FROM NATURE TO DATA: INSIGHTS INTO ANTARCTIC FUR SEAL NAVIGATION

by

Zac David Warham

B.Mar.Sc., Macquarie University, 2017
B.IT., Macquarie University, 2018

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate and Postdoctoral Studies

(Oceans and Fisheries)

The University of British Columbia

(Vancouver)

August 2024

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

From Nature to Data: Insights into Antarctic Fur Seal Navigation

Submitted by Zac David Warham in partial fulfilment of the requirements for

the degree of Master of Science

in Oceans and Fisheries

Examining Committee:

David Rosen, Professor, Institute for Oceans and Fisheries, UBC

Supervisor

Nancy Heckman, Professor, Department of Statistics, UBC

Supervisory Committee Member

Scott Hinch, Professor, Department of Forest and Conservation Sciences, UBC

Additional Examiner

Additional Supervisory Committee Members:

Duncan Leitch, Professor, Department of Zoology, UBC

Supervisory Committee Member

Abstract

Marine mammals routinely travel vast distances across the oceans, but there is ongoing debate on how they manage these feats. The typical approach to investigating how animals navigate involves first identifying a specific cue or environmental correlate that may influence their decisions, and then looking for behavioural evidence to support this hypothesis. However, selecting a cue to investigate without *a priori* knowledge of an animal's movements can be inefficient and lead to confirmation bias. Therefore, I set out to demonstrate how a meta-analysis that identifies movement patterns in previously collected data is a valuable tool to generate useful hypotheses in navigation research. I illustrate this process using telemetry data from the Antarctic fur seal (AFS; *Arctocephalus gazella*), a model species due to its extensive at-sea movement and the vast amount of existing publicly available location data spanning two decades. The data was pre-processed to split journeys into outbound and inbound transiting periods whilst discarding intermediary foraging behaviour. In total, reliable data was obtained from 86 individuals from Marion Island and 132 individuals from Bird Island for my analyses. I then deployed four different approaches to quantify potential patterns in their movement to ultimately determine how they navigate. Circular statistics identified consistent preference for island-specific dispersal direction during outbound legs from their home island, as well as similarities between outbound and inbound directions. Area analysis identified potential corridor usage during return legs suggesting regular travel routes that can be potentially used to identify environmental cues. Block design statistics showed the seals' ability to gradually correct their heading during inbound legs to the island. However, I did not find any correlation between lunar fraction and the timing of the start of these return legs. This study with AFS also provides a demonstration of how clearly delineating patterns in movement will permit unbiased testing of environmental correlates to determine how marine mammals navigate. For marine mammals in general, this research illustrates the importance of meta-analyses as an efficient, informative approach to analysis decision-making.

Lay Summary

For years, scientists have been collecting at-sea location data on the Antarctic fur seal. Most of this data has been aimed at describing where they feed in the ocean. As a result, we know that they travel a great distance to feed, but we do not know how they manage to get there or to successfully return to their home islands. I utilised various analyses of existing data to identify patterns in how they travel. My results found previously unknown patterns in Antarctic fur seal movement that provide important clues to their methods of navigation. My study also highlights the value of previously gathered data to answer novel questions as a part of the scientific method.

Preface

The broad topic of pinniped navigation was proposed by my supervisor Prof. David Rosen. My contribution to the topic was to identify a data-rich species and acquire the raw telemetry data for my research. My data originated from publicly available datasets from the Australian Antarctic Data Centre and British Antarctic Survey. I then adapted existing methodologies from the fields of navigation and statistics to suit the data and investigate critical patterns in their movement. Further environmental datasets were utilised from the CSIRO (Commonwealth Scientific and Industrial Research Organisation), NOAA (National Oceanic and Atmospheric Administration), and the USNO (United States Naval Observatory), which have been cited accordingly. At this stage, no publications have arisen from this research; however, a future paper will be based upon the methodologies and findings detailed in Chapter Two. The writing in this thesis is my own and was reviewed and adapted with thanks to my supervisor Prof. David Rosen and my committee members Prof. Duncan Leitch and Prof. Nancy Heckman.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to acknowledge the host of individuals who helped me along the way, both professionally and personally. First, I must acknowledge Professor David Rosen, who has been my supervisor throughout this project. I am both thankful for the opportunity he has provided me and the immense educational support in writing this body of work. Completing this thesis under his guidance has been a great learning endeavour. I must also acknowledge the financial support provided by Professor David Rosen and David Green. This funding was the backbone of my relocation to Canada and a great aid in completing this research. Next, I would like to thank my committee members, Professor Nancy Heckman, and Professor Duncan Leitch. It is with great thanks that through both a retirement and an international relocation that they continued to provide a fresh pair of eyes and guidance for this project.

Furthermore, I would like to acknowledge the numerous contributors to the R and MATLAB packages I utilised throughout this research. There are too many developers behind these packages to list here, but each package will be cited appropriately. I would also like to thank the teams that manage the publicly available datasets that I used in my research. Without this available data, meta-analyses like mine would never get off the ground. The specifics of which will be mentioned in the Preface of this thesis, as well as cited throughout the text.

Finally, I would like to acknowledge my family, friends, and teachers who have supported me throughout my education. For their privacy, I will keep them anonymous from this work, but I look forward to celebrating this achievement with you all as the time comes.

1 General Introduction

Almost all animals have some form of oriented movement, meaning movement that is not completely random. This can be referred to as directed movement. For directed movement to be considered navigation, it must be planned movement toward a goal (Gould, 2004). This simple definition obfuscates an incredibly complex scientific question: *how do animals navigate*? It has been stated that 'there is no one answer to the mysteries of animal navigation' (Gould & Gould, 2012) and with literature spanning over 100 years (Wiltschko & Wiltschko, 2022), scientists have shown that navigation research is a mixture of surprising discoveries and confounding mysteries.

At this point, it is known that there are a multitude of navigation systems among animals. The scale of navigation ranges from distances of millimetres to thousands of kilometres (Freas & Cheng, 2022). These can be categorized into numerous types of navigation, including compass orientation, piloting, and true navigation (Gould & Gould, 2012). Navigation almost always requires some form of environmental cue, but the list of potential cues is extensive, including geomagnetic forces, polarised light, sunlight, celestial bodies, geography, smells, tidal forces, sounds, etc. (Gould & Gould, 2012; Sequeira, 2020). All of these different cues require appropriate sensory and processing systems if they are to be utilised in a navigational sense. Additionally, the knowledge to use these sensory inputs to navigate must come from somewhere, whether it is learned, observed, or even passed on via genetic memory. All of this variety creates the challenging goal for scientists: to try and determine *how* a specific animal successfully navigates in their environment.

1.1 General Navigation Background

Whilst finding the exact origins of navigation research would be a scholarly nightmare, it is safe to say that it has been on the minds of scientists for at least 150 years. One of the early examples of animal navigation was labelled as 'dead reckoning,' defined as keeping track of a journey whilst being able to calculate for possible deviations (Darwin, 1873; Gould & Gould, 2012). Although Darwin did not provide direct evidence of this ability in animals (only likening the possibility to that of human behaviour), specific examples of environmental cues and types of navigation began to emerge shortly afterwards. For example, early studies by Felix Santschi found that ants change course after a mirror was used to reflect the sun (Schweizerische Zoologische & Muséum d'histoire naturelle de, 1911). As the study of animal navigation has progressed, navigation strategies are commonly defined as belonging to one of six categories: taxis, compass orientation, vector navigation, piloting, inertial navigation, and true navigation (Gould & Gould, 2012). The remainder of this subsection will present examples from the animal kingdom of each of these strategies and provide insight into what is known and unknown regarding the cues that may drive these strategies.

Defined as the simplest strategy, *taxis orientation* is when the cue is aligned with the goal (Gould & Gould, 2012). In other words, the animal moves toward, away from, or at a fixed angle to the source of the stimulus. The stimulus itself can be anything that the animal can sense, whether stationary or not. Due to its simplicity, this form of cue orientation and movement is observed in a range of kingdoms using a number of different cues, such as magnetotaxis (magnetic alignment) in prokaryotes (Bennet & Eder, 2016) or invertebrates such as worms pirouetting to orient themselves using chemotaxis (chemical gradient) orientation (Jonathan et al., 1999). Taxis orientation however, is by no means a method solely used by “simpler” organisms; phototaxis (light or brightness) has been associated with how hatchling green sea turtles orient themselves towards the ocean after hatching (Mrosovsky & Shettleworth, 1968).

Compass orientation is the next method, in terms of complexity of navigation. This involves the use of a cue to maintain a constant bearing (or adjust the bearing in response to changes in the strength or position of the cue). The specific direction of movement required in a given situation is a function of the geographic relationship between the animal’s present location and its goal (which is situationally specific). As humans, we are most familiar with compass orientation using geomagnetic signals. A large number of other animals have also demonstrated geomagnetic compass orientation. For example, this system has been demonstrated in sea turtles crossing oceans and various species of birds as they fly home or migrate (Lohmann et al., 2022). However, compass orientation also refers to the use of a solar compass, as previously mentioned in ants (Schweizerische Zoologische & Muséum d'histoire naturelle de, 1911). Glass eels use a lunar compass that allows them to orient their swim direction in line with the direction of the moon (Cresci et al., 2021).

Individuals using *vector navigation* rely on a sequence of compass bearings (and associated distances) to navigate (Berthold, 1991; Gould & Gould, 2012). It is different from compass orientation in the sense that compass orientation is a fixed orientation based on a cue, whereas vector navigation is a sequence of compass bearings that could have different strengths or positions as the animal navigates. In essence, it is an advanced form of dead-reckoning sometimes referred to as path integration, whereby an individual can determine its location based on its previous movements (Asem & Fortin, 2017; Wahl et al., 2015). The monarch butterfly is considered to use vector navigation as it has been shown to head southwest both before and after translocation (Gould & Gould, 2012; Mouritsen et al., 2013). However, it is also suggested (and disputed) that they could use true navigation (Gould & Gould, 2012; Oberhauser et al., 2013). A less contentious example comes from studies of the hooded crow that followed the same vector after translocation as if they were home and, and consequently arrived in an abnormal location from their normal breeding range as a result of migration from this new location (Zimmerman & Peterson, 1998). One drawback of vector navigation is that it is (on its own) unable to compensate for changes in position, bearing, or speed caused by drift from air or water currents.

Rats and other rodent species such as hamsters (Etienne, 1992) display inertial navigation when foraging. *Inertial navigation* is when an animal is able to keep track of their journey and determine their location from this information (Gould & Gould, 2012). Honeybees are another species that have shown the ability to use inertial navigation. After a hive was translocated, they showed the ability to determine direction and distance of a food source with an error rate of less than 5% due to their ability to keep track of their journeys (Gould & Gould, 2012).

Unlike vector navigation (also inertial and true navigation, see below), which is generally independent of landmarks, *piloting* is the act of navigating in relation to (a series of) landmarks (Gould & Gould, 2012). This is different from taxis as it typically requires some sort of memory map to be able to recognise the landmark and steer accordingly. By incorporating knowledge of specific landmarks as guides, pilotage is essentially a refinement of simple dead reckoning navigation. Whilst ants have been shown to use piloting, even for landmarks in their peripheral vision (Wehner & Muller, 2010), the homing pigeon is probably represents the most well-known example of piloting (Biro et al., 2004). These findings suggest that visual attraction may be the only requirement to navigate. While there is conflicting research regarding whether it takes more than just visual cues (Wallraff et al., 2005), both works agree that pigeons employ pilotage in their navigation.

True navigation is seen as navigating with an apparent knowledge of the location of the goal (Gould & Gould, 2012). It differs from the other types of navigation as it involves a sophisticated map sense to determine where an individual is in relation to "home". This mode of navigation is most evident when the animal is in unfamiliar territory (Gould & Gould, 2012). As mentioned above, the homing pigeon is the most well-known example of true navigation (Wallraff et al., 2005). Another example is the green sea turtle, which will be discussed further below.

It is important to note that just because a species has been shown to use one type of navigation, it is not excluded from exhibiting other types as well. Animals have been shown to use multiple cues at the same time to create a better navigational sense. Often, this involves a preferred navigation mode or cue, with increased reliance on alternatives when the primary mode is inadequate or disrupted (e.g., cloudy skies interfering with solar or celestial cues). For example, whilst piloting has been shown as the primary method of navigation in rats, studies have also indicated an ability to use inertial navigation (Wallace et al., 2002; Whishaw et al., 2001).

Another (often complimentary) strategy for navigation is the "phase model" (Bingman & Cheng, 2005) where different cues are used at separate stages of navigation. For example, salmon are thought to use geomagnetic cues for open-ocean navigation in conjunction with an inherited magnetic map (Putman et al., 2014). However, for near shore and in-river navigation they have also been shown to use highly specific olfactory cues (imprinted during early development) (Groves et al., 1968) as a key sense of navigation.

The green sea turtle is another example of a species that uses more than just one navigation strategy, with evidence for both true navigation and limited compass orientation in their open-ocean migrations. Research has shown that green sea turtles in the Indian Ocean are able to reorient themselves at sea if they are off course due to current drift, or other factors that could result in incorrect headings (Hays et al., 2020). Whilst this study was only able to speculate on what cues may lead to their reorientation, it was able to highlight the likelihood of a phased approach to navigation that requires multiple methods of navigation to provide varying degrees of accuracy, from coarse to fine detail. The proposed model by Bingman and Cheng (2005) suggests that at a global scale such as Hays et al.'s turtles (or in their case, a hypothetical nocturnal migrant), an individual may start with a rough global memory map based on magnetic field and then hone in on their location through odours, landmarks, and then other beacons (such as roads, bodies of water, buildings, etc.) as the scale of their remaining journey is reduced.

1.2 Methods of Studying Navigation

The previous section provided various definitions for the different methods of navigation used by animals. Scientists can deploy a variety of scientific approaches to determine which method (or methods) are in use for navigation for a given species at a given time. One basic technique is to observe the movement patterns of individual animals to shed light on what general navigation strategies and associated environmental cues they might be using. Observation studies can be entirely void of animal interaction, such as detection of foxes from cameras fixed in their surrounding environment (Matuska et al., 2016), acoustic observations by listening to sounds of migratory birds (Kearney, 2023), and drone monitoring of shallow water shark species (Raoult et al., 2018). Slightly more invasive techniques involve application of artificial identifying markers onto the animal, combined with some sort of resight effort such as iron branding unique IDs on Steller sea lions (Shuert et al., 2015). However, all of these studies can only usually provide animal presence/absence information at pre-chosen observation locations, and movement must be inferred over long durations and/or distances.

Most modern observational movement studies use telemetry tracking. This involves attaching a device to the subject animal and gathering time/location data through re-capture or remote collection via satellite networks or aerial, aquatic, fixed, or drone receivers. Telemetry data can either directly or indirectly be used to determine the position of an animal at any given time. Direct measurements involve methods such as GPS (global positioning service) or ARGOS (Advanced Research and Global Observation Satellite) systems to archive or transmit the geographic location of an animal during the tag deployment (Harcourt et al., 2019). Indirect measurements use environmental features and/or computer models to estimate position, such as using light levels and day length to estimate latitude and longitude (Lisovski et al., 2020). Studies using remote telemetry light loggers have shown that the artic tern can fly over 81,000 kilometres during its yearly migration (Egevang et al., 2010). Acoustic tags and accelerometers have shown the migration patterns of translocated fish, information which was subsequently used to assist with conservation actions (Taylor et al., 2016). While most real-time (non-archival) tags use satellites to transmit their data, drones have recently been used to gather data from smaller tags that are not powerful enough to transmit to satellite networks (Hui et al., 2021).

Most observation studies focus on the *where* of animal navigation. To determine *how* an animal navigates many studies manipulate a variable to determine if the controlled change impacts the individual's ability to navigate. These studies require an increased ethical consideration due to the higher potential of harm that can be caused to the animal. Altering the environment is a common approach, although it has been primarily applied to individuals in captivity. For example, a mirror has been used to change the viewing direction of the sun for ants (Schweizerische Zoologische & Muséum d'histoire naturelle de, 1911) whilst birds have been time-shifted (change night/day cycle through artificial lighting) to offset their internal clocks (Åkesson et al., 2017). A common method to test for cues are 'Emlen funnels.' These typically involve quantifying the movement of an animal in a cage where an environmental cue is manipulated (such as artificial sun direction or change in geomagnetic field). They are commonly used in studies of migrating birds (Bianco et al., 2016) but can be applied to a variety of animals. For example, salmon were placed in covered enclosures surrounded in coils to control for both celestial cues and magnetic direction (Quinn & Brannon, 1982).

Alternately, sometimes it is possible to manipulate the animal directly. The primary form of manipulation is translocation of an individual. The purpose of translocation is to determine how any animal navigates in a novel environment. One potential outcome of translocation is that an animal follows the same movement patterns it would typically take from its home environment, thereby becoming "lost" when it cannot find its target destination, although it may also ultimately settle in the new location. Alternately, the animal may use some combination of cues and/or knowledge to successfully reach its target destination. Each outcome provides clues into how an animal navigates and what cues it may use to do so. For example, successfully moving along an atypical route to reach a preferred destination might indicate true navigation, while following an "incorrect" route will provide clues as to what environmental stimuli might be important in guiding the animal's path. Successful translocation studies have demonstrated potential magnetic homing in bats (Guilbert, 2003), that skin colouration can impact the success of homing in a lizard species (Scali et al., 2013), and – in what is likely to be the most well-known example – the map and compass abilities of homing pigeons (Gould & Gould, 2012). While translocation studies are good indicators for the *how* of navigation, they have been shown to involve a high degree of risk, and are logistically difficult for large animals. Mortality (Villaseñor et al., 2013), potential disease introduction (Chipman et al., 2008), and other social and physiological disruptions (Letty et al., 2007) are all potential negative impacts associated with translocation studies.

Another type of animal manipulation involves temporarily or permanently impeding the suspected sensory systems involved. While this can involve temporary measures such as attaching magnets to the heads of birds (Packmor et al., 2021), this typically involves more permanent manipulations. One course of invasive manipulation is to destroy suspected sensory organs. Such studies include suturing the eyes of rats to determine geomagnetic ability (Norimoto & Ikegaya, 2015) and removing bumblebee wings to show vector navigation (Patel et al., 2022). Sensory input can also be disrupted at the neurological level through manipulation of a particular portion of the animal's brain or neural functions. Severing the trigeminal nerve of reed warblers stopped the ability of migratory weed warblers to orient themselves after translocation (Kishkinev et al., 2013), whilst multiple studies have shown the impact of various brain lesions in homing pigeons (Bingman et al., 1990; Gagliardo et al., 1997; Papi & Casini, 1990). Similarly, drug administration has been shown to impact the navigational ability of learned navigation routes in adult turtles (juveniles were unaffected) (Roth & Krochmal, 2016).

Genetic studies can also be used to look at generational defects or memory to determine if there are specific drivers of navigation. One example of this is the impact of chromosome translocation on the ability of mice to navigate whilst swimming (Leitinger et al., 1994). Neurological approaches may also involve quantifying gene expression, such as that of the suspected light-dependent magnetic receptor, Cry4, in zebra finches (Pinzon-Rodriguez et al., 2018).

Studying navigation in aquatic animals is even more complicated. Unlike in the terrestrial environment where it is (relatively) easy to control for different variables, research in the marine environment brings a whole new host of challenges that renders a lot of the typical investigative methods redundant. In the next subsection, I will discuss how modern advancements in technology have attributed to an influx in animal movement data and how they might be useful for overcoming the difficulties of implementing traditional approaches in the study of marine mammal navigation.

1.3 Challenges in Studying Marine Mammals

Up until this point, I have avoided examples of navigation in marine mammals. This is despite the fact that marine mammals regularly undertake spectacular feats of navigation. For example, the longest recorded migration of any mammal is a grey whale who travelled 22,511km from Russia to Mexico, and back (Mate et al., 2015). However, gathering the data to demonstrate and understand marine mammal navigation is hindered by a unique and challenging set of circumstances induced by their behaviour, size, and their environment. Marine mammals travel through a marine environment that is vast with few apparent natural "borders", largely free of discernible surface features, and can move in three dimensions. Their subsurface movements make them difficult to observe directly, present a challenge for (re)capturing for tagging, and make tracking logistically and technologically difficult. Their large size and elevated social status (they typically gain more attention in the public eye than terrestrial mammals) limit the experimental manipulations that can be performed and sample sizes that can be obtained.

Logically, the safest way to monitor an individual would be remotely from afar through the use of cameras, remote sensing, and drones. Studies have traditionally relied on directed identification efforts by researchers, but more recently studies have used drones to photograph and identify marine mammal species (Aniceto et al., 2018), and image recognition software from public sightings have been used to show whale distribution (Cheeseman et al., 2017). Unfortunately, whilst useful tools in their own rights, these methods usually result in single location data points or, at best, a series of sporadic data if multiple sightings occur of the same individual. This means if scientists wish to gather sequential data for research into habitat use, home range or navigation, animal telemetry must be an integral component of observation (Harcourt et al., 2019).

In studies of terrestrial animals, artificial intelligence is already being utilised for animal monitoring with cameras set up in the environment, including cameras used in zoos to analyse orangutan behaviour (Congdon et al., 2022). Alternately, cameras can be attached to individual animals, such as in studies of vultures that use animal-borne cameras to monitor the ecosystem they live in (Melzheimer et al., 2022). However, it would be impossible to deploy cameras across the ocean for monitoring animals in the same way as in captivity or in terrestrial systems (game trail cameras). Further, animal borne cameras on marine mammals face additional technological challenges. For example, radio transmissions are blocked by salt water (Holton et al., 2021), meaning only species that resurface for extended periods can be used to transmit data remotely. For species that do not resurface regularly (or at all), scientists typically must recover the tags manually, or rely upon a series of underwater receivers and transmitters than can relay the information, which adds further cost and complexity to the tagging process. These same restrictions apply to the use of tags that record geolocation of marine mammals at sea.

An analysis of the costs and benefits of tracking devices must also include consideration of the trade-offs between effective data collection and animal welfare. When designing an experiment that requires animal telemetry, the device attachment method must be taken into consideration not just in terms of logistics and data quality, but also in terms of the impact of telemetry devices on an individual. It is a balancing act between animal welfare and completeness of data due to the high risk to the animal and potential for early tag termination through environmental damage, detachment, and predation (Hays et al., 2007). In recent years, reviews have been published for both internal (Horning et al., 2017) and external (Horning et al., 2019) devices, detailing the best practices for both methods. For marine mammals, the general means of attachment for collecting navigation data is an external device that is a specific percentage of total body mass, although a consensus has not been reached on an acceptable mass threshold (Horning et al., 2019; Rosen et al., 2018; Walker et al., 2012). Research focusing on *Arctocephalus gazella* shows conflicting information, with some studies stating no difference on foraging length with device size (Robinson et al., 2023) compared to other research that has shown a difference (Walker & Boveng, 1995). Whilst scientists are still debating the impacts of device size and weight on marine mammals, terrestrial research into zebra finches shows that the colour of tags can impact mating success (Burley et al., 1982). This could indicate that we must consider not only the impact of energy expenditure from increased drag, but potential social (or predation) effects as well.

Amidst the technological boom of telemetry devices, a whole host of new data has been collected. As a result, this has facilitated an alternative method of studying navigation by utilising this data as a part of a meta-analysis. The typical reasoning for combining data in this fashion is for increased statistical power, however an equally important consideration is that using pre-existing data avoids the welfare concerns imposed on individuals that were discussed above. To date, most tracking research has sought to determine *where* individual marine mammals navigate to, and their behaviour once they have arrived. Specifically, the vast majority of research across species has delineated foraging grounds at the individual (Baylis et al., 2015; Hindell et al., 1991; Lee et al., 2017), population (Kirkwood et al., 2006; Wege et al., 2016), and species level (Nordstrom et al., 2013). These studies typically ask questions such as *where* they forage, *where* are the breeding grounds, or *where* do they migrate, but not *how* do they get there? The following subsection will summarize what we have learnt so far in terms of the *how* of navigation in marine mammals.

1.4 Navigation in Marine Mammals

Despite the challenges in studying marine mammals (see 1.3), scientists have been able to show a variety of mechanisms of navigation in marine mammals, as well as some of the cues that drive their navigation. For example, marine animal tracking has shown that some species such as the northern elephant seal and the humpback whale have migration lengths of thousands of kilometres (Condit et al., 2021; Rasmussen et al., 2007). In addition, some fur seals can successfully navigate a return trip to their island of origin after being translocated over 400km away (Hume et al., 2002). Navigational behaviour can also be affected by short term cues, such as those associated with predator avoidance. One study showed how a harbour seal will alter its typical movement patterns during a transient killer whale encounter (Womble et al., 2007). These examples illustrate the complexity of interpreting the causes of different patterns in animal movement from telemetry data.

Among species of marine mammals, the most researched (in terms of available data) are the northern and southern elephant seals. This is likely due to their long migrations, which reach 10,000 km in some cases (Beltran et al., 2022), alongside their relative ease of tagging due to the large amount of time spent on land during mating and moulting seasons, as opposed to fully aquatic species. A comparison of turn around date to distance from shore showed that northern elephant seals have a map sense (in terms of a sense of distance to end location), but the study was unable to determine the cues of their navigation (Beltran et al., 2022). Another study has shown the ability of northern elephant seals to both maintain their heading during dives and correct their heading after surfacing (Matsumura et al., 2011). They hypothesise that this behaviour is indicative of visual cue orientation above the surface whilst other cues must be in use when below the surface (a magnetic cue was suggested). Due to the abundance of elephant seal data and the need for large sample sizes (Sequeira et al., 2019), meta-analysis studies are also being conducted on elephant seals. As a result, directed movement and corridor usage has been identified in southern elephant seals (Rodríguez et al., 2017).

In some studies, a ‘cue first’ approach is implemented. In other words, scientists introduce a novel cue to determine the animal’s interaction as opposed to measuring existing cues. Studies have shown that captive harbour seals are able to learn and orient themselves relative to underwater landmarks (Maaß & Hanke, 2022; Maaß et al., 2022). This shows potential for piloting in the species as one possible strategy used for navigation. It has also been hypothesised that harbor seals can orient themselves using olfactory cues due to their sensitivity towards dimethyl sulphide (Kowalewsky et al., 2006), which is produced by phytoplankton (Luschi, 2013). Other research has shown that harbor seals can orient themselves relative to a ‘lodestar’ in an artificial celestial environment (Mauck et al., 2008). Similar indications of landmark recognition have been shown in wild marine mammals. Weddell seals after a translocation during ice haul-outs have displayed directed movement (Fuiman et al., 2021), although there is debate regarding the visual cues they employ. Antarctic fur seal pups have shown directed movement on land prior to their first moult (Nagel et al., 2021). Both the Weddell and Antarctic fur seal pup studies provide valuable insight by showing the ability to navigate out of the water as opposed to just within it. On the other hand, northern fur seal migrations have been shown to relate to the movement of oceanic currents (Ream et al., 2005), although it is unclear the degree to which this is directed locomotion or environmental drift.

Outside of the pinniped taxonomic group, it has been suggested that cetaceans have geomagnetic sensitivity due to a correlation between strandings and local magnetic minima (Kirschvink et al., 1986). This sensitivity to geomagnetism has been proposed in fin whales (Walker et al., 1992) and bottlenose dolphins (Kremers et al., 2014) but suggested to be unlikely in humpback whales (Horton et al., 2011; Horton et al., 2020). However, potential links have been found to the lunar cycle and departure time from breeding grounds in humpback whales (Horton et al., 2017). Additional humpback whales research has also shown a correlation between temperature and migration patterns, suggesting that they migrate by following waters within a desired range (Rasmussen et al., 2007). Similar findings in relation to temperature have also been found in both manatees and dugongs (Sheppard et al., 2006). One species of manatee has been shown to utilise shipping channels as navigational corridors (Cloyed et al., 2019). Unlike with harbor seals, a study testing for orientation towards dimethyl sulfide in bottlenose dolphins found no significant indications of this ability (Bouchard et al., 2022). This finding is of particular importance because it highlights that navigation research is not a "one size fits all" approach. Each species must be considered individually when determining the *how* of their navigation.

Clearly, there is still much to learn regarding the cues and methods used by various marine mammals. For my study I have chosen the model species *Arctocephalus gazella*. This species was chosen as they undertake extensive trips to and from specific island rookeries, and there is an extensive amount of available tracking data from the last few decades.

1.5 Antarctic Fur Seal Navigation

The Antarctic fur seal (AFS), *Arctocephalus gazella*, is an otariid (eared) seal that shows significant sexual dimorphism (Staniland & Robinson, 2008), with males approximately 1.5 times the length and 4 times the weight of females (Forcada & Staniland, 2018). Population estimates vary wildly, from a little under one million to over three million globally. The population is considered to be declining, but maintains a conservation status of 'least concern' level in the IUCN Red List (Forcada et al., 2023; Hofmeyr, 2016). Whilst pups are known to be preyed upon by the leopard seal in some locations (Schwarz et al., 2013), the general consensus for the decline in population is considered to be due to climate change reducing prey availability (Forcada & Hoffman, 2014). These effects could also be magnified by the low genetic diversity of the species (Hofmeyr, 2016). Breeding colonies are found in at least six locations, mainly the South Georgia Islands but also at various islands in the Southern Indian Ocean and Southern Atlantic Ocean (Forcada & Staniland, 2018; Hofmeyr, 2016; Krause et al., 2022). Females give birth around late November to early December before a four-month lactation period (Doidge & Croxall, 1989).

Contrary to its name, the AFS predominantly occupies the subantarctic region (Forcada & Staniland, 2018) (Figure 1). Males have a reported maximum dive depth of 350 m and 10 minutes in duration, with females diving up to 210 m and 5 minutes. However, they predominantly dive a lot shallower than this, with most dives less than 60 m (Forcada & Staniland, 2018; Viviant et al., 2016). There are mixed reports on diet due to it being heavily region dependent, with a largely krill diet in the Atlantic Ocean (Reid & Arnould, 1996) and mostly fish in the Indian Ocean (Cherel et al., 1997; Lea et al., 2002). However, there are observations of squid and penguins being consumed as well (Forcada & Staniland, 2018; Reid & Arnould, 1996). The separation in diet among populations is suspected to be as the result of environmental factors, not phylogenetic ones (Robinson et al., 2023). Males usually forage further from shore, which is considered to be due to the need to avoid direct resource competition with the more geographically restricted lactating females (Boyd et al., 1998), with post-weaning females foraging at much greater distances (Boyd et al., 2002). It is hypothesised that seals will remember their previous foraging direction and repeatedly return in future trips (Bonadonna et al., 2001), allowing individuals to capitalise on areas of known prey availability (Iwata et al., 2015). It has also been shown that dispersal direction during foraging trips is relatively uniform within populations (Bonadonna et al., 2001), likely as a strategy to avoid competition on foraging grounds with other species from the same home islands (Wege et al., 2019).

During longer migration movements, it has been shown that the South Georgia AFS population will migrate both south (Boyd et al., 1998; Boyd et al., 2002) and north (Staniland et al., 2011) during the winter, post breeding season. Some females have been shown to migrate as far as the Patagonian shelf, 2000 km away from South Georgia (Staniland et al., 2011). Given that the males spend the majority of their time foraging further afield than the females (Boyd et al., 1998), it is an impressive feat to see females spending substantial time at sea, given the significant sexual dimorphism of the species (Forcada & Staniland, 2018; Staniland & Robinson, 2008). Females from the South Shetland islands colony will migrate north to forage during winter, presumably due to increased availability of krill (March et al., 2021). Similar findings were found for both the South Shetland and South Georgia populations (Arthur et al., 2017; Bamford et al., 2021). Arthur et al. (2017) also showed that the Marion Island population will consistently forage at the Del Cano Rise (to the east/northeast). All this information describes *where* AFS navigate to. In the next subsection I will explain how my research is the first step for determining *how* they navigate there and back again.

1.6 My Research Focus

Most of the examples discussed to this point have shown what we know about marine mammal navigation, with the majority of research questions pertaining to *where* animals navigate to, not *how* they navigate. My goal is to demonstrate how we can use already available data from marine mammals to provide the stepping stones for deciphering the cues and methods used by marine mammals for at-sea navigation using AFS as a model system. To provide the first steps in this research pursuit, I sought to answer the question: what movement patterns exist in Antarctic fur seal open-ocean navigational periods? It is hoped that if any patterns are found, further studies can be conducted that focus on the potential cues that could lead to them.

My general approach followed three steps: (i) data cleaning and validation (quality control), (ii) identification of transiting movement (to focus on directed travel vs foraging movement), and (iii) identification of movement patterns. Thanks to the Australian Antarctic Data Centre (Ropert-Coudert et al., 2020), and the British Antarctic Survey (Staniland et al., 2020), a combined 786 AFS tracks were utilised in my research. This data consisted of tracks from two study sites, Marion Island and Bird Island. Other islands that were included in the original dataset were not included in this study due to the significantly smaller sample sizes at each of these islands. This was to follow the recommendations proposed by Sequeira et al. (2019) on maintaining a sufficient sample size for each type of question. Data from Marion Island spanned a five-year period, whilst data from Bird Island spanned 15 (with one-year of missing data).

Quality control is an important step in using animal tracking data, given the technical issues inherent in satellite tracking and data archiving. It is also important for ensuring the type of data matches the study objectives. The details of this process are discussed briefly in Chapter 2, but detailed extensively in the Appendix. Given that my study focussed on periods when the seals appeared to use directed movement towards a goal, it was important to separate the different types of at-sea behaviour. To accomplish this, the data was first divided into the two island subsets before being fitted to a predictive model called move persistence (Jonsen et al., 2023). This provided a means of distinguishing between periods of foraging or rest from those of transiting behaviour. This also allowed me to differentiate between episodes when the fur seals were transiting away from and towards the island. Finally, I was able to look for patterns within the movement data. There are countless ways of describing and analysing patterns in movement data. I chose analyses that I felt would most likely lead to a better understanding of the underlying cues and methods used for at-sea navigation. Therefore, four subsequent analyses were performed on the predicted transiting periods. I used Rao's Spacing Test to check for non-uniform dispersal pattern of individuals on their outbound journey. This was followed by a visual analysis of inbound corridor usage which could be used to test against potential environmental correlates, such as bathymetric features in the region. For my next analysis, I wanted to determine if individuals got progressively more direct in their approach to their home rookery. To determine this, the data was binned and put through a Skillings Mack test (Chatfield & Mander, 2009) to check for a change in bearing offset along each return journey. Finally, the starting point of their return journey was then compared against the lunar fraction to determine if a relationship exists between the two.

The goal of Chapter 2 is to define and discuss specific movement patterns in AFS. In Chapter 3 I also investigate and discuss the strengths and weaknesses of trying to understand the relationship between AFS movement patterns identified in Chapter 2 – namely movement corridors – and several different potential environmental correlates. The final chapter of this thesis provides an overview of what these findings mean for future research and considers any potential limitations in these approaches in relation to other previous literature.

Overall, the purpose of this research is to outline and test an approach for studying navigation in a species using pre-existing data. It emphasises the importance of switching from asking *where* an animal navigates, to *how* an animal navigates.

2 From Nature to Data: Insights into Antarctic Fur Seal Navigation

2.1 Introduction

Animals need to accurately move between locations for various reasons, including to seek shelter or food, to breed, or to avoid predation. Depending on the species, the scale of this movement might be as small as a spider crossing its web to eat or as large as a whale traveling across an ocean to breed. Navigation is when an individual's movement is oriented or, in other words, non-random (Gould, 2004). Unlike humans, who have a vast array of technological advancements to help us navigate, animals must do so using natural instinct and/or ability. Many of these abilities broadly mimic human technology. For example, animals have evolved sensors to detect a wide variety of environmental signals. Whilst, to the lay person, a map may seem like a strictly human tool, scientists have discovered that some animals keep a mental map to navigate their surroundings (Bingman & Cheng, 2005; Putman et al., 2014).

For over 100 years (Wiltschko & Wiltschko, 2022), scientists have studied animal navigation to define home ranges, assist in policy making, inform conservation methods, or simply out of curiosity and wonderment. This research has uncovered a vast array of navigation techniques used by animals, such as piloting, inertial navigation, and taxis orientation (Gould & Gould, 2012) (see 1.1). Navigation in marine mammals is of particular interest given its apparent scale and challenges. Compared to what humans can sense and in contrast to the apparent multitude of environmental cues available to terrestrial animals, marine mammals are able to precisely navigate long distances over a seemingly featureless physical environment. Understanding how they are able to navigate with such little information, or perhaps by using information sources we are not aware of, is a scientific challenge.

However, conducting navigational studies with marine mammals presents unique challenges. Their aquatic existence means that unlike terrestrial animals that can be more readily visually tracked (Congdon et al., 2022; Matuska et al., 2016), translocated (Thorup et al., 2007; Zimmerman & Peterson, 1998), and otherwise manipulated (Patel et al., 2022; Schweizerische Zoologische & Muséum d'histoire naturelle de, 1911), marine mammals require more effort to study. Marine mammals also tend towards larger body size, making them difficult to handle, and enabling them to typically travel further distances to seek out richer food supplies than smaller marine life (Boyd, 2004). Whilst smaller marine vertebrates also undertake long migrations (Dutton et al., 2008;

Guzman et al., 2018; Ortiz et al., 2003), the grey whale tops all of them at over 22,500 km for a return migration journey (Mate et al., 2015). Consideration must also be given to the environment they are navigating. Marine mammals navigate within a three-dimensional environment and many of the physical features of their environment that they may be using to navigate are transient in nature, adding a critical temporal restriction to potential studies. Furthermore, given the distances they travel, they may typically cross into international borders of multiple countries, increasing the complexity of (re)capture efforts and data collection. The combination of physical, environmental, and behavioural characteristics contributes to an increased cost to track individual animals. This cost can be prohibitive given the need to have a sufficient sample size to answer the complex questions that navigational studies aim to address (Sequeira et al., 2019).

Fortunately, marine mammal location information from telemetry tags is often collected as a part of other studies. Despite the inherent costs of telemetry studies in marine mammals, advancements in the technological capabilities of tracking devices in the recent decades (Harcourt et al., 2019) have led to a boom in studies tracking different species of marine mammals. This has led to a great catalogue of existing data on the location of many species of marine mammals that was gathered for studies, focussed on topics such as foraging ecology and oceanographic monitoring, but normally not on navigation. Data gathered typically involves the time-stamped coordinates of an animal as latitude and longitude and – depending on the capabilities of the equipment – other environmental factors such as depth, light level, temperature, and salinity etc. This data can be repurposed to investigate aspects of navigation, including understanding what environmental cues a particular animal utilises during their journeys. A few studies have used such data to provide insight into how a limited number of marine mammal species navigate. Studies investigating specific navigation strategies include those of a suspected map sense in the northern elephant seal (Beltran et al., 2022) and possible corridor usage in southern elephant seals (Rodríguez et al., 2017). Other studies have focused on the use of potential environmental cues such as temperature (Rasmussen et al., 2007; Sheppard et al., 2006), currents (Ream et al., 2005), and olfaction (Bouchard et al., 2022; Kowalewsky et al., 2006), although research involving marine mammals in the wild are few and far between.

One species that undergoes impressive feats of navigation and for whom there is a vast amount of readily available tracking data is Antarctic fur seal (AFS; *Arctocephalus gazella*). The AFS has a circumpolar distribution (Forcada & Staniland, 2018) (Figure 1) and breeds on various islands in the subantarctic and Antarctic regions (Forcada & Staniland, 2018; Hofmeyr, 2016; Krause et al., 2022). During lactation, females typically spend four to five days at sea searching for food. Their search is relatively near to shore when compared to the males who will forage over a larger area, presumably due to competition with females (Boyd et al., 1998), and disperse from the island after mating (Forcada & Staniland, 2018). Beyond the breeding season, AFS will travel throughout a vast portion of their home range, with potential immigration to new locations (Hofmeyr, 2016). However, their specific distribution during this period is relatively unknown (Hofmeyr, 2016).

This study focuses on AFS individuals breeding on Marion Island and Bird Island tagged throughout the year. Marion Island is approximately 300 square kilometres in area (Van Zinderen Bakker, 1967) and is located in the subantarctic zone of the Indian Ocean. Bird Island is a much smaller island at only four square kilometres (Bonner & Croxall, 1988) that is situated in the subantarctic region of the Atlantic Ocean. Both islands are the result of volcanic formation; however Marion Island has a higher maximum altitude, more than three times that of Bird Island at 3890 ft. compared to 1198 ft. (Bonner & Croxall, 1988; Van Zinderen Bakker, 1967). In stark contrast to the drastic difference in island sizes, AFS from Marion Island account for less than 1% of the total population (Hofmeyr et al., 2017), whereas Bird Island accounts for 5.2% (Forcada et al., 2023).

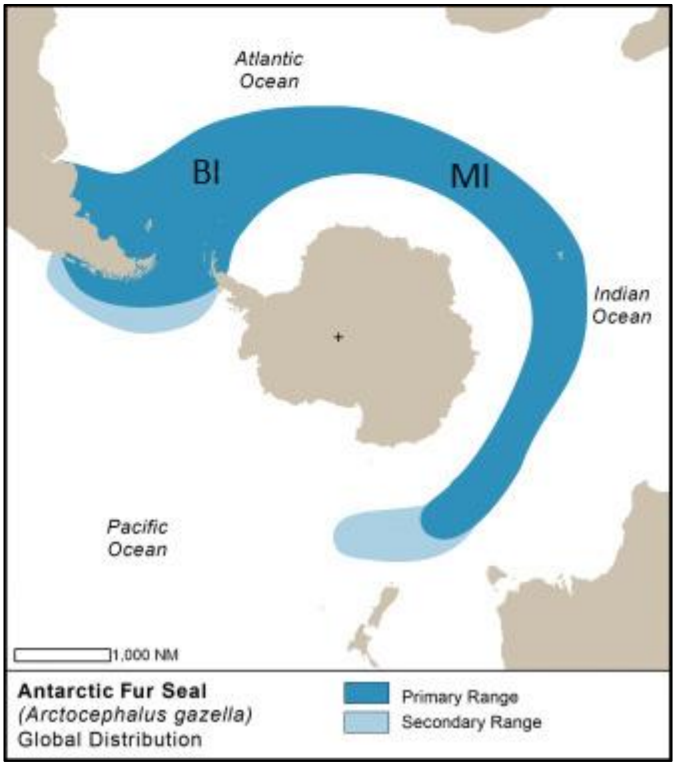

Figure 1. Map of AFS home range. Figure modified from Forcada and Staniland (2018). The approximate locations of the two study sites are marked: Marion Island (MI) and Bird Island (BI).

Although there has been extensive research on the AFS, only a small portion of this relates to navigation. Research at the Kerguelen Archipelago has suggested that individuals choose their outbound dispersal direction based on a 'colony-preferred direction' and also have directional consistency between journeys at the individual level (Bonadonna et al., 2001). It is also known that AFS will tend to avoid travelling to foraging grounds of other species (Wege et al., 2019). Genetic tagging showed that males returning to the breeding grounds are able to return to 'within little more than one body length of where they were in previous seasons' (Hoffman et al., 2006). Beyond these studies, the remaining literature focusses on where they travel to (i.e., their destinations), with suggested links between foraging and/or migration routes and bathymetric features and meteorological conditions for AFS originating on South Georgia (Boyd et al., 2002; Staniland et al., 2011), the Shetland Islands (Arthur et al., 2017; March et al., 2021) and Marion Island (Arthur et al., 2017). These studies show that we know *where* the AFS travels *to*, but they do not explain *how* they get there.

The importance of the question of *how* the AFS navigates cannot be under-estimated. Scientists do not know what techniques or environmental correlates fur seals use to find their foraging grounds or return to shore with such high fidelity. Hence, we may be inadvertently hindering their ability by altering unknown environmental correlates that inform their navigation systems. Whilst current research is being used to inform decisions on marine protected areas for foraging, we need to know what cues they use to navigate to ensure that our decisions allow the species to continue accessing these resources. Furthermore, due to the top-down effects marine mammals have (Bowen, 1997), it is crucial to ensure we understand how they navigate if we are to maintain our ocean systems and make decisions towards species conservation.

Due to the aforementioned challenges in determining methods of navigation in marine mammals with classical techniques used for terrestrial animals (see 1.3), it is crucial to take other approaches. The purpose of my research is to start the process of determining *how* marine mammals navigate through the meta-analysis of the existing tracking data using AFS as a model species. Specifically, I will quantify and describe trends or patterns in the movement of AFS from existing location data, based on the hypothesis that patterns must exist in their movement both at an individual and at the population level. To investigate this hypothesis, I adopted four separate analyses with the following predictions.

(i) Dispersal direction around the island is non-uniform. This is based on the knowledge that AFS are travelling to known, distinct foraging grounds and that they should travel directly to these areas from their rookeries.

(ii) Individuals follow common routes to return home. If AFS are effectively travelling (to and) from a limited number of foraging grounds, this should result in a small number of common routes directly back to the rookery.

(iii) Individuals increase their accuracy as they approach shore. Due to environmental factors such as drift, AFS must be able to improve its accuracy to adjust for the unintended changes in their movements in order to successfully return to the island.

(iv) That lunar fraction is used as an indicator for when to return home. Other marine mammals have shown the potential to use lunar cues for timing their trips, and therefore AFS may also be using this technique.

Identifying if these predictions hold true is the first step to test for specific environmental cues in long-range navigation in this species. One specific approach to such an analysis is explored in Chapter 3. Together, this thesis not only investigates the movement of AFS but develops a methodology of data analysis that can be adapted to other marine mammalian species as well. Such a template will facilitate navigation studies using existing telemetry data on a range of novel marine mammal species.

2.2 Methods

2.2.1 Data Sources

Data describing animal location at sea was collated from publicly available datasets provided by the Australian Antarctic Data Centre (AADC) (Ropert-Coudert et al., 2020) and the British Antarctic Survey (BAS) (Staniland et al., 2020). These time-stamped location data represent animal tracks. I define a "track" as a sequence of GPS coordinates (describing individual locations) collected from one animal during a single deployment of a telemetry device. The AADC dataset as described by the original authors represented "consolidated tracking data for multiple species of Antarctic meso- and top-predators to identify Areas of Ecological Significance" (Ropert-Coudert et al., 2020). Out of the entire dataset, 553 tracks were from AFS obtained using ARGOS (Advanced Research and Global Observation Satellite) technology. These were collected using either global location sensors (GLS) or platform transmitting terminal (PTT) tags. The unfiltered AADC data spanned 17 years between 1995 and 2015 (not all years had data) and was collected across ten different islands. The scope of the BAS dataset was more focused, with all tags originating from adult female AFS during the summer breeding season from Bird Island, South Georgia. A total of 492 tracks were collected using ARGOS-certified PTT tags, obtained between 1995 and 2010. Combining both data sources and filtering to my two focal study sites resulted in data available from a total of 786 tracks from tagged AFS, representing tracks from both during, and after the breeding season. (Table 1).

Table 1. Track Dataset. The number of tracks available for the study after combining the AADC and BAS data sources. Sample sizes are broken down by age and sex of the animal (if identified in dataset**)** and the type of tag deployed, either *Platform Transmitting Terminals* (PTT) or *Global Location Sensor* (GLS).

Study	Island	Tag Type	Age	Sex	Count
BAS	Bird Island	PTT	Adult	Female	492
AADC	Marion Island	PTT	Unknown	Unknown	104
AADC	Marion Island	GLS	Adult	Female	32
AADC	Bird Island	PTT	Unknown	Female	139
AADC	Bird Island	GLS	Adult	Female	12
AADC	Bird Island	PTT	Unknown	Unknown	7

2.2.2 Data Preparation

The goal of this research was to investigate navigation in terms of complete journeys. A complete journey is defined as the portion of a track that starts with the departure from an island and ends with the return to the same island by an individual. Other types of trips that also require navigation (such as between island immigration) and journeys with track termination events (typically power loss, predation, or equipment error) are not considered "complete" under my definition. Therefore, a series of pre-filtering steps were required to exclude these journeys. For the pre-filtering, unless specified otherwise, all data was processed using the R Programming Language (R Core Team, 2022) in the R Studio IDE (Posit team, 2023). A list of all packages used can be found in appendices one through four. After the initial collation of the 786 tracks from tagged AFS, a *Shiny* app (Chang et al., 2023) was developed to visualise all tracks to facilitate the first stage of data cleaning. A manual inspection of all tracks was performed to provide a subset that only included complete journeys. Due to the duration of tag deployment, a track could have multiple complete journeys if it departed from and returned to the same island multiple times. All incomplete journeys were discarded from my analysis, that is any journey that was missing departure information, did not return to the same island, or did not return at all (most typical situation due to termination of the data stream). This resulted in a total of 1056 complete journeys, referred to going forward simply as "journeys".

Due to using the raw data provided by the AADC and BAS datasets, the data was not always complete and accurate for each journey. Therefore, data was further filtered to remove inaccurate data within a journey as well as entire journeys that did not meet data standards. In particular, location class was missing from a significant portion of the journeys. Location class is a classification method employed by ARGOS to provide an estimate of error for each location (Collecte Localisation Satellites, 2016). Other characteristics that led to exclusion of data from journeys involved (but are not limited to) missing data points, positioning areas (e.g. middle of a continent), and duplicated data. These characteristics were compounded due to the errors in original collation for each data source and then combining the two different sources into one. Therefore, further cleaning was required to address the inconsistencies in the various methods of data collection. The first stage was to eliminate duplicate timestamps within a journey. If an individual had two reported locations at the same time, the first occurrence in the data was kept for the purpose of my analysis. The next stage was to remove duplicate tracks as some scientists had submitted their data to both data sources. Duplicate tracks were identified through a combination of visualise inspection and cross-referencing metadata values (mainly individual ID). Three journeys were then removed due to instrumentation error in the dataset as these tags had adjacent locations (in terms of time) that were across the ocean from each other. The final stage of cleaning was to filter to only include journeys that had complete location class information. This is due to the fact that complete location class information is a strict requirement for the move persistence modelling conducted in *aniMotum* (Jonsen et al., 2023) (see 2.2.3).

2.2.3 Identifying Travelling Segments

This study concentrated on the transiting legs (as opposed to the foraging legs) of journeys. This was defined as the leg between departure from the island and the first foraging period (outbound leg) and the portion between the last foraging period and return arrival to the island (inbound leg). A move persistence model (Jonsen et al., 2023) was used to provide a 'behavioural estimation' for each location to categorise sections of each journey into times of foraging and transiting. Briefly, a move persistence model scales (0-1.0) the likelihood of individual portions of a track corresponding to one of two types of movement, with higher move persistence values indicating movement of higher velocity and lower turning angles. As such, higher values are more likely to represent travelling with lower values indicating foraging or resting. The model produces predicted locations and corresponding move persistence values at regular, pre-set intervals. A speed, distance, and angle (SDA) filter was used to remove invalid points in a journey. The SDA removed points that were beyond the physiological capabilities of the species (e.g. impossible speeds and distance travelled over time). The maximum velocity used for the SDA filter was 3m/s with the predicted values at a time step of every six hours. These values were chosen to meet the established standards from previous research for SDA filters on this species (J. Sterling., personal communication, 2021). The move persistence ratio (mp) was then extracted alongside the predicted coordinates for each journey. The mp value was normalised separately for each individual to maintain consistency across the track when a track contained more than one journey. The model was fitted separately for both Bird Island and Marion Island using the *aniMotum* (Jonsen et al., 2023) package in R (Figure 2, Figure 3).

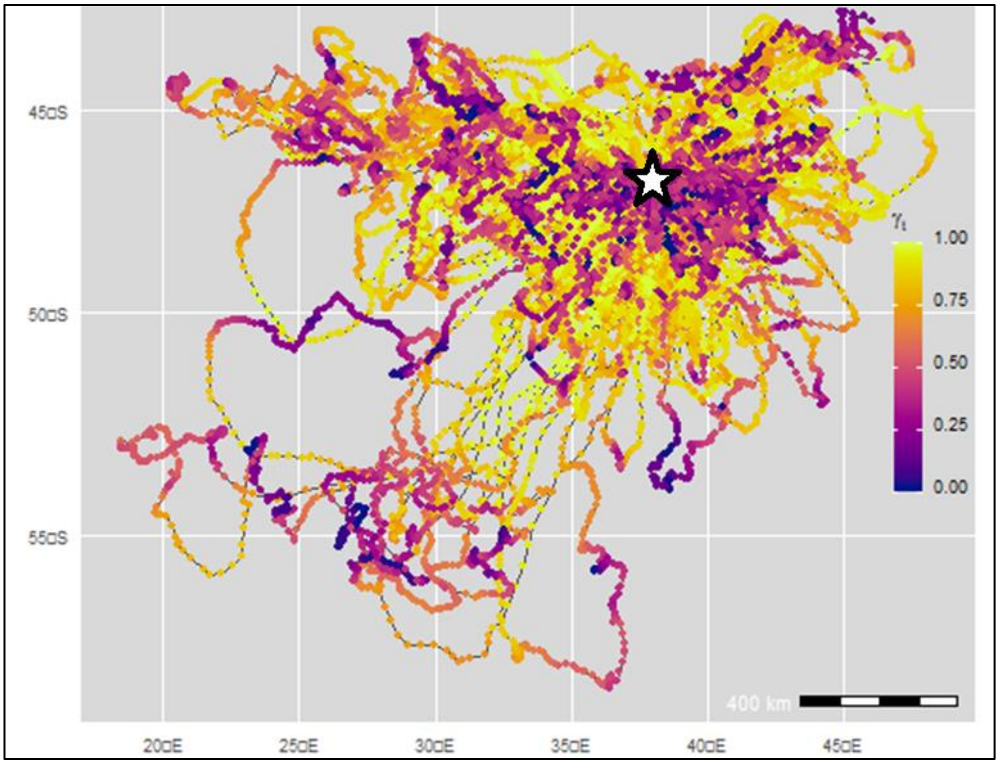

Figure 2. *aniMotum* output from Marion Island. Map of usable journeys from Marion Island indicating their corresponding move persistence values at each section. Move persistence decreases along the gradient of yellow (likely travelling) to purple (likely foraging or resting). Marion Island location is indicated by the star. Map attribution: Jonsen et al. (2023).

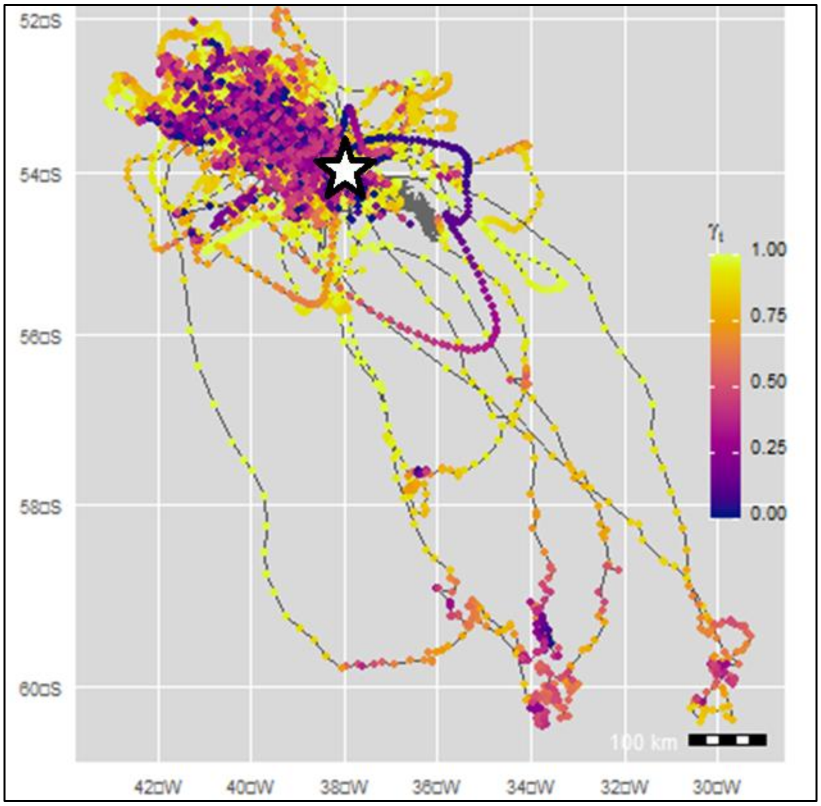

Figure 3. *aniMotum* output from Bird Island. Map of usable journeys from Bird Island indicating their corresponding move persistence values at each section. Move persistence decreases along the gradient of yellow (likely travelling) to purple (likely foraging or resting). Bird Island location is indicated by the star. Map attribution: Jonsen et al. (2023).

Locations within individual journeys were then filtered to only include assumed travelling portions by only retaining locations with mp >0.75 to exclude assumed foraging and resting behaviours. The dataset was then modified to include the distance of each at-sea location while traveling. For each location, the displacement from the island was calculated using the Vincent Ellipsoid distance from the *geosphere* (Hijmans, 2022) package. The distance from Marion Island was calculated based on coordinates to the centre of the island, 46°55'26.3"S, 37°43'59.5"E. Each distance to Marion Island was then adjusted to account for an average distance (20km) to the ocean from this position. This is so that all subsequent analyses are based upon approximate distance to shore as opposed to the centre of the island. For Bird Island, the reference position was 54°S, 38°W with no distance adjustment to the ocean given the drastically smaller island size.

At this stage, the dataset for each journey only contained the locations that corresponded to predicted transiting movement behaviour (mp > 0.75). By filtering out the foraging periods, the remaining data contained a series of separate transiting legs. These individual transiting legs were split by any two subsequent locations that had a time gap greater than six hours where a period of slower foraging had been removed from the journey (this is due to the six-hour smoothing of the data applied during the move persistence modelling). All the remaining transiting legs were then given a unique identifier to distinguish the individual periods of transiting within that journey based upon these gaps (Figure 4). Since my research is specifically only analysing the outbound and inbound transiting legs, additional data filtering was required to address instances of (i) journeys with no foraging periods (i.e., were composed of only a single, continuous leg) and (ii) legs that occurred between different foraging patches (as opposed to the initial departure and last return legs of a journey). To filter out these other transiting legs, a proximity filter was applied to include only legs that had a location within 40km of the island was performed, alongside a filter to keep legs with a duration of at least 2.5 days. This removed short duration legs and/or those that did not start or end in proximity to the island, that may not indicate long-range navigation to and from the island.

To differentiate outbound and inbound legs, a custom function was then fitted to automatically determine the direction of travel. This function calculated a rolling mean for the distance of every four consecutive locations using the *slopeEveryN* function from *Thermimage* (Tattersall, 2021). A mean distance that is increasing indicates a departing leg as the individual is getting further from the island and a decreasing mean indicates a leg where the individual is returning to shore. In contrast, a mean that approaches zero indicates a journey that is looping away and then towards the island. All legs were then visually inspected to ensure they were correctly categorised, and any remaining invalid legs were removed (e.g., at-sea legs that were missed through the proximity cutoff). All looping journeys were split at the halfway point (in terms of the number of data points) into separate inbound and outbound legs (Table 2).

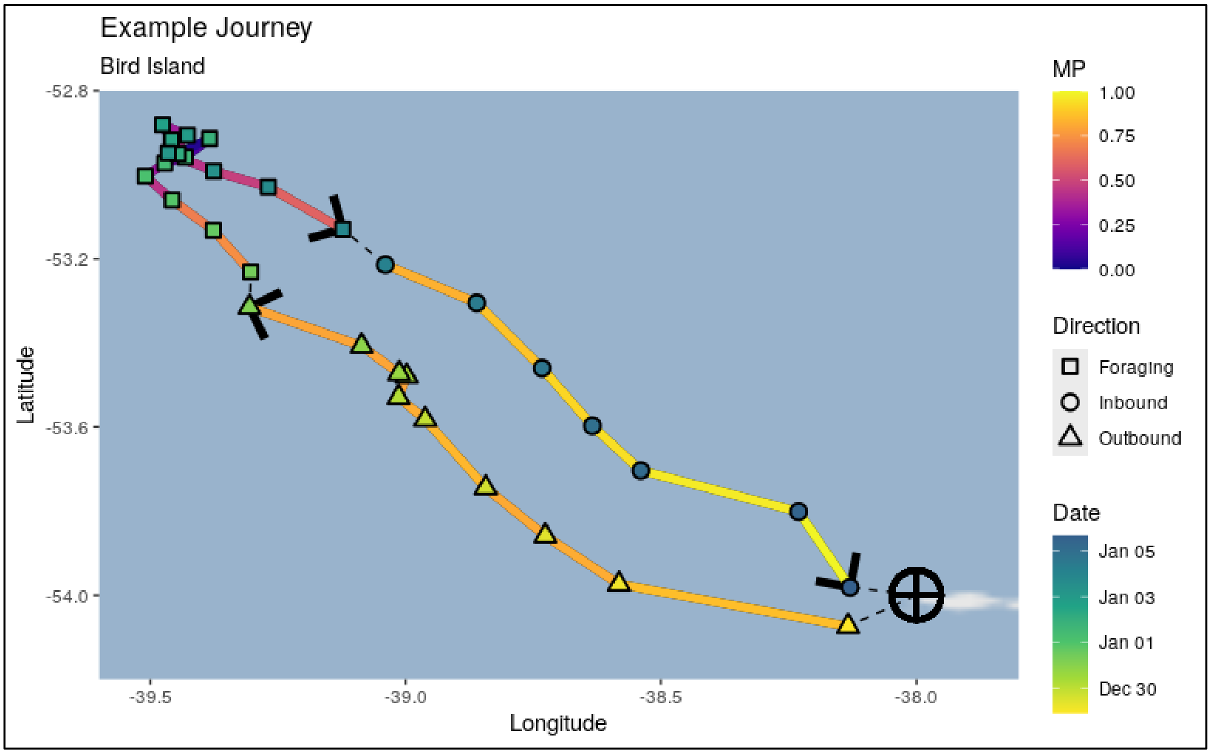

Figure 4. Defining a Journey. A single journey extracted from a series of journeys in an individual AFS track. The gradient-coloured arrows indicate the move persistence (mp) value at each point as determined by the *aniMotum* analysis with a high value meaning travelling. The shape of the points indicates the type of movement for each leg as categorized by the cut-off point of 0.75 mp. The colour of a point indicates the date of the location. Gaps between the first and last point and the island indicate the near-shore proximity cut-off. The gaps between the legs indicate the point where the threshold switches between foraging and transiting behaviour categorisation. Due to removing the foraging period, the time gap between the two transiting legs is greater than six hours and this criteria was used to provide unique labels to each of the transiting legs.

Table 2. Counts of Usable Legs at Each Island. Counts of outbound (leaving from island) and inbound (returning to island) legs at each island after data cleaning. All legs represent complete journeys, but some legs subsequently removed after move persistence analysis.

Island	Total AFS Count	Outbound Legs	Inbound Legs
Bird Island	132	116	128
Marion Island	86	190	194

2.2.4 Dispersal Direction

I examined the level of uniformity in dispersal heading as a measure of preferred direction in both departure from and arrival to the island, and the relationship between the dispersal patterns for the outbound and inbound legs. If the dispersal pattern was uniform (null hypothesis), it would indicate seals made a random choice of direction. For each individual, the location used to calculate their heading was taken from the closet datapoint to a fixed radius from each island. For Marion Island, I measured the angle of dispersal (i.e., compass heading relative to the island) of the track at a distance of 50 km from the island. Given the size of the island, the analysis used an effective distance of 62.5 km to account for the average distance from the centre of the island to the shore. Bird Island used the same 50 km radius; however, since the coordinates were used to measure the distance from the rookery rather than the centre of the island, no additional offset was added.

For each inbound/outbound leg of a journey, the predicted location data derived from the move persistence model was filtered to only include the nearest point to the target dispersal radius for each island. Any legs that did not include a point that fell within 10km of the radius were then discarded for this analysis. This 10km threshold was to allow for the data smoothing that occurred in the move persistence model that fitted the data to six-hour intervals. A further filter was then performed to keep only journeys that had both a matching outbound and inbound leg within the range of the dispersal radius. This left data for 72 paired legs for Marion Island and 27 for Bird Island. Paired data was utilised for this analysis to ensure that the heading was taken from the same known position for both inbound and outbound legs of the same journey.

The more common Rayleigh test was originally considered to test the null hypothesis that the data is of a uniform distribution, but due to the possibility of a multimodal distribution (Rayleigh has low power to detect multimodal distributions), a Rao's Spacing Test (Landler et al., 2019) using the *Circular* (Lund et al., 2022) package was calculated instead, with four analyses, testing the dispersal for each island separately for both outbound and inbound legs. The data were also exported into MATLAB (The Mathworks Inc, 2022) to calculate whether there was a significant difference in the dispersal patterns between outbound and inbound journeys for each island using the *Circular Statistic Toolbox* (Berens, 2009). Finally, I calculated the deviation of paired dispersal legs, a descriptive statistical analysis of the difference in the angle between the outbound and inbound leg for each set of paired legs. These were measured at the same reference distance of 50 ± 10 km from the island.

2.2.5 Corridor Usage During Inbound Legs

I wanted to determine if seals used specific routes when returning to the island to provide an insight into high usage areas. This would allow for regional comparisons of the environment for future analysis into what environmental correlates might result in heavily trafficked corridors. Corridor analysis was used to visualise the long-range movements of individuals over the inbound legs (194 Marion, 128 Bird) of their journeys. This visualization is an adaptation of the common meteorological "Quiver plots" used to show wind direction (O'Hara-Wild, 2021).

The first stage involved calculating the bearing of every individual for all locations along each leg. Each bearing was then converted into its vector components of u and v. In this context, u is defined as the direction parallel to the x-axis [u = mp x cos(bearing)] and v is defined as the direction parallel to the y-axis [v = mp x sin(bearing)]. Data must be in this format for software packages that produce weather maps to visualize the results (CLIM301, n.d.). In this case, the *ggquiver* (O'Hara-Wild, 2021) package calculated the average heading and velocity. The heading and velocity components from all locations of all seals were then averaged across a series of 0.25 by 0.25 degrees grid cells to provide an average vector within each cell. These were mapped for inspection onto a Quiver plot that indicated both average heading and velocity within each cell, as well as total number of data points within that cell.

2.2.6 Bearing Offset at Various Distances

As the individuals return closer to shore, I theorised that they should gradually become more directed in their navigation. That is, as they approach shore, the difference between their actual direction of travel and the true bearing to the island – the bearing offset – should decrease. The bearing offset for any given location is therefore defined as the absolute difference in the number of degrees between their current heading and the heading that would take them directly to the island. This was calculated for each location on their return leg, and each datapoint was then paired with the displacement from the island destination. Hence, each return leg produced a number of bearing offset-distance datapoints equal to the number of at-sea locations.

To determine if there was a significant change in bearing offset with displacement from the island, the data was categorised into five distance intervals for a block design approach with each individual journey appearing in multiple categories (typically referred to as treatments). When an individual was represented in a particular interval more than once, the median bearing was used. The selected intervals were based on the pattern of displacement of seals from each island. Specifically, for Marion Island, 50 km intervals were used whereas intervals of 25 km were used for Bird Island to match the 50% shorter mean journey displacements. The first interval (0-50 km Marion, 0-25 km Bird) was then dropped for each island to exclude the potential for atypical behaviour due to island proximity biasing the results. For both islands, the remaining four intervals were then used for the block design. The original intention was to use a Friedman-type statistic; however, this type of statistic is not suitable when some data is missing. Due to the nature of the dataset and preprocessing, individuals often were missing values attributable to "skipping" an interval or travelling shorter distances (i.e., did not travel to the further intervals).

Therefore, to determine if we can reject the null hypothesis that the median of difference in bearings is zero, an alternative test, the Skillings-Mack test, was deemed more appropriate (Chatfield & Mander, 2009) as it has been specifically designed to handle missing values within categories. Four categories were still appropriate since each category contained no more than 20% missing values. A further filter was applied to exclude individual legs that had more than 30% missing median values. The data was then processed using the *Skillings-Mack* (Srisuradetchai, 2023) R package which gave a base p-value. Due to the missing values, a Monte Carlo permutation was then performed with 10000 replicates to provide a simulated p-value.

The first interval (50-100 km Marion, 25-50 km Bird) was then dropped, and the test was performed again with the remaining three intervals to check for bias being introduced through the higher data availability of the first interval compared to subsequent intervals. In addition to the Skillings-Mack statistic, the data was subset into the same four intervals as the first Skillings-Mack test but included only legs which were represented in all four distance categories (i.e., no missing values). This resulted in the analysis of 164 complete legs (112 Marion, 52 Bird) which accommodates the assumptions of the more common Friedman test. This was performed to provide additional confidence to the results of the Skillings-Mack statistic. Finally, a Wilcoxon signed rank test was performed on the same subset used in the Friedman test for each consecutive pair of distance intervals to indicate any significant overall differences between intervals. This was then used to identify any potential pattern of change (which intervals were different from each other) in the data.

2.2.7 Lunar Fraction and Inbound Timing

Given the potential use of lunar cues in navigation exhibited in other animals, I wanted to determine whether lunar phase is related to the timing of the start of the return journey back to the island. A publicly available resource (U.S. Navy, 2023) that detailed the lunar fraction for each day in a specific time zone was used to retroactively calculate the lunar fraction at the start of each return leg. This was performed by taking the time stamp of the first point of each return leg (194 Marion, 128 Bird) and linking that to the lunar fraction of 12:00 am at the start of that day. The resulting data was first visualised as a histogram to show the distribution of data during each phase of the moon. The data was then put into a contingency table and a Chi-squared goodness of fit (with seven bins) was used to test the null hypothesis that the distribution of the lunar phase at the start of the return trip followed the expected sinusoidal distribution exhibited by the different phases of the moon. Expected proportions were calculated by collating every lunar fraction between January 1st of the first year of data analysis (2009 Marion, 1996 Bird) and December 31st of the last year of data analysis (2013 Marion, 2010 Bird) inclusive. Years that were not represented in the data set for each island were excluded from the expected proportion calculation. The resulting counts were then converted into proportions and compared against the observed proportions from the data. The test was carried out separately for both islands.

2.3 Results

2.3.1 Dispersal Direction

The dispersal direction analysis set out to answer two questions. First, is the pattern of outbound and/or inbound dispersal direction non-uniform? The second question was whether dispersal direction is the same between outbound and inbound legs. Whilst the Rao Spacing Test does not require paired legs, the intention was to compare between outbound and inbound directions with the same subset of data, therefore, the data was filtered to only include paired legs (matching outbound and inbound journeys). Marion Island had a total of 72 paired legs (Figure 5), while Bird Island had only 27 paired legs that fit the same criteria (Figure 6).

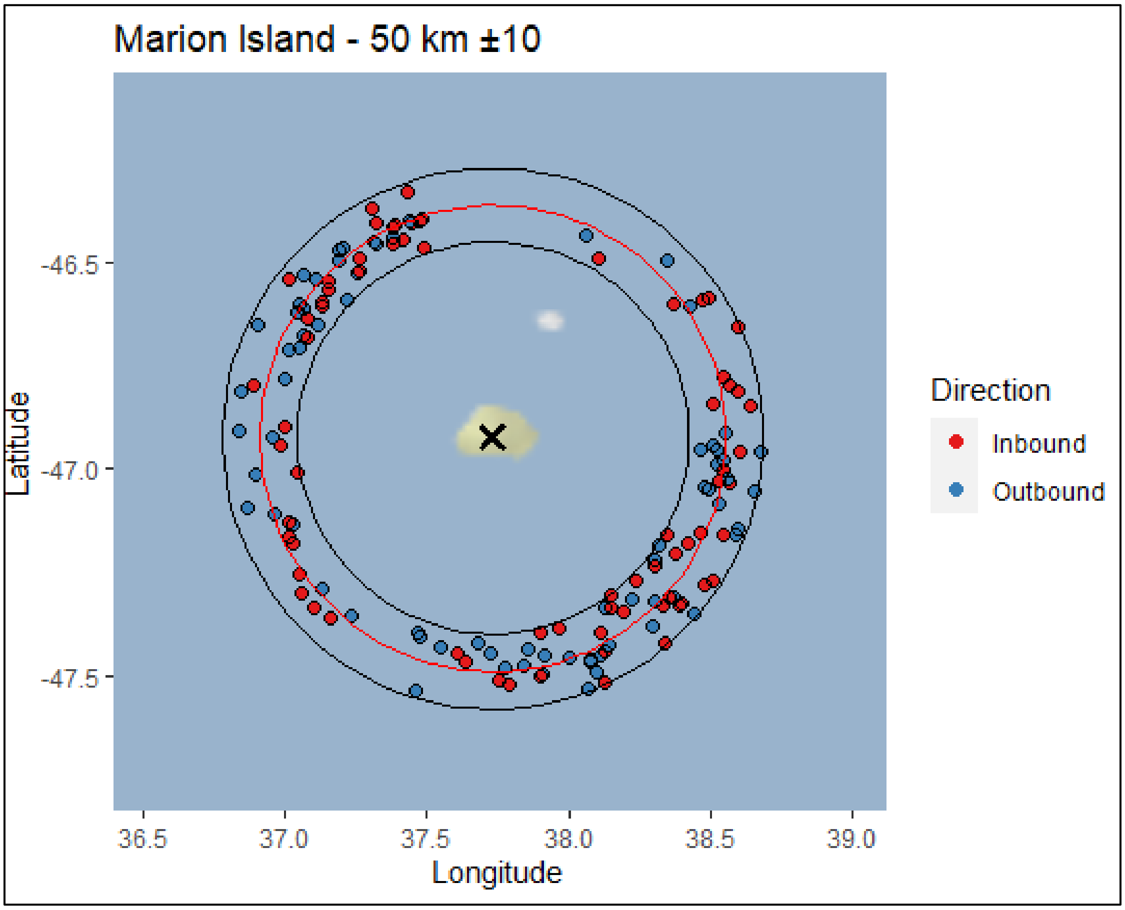

Figure 5. Heading of trips to and from Marion Island. The locations are given of each outbound (blue dots) and inbound leg (red dots) within the 50 km ± 10 km threshold. Only data from paired legs (representing a single matched inbound and outbound portion) was used. The "X" marks the centre of the threshold circle. Background map attribution: Stadia Maps (Stadia Maps, 2023).

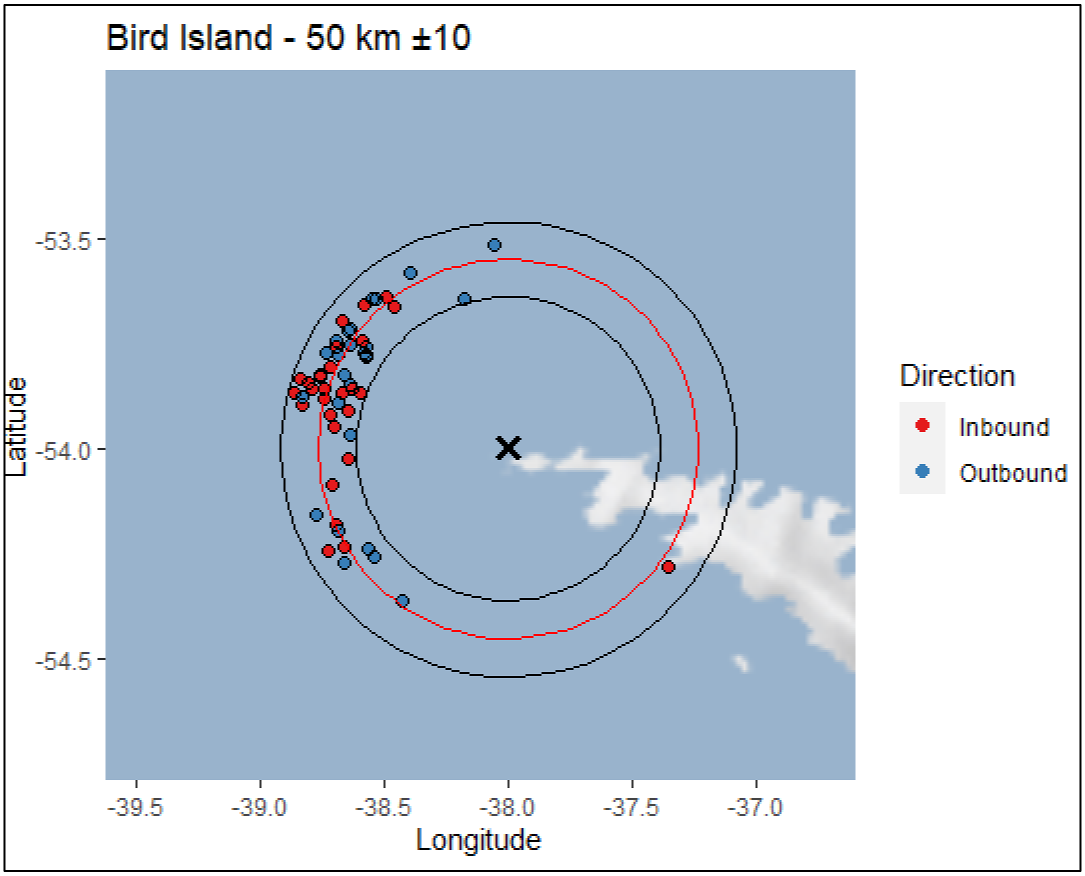

Figure 6. Heading of trips to and from Bird Island. The locations are given of each outbound (blue dots) and inbound leg (red dots) within the 50 km ± 10 km threshold. Only data from paired legs (representing a single matched inbound and outbound portion) was used. The "X" marks the centre of the threshold circle. Background map attribution: Stadia Maps (Stadia Maps, 2023).

Rao Spacing test for uniformity of the outbound dispersal at Marion Island was significant (T = 148.5042, $0.01 < P\text{-value} < 0.05$), meaning that the dispersal pattern was non-uniform. There was a clear bimodal distribution with one peak to the northwest and another to the southeast (Figure 7). The inbound dispersal however was not significantly directional, reflecting a more uniform distribution of headings (T = 148.2841, $0.05 < P\text{-value} < 0.10$). The results for Bird Island showed a significant departure from uniformity for both inbound and outbound legs (Inbound T = 251.239, Outbound T = 219.0984, both P-values < 0.001) with a unimodal distribution that was wider than either of Marion Island's distributions (Figure 7). The distribution of both legs was predominantly clumped to the west of the island (Figure 7).

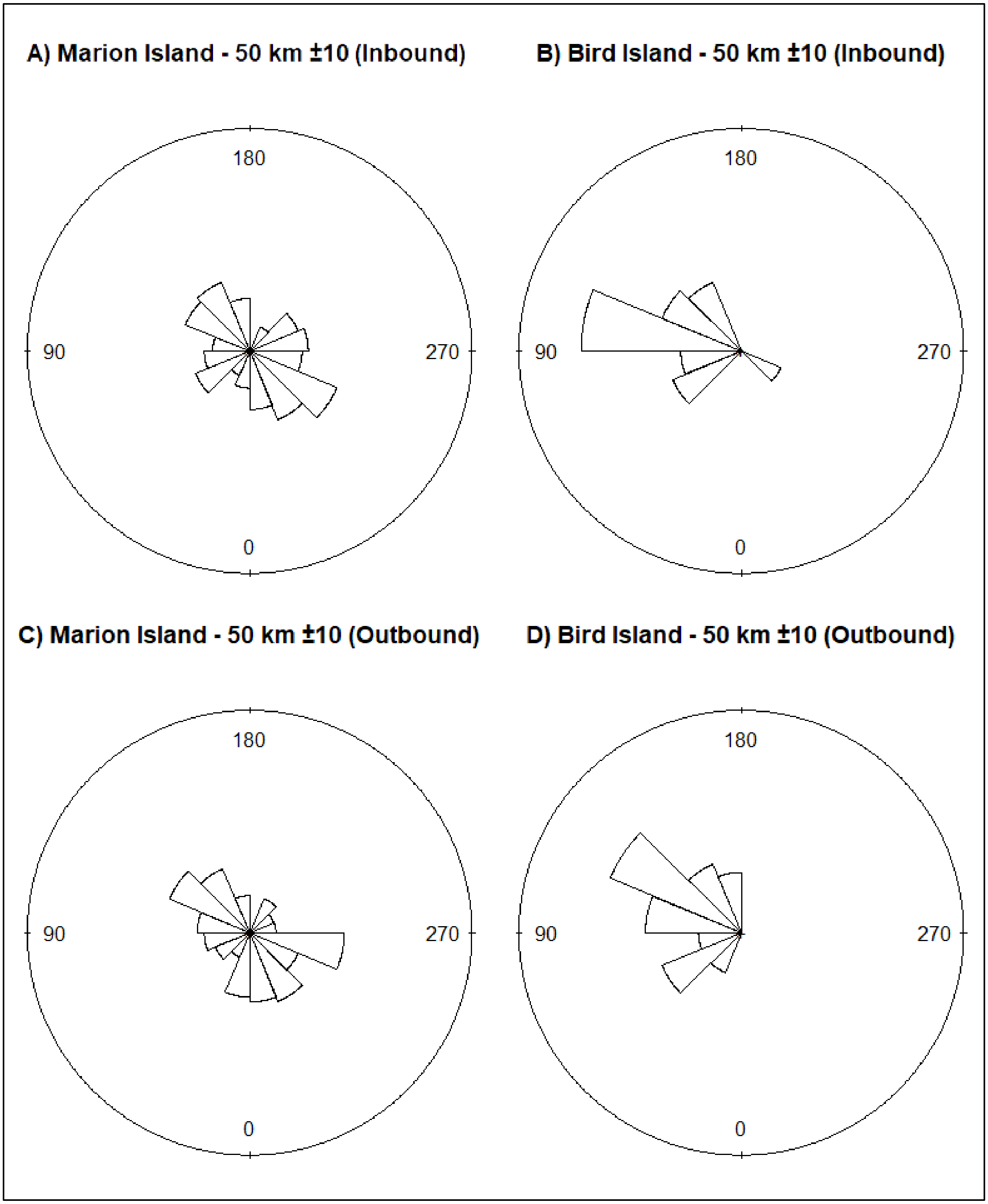

Figure 7. Rose plots for Rao Spacing Test. Plots show the distribution of headings relative to Marion and Bird Islands at 50 ± 10 km distance. Data is represented separately for inbound and outbound legs at both islands. Rao Spacing Test showed significantly non-uniform distributions for data represented in plots B, C and D, but not for plot A.

Given that we want to know if there is a relationship between outbound and inbound bearing deviation, the data was first represented in a scatterplot alongside the line of *outbound* = *inbound* for each island (Figure 8, Figure 9).

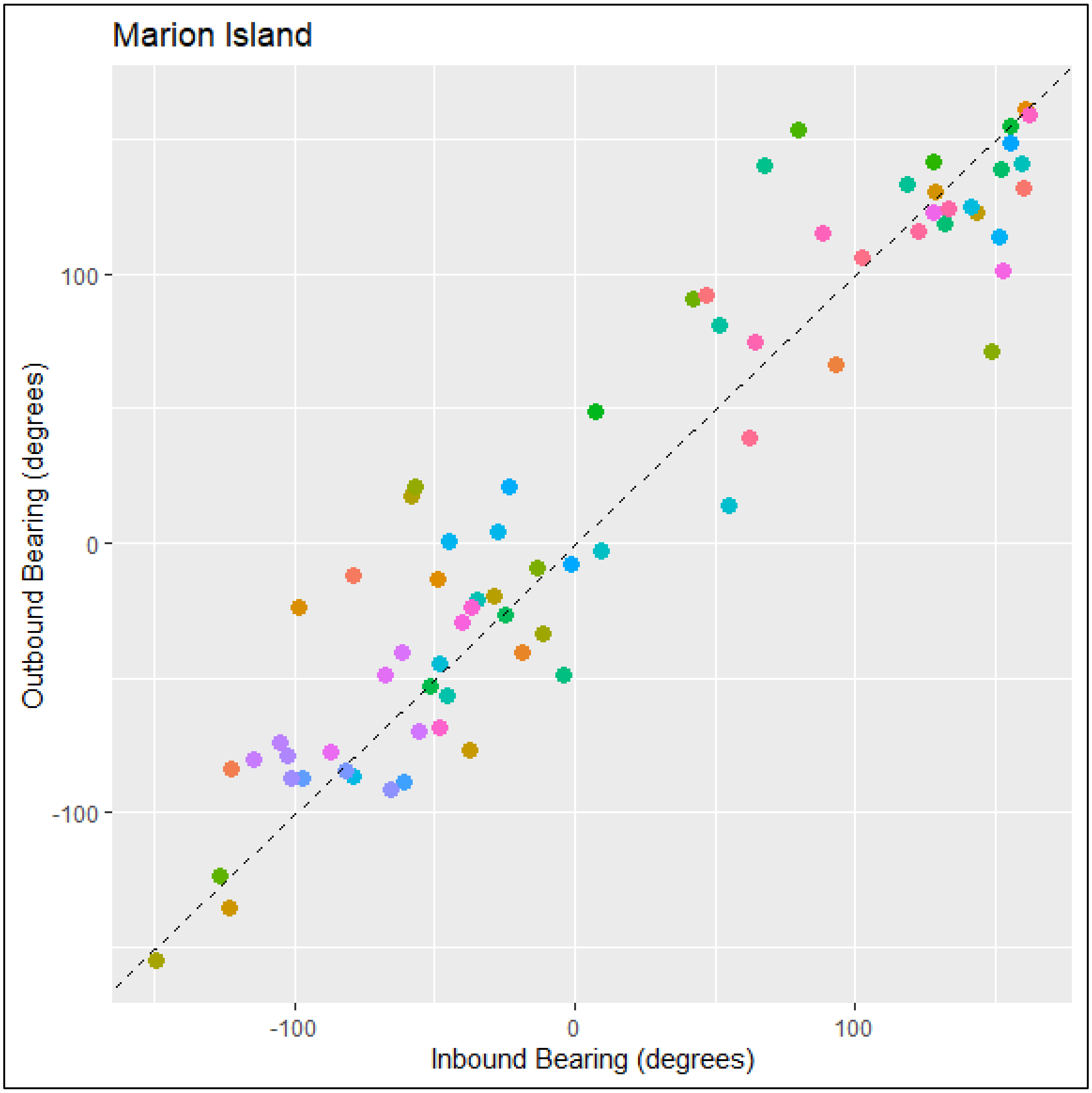

Figure 8. Marion Island Deviation Scatterplot. The scatterplot shows the data for the outbound and inbound bearings for each paired journey at Marion Island. Each point is a unique journey in the dataset. The black dotted line indicates the slope of the data if the outbound bearing is equal to the inbound bearing.

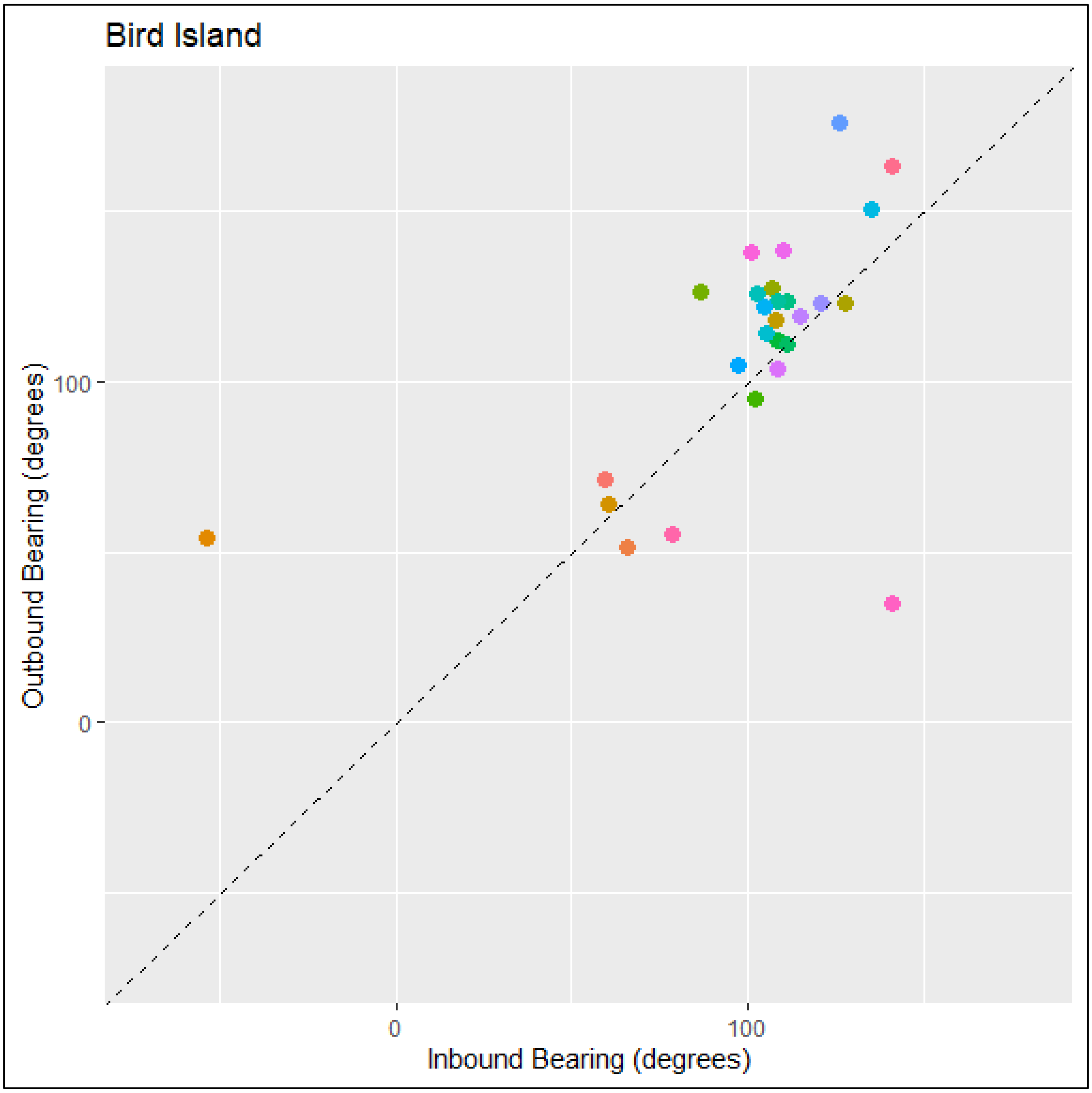

Figure 9. Bird Island Deviation Scatterplot. The scatterplot shows the data for the outbound and inbound bearings for each paired journey at Bird Island. Each point is a unique journey in the dataset. The black dotted line indicates the slope of the data if the outbound bearing is equal to the inbound bearing.

A *circular-circular correlation test* was then performed for both islands comparing the median of each island's inbound and outbound dispersals. For both islands, the median inbound and outbound dispersal directions were not significantly correlated, suggesting different dispersals patterns between the two inbound and outbound legs (Marion p = 0.45, Bird p = 0.61).

As a further descriptive analysis, I also calculated the difference in the angle between the outbound and inbound leg for each set of paired legs at a distance of 50 ± 10 km from the island. The deviation between paired dispersal legs for Bird Island showed a mean and median deviation of 22.3 and 14.6 degrees respectively. Marion Island showed a slightly greater difference in direction with mean of 24.6 degrees and a median of 18.5 degrees (Figure 10, Figure 11).

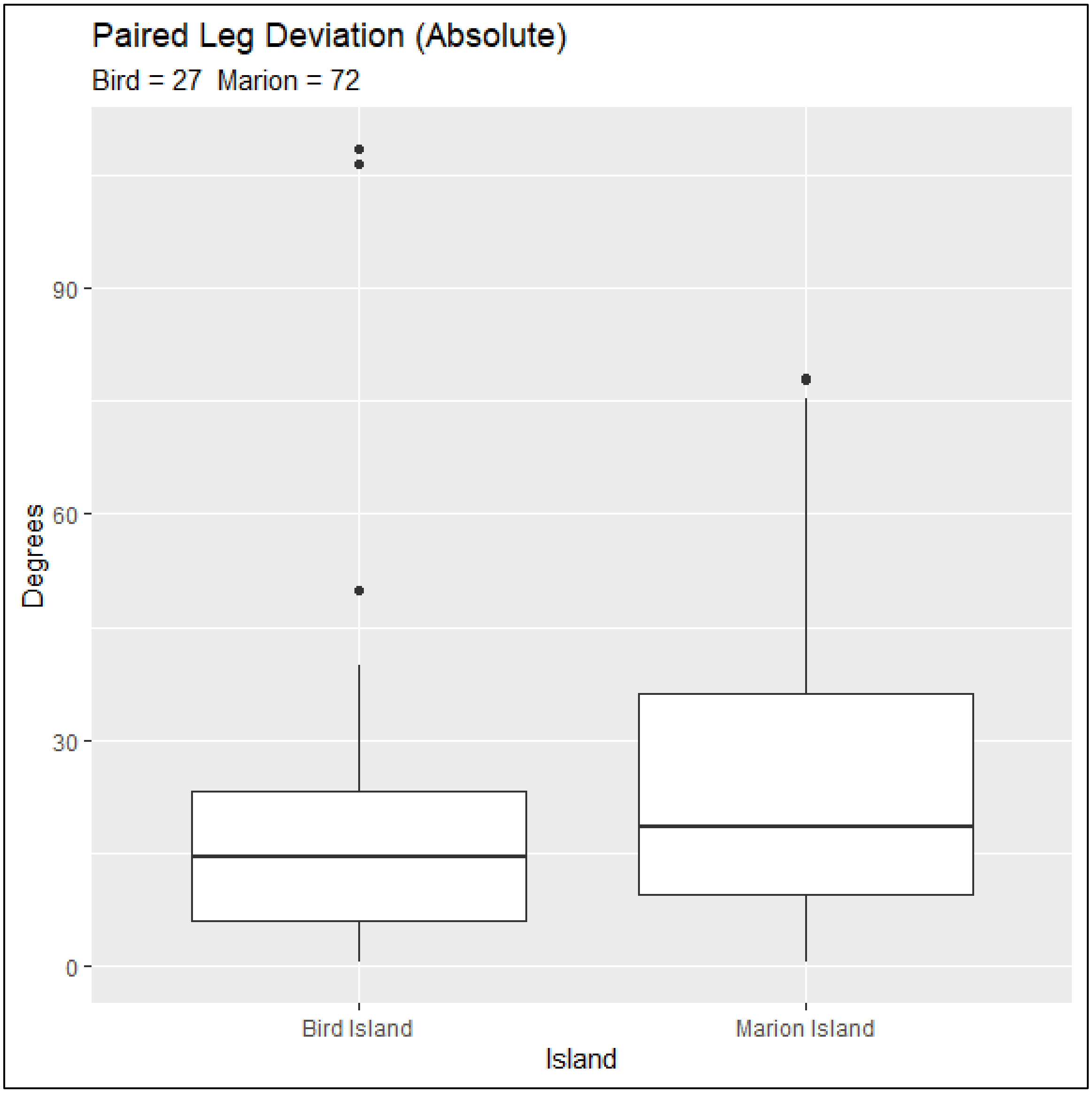

Figure 10. Absolute Bearing Deviation of Paired Dispersal Legs. The box plot illustrates the difference in headings (in degrees) between paired outbound and inbound legs for both islands. This plot uses absolute values in the differences to show the change regardless of direction.

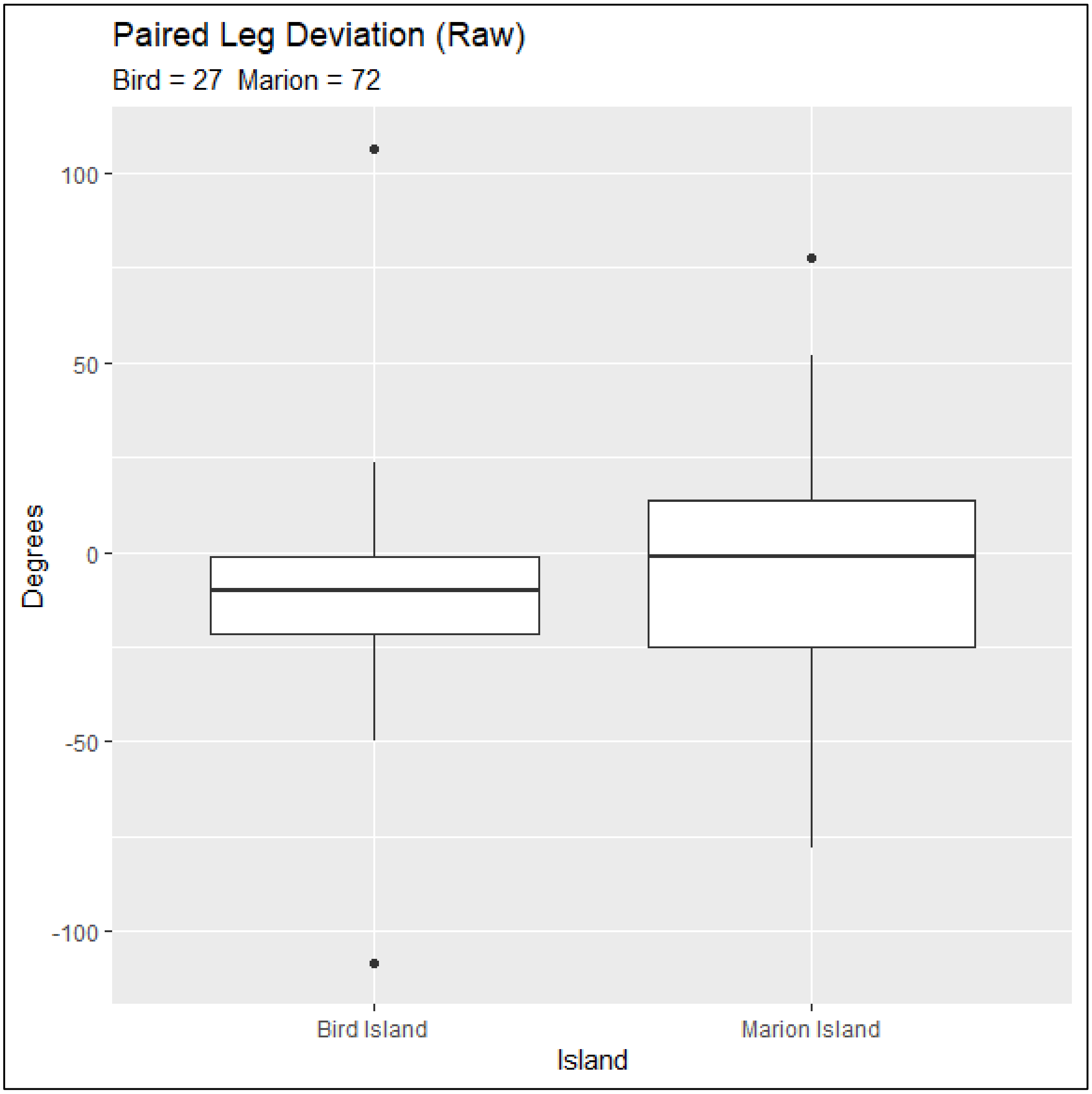

Figure 11. Raw Bearing Deviation of Paired Dispersal Legs. The box plot illustrates the difference in headings (in degrees) between paired outbound and inbound legs for both islands. This plot uses raw values to show direction.

2.3.2 Corridor Usage During Inbound Legs

Two visualizations were produced for the corridors analysis, one for each island. Marion Island showed two general corridor directions, one to the west, north-west and another to the east, south-east (Figure 12). This distribution aligns with the results shown in the previous dispersal direction analysis at 50 km for this island. As for Bird Island, there was one significant corridor to the West which matched the 50 km dispersal direction results as well (Figure 13). To calculate the utilization rate, the smallest rectangle that could be made using the furthest occurrence in each N/S and W/E direction was constructed based upon the inbound leg data. The number of grid points with occurrences was then divided by the total cells in the rectangle. By this metric, Marion Island may have had a higher utilization rate than Bird Island; however, Bird Island had a higher mean and median value (Table 3).

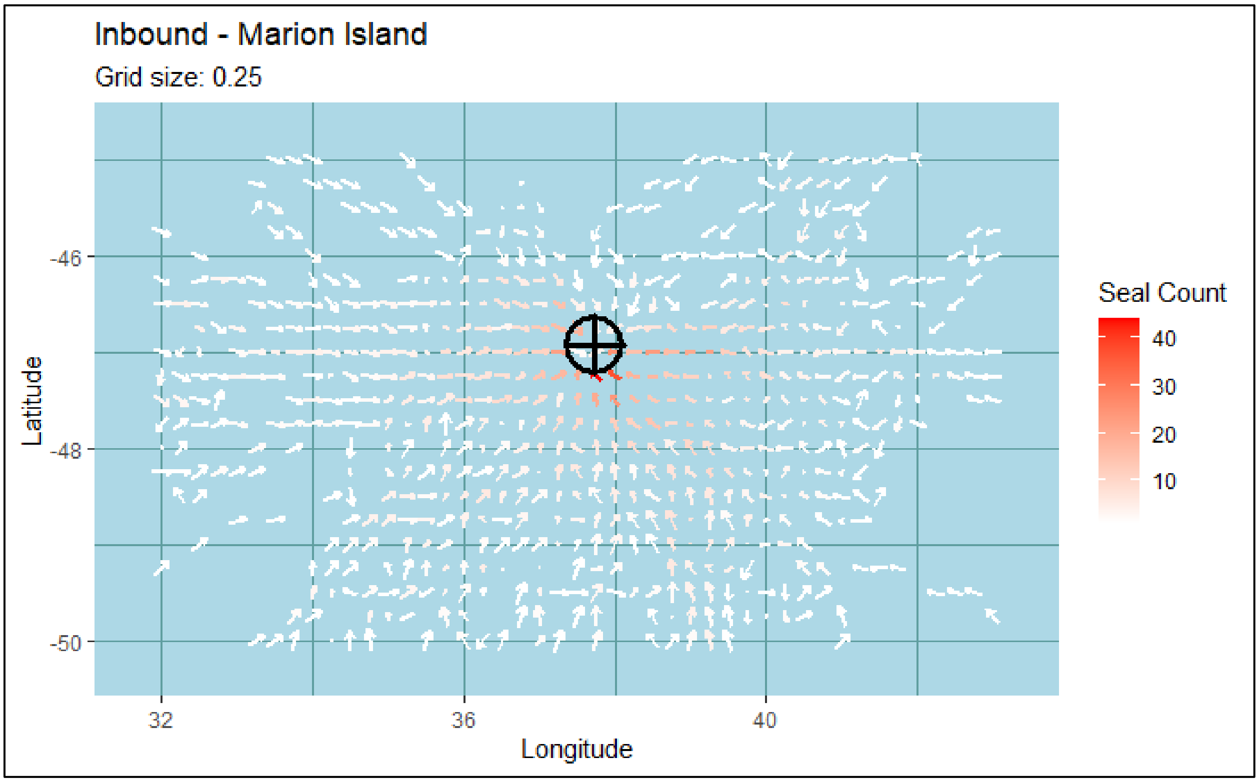

Figure 12. Marion Island Quiver Plot of Area Use. The plot illustrates areas of common usage (potential travel corridors) on return journeys to the island (indicated by the target icon). The colour of the arrow indicates the number of seal occurrences in each 0.25 Lat x 0.25 Lon grid. The direction of the arrow is the average heading. The size of the arrow is proportional to the average speed (larger arrow is faster movement).

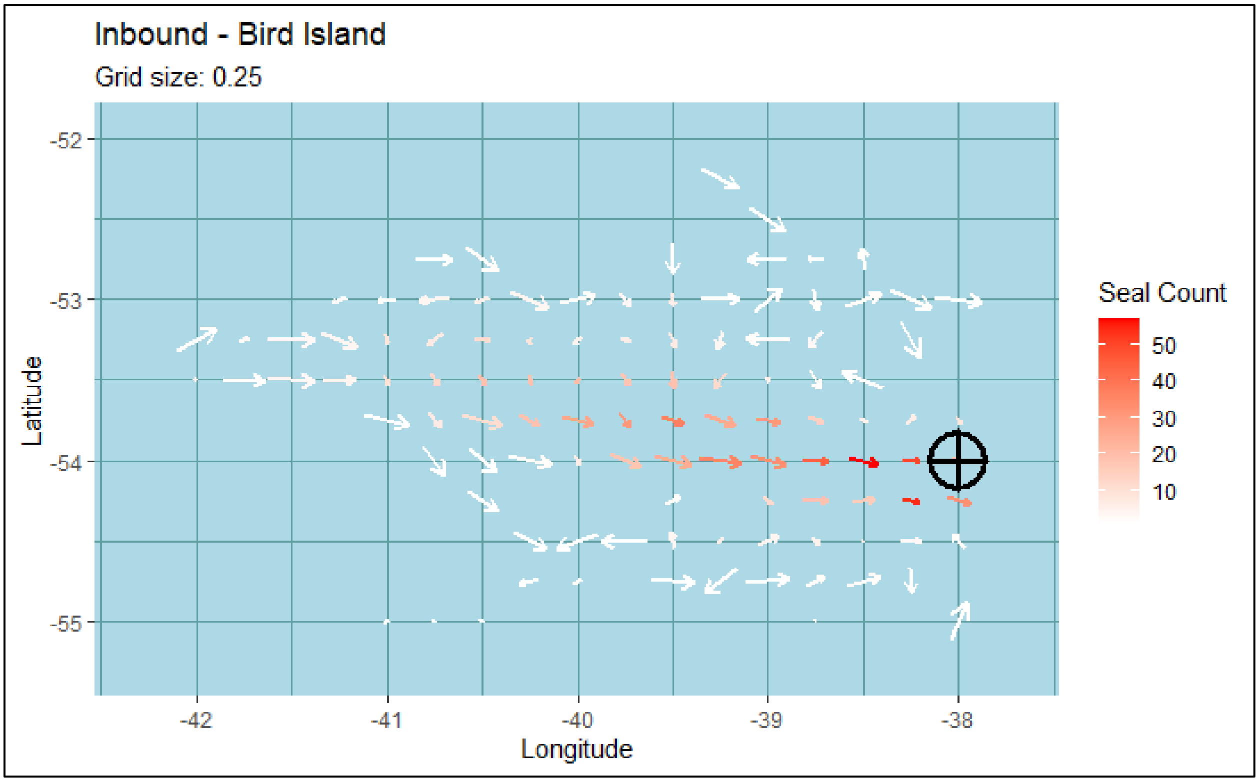

Figure 13. Bird Island Quiver Plot of Area Use. The plot illustrates areas of common usage (potential travel corridors) on return journeys to the island (indicated by the target icon). The colour of the arrow indicates the number of seal occurrences in each 0.25 Lat x 0.25 Lon grid. The direction of the arrow is the average heading. The size of the arrow is proportional to the average speed (larger arrow is faster movement).

Table 3. Corridor Grid Details. *Total Grids* is the maximum number of grid cells (0.25 Lat by 0.25 Lon) that enclose the return leg data boundaries, based upon the maximum and minimum values represented in the data. *Utilised Grid* is how many of the aforementioned grid cells have at least one seal occurrence. *Ratio* is the percentage of the total number of grid cells that were utilised. *Max*, *Mean* and *Median* refer to the number of seal counts (known seal locations) in each grid.

Island	Total Grids	Utilised Grid	Ratio	Max	Mean	Median
Bird Island	176	109	61.9%	59	9.0	4
Marion Island	880	611	69.4%	44	3.5	2

2.3.3 Bearing Offset at Various Distances

For both islands (Figure 14 & Figure 15), a series of statistical tests was undertaken to test whether the bearing offset for inbound legs changed with distance to the island. First, a Skillings Mack test

(10,000 replicates for simulated p-values) was performed on the first four intervals (50 km intervals at Marion from 50-250 km, 25 km intervals at Bird from 25-125 km). A second Skillings Mack test was carried out using the 2nd through 4th intervals (100-250 km Marion, 50-125 km Bird). These were followed by a Friedman Test on fully represented data only (no missing values) and a series of three Wilcoxon Rank Sum tests for testing differences in the bearing offsets between sequential interval pairs.

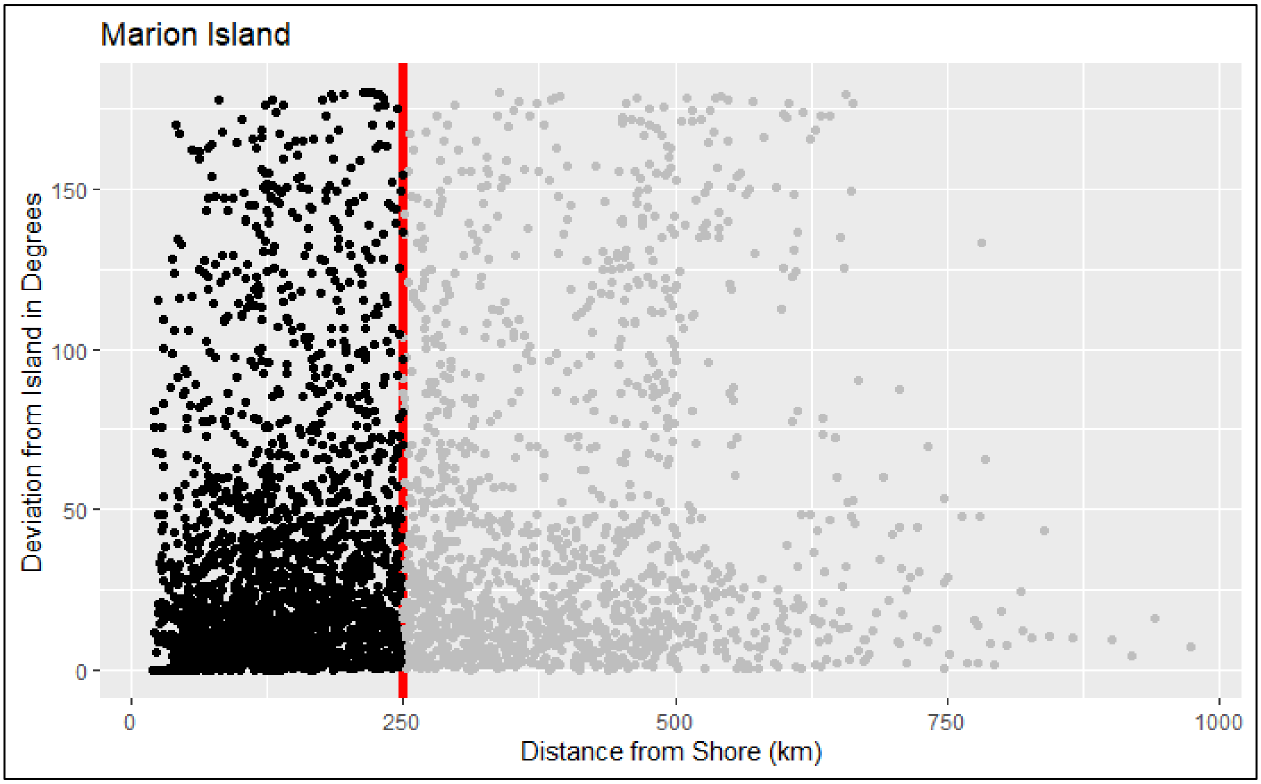

Figure 14. Distance vs Bearing Offset at Marion Island. The absolute value of the difference between an individual's heading and the direction to the island (in degrees) for each location up to 1000 km from Marion Island. Analysis was limited to a distance of 250 km (red line) to maintain the 70% completeness threshold.

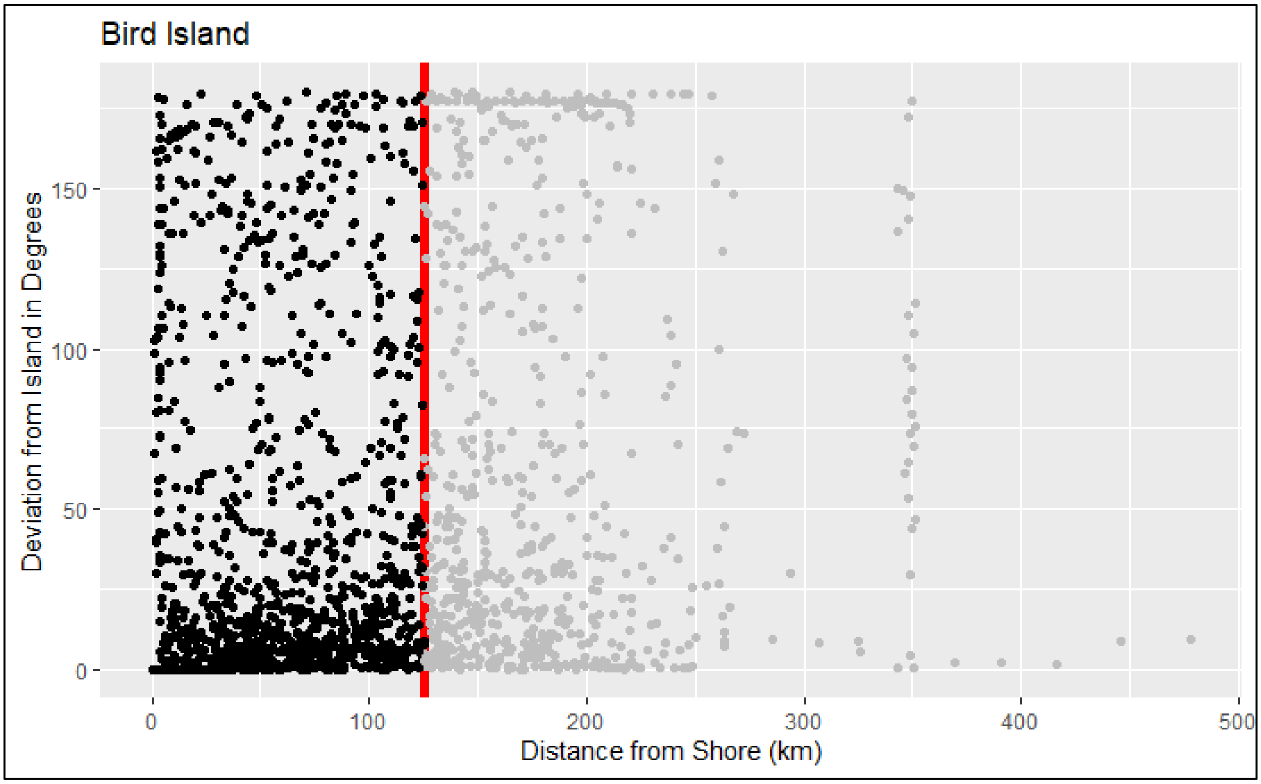

Figure 15. Distance vs Bearing Offset at Bird Island. The absolute value of the difference between an individual's heading and the direction to the island (in degrees) for each location up to 500 km from Bird Island. Analysis was limited to a distance of 125 km (red line) to maintain the 70% completeness threshold.

Due to the missing data, only the simulated p-value from the Skillings Mack test was considered as a statistical result. For Marion Island (Figure 16), the first Skillings Mack test was significant (SM = 47.43, df = 3) with a simulated p-value < 0.001. After removing the first interval, the result was still significant (SM = 20.079, df = 2) with the simulated p-value < 0.001. Together, these results suggest that there were significant differences in mean bearing offset among the intervals, and that this result was not solely the result of the proximity of the seals to the island within the first interval. The Friedman test on the subset of data with no missing values was also significant (FM = 233.748, df = 111, p < 0.001). The Wilcoxon tests for pairs of intervals suggested a significant difference between intervals 3 and 4 (distances 150-200 and 200-250 km, median offset = 16.55 and 19.81 degrees respectively) (W = 2429.5, p = 0.014), but not between intervals one and two or between intervals two and three (W Statistic = 2655.5, p = 0.893 and W =2768.5 p = 0.186, respectively).

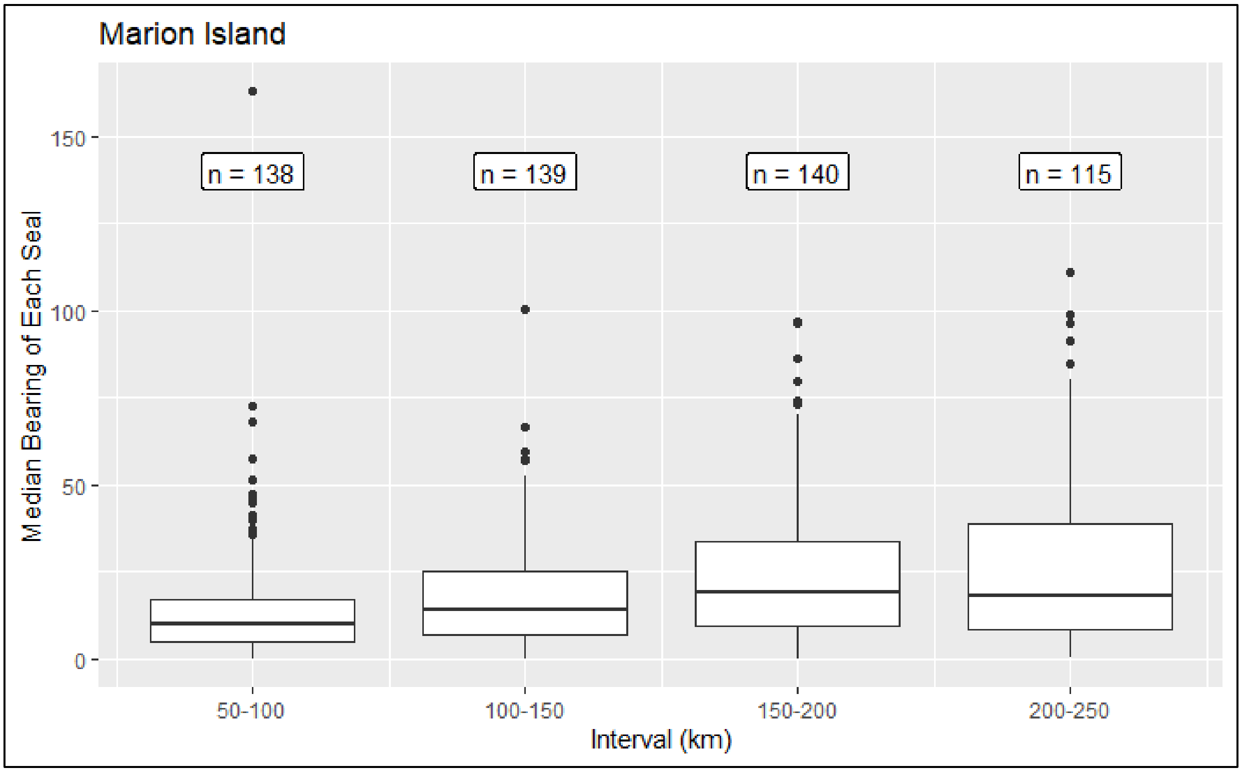

Figure 16. Bearing Offset by Interval at Marion Island. The plots show the distribution of the median value of the absolute difference between an individual's heading and the direction to the island within each interval at Marion Island. Each interval is labelled with the number of AFS represented in that interval.

Similar results were found for Bird Island (Figure 17), with the Skillings Mack test on all intervals being significant (SM = 28.072, df = 3, simulated $p < 0.001$). Significant differences remained after the removal of the first interval (SM = 7.915, df = 2, simulated p = 0.021). The result of the Friedman test was also significant (FM = 95.31, df = 51, $p < 0.001$). Similar to Marion Island, the Wilcoxon test for Bird Island only showed a significant difference between the last two sets of intervals (distances 75-100 and 100-125 km, median offset 10.26 and 25.94 respectively) (W = 450.5, p = 0.013), but not between intervals one and two or between intervals two and three (W Statistic = 609.5, p = 0.408 and W = 556.5 p = 0.167, respectively).

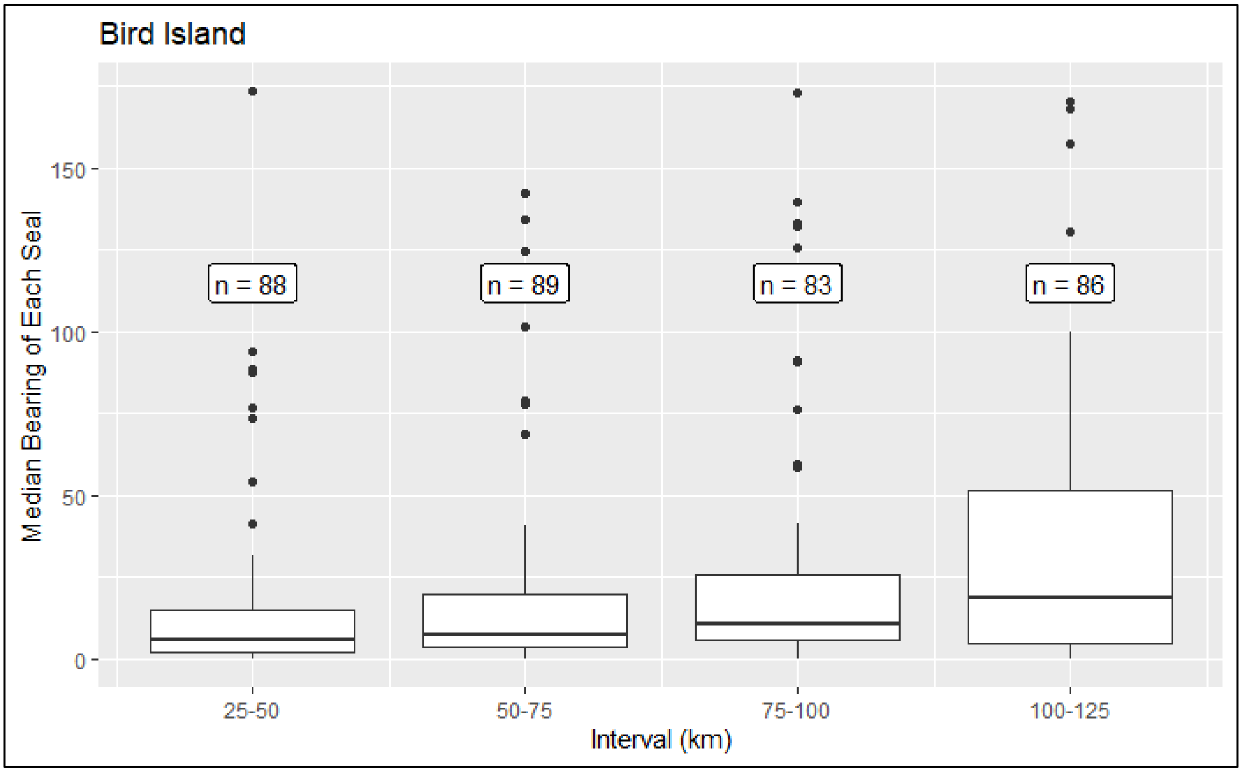

Figure 17. Bearing Offset by Interval at Bird Island. The plots show the distribution of the median value of the absolute difference between an individual's heading and the direction to the island within each interval at Bird Island. Each interval is labelled with the number of AFS represented in that interval.

2.3.4 Lunar Fraction and Inbound Timing

To test whether seal return departure date and lunar fraction were related, a Chi-square goodness of fit test was conducted at both islands (Marion Figure 18, Bird Figure 19), with the null hypothesis being that departure dates do not favour certain lunar fractions. The analysis uses counts of return and departure dates pooled across years. Lunar cycles at the start of the inbound leg were assigned to seven bins. Taking into account the underlying sinusoidal nature of the lunar cycle, the expected proportion of days in a particular lunar fraction (that is, in a particular bin) were calculated based on the years of matching data (Table 4). The distribution for seals returning to both Marion Island (X-squared = 11.183, df = 6, p-value = 0.083) and Bird Island (X-squared = 1.924, df = 6, p-value = 0.927) were not significantly different than expected. At Bird Island, there was very little deviation from the expected values. The lunar fraction present at the start of inbound journeys for individuals at Marion Island showed a higher peak towards the quarter moons (0.5) but there was not enough of a difference to be significant.

Table 4. Expected Proportions of Lunar Fractions. Expected proportions for the data (assuming random distribution of initiation of return trips) for each lunar fraction bin based upon the years of collected telemetry data. Fractions based on five years of data for Marion Island and 14 years for Bird Island.

Island	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7
Marion	0.182	0.151	0.111	0.112	0.118	0.142	0.185
Bird	0.185	0.148	0.109	0.112	0.118	0.141	0.186

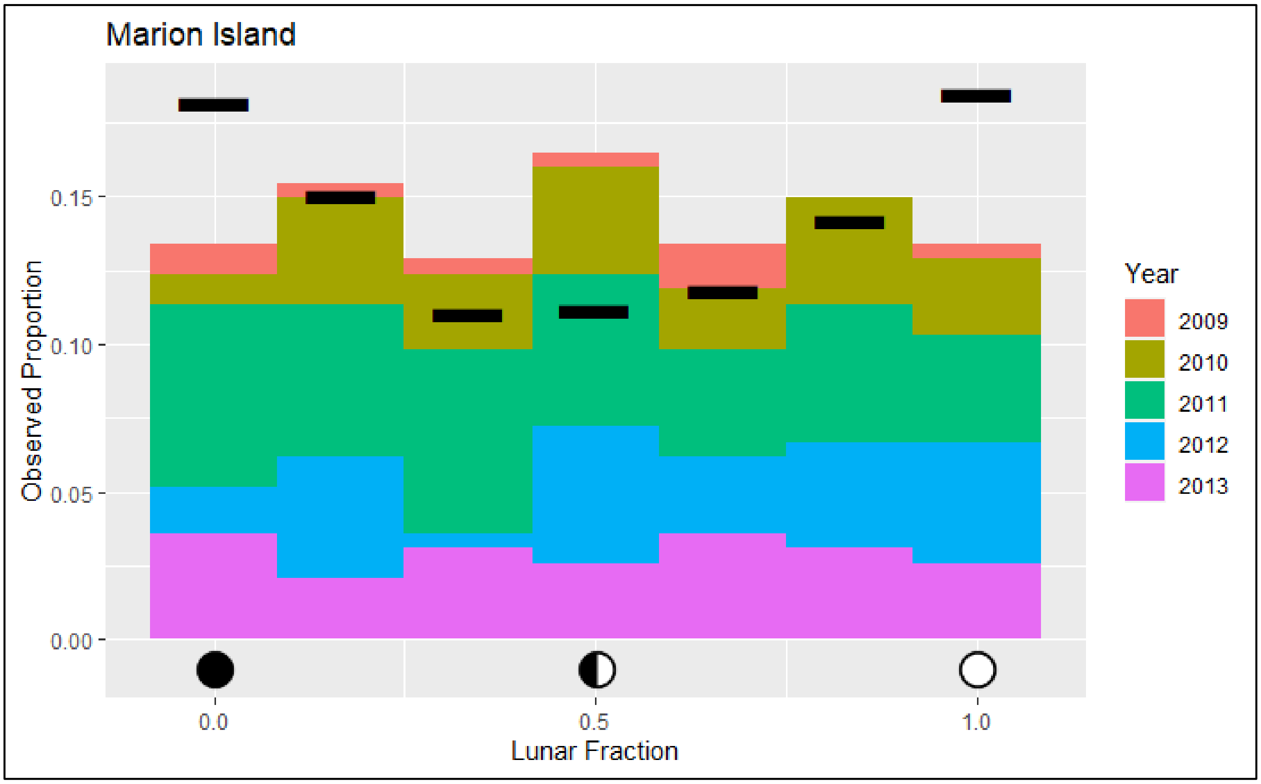

Figure 18. Marion Island Return Trip Lunar Fractions. The bars indicate the proportions of the observed lunar fraction values at the start of all return legs from Marion Island, differentiated by study year. The black lines indicate the expected proportions of lunar fractions if return was randomly distributed across the lunar cycle. These were calculated from the known lunar phase values for every day during the years the data was collected (Table 4). Statistically, the observed values were not significantly different from the expected values (p-value = 0.083).

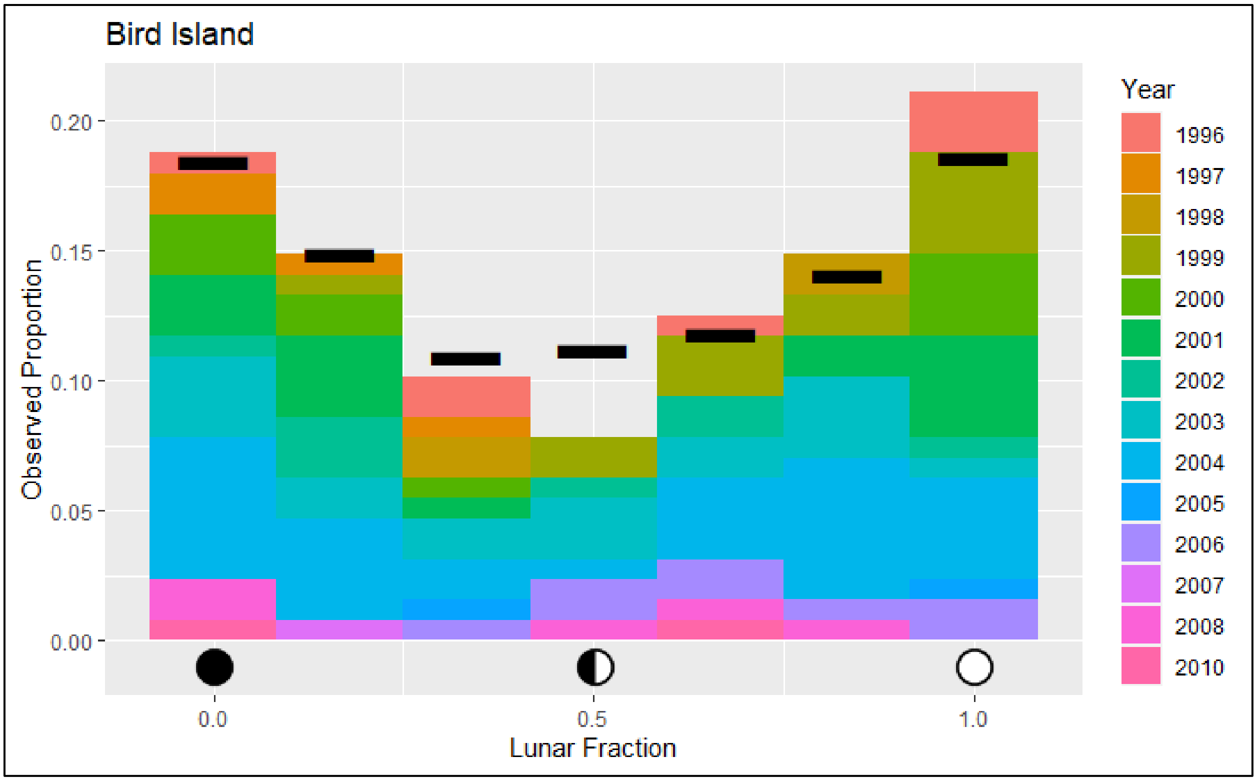

Figure 19. Bird Island Return Trip Lunar Fractions. The bars indicate the proportions of the observed lunar fraction values at the start of all return legs from Bird Island, differentiated by study year. The black lines indicate the expected proportions of lunar fractions if return was randomly distributed across the lunar cycle. These were calculated from the known lunar phase values for every day during the years the data was collected (Table 4). Statistically, the observed values were not significantly different from the expected values (p-value = 0.927).

2.4 Discussion

Existing research into AFS movement has focused primarily on identifying key foraging locations. It has been demonstrated that AFS have a high degree of site fidelity both for foraging locations (Arthur et al., 2015; Bonadonna et al., 2001; Bonadonna et al., 2000; Staniland et al., 2004) and their originating haul-out location (Hoffman et al., 2006). It has been suggested 'that it is where an individual forages, not who that individual is' (Staniland et al., 2004) that dictates dive behaviour, meaning that the location has a potentially larger impact on decision-making than phylogenetic characteristics. A previous study at Marion Island (as well as Cape Shirreff) showed that individuals foraged more in regions of windier conditions (Arthur et al., 2017), while individuals from Bird Island foraged in areas of lower wind speeds. Studies that investigate AFS navigation beyond foraging location are scarce. One study that investigated directed movement in pups prior to moulting showed that individuals follow a star shape as they explore the area surrounding their home location during this narrow timeframe (Nagel et al., 2021). Little more is known about this species' navigation methods beyond this isolated study. In other words, existing research has primarily shown that we know *where* they navigate to but not *how* they successfully navigate there and back again.

I set out to show how we can use existing research in a scientific process that asks a set of questions that will characterize the movement of AFS navigation and eventually lead to a better understanding of what methods and environmental cues they are using. This relates to the concept of switching the investigation from 'understanding variability' to 'defining the norm' (Sequeira et al., 2019). Given there are a myriad number of ways to characterize and quantify the movement patterns of animals such as the AFS, I selected the following series of questions based upon their suitability to my specific dataset. Although related, these questions are addressed somewhat separately given the spatial and temporal differences inherent in each question.

i. Dispersal Direction - Is near-island dispersal direction away from and towards the island directed or uniform, and are the inbound and outbound legs correlated?
ii. Corridor Usage During Inbound Legs - Are multiple individuals following common route(s) home?
iii. Bearing Offset at Various Distances - Does the angle of change required to be facing the island decrease as individuals get closer to shore?

iv. Lunar Fraction and Inbound Timing - Do AFS start their return leg during a specific lunar phase?

In answering these questions, I have also shown that – through the power of meta-analysis and data sharing – we can re-use existing data to identify patterns in navigation beyond just *where* AFS travel. My findings provide the stepping stones needed to begin investigating the potential cues and driving forces of navigation. Specifically, by identifying potential navigational patterns in AFS, future research can be directed specifically towards testing likely mechanisms of navigation that fit those patterns.

2.4.1 Dispersal Direction

Given the seemingly featureless surface and endless horizon of the open ocean, my first question focused on the near shore dispersal of individuals. The question relates to whether seals head off in a specific direction in pursuit of likely prey concentrations, or whether they randomly head away from the islands in search of food. It is also important to consider that a difference in an individual's heading by just a few degrees near shore can lead to a significant difference in displacement across a single journey. To determine if dispersal direction near the island is directed or random, each journey was split into outbound and inbound legs to investigate dispersal for each direction, as well as to compare outbound and inbound dispersal patterns. This resulted in three specific null hypotheses. The first two null hypotheses were that individuals show an even distribution when departing or returning to the island, respectively. The third null hypothesis was if these two inbound and outbound legs of a journey showed no correlation between each other. These hypotheses were tested independently for each island.

For departure, the null hypothesis of uniform distribution in their outbound trips was rejected for both islands. Seals from Marion Island showed a bimodal dispersal distribution, with one cluster of departures to the northwest and another to the southeast. Seals from Bird Island displayed a unimodal dispersal distribution in their departing trips, with a significant cluster of trips departing to the west of the island. My findings show that for the outbound dispersal from both islands, the individuals are following directed movement. It is logical to assume that this distribution is a result of them heading in the direction of their foraging grounds. However, this also means that the individuals must have some knowledge or cue they follow to know *how* to get to these foraging grounds. As seals obviously cannot directly detect the prey fields at this distance, this implies that seals require (i) some sort of knowledge of where to head and (ii) some means of orienting to that heading. This further implies a location or directional memory (with or without and internal map) and some mode of sensory input to result in the individuals heading in this direction. The advantage of the dataset I am using is that it only includes complete journeys, that is, journeys that successfully return to shore. This means I was able to study the successful attempts at navigating as opposed to potentially unsuccessful ones. While this may have biased the results, that is, it is possible that "unsuccessful" individuals have a more random heading distribution, logically their numbers within a population would be very small so any bias would be negligible.

To get a more complete picture of their distribution pattern, I also needed to examine the characteristics of their return journeys and how these compare to their departure trips. Upon analysing the inbound legs, Bird Island was found to have a significantly non-uniform distribution. This means their distribution was not evenly spread around the island but instead was characterised by a preferred specific direction or directions. However, unlike outbound legs, there was insufficient evidence to reject the null hypothesis for the Marion Island inbound legs. Although it was not reflected by the statistics, visually the inbound dispersal for Marion Island still approximately follows the same bimodal distribution as the outbound legs. The wider distribution for the inbound legs is to be expected due to the difference in the types and level of movement of the return trips prior to reaching the 50 km threshold. Unlike individuals leaving from shore who have travelled very short distances to reach this analysis point, individuals returning to the island have undertaken journeys that have spanned many days and hundreds of kilometres. Therefore, it would make logical sense that they are more dispersed as a population, leading to the decreased observed degree of clumping. The less distinct patterns of arrival heading on return journeys may also be a result of correction for their heading as they progress towards the island. The idea of the potential for course correction as they approach shore will be discussed further below (2.4.3).

Perhaps more importantly, the *circular-circular correlation test* did not provide sufficient evidence to reject the third null hypothesis for both islands (outbound is equal to inbound bearing), despite visual analysis seemingly indicating that the departing dispersal was similar to the return journey at both islands (Figure 7). This contradicts the expected result, presuming that there would be a relationship between the patterns in departing and return headings. It was assumed that individuals heading off in one direction to reach a foraging ground are more likely to return from that direction, however the statistics did not support this. The summary values of paired leg deviation for Bird Island showed a median deviation 14.6 degrees, while the median deviation for Marion Island was 18.5 degrees. Such differences are not insignificant. At a distance of 200 km from the island, these median differences in headings between outbound and inbound legs work out to a displacement of a little over 50 km for Bird Island and a little over 64 km for Marion Island respectively (see Figure 20 for an example). Not only does that suggest that AFS are following substantially different outbound and inbound routes, but the potential environmental correlates that they are using could change drastically given this level of displacement. This also implies that the individuals could have the ability to keep track of their movements over time (see inertial navigation, 1.1) or are following some kind of map or environmental correlate that exists across this scale.

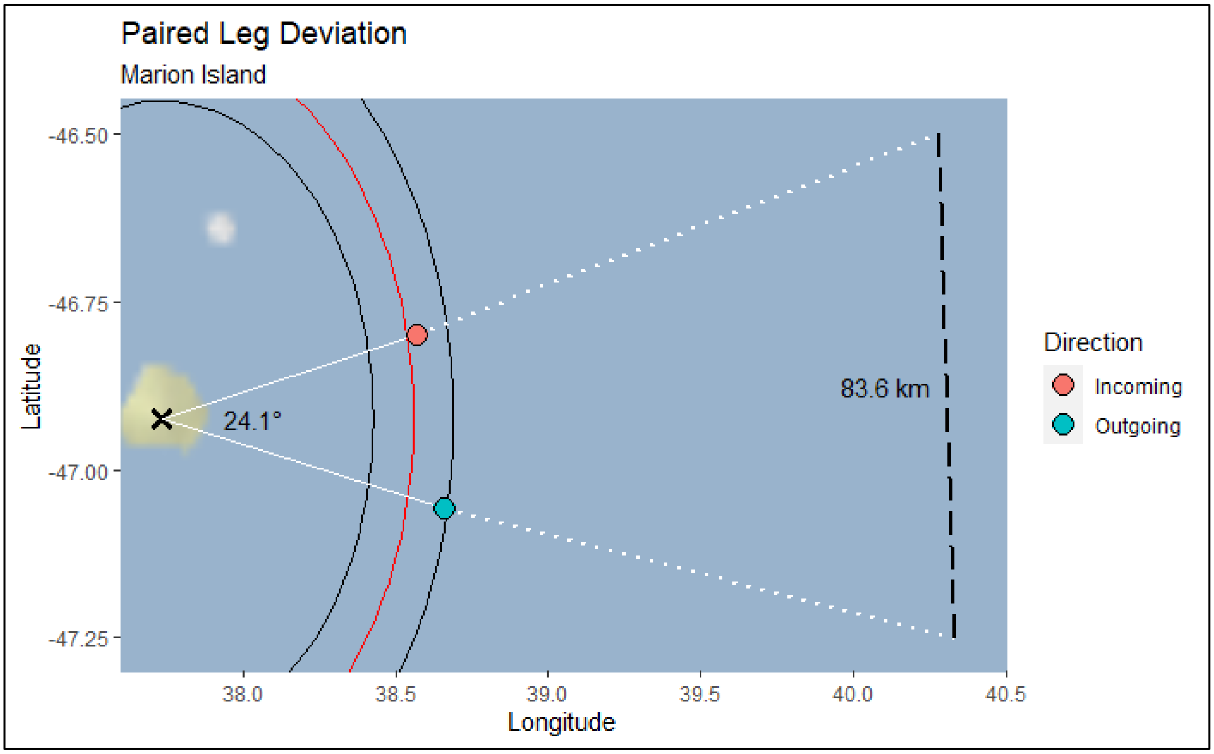

Figure 20. Illustration of Paired Leg Deviation. This schematic shows an example of the effect of difference in headings on paired outbound and inbound headings for an AFS journey. The solid white lines connect the location of the rookery with the seals' location for inbound and outbound journey legs at the 50 km ± 10 km distance. The angular difference between the two points is labelled at the island. The white dotted lines extend in the same direction to the distance of 200 km from the island. The black dashed line indicates the potential difference in distance between outbound and inbound legs at this distance.

These findings show that individuals are displaying directed movement. This provides a basis for further studies using region-specific dispersal directions to try and determine what factors cause these patterns.

2.4.2 Corridor Usage During Inbound Legs

The results of the dispersal analysis led to my next question: are multiple individuals following common routes home? This was investigated via a visual inspection of area utilisation to identify potential corridors. A corridor in this instance is defined as an area of high utilisation rates, higher speeds and of consistent heading. This specifically excludes foraging grounds as foraging is typically a slower movement rate than transiting behaviour. Given the complexity of quantifying a corridor, I focused on identifying a method of visualising the data so that any potential patterns could drive subsequent research into specific correlates (see 3.2 for potential adaptations of this methodology).

After research into various methods, I settled on an adaptation of the quiver plot. Typically used for meteorological studies in wind analysis, a quiver plot uses a grid system to provide a summary of the wind conditions in a region. An arrow in each grid cell is used to indicate the wind characteristics such as direction and intensity. In my case, the length of the arrow for each cell indicated the average speed, the direction of the arrow was the mean bearing, and the colour of the arrow showed the number of legs that are represented by the data. One quiver plot was generated to display the return legs for each island. Visually, one looks for high use "corridors" of adjacent cells that are characterised by a common, high speed directed movement back towards the island.

The return journey plot for Marion Island (Figure 12) does appear to show some potential corridor usage, particularly in the south-east direction. Whilst not statistically tested, this corridor roughly overlaps with the rose diagrams from the dispersal analysis (Figure 7). The other mode in the bimodal distribution also appears to be present in the quiver plot, albeit to a lesser extent. The return leg plot of Bird Island (Figure 13) also appeared to be visually related to the findings from its respective rose diagram (Figure 7). Of particular interest is how highly directed the arrows become in the area of high AFS presence at Bird Island. The arrows on the fringes of the utilisation area are a little haphazard in direction. This is due to individuals travelling in opposing directions within this cell. However, for the stream of red cells heading west from the island, the arrows are consistently funnelling towards the island. To me, this indicates a high chance of corridor usage. Marion Island on the other hand, does show similar funnelling as seal presence increases, but the funnelling is not as narrow. This is not surprising given the broader dispersal pattern demonstrated in my previous analysis. Therefore, my results suggest that individuals are indeed following common routes home specific to each island. Subsequent research is required to determine what environmental cues and sensory information they are utilising to follow these corridors.

It is worth mentioning that, whilst quiver plots are not a statistical test, the utilisation rate calculated from these plots likely have sufficient statistical power to accurately describe area usage. Arthur et al. (2017) demonstrated that 50 individuals are enough to asymptote the curve of new seals against newly visited cells at Marion Island. Hence, I have confidence in using this visual inspection as a starting ground for investigating environmental correlates.

2.4.3 Bearing Offset at Various Distances

Given the general visual trend of the arrows becoming more directed as the individuals approach the island, my next analysis was to determine if this apparent pattern is statistically significant. Specifically, I asked the question: does the bearing offset to the island decrease as individuals get closer to shore? That is, when an individual is on their return leg, does the difference between the heading they are travelling and the direction to the island decrease as they get closer to their rookery? Since seals inevitably successfully reach the island, if we see no differences in their travel heading at different distances, this would indicate that an individual's heading is highly accurate in reference to the island's position throughout their journey. Looking for changes in directedness will not only provide insight into their travelling behaviour but can also indicate potential notable geographic locations and/or changes in behaviours for further analysis. For example, if there is a sudden change in directedness at a certain point, it might indicate that they are correcting their heading using a specific cue. This could involve landmark recognition or changes in (or introduction of) an existing cue such as, geomagnetic signals, gravity fluctuations, or ocean currents. If the change in heading is more gradual, this could indicate that individuals are progressively correcting their position through some degree of map sense.

To answer the question of whether directedness changes during a return trip, I used a block design that divided the distance home and compared the median heading each fur seal within the intervals. However, this was not straight forward given the nature of the tracking data and required three types of statistical tests using different subsets of the data to ensure efficacious results. The results of the Skillings Mack tests – with and without the first distance interval – indicated a change in median heading between some interval or intervals for both islands, a result confirmed using the Friedman Rank Sum test with a subset of the data.

However, both of these tests can only show that a change occurs, not where or in what direction that change occurs. To provide a cautious look into this question (given the risk of Type 1 error in multiple testing), a Wilcoxon Rank Sum test was then performed between consecutive interval pairs. For both islands, the Wilcoxon Rank Sum reported insignificant differences in mean bearing offset between the first two interval pairs (intervals 1-2 and 2-3), indicating there was no evidence of change in mean directedness for both hypotheses. That is up to a distance of 100 km from Bird Island, and up to a distance of 200 km from Marion Island. The finding that the pairing between intervals three and four was significant for both Marion Island and Bird Island indicates a change in directness past these distances for both islands. A cursory look at the physical environmental characteristics defined by these intervals for both islands suggests that the individuals are too far from shore and the water is too deep for likely landmark recognition. This could indicate the seals are switching between navigational methods used due to an unknown cue in this region. While there is not enough evidence at this stage to suggest what methodology the fur seals might be using, it does seem plausible that multiple methods are being employed for navigating.

By looking at the median values of each interval, I can infer that the there is a generally decreasing trend indicating they are moving in a more directed fashion towards the island. Given that it is a population wide trend, it is possible that the individuals are utilizing the suggested corridors highlighted in 2.4.2 as they return to their island rookeries. Due to the need for adequate statistical power when answering these questions (Sequeira et al., 2019), future studies would benefit from a higher sample size to ensure there is adequate information at various distances to the island to include more intervals that defined a finer scale. The bearing offset analysis shows that whilst there is directedness in their approach to the shore, it is just the first step in determining *how* the AFS navigates.

2.4.4 Lunar Fraction and Inbound Timing

Unlike the other analyses which primarily focused on finding patterns in the movement of the fur seals through their environment, the lunar fraction analysis was performed to try and provide insight into one of the potential cues that the AFS could be using. Studies with northern elephant seals indicate that return journeys commence at specific times relative to distance from the island, although what cues this behaviour is unknown. The timing of migrations in humpback whales may be affected by lunar cycle (Horton et al., 2020). Hence, I hypothesised that the timing of an individual's return could be dependent on the amount of the moon that is visible. It has been shown in other studies (Mauck et al., 2005; Mauck et al., 2008) that harbour seals can orient themselves relative to a lodestar. Therefore, it is at least plausible that AFS could also use it in their navigational bag of tricks. If there is a relationship between time of return and lunar fraction, this could indicate a potential clock and map sense that informs an individual how far it has travelled and how long it takes to get there or return home.

Fortunately, lunar fraction data is readily available as it is derived from an algorithm that considers the date and the animal's position on the earth to calculate the lunar fraction in that region. It does not, however, consider whether the moon is actually visible or not (i.e., inclement weather). After determining the lunar fraction for each return leg, the Grouped Chi-square goodness of fit test for both islands were insignificant, showing insufficient evidence to reject the null hypothesis. In fact, it is interesting how similar Marion Island was to the expected values (Figure 18). Biologically this is understandable as, due to the inconsistencies of moon visibility, it would likely be too unreliable to use as a calendar. Bird Island does show some, albeit insignificant, deviations from the expected distribution (Figure 19); however given the same reasoning as for Marion Island, there is likely no biological significance of the deviation. This is likely due to the return time being more dependent on physiological factors (particularly whilst lactating) or that other environmental cues are contributing to the timing of the start of the return trip.

Furthermore, it is not known whether individuals can distinguish between the two quarter moons. Lunar fraction does not take into account the difference between first and last quarter and therefore, for the purpose of this analysis, it was assumed that AFS are unable to distinguish the difference as well. If they could, it would have led to greater change in values around the 0.5 lunar fraction value. However, due to the observed pattern in the initial results, this idea was only briefly considered and there were no indications of changes in the results with these alternative assumptions. It is important to note that this result does not rule out the use of the lunar cues in the AFS navigation. It just highlights that, if they do use lunar cues, they use them in a different way to what I hypothesised.

2.4.5 Context within Existing Literature

There have been previous studies of at-sea movement of AFS that have touched on aspects of navigation. For example, it has been shown that there are potential bathymetric links with route of travel (Bonadonna et al., 2000), as well as bathymetric linking to foraging areas (Arthur et al., 2017). Individual sex also plays a role in distance travelled to reach foraging areas with males foraging further from shore than females during the breeding season (Boyd et al., 1998). Furthermore, it is hypothesised that memory plays a role in destination choice (Bonadonna et al., 2001). In terms of migration, we know AFS migration destinations during the winter, post breeding season period (Bamford et al., 2021; Boyd et al., 1998; Boyd et al., 2002; March et al., 2021; Staniland et al., 2011). Typically, the selection of these at-sea locations is driven by the presence of a food source in the area. In general, the majority of current AFS research revolves around where the AFS navigates to as opposed to how they navigate. It is difficult to determine the cause of this. Partly, this reflects current scientific priorities, but I suspect the lack of focus into how they navigate is a result of the inherent difficulties that are present when studying marine mammals (see 1.3), combined with the difficulties in publishing studies with statistically non-significant findings (Murad et al., 2018). This research itself had its own insignificant result in relation to the findings of my lunar fraction analysis. However, the novel findings found through my meta-analysis are crucial in taking the next steps towards understanding how AFS navigates.

My finding that AFS exhibit distinct dispersal patterns and directed movement has also been seen in other pinnipeds such as the Subantarctic fur seal (Wege et al., 2016), which displays a strong relationship between season and direction of foraging. The Weddell seal is another species that has displayed directed movement. They have been shown to display directed movement both with and without the presence of landmarks, albeit on pack ice as opposed to underwater (Fuiman et al., 2021).

Corridor usage in terrestrial animals, in particular various species of bird (Biro et al., 2004; Gill et al., 2009; Zimmerman & Peterson, 1998), has been well documented for multiple decades now, and has started to gain more attention in recent years in marine animals. The majority of research of corridor usage in the ocean environment has primarily focused on long distant migrants (Gavrilov et al., 2018; Horton et al., 2017; Pendoley et al., 2014; Shillinger et al., 2008). A particular outcome of these studies has been to utilise potential corridors to define important regions for marine protected areas. Other examples of corridor usage in marine mammals include the channel usage of West Indian manatees (Cloyed et al., 2019), and the foraging preferences of the harbour seal (Hastie et al., 2016). My study is likely the first example of potential corridor usage in long-distance movement in an otariid.

Whilst my research did not find a significant result pertaining to preference for certain lunar fractions for starting their return journey, lunar cues such as a relationship between dive depth and lunar fraction have been shown in other species of pinnipeds such as the Galapagos fur seal (Horning & Trillmich, 1999) and the northern fur seal (Lea et al., 2010; Sterling et al., 2014; Zeppelin et al., 2019). However, in these cases it was primarily suspected to be as a result of following prey availability, such as has been shown in a species of shark (Vedor et al., 2021). One specific study focusing on the humpback whale asked a similar question to mine and subsequently found that humpbacks will initiate their migration in relation to the lunar phase (Horton et al., 2020). In terrestrial species, there has been evidence of a migration specific relationship to the lunar fraction for the European nightjar birds (Norevik et al., 2019). Therefore, the lack of significant results relating start of return trip to lunar phase does not preclude lunar phase from serving as either a temporal or directional environmental navigation cue in AFS long-distance movements.

Some studies have shown more complex forms of navigation strategies in other pinnipeds. The ability to return to feeding grounds after translocation has been displayed by Australian fur seals (Hume et al., 2002), whilst some species have displayed both memory of themselves and their social inhabitants (Pitcher et al., 2010; Smeele et al., 2019). The harbour seal has shown a potential cognitive map as a result of its ability for landmark recognition (Maaß et al., 2022) and the northern elephant seal is suspected to use a map sense to navigate (Beltran et al., 2022). The presence of directed movement through the dispersal and bearing offset analyses could indicate more advanced forms of navigation such as inertial navigation or a map sense. As a result, this opens the door to future research studying these potential strategies in AFS.

2.4.6 Conclusions

My research set out to provide the starting blocks for investigating AFS navigation using pre-existing location data. Specifically, I sought to identify movement patterns that would provide clues as to what cues they might use in long-range navigation. This process involved two main stages of work, these were (i) through the use of meta-analysis, design a methodology for identifying patterns in navigation from existing data, and (ii) demonstrate the effectiveness of this approach through a novel species, the Antarctic fur seal. As a result of this research, there is evidence that multiple patterns do exist in how the AFS navigates. More explicitly, I have shown that, (i) through the use of dispersal direction, AFS displays highly characteristic directed movement away from the rookeries, (ii) bearing offset analysis indicates that their directed movement allows them to course correct as they return to the island, and (iii) that AFS display corridor behaviour when navigating. This opens the door to future studies to specifically test for the mechanisms utilised by AFS that allows them to successfully navigate in the open ocean. Ultimately, this research has provided more questions than answers. However, these findings are a necessary step leading to informed decision-making on the avenues of research going forward with the goal of answering *how* the Antarctic fur seal navigates.

3 Exploratory Data Analysis

3.1 Movement vs Environment

One of the objectives of my research was to quantify patterns in AFS movement that would help to identify the potential environmental cues that are responsible for these patterns. This section will provide a high-level example of one potential approach. This section does not delve into a statistical analysis to quantitatively evaluate any potential relationship between environmental cues and AFS movement, but rather provides a starting point by comparing environmental data – specifically bathymetry, ocean temperature, surface current and magnetic inclination – to the trends observed in my corridor analysis (2.2.5).

Specifically, I found a publicly available dataset for each one of the potential environmental correlates and overlayed the information onto the corridors plot for both islands. This approach has the benefit of rapidly comparing multiple different environmental and movement patterns through a visual representation of the data. Whilst each potential relationship would ultimately require a more rigorous, in-depth quantitative analysis, this precursory overview will serve to highlight some of the advantages and potential disadvantages of using this approach with specific environmental cues that could be potentially used by AFS for at-sea navigation.

3.2 Potential Environmental Correlates

3.2.1 Bathymetry

It is suspected that some cetaceans (Zapetis & Szesciorka, 2022) and sirenians (Burgess & Evans, 2022) use visual cues such as bathymetry and landmarks to navigate through their environment. Both of these methods are examples of piloting behaviour (detailed in 1.1). While above-water landmarks (such as islands) could serve as a potential cue in specific instances, bathymetry is a more pervasive physical characteristic of the ocean environment and could potentially be used as a navigational cue by AFS. It has already been shown that foraging grounds of the AFS often include areas with distinguishing bathymetric features (Arthur et al., 2017), so it is not unreasonable that bathymetry could also serve in their navigational behaviour. In theory, AFS could use bathymetry by following a change in slope or specific features and landmarks such as canyons and seamounts. However, overlaying the corridors plot with the data provided by the R package, *marmap* (Pante & Simon-Bouhet, 2013) shows mixed results. For example, Marion Island which was the island where the role of the bathymetric information for foraging grounds was previously identified, shows little apparent relationship between bathymetric features or change in bathymetry and the corridors utilised by individuals (Figure 21). Bird Island, on the other hand, does appear to show some degree of relationship between bathymetry and corridor usage. It appears that individuals are heading up the bathymetric gradient, towards shallower waters while returning to the island (Figure 22). However, I hypothesise that this is likely a coincidental relationship given that the AFS only dive a maximum of a couple hundred of metres (Forcada & Staniland, 2018). Given that the corridors displayed are in waters far deeper than their potential visual acuity would allow, if corridors are being utilized in the open ocean, they must be using a different cue to navigate through them. Granted that while this is not a conclusive analysis, I suspect that bathymetry is unlikely a significant driving force of how the AFS navigates.

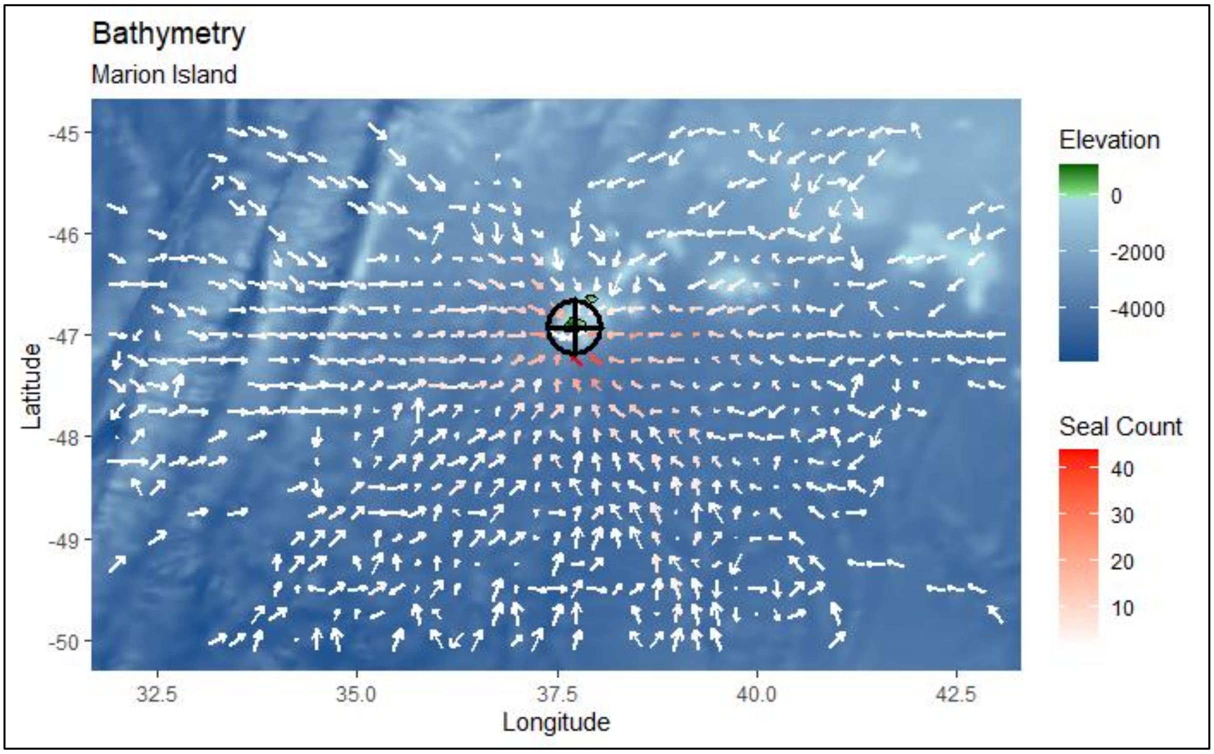

Figure 21. Marion Island Bathymetry Corridors. Quiver plot overlayed on top of the surrounding bathymetric features for Marion Island. Note, water depth is provided as negative elevation from sea level. Marion Island is indicated by the target. Bathymetry data from the *marmap* package (Pante & Simon-Bouhet, 2013).

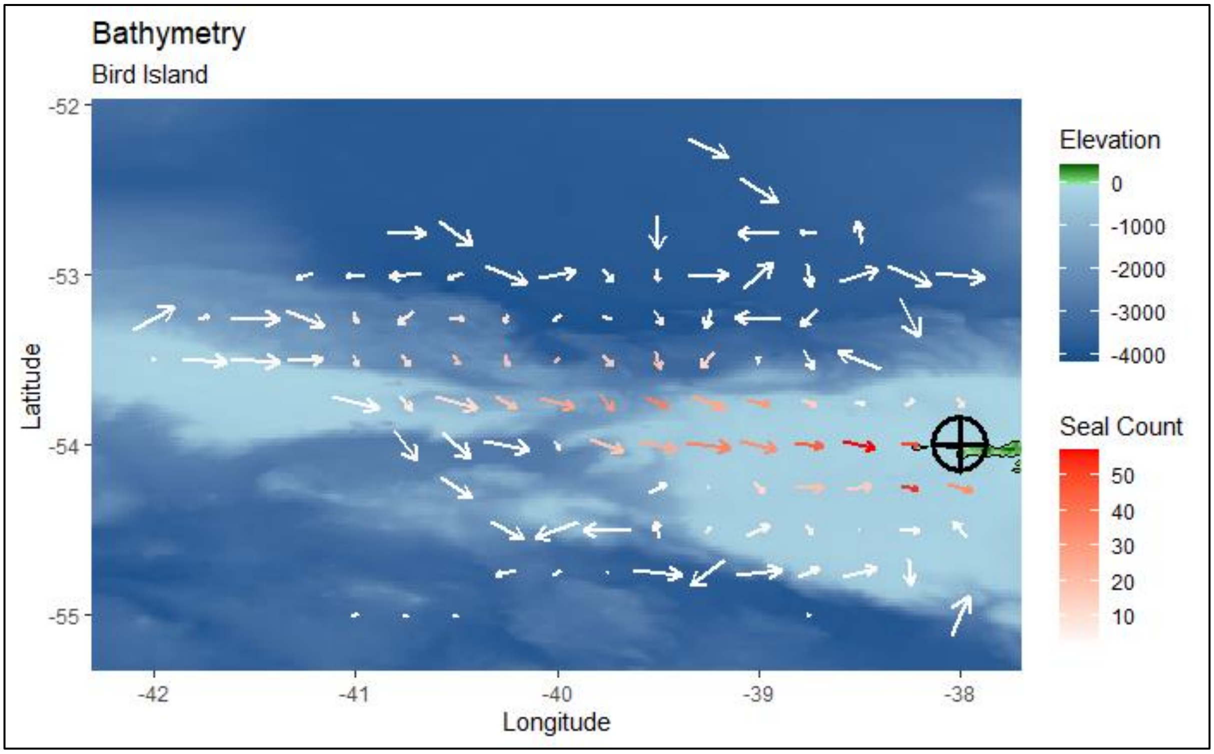

Figure 22. Bird Island Bathymetry Corridors. Quiver plot overlayed on top of the surrounding bathymetric features for Bird Island. Note, water depth is provided as negative elevation from sea level. Bird Island is indicated by the target. Bathymetry data from the *marmap* package (Pante & Simon-Bouhet, 2013).

3.2.2 Water Temperature

Since AFS spend the majority of their time in the surface layer, I hypothesize that two other oceanographic variables more likely serve as cues for navigation: water temperature and ocean currents (See 3.2.3). Surface ocean temperature data is readily available online and similar studies have shown that the temperature can drive navigation in various marine mammalian species (Rasmussen et al., 2007; Sheppard et al., 2006). Surface temperature has already been shown to be a predictor of foraging effort for AFS (Arthur et al., 2017) and is considered ‘critical for thermoregulation’ (March et al., 2021). Therefore, it is possible that it could be utilised in other aspects of their behaviour such as navigation. I predict that, if a relationship occurs, it is likely that they would follow a thermal gradient, given that surface water temperature is correlated with latitude in the Southern Ocean (Arthur et al., 2017). However, despite the availability of data, it is a little more difficult to investigate the relationship of temperature to movement than bathymetry.

This is due to the constant changing nature of temperature as a result of seasonal variation and bathymetry (Boyd et al., 2001), as opposed to the consistency of bathymetric features across time. Therefore, to provide a conclusive analysis of the impacts of temperature, the data must correspond to the time and location of each data point to ensure that the temperature experienced by the individual is impacting their decisions, rather than relying on a historical average. Despite this added complexity, it is crucial to understand if temperature is an environmental correlate used by AFS due to the changes induced as a result of climate change. For example, humpback whales have been shown to navigate by following waters of similar temperature, and increased temperature could drastically change their typical journey. To simplify the efforts required for the example I have provided, I utilised the 2020 annual ocean temperature data at 2.5 m depth as provided by Bluelink Reanalysis (BRAN). For Marion Island (Figure 23), the two dispersal directions roughly line up with a change in gradient. Of particular interest is the observation that one direction is an increasing gradient whilst the other is decreasing, which is an impressive feat if they are following specific disparate gradients based on the direction of their dispersal. Bird Island (Figure 24) also appears to show a rough link between return travel corridors and temperature gradient, with the corridor roughly aligning with an area of consistent temperature. However, despite the shorter time scale of the data used in this example, future research that wishes to explore ocean temperature as an environmental correlate will need to ensure the data is more attuned to the timing of the individuals. This also presumes that the AFS can detect a temperature change at this level of precision. Preliminary studies determining their thermal detection sensitivity would help to indicate if a higher resolution of temperature data is required or if they could even use this type of environmental cue.

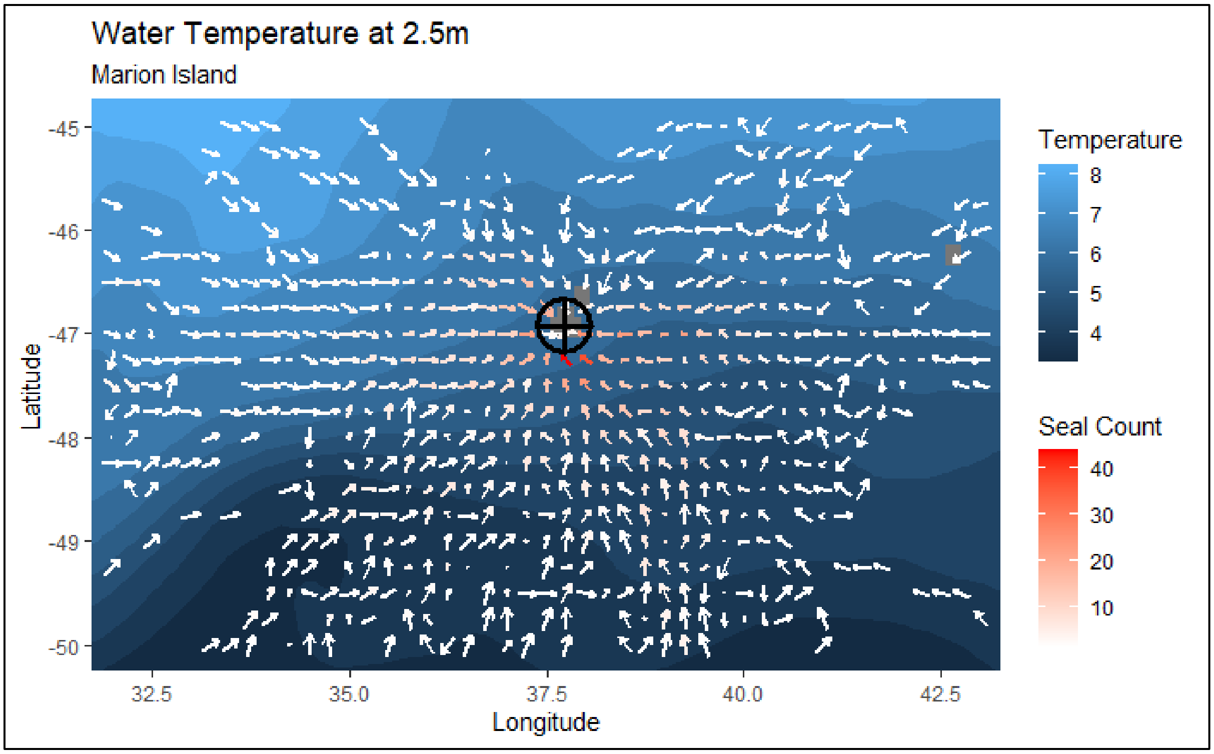

Figure 23. Marion Island Ocean Temperature. Quiver plot overlayed on top of the surrounding ocean temperature for Marion Island. Marion Island is indicated by the target. Ocean temperature data from BRAN2020 (Chamberlain et al., 2021).

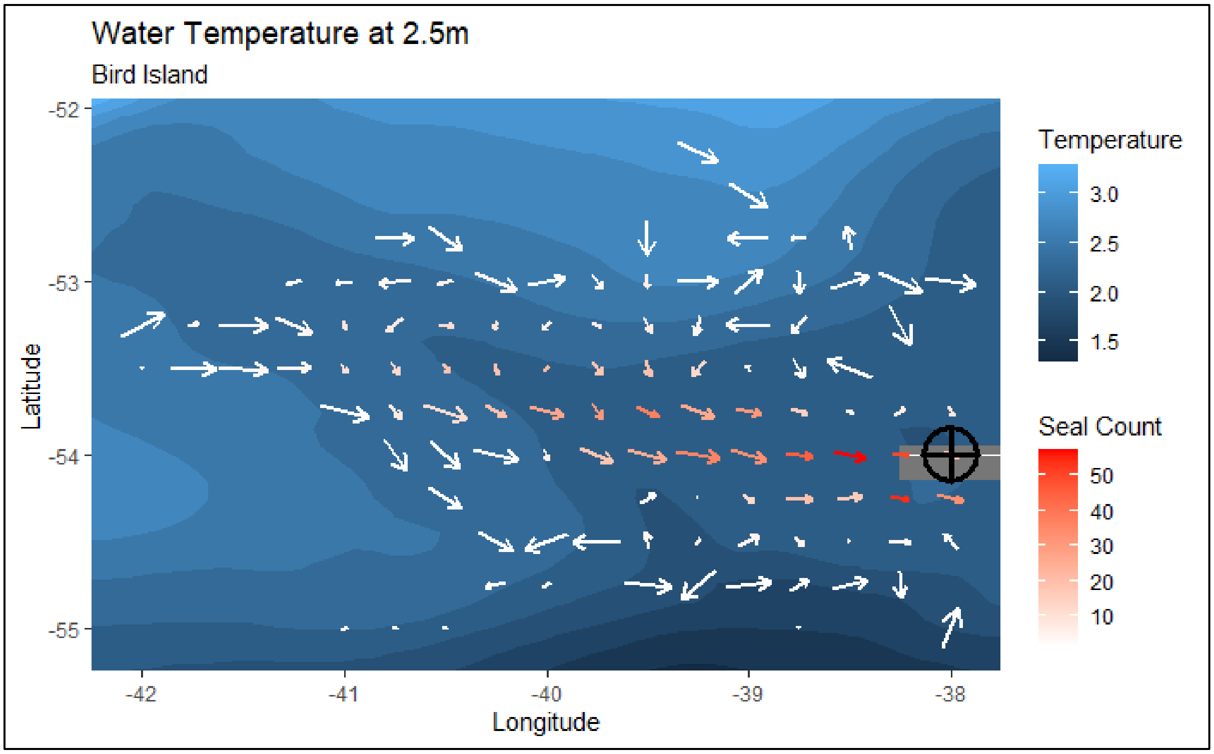

Figure 24. Bird Island Ocean Temperature. Quiver plot overlayed on top of the surrounding ocean temperature for Bird Island. Bird Island is indicated by the target. Ocean temperature data from BRAN2020 (Chamberlain et al., 2021).

3.2.3 Ocean Currents

Currents are another environmental feature in the ocean's surface layer that may be important to investigate for its role in AFS navigation. Currents are often considered in navigational studies due to their ability to concentrate food sources within a region as well as the drift they can impose on the movement of individuals that navigate through them. Previous studies have shown that multiple marine mammals make use of currents to navigate, such as northern fur seals as they migrate to winter foraging grounds (Ream et al., 2005), and dugongs who use currents to travel between foraging grounds and resting areas (Sheppard et al., 2006). Currents are also used to help explain the distribution of some species, such as the minke whale who presumably resides within the Antarctic Circumpolar Current due to the higher prey abundance (Lee et al., 2017). Ocean current utilisation is important to understand as it can be used as an indicator to predict the presence of individuals in a given region.

Ocean currents, however, share a similar drawback to temperature studies in terms of data scale and availability. The change in currents (albeit less significant than temperature) in a given year or from year to year requires data to be available over various timescales and for multiple periods when conducting any meta-analyses. For this example, I utilised averaged data gathered since 1979 (Lumpkin & Centurioni, 2010) from drifter buoys and excluded values below 0.2 m/s. The currents surrounding Marion Island (Figure 25) showed a general northward trend which does not correspond well with the travel corridors displayed by AFS. Bird Island (Figure 26), on the other hand, does show some similarities between the direction of navigation and the predominant flow of the currents. However, further research is required to determine if the speed of the currents are strong enough to directly displace the animals or act as environmental cues.

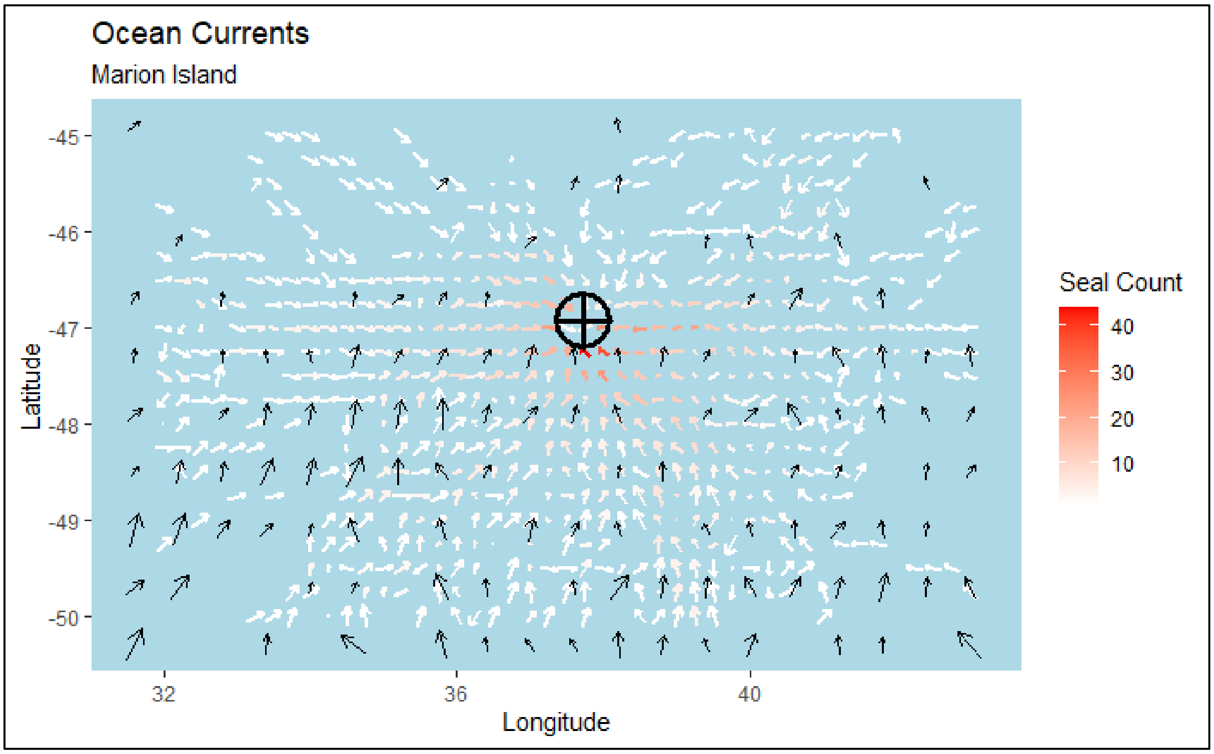

Figure 25. Marion Island Ocean Currents. Dual quiver plot view showing the corridor usage (white to red gradient arrows) alongside the ocean currents (black arrows). To increase readability of the graph, currents below 0.2 m/s have been removed. Larger arrows indicate higher speeds. Marion Island is indicated by the target. Ocean current data from (Lumpkin & Centurioni, 2010).

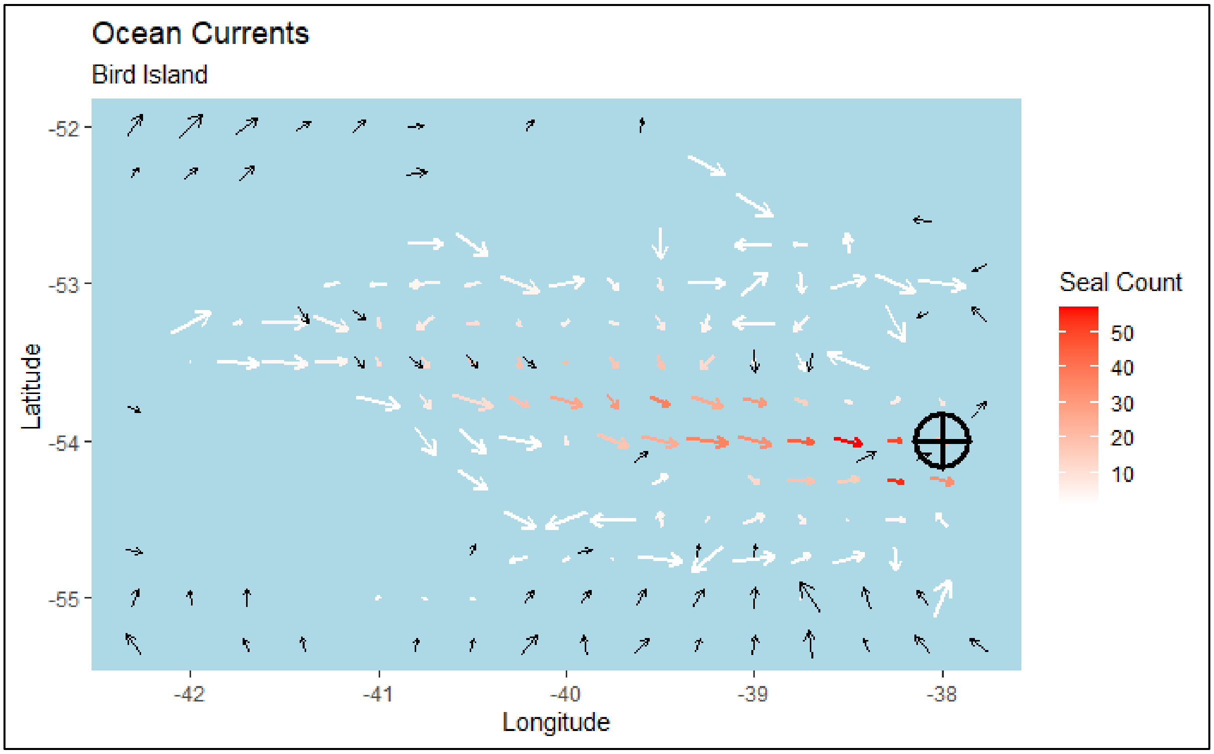

Figure 26. Bird Island Ocean Currents. Dual quiver plot view showing the corridor usage (white to red gradient arrows) alongside the ocean currents (black arrows). To increase readability of the graph, currents below 0.2 m/s have been removed. Larger arrows indicate higher speeds. Bird Island is indicated by the target. Ocean current data from (Lumpkin & Centurioni, 2010).

3.2.4 Geomagnetic Cues

Given its current prominence in navigational studies (Cresci et al., 2021; Lohmann et al., 2022; Wiltschko & Wiltschko, 2022), it would be amiss not to consider geomagnetic signals as one potential cue that could be utilised by the AFS. Whilst there has been behavioural observations (Kirschvink et al., 1986; Kremers et al., 2014) and psychophysical testing (Walker et al., 1992) of potential geomagnetic sensitivity in multiple cetacean species, evidence in pinnipeds is only hypothesised at this stage (Fuiman et al., 2021; Hüttner et al., 2023). There are several types of geomagnetic cues: inclination, declination, total field strength, magnetic anomalies. For this analysis I focussed on the commonly studied geomagnetic inclination. Data for this overlay exercise was included as part of the R package *igrf* (Hufkens, 2022) and for the purpose of this visualisation was delineated into 0.25 degree intervals. Bird Island (Figure 28) appears to correlate with the decrease in magnetic inclination towards the island. If it is being utilised, it could potentially explain some of the inconsistent navigation patterns displayed to the north of the island. Hypothetically, these individuals could be following the gradient in the wrong location and are subsequently having to rely on an alternative cue to get back on track to return home. For Marion Island (Figure 27), the corridors displayed typically only cross one inclination gradient before ending. The direction to the island is also both up and down the gradient depending on the corridor. Therefore, if it is being utilised, it is likely at a scale not visible at this resolution of data. It is also hypothesised that the AFS (among other species) typically swim in circular patterns to learn the magnetic footprint of an area (Narazaki et al., 2021). Such a hypothesis is just one of the potential avenues of research that can be explored in terms of geomagnetism and highlights the complexity of such studies.

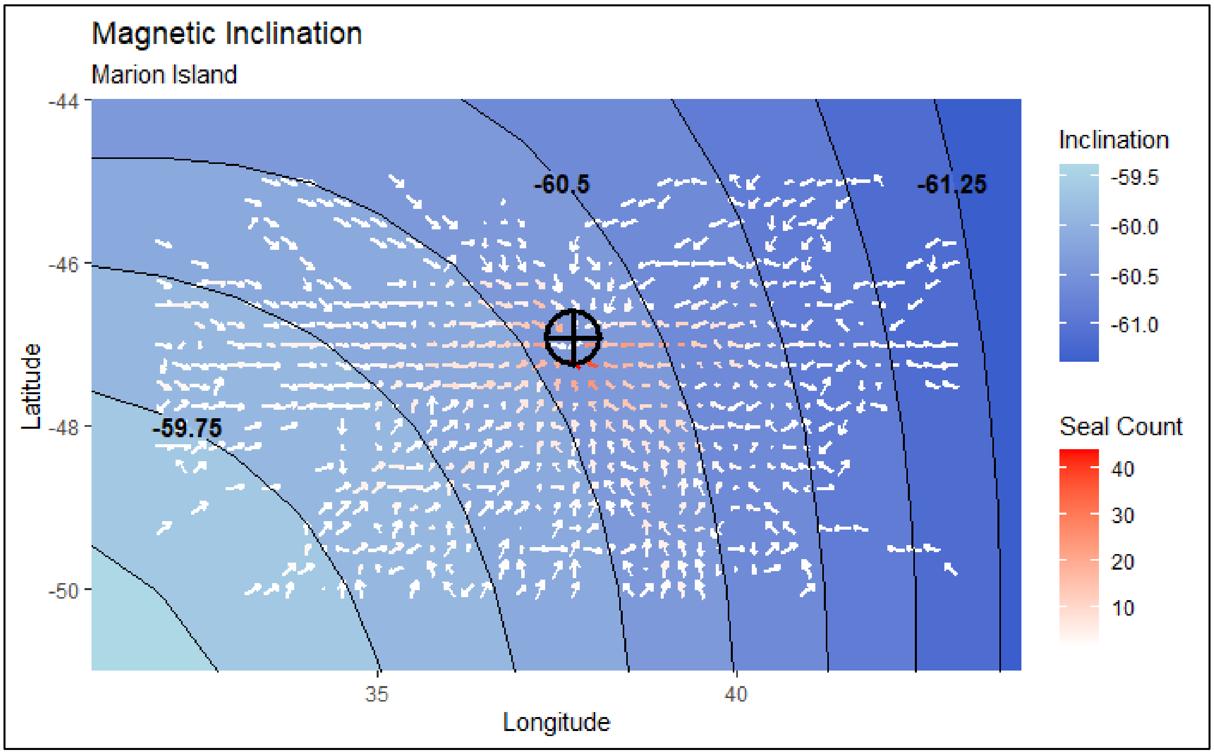

Figure 27. Marion Island Magnetic Inclination. Quiver plot overlayed on top of the surrounding magnetic inclination for Marion Island. Marion Island is indicated by the target. Magnetic inclination data from the *igrf* package (Hufkens, 2022).

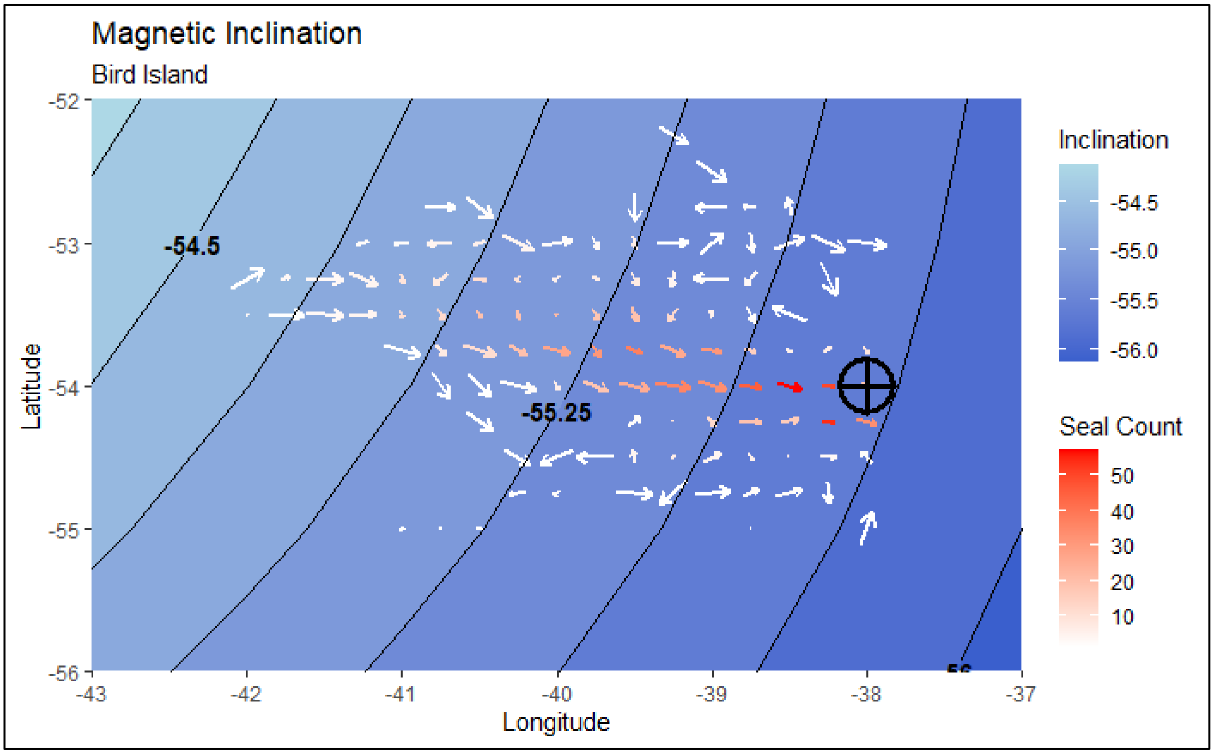

Figure 28. Bird Island Magnetic Inclination. Quiver plot overlayed on top of the surrounding magnetic inclination for Bird Island. Bird Island is indicated by the target. Magnetic inclination data from the *igrf* package (Hufkens, 2022).

3.3 Conclusions

To this day, scientists are unsure of *how* the AFS is able to navigate in the open ocean. Therefore, before to determine if cues such as the four investigated in this chapter are utilised by the Antarctic Fur Seal, it is important to investigate what patterns exist in their movement. Of course, the examples presented here do not encompass the entire range of possible questions one could ask in comparing environmental cues to movement corridors. It should also be acknowledged that corridors are not always the best metric for trying to identify links to environmental correlates. Instead, it is hoped that these examples can instil discussion among scientists into the possibilities of navigation research across the marine mammalian species that my work has only touched the surface of with the AFS.

4 Overall Conclusions

4.1 Specific Findings

The extensive data pre-processing I performed on the telemetry data obtained from the Australian Antarctic Data Centre (Ropert-Coudert et al., 2020) and the British Antarctic Survey (Staniland et al., 2020) yielded a total of 553 Antarctic fur seal (AFS) tracks that were utilised to investigate four different avenues of inquiry on how fur seals successfully navigate to and from Marion Island and Bird Island.

The initial set of analyses determined the relationship between dispersal direction away from and towards each island within each foraging trip. First, I examined whether dispersal direction was directed or random. It was shown that at a distance of 50 km (±10 km) from the island, individuals displayed preferred directional movement behaviour in some instances. Outbound travel from both islands and inbound travel to Bird Island showed a significantly non-uniform distribution of dispersal, indicating a preferred directed travel path and, assumedly, some underlying means of choosing this direction that is leading to this distribution. This similarity might, at first glance, suggest that animals are simply retracing their path back to the island. However, a direct comparison of outbound and inbound bearing showed a notable difference in angle that implied animals were likely finding their way back to the island using a different path, implying a more complex form of navigation. This led to my second line of inquiry which looked at the entire inbound legs with the goal of visualising any potential corridor usage at a per island level. I was able to visualise their movements with the aid of mapping data and highlight potential areas of preferred area use indicative of corridor usage. In Chapter 3, I also demonstrated how these results could be integrated with other environmental data to indicate drivers of these. In my next analysis, I calculated the bearing offset of each location for the inbound legs to identify the scale of their directed movements. I found that there was a general increase in directedness as individuals approached the island. This is a step forward for understanding the navigational ability of the AFS as we now know the scale at which any potential cues must exist, given their ability to navigate across the range of intervals tested in this analysis. My final analysis focused specifically on the timing of the start of the inbound legs. This analysis aimed to determine if individuals use the current lunar phase as a cue when determining when they should return to the island. Whilst no significant results were found, it does not rule out the possibility of lunar cues in their navigation as a whole.

Also, as a result of these analyses I was able to show that the use of meta-analysis is a valuable tool for identifying patterns in AFS movement through pre-existing telemetry data. These findings provide an important step in the process of understanding AFS navigation as a whole and is a crucial step in the search for the key factors of navigation such as cues and environmental correlates that can complete the puzzle of *how* the AFS is able to navigate in the open ocean.

4.2 Strengths and Limitations

Scientific research is intended as a methodology to address a specific question or multiple generally related questions. In theory this should be a perfectly balanced process that contains no semblance of limitations or inaccuracies. Unfortunately, in practice it is near impossible to achieve the ideal circumstances for scientific research. Each methodology must therefore include a degree of consideration into what these limitations may be and the potential strengths it has compared to other approaches. In this section I will discuss some of the trade-offs I faced when designing this study and provide the reasoning behind these decisions.

If we consider the study of animal navigation as a spectrum of understanding, there is a considerable distance between the initial pilot studies and understanding *how* an individual is able to navigate. As discussed in 1.5, most studies focused on the AFS have set their goal for understanding *where* they navigate to as opposed to *how* they navigate there. Whilst this is a crucial step in the process, determining what sensory modalities and environmental cues they use would be a big leap forward from this point. Therefore, my research set out to close the gap between these two points in a cost-efficient manner. To begin my research, I had to sift through a disorganised accumulation of data from various sources and convert it into a coherent dataset that can be studied appropriately. Commonly referred to as data tidying, this process was by far the most time-consuming period of my research. I will discuss this process in two parts: the unavoidable data processing and the preventable inconsistencies.

Preventable inconsistencies are a two-fold process. The first cause is related to inconsistencies in the original data collection. This typically involves mistakes in archiving, such as duplicate entries, missing data points, or invalid data. These problems are usually a result of human error and can only be fixed through more thorough data validation during the course of scientific research. Due to the nature of a meta-analysis, I cannot prevent such errors but had to deal with them accordingly.

On the other hand, unpreventable steps are cross referencing data, mismatched standards, and instrumentation error. Cross referencing data is the process of determining if data from the same study has been submitted to multiple sources and has been duplicated through the combining of these sources in a meta-analysis. Unfortunately, due to standards imposed by the parties involved in storing the data, typical distinct values such as individual ID can be modified and whilst they may appear as unique animals, they are actually exact copies of the data. This was a common error in my original dataset, and I will discuss how I accounted for this later in this section. In terms of mismatched standards, this could be anything from the units used (such as how latitude and longitude is formatted), the time zone used by the instruments, to the way that data is stored (file types and associated meta data). Instrumentation error is the most inconsistent of errors in data handling and typically involves algorithmic data processing or tedious manual manipulation.

Given the apparent imperfections in the data, it was crucial I set up a thorough process of ingesting data and preparing it for analysis. Proper data handling was required to both ensure that I could easily identify and correct any mistakes, and to prevent introducing errors of my own. This resulted in the creation of two data processing workflows. The first involved creating an *R* (R Core Team, 2022) package that handled the intake and combining of all data files. Given that data was to be added progressively to my dataset, it was important to ensure that I followed the same process each time. My package provided the ability to map the data from the source files to a new combined file, as well as providing a documented file of the process. This meant that data could always be followed back up the chain to where it came from if an inconsistency was found. It also had specific checks in place that would throw errors if user input was invalid. After ensuring the data was consistently combined into a new source file, I decided that manual inspection would be the best means of handling potential errors due to the limited availability of algorithmic solutions in this field. At this stage I developed a *Shiny* app (Chang et al., 2023) to systematically visualise all tracks individually and split them into journeys. This had the benefit of identifying duplicate tracks as well as highlighting potential invalid data points (for example, those that appeared on land). It also highlighted the tendency for tracks to have large clusters of points adjacent to shore. This was a previously unknown feature of the dataset and its identification allowed for furthering data tidying before the analysis stage.

Whilst this heavy stage of tidying the data was very time consuming and is sometimes skipped as a result, it is absolutely essential to ensure valid results in the remaining processes. In addition to this, it allows for unbiased filtering of the data during the visualisation stage. It was then a case of determining valid questions and potential statistical analyses that could answer them. If the questions were predetermined, then there is a risk of only keeping data that supports the hypothesis as opposed to maintaining integrity of the data. This is where the first algorithmic filtering was implemented through the use of the *aniMotum* (Jonsen et al., 2023) package. This package allowed for categorisation of the data in a repeatable and efficient manner whilst matching using the same values for ARGOS filtering (3 m/sec) that have been used in previous research (J. Sterling, personal communication, December 1, 2022). This analysis stage highlighted that the largest factor in data reliability was due to the missing location class (LC) values in the dataset. Location class is a categorical measure of accuracy for each data point. The *aniMotum* package heavily relies on this information during its classification and therefore the majority of tracks were discarded before the analysis stage due to this information not being present for the entirety of the track. It is hoped that the importance of this information is not overlooked by future researchers to allow for greater sample sizes in studies such as mine.

As for the analysis portion of my research – specifically the dispersal, bearing offset, and lunar fraction methodologies – the limitations that were present were relatively minor in comparison to the time that was required to solve those in the data pre-filtering stages. The dispersal analysis had two potential faults, the first being that rookery information was typically missing from the dataset. Whilst I was able to show the dispersal patterns that exist, I am unable to comment on the impact of rookery location on their dispersal. However, this does not limit the functionality of the analysis, particularly in terms of the corridor visualisation. If rookery location is a factor, that is more relevant of *where* they navigate to/from as opposed to *how* they navigate. The benefit of my approach is relating the data back to *how* they navigate and determining potential environmental correlates, such as those discussed above in Section 4.2. The other limit to this analysis is less of a fault and more of a matter of defining standards. The thresholds for this analysis were chosen at 50 km based upon the greatest availability of data as opposed to specific biological significance for cues. It is hoped that, as a result of identifying these patterns, this research can open discussion to determine what thresholds would be the most relevant from a biological perspective.

As discussed in 2.4.3, the Skillings Mack statistic used in testing the differences across blocks for separate individuals is a relatively robust test for this type of analysis, but one that has rarely been utilised in navigational studies before. It is hard to state the reason for this; however, I theorise that this is due to a combination of its obscurity, and its lack of ability to determine direction. Whilst only changes between distance categories can be demonstrated with this method and not any particular trend, through the cautious observation of the descriptive statistics in this analysis I am confident in my results that the direction is of increasing directedness towards the island. This is a logical conclusion due to the utilisation of complete journeys. If individuals bearing offset was increasing the closer they got to shore, they would ultimately be travelling in the wrong direction and would not be represented in the dataset due to their failure to successfully navigate during their inbound legs. The final potential limitation I will address is that of the reliability of lunar fraction as a metric. Lunar fraction does not consider the difference in which side of the moon was visible; however, at the same time, there has been no research to determine if the AFS itself can determine the difference. Therefore, I hypothesise that if they are using lunar cues, it is more likely related to light level than shape of the moon that is visible. However, light level itself is only relevant on a night of clear skies and this could indicate for or against the usage of lunar fraction as a cue given the insignificant correlation found by my analysis. Despite the insignificant results, we cannot rule out lunar cues. It is likely that, if they are being used, they are not the sole cue for navigation.

Whilst my study demonstrates the clear value of telemetry data to navigation research, it is worth noting that telemetry data is typically collected by attaching equipment to an individual for an extended period. In recent years a plethora of marine animal telemetry data has become available thanks to the advancements of telemetry equipment (Holton et al., 2021), combined with the increased awareness of telemetry capabilities (Harcourt et al., 2019). However, as the prevalence of telemetry studies increases, so does the risk of a negative impact on individuals that are tagged by the equipment (Blanchet et al., 2014; Cleasby et al., 2021; Rosen et al., 2018; Walker & Boveng, 1995; Walker et al., 2012). This has led to a series of studies trying to standardise the methods of tagging animals to reduce the impacts on individuals (Andrews et al., 2019; Horning et al., 2019; Horning et al., 2017), as well as the adoption of new techniques (Aniceto et al., 2018; Fontes et al., 2022; Fregosi et al., 2016; Raoult et al., 2018; Wiley et al., 2023). The advantage of a meta-analysis means that further risk was not put on individuals, however this risk was of course in place for the individuals that the data originally came from. Whilst I have no concerns with the data I have utilised, it is important to consider the sourcing of data to ensure that the data was obtained with proper ethical practices. It is hoped that the advancements described in the studies above will lead to safer and humane data collection in the future.

Finally, the important take-away of the strengths and limitations of not just this study, but any meta-analysis is to consider the saying "garbage in, garbage out". In my mind the most crucial component for any research is to ensure that the data is as accurate as possible, both during the collection and analysis phases of research. My research has shown, that through the use of meta-analysis, a transitional study between *where* and *how* is an effective means of bridging the gap in our understanding of navigation. Provided of course, that the proper care and attention is given to maintaining accurate and coherent 'tidy data'.

4.3 Closing Remarks

My research was designed to begin to bridge the knowledge void in terms of AFS navigation. Even after decades of research, we as scientists are unable to determine *how* this species navigates, nor do we know if navigation is learned or driven by external factors. I sought to identify movement patterns through a series of four different analyses. As a whole, I have shown evidence for directed movement that could indicate specific types of navigation being utilised by the species such as inertial navigation and some degree of a map sense. These findings are a significant step forward for navigational research both in AFS and marine mammals in general. I maintain that the methodologies proposed in this study are sound as they are backed by thoroughly organised data that can only be achieved through the cooperative effort of maintaining public datasets such as those from the Australian Antarctic Data Centre and British Antarctic Survey, and the methodologies of data preparation that were undertaken in my research. Whilst it may seem now that there are more questions than answers, research like that represented by my study is crucial as it provides insight into what questions we should be asking. Through this study I have achieved my goal of showing the necessity of identifying patterns in movement behaviour if we are to answer *how* the AFS navigates, and I look forward to seeing where the scientific community explores next.

References

Åkesson, S., Ilieva, M., Karagicheva, J., Rakhimberdiev, E., Tomotani, B., & Helm, B. (2017). Timing avian long-distance migration: from internal clock mechanisms to global flights. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1734), 20160252. https://doi.org/10.1098/rstb.2016.0252

Andrews, R. D., Baird, R. W., Calambokidis, J., Goertz, C. E. C., Gulland, F. M. D., Heide-Jorgensen, M. P., Hooker, S. K., Johnson, M., Mate, B., Mitani, Y., Nowacek, D. P., Owen, K., Quakenbush, L. T., Raverty, S., Robbins, J., Schorr, G. S., Shpak, O. V., Townsend Jr, F. I., Uhart, M., . . . Zerbini, A. N. (2019). Best practice guidelines for cetacean tagging. *J. Cetacean Res. Manage.*, *20*(1), 27-66. https://doi.org/10.47536/jcrm.v20i1.237

Aniceto, A. S., Biuw, M., Lindstrøm, U., Solbø, S. A., Broms, F., & Carroll, J. (2018). Monitoring marine mammals using unmanned aerial vehicles: quantifying detection certainty. *Ecosphere*, *9*(3), e02122. https://doi.org/https://doi.org/10.1002/ecs2.2122

Arthur, B., Hindell, M., Bester, M., De Bruyn, P. J. N., Trathan, P., Goebel, M., & Lea, M.-A. (2017). Winter habitat predictions of a key Southern Ocean predator, the Antarctic fur seal (Arctocephalus gazella). *Deep Sea Research Part II: Topical Studies in Oceanography*, *140*, 171-181. https://doi.org/10.1016/j.dsr2.2016.10.009

Arthur, B., Hindell, M., Bester, M., Trathan, P., Jonsen, I., Staniland, I., Oosthuizen, W. C., Wege, M., & Lea, M. A. (2015). Return customers: foraging site fidelity and the effect of environmental variability in wide-ranging antarctic fur seals. *PLoS One*, *10*(3), e0120888. https://doi.org/10.1371/journal.pone.0120888

Asem, J. S. A., & Fortin, N. J. (2017). 1.15 - Memory for Space, Time, and Episodes☆. In J. H. Byrne (Ed.), *Learning and Memory: A Comprehensive Reference (Second Edition)* (pp. 255-283). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-809324-5.21015-8

Bamford, C. C. G., Warwick-Evans, V., Staniland, I. J., Jackson, J. A., & Trathan, P. N. (2021). Wintertime overlaps between female Antarctic fur seals (Arctocephalus gazella) and the krill fishery at South Georgia, South Atlantic. *PLoS One*, *16*(3), e0248071. https://doi.org/10.1371/journal.pone.0248071

Baylis, A. M., Orben, R. A., Arnould, J. P., Peters, K., Knox, T., Costa, D. P., & Staniland, I. J. (2015). Diving deeper into individual foraging specializations of a large marine predator, the southern sea lion. *Oecologia*, *179*(4), 1053-1065. https://doi.org/10.1007/s00442-015-3421-4

Beltran, R. S., Yuen, A. L., Condit, R., Robinson, P. W., Czapanskiy, M. F., Crocker, D. E., & Costa, D. P. (2022). Elephant seals time their long-distance migrations using a map sense. *Curr Biol*, *32*(4), R156-R157. https://doi.org/10.1016/j.cub.2022.01.031

Bennet, M. A., & Eder, S. H. K. (2016). Magnetoreception and Magnetotaxis. In *Iron Oxides* (pp. 567-590). https://doi.org/10.1002/9783527691395.ch22

Berens, P. (2009). CircStat: A MATLAB Toolbox for Circular Statistics. *Journal of Statistical Software*, *31*(10), 1 - 21. https://doi.org/10.18637/jss.v031.i10

Berthold, P. (1991). Genetic control of migratory behaviour in birds. *Trends in Ecology & Evolution*, *6*(8), 254-257. https://doi.org/https://doi.org/10.1016/0169-5347(91)90072-6

Bianco, G., Ilieva, M., Veibäck, C., Öfjäll, K., Gadomska, A., Hendeby, G., Felsberg, M., Gustafsson, F., & Åkesson, S. (2016). Emlen funnel experiments revisited: methods update for studying compass orientation in songbirds. *Ecol Evol*, *6*(19), 6930-6942. https://doi.org/10.1002/ece3.2383

Bingman, V. P., & Cheng, K. (2005). Mechanisms of animal global navigation: comparative perspectives and enduring challenges. *Ethology Ecology & Evolution*, *17*(4), 295-318. https://doi.org/10.1080/08927014.2005.9522584

Bingman, V. P., Ioalé, P., Casini, G., & Bagnoli, P. (1990). The avian hippocampus: Evidence for a role in the development of the homing pigeon navigational map. *Behavioral Neuroscience*, *104*(6), 906-911. https://doi.org/10.1037/0735-7044.104.6.906

Biro, D., Meade, J., & Guilford, T. (2004). Familiar route loyalty implies visual pilotage in the homing pigeon. *Proc Natl Acad Sci U S A*, *101*(50), 17440-17443. https://doi.org/10.1073/pnas.0406984101

Blanchet, M.-A., Lydersen, C., Biuw, M., Nico de Bruyn, P. J., Hofmeyr, G., Krafft, B. A., & Kovacs, K. M. (2014). Instrumentation and handling effects on Antarctic fur seals (Arctocephalus gazella). *Polar Research*, *33*(1). https://doi.org/10.3402/polar.v33.21630

Bonadonna, F., Lea, M.-A., Dehorter, O., & Guinet, C. (2001). Foraging ground fidelity and route-choice tactics of a marine predator: the Antarctic fur seal Arctocephalus gazella. *Marine Ecology Progress Series*, *223*, 287-297. https://doi.org/10.3354/meps223287

Bonadonna, F., Lea, M.-A., & Guinet, C. (2000). Foraging routes of Antarctic fur seals (Arctocephalus gazella) investigated by the concurrent use of satellite tracking and time-depth recorders. *Polar Biology*, *23*(3), 149-159. https://doi.org/10.1007/s003000050021

Bonner, W. N., & Croxall, J. P. (1988). *An assessment of environmental impacts arising from scientific research and its logistic support at Bird Island, South Georgia*. British Antarctic Survey.

Bouchard, B., Barnagaud, J.-Y., Verborgh, P., Gauffier, P., Campagna, S., & Célérier, A. (2022). A field study of chemical senses in bottlenose dolphins and pilot whales. *The Anatomical Record*, *305*(3), 668-679. https://doi.org/10.1002/ar.24703

Bowen, W. D. (1997). Role of marine mammals in aquatic ecosystems. *Marine Ecology Progress Series*, *158*, 267-274. https://doi.org/10.3354/meps158267

Boyd, I. L. (2004). Migration of Marine Mammals. Biological Resources and Migration, Berlin, Heidelberg.

Boyd, I. L., Hawker, E. J., Brandon, M. A., & Staniland, I. J. (2001). Measurement of ocean temperatures using instruments carried by Antarctic fur seals. *Journal of Marine Systems*, *27*(4), 277-288. https://doi.org/10.1016/S0924-7963(00)00073-7

Boyd, I. L., McCafferty, D. J., Reid, K., Taylor, R., & Walker, T. R. (1998). Dispersal of male and female Antarctic fur seals (Arctocephalus gazella). *Canadian Journal of Fisheries and Aquatic Sciences*, *55*(4), 845-852. https://doi.org/10.1139/f97-314

Boyd, I. L., Staniland, I. J., & Martin, A. R. (2002). Distribution of foraging by female Antarctic fur seals. *Marine Ecology Progress Series*, *242*, 285-294. https://doi.org/10.3354/meps242285

Burgess, M., & Evans, K. O. (2022). Sirenia Navigation. In J. Vonk & T. K. Shackelford (Eds.), *Encyclopedia of Animal Cognition and Behavior* (pp. 6446-6451). Springer International Publishing. https://doi.org/10.1007/978-3-319-55065-7_1332

Burley, N., Krantzberg, G., & Radman, P. (1982). Influence of colour-banding on the conspecific preferences of zebra finches. *Animal Behaviour*, *30*(2), 444-455. https://doi.org/10.1016/S0003-3472(82)80055-9

Chamberlain, M. A., Oke, P. R., Fiedler, R. A. S., Beggs, H. M., Brassington, G. B., & Divakaran, P. (2021). Next generation of Bluelink ocean reanalysis with multiscale data assimilation: BRAN2020. *Earth Syst. Sci. Data*, *13*(12), 5663-5688. https://doi.org/10.5194/essd-13-5663-2021

Chang, W., Cheng, J., Allaire, J. J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert, A., & Borges, B. (2023). *shiny: Web Application Framework for R*.

Chatfield, M., & Mander, A. (2009). The Skillings–Mack Test (Friedman Test when There are Missing Data). *The Stata Journal*, *9*(2), 299-305. https://doi.org/10.1177/1536867x0900900208

Cheeseman, T., Johnson, T., Southerland, K., & Muldavin, N. (2017). Happywhale: Globalizing marine mammal photo identification via a citizen science web platform. *Happywhale, Santa Cruz, CA, USA, Rep. SC/67b/PH/02*.

Cherel, Y., Guinet, C., & Tremblay, Y. (1997). Fish prey of Antarctic fur seals Arctocephalus gazella at Ile de Croy, Kerguelen. *Polar Biology*, *17*(1), 87-90. https://doi.org/10.1007/s003000050109

Chipman, R., Slate, D., Rupprecht, C., & Mendoza, M. (2008). Downside risk of wildlife translocation. *Dev Biol (Basel)*, *131*, 223-232.

Cleasby, I. R., Morrissey, B. J., Bolton, M., Owen, E., Wilson, L., Wischnewski, S., & Nakagawa, S. (2021). What is our power to detect device effects in animal tracking studies? *Methods in Ecology and Evolution*, *12*(7), 1174-1185. https://doi.org/doi.org/10.1111/2041-210X.13598

CLIM301. (n.d.). *Wind: u and v Components*. George Mason University. Retrieved 2023/10/01 from http://colaweb.gmu.edu/dev/clim301/lectures/wind/wind-uv

Cloyed, C. S., Hieb, E. E., Collins, M. K., DaCosta, K. P., & Carmichael, R. H. (2019). Linking Use of Ship Channels by West Indian Manatees (Trichechus manatus) to Seasonal Migration and Habitat Use. *Frontiers in Marine Science*. https://doi.org/10.3389/fmars.2019.00318

Collecte Localisation Satellites. (2016). *CLS Argos System User Manual*

Condit, R., Beltran, R. S., Robinson, P. W., Crocker, D. E., & Costa, D. P. (2021). Birth timing after the long feeding migration in northern elephant seals. *Marine Mammal Science*, *38*(3), 931-940. https://doi.org/10.1111/mms.12896

Congdon, J. V., Hosseini, M., Gading, E. F., Masousi, M., Franke, M., & MacDonald, S. E. (2022). The Future of Artificial Intelligence in Monitoring Animal Identification, Health, and Behaviour. *Animals*, *12*(13), 1711. https://doi.org/10.3390/ani12131711

Cresci, A., Sandvik, A. D., Sævik, P. N., Ådlandsvik, B., Olascoaga, M. J., Miron, P., Durif, C. M. F., Skiftesvik, A. B., Browman, H. I., & Vikebø, F. (2021). The lunar compass of European glass eels (Anguilla anguilla) increases the probability that they recruit to North Sea coasts. *Fisheries Oceanography*, *30*(3), 315-330. https://doi.org/10.1111/fog.12521

Darwin, C. (1873). Origin of Certain Instincts. *Nature*, *7*(179), 417-418. https://doi.org/10.1038/007417a0

Doidge, D. W., & Croxall, J. P. (1989). Factors affecting weaning weight in Antarctic fur seals Arctocephalus gazella at South Georgia. *Polar Biology*, *9*(3), 155-160. https://doi.org/10.1007/BF00297170

Dutton, P., Benson, S., & Hitipeuw, C. (2008). *Pacific Leatherback Sets Long-Distance Record* (The State of the World's Sea Turtles, Issue 3).

Egevang, C., Stenhouse, I. J., Phillips, R. A., Petersen, A., Fox, J. W., & Silk, J. R. D. (2010). Tracking of Arctic terns Sterna paradisaea reveals longest animal migration. *Proceedings of the National Academy of Sciences*, *107*(5), 2078-2081. https://doi.org/10.1073/pnas.0909493107

Etienne, A. S. (1992). Navigation of a Small Mammal by Dead Reckoning and Local Cues. *Current Directions in Psychological Science*, *1*(2), 48-52. https://doi.org/10.1111/1467-8721.ep11509737

Fontes, J., Macena, B., Solleliet-Ferreira, S., Buyle, F., Magalhães, R., Bartolomeu, T., Liebsch, N., Meyer, C., & Afonso, P. (2022). The advantages and challenges of non-invasive towed PILOT tags for free-ranging deep-diving megafauna. *Animal Biotelemetry*, *10*(1), 39. https://doi.org/10.1186/s40317-022-00310-1

Forcada, J., & Hoffman, J. I. (2014). Climate change selects for heterozygosity in a declining fur seal population. *Nature*, *511*(7510), 462-465. https://doi.org/10.1038/nature13542

Forcada, J., Hoffman, J. I., Gimenez, O., Staniland, I. J., Bucktrout, P., & Wood, A. G. (2023). Ninety years of change, from commercial extinction to recovery, range expansion and decline for Antarctic fur seals at South Georgia. *Global Change Biology*, *29*(24), 6867-6887. https://doi.org/10.1111/gcb.16947

Forcada, J., & Staniland, I. J. (2018). Antarctic Fur Seal: Arctocephalus gazella. In B. Würsig, J. G. M. Thewissen, & K. M. Kovacs (Eds.), *Encyclopedia of Marine Mammals (Third Edition)* (pp. 25-27). Academic Press. https://doi.org/10.1016/B978-0-12-804327-1.00046-7

Freas, C. A., & Cheng, K. (2022). The Basis of Navigation Across Species. *Annual Review of Psychology*, *73*(1), 217-241. https://doi.org/10.1146/annurev-psych-020821-111311

Fregosi, S., Klinck, H., Horning, M., Costa, D. P., Mann, D., Sexton, K., Hückstädt, L. A., Mellinger, D. K., & Southall, B. L. (2016). An animal-borne active acoustic tag for minimally invasive behavioral response studies on marine mammals. *Animal Biotelemetry*, *4*(1), 9. https://doi.org/10.1186/s40317-016-0101-z

Fuiman, L. A., Williams, T. M., & Davis, R. W. (2021). On the straight and narrow: directed movement by Weddell seals (Leptonychotes weddellii) during on-ice travel. *Polar Biology*, *44*(3), 601-606. https://doi.org/10.1007/s00300-021-02811-w

Gagliardo, A., Mazzotto, M., & Bingman, V. P. (1997). Piriform cortex ablations block navigational map learning in homing pigeons. *Behavioural Brain Research*, *86*(2), 143-148. https://doi.org/10.1016/S0166-4328(96)02253-X

Gould, J. L. (2004). Animal navigation. *Curr Biol*, *14*(6), R221-224. https://doi.org/10.1016/j.cub.2004.02.049

Gould, J. L., & Gould, C. G. (2012). *Nature's compass: the mystery of animal navigation*. Princeton University Press.

Groves, A. B., Collins, G. B., & Trefethen, P. S. (1968). Roles of Olfaction and Vision in Choice of Spawning Site by Homing Adult Chinook Salmon (Oncorhynchus tshawytscha). *Journal of the Fisheries Research Board of Canada*, *25*(5), 867-876. https://doi.org/10.1139/f68-082

Guilbert, J. M. (2003). *Translocation of bats : a study investigating the magnetic sense as a mechanism for homing by bats and the suitability of Tiritiri Matangi Island as a translocation site* Thesis (MSc--Biological Sciences)--University of Auckland, 2003.].

Guzman, H. M., Gomez, C. G., Hearn, A., & Eckert, S. A. (2018). Longest recorded trans-Pacific migration of a whale shark (Rhincodon typus). *Marine Biodiversity Records*, *11*(1), 8. https://doi.org/10.1186/s41200-018-0143-4

Harcourt, R., Sequeira, A. M. M., Zhang, X., Roquet, F., Komatsu, K., Heupel, M., McMahon, C., Whoriskey, F., Meekan, M., Carroll, G., Brodie, S., Simpfendorfer, C., Hindell, M., Jonsen, I., Costa, D. P., Block, B., Muelbert, M., Woodward, B., Weise, M., . . . Fedak, M. A. (2019). Animal-Borne Telemetry: An Integral Component of the Ocean Observing Toolkit [Systematic Review]. *Frontiers in Marine Science*, *6*. https://doi.org/10.3389/fmars.2019.00326

Hays, G. C., Bradshaw, C. J. A., James, M. C., Lovell, P., & Sims, D. W. (2007). Why do Argos satellite tags deployed on marine animals stop transmitting? *Journal of Experimental Marine Biology and Ecology*, *349*(1), 52-60. https://doi.org/10.1016/j.jembe.2007.04.016

Hays, G. C., Cerritelli, G., Esteban, N., Rattray, A., & Luschi, P. (2020). Open Ocean Reorientation and Challenges of Island Finding by Sea Turtles during Long-Distance Migration. *Current Biology*, *30*(16), 3236-3242.e3233. https://doi.org/10.1016/j.cub.2020.05.086

Hijmans, R. J. (2022). *geosphere: Spherical Trigonometry*.

Hindell, M., Burton, H., & Slip, D. (1991). Foraging areas of southern elephant seals, Mirounga leonina, as inferred from water temperature data. *Marine and Freshwater Research*, *42*(2), 115-128. https://doi.org/10.1071/MF9910115

Hoffman, J. I., Trathan, P. N., & Amos, W. (2006). Genetic tagging reveals extreme site fidelity in territorial male Antarctic fur seals Arctocephalus gazella. *Mol Ecol*, *15*(12), 3841-3847. https://doi.org/10.1111/j.1365-294X.2006.03053.x

Hofmeyr, G., De Bruyn, N., Wege, M., & Bester, M. (2017). A conservation assessment of Arctocephalus gazella. In.

Hofmeyr, G. J. G. (2016). Arctocephalus gazella. *The IUCN Red List of Threatened Species*. https://doi.org/10.2305/IUCN.UK.2016-1.RLTS.T2058A66993062.en

Holton, M. D., Wilson, R. P., Teilmann, J., & Siebert, U. (2021). Animal tag technology keeps coming of age: an engineering perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1831), 20200229. https://doi.org/10.1098/rstb.2020.0229

Horning, M., Andrews, R. D., Bishop, A. M., Boveng, P. L., Costa, D. P., Crocker, D. E., Haulena, M., Hindell, M., Hindle, A. G., Holser, R. R., Hooker, S. K., Hückstädt, L. A., Johnson, S., Lea, M.-A., McDonald, B. I., McMahon, C. R., Robinson, P. W., Sattler, R. L., Shuert, C. R., . . . Womble, J. N. (2019). Best practice recommendations for the use of external telemetry devices on pinnipeds. *Animal Biotelemetry*, *7*(1). https://doi.org/10.1186/s40317-019-0182-6

Horning, M., Haulena, M., Tuomi, P. A., Mellish, J.-A. E., Goertz, C. E., Woodie, K., Berngartt, R. K., Johnson, S., Shuert, C. R., Walker, K. A., Skinner, J. P., & Boveng, P. L. (2017). Best practice recommendations for the use of fully implanted telemetry devices in pinnipeds. *Animal Biotelemetry*, *5*(1), 13. https://doi.org/10.1186/s40317-017-0128-9

Horton, T. W., Hauser, N., Zerbini, A. N., Francis, M. P., Domeier, M. L., Andriolo, A., Costa, D. P., Robinson, P. W., Duffy, C. A. J., Nasby-Lucas, N., Holdaway, R. N., & Clapham, P. J. (2017). Route Fidelity during Marine Megafauna Migration. *Frontiers in Marine Science*, *4*. https://doi.org/10.3389/fmars.2017.00422

Horton, T. W., Holdaway, R. N., Zerbini, A. N., Hauser, N., Garrigue, C., Andriolo, A., & Clapham, P. J. (2011). Straight as an arrow: humpback whales swim constant course tracks during long-distance migration. *Biology letters*, *7*(5), 674-679. https://doi.org/doi:10.1098/rsbl.2011.0279

Horton, T. W., Zerbini, A. N., Andriolo, A., Danilewicz, D., & Sucunza, F. (2020). Multi-Decadal Humpback Whale Migratory Route Fidelity Despite Oceanographic and Geomagnetic Change [Original Research]. *Frontiers in Marine Science*, *7*. https://doi.org/10.3389/fmars.2020.00414

Hufkens, K. (2022). *igrf: The 13th Generation International Geomagnetic Reference Field (IGRF) implemented as an R package*.

Hui, N. T., Lo, E. K., Moss, J. B., Gerber, G. P., Welch, M. E., Kastner, R., & Schurgers, C. (2021). A more precise way to localize animals using drones. *Journal of Field Robotics*, *38*(6), 917-928. https://doi.org/10.1002/rob.22017

Hume, F., Pemberton, D., Gales, R., Brothers, N., & Greenwood, M. (2002). Trapping and relocating seals from salmonid fish farms in Tasmania, 1990-2000: was it a success? *Papers and Proceedings of the Royal Society of Tasmania*, 1-6. https://doi.org/10.26749/rstpp.136.1

Hüttner, T., von Fersen, L., Miersch, L., & Dehnhardt, G. (2023). Passive electroreception in bottlenose dolphins (Tursiops truncatus): implication for micro- and large-scale orientation. *Journal of Experimental Biology*, *226*(22). https://doi.org/10.1242/jeb.245845

Iwata, T., Sakamoto, K. Q., Edwards, E. W., Staniland, I. J., Trathan, P. N., Goto, Y., Sato, K., Naito, Y., & Takahashi, A. (2015). The influence of preceding dive cycles on the foraging decisions of Antarctic fur seals. *Biol Lett*, *11*(7). https://doi.org/10.1098/rsbl.2015.0227

Jonathan, T. P.-S., Thomas, M. M., & Shawn, R. L. (1999). The Fundamental Role of Pirouettes in Caenorhabditis elegans Chemotaxis. *The Journal of Neuroscience*, *19*(21), 9557. https://doi.org/10.1523/JNEUROSCI.19-21-09557.1999

Jonsen, I., Grecian, J., Phillips, L., Carroll, G., McMahon, C., Harcourt, R., Hindell, M., & Patterson, T. (2023). aniMotum, an R package for animal movement data: Rapid quality control, behavioural estimation and simulation. *Methods in Ecology and Evolution*, *14*(3), 806-816. https://doi.org/10.1111/2041-210x.14060

Kearney, J. (2023). *Protecting Bird Migration Habitat in Kespukwitk Autumn 2022*.

Kirkwood, R., Lynch, M., Gales, N., Dann, P., & Sumner, M. (2006). At-sea movements and habitat use of adult male Australian fur seals (Arctocephalus pusillus doriferus). *Canadian Journal of Zoology*, *84*(12), 1781-1788. https://doi.org/10.1139/z06-164

Kirschvink, J. L., Dizon, A. E., & Westphal, J. A. (1986). Evidence From Strandings for Geomagnetic Sensitivity in Cetaceans. *Journal of Experimental Biology*, *120*(1), 1-24. https://doi.org/10.1242/jeb.120.1.1

Kishkinev, D., Chernetsov, N., Heyers, D., & Mouritsen, H. (2013). Migratory Reed Warblers Need Intact Trigeminal Nerves to Correct for a 1,000 km Eastward Displacement. *PLoS One*, *8*(6), e65847. https://doi.org/10.1371/journal.pone.0065847

Kowalewsky, S., Dambach, M., Mauck, B., & Dehnhardt, G. (2006). High olfactory sensitivity for dimethyl sulphide in harbour seals. *Biol Lett*, *2*(1), 106-109. https://doi.org/10.1098/rsbl.2005.0380

Krause, D. J., Bonin, C. A., Goebel, M. E., Reiss, C. S., & Watters, G. M. (2022). The Rapid Population Collapse of a Key Marine Predator in the Northern Antarctic Peninsula Endangers Genetic Diversity and Resilience to Climate Change [Original Research]. *Frontiers in Marine Science*, *8*. https://doi.org/10.3389/fmars.2021.796488

Kremers, D., López Marulanda, J., Hausberger, M., & Lemasson, A. (2014). Behavioural evidence of magnetoreception in dolphins: detection of experimental magnetic fields. *Naturwissenschaften*, *101*(11), 907-911. https://doi.org/10.1007/s00114-014-1231-x

Landler, L., Ruxton, G. D., & Malkemper, E. P. (2019). Circular statistics meets practical limitations: a simulation-based Rao's spacing test for non-continuous data. *Movement Ecology*, *7*(1), 15. https://doi.org/10.1186/s40462-019-0160-x

Lea, M., Cherel, Y., Guinet, C., & Nichols, P. (2002). Antarctic fur seals foraging in the Polar Frontal Zone: inter-annual shifts in diet as shown from fecal and fatty acid analyses. *Marine Ecology Progress Series*, *245*, 281-297. https://doi.org/10.3354/meps245281

Lee, J. F., Friedlaender, A. S., Oliver, M. J., & DeLiberty, T. L. (2017). Behavior of satellite-tracked Antarctic minke whales (Balaenoptera bonaerensis) in relation to environmental factors around the western Antarctic Peninsula. *Animal Biotelemetry*, *5*(1). https://doi.org/10.1186/s40317-017-0138-7

Leitinger, B., Poletaeva, I. I., Wolfer, D. P., & Lipp, H.-P. (1994). Swimming navigation, open-field activity, and extrapolation behavior of two inbred mouse strains with Robertsonian translocation of chromosomes 8 and 17. *Behavior Genetics*, *24*(3), 273-284. https://doi.org/10.1007/BF01067194

Letty, J., Marchandeau, S., & Aubineau, J. (2007). Problems encountered by individuals in animal translocations: Lessons from field studies. *Écoscience*, *14*(4), 420-431. https://doi.org/10.2980/1195-6860(2007)14[420:PEBIIA]2.0.CO;2

Lisovski, S., Bauer, S., Briedis, M., Davidson, S. C., Dhanjal-Adams, K. L., Hallworth, M. T., Karagicheva, J., Meier, C. M., Merkel, B., Ouwehand, J., Pedersen, L., Rakhimberdiev, E., Roberto-Charron, A., Seavy, N. E., Sumner, M. D., Taylor, C. M., Wotherspoon, S. J., & Bridge, E. S. (2020). Light-level geolocator analyses: A user's guide. *Journal of Animal Ecology*, *89*(1), 221-236. https://doi.org/10.1111/1365-2656.13036

Lohmann, K. J., Goforth, K. M., Mackiewicz, A. G., Lim, D. S., & Lohmann, C. M. F. (2022). Magnetic maps in animal navigation. *Journal of Comparative Physiology A*, *208*(1), 41-67. https://doi.org/10.1007/s00359-021-01529-8

Lumpkin, R., & Centurioni, L. (2010). *NOAA Global Drifter Program quality-controlled 6-hour interpolated data from ocean surface drifting buoys*. https://doi.org/10.25921/7ntx-z961

Lund, U., Agostinelli, C., Arai, H., Gagliardi, A., García-Portugués, E., Giunchi, D., Irisson, J.-O., Pocernich, M., & Rotolo, F. (2022). *circular: Circular Statistics*.

Luschi, P. (2013). Long-Distance Animal Migrations in the Oceanic Environment: Orientation and Navigation Correlates. *ISRN Zoology*, *2013*, 631839. https://doi.org/10.1155/2013/631839

Maaß, E., & Hanke, F. D. (2022). How harbour seals (Phoca vitulina) encode goals relative to landmarks. *Journal of Experimental Biology*, *225*(5). https://doi.org/10.1242/jeb.243870

Maaß, E., Miersch, L., Pfuhl, G., & Hanke, F. D. (2022). A harbour seal (Phoca vitulina) can learn geometrical relationships between landmarks. *Journal of Experimental Biology*, *225*(24). https://doi.org/10.1242/jeb.244544

March, D., Drago, M., Gazo, M., Parga, M., Rita, D., & Cardona, L. (2021). Winter distribution of juvenile and sub-adult male Antarctic fur seals (Arctocephalus gazella) along the western Antarctic Peninsula. *Sci Rep*, *11*(1), 22234. https://doi.org/10.1038/s41598-021-01700-w

Mate, B. R., Ilyashenko, V. Y., Bradford, A. L., Vertyankin, V. V., Tsidulko, G. A., Rozhnov, V. V., & Irvine, L. M. (2015). Critically endangered western gray whales migrate to the eastern North Pacific. *Biol Lett*, *11*(4), 20150071. https://doi.org/10.1098/rsbl.2015.0071

Matsumura, M., Watanabe, Y. Y., Robinson, P. W., Miller, P. J., Costa, D. P., & Miyazaki, N. (2011). Underwater and surface behavior of homing juvenile northern elephant seals. *J Exp Biol*, *214*(Pt 4), 629-636. https://doi.org/10.1242/jeb.048827

Matuska, S., Hudec, R., Kamencay, P., Benco, M., & Radilova, M. (2016). A novel system for non-invasive method of animal tracking and classification in designated area using intelligent camera system. *Radioengineering*, *25*(1), 161-168.

Mauck, B., Brown, D., Schlosser, W., Schaeffel, F., & Dehnhardt, G. (2005). HOW A HARBOR SEAL SEES THE NIGHT SKY. *Marine Mammal Science*, *21*(4), 646-656. https://doi.org/10.1111/j.1748-7692.2005.tb01257.x

Mauck, B., Gläser, N., Schlosser, W., & Dehnhardt, G. (2008). Harbour seals (Phoca vitulina) can steer by the stars. *Animal Cognition*, *11*(4), 715-718. https://doi.org/10.1007/s10071-008-0156-1

Melzheimer, J., Aschenborn, O., Rast, W., & Zwilling, J. (2022, 2022/09/13). *Early warning system for environmental changes: Novel animal tags with camera and AI complete maiden flight at Tierpark Berlin*

Mouritsen, H., Derbyshire, R., Stalleicken, J., Mouritsen, O. Ø., Frost, B. J., & Norris, D. R. (2013). An experimental displacement and over 50 years of tag-recoveries show that monarch butterflies are not true navigators. *Proceedings of the National Academy of Sciences*, *110*(18), 7348-7353. https://doi.org/10.1073/pnas.1221701110

Mrosovsky, N., & Shettleworth, S. J. (1968). Wavelength Preferences and Brightness Cues in the Water Finding Behaviour of Sea Turtles. *Behaviour*, *32*(4), 211-257. https://doi.org/10.1163/156853968X00216

Nagel, R., Mews, S., Adam, T., Stainfield, C., Fox-Clarke, C., Toscani, C., Langrock, R., Forcada, J., & Hoffman, J. I. (2021). Movement patterns and activity levels are shaped by the neonatal environment in Antarctic fur seal pups. *Scientific Reports*, *11*(1), 14323. https://doi.org/10.1038/s41598-021-93253-1

Narazaki, T., Nakamura, I., Aoki, K., Iwata, T., Shiomi, K., Luschi, P., Suganuma, H., Meyer, C. G., Matsumoto, R., Bost, C. A., Handrich, Y., Amano, M., Okamoto, R., Mori, K., Ciccione, S., Bourjea, J., & Sato, K. (2021). Similar circling movements observed across marine megafauna taxa. *iScience*, *24*(4), 102221. https://doi.org/10.1016/j.isci.2021.102221

Nordstrom, C. A., Battaile, B. C., Cotté, C., & Trites, A. W. (2013). Foraging habitats of lactating northern fur seals are structured by thermocline depths and submesoscale fronts in the eastern Bering Sea. *Deep Sea Research Part II: Topical Studies in Oceanography*, *88-89*, 78-96. https://doi.org/10.1016/j.dsr2.2012.07.010

Norimoto, H., & Ikegaya, Y. (2015). Visual Cortical Prosthesis with a Geomagnetic Compass Restores Spatial Navigation in Blind Rats. *Current Biology*, *25*(8), 1091-1095. https://doi.org/10.1016/j.cub.2015.02.063

O'Hara-Wild, M. (2021). *ggquiver: Quiver Plots for ggplot2*.

Oberhauser, K. S., Taylor, O. R., Reppert, S. M., Dingle, H., Nail, K. R., Pyle, R. M., & Stenoien, C. (2013). Are monarch butterflies true navigators? The jury is still out. *Proceedings of the National Academy of Sciences - PNAS*, *110*(39), E3680-E3680. https://doi.org/10.1073/pnas.1308369110

Ortiz, M., Prince, E. D., Serafy, J. E., Holts, D. B., Davy, K. B., Pepperell, J. G., Lowry, M. B., & Holdsworth, J. C. (2003). Global overview of the major constituent-based billfish tagging programs and their results since 1954. *Marine and Freshwater Research*, *54*(4), 489-507. https://doi.org/10.1071/MF02028

Packmor, F., Kishkinev, D., Bittermann, F., Kofler, B., Machowetz, C., Zechmeister, T., Zawadzki, L. C., Guilford, T., & Holland, R. A. (2021). A magnet attached to the forehead disrupts magnetic compass orientation in a migratory songbird. *Journal of Experimental Biology*, *224*(22). https://doi.org/10.1242/jeb.243337

Pante, E., & Simon-Bouhet, B. (2013). *marmap: A Package for Importing, Plotting and Analyzing Bathymetric and Topographic Data in R*. *PLoS One* Public Library of Science. https://doi.org/10.1371/journal.pone.0073051

Papi, F., & Casini, G. (1990). Pigeons with ablated pyriform cortex home from familiar but not from unfamiliar sites. *Proceedings of the National Academy of Sciences*, *87*(10), 3783-3787. https://doi.org/10.1073/pnas.87.10.3783

Patel, R. N., Kempenaers, J., & Heinze, S. (2022). Vector navigation in walking bumblebees. *Current Biology*, *32*(13), 2871-2883.e2874. https://doi.org/10.1016/j.cub.2022.05.010

Pinzon-Rodriguez, A., Bensch, S., & Muheim, R. (2018). Expression patterns of cryptochrome genes in avian retina suggest involvement of Cry4 in light-dependent magnetoreception. *Journal of The Royal Society Interface*, *15*(140), 20180058. https://doi.org/10.1098/rsif.2018.0058

Posit team. (2023). *RStudio: Integrated Development Environment for R*.

Putman, Nathan F., Scanlan, Michelle M., Billman, Eric J., O'Neil, Joseph P., Couture, Ryan B., Quinn, Thomas P., Lohmann, Kenneth J., & Noakes, David L. G. (2014). An Inherited Magnetic Map Guides Ocean Navigation in Juvenile Pacific Salmon. *Current Biology*, *24*(4), 446-450. https://doi.org/10.1016/j.cub.2014.01.017

Quinn, T. P., & Brannon, E. L. (1982). The use of celestial and magnetic cues by orienting sockeye salmon smolts. *Journal of comparative physiology*, *147*(4), 547-552. https://doi.org/10.1007/BF00612020

R Core Team. (2022). *R: A Language and Environment for Statistical Computing*.

Raoult, V., Tosetto, L., & Williamson, J. E. (2018). Drone-Based High-Resolution Tracking of Aquatic Vertebrates. *Drones*, *2*(4), 37. https://doi.org/10.3390/drones2040037

Rasmussen, K., Palacios, D., Calambokidis, J., Saborío, M., Dalla Rosa, L., Secchi, E., Steiger, G., Allen, J., & Stone, G. (2007). Southern hemisphere Humpback whales wintering off Central America: Insights from water temperature into the longest mammalian migration. *Biology letters*, *3*, 302-305. https://doi.org/10.1098/rsbl.2007.0067

Ream, R. R., Sterling, J. T., & Loughlin, T. R. (2005). Oceanographic features related to northern fur seal migratory movements. *Deep Sea Research Part II: Topical Studies in Oceanography*, *52*(5-6), 823-843. https://doi.org/10.1016/j.dsr2.2004.12.021

Reid, K., & Arnould, J. P. Y. (1996). The diet of Antarctic fur sealsArctocephalus gazella during the breeding season at South Georgia. *Polar Biology*, *16*(2), 105-114. https://doi.org/10.1007/BF02390431

Robinson, S., Goldsworthy, S. D., van den Hoff, J., & Hindell, M. (2023). The foraging ecology of two sympatric fur seal species, Arctocephalus gazella and Arctocephalus tropicalis, at Macquarie Island during the austral summer. https://doi.org/10.1071/MF01218

Rodríguez, J. P., Fernández-Gracia, J., Thums, M., Hindell, M. A., Sequeira, A. M. M., Meekan, M. G., Costa, D. P., Guinet, C., Harcourt, R. G., McMahon, C. R., Muelbert, M., Duarte, C. M., & Eguíluz, V. M. (2017). Big data analyses reveal patterns and drivers of the movements of southern elephant seals. *Scientific Reports*, *7*(1), 112. https://doi.org/10.1038/s41598-017-00165-0

Ropert-Coudert, Y., Van de Putte, A. P., & Reisinger, R. R. (2020). *The retrospective analysis of Antarctic tracking data project* https://doi.org/10.1038/s41597-020-0406-x

Rosen, D. A. S., Gerlinsky, C. G., & Trites, A. W. (2018). Telemetry tags increase the costs of swimming in northern fur seals,Callorhinus ursinus. *Marine Mammal Science*, *34*(2), 385-402. https://doi.org/10.1111/mms.12460

Roth, T. C., & Krochmal, A. R. (2016). Pharmacological evidence is consistent with a prominent role of spatial memory in complex navigation. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1824), 20152548. https://doi.org/10.1098/rspb.2015.2548

Scali, S., Sacchi, R., Azzusi, M., Daverio, S., Oppedisano, T., & Mangiacotti, M. (2013). Homeward bound: factors affecting homing ability in a polymorphic lizard. *Journal of Zoology*, *289*(3), 196-203. https://doi.org/10.1111/j.1469-7998.2012.00977.x

Schwarz, L. K., Goebel, M. E., Costa, D. P., & Kilpatrick, A. M. (2013). Top-down and bottom-up influences on demographic rates of Antarctic fur seals Arctocephalus gazella. *J Anim Ecol*, *82*(4), 903-911. https://doi.org/10.1111/1365-2656.12059

Schweizerische Zoologische, G., & Muséum d'histoire naturelle de, G. (1911). *Revue suisse de zoologie* (Vol. t.19 (1911)). Kundig [etc.].

Sequeira, A. M. M. (2020). Animal Navigation: The Mystery of Open Ocean Orientation. *Curr Biol*, *30*(18), R1054-R1056. https://doi.org/10.1016/j.cub.2020.07.049

Sequeira, A. M. M., Heupel, M. R., Lea, M. A., Eguiluz, V. M., Duarte, C. M., Meekan, M. G., Thums, M., Calich, H. J., Carmichael, R. H., Costa, D. P., Ferreira, L. C., Fernandez-Gracia, J., Harcourt, R., Harrison, A. L., Jonsen, I., McMahon, C. R., Sims, D. W., Wilson, R. P., & Hays, G. C. (2019). The importance of sample size in marine megafauna tagging studies. *Ecol Appl*, *29*(6), e01947. https://doi.org/10.1002/eap.1947

Sheppard, J. K., Preen, A. R., Marsh, H., Lawler, I. R., Whiting, S. D., & Jones, R. E. (2006). Movement heterogeneity of dugongs, Dugong dugon (Müller), over large spatial scales. *Journal of Experimental Marine Biology and Ecology*, *334*(1), 64-83. https://doi.org/10.1016/j.jembe.2006.01.011

Shuert, C., Horning, M., & Mellish, J.-A. (2015). The Effect of Novel Research Activities on Long-term Survival of Temporarily Captive Steller Sea Lions (Eumetopias jubatus). *PLoS One*, *10*(11), e0141948. https://doi.org/10.1371/journal.pone.0141948

Srisuradetchai, P. (2023). *Skillings.Mack.Suppress: The Skillings-Mack Test Statistic for Block Designs with Missing Observations*.

Stadia Maps, OpenMapTiles, OpenStreetMap contributors, Stamen Design. (2023). *Stamen Map*. Stadia Maps.

Staniland, I., Waluda, C., & Forcada, J. (2020). *Platform Transmitting Terminal (PTT) tracking of Antarctic Fur Seals at Bird Island, South Georgia, from 1995 - 2010* https://doi.org/10.5285/224003CC-8F77-4980-8CAC-5574E179E4F8

Staniland, I. J., Reid, K., & Boyd, I. L. (2004). Comparing individual and spatial influences on foraging behaviour in Antarctic fur seals Arctocephalus gazella. *Marine Ecology Progress Series*, *275*, 263-274. https://doi.org/10.3354/meps275263

Staniland, I. J., & Robinson, S. L. (2008). Segregation between the sexes: Antarctic fur seals, Arctocephalus gazella, foraging at South Georgia. *Animal Behaviour*, *75*(4), 1581-1590. https://doi.org/10.1016/j.anbehav.2007.10.012

Staniland, I. J., Robinson, S. L., Silk, J. R. D., Warren, N., & Trathan, P. N. (2011). Winter distribution and haul-out behaviour of female Antarctic fur seals from South Georgia. *Marine Biology*, *159*(2), 291-301. https://doi.org/10.1007/s00227-011-1807-3

Tattersall, G. J. (2021). *Thermimage: Thermal Image Analysis*.

Taylor, M. D., Payne, N. L., Becker, A., & Lowry, M. B. (2016). Feels like home: homing of mature large-bodied fish following translocation from a power-station canal. *ICES Journal of Marine Science*, *74*(1), 301-310. https://doi.org/10.1093/icesjms/fsw168

The Mathworks Inc. (2022). *MATLAB version: 9.13.0 (R2022b)*. The MathWorks Inc.

Thorup, K., Bisson, I.-A., Bowlin, M. S., Holland, R. A., Wingfield, J. C., Ramenofsky, M., & Wikelski, M. (2007). Evidence for a navigational map stretching across the continental U.S. in a migratory songbird. *Proceedings of the National Academy of Sciences*, *104*(46), 18115-18119. https://doi.org/10.1073/pnas.0704734104

U.S. Navy. (2023). *Fraction of the Moon Illuminated*. Astronomical Applications Department. Retrieved 2023/12/01 from https://aa.usno.navy.mil/data/MoonFraction

Van Zinderen Bakker, E. M. (1967). Marion and Prince Edward Islands: Biological Studies. *Nature*, *213*(5073), 230-231. https://doi.org/10.1038/213230a0

Villaseñor, N. R., Escobar, M. A. H., & Estades, C. F. (2013). There is no place like home: high homing rate and increased mortality after translocation of a small mammal. *European Journal of Wildlife Research*, *59*(5), 749-760. https://doi.org/10.1007/s10344-013-0730-y

Viviant, M., Jeanniard-du-Dot, T., Monestiez, P., Authier, M., Guinet, C., & Costa, D. (2016). Bottom time does not always predict prey encounter rate in Antarctic fur seals. *Functional Ecology*, *30*(11), 1834-1844. https://doi.org/10.1111/1365-2435.12675

Wahl, V., Pfeffer, S. E., & Wittlinger, M. (2015). Walking and running in the desert ant Cataglyphis fortis. *Journal of Comparative Physiology A*, *201*(6), 645-656. https://doi.org/10.1007/s00359-015-0999-2

Walker, B. G., & Boveng, P. L. (1995). Effects of time–depth recorders on maternal foraging and attendance behavior of Antarctic fur seals (Arctocephalus gazella). *Canadian Journal of Zoology*, *73*(8), 1538-1544. https://doi.org/10.1139/z95-182

Walker, K. A., Trites, A. W., Haulena, M., & Weary, D. M. (2012). A review of the effects of different marking and tagging techniques on marine mammals. *Wildlife Research*, *39*(1). https://doi.org/10.1071/wr10177

Walker, M. M., Kirschvink, J. L., Ahmed, G., & Dizon, A. E. (1992). Evidence That Fin Whales Respond to the Geomagnetic Field During Migration. *Journal of Experimental Biology*, *171*(1), 67-78. https://doi.org/10.1242/jeb.171.1.67

Wallace, D. G., Hines, D. J., Pellis, S. M., & Whishaw, I. Q. (2002). Vestibular information is required for dead reckoning in the rat. *J Neurosci*, *22*(22), 10009-10017. https://doi.org/10.1523/jneurosci.22-22-10009.2002

Wallraff, H. G., Springer, B., Life Sciences eBooks, E. I., & Ebook, C. (2005). *Avian navigation: pigeon homing as a paradigm* (1. Aufl. ed.). Springer. https://doi.org/10.1007/b137573

Wege, M., de Bruyn, P. J. N., Hindell, M. A., Lea, M. A., & Bester, M. N. (2019). Preferred, small-scale foraging areas of two Southern Ocean fur seal species are not determined by habitat characteristics. *BMC Ecol*, *19*(1), 36. https://doi.org/10.1186/s12898-019-0252-x

Wege, M., Tosh, C. A., de Bruyn, P. J. N., & Bester, M. N. (2016). Cross-seasonal foraging site fidelity of subantarctic fur seals: implications for marine conservation areas. *Marine Ecology Progress Series*, *554*, 225-239. https://doi.org/10.3354/meps11798

Wehner, R., & Muller, M. (2010). Piloting in desert ants: pinpointing the goal by discrete landmarks. *Journal of Experimental Biology*, *213*(24), 4174-4179. https://doi.org/10.1242/jeb.050674

Whishaw, I. Q., Hines, D. J., & Wallace, D. G. (2001). Dead reckoning (path integration) requires the hippocampal formation: evidence from spontaneous exploration and spatial learning tasks in light (allothetic) and dark (idiothetic) tests. *Behavioural Brain Research*, *127*(1), 49-69. https://doi.org/10.1016/S0166-4328(01)00359-X

Wiley, D. N., Zadra, C. J., Friedlaender, A. S., Parks, S. E., Pensarosa, A., Rogan, A., Alex Shorter, K., Urbán, J., & Kerr, I. (2023). Deployment of biologging tags on free swimming large whales using uncrewed aerial systems. *Royal Society Open Science*, *10*(4), 221376. https://doi.org/10.1098/rsos.221376

Wiltschko, R., & Wiltschko, W. (2022). Animal navigation: how animals use environmental factors to find their way. *The European Physical Journal Special Topics*. https://doi.org/10.1140/epjs/s11734-022-00610-w

Womble, J. N., Gende, S. M., & Blundell, G. M. (2007). Dive Behavior of a Harbor Seal (Phoca Vitulina Richardii) in the Presence of Transient Killer Whales (Orcinus Orca) in Glacier Bay National Park, Alaska. *Marine Mammal Science*, *23*(1), 203-208. https://doi.org/10.1111/j.1748-7692.2006.00089.x

Zapetis, M., & Szesciorka, A. (2022). Cetacean Navigation. In J. Vonk & T. K. Shackelford (Eds.), *Encyclopedia of Animal Cognition and Behavior* (pp. 1263-1270). Springer International Publishing. https://doi.org/10.1007/978-3-319-55065-7_986

Zimmerman, J. L., & Peterson, S. R. (1998). *Migration of Birds*. Fish & Wildlife Service, U.S. Department of the Interior.

Appendices

Appendix One: Marion Island Code

This appendix is a collection of all the R code produced for the analysis of Marion Island (excluding helper functions and MATLAB code). Comments are provided where relevant to provide brief details as warranted. Headings are approximately linked to those used in the data chapter of this thesis. The provision of this code is to grant the ability for potential feedback as well as the option to use it as a starting point for future analyses. All code provided as is.

Project Setup

Required Libraries

Library calls for all packages used in my thesis

```
library(tidyverse)

library(ggspatial)

library(sf)

library(aniMotum)

library(ggmap)

library(raster, exclude="select")

library(geosphere)

library(circular)

library(oce)

library(scales)

library(stats)

library(rCAT)

library(ggquiver)

library(Thermimage)

library(rvest)
```

```
library(Skillings.Mack.Suppress) # Modified from Skilling.Mack
library(ggsvg)
library(marmap)
library(DescTools)
library(igrf)
library(metR)
library(rnaturalearth)
library(ncdf4)
```

Custom Functions

Functions that I have created (by hand or with OpenAI ChatGPT) or sourced from online w/ attributions

```
source("imports/required_functions.R")
```

Load Raw Data

Load data

```
# Read filtered tracks
journeys <- readRDS("data/shiny-tracks-999.rds")
```

Filtering the Data

Standardizing Between Sources

Filtering using ThesisCrawler, remove duplicate locations, remove duplicate tracks

```
# Filter by manual complete journeys
jourComp <- journeys %>% filter(journeyStatus == "Complete")
```

```r
# Some timestamps are duplicated with different lat and lon, just keeping unique (first) values
jourComp <- jourComp %>%
  distinct(uniqueID, .keep_all = TRUE)


# There are duplicate IDs between BAS and AADC; this removes them
bas_dupes <- read.csv("data/raw/BAS_metadata.csv")
aadc_dupes <- read.csv("data/raw/RAATD_metadata.csv")

bas_dupes <- bas_dupes %>% select(PTT, DEP_ID)
bas_dupes$jointID <- str_c(bas_dupes$DEP_ID, "_", bas_dupes$PTT)
bas_dupes <- bas_dupes %>% select(jointID)
aadc_dupes <- aadc_dupes %>% filter(abbreviated_name == "ANFS") %>% select(individual_id)

dupes <- inner_join(aadc_dupes, bas_dupes, by = c("individual_id" = "jointID"))

to_remove <- unique(
    jourComp %>% filter(id %in% as.vector(dupes$individual_id)) %>% select(id)
  )$id

jourComp <- jourComp %>% filter(!(id %in% to_remove))
##
```

Subsetting to Marion Island

Subset - Marion island

```
marion_rookery_names <- c("Trypot Beach, Marion Island", "Watertunnel, Marion Island")


selected_island <- jourComp %>%

  filter(island %in% marion_rookery_names)


marion_coords <- c("lon" = 37.73319, "lat" = -46.92397)


selected_island$datetime <- as.POSIXct(

  selected_island$datetime, format = "%Y-%m-%d %H:%M:%S", tz = "UTC")


# Code to make sure there are no tracks with lc missing for some values

# group_by(assignedID) %>%

#   filter(!any(lc == ""))

marion_island <- selected_island %>% filter(lc != "")
```

Move Persistence Modelling

Animotum Parameters

Animtoum - Plot features

```
animotum_aes <- aes_lst(conf = FALSE,

                  line = TRUE)

animotum_aes$df$size[1] <- 1.5

animotum_aes$df$size[3] <- 0.1
```

```
animotum_aes$df$col[3] <- grey(0.3)

animotum_aes$df$fill[5] <- grey(0.4)
```

Prepare Data

Prepare data for animotum

```
marion_premotum <- marion_island %>% ungroup() %>%

  select(id, datetime, lat, lon, lc) %>%

  mutate(date = as.character(datetime)) %>%

  select(id, date, lc, lon, lat)
```

Fitting the Model

Part 1: State Space Model

Fit move persistence state space model (fit_ssm)

```
future::plan("multisession")



marion_fit <- marion_premotum %>%

  split(.$id) %>%

  furrr::future_map(~ try(fit_ssm(

    x = .x,

    vmax = 3,

    model = "mp",

    time.step = 6,

    control = ssm_control(verbose = 0)

  ), silent = TRUE),

  .progress = FALSE,

  .options = furrr::furrr_options(seed = TRUE)
```

```
 ) %>%
 bind_rows(.)
future::plan("sequential")
```

Part 2: Move Persistence Model

Fit and plot

```
marion_animotum <- aniMotum::map(marion_fit,
                    what = "p",
                    aes = animotum_aes,
                    normalise = TRUE,
                    group = FALSE,
                    silent = TRUE) +
 xlab(element_blank()) +
 ylab(element_blank()) +
 ggspatial::annotation_scale(height = unit(1.25, "mm"),
                             aes(width_hint = 0.2,
                                 location = "br",
                                 text_col = "white")) +
 theme(legend.position = c(0.95,0.5),
       legend.direction = "vertical",
       legend.key.width = unit(4, "mm"),
       legend.key.height = unit(7, "mm"),
       legend.title = element_text(size = 9),
       legend.text = element_text(size = 7),
```

```
        axis.text = element_text(size = 7),
        panel.grid = element_line(colour = "white"))
```

Results

MPM Values as Tracks

```
## Scale on map varies by more than 10%, scale bar may be inaccurate
```

Extracting MPM Values

Extract Move Persistence (animotum)

```
latlon <- data.frame(st_coordinates(
  st_transform(
    marion_animotum$plot_env$loc_sf$geometry,
    "+proj=longlat +datum=WGS84"
  )
))


marion_mp <- cbind(
  id = marion_animotum$plot_env$loc_sf$id,
  latlon,
  date = marion_animotum$plot_env$loc_sf$date,
  mpm = marion_animotum$plot_env$loc_sf$g
) %>%
  rename(lon = X,
         lat = Y)
```

Environment Cleaning

Clean environment

```
# Removes temporary variables to make analysis easier

tempVariables <- c("bas_dupes", "aadc_dupes", "dupes", "to_remove", "selected_island",

                   "animotum_aes")

tempVariables <- append(tempVariables, "tempVariables")

rm(list = tempVariables)
```

Reformating the Data

Split into Journeys

Get only high speed legs

```
fastTracks <- marion_mp %>% filter(mpm > 0.75)
```

Which individuals have multiple journeys (?) Calculating number of journeys per seal and appending the count to the seal ID

```
toReorder <- marion_island %>%

  group_by(id) %>%

  summarize(num_journeys_per_seal = n_distinct(journeyID)) %>%

  arrange(desc(num_journeys_per_seal))



fastTracks <- left_join(fastTracks, toReorder, by = join_by(id)) %>%

  arrange(desc(num_journeys_per_seal))


fastTracks$id <- str_c(fastTracks$id, '-', fastTracks$num_journeys_per_seal)
```

Subsetting Data

These are not paired legs (ie outgoing may not have matching incoming based on distance and speed filters) filtering out all points within certain distance of island this is to ignore points that are considered to be on land

```
marion_journeys <- fastTracks



marion_journeys$distance <- mapply(function(lon, lat) {

  distVincentyEllipsoid(marion_coords, c(lon, lat))

}, marion_journeys$lon, marion_journeys$lat)



marion_journeys <- marion_journeys %>%

  filter(distance > 20000)
```

Create a fast_leg_id purely for grouping by for time. New fast_leg_id is calculated by determining if there is more than 6 hours between two points. This is because of the animotum modeling to 6 hour intervals This is to calculate each leg

```
marion_journeys <- marion_journeys %>%

  arrange(id, date) %>%

  group_by(id) %>%

  mutate(timeDelay = date - lag(date, default = first(date))) %>%

  ungroup() %>%

  mutate(fast_leg_id = cumsum(timeDelay > 21600 | timeDelay == 0))
```

For each fast_leg_id, take the first and last set with six hours time difference

Ammendment; take the sets that have a point within 40km of island, this should filter out long distance small foraging trips

```
marion_journeys <- marion_journeys %>%

  group_by(fast_leg_id) %>%

  filter(min(distance) < 40000) %>%

  ungroup()
```

Then check length of each legID so that it has a minimum number of points

Ammendment just using fast_leg_id as hard to tell if outgoing or incoming, so just looking at legs Only look at legs that are longer than 2.5 days

```
marion_journeys <- marion_journeys %>%

  group_by(fast_leg_id) %>%

  filter(n() >= 10) %>%

  ungroup()
```

Split into incoming and outgoing

Custom function to split tracks based on certain thresholds This defines the direction of the journey by looking at changes in means A rolling mean that is decreasing in distance is incoming A rolling mean that is increasing in distance is outgoing A rolling mean that is approaching zero is either looping or parralel to island

```
marion_incomings <- marion_journeys %>%

  group_by(fast_leg_id) %>%

  filter(trackDirection(slopeEveryN(distance, 4), threshold = 0.5) == "Incoming") %>%

  ungroup()



marion_looping <- marion_journeys %>%

  group_by(fast_leg_id) %>%

  filter(trackDirection(slopeEveryN(distance, 4), threshold = 0.5) == "Looping") %>%
```

```
 ungroup()


marion_outgoings <- marion_journeys %>%
 group_by(fast_leg_id) %>%
 filter(trackDirection(slopeEveryN(distance, 4), threshold = 0.5) == "Outgoing") %>%
 ungroup()
```

Tracks were then visually inspected for accuracy of classification

Environment Cleaning

Clean environment

```
# Removes temporary variables to make analysis easier
tempVariables <- c("fastTracks", "latlon", "toReorder")
tempVariables <- append(tempVariables, "tempVariables")
rm(list = tempVariables)
```

Dispersal

Dispersal Parameters

Circle dimensions and thresholds

```
thresholdAdjustment <- 12500
threshold <- thresholdAdjustment + 50000
thresholdLimit <- 10000
```

Filter and Subset

Find Nearest Finds data that is closest to the threshold and still within a minimum distance from it This is to provide a snapshot of a specific region instead of the whole range of distances

```r
add_thresholds <- marion_journeys
add_thresholds <- add_thresholds %>%
  group_by(fast_leg_id) %>%
  mutate(groupCount = 1:n()) %>%
  mutate(max_thresh = max(distance)) %>%
  filter(max_thresh > threshold) %>%
  mutate(first_thresh = distance[1]) %>%
  filter(first_thresh < threshold) %>%
  mutate(last_thresh = distance[length(distance)]) %>%
  filter(last_thresh < threshold) %>%
  ungroup()

ids <- unlist(unique(add_thresholds[,"fast_leg_id"]))

points <- data.frame()

for(id in ids) {

  values <- add_thresholds %>%
    group_by(fast_leg_id) %>%
    filter(fast_leg_id == UQ(id))

  points <- rbind(points, nearest(pull(values, distance), threshold))
```

```
}


add_thresholds <- add_thresholds %>%
  select(-c(max_thresh, first_thresh, last_thresh))


points <- points %>%
  mutate(id = ids) %>%
  rename(first = 1, last = 2) %>%
  select(id, first, last)
```

Filter to Nearest

```
nearest_two <- data.frame()


for(i in ids) {
  inside <- add_thresholds %>%
    filter(fast_leg_id == points[points$id == i,]$id) %>%
    filter(
      groupCount == points[points$id == i, 2] | groupCount == points[points$id == i, 3]
    )


  nearest_two <- rbind(nearest_two, inside)
}
```

```
nearest_two <- nearest_two %>%

  group_by(fast_leg_id) %>%

  mutate(leg = ifelse(groupCount == min(groupCount), "outgoing", "incoming"))
```

Create Threshold and Limit Circles to draw on map

```
circles = data.frame(

    ID = as.numeric(c(1:1)),

    longitude = as.numeric(c(marion_coords["lon"])),

    latitude = as.numeric(c(marion_coords["lat"]))

)



thresholdCircle <- make_circles(circles, threshold/1000)

outterCircle <- make_circles(circles, (threshold+thresholdLimit)/1000)

innerCircle <- make_circles(circles, (threshold-thresholdLimit)/1000)
```

Filter to Limit Circles

```
# Any points within the limit



limited_points <- nearest_two %>%

  filter(abs(threshold - distance) < thresholdLimit)



# Any paired points within the limit (must be from the same journey to be a pair)



limited_points <- limited_points %>%

  group_by(fast_leg_id) %>%
```

```
  filter(n() == 2) %>%
  ungroup()
```

Stats

Descriptive stats for this island

View Data

```
lonRange <- extendrange(limited_points$lon, f = 0.24)
latRange <- extendrange(limited_points$lat, f = 0.24)

# Factored by leg

dispersal <-
  ggmap(get_stadiamap(
    bbox = c(lonRange[1], latRange[1], lonRange[2], latRange[2]),
    zoom = 6,
    maptype = "stamen_terrain_background"
  )) +
  geom_point(
    data = limited_points,
    aes(x = lon, y = lat, colour = as.factor(str_to_title(leg))),
    size = 2,
    alpha = 1
  ) +
  geom_point(
    data = limited_points,
```

```r
  shape = 1,
  size = 2,
  colour = "black"
) +
geom_polygon(
  data = thresholdCircle,
  aes(lon, lat, group = ID),
  color = "red",
  alpha = 0
) +
geom_polygon(
  data = outterCircle,
  aes(lon, lat, group = ID),
  color = "black",
  alpha = 0
) +
geom_polygon(
  data = innerCircle,
  aes(lon, lat, group = ID),
  color = "black",
  alpha = 0
) +
labs(
```

```
  colour = 'Direction',
  x = "Longitude",
  y = "Latitude"
) +
ggtitle(paste0(
  colloquial_name,
  " - ",
  (threshold - thresholdAdjustment) / 1000,
  " km ±",
  thresholdLimit / 1000
)) +
geom_point(
  aes(x = marion_coords[1], y = marion_coords[2]),
  shape = 4,
  size = 2,
  stroke = 2
) +
scale_color_brewer(palette="Set1", labels = c("Inbound", "Outbound"))
```

Results

Distribution of bearings (test)

Used rao.spacing.test() as it is better for circular statistics. Chosen over rayleigh because it could be multimodal. Null hypothesis, the data is uniformly distributed, reject if p-value is low

This is to be used at a factor level. That is, do once for outgoing and once for incoming. MATLAB is required to compare between them

Calculate bearings

```
with_bearings <- limited_points %>%
  mutate(bearing = bearing(
    cbind(lon, lat), c(marion_coords["lon"], marion_coords["lat"])
  ))

outgoing_bearing <- with_bearings %>% filter(leg == "outgoing") %>% select(bearing)

incoming_bearing <- with_bearings %>% filter(leg == "incoming") %>% select(bearing)
```

Describe Paired Deviation

This section calculates the angular difference between the outgoing and incoming points and determines the distance between the two points if at extrapolated to the threshold distance from shore

```
# Calculate the angle between the two points for each pair of legs
paired_deviation <- with_bearings %>%
    group_by(fast_leg_id) %>%
    summarize(smallest_angle = min(abs(diff(bearing))))

# Calculations to determine the separation at the threshold distance (in kilometers)
dispersal_distance <- (threshold - thresholdAdjustment) / 1000
paired_deviation$smallest_angle_rad <- paired_deviation$smallest_angle * (pi / 180)
paired_deviation$separation <- 2 * dispersal_distance * sin(paired_deviation$smallest_angle_ra
```

```
d / 2)

paired_deviation$separation_200km <- 2 * 200 * sin(paired_deviation$smallest_angle_rad / 2)

paired_deviation <- paired_deviation %>% mutate(island = "Marion Island")


# Write to csv for comparison between islands

write.csv(paired_deviation, "outputs/marion_deviation.csv")


# Plot the deviation

ggplot(paired_deviation, aes(x = island, y = smallest_angle)) +

  geom_boxplot() +

  labs(

    title = paste0("Marion Paired Leg Deviation (", dispersal_distance, " km)"),

    subtitle = paste0("Pairs = ", nrow(paired_deviation)),

    x = "Island",

    y = "Degrees"

  )
```

```
# Plot the separation

ggplot(paired_deviation, aes(x = island, y = separation)) +

  geom_boxplot() +

  labs(

    title = paste0("Marion Paired Leg Separation (", dispersal_distance, " km)"),

    subtitle = paste0("Pairs = ", nrow(paired_deviation)),
```

```
    x = "Island",

    y = "Kilometers"

  )
```

Combined

Look at both incoming and outgoing

```
some_bearings <-

  circular(with_bearings$bearing, type = "angles", units = "degrees")



rao_result <- rao.spacing.test(some_bearings)



rao_sub <- capture.output(print(rao_result))[5]



rose.diag(

  some_bearings,

  bins = 16,

  main = paste0(

    colloquial_name,

    " - ",

    (threshold - thresholdAdjustment) / 1000,

    " km ±",

    thresholdLimit / 1000,

    " (Combined)"

  ),
```

```
  sub = rao_sub,

  ticks = FALSE,

  zero = 4.71239,

  rotation = c("clock")

)
```

```
write.csv(with_bearings$bearing, "outputs/marion_combined.csv", row.names = FALSE)



rao_result

##

##        Rao's Spacing Test of Uniformity

##

## Test Statistic = 149.6018

## 0.001 < P-value < 0.01

##
```

Outgoing

Look at just outgoing

```
some_bearings <-

  circular(outgoing_bearing, type = "angles", units = "degrees")



rao_result <- rao.spacing.test(some_bearings)



rao_sub <- capture.output(print(rao_result))[5]
```

```
rose.diag(
  some_bearings,
  bins = 16,
  main = paste0(
    "C) ",
    colloquial_name,
    " - ",
    (threshold - thresholdAdjustment) / 1000,
    " km ±",
    thresholdLimit / 1000,
    " (Outbound)"
  ),
  sub = rao_sub,
  ticks = FALSE,
  zero = 4.71239,
  rotation = c("clock")
)
```

```
rao_result
##
##       Rao's Spacing Test of Uniformity
##
## Test Statistic = 148.5042
```

```
## 0.01 < P-value < 0.05
##
write.csv(outgoing_bearing, "outputs/marion_outgoing.csv", row.names = FALSE)
```

Incoming

Look at just incoming

```
some_bearings <-
  circular(incoming_bearing, type = "angles", units = "degrees")

rao_result <- rao.spacing.test(some_bearings)

rao_sub <- capture.output(print(rao_result))[5]

rose.diag(
  some_bearings,
  bins = 16,
  main = paste0(
    "A) ",
    colloquial_name,
    " - ",
    (threshold - thresholdAdjustment) / 1000,
    " km ±",
    thresholdLimit / 1000,
    " (Inbound)"
  ),
```

```
  sub = rao_sub,

  ticks = FALSE,

  zero = 4.71239,

  rotation = c("clock")

)
```

```
rao_result

##

##        Rao's Spacing Test of Uniformity

##

## Test Statistic = 148.2841

## 0.05 < P-value < 0.10

##

write.csv(incoming_bearing, "outputs/marion_incoming.csv", row.names = FALSE)
```

Outgoing vs Incoming

Read the data from MATLAB See circular.m appendix for MATLAB code

```
circ_cmtest <- read.csv("outputs/marion_circ_cmtest.csv", header = TRUE)

print(paste0("Test Statistic: ", circ_cmtest$tstat))

## [1] "Test Statistic: 0.000297111363166014"

print(paste0("P Value: ", circ_cmtest$pval))

## [1] "P Value: 0.451857491028 71"

## Warning: Your circ_cmtest may be out of date;

##  re-run circular.m in MATLAB to be safe
```

Environment Cleaning

Clean environment

```
# Removes temporary variables to make analysis easier

tempVariables <- c("dispersal_total_journeys", "dispersal_total_seals", "i", "ids",

                   "print_kept", "values", "thresholdCircle", "outterCircle",

                   "innerCircle", "circles", "inside")

tempVariables <- append(tempVariables, "tempVariables")

rm(list = tempVariables)
```

Corridors

Filter and Subset

Split loopings into incomings and outgoings (based on midpoint to make easier)

```
marion_looping_outgoing <- marion_looping %>%

  group_by(fast_leg_id) %>%

  slice(seq(1, floor(n()/2))) %>%

  ungroup()


marion_looping_incoming <- marion_looping %>%

  group_by(fast_leg_id) %>%

  slice(seq(floor(n()/2), n())) %>%

  ungroup()
```

Vector Analysis

Get the bearing they are currently facing at each point Then calculate the u and v for geom_quiver

```r
marion_incomings_all <- rbind(marion_incomings, marion_looping_incoming)


# Create a new column to store the bearing

marion_incomings_all$bearing_BfromA = NA


# Loop through each fast_leg_id

all_unique_legs = unique(marion_incomings_all$fast_leg_id)

for (fast_leg_id in all_unique_legs) {

  current_leg = marion_incomings_all[marion_incomings_all$fast_leg_id == fast_leg_id, ]

  current_leg = current_leg[order(current_leg$date), ] # Sort the rows by datetime


  # Set the first row's bearing_BfromA value to NULL

  marion_incomings_all[

    marion_incomings_all$fast_leg_id == fast_leg_id &

      marion_incomings_all$date == current_leg$date[1],

    "bearing_BfromA"] = NA


  # Loop through each row, excluding the first one,

  # and calculate the bearing from the previous row

  for (i in 2:nrow(current_leg)) {

    current_row = current_leg[i, ]

    prev_row = current_leg[i-1, ]
```

```
    bearing_val <- bearing(
      c(prev_row$lat, prev_row$lon), c(current_row$lat, current_row$lon)
    )

    # Transform bearing to the range of 0 to 360 degrees
    bearing_val <- (bearing_val + 360) %% 360

    marion_incomings_all[
      marion_incomings_all$fast_leg_id == fast_leg_id &
        marion_incomings_all$date == current_row$date,
      "bearing_BfromA"] <- bearing_val
  }
}

marion_incomings_all_ends <- na.omit(marion_incomings_all)

corridor_bearings <- marion_incomings_all_ends %>%
  mutate(bearing_BfromA = deg2rad(bearing_BfromA))

corridor_bearings <- corridor_bearings

corridor_bearings$u <- corridor_bearings$mpm * cos(corridor_bearings$bearing_BfromA)
corridor_bearings$v <- corridor_bearings$mpm * sin(corridor_bearings$bearing_BfromA)
```

Half grid size

```
grid_size <- 0.5


corridor_half <- corridor_bearings %>%
  mutate(
    rounded_lon = floor(lon / grid_size) * grid_size,
    rounded_lat = floor(lat / grid_size) * grid_size
  ) %>%
  group_by(rounded_lon, rounded_lat) %>%
  summarise(
    seal_count = n_distinct(fast_leg_id),
    avg_u = mean(u),
    avg_v = mean(v),
    .groups = "drop"
  )
```

Results

Without background

```
# Define the color gradient
colors <- c(
  "red", "red", "green", "green", "blue", "blue",
  "black", "black", "yellow", "yellow"
)


breaks <- c(-180, -135, -134, -45, -44,  44, 45, 134, 135, 180)
```

```r
my_gradient <- scale_color_gradientn(
  colors = colors, values = rescale(breaks, to = c(0, 1))
)

# Create the grid plot
grid_plot <-
  ggplot(
    corridor_half,
    aes(
      x = rounded_lon,
      y = rounded_lat,
      u = avg_u,
      v = avg_v,
      size = 10,
      col = seal_count
    )
  ) +
  geom_quiver(center = TRUE, size = 1) +
  my_gradient +
  scale_x_continuous(limits = c(
    min(corridor_half$rounded_lon) - 1,
    max(corridor_half$rounded_lon) + 1
  )) +
```

```
  scale_y_continuous(limits = c(
    min(corridor_half$rounded_lat) - 1,
    max(corridor_half$rounded_lat) + 1
  )) +
  labs(x = "Longitude", y = "Latitude") +
  theme_bw() +
  geom_point(
    shape = 10,
    stroke = 2,
    size = 10,
    color = "darkviolet",
    aes(x = marion_coords[1], y = marion_coords[2])
  ) +
  ggtitle("Inbound", subtitle = paste0("Grid size: ", grid_size))
```

With Background

```
lonRange <- extendrange(corridor_half$rounded_lon, f = 0.05)
latRange <- extendrange(corridor_half$rounded_lat, f = 0.05)


# Factored by journey


print_map <-
  ggmap(get_stadiamap(
    bbox = c(lonRange[1], latRange[1], lonRange[2], latRange[2]),
```

```
  zoom = 6,

  maptype = "stamen_toner_lite"

)) +

geom_quiver(

  data = corridor_half,

  center = TRUE,

  size = 1,

  aes(

    x = rounded_lon,

    y = rounded_lat,

    u = avg_u,

    v = avg_v,

    size = 10,

    col = seal_count

  )

) +

my_gradient +

labs(x = "Longitude", y = "Latitude") +

theme_bw() +



ggtitle("Inbound", subtitle = paste0("Grid size: ", grid_size))
```

Reduced Corridor Lat and Lon

Reduce the range to more complete region

```
corridor_restrict_coord <- corridor_half %>%
  filter(between(rounded_lon, 32, 43)) %>%
  filter(between(rounded_lat, -50, -45))


# Create the grid plot
grid_plot <-
  ggplot(
    corridor_restrict_coord,
    aes(
      x = rounded_lon,
      y = rounded_lat,
      u = avg_u,
      v = avg_v,
      size = 10,
      col = seal_count
    )
  ) +
  geom_quiver(center = TRUE, size = 1) +
  scale_color_gradient(low = "white", high = "red") +
  scale_x_continuous(limits = c(
    min(corridor_restrict_coord$rounded_lon) - 0.3,
    max(corridor_restrict_coord$rounded_lon) + 0.3
  )) +
```

```
  scale_y_continuous(limits = c(
    min(corridor_restrict_coord$rounded_lat) - 0.3,
    max(corridor_restrict_coord$rounded_lat) + 0.3
  )) +
  labs(x = "Longitude", y = "Latitude") +
  theme(
    panel.background = element_rect(fill = "lightblue"),
    panel.grid.major = element_line(color = "cadetblue"),
    panel.grid.minor = element_line(color = "cadetblue")
  ) +
  geom_point(
    shape = 10,
    stroke = 2,
    size = 10,
    color = "black",
    aes(x = marion_coords[1], y = marion_coords[2])
  ) +
  ggtitle("Inbound", subtitle = paste0("Grid size: ", grid_size))
```

Reduced Corridor; higher resolution

Change grid size to quarters with the reduced range

```
grid_size <- 0.25

corridor_higher_res <- corridor_bearings %>%
```

```
  mutate(
    rounded_lon = floor(lon / grid_size) * grid_size,
    rounded_lat = floor(lat / grid_size) * grid_size
  ) %>%
  group_by(rounded_lon, rounded_lat) %>%
  summarise(
    seal_count = n_distinct(fast_leg_id),
    avg_u = mean(u),
    avg_v = mean(v),
    .groups = "drop"
  )



corridor_higher_restricted <- corridor_higher_res %>%
  filter(between(rounded_lon, 32, 43)) %>%
  filter(between(rounded_lat,-50,-45))


# Create the grid plot
grid_plot <-
  ggplot(
    corridor_higher_restricted,
    aes(
      x = rounded_lon,
```

```
    y = rounded_lat,

    u = avg_u,

    v = avg_v,

    size = 10,

    col = seal_count

  )

) +

geom_quiver(center = TRUE, size = 1) +

scale_color_gradient(low = "white", high = "red") +

scale_x_continuous(limits = c(

  min(corridor_higher_restricted$rounded_lon) - 0.3,

  max(corridor_higher_restricted$rounded_lon) + 0.3

)) +

scale_y_continuous(limits = c(

  min(corridor_higher_restricted$rounded_lat) - 0.3,

  max(corridor_higher_restricted$rounded_lat) + 0.3

)) +

labs(x = "Longitude", y = "Latitude") +

theme(

  panel.background = element_rect(fill = "lightblue"),

  panel.grid.major = element_line(color = "cadetblue"),

  panel.grid.minor = element_line(color = "cadetblue")

) +
```

```
  geom_point(
    shape = 10,
    stroke = 2,
    size = 10,
    color = "black",
    aes(x = marion_coords[1], y = marion_coords[2])
  ) +
  ggtitle("Inbound - Marion Island", subtitle = paste0("Grid size: ", grid_size)) +
  labs(colour = "Seal Count")
```

Bathymetry

```
marion_bathymetry <- getNOAA.bathy(
  lon1 = min(corridor_higher_restricted$rounded_lon) - 0.3,
  lon2 = max(corridor_higher_restricted$rounded_lon) + 0.3,
  lat1 = min(corridor_higher_restricted$rounded_lat) - 0.3,
  lat2 = max(corridor_higher_restricted$rounded_lat) + 0.3,
  resolution = 1
)
## Querying NOAA database ...
## This may take seconds to minutes, depending on grid size
## Building bathy matrix ...
rescaled <- LinScale(0, min(marion_bathymetry), max(marion_bathymetry), 0, 1)[1]


corridor_plot <- autoplot.bathy(marion_bathymetry, geom = c("tile")) +
```

```
scale_fill_gradientn(
  values = c(0, rescaled - 0.001, rescaled, rescaled + 0.001, 1),
  colours = c("dodgerblue4", "lightblue", "cyan", "lightgreen", "darkgreen")
) +
labs(y = "Latitude", x = "Longitude", fill = "Elevation") +
coord_cartesian(expand = 0) +
geom_quiver(
  data = corridor_higher_restricted,
  aes(
    x = rounded_lon,
    y = rounded_lat,
    u = avg_u,
    v = avg_v,
    size = 10,
    col = seal_count
  ),
  center = TRUE,
  size = 1
) + scale_color_gradient(low = "white", high = "red") +
geom_point(
  shape = 10,
  stroke = 2,
  size = 10,
```

```
    color = "black",
    aes(x = marion_coords[1], y = marion_coords[2])
  ) +
  ggtitle("Bathymetry", subtitle = "Marion Island") +
  labs(colour = "Seal Count")  +
  guides(fill = guide_colorbar(order = 1), colour = guide_colorbar(order = 2))
## Coordinate system already present. Adding new coordinate system, which will
## replace the existing one.
corridor_plot
```

Magnetic Inclination

Create magnetic grid based on the 2020 published values

```
magnetic_grid <- igrf::igrf_grid(
  field = "main",
  year = 2020,
  altitude = 0,
  resolution = 1
)
```

Code Defining boundaries and plot

```
lat_min <- (min(corridor_higher_restricted$rounded_lat) - 1)
lat_max <- (max(corridor_higher_restricted$rounded_lat) + 1)


lon_min <- (min(corridor_higher_restricted$rounded_lon) - 1)
lon_max <- (max(corridor_higher_restricted$rounded_lon) + 1)
```

```
subset_magnetic_grid <- magnetic_grid[magnetic_grid$lat >= lat_min &
                                    magnetic_grid$lat <= lat_max &
                                    magnetic_grid$lon >= lon_min &
                                    magnetic_grid$lon <= lon_max,]

base <- ggplot() +
  geom_contour_fill(
    data = subset_magnetic_grid,
    aes(x = lon, y = lat, z = I),
    breaks = MakeBreaks(0.25),
    size = 0.2
  ) +
  scale_fill_gradient(name = "Inclination",
                      low = "royalblue3",
                      high = "lightblue") +
  xlim(lon_min, lon_max) +
  ylim(lat_min, lat_max)  +
  labs(
    title = "Magnetic Inclination",
    subtitle = "Marion Island",
    x = "Longitude",
    y = "Latitude",
```

```r
  colour = "Seal Count"
) +
geom_quiver(
  data = corridor_higher_restricted,
  aes(
    x = rounded_lon,
    y = rounded_lat,
    u = avg_u,
    v = avg_v,
    size = 10,
    col = seal_count
  ),
  center = TRUE,
  size = 1
) +
scale_color_gradient(low = "white", high = "red") +
guides(fill = guide_colorbar(order = 1), colour = guide_colorbar(order = 2)) +
geom_point(
  shape = 10,
  stroke = 2,
  size = 10,
  color = "black",
  aes(x = marion_coords["lon"], y = marion_coords["lat"])
```

```
  ) +
  geom_contour2(
    data = subset_magnetic_grid,
    aes(x = lon, y = lat, z = I, label = stat(level), label_colour = "black", fontface = "bold"),
    label.placer = label_placer_fraction(0.2, isoband::angle_fixed(theta = 0)),
    skip = 2,
    breaks = MakeBreaks(0.25)
  ) +
  scale_x_continuous(expand = c(0,0)) +
  scale_y_continuous(expand = c(0,0))
## Scale for x is already present.
## Adding another scale for x, which will replace the existing scale.
## Scale for y is already present.
## Adding another scale for y, which will replace the existing scale.
```

Plot Magnetic Inclination

```
base
```

Ocean Currents

```
drift <- nc_open("gdp6h_ragged_current.nc")


current <- data.frame(id = ncvar_get(drift, "ids"), lon = ncvar_get(drift, "lon"), lat = ncvar_get(drift, "lat"), time = ncvar_get(drift, "time"), v = ncvar_get(drift, "ve"), u = ncvar_get(drift, "vn"))
```

```
current_sf <- current %>%

  filter(lon > min(corridor_higher_restricted$rounded_lon) - 0.3) %>%

  filter(lon < max(corridor_higher_restricted$rounded_lon) + 0.3) %>%

  filter(lat > min(corridor_higher_restricted$rounded_lat) - 0.3) %>%

  filter(lat < max(corridor_higher_restricted$rounded_lat) + 0.3) %>%

  st_as_sf(coords = c("lon", "lat")) %>%

  st_set_crs(4326)


current_grid <- current_sf %>%

  st_make_grid(n = c(70,60)) %>%

  st_sf()



drifter.split.sf.se = current_sf


drifter.gridded = current_grid %>%

  mutate(id = 1:n(), contained = lapply(st_contains(st_sf(geometry),drifter.split.sf.se),identity),

         obs = sapply(contained, length),

         u = sapply(contained, function(x) {median(drifter.split.sf.se[x,]$u, na.rm = TRUE)}),

         v = sapply(contained, function(x) {median(drifter.split.sf.se[x,]$v, na.rm = TRUE)}))



drifter.gridded = drifter.gridded %>% select(obs, u, v) %>% na.omit()
```

```
## obtain the centroid coordinates from the grid as table
coordinates = drifter.gridded %>%
  st_centroid() %>%
  st_coordinates() %>%
  as_tibble() %>%
  rename(x = X, y = Y)
## Warning: st_centroid assumes attributes are constant over geometries
## remove the geometry from the simple feature of gridded drifter dataset
st_geometry(drifter.gridded) = NULL

## stitch together the extracted coordinates and drifter information int a single table for SE m
onsoon season
current.gridded.se = coordinates %>%
  bind_cols(drifter.gridded) %>%
  mutate(season = "SE")


drifter.current.gridded = current.gridded.se


## select grids for SE season only
drf.se = drifter.current.gridded
```

```
## interpolate the U component
u.se = interpBarnes(x = drf.se$x, y = drf.se$y, z = drf.se$u)


## obtain dimension that determine the width (ncol) and length (nrow) for tranforming wide i
nto long format table
dimension = data.frame(lon = u.se$xg, u.se$zg) %>% dim()


## make a U component data table from interpolated matrix
u.tb = data.frame(lon = u.se$xg,
                  u.se$zg) %>%
  gather(key = "lata", value = "u", 2:dimension[2]) %>%
  mutate(lat = rep(u.se$yg, each = dimension[1])) %>%
  select(lon,lat, u) %>% as.tibble()
## Warning: `as.tibble()` was deprecated in tibble 2.0.0.
## ℹ Please use `as_tibble()` instead.
## ℹ The signature and semantics have changed, see `?as_tibble`.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
## interpolate the V component
v.se = interpBarnes(x = drf.se$x,
                    y = drf.se$y,
                    z = drf.se$v)
```

```
## make the V component data table from interpolated matrix
v.tb = data.frame(lon = v.se$xg, v.se$zg) %>%
  gather(key = "lata", value = "v", 2:dimension[2]) %>%
  mutate(lat = rep(v.se$yg, each = dimension[1])) %>%
  select(lon,lat, v) %>%
  as.tibble()

## stitch now the V component intot the U data table and compute the velocity
uv.se = u.tb %>%
  bind_cols(v.tb %>% select(v)) %>%
  mutate(vel = sqrt(u^2+v^2))
ocean_plot <- ggplot() +
  geom_quiver(
    data = corridor_higher_restricted,
    aes(
      x = rounded_lon,
      y = rounded_lat,
      u = avg_u,
      v = avg_v,
      size = 10,
      col = seal_count
    ),
```

```
  center = TRUE,
  size = 1
) +
metR::geom_vector(
  data = uv.se,
  aes(
    x = lon,
    y = lat,
    dx = u,
    dy = v
  ),
  min.mag = 0.2,
  skip.y = 5,
  skip.x = 2,
  arrow.angle = 30,
  arrow.type = "open",
  arrow.length = 0.5,
  preserve.dir = TRUE,
  direction = "ccw",
  colour = "black",
  show.legend = TRUE
) +
labs(x = "", y = "")  +
```

```
 scale_color_gradient(low = "white", high = "red") +

 theme(

  panel.background = element_rect(fill = "lightblue"),

  panel.grid.major = element_line(color = "lightblue"),

  panel.grid.minor = element_line(color = "lightblue")

 ) +

 geom_point(

  shape = 10,

  stroke = 2,

  size = 10,

  color = "black",

  aes(x = marion_coords["lon"], y = marion_coords["lat"])

 ) +

 labs(

  title = "Ocean Currents",

  subtitle = "Marion Island",

  x = "Longitude",

  y = "Latitude",

  colour = "Seal Count"

 )

ocean_plot
```

Water Temperature

```
sst <- nc_open("ocean_temp_ann_2020.nc")


sst_lon  <- ncvar_get(sst, "xt_ocean" )

sst_lat  <- ncvar_get(sst, "yt_ocean" )

sst_temp <- ncvar_get(sst, "temp")


# Which depth slice to get

depth_slice <- 1


depth_max <- ncvar_get(sst, "st_ocean")[depth_slice]

surface_slice <- sst_temp[, , depth_slice]


sst_matrix <- as.matrix(expand.grid(sst_lon,sst_lat))


sst_vector <- as.vector(surface_slice)


sst_df <- data.frame(cbind(sst_matrix,sst_vector))

names(sst_df) <- c("lon","lat","temps")


# Adjust lon values if needed

sst_df$lon <- ifelse(sst_df$lon > 180, sst_df$lon - 360, sst_df$lon)


# Filter sst_df based on the latitudinal and longitudinal ranges
```

```
sst_df_filtered <- sst_df %>%
  filter(
    lat >= min(corridor_higher_restricted$rounded_lat) - 0.3,
    lat <= max(corridor_higher_restricted$rounded_lat) + 0.3,
    lon >= min(corridor_higher_restricted$rounded_lon) - 0.3,
    lon <= max(corridor_higher_restricted$rounded_lon) + 0.3
  )
# Plot with filtered data
plot_temperature <- ggplot() +
  geom_contour_fill(data = sst_df_filtered, aes(lon, lat, z = temps)) +
  geom_quiver(
    data = corridor_higher_restricted,
    aes(
      x = rounded_lon,
      y = rounded_lat,
      u = avg_u,
      v = avg_v,
      size = 10,
      col = seal_count
    ),
    center = TRUE,
    size = 1
  ) +
```

```
  scale_color_gradient(low = "white", high = "red") +
  labs(
    title = paste0("Water Temperature at ", depth_max, "m"),
    subtitle = "Marion Island",
    x = "Longitude",
    y = "Latitude",
    colour = "Seal Count",
    fill = "Temperature"
  ) +
  guides(fill = guide_colorbar(order = 1), colour = guide_colorbar(order = 2)) +
  geom_point(
    shape = 10,
    stroke = 2,
    size = 10,
    color = "black",
    aes(x = marion_coords["lon"], y = marion_coords["lat"])
  ) +
  scale_x_continuous(expand = c(0,0)) +
  scale_y_continuous(expand = c(0,0)) +
  theme(
    panel.background = element_rect(fill = "#777777")
  )
plot_temperature
```

Bearings

Prepare Data

Collect data from bearing A to B

```
bearings_data <- na.omit(marion_incomings_all)
```

Calculate bearing B to C as extra column

```
predefined_lat = marion_coords[[2]]

predefined_lon = marion_coords[[1]]



# Create a new column to store the bearing

bearings_data$bearing_BtoC = NA



# Loop through each fast_leg_id

all_unique_legs = unique(bearings_data$fast_leg_id)

for (fast_leg_id in all_unique_legs) {

  # Sort the rows by datetime

  journey_bearings_data = bearings_data[bearings_data$fast_leg_id == fast_leg_id, ]

  journey_bearings_data = journey_bearings_data[order(journey_bearings_data$date), ]



  # For the last row, calculate the bearing from the predefined lat and lon

  last_row = journey_bearings_data[nrow(journey_bearings_data), ]

  bearings_data[bearings_data$fast_leg_id == fast_leg_id &

                bearings_data$date == last_row$date,

               "bearing_BtoC"] = bearing(c(last_row$lon, last_row$lat),

                                         c(predefined_lon, predefined_lat))
```

```
 # Loop through rows, excluding the last one, and calculate the bearing to the next row
 for (i in 1:(nrow(journey_bearings_data) - 1)) {
   current_row = journey_bearings_data[i, ]
   next_row = journey_bearings_data[i + 1, ]
   bearings_data[bearings_data$fast_leg_id == fast_leg_id &
                  bearings_data$date == current_row$date,
                "bearing_BtoC"] = bearing(c(current_row$lon, current_row$lat),
                                          c(next_row$lon, next_row$lat))
 }
}
```

Absolute bearing change & bearing from island

```
bearings_data <- bearings_data %>%
  mutate(bearing_angleAtB = angle_diff(bearing_BfromA, bearing_BtoC))


bearings_data <- bearings_data %>%
  mutate(bearing_BtoIsland = bearing(cbind(lon, lat), c(predefined_lon, predefined_lat)))


bearings_data <- bearings_data %>%
  mutate(bearing_abs_island = angle_diff(bearing_BtoIsland, bearing_BtoC))
```

Results

Plot angle and distance

```
ggplot(bearings_data %>% filter(distance/1000 < 1000)) +
  geom_vline(xintercept = 250, colour = "red", size = 2) +
```

```
  geom_point(aes(x = distance/1000, y = bearing_abs_island, color = ifelse(distance > 250000, "
beyond_250", "within_250"))) +
  scale_color_manual(values = c("within_250" = "black", "beyond_250" = "gray")) +
  xlab("Distance from Shore (km)") +
  ylab("Deviation from Island in Degrees") +
  labs(
    title = "Marion Island"
  ) +
  theme(legend.position="none")
```

Stats on groups of distance vs angle (calculate distance intervals)

```
distance_interval <- bearings_data %>%
  select(distance, bearing_abs_island, fast_leg_id) %>%
  rename(angle_diff = bearing_abs_island) %>%
  mutate(distance = distance/1000)
```

Filter to intervals less than 1000km

```
distance_interval <- distance_interval %>% filter(distance < 1000)


# Calculate the minimum and maximum distances in your dataframe
min_distance <- min(distance_interval$distance)
max_distance <- max(distance_interval$distance)


# Create intervals of 200km each
```

```
intervals <- seq(min_distance, max_distance, by = 200)

distance_interval$distance_interval <- cut(distance_interval$distance,

                                    breaks = intervals,

                                    include.lowest = TRUE)
```

Skillings Mack

Create Ranks

Create ranks and define interval size for skillings mack

```
# Distance grouping in kilometers (convert to meters)

skillings_size <- 50

skillings_size <- skillings_size * 1000

times_to_simulate <- 10000



# Keep journeys with this many intervals

# If size is 50km and interval is 3, we keep up to 150km

# However we exclude first interval (in this case, 50km)

# Therefore we only look at 50-100 and 100-150

skillings_interval <- 5



# Keep the journeys that have at least 70% complete values

skillings_limits <- (skillings_interval - 1) * 0.7



skillings <- bearings_data %>%

  mutate(

    interval = cut(
```

```
      distance,
      breaks = seq(0,
                  ceiling(max(distance) / skillings_size) * skillings_size,
                  by = skillings_size
                ),
      include.lowest = TRUE,
      labels = FALSE
    )
  ) %>%
  group_by(fast_leg_id, interval) %>%
  summarise(median_bearing = median(bearing_abs_island, na.rm = TRUE)) %>%
  ungroup()
## `summarise()` has grouped output by 'fast_leg_id'. You can override using the
## `.groups` argument.
```

Display number of individuals in each interval

```
skillings %>%
  group_by(interval) %>%
  count(interval) %>%
  mutate(max_km = interval*skillings_size/1000)
## # A tibble: 27 × 3
## # Groups:   interval [27]
##    interval     n max_km
##       <int> <int>  <dbl>
```

```
## 1        1  182     50
## 2        2  187    100
## 3        3  168    150
## 4        4  142    200
## 5        5  115    250
## 6        6   93    300
## 7        7   74    350
## 8        8   58    400
## 9        9   43    450
## 10      10   35    500
## # ℹ 17 more rows
```

Remove tracks that are biasing results as per comments

```
# Remove tracks that are less than first interval
# Remove tracks above certain interval size
# Shift interval range by 1 to start at 1 instead of 2
skillings_raw <- skillings %>%
  filter(interval >= 2 & interval <= skillings_interval) %>%
  group_by(fast_leg_id) %>%
  mutate(interval = interval - 1) %>%
  rename(interval_shifted = interval) %>%
  ungroup()
```

Run tests

Run Skillings Mack Focus on tracks with more than Skillings_limits percentage of complete values

```
# fast_leg_id expands to all rows

# interval_shifted fills in blanks from 1 to range

# median_bearing is filled with NA values


group <- unique(skillings_raw$fast_leg_id)

n <- max(skillings_raw$interval_shifted)


skillings_blank <- data.frame(

  fast_leg_id = rep(group, each = n),

  interval_shifted = unlist(lapply(split(

    1:(n * length(group)), group

  ), function(x)

    1:length(x))),

  row.names = NULL

)


skillings_test <-

  full_join(skillings_blank,

            skillings_raw,

            by = join_by(fast_leg_id, interval_shifted))


# Remove journeys that have less than skillings_limits (percentage) of complete values
```

```
skillings_NA <- skillings_test %>%
  group_by(fast_leg_id) %>%
  mutate(non_na_count = sum(!is.na(median_bearing))) %>%
  ungroup() %>%
  filter(non_na_count > skillings_limits) %>%
  select(-non_na_count)

skillings_matrix <- skillings_NA %>%
  pivot_wider(names_from = interval_shifted, values_from = median_bearing) %>%
  select(-fast_leg_id)

# Switch blocks and treatments order
skillings_run <- t(skillings_matrix)
Ski_total <-
  Ski.Mack(
    y = data.matrix(skillings_run),
    simulate.p.value = TRUE,
    B = times_to_simulate,
    suppress = TRUE
  )
##
## Skillings-Mack Statistic =  47.430200 , p-value =  0.000000
```

```
## Note: the p-value is based on the chi-squared distribution with d.f. =  3

## Based on B =  10000 , Simulated p-value =  0.000000
```

Plot Intervals

```
plot_skillings <- drop_na(skillings_NA)



ggplot(plot_skillings, aes(interval_shifted, median_bearing)) +

  geom_boxplot(aes(group = interval_shifted)) +

  xlab("Interval (km)") +

  ylab("Median Bearing of Each Seal") +

  scale_x_continuous(breaks = 1:4,  labels=c("50-100", "100-150", "150-200", "200-250")) +

  labs (

    title = "Marion Island"

  ) +

  geom_label(data = plot_skillings %>%

              group_by(interval_shifted) %>%

              summarise(count = n()),

            aes(x = interval_shifted, label = paste0("n = ", count)), y = 140)
```

Remove First Interval

```
skillings_run_tail <- skillings_run[2:nrow(skillings_run),]



Ski_ignore_first <-

  Ski.Mack(

    y = data.matrix(skillings_run_tail),
```

```
    simulate.p.value = TRUE,
    B = times_to_simulate,
    suppress = TRUE
  )
##
## Skillings-Mack Statistic =  20.078728 , p-value =  4.4e-05
## Note: the p-value is based on the chi-squared distribution with d.f. =  2
## Based on B =  10000 , Simulated p-value =  0.000000
```

Calculate percentage of NA for each interval

```
## Percentage of NULL at Each Interval
## Interval: 50km
##     1     2     3     4
##  1.43  0.71  0.00 17.86
```

Consecutive T Tests

T-test to show general but less accurate method

```
t.test(skillings_run[1,], skillings_run[2,], alternative = "l")
##
##  Welch Two Sample t-test
##
## data:  skillings_run[1, ] and skillings_run[2, ]
## t = -1.6632933, df = 269.32893, p-value = 0.04870844
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
```

```
##          -Inf -0.02683371405
## sample estimates:
##   mean of x   mean of y
## 14.91348690 18.41066237
t.test(skillings_run[2,], skillings_run[3,], alternative = "l")
##
##  Welch Two Sample t-test
##
## data:  skillings_run[2, ] and skillings_run[3, ]
## t = -2.693124, df = 261.92574, p-value = 0.003767686
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##         -Inf -2.334918359
## sample estimates:
##   mean of x   mean of y
## 18.41066237 24.44292575
t.test(skillings_run[3,], skillings_run[4,], alternative = "l")
##
##  Welch Two Sample t-test
##
## data:  skillings_run[3, ] and skillings_run[4, ]
## t = -1.1326268, df = 220.75493, p-value = 0.1292999
## alternative hypothesis: true difference in means is less than 0
```

```
## 95 percent confidence interval:
##        -Inf 1.529803785
## sample estimates:
##   mean of x   mean of y
## 24.44292575 27.78043508
```

Filter to complete intervals

Filter to only complete so we can use well known tests on subset

```
skillings_complete <- skillings_run[, colSums(is.na(skillings_run)) == 0]
```

Sign Test R

```
skillings_sign_test <- sign(skillings_complete[1,] - skillings_complete[2,])
wilcox.test(skillings_sign_test, mu = 0)
##
##  Wilcoxon signed rank test with continuity correction
##
## data:  skillings_sign_test
## V = 2655.5, p-value = 0.08928759
## alternative hypothesis: true location is not equal to 0


skillings_sign_test <- sign(skillings_complete[2,] - skillings_complete[3,])
wilcox.test(skillings_sign_test, mu = 0)
##
##  Wilcoxon signed rank test with continuity correction
##
```

```
## data:  skillings_sign_test
## V = 2768.5, p-value = 0.1864336
## alternative hypothesis: true location is not equal to 0

skillings_sign_test <- sign(skillings_complete[3,] - skillings_complete[4,])
wilcox.test(skillings_sign_test, mu = 0)
##
##  Wilcoxon signed rank test with continuity correction
##
## data:  skillings_sign_test
## V = 2429.5, p-value = 0.01408467
## alternative hypothesis: true location is not equal to 0
```

Friedman Test

```
friedman.test(skillings_complete)
##
##  Friedman rank sum test
##
## data:  skillings_complete
## Friedman chi-squared = 233.74831, df = 111, p-value = 9.275569e-11
```

Lunar

Download Lunar Fractions

Download lunar fractions based on year range of the dataset

```
years <- c(year(min(marion_incomings_all$date)):year(max(marion_incomings_all$date)))


lunar_raw <- NULL


for(lunaryear in years) {


  lunarl <- paste0(
    "https://aa.usno.navy.mil/calculated/moon/fraction?year=",
    lunaryear,
    "&task=00&tz=3&tz_sign=1&tz_label=false&submit=Get+Data"
  )


  lunar_html <- read_html(lunarl)


  textData <- lunar_html %>%
    html_node("body") %>%
    html_table(header = TRUE)



  textData <- textData[-1, ]


  colnames(textData) <- textData[1, ]
```

```
  textData <- textData[-1, ]


  textData[] <- lapply(textData, as.numeric)


  textData <- textData %>%
    pivot_longer(cols = -Day, names_to = "month", values_to = "lunar")


  textData <- na.omit(textData)


  textData <- textData %>%
    mutate(year = lunaryear)


  lunar_raw <- rbind(lunar_raw, textData)


}
```

Convert HTML format to Posix compatible Store lunar fraction for every day in the years period

```
lunar_values <- lunar_raw


lunar_values$month <- sub("July", "Jul.", lunar_values$month)
lunar_values$month <- sub("June", "Jun.", lunar_values$month)
lunar_values$month <- gsub("\\.", "", lunar_values$month)


month_mapping <- setNames(1:12, month.abb)
```

```
lunar_values <- lunar_values %>%
  mutate(month = month_mapping[month])


lunar_values <- lunar_values %>%
  mutate(Date = as.Date(paste(year, month, Day, sep = "-")))


# Convert to POSIXct with GMT+3 timezone and set time to midday
lunar_values$DateTime <-
  as.POSIXct(paste(lunar_values$Date, "12:00:00"),
             format = "%Y-%m-%d %H:%M:%S",
             tz = "Africa/Nairobi")


# Remove the Date column if no longer needed
lunar_values <- lunar_values %>% select(DateTime, lunar) %>% rename(date = DateTime)
```

Regrab clean bearings dataset

```
lunar_journeys <- na.omit(marion_incomings_all) %>%
  mutate(bearing_BfromA = deg2rad(bearing_BfromA))
```

Correlate Fractions

Convert dates of data to correct timezone

```
# Get start of journeys
lunar_times <- lunar_journeys %>%
  group_by(fast_leg_id) %>%
  filter(row_number() == 1)
```

```
# Get lat/lon for timezones and date for time
lunar_times <- lunar_times %>%
  select(c("date", "lat", "lon")) %>% ungroup()
## Adding missing grouping variables: `fast_leg_id`
# Create column with Marion Island Time (GMT+3)
lunar_times <- lunar_times %>%
  mutate(date_gmt3 = with_tz(date, tzone = "Africa/Nairobi"))
```

Results

add lunar fraction to each point from lunar values

```
lunar_values <- lunar_values %>% rename(lunar_date = date)
lunar_values$lunar_date <- as.Date(ymd_hms(lunar_values$lunar_date))

lunar_times$date_gmt3 <- as.POSIXct(format(lunar_times$date_gmt3, "%Y-%m-%d 12:00:00"))
lunar_times$date_gmt3 <- as.Date(ymd_hms(lunar_times$date_gmt3))

lunar_reading <-
  inner_join(lunar_values, lunar_times, by = join_by(lunar_date == date_gmt3)) %>%
  select(fast_leg_id, date, lunar, lat, lon) %>%
  mutate(year = as.character(lubridate::year(date)))
```

Visualize the spread of data

```
# Number of bins
bins_count <- 7
```

```
# Expected proportion values
expected_proportion <- layer_data(ggplot(lunar_raw) + geom_histogram(aes(x=lunar), bins =
bins_count))$count/nrow(lunar_raw)

# Navigation distribution (by year) - proportion
nav_plot <- ggplot(lunar_reading) +
  geom_histogram(aes(x=lunar, fill = year, group = year, y = after_stat(count)/sum(after_stat(
count))), bins = bins_count) +
  scale_x_continuous(breaks = c(0, 0.5, 1)) +
  ylab("Observed Proportion") +
  xlab("Lunar Fraction") +
  labs(
    title = "Marion Island",
    fill = "Year"
  )

nav_lines <- lunar_raw %>%
  filter(year %in% lunar_reading$year)

pre_moons <- nav_plot +
  geom_point(data = nav_lines, shape = "\u25AC", stat = "bin", bins = bins_count, aes(x = lunar,
y = after_stat(count)/sum(after_stat(count))), color = "black", size = 10)
```

```
# Message = false to avoid errors in knitting; converts line to dotted line


# Suppress warnings to hide SVG code

suppressWarnings({moons(pre_moons, fill = "white")})
```

Grouped Chi-square

Group to increase count size in each bin (adjusted probability due to both quarters)

```
# data.frame(cut(lunar_reading$lunar, 3, labels = c("New Moon", "Quarter", "Full")))


# Create a contingency table

contingency_table <- layer_data(ggplot(lunar_reading) + geom_histogram(aes(x=lunar), bins
= bins_count))$count


# Grab the counts of each bin and convert to proportion

adjusted_prob <- layer_data(ggplot(lunar_raw) + geom_histogram(aes(x=lunar), bins = bins_c
ount))$count/nrow(lunar_raw)


# Perform chi-squared test

chi_squared_test <- chisq.test(contingency_table, p = adjusted_prob)


# Print the results

print(chi_squared_test)

##

##  Chi-squared test for given probabilities
```

```
##
## data:  contingency_table
## X-squared = 11.182848, df = 6, p-value = 0.08288704
```

Raw expected proportions

```
round(adjusted_prob, 3)
## [1] 0.182 0.151 0.111 0.112 0.118 0.142 0.185
```

Descriptive

Prepare Data

Regrab clean dataset

```
marion_incomings_all_ends <- na.omit(marion_incomings_all)


descriptive_filtered <- marion_incomings_all_ends %>%
  mutate(bearing_BfromA = deg2rad(bearing_BfromA))
```

Arrival and Duration

Return Month

Bar chart for arrival month

```
arrivalMonth <- descriptive_filtered %>%
  group_by(fast_leg_id) %>%
  slice(n()) %>%
  mutate(month = month(date, label = TRUE, abbr = FALSE)) %>%
  ungroup()


ggplot(arrivalMonth, aes(month)) +
```

```
  geom_bar() +
  scale_x_discrete(limits = month.name)
```

> Gender (answer is undefined)

```
animalSex <- marion_island %>%
  select(id, sex) %>%
  distinct(id, .keep_all = TRUE)


animalID <- arrivalMonth %>%
  separate_wider_delim(id, "-", cols_remove = FALSE, names = c("animalID", NA)) %>%
  select(animalID, id) %>%
  distinct(animalID, .keep_all = TRUE)


genders <- right_join(animalSex, animalID, by=join_by(id == animalID)) %>%
  rename(longID = `id.y`)


gendered <- left_join(arrivalMonth, genders %>% select(-id), by=join_by(id == longID))


unique(gendered$sex)
## [1] "U"
```

Return Date vs Return Duration

Box and whisker for arrival and duration

```
maxTime <- descriptive_filtered %>%
  group_by(fast_leg_id) %>%
  summarise(maxDate = max(date))
```

```
minTime <- descriptive_filtered %>%

  group_by(fast_leg_id) %>%

  summarise(minDate = min(date))


return_times <- full_join(maxTime, minTime, by = join_by(fast_leg_id))


return_times <- return_times %>%

  mutate(returnTime = as.numeric(difftime(maxDate, minDate, units = "days")))


return_times <- return_times %>%

  mutate(month = month(maxDate, label = TRUE, abbr = FALSE))



ggplot(return_times, aes(month, returnTime)) +

  geom_boxplot()  +

  scale_x_discrete(limits = month.name) +

  geom_jitter(width = 0.2)
```

Offset Distance for Return Start

```
offset_dist <- descriptive_filtered %>%

  group_by(fast_leg_id) %>%

  mutate(month = month(date, label = TRUE, abbr = FALSE)) %>%

  filter(row_number()==1) %>%
```

```
  mutate(distance_km = distance/1000)


ggplot(offset_dist, aes(month, distance_km)) +
  geom_boxplot()  +
  scale_x_discrete(limits = month.name) +
  geom_jitter(width = 0.2)
```

Other Descriptive Stats

```

print(paste0("Return Leg Count: ", length(unique(offset_dist$fast_leg_id))))
## [1] "Return Leg Count: 194"


print(paste0("Mean Starting Return Distance: ", round(mean(offset_dist$distance_km), 2), "km"))
## [1] "Mean Starting Return Distance: 273.9km"


print(paste0("Mean Return Time: ", round(mean(return_times$returnTime), 2), " day(s)"))
## [1] "Mean Return Time: 4.59 day(s)"
```

Appendix Two: Bird Island Code

This appendix is a collection of all the R code produced for the analysis of Bird Island (excluding helper functions and MATLAB code). Comments are provided where relevant to provide brief details as warranted. Headings are approximately linked to those used in the data chapter of this thesis. The provision of this code is to grant the ability for potential feedback as well as the option to use it as a starting point for future analyses. All code provided as is.

Project Setup

Required Libraries

Library calls for all packages used in my thesis

```r
library(tidyverse)
library(ggspatial)
library(sf)
library(aniMotum)
library(ggmap)
library(raster, exclude="select")
library(geosphere)
library(circular)
library(oce)
library(scales)
library(stats)
library(rCAT)
library(ggquiver)
library(Thermimage)
library(rvest)
library(Skillings.Mack.Suppress) # Modified from Skilling.Mack
```

```
library(ggsvg)

library(marmap)

library(DescTools)

library(igrf)

library(metR)

library(rnaturalearth)

library(ncdf4)
```

Custom Functions

Functions that I have created (by hand or with OpenAI ChatGPT) or sourced from online w/ attributions

```
source("imports/required_functions.R")
```

Load Raw Data

Load data

```
# Read filtered tracks

journeys <- readRDS("data/shiny-tracks-999.rds")
```

Filtering the Data

Standardizing Between Sources

Filtering using ThesisCrawler, remove duplicate locations, remove duplicate tracks

```
# Filter by manual complete journeys

jourComp <- journeys %>% filter(journeyStatus == "Complete")




# Some timestamps are duplicated with different lat and lon, just keeping unique (first) values
```

```
jourComp <- jourComp %>%
  distinct(uniqueID, .keep_all = TRUE)


# There are duplicate IDs between BAS and AADC; this removes them
bas_dupes <- read.csv("data/raw/BAS_metadata.csv")
aadc_dupes <- read.csv("data/raw/RAATD_metadata.csv")

bas_dupes <- bas_dupes %>% select(PTT, DEP_ID)
bas_dupes$jointID <- str_c(bas_dupes$DEP_ID, "_", bas_dupes$PTT)
bas_dupes <- bas_dupes %>% select(jointID)
aadc_dupes <- aadc_dupes %>% filter(abbreviated_name == "ANFS") %>% select(individual_id)

dupes <- inner_join(aadc_dupes, bas_dupes, by = c("individual_id" = "jointID"))

to_remove <- unique(
    jourComp %>% filter(id %in% as.vector(dupes$individual_id)) %>% select(id)
  )$id

jourComp <- jourComp %>% filter(!(id %in% to_remove))
##
```

Remove broken tracks (check if required)

```
jourComp <- jourComp %>%

  filter(id != 190) %>%

  filter(id != 1022) %>%

  filter(id != 319)
```

Subsetting to Bird Island

Subset - Bird island

```
bird_rookery_names <- c("Bird Island, South Georgia")



selected_island <- jourComp %>%

  filter(island %in% bird_rookery_names)



bird_coords <- c("lon" = -38, "lat" = -54)



selected_island$datetime <- as.POSIXct(

  selected_island$datetime, format = "%Y-%m-%d %H:%M:%S", tz = "UTC")



# Code to make sure there are no tracks with lc missing for some values

# group_by(assignedID) %>%

#   filter(!any(lc == ""))

bird_island <- selected_island %>% filter(lc != "")
```

Move Persistence Modelling

Animotum Parameters

Animtoum - Plot features

```
animotum_aes <- aes_lst(conf = FALSE,

                  line = TRUE)

animotum_aes$df$size[1] <- 1.5

animotum_aes$df$size[3] <- 0.1

animotum_aes$df$col[3] <- grey(0.3)

animotum_aes$df$fill[5] <- grey(0.4)
```

Prepare Data

Prepare data for animotum

```
bird_premotum <- bird_island %>% ungroup() %>%

  select(id, datetime, lat, lon, lc) %>%

  mutate(date = as.character(datetime)) %>%

  select(id, date, lc, lon, lat)
```

Fitting the Model

Part 1: State Space Model

Fit move persistence state space model (fit_ssm)

```
future::plan("multisession")



bird_fit <- bird_premotum %>%

  split(.$id) %>%

  furrr::future_map(~ try(fit_ssm(

    x = .x,

    vmax = 3,

    model = "mp",

    time.step = 6,
```

```
   control = ssm_control(verbose = 0)
 ), silent = TRUE),
 .progress = FALSE,
 .options = furrr::furrr_options(seed = TRUE)
 ) %>%
 bind_rows(.)
future::plan("sequential")
```

Part 2: Move Persistence Model

Fit and plot

```
bird_animotum <- aniMotum::map(bird_fit,
                    what = "p",
                    aes = animotum_aes,
                    normalise = TRUE,
                    group = FALSE,
                    silent = TRUE) +
 xlab(element_blank()) +
 ylab(element_blank()) +
 ggspatial::annotation_scale(height = unit(1.25, "mm"),
                             aes(width_hint = 0.2,
                                 location = "br",
                                 text_col = "white")) +
 theme(legend.position = c(0.95,0.5),
       legend.direction = "vertical",
       legend.key.width = unit(4, "mm"),
```

```
    legend.key.height = unit(7, "mm"),

    legend.title = element_text(size = 9),

    legend.text = element_text(size = 7),

    axis.text = element_text(size = 7),

    panel.grid = element_line(colour = "white"))
```

Results

MPM Values as Tracks

```
## Scale on map varies by more than 10%, scale bar may be inaccurate
```

Extracting MPM Values

Extract Move Persistence (animotum)

```
latlon <- data.frame(st_coordinates(

  st_transform(

    bird_animotum$plot_env$loc_sf$geometry,

    "+proj=longlat +datum=WGS84"

  )

))



bird_mp <- cbind(

  id = bird_animotum$plot_env$loc_sf$id,

  latlon,

  date = bird_animotum$plot_env$loc_sf$date,

  mpm = bird_animotum$plot_env$loc_sf$g

) %>%
```

```
  rename(lon = X,
         lat = Y)
```

Environment Cleaning

Clean environment

```
# Removes temporary variables to make analysis easier
tempVariables <- c("bas_dupes", "aadc_dupes", "dupes", "to_remove", "selected_island",
                   "animotum_aes")
tempVariables <- append(tempVariables, "tempVariables")
rm(list = tempVariables)
```

Reformating the Data

Split into Journeys

Get only high speed legs

```
fastTracks <- bird_mp %>% filter(mpm > 0.75)
```

Which individuals have multiple journeys (?) Calculating number of journeys per seal and appending the count to the seal ID

```
toReorder <- bird_island %>%
  group_by(id) %>%
  summarize(num_journeys_per_seal = n_distinct(journeyID)) %>%
  arrange(desc(num_journeys_per_seal))



fastTracks <- left_join(fastTracks, toReorder, by = join_by(id)) %>%
  arrange(desc(num_journeys_per_seal))
```

```
fastTracks$id <- str_c(fastTracks$id, '-', fastTracks$num_journeys_per_seal)
```

Subsetting Data

These are not paired legs (ie outgoing may not have matching incoming based on distance and speed filters) unlike marion we do not filter out as the lat and lon is for the beach as opposed to middle of the island

```
bird_journeys <- fastTracks


bird_journeys$distance <- mapply(function(lon, lat) {
  distVincentyEllipsoid(bird_coords, c(lon, lat))
}, bird_journeys$lon, bird_journeys$lat)


# bird_journeys <- bird_journeys %>%
#   filter(distance > 20000)
```

Create a fast_leg_id purely for grouping by for time. New fast_leg_id is calculated by determining if there is more than 6 hours between two points. This is because of the animotum modeling to 6 hour intervals This is to calculate each leg

```
bird_journeys <- bird_journeys %>%
  arrange(id, date) %>%
  group_by(id) %>%
  mutate(timeDelay = date - lag(date, default = first(date))) %>%
  ungroup() %>%
  mutate(fast_leg_id = cumsum(timeDelay > 21600 | timeDelay == 0))
```

For each fast_leg_id, take the first and last set with six hours time difference

Ammendment; take the sets that have a point within 40km of island, this should filter out long distance small foraging trips

```
bird_journeys <- bird_journeys %>%
  group_by(fast_leg_id) %>%
  filter(min(distance) < 40000) %>%
  ungroup()
```

Then check length of each legID so that it has a minimum number of points

Ammendment just using fast_leg_id as hard to tell if outgoing or incoming, so just looking at legs Only look at legs that are longer than 2.5 days

```
bird_journeys <- bird_journeys %>%
  group_by(fast_leg_id) %>%
  filter(n() >= 10) %>%
  ungroup()
```

Split into incoming and outgoing

Custom function to split tracks based on certain thresholds This defines the direction of the journey by looking at changes in means A rolling mean that is decreasing in distance is incoming A rolling mean that is increasing in distance is outgoing A rolling mean that is approaching zero is either looping or parralel to island

```
bird_incomings <- bird_journeys %>%
  group_by(fast_leg_id) %>%
  filter(trackDirection(slopeEveryN(distance, 4), threshold = 0.5) == "Incoming") %>%
  ungroup()


bird_looping <- bird_journeys %>%
  group_by(fast_leg_id) %>%
```

```
  filter(trackDirection(slopeEveryN(distance, 4), threshold = 0.5) == "Looping") %>%
  ungroup()


bird_outgoings <- bird_journeys %>%
  group_by(fast_leg_id) %>%
  filter(trackDirection(slopeEveryN(distance, 4), threshold = 0.5) == "Outgoing") %>%
  ungroup()
```

Tracks were then visually inspected for accuracy of classification

```
excludes <- c(8, 166, 216, 241, 248, 423, 429, 432, 435, 436, 458, 594, 612, 614, 631)


bird_looping <- bird_looping %>%
  filter(!fast_leg_id %in% excludes)
```

Environment Cleaning

Clean environment

```
# Removes temporary variables to make analysis easier
tempVariables <- c("fastTracks", "latlon", "toReorder")
tempVariables <- append(tempVariables, "tempVariables")
rm(list = tempVariables)
```

Dispersal

Dispersal Parameters

Circle dimensions and thresholds

```
thresholdAdjustment <- 0

threshold <- thresholdAdjustment + 50000

thresholdLimit <- 10000
```

Filter and Subset

Find Nearest Finds data that is closest to the threshold and still within a minimum distance from it This is to provide a snapshot of a specific region instead of the whole range of distances

```
add_thresholds <- bird_journeys

add_thresholds <- add_thresholds %>%

  group_by(fast_leg_id) %>%

  mutate(groupCount = 1:n()) %>%

  mutate(max_thresh = max(distance)) %>%

  filter(max_thresh > threshold) %>%

  mutate(first_thresh = distance[1]) %>%

  filter(first_thresh < threshold) %>%

  mutate(last_thresh = distance[length(distance)]) %>%

  filter(last_thresh < threshold) %>%

  ungroup()


ids <- unlist(unique(add_thresholds[,"fast_leg_id"]))


points <- data.frame()


for(id in ids) {
```

```
  values <- add_thresholds %>%

    group_by(fast_leg_id) %>%

    filter(fast_leg_id == UQ(id))


  points <- rbind(points, nearest(pull(values, distance), threshold))


}


add_thresholds <- add_thresholds %>%

  select(-c(max_thresh, first_thresh, last_thresh))


points <- points %>%

  mutate(id = ids) %>%

  rename(first = 1, last = 2) %>%

  select(id, first, last)
```

Filter to Nearest

```
nearest_two <- data.frame()


for(i in ids) {

  inside <- add_thresholds %>%

    filter(fast_leg_id == points[points$id == i,]$id) %>%

    filter(

      groupCount == points[points$id == i, 2] | groupCount == points[points$id == i, 3]
```

```
   )


 nearest_two <- rbind(nearest_two, inside)

}


nearest_two <- nearest_two %>%

 group_by(fast_leg_id) %>%

 mutate(leg = ifelse(groupCount == min(groupCount), "outgoing", "incoming"))
```

Create Threshold and Limit Circles to draw on map

```
circles = data.frame(

   ID = as.numeric(c(1:1)),

   longitude = as.numeric(c(bird_coords["lon"])),

   latitude = as.numeric(c(bird_coords["lat"]))

)


thresholdCircle <- make_circles(circles, threshold/1000)

outterCircle <- make_circles(circles, (threshold+thresholdLimit)/1000)

innerCircle <- make_circles(circles, (threshold-thresholdLimit)/1000)
```

Filter to Limit Circles

```
# Any points within the limit


limited_points <- nearest_two %>%

 filter(abs(threshold - distance) < thresholdLimit)
```

```
# Any paired points within the limit (must be from the same journey to be a pair)

limited_points <- limited_points %>%
  group_by(fast_leg_id) %>%
  filter(n() == 2) %>%
  ungroup()
```

Stats

Descriptive stats for this island

View Data

```
lonRange <- extendrange(limited_points$lon, f = 0.5)
latRange <- extendrange(limited_points$lat, f = 0.5)

# Factored by leg

dispersal <-
  ggmap(get_stadiamap(
    bbox = c(lonRange[1], latRange[1], lonRange[2], latRange[2]),
    zoom = 6,
    maptype = "stamen_terrain_background"
  )) +
  geom_point(
    data = limited_points,
    aes(x = lon, y = lat, colour = as.factor(str_to_title(leg))),
```

```
  size = 2,
  alpha = 1
) +
geom_point(
  data = limited_points,
  shape = 1,
  size = 2,
  colour = "black"
) +
geom_polygon(
  data = thresholdCircle,
  aes(lon, lat, group = ID),
  color = "red",
  alpha = 0
) +
geom_polygon(
  data = outterCircle,
  aes(lon, lat, group = ID),
  color = "black",
  alpha = 0
) +
geom_polygon(
  data = innerCircle,
```

```
  aes(lon, lat, group = ID),
  color = "black",
  alpha = 0
) +
labs(
  colour = 'Direction',
  x = "Longitude",
  y = "Latitude"
) +
ggtitle(paste0(
  colloquial_name,
  " - ",
  (threshold - thresholdAdjustment) / 1000,
  " km ±",
  thresholdLimit / 1000
)) +
geom_point(
  aes(x = bird_coords[1], y = bird_coords[2]),
  shape = 4,
  size = 2,
  stroke = 2
) +
scale_color_brewer(palette="Set1", labels = c("Inbound", "Outbound"))
```

Results

Distribution of bearings (test)

Used rao.spacing.test() as it is better for circular statistics. Chosen over rayleigh because it could be multimodal. Null hypothesis, the data is uniformly distributed, reject if p-value is low

This is to be used at a factor level. That is, do once for outgoing and once for incoming. MATLAB is required to compare between them

Calculate bearings

```
with_bearings <- limited_points %>%

  mutate(bearing = bearing(

    cbind(lon, lat), c(bird_coords["lon"], bird_coords["lat"])

  ))



outgoing_bearing <- with_bearings %>% filter(leg == "outgoing") %>% select(bearing)



incoming_bearing <- with_bearings %>% filter(leg == "incoming") %>% select(bearing)
```

Describe Paired Deviation

This section calculates the angular difference between the outgoing and incoming points and determines the distance between the two points if at extrapolated to the threshold distance from shore

```
# Calculate the angle between the two points for each pair of legs

paired_deviation <- with_bearings %>%

     group_by(fast_leg_id) %>%

     summarize(smallest_angle = min(abs(diff(bearing))))



# Calculations to determine the separation at the threshold distance (in kilometers)
```

```
dispersal_distance <- (threshold - thresholdAdjustment) / 1000
paired_deviation$smallest_angle_rad <- paired_deviation$smallest_angle * (pi / 180)
paired_deviation$separation <- 2 * dispersal_distance * sin(paired_deviation$smallest_angle_rad / 2)
paired_deviation$separation_200km <- 2 * 200 * sin(paired_deviation$smallest_angle_rad / 2)
paired_deviation <- paired_deviation %>% mutate(island = "Bird Island")

# Write to csv for comparison between islands
write.csv(paired_deviation, "outputs/bird_deviation.csv")

# Plot the deviation
ggplot(paired_deviation, aes(x = island, y = smallest_angle)) +
  geom_boxplot() +
  labs(
    title = paste0("Bird Paired Leg Deviation (", dispersal_distance, " km)"),
    subtitle = paste0("Pairs = ", nrow(paired_deviation)),
    x = "Island",
    y = "Degrees"
  )
```

```
# Plot the separation
ggplot(paired_deviation, aes(x = island, y = separation)) +
  geom_boxplot() +
```

```
labs(
    title = paste0("Bird Paired Leg Separation (", dispersal_distance, " km)"),
    subtitle = paste0("Pairs = ", nrow(paired_deviation)),
    x = "Island",
    y = "Kilometers"
  )
```

Combined

Look at both incoming and outgoing

```
some_bearings <-
  circular(with_bearings$bearing, type = "angles", units = "degrees")

rao_result <- rao.spacing.test(some_bearings)

rao_sub <- capture.output(print(rao_result))[5]

rose.diag(
  some_bearings,
  bins = 16,
  main = paste0(
    colloquial_name,
    " - ",
    (threshold - thresholdAdjustment) / 1000,
    " km ±",
```

```
    thresholdLimit / 1000,
    " (Combined)"
  ),
  sub = rao_sub,
  ticks = FALSE,
  zero = 4.71239,
  rotation = c("clock")
)
```

```
write.csv(with_bearings$bearing, "outputs/bird_combined.csv", row.names = FALSE)


rao_result
##
##       Rao's Spacing Test of Uniformity
##
## Test Statistic = 234.7145
## P-value < 0.001
##
```

Outgoing

Look at just outgoing

```
some_bearings <-
  circular(outgoing_bearing, type = "angles", units = "degrees")


rao_result <- rao.spacing.test(some_bearings)
```

```r
rao_sub <- capture.output(print(rao_result))[5]

rose.diag(
  some_bearings,
  bins = 16,
  main = paste0(
    "D) ",
    colloquial_name,
    " - ",
    (threshold - thresholdAdjustment) / 1000,
    " km ±",
    thresholdLimit / 1000,
    " (Outbound)"
  ),
  sub = rao_sub,
  ticks = FALSE,
  zero = 4.71239,
  rotation = c("clock")
)
```

```r
rao_result
```

```
##
##        Rao's Spacing Test of Uniformity
##
## Test Statistic = 219.0984
## P-value < 0.001
##
write.csv(outgoing_bearing, "outputs/bird_outgoing.csv", row.names = FALSE)
```

Incoming

Look at just incoming

```
some_bearings <-
  circular(incoming_bearing, type = "angles", units = "degrees")

rao_result <- rao.spacing.test(some_bearings)

rao_sub <- capture.output(print(rao_result))[5]

rose.diag(
  some_bearings,
  bins = 16,
  main = paste0(
    "B) ",
    colloquial_name,
    " - ",
    (threshold - thresholdAdjustment) / 1000,
```

```
    " km ±",
    thresholdLimit / 1000,
    " (Inbound)"
  ),
  sub = rao_sub,
  ticks = FALSE,
  zero = 4.71239,
  rotation = c("clock")
)
```

```
rao_result
##
##        Rao's Spacing Test of Uniformity
##
## Test Statistic = 251.239
## P-value < 0.001
##
write.csv(incoming_bearing, "outputs/bird_incoming.csv", row.names = FALSE)
```

Outgoing vs Incoming

Read the data from MATLAB See circular.m appendix for MATLAB code

```
circ_cmtest <- read.csv("outputs/bird_circ_cmtest.csv", header = TRUE)
print(paste0("Test Statistic: ", circ_cmtest$tstat))
## [1] "Test Statistic: 0.0309533235295281"
```

```
print(paste0("P Value: ", circ_cmtest$pval))

## [1] "P Value: 0.610450679081531"

## Warning: Your circ_cmtest may be out of date;

##  re-run circular.m in MATLAB to be safe
```

Environment Cleaning

Clean environment

```
# Removes temporary variables to make analysis easier

tempVariables <- c("dispersal_total_journeys", "dispersal_total_seals", "i", "ids",

                   "print_kept", "values", "thresholdCircle", "outterCircle",

                   "innerCircle", "circles", "inside")

tempVariables <- append(tempVariables, "tempVariables")

rm(list = tempVariables)
```

Corridors

Filter and Subset

Split loopings into incomings and outgoings (based on midpoint to make easier)

```
bird_looping_outgoing <- bird_looping %>%

  group_by(fast_leg_id) %>%

  slice(seq(1, floor(n()/2))) %>%

  ungroup()



bird_looping_incoming <- bird_looping %>%

  group_by(fast_leg_id) %>%
```

```
  slice(seq(floor(n()/2), n())) %>%

  ungroup()
```

Vector Analysis

Get the bearing they are currently facing at each point Then calculate the u and v for geom_quiver

```
bird_incomings_all <- rbind(bird_incomings, bird_looping_incoming)


# Create a new column to store the bearing

bird_incomings_all$bearing_BfromA = NA


# Loop through each fast_leg_id

all_unique_legs = unique(bird_incomings_all$fast_leg_id)

for (fast_leg_id in all_unique_legs) {

  current_leg = bird_incomings_all[bird_incomings_all$fast_leg_id == fast_leg_id, ]

  current_leg = current_leg[order(current_leg$date), ] # Sort the rows by datetime


  # Set the first row's bearing_BfromA value to NULL

  bird_incomings_all[

    bird_incomings_all$fast_leg_id == fast_leg_id &

      bird_incomings_all$date == current_leg$date[1],

    "bearing_BfromA"] = NA


  # Loop through each row, excluding the first one,

  # and calculate the bearing from the previous row
```

```
  for (i in 2:nrow(current_leg)) {
    current_row = current_leg[i, ]
    prev_row = current_leg[i-1, ]

    bearing_val <- bearing(
      c(prev_row$lat, prev_row$lon), c(current_row$lat, current_row$lon)
    )

    # Transform bearing to the range of 0 to 360 degrees
    bearing_val <- (bearing_val + 360) %% 360

    bird_incomings_all[
      bird_incomings_all$fast_leg_id == fast_leg_id &
        bird_incomings_all$date == current_row$date,
      "bearing_BfromA"] <- bearing_val
  }
}

bird_incomings_all_ends <- na.omit(bird_incomings_all)

corridor_bearings <- bird_incomings_all_ends %>%
  mutate(bearing_BfromA = deg2rad(bearing_BfromA))
```

```
corridor_bearings <- corridor_bearings


corridor_bearings$u <- corridor_bearings$mpm * cos(corridor_bearings$bearing_BfromA)

corridor_bearings$v <- corridor_bearings$mpm * sin(corridor_bearings$bearing_BfromA)
```

Half grid size

```
grid_size <- 0.5


corridor_half <- corridor_bearings %>%

  mutate(

    rounded_lon = floor(lon / grid_size) * grid_size,

    rounded_lat = floor(lat / grid_size) * grid_size

  ) %>%

  group_by(rounded_lon, rounded_lat) %>%

  summarise(

    seal_count = n_distinct(fast_leg_id),

    avg_u = mean(u),

    avg_v = mean(v),

    .groups = "drop"

  )
```

Results

Without background

```
# Define the color gradient

colors <- c(

  "red", "red", "green", "green", "blue", "blue",
```

```
  "black", "black", "yellow", "yellow"
)

breaks <- c(-180, -135, -134, -45, -44,  44, 45, 134, 135, 180)
my_gradient <- scale_color_gradientn(
  colors = colors, values = rescale(breaks, to = c(0, 1))
)

# Create the grid plot
grid_plot <-
  ggplot(
    corridor_half,
    aes(
      x = rounded_lon,
      y = rounded_lat,
      u = avg_u,
      v = avg_v,
      size = 10,
      col = seal_count
    )
  ) +
  geom_quiver(center = TRUE, size = 1) +
  my_gradient +
```

```
  scale_x_continuous(limits = c(
    min(corridor_half$rounded_lon) - 1,
    max(corridor_half$rounded_lon) + 1
  )) +
  scale_y_continuous(limits = c(
    min(corridor_half$rounded_lat) - 1,
    max(corridor_half$rounded_lat) + 1
  )) +
  labs(x = "Longitude", y = "Latitude") +
  theme_bw() +
  geom_point(
    shape = 10,
    stroke = 2,
    size = 10,
    color = "darkviolet",
    aes(x = bird_coords[1], y = bird_coords[2])
  ) +
  ggtitle("Inbound", subtitle = paste0("Grid size: ", grid_size))
```

With Background

```
lonRange <- extendrange(corridor_half$rounded_lon, f = 0.05)
latRange <- extendrange(corridor_half$rounded_lat, f = 0.05)

# Factored by journey
```

```
print_map <-
  ggmap(get_stadiamap(
    bbox = c(lonRange[1], latRange[1], lonRange[2], latRange[2]),
    zoom = 6,
    maptype = "stamen_toner_lite"
  )) +
  geom_quiver(
    data = corridor_half,
    center = TRUE,
    size = 1,
    aes(
      x = rounded_lon,
      y = rounded_lat,
      u = avg_u,
      v = avg_v,
      size = 10,
      col = seal_count
    )
  ) +
  my_gradient +
  labs(x = "Longitude", y = "Latitude") +
  theme_bw() +
```

```
  ggtitle("Inbound", subtitle = paste0("Grid size: ", grid_size))
```

Reduced Corridor Lat and Lon

Reduce the range to more complete region

```
corridor_restrict_coord <- corridor_half %>%
  filter(between(rounded_lon, -42, -38)) %>%
  filter(between(rounded_lat, -55, -52))

# Create the grid plot
grid_plot <-
  ggplot(
    corridor_restrict_coord,
    aes(
      x = rounded_lon,
      y = rounded_lat,
      u = avg_u,
      v = avg_v,
      size = 10,
      col = seal_count
    )
  ) +
  geom_quiver(center = TRUE, size = 1) +
  scale_color_gradient(low = "white", high = "red") +
```

```
scale_x_continuous(limits = c(

  min(corridor_restrict_coord$rounded_lon) - 0.3,

  max(corridor_restrict_coord$rounded_lon) + 0.3

)) +

scale_y_continuous(limits = c(

  min(corridor_restrict_coord$rounded_lat) - 0.3,

  max(corridor_restrict_coord$rounded_lat) + 0.3

)) +

labs(x = "Longitude", y = "Latitude") +

theme(

  panel.background = element_rect(fill = "lightblue"),

  panel.grid.major = element_line(color = "cadetblue"),

  panel.grid.minor = element_line(color = "cadetblue")

) +

geom_point(

  shape = 10,

  stroke = 2,

  size = 10,

  color = "black",

  aes(x = bird_coords[1], y = bird_coords[2])

) +

ggtitle("Inbound", subtitle = paste0("Grid size: ", grid_size))
```

Reduced Corridor; higher resolution

Change grid size to quarters with the reduced range

```
grid_size <- 0.25


corridor_higher_res <- corridor_bearings %>%
  mutate(
    rounded_lon = floor(lon / grid_size) * grid_size,
    rounded_lat = floor(lat / grid_size) * grid_size
  ) %>%
  group_by(rounded_lon, rounded_lat) %>%
  summarise(
    seal_count = n_distinct(fast_leg_id),
    avg_u = mean(u),
    avg_v = mean(v),
    .groups = "drop"
  )



corridor_higher_restricted <- corridor_higher_res %>%
  filter(between(rounded_lon, -42, -38)) %>%
  filter(between(rounded_lat, -55, -52))


# Create the grid plot
grid_plot <-
```

```
ggplot(
  corridor_higher_restricted,
  aes(
    x = rounded_lon,
    y = rounded_lat,
    u = avg_u,
    v = avg_v,
    size = 10,
    col = seal_count
  )
) +
geom_quiver(center = TRUE, size = 1) +
scale_color_gradient(low = "white", high = "red") +
scale_x_continuous(limits = c(
  min(corridor_higher_restricted$rounded_lon) - 0.3,
  max(corridor_higher_restricted$rounded_lon) + 0.3
)) +
scale_y_continuous(limits = c(
  min(corridor_higher_restricted$rounded_lat) - 0.3,
  max(corridor_higher_restricted$rounded_lat) + 0.3
)) +
labs(x = "Longitude", y = "Latitude") +
theme(
```

```
   panel.background = element_rect(fill = "lightblue"),
   panel.grid.major = element_line(color = "cadetblue"),
   panel.grid.minor = element_line(color = "cadetblue")
  ) +
  geom_point(
   shape = 10,
   stroke = 2,
   size = 10,
   color = "black",
   aes(x = bird_coords[1], y = bird_coords[2])
  ) +
  ggtitle("Inbound - Bird Island", subtitle = paste0("Grid size: ", grid_size)) +
  labs(colour = "Seal Count")
```

Bathymetry

```
bird_bathymetry <- getNOAA.bathy(
  lon1 = min(corridor_higher_restricted$rounded_lon) - 0.3,
  lon2 = max(corridor_higher_restricted$rounded_lon) + 0.3,
  lat1 = min(corridor_higher_restricted$rounded_lat) - 0.3,
  lat2 = max(corridor_higher_restricted$rounded_lat) + 0.3,
  resolution = 1
)
## Querying NOAA database ...
## This may take seconds to minutes, depending on grid size
```

```
## Building bathy matrix ...
rescaled <- LinScale(0, min(bird_bathymetry), max(bird_bathymetry), 0, 1)[1]


corridor_plot <- autoplot.bathy(bird_bathymetry, geom = c("tile")) +
  scale_fill_gradientn(
    values = c(0, rescaled - 0.001, rescaled, rescaled + 0.001, 1),
    colours = c("dodgerblue4", "lightblue", "cyan", "lightgreen", "darkgreen")
  ) +
  labs(y = "Latitude", x = "Longitude", fill = "Elevation") +
  coord_cartesian(expand = 0) +
  geom_quiver(
    data = corridor_higher_restricted,
    aes(
      x = rounded_lon,
      y = rounded_lat,
      u = avg_u,
      v = avg_v,
      size = 10,
      col = seal_count
    ),
    center = TRUE,
    size = 1
  ) + scale_color_gradient(low = "white", high = "red") +
```

```
  geom_point(
    shape = 10,
    stroke = 2,
    size = 10,
    color = "black",
    aes(x = bird_coords[1], y = bird_coords[2])
  ) +
  ggtitle("Bathymetry", subtitle = "Bird Island") +
  labs(colour = "Seal Count") +
  guides(fill = guide_colorbar(order = 1), colour = guide_colorbar(order = 2))
## Coordinate system already present. Adding new coordinate system, which will
## replace the existing one.
corridor_plot
```

Magnetic Inclination

Create magnetic grid based on the 2020 published values

```
magnetic_grid <- igrf::igrf_grid(
  field = "main",
  year = 2020,
  altitude = 0,
  resolution = 1
)
```

Code Defining boundaries and plot

```r
lat_min <- (min(corridor_higher_restricted$rounded_lat) - 1)
lat_max <- (max(corridor_higher_restricted$rounded_lat) + 1)

lon_min <- (min(corridor_higher_restricted$rounded_lon) - 1)
lon_max <- (max(corridor_higher_restricted$rounded_lon) + 1)

subset_magnetic_grid <- magnetic_grid[magnetic_grid$lat >= lat_min &
                                      magnetic_grid$lat <= lat_max &
                                      magnetic_grid$lon >= lon_min &
                                      magnetic_grid$lon <= lon_max,]

base <- ggplot() +
  geom_contour_fill(
    data = subset_magnetic_grid,
    aes(x = lon, y = lat, z = I),
    breaks = MakeBreaks(0.25),
    size = 0.2
  ) +
  scale_fill_gradient(name = "Inclination",
                      low = "royalblue3",
                      high = "lightblue") +
  xlim(lon_min, lon_max) +
  ylim(lat_min, lat_max)  +
```

```
labs(
  title = "Magnetic Inclination",
  subtitle = "Bird Island",
  x = "Longitude",
  y = "Latitude",
  colour = "Seal Count"
) +
geom_quiver(
  data = corridor_higher_restricted,
  aes(
    x = rounded_lon,
    y = rounded_lat,
    u = avg_u,
    v = avg_v,
    size = 10,
    col = seal_count
  ),
  center = TRUE,
  size = 1
) +
scale_color_gradient(low = "white", high = "red") +
guides(fill = guide_colorbar(order = 1), colour = guide_colorbar(order = 2)) +
geom_point(
```

```
    shape = 10,
    stroke = 2,
    size = 10,
    color = "black",
    aes(x = bird_coords["lon"], y = bird_coords["lat"])
  ) +
  geom_contour2(
    data = subset_magnetic_grid,
    aes(x = lon, y = lat, z = I, label = stat(level), label_colour = "black", fontface = "bold"),
    label.placer = label_placer_fraction(0.6, isoband::angle_fixed(theta = 0)),
    skip = 2,
    breaks = MakeBreaks(0.25)
  ) +
  scale_x_continuous(expand = c(0,0)) +
  scale_y_continuous(expand = c(0,0))
```

```
## Scale for x is already present.
## Adding another scale for x, which will replace the existing scale.
## Scale for y is already present.
## Adding another scale for y, which will replace the existing scale.
```

Plot Magnetic Inclination

```
base
```

Ocean Currents

```
drift <- nc_open("gdp6h_ragged_current.nc")


current <- data.frame(id = ncvar_get(drift, "ids"), lon = ncvar_get(drift, "lon"), lat = ncvar_ge
t(drift, "lat"), time = ncvar_get(drift, "time"), v = ncvar_get(drift, "ve"), u = ncvar_get(drift, "v
n"))


current_sf <- current %>%
  filter(lon > min(corridor_higher_restricted$rounded_lon) - 0.3) %>%
  filter(lon < max(corridor_higher_restricted$rounded_lon) + 0.3) %>%
  filter(lat > min(corridor_higher_restricted$rounded_lat) - 0.3) %>%
  filter(lat < max(corridor_higher_restricted$rounded_lat) + 0.3) %>%
  st_as_sf(coords = c("lon", "lat")) %>%
  st_set_crs(4326)


current_grid <- current_sf %>%
  st_make_grid(n = c(70,60)) %>%
  st_sf()




drifter.split.sf.se = current_sf


drifter.gridded = current_grid %>%
  mutate(id = 1:n(), contained = lapply(st_contains(st_sf(geometry),drifter.split.sf.se),identity),
```

```
      obs = sapply(contained, length),

      u = sapply(contained, function(x) {median(drifter.split.sf.se[x,]$u, na.rm = TRUE)}),

      v = sapply(contained, function(x) {median(drifter.split.sf.se[x,]$v, na.rm = TRUE)}))



drifter.gridded = drifter.gridded %>% select(obs, u, v) %>% na.omit()


## obtain the centroid coordinates from the grid as table

coordinates = drifter.gridded %>%

  st_centroid() %>%

  st_coordinates() %>%

  as_tibble() %>%

  rename(x = X, y = Y)

## Warning: st_centroid assumes attributes are constant over geometries

## remove the geometry from the simple feature of gridded drifter dataset

st_geometry(drifter.gridded) = NULL


## stitch together the extracted coordinates and drifter information int a single table for SE m
onsoon season

current.gridded.se = coordinates %>%

  bind_cols(drifter.gridded) %>%

  mutate(season = "SE")
```

```r
drifter.current.gridded = current.gridded.se

## select grids for SE season only
drf.se = drifter.current.gridded

## interpolate the U component
u.se = interpBarnes(x = drf.se$x, y = drf.se$y, z = drf.se$u)

## obtain dimension that determine the width (ncol) and length (nrow) for tranforming wide into long format table
dimension = data.frame(lon = u.se$xg, u.se$zg) %>% dim()

## make a U component data table from interpolated matrix
u.tb = data.frame(lon = u.se$xg,
                  u.se$zg) %>%
  gather(key = "lata", value = "u", 2:dimension[2]) %>%
  mutate(lat = rep(u.se$yg, each = dimension[1])) %>%
  select(lon,lat, u) %>% as.tibble()
## Warning: `as.tibble()` was deprecated in tibble 2.0.0.
## ℹ Please use `as_tibble()` instead.
## ℹ The signature and semantics have changed, see `?as_tibble`.
## This warning is displayed once every 8 hours.
```

```
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
## interpolate the V component
v.se = interpBarnes(x = drf.se$x,
                    y = drf.se$y,
                    z = drf.se$v)

## make the V component data table from interpolated matrix
v.tb = data.frame(lon = v.se$xg, v.se$zg) %>%
  gather(key = "lata", value = "v", 2:dimension[2]) %>%
  mutate(lat = rep(v.se$yg, each = dimension[1])) %>%
  select(lon,lat, v) %>%
  as.tibble()

## stitch now the V component intot the U data table and compute the velocity
uv.se = u.tb %>%
  bind_cols(v.tb %>% select(v)) %>%
  mutate(vel = sqrt(u^2+v^2))
ocean_plot <- ggplot() +
  geom_quiver(
    data = corridor_higher_restricted,
    aes(
      x = rounded_lon,
```

```
    y = rounded_lat,
    u = avg_u,
    v = avg_v,
    size = 10,
    col = seal_count
  ),
  center = TRUE,
  size = 1
) +
metR::geom_vector(
  data = uv.se,
  aes(
    x = lon,
    y = lat,
    dx = u,
    dy = v
  ),
  min.mag = 0.2,
  skip.y = 5,
  skip.x = 2,
  arrow.angle = 30,
  arrow.type = "open",
  arrow.length = 0.5,
```

```
  preserve.dir = TRUE,
  direction = "ccw",
  colour = "black",
  show.legend = TRUE
) +
labs(x = "", y = "")  +
scale_color_gradient(low = "white", high = "red") +
theme(
  panel.background = element_rect(fill = "lightblue"),
  panel.grid.major = element_line(color = "lightblue"),
  panel.grid.minor = element_line(color = "lightblue")
) +
geom_point(
  shape = 10,
  stroke = 2,
  size = 10,
  color = "black",
  aes(x = bird_coords["lon"], y = bird_coords["lat"])
) +
labs(
  title = "Ocean Currents",
  subtitle = "Bird Island",
  x = "Longitude",
```

```
  y = "Latitude",
  colour = "Seal Count"
 )
ocean_plot
```

Water Temperature

```
sst <- nc_open("ocean_temp_ann_2020.nc")


sst_lon  <- ncvar_get(sst, "xt_ocean" )
sst_lat  <- ncvar_get(sst, "yt_ocean" )
sst_temp <- ncvar_get(sst, "temp")


# Which depth slice to get
depth_slice <- 1


depth_max <- ncvar_get(sst, "st_ocean")[depth_slice]
surface_slice <- sst_temp[, , depth_slice]


sst_matrix <- as.matrix(expand.grid(sst_lon,sst_lat))


sst_vector <- as.vector(surface_slice)


sst_df <- data.frame(cbind(sst_matrix,sst_vector))
names(sst_df) <- c("lon","lat","temps")
```

```
# Adjust lon values if needed
sst_df$lon <- ifelse(sst_df$lon > 180, sst_df$lon - 360, sst_df$lon)

# Filter sst_df based on the latitudinal and longitudinal ranges
sst_df_filtered <- sst_df %>%
  filter(
    lat >= min(corridor_higher_restricted$rounded_lat) - 0.3,
    lat <= max(corridor_higher_restricted$rounded_lat) + 0.3,
    lon >= min(corridor_higher_restricted$rounded_lon) - 0.3,
    lon <= max(corridor_higher_restricted$rounded_lon) + 0.3
  )
# Plot with filtered data
plot_temperature <- ggplot() +
  geom_contour_fill(data = sst_df_filtered, aes(lon, lat, z = temps)) +
  geom_quiver(
    data = corridor_higher_restricted,
    aes(
      x = rounded_lon,
      y = rounded_lat,
      u = avg_u,
      v = avg_v,
      size = 10,
```

```
    col = seal_count
  ),
  center = TRUE,
  size = 1
) +
scale_color_gradient(low = "white", high = "red") +
labs(
  title = paste0("Water Temperature at ", depth_max, "m"),
  subtitle = "Bird Island",
  x = "Longitude",
  y = "Latitude",
  colour = "Seal Count",
  fill = "Temperature"
) +
guides(fill = guide_colorbar(order = 1), colour = guide_colorbar(order = 2)) +
geom_point(
  shape = 10,
  stroke = 2,
  size = 10,
  color = "black",
  aes(x = bird_coords["lon"], y = bird_coords["lat"])
) +
scale_x_continuous(expand = c(0,0)) +
```

```
  scale_y_continuous(expand = c(0,0)) +

  theme(

    panel.background = element_rect(fill = "#777777")

  )

plot_temperature
```

Bearings

Prepare Data

Collect data from bearing A to B

```
bearings_data <- na.omit(bird_incomings_all)
```

Calculate bearing B to C as extra column

```
predefined_lat = bird_coords[[2]]

predefined_lon = bird_coords[[1]]



# Create a new column to store the bearing

bearings_data$bearing_BtoC = NA



# Loop through each fast_leg_id

all_unique_legs = unique(bearings_data$fast_leg_id)

for (fast_leg_id in all_unique_legs) {

  # Sort the rows by datetime

  journey_bearings_data = bearings_data[bearings_data$fast_leg_id == fast_leg_id, ]

  journey_bearings_data = journey_bearings_data[order(journey_bearings_data$date), ]
```

```
# For the last row, calculate the bearing from the predefined lat and lon
last_row = journey_bearings_data[nrow(journey_bearings_data), ]
bearings_data[bearings_data$fast_leg_id == fast_leg_id &
                bearings_data$date == last_row$date,
              "bearing_BtoC"] = bearing(c(last_row$lon, last_row$lat),
                                        c(predefined_lon, predefined_lat))

# Loop through rows, excluding the last one, and calculate the bearing to the next row
for (i in 1:(nrow(journey_bearings_data) - 1)) {
  current_row = journey_bearings_data[i, ]
  next_row = journey_bearings_data[i + 1, ]
  bearings_data[bearings_data$fast_leg_id == fast_leg_id &
                  bearings_data$date == current_row$date,
                "bearing_BtoC"] = bearing(c(current_row$lon, current_row$lat),
                                          c(next_row$lon, next_row$lat))
 }
}
```

Absolute bearing change & bearing from island

```
bearings_data <- bearings_data %>%
  mutate(bearing_angleAtB = angle_diff(bearing_BfromA, bearing_BtoC))


bearings_data <- bearings_data %>%
  mutate(bearing_BtoIsland = bearing(cbind(lon, lat), c(predefined_lon, predefined_lat)))
```

```
bearings_data <- bearings_data %>%
  mutate(bearing_abs_island = angle_diff(bearing_BtoIsland, bearing_BtoC))
```

Results

Plot angle and distance

```
ggplot(bearings_data %>% filter(distance/1000 < 500)) +
  geom_vline(xintercept = 125, colour = "red", size = 2) +
  geom_point(aes(x = distance/1000, y = bearing_abs_island, color = ifelse(distance > 125000, "
beyond_125", "within_125"))) +
  scale_color_manual(values = c("within_125" = "black", "beyond_125" = "gray")) +
  xlab("Distance from Shore (km)") +
  ylab("Deviation from Island in Degrees") +
  labs(
    title = "Bird Island"
  ) +
  theme(legend.position="none")
```

Stats on groups of distance vs angle (calculate distance intervals)

```
distance_interval <- bearings_data %>%
  select(distance, bearing_abs_island, fast_leg_id) %>%
  rename(angle_diff = bearing_abs_island) %>%
  mutate(distance = distance/1000)
```

Filter to intervals less than 1000km

```
distance_interval <- distance_interval %>% filter(distance < 1000)


# Calculate the minimum and maximum distances in your dataframe
min_distance <- min(distance_interval$distance)
max_distance <- max(distance_interval$distance)


# Create intervals of 200km each
intervals <- seq(min_distance, max_distance, by = 200)
distance_interval$distance_interval <- cut(distance_interval$distance,
                                           breaks = intervals,
                                           include.lowest = TRUE)
```

Skillings Mack

Create Ranks

Create ranks and define interval size for skillings mack

```
# Distance grouping in kilometers (convert to meters)
skillings_size <- 25
skillings_size <- skillings_size * 1000
times_to_simulate <- 10000


# Keep journeys with this many intervals
# If size is 50km and interval is 3, we keep up to 150km
# However we exclude first interval (in this case, 50km)
# Therefore we only look at 50-100 and 100-150
skillings_interval <- 5
```

```
# Keep the journeys that have at least 70% complete values
skillings_limits <- (skillings_interval - 1) * 0.7


skillings <- bearings_data %>%
  mutate(
    interval = cut(
      distance,
      breaks = seq(0,
                   ceiling(max(distance) / skillings_size) * skillings_size,
                   by = skillings_size
                 ),
      include.lowest = TRUE,
      labels = FALSE
    )
  ) %>%
  group_by(fast_leg_id, interval) %>%
  summarise(median_bearing = median(bearing_abs_island, na.rm = TRUE)) %>%
  ungroup()
## `summarise()` has grouped output by 'fast_leg_id'. You can override using the
## `.groups` argument.
```

Display number of individuals in each interval

```
skillings %>%
  group_by(interval) %>%
  count(interval) %>%
  mutate(max_km = interval*skillings_size/1000)
## # A tibble: 21 × 3
## # Groups:   interval [21]
##    interval     n max_km
##       <int> <int>  <dbl>
##  1        1   113     25
##  2        2   109     50
##  3        3   106     75
##  4        4    87    100
##  5        5    90    125
##  6        6    72    150
##  7        7    53    175
##  8        8    43    200
##  9        9    24    225
## 10       10    10    250
## # ℹ 11 more rows
```

Remove tracks that are biasing results as per comments

```
# Remove tracks that are less than first interval
# Remove tracks above certain interval size
# Shift interval range by 1 to start at 1 instead of 2
```

```
skillings_raw <- skillings %>%
  filter(interval >= 2 & interval <= skillings_interval) %>%
  group_by(fast_leg_id) %>%
  mutate(interval = interval - 1) %>%
  rename(interval_shifted = interval) %>%
  ungroup()
```

Run tests

Run Skillings Mack Focus on tracks with more than Skillings_limits percentage of complete values

```
# fast_leg_id expands to all rows
# interval_shifted fills in blanks from 1 to range
# median_bearing is filled with NA values

group <- unique(skillings_raw$fast_leg_id)
n <- max(skillings_raw$interval_shifted)

skillings_blank <- data.frame(
  fast_leg_id = rep(group, each = n),
  interval_shifted = unlist(lapply(split(
    1:(n * length(group)), group
  ), function(x)
    1:length(x))),
  row.names = NULL
)
```

```
skillings_test <-
  full_join(skillings_blank,
            skillings_raw,
            by = join_by(fast_leg_id, interval_shifted))

# Remove journeys that have less than skillings_limits (percentage) of complete values

skillings_NA <- skillings_test %>%
  group_by(fast_leg_id) %>%
  mutate(non_na_count = sum(!is.na(median_bearing))) %>%
  ungroup() %>%
  filter(non_na_count > skillings_limits) %>%
  select(-non_na_count)

skillings_matrix <- skillings_NA %>%
  pivot_wider(names_from = interval_shifted, values_from = median_bearing) %>%
  select(-fast_leg_id)

# Switch blocks and treatments order
skillings_run <- t(skillings_matrix)
Ski_total <-
  Ski.Mack(
```

```
    y = data.matrix(skillings_run),
    simulate.p.value = TRUE,
    B = times_to_simulate,
    suppress = TRUE
  )
##
## Skillings-Mack Statistic =  28.071977 , p-value =  4e-06
## Note: the p-value is based on the chi-squared distribution with d.f. =  3
## Based on B =  10000 , Simulated p-value =  1e-04
```

Plot Intervals

```
plot_skillings <- drop_na(skillings_NA)


ggplot(plot_skillings, aes(interval_shifted, median_bearing)) +
  geom_boxplot(aes(group = interval_shifted)) +
  xlab("Interval (km)") +
  ylab("Median Bearing of Each Seal") +
  scale_x_continuous(breaks = 1:4,  labels=c("25-50", "50-75", "75-100", "100-125")) +
  labs (
    title = "Bird Island"
  ) +
  geom_label(data = plot_skillings %>%
              group_by(interval_shifted) %>%
              summarise(count = n()),
            aes(x = interval_shifted, label = paste0("n = ", count)), y = 115)
```

Remove First Interval

```
skillings_run_tail <- skillings_run[2:nrow(skillings_run),]


Ski_ignore_first <-
  Ski.Mack(
    y = data.matrix(skillings_run_tail),
    simulate.p.value = TRUE,
    B = times_to_simulate,
    suppress = TRUE
  )
##
## Skillings-Mack Statistic =  7.914585 , p-value =  0.019115
## Note: the p-value is based on the chi-squared distribution with d.f. =  2
## Based on B =  10000 , Simulated p-value =  0.019500
```

Calculate percentage of NA for each interval

```
## Percentage of NULL at Each Interval
## Interval: 25km
##     1     2     3     4
## 10.20  9.18 15.31 12.24
```

Consecutive T Tests

T-test to show general but less accurate method

```

t.test(skillings_run[1,], skillings_run[2,], alternative = "l")
```

```
##
##  Welch Two Sample t-test
##
## data:  skillings_run[1, ] and skillings_run[2, ]
## t = -0.55354822, df = 174.99178, p-value = 0.2902973
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##        -Inf 4.668495494
## sample estimates:
##   mean of x   mean of y
## 16.19878121 18.54796055
t.test(skillings_run[2,], skillings_run[3,], alternative = "l")
##
##  Welch Two Sample t-test
##
## data:  skillings_run[2, ] and skillings_run[3, ]
## t = -1.1951365, df = 157.46658, p-value = 0.1169153
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##       -Inf 2.24520433
## sample estimates:
##   mean of x   mean of y
## 18.54796055 24.38824256
```

```
t.test(skillings_run[3,], skillings_run[4,], alternative = "l")

##

##  Welch Two Sample t-test

##

## data:  skillings_run[3, ] and skillings_run[4, ]

## t = -1.584498, df = 165.90564, p-value = 0.05749237

## alternative hypothesis: true difference in means is less than 0

## 95 percent confidence interval:

##          -Inf 0.4000617811

## sample estimates:

##   mean of x   mean of y

## 24.38824256 33.49700151
```

Filter to complete intervals

Filter to only complete so we can use well known tests on subset

```
skillings_complete <- skillings_run[, colSums(is.na(skillings_run)) == 0]
```

Sign Test R

```
skillings_sign_test <- sign(skillings_complete[1,] - skillings_complete[2,])

wilcox.test(skillings_sign_test, mu = 0)

##

##  Wilcoxon signed rank test with continuity correction

##

## data:  skillings_sign_test

## V = 609.5, p-value = 0.4083406
```

```
## alternative hypothesis: true location is not equal to 0

skillings_sign_test <- sign(skillings_complete[2,] - skillings_complete[3,])
wilcox.test(skillings_sign_test, mu = 0)
##
##  Wilcoxon signed rank test with continuity correction
##
## data:  skillings_sign_test
## V = 556.5, p-value = 0.1671199
## alternative hypothesis: true location is not equal to 0

skillings_sign_test <- sign(skillings_complete[3,] - skillings_complete[4,])
wilcox.test(skillings_sign_test, mu = 0)
##
##  Wilcoxon signed rank test with continuity correction
##
## data:  skillings_sign_test
## V = 450.5, p-value = 0.01274136
## alternative hypothesis: true location is not equal to 0
```

Friedman Test

```
friedman.test(skillings_complete)
##
##  Friedman rank sum test
##
```

```
## data:  skillings_complete

## Friedman chi-squared = 95.310092, df = 51, p-value = 0.0001676726
```

Lunar

Download Lunar Fractions

Download lunar fractions based on year range of the dataset

```
years <- c(year(min(bird_incomings_all$date)):year(max(bird_incomings_all$date)))


lunar_raw <- NULL


for(lunaryear in years) {


  lunarl <- paste0(
    "https://aa.usno.navy.mil/calculated/moon/fraction?year=",
    lunaryear,
    "&task=00&tz=3&tz_sign=-1&tz_label=false&submit=Get+Data"
  )


  lunar_html <- read_html(lunarl)


  textData <- lunar_html %>%
    html_node("body") %>%
    html_table(header = TRUE)
```

```
 textData <- textData[-1, ]

 colnames(textData) <- textData[1, ]

 textData <- textData[-1, ]

 textData[] <- lapply(textData, as.numeric)

 textData <- textData %>%
   pivot_longer(cols = -Day, names_to = "month", values_to = "lunar")

 textData <- na.omit(textData)

 textData <- textData %>%
   mutate(year = lunaryear)

 lunar_raw <- rbind(lunar_raw, textData)

}
```

Convert HTML format to Posix compatible Store lunar fraction for every day in the years period

```
lunar_values <- lunar_raw
```

```
lunar_values$month <- sub("July", "Jul.", lunar_values$month)

lunar_values$month <- sub("June", "Jun.", lunar_values$month)

lunar_values$month <- gsub("\\.", "", lunar_values$month)


month_mapping <- setNames(1:12, month.abb)


lunar_values <- lunar_values %>%

  mutate(month = month_mapping[month])


lunar_values <- lunar_values %>%

  mutate(Date = as.Date(paste(year, month, Day, sep = "-")))


# Convert to POSIXct with GMT-3 timezone and set time to midday

lunar_values$DateTime <-

  as.POSIXct(paste(lunar_values$Date, "12:00:00"),

             format = "%Y-%m-%d %H:%M:%S",

             tz = "Etc/GMT-3")


# Remove the Date column if no longer needed

lunar_values <- lunar_values %>% select(DateTime, lunar) %>% rename(date = DateTime)
```

Regrab clean bearings dataset

```
lunar_journeys <- na.omit(bird_incomings_all) %>%

  mutate(bearing_BfromA = deg2rad(bearing_BfromA))
```

Correlate Fractions

Convert dates of data to correct timezone

```
# Get start of journeys
lunar_times <- lunar_journeys %>%
  group_by(fast_leg_id) %>%
  filter(row_number() == 1)


# Get lat/lon for timezones and date for time
lunar_times <- lunar_times %>%
  select(c("date", "lat", "lon")) %>% ungroup()
## Adding missing grouping variables: `fast_leg_id`
# Create column with Bird Island Time (GMT-3)
lunar_times <- lunar_times %>%
  mutate(date_gmt3 = with_tz(date, tzone = "Etc/GMT-3"))
```

Results

add lunar fraction to each point from lunar values

```
lunar_values <- lunar_values %>% rename(lunar_date = date)
lunar_values$lunar_date <- as.Date(ymd_hms(lunar_values$lunar_date))


lunar_times$date_gmt3 <- as.POSIXct(format(lunar_times$date_gmt3, "%Y-%m-%d 12:00:00"))
lunar_times$date_gmt3 <- as.Date(ymd_hms(lunar_times$date_gmt3))


lunar_reading <-
```

```
 inner_join(lunar_values, lunar_times, by = join_by(lunar_date == date_gmt3)) %>%

 select(fast_leg_id, date, lunar, lat, lon) %>%

 mutate(year = as.character(lubridate::year(date)))
```

Visualize the spread of data

```
# Number of bins

bins_count <- 7



# Expected proportion values

expected_proportion <- layer_data(ggplot(lunar_raw) + geom_histogram(aes(x=lunar), bins =

bins_count))$count/nrow(lunar_raw)



# Navigation distribution (by year) - proportion

nav_plot <- ggplot(lunar_reading) +

 geom_histogram(aes(x=lunar, fill = year, group = year, y = after_stat(count)/sum(after_stat(

count))), bins = bins_count) +

 scale_x_continuous(breaks = c(0, 0.5, 1)) +

 ylab("Observed Proportion") +

 xlab("Lunar Fraction") +

 labs(

  title = "Bird Island",

  fill = "Year"

 )
```

```
nav_lines <- lunar_raw %>%
  filter(year %in% lunar_reading$year)

pre_moons <- nav_plot +
  geom_point(data = nav_lines, shape = "\u25AC", stat = "bin", bins = bins_count, aes(x = lunar, y = after_stat(count)/sum(after_stat(count))), color = "black", size = 10)
# Message = false to avoid errors in knitting; converts line to dotted line

# Suppress warnings to hide SVG code
suppressWarnings({moons(pre_moons, fill = "white")})
```

Grouped Chi-square

Group to increase count size in each bin (adjusted probability due to both quarters)

```
# data.frame(cut(lunar_reading$lunar, 3, labels = c("New Moon", "Quarter", "Full")))

# Create a contingency table
contingency_table <- layer_data(ggplot(lunar_reading) + geom_histogram(aes(x=lunar), bins = bins_count))$count

# Grab the counts of each bin and convert to proportion
adjusted_prob <- layer_data(ggplot(lunar_raw) + geom_histogram(aes(x=lunar), bins = bins_count))$count/nrow(lunar_raw)

# Perform chi-squared test
```

```
chi_squared_test <- chisq.test(contingency_table, p = adjusted_prob)


# Print the results
print(chi_squared_test)
##
##  Chi-squared test for given probabilities
##
## data:  contingency_table
## X-squared = 1.9239378, df = 6, p-value = 0.9265593
```

Raw expected proportions

```
round(adjusted_prob, 3)
## [1] 0.185 0.148 0.109 0.112 0.118 0.141 0.186
```

Descriptive

Prepare Data

Regrab clean dataset

```
bird_incomings_all_ends <- na.omit(bird_incomings_all)


descriptive_filtered <- bird_incomings_all_ends %>%
  mutate(bearing_BfromA = deg2rad(bearing_BfromA))
```

Arrival and Duration

Return Month

Bar chart for arrival month

```
arrivalMonth <- descriptive_filtered %>%
  group_by(fast_leg_id) %>%
```

```
  slice(n()) %>%
  mutate(month = month(date, label = TRUE, abbr = FALSE)) %>%
  ungroup()


ggplot(arrivalMonth, aes(month)) +
  geom_bar() +
  scale_x_discrete(limits = month.name)
```

> Gender (answer is undefined)

```
animalSex <- bird_island %>%
  select(id, sex) %>%
  distinct(id, .keep_all = TRUE)


animalID <- arrivalMonth %>%
  separate_wider_delim(id, "-", cols_remove = FALSE, names = c("animalID", NA)) %>%
  select(animalID, id) %>%
  distinct(animalID, .keep_all = TRUE)


genders <- right_join(animalSex, animalID, by=join_by(id == animalID)) %>%
  rename(longID = `id.y`)


gendered <- left_join(arrivalMonth, genders %>% select(-id), by=join_by(id == longID))


unique(gendered$sex)
```

```
## [1] "F" "M" ""
```

Return Date vs Return Duration

Box and whisker for arrival and duration

```
maxTime <- descriptive_filtered %>%

  group_by(fast_leg_id) %>%

  summarise(maxDate = max(date))



minTime <- descriptive_filtered %>%

  group_by(fast_leg_id) %>%

  summarise(minDate = min(date))



return_times <- full_join(maxTime, minTime, by = join_by(fast_leg_id))



return_times <- return_times %>%

  mutate(returnTime = as.numeric(difftime(maxDate, minDate, units = "days")))



return_times <- return_times %>%

  mutate(month = month(maxDate, label = TRUE, abbr = FALSE))




ggplot(return_times, aes(month, returnTime)) +

  geom_boxplot()  +

  scale_x_discrete(limits = month.name) +

  geom_jitter(width = 0.2)
```

Offset Distance for Return Start

```
offset_dist <- descriptive_filtered %>%
  group_by(fast_leg_id) %>%
  mutate(month = month(date, label = TRUE, abbr = FALSE)) %>%
  filter(row_number()==1) %>%
  mutate(distance_km = distance/1000)


ggplot(offset_dist, aes(month, distance_km)) +
  geom_boxplot()  +
  scale_x_discrete(limits = month.name) +
  geom_jitter(width = 0.2)
```

Other Descriptive Stats

```
print(paste0("Return Leg Count: ", length(unique(offset_dist$fast_leg_id))))
## [1] "Return Leg Count: 128"


print(paste0("Mean Starting Return Distance: ", round(mean(offset_dist$distance_km), 2), "k
m"))
## [1] "Mean Starting Return Distance: 131.43km"


print(paste0("Mean Return Time: ", round(mean(return_times$returnTime), 2), " day(s)"))
## [1] "Mean Return Time: 3.28 day(s)"
```

Appendix Three: Helper Functions

This appendix is a collection of R functions used to reduce some of the code redundancy between analyses of each of the islands. Some of this information has been taken from Stack Overflow and Wikimedia Commons. The links to these files and documents have been provided when used. All code provided as is.

```
nearest <- function(list, threshold) {
  if(min(list) >= threshold)
    stop("Threshold must be larger than the smallest value in the list")
  if(max(list) <= threshold)
    stop("Threshold must be smaller than the largest value in the list")
  if(list[1] >= threshold)
    stop("First element of list must be smaller than threshold")
  if(list[length(list)] >= threshold)
    stop("Last element of list must be smaller than threshold")
  for(i in 1:length(list)) {
    if(list[i] > threshold) {
      if(threshold - list[i - 1] <= list[i] - threshold) {
        first <- (i-1)
        break
      } else {
        first <- i
        break
      }
    }
```

```
  }

  for(j in length(list):1) {
    if(list[j] > threshold) {

      if(threshold - list[j + 1] <= list[j] - threshold) {
        last <- (j+1)
        break
      } else {
        last <- j
        break
      }
    }
  }
  return(c(first, last))
}
# https://stackoverflow.com/a/34187454/6186212
make_circles <- function(centers, radius, nPoints = 100){
    meanLat <- mean(centers$latitude)
    radiusLon <- radius /111 / cos(meanLat/57.3)
    radiusLat <- radius / 111
    circleDF <- data.frame(ID = rep(centers$ID, each = nPoints))
    angle <- seq(0,2*pi,length.out = nPoints)
```

```
    circleDF$lon <- unlist(lapply(centers$longitude, function(x) x + radiusLon * cos(angle)))
    circleDF$lat <- unlist(lapply(centers$latitude, function(x) x + radiusLat * sin(angle)))
    return(circleDF)
}

angle_diff <- function(theta1, theta2){
  theta <- abs(theta1 - theta2) %% 360
  return(ifelse(theta > 180, 360 - theta, theta))
}
# Rolling slope
trackDirection <- function(col, threshold = 0.5) {
  colMean <- mean(col)
  absMean <- mean(abs(col))
  absColMean <- abs(colMean)

  if(colMean > 0) {
    direction <- "Outgoing"
  } else {
    direction <- "Incoming"
  }

  if(absMean * threshold > absColMean) {
    direction <- "Looping"
```

```
  }

  return(direction)
}

moons <- function(plot, offset = -0.01, fill = "white") {
  moon_new <- data.frame(
    x = 0,
    y = offset,
    svg = paste(readLines("https://upload.wikimedia.org/wikipedia/commons/archive/5/59/20211115011522%21New_moon_symbol.svg"), collapse = "\n")
  )

  moon_quarter <- data.frame(
    x = 0.5,
    y = offset,
    svg = paste(readLines("https://upload.wikimedia.org/wikipedia/commons/archive/a/af/20211122091343%21First_quarter_moon_symbol.svg"), collapse = "\n")
  )

  moon_full <- data.frame(
    x = 1,
    y = offset,
```

```
    svg = paste(readLines("https://upload.wikimedia.org/wikipedia/commons/archive/5/57/2021
1115011427%21Full_moon_symbol.svg"), collapse = "\n")
  )


  plot <- plot +
    geom_point_svg(data = moon_new, aes(x, y), svg = moon_new$svg)  +
    geom_point_svg(data = moon_quarter, aes(x, y), svg = moon_quarter$svg, css("circle", 'fill'=
UQ(fill)))  +
    geom_point_svg(data = moon_full, aes(x, y), svg = moon_full$svg, css("circle", 'fill'=UQ(fill
)))


  return(plot)
}
```

Appendix Four: MATLAB Code

This appendix is a collection of all the code produced for the between island analysis that was performed in MATLAB. All code provided as is.

```
% Required Add-on: Circular Statistics Toolbox (Directional Statistics)
% Link to download: https://www.mathworks.com/matlabcentral/fileexchange/10676-circular-
statistics-toolbox-directional-statistics
% Instructions: Download and install the add-on before running this script.

% Check if the required add-on is installed
if ~exist('circ_corrcc', 'file')
    error('Error: The required add-on is not installed. Please download and install it.');
end

clc
clear

outgoing = deg2rad(table2array(readtable("outputs\bird_outgoing.csv")));
incoming = deg2rad(table2array(readtable("outputs\bird_incoming.csv")));

% Perform Circular-Circular Correlation Test
[pval, tstat] = circ_corrcc(outgoing, incoming);

% Display the result
% disp(details)
disp(['p-value: ', num2str(pval)])
disp(['Test statistic: ', num2str(tstat)])

% Create a matrix with pval and z
resultMatrix = [pval, tstat];

% Define the header names
headerNames = {'pval', 'tstat'};

% Create a table from the matrix and header names
resultTable = array2table(resultMatrix, 'VariableNames', headerNames);

% Write the table to a CSV file
writetable(resultTable, 'outputs\bird_circ_cmtest.csv');

outgoing = deg2rad(table2array(readtable("outputs\marion_outgoing.csv")));
incoming = deg2rad(table2array(readtable("outputs\marion_incoming.csv")));

% Perform Circular-Circular Correlation Test
[pval, tstat] = circ_corrcc(outgoing, incoming);

% Display the result
% disp(details)
disp(['p-value: ', num2str(pval)])
disp(['Test statistic: ', num2str(tstat)])
```

```
% Create a matrix with pval and z
resultMatrix = [pval, tstat];

% Define the header names
headerNames = {'pval', 'tstat'};

% Create a table from the matrix and header names
resultTable = array2table(resultMatrix, 'VariableNames', headerNames);

% Write the table to a CSV file
writetable(resultTable, 'outputs\marion_circ_cmtest.csv');
```

```
p-value: 0.61045
Test statistic: 0.030953
p-value: 0.45186
Test statistic: 0.00029711
```