 5 on 31 August 2010. The next stage was to truncate the data to include only detections that exceed the highest sampled background noise level of 96dB. This left 443 calls after truncation that were used in the analysis reported in the main paper.

A.1.3 (In)accurate bearings

When visually exploring our data, we noticed that not all bearings were equally precise. Plotting a subset of the bearings suggested a split in bearings of mostly accurate and some inaccurate ones. It was there hypothesised to best allow for a two-part mixture model on bearing accuracy, as presented in the main paper. An example of inaccurate versus accurate bearings is presented in Figure A.1.

A.2 Spatial mesh

In spatial capture-recapture, and therefore also in acoustic spatial capture-recapture, we estimate a density surface. Total abundance is derived by integrating over space, giving a total abundance for the area. In practice, we approximate the integral by summing the function over a discrete mesh. We cannot compute the approximation over infinite 2-dimensional space, so we only approximate the integral over our study

area $\mathcal{A} \subset \mathbb{R}^2$. The condition on this study area is that the probability of a call originating at the bounds or farther out has a probability of being detected on at least two sensors is at most 1%. This makes the integration computationally feasible whilst inducing negligible bias.

In our case study, loud calls could potentially be audible hundreds of kilometres from the source. To keep computer run-time low, we had to limit the number of grid cells in the order of 500, which limits the resolution of our mesh. As our study site did not contain many steep gradients in the covariates, we did not need a high resolution to capture this. However, we did need a finer resolution closer to the sensor array to accommodate the bearings. Recall that we included bearing data with errors in our model. If we used a low resolution grid close to our array, we would have greatly reduced the probability that one of the midpoints of our grid cells was close to what the recorded bearings point towards, and thus it would have been hard for the model to estimate the bearing accuracy. Our solution was to use a mesh with increasing grid spacing the farther the cells are from the sensor array.

This mesh had a two-stage grid spacing: the 'high resolution' cells within a 10,000 meter radius from the outer sensors of the array had an area of 6.25 km^2 each; the 'low resolution' cells between 10,000 and 50,000 meters from the outer sensors had an area of 25 km^2. This way we created higher resolution near the sensors whilst keeping the number of grid cells in the mesh to 438. Every grid cell was rectangular on the Albers projection, and the associated covariate information (i.e., ocean depth and distance to the coast) was derived for the midpoint of the cell. Plots of the mesh with detection probabilities associated with the simulation parameters (which were based on parameter estimates from the real data) are presented in Appendix D. These plots that the detection probabilities are well below 1% at the bounds.

A.3 Simulation study

Simulation studies were performed to assess the general performance of the implemented method. Data were simulated in the following way:

1. Derive the expected density for every grid point and simulate the number of emitted calls for every mesh cell $m \in 1, ..., M$ from Poisson($\lambda = a_m D_m$), where a_m and D_m are area and density per unit area for cell m, respectively.
2. Ensemble a dataset of all emitted calls with all their characteristic information.
3. Calculate the distances for the calls to all sensors.
4. For all calls, derive the bearings from the sensors to their locations, and add error from VonMises($0, \kappa$).
5. For every call, simulate a source level from a zero-truncated Normal distribution $\mathcal{N}_0^\infty(\mu_s, \sigma_s^2)$.
6. Using the distances and the transmission loss parameter β_r, derive the received levels for all calls at all sensors, and add measurement error from $\mathcal{N}(0, \sigma_r^2)$.
7. Assign detection probability g_0 to every call at a sensor when r_{ij} is at least truncation value t_r, and 0 otherwise.
8. Simulate from $U(0, 1)$ and accept as a positive detection if the value simulated is less than or equal to the detection probability. Then, remove all calls and associated data if the total number detections for that call is less than two.
9. Return the detection histories, received levels, and bearings for the remaining calls.

Alternatively, if a mixture on the bearing precision is desired, step (4) is replaced by:

4a) For all calls, derive the bearings from the sensors to the call.

4b) Simulate from $U(0, 1)$ and label the detection of a call on a sensor as 'bad/low-precision' if the value simulated is less than or equal to ψ_κ. To the bearings for these sensor-call interactions, add error from VonMises(0, κ).

4c) For the remaining bearings (the 'good/high-precision' bearings) add an error from VonMises(0, $\kappa + \delta_\kappa$). We link the two distributions through their dispersion parameters to ensure identifiability by making it explicit that the second distribution always has a larger dispersion parameter, and thus lower variance.

TABLE A.1: Parameter estimates for the variable source level model with a homogeneous or inhomogeneous density model, and the simulation parameters.

Parameter	Homogeneous density	Inhomogeneous density	Simulation
g_0	0.53164	0.55022	0.6
β_r	19.14684	17.65528	18.0
σ_r	2.62344	2.70003	2.7
μ_s	166.62167	162.46326	163.0
σ_s	5.13998	4.98272	5.0
κ	0.30208	0.23774	0.3
δ_κ	36.94962	37.45905	36.7
ψ_κ	0.10337	0.09690	0.1
β_0	1.20156	-12.08586	-12.0
β_d	-	45.48018	45.0
β_{d^2}	-	-53.83861	-53.0

A.3.1 Simulation studies for functionality

We first performed several simulation studies to evaluate the bias and variance of the estimates. The code used to simulate data can be found in `simulate_data_Rcpp.R`. The parameters values used for the simulations were based on exploratory fits on the real data. Using a truncation value of 96 dB for the detection function, we fit a homogeneous and non-homogeneous density model. Both models included a mixture on bearing precision and source level as a latent variable.

We did not centre the covariates because this makes the spatial density covariates harder to interpret, and an exploratory analysis with centred covariates yielded almost identical results (Table A.1).

Our second simulation scenario involved simulating from a model with a fixed source level. Fitting a fixed source level model to the data, we got the following parameter estimates (Table A.2).

A.3.2 Checking the buffer width

Figures A.2–A.5 are associated with the parameters from Table A.1, corresponding to the model with a variable source level. Figures A.6–A.9 are associated with the parameters from Table A.2, corresponding to the model with a fixed source level.

TABLE A.2: Parameter estimates for the fixed source level model with a homogeneous or inhomogeneous density model, and the simulation parameters.

Parameter	Homogeneous density	Inhomogeneous density	Simulation
g_0	0.61287	0.65202	0.6
β_r	17.00578	14.54507	14.5
σ_r	4.20759	4.51637	4.5
μ_s	163.73203	153.98830	155.0
κ	0.40459	0.31278	0.3
δ_κ	33.16735	35.88457	34.7
ψ_κ	0.11720	0.10118	0.1
β_0	0.01268	-16.09090	-16.0
β_d	-	56.95405	57.0
β_{d^2}	-	-68.37234	-68.5

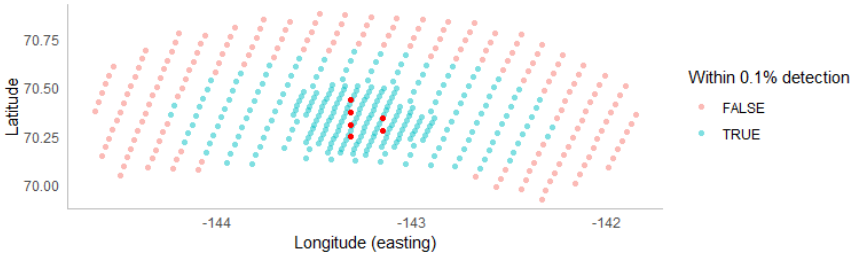


FIGURE A.2: Plot of midpoints of the grid cells. Blue indicates that the expected probability of detecting a call produced at that point at least twice is at least the threshold, and pink indicates otherwise. Detection probabilities were derived from the homogeneous density parameters for the variable source level model. The sensor array is indicated in red.

They show the midpoints of the grid cells, and whether a call produced at that point is expected to surpass the multiply-detection threshold (blue) or not (pink), which is either 1% or 0.1%. We used a threshold of 0.1% for our study, but these figures show how we could have reduced the size or our mesh by increasing this threshold to 1%. The red dots represent the sensor array.

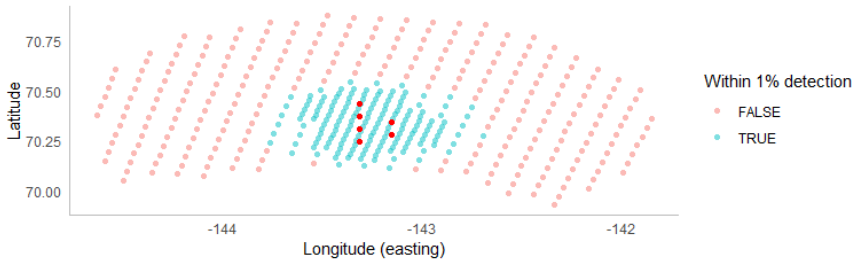


FIGURE A.3: Plot of midpoints of the grid cells. Blue indicates that the expected probability of detecting a call produced at that point at least twice is at least the threshold, and pink indicates otherwise. Detection probabilities were derived from the homogeneous density parameters for the variable source level model. The sensor array is indicated in red.

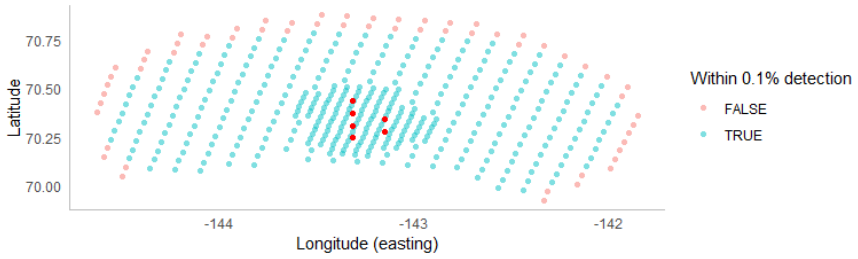


FIGURE A.4: Plot of midpoints of the grid cells. Blue indicates that the expected probability of detecting a call produced at that point at least twice is at least the threshold, and pink indicates otherwise. Detection probabilities were derived from the inhomogeneous density parameters for the variable source level model. The sensor array is indicated in red.

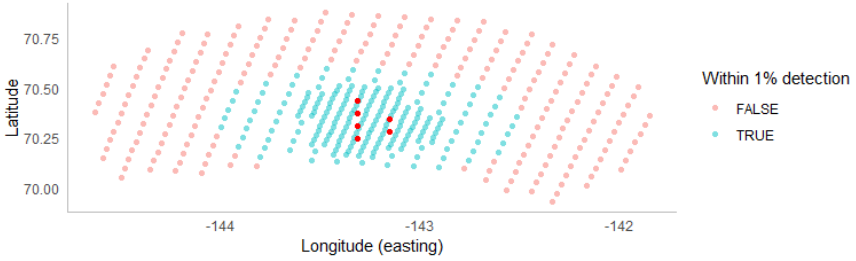


FIGURE A.5: Plot of midpoints of the grid cells. Blue indicates that the expected probability of detecting a call produced at that point at least twice is at least the threshold, and pink indicates otherwise. Detection probabilities were derived from the inhomogeneous density parameters for the variable source level model. The sensor array is indicated in red.

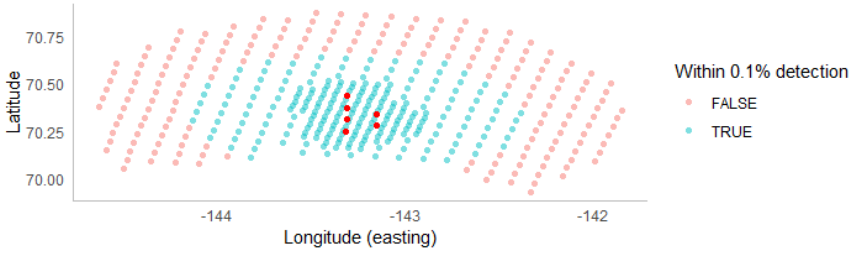


FIGURE A.6: Plot of midpoints of the grid cells. Blue indicates that the expected probability of detecting a call produced at that point at least twice is at least the threshold, and pink indicates otherwise. Detection probabilities were derived from the homogeneous density parameters for the fixed source level model. The sensor array is indicated in red.

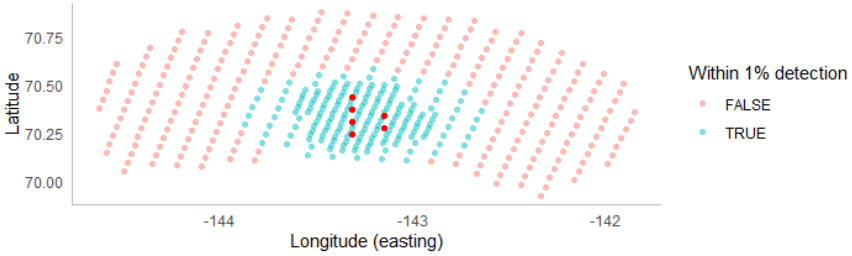


FIGURE A.7: Plot of midpoints of the grid cells. Blue indicates that the expected probability of detecting a call produced at that point at least twice is at least the threshold, and pink indicates otherwise. Detection probabilities were derived from the homogeneous density parameters for the fixed source level model. The sensor array is indicated in red.

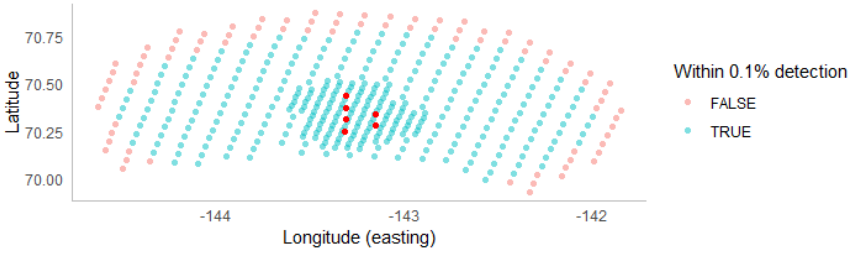


FIGURE A.8: Plot of midpoints of the grid cells. Blue indicates that the expected probability of detecting a call produced at that point at least twice is at least the threshold, and pink indicates otherwise. Detection probabilities were derived from the inhomogeneous density parameters for the fixed source level model. The sensor array is indicated in red.

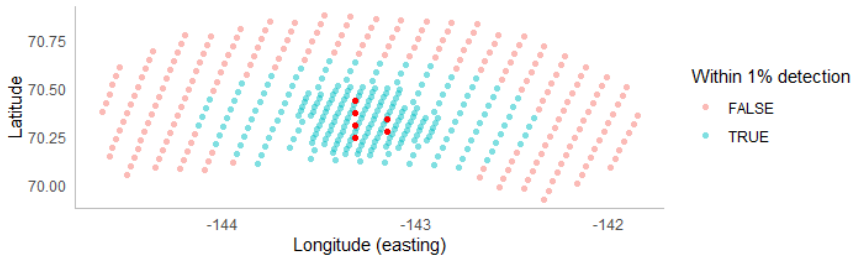


FIGURE A.9: Plot of midpoints of the grid cells. Blue indicates that the expected probability of detecting a call produced at that point at least twice is at least the threshold, and pink indicates otherwise. Detection probabilities were derived from the inhomogeneous density parameters for the fixed source level model. The sensor array is indicated in red.

A.4 Candidate density models

We fitted a wide range of models to the data from our case study. In total, 35 models were tried (including a naive constant density model, for completeness). Earlier studies found that the bowhead whales tend to migrate between 10 and 75 kilometres from the coast, over a plateau that is only between 10 and 35 meters deep (Greene et al., 2004). Thus, we limited our models to combinations of those variables, with and without interaction. The density models are presented using the common model specification standard for `R`. Besides a constant model and polynomial models, we also tried more flexible relationships between call density and the explanatory variables. This was achieved with generalised additive models using the `mgcv`-package in `R` (Wood, 2017). Through the smooth function `mgcv::s()` we manually set the number of knots `k`, used thin plate regression splines as the basis function (which is the default), and did not allow for any penalisation (`fx = TRUE`). We present the models that were considered in Figure A.10.

An overview of the model fits is presented in Table A.3. We derived the AIC for all models (Akaike, 1998). AIC assumes independence between the observations (as it is a likelihood based statistic) and penalises for the number of parameters in the model.

TABLE A.3: Overview of tested density models and results. 'ID' is a unique identifier, 'Density Model' is the R specification of the density model, 'N_hat' is the estimated abundance, '#par' is the number of model parameters, and 'AIC' is the AIC score.

ID	Density Model	N_hat	#par	AIC
33	D ~ s(depth, k = 6, fx = TRUE) + s(distance_to_coast, k = 6, fx = TRUE)	5741.39	19	5894.56
30	D ~ s(distance_to_coast, k = 6, fx = TRUE) + depth	5141.57	15	5907.64
31	D ~ s(distance_to_coast, k = 6, fx = TRUE) + depth + depth2	4029.96	16	5914.61
32	D ~ s(depth, k = 4, fx = TRUE) + s(distance_to_coast, k = 4, fx = TRUE)	5974.55	15	5920.34
8	D ~ logdepth + distance_to_coast + distance_to_coast2 + distance_to_coast3	7270.06	13	5944.46
28	D ~ s(distance_to_coast, k = 7, fx = TRUE)	6318.69	15	5948.72
29	D ~ s(distance_to_coast, k = 8, fx = TRUE)	6274.61	16	5950.36
15	D ~ depth + depth2 + distance_to_coast + distance_to_coast2 + distance_to_coast3	6263.86	14	5970.17
27	D ~ s(distance_to_coast, k = 6, fx = TRUE)	4977.06	14	5976.51
26	D ~ s(distance_to_coast, k = 5, fx = TRUE)	6112.98	13	5989.40
25	D ~ s(distance_to_coast, k = 4, fx = TRUE)	6647.31	12	5991.38
3	D ~ distance_to_coast + distance_to_coast2 + distance_to_coast3	6012.47	12	6021.90
14	D ~ depth + depth2 + distance_to_coast + distance_to_coast2	5053.38	13	6054.79
35	D ~ distance_to_coast + distance_to_coast2 + depth + depth2 + depth:distance_to_coast	5117.73	14	6056.06
23	D ~ s(depth, k = 6, fx = TRUE) + distance_to_coast + distance_to_coast2	5286.00	16	6059.02
9	D ~ logdepth + depth + distance_to_coast + distance_to_coast2	4959.21	13	6059.69
12	D ~ depth + distance_to_coast + distance_to_coast2	5185.96	12	6061.16
24	D ~ s(distance_to_coast, k = 3, fx = TRUE)	5912.88	11	6061.53
2	D ~ distance_to_coast + distance_to_coast2	6050.36	11	6084.48
34	D ~ distance_to_coast + depth + depth:distance_to_coast	5024.66	12	6109.61
22	D ~ s(depth, k = 6, fx = TRUE) + distance_to_coast	6015.53	15	6173.51
10	D ~ logdepth + depth + depth2 + distance_to_coast	8430.18	13	6178.89
20	D ~ s(depth, k = 7, fx = TRUE)	8826.32	15	6246.69
19	D ~ s(depth, k = 6, fx = TRUE)	10760.82	14	6250.61

21	D ˜ s(depth, k = 8, fx = TRUE)	9122.84	16	6252.56
13	D ˜ depth + depth2 + distance_to_coast	11339.95	12	6273.32
7	D ˜ logdepth + depth + depth2	9691.47	12	6277.54
18	D ˜ s(depth, k = 5, fx = TRUE)	23712.90	13	6305.30
5	D ˜ depth + depth2	113059.51	11	6326.00
17	D ˜ s(depth, k = 4, fx = TRUE)	25480.35	12	6349.51
16	D ˜ s(depth, k = 3, fx = TRUE)	30564.71	11	6362.47
6	D ˜ logdepth	37011.68	10	6381.44
1	D ˜ 1	26638.36	9	6398.78
11	D ˜ depth + distance_to_coast	27432.40	11	6400.66
4	D ˜ depth	27365.89	10	6400.74

A.4.1 Results from the 999 bootstraps

To find the uncertainty around our point estimates from our best model (33), we bootstrapped (i.e., re-sampled with replacement) the original calls 999 times. This created 999 'equally likely' data sets, assuming that calls are independent. The uncertainty estimates from these bootstraps are presented in Table A.4.

A surprising result is the high CV of $\widehat{\log(\psi_\kappa)}$ (1423%). This is caused by a negative outlier (-234.24) and a mean close to 0. As these estimates are on the log scale and the outlier is $<< 0$, these effects are dampened when converting to the real scale, which is why $\text{CV}(\widehat{\kappa}) = 22\%$ (see Table 2.3 in Chapter 2).

TABLE A.4: Uncertainty estimates from 999 bootstraps. We present the parameter name, the link function, the estimate, the lower and upper limits of the absolute (LCL and UCL) and relative (LRCL and URCL) 95% confidence intervals, the standard deviation (SD) and the variance, and the coefficient of variation (CV). Estimates are rounded to two significant figures.

Parameter	Link	Est.	LCL	UCL	LRCL	URCL	SD	CV (%)
N	iden.	5741	3877	5935	-0.32	0.034	576	11
g_0	logit	0.20	0.072	0.36	-0.65	0.74	0.061	29
β_r	log	2.8	2.8	2.9	-0.017	0.017	0.019	0.67
σ_r	log	1.0	0.92	1.1	-0.093	0.089	0.040	4.0
μ_s	log	5.1	5.1	5.1	-0.0030	0.0054	0.0094	0.19
σ_s	log	1.7	1.5	1.8	-0.096	0.040	0.051	3.0
κ	log	-0.27	-0.94	0.18	-2.5	1.7	7.4	1423
δ_κ	log	3.8	3.7	4.0	-0.028	0.044	0.064	1.7
ψ_κ	logit	-2.2	-2.4	-1.9	-0.12	0.12	0.11	4.9
$\beta_{\text{(Intercept)}}$	log	-266	-267	-241	-0.0062	0.092	7.9	3.0
$\beta_{\text{s(depth).1}}$	log	42	20	44	-0.54	0.035	18	46
$\beta_{\text{s(depth).2}}$	log	302	293	304	-0.032	0.0069	6.7	2.2
$\beta_{\text{s(depth).3}}$	log	41	-24	50	-1.6	0.22	22	58
$\beta_{\text{s(depth).4}}$	log	-142	-145	-124	-0.026	0.12	6.9	4.9
$\beta_{\text{s(depth).5}}$	log	155	136	158	-0.12	0.020	6.2	4.0
$\beta_{\text{s(distance\_to\_coast).1}}$	log	159	145	162	-0.089	0.016	4.5	2.8
$\beta_{\text{s(distance\_to\_coast).2}}$	log	-450	-453	-407	-0.0083	0.095	16	3.5
$\beta_{\text{s(distance\_to\_coast).3}}$	log	101	90	104	-0.11	0.029	4.3	4.3
$\beta_{\text{s(distance\_to\_coast).4}}$	log	-109	-110	-97	-0.011	0.12	5.1	4.7
$\beta_{\text{s(distance\_to\_coast).5}}$	log	-225	-231	-199	-0.029	0.12	11	5.0

```
models <- list(
  D ~ 1,
  D ~ distance_to_coast + distance_to_coast2,
  D ~ distance_to_coast + distance_to_coast2 + distance_to_coast3,
  D ~ depth,
  D ~ depth + depth2,
  D ~ logdepth,
  D ~ logdepth + depth + depth2,
  D ~ logdepth + distance_to_coast + distance_to_coast2 +
    distance_to_coast3,
  D ~ logdepth + depth + distance_to_coast + distance_to_coast2,
  D ~ logdepth + depth + depth2 + distance_to_coast,
  D ~ depth + distance_to_coast,
  D ~ depth + distance_to_coast + distance_to_coast2,
  D ~ depth + depth2 + distance_to_coast,
  D ~ depth + depth2 + distance_to_coast + distance_to_coast2,
  D ~ depth + depth2 + distance_to_coast + distance_to_coast2 +
    distance_to_coast3,
  D ~ s(depth, k = 3, fx = TRUE),
  D ~ s(depth, k = 4, fx = TRUE),
  D ~ s(depth, k = 5, fx = TRUE),
  D ~ s(depth, k = 6, fx = TRUE),
  D ~ s(depth, k = 7, fx = TRUE),
  D ~ s(depth, k = 8, fx = TRUE),
  D ~ s(depth, k = 6, fx = TRUE) + distance_to_coast,
  D ~ s(depth, k = 6, fx = TRUE) + distance_to_coast +
    distance_to_coast2,
  D ~ s(distance_to_coast, k = 3, fx = TRUE),
  D ~ s(distance_to_coast, k = 4, fx = TRUE),
  D ~ s(distance_to_coast, k = 5, fx = TRUE),
  D ~ s(distance_to_coast, k = 6, fx = TRUE),
  D ~ s(distance_to_coast, k = 7, fx = TRUE),
  D ~ s(distance_to_coast, k = 8, fx = TRUE),
  D ~ s(distance_to_coast, k = 6, fx = TRUE) + depth,
  D ~ s(distance_to_coast, k = 6, fx = TRUE) + depth + depth2,
  D ~ s(depth, k = 4, fx = TRUE) +
    s(distance_to_coast, k = 4, fx = TRUE),
  D ~ s(depth, k = 6, fx = TRUE) +
    s(distance_to_coast, k = 6, fx = TRUE),
  D ~ distance_to_coast + depth + depth:distance_to_coast,
  D ~ distance_to_coast + distance_to_coast2 + depth + depth2 +
    depth:distance_to_coast
)
```

FIGURE A.10: R code chunk with the 35 candidate spatial density models.

Appendix B

Supplementary materials for Chapter 4

B.1 Performance metrics

B.1.1 Definitions

Every simulation had a different realised sex-specific adult abundance, due to stochasticity in the demographic processes such as fecundity. Therefore, the error in the abundance estimate for a specific simulation was defined as the estimated abundance minus the true abundance for said simulation, such that

$$\varepsilon_{s,m} = \widehat{N}_s^{(m)} - N_{s,100}^{(m)}, \tag{B.1}$$

where $m \in 1, ..., M$ denotes the simulation index and $s \in \{♀, ♂\}$ denotes the sex. We define the relative error as relative to the true adult abundance in year 100 of the simulation, i.e.,

$$\frac{\varepsilon_{s,m}}{N_{s,100}^{(m)}}.$$

The mean error, median error, mean relative error, median relative error, and the mean absolute error all easily follow from this, where 'mean' always refers to the arithmetic

mean. Finally, the root mean square error (RMSE) is defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{M}\sum_{m=1}^{M} \varepsilon_{s,m}^2}.$$

All of these definitions match the definitions common in other literature.

B.1.2 Results

TABLE B.1: Performance metrics for the estimation of parameter N_{φ}, extracted from 1000 simulations of the simple shark population. The columns are, from left to right: scenario, median relative error, mean relative error, median error, mean error, mean absolute error, and root mean square error. Scenarios ME-67:GC-10, ME-67:GC-5, ME-67:GC+0, ME-33:GC-10, ME-33:GC-5, and ME+0:GC-10 were not included as (most of) the simulations did not fit successfully. Scenario 3-3 uses the correct measurement error (2.89) and growth curve specification.

Scenario	Mdn. RE	Mean RE	Mdn. Err.	Mean Err.	MAE	RMSE
ME-67:GC+5	-28.48	-25.66	-223.83	-203.94	220.61	245.14
ME-67:GC+10	-62.69	-61.36	-495.14	-487.72	487.72	495.50
ME-33:GC+0	2.88	7.32	22.72	58.23	142.91	201.11
ME-33:GC+5	-28.54	-25.71	-224.94	-204.40	220.93	245.38
ME-33:GC+10	-62.78	-61.34	-495.26	-487.58	487.58	495.28
ME+0:GC-5	31.80	37.56	252.60	298.59	304.39	390.33
ME+0:GC+0	2.57	7.14	20.83	56.78	142.47	200.48
ME+0:GC+5	-28.72	-25.85	-227.03	-205.50	221.78	246.05
ME+0:GC+10	-62.70	-61.23	-494.06	-486.74	486.74	494.33
ME+33:GC-10	43.85	50.21	343.32	399.15	400.98	488.70
ME+33:GC-5	30.59	36.34	243.51	288.88	295.53	381.69
ME+33:GC+0	2.23	6.76	18.13	53.74	141.65	199.10
ME+33:GC+5	-29.21	-26.29	-230.69	-209.01	224.53	248.47
ME+33:GC+10	-62.81	-61.25	-494.00	-486.85	486.85	494.26
ME+67:GC-10	41.04	47.21	320.66	375.24	377.73	466.13
ME+67:GC-5	29.16	34.90	232.41	277.39	285.22	371.65
ME+67:GC+0	1.79	6.12	13.53	48.64	140.45	196.86
ME+67:GC+5	-29.80	-26.97	-235.30	-214.35	228.79	252.27
ME+67:GC+10	-62.85	-61.37	-495.30	-487.84	487.84	495.04

TABLE B.2: Performance metrics for the estimation of parameter $N_{♂}$, extracted from 1000 simulations for the simple shark population. The columns are, from left to right: scenario, median relative error, mean relative error, median error, mean error, mean absolute error, and root mean square error. Scenarios ME-67:GC-10, ME-67:GC-5, ME-67:GC+0, ME-33:GC-10, ME-33:GC-5, and ME+0:GC-10 were not included as (most of) the simulations did not fit successfully. Scenario 3-3 used the correct measurement error (2.89) and growth curve specification.

Scenario	Mdn. RE	Mean RE	Mdn. Err.	Mean Err.	MAE	RMSE
ME-67:GC+5	-28.40	-25.93	-223.56	-206.14	217.13	243.78
ME-67:GC+10	-62.28	-61.38	-488.98	-487.42	487.42	494.97
ME-33:GC+0	2.96	6.74	24.43	52.92	144.84	189.98
ME-33:GC+5	-28.53	-25.99	-224.78	-206.64	217.45	244.05
ME-33:GC+10	-62.41	-61.37	-489.77	-487.34	487.34	494.79
ME+0:GC-5	31.59	36.63	247.66	289.87	297.53	374.35
ME+0:GC+0	2.79	6.55	22.52	51.44	144.49	189.38
ME+0:GC+5	-28.72	-26.14	-226.29	-207.83	218.33	244.83
ME+0:GC+10	-62.30	-61.28	-489.39	-486.58	486.58	493.89
ME+33:GC-10	43.72	49.00	341.71	387.75	390.42	470.34
ME+33:GC-5	30.62	35.41	238.47	280.12	289.00	365.82
ME+33:GC+0	2.34	6.16	19.50	48.30	143.76	188.06
ME+33:GC+5	-29.16	-26.59	-228.60	-211.41	221.07	247.38
ME+33:GC+10	-62.32	-61.30	-489.56	-486.78	486.78	493.90
ME+67:GC-10	40.98	46.03	320.23	364.13	367.87	448.20
ME+67:GC-5	29.36	33.96	229.31	268.62	279.08	355.90
ME+67:GC+0	1.72	5.51	13.73	43.13	142.59	185.95
ME+67:GC+5	-29.89	-27.27	-233.68	-216.81	225.38	251.36
ME+67:GC+10	-62.58	-61.44	-490.18	-487.86	487.86	494.77

B.2 Variance estimates

B.2.1 Definitions

We used the standard deviation of the error term, i.e., standard error, as a measure of uncertainty for estimated adult abundance. We did not use standard deviation of the adult abundance estimates themselves, as these contain stochasticity from the simulation process (every simulation had a slightly different true adult abundance in the final year). We define the simulation-based or empirical standard error of the adult abundance estimate as

$$\sigma_s^N = \sqrt{\frac{1}{M}\sum_{m=1}^{M}(\varepsilon_{s,m} - \bar{\varepsilon}_{s,m})^2}, \tag{B.2}$$

TABLE B.3: Performance metrics for the estimation of parameter $N_{♀}$, extracted from 1000 simulations of the complex shark population. The columns are, from left to right: scenario, median relative error, mean relative error, median error, mean error, mean absolute error, and root mean square error. Scenarios ME-67:GC-10, ME-67:GC-5, ME-67:GC+0, ME-33:GC-10, ME-33:GC-5, and ME+0:GC-10 were not included as (most of) the simulations did not fit successfully. Scenario 3-3 used the correct measurement error (2.89) and growth curve specification.

Scenario	Mdn. RE	Mean RE	Mdn. Err.	Mean Err.	MAE	RMSE
ME-67:GC+5	-38.21	-35.96	-194.00	-185.43	187.72	199.33
ME-67:GC+10	-69.41	-68.34	-355.48	-352.09	352.09	356.40
ME-33:GC+0	1.89	5.30	9.12	26.84	82.06	112.47
ME-33:GC+5	-38.37	-36.07	-194.96	-185.99	188.23	199.71
ME-33:GC+10	-69.34	-68.26	-354.98	-351.70	351.70	355.94
ME+0:GC-5	55.54	61.97	282.61	318.31	318.38	365.22
ME+0:GC+0	1.35	4.64	6.48	23.44	81.55	111.32
ME+0:GC+5	-38.63	-36.31	-195.61	-187.25	189.38	200.68
ME+0:GC+10	-69.19	-68.16	-353.91	-351.15	351.15	355.29
ME+33:GC-10	109.97	119.20	556.22	612.37	612.37	666.21
ME+33:GC-5	53.20	59.79	270.34	307.08	307.24	354.45
ME+33:GC+0	0.57	3.70	2.63	18.60	80.97	109.82
ME+33:GC+5	-39.05	-36.72	-197.27	-189.35	191.32	202.37
ME+33:GC+10	-69.23	-68.08	-353.09	-350.72	350.72	354.71
ME+67:GC-10	105.31	114.40	534.45	587.65	587.65	641.11
ME+67:GC-5	50.39	57.13	258.11	293.33	293.59	341.34
ME+67:GC+0	-0.63	2.49	-3.03	12.36	80.40	108.17
ME+67:GC+5	-39.51	-37.25	-200.26	-192.09	193.91	204.65
ME+67:GC+10	-69.15	-68.02	-352.44	-350.42	350.42	354.27

where $m \in 1, ..., M$ denotes the simulation index, $s \in \{♀, ♂\}$ denotes the sex, and $\bar{\varepsilon}_{s,m}$ is the arithmetic mean of $\varepsilon_{s,m}$ (see Equation (B.1.1)).

We define the model-based or estimated standard error for a single simulation m as

$$\widehat{\sigma}^{N}_{s,m} = \sqrt{\widehat{\mathrm{Var}}(\widehat{N}^{(m)}_{s})}, \tag{B.3}$$

where $\widehat{\mathrm{Var}}(\widehat{N}^{(m)}_{s})$ is the estimated variance extracted from the Hessian matrix for sex s and simulation m.

We also present the confidence interval (CI) coverage to evaluate whether variance is correctly estimated by the model. The 95% CI coverage is derived as the ratio of

TABLE B.4: Performance metrics for the estimation of parameter $N_{♂}$, extracted from 1000 simulations for the complex shark population. The columns are, from left to right: scenario, median relative error, mean relative error, median error, mean error, mean absolute error, and root mean square error. Scenarios ME-67:GC-10, ME-67:GC-5, ME-67:GC+0, ME-33:GC-10, ME-33:GC-5, and ME+0:GC-10 were not included as (most of) the simulations did not fit successfully. Scenario 3-3 used the correct measurement error (2.89) and growth curve specification.

Scenario	Mdn. RE	Mean RE	Mdn. Err.	Mean Err.	MAE	RMSE
ME-67:GC+5	-31.74	-29.79	-203.73	-194.20	201.00	220.13
ME-67:GC+10	-61.34	-60.03	-393.89	-390.66	390.66	397.78
ME-33:GC+0	2.13	5.35	12.74	34.22	113.74	152.24
ME-33:GC+5	-31.95	-29.89	-204.63	-194.80	201.48	220.51
ME-33:GC+10	-61.30	-59.99	-393.09	-390.44	390.44	397.48
ME+0:GC-5	45.75	51.88	297.86	336.61	337.30	401.99
ME+0:GC+0	1.59	4.91	10.28	31.35	113.06	151.03
ME+0:GC+5	-32.33	-30.13	-206.50	-196.36	202.81	221.63
ME+0:GC+10	-61.23	-59.98	-392.55	-390.36	390.36	397.27
ME+33:GC-10	88.97	98.09	579.49	636.75	636.75	703.85
ME+33:GC-5	44.11	50.38	288.00	326.83	327.73	392.72
ME+33:GC+0	1.05	4.24	6.59	27.01	112.18	149.32
ME+33:GC+5	-32.61	-30.56	-208.75	-199.14	205.21	223.70
ME+33:GC+10	-61.30	-60.04	-392.71	-390.76	390.76	397.51
ME+67:GC-10	85.75	94.83	559.04	615.54	615.54	682.57
ME+67:GC-5	42.24	48.54	276.01	314.85	316.10	381.44
ME+67:GC+0	0.14	3.35	0.95	21.20	111.13	147.24
ME+67:GC+5	-33.16	-31.13	-211.30	-202.85	208.45	226.55
ME+67:GC+10	-61.38	-60.19	-394.36	-391.68	391.68	398.23

simulations in which the 95% log-normal CI (which is normal on the link scale) contains the true adult abundance.

B.2.2 Results

TABLE B.5: Model-based estimates for the standard error of adult abundance estimates, averaged over 1000 simulations ($\overline{\sigma}_s^N$), the empirical standard errors of adult abundance estimates, derived from 1000 simulations (σ_s^N), the difference between the two ($\Delta\sigma_s^N$), and the difference relative to the empirical standard error ($\Delta\sigma_s^N/\sigma_s^N$; %), for both sexes of the simple species. The first number in a scenario label refers to the assumed measurement error on length; the second number refers to the used growth function. Scenarios ME-67:GC-10, ME-67:GC-5, ME-67:GC+0, ME-33:GC-10, ME-33:GC-5, and ME+0:GC-10 were not included as (most of) the simulations did not successfully fit. Scenario 3-3 used the correct measurement error standard deviation (2.89) and growth curve specification.

	Males				Females			
Scenario	$\overline{\sigma}_{\male}^N$	$\sigma_{\male}^N$	$\Delta\sigma_{\male}^N$	$\Delta\sigma_{\male}^N/\sigma_{\male}^N$	$\overline{\sigma}_{\female}^N$	$\sigma_{\female}^N$	$\Delta\sigma_{\female}^N$	$\Delta\sigma_{\female}^N/\sigma_{\female}^N$
ME-67:GC+5	126.84	130.20	-3.36	-2.58	128.27	136.09	-7.82	-5.75
ME-67:GC+10	65.96	86.14	-20.18	-23.43	66.49	87.54	-21.05	-24.05
ME-33:GC+0	183.04	182.56	0.49	0.27	185.51	192.59	-7.08	-3.68
ME-33:GC+5	126.75	129.92	-3.17	-2.44	128.19	135.83	-7.65	-5.63
ME-33:GC+10	65.99	85.56	-19.57	-22.87	66.53	87.02	-20.49	-23.54
ME+0:GC-5	234.63	237.00	-2.37	-1.00	238.17	251.53	-13.36	-5.31
ME+0:GC+0	182.76	182.35	0.41	0.22	185.23	192.36	-7.14	-3.71
ME+0:GC+5	126.51	129.46	-2.96	-2.28	127.97	135.38	-7.41	-5.47
ME+0:GC+10	66.17	84.72	-18.55	-21.89	66.74	86.31	-19.57	-22.67
ME+33:GC-10	256.19	266.36	-10.17	-3.82	260.46	282.11	-21.66	-7.68
ME+33:GC-5	232.56	235.40	-2.84	-1.21	236.10	249.59	-13.49	-5.41
ME+33:GC+0	182.11	181.85	0.27	0.15	184.61	191.81	-7.20	-3.76
ME+33:GC+5	125.76	128.52	-2.77	-2.15	127.24	134.44	-7.20	-5.35
ME+33:GC+10	66.15	83.65	-17.50	-20.92	66.74	85.30	-18.56	-21.76
ME+67:GC-10	251.11	261.46	-10.35	-3.96	255.29	276.67	-21.38	-7.73
ME+67:GC-5	230.12	233.59	-3.46	-1.48	233.65	247.46	-13.81	-5.58
ME+67:GC+0	181.04	180.97	0.07	0.04	183.54	190.85	-7.31	-3.83
ME+67:GC+5	124.62	127.23	-2.62	-2.06	126.11	133.08	-6.97	-5.24
ME+67:GC+10	65.94	82.44	-16.50	-20.02	66.55	84.13	-17.58	-20.89

TABLE B.6: Model-based estimates for the standard error of adult abundance estimates, averaged over 1000 simulations ($\overline{\sigma}_s^N$), the empirical standard errors of adult abundance estimates, derived from 1000 simulations (σ_s^N), the difference between the two ($\Delta\sigma_s^N$), and the difference relative to the empirical standard error ($\Delta\sigma_s^N/\sigma_s^N$; %), for both sexes of the complex species. The first number in a scenario label refers to the assumed measurement error on length; the second number refers to the used growth function. Scenarios ME-67:GC-10, ME-67:GC-5, ME-67:GC+0, ME-33:GC-10, ME-33:GC-5, and ME+0:GC-10 were not included as (most of) the simulations did not successfully fit. Scenario 3-3 used the correct measurement error standard deviation (2.89) and growth curve specification.

	Males				Females			
Scenario	$\overline{\sigma}_{\male}^N$	$\sigma_{\male}^N$	$\Delta\sigma_{\male}^N$	$\Delta\sigma_{\male}^N/\sigma_{\male}^N$	$\overline{\sigma}_{\female}^N$	$\sigma_{\female}^N$	$\Delta\sigma_{\female}^N$	$\Delta\sigma_{\female}^N/\sigma_{\female}^N$
ME-67:GC+5	94.60	103.71	-9.11	-8.78	64.03	73.15	-9.12	-12.47
ME-67:GC+10	53.66	74.95	-21.29	-28.40	31.49	55.27	-23.78	-43.02
ME-33:GC+0	142.41	148.42	-6.01	-4.05	105.71	109.27	-3.56	-3.26
ME-33:GC+5	94.49	103.38	-8.90	-8.61	63.93	72.79	-8.85	-12.16
ME-33:GC+10	53.72	74.51	-20.80	-27.91	31.58	54.80	-23.22	-42.37
ME+0:GC-5	205.94	219.86	-13.92	-6.33	163.31	179.16	-15.85	-8.85
ME+0:GC+0	141.84	147.82	-5.98	-4.05	105.08	108.88	-3.80	-3.49
ME+0:GC+5	94.18	102.81	-8.63	-8.40	63.71	72.22	-8.51	-11.78
ME+0:GC+10	53.75	73.85	-20.10	-27.22	31.70	54.06	-22.36	-41.36
ME+33:GC-10	269.24	300.08	-30.84	-10.28	221.69	262.50	-40.81	-15.55
ME+33:GC-5	203.93	217.85	-13.92	-6.39	161.13	177.10	-15.97	-9.02
ME+33:GC+0	140.97	146.93	-5.97	-4.06	104.17	108.29	-4.12	-3.81
ME+33:GC+5	93.63	101.96	-8.33	-8.17	63.33	71.47	-8.14	-11.39
ME+33:GC+10	53.68	72.97	-19.29	-26.44	31.80	53.13	-21.33	-40.15
ME+67:GC-10	264.83	295.13	-30.30	-10.27	216.85	256.41	-39.57	-15.43
ME+67:GC-5	201.47	215.43	-13.96	-6.48	158.47	174.64	-16.17	-9.26
ME+67:GC+0	139.79	145.77	-5.98	-4.10	102.99	107.51	-4.53	-4.21
ME+67:GC+5	92.89	100.93	-8.05	-7.97	62.82	70.62	-7.81	-11.05
ME+67:GC+10	53.51	71.95	-18.44	-25.62	31.88	52.10	-20.22	-38.81

B.3 Modelling population growth

In an alternative version of this model we included a growth parameter r, which allowed for the estimation of exponential growth or decline in the population size. Including growth in the model meant that the distribution of age given measured length $f(a|\ell^*)$ is no longer defined exclusively by survival parameter ϕ but rather by a combination of ϕ and r (Caswell, 2006, Section 4.5.2.1), given that the population reached a new stable age distribution. Hillary et al. (2018) estimated $f(a|\ell^*)$ by grouping the data and fitting a multinomial distribution, however this was beyond the scope of our research. We could have assumed that the population had settled into a new equilibrium, in which case the age distribution $f(a)$ would be proportional to the dominant eigenvector from the associated Leslie matrix. However, as our population growth/decline was stochastic rather than systematic, this did not seem appropriate (see the Discussion for more detail). We did run our simulation study with yearly growth parameters $r_{♀}$ and $r_{♂}$ assuming that the $f(a|\ell^*)$ as presented in the main text was *approximately* correct, i.e., ignoring population growth in the age distribution formulation. We decided not to include this part of the study in the main body of this thesis, as we did not believe that the results could be used to accurately assess the effect of incorrect ageing on parameter estimation. Nonetheless, we included these results here for completeness as they could contain some valuable insight and form the basis for future research. As we considered abundance for both sexes separately, we estimated the following four parameters: $N_{♀,t_0}$, $N_{♂,t_0}$, $r_{♀}$ and $r_{♂}$, where t_0 is some reference year.. The kinship probabilities remained the same as presented in the Equations (4.4)–(4.7). As population size was no longer assumed equal for all years, abundances in different years are linked through a geometric population dynamics model:

$$N_t = N_{t_0} r^{t-t_0}, \tag{B.4}$$

where $r \in (0, \infty)$ denotes the yearly growth rate. We set $t_0 = 2014$ to match Bradley et al. (2017a) as closely as possible.

TABLE B.7: The 95% confidence interval (CI) coverage for the sex-specific adult abundance for the successfully converging scenarios, denoted in percentages (%). The 95% CIs were estimated for every simulation using the model-based standard error assuming a normal distribution on the link scale, which results in a log-normal distribution on the real scale.

	Simple species		Complex species	
Scenario	$\widehat{N}_{♂}$	$\widehat{N}_{♀}$	$\widehat{N}_{♂}$	$\widehat{N}_{♀}$
ME-67:GC+5	58.5	60.0	49.0	30.5
ME-67:GC+10	1.6	2.7	2.7	0.8
ME-33:GC+0	96.3	96.4	95.9	95.8
ME-33:GC+5	58.0	59.9	48.6	30.3
ME-33:GC+10	1.5	2.6	2.7	0.8
ME+0:GC-5	73.2	74.4	49.7	28.4
ME+0:GC+0	96.1	96.4	95.9	95.4
ME+0:GC+5	57.7	59.7	48.3	29.9
ME+0:GC+10	1.7	2.5	2.6	0.7
ME+33:GC-10	56.4	56.8	4.2	0.7
ME+33:GC-5	74.7	76.4	51.6	31.0
ME+33:GC+0	96.1	96.2	95.7	94.8
ME+33:GC+5	56.3	58.4	47.5	29.0
ME+33:GC+10	1.7	2.6	2.4	0.6
ME+67:GC-10	60.7	62.1	5.3	1.1
ME+67:GC-5	77.1	78.2	54.6	35.4
ME+67:GC+0	96.3	96.2	95.5	94.4
ME+67:GC+5	54.4	55.9	46.4	27.8
ME+67:GC+10	1.5	2.6	2.4	0.4

B.3.1 Estimated abundance through time

We fit our 25 scenarios, consisting of all combinations of 5 different measurement errors and 5 different growth curves, to both populations. We modelled the male and female side of the population separately resulting in four figures, each containing 19 population history plots—six plots are blank since the models in these scenarios did not (all) fit correctly. In Figures B.1–B.4 we notice a similar pattern of over- and underestimation related to shifting the growth curves. However, as we also model exponential growth or decline, we also notice effects of shifting the growth curves on the direction and magnitude of this trend. Albeit potentially informative, due to the inconsistency between modelling growth and the assumed age distribution we believe that these results can not be directly used for inference.

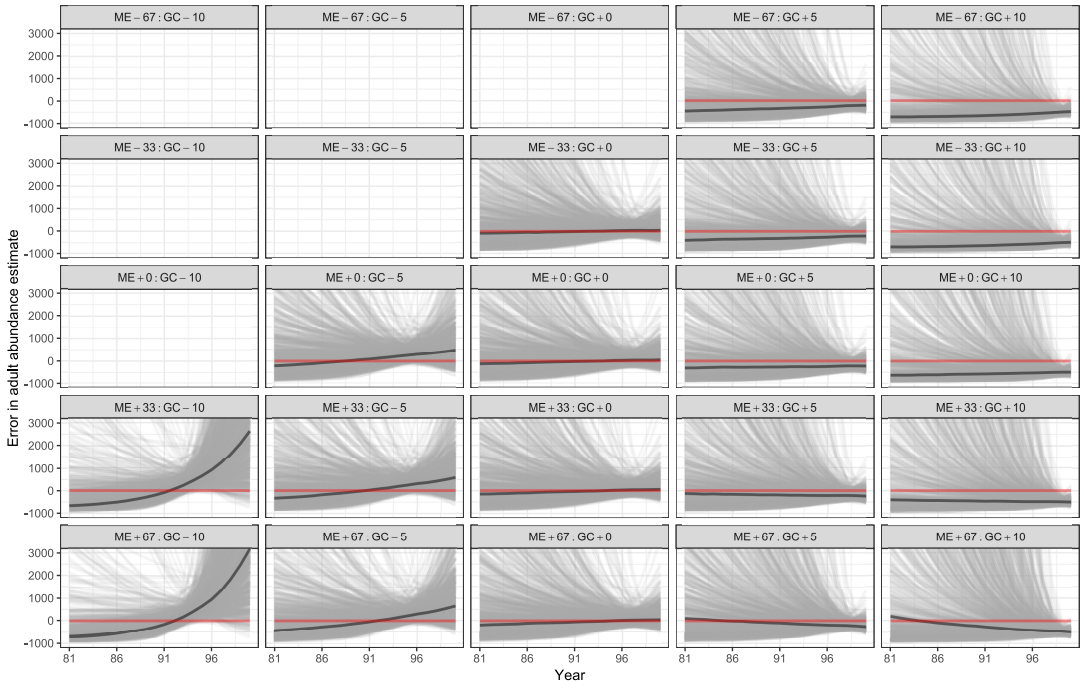


FIGURE B.1: Plots of the 1000 simple female adult population trends for the simple species for the last twenty years of the simulation, for the 19 scenarios that resulted in successful fits. The median of these 1000 trends is indicated in dark grey, and the truth adult abundance is indicated in red. The scenarios were labelled using the format 'ME±XX:GC±YY', where ME refers to the measurement error, XX denotes the percentage over- or underestimate, GC stands for growth curve, and YY denotes the percentage of up- or downwards shifting; for example, the scenario with a 33% overestimated standard deviation of length measurement error and a 5% downshifted growth curve had label ME+33:GC-5.

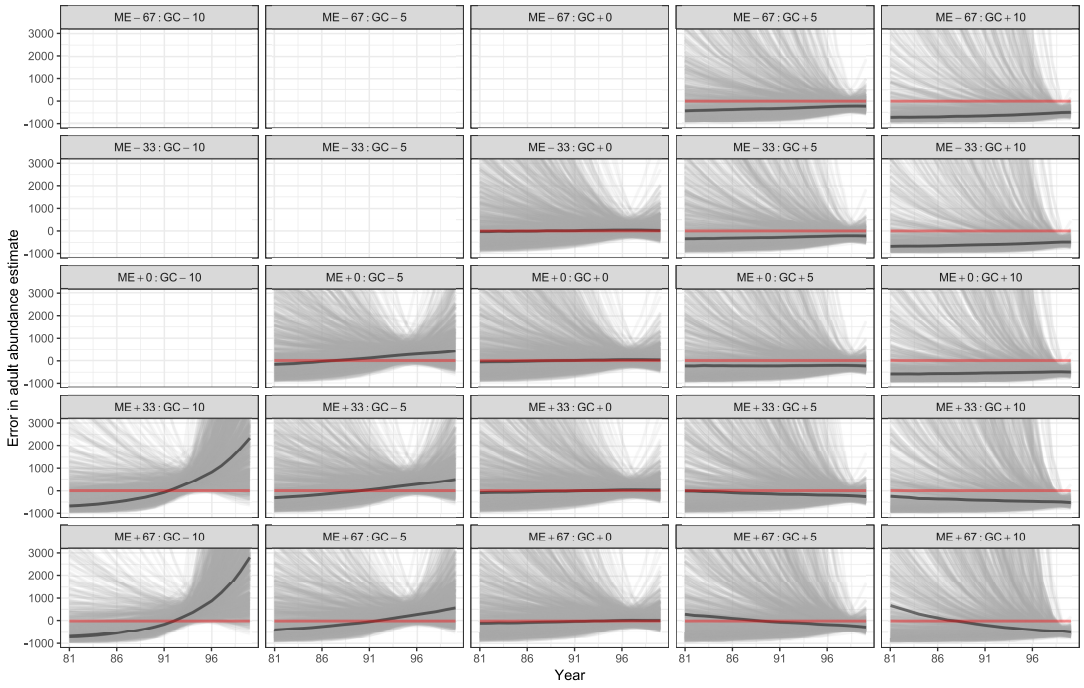


FIGURE B.2: Plots of the 1000 estimated male adult population trends for the simple species for the last twenty years of the simulation, for the 19 scenarios that resulted in successful fits. The median of these 1000 trends is indicated in dark grey, and the truth adult abundance is indicated in red. The scenarios were labelled using the format 'ME±XX:GC±YY', where ME refers to the measurement error, XX denotes the percentage over- or underestimate, GC stands for growth curve, and YY denotes the percentage of up- or downwards shifting; for example, the scenario with a 33% overestimated standard deviation of length measurement error and a 5% downshifted growth curve had label ME+33:GC-5.

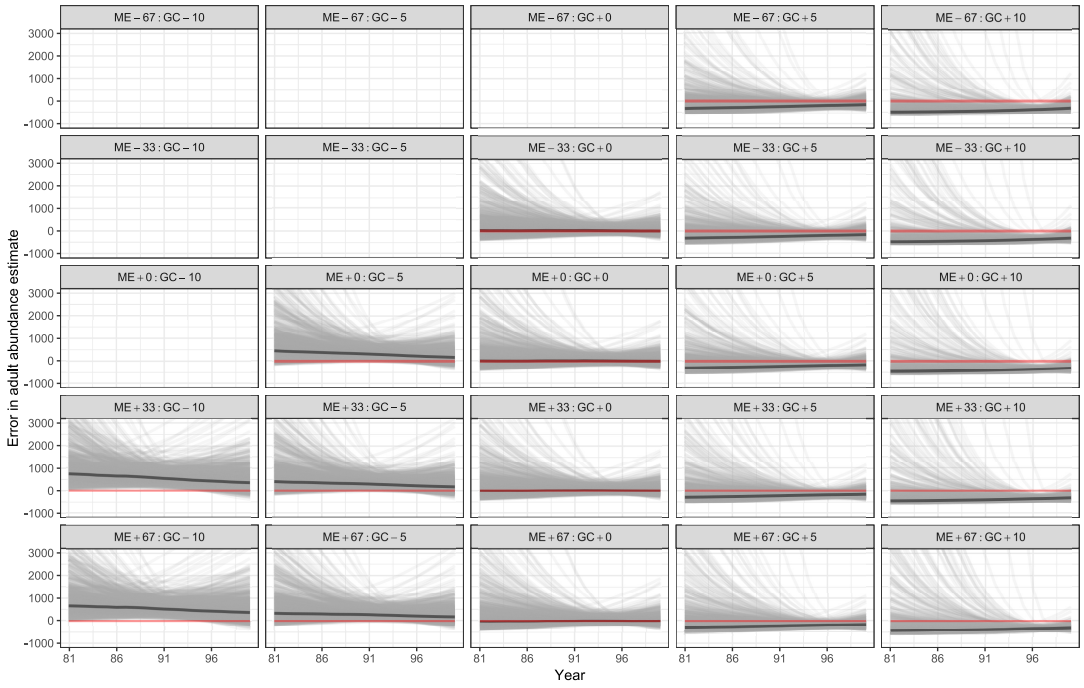


FIGURE B.3: Plots of the 1000 estimated female adult population trends for the complex species for the last twenty years of the simulation, for the 19 scenarios that resulted in successful fits. The median of these 1000 trends is indicated in dark grey, and the truth adult abundance is indicated in red. The scenarios were labelled using the format 'ME±XX:GC±YY', where ME refers to the measurement error, XX denotes the percentage over- or underestimate, GC stands for growth curve, and YY denotes the percentage of up- or downwards shifting; for example, the scenario with a 33% overestimated standard deviation of length measurement error and a 5% downshifted growth curve had label ME+33:GC-5.

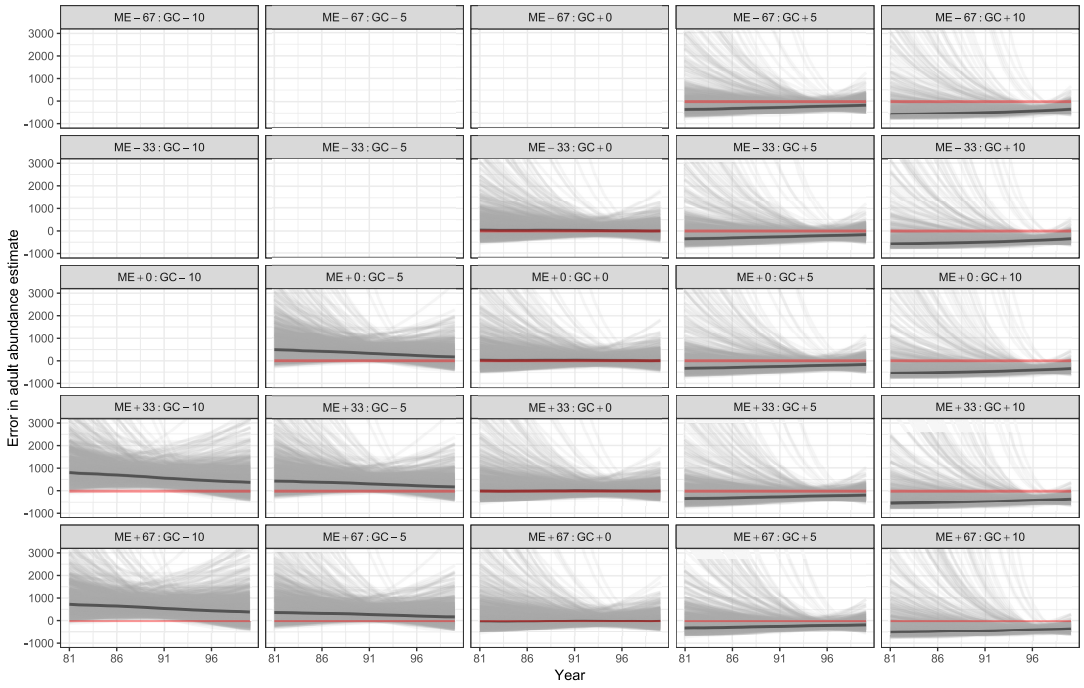


FIGURE B.4: Plots of the 1000 estimated male adult population trends for the complex species for the last twenty years of the simulation, for the 19 scenarios that resulted in successful fits. The median of these 1000 trends is indicated in dark grey, and the truth adult abundance is indicated in red. The scenarios were labelled using the format 'ME±XX:GC±YY', where ME refers to the measurement error, XX denotes the percentage over- or underestimate, GC stands for growth curve, and YY denotes the percentage of up- or downwards shifting; for example, the scenario with a 33% overestimated standard deviation of length measurement error and a 5% downshifted growth curve had label ME+33:GC-5.

Appendix C

Ethics approvals

Date submitted	16/06/2023
Secondary data	1
CREEM approve	1
Retrospective	1
Investigator	Len Thomas, Danielle Harris, Felix Petersma, Tiago Marques, Aaron Thode*, Gisela Cheoo*, Katherine Kim
Project title	Accommodating false positives within acoustic spatial capture-recapture, with variable source levels, noisy bearings and an inhomogeneous spatial density
C1 initiation	**1**
C1 justification	CREEM/USTAN staff or students did not motivate, initiate, or fund the data collection - data collected prior to conception of this desk-based study
C2 involvement	**1**
C2 justification	CREEM/USTAN staff or students were not involved in data collection - data collected prior to conception of this desk-based study
C3 impact	**1**
C3 justification	PAM study using stationary recorders, animals were not approached or followed for recording of the used data
C4 reputation	**1**
C4 justification	Results from data collection published elsewhere with ethics statement; approved monitoring program around activities conducted under incidental harassment authorizations (IHAs) issued by the U.S. National Marine Fisheries Service
Total criteria	**4**
Outcome	Approve
Outcome date	27/06/2023
Notes	Justification confirmed via email by applicant

FIGURE C.1: Secondary data approval for the work performed in relation to the research presented in Chapters 2 and 3.

University of St Andrews

School of Biology Ethics Committee

Project Title:	Estimating population density of cartilaginous fish populations
Researchers Name(s):	Felix Petersma and Prof Len Thomas
Supervisor(s):	Prof Len Thomas
Approval Date:	14-SEP-2022
Biology SEC Ref:	BL16504

Thank you for submitting your animal ethics application, which was considered by the School of Biology Ethics Committee.

The School of Biology Ethics Committee approves this study from an ethical point of view.

Approval is given for five years. Projects, which have not commenced within two years of original approval, must be re-submitted to the School Ethics Committee.

You must inform the School Ethics Committee when the research has been completed. If you are unable to complete your research within the five-year validation period, you will be required to write to the School Ethics Committee to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the condition that local permits are obtained prior to the project starting, and also on the understanding that the ASAB Guidelines for the Treatment of Animals in Behavioural Research and Teaching published in Animal Behaviour, 2003, 65, 249-255, are adhered to.

Yours sincerely,

Convenor of the School Ethics Committee

Ccs School Ethics Committee
(Home Office Liaison Officer)

SEC Convenor, Scottish Oceans Institute, St Andrews, Fife KY16 8LB, Scotland
Email: bioethics@st-andrews.ac.uk Tel: 01334 463554

FIGURE C.2: Ethics approval for the work performed in relation to the research presented in Chapter 4.

Bibliography

Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected papers of Hirotugu Akaike*, pages 199–213. Springer.

Baylis, S. M. (2022). fishSim: simulate populations with multiple stocks, movement, mating, mortality, and parentage.

Blackwell, S. B., McDonald, T. L., Kim, K. H., Aerts, L. A. M., Richardson, W. J., Greene Jr., C. R., and Streever, B. (2012). Directionality of bowhead whale calls measured with multiple sensors. *Marine Mammal Science*, 28(1):200–212.

Blackwell, S. B., Richardson, W., Greene, Jr., C., and Streever, B. (2007). Bowhead whale (*Balaena mysticetus*) migration and calling behaviour in the Alaskan Beaufort Sea, autumn 2001–04: an acoustic localization study. *ARCTIC*, 60(3):255–270.

Blackwell, S. B., Thode, A. M., Conrad, A. S., Ferguson, M. C., Berchok, C. L., Stafford, K. M., Marques, T. A., and Kim, K. H. (2021). Estimating acoustic cue rates in bowhead whales, *Balaena mysticetus*, during their fall migration through the Alaskan Beaufort Sea. *The Journal of the Acoustical Society of America*, 149(5):3611–3625.

Borchers, D. L., Buckland, S. T., and Zucchini, W. (2002). *Estimating Animal Abundance: Closed Populations*. Springer London, London, 1 edition.

Borchers, D. L. and Efford, M. G. (2008). Spatially explicit maximum likelihood methods for capture-recapture studies. *Biometrics*, 64(2):377–385.

Borchers, D. L., Stevenson, B. C., Kidney, D., Thomas, L., and Marques, T. A. (2015). A unifying model for capture–recapture and distance sampling surveys of

wildlife populations. *Journal of the American Statistical Association*, 110(509):195–204.

Bouchet, P. J. and Meeuwig, J. J. (2015). Drifting baited stereo-videography: a novel sampling tool for surveying pelagic wildlife in offshore marine reserves. *Ecosphere*, 6(8):1–29.

Braczkowski, A. R., Balme, G. A., Dickman, A., Fattebert, J., Johnson, P., Dickerson, T., Macdonald, D. W., and Hunter, L. (2016). Scent lure effect on camera-trap based leopard density estimates. *PLOS ONE*, 11(4):e0151033.

Bradford, R. W., Thomson, R., Bravington, M. V., Foote, D., Gunasekera, R., Bruce, B. D., Harasti, D., Otway, N., and Feutry, P. (2018). A close-kin mark-recapture estimate of the population size and trend of east coast grey nurse. Technical Report July, CSIRO, Hobart.

Bradley, D., Conklin, E., Papastamatiou, Y. P., McCauley, D. J., Pollock, K., Kendall, B. E., Gaines, S. D., and Caselle, J. E. (2017a). Growth and life history variability of the grey reef shark (*Carcharhinus amblyrhynchos*) across its range. *PLoS ONE*, 12(2):1–20.

Bradley, D., Conklin, E., Papastamatiou, Y. P., McCauley, D. J., Pollock, K., Pollock, A., Kendall, B. E., Gaines, S. D., and Caselle, J. E. (2017b). Resetting predator baselines in coral reef ecosystems. *Scientific Reports*, 7(February):1–9.

Bravington, M. V., Feutry, P., Pillans, R. D., Johnson, G., Saunders, T., Gunasekera, R., Bax, N. J., and Kyne, P. M. (2019). Close-kin mark-recapture population size estimate of *Glyphis garricki* in the Northern Territory. Technical Report November, CSIRO, Hobart.

Bravington, M. V., Grewe, P. M., and Davies, C. R. (2016a). Absolute abundance of southern bluefin tuna estimated by close-kin mark-recapture. *Nature Communications*, 7(1):1–8.

Bravington, M. V., Skaug, H. J., and Anderson, E. C. (2016b). Close-kin mark-recapture. *Statistical Science*, 31(2):259–274.

Brooks, E. J., Sloman, K. A., Sims, D. W., and Danylchuk, A. J. (2011). Validating the use of baited remote underwater video surveys for assessing the diversity,

distribution and abundance of sharks in the Bahamas. *Endangered Species Research*, 13(3):231–243.

Buchmann, S. and Ascher, J. S. (2005). The plight of pollinating bees. *Bee World*, 86(3):71–74.

Buckland, S. T., Borchers, D. L., Marques, T. A., and Fewster, R. M. (2023). Wildlife population assessment: changing priorities driven by technological advances. *Journal of Statistical Theory and Practice*, 17(2):1–22.

Buckland, S. T., Rexstad, E. A., Marques, T. A., and Oedekoven, C. S. (2015). *Distance Sampling: Methods and Applications*. Methods in Statistical Ecology. Springer International Publishing, Cham.

Burgar, J. M., Stewart, F. E., Volpe, J. P., Fisher, J. T., and Burton, A. C. (2018). Estimating density for species conservation: comparing camera trap spatial count models to genetic spatial capture-recapture models. *Global Ecology and Conservation*, 15:1–10.

Burke, P. J., Raoult, V., Natanson, L. J., Murphy, T. D., Peddemors, V., and Williamson, J. E. (2020). Struggling with age: common sawsharks (*Pristiophorus cirratus*) defy age determination using a range of traditional methods. *Fisheries Research*, 231(August):105706.

Burt, M. L., Borchers, D. L., Jenkins, K. J., and Marques, T. A. (2014). Using mark-recapture distance sampling methods on line transect surveys. *Methods in Ecology and Evolution*, 5(11):1180–1191.

Burton, A. C., Neilson, E., Moreira, D., Ladle, A., Steenweg, R., Fisher, J. T., Bayne, E., and Boutin, S. (2015). Wildlife camera trapping: a review and recommendations for linking surveys to ecological processes. *Journal of Applied Ecology*, 52(3):675–685.

Campana, S. E. (2001). Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation methods. *Journal of Fish Biology*, 59(2):197–242.

Campos-Candela, A., Palmer, M., Balle, S., and Alós, J. (2018). A camera-based method for estimating absolute density in animals displaying home range behaviour. *Journal of Animal Ecology*, 87(3):825–837.

Caravaggi, A., Zaccaroni, M., Riga, F., Schai-Braun, S. C., Dick, J. T., Montgomery, W. I., and Reid, N. (2016). An invasive-native mammalian species replacement process captured by camera trap survey random encounter models. *Remote Sensing in Ecology and Conservation*, 2(1):45–58.

Casas, L. and Saborido-Rey, F. (2023). A review of an emerging tool to estimate population parameters: the close-kin mark-recapture method. *Frontiers in Marine Science*, 10(July):1–18.

Caswell, H. (2006). *Matrix Population Models*. Sinaeur, 2nd edition.

Chandler, R. B. and Royle, J. A. (2013). Spatially explicit models for inference about density in unmarked or partially marked populations. *The Annals of Applied Statistics*, 7(2):936–954.

Cheoo, G. V. (2019). *Estimation of bowhead whale (Balaena mysticetus) population density using spatially explicit capture-recapture (SECR) methods*. PhD thesis, Universidade De Lisboa.

Collins, M. A., Yau, C., Guilfoyle, F., Bagley, P., Everson, I., Priede, I. G., and Agnew, D. (2002). Assessment of stone crab (*Lithodidae*) density on the South Georgia slope using baited video cameras. *ICES Journal of Marine Science*, 59(2):370–379.

Conn, P. B., Bravington, M. V., Baylis, S., and Ver Hoef, J. M. (2020). Robustness of close-kin mark–recapture estimators to dispersal limitation and spatially varying sampling probabilities. *Ecology and Evolution*, 10(12):5558–5569.

Conn, P. B., Gorgone, A. M., Jugovich, A. R., Byrd, B. L., and Hansen, L. J. (2011). Accounting for transients when estimating abundance of bottlenose dolphins in Choctawhatchee Bay, Florida. *Journal of Wildlife Management*, 75(3):569–579.

Conn, P. B., Ver Hoef, J. M., Mcclintock, B. T., Moreland, E. E., London, J. M., Cameron, M. F., Dahle, S. P., and Boveng, P. L. (2014). Estimating multispecies abundance using automated detection systems: ice-associated seals in the Bering Sea. *Methods in Ecology and Evolution*, 5(12):1280–1293.

Constantine, R., Jackson, J. A., Steel, D., Baker, C. S., Brooks, L., Burns, D., Clapham, P., Hauser, N., Madon, B., Mattila, D., Oremus, M., Poole, M., Robbins,

J., Thompson, K., and Garrigue, C. (2012). Abundance of humpback whales in Oceania using photo-identification and microsatellite genotyping. *Marine Ecology Progress Series*, 453:249–261.

Cormack, R. M. (1964). Estimates of survival from the sighting of marked animals. *Biometrika*, 51(3/4):429.

Creel, S., Spong, G., Sands, J. L., Rotella, J., Zeigle, J., Joe, L., Murphy, K. M., and Smith, D. (2003). Population size estimation in Yellowstone wolves with error-prone noninvasive microsatellite genotypes. *Molecular Ecology*, 12(7):2003–2009.

Dahlsten, D. L., Six, D. L., Rowney, D. L., Lawson, A. B., Erbilgin, N., and Raffa, K. F. (2004). Attraction of *Ips pini* (Coleoptera: Scolytinae) and its predators to natural attractants and synthetic semiochemicals in Northern California: implications for population monitoring. *Environmental Entomology*, 33(6):1554–1561.

Dawson, D. K. and Efford, M. G. (2009). Bird population density estimated from acoustic signals. *Journal of Applied Ecology*, 46(6):1201–1209.

De Paoli-Iseppi, R., Deagle, B. E., McMahon, C. R., Hindell, M. A., Dickinson, J. L., and Jarman, S. N. (2017). Measuring animal age with DNA methylation: from humans to wild animals. *Frontiers in Genetics*, 8:2010–2017.

Delaval, A., Bendall, V., Hetherington, S. J., Skaug, H. J., Frost, M., Jones, C. S., and Noble, L. R. (2023). Evaluating the suitability of close-kin mark-recapture as a demographic modelling tool for a critically endangered elasmobranch population. *Evolutionary Applications*, 16(2):461–473.

DiRenzo, G. V., Hanks, E., and Miller, D. A. (2023). A practical guide to understanding and validating complex models using data simulations. *Methods in Ecology and Evolution*, 14(1):203–217.

Dunlop, K., Ruxton, G., Scott, E., and Bailey, D. (2015). Absolute abundance estimates from shallow water baited underwater camera surveys; a stochastic modelling approach tested against field data. *Journal of Experimental Marine Biology and Ecology*, 472:126–134.

Durmaz, A. A., Karaca, E., Demkow, U., Toruner, G., Schoumans, J., and Cogulu, O. (2015). Evolution of genetic techniques: past, present, and beyond. *BioMed Research International*, 2015:1–7.

Eddelbuettel, D. (2013). *Seamless R and C++ Integration with Rcpp*. Springer New York, New York, NY.

Efford, M. (2004). Density estimation in live-trapping studies. *Oikos*, 106(3):598–610.

Efford, M. G., Dawson, D. K., and Borchers, D. L. (2009). Population density estimated from locations of individuals on a passive detector array. *Ecology*, 90(10):2676–2682.

Ferretti, F., Curnick, D., Liu, K., Romanov, E. V., and Block, B. A. (2018). Shark baselines and the conservation role of remote coral reef ecosystems. *Science Advances*, 4(3):1–13.

Fewster, R. M., Stevenson, B. C., and Borchers, D. L. (2016). Trace-contrast models for capture–recapture without capture histories. *Statistical Science*, 31(2):245–258.

Flowerdew, J. R., Shore, R. F., Poulton, S. M. C., and Spars, T. H. (2004). Live trapping to monitor small mammals in Britain. *Mammal Review*, 34(1-2):31–50.

Foster, R. J. and Harmsen, B. J. (2012). A critique of density estimation from camera-trap data. *Journal of Wildlife Management*, 76(2):224–236.

Fournier, D. and Archibald, C. P. (1982). A general theory for analyzing catch at age data. *Canadian Journal of Fisheries and Aquatic Sciences*, 39(8):1195–1207.

Francis, R. (1988). Maximum likelihood estimation of growth and growth variability from tagging data. *New Zealand Journal of Marine and Freshwater Research*, 22(1):43–51.

Garvey, R. S. and Fowler, A. C. (2023). On the mathematical theory of plumes. *Geophysical & Astrophysical Fluid Dynamics*, 117(2):79–106.

George, J. C., Clark, C., And, G. M. C., and Ellison, W. T. (1989). Observations on the ice-breaking and ice navigation behavior of migrating bowhead whales (*Balaena mysticetus*) near Point Barrow, Alaska, spring 1985. *Arctic*, 42(1):24–30.

Gilbert, N. A., Clare, J. D. J., Stenglein, J. L., and Zuckerberg, B. (2021). Abundance estimation of unmarked animals based on camera-trap data. *Conservation Biology*, 35(1):88–100.

Graham, N. A. J., Spalding, M. D., and Sheppard, C. R. C. (2010). Reef shark declines in remote atolls highlight the need for multi-faceted conservation action. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 20(5):543–548.

Greene, C. R., McLennan, M. W., Norman, R. G., McDonald, T. L., Jakubczak, R. S., and Richardson, W. J. (2004). Directional frequency and recording (DIFAR) sensors in seafloor recorders to locate calling bowhead whales during their fall migration. *The Journal of the Acoustical Society of America*, 116(2):799–813.

Hartl, D. L. (2020). *A Primer of Population Genetics and Genomics*. Oxford University Press.

Harvey, E. S., Cappo, M., Butler, J. J., Hall, N., and Kendrick, G. A. (2007). Bait attraction affects the performance of remote underwater video stations in assessment of demersal fish community structure. *Marine Ecology Progress Series*, 350:245–254.

Harvey, E. S., Newman, S. J., McLean, D. L., Cappo, M., Meeuwig, J. J., and Skepper, C. L. (2012). Comparison of the relative efficiencies of stereo-BRUVs and traps for sampling tropical continental shelf demersal fishes. *Fisheries Research*, 125-126:108–120.

Heagney, E., Lynch, T., Babcock, R., and Suthers, I. (2007). Pelagic fish assemblages assessed using mid-water baited video: standardising fish counts using bait plume size. *Marine Ecology Progress Series*, 350:255–266.

Hildebrand, J. A., Baumann-Pickering, S., Frasier, K. E., Trickey, J. S., Merkens, K. P., Wiggins, S. M., McDonald, M. A., Garrison, L. P., Harris, D., Marques, T. A., and Thomas, L. (2015). Passive acoustic monitoring of beaked whale densities in the Gulf of Mexico. *Scientific Reports*, 5(1):16343.

Hillary, R. M., Bravington, M. V., Patterson, T. A., Grewe, P., Bradford, R., Feutry, P., Gunasekera, R., Peddemors, V., Werry, J., Francis, M. P., Duffy, C. A. J., and Bruce, B. D. (2018). Genetic relatedness reveals total population size of white sharks in eastern Australia and New Zealand. *Scientific Reports*, 8(1):2661.

Hohn, A. A. (2009). Age Estimation. In *Encyclopedia of Marine Mammals*, chapter 9, pages 11–17. Elsevier.

Howe, E. J., Buckland, S. T., Després-Einspenner, M., and Kühl, H. S. (2017). Distance sampling with camera traps. *Methods in Ecology and Evolution*, 8(11):1558–1565.

Hulson, P.-J. F. and Williams, B. C. (2024). Inclusion of ageing error and growth variability using a bootstrap estimation of age composition and conditional age-at-length input sample size for fisheries stock assessment models. *Fisheries Research*, 270(June 2023):106894.

Hutchinson, J. M. and Waser, P. M. (2007). Use, misuse and extensions of "ideal gas" models of animal encounter. *Biological Reviews*, 82(3):335–359.

IUCN (2012). Guidelines for Application of IUCN Red List Criteria at Regional and National Levels: Version 4.0. Technical report, IUCN, Gland, Switzerland.

Jacinto, M. R., Songcuan, A. J. G., Yip, G. V., and Santos, M. D. (2015). Development and application of the fisheries vulnerability assessment tool (Fish Vool) to tuna and sardine sectors in the Philippines. *Fisheries Research*, 161:174–181.

Jackson, R. M., Roe, J. D., Wangchuk, R., and Hunter, D. O. (2006). Estimating snow leopard population abundance using photography and capture–recapture techniques. *Wildlife Society Bulletin*, 34(3):772–781.

Janoschek, A. (1957). Das reaktionskinetische Grundgesetz und seine Beziehungen zum Wachstums-und Ertragsgesetz. *Statistische Vierteljahresschrift*, 10:25–37.

Jensen, F. B., Kuperman, W. A., Porter, M. B., and Schmidt, H. (2011). *Computational Ocean Acoustics*. Springer New York, New York, NY.

Jepsen, N., Schreck, C., Clements, S., and Thorstad, E. B. (2005). A brief discussion on the 2% tag/bodymass rule of thumb. In Marmulla, M. S., Lembo, G., and G., editors, *Aquatic telemetry: advances and applications*, number June 2003, page 295. Food and Agriculture Organization of the United Nations.

Jiménez, J., Cara, D., García-Dominguez, F., and Barasona, J. A. (2023). Estimating wolf (*Canis lupus*) densities using video camera traps and spatial capture–recapture analysis. *Ecosphere*, 14(7):1–14.

Jolly, A. G. M. (1965). Explicit estimates from capture-recapture data with both death and immigration-stochastic model. *Biometrika*, 52(1):225–247.

Jourdain, N. O. A. S., Cole, D. J., Ridout, M. S., and Rowcliffe, J. M. (2020). Statistical sevelopment of animal density estimation using random encounter modelling. *Journal of Agricultural, Biological and Environmental Statistics*, 25(2):148–167.

Karanth, K. U. and Nichols, J. D. (1998). Estimation of tiger densities in India using photographic capture and recaptures. *Ecology*, 79(8):2852–2862.

Kidney, D., Rawson, B. M., Borchers, D. L., Stevenson, B. C., Marques, T. A., and Thomas, L. (2016). An efficient acoustic density estimation method with human detectors applied to gibbons in Cambodia. *PLoS ONE*, 11(5).

Kilfoil, J. P., Campbell, M. D., Heithaus, M. R., and Zhang, Y. (2021). The influence of shark behavior and environmental conditions on baited remote underwater video survey results. *Ecological Modelling*, 447(February):109507.

Klavitter, J. L., Marzluff, J. M., and Vekasy, M. S. (2003). Abundance and demography of the Hawaiian hawk: is delisting warranted? *The Journal of Wildlife Management*, 67(1):165–176.

Kodi, A. R., Howard, J., Borchers, D. L., Worthington, H., Alexander, J. S., Lkhagvajav, P., Bayandonoi, G., Ochirjav, M., Erdenebaatar, S., Byambasuren, C., Battulga, N., Johansson, Ö., and Sharma, K. (2024). Ghostbusting—Reducing bias due to identification errors in spatial capture-recapture histories. *Methods in Ecology and Evolution*, 00(March):1–11.

Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., and Bell, B. M. (2016). TMB: automatic differentiation and laplace approximation. *Journal of Statistical Software*, 70(5).

Küsel, E. T., Mellinger, D. K., Thomas, L., Marques, T. A., Moretti, D., and Ward, J. (2011). Cetacean population density estimation from single fixed sensors using

passive acoustics. *The Journal of the Acoustical Society of America*, 129(6):3610–3622.

Langlois, T., Goetze, J., Bond, T., Monk, J., Abesamis, R. A., Asher, J., Barrett, N., Bernard, A. T. F., Bouchet, P. J., Birt, M. J., Cappo, M., Currey-Randall, L. M., Driessen, D., Fairclough, D. V., Fullwood, L. A. F., Gibbons, B. A., Harasti, D., Heupel, M. R., Hicks, J., Holmes, T. H., Huveneers, C., Ierodiaconou, D., Jordan, A., Knott, N. A., Lindfield, S., Malcolm, H. A., McLean, D., Meekan, M., Miller, D., Mitchell, P. J., Newman, S. J., Radford, B., Rolim, F. A., Saunders, B. J., Stowar, M., Smith, A. N. H., Travers, M. J., Wakefield, C. B., Whitmarsh, S. K., Williams, J., and Harvey, E. S. (2020). A field and video annotation guide for baited remote underwater stereo-video surveys of demersal fish assemblages. *Methods in Ecology and Evolution*, 11(11):1401–1409.

Larison, B., Pinho, G. M., Haghani, A., Zoller, J. A., Li, C. Z., Finno, C. J., Farrell, C., Kaelin, C. B., Barsh, G. S., Wooding, B., Robeck, T. R., Maddox, D., Pellegrini, M., and Horvath, S. (2021). Epigenetic models developed for plains zebras predict age in domestic horses and endangered equids. *Communications Biology*, 4(1):1412.

Letessier, T. B., Mouillot, D., Bouchet, P. J., Vigliola, L., Fernandes, M. C., Thompson, C., Boussarie, G., Turner, J., Juhel, J.-B., Maire, E., Caley, M. J., Koldewey, H. J., Friedlander, A., Sala, E., and Meeuwig, J. J. (2019). Remote reefs and seamounts are the last refuges for marine predators across the Indo-Pacific. *PLOS Biology*, 17(9):1–20.

Link, W. A., Schofield, M. R., Barker, R. J., and Sauer, J. R. (2018). On the robustness of N-mixture models. *Ecology*, 99(7):1547–1551.

Ljungblad, D. K., Thompson, P. O., and Moore, S. E. (1982). Underwater sounds recorded from migrating bowhead whales, *Balaena mysticetus*, in 1979. *The Journal of the Acoustical Society of America*, 71(2):477–482.

Lloyd-Jones, L. R., Bravington, M. V., Armstrong, K. N., Lawrence, E., Feutry, P., Todd, C. M., Dorrestein, A., Welbergen, J. A., Martin, J. M., Rose, K., Hall, J., Phalen, D. N., Peters, I., Baylis, S. M., Macgregor, N. A., and Westcott, D. A. (2023). Close-kin mark-recapture informs critically endangered terrestrial mammal status. *Scientific Reports*, 13(1):1–16.

Lucas, T. C. D., Moorcroft, E. A., Freeman, R., Rowcliffe, J. M., and Jones, K. E. (2015). A generalised random encounter model for estimating animal density with remote sensor data. *Methods in Ecology and Evolution*, 6(5):500–509.

Lukacs, P. M. and Burnham, K. P. (2005). Review of capture-recapture methods applicable to noninvasive genetic sampling. *Molecular Ecology*, 14(13):3909–3919.

MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Royle, A. A., and Langtimm, C. A. (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, 83(8):2248–2255.

Mallet, D. and Pelletier, D. (2014). Underwater video techniques for observing coastal marine biodiversity: a review of sixty years of publications (1952–2012). *Fisheries Research*, 154(February 2022):44–62.

Marques, T., Munger, L., Thomas, L., Wiggins, S., and Hildebrand, J. (2011). Estimating North Pacific right whale *Eubalaena japonica* density using passive acoustic cue counting. *Endangered Species Research*, 13(3):163–172.

Marques, T. A., Thomas, L., Martin, S. W., Mellinger, D. K., Ward, J. A., Moretti, D. J., Harris, D., and Tyack, P. L. (2013). Estimating animal population density using passive acoustics. *Biological Reviews*, 88(2):287–309.

Mayne, B., Berry, O., and Jarman, S. (2023). Calibrating epigenetic clocks with training data error. *Evolutionary Applications*, 16(8):1496–1502.

McClintock, B. T., White, G. C., Antolin, M. F., and Tripp, D. W. (2009). Estimating abundance using mark-resight when sampling is with replacement or the number of marked individuals is unknown. *Biometrics*, 65(1):237–246.

McCoy, E., Burce, R., David, D., Aca, E. Q., Hardy, J., Labaja, J., Snow, S. J., Ponzo, A., and Araujo, G. (2018). Long-term photo-identification reveals the population dynamics and strong site fidelity of adult whale sharks to the coastal waters of Donsol, Philippines. *Frontiers in Marine Science*, 5:1–11.

McNeil, J. N. (1991). Behavioral ecology of pheromone-mediated communication in moths and its importance in the use of pheromone traps. *Annual Review of Entomology*, 36(1):407–430.

Mills, L. S., Citta, J. J., Lair, K. P., Schwartz, M. K., and Tallmon, D. A. (2000). Estimating animal abundance using noninvasive DNA sampling: Promise and pitfalls. *Ecological Applications*, 10(1):283–294.

Morris, T. P., White, I. R., and Crowther, M. J. (2019). Using simulation studies to evaluate statistical methods. *Statistics in Medicine*, 38(11):2074–2102.

Mowat, G. and Paetkau, D. (2002). Estimating marten *Martes americana* population size using hair capture and genetic tagging. *Wildlife Biology*, 8(3):201–209.

Nakashima, Y., Fukasawa, K., and Samejima, H. (2018). Estimating animal density without individual recognition using information derivable exclusively from camera traps. *Journal of Applied Ecology*, 55(2):735–744.

Nichols, J. D. and Williams, B. K. (2006). Monitoring for conservation. *Trends in Ecology & Evolution*, 21(12):668–673.

Otis, D. L., Burnham, K. P., White, G. C., and Anderson, D. R. (1978). Statistical inference from capture data on closed animal populations. *Wildlife Monographs*, 62:3–135.

Papastamatiou, Y. P., Bodey, T. W., Friedlander, A. M., Lowe, C. G., Bradley, D., Weng, K., Priestley, V., and Caselle, J. E. (2018). Spatial separation without territoriality in shark communities. *Oikos*, 127(6):767–779.

Parmenter, R. R., Yates, T. L., Anderson, D. R., Burnham, K. P., Dunnum, J. L., Franklin, A. B., Friggens, M. T., Lubow, B. C., Miller, M., Olson, G. S., Parmenter, C. A., Pollard, J., Rexstad, E., Shenk, T. M., Stanley, T. R., and White, G. C. (2003). Small-mammal density estimation: a field comparison of grid-based vs. web-based density estimators. *Ecological Monographs*, 73(1):1–26.

Patterson, T. A., Hillary, R., Feutry, P., Gunasekera, R., Marthick, J., and Pillans, R. D. (2022). Rapid estimation of cryptic adult abundance and breeding dynamics in a critically endangered elasmobranch from close-kin mark recapture. *bioRxiv Ecology*, pages 1–34.

Phillips, G. T. (2016). *Passive acoustics: a multifaceted tool for marine mammal conservation*. PhD thesis, Duke University.

Pledger, S., Pollock, K. H., and Norris, J. L. (2010). Open capture-recapture models with heterogeneity: II. Jolly-Seber model. *Biometrics*, 66(3):883–890.

Polanowski, A. M., Robbins, J., Chandler, D., and Jarman, S. N. (2014). Epigenetic estimation of age in humpback whales. *Molecular Ecology Resources*, 14(5):976–987.

Pollock, K. H. (1982). A capture-recapture design robust to unequal probability of capture. *The Journal of Wildlife Management*, 46(3):752.

Poutanen, J., Pusenius, J., Wikström, M., and Brommer, J. E. (2019). Estimating population density of the white-tailed deer in Finland using non-invasive genetic sampling and spatial capture–recapture. *Annales Zoologici Fennici*, 56(1-6):1.

Prado, N. A., Brown, J. L., Zoller, J. A., Haghani, A., Yao, M., Bagryanova, L. R., Campana, M. G., E. Maldonado, J., Raj, K., Schmitt, D., Robeck, T. R., and Horvath, S. (2021). Epigenetic clock and methylation studies in elephants. *Aging Cell*, 20(7):1–11.

Priede, I., Bagley, P., Smith, A., Creasey, S., and Merrett, N. (1994). Scavenging deep demersal fishes of the Porcupine Seabight, north-east Atlantic: observations by baited camera, trap and trawl. *Journal of the Marine Biological Association of the United Kingdom*, 74(3):481–498.

Priede, I. G. and Bagley, P. M. (2000). In situ studies on deep-sea demersal fishes using autonomous unmanned lander platforms. *Oceanography and Marine Biology*, 38:357–392.

Priede, I. G. and Merrett, N. R. (1996). Estimation of abundance of abyssal demersal fishes; a comparison of data from trawls and baited cameras. *Journal of Fish Biology*, 49(sa):207–216.

Priede, I. G., Smith, K. L., and Armstrong, J. D. (1990). Foraging behavior of abyssal grenadier fish: inferences from acoustic tagging and tracking in the North Pacific Ocean. *Deep Sea Research Part A, Oceanographic Research Papers*, 37(1):81–101.

Prystupa, S., McCracken, G. R., Perry, R., and Ruzzante, D. E. (2021). Population abundance in arctic grayling using genetics and close-kin mark-recapture. *Ecology and Evolution*, 11(9):4763–4773.

R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Reid, J. A., Horn, R. B., and Forsman, E. D. (1999). Detection rates of spotted owls based on acoustic-lure and live-lure surveys. *Wildlife Society Bulletin*, 27(4):986–990.

Richards, L. J., Schnute, J. T., Kronlund, A. R., and Beamish, R. J. (1992). Statistical models for the analysis of ageing error. *Canadian Journal of Fisheries and Aquatic Sciences*, 49(9):1801–1815.

Roman, J., Estes, J. A., Morissette, L., Smith, C., Costa, D., McCarthy, J., Nation, J., Nicol, S., Pershing, A., and Smetacek, V. (2014). Whales as marine ecosystem engineers. *Frontiers in Ecology and the Environment*, 12(7):377–385.

Rovero, F., Zimmermann, F., Berzi, D., and Meek, P. (2013). "Which camera trap type and how many do I need?" A review of camera features and study designs for a range of wildlife research applications. *Hystrix*, 24(2):148–156.

Rowcliffe, J. M. and Carbone, C. (2008). Surveys using camera traps: are we looking to a brighter future? *Animal Conservation*, 11(3):185–186.

Rowcliffe, J. M., Carbone, C., Kays, R., Kranstauber, B., and Jansen, P. A. (2012). Bias in estimating animal travel distance: the effect of sampling frequency. *Methods in Ecology and Evolution*, 3(4):653–662.

Rowcliffe, J. M., Field, J., Turvey, S. T., and Carbone, C. (2008). Estimating animal density using camera traps without the need for individual recognition. *Journal of Applied Ecology*, 45(4):1228–1236.

Rowcliffe, J. M., Jansen, P. A., Kays, R., Kranstauber, B., and Carbone, C. (2016). Wildlife speed cameras: measuring animal travel speed and day range using camera traps. *Remote Sensing in Ecology and Conservation*, 2(2):84–94.

Rowcliffe, J. M., Kays, R., Carbone, C., and Jansen, P. A. (2013). Clarifying assumptions behind the estimation of animal density from camera trap rates. *Journal of Wildlife Management*, 77(5):876.

Roy, D. (2003). The discrete normal distribution. *Communications in Statistics - Theory and Methods*, 32(10):1871–1883.

Royle, J. A. (2004). N-mixture models for estimating population size from spatially replicated counts. *Biometrics*, 60(1):108–115.

Royle, J. A. (2018). Modelling sound attenuation in heterogeneous environments for improved bioacoustic sampling of wildlife populations. *Methods in Ecology and Evolution*, 9(9):1939–1947.

Royle, J. A., Nichols, J. D., Karanth, K. U., and Gopalaswamy, A. M. (2009). A hierarchical model for estimating density in camera-trap studies. *Journal of Applied Ecology*, 46(1):118–127.

Royle, J. A. and Young, K. V. (2008). A hierarchical model for spatial capture-recapture data. *Ecology*, 89(8):2281–2289.

Ruzzante, D. E., McCracken, G. R., Førland, B., MacMillan, J., Notte, D., Buhariwalla, C., Mills Flemming, J., and Skaug, H. (2019). Validation of close-kin mark–recapture (CKMR) methods for estimating population abundance. *Methods in Ecology and Evolution*, 10(9):1445–1453.

Schobernd, Z. H., Bacheler, N. M., and Conn, P. B. (2014). Examining the utility of alternative video monitoring metrics for indexing reef fish abundance. *Canadian Journal of Fisheries and Aquatic Sciences*, 71(3):464–471.

Schofield, G., Katselidis, K. A., Dimopoulos, P., and Pantis, J. D. (2008). Investigating the viability of photo-identification as an objective tool to study endangered sea turtle populations. *Journal of Experimental Marine Biology and Ecology*, 360(2):103–108.

Schwarz, C. J. (2001). The Jolly-Seber model: more than just abundance. *Journal of Agricultural, Biological, and Environmental Statistics*, 6(2):195–205.

Schwarz, C. J. and Arnason, A. N. (1996). A general methodology for the analysis of capture-recapture experiments in open populations. *Biometrics*, 52(3):860–873.

Seber, A. G. A. F. (1965). A note on the multiple-recapture census. *Biometrika*, 52(1):249–259.

Seber, G. A. F. (1982). *The Estimation of Animal Abundance: And Related Parameters*. Macmillan Pub. Co., New York, 2nd ed edition.

Sévêque, A., Lonsinger, R. C., Waits, L. P., Brzeski, K. E., Komoroske, L. M., Ott-Conn, C. N., Mayhew, S. L., Norton, D. C., Petroelje, T. R., Swenson, J. D., and Morin, D. J. (2024). Sources of bias in applying close-kin mark–recapture to terrestrial game species with different life histories. *Ecology*, 105(3):1–16.

Sharma, Y., Bennett, J. B., Rašić, G., and Marshall, J. M. (2022). Close-kin mark-recapture methods to estimate demographic parameters of mosquitoes. *PLOS Computational Biology*, 18(12).

Shelden, K. E. and Rugh, D. J. (1995). The bowhead whale, *Balaena mysticetus*: its historic and current status. *Marine Fisheries Review*, 57(3-4):1–20.

Sherman, C. S., Chin, A., Heupel, M. R., and Simpfendorfer, C. A. (2018). Are we underestimating elasmobranch abundances on baited remote underwater video systems (BRUVS) using traditional metrics? *Journal of Experimental Marine Biology and Ecology*, 503(September 2017):80–85.

Skaug, H. J. (2001). Allele-sharing methods for estimation of population size. *Biometrics*, 57(3):750–756.

Skaug, H. J. (2017). The parent–offspring probability when sampling age-structured populations. *Theoretical Population Biology*, 118:20–26.

Speed, C., Meekan, M., Field, I., McMahon, C., Stevens, J., McGregor, F., Huveneers, C., Berger, Y., and Bradshaw, C. (2011). Spatial and temporal movement patterns of a multi-species coastal reef shark aggregation. *Marine Ecology Progress Series*, 429:261–275.

Stevenson, B. C., Borchers, D. L., Altwegg, R., Swift, R. J., Gillespie, D. M., and Measey, G. J. (2015). A general framework for animal density estimation from acoustic detections across a fixed microphone array. *Methods in Ecology and Evolution*, 6(1):38–48.

Stevenson, B. C., Borchers, D. L., and Fewster, R. M. (2019). Cluster capture-recapture to account for identification uncertainty on aerial surveys of animal populations. *Biometrics*, 75(1):326–336.

Stevenson, B. C., Dam-Bates, P., Young, C. K. Y., and Measey, J. (2021). A spatial capture–recapture model to estimate call rate and population density from passive acoustic surveys. *Methods in Ecology and Evolution*, 12(3):432–442.

Sugai, L. S. M., Silva, T. S. F., Ribeiro, J. W., and Llusia, D. (2019). Terrestrial passive acoustic monitoring: review and perspectives. *BioScience*, 69(1):15–25.

Swenson, J. D., Brooks, E. N., Kacev, D., Boyd, C., Kinney, M. J., Marcy-Quay, B., Sévêque, A., Feldheim, K. A., and Komoroske, L. M. (2024). Accounting for unobserved population dynamics and aging error in close-kin mark-recapture assessments. *Ecology and Evolution*, 14(2):1–24.

Thode, A. M., Blackwell, S. B., Conrad, A. S., Kim, K. H., Marques, T., Thomas, L., Oedekoven, C. S., Harris, D., and Bröker, K. C. (2020). Roaring and repetition: how bowhead whales adjust their call density and source level (Lombard effect) in the presence of natural and seismic airgun survey noise. *The Journal of the Acoustical Society of America*, 147(3):2061–2080.

Thode, A. M., Kim, K. H., Blackwell, S. B., Greene, C. R., Nations, C. S., McDonald, T. L., and Macrander, A. M. (2012). Automated detection and localization of bowhead whale sounds in the presence of seismic airgun surveys. *The Journal of the Acoustical Society of America*, 131(5):3726–3747.

Thomas, L., Jaramillo-Legorreta, A., Cardenas-Hinojosa, G., Nieto-Garcia, E., Rojas-Bracho, L., Ver Hoef, J. M., Moore, J., Taylor, B., Barlow, J., and Tregenza, N. (2017). Last call: passive acoustic monitoring shows continued rapid decline of critically endangered vaquita. *The Journal of the Acoustical Society of America*, 142(5):512–517.

Tourani, M. (2022). A review of spatial capture–recapture: ecological insights, limitations, and prospects. *Ecology and Evolution*, 12(1):1–13.

Trenkel, V. M., Charrier, G., Lorance, P., and Bravington, M. V. (2022). Close-kin mark–recapture abundance estimation: practical insights and lessons learned. *ICES Journal of Marine Science*, 79(2):413–422.

Tsukahara, Y., Nakamichi, R., Matsuura, A., and Akita, T. (2023). Comparison of kinship-identification methods for robust stock assessment using close-kin mark-recapture data for Pacific bluefin tuna. *Authorea Preprints*.

Vasilakopoulos, P., Jardim, E., Konrad, C., Rihan, D., Mannini, A., Pinto, C., Casey, J., Mosqueira, I., and O'Neill, F. G. (2020). Selectivity metrics for fisheries management and advice. *Fish and Fisheries*, 21(3):621–638.

von Bertalanffy, L. (1938). A quantitative theory of organic growth. *Human Biology*, 10(2):181–213.

Wacker, S., Skaug, H. J., Forseth, T., Solem, Ø., Ulvan, E. M., Fiske, P., and Karlsson, S. (2021). Considering sampling bias in close-kin mark–recapture abundance estimates of Atlantic salmon. *Ecology and Evolution*, 11(9):3917–3932.

Waples, R. S. and Feutry, P. (2022). Close-kin methods to estimate census size and effective population size. *Fish and Fisheries*, 23(2):273–293.

Westerberg, H. and Westerberg, K. (2011). Properties of odour plumes from natural baits. *Fisheries Research*, 110(3):459–464.

Wetterstrand, K. A. (2023). DNA sequencing costs: Data from the NHGRI Genome Sequencing Program (GSP). `https://www.genome.gov/sequencingcostsdata`. Last accessed on May 09, 2023.

White, C. R., Cassey, P., Schimpf, N. G., Halsey, L. G., Green, J. A., and Portugal, S. J. (2012). Implantation reduces the negative effects of bio-logging devices on birds. *Journal of Experimental Biology*, 216(4):537–542.

Wilhelm, T. A., Sheppard, C. R. C., Sheppard, A. L. S., Gaymer, C. F., Parks, J., Wagner, D., and Lewis, N. (2014). Large marine protected areas—advantages and challenges of going big. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 24:24–30.

Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC, 2nd edition.

Woods, J., Paetkau, D., Lewis, D., McLellan, B., Proctor, M., and Strobeck, C. (1999). Genetic tagging of free-ranging black and brown bears. *Wildlife Society Bulletin*, 27(3):616–627.

Wursig, B., Dorset, E. M., Fraker, M. A., Payne, R. S., and Richardson, W. J. (1985). Behavior of bowhead whales, *Balaena mysticetus*, summering in the Beaufort Sea: a description. *Fishery Bulletin*, 83(3):357–377.

Yack, T. M., Barlow, J., Calambokidis, J., Southall, B., and Coates, S. (2013). Passive acoustic monitoring using a towed hydrophone array results in

identification of a previously unknown beaked whale habitat. *The Journal of the Acoustical Society of America*, 134(3):2589–2595.

Yoshizaki, J. (2007). *Use of natural tags in closed population capture-recapture studies: modeling misidentification*. PhD thesis, North Carolina State University.

Zimmer, W. (2011). *Passive Acoustic Monitoring of Cetaceans*. Cambridge University Press, Cambridge.