On Kiwi (*Apteryx mantelli*) vocal behaviour and activity: relations to population densities and applications to conservation

A dissertation presented in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

in

Zoology

at Massey University, Manawatū, Aotearoa New Zealand.

Alberto De Rosa

2021

Kia whakarongo ake au
Ki te tangi a te manu nei
a te Kiwi
Tui-i-i, tui-i-i, tuituia
Tuia i runga
Tuia raro
Tuia i waho
Tuia i roto
Tuia i te here tangata

Ka rongo te po
Ka ronga te ao.
Tuia i te muka tangata
I takea mai i Hawaiki-Nui
I Hawaiki-Roa,
I Hawaiki-Pāmamao
Oti rā me ērā atu ano Hawaiki
Te hono a wairua
Whakaputa ki Te Whaiao
Ki Te Ao Mārama.
Tihe mauri ora!
Tui, tui, tuituia

My attention is drawn
to the cry of the bird nearby,
to the Kiwi
Calling “Bind, join, be united as one”
May it be woven above,
Enmeshed below,
Entwined outside.
and within our very beings,
Interlaced by threads of human love

May there be peace at night (in death)
and peace by day. (in our lives)
Intertwined with the cords of humankind
Originating from the great homeland
From the far homeland,
From the remote homeland
And from all other ancestral lands
Merging with the spirits there
Then coming out into the Light
Out into the World of Consciousness.
The living spirit is within us!

Summary

According to the International Union for the Conservation of Nature (IUCN), over 38,500 species of living organisms assessed (27.8%) are currently threatened with extinction. Reducing this startling percentage requires cost–effective monitoring of populations of many and varied species.

Information regarding population trends is crucial to allow decision makers to judiciously allocate unavoidably limited resources.

Acoustic monitoring has long been employed to document the presence and estimate populations of vocal species for conservation purposes. Determining populations trends without the need of sighting or capturing animals can drastically reduce costs and improve welfare. However, as with many other indirect monitoring practices, acoustic surveys impose a series of assumptions about the detectability of the observed animals and their vocal behaviour. Whereas the variability in detection distances and other observer–induced effects can be minimised using acoustic recorders, enabling the delivery of animal abundances using acoustic monitoring requires detailed knowledge of the target species' behaviours to relate numbers of detected acoustic cues to those of animals in an area.

The iconic North Island Brown Kiwi (*Apteryx mantelli*, Bartlett 1851) is a flightless nocturnal bird species endemic to Aotearoa New Zealand, fragmentedly distributed across its mainland range and some of its offshore islands. North Island Brown Kiwi are known for their characteristic vocalisations which differ between sexes, with males emitting series of whistle-like syllables, and females producing series of hoarser and lower frequency syllables. Indeed, acoustic surveys are routinely employed by conservation groups and the Department of Conservation Te Papa Atawhai to monitor North Island Brown Kiwi. These surveys, known as Kiwi Call Counts, require observers to annotate sex, direction of arrival, and distance of the detected Kiwi vocalisations over a set period of time. However, little is known of North Island Brown Kiwi vocal behaviour and how this may relate to animal abundance and the development of more accurate and objective monitoring practices is included among the objectives of the Kiwi Recovery Plan (Germano et al., 2018).

This thesis aimed to investigate North Island Brown Kiwi vocal behaviour and activity to build more objective and accurate acoustic monitoring protocols. Firstly, results from an extensive literature review on the acoustic playback technique — which has been shown to have the potential to enhance acoustic surveys in other species — led to the development of a set of recommendations to enable reproducibility when using playback. Secondly, results from playback experiments showed how single microphone acoustic recording units (ARUs) can be used to localise sound sources with reasonable degrees of uncertainty. This enables the potential transition of Kiwi Call Counts from relying on human observers to ARUs, which would allow for objective interpretation of the data while creating a potentially perpetual record.

One of the thesis aims was to ascertain the potential of using playback to standardise the response of Kiwi populations. The results of experiments testing the effect of playback and environmental factors on kiwi vocal response show that there is no real relationship between the vocal activity of the target Kiwi community and playback. However, they corroborate and add to existing knowledge of Kiwi vocal behaviour by identifying relationships between the latter and external factors, such as lunar illumination and weather conditions. This thesis finally concentrated on the issue of relating vocal activity to animal abundance by developing and trialling the use of animal-borne acoustic recorders in conjunction to fixed ARUs. Since using animal-borne acoustic recorders entails handling target animals, we first performed an experiment on post–handling vocal behaviour to ascertain whether the vocal activity of handled birds of our target community differed from that of birds that had never been handled. The results from this experiment showed that the vocal activity recorded from a gully inhabited by never handled Kiwi did not differ from that of a gully inhabited by birds that were handled during the survey — and have been regularly handled over the last 17 years — in any detectable way. This is encouraging both for animal welfare purposes, and for comparing acoustic surveys from both managed and more wild Kiwi populations.

Finally, the results from employing the animal-borne acoustic recorders to inform density estimates showed how information about individual vocal activity informs more realistic and consistent population estimates than methods based only on community–level vocalisations. On all the occasions sampled, results of population estimates only accounting for environmentally recorded vocalisations delivered lower abundance expectations for both males and females. Repeated sampling results show how estimates that account for individual vocal activity are both more consistent and closer to real densities than traditional methods, as estimated by paired sampling with a specialised dog survey. Lastly, information from individual vocal activity in some populations informed more accurate estimates for other populations without individually tagged animals. Taking advantage of having multiple populations with tagged individuals, we estimated abundances of a target population with three different models: unmarked, tagged with animal-borne acoustic recorders, and with information from other populations' tagged individuals. This last estimate was in between the unmarked and with animal-borne acoustic recorders and apparently more accurate than the unmarked model.

This thesis provides methods and shows encouraging results to eventually employ passive acoustic monitoring to infer Kiwi abundance in a cost-effective and non–invasive fashion at large scale, and invites further employment of animal-borne acoustic recorders to confidently deliver abundance estimates, crucial information for conservation decision makers. Using animal-borne acoustic recorders and ARUs together as a way to estimate populations does involve some invasive trials, but has the potential to lead to fully non–invasive robust abundance estimates though passive acoustic monitoring.

Acknowledgements

Sincere thanks to my supervisors Isabel Castro (Prof.) and Stephen Marsland (Prof.) for the immense support during this project, and Weihong Ji (A/P) for always being ready to get involved if need be. I am grateful for the opportunity to work with you and for entrusting me with being your student. There were challenges along the road, but it is thanks to your advice, support, and encouragement, that I am now writing this page. I want to express my gratitude to Te Pūnaha Matatini for sponsoring my doctoral fees and salary, to KIWIS for KIWI for supporting my fieldwork expenes, and the Ornithological Society of New Zealand (Birds NZ) for awarding me the David Medway Scholarship. I would like to also thank Manaaki Whenua Landcare Research Kiwi Rescue MBIE Programme (C09X1609) as well as the Royal Society of New Zealand Marsden Fund (17-MAU-154). I want to thank the Massey University School of Agriculture and Environment technical and administrative staff, in last–name alphabetical order Jane Algar, Paul Barrett, Sharon Bishop, Sandra Dunkinson, Tracy Harris, Shaun Nielsen, Tara Penketh, Cleland Wallace. I would like to express my most sincere thanks to David and all the Chamberlin family for both allowing me to conduct my research on their land, and treating me like family. I would like to thank Ngai Tai for supporting my work with their taonga on Ponui Island. Thank you to Ngati Kuta and Te Patukeha hapū for their support of the greater Kiwi research project, and mine in particular in the Bay of Islands, and Ngati Rehia and Ngati Torrehina and The Landing for allowing us to work with Kiwi at Purerua Peninsula.

I want to wholeheartedly thank Daryl Olsen for the extremely pleasant collaboration throughout the research project and I am extremely grateful to all those who helped and supported me in my research; specifically and in last–name alphabetical order: Tim Arnold, BJ Black, Natasha Bansal, Phil Battley, Richard Brown, Emma Feenstra, Dane Hawker, Ange Hura, Te Taonga Hook, Julius Juodakis, Janneke Klute, Virginia Listanti, Yi Luo, Maaka McKinney, Lucas Mugnier, Monika Nowicki, Raffaele Paduano, Nirosha Priyadarshani, Alvin Rewha, Tim Robinson, Kathryn Strang, Emily Tamihana, Malin Undin, Alex White, Andre Witehira, Richard "Blandy" Witehira, Shogo Yarita, and Yuki Yarita.

I want to thank my family for the continuous support, my partner Cristabel, my parents Loredana and Marino, my siblings Nicoletta and Giuseppe, and all my aunties, uncles, and cousins. I want to thank all my friends from Massey University, especially Jorge Bresciano, Lucy Cramps, Evans Effah, Michelle Guerrero, Tyler Jarman, Leonardo Meza Joya, Kyaw Min Tun, and Yvonne Walker, whom I have shared an office with during my time at Massey University. I am very grateful to all my friends in Aotearoa New Zealand and overseas, whom I am not going to name individually because of their overwhelming number. I finally want to thank the mentors that have helped me throughout the years leading to this doctoral journey, in chronological order Gianfranco Valori, Francesco Zaccanti, Massimo Dell'Orso, Maria Vallisneri, and Zachariah Gompert.

Contents

List of Figures

List of Tables

Declaration

This thesis is based on publications, and the formatting style of each chapter follows the guidelines for the journal to which it will be submitted or in which it has been published or accepted for publication. Hence, there are inconsistencies in writing style and some repetitions in the chapters' methods sections presented in this thesis. Contributions of authors are specified in each chapter when necessary.

Chapter 1

General Introduction

Figure 1.1: Clea the Kiwi gets her transmitter changed, picture by Yuki Yarita

1.1 What is Passive Acoustic Monitoring and Why Endeavour To Improve It

Monitoring animal population trends is a crucial practice in informing conservation decisions (Temple and Wiens, 1989). Since the resources available for conservation are inherently limited, it is important to optimise their use in order to maximise the effectiveness of conservation actions, hence to employ and develop monitoring practices that ensure objectivity and reliability of the information relayed.

Although the following sections explicitly focus on birds (class *Aves*) to more clearly define the context of this thesis, similar points can be made in regard to almost any terrestrial sound–producing taxon.

1.1.1 Monitoring of Avian Populations: Before Acoustic Monitoring

There are many different methods that are available to researchers to infer animal abundance, and it is desirable to choose those that appropriately trade off accuracy against efficiency for a given task. We are here referring to absolute abundance, as although relative abundance may have historically enabled researchers to cover larger areas, it is only suited to limited purposes, and inferences of actual populations are preferred (Kendeigh, 1944, scrupulously addresses the advantages and disadvantages of both estimates). Emlen (1971) lists seven general methods to estimate bird abundance, together with a set of variables that generally affect the observations and special sources of error that separate the then–available sampling methods from actual population density estimates of non-flocking terrestrial birds (see Emlen, 1971, Figure 1). Since then, we have gone through quite a leap in available technologies (e.g. portable recording devices, remote sensors of many sorts such as camera traps and autonomous acoustic recording units) and techniques (e.g. statistical methods such as distance sampling and spatially explicit capture recapture, see section 1.1.5), although most field sampling methods (van Heezik and Seddon, 2017; Krebs, 1998) can relate to those clearly described by Emlen (1971), which were:

1. MARK & RETALLY FRACTION
2. FIND & COUNT ALL NESTS
3. MAP & COUNT SINGING MALES
4. COUNT IN A NARROW STRIP
5. COUNT TO FLUSHING DISTANCE
6. COUNT X EFFECTIVITY

7. COUNT X DETECTABILITY

In the first method (1), after capturing a number of individuals in an area, the ratio of marked to total individuals captured or observed at repeated tallies would be representative of the population size (Petersen, 1896). Methods (2) to (5) heavily depend on thorough searches and whereas the author indicates that in (2) nests can be replaced by actual birds, this method entails great effort and time investment. In method (3), the locations of singing males are plotted on maps to approximate individual territories, unfortunately missing silent animals and sex ratio. Method (4) attempts to give detailed account of all present birds by taking the number of animals detected in a sample strip as approximate for a greater area, assuming similar distribution of animals. Method (5) involves annotating all the animals one would encounter and their distance from the observer while walking a transect line and, likewise the last two methods, is reported to be possibly biased by missed detections, misjudging of distance, and unreliable conversion factors from detected birds to population size. The last two methods introduce coefficients to correct for effort (6) and probability of detection (7). Specifically, in (6) the additional piece of information regards the degree to which the observer surveyed each section of the transect, whilst (7) introduces the idea of a species–specific multiplier to relate the number of detected individuals to that of the actual birds in the area (Emlen, 1971).

1.1.2 Towards Reliable Acoustic Monitoring

Several studies strive to improve avian population sampling techniques, reporting advantages and shortcomings of aural surveys. For instance, Gates and Smith (1972) describe how using appropiate analytical techniques helps correct for the high variability in the vocal behaviour of Mourning Doves (*Zenaida macroura*, Linnaeus) in acoustic surveys. The authors however report individual differences among the obsevers, including hearing acuity and experience, as possible inherent sources of bias.

Wilson and Bart (1985), and Gibbs and Wenny (1993) introduce the problem of individual variability in avian vocal behaviour, which could possibly be linked to population density. Although this was not the case for Wilson and Bart (1985), Gibbs and Wenny (1993) reported that paired males of some species may be under–represented in aural surveys if unpaired males were more vocal, suggesting that reliable abundance estimates based on vocal cues would ideally need to be able to account for variability among individuals, whether based on actual individuality, age groups, pairing status, or other groupings. It is important to highlight how, without sensible multiplying factors to relate audible cues to animal abundance, acoustic monitoring could not relay the latter (Thomas and Marques, 2012). Although acoustic monitoring can always indicate occupancy, accurate abundance estimates are crucial for conservation purposes (e.g. detailed population trends, translocations).

Over the past few decades, several studies address how acoustic monitoring fares against more traditional methods. Evans et al. (2007) examine the viability of using playback to standardise call counts of red grouse (*Lagopus lagopus scoticus*, Latham), comparing the latter to the customary censuses, which employed specially trained dogs to find the animals. Their results indicate that acoustic surveys provide similar results to the dog–reliant census and represent a valid alternative method when necessary.

Other studies, such as Simons et al. (2007) and Alldredge et al. (2007) explore the reliability of information from observers, finding that there is variability among observers in terms of both number of birds counted and distances estimated. Moreover, other studies find that observers with different levels of experience tend to report varying numbers of birds, with less experienced observers more likely to report false positives of common species, and more experienced ones more likely to report false positives of rare species (Farmer et al., 2012). One of the proposed solutions is to perform simulated experiments to derive correcting factors for actual estimates from observers with different skills levels (Alldredge et al., 2008).

1.1.3 Passive Acoustic Monitoring

Passive Acoustic Monitoring (PAM Crunchant et al., 2021), which consists of deploying acoustic recording units (ARUs) in an environment, and drawing conclusion from the collected sounds, has long been employed in marine systems (Mellinger et al., 2007) and more recently gained popularity with terrestrial systems (Sugai et al., 2019). This can be because PAM is relatively inexpensive (Williams et al., 2018), constitutionally non–invasive (Curzer et al., 2013), and generates a potentially permanent record of data that can be reanalysed to reduce observer bias (Castro et al., 2019). In avian systems, there is growing interest in density and abundance estimates from PAM data (Pérez-Granados and Traba, 2021). However, in order to obtain estimates of animal density rather than call density, it is imperative to have detailed knowledge of individual and collective vocal behaviour of the target system (Marques et al., 2013), and whether these can be standardised with the use of external stimuli, such as playback (e.g. Okahisa et al., 2016; Perkins et al., 2017).

With rapidly evolving technologies and reduction in cost, PAM promises to be a cheap and widely applicable alternative to more expensive, invasive, and time–consuming monitoring practices.

1.1.4 ARUs as alternatives to human observers

Thanks to improvements in techology, self–powered remotely deployable acoustic recording units (ARUs) have made their appearance as powerful sources of potentially unbiased data. Researchers have thus endeavoured to evaluate the available hardware (Brandes, 2008; Pérez-Granados et al., 2019a; Venier et al., 2012), how these fared in different environments (Celis-Murillo et al., 2012), and their suitability for PAM (e.g. Cragg et al., 2015). Applications experimenting with social media citizen science involvement have also been considered (Saito et al., 2015).

Many researchers addressed the question of whether ARUs could replace human observers as sources of information for acoustic monitoring purposes, frequently to estimate species richness (Alquezar and Machado, 2015; Bateman and Uzal, 2021; Haselmayer and Quinn, 2000), but also directly comparing estimates of abundance from point counts (Bombaci and Pejchar, 2019; Sedláček et al., 2015; Van Wilgenburg et al., 2017) and, more fundamentally, detection probability (Castro et al., 2019; Yip et al., 2017).

Overall, researchers report that ARUs are an effective alternative to human observers, although they are subject to different sorts of bias (Castro et al., 2019; Digby et al., 2013), and some have proposed ways to relate the results from ARUs to those of human observers. Among these, because ARUs generally have smaller detection ranges than human observers, Darras et al. (2018a) show how standardising the detection ranges between ARUs and humans can lead to statistically indistinguishable estimates. Another solution, from the work of Van Wilgenburg et al. (2017), is to employ paired sampling, that is the concurrent performing of human–reliant point counts with the deployment of ARUs, to derive coefficients to correct for the difference in detection range.

One of the main reported shortcomings of ARUs is that they cannot be used to estimate distances to the detected birds, which are sometimes necessary to compute density, hence abundance, estimates.

Microarrays of Microphones for Sound Localisation

Relatively recently, miniature arrays of microphones have been proposed as a possible solution to sound localisation. The basic idea behind this solution is that sounds being detected by closely positioned microphones would arrive sooner to the microphone closest to their source location, and the time difference of arrival among microphones can be computed to accurately estimate their direction of arrival (Spiesberger and Fristrup, 1990). In practice, applications with terrestrial animals have developed quite rapidly over the last decade, with studies reporting proof of concept localisations little less than ten years ago (Cai et al., 2013; Stepanian et al., 2016). Applications with multiple microarrys for sound localisation quickly followed (Suzuki et al., 2017). These include wider field applications (Stevenson et al., 2015; Wijers et al., 2021) with the possibility of actually following individuals through the localisation of their vocalisations (Wijers et al., 2020).

1.1.5 Estimating abundance from PAM data in terrestrial animals

There are several statistical methods that can be used to estimate abundance from passively collected acoustic data, although they usually depend on the characteristics of the target study system. A recent review by Pérez-Granados and Traba (2021) evaluated articles estimating abundance from PAM data and provided valuable recommendations for future improvements. For animals that are acoustically individually identifiable, spatially explicit capture recapture methods (SECR — Efford et al., 2009; Royle et al., 2013) provide a solid framework for density estimates. Using the capture history, that is the history of detections at different ARUs of each indentified animal, SECR provides accurate estimates of the number animals in an area. Alternatively, if the acoustic recording system allows for localisation of vocalising animals, distance sampling (Buckland et al., 2012) takes advantage of spatially locating vocalising animals assuming that vocalisations coming from different locations correspond to different animals. Moreover, specialised versions of SECR (Stevenson et al., 2015) can integrate sound localisation to improve precision (Juodakis et al., 2021a).

Without spatial localisation or individual identification, and assuming that vocal activity is dependent on density, approaches include relating the detected vocal activity rate (DVAR, Pérez-Granados et al., 2019c) to animal density, possibly using paired sampling (Van Wilgenburg et al., 2017). Finally, if information of individual vocal activity, or cue rate, of the focal species is available, this can be used to estimate abundance in the proximity of single ARUs (Sebastián-González et al., 2018).

1.2 North Island Brown Kiwi

North Island Brown Kiwi (NIBK, *Apteryx mantelli*, Bartlett, 1851) are flightless nocturnal cryptic birds belonging to the only Apterigiformes genus, *Apteryx*. NIBK are considered to be a somewhat territorial and vocal species (McLennan et al., 1987; Taborsky and Taborsky, 1992, 1995, 1999), but other studies have shown that territoriality may be dependent on population density (Dixon, 2015; Potter, 1989).

In the red list of threatened species redacted by the International Union for the Conservation of Nature (IUCN), NIBK are classified as *Vulnerable*, with a population between 10,000 and 19,999 which is estimated to have undergone an important reduction (30–49%) over the last three generations (26 years, BirdLife International, 2017). According to the New Zealand Department of Conservation Te Papa Atawhai, in 2018 NIBK tallied up to 25,100 individuals and this species is considered *At Risk - Declining* (Germano et al., 2018, Tables 2 and 3). NIBK are endemic to Aotearoa New Zealand and have a fragmented distribution across the North Island of Aotearoa New Zealand; NIBK are also found in some offshore islands (Colbourne, 2005). NIBK are classified in four recognised management units (taxa, Germano et al., 2018; Undin et al., 2021).

Vocalisations of male and female NIBK are quite distinct; Corfield et al. (2008) describes males as producing long–travelling whistles which could be grouped into phrases, and with fundamental frequencies around $1.5kHz$ (and overtones up to approximately $13kHz$), and female as producing hoarser and lower vocalisations with intensity concentrated in two formants (see Figure 1.2); duets have been suggested to possibly have territorial implications (Corfield et al., 2008).

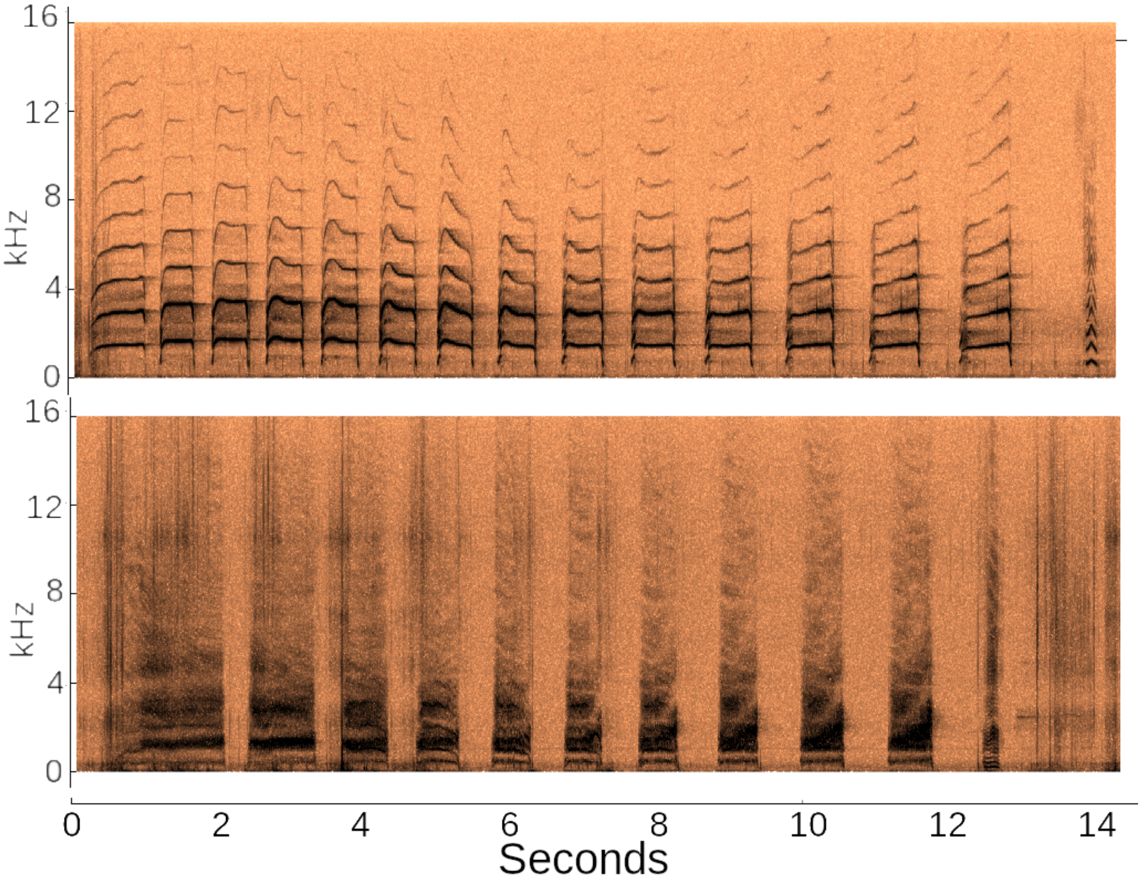


Figure 1.2: Example spectrograms of male (top) and female (bottom) NIBK.

In the current Kiwi Recovery Plan, the New Zealand Department Of Conservation Te Papa Atawhai identifies inconsistencies between NIBK population densities and the current monitoring method employed for surveying populations (issue 4.6 Germano et al., 2018), as well as the need for implementing a long–term monitoring programme to evaluate the changes in NIBK population abundances (objective 4.1, action 4.2 Germano et al., 2018). The already in–place acoustic monitoring framework known as “Kiwi Call Counts” (KCC, Craig, 2018; Robertson et al., 2017) requires observers to listen for NIBK vocalisations for two hours over 4 to 8 nights (depending on the number of listening stations per site, totalling 48-50 hours) and annotate each vocalisation detected, the sex of the animal which emitted it, and its estimated distance and direction of arrival.

For this thesis, which aims to enable more objective and cost–effective population estimates, NIBK represents an ideal study system to employ PAM to identify a reliable way of estimating abundance with the genuine potential for real world applications for three simple reasons:

- NIBK are cryptic and nocturnal, yet vocal, hence acoustic monitoring is a good fit and PAM provides a more cost-effcient way to survey vocal animals than human observers (Williams et al., 2018).
- The current monitoring scheme, based on specialised acoustic surveys, needs more reliable and efficient developments to translate vocal activity into abundance estimates (Germano et al., 2018).
- Other *Apteryx* species have been found to either respond well to playback (Digby et al., 2013), or are individually recognisable from their vocalisations (Bedoya and Molles, 2021), aspects that would very well work with robust statistial estimates of abundance (Efford et al., 2009; Royle et al., 2013).

1.2.1 NIBK on Ponui Island

The majority of birds considered in this thesis reside on Ponui Island (Figure 1.3), where the Kiwi research team led by Isabel Castro has worked since 2004. Fourteen NIBK were released on Ponui in 1964 (Miles and Castro, 2000),

Ponui Island nowadays hosts a high-density NIBK population (Colbourne, 2005; Cunningham et al., 2007; Ziesemann, 2011). Over the last 17 years, between 30 and 50 birds have been equipped with very high frequency (VHF) radio transmitters that are monitored at least monthly. Having record of the history of these animals, including morphological measurements, usual resting areas, nesting sites, and social relationships (Cunningham and Castro, 2011; Dixon, 2015; Undin et al., 2021), makes Ponui Island an ideal site to conduct this research. Previous studies, including MSc and doctoral theses (e.g. Bansal, 2020; Dixon, 2015; Vattiato, 2021; Vieco Gálvez, 2019; Ziesemann, 2011), provide essential information on the animals' abundance and habits, that has been crucial in designing all of the experiments herein described. For instance, the works of Ziesemann (2011) and Dixon (2015) include essential information on the animals' home ranges and usual roosting sites, representing for us realistic expectations on the number of birds potentially in range of our playback experiment in chapter 4.

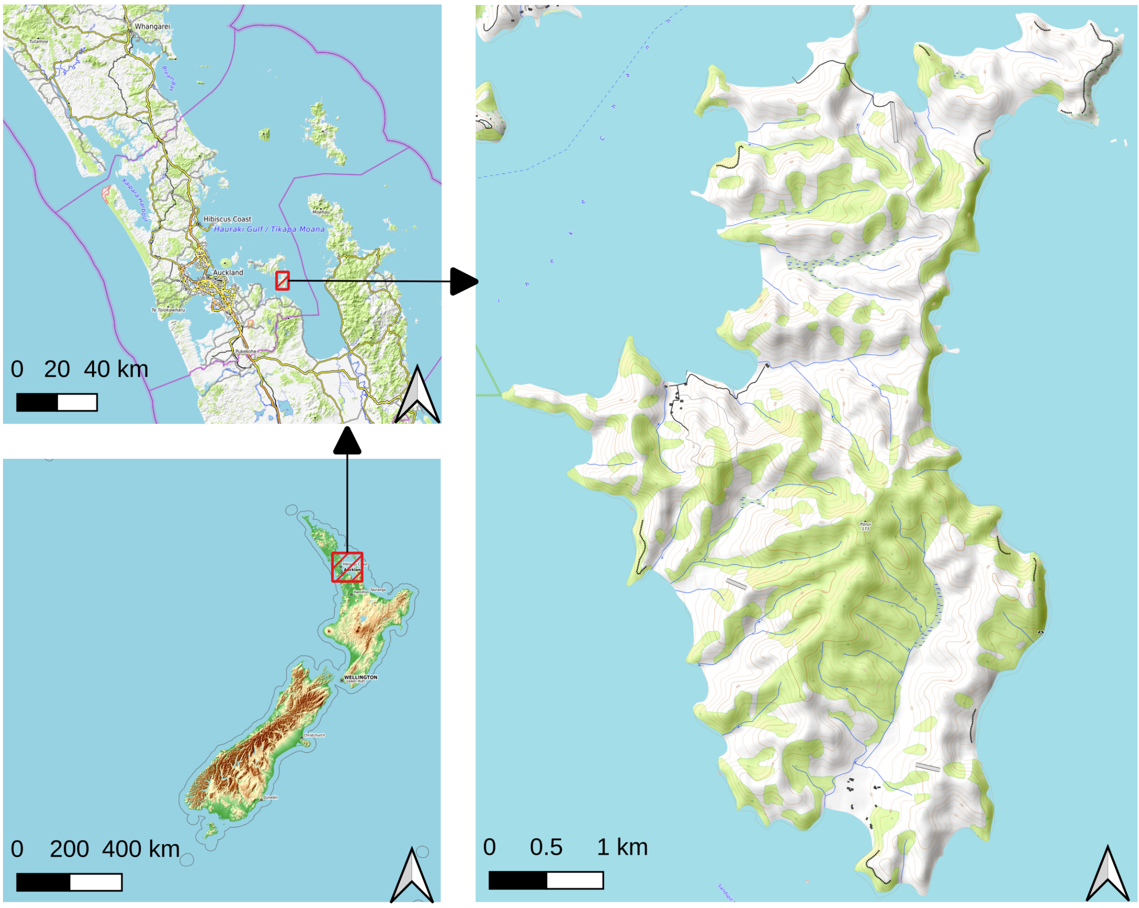


Figure 1.3: Ponui (Chamberlin's) Island (175.14860,−36.90422 : 175.21860,-36.82961 – Coordinate Reference System WGS84) is a $\sim 18Km^2$ privately owned island located in the Hauraki Gulf, Auckland region, Aotearoa New Zealand. Ponui Island is divided in three properties. In this study we worked within the South Ponui Island area, where David Chamberlin and his family have a farm. Ponui Island has been inhabited for several centuries, with evidence of Māori settlements from the XIV century. The island is mostly covered by pasture but retains primary and secondary native forest patches, thus providing a thriving environment for NIBK (see Atkinson, 1959; Cameron and de Lange, 2006; Hynes, 1955; Irwin, 2020, for details on Ponui Island's history and vegetation).

On Ponui Island, marked birds hold overlapping home ranges of between five and six hectares (Ziesemann, 2011, Chapter 3). Birds roost in burrows or on the surface in forest patches within their range and the choice of roost site is dependent on proximity to foraging areas (Dixon, 2015, Chapter 4). At night NIBK move from the roost sites to foraging sites and are mostly found in forested and scrub areas (Dixon, 2015, Chapter 4, figure 3.6). In the work of Dixon (2015), over 2377 nocturnal localisations of Ponui NIBK, animals displayed variability in habitat utilisation on the temporal, breeding and social status scales (e.g. birds that tended to shelter in pairs were found more often in forest environments than those sheltering alone or in groups).

1.3 Aims

NIBK are a vulnerable endemic taxon with fragmented populations (BirdLife International, 2017), in need of more robust and efficient long–term monitoring solutions (Germano et al., 2018). Acoustic surveys, either reliant on observers or automated (Castro et al., 2019; Yip et al., 2017), have the potential to provide crucial insight into population trends in non–invasive and cost–efficient fashion (Pérez-Granados and Traba, 2021). However, in order to provide realistic abundance estimates, we need a way to translate number of vocalisations into number of animals (Marques et al., 2013). This could be done by:

- Localising vocal animals, assuming that vocalisations from different locations are from different animals and that all animals vocalise (Buckland et al., 2012; Stevenson et al., 2015).
- Recognising individuals based on acoustic properties (Bedoya and Molles, 2021; Efford et al., 2009).
- Estimating how a detected vocal activity rate relates to abundance with paired (concurrent) sampling (Van Wilgenburg et al., 2017).
- deriving an individual cue rate (Sebastián-González et al., 2018).
  All of these methods could also benefit from soliciting vocalisations using acoustic playback as a mean to standardise vocal response (Okahisa et al., 2016). This thesis aims to contribute to the development of acoustic–based abundance estimates with the following specific aims:

1. **Identify an inexpensive way to localise birds to potentially employ PAM-KCC**. Replacing human observers with ARUs can reduce costs towards more efficient KCCs. However, KCC require the observers to include the direction of arrival, estimated distance, and sex of each vocalising animal and, although this information is not currently being used, it is being collected and, to cut costs, replacing human observers with remotely deployed ARUs would require a way to include such information while only employing ARUs.
2. **Gauge the effectiveness of acoustic playback as a means to enhance acoustic surveys in NIBK**. Acoustic Playback has been reported as a tool for eliciting vocalisations from NIBK that would have possibly not vocalised otherwise (McLennan, 1992), but is not employed in KCC during the National KCC scheme (see page 65). Nevertheless, a formal test of NIBK response to acoustic playback has not to date been reported, and this technique has been shown to improve acoustic monitoring of other avian species (e.g. Hart, 2020; Okahisa et al., 2016).
3. **Investigate potential post–handling changes in the vocal behaviour NIBK**. One way of both obtaining an indication of individual vocalisation rate, and provide data for

investigating individual identification is to employ animal–borne acoustic recorders (Stowell et al., 2017). Since using these devices would entail more handling of wild NIBK than usual, understanding whether birds that have been handled differ in vocal behaviour due to, for example, stress, is important both to evaluate potential animal welfare implications of the use of animal borne recorders, and to check whether the data obtained would be reflective of normal individual vocal activity.

4. **Evaluate the use of animal-borne recorders to inform effective NIBK PAM acoustic**. In order to relate the number of vocal cues detected to the number of birds we can use animal borne recorders to acquire individual vocal activity and compare this to concurrently recorded community vocal activity. We can then use the relationship between individual activity and community detected vocalisations for abundance estimates.

1.3.1 Thesis Structure and relationship between chapters and aims

Besides chapters regarding the aforementioned aims, this thesis includes one literature review regarding the use of the acoustic playback technique in ornithological field contexts. This was necessary as a standardised fashion of reporting the use of this technique with wild birds was lacking and, concurrently, necessary for this study. As such, this thesis includes the following chapters:

Chapter 1 This chapter. Provides background information regarding the ideas, context, methods and scopes of this thesis.

Chapter 2 This is a literature review regarding the playback technique used in chapters 3 and 4. This has been accepted by the journal IBIS. Together with chapter 3, it addresses aim #1.

Chapter 3 Describes a playback experiment that represents the methodological backbone of this thesis. It provides experimental evidence for the spatial deployment of the ARUs as well as an inexpensive alternative to localise sound source using sound intensity rather than time of arrival.

Chapter 4 Explores the response of Kiwi to playback of other NIBK vocalisation as a way to standardise acoustic surveys. We conducted two series of playback trials on Ponui Island, one within the breeding season and the other outside it. This chapter addresses aim #2.

Chapter 5 Compares the vocal activity of a community Kiwi from an area not visited by Kiwi practitioneers versus a Kiwi community undergoing transmitter change; this includes researchers actively searching and capturing birds and of course birds that are handled. This chapter addresses aim #3.

Chapter 6 Evaluates the use of miniature animal borne acoustic recording units in conjunction to ARUs to relate the number of animals in an area to the number of vocalisations detected therein. This chapter addresses aim #4 as well as providing data for future individual identification research.

Chapter 7 Summarises the results from all the previous chapters and ties them together. This chapter provides insight for future studies on the subject and gives the next steps in completing the development of ABARs/ARUs method for estimating Kiwi abundance from acoustic surveys.

This thesis is strongly based on fieldwork conducted on Ponui Island and in Northland between 2018 and 2021. In order to optimise time and resources, multiple experiments were conducted in a relatively small timeframe and have at times overlapped. Table 1.1 summarises how this took place and clarifies where and when sampling occurred.

1.4 Ethical issues

In order to carry out the fieldwork to gather the data for this thesis we received approval from the Massey University Animal Ethics Committee `Protocol ID 20/26` and we operated in the field under the Department of Conservation permit `63716-FAU`.

Table 1.1: Summary of field trips related to the chapters of this thesis. The *Location* column refers to broadcasting sites used in chapter 3 (Massey Grounds in Palmerston North and Totara Reserve in the Pohangina Valley), or gullies on Ponui Island (Homestead is Homestead Gully, Kauri is Kauri Bush, RSHG is Red Stony Hill Gully, Pipe is Pipe Gully) and Northland (The Landing). ABAR = Animal–Borne Acoustic Recorder. ARU = Acoustic Recording Unit. NIBK = North Island Brown Kiwi (*Apteryx mantelli*).

Year	Month	Action	Chapters	Location
2018	May	First trip Ponui Island	2-6	Kauri, RSHG, Pipe
	November	Second trip to Ponui Island	3	Kauri, PKG
2019	February	Playback Experiment	3,6	Massey Grounds
	March	Kiwi Practitioneer Training	4-6	Ponui
	October	Playback Experiment	3	Totara Reserve
2020	May-June	ABARs and ARUs deployment on Ponui Island #1	6	Kauri, RSHG
2021	March	ABARs and ARUs deployment on Ponui Island #2	4-6	Kauri, Homestead
		Post-handling data collected	4	Kauri, Homestead
		Playback with NIBK	4,5	Kauri, Homestead
		ABARs and ARUs deployment Northland	6	The Landing
	July	Playback with NIBK	5	Kauri

Chapter 2

The acoustic playback technique in avian fieldwork contexts: a systematic review and recommendations for best practice

Figure 2.1: Isabel in the field with Tane. Picture by Juan Carlos Garcia Ramirez

Abstract

Acoustic playback is commonly used to study wild birds, with applications as diverse as investigating behaviours, ascertaining the presence of rare and elusive species, and attracting individuals to a location. The number of studies employing playback grows larger every year because it is easy to apply, increasingly affordable, and very effective. However, the way that it is used and reported varies significantly across researchers and species. This lack of a protocol for reporting acoustic playbacks inevitably slows the progress of the field, since studies cannot be easily compared.

In 1991, some of the most knowledgeable researchers in the field of animal communication met at a North Atlantic Treaty Organization (NATO) Advanced Research Workshop (ARW) at Thornbridge Hall in the UK to consider the design of playback experiments. The proceedings of the meeting were published in 1992 and contain crucial guidelines regarding this use of playback. In this paper we review the literature for papers published since that milestone that use acoustic playback in ornithological fieldwork contexts. We use the RepOrting standads for Systematic Evidence Syntheses (ROSES) to evaluate the description of the methods used. The main goal of this review paper is to identify a shared set of rules for employing and reporting the playback technique in such contexts to promote reproducibility and comparability. We found 625 peer–reviewed articles in three on–line databases, of which 419 reported field (rather than captive) studies employing avian playback.

The biggest hindrance to reproducibility is the availability of the acoustic tracks used by the authors; four percent (15 articles) of our sample made their tracks publicly available. We found that only one article provided enough details for their playback application to be fully reproducible. Further, only four articles (0.92%) provided enough information for reproducibility, even if we assume information about track preparation and recording details to be unnecessary when tracks are available.

Based on our synthesis of the literature, we provide a set of recommendations for the reporting of playback uses to promote reproducibility, including sample paragraphs of description as supplementary material. We strongly recommend that tracks used for such experiments are deposited in dedicated on-line repositories for the use of other researchers.

While our focus is avian fieldwork applications of the playback technique, we believe that they can be easily transferred to other animal systems subject to acoustic playback.

2.1 Introduction

While crucial experimental design aspects of studies involving acoustic playback have justly received extensive coverage over the years (e.g., pseudoreplication: Hurlbert, 1984; Kroodsma, 1989a,b, 1990, 1992, 2017; McGregor et al., 1992; Searcy, 1989; Weary and Krebs, 1992), the comparability and reproducibility of the reported experiments to our knowledge have not been investigated. However, for scientific experiments it is important to use comparable methods (Barker, 2008) and to report them clearly (McGregor, 2000; Wiley, 2003), enabling fellow researchers to critique the method, reproduce the experiment, or perform their own experiments in comparable fashion.

The primary driver of this review was to ascertain whether it is possible to reproduce the application of the playback technique from avian field context studies relying solely on the published content. Many different animals respond to playback of pre-recorded sounds, for example by vocalising in turn, or approaching the speakers and sometimes interacting with them by closely examining, and even attacking them. Historically, a useful source of suggestions for the design of studies involving acoustic playback is the book resulting from the Thornbridge Hall North Atlantic Treaty Organization (NATO) Advanced Research Workshop (ARW) Consensus, which includes a table listing features considered important in such experiments (McGregor et al., 1992). This publication arose from a meeting in August 1991, when experts in animal communication congregated t odiscuss experimental design and analyses of the rapidly growing number of acoustic playback experiments. The table provides a very useful list of elements to consider regarding different aspects of experimental design using playback, specifically test sounds, environmental conditions, test animals, playback equipment, and procedure. However, as more numerous and advanced recording, broadcasting, and analytical tools have become widely available, increasingly detailed investigations are possible (Freeman et al., 2017).

Whenever playback is used to seek a behavioural response, as opposed to increasing detectability or testing tolerance to noises, the methods used should aim to make the target audience mistake the broadcast stimuli for real animals. Nevertheless, the equipment used for playback is of varying quality, and given that a number of species can discriminate altered sounds (e.g., Luther et al., 2017; Wong and Gall, 2015), accounts of playback procedures should include sufficient detail to allow for reproducibility and assessing of playback quality. Reproducibility, one of the vital elements of the scientific method (Baker, 2016), delivered mainly via the *material and methods* section within peer-reviewed publications, although it can also include on-line access to data or other material.

However, information that allows reproducibility is also important because we do not really know what the birds are hearing and reacting to, and because carrying out the same experiment with different devices (e.g., broadcasting the same stimuli with speakers covering different frequency ranges) may produce different results. While accepting that not all variables are measurable, nor all experimental conditions (e.g., weather) potentially replicable, we should strive to report what is possible to enable readers to paint the most accurate picture of our applications of the technique as possible. There are two main reasons for focusing on avian fieldwork contexts. Firstly, the number of articles published yearly on playback in avian field contexts is increasing (Figure 2.2) and, secondly, the use of playback in many other systems and contexts has been recently covered by extensive reviews. These cover marine mammals (Deecke, 2006; Ichikawa et al., 2011; Janik and Sayigh, 2013; Tyack, 2009; Weilgart, 2007), amphibians (Narins, 2018), bats (Gager, 2019; Jones and Siemers, 2011), rodents (Wöhr, 2018), primates (Fischer et al., 2013), and insects (pest management – Mankin et al., 2012), as well as considering several specific applications such as cognitive research (Gentry et al., 2020), interactive playback (King, 2015), gender identification (Volodin et al., 2015), effectiveness of heterospecific versus conspecific vocalisations in attracting animals (Putman and Blumstein, 2019), duets (Douglas and Mennill, 2010), and eavesdropping (Magrath et al., 2015). Compared to wild environments, studies involving acoustic playback with captive birds have been rightfully reviewed using different criteria (e.g., Beecher et al., 1986; Derégnaucourt, 2011; Derégnaucourt and Bovet, 2016), taking into account how drastically different the conditions under which we can perform acoustic playback in captive conditions.

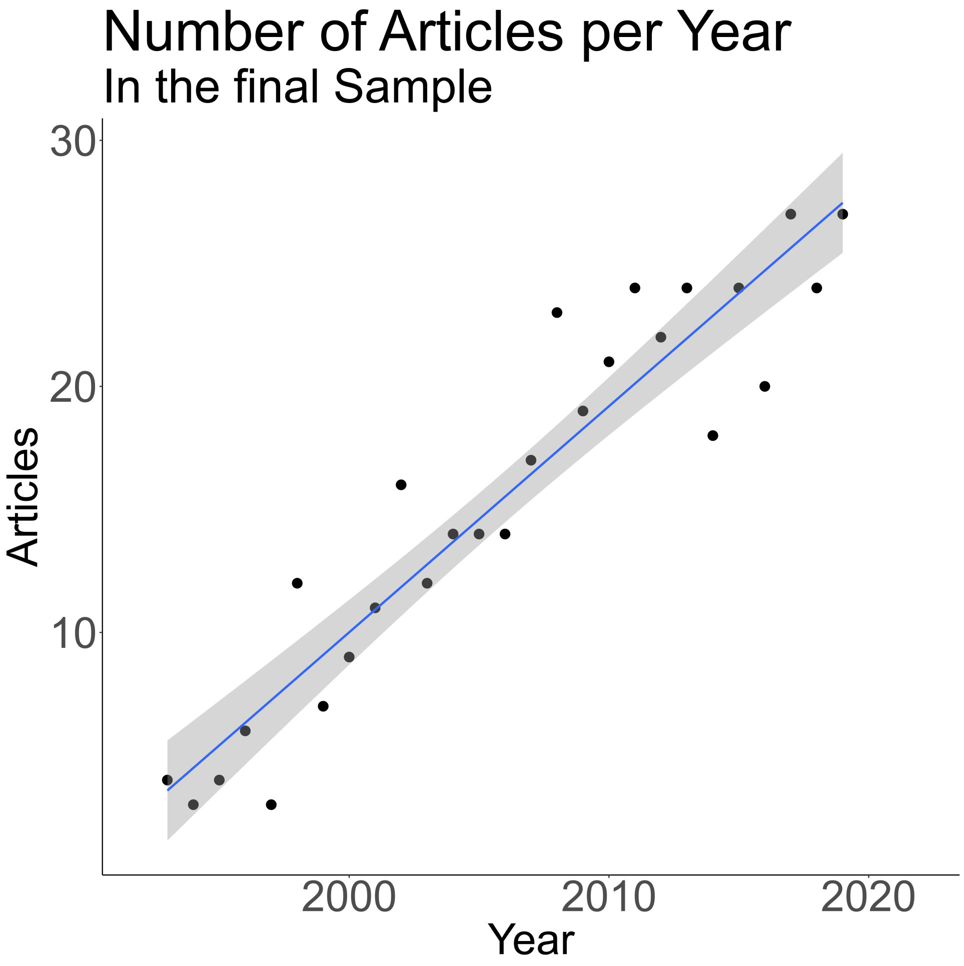


Figure 2.2: Number of articles in our sample by year of publication.

Some examples of the applications of acoustic playback in ornithological field contexts include the detection of elusive species (e.g., Marion et al., 1981; Okahisa et al., 2016), attract birds (e.g., Martin et al., 1995; Podolsky and Kress, 1989), study predator–prey interactions (e.g., Abbey–Lee et al., 2018; Mougeot and Bretagnolle, 2000), territoriality (e.g., Akçay et al., 2011; Leedale et al., 2015), effects of changes in vocalisations' syntax (e.g., Hedley et al., 2017; Taylor et al., 2017), eavesdropping (e.g., Niederhauser et al., 2018; Vargas–Castro et al., 2017), genetic/acoustic discrimination (e.g., Lipshutz et al., 2017; Provost et al., 2018), and neighbour/stranger recognition (e.g., Dadwal and Bhatt, 2017; Moskát et al., 2017).

Avian Perception of Sound, and Relevance to Playback

Objective parameters to allow for reproducibility of playback are needed because human perception of sound differs from that of birds, and this has implications for the recording and broadcasting of avian vocalisations. Sounds emitted from vocalising animals are, more often than not, directional (Patricelli et al., 2008), and are affected by attenuation, reverberations, amplitude fluctuations, and atmospheric effects (Catchpole and Slater, 2003). The combination of these various degradations means that sound changes over distance in a non-linear way. The ability of birds to identify the direction from which a signal arrives can be directly linked to physiological cues (Bradbury and Vehrencamp, 2011; Dent and Dooling, 2003; Larsen et al., 2006; Schnyder et al., 2014). Additionally, birds' ability to estimate the distance from a signaller through degradational cues, or to "range" a sound, has been extensively reported.

Hence, to avert misinterpreting the responses observed when performing playback, it is important to consider how birds localise sound sources, and whether they may identify degradation cues, and potentially behave differently because of these.

Ranging

Richards (1981), Richards reported recording two different vocalisations, respectively ten and fifty metres away from a singing Carolina wren *Thryothorus ludovicianus*. These two vocalisations were then played back to conspecifics at a standardised amplitude. The wrens exposed to this playback responded by counter-singing to the "far" song and by attacking the "near" song. This was, to our knowledge, the first experiment leading to what Morton (1983) formalised as the "Ranging Hypothesis". The hypothesis states that birds assess distances from one another using sound degradation, by comparing newly heard sounds to an internal representation of the undistorted sound. The use of degradation cues to estimate signaller distance has been tested and reported multiple times (Fotheringham et al., 1997; McGregor, 1994; Naguib, 1996, 1998). However, some results were thought to be unreliable, as a bird could potentially move position and triangulate to estimate distance to a loudspeaker if tracks were played for a sufficiently long time (Naguib et al., 2000; Wiley, 1998). Further studies have shown that birds are able to range vocalisations by relying only on degradation cues (Holland et al., 2001; Naguib and Wiley, 2001; Zahorik, 2002).

The detection of sound alteration

Birds can sense differences in amplitude and frequency modulated sounds (Bradbury and Vehrencamp, 2011; Stebbins, 1983). Avian amplitude (Cynx et al., 1990; Dooling and Searcy, 1981) and frequency discrimination thresholds (Dooling et al., 2000; Saunders et al., 1978) have been found to be similar to those of humans.Beason (2004) suggests that birds generally hear best between 1000 and 5000 Hz, with upper and lower auditory limits of 25,000 Hz (25 KHz) and 5 Hz respectively. However, some studies report upper limits closer to those typical of humans (20 kHz, after conditioning), and theoretical ones up to 30 kHz (Dooling, 2004; Li et al., 2011; Martin, 2017; Schwartzkopff, 1955). Some bird vocalisations include high-frequency harmonic components extending into the ultrasonic range (Brand, 1938; Narins et al., 2004; Tanimoto et al., 2017; Thorpe and Griffin, 1962) nd in some species, the energy carried by these high-frequency components is like that of the vocalisations' audible ones (Li et al., 2011). Although high-frequency components could be byproducts of vocalisation, they may instead have a specific role in signalling conspecifics (Li et al., 2011; Thorpe and Griffin, 1962; Wilson and Hare, 2004). It has therefore been suggested that assumed avian auditory system limits may need to be re-evaluated (Li et al., 2011; Narins et al., 2004). Experiments concerning noise (Lohr et al., 2003; Scharf, 1970) and frequency alterations over time (Ratcliffe and Weisman, 1985, 1986) suggest that the way that birds perceive differences in sounds could be quite different from ours. Of course, such perceptions could be specific to genus or species. Lohr et al. (1998), and Dooling and Prior (2017) showed how the ability to discriminate sound alterations in the *fine structure* of the sound (local timing, accounting for spectral and temporal cues within individual harmonic syllables), rather than just in the *temporal envelope* (accounting for the global features of a sound) in some birds is significantly better than in humans.

Reaction to sound alterations

Several studies have addressed the matter of differential responses to altered stimuli. Some species, such as Field Sparrows *Spizella pusilla* and White-crowned Sparrows *Zonotrichia leucophrys* show stronger responses at lowered frequencies (Luther et al., 2017; Nelson, 1989), while other species, such as Northern Cardinals *Cardinalis cardinalis* and Black-capped Chickadees *Poecile atricapillus* respond more vigorously than average to shifted frequency vocalisations (providing that background noise is not too high) or when presented with note-specific alterations (Luther and Magnotti, 2014; Wong and Gall, 2015). Average amplitudes of natural vocalisations at source are not known for most birds, and there is huge individual variation and modulation potential (Naguib, 1995; Tanimoto et al., 2017). Higher broadcasting amplitudes usually elicit altered (often stronger) responses from wild birds (Dabelsteen, 1981; Nelson, 2000; Ritschard et al., 2010). Amplitude differences are thus double-edged swords, as the species' differential response to them could be advantageous when using playback to detect rare or elusive species (Marion et al., 1981; Okahisa et al., 2016)but as a source of bias in behavioural contexts. Finally, changes in the quality of a broadcast vocalisation can have significant impact on the birds' response. Several species respond more aggressively to playbacks of clear versus degraded vocalisations broadcast at the same amplitude (McGregor and Krebs, 1984; McGregor et al., 1983; Richards, 1981; Slabbekoorn, 2004).

In this review we examine a sample set of papers that use the playback technique in avian field studies to investigate how reproducible the studies are. We synthesise our findings to provide a set of features that we deem important to report in published articles involving the playback technique in order to improve reproducibility; these come from both the reviewing of the literature and experience. A consistent approach to performing and reporting experiments employing the playback technique would make it easier to understand, reproduce, and compare the results of studies, thus enhancing advancement of this field. We also provide a series of practical recommendations, including suggestions on how to measure amplitudes more easily in the field.

2.2 Methods

We performed a systematic literature review following the RepOrting standards for Systematic Evidence Syntheses (ROSES) protocol (Haddaway et al., 2018). We performed a literature search using three different on-line literature databases: Scopus (`https://scopus.com`), PubMed Central (`https://www.ncbi.nlm.nih.gov/pmc/`), and Web of Science (`https://webofknowledge.com`) on 18 December 2019. Our search was limited to original research articles (i.e., not reviews) from 1993 onwards (that is, published after the 1992 Thornbridge Hall Nato ARW consensus). We searched using the followin logical combination of keywords: `(*bird* OR avian) AND playback AND (*acoustic* OR sound) AND (natur* OR field)` where asterisks represent the possiblity of any prefix/appendix to be attached to the search word (e.g., *natur** would include *natur*`e` and *natur*`al`).

We only included articles written in English in our sample. We also excluded all the search engine results that did not involve birds (although broadcast of avian vocalisations to non-birds was accepted, as well as broadcast of non-avian sounds to birds), all non-peer-reviewed articles (e.g., grey literature, theses), and articles only involving captive birds. All the articles remaining after these exclusion criteria were applied are included in our sample; a listing of the whole sample is available as supplementary material.

Focusing exclusively on the technical aspects of employing the acoustic playback technique, we prepared a checklist to evaluate the extent to which the reporting in each paper was sufficient to reproduce it (Table 2.1). This list included all the information we expected to see in the *Methods* section and included both information that is important to reproduce the research, and to interpret the results.

Table 2.1: Scoring checklist, apart for the Context variable, which is point based (1 to 8), each question placed articles into categories. Whereas the first question is point based (the higher, the more detailed), all the others assign papers to categories (numberic values are not important).

Feature		Points
Context		1 to 8
Location		
	w/co–ordinates	2
	w/o co–ordinates	1
Vegetation		
	Nothing	0
	Description	1
	References	2
Time		
	of Year	1
	of Day	1
Extra	(Fauna/Geology/Weather)	
	Nothing	0
	Description	1
	References	2
Feature		Category
Tracks Origin		
	Not Specified	1
	Recorded by the authors	2
	Obtained from a third party	3
	Some recorded by the authors and some obtained from a third party	4
Recording devices		
	Not Specified	1
	Brand & Model	2
	Additional information	3
	Frequency Response Charts	4
	Not Applicable	5
Tracks' Availability		
	Not Available	1
	Some Available	2
	Available	3
Recordings' Details	(Frequency range / bit depth)	
	Not Applicable	1
	Not Specified	2
	Present	3
From tracks to stimuli		
	No Editing	1
	Outlined but not reproducible	2
	Reproducible	3
	Not Specified	4
Playback Procedure		
	Not Specified	1
	Outlined but not reproducible	2
	Reproducible	3
Obsevers' Behaviour		
	Not Specified	1
	Outlined but not reproducible	2
	Reproducible	3
	Not Applicable	4
Broadcasting SPL		
	Not Specified	1
	Not Shown, Explained	2
	Shown, Not Explained	3
	Shown, Explained	4
	Shown, Referenced	5
SPL Measurement		
	Not Specified	1
	Not Measured	2
	Measured, Not Described	3
	Described, Not Reproducible	4
	Reproducible	5
Broadcasting devices		
	Not Specified	1
	Brand & Model	2
	Additional information	3
	Frequency Response Charts	4

The first criterion on the checklist considered how comprehensively the authors described the context in which they employed the playback technique by including relevant geographic, climatic, and habitat information. Although not as objective as the other criteria, we first considered the amount of information regarding the playback context reported by articles in our sample. We used a point-based system according to which each article could potentially receive up to eight points, whereas all the remaining criteria assigned each article to a category. The remaining criteria ascribed articles to different categories according to the origin of their audio tracks (and details of the recording devices if not acquired from a third party), the availability of such tracks, details about their properties (frequency range and bit depth), whether and how these were processed, the actual playback procedure, observers' behaviour, and broadcasting details (including sound pressure level, see details in Table 2.1.

The standard screening strategy was to first assess eligibility as follows: read title and abstract, if not conclusive read methods section, if not conclusive read entire article and supplementary material. Subsequently, we read each eligible article and relative supplementary material to assign it to the different categories, according to the checklist.

For consistency, one person (ADR) performed the bulk of the evaluation. ADR also performed a second evaluation one year after the first, and blind to the first result: the second evaluation matched 96.7% of the scores (see supplementary material). We further verified the evaluation by asking five additional people (see Acknowledgements), blind to the main reviewer's scoring, to score three randomly selected unique articles already scored by ADR, plus two shared among all, for a total of 17 (4% of the sample) unique articles.

The ROSES protocol also includes procedures for obtaining missing data. However, we did not use these procedures because we consider them unnecessary; the main point of this review was to identify the degree to which we would be able to reproduce applications of the acoustic playback technique in avian field contexts wholly from the published material (including supplementary material).

2.3 Results

2.3.1 Analysis of Reporting in the Literature

The different on-line databases returned very different numbers of articles (Table I.i). The total number of articles returned by the various searches was 1087, of which 625 were individual articles and the rest repetitions between the databases; we found 601 articles through Web of Science (of which we included 406 in this study, 294 were only found through Web of Science, see figure I.i), 177 through Scopus (115 included, 9 were only found through Scopus), and 73 through PubMed Central (38 included, 4 unique in PubMed Central, see Figure S1). We included about two-thirds (419) of the total 625 unique articles resulting from the search in the sample. Of the 206 that we excluded, 62 were not about birds, 90 involved the use of captive birds in aviaries or laboratories, three were not classifiable as "articles" (e.g., books, technical reports), 27 did not employ playback, 21 were not written in English, and three were duplicates. The number of articles per year in the final sample of evaluated papers showed an increase from as few as three in the 1990s to high twenties from 2008 onwards (Figure 2.2). We next discuss the distribution of articles according to each checklist criterion (Figure 2.3).

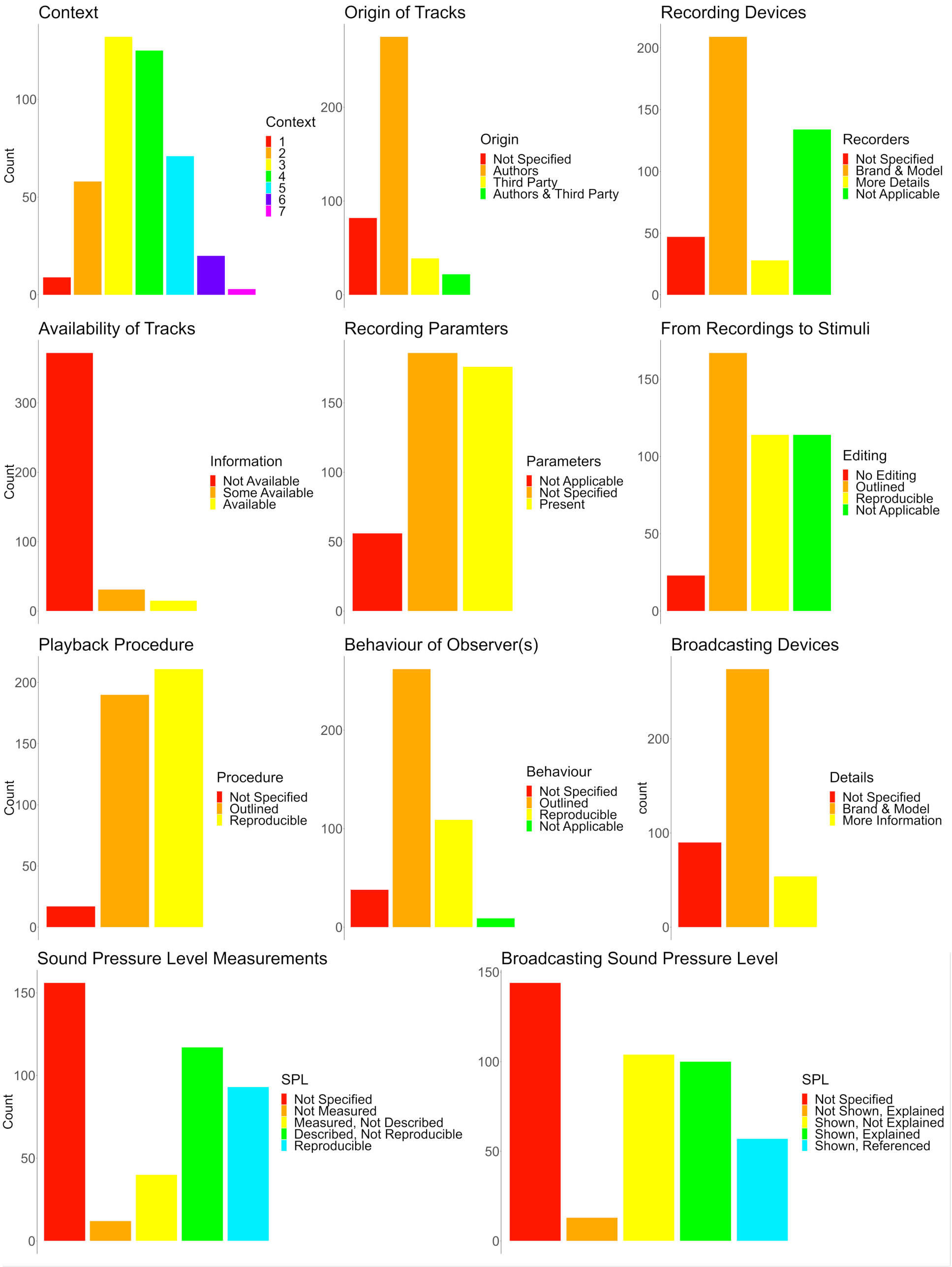


Figure 2.3: Proportions of the sample articles resulting from our literature search according to our review questions. The colours respectively represent the numbers of articles belonging to each category as per individual legends.

Playback procedure, context, and observers' behaviour

Half (52%) of the articles scored 4 or more on the context criterion (the maximum score was 8), which usually meant that coordinates for the study site, date and time of broadcast were provided, but see table 2.1 for detail. Overall, the most represented value (median) was 4, with a mean of 3.60 and a standard deviation of 1.18. Only 4% (17) did not specify any detail about the playback procedure (stimuli order and playback duration), whilst 45% and 50% provided some (190) or detailed (211) information. Nine percent (38) of the sample did not specify the activities of the observer(s) during playback, whilst 26% (109) provided enough information to effectively reproduce what they were doing. Most (62%, 262), of the sample gave some information on what the observer(s) was doing, but not enough to faithfully reproduce it.

Recording Devices and Settings

Most of the articles in the sample (66%, 275) used custom tracks (either recorded or generated by the authors). Overall, tracks obtained from third parties were used in 9% (39) of the articles, while 5% (22) used a combination of custom and third-party tracks. The remaining 20% (82) did not provide enough information to identify the origin of their tracks. After excluding the 39 articles that obtained their recordings exclusively from third-parties and the 104 for which we had no information about track origins, 297 articles that used recording devices remained. Out of these, 70% (209) provided brand and model of their recording devices, and 9% (28) included additional details (e.g., frequency response). Forty-seven articles (16%) did not provide information about the recording devices employed. None of the articles provided frequency response charts for their recording devices. Fifty-nine percent (176) of the articles in which the tracks were recorded by the authors provided both sampling rate and audio bit depth, while the remaining 121 articles did not.

Sound file preparation

Following recording, most tracks were post-processed to prepare them for playback. This can include various digital filters such as denoising, amplitude correction, segmentation, and frequency range reduction. Most of the considered articles (73%, 305) reported something about their track preparation procedure. Of these, 7% (23) did not alter the tracks, 55% (167) gave some information about their procedure, and 37% (114) provided enough information to reproduce their track editing.

Broadcast

Whenever possible, we acquired technical specifications for the loudspeakers used by the authors. Most articles (65%, 274) reported at least brand and model of their devices; 13% (54) also provided additional specifications occasionally including frequency response intervals (e.g. "flat response

$\pm 3dB$"), whilst 90 articles (21%) did not specify the broadcasting device employed. No article reported frequency response charts for their broadcasting devices. Most of the articles in our sample (60%, 250) reported measuring their broadcasting sound pressure level (SPL). Twenty-two percent (93) provided a detailed report that included enough information to reproduce the same exact measurement, and 28% (117) reported partial information about the measurement. Twenty-five percent (104) reported the SPL of broadcast without explaining how it was measured; 37% (157) reported and explained their broadcasting SPL (57 of them also included references), while 34% (144) did not report it at all; 3% (13) explained their broadcasting SPL to the reader without reporting it. For a playback experiment to be fully reproducible, the tracks used by the authors need to be publicly available. This is not yet common, with tracks from only 3.6%, 15) of our sample being easily available (without having to contact the authors) whilst for some of the articles (7%, 31) a subset of the tracks were available. 372 articles (89%) did not provide any means of retrieving any of the acoustic tracks employed.

Objectiveness of the review and reproducibility of the articles

Overall, we feel confident we conducted an objective review, in both the selection and scoring of the articles. Of the two shared articles that we gave to five extra reviewers, one was always discarded as non-avian, whilst the other was consistently scored by all, although some minor differences arose. Most differences both among external reviewers and between the main reviewer and the external reviewers, related to SPL measurements.

Reproducibility of methods

If we define a reproducible article as one that faithfully describes how playback is performed, how the stimuli are prepared, how the observers behave, at what SPL sounds are broadcast, and that provides access to the employed tracks (although digital on-line archiving may not have been available for the entire sample, archives and institutions such as the Macaulay Library —– established in 1929 — have existed since long before acoustic playback), we would only have one such article in our sample. This is a study of the low frequency vocalisations of the North African Houbara Bustard *Chlamydotis undulata undulata* by Cornec et al. (2017). This definition of reproducibility specifically relates to criteria 4, 6, 7, 8, and 9 (Table 2.1), leaving the remaining questions for comparison purposes. For example, criterion 1, which relates to context, might be essential for reproducing an experiment in similar conditions or conditions that differed in known degree from those of the initial study. Assuming that we would not need to know how the stimuli were prepared, because they would be available, and not accounting for the SPL measurement parameter, we could drop criteria 6 and 9 and would have five reproducible articles in our sample of 419 articles (0.92%; Cornec et al., 2017; Dufty and Crandall, 2005; Krieg and Getty, 2016; Li et al., 2011; Moskát et al., 2018).

2.4 Discussion and Recommendations

Our literature review shows that as a community we are still not providing all the information that is needed to faithfully reproduce the research we carry out. In the next sections we provide some recommendations regarding the use of the playback technique and discuss aspects of avian physiology and behaviour that are likely to affect bird response to playback and hence highlight the need for reproducibility.

2.4.1 Recommendations (Table 2.2)

Report recording conditions. Ideally, high-quality, almost undegraded (e.g., recorded less than a metre away from the subject) vocalisations should be used to make playback stimuli, to avoid conditioning responses with unwanted distance cues (Lambrechts and Dhondt, 1995; Ríos–Chelén et al., 2017). Since this is not always possible, the best compromise we can currently suggest is to account for at least the SPL and meteorological variables measured when the sounds were recorded, and the distance from the bird at which these were recorded. If this is not possible, we would recommend broadcasting or otherwise producing a reference sound (reproducible elsewhere) at a known distance from the speaker; this will help in determining, albeit more approximately, the SPL of the recorded vocalising animal. By broadcasting vocalisations at the SPL measured at the time of recording, the possibility of broadcasting ambiguous signals, such as high-amplitude degraded songs, should be minimised.

We suggest that all details of the mechanics of the recording are presented: recording device details (brand and model, settings), environmental conditions (vegetation structure, base noise SPL, temperature, humidity, wind speed and direction), SPL (including detailed measurements, see below), and bird orientation (relative to the recorder, e.g., "facing the recorder") and distance whilst recording, are all desirable information to include in any report concerning the use of self-recorded stimuli. If tracks are obtained from third parties, the on-line repository identifiers are needed, together with citation of the original source reporting them. We suggest including all the details as comments accompanying the tracks uploaded to on-line repositories.

Report post-processing methods. Post-processing for track creation also needs to be reported because performing the same experiment using tracks with different dynamic range (because of the bit depth), frequency response (because of the sampling rate), or altered using different filters, denoising algorithms and/or more advanced edits, could all lead to different results.

Understanding the degree to which a target species can discriminate amplitude and frequency modulations is a decisive step in conducting effective broadcasting experiments.

Make tracks publicly available. We strongly recommend that high-quality, 'losslessly'–encoded versions, such as Free Lossless Audio Codec (FLAC) and Waveform Audio File (WAV) files, of the final tracks are used to be made publicly available on the internet. This enables readers to evaluate the quality of the sounds used, and other experimenters to directly use them, if applicable. This could be useful even when counterintuitive. For example, if we broadcast neighbour and stranger vocalisations to subjects, we may think them to be useful just to ourselves. However, our neighbour tracks could be employed by other researchers in need of a stranger track and so on.

There already are several online repositories that are, either partially or wholly, dedicated to avian vocalisations, such as Xeno-canto (`https://www.xeno-canto.org/`) and the Macaulay Library of the Cornell Laboratory of Ornithology (`https://www.macaulaylibrary.org`).

Whenever sharing the final versions of the employed tracks is not possible, as in the cases of commercial recordings, the latter can be referenced, and the references accompanied by a detailed description of post-processing procedures.

Select playback technology appropriately. The most prominent aspect to consider regarding the use of a loudspeaker to broadcast bird vocalisations is probably how these will be perceived by the birds. It is advisable, when possible, to consider known hearing and vocal frequency bandwidth limits of target species, and to choose devices accordingly. Some articles in our sample mentioned the frequency response of their loudspeakers; we add that it always is desirable to include frequency response charts for all the employed equipment. Unfortunately, this information is not always included in the equipment's specification sheets. However, a reasonable approximation can be generated by broadcasting a frequency sweep at normalised amplitude (which can be generated using free or commercial acoustic software) in front of the SPL meter in logging mode; plotting the resulting amplitude will result in a rough, albeit informative, frequency response chart. The distance between loudspeaker and SPL meter, conditions (e.g., soundproof room), and method of generating the frequency sweep for completeness should also be reported.

Report playback parameters. The SPL (or SPL range) at broadcast should be measured and device specifications and settings for both loudspeaker(s) and SPL meter, including brand, model, signal to noise ratio, frequency response chart, positioning, and eventual settings when applicable for loudspeakers should all be reported. Since measuring the SPL of natural sounds in the field is often harder than anticipated, we suggest generating a long (e.g., 30 seconds) pure tone at the peak amplitude of the vocalisation to broadcast, to normalise its amplitude with that of the natural sound at peak amplitude, and to more easily measure this instead of the more volatile natural sound. In addition, for the SPL meter, it is important to include distance at time of measurement(s), brand, model, scale (dBA/dBC/dBZ/other), response (fast/slow), range (e.g., 30 – 130 dB), and frequency response (possibly including the chart). Returning to the playback itself, we recommend reporting coordinates and time of year and day, as well as environmental variables such as altitude, temperature, humidity, wind speed and direction and a description of the area including surrounding vegetation community and geophysical information such as an indication of soil type and slope. The relevance of this contextual information is that although it may not directly influence the ability of other researchers to perform the experiment, it is helpful in interpreting results. Additional important features to report include *procedural* aspects, such the duration of playback, number of loudspeakers used, use of static/interactive playback, and observer(s) position and behaviour before and during playback.

Table 2.2: Useful details to report to facilitate reproducibility of ornithological playback technique applications.

Feature		points
context		1 to 8
location		
	w/ co-ordinates	2
	w/o co-ordinates	1
vegetation		
	nothing	0
	description	1
	references	2
time		
	of year	1
	of day	1
extra	(fauna/geology/weather)	
	nothing	0
	description	1
	references	2

Feature		category
tracks origin		
	not specified	1
	recorded by the authors	2
	obtained from a third party	3
	some recorded by the authors and some obtained from a third party	4
recording devices		
	not specified	1
	brand & model	2
	additional information	3
	frequency response charts	4
	not applicable	5
tracks' availability		
	not available	1
	some available	2
	available	3
recordings' details	(frequency range / bit depth)	
	not applicable	1
	not specified	2
	present	3
from tracks to stimuli		
	no editing	1
	outlined but not reproducible	2
	reproducible	3
	not specified	4
playback procedure		
	not specified	1
	outlined but not reproducible	2
	reproducible	3
obsevers' behaviour		
	not specified	1
	outlined but not reproducible	2
	reproducible	3
	not applicable	4
broadcasting spl		
	not specified	1
	not shown & explained	2
	shown & not explained	3
	shown & explained	4
	shown & referenced	5
spl measurement		
	not specified	1
	not measured	2
	measured & not described	3
	described & not reproducible	4
	reproducible	5
broadcasting devices		
	not specified	1
	brand & model	2
	additional information	3
	frequency response charts	4

2.5 Conclusions

The playback technique represents a very powerful tool in conservation contexts (e.g., eliciting vocalisations of rare or cryptic species), and can provide invaluable understanding of several aspects of a target taxon's behaviour. It is easy to use, relatively non-invasive and can provide directly interpretable outcomes. This makes it an excellent choice to pursue research questions in areas such as ecology, ethology, and evolutionary biology. While considering the development of our own studies involving acoustic playback, we were surprised to be unable to find a consistent set of guidelines for the design, execution, and reporting of such experiments. To our knowledge, the table provided by McGregor et al. (1992) whilst not intended as an instruction list, represented the only available guideline authors could follow to better design their experiments until now. A recent paper by Gentry et al. (2020), while not specifically about ornithological field applications, also provides very useful guidelines that correspond well to those we have identified here. Based on the literature and our own experience we provide, in Table 2.2, a list of features to consider and report to allow any reader to reproduce future applications of the playback technique. Of course, as technology advances, and for particular environments and experimental aims, there may be other factors that need to be considered. We also provide two sample paragraphs (Appendices I.i and I.ii) that, in our opinion, would represent exhaustive reports regarding potential studies involving acoustic playback. By consistently reporting methods in a community-wide fashion, and by making publicly available the acoustic stimuli we employ, the research community would not only allow fellow researchers to reproduce experiments, but also to compare and evaluate the experiments reported, hence improving the reliability of future developments in bioacoustics more easily.

Chapter 3

A, R Us There?
energy–based sound source location with single microphone acoustic recording units

Figure 3.1: Raffaele preparing for playback at Totara Reserve, Pohangina Valley

Abstract

Passive acoustic monitoring (PAM) using remotely deployable acoustic recording units can produce a long term permanent record of the soundscape of natural sites without significantly affecting the animals present. Applications of PAM have been expanding as technology improves.

Knowledge about the location of sound sources using passive acoustic methods could be useful in ecology and animal behaviour studies. However, to infer sound source location is computationally intense and requires more advanced hardware than a simple recorder. Another way to estimate the distance to sound source is using energy-based methods which use sturdier and cheaper devices at some cost in accuracy.

In this experiment we tested the applicability and accuracy of energy–based sound source localisation using single microphone Acoustic Recording Units, by broadcasting, recording, and localising multiple sounds, broadcast at different heights, directions, in three distinct environments.

We analysed the localisation errors obtained using five different configurations of single–microphone acoustic recording units and, for reference using a subset of data, by three microphone arrays using time difference of arrival (TDOA). As expected, the TDOA method outperformed the energy–based method with median localisation errors below 7 metres. However, although the overall energy–based median estimate error was 25 metres, optimising the spatial configuration of ARUs reduced this error to approximately 10 metres.

These results from using single microphone acoustic recording units to infer sound source locations have reasonable uncertainty. Although energy–based methods are less accurate than time of arrival ones, we found that errors could be reduced by calibrating the acoustic recording units configuration to one's target species.

The statistical analyses showed that some sounds (Kiwi female and a G5 pure tone) were easier to localise than others, and that one particular spatial configuration of ARUs (four ARUs at the vertices of a 100 m sided cross) generally delivered median estimates 5 m more accurate than with other configurations. Conversely, using only two ARUs 100 m apart (configuration #2) delivered median estimates with errors almost thirty metres worse than using other ARU configurations. Finally, estimating sounds in a level forest involved higher than average errors (by ∽ 4 metres), and very high (equal or higher than 8,000 Hz) and low (lower than 300 Hz) pure tones also involved higher median errors (around 3 metres higher than other sounds). Unexpectedly, weather, height/angle combinations, and directions of broadcast did not seem to affect this localisation method to an appreciable degree.

Although time of arrival methods may eventually be routinely used as they indeed provide more accurate estimates, their applicability to large scale monitoring may not yet be practical or affordable. As some statistical methods and estimates can benefit from sound source location information, we hope to enable their application using energy–based methods besides the already–established time of arrival ones. We conclude that teams wishing to use energy–based methods for localisation of

animals from vocalisations need to carry out similar experiments to the one presented herein with their own equipment to select the ARU configuration that best suit their research or management questions.

3.1 Introduction

Acoustic Recording Units (ARUs) allow the collection of large datasets with the potential for great value in conservation (Priyadarshani et al., 2018a). They are particularly useful when monitoring species of conservation concern with small populations in fragmented and often difficult to reach locations (e.g. Williams et al., 2018). A large number of different ARUs are available to researchers (Blumstein et al., 2011; Pérez-Granados et al., 2019a), and Passive Acoustic Monitoring (PAM) using ARUs has an important role in soundscape studies (Farina, 2013) and studies of different animal taxa, such as invertebrates, mammals, amphibians, and birds (Sousa-Lima et al., 2013; Sugai et al., 2019).
One particular aspect of acoustic monitoring involves sound source localisation, which aims at locating animals mostly using time difference of arrival (TDOA) of sounds to arrays of synchronised microphones (Marques et al., 2013; Rhinehart et al., 2020). Conversely, energy–based estimates use power, scaling its estimates to a reference sound with known location, to approximate distance (Li and Hu, 2003). At a minimum, spatial localisation helps to disambiguate whether the detections of different sets of devices are of the same vocalisation or not. However, in frameworks such as Spatially Explicit Capture-Recapture (SECR — Efford, 2004; Efford and Boulanger, 2019; Royle et al., 2013) location information can also be added as supporting variables (Stevenson et al., 2015). Being able to use PAM to obtain sound source location will improve monitoring practices, for example in the cases of point counts using distance sampling (Buckland et al., 2012; Ralph et al., 1995), which require estimates of location, and as such are currently limited to human observers rather than ARUs. Identifying sound source location with PAM allows for the extension and possible improvement of these and similar practices.

The ability to estimate sound source location has been shown using arrays of microphones or recorders (Darras et al., 2018a; Fan et al., 2010; Rhinehart et al., 2020; Suzuki et al., 2017). These are more expensive, and power–hungry than traditional single microphone ARUs. PAM often requires sturdy and weather–proof devices able to be deployed in the wilderness for long periods of time, relying on mimimal power supply (which is also important in wireless synchronisation networks, see Cobos et al., 2017). Using TDOA methods requires perfect synchronisation among the recording channels. While this can be achieved by physical, wireless, or acoustic means (Rhinehart et al., 2020), it is often impractical (e.g. periodic synchronising sound broadcast, expensive wireless synchronisation), or environmentally invasive (e.g. physical synchronisation via cables) for long term monitoring. For instance, VoxNet emerged as a very promising field–deployable TDOA system, but is still very expensive and difficult to maintain (Allen et al., 2008; Rhinehart et al., 2020; Taylor et al., 2016). Eventually, technology will surely allow us to effortlessly use high precision TDOA methods routinely (Hart and Martinez, 2006). Remotely deployable multi–channel ARUs are emerging (e.g. Wijers et al., 2021), but these recently developed devices have not replaced single–microphone ARUs in everyday conservation surveys, and energy–based localisation could always be employed to analyse historically collected ARU data. As such, PAM can benefit from energy–based sound source localisation (Li and Hu, 2003), which uses the strength of the received signal to approximate distance. Recordings of sound sources closer to a recording device will display higher amplitude, hence energy, than more remote equally loud ones. A reference sound from a known location serves to thus convert inverse energies to distances in metres. Estimated distances from multiple detectors can finally be integrated to localise the sound source.

In this study we used playback to evaluate the accuracy of different configurations of ARUs at estimating sound source location; we further compared a subset of these estimates to estimates made using three microphone arrays and TDOA. The benefits of a playback experiment are that we know the true locations of both the sound source and the ARUs or arrays, and thus have ground truth for the spatial location of a sound. We use Euclidean distance between known and estimated sound source locations (i.e., localisation error) as our response variable, seeking to identify variables that are consistently associated with more accurate localisation of the playback speaker. We used a statistical model considering different ARU configurations, height/vertical angle of broadcast, trial (environment), direction of broadcast, and sounds played, as covariates that might affect the the sound localisation errors.

3.2 Materials and Methods

3.2.1 Track Preparation

Our playback track was 1 minute, 16 seconds long, and comprised noise, pure tones, and avian vocalisations, each separated by silence (Table 3.1). Although this thesis is chiefly focussed on North Island Brown Kiwi (*Apteryx mantelli*, Bartlett, 1851) as main taxon of interest, we included several Aotearoa New Zealand endemic taxa and a commonly–studied North American taxon (Black Capped Chickadee — *Poecile atricapillus*, Linnaeus, 1766) to give an heftier contribution to acoustic monitoring and broaden the applications of the method. We prepared the track using Audacity (Audacity Team, 2018). We normalised all the track segments (except the initial noise) to have the same peak amplitude (so that the most intense sound of each segment had the same intensity/sound pressure level). We used a second track, comprising a 30 seconds long cyclic repetition of a standard 440 Hz A (at the same amplitude as in the main track) to calibrate the speaker's sound pressure level in the field. Both tracks, in reproducible (mp3 extension 44100 Hz rate, 32 bit depth) and editable (Audacity project `.aup`) forms, are available within the supplementary material. We here used uncompressed WAV tracks for our sound collection and analyses, but are providing mp3 and aup versions for ease of download.

The first two seconds of the playback track included software–generated Brownian noise; after 0.25 seconds of silence, an equal tempered (Hagenow, 1934) C Major arpeggio from C3 to C6 of generated tunes followed (Suits, 1998). Each tone of the arpeggio lasted one second and was separated from the next by a pause (silence) of 0.25 seconds. We chose to include a C Major arpeggio, rather than a different scale, as it is easily reproducible with a variety of musical instruments. After a 0.5 seconds pause (silence) following C6 the avian vocalisations started. Each avian vocalisation segment lasted five seconds and was separated by 0.5 seconds of computer–generated silence from the next.

The first avian vocalisation in our track was that of a Ruru (Morepork – *Ninox novaeseelandiae*, Gmelin – 1788), followed by Hihi (Stitchbird – *Notiomystis cincta*, Du Bus & Gisignies, 1839), Kea (*Nestor notabilis*, Gould, 1856), Matuku (Australasian Bittern – *Botaurus poiciloptilus*, Wagler, 1827), female North Island Brown Kiwi, Toutouwai (North Island Robin – *Petroica longipes*, Lesson & Garnot, 1827), North Island Brown Kiwi male, Black–Capped Chickadee, and Tīeke (North Island Saddleback *Philesturnus rufusater*, Lesson, 1828). All scientific names and authorities were confirmed on the BirdLife International website on the 5th of February, 2020. A sequence of pure tones (500, 2,000, 4,000, 8,000, 12,000, and 16,000 Hz; Darras et al., 2018b) lasting 1 second each and 0.25 seconds apart followed 0.5 seconds after the last bird sound. A pause of 0.25 seconds was followed by a 440 Hz pure tone and 2 seconds of silence. This last pause allowed the rotation between broadcasting directions (see below for details).

Table 3.1: Broadcasting track summary, all the gaps between segments are filled with software–generated silence, as well as the last two seconds from 01:14.000 to 01:16.000 to allow us to rotate the structure between tracks' repetitions.

From	To	Sound	Source file contributor (Source — Portion on Source File)
00:00.000	00:02.000	Brownian noise	Audacity, generated
00:02.250	00:3.250	C3 — 130.81 Hz	Audacity, generated
00:03.500	00:04.500	E3 — 164.81 Hz	Audacity, generated
00:04.750	00:05.750	G3 — 196.00 Hz	Audacity, generated
00:06.000	00:07.000	C4 — 261.63 Hz	Audacity, generated
00:07.250	00:08.250	E4 — 329.63 Hz	Audacity, generated
00:08.500	00:09.500	G4 — 392.00 Hz	Audacity, generated
00:09.750	00:10.750	C5 — 523.25 Hz	Audacity, generated
00:11.000	00:12.000	E5 — 659.25 Hz	Audacity, generated
00:12.250	00:13.250	G5 — 783.99 Hz	Audacity, generated
00:13.500	00:14.500	C6 — 1046.50 Hz	Audacity, generated
00:15.000	00:20.000	Ruru	Wagner (XC114338 — 2012, $0m16s$ to $0m21s$)
00:20.500	00:25.500	Hihi	Slaymaker (XC436719 — 2018, $0m19s$ to $0m24s$)
00:26.000	00:31.000	Kea	van Bemmelen (XC198866 — 2014, $0m03s$ to $0m08s$)
00:31.500	00:36.500	Matuku	Graff (XC165265 — 2013, $3m17s$ to $3m22s$)
00:37.000	00:42.000	North Island Brown Kiwi female	Moon et al. (CD Track 1 — 2012, $1m22.5s$ to $1m27.5s$)
00:42.500	00:47.500	Toutouwai	Fulford (XC348754 — 2016a, $1m16.5s$ to $1m21.5s$)
00:48.000	00:53.000	North Island Brown Kiwi male	Moon et al. (CD Track 1 — 2012, $2m18s$ to $2m23s$)
00:53.500	00:58.500	Black–Capped Chickadee	Leite (XC420822 — 2018, $0m08.5s$ to $0m13.5s$)
00:59.000	01:04.000	Tīeke	Fulford (XC348779 — 2016b, $0m2.234s$ to $0m7.234s$)
01:04.500	01:05.500	500 Hz as in Darras et al. (2018b)	Audacity, generated
01:05.750	01:06.750	2000 Hz as in Darras et al. (2018b)	Audacity, generated
01:07.000	01:08.000	4000 Hz as in Darras et al. (2018b)	Audacity, generated
01:08.250	01:09.250	8000 Hz as in Darras et al. (2018b)	Audacity, generated
01:09.500	01:10.500	12000 Hz as in Darras et al. (2018b)	Audacity, generated
01:10.750	01:11.750	16 000Hz as in Darras et al. (2018b)	Audacity, generated
01:12.000	01:14.000	440 Hz A as in calibrating track	Audacity, generated

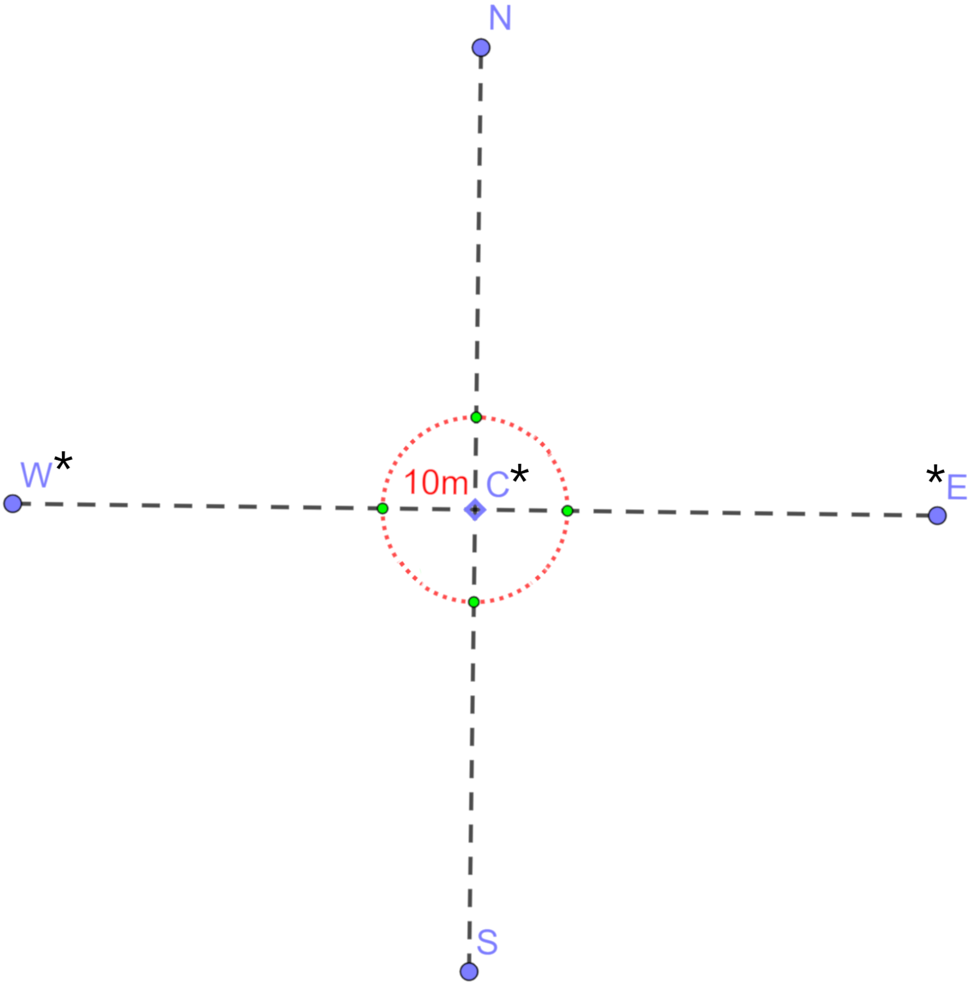


Figure 3.2: Schematic representation of the experimental setup. The cardinal points and blue dots represent four ARU locations, the capital C and blue rhomboid represent the central ARU and weather station's location, the green dots on the dotted red circumference represent the broadcasting stations. Asterisks (*) represent ARU locations where microphone arrays were deployed for the comparison. Each of the two perpendicular black dashed lines is 100 metres long, so the distance between C and any of the vertices is 50 metres.

3.2.2 Recording setup

We deployed 5 ARUs at the vertices and intersection of an equilateral 100 m cross–shaped experimental field, with one ARU at the centre and one at each vertex of the cross (0 ° /360 ° **N**orth, 90 ° **E**ast, 180 ° **S**outh, and 270 ° **W**est, Figure 3.2).

When recording data with microarrays we deployed three laptop computers at the central, eastern, and western recording points respectively. We connected a miniDSP UMA-8 multi–channel device, comprising an array of seven Knowles SPH1668LM4H-1 microphones. These commercial microarrays comprise one central microphone surrounded by six more and a mini-USB port along circumference. The mini-USB ports faced North. Both the frequency response chart and microarray schematics are available n the manufacturer's data sheet.

We used two different types of ARU: the New Zealand Department of Conservation (DOC) "AR4" (Frequency response curve in Appendix II, Figure II.iii, `electronics@doc.govt.nz`), and the Cornell Laboratory of Ornithology "Swift" (Frequency response curve in Appendix II, Figure II.iv, `BRP@cornell.edu`). Each ARU was positioned on a tripod so that the centre of the microphone was 120 cm above the ground, and pointing towards the ground as to resemble field applications (upward facing ARUs would be exposed to both additional noise from rain, and possible accumulation of debris in the microphone's cage). This height was chosen to match the conventional use of chest height for ARU deployment (e.g. Towsey et al., 2014).

We randomly assigned the ARUs to the five ARU–locations at each change of broadcasting location to minimise possible device–related bias.

3.2.3 Experimental Procedure

The ground preparation phase can be summarised as follows:

1. Arrival on site.
2. Identification of central point using GPS device.
3. Identification of ARUs points (N, S, W, and E) using GPS device, tape measure, and compass.
4. Positioning of tripods and ARUs at the five ARUs locations (plus laptops with microarrays when required).
5. Positioning of weather station at the centre of the experimental field, close to the central ARU.
6. Identification and marking of broadcasting locations of the day (depending on schedule).
7. Adjusting of sound pressure level of the calibration track (see Track Preparation, page 39) to be within 75 ± 1 dbA at 1 metre horizontal distance between the speaker and the sound pressure level meter (see broadcasting, page 44), the former facing the latter, which we set for fast response and range $[30, 130]$ db.
8. Start weather logging. We logged temperature, relative humidity, atmospheric pressure, and wind speed and direction every 10 seconds using a Kestrel 4500BT Weather Meter (see `https://kestrelmeters.com/products/kestrel-4500-weather-meter`) mounted on a tripod at the centre of the experimental field.

For each broadcasting trial, the playback phase of the experiment comprised: blowing a sport whistle three times from the centre for distance calibration and acoustic synchronisation purposes, setting current height/angle configuration, and cyclically broadcasting the track towards each of the directions, starting from N, to E, S and finally W.

The sport whistle blowing served as auditory aligning marker (to align the distance calibration signal, as we blew it from right above the central ARU and, later on, used it as a reference sound 50 m away sound for each vertex–ARU. Additionally, we used this as an alignment marker to pan all the five ARU recordings to the same starting point as their clock could have had small offsets between each other.

3.2.4 Field Sites and Trials

The field trials were conducted on three different experimental fields: *open field* (hereafter open), *level land forest* (hereafter level), and *uneven ground forest* (hereafter uneven). The *open* trial took place on a soccer/rugby field at Massey University, Palmerston North, Manawatu–Whanganui, Aotearoa New Zealand. The centre of the experimental field was located at S40.38728 ° E175.62515 ° (geodetic reference datum for the coordinates is WGS 84). These trials were performed between the 12th and the 23rd of February, 2019. During the *open* trial we additionally found out the distance at which ARUs and microarrays would not detect broadcasts (measured as not being able to see the vocalisations on a spectrogram).

The *level* and *uneven* field trials took place between 2nd and 6th of October, 2019 at the Tōtara Reserve, in the Pohangina Valley, Manawatu–Whanganui, Aotearoa New Zealand. The centre of the *level* experimental field was located at S40.13218 ° E175.84843 ° whereas that of the *uneven* experimental field was located at S40.15131 ° E175.85342°. The Tōtara Reserve lies approximately fifteen kilometres North to Northeast of the small town of Ashhurst, in the Manawatu–Whanganui region. The reserve is crossed, generally Northeast to Southwest, by the Pohangina River, which is a tributary of the Manawatu River. The vegetation, in both the *level* and *uneven* sites was fairly dense (with trees' basal area covering around $60\frac{m^2}{ha}$ — Priyadarshani et al., 2018a, see Fig. 3.3) and mainly comprised native trees and understory, including flagship species such as Northern rātā (*Metrosideros robusta*, Cunn — Hassler, 2020b) and Ponga ferns (*Alsophila dealbata*, Forst — Hassler, 2020a). For more detailed information about the vegetation and geological structures of the area see Lusk, (1984) and Challands, (1976).

3.2.5 Broadcasting

We established broadcasting locations every 90 ° along a circle of 10 metre radius centred on the middle ARU (C), starting from due North. We deployed the speaker in three different configurations: (i) at height 350 cm above the ground, and horizontal, (ii) at height 350 cm above the ground, and facing upwards at 30 ° to the horizontal, and (iii) at ground level and facing upwards at 30 ° to the horizontal. This deployment was facilitated by an adjustable (telescopic) supporting

Figure 3.3: Performing playback at the Massey University Grounds (A) and Totara Reserve, Pohangina (B). This figure further shows the structure operating elevated flat configuration on (A) and at ground level on (B).

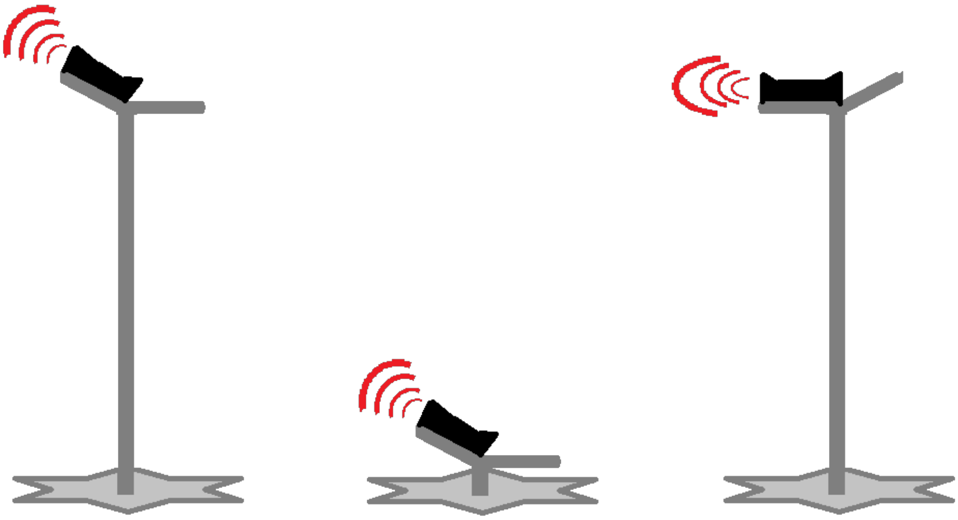


Figure 3.4: Schematic representation of the broadcasting structure in its three different configurations. The structure is represented hosting the speaker.

aluminium structure (Figure 3.4). The compass direction of the call was also varied during the recording, in 90° increments from looking North (i.e., first looking North, then East, then South, and finally West).

The broadcasts were made using a Foxpro NX4 speaker, for which we experimentally derived the frequency response (Figure II.ii), by playing a frequency sweep we generated using Audacity. This included a calibrating introduction (440 A at 0 — full — amplitude, 20 s) followed by a 5 seconds pause and a 120 seconds $[20, 20000]$ 0 — full — amplitude linear frequency sweep (Pope, 2008). We then used the speaker's curve (from the SPL meter) to adjust the acoustic recording devices' frequency response (figures II.iv and II.iii) curves obtained by recording the same track played by the same speaker using REW (Mulcahy, 2004). These frequency response charts are specific of our devices and do not necessarily represent the performance of other speakers or recording devices of the same brands and models.

All the sound pressure level (SPL) measurements were made using a Digitech QM1592 Professional Sound Level Meter (see `https://www.jaycar.com.au/pro-sound-level-meter-with-calibrator/p/QM1592` or `https://www.techbrands.com`), which was calibrated before usage with the provided calibrator according to the user manual.

3.2.6 Location Estimation

Time Difference of Arrival

We used an implementation of the generalised cross–correlation algorithm (Knapp and Carter, 1976, Appendix II.ii) written by Richard Brown and Stephen Marsland, to estimate the direction of arrival of each sound, and subsequently localise the sound sources. We estimated the direction of arrival of all the sounds from one location (North) in the *open* trial, estimating directions of arrival of each sound to the three microphone microarray in a 0.3 seconds window. Figure 3.5 illustrates the output of the algorithm for one of the localised sounds.

Energy–based estimation

We prepared multi–channel Audacity projects (Audacity Team, 2018) with the recordings of the broadcast from each of the five ARUs. Each project comprised the five (one per ARU) unsynchronised parallel channels. We aligned the tracks using the sport whistle sound at each location/trial. Using the weather data from the portable weather station, we calculated the specific speed of sound for each synchronising point roughly every 15 minutes. We used this speed to calculate how much time would it take for sound to travel 50 metres and we thus manually shifted the North, East, South, and West channels accordingly, using the Centre track as reference. Using the central channel as a reference (because it was the closer to all broadcasting location, hence the most likely to always hear the playbacks), we then isolated a 0.1 s to 0.3 s fragment for each part of the playback track in the current section and extracted it in five-channels wav files. We used AviaNZ (Marsland et al., 2019, `http://avianz.net`) to compute the signal power of each channel using equation (3.1). To investigate whether a particular configuration of ARUs, out of the following five, resulted in smaller localisation errors:

Configuration #1 Only Central and Eastern tracks included

Configuration #2 Only Western and Eastern tracks included

Configuration #3 Central, Northern, and Eastern Included

Configuration #4 Northern, Eastern, Southern, and Western tracks included

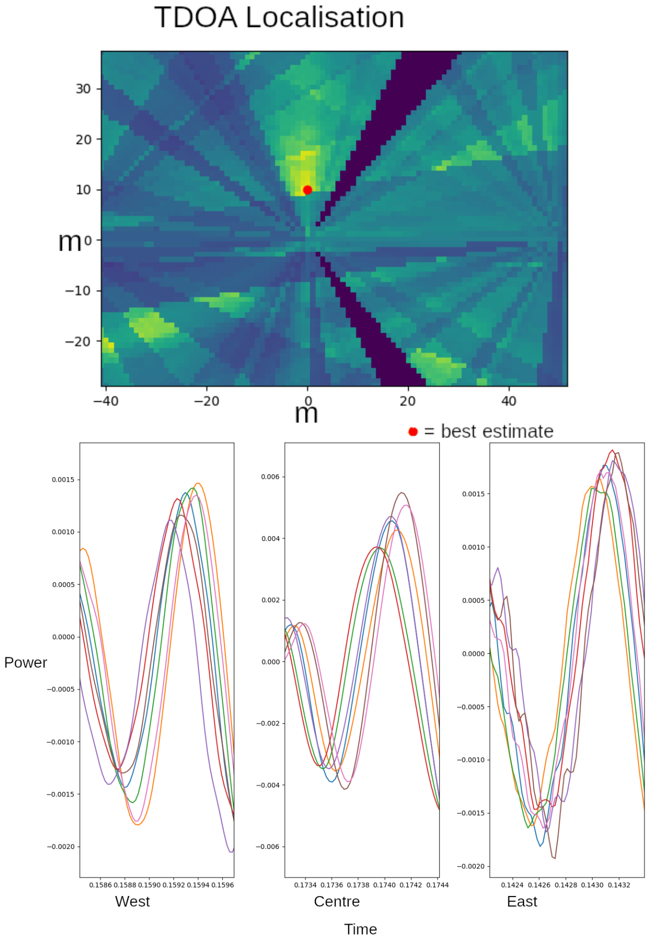


Figure 3.5: Example output of the GCC implementation. The upper plot shows with a red dot the best estimate of the current sound source whilst the three lower panels show representations of the signals from the seven channels of each of the microarrays. These are, from left to right, the Western, Central, and Eastern microarrays.

Configuration #5 All five ARUs

We then used the least squares method (Li and Hu, 2003) to infer the most likely sound source location (z) with different ARU configurations (using equation 3.2), where C is the current set of ARUs considered. Finally, we calculated the Euclidean distances between known and estimated locations as in Equation (3.3). This resulted in a total of 7,795 distances between known and estimated locations.

$$\text{Power} = \frac{\sum_{t=0}^{T} s_t^2}{T} \quad \text{where T} = \text{length of file} \tag{3.1}$$

$$\binom{x_{est}}{y_{est}} = \binom{x}{y} - \binom{x_c}{y_c} - \frac{r_i \times \left[\begin{smallmatrix} x - x_c \\ y - y_c \end{smallmatrix}\right]}{(x - x_c)^2 - (y - y_c)^2} \tag{3.2}$$

$$\forall\, c \in C$$

$$\Delta l_{est}^{known} = \sqrt{(x_{known} - x_{est})^2 + (y_{known} - y_{est})^2} \tag{3.3}$$

3.2.7 Statistical Analyses

We modelled the Euclidian distance between the known broadcasting locations and the estimated coordinates with a normal likelihood (Δl) with lower bound 0 — Equation (3.4) — using Stan (`https://mc-stan.org`), through the `R` (`R` Core Team, 2013) package `rstan` (Stan Development Team et al., 2016). Here we worked within a Bayesian framework for three interconnected reasons. First, Bayesian statistics makes it easy to discern how different factors may have impacted the estimates. By specifying zero–centered prior distributions (before the simulations) for all the covariates, the posterior distributions (after the simulations) that fell completely above zero consistently contributed to larger mean euclidean distances (bigger errors). Conversely, the posterior distributions that fell completely below zero consistently contributed to smaller mean euclidean distances (smaller errors), and those that remained to some extent centered around zero did not consistently contibute to either larger or smaller errors. Secondly, by choosing a normal likelihood (i.e., not employing a link function), the contributions of each different factor, represented by a covariate (β) in the linear expression in Equation (3.5), are expressed in the same unit of measure of the response variable, that is metres, hence are easily interpretable. Finally this method allows us to achieve a quantitative understanding of the contribution of different factors to the euclidean distances (errors), as posterior distributions that fell closer to zero represented smaller contributions to the response variable's mean than those that fell farther from zero. Each occurrence was modelled with hierarchical mean and standard deviation.

$$\Delta l_q \sim Normal(\mu = m_q, \sigma = sig_0) \, \forall \, q \in [1:Q] \tag{3.4}$$

$$\begin{aligned} m_q = &\beta^0 + \beta^{tri}_{tri_q} + \beta^{ha}_{ha_q} + \beta^{tra}_{tra_q} + \beta^{dob}_{dob_q} + \beta^{ARUc}_{ARUcq} + \\ &+ \beta^{AP} \times AP^Z_q + \beta^{RH} \times RH^Z_q + \beta^{T} \times T^Z_q + \beta^{WS} \times WS^Z_q \end{aligned} \tag{3.5}$$

$$\ln(sig_w) \sim Normal(\mu = 0, \sigma = 1) \, \forall \, w \in [0:W] \tag{3.6}$$

$$\beta^*_w \sim Normal(\mu = 0, \sigma = sig_w) \, \forall \, w \in [1:W]$$

$$z = \frac{x - \mu}{\sigma} \tag{3.7}$$

The mean was further dependent on a linear expression — Equation (3.5) — that included a global offset parameter (or intercept, β^0), the following categorical covariates: trial/environment (open, level, uneven — `tri`), height (elevated flat, elevated angled, ground level — `ha`), direction of broadcast (N, E, S, W — `dob`), track (26 categories, see Track Preparation, page 39 — `tra`) and ARUs configuration (1, 2, 3, 4, 5 — `ARUc`), and the following non–categorical weather covariates, each multiplying its z–standardised value — Equation (3.7): Atmospheric Pressure (`AP`), Relative Humidity (`RH`), Temperature `T`, and Wind Speed (`WS`). Since we did not obtain reliable estimates of wind direction for all (open, level, uneven) trials, we did not include this covariate in the analysis. All the natural logarithms of the standard deviations were modelled as normally distributed with mean equal to zero and standard deviation equal to one — Equation (3.6) — 47 σ were modelled: sig_0 as hierarchical prior of the normal likelihood and sig_1 through sig_{46} for the β covariates (the superscript asterisk * on the β is to indicate that any and all the β_w have normal priors with mean equal to 0 and hierarchically modelled standard deviations). We chose to have all hierarchical standard deviations thus modelled to minimise constraints on covariates and decided to model the logarithms of the standard deviations to both allow for easier interpretation and ensure that these would be positive real numbers.

$$\Delta l_q \sim Normal(\mu = m_q, \sigma = sig_0) \, \forall \, q \in [1:Q] \tag{3.8}$$

$$m_q = \beta^0 + \beta^{method}_{method_q} \tag{3.9}$$

Additionally, we ran a simpler model — Equations (3.8) and (3.9) — to compare the TDOA and energy–based estimates. This only included a global intercept and a categorical *method* covariate.

We ran both models for $20,000$ iterations, discarding the first half as warm-up/burn-in, a thinning of 4 (that is, retaining one every 4 MCMC draws), a proposal acceptance probability (`adapt_delta`) of 0.999 and a maximum treedepth of 15 in six independent chains. This resulted in a final sample of 15,000 draws per variable.

3.3 Results

The average (mean squared) error with the energy–based method was 21.49 m, while the TDOA one was 6.65 m (Figure 3.10, Table II.i).

Figure 3.6 collectively portrays the distribution of the Euclidian distances between real and estimated sound source locations and the distribution of the z–standardised weather variables. The abscissæ in plots (A) through (E) are in metres, whereas those in plots (F) to (I) are in standard deviations terms, being the z distribution standardised with $\mu = 0$ and $\sigma = 1$. We found that no avian vocalisations were detectable from visual inspection of the spectrograms farther than 300 m from the ARUs, and that only the C6 pure tone would still be detectable at 350 m.

Table 3.2: Summary of the distribution of the Euclidian distances (in metres) in different subsets of the recorders.

Subgroup	%2.5%	50%	97.5%
All data	14.06	25.00	34.94
Configuration #1	14.70	25.30	31.79
Configuration #2	40.49	47.82	58.18
Configuration #3	18.84	24.41	32.26
Configuration #4	8.14	9.98	13.05
Configuration #5	17.99	22.12	33.98

The MCMC simulation terminated without warnings and the Gelman–Rubin convergence statistic $\hat{R}$ was close to, if not exactly, one, for all the simulated variables.
No direction of broadcast, nor height/angle configuration or weather covariate was associated with smaller or larger Euclidean distances between known and estimated locations (i.e., they were largely centred around zero).

We found that the best configuration to estimate the distance to the broadcast was Configuration #4, with four ARUs placed at the four vertices of the area, and the worst configuration #2, where there were just two ARUs 100m apart (Figure 3.7).

The covariate for Configuration #2 was associated with errors in the sound source location at times 30 m (see Table 3.3, 97.5% Quantile column) larger than average while the location errors associated with the covariate for configuration #4 were mostly $\sim$ 11 m (see Table 1, 50% Quantile column – median) smaller than average.

There was also an effect of trial on our ability to accurately locate sound source. Broadcasts on the Level trial were consistently associated with slightly larger errors than average while broadcasts on the uneven trial were mostly associated with smaller than average errors (Figure 3.8).

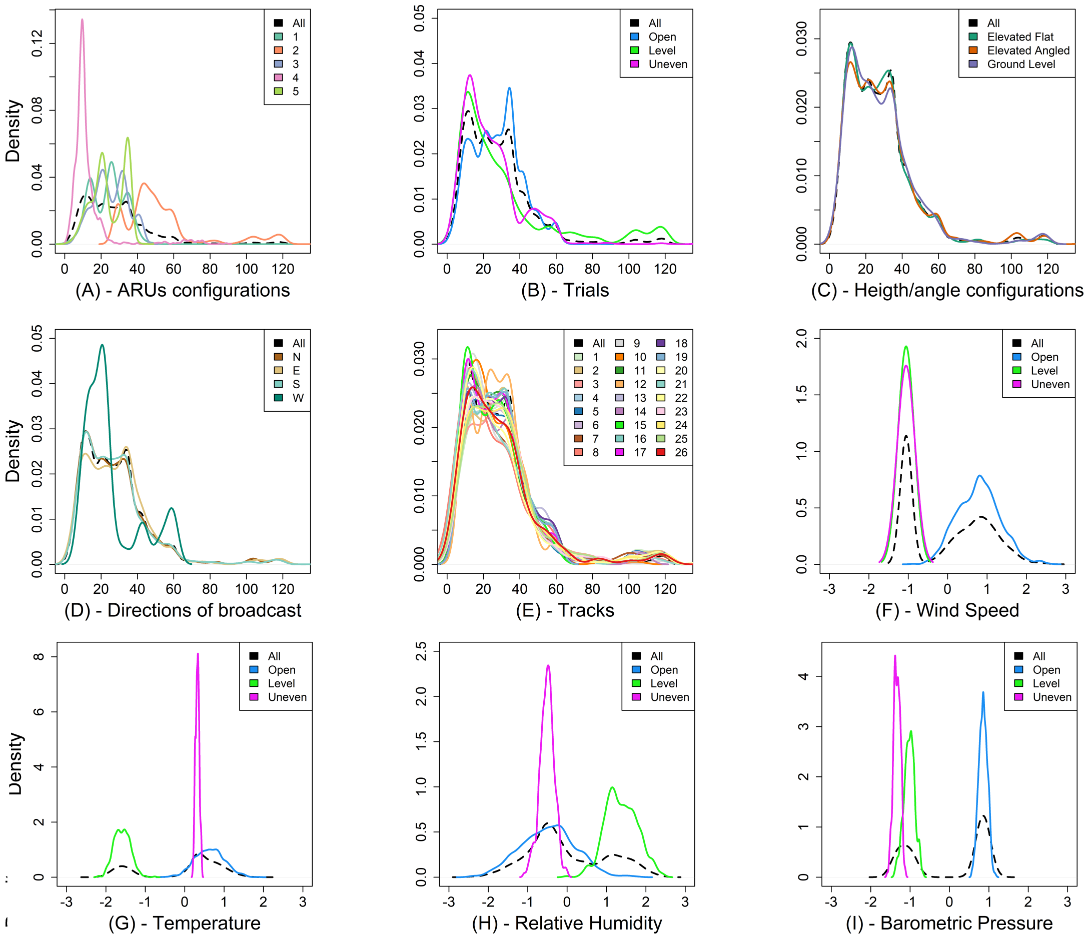


Figure 3.6: Plots A through E depict density curves of the distributions of the Euclidean distances from known to estimated locations grouped by the different considered categories whereas plots F through I likewise present the z–standardised values for the weather variables. All the plots include a black dashed line representing the distribution of the considered variable regardless of any grouping and a differently coloured continuous line for each of the groups. For example, see how in plot (C) the distribution of the different height/angle broadcasting configurations almost retrace the overall distribution of Euclidian distances, whereas in plot (H) the overall distribution of standardised temperature values markedly differs from both the Level and Uneven trials. For most of the Euclidean distances plots (B—E), the distributions of different groupings are fairly similar to the overall distribution of data, exception made for plot D where sounds broadcast towards the West have overall smaller errors. Configuration # 4 presented the lowest values together with a narrower distribution width (Table 3.2); in plot (A) we see that the pink line, corresponding to ARUs conguration # 4, indicates a high density of small Euclidian distances peaking around 10 m and barely going over. In contrast, conguration # 2 (orange line) shows overall 20–60 m Euclidian distances/errors with some occurrences up to over 100 m, and the remaining configurations fluctuate within the 15–40 m range.

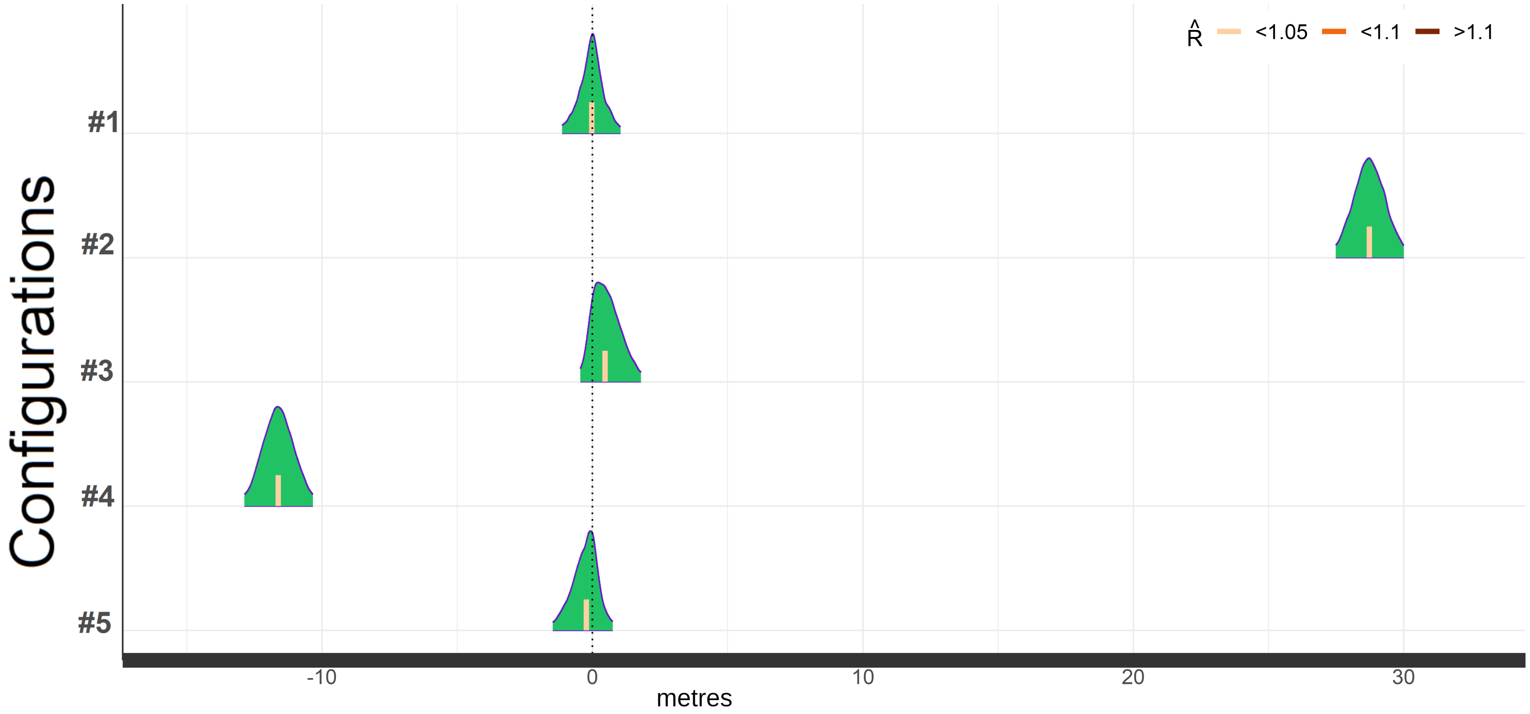


Figure 3.7: Posterior distributions for different ARU configurations. Configuration #2 is consistently associated with larger Euclidian distances between known and estimated locations whereas configuration #4 shows smaller ones.

Table 3.3: Gelman-Rubin convergence statistic ($\hat{R}$), standard deviation, 2.5%, 50% (median), and 97.5% of the posterior distribution for the considered covariates (parameters).

Parameter	$\hat{R}$	sd	2.5%	50%	97.5%
Intercept	1.00	8.00	-2.64	0.44	24.79
$\beta height/angle$ Elevated Angled	1.00	0.62	-1.81	-0.38	0.60
$\beta height/angle$ Elevated Flat	1.00	0.65	-0.29	0.92	2.20
$\beta height/angle$ Ground Level	1.00	0.59	-1.37	-0.04	1.04
$\beta ARUs$ #1	1.00	0.52	-1.12	-0.02	1.04
$\beta ARUs$ #2	1.00	0.64	27.49	28.73	30.00
$\beta ARUs$ #3	1.00	0.58	-0.45	0.47	1.79
$\beta ARUs$ #4	1.00	0.64	-12.87	-11.62	-10.34
$\beta ARUs$ #5	1.00	0.55	-1.47	-0.23	0.75
βDOB N	1.00	0.47	-1.08	-0.12	0.82
βDOB E	1.00	0.52	-0.41	0.39	1.61
βDOB S	1.00	0.47	-0.77	0.07	1.13
βDOB W	1.00	1.01	-3.18	-0.53	0.75
$\beta Trial$ Open	1.00	1.52	-3.34	0.33	3.17
$\beta Trial$ Level	1.00	1.57	0.31	4.16	6.96
$\beta Trial$ Uneven	1.00	1.57	-5.76	-1.88	0.58
$\beta Sound$ C3	1.00	1.08	1.33	3.52	5.59
$\beta Sound$ E3	1.00	1.10	0.82	3.13	5.15
$\beta Sound$ G3	1.00	1.10	0.66	2.89	5.02
$\beta Sound$ C4	1.00	1.08	0.06	2.11	4.21
$\beta Sound$ E4	1.00	1.02	-0.16	1.49	3.62
$\beta Sound$ G4	1.00	1.02	-0.13	1.55	3.63
$\beta Sound$ C5	1.00	0.66	-0.87	0.21	1.84
$\beta Sound$ E5	1.00	1.01	-3.61	-1.44	0.13
$\beta Sound$ G5	1.00	1.05	-4.09	-2.01	-0.05
$\beta Sound$ C6	1.00	1.01	-3.73	-1.62	0.07
$\beta Sound$ Ruru	1.00	0.75	-2.36	-0.49	0.56
$\beta Sound$ Hihi 5	1.00	0.72	-2.19	-0.35	0.67
$\beta Sound$ Kea	1.00	0.69	-2.05	-0.29	0.72
$\beta Sound$ Matuku	1.00	0.63	-1.68	-0.13	0.95
$\beta Sound$ Kiwi Female	1.00	1.02	-4.12	-2.09	-0.17
$\beta Sound$ Toutoutwai	1.00	0.64	-1.74	-0.18	0.92
$\beta Sound$ Kiwi Male	1.00	0.93	-3.27	-1.20	0.21
$\beta Sound$ Black-Capped Chickadee	1.00	0.61	-1.58	-0.12	0.97
$\beta Sound$ Tīeke	1.00	0.67	-1.93	-0.25	0.82
$\beta Sound$ 500 Hz	1.00	0.99	-0.08	1.60	3.59
$\beta Sound$ 1000 Hz	1.00	0.98	-3.52	-1.47	0.10
$\beta Sound$ 2000 Hz	1.00	0.72	-0.65	0.40	2.17
$\beta Sound$ 4000 Hz	1.00	1.01	1.21	3.34	5.20
$\beta Sound$ 8000 Hz	1.00	1.00	1.41	3.47	5.35
$\beta Sound$ 16000 Hz	1.00	1.00	1.89	3.89	5.79
$\beta Sound$ A 440 Hz	1.00	0.70	-0.77	0.29	2.06
$\beta Weather$ Temperature	1.01	6.61	-4.01	0.05	8.20
$\beta Weather$ Relative Humidity	1.00	8.18	-28.27	-0.29	2.86
$\beta Weather$ Barometric Pressure	1.00	6.92	-18.06	-5.44	1.65
$\beta Weather$ Wind Speed	1.00	8.12	-24.03	-0.53	2.53

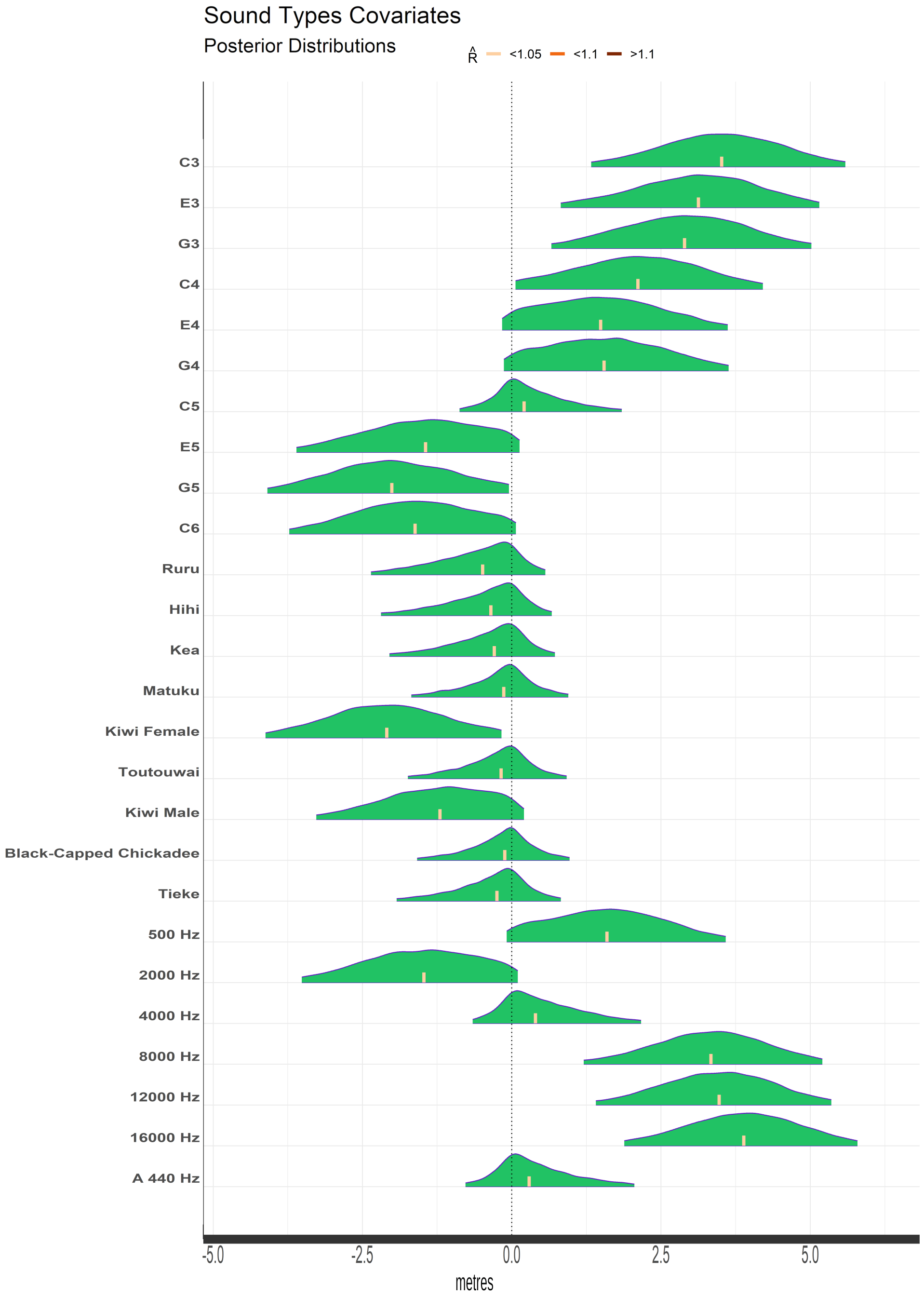


Figure 3.8: Posterior distributions for different sound type covariates. Extremely high and low pure tones are mostly associated with larger Euclidean distances between known and estimated locations whereas most avian sounds are average, except for the Kiwi female, which were associated with smaller errors

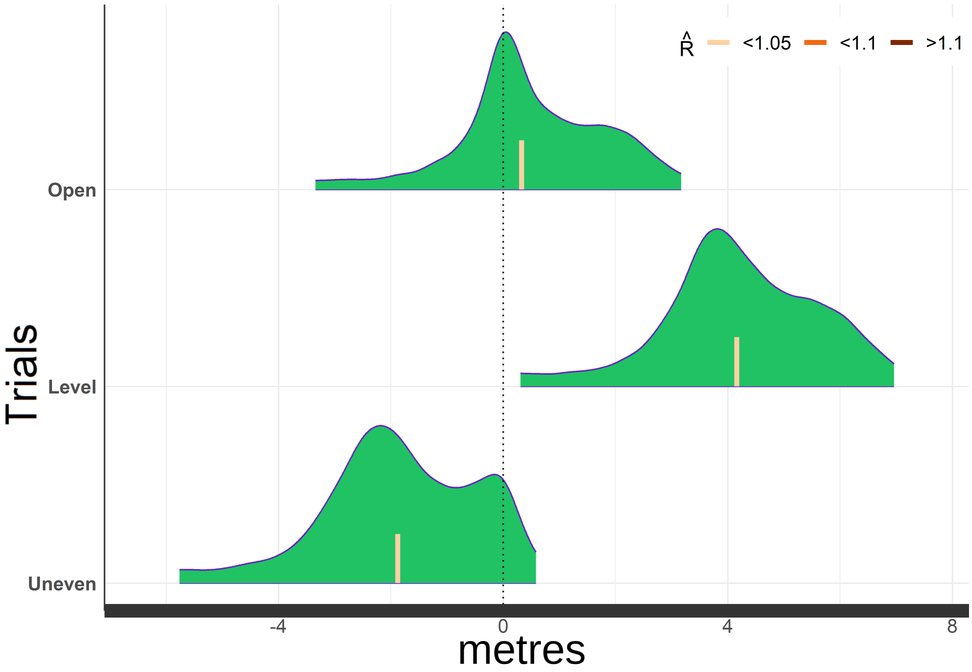


Figure 3.9: Posterior distributions for different trials' covariates. Estimates made during the level trial are mostly associated with slightly larger Euclidean distances between known and estimated locations

Both extremely low (from C3 to G4 and 500 Hz) and high pure tones (4,000, 8,000, and 16,000 Hz) covariates were consistently associated with larger errors while those of some pure tones in the middle range (C5, E5, G5, and 1,000 HZ) were consistently associated with smaller errors (Figure 3.9). Among the avian vocalisations the Kiwi covariates, especially the female, were more consistently associated with smaller errors when compared to other bird sounds.

The quantitative side of the analysis further puts these covariates into perspective, as different covariates have stronger effects on the average Euclidan distance modelled: for instance, employing ARUs configuration #4 is associated with errors mostly ∼ 11m (see Table 3.3, 50% Quantile column) smaller than average whilst trying to estimate the sound source location in a level forested area was mostly associated with errors up to 7 m (see Table 3.3, 97.5% Quantile column) larger than average. Finally, most posterior distributions of the hierarchically modelled natural logarithms of the standard deviations were centred around zero and comprised very small numbers with three (normal likelihood, ARU#2, and ARU#4) between 2 and 3. Given the impractical size of the picture we include it as supplementary material. Following the results from the statistical models, Figure 3.11 and Table II.ii illustrate the difference between the energy–based and TDOA localisations using the data from configuration #4. Finally, the simple comparison model, which included only a global offset and a energy/TDOA categorical covariate, (figure 3.12),table III) confirmed this by showing clear differences between the different levels of the methods' covariate.

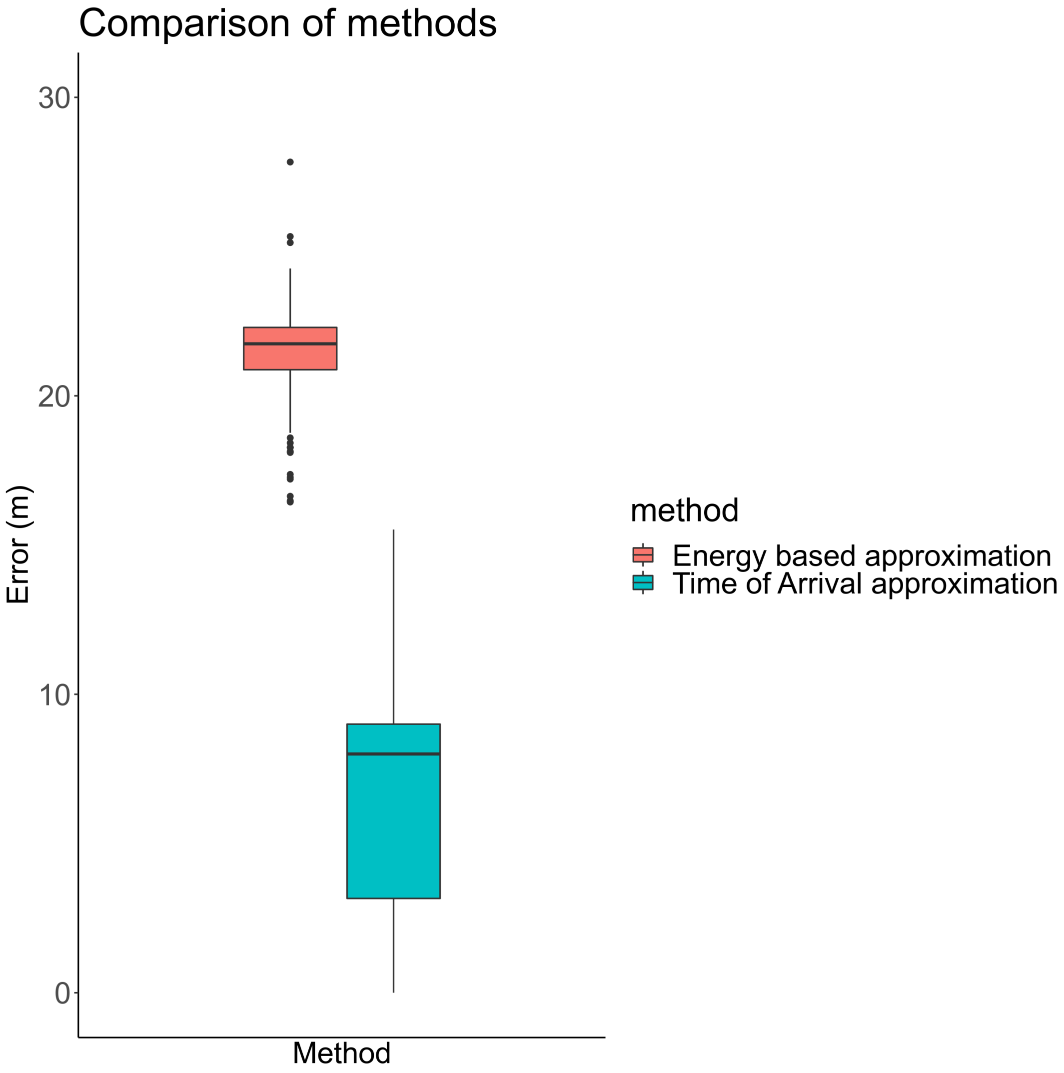


Figure 3.10: Distribution of the Euclidean distances (errors) obtained by localising sounds from one location during the *open* trial using the energy–based method with all the five ARUs (left) and using TDOA with the three microarrays (right).

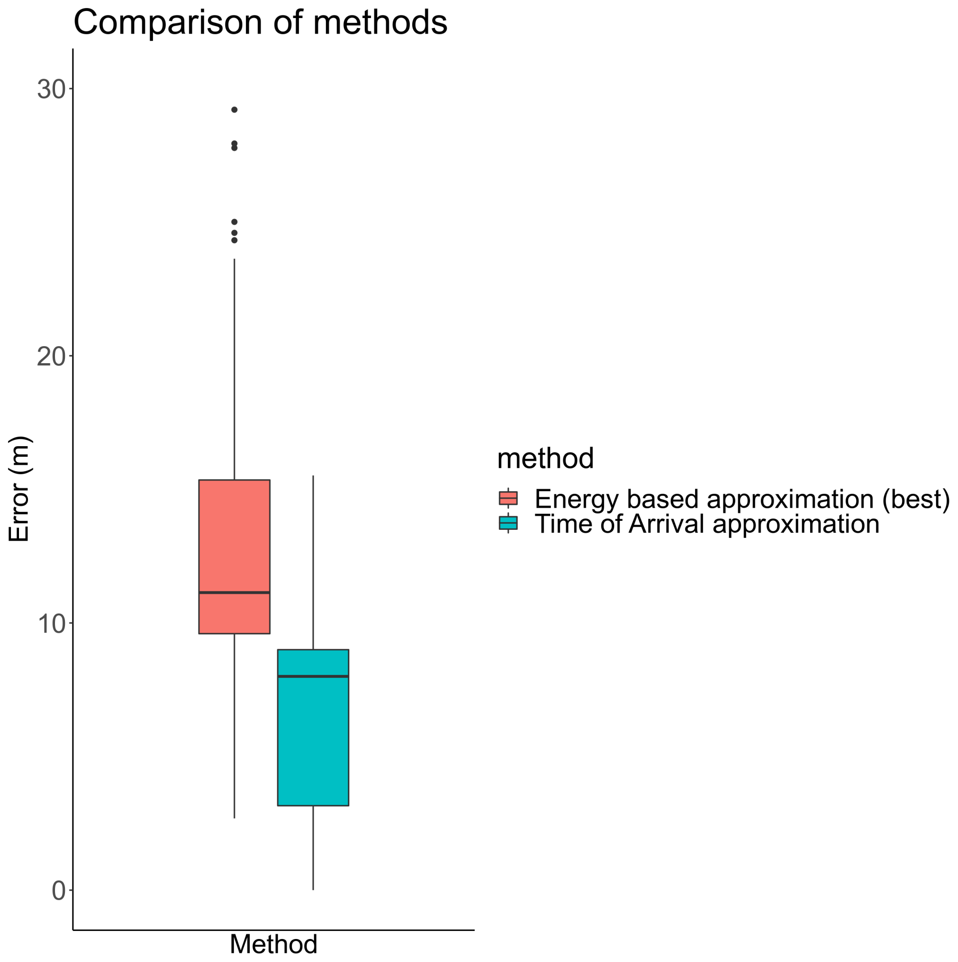


Figure 3.11: Distribution of the Euclidean distances (errors) obtained by localising sounds from one location during the *open* trial using the energy–based method with the ARU configuration #4 (left) and using TDOA with the three microarrays (right). Note that the gap between these is noticably smaller than when using all the ARUs (Figure 3.10).

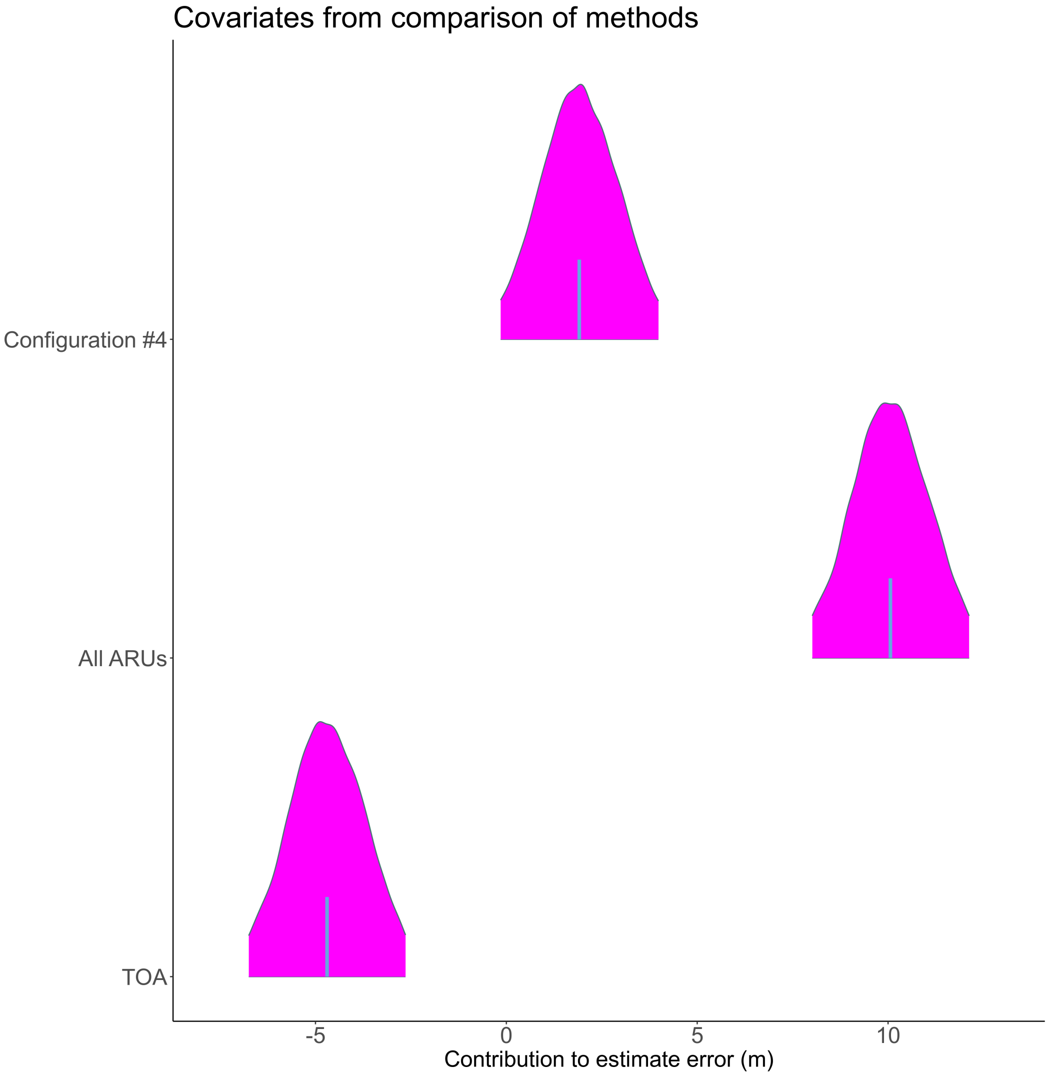


Figure 3.12: Posterior distributions from the simple methods comparison model. See how the covariate from configuration #4 has an intermediate contrubution to error between the TDOA method and the energy–based method employing all ARUs.

3.4 Discussion and Conclusions

This experiment aimed at (1) evaluating whether energy–based sound source localisation could be applicable to PAM using unsynchronised single microphone ARUs, (2) investigating whether uncontrollable factors, such as height of vocalisation, vocalising direction, weather variables, and environment (vegetation, slope) would affect sound source localisation when using non self–synchronised devices, and (3) providing information about whether using different configurations of ARUs would result in more consistent sound source localisations. We found that sound source localisation with energy–based methods is feasible with errors closer to those resulting from TDOA estimates when applying knowledge of which configuration worked best. Although some uncontrollable factors, such as environment (e.g., level forested area), can have an impact on the sound source location, the use of an appropriate configuration of ARUs can potentially make up for such drawbacks. Both when comparing energy–based to TDOA, and among different ARU configurations over different trials, ARUs configuration #4 (four ARUs on the vertices of the cross–shaped experimental field) definitely resulted in the smaller overall discrepancies between real broadcasting locations versus estimates, emerging as the most reliable configuration regardless of the broadcast sound. The fact that configuration #5 performed worse than configuration #4 may be due to sound clipping due to high amplitude at the central ARU. This phenomenon would have had the central ARU relaying higher energies and, in turn, smaller estimated distances which might have led to bigger errors. The differences in localisation error among different forested sites may be more related to vegetation density rather than wind as the latter was quite consistently absent in both forested environments whilst the former was higher in the level site. In support of this, a previous sound propagation study taking place in the same general sites, showed that birdsong transmission is better in forests than open environments and that reverberation due to reflective surfaces can mask frequencies higher than 8 KHz Priyadarshani et al. (2018a). In our study the higher pure tones (8, 12, and 16 KHz) weres harder to localise than the lower ones.

It is positive not to have found any consistent effects of height/angle configurations and directions of broadcast, as the uncontrollable nature of these factors would have been a challenge for developing sound source locating systems. We however encourage any researcher interested in employing energy–based localisation with a target species to trial the method beforehand. They could do so by running a smaller and simplified version of this experiment with two reference sounds (e.g. one sound we found easier and one we found harder to localise) and vocalisations from their target system.

There are still gaps in both field and analytic methodologies. To fill these gaps, any existing methods need to be tested with actual birds and paired sampling (involving multiple sampling methods at the same time). Experimental testing will strenghten the confidence we can have in PAM methods using energy as a proxy for distance while TDOA methods become more widely applicable. We believe that passive acoustic monitoring can and will be a valid alternative to account for animal presence and abundance, if objective and standardised methods are employed.

Chapter 4

Experimental test of acoustic playback as a monitoring enhancing tool in North Island Brown Kiwi (*Apteryx mantelli*, Bartlett)

Figure 4.1: Virginia and Cristabel preparing for Playback.

Abstract

Bioacoustic and ecoacoustic monitoring surveys in terrestrial contexts have gained popularity as they are both cost–effective and produce potentially perpetual data records. However, animals vocalise at different rates depending on several factors, including geographical, seasonal and social ones, potentially introducing bias in abundance estimates.

Acoustic playback is a widespread technique that has proven effective as a tool to standardise response in acoustic monitoring surveys in different taxa and could be of help when working with territorial, rare and elusive species. For the technique to work, individuals of the target species need to respond to playback in a predictable way. The Aotearoa New Zealand iconic North Island Brown Kiwi (*Apteryx mantelli*), a nocturnal and cryptic flightless bird, is routinely monitored through call surveys that, due to the variability of this species' vocal activity, take place over several nights. Acoustic playback has been previously employed with North Island Brown Kiwi to solicit vocalisations within acoustic surveys, although this is not done in national censuses and a formal test of responsiveness to acoustic playback has not been reported.

We aimed to evaluate the potential of playback as an enhancement tool for monitoring, and to advance our understanding of this species' behaviour. We have broadcast vocalisations of strangers to a high density population of North Island Brown Kiwi within and outside the breeding season to account for possible variation due to season. We found that North Island Brown Kiwi did not respond consistently to playback, and that the number of vocalisations per unit of time was instead related to season and external factors. Specifically, the animals tended to vocalise more at higher temperatures and relative humidities, but less in more illuminated or windier nights. In addition, we found an indication that, regardless of the season, the time between vocalisations tended to increase with the progressing of the playback. Our results suggest that either North Island Brown Kiwi are able to very quickly learn that playback calls are not real conspecifics, or that they habituate and either way they stop responding. Our findings discourage the use of playback for monitoring purposes in North Island Brown Kiwi. We suggest using carefully designed experiments of neighbour/stranger broadcasts, testing individual responses to playback, and using broadcasts of synthetic stimuli, to disentangle the effects of habituation versus other types of learning.

4.1 Introduction

Many animals produce sounds that can provide information to answer a wide range of research questions (Blumstein et al., 2011). Among these, acoustic surveys can represent an efficient (Williams et al., 2018) and non–invasive way of monitoring terrestrial animals (Mossman and Weir, 2005; Rosenstock et al., 2002; Sugai et al., 2019, see also section 1.1 of this thesis). These take advantage of animal vocal activity to ascertain presence of a certain species, and can also provide information about animal density in a certain area when multiplying factors to translate acoustic cue numbers to animal numbers are available (Marques et al., 2013).

Within acoustic surveys, acoustic playback has been employed to standardise detectability for several decades (e.g. Gibbs and Melvin, 1997; Johnson and Dinsmore, 1986); this technique is used in a wide variety of animal groups (Deecke, 2006; Gager, 2019; Narins, 2018, chapter 2), and has been shown to be a viable tool to improve surveys' efficiency with a variety of taxa (e.g., Ferreira et al., 2021; Hart, 2020; Okahisa et al., 2016; Schlaepfer et al., 2021; Wynia et al., 2019). Chapter 2 further provides indications on avian perception of sounds and on how birds may thus respond differently to playback (see page 20).

The iconic North Island Brown Kiwi (NIBK — *Apteryx mantelli*, Bartlett) are flightless and nocturnal birds with very distinctive vocalisations (Corfield et al., 2008), which further differ between sexes (see Figure 4.2). According to official reports high density populations produce tens of vocalisations per hour, whereas very few are heard at low densities, if any (Craig, 2018, 2021). NIBK vocalisations have been deemed relevant for monitoring purposes (Keast et al., 2010; Pierce and Westbrooke, 2003) and specialised acoustic surveys known as Kiwi Call Counts (KCC, Glossary table VII.i, page XXIII) are routinely used to monitor NIBK populations at distinct locations around Aotearoa New Zealand. The KCC scheme involves listening for Kiwi vocalisations for two hours at a time. The scheme requires 48–50 hours of listening from each site (station). Since each station comprises different numbers of listening locations, stations with fewer listening locations will need more listening nights to reache the total number of required listening hours. This many hours are required because there is high variability in vocal activity of Kiwi species, even in seemingly identical conditions (McLennan, 1992, section 3.2.3), hence multiple nights can be necessary to obtain a reliable number (Robertson et al., 2017).

KCC are used to identify changes in call rate in target populations as approximates of populations' and behavioural trends and, originally, McLennan (1992) suggested that using playback every 10–12 minutes during the second hour would help soliciting vocalisations from birds that had not vocalised thus far (McLennan, 1992, section 3.3). The use of playback is normally allowed during the second hour of listening according to Robertson et al. (2017), but only to check for presence of NIBK in sparsely populated areas. It is prohibited to use playback during the Na-

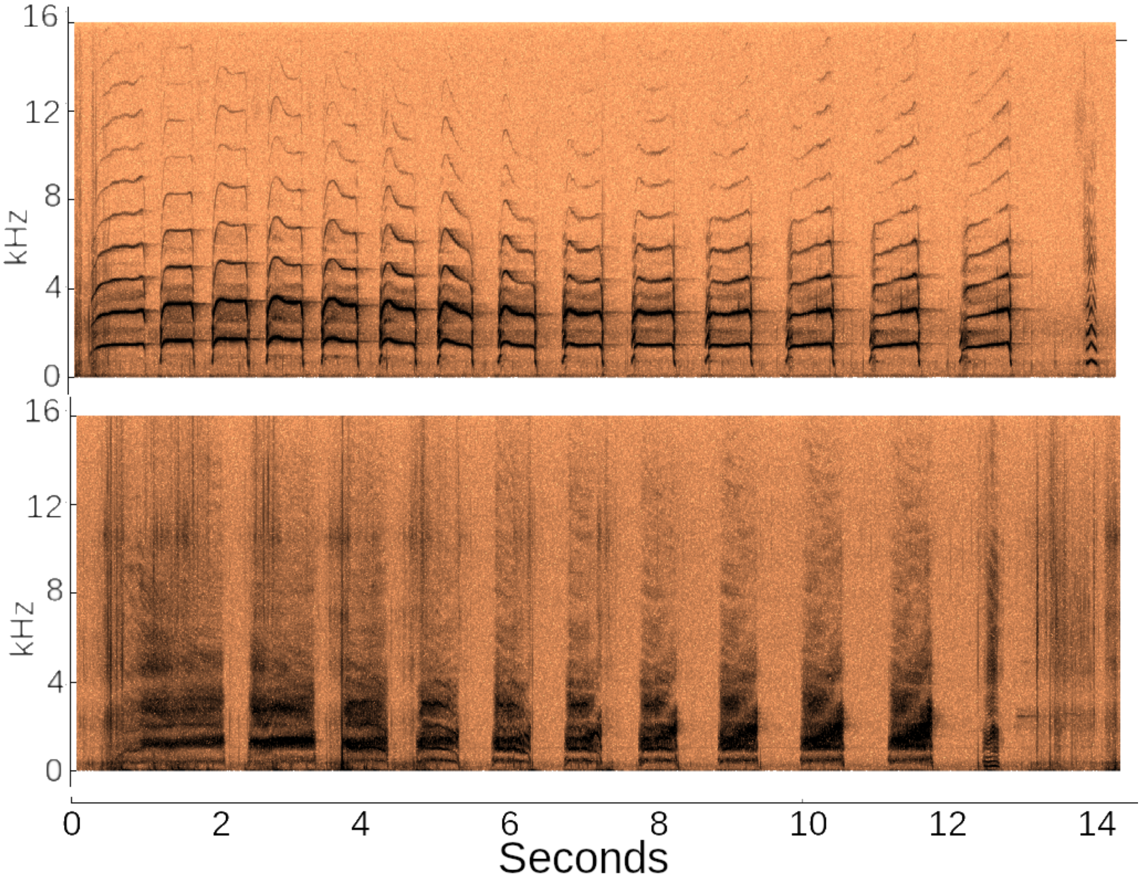


Figure 4.2: Example spectrograms of male (top) and female (bottom) NIBK.

tionwide KCC Scheme (Robertson et al., 2017, section 7.2.1), which usually comprises one shared window of listening during May (this included a backup window during June in 2020, see Craig, 2021) during which listeners throughout the country perform acoustic surveys in predetermined locations (Craig, 2021; Robertson et al., 2017).

These views in respect to KCC and the acoustic playback technique may be due to the lack of a formal test of acoustic playback with NIBK. There have been formal tests of acoustic playback with congeneric taxa. In the work of Digby (2013), the congeneric Kiwi Pukupuku (Little Spotted Kiwi, *Apteryx owenii*, Gould) also shows a diverse range of vocal behaviours, with males vocalising up to three times as much as females, and differences in vocal activity at different times of the year and of night (Digby et al., 2014b). Nonetheless, Kiwi Pukupuku were shown to be quite territorial and responsive to playback, often approaching the speaker and vocalising in response to conspecific vocalisations (Digby et al., 2014a).

Here we present the first formal test of NIBK responsiveness to playback for monitoring purposes. By testing whether playback could affect the vocal behaviour of NIBK we aimed to (1) provide information to improve monitoring practices, and (2) improve our understanding of NIBK vocal behaviour.

4.2 Materials and Methods

4.2.1 Context

We conducted this study on Ponui (Chamberlin's) Island (175.14860, −36.90422 : 175.21860,-36.82961 – Coordinate Reference System WGS84 for all coordinates herein; Figure 4.3), in the Hauraki Gulf, Auckland region, Aotearoa New Zealand.

Fourteen NIBK were translocated to Ponui in 1964 (Miles and Castro, 2000), and the island now hosts a thriving high density population of NIBK (Cunningham et al., 2007). Ponui Island hosts the Kiwi Research Group led by Isabel Castro, and NIBK (as well as other species) have been part of several researches since 2004. For this study, we worked in one of the remnant kauri (*Agathis australis*) forest patches in Southern Ponui known as Kauri Bush (175.177635, −36.887127 : 175.182262, −36.882096; Figure 4.3). Approximately 50 birds have been equipped with very high frequency (VHF) radio transmitters that are monitored at least monthly. Having record of the history of these animals, including morphological measurements, usual resting areas and nesting sites (Dixon, 2015; Jamieson et al., 2016), and social relationships (Undin et al., 2021) among others (Cunningham and Castro, 2011), made Ponui Island an ideal site to conduct this study.

As indicated by McLennan (1992), Colbourne and Digby (2016), and Robertson et al. (2017), we started performing playback no earlier than 30 minutes after sunset, yet within the first two hour of darkness, both outside (20-23 March 2021, sunset ranged from 19:31 to 19:26, — `https://www.timeanddate.com/`, and playback started at 21:00) and within (30 July - 1 August 2021, sunset ranged from 17:32 to 17:34, — `https://www.timeanddate.com/`, and playback started at 18:30) NIBK breeding season (Robertson et al., 2017; Undin, 2021, Chapter 2.3).

Average temperatures were 17.6° C (standard deviation, $\sigma = 1.28$) and 12.5° C ($\sigma = 1.59$) during the first (March) and second (July) data collection period respectively. Average relative humidities were 70% ($\sigma = 3.25$) and 89% ($\sigma = 8.81$) respectively. Average prevailing wind speeds were 11 ($\sigma = 2.55$) and 10 ($\sigma = 2.89$) Km/h respectively. The average percentage of moon illumination was 52% ($\sigma = 0.12$) and 67% ($\sigma = 0.08$) respectively. We decided to account for lunar illumination rather than phase following the results from Colbourne and Digby (2016).

Additional information regarding Ponui Island, its NIBK population, vegetation, and related studies can be found in e.g., the following theses: Bansal (2020); Dixon (2015); Strang (2018); Vattiato (2021); Vieco Gálvez (2019).

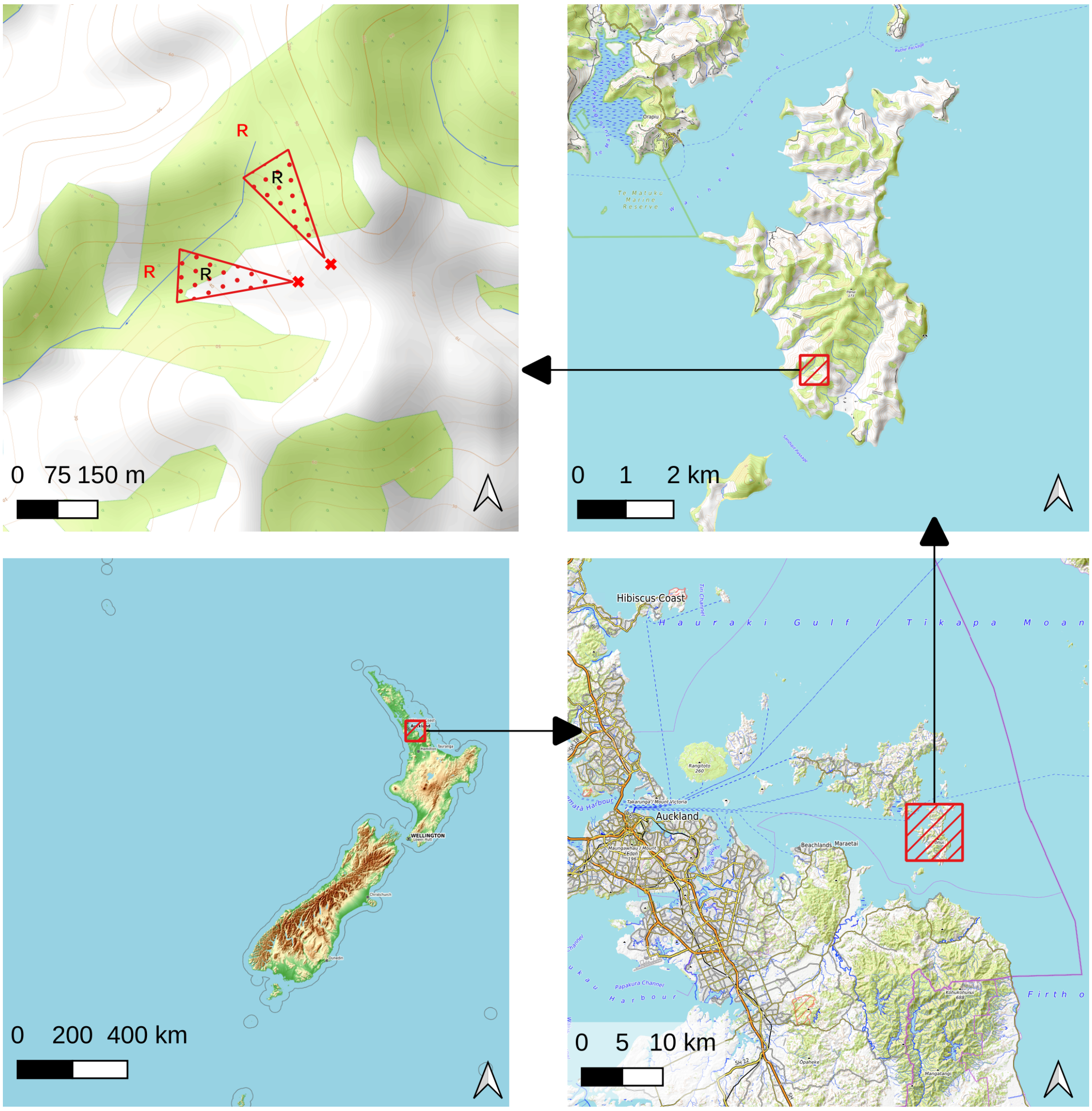


Figure 4.3: Location of the study site relative to Aotearoa New Zealand (bottom left), the greater Auckland region (bottom right), and Ponui Island (top right). The top left quadrant shows the playback setup with letters “R” representing the locations of the ARUs (red at 300 m and black at 200 m from broadcasting stations respectively), and “x” mark the spots from where we broadcast towards the two ARUs as represented by the dotted triangles. The size of the triangles is not prportionate to the sound propagating distance of the loudspeakers, rather an indication of direction of broadcast.

Vocal Behaviour And Territoriality Of Ponui Island North Island Brown Kiwi

NIBK are flightless and nocturnal ground dwelling cryptic birds belonging to the only Apterigiformes genus, *Apteryx*. Like many endemic animal species in Aotearoa New Zealand (e.g., the Toutouwai, *Petroica longipes*, see Boulton et al., 2008), NIBK distribution is to some extent consequential of habitat fragmentation (Potter, 1990; Undin et al., 2021) and interactions with relatively recently introduced mammals (Basse et al., 1999; McLennan et al., 1996; Parlato et al., 2015). Over the past thirty years, several studies reported NIBK as relatively territorial (McLennan et al., 1987; Taborsky and Taborsky, 1992, 1995, 1999), but other studies have shown that territoriality may vary with population density (Dixon, 2015; Potter, 1989; Ziesemann, 2011).

On Ponui Island, marked birds hold overlapping home ranges of between five and six hectares (Ziesemann, 2011, Chapter 3). Birds roost in burrows or on the surface in forest patches within their range and the choice of roost site is dependent on proximity to foraging areas (Dixon, 2015, Chapter 4). At night NIBK move from the roost sites to foraging sites and are mostly found in forested and scrub areas (Dixon, 2015, Chapter 4, figure 3.6). Vocalisations of NIBK males and females NIBK are quite distinct; Corfield et al. (2008) describes males as producing long–travelling whistles which could be grouped into phrases, and with fundamental frequencies around $1.5kHz$ (and overtones up to approximately $13kHz$), whereas female produce hoarser and lower (fundamental frequency around 0.1 kHz) vocalisations with intensity concentrated in two formants. Duets are relatively common, and can be initiated by the male or the female; Corfield et al. (2008) suggested that duets can be territorial calls. Working in mainland Aotearoa New Zealand, Miles et al. (1997) found that call rate did not match population density in two low–density neighbouring NIBK communities, and that communities with relatively higher density were more responsive to acoustic playback (Miles et al., 1997).

4.2.2 Experimental design, stimuli and devices

We planned for four nights of non interactive sessions from two broadcasting points (to minimise habituation) in each season, during which we performed playback for one hour. Each hour included a total of six stimuli, one every ten minutes, of which two were control sounds (Brownian Noise and Ruru — *Ninox novaeseelandiae*, Gmelin — vocalisations) and four were NIBK vocalisations. The Brownian noise was generated in Audacity (Audacity Team, 2018) at full amplitude (i.e., setting Amplitude to 1). We did not use the same vocalisation twice as a stimulus (Kroodsma, 1989b; McGregor et al., 1992) and thus employed a total of thirty-two different NIBK stimuli (twelve male vocalisations, twelve female vocalisations and eight duets) as well as eight Ruru vocalisations. For duets, we always presented the same structure (male initiates the duet and female follows); this is because there is not enough knowledge of whether there is meaning in the order of the duet. All the stimuli are available at `https://avianz.net/playback`.

We selected and extracted high (> 10 dB) signal to noise ratio vocalisations that we had previously recorded at several locations throughout the North Island using AviaNZ (Marsland et al., 2019). All the vocalisations were recorded using New Zealand Department of Conservation Te Papa Atawhai (hereafter DOC) AR4 ARUs (sampling rate 32 KHz, bit depth 16, frequency response chart in figure II.iii). We then normalised the peak amplitude of all tracks to be the same (0 dB) and built the playback tracks using Audacity. All the tracks were exported from Audacity and loaded to the louspeaker in Waveform Audio File format (WAV).

In order to properly test the use of playback, we broadcast from two locations (175.181534, -36.886633 – 63 m asl and 175.181099, -36.886982 – 58 m asl, figure 4.3) in the middle of the pasture overlooking Kauri Bush to the North and Horse's Mane to the South. These are easily accessible sites and are centrally located for listening to the vocalisations coming from the sourrounding forest patches. We deployed a total of four ARUs per data collection period. One ARU was located at approximately 300 m from each broadcasting point (which is the detection limit we derived for our ARUs, see page 51), and a second one at about 200 m from the speaker (on the same imaginary line) also to make sure to always detect the stimuli in at least one ARUs and the calls of all possible responding birds. We programmed the ARUs to start recording thirty minutes before the experiment started and stop thirty minutes after its end. This both served as a safety net in case we could not start the experiment as scheduled, and allowed us to potentially analyse the vocal activity before, during, and after the experiment if deemed necessary. We randomised the order of broadcasting from the two loudspeakers, with the only condition that each of them emitted three stimuli per night (e.g., speaker one could play stimuli at 0, 20 and 30 minutes from the start and speaker two would play the other three at 10, 40, and 50 minutes from the start). Because of this, we produced two different one hour long tracks per night, one for each speaker, collectively comprising a stimulus (control or NIBK) every ten minutes, yet with only three stimuli per track. We produced a total of eight pairs of tracks.

We here refer to each 10 minutes period after a broadcast during which we recorded NIBK vocalisations as a "time slot" (i.e., each playback session comprised six time slots). This is important to understand the different aspects of vocal activity that we consider to investigate possible effects of playback on the Ponui Island NIBK. We chose 10 minutes time slots following McLennan (1992) and specifically looked at:

1. TOTAL ACTIVITY: this is total number of vocalisations over the 10 minutes; we looked at this to grossly compare vocal activity following different simuli (NIBK versus control).

2. RATE OF ACTIVITY: this is time between vocalisations for each time slot; we looked at this to evince whether vocalsiations would be closer to each other after different stimuli.

3. EVENNESS OF ACTIVITY: this is the number of early versus late vocalisations following each stimulus; we looked at this to investigate whether a higher number of vocalisations would concentrate shortly after a broadcast so as to represent an actual response.

Details of how we analysed these aspects of vocal behaviour follow in section 4.2.4.

4.2.3 Playback Protocol

SPL We adjusted the speakers' sound pressure level (SPL) using a reference pure tone track one and a half metres away from a Digitech QM1592 Professional Sound Level Meter (see `https://www.jaycar.com.au/pro-sound-level-meter-with-calibrator/p/QM1592` or `https://www.techbrands.com`), in fast response mode within the 30–130 dBA range, which we calibrated before usage according to the user manual. We noted the sound pressure level of the reference pure tone (440 Hz A) at one and a half metres away from the speaker for a NIBK call at natural amplitude, that is male NIBK vocalisations in the 75–79 dBA range, and female ones in the 61–63 dBA range respectively (Castro et al., 2019); we then used the reference sound to adjust the speakers in the field (since this is a pure tone, it maintains a steady SPL and so is easier to adjust to a desired SPL), hence easier to work with in field conditions. We had the SPL meter on a tripod 20 cm high from the ground and facing the loudspeaker.

Setup We deployed the weather station at base camp (175.187383, -36.894460, 0 m asl) and walked to the site, arriving at least ten minutes before starting each night. One to two observers ran each station. The observers placed the speakers on the same spot on the ground each night, with the forward side of the device slightly elevated ($\sim 30^\circ$) to imitate a calling NIBK. Combining previous knowledge of Ponui Island NIBK from Dixon (2015), where transmitterised birds from distinct areas were reported to consistently frequent their roosting areas and our own dog surveys (chapter 6) we could expect more than 18 birds to be within earshot of the loudspeakers.

Protocol and Observers' Behaviour We turned the loudspeakers on at 21:00 or 18:30 during the March and July data collection periods respectively. The observer(s) then moved back into the forest patch a few metres (see figure 4.3) to monitor any particular event (e.g., a bird approaching the loudspeaker), and waited in silence until the end of the session. At the end of the session the observers regrouped and walked back to base camp and turned off the weather station.

4.2.4 Computer work

Data extraction

We processed the audio files from the ARUs using AviaNZ (Marsland et al., 2019, Version 3.2), which uses a combination of filters (wavelet package decomposition and convolutional neural network) to segment vocalisations of interest from the files in a given folder. We subsequently reviewed the results to correct for both false positives and false negatives. We also manually changed the labels that the program had assigned to the stimuli to custom *ad hoc* stimuli labels. We visually compared the spectrograms of the stimuli from both the original track and the ARU recordings, as well as the time of recording, to ensure not to mistake them for real birds' vocalisations. We joined the AviaNZ–generated spreadsheets from each pair of ARU and removed duplicate detections (when unsure we further checked aurally and visually that they were indeed duplicates and not concurrent vocalisations). We finally combined the spreasheets generated through AviaNZ, the moonlight data downloaded from `https://timendate.com` and the comma separate value files (csv) from the weather station, in the three csv files that we used for the statistical analyses.

Statistical analyses

We devised three models to investigate the vocal behaviour of NIBK in relation to acoustic playback for monitoring purposes. We modelled:

1. the number of vocalisations detected per stimulus' time slot (ten minutes) with a Poisson likelihood (with log parameterisation / link function).
2. the times between vocalisations during each stimulus' time slot with an exponential likelihood (logit link function).
3. the number of vocalisations within the first three minutes of each time slot related to the total number of vocalisations in that time slot with a binomial likelihood (with logit parameterisation / link function).

We chose these three models to thoroughly explore NIBK vocal activity after playback. In the Poisson model, (1) the sheer number of vocalisations per time slot would be a gross indicator of higher or lower overall vocal activity. In the exponential model, (2) by looking at the waiting times between vocalisations, we would have an indication of whether strangers' vocalisations would solicit more clustered vocalisations than expected by chance, and with the binomial model, (3) we would get a finer–scale look at the distribution of the vocalisations per time slot. A higher probability of success would indicate higher numbers of vocalisations occurring shortly after broadcast rather than in the rest of time slot than expected by chance. We used the number of vocalisations in the first three minutes of each time slot as *success* in our binomial model to have a clearer understanding on whether the probability of detecting NIBK vocalisations immediately after broadcast was any different from the rest of the time slot. Although different choices could have also been acceptable, for instance using the vocalisations detected during half of the time slot as *success*, we deemed three minutes a short enough time for pointing at genuinely solicited vocalisations, yet not so short as to limit the detection of potentially solicited duets or even more complex chains of vocalisations.

Here we worked within a Bayesian framework for two reasons. First, to be able to discern positive versus negative changes in vocal activity. By specifying zero–centred prior distributions (before the simulations) for all the covariates, the posterior distributions (after the simulations) that fell completely above zero consistently contributed to higher rates of vocalisations per time slot (meaning more vocalisations in the Poisson model and shorter waiting times between vocalisations for the exponential model), and higher probability of vocalisations in the first three minutes (binomial). Conversely, the posterior distributions that fell completely below zero consistently contributed to lower rates, and those that remained somewhat centred around zero did not consistently contibuted to either higher or lower rates.

We ran all the models in Stan (Carpenter et al., 2017) through the `rstan` (Stan Development Team et al., 2016) package in `R` environment (`R` Core Team, 2013). All the models and data are available at `http://avianz.net/playback`.

We modelled the number of vocalisations detected during the time slot following each broadcast [θ in Equation (4.1)] and the number of vocalisations within the first three minutes of each time slot against the time slot's total (binomial), including a global offset (β^0), a series of categorical covariates according to sex (β^s, three levels: male, female, total number), data collection period (β^t, two levels: non-breeding, breeding), stimulus (β^{st}, five levels: female, male, noise, Ruru, duet), area (β^a, three levels: upper gully, swamp, total number), day (β^d, four levels: first, second, third, forth), and a series of weather related covariates — Equation (4.2) — which multiplied the Z-standardised ($z = \frac{x-\mu}{\sigma}$) weather values; these included moonlight (ml), temperature (te), prevailing wind speed (ws), prevailing wind direction (wd), relative humidity (rh), and atmospheric pressure (ap). When modelling the times between vocalisations we applied the same linear expression, with the only difference that as the times belonged to a single vocalisation the levels of the area and sex covariate were reduced to two (i.e., removing the *sum* level). The following equations exemplify the structure of the models using the Poisson model as an example; the only difference between these and the binomial and exponential one would be in the likelihood and link function.

$$\theta_q \sim Poisson(\lambda_q) \tag{4.1}$$

$$\begin{aligned} log(\lambda_q) = &\beta^0 + \\ &\beta^s_{s_q} + \beta t_{t_q} + \beta^{st}_{st_q} + \beta^a_{a_q} + \beta^d_{d_q} + \\ &\beta^{ml} \times ml_q + \beta^{te} \times te_q + \beta^{ws} \times ws_q + \beta^{wd} \times wd_q + \beta^{rh} \times rh_q + \beta^{ap} \times ap_q \end{aligned} \tag{4.2}$$

$$where\, q \in [1..N]$$

Each model ran for 60,000 iterations over three independent Hamiltonian Monte Carlo chains (Carpenter et al., 2017), retaining only one in every three simulated values to prevent autocorrelation. We discarded the first half of the simulated values as burn–in, resulting in 30,000 final draws per variable per model. After testing that weakly informative priors would produce similar results to uninformative priors (yet in tremendously shorter times), we specified weakly informative priors for all covariates across the three models ($Normal(\mu = 0, \sigma = 2)$). In practice, we first ran a few models for a reduced number of iterations (10,000) with hierarchically modelled standard deviations (uninformative priors) on all the random variables and, after noting that all of them would converge to median values no greater than 2, we removed the hierarchical part for the standard deviations and replaced them with 2 (weakly informative prior).

4.3 Results

4.3.1 Vocalisations detected

Although we planned for two sets of four nights of playback (20 to 23 March and 29 July to 1 August 2021), due to external factors we could not land on the island early enough and could only perform three nights of playback during breeding season (July). All the four ARUs (two in the central uphill area of the gully and two in the downhill western swamp, figure 4.3) worked properly during the first (March) data collection period, but one of the recorders in the downhill swamp area failed completely during the second (July) data collection period. Although some birds were also equipped with individual acoustic recorders (see chapter 6), none of them vocalised during this experiment. However, the detection area of the two pairs was intentionally largely overlapping within each pair as failsafe. During the March data collection period, in the central uphill area, one ARU (North) detected 16 vocalisations, the other (East) 13 vocalisations, for a total of 17 unique NIBK vocalisations (without counting the stimuli). Similarly, still during the March data collection period, in the downhill swamp area, one recorder (West), detected 41 vocalisations, the other (South) 35 vocalisations, for a total of 54 unique NIBK vocalisations (without counting the stimuli). Likewise, during the July data collection period, in the central upper area, one

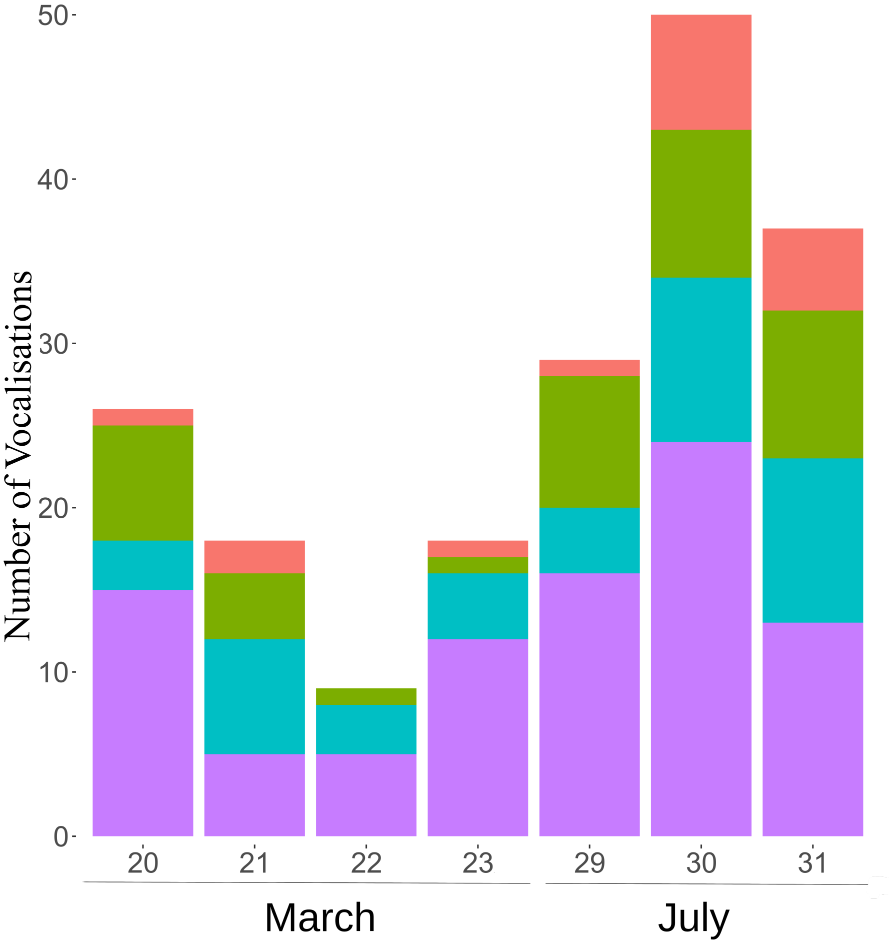


Figure 4.4: Summary of the number of vocalisations detected by the ARUs over the survey. Each bar depicts the total number of vocalisations detected each night fractioned as male vocalsiations from the downhill ARUs (purple), female vocalsiations from the downhill ARUs (blue), male vocalisations from the uphill ARUs (green), and female vocalisations from the uphill ARUs (pink).

ARU (North) detected 38 vocalisations, the other (East) 26 vocalisations, for a total of 39 unique NIBK vocalisations (without counting the stimuli). Conversely, during the July data collection period, in the downhill swamp area, one ARU (West) detected 74 vocalisations, whereas the other (South) failed to detect anything (the ARU worked as per the programmed schedule, but only recorded noise, probably due to a microphone malfunction). Figure 4.4 and Table 4.1 summarise the distribution of the unique vocalisations detected over the seven playback nights.

Table 4.1: Number of vocalisations detected in each area over the survey. This table reports the absolute number of vocalisations detected by each single ARU over the two data collection periods, as well as the number of unique vocalisations not shared between the uphill and downhill ARUs pairs.

Day	Month	Uphill Gully			Downhill Swamp		
		North	East	Unique	South	West	Unique
20	March	8	5	8	13	15	18
21		4	4	6	10	11	12
22		1	1	2	3	5	8
23		2	2	2	9	10	16
Subtotal March		15	12	17	35	39	54
29	July	9	4	9	-	20	20
30		15	10	16	-	32	32
31		14	12	14	-	22	22
Subtotal July		38	26	39	-	77	77
Total		53	38	56	35	113	131

4.3.2 Statistical analyses

All the models performed well, with high number of independent draws (at least 240,000 independent draws per variable, or 80% of the total), no relevant Monte Carlo errors (all smaller than 5% of the posterior standard deviations), and no $\hat{R}$ greater than 1.01 (Carpenter et al., 2017). This is the Gelman & Rubin potential scale reduction statistic, an indicator of the variance explored by each chain compared to the rest of the chains, and should desirably be close to one. See tables IV.i, IV.ii, and IV.iii for detailed accounts.

In the following paragraphs, we report the results for each individual model. We are giving median and 95% posterior high density intervals (hdi) for the covariates that resulted (positively or negatively) consistently associated with the response variables. We are additionally reporting the percentages of the posterior distributions above or below zero of the other covariate, for visualisation purposes.

Number of vocalisations per time slot. The only covariates consistently associated with the response were weather related (figure 4.5). Specifically, the moonlight (median -0.59, 95% hdi from -0.82 to -0.38) and wind speed (median -0.37, 95% hdi from -0.56 to -0.18) covariates were consistently negatively associated with the number of vocalisations per unit of time, whereas temperature (median 0.66, 95% hdi from 0.40 to 0.92) and humidity (median 0.47, 95% hdi from 0.15 to 0.80) were positively associated with the number of vocalisations per unit of time.

The other covariates' posterior distributions resulted in no consistent association (either positive or negative) with the number of vocalisations per unit of time. We however report that most of the data collection period/season covariates posterior distributions resulted on opposite sides of zero, with approximately 75% of the March covariate being on the negative side and 75% of the July one on the positive side, indicating that higher numbers of vocalisations in July may be related to seasonal variation, although (figure 4.6).

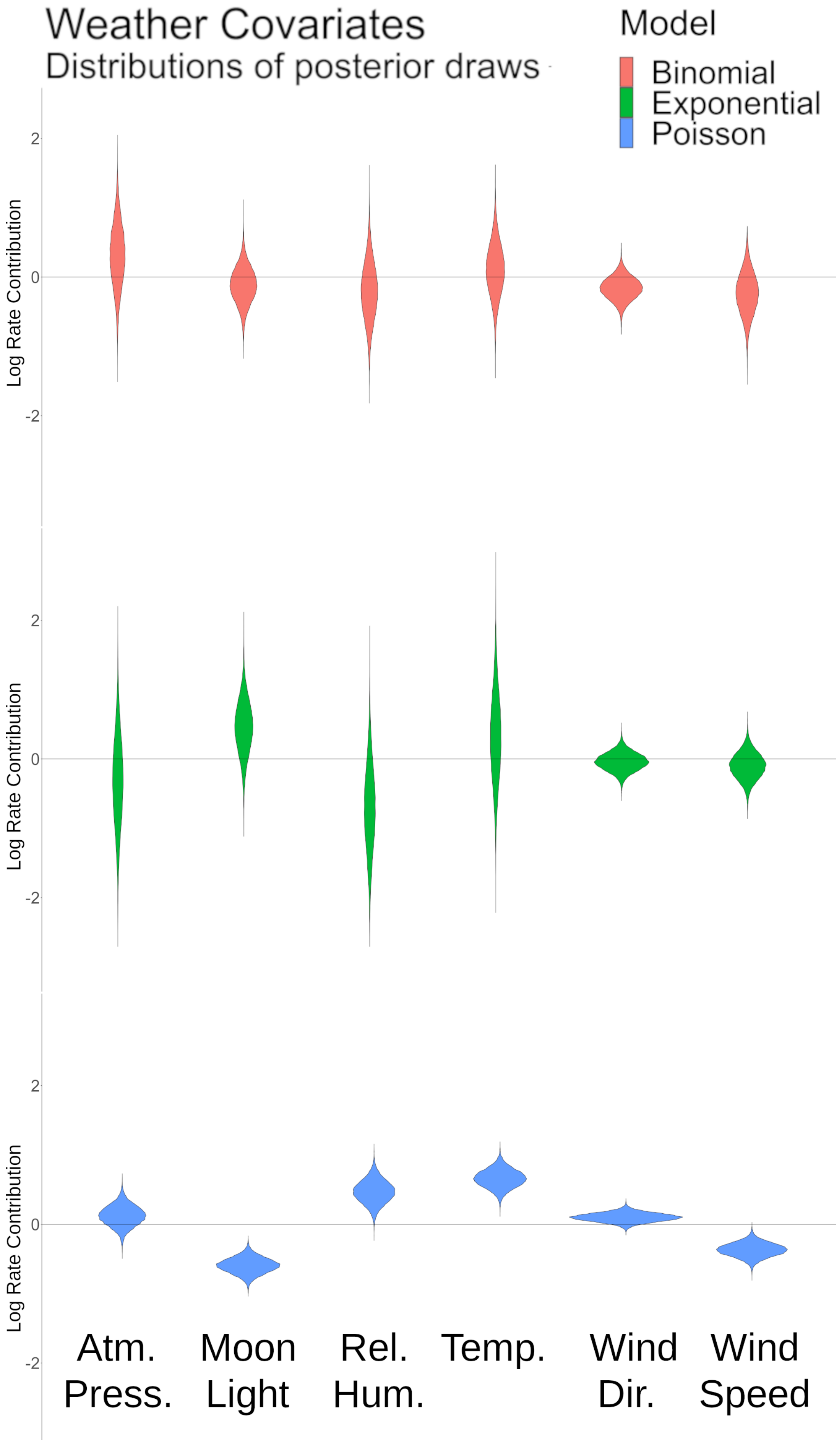


Figure 4.5: Posterior distributions for the weather covariates across the three models. Distributions completely above or below the horizontal zero line can be considered as having a consistent effect on the response variable.

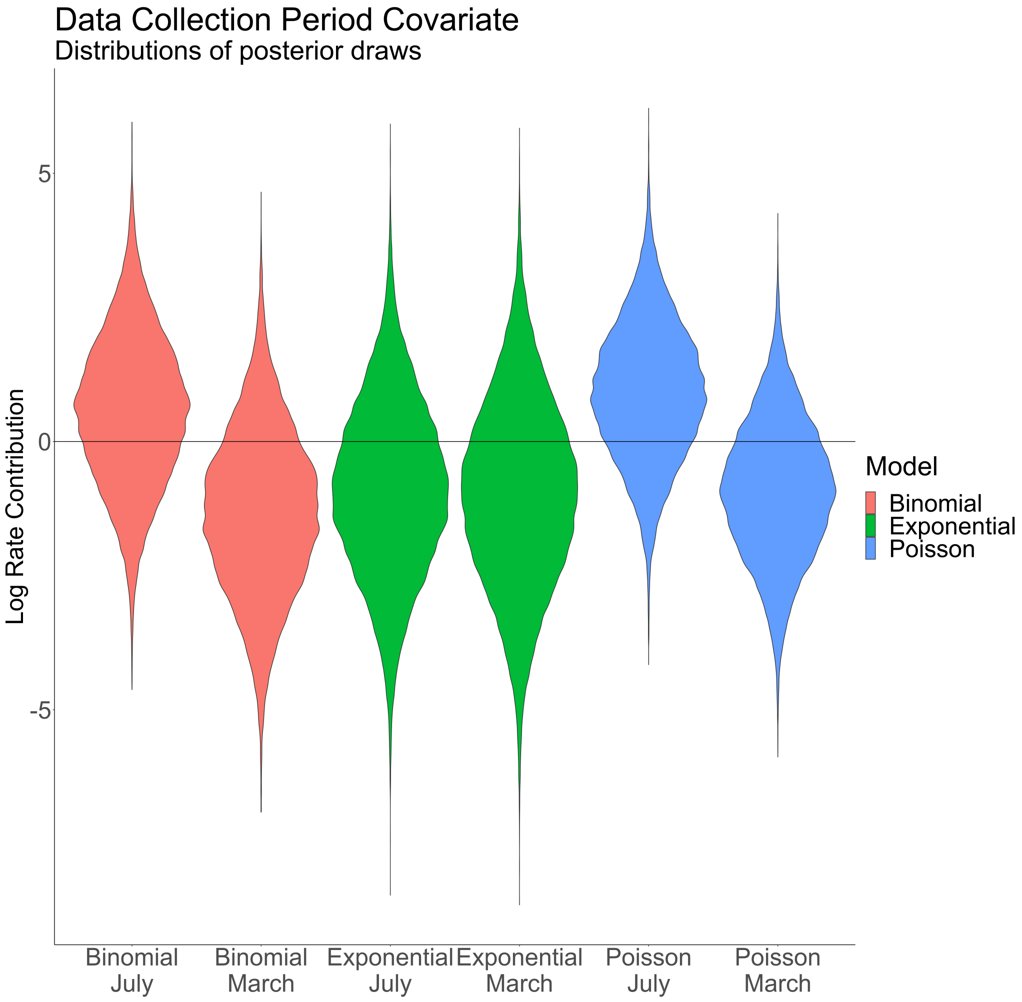


Figure 4.6: Posterior distributions for the data collection period covariates across the three models. Distributions completely above or below the horizontal zero line can be considered as having a consistent effect on the response variable.

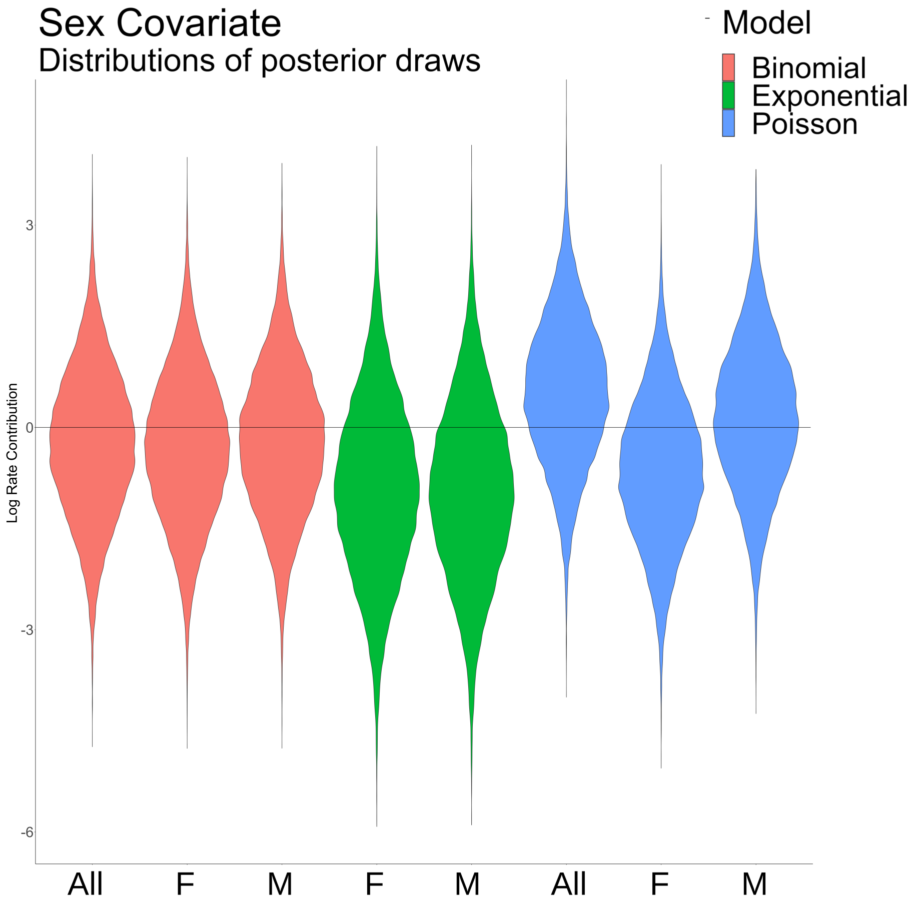


Figure 4.7: Posterior distributions for the sex covariates across the three models. Distributions completely above or below the horizontal zero line can be considered as having a consistent effect on the response variable.

Less distinctively, approximately 75% of the posterior distribution of the female category (sex covariate) was below 0, while that of the male category ended up quite centred around zero, and most (∼ 70%) of the combined sex cateogry covariate posterior distribution was above zero (figure 4.7); although not consistently associated with the response variables this indicates that higher number of vocalisations are likely to be detected from males (as in Miles, 1995).

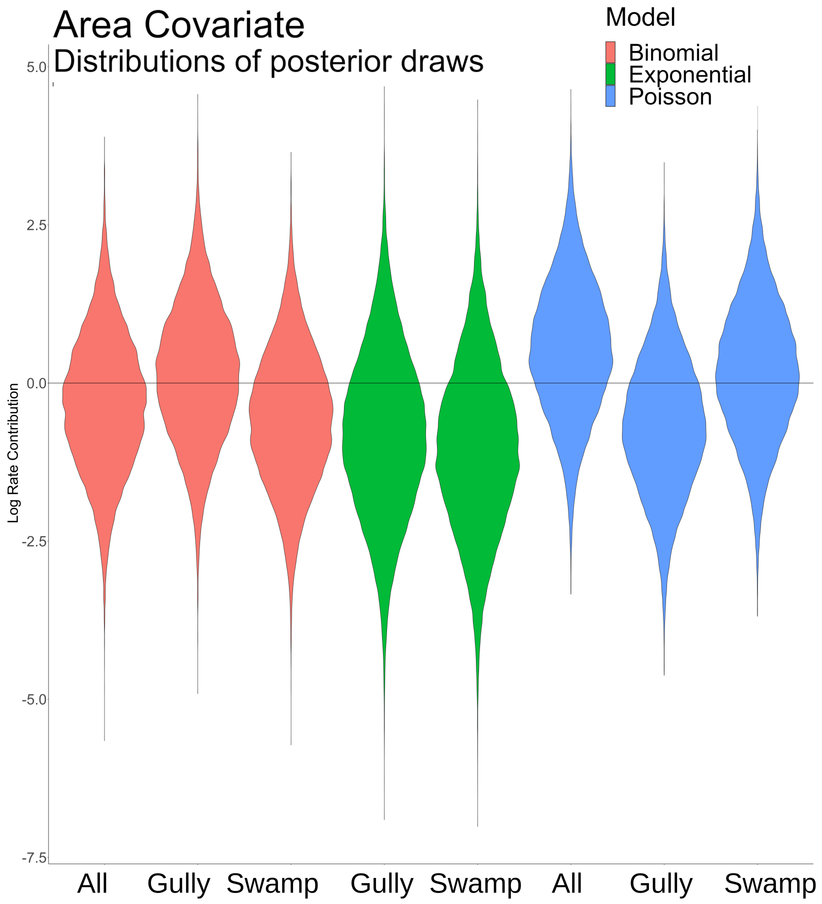


Figure 4.8: Posterior distributions for the area covariates across the three models. Distributions completely above or below the horizontal zero line can be considered as having a consistent effect on the response variable.

The area covariates relayed a similar result, with the upper gully posterior distribution mostly ($\sim$ 70%) below zero, the swamp one centred around zero, and the combined one mostly ($\sim$ 70%) above zero (figure 4.8); again, although not consistently associated with the response variable, this indicates how higher numbers of vocalisations are likely to be detected in the downhill swamp area.

Of the stimuli covariates, $\sim$ 70% of the posterior distribution for the female vocalisation category sat on the positive side, the male stimulus and control sounds' ones were centred around zero and $\sim$ 60% of the duet simulus one was negative; none of these are consistently associated with higher or lower number of vocalisations (figure 4.9), although there seem to be an higher chance of detecting more vocalisations after a female stimulus, which corroborates the results of (Miles et al., 1997). Finally, all the categories (1 to 4) of the day covariate were centred around zero (figure 4.10).

Time between vocalisations per time slot. There were no covariates consistently associated with the response variable. We report that, of posterior distributions of the weather covari-

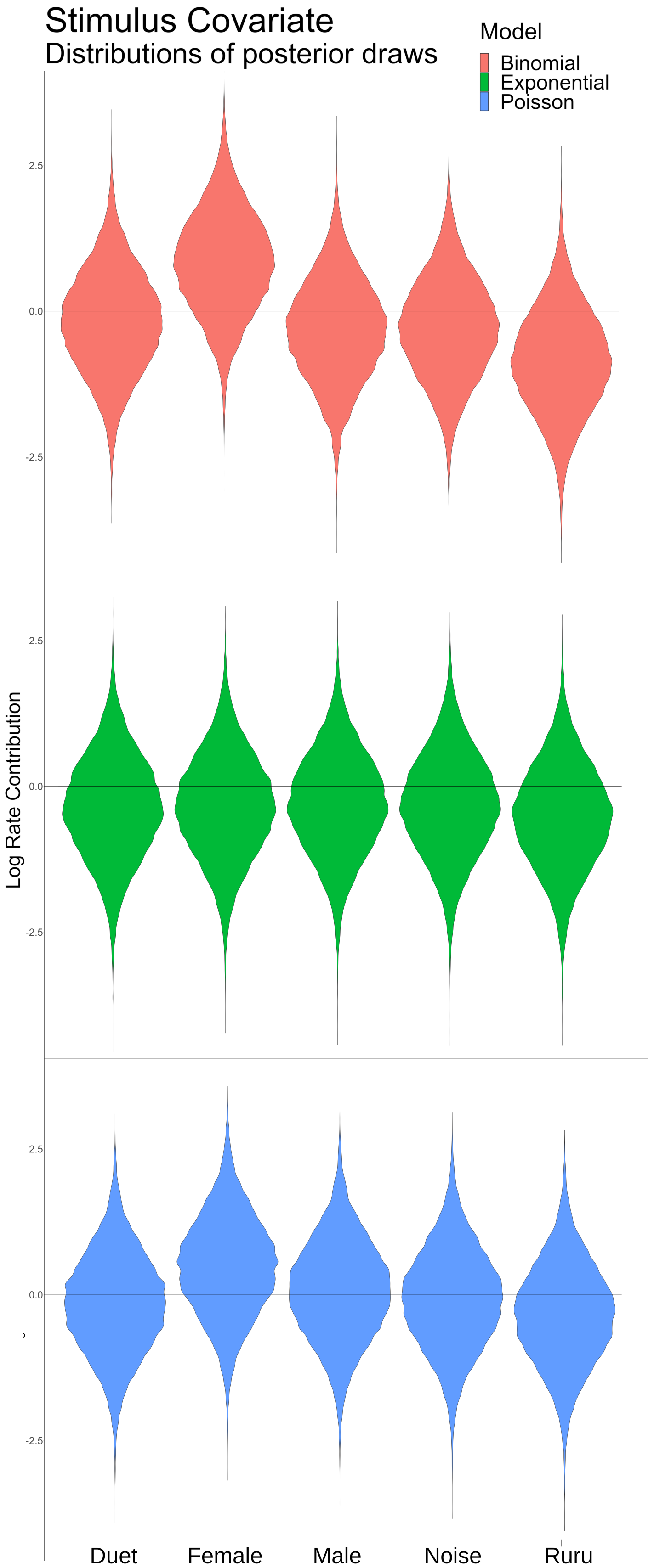


Figure 4.9: Posterior distributions for the stimuli covariates across the three models. Distributions completely above or below the horizontal zero line can be considered as having a consistent effect on the response variable.

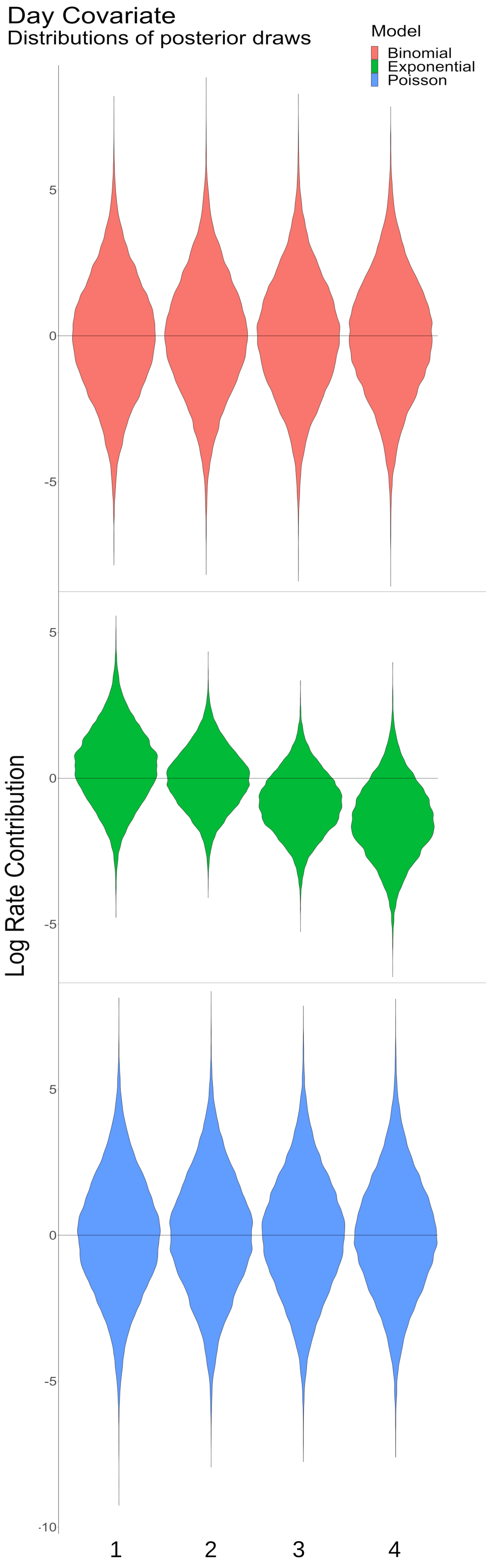


Figure 4.10: Posterior distributions for the day covariates across the three models. Distributions completely above or below the horizontal zero line can be considered as having a consistent effect on the response variable.

ates, the moonlight and temperature ones lay mostly on the positive side, the wind related covariates (speed and direction) were mostly centred around zero, and the atmospheric pressure and relative humidity ones were mostly negative (figure 4.5). The posterior distributions for the two data collection periods, sex, and area categories were quite similarly distributed with $\sim$ 60% of them below zero (figure 4.6, 4.7, and 4.8). The posterior distributions of the stimuli categories were all similarly distributed, with slightly more than 60% below zero (figure 4.9). Finally, $\sim$ 60% of the posterior distribution for the first day category lay above zero, the second day one was almost perfectly centred around zero, $\sim$ 75% of the third day one lay below zero, and a little more than 85% of the fourth day one was below zero (figure 4.10).

Number of early versus total vocalisations per time slot. There were no covariates consistently associated with the response variable. We report that, among the posterior distributions of the weather covariates, $\sim$ 70% of the moonlight, prevailing wind speed, prevailing wind direction, and relative humidity ones were below zero, while the temperature and atmospheric pressure were $\sim$ 60% and $\sim$ 75% above zero, respectively (figure 4.5). Similarly to the Poisson model, most of the data collection period covariates posterior distributions resulted on opposite sides of zero, with $\sim$ 80% of the March covariate being on the negative side and 60% of the July one on the positive side (figure 4.6). The three categories of the sex covariate (Female, Male, and Sum) were almost identically distributed, with $\sim$ 60% of their posterior distributions below zero (figure 4.7). The posterior distributions of the three categories of the area covariate resulted with the upper gully one being mostly centred around zero, the swamp one with $\sim$ 70% below zero, and the combined one with $\sim$ 60% below zero (figure 4.8). Of the five categories of the stimuli covariate, the posterior distribution of Female stimulus one resulted $\sim$ 80% above zero, the Male stimulus and controls around $\sim$ 65% below zero, and the duet stimulus $\sim$ 80% below zero (figure 4.9). Finally, likewise the Poisson model, all the day categories were similarly centred around zero (figure 4.10).

4.4 Discussion and Conclusions

Our results show no differences in the vocal behaviour of a high density population of NIBK when presented with stranger conspecific playback, regardless of season. This has several possible explanations, related to territoriality or hearing acuity. We know from marked birds that home ranges of NIBK on Ponui Island greatly overlap (Ziesemann, 2011) and that birds from Kauri Bush spend most of their time within or around the Gully (Dixon, 2015), indicating overlapping home ranges between groups of birds. Nests from individual birds, however, are always placed in the same area within bush patches. Therefore, although the broadcasts were definitely detectable by NIBK, as indicated by the fact that our ARUs detected them as well, it could be that the simulated individuals from the pasture (i.e., the loudspeakers) were far away enough as not to be perceived as a threat to respond to by NIBK. Other options include time–related responsiveness, such that animals would respond only during a certain time–window (different from the one employed here), or that they would not respond to strangers, rather to known individuals. These explanations could be tested by performing an experiment with strangers vocalisations changing only the time window and, on the other hand, by broadcasting known individuals, for which we already have high quality vocalisations (from chapter 6). A completely different perspective could be that thanks to high hearing acuity (Corfield et al., 2011) and a specialisation to high frequencies (Corfield et al., 2012), NIBK were able to recognise our broadcasts as fake, and so ignored them. This would however be interesting to test, as NIBK, even on Ponui Island, are known to respond to imitations, for instance using a shepherd whistle for capturing purposes (McLennan and Potter, 1992). This could be done with playback synthetic vocalisations of different degrees of complexity, although such a test would require different settings, chiefly involving playback from within gullies or to populations known to respond to broadcast.The fact that birds equipped with individual acoustic recorders did not vocalise during this experiment may be related to a specific response to handling. As speculated in chapter 5, this could be clarified by an experiment involving a first handling to equip a number of birds with acoustic devices around one month prior to the start of recording. A mock–transmitter change for half of these individuals would then ensue a few days after the turning–on of the devices. This would result in having individual vocal activity response both before and afer handling. An opportunistic playback experiment could also take place then. It is also possible that these birds may have just not vocalised at all for reasons completely unrelated to the devices as birds with ABARs at several other times vocalised during the first two hours of darkness (which would include the playback time, see chapter 6).

The first aim of this study was to provide information to improve monitoring practices. Our results reinforce the guidelines from the Kiwi Best Practice Manual in discouraging the use of playback for monitoring purposes. However, in addition to the reason given by Robertson et al. (2017) that evolving technology which would make data gathered at years' distance not comparable), we add that, at least in some populations there seem to be no difference in the animals' vocal behaviour when broadcasting calls compared to when unstimulated. We cannot provide any advice on the use of playback to enhance the detection of animals in sparsely populated areas due to the high density of our target population. Further experiments could simultaneously investigate differences in vocal activity following playback in a low and a high density managed populations. The simultaneous study would help to minimise potential seasonal effects and by using managed populations there would be additional information (including localisation using telemetry before, during, and after playback) of marked individuals. Despite not being consistently associated with the distribution of the vocalisations, it would also be interesting to further investigate the possible trend indicated by the “day” covariate from the exponential model (figure 4.10), where differently from the other response variables, times between vocalisations per time slot might be more likely longer in the later days of playback, regardless of the data collection period (figure 4.6); although none of these covariates were consistently associated with the response, if confirmed by further experiments, this would also reinforce the idea that playback would not be a desirable tool for acoustic monitoring, as KCC are performed over multiple nights (Craig, 2018).

The second aim of this study was to improve our understanding of NIBK vocal behaviour. The results of our statistical analyses which are consistently associated with the response variable in the Poisson model add to previously reported relationships of number of NIBK vocalisations over time with external factors, such as moonlight and weather conditions (Colbourne and Digby, 2016). Differently from chapter 5, where the presence of a `day` covariate would basically nullify the contribution of these external factors to the response variables (see page 108), here we have both `day` and `data collection period` (representing different seasons) covariates, which points to an actual effect of lunar illumination, temperature, and relative humidity with the number of vocalisations per time slot. The relationship of wind speed with fewer number of vocalisations could be due to reduced detection of the ARUs (e.g., Colbourne and Digby, 2016), however other studies had previously reported fewer vocalisations in more illuminated nights (Colbourne and Digby, 2016; Colbourne and Kleinpaste, 1984, although in the 2016 study this was only relevant for females in a particular site). The positive relationship of temperature and atmospheric pressure with the number of vocalisations could indicate that more favourable weather conditions entail a higher number of social encounters, hence of vocalisations, and may be worth further investigation. This could be integrated with the use of camera traps to record such potential interactions.

To conclude, we here presented the results of the first formal evaluation of playback as an acoustic monitoring enhancing tool with NIBK. Acoustic playback may be well suited to elicit responses in some populations (e.g., Taborsky and Taborsky, 1991). However, as long as the number of vocalisations per unit of time is used as approximation of abundance nationwide, we would deter its usage in formal monitoring. This is because it could possibly lead to biased estimates for populations that responded more to playback than others. In addition, as NIBK vocal behaviour is still quite unknown, we endeavour to investigate alternative possibilities, such as individual vocal activity, to contribute to more accurate acoustic monitoring and, in turn, more informed conservation efforts. On top of clarifying how individuals at different densities respond to playback, obtaining individual–level data on vocal activity would provide crucial information to accurately relate numbers of vocalisations obtained through KCC to abundance estimates. However, NIBK (at least on Ponui Island) home ranges greatly overlap (Ziesemann, 2011), and deploying ARUs in the proximities of a known burrow would not warrant the recording of just one pair as in congeneric species (Bedoya and Molles, 2021; Digby, 2013, could confidently record individuals from ARUs from territories of Kiwi Pukupuku and Roroa). Animal–borne acoustic recorders (Stowell et al., 2017) could represent a valid alternative to obtain such data and a way of comparing individual versus community level vocal behaviours (see Chapter 6).

Chapter 5

North Island Brown Kiwi (*Apteryx mantelli*, Bartlett) vocal activity in relationship to handling during annual transmitter change

Figure 5.1: Te Taonga handling Clea the Kiwi. Picture by Vilma Rodriguez

Abstract

Capturing and handling wild birds has been reported to have a variety of effects on the animals. This undoubtedly stressful event may cause short and long term behavioural changes. Whereas some studies report long-term observable differences in behaviour, some others report variable short-term effects, and some report no observable effects.

It is important to evaluate such potential changes in target individuals to ensure both animals' welfare and the collection of subsequent unbiased data. However, nocturnal and cryptic animals are hard to visually monitor after release.

In this study, we compared the vocal activity of a long–term studied population of North Island Brown Kiwi (*Apteryx mantelli*, Bartlett) before and after the annual transmitter change, with that of a neighbouring population that are naive to investigator disturbance.

We found no differences in vocal activity between the two considered populations in relation with either long–term nor short–term post–handling effects.

We additionally report how the vocal activities of both populations were consistently associated with exogenous factors. Specifically birds were less vocal at higher lunar illumination and atmospheric pressures, although the effect of moonlight disappeared when including a time–related (day) covariate. Similarly there were more duets at lower atmospheric pressures, and certain prevailing wind directions also affected the number of duets.

These results cannot be generalised to all North Island Brown Kiwi populations. However, since our approach is relatively inexpensive and easy to apply (as it operates with passive acoustic monitoring and free open–source software), we invite fellow researchers to carry out similar trials whenever possible to improve our understanding of post–handling behaviours of cryptic and elusive species.

5.1 Introduction

Capturing and handling wild birds is an often unavoidable process, as it allows us to get otherwise unobtainable information directly from the animals (e.g., measurements, blood samples and the like), as well as to mark them for individual recognition. This information is indispensable to monitor populations and to understand the animals' biology. Along with coloured and metal bands, the rapid development of animal–borne tracking devices and data loggers over the past few decades has provided a wealth of information to researchers (Bridge et al., 2011; Wilmers et al., 2015). However, we generally expect these invasive procedures to elicit behavioural and physiological changes in the animal we handle (Mortensen and Rosell, 2020). It is thus important to monitor the possible effects of investigator disturbance (Carney and Sydeman, 1999), in order to ensure both good animal welfare (Curzer et al., 2013; Soulsbury et al., 2020) and unbiased results (Powell and Proulx, 2003).

Capturing and handling wild birds has been reported to have a variety of effects on the animals (Weiser et al., 2016). Whereas some studies report long-term observable differences in behaviour (e.g., Linhart et al., 2012), some others report variable short-term effects (Budka et al., 2019; Schlicht and Kempenaers, 2015; Seress et al., 2017), and some report no effects (Petrusková et al., 2021, although short-term effects could not be assessed). In this study we compared the vocal behaviour of two neighbouring communities of North Island Brown Kiwi (NIBK — *Apteryx mantelli*, Bartlett) in a long–term study site before and after manipulations.

Relatively little is known about the response of wild kiwi species (*Apteryx* spp.) to handling. However, wild NIBK showed significantly higher costicosterone levels immediately after capture when compared to captive conspecifics suggesting higher levels of stress (Adams, 2000), and some Roroa (Great Spotted Kiwi, *Apteryx haastii*, Potts) individuals abandoned their usual roosting sites, at times for months, following handling (Toy and Toy, 2021). Both these observations beg the question of how impactful handling is to *Apteryx* populations and how research and conservation efforts are affected by them.

Apteryx are mainly nocturnal, making it hard to observe them directly to record their behaviour. However, their sexually dimorphic calls (Corfield et al., 2008) are very loud and therefore easy to detect without interfering with the callers using acoustic recorders. NIBK calling behaviour has been studied several times, including monitoring the effects of some environmental variables. Previous studies have reported lower numbers of vocalisations in more moonlit nights (Colbourne and Digby, 2016; Colbourne and Kleinpaste, 1984, although in the latter this only was true for females in Whanganui).

In this study we took advantage of a long–term high–density study population of NIBK with marked birds to examine the calling behaviour of NIBK before and after manipulations when considering environmental factors in addition to handling. In this population, individuals are captured at least annually to replace transmitters (batteries last approximately one year). We placed Autonomous Recording Units (ARUs) in two separate areas, one with birds wearing transmitters and another with no marked (nor handled) birds, for 16 days which included the handling period for all tagged birds. We also recorded environmental conditions reported to possibly affect calling in kiwi and other species. Our hypothesis was that if handling affected calling behaviour we should see differences in the calling of birds in the two gullies after considering environmental effects.

5.2 Materials and Methods

5.2.1 Study Site

We conducted this study on Ponui (Chamberlin's) Island (Figure 5.2, 175.14860,−36.90422 : 175.21860,-36.82961 – Coordinate Reference System WGS84 for all coordinates herein), in the Hauraki Gulf, Auckland region, Aotearoa New Zealand.

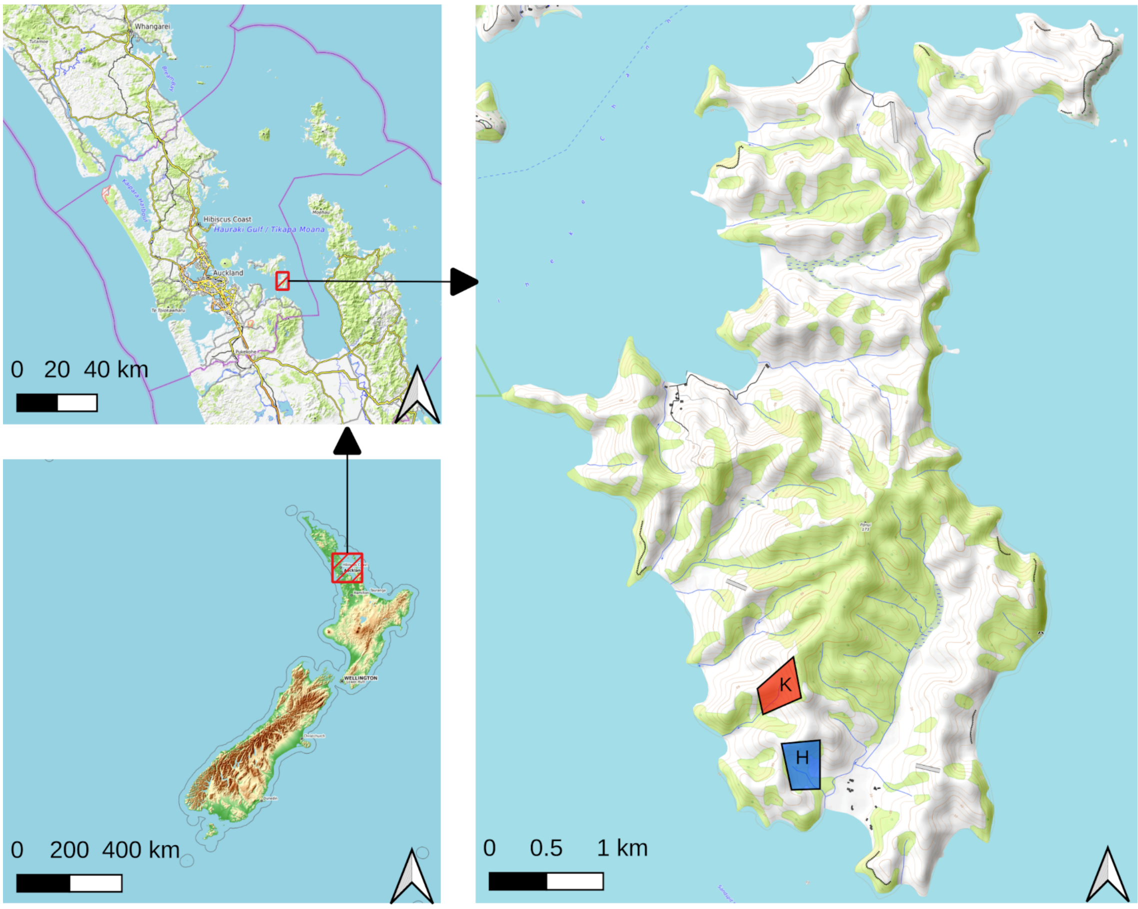


Figure 5.2: Location of the study site relative to Aotearoa New Zealand (bottom left), the greater Auckland region (top left), and Ponui Island (main, right). The coloured areas represent the gullies that have been involved in this study with their initials to help identification. Shaded in red is **K**auri Bush, and shaded in blue is **H**omestead Gully, which we chose as control for our analysis of vocal activity against human disturbance as it does not give residence to NIBK with transmitters.

Fourteen NIBK were translocated on Ponui in 1964 (Miles and Castro, 2000) and, as of 2021, Ponui Island hosts a high-density NIBK population, which has been under study since 2004 (Colbourne, 2005; Cunningham et al., 2007; Ziesemann, 2011). Over the last 17 years, from 30 to 50 birds have been equipped with very high frequency (VHF) radio transmitters that are monitored at least monthly. Having record of the history of these animals, including morphological measurements, usual resting areas, nesting sites, and social relationships (Cunningham and Castro, 2011; Dixon, 2015; Undin et al., 2021), make Ponui Island an ideal site to conduct this trial.
We used NIBK in two of the remnant kauri (*Agathis australis*) forest patches in Southern Ponui. Kauri Bush (175.177635,−36.887127 : 175.182262,−36.882096) and Homestead Gully (175.180982,−36.892250 : 175.184780,−36.888938). Kauri Bush is home to 14 birds with transmitters, but we had no marked birds in Homestead Gully (Figure 5.2) and we so used it as a control for handled versus unhandled birds. However, given the proximity of the two gullies, the overlap in NIBK home ranges on Ponui, and that some birds have been marked with passive integrated transponders but not radio transmitterised, there is a slim possibility that birds previously handled migrated to this gully. However, none of the project's transmitterised birds has ever been localised in Homestead Gully (Dixon, 2015). This comes as a trade off with the high degree of similarity in vegetation and geography between the two gullies. The birds of Kauri Bush equipped with transmitters at the time of this study had been in the study for on average 10.36 years (median = 12, min = 1, max = 15), and had been handled between 3 and 17 times overall (mean = 12, median = 14), at an average yearly rate of 1.34 times per year (median 1.25, min = 0.58, max = 3.00).

More detailed information regarding Ponui Island, its NIBK population, vegetation, and related studies can be found in the following theses: Bansal (2020); Dixon (2015); Strang (2018); Undin (2021); Ziesemann (2011).

5.2.2 Study System

NIBK are flightless and nocturnal cryptic birds belonging to the only Apterigiformes genus, *Apteryx*. NIBK are considered to be a somewhat territorial and vocal species (McLennan et al., 1987; Taborsky and Taborsky, 1992, 1995, 1999), but other studies have shown that territoriality may be dependent on population density (Potter, 1989). On Ponui Island, marked birds hold overlapping home ranges of between five and six hectares (Ziesemann, 2011, Chapter 3). Birds roost in burrows or on the surface in forest patches within their range and the choice of roost site is dependent on proximity to foraging areas (Dixon, 2015, Chapter 4). At night kiwi move from the roost sites to foraging sites and are mostly found in forested and scrub areas (Dixon, 2015, Chapter 4, figure 3.6). Additionally, in the work of Dixon (2015, which included 2377 nocturnal localisations of Ponui NIBK), although animals displayed variability in habitat utilisation on the temporal, breeding and social status scales (e.g., birds that tended to shelter in pairs were found more often in forest environments than those sheltering alone or in groups), none of the transmitterised birds was ever recorded close or within Homestead gully. vocalisations of males and females NIBK are quite distinct; Corfield et al. (2008) describes males as producing long–travelling whistles which could be grouped into phrases, and with fundamental frequencies around $1.5kHz$ (and overtones up to approximately $13kHz$), whereas female produce hoarser and lower (fundamental frequency around 0.1 kHz) vocalisations with intensity concentrated in two formants (Figure 1.2); NIBK are also well known for their duets, which are multiple vocalisations in short succession and have been suggested to possibly have territorial implications (Corfield et al., 2008). Duets often include male and female vocalisations which can be overlapping, or a few seconds apart from each other: we here adopt a stricter definition (for the binomial model, see Methods), considering only overlapping vocalisations (of any sex) as choral.

5.2.3 Experimental Design

We chose two areas of approximately equal size and vegetation and deployed AR4 ARUs from the New Zealand Department of Conservation Te Papa Atawhai (32 KHz frequency range, 16 bit depth, frequency response chart in figure II.iii). We deployed one device in the heart of Kauri Bush (175.179924,−36.884614) and in that of Homestead Gully (175.182917,−36.890594) respectively on the morning of the 8th of March, 2021 and retrieved them on the 25th of March, 2021 during the daytime. One area, Kauri Bush, has been part of the study since 2006 and many of the birds roosting there are marked with radio transmitters (see chapter 6). The second area, Homestead Gully has not been part of the study and NIBK using this gully as roosting sites have had no researcher interference.

We recorded from 19:00 each night for 5 hours (sunset was at 19:50 on the first day of this study , 08/03/2021, and at 19:27 on the last, 24/03/2021).

We logged weather variables using a Kestrel 4500 portable weather station (`https://nz-kestrelmeters.glopalstore.com/products/kestrel-4500-weather-meter`, located at base camp, 175.187373,−36.894459) and obtained moonlight illumination data from `https://https://www.timeanddate.com/moon/@2184425`. We captured, handled, measured, and replaced the radio transmitters of birds from the 15th to the 20th of March. No blood samples were taken on this trip. We captured most birds in Kauri Bush on the 15th of March, hence we used this date as threshold for grouping the data in before and after manipulation (*before* and *after* categorical covariate in the statistical analyses).

5.2.4 Analyses

Using AviaNZ (Marsland et al., 2019), we processed all the recordings to detect NIBK vocalisations. We then manually reviewed the results to correct for false positives (segments wrongfully labelled as NIBK). We also skimmed the recordings in manual mode to amend any false negatives (unlabelled NIBK vocalisations).

We then modelled three aspects of the distributions of NIBK vocalisations, specifically the number of vocalisations per hour, the time between vocalisations, and the number of vocalisations involved in duets. In this study we set a very strict criterion for duets, as we considered duets only those vocalisations that to any extent overlapped with others, meaning that if a vocalisation started before the end of another, we would label them both as being part of a duet.

We used a model with Poisson likelihood (with Log link function) to examine the number of vocalisations per unit of time (hour), a model with Exponential likelihood (with Logit link function) for the time between vocalisations, and a model with Binomial likelihood for the number of vocalisations involved in choruses (i.e., any overlapping set of vocalisations) against the single calls per hour. We worked within a Bayesian framework for two reasons. First, to be able to discern positive versus negative changes in vocal activity. By specifying zero–centred prior distributions (before the simulations) for all the covariates, the posterior distributions (after the simulations) that fell completely above zero consistently contributed to higher rates of vocalisations per night (meaning more vocalisations in the Poisson model and shorter waiting times between vocalisations for the exponential model), and higher probability of overlapping vocalisations (binomial). Conversely, the posterior distributions that fell completely below zero consistently contributed to lower rates, and those that remained somewhat centred around zero did not consistently contibute to either higher or lower rates. Table 5.1 summarises the composition of all the models. We ran two rounds of these models, one including a `day` categorical covariate (*i.e.* with a level for each day), and one without this.

We ran all the models through `R` (`R` Core Team, 2013) and `rstan` (Stan Development Team et al., 2016). All models ran on three independent chains, each for 60,000 iterations, discarding the first half as burn–in, and retaining one every fifteen values with an adapt delta of 99.9% and maximum treedepth of 15, for a final sample of six thousand draws per covariate per model.

5.3 Results

In total, the ARUs detected 1131 vocalisations, of which 490 were detected in Homestead Gully and 641 in Kauri Bush. Males called more than females in both gullies, with 104 vocalisations from females in Homestead against 386 from males. Only 29 females vocalisations were detected in Kauri Bush, against the 612 from males (figure 5.3 and table 5.2).

All the models ran smoothly and without post–sampling warnings and the chains reached good convergence values. In fact, having set all diagnostic parameters (standard error of the mean of

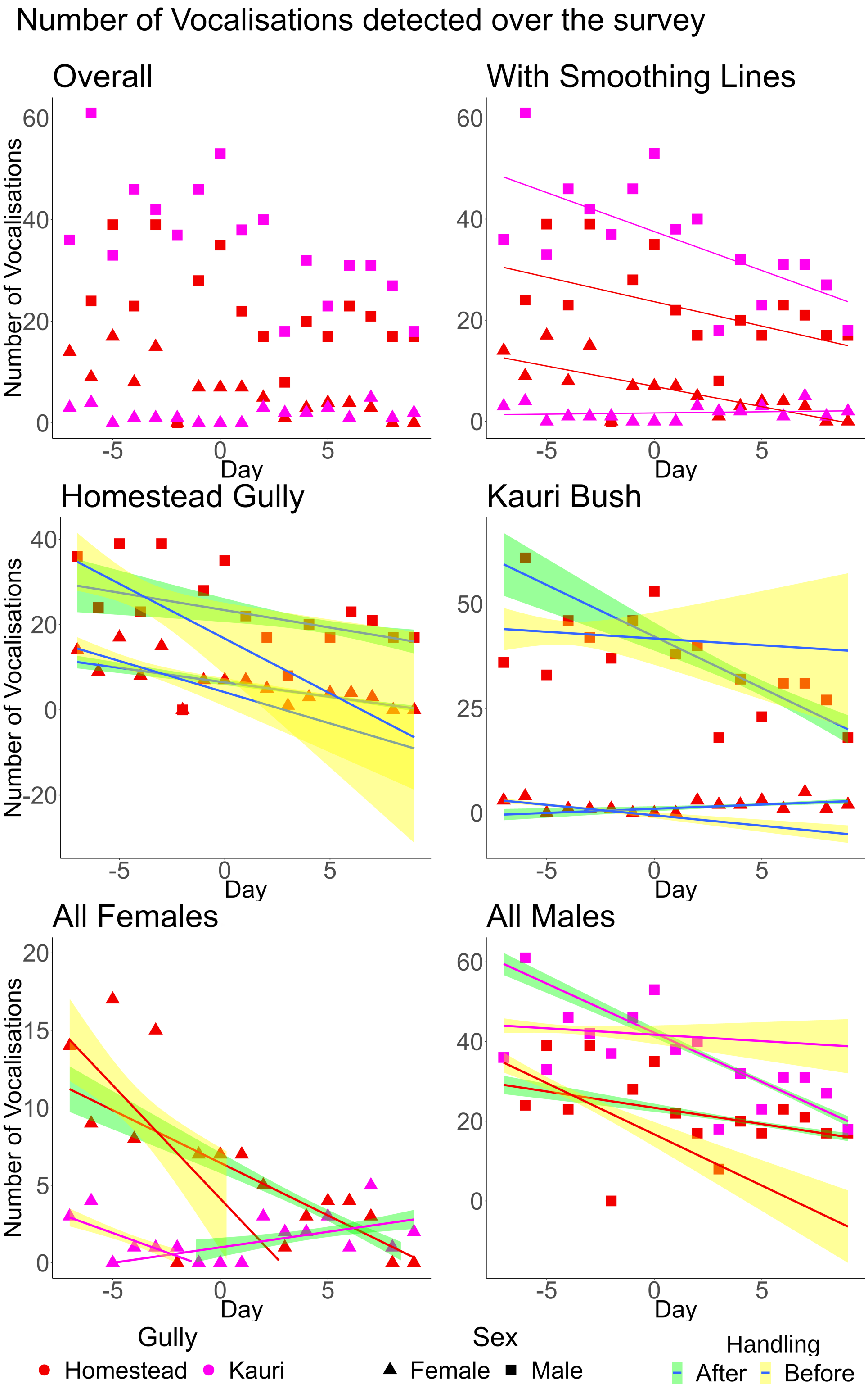


Figure 5.3: Distribution of the vocalisations detected over the duration of the survey from the two gullies. The green and yellow shades around the regression lines represent regressions accounting for only the data before (yellow) or after (green) the arrival of the research team on the island

the posterior draws, effective number of independent draws, and Gelman and Rubin potential scale reduction statistic $\hat{R}$) to their minimum warning thresholds (5%, 5%, and 1.01 respectively), there were no variables exceeding these in any of the models.

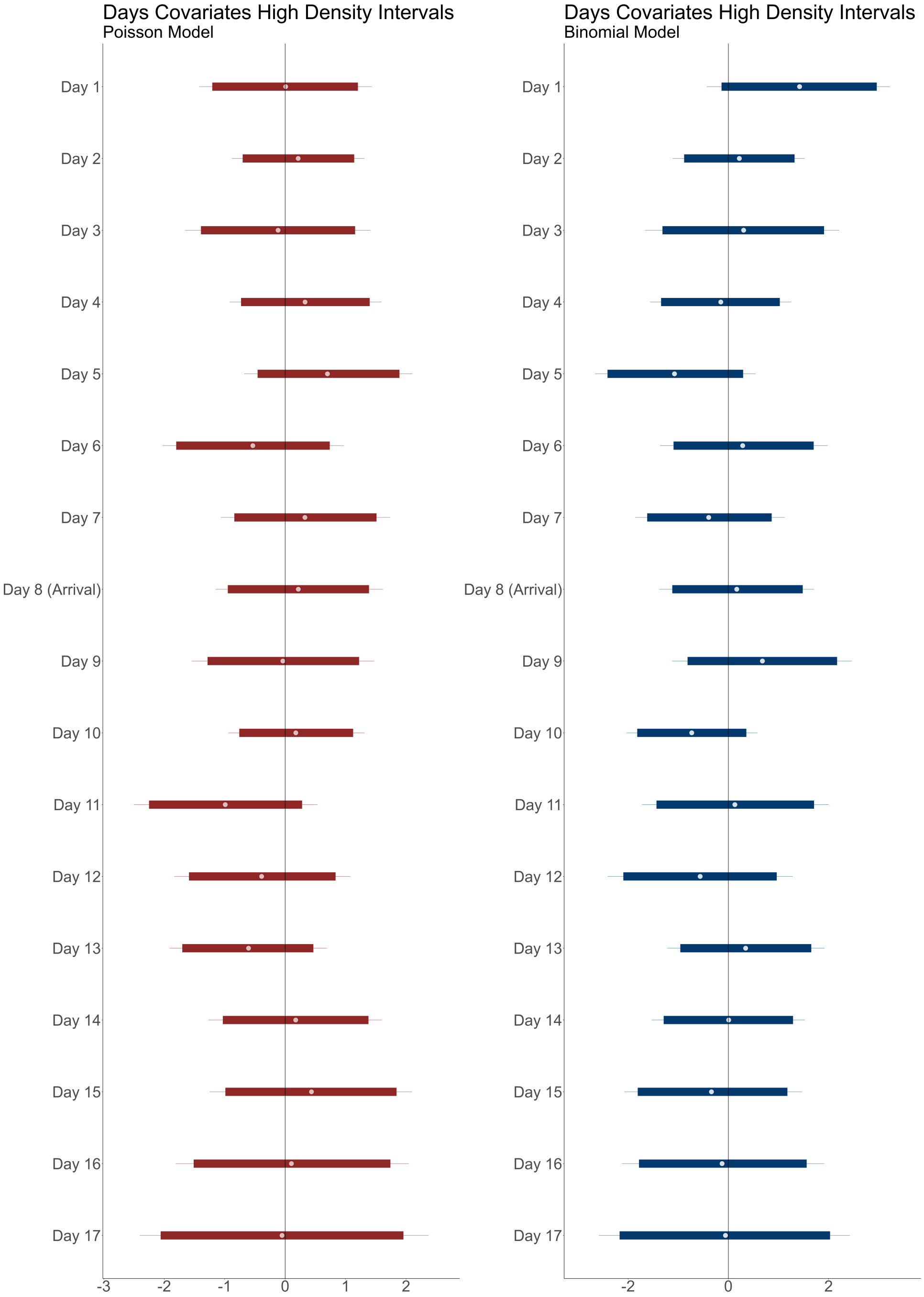
In the models which included the `day` covariate, this never was consistently associated with higher or lower response variables, although there was quite a lot of variability (figures 5.4). The only covariates consistently associated with the response variables in these models were Prevailing Wind Speed and Direction in the Poisson Model, and Relative Humidity in the binomial model.

Table 5.1: Summary of the covariates included in each model.

Covariate	Type (**C**ategorical, **N**umerical)	Categories (if any)	Models		
			Poisson	Exponential	Binomial
Intercept	C	global	✓	✓	✓
Sex	C	Female, Male	✓	✓	
Relative Sex	C	Male after Male Male after Female Female after Male Female after Female		✓	
Gully	C	Kauri, Homestead	✓	✓	✓
Hour	C	21, 22, 23, 24	✓	✓	✓
Handling	C	Before, After	✓	✓	✓
Moonlight	N	-	✓	✓	✓
Temperature	N	-	✓	✓	✓
Prevailing Wind Speed	N	-	✓	✓	✓
Prevailing Wind Direction	N	-	✓	✓	✓
Relative Humidity	N	-	✓	✓	✓
Atmospheric Pressure	N	-	✓	✓	✓

Table 5.2: Numbers of vocalisations detected by the ARUs between 19:00 and 24:00 each day, divided by sex and gully. Handling started on day 15 and stopped on day 20 (highlighted in blue).

Day	Homestead		Kauri	
	F	M	F	M
8	14	36	3	36
9	9	24	4	61
10	17	39	0	33
11	8	23	1	46
12	15	39	1	42
13	0	0	1	37
14	7	28	0	46
15	7	35	0	53
16	7	22	0	38
17	5	17	3	40
18	1	8	2	18
19	3	20	2	32
20	4	17	3	23
21	4	23	1	31
22	3	21	5	31
23	0	17	1	27
24	0	17	2	18



Days Covariates High Density Intervals
Poisson Model

Days Covariates High Density Intervals
Binomial Model

Figure 5.4: 95% High Density Intervals for the Days covariates.

All the results reported hereafter are from the models without `day` covariate. Using the Poisson model, we found that the only covariates consistently affecting the probability of a Poisson random variable representing the number of vocalisations per hour were moonlight and atmospheric pressure (figure V.i and table 5.3).

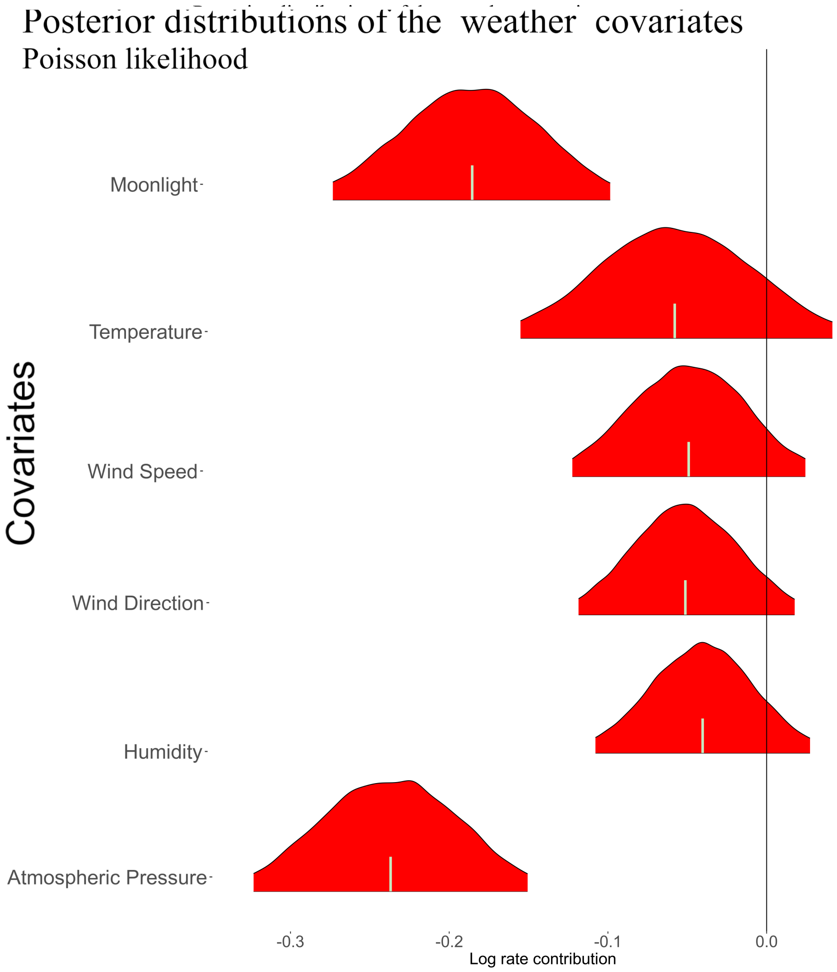


Figure 5.5: Weather covariates for the Poisson model. Covariates with posterior distributions overlapping zero (vertical line) are not consistently associated with the response variable.

Table 5.3: Weather covariates high density intervals for the Poisson model. The columns show the Gelman and Rubin potential scale reduction statistic ($\hat{R}$), the effective number of independent samples per variable, mean and standard deviation of the MCMC draws, and three representative quantiles (2.5%, 50%, and 97.5%) of the same.

Parameter	$\hat{R}$	n_{eff}	mean	sd	2.5%	50%	97.5%
Moonlight	1.000	13722	-0.186	0.045	-0.273	-0.186	-0.099
Temperature	1.000	13515	-0.057	0.050	-0.155	-0.058	0.041
Wind Speed	1.000	14982	-0.049	0.038	-0.122	-0.049	0.024
Wind Direction	1.000	15040	-0.051	0.035	-0.118	-0.051	0.018
Humidity	1.000	14985	-0.040	0.034	-0.108	-0.040	0.027
Atmospheric Pressure	1.000	14735	-0.237	0.045	-0.323	-0.237	-0.151

The categorical sex covariate shows expected differences between the sexes, with males more consistently associated with higher numbers of vocalisations per hour than females (figure V.ii). The posterior distributions of both levels of the handling and gully covariates were almost identically distributed (centred around 0, figures V.iii and V.iv), and showed no effect of these on the probability of detecting more or less vocalisations.

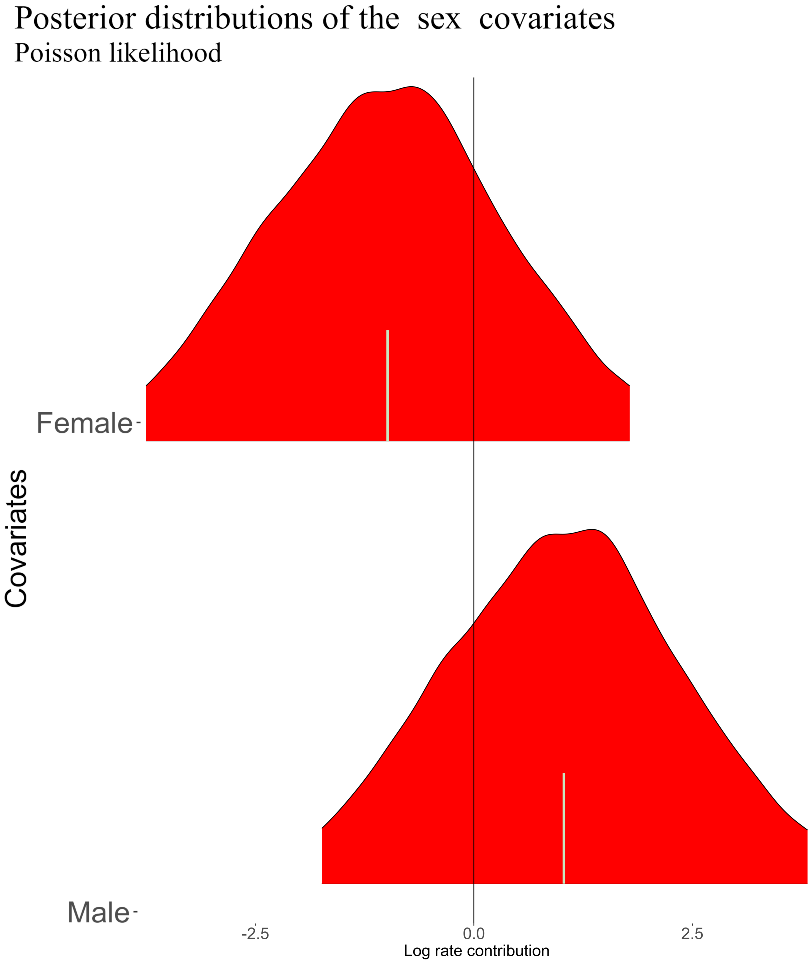


Figure 5.6: Sex covariate for the Poisson model. Covariates with posterior distributions overlapping zero (vertical line) are not consistently associated with the response variable.

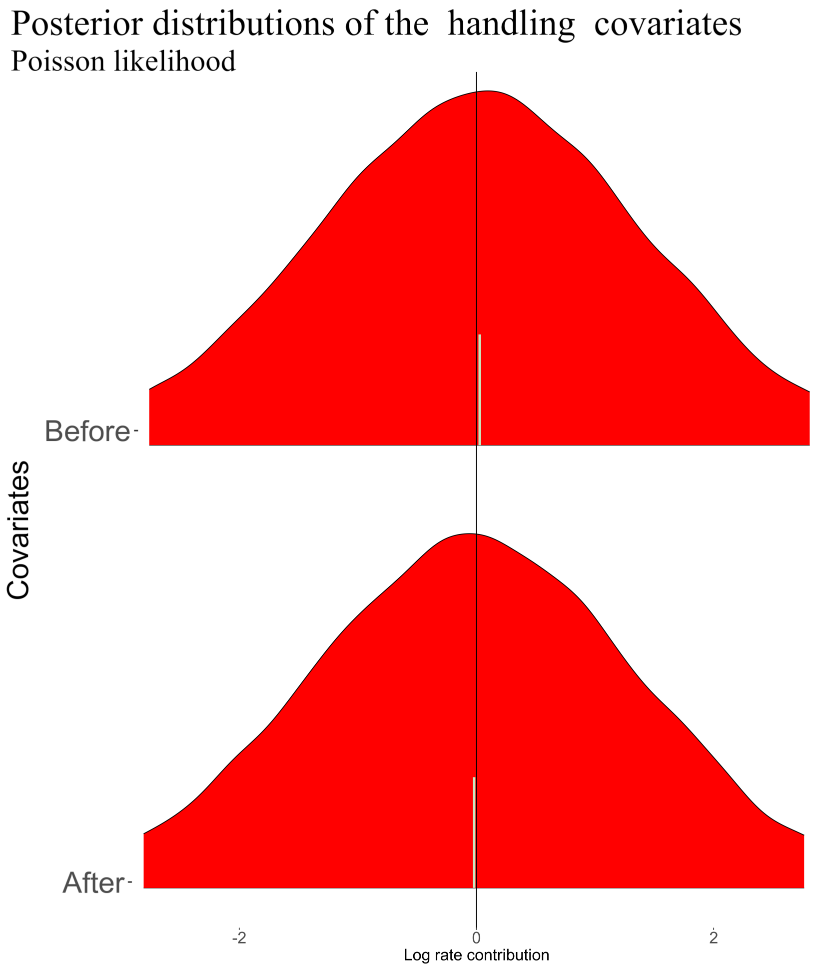


Figure 5.7: Handling covariate for the Poisson model. Covariates with posterior distributions overlapping zero (vertical line) are not consistently associated with the response variable.

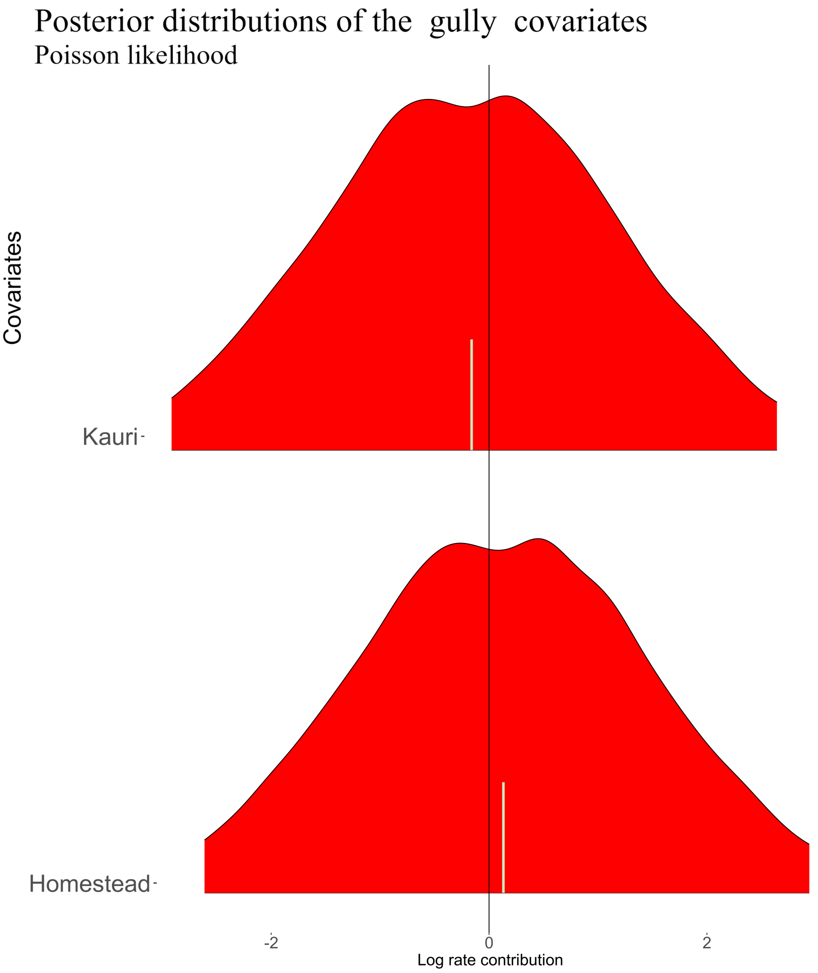


Figure 5.8: Gully covariate for the Poisson model. Covariates with posterior distributions overlapping zero (vertical line) are not consistently associated with the response variable.

From the Binomial model, we again found that the sole covariates consistently affecting the probability of a Binomial random variable representing the number of vocalisations involved in duets in an hour (success) over the total number of vocalisations per hour were prevailing wind direction, and atmospheric pressure (figure 5.9 and table 5.4).

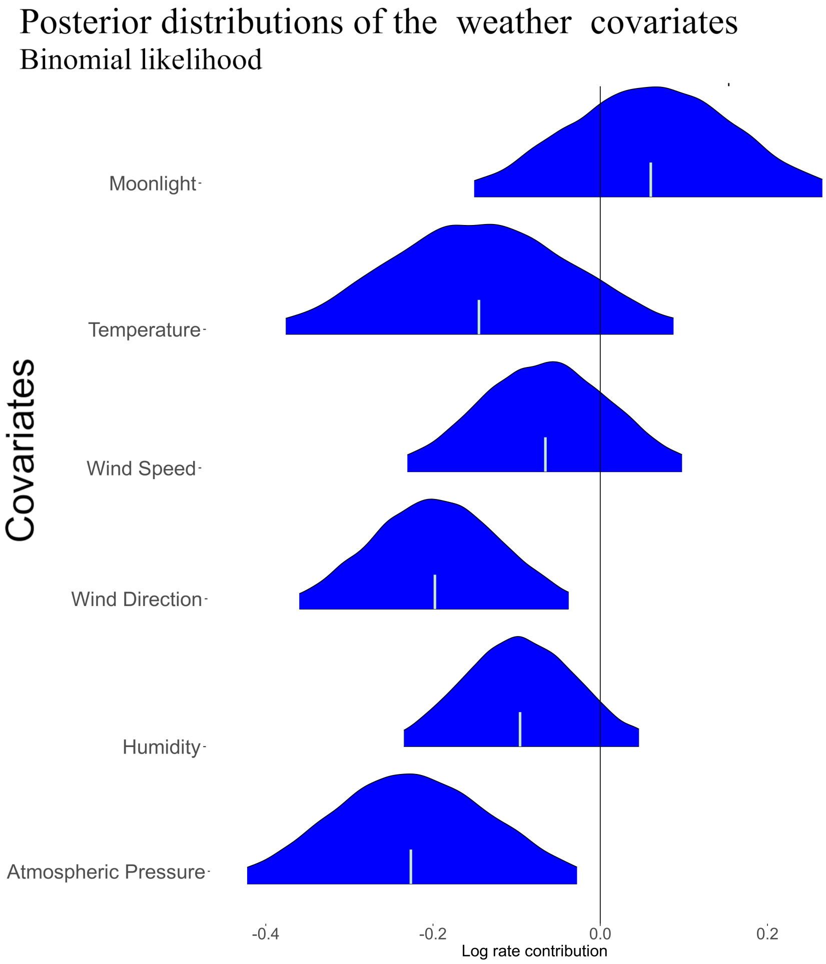


Figure 5.9: Weather covariates for the Binomial model. Covariates with posterior distributions overlapping zero (vertical line) are not consistently associated with the response variable.

Table 5.4: Weather covariates high density intervals for the Binomial model. The columns show the Gelman and Rubin potential scale reduction statistic ($\hat{R}$), the effective number of independent samples per variable, mean and standard deviation of the MCM draws, and three representative quantiles (2.5%, 50%, and 97.5%) of the same.

Parameter	$\hat{R}$	n_{eff}	mean	sd	2.5%	50%	97.5%
Moonlight	1.000	14307	0.060	0.106	-0.151	0.060	0.266
Temperature	1.000	14967	-0.144	0.118	-0.376	-0.145	0.088
Wind Speed	1.000	14639	-0.066	0.084	-0.231	-0.066	0.098
Wind Direction	1.000	15039	-0.198	0.082	-0.360	-0.198	-0.038
Humidity	1.000	14368	-0.095	0.072	-0.235	-0.096	0.046
Atmospheric Pressure	1.000	15064	-0.226	0.101	-0.422	-0.227	-0.028

Similarly to the Poisson model, neither the gully nor the handling categorical covariates showed any association with higher or lower probabilities for the Binomial model (figures 5.10 and 5.11). In the exponential model there were no covariates significantly associated with the probability of more sparsely or clustered vocalisations throughout the first four hours of darkness.

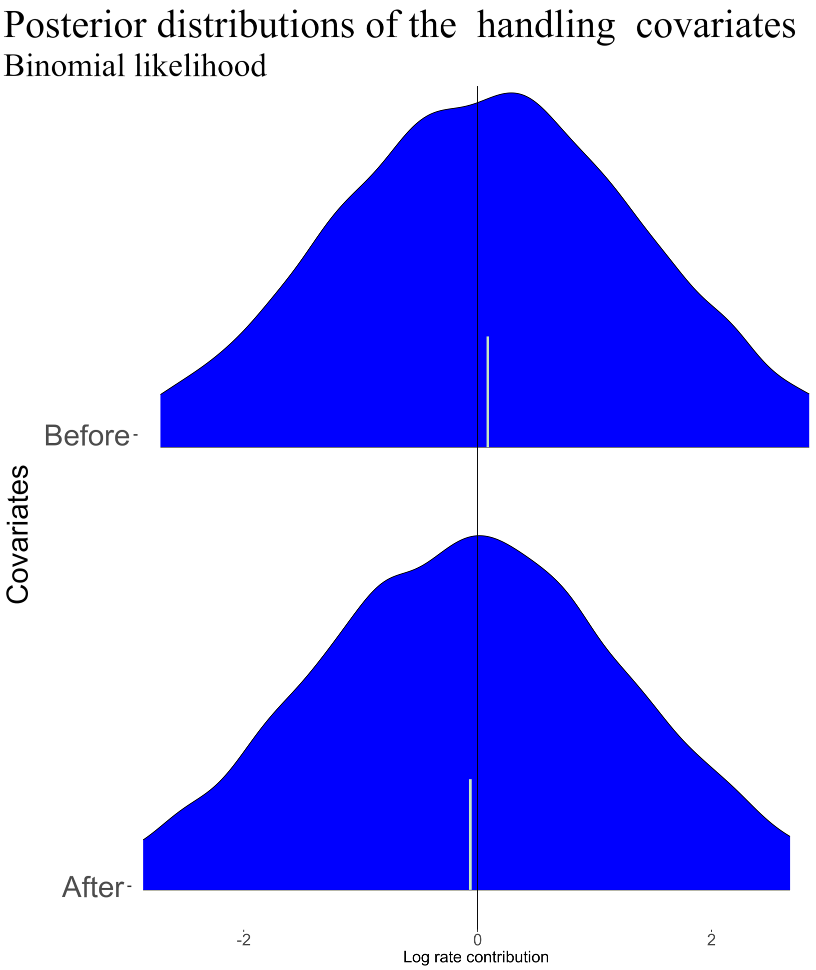


Figure 5.10: Handling covariate for the Binomial model. Covariates with posterior distributions overlapping zero (vertical line) are not consistently associated with the response variable.

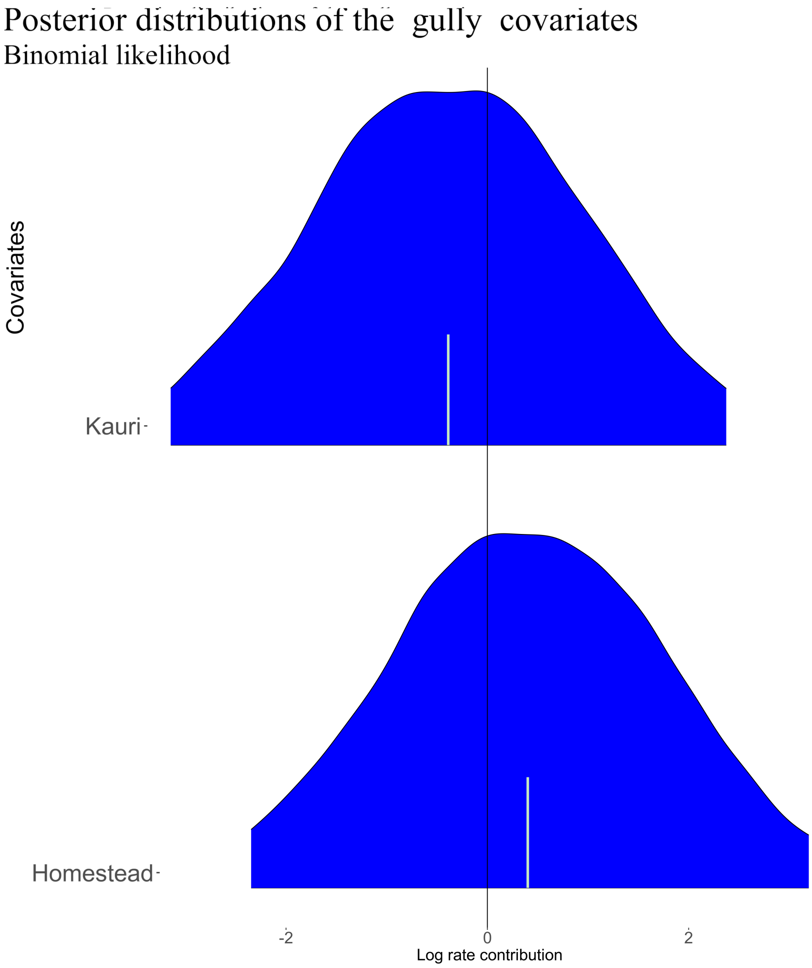


Figure 5.11: Gully covariate for the Binomial model. Covariates with posterior distributions overlapping zero (vertical line) are not consistently associated with the response variable.

5.4 Discussion and Conclusions

One model included a categorical covariate with different levels for each day, and a second did not. We used the first to ascertain whether there was any particular day (specifically, the 8th day, which was the day of transmitter changing for most birds) when the response variable (of any given model) was consistently different from the others: we found this not to be the case (figure 5.4). However, since the `day` covariate showed great variability, we decided to run a second set of models without this date covariate to investigate the possibility of this variability being explained by other covariates (e.g., the handling covariate).

Apart for some oddities, for example that there were no detected vocalisations from Homestead on the 13th of March, the vocalisations seemed to be distributed similarly between the two gullies; this was also confirmed by the gully covariates across the models. If the Kauri birds had changed their vocal behaviour due to the years of handling, we would expect to see differences in the distribution of the gully covariate. Similarly, if there were differences common to both gullies in the vocal behaviour before and after handling, we expected the handling covariate to show some differences. Finally, if the birds in Kauri bush changed their vocal behaviour post–handling, we expected to see this in the gully/handling interaction covariate.

In a congeneric species, the Roroa, Toy and Toy (2021) found that capturing events were at times associated with birds abandoning their usual roosting sites. If this were the case for the Ponui NIBK, we would have expected the Kauri vocalisations to noticeably drop, as during a previously conducted sweep survey with a specialised Kiwi dog for a different experiment (chapter 6, see page 117), we only found one additional bird to those with transmitters in this gully. That is, even if this only sets a lower bound for the Kauri population, we would still expect the fourteen birds with transmitters to represent a considerable part of the population and as such, if they abandoned their usual gully (even if for just a few nights) we expected to hear noticeably fewer vocalisations from Kauri bush.

We found the association of external factors with both the number of vocalisations per hour and the duets per hour interesting. Previous studies have found that NIBK may behave differently according to the lunar phase. Colbourne and Kleinpaste (1984) reported how NIBK would vocalise less in more illuminated nights, while in the study by Colbourne and Digby (2016) only females from one of the sites (Whanganui) would vocalise differently according to moonlight; they also found that Roroa *Apteryx haastii* females were affected by moonlight. This lower level of vocal activity during moonlit nights may be related to visibility. We can speculate that since producing vocalisations requires energy like any other activity, moonlit nights represent opportunities for the animals to conserve energy and simply see each other rather than having to rely on acoustic signals. Alternatively, although unlikely, it could be that during a time when NIBK sill had natural native predators such as the North Island Adzebill (*Aptornis otidiformis*, Owen 1844 — see Tennyson and Martinson, 2006), NIBK evolved such a behaviour to reduce the possiblity of being preyed upon. While we could interpret the prevailing wind direction as a bias of passive acoustic monitoring (ARUs are fixed to trees and sounds coming from certain directions are less detectable, see Castro et al., 2019), it is interesting to see that higher atmospheric pressure, often indicating good weather, was associated with fewer vocalisations and duets. Together with the indication that there may be fewer vocalisations with high temperatures and lunar illumination, this could point to some relationship between vocal activity and ease of foraging, and would be quite interesting to pair a food availability survey (e.g., invertebrate census) with an acoustic survey to see whether NIBK vocalise differently when more or less food is available. However, having here only one (not even complete) lunar cycle, we cannot exclude that moonlight could in this instance approximate time or some other time–related undetected cause of variability in NIBK vocal activity, especially because the moonlight covariate was not consistently associated with the response variables when the day covariate was present. Thus, we recommend both repeating acoustic sampling on Ponui island during different lunar phases to ascertain this apparent association, and analysing data from multiple locations concurrently, to minimise undetected local effects and confirm or dispute the association between lunar illumination and vocal activity.

Here we show how, for a single high densiy population of NIBK that are frequently handled, handling does not appear to affect the number of vocalisations detected in an area. However we did detect distinct changes in vocal activity and further investigation would be interesting. Nevertheless, in the context of this thesis this reassures us that data from ABARs can be used to inform abundance estimates (chapter 6), as the community vocal behaviour does not appear to be affected by the attachment of the devices.

Chapter 6

Animal-borne acoustic recorders can inform consistent abundance estimates based on passive acoustic monitoring

Figure 6.1: Emily, BJ, Alvin, and Andre (taking the picture) with one of the Northland birds.

Abstract

Monitoring animal populations is important to allow for appropriate conservation actions, and monitoring methods need to be as objective, efficient, and accurate as possible. Passive Acoustic Monitoring, whereby remotely deployed acoustic recorders collect data in unsupervised fashion, is relatively cheap and virtually impervious to observer's bias, whilst also producing potentially perpetual data sets.

Already prominent in marine environments, this monitoring method has become popular in the terrestrial scene, however quantitative estimates are seldom possible as we lack proper scaling factors from cue numbers to animal abundance. Moreover, thanks to technological developments, miniature animal–borne loggers of many sorts, can nowadays deliver information on individual animal features, including vocal behaviour.

Here we show how concurrently equipping a portion of individual animals with miniature animal-borne acoustic recorders (ABARs), and deploying environmental recorders can inform abundance estimates. Studying North Island Brown Kiwi (*Apteryx mantelli*, Bartlett 1851), we found how statistical models informed with individual vocal behaviour recorded through individual loggers convey more consistent and realistic abundance estimates than models lacking this information. Although attaching transmitters is an inherently invasive method, we further detail how iterating its application can lead to fully uninvasive passive acoustic monitoring abundance. This would be possible by informing unmarked populations' estimates with individual behavioural data from marked populations while accounting for environmental covariates. We however invite researchers interested in developing such a monitoring framework to carefully consider the ethical implications of repeated handling for their target species, and provide recommendations on trial experiments to evaluate post–handling vocal behaviour at the individual level.

Overall, combining data from miniature animal–borne acoustic recorders and environmental recorders can deliver more consistent, informed, hence realistic abundance estimates. Furthermore, the use of ABARs provides data for investigating individual identification which, combined with the individual vocal activity, would inform ever more accurate passive acoustic abundance estimates.

6.1 Introduction

Reliable estimates of animal abundance, or population densities, are very important in conservation, as knowing whether populations are growing or declining informs decision makers to better manage the available resources. However, while monitoring animal populations can be challenging and/or very expensive, it should ideally be non-invasive, standardised, and scalable.

As recently as the last couple of decades, we have witnessed the establishment, expansion and consolidation of terrestrial passive acoustic monitoring (PAM, see Sugai et al., 2019), whereby remotely deployed acoustic recording units (ARUs) are programmed to store soundscapes in unsupervised fashion. PAM methods are non-invasive, provide objectively analysable data and, although they potentially produce enormous amounts of data, progress in automatic segmentation (Potamitis et al., 2014; Priyadarshani et al., 2018b) allow for the unsupervised processing of field recordings with continuously improving results (Juodakis et al., 2021b; Priyadarshani et al., 2020). PAM methods can potentially be used to investigate activity patterns, behaviour, species occupancy, richness, distribution, or abundance among others (Browning et al., 2017).

However, to reliably estimate population densities and abundances, we need a way to relate the recorded data to the number of individuals, while accounting for silent individuals (Efford et al., 2009; Pérez-Granados et al., 2019b; Royle et al., 2013; Sebastián-González et al., 2018). There are several statistical methods that can be used to estimate abundance from passively collected acoustic data, although they usually depend on the characteristics of the target study system. A recent review by Pérez-Granados and Traba (2021) evaluated articles estimating abundance from PAM data and provided valuable recommendations for future improvements. For animals that are acoustically individually identifiable, spatially explicit capture recapture methods (SECR — Efford et al., 2009; Royle et al., 2013) provide a solid framework for density estimates. Using the capture history, that is the history of detections at different ARUs of each indentified animal, SECR provides accurate estimates of the number animals in an area. Alternatively, if the acoustic recording system allows for localisation of vocalising animals, distance sampling (Buckland et al., 2012) takes advantage of spatially locating vocalising animals assuming that vocalisations coming from different locations correspond to different animals. Moreover, specialised versions of SECR (Stevenson et al., 2015) can integrate sound localisation to improve precision (Juodakis et al., 2021a).

Without spatial localisation or individual identification, and assuming that vocal activity is dependent on density, approaches include relating the detected vocal activity rate (DVAR, Pérez-Granados et al., 2019c) to animal density, possibly using paired sampling (Van Wilgenburg et al., 2017). Finally, if information of individual vocal activity, or cue rate, of the focal species is available, this can be used to estimate abundance around single ARUs (Sebastián-González et al., 2018).

It is thus important to identify an appropriate relationship between the detected vocal activity rate and the actual number of individuals in the surveyed area (Juodakis et al., 2021a; Marques et al., 2013; Thomas and Marques, 2012).

In the last few years animal–borne microphones (Gill et al., 2016) soon thereafter followed by animal–borne acoustic recorders (ABARs) have entered the terrestrial scene (Stowell et al., 2017), first within behavioural studies (Greif and Yovel, 2019; Wijers et al., 2018), and more recently in conservation–related studies (Reid et al., 2021). Animal-borne sensors have been around since the late 1930s (Kooyman, 2007). ABARs acquire acousic recordings directly from the equipped individuals and have emerged during the so–called golden era of animal-borne sensors, although they are somewhat under–represented compared to other sensors in the literature (Wilmers et al., 2015). Deploying ARUs together with ABARs thus enables the concurrent recording of vocalisations at the individual and community levels. We can then employ modern statistical methods (Royle et al., 2013) to model individual level cues and fathom more informed community abundances than would have been possible in a non–tagged population. Whereas a single use of these devices can provide an abundance estimate for the particular period of sound collection, iteratively employing ABARs can potentially make PAM sufficiently accurate to be the main monitoring method.

In this study we describe a protocol to employ ABARs on potentially any vocal animal system capable of being equipped with similar devices. We exemplify this by presenting a case study featuring North Island Brown Kiwi (*Apteryx mantelli*, Bartlett — hereafter NIBK). NIBK are routinely monitored through specialised aural surveys known as "Kiwi Call Counts" (KCC, Robertson et al., 2017). During KCC, observers listen for two hours per night (at least 30 minutes after sunset but within the first two hours of darkness) over 4-8 nights (Craig, 2018). In the current Kiwi Recovery Plan, the New Zealand Department Of Conservation Te Papa Atawhai identifies inconsistencies between NIBK population densities and KCC results (issue 4.6, Germano et al., 2018), as well as the need for implementing a long–term monitoring programme to evaluate changes in NIBK population abundance (objective 4.1, action 4.2, Germano et al., 2018).

In this study, we present results from abundance estimates on four sampling occasions of three high density NIBK communities (two sampling occasions were in the same population) where we equipped individuals with ABARs, and deployed ARUs. We compared the estimates from two different years and seasons on the same population using three different statistical methods (unmarked populations mark resight, distance sampling, and ABAR–informed mark resight, see section 6.2.7). We further tested whether using information from populations with individuals tagged with ABARs could improve estimates on unmarked populations. We finally discuss current limitations and future perspectives using this method.

6.2 Material and Methods

6.2.1 Ethical Statement

This study has been possible thanks to the approval of Iwi and permits from DOC (Authority Number 63716-FAU), and the Massey University Animal Ethics Committee (protocol ID 20/26)

6.2.2 Study System and Site

We conducted this study at two locations within Aotearoa New Zealand: Ponui (Chamberlin's) Island (Figure 6.2, 175.14860,−36.90422 : 175.21860,-36.82961 − Coordinate Reference System WGS84 for all coordinates herein), in the Hauraki Gulf, Auckland region, and at The Landing (174.079718,−35.174519 : 174.084968,−35.177706), on the Purerua Peninsula, Northland.

Fourteen NIBK were translocated to Ponui Island in 1964 (Miles and Castro, 2000), and the island hosts a high-density NIBK population, which has been under study since 2004 (Colbourne, 2005; Cunningham et al., 2007). On Ponui Island, over the last 17 years, from 30 to 50 birds have been equipped with very high frequency (VHF) radio transmitters that are monitored at least monthly. Having record of the history of these animals, including morphological measurements, usual resting areas, nesting sites, and social relationships (Cunningham and Castro, 2011; Dixon, 2015; Undin et al., 2021) makes of Ponui Island the ideal site to trial the use of ABARs to improve acoustic abundance estimates in NIBK.

In the last DOC KCC report (Craig, 2018), The Landing had the closest number of vocalisations to that of Ponui Island. The NIBK at The Landing are not usually equipped with VHF transmitters. On Ponui Island, we worked in two of the remnant kauri (*Agathis australis*) forest patches located on South Ponui farm: Kauri Bush (175.177635,−36.887127 : 175.182262,−36.882096) and Red Stony Hill Gully (RSHG — 175.180982,−36.892250 : 175.184780,−36.888938). At the time of the surveys, Kauri Bush was home to 14 birds with transmitters, whereas RSHG hosted 17 transmitterised birds (Figure 6.2). Detailed information regarding Ponui Island, its NIBK population, vegetation, and related studies can be found in the following theses: Bansal (2020); Dixon (2015); Strang (2018); Vattiato (2021); Ziesemann (2011).

NIBK are flightless nocturnal cryptic birds belonging to the only Apterigiformes genus, *Apteryx.* Vocalisations of males and females NIBK are quite distinct (see figure 1.2; NIBK tend to be more vocal during the first and last hours of darkness, male vocalisations can be three times as numerous as females ones, and number of vocalisations detected are not linearly related to animal abundance (Miles, 1995). Corfield et al. (2008) describes males as producing long–travelling whistles which could be grouped into phrases, and with fundamental frequencies around $1.5kHz$ (and overtones up to approximately $13kHz$), and females as producing hoarser and lower (fundamental frequency

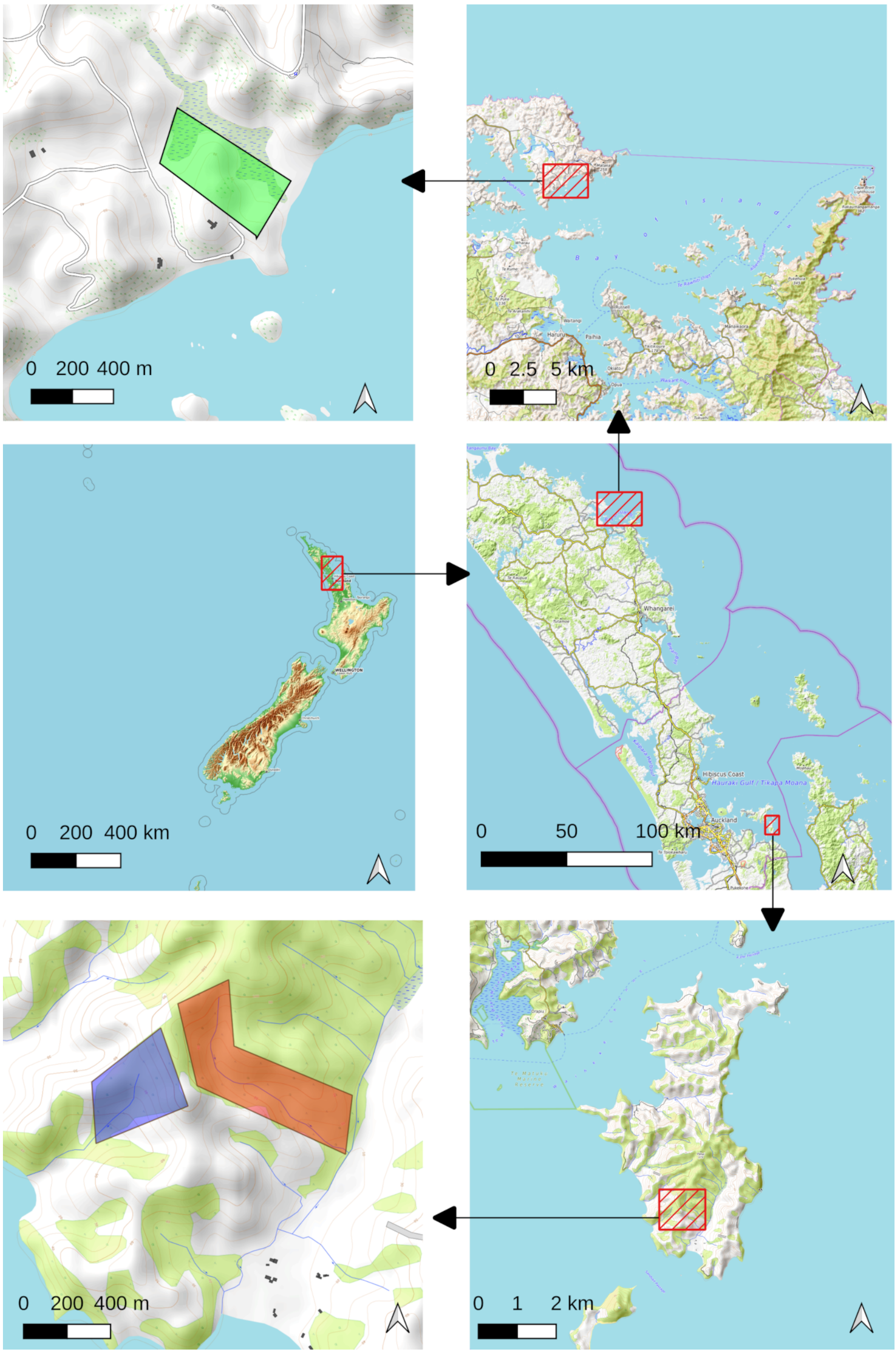


Figure 6.2: Maps indicating the location of the sample populations in relation to Aotearoa New Zealand. The arrow from the central left map leads to a map showing the greather Auckland and Northland regions from which one arrow leads upwards towards the Ipipiri (Bay of Islands) Map and thereby left to the map of The Landing (green shaded polygon), and one arrow leads downwards towards the Ponui Island map and thereby left to the map comprising Kauri Bush (blue shaded polygon) and RSHG (red shaded polygon).

around 0.1 kHz) vocalisations with intensity concentrated in two formants; duets have been suggested to possibly have territorial implications (Corfield et al., 2008).
However, some studied consider NIBK more strictly territorial (McLennan et al., 1987; Taborsky and Taborsky, 1992, 1995, 1999), but others have shown that territoriality may be dependent on population density (Potter, 1989). On Ponui Island, marked birds hold overlapping home ranges of between five and six hectares (Ziesemann, 2011, Chapter 3). Birds roost in burrows or on the surface in forest patches within their range and the choice of roost site is dependent on proximity to foraging areas (Dixon, 2015, Chapter 4). At night kiwi move from the roost sites to foraging sites and are mostly found in forested and scrub areas (Dixon, 2015, Chapter 4, figure 3.6). Over 2377 nocturnal localisations of Ponui NIBK, animals displayed variability in habitat utilisation on the temporal, breeding and social status scales (e.g. birds that tended to shelter in pairs were found more often in forest environments than those sheltering alone or in groups).

6.2.3 Equipment

ABARs

We developed and tested our ABARs starting from a SOROKA 15E (`https://ts-market.com/products/models/19058/`, figure VI.i) digital recorder, to which we soldered an external waterproof microphone, and an Adult Kiwi VHF transmitter (`https://kiwitrack.weebly.com`).

Before assembling the ABARs, we tested the detection range of the ABARs at the natural average amplitude of NIBK (figure 6.3, see Castro et al., 2019, for additional information as well as the test tracks; we measured fast response dbA, 1.5 metres away from the loudspeaker in the 30-130 db range); we played back the vocalisations from on top of the device, one metre, two metres, and five metres away from it, always facing the device as shown in Figure 6.3. We then tested different possible protection materials for the external microphone: neoprene, clear plastic, epoxy resin, and silicone.

The ABARs weigh 21 grams, which is within the acceptable weight of NIBK transmitters (Robertson et al., 2017), and can last up to two weeks in continuous recording setting (12 hrs per day) at 32 KHz frequency range and 16 bit resolution. Although the VHF transmitter and recorder are connected to the same battery, the acoustic recording component (from the SOROKA 15E) conveniently shuts down when the (3.5 V) battery drops below 3 V of produced voltage, whilst the VHF transmitters keep working until the voltage drops to 2.5 V, potentially allowing for the safe recovery of the device even from the more elusive animals. The ABARs were reprogrammed before each fieldtrip.

We tested the detection distance of the ARUs (DOC AR4, 32 KHz frequency range and 16 bit resolution, see a representative frequency response chart in figure II.iii) during a previous

Figure 6.3: Testing the detection range of an ABAR. On the the leftmost leg of the big tripod we secured a SOROKA 15E with external battery and microphone. Beside this, the smaller tripod holds the SPL meter in place whilst a tape mesure runs between a movable aluminium platform (the same as in chapter 3) onto which lays a FoxPro NX4 speaker. We used a compass to have the tape measure always in the same direction for reproducibility in case we needed to repeat the tests. The testing track included two NIBK vocalisations (one per sex). In the actual test we broadcast these vocalisations at natural amplitude (Castro et al., 2019) at increasing distances along the tape measure to ascertain the distance at which a NIBK vocalisation would not be detected by the ABAR.

experiment (Chapter 3) by broadcasting NIBK female and male vocalisations from increasingly larger distances until said vocalisations were not visible on the spectrogram.

We used a FoxPro NX4 speaker for all the broadcasts (frequency response graph figure II.ii) and made all the sound pressure level (SPL) measurements using a Digitech QM1592 Professional Sound Level Meter (see `https://www.jaycar.com.au/pro-sound-level-meter-with-calibrator/p/QM1592` or `https://www.techbrands.com`), which was calibrated before usage with the provided calibrator according to user manual.

6.2.4 Acoustic Sampling

We sampled three areas in two NIBK high density populations in 2020 and 2021. Kauri Bush and RSHG, on Ponui Island, were sampled from the 21st to the 24th of May and from the 6th to the 10th of June 2020 respectively. In 2021 we concurrently sampled Kauri Bush and The Landing (Purerua Peninsula) from the 18th to the 24th of March. Surveys included three types of sound sampling: 1. human listening following KCC (Robertson et al., 2017); 2. four ARUs per area, on the vertices of a 100 m equilateral cross following the results of a previous experiment (Chapter 3); and 3. eight (2020) and seven (2021) birds equipped with ABARs. We followed Colbourne and Digby (2016) and Robertson et al. (2017) and sampled for two hours from 45 minutes after sunset; this period is hereafter referred to as a *night*.

6.2.5 Kauri Bush Dog Sweep Survey

A dog-handler team (Eko and Isabel Castro = the pair) zigzagged Kauri Bush from top to bottom searching for kiwi on June, the 5th 2020. A second person (Stephen Marsland) maintained visual contact with the pair at a distance to avoid interfering with the search, but to be able to indicate to the dog handler any areas that may have been missed by the pair; these areas were then checked to fully examine all the gully. Each time the dog indicated the presence of a bird, this was visually observed and a telemetry set was used to ascertain whether the bird had a transmitter. All transmitter birds were accounted for using radio telemetry to check the accuracy of the dog in finding the kiwi.

6.2.6 Data extraction and ABARs vocal ownership

We used AviaNZ (Marsland et al., 2019) to process all the data, both from the ABARs and ARUs. We then reviewed the resulting labels to ascertain that all the labelled segments were actual NIBK vocalisations, removing the segments incorrectly labelled as NIBK (false positives). As ABARs would potentially record vocalisations from both the bird wearing them and nearby ones, we needed an objective way of assigning vocalisation ownership. We adapted an AviaNZ function (as the program is open source) to perform the following steps:

1. within each ABAR detected vocalisations:
   (a) we removed any vocalisation that did not match the sex of the bird wearing the ABAR
   (b) for the remaining vocalisations, we calculated each individual syllable's signal to noise ratio (SNR)
   (c) we calculated the mean minimum SNR (A)
   (d) we identified the strongest (highest median SNR) vocalisation and saved its minimum SNR as (B)
   (e) we then discarded all the vocalisations which maximum SNR was both below (B), and which median was below (A)
2. since this would have left at least one vocalisation per device, regardless of overall quality (i.e., even if a device only recorded strangers' vocalisations), looking at all the vocalisations left (among ABARs), we kept only those where median was greater than the overall mean minimum and where minimum was above 0
3. we additionally asked a NIBK expert (Isabel Castro), blind of our division, to allocate calls to the bird wearing the ABAR.

6.2.7 Analyses

In order to evaluate whether using ABARs informed more accurate abundance estimates we designed a three stages analysis.

We first employed two separate models for the community level vocalisations from each sampling occasion; we used custom versions of spatial capture recapture (SCR) models, specifically one for unmarked populations and the other for populations with a known number of marked individuals in the population (see Chapters 18 and 19 of Royle et al., 2013, for detailed examples). The former comprises a first loop to *recruit* a number of individuals through a dedicated Bernoullian variable, and simulate a Poisson distributed number of cues per individual, and a second loop to limit the number of total cues to that of the actually observed ones: at the end of the simulation, the distribution of the number of *recruited* individuals is informative of an expected abundance. The latter additionally has a loop to model the observed cues from marked individuals, which covariates will then affect the *recruited* individuals, then proceeding like the previous model, hence effectively accounting for the distribution of observed individual cues to inform that of simulated/*recruited* ones. The main assumption imposed by the unmarked population model is that all animals in the area vocalise; the removing of this by the added loop in the ABAR–informed model (where animals that do not vocalise exist), additionally allows for the simulation of silent recruited individuals, which should enable more realistic estimates of abundance.

Secondly, we estimated distances from vocalising animals from the same location on two different sampling occasions (May, 2020 and March, 2021) and compared the estimates from the previous models to those from a distance sampling model (Buckland et al., 2012; Royle et al., 2013, Chapter 4). We did so by using the method detailed in Chapter 3 which, using energy as an approximate for distance, provides localisation estimates using single–microphone ARUs.

Finally, we further customised the ABARs–informed SCR model to utilise the information from populations with marked individuals (reference populations) to infer abundances in unmarked populations (target populations). To do this, the first loop of the model ran with data from individuals with ABARs from the reference populations, whilst the second loop (to recruit individuals), ran for both the reference and target populations. At the end of the model the numbers of recruited individuals from the reference populations were added to the numbers of tagged individuals, whilst those from the target populations were not (technically they were added to zeroes, see appendix VI.iii for the full model). All the models are provided as supplementary material and both models and data are available at `http://github.com/albertodiegoderosa/abars`.

We ran all the models using `JAGS` (Plummer et al., 2003). After 400,000 burn-in iterations, each model ran for 800,000 iterations retaining one every 40 values on four independent chains, effectively resulting in 80,000 final draws per observed variable.

6.3 Results

All eight ABARs worked properly during the 2020 sampling in Kauri 2020, while two failed in RSHG. These two ABARs worked as per programmed schedule, but one of them recorded only noise, while the memory card of the other one was unreadable, possibly damaged by water (the device still worked afterwards). The protocol followed to ascribe vocalisations recorded by the ABARs to the bird wearing the ABAR proved very effective as:

- it succeeded in selecting the same vocalisations as the NIBK expert (Isabel Castro) did.
- the number of vocalisations from the kiwi wearing the ABAR recorded by each device differed between animals (figure VI.ii shows proportions of owner versus other kiwi vocalisation in 2020).

In 2021, three ABARs failed in Kauri bush, while two failed at The Landing. All of these recorded only noise. This left us with a total of twenty-four ABARs that delivered individual recordings.

There were no problems with the ARUs.

During the survey with the help of the kiwi dog, we found a total of seventeen birds in Kauri Bush; the dog missed one of the transmitterised birds. This establishes the lower bound for the expected number of birds in Kauri Bush at eighteen.

At the community level, the ARUs on average detected 2.75 (standard deviation, $\sigma = 1.71$) female vocalisations and 6.75 ($\sigma = 4.65$) male vocalisations from Kauri Bush in 2020, 6.14 ($\sigma = 3.24$) female vocalisations and 9.86 ($\sigma = 4.74$) male vocalisations from Kauri Bush in 2021, 9.5 ($\sigma = 5.97$) female vocalisations and 16.5 ($\sigma = 5.2$) male vocalisations from RSHG, and 0 ($\sigma = 0$) female vocalisations and 0.43 ($\sigma = 0.79$) male vocalisation from The Landing (Table 6.1 and figure 6.4).

On the individual level, each ABAR on average detected 0.31 ($\sigma = 0.48$) female vocalisations and 0.38 ($\sigma = 0.5$) male vocalisations from Kauri Bush in 2020, 1.21 ($\sigma = 0.89$) female vocalisations and 0.5 ($\sigma = 0.65$) male vocalisations from Kauri Bush in 2021, 0.25 ($\sigma = 0.46$) female vocalisations and 0.56 ($\sigma = 0.51$) male vocalisations from RSHG, and 0 ($\sigma = 0$) female vocalisations and 0.04 ($\sigma = 0.19$) male vocalisations from The Landing (table 6.2 and figure 6.5).

Table 6.1: Number of vocalisations detected by the ARUs per occasion per sex per night.

Occasion	Sex	Night						
		1	2	3	4	5	6	7
Kauri 2020	F	2	3	1	5			
	M	7	5	2	13			
Kauri 2021	F	12	4	8	6	7	3	3
	M	17	10	15	10	6	7	4
RSHG	F	17	11	7	3			
	M	20	17	20	9			
The Landing	F	0	0	0	0	0	0	0
	M	1	0	0	0	0	2	0

All the models reached satisfactory chain convergence as confirmed by both visual inspection and the Gelman-Rubin scale reduction factors ($\hat{R}$). All the $\hat{R}$ were in fact very close to, or exactly, one, with no value above 1.05.

In the first stage of our analyses (figures 6.6 to 6.9), all the ABAR–informed models returned higher expected numbers of individuals (both males and females) than the Unmarked populations models, with all posterior distributions for the estimated number of animals to some extent overlapping between each pair of models (figures 6.6 to 6.9, table 6.3).

In the second stage of our analyses, comparing the estimated number of birds in the same population over two different sampling occasions with three methods, the ABAR–informed models were the most consistent, followed by the unmarked population models and finally by the distance sampling models estimates (Figure 6.10 and table 6.4). When combining data from both sexes, the ABARs–informed models further resulted as the more realistic, with median estimated values closer to the findings from the dog sweep survey (Figure 6.11).

The numbers of males and females estimated from the models in stage three in RSHG were at

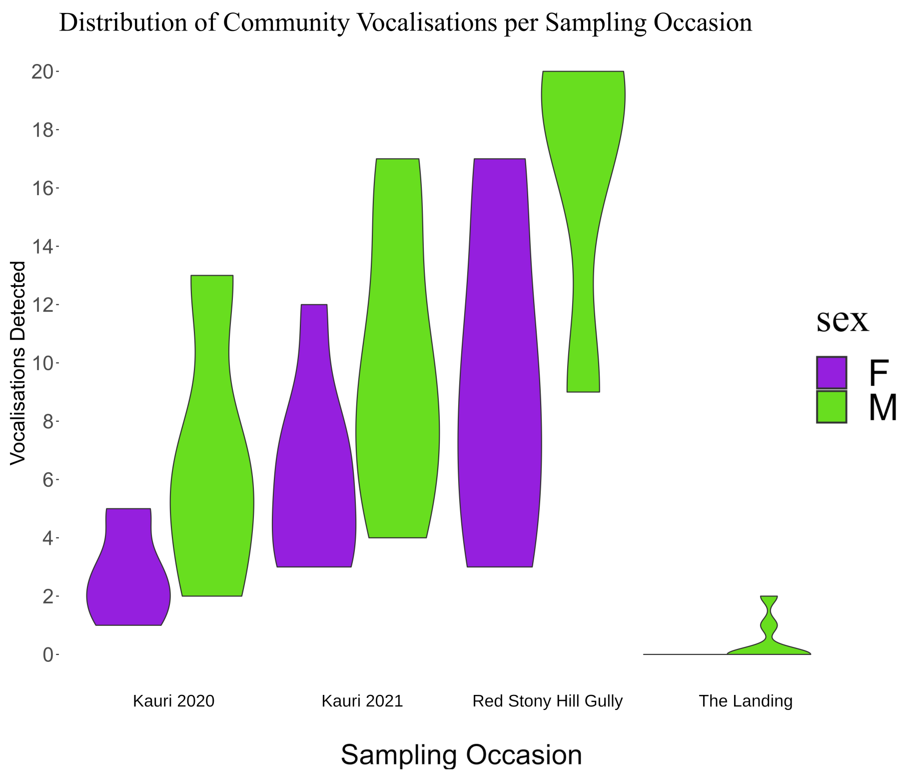


Figure 6.4: Community Female and Male Vocalisations detected over the four surveys. The higher the number of vocalisations detected per night by the ARUs, the wider the shape in the plot. For instance in the 2020 Kauri bush survey, there were more nights with only two female vocalisations per night than four.

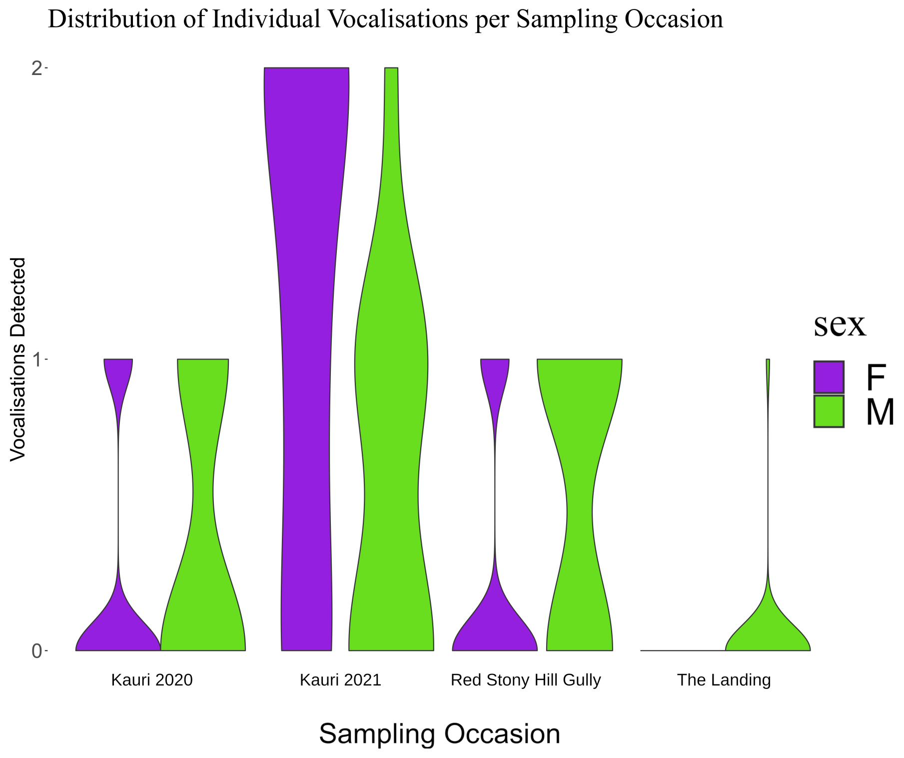


Figure 6.5: ABARs Female and Male Vocalisations detected over the four surveys.

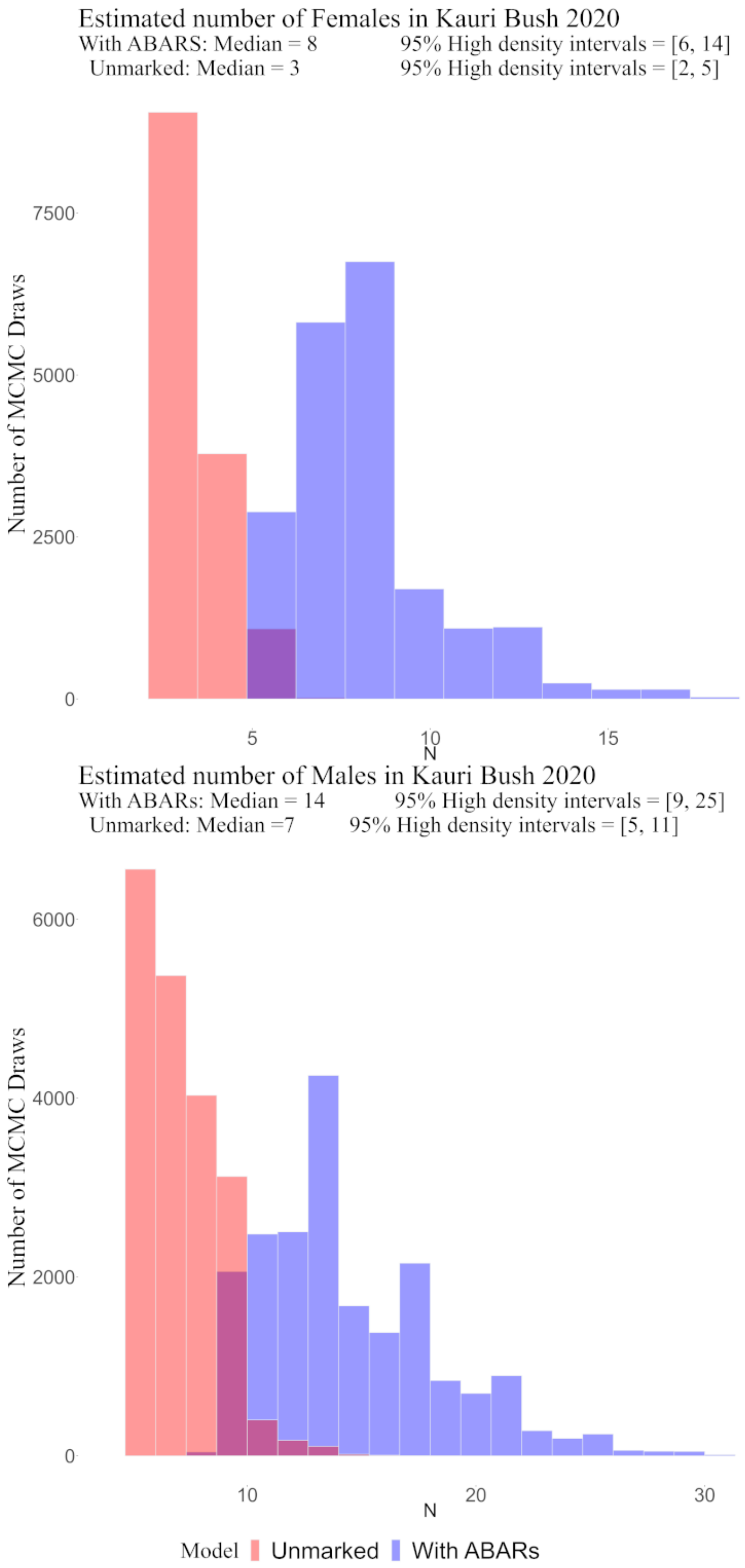
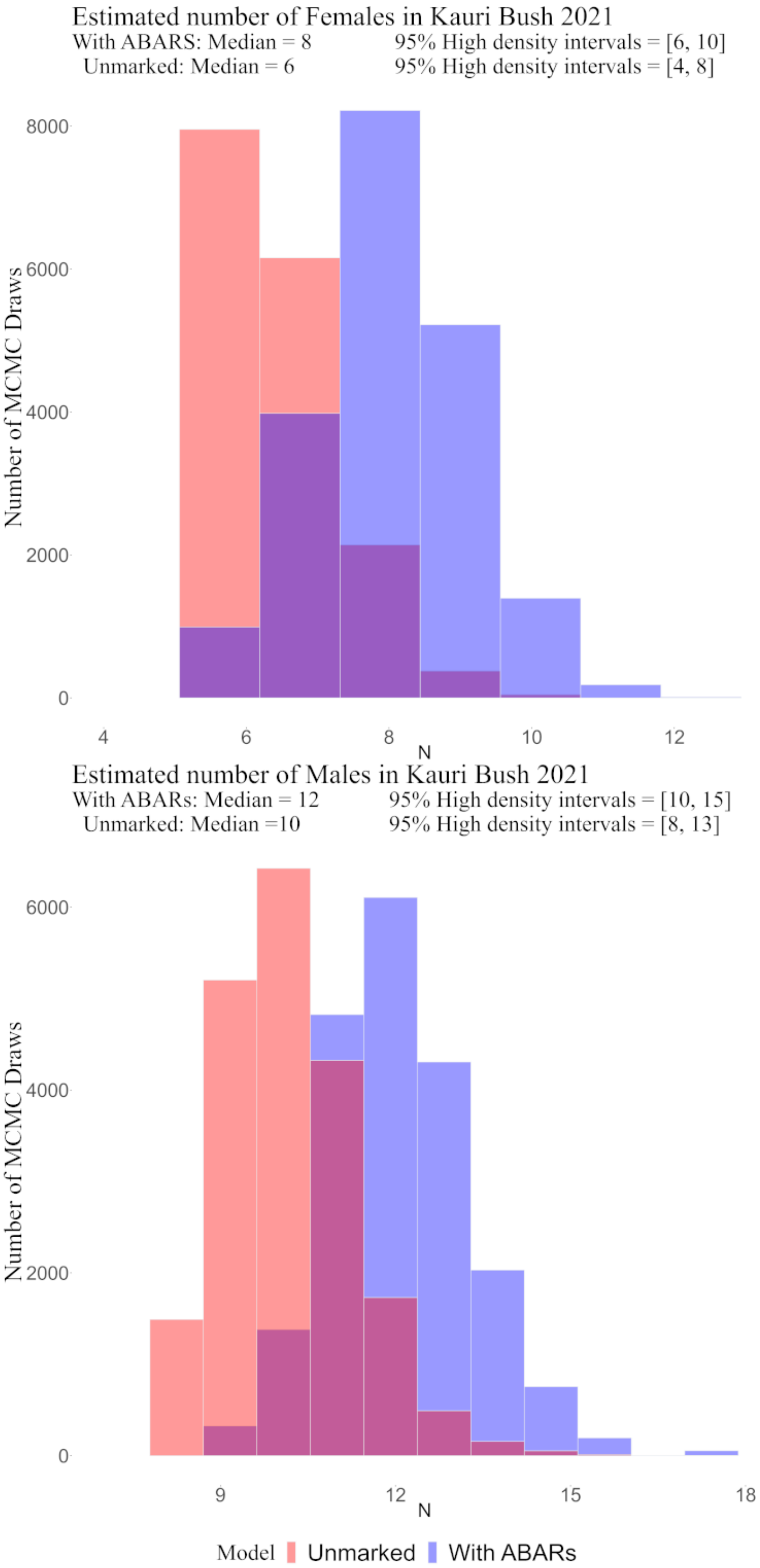


Figure 6.6: Single population models: Unmarked versus ABAR–informed estimates for Kauri Bush 2020. Where the two models overlap in estimated population size, the model with the smaller occurence is represented in a darker intermediate colour. ABAR = Animal–Borne Acoustic Recorder, MCMC = Markov chain Monte Carlo.



Estimated number of Females in Kauri Bush 2021

With ABARS: Median = 8 95% High density intervals = [6, 10]

Unmarked: Median = 6 95% High density intervals = [4, 8]

Estimated number of Males in Kauri Bush 2021

With ABARs: Median = 12 95% High density intervals = [10, 15]

Unmarked: Median =10 95% High density intervals = [8, 13]

Figure 6.7: Single population models: Unmarked versus ABAR–informed estimates for Kauri Bush 2021. Where the two models overlap in estimated population size, the model with the smaller occurence is represented in a darker intermediate colour. ABAR = Animal–Borne Acoustic Recorder, MCMC = Markov chain Monte Carlo.

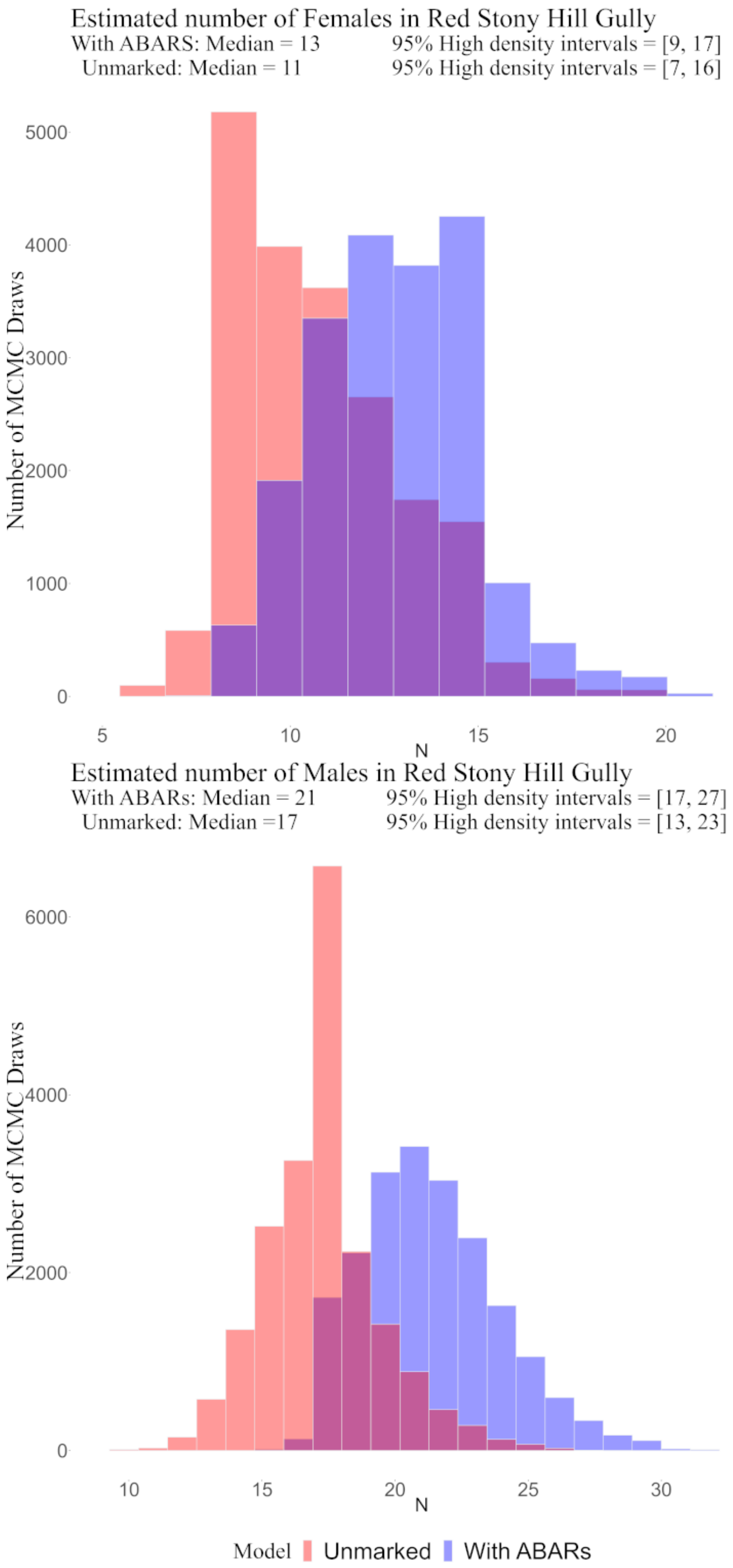


Figure 6.8: Single population models: Unmarked versus ABAR–informed estimates for Red Stony Hill Gully. Where the two models overlap in estimated population size, the model with the smaller occurence is represented in a darker intermediate colour. ABAR = Animal–Borne Acoustic Recorder, MCMC = Markov chain Monte Carlo.

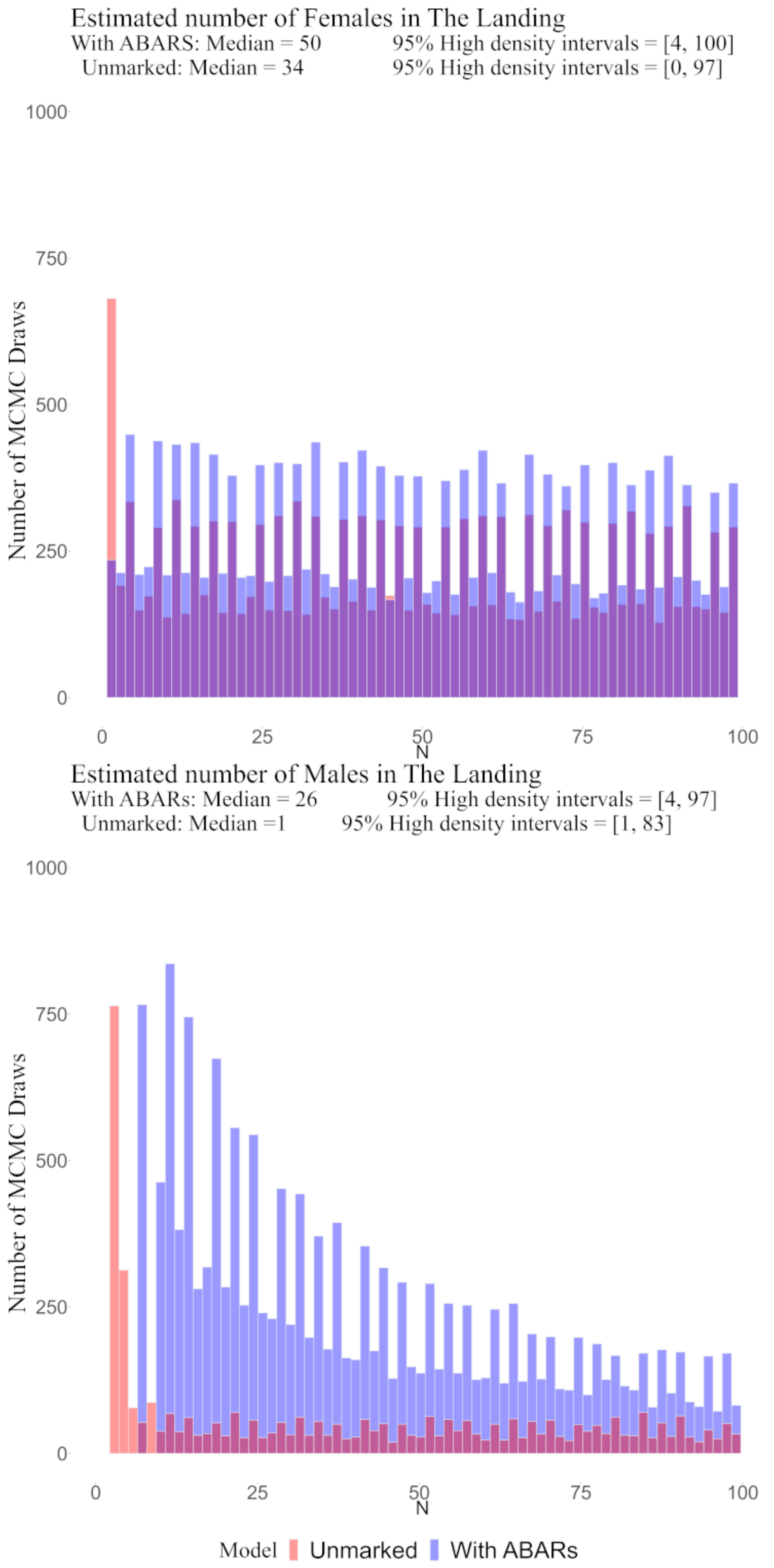


Figure 6.9: Single population models: Unmarked versus ABAR–informed estimates for The Landing. Where the two models overlap in estimated population size, the model with the smaller occurence is represented in a darker intermediate colour. ABAR = Animal–Borne Acoustic Recorder, MCMC = Markov chain Monte Carlo.

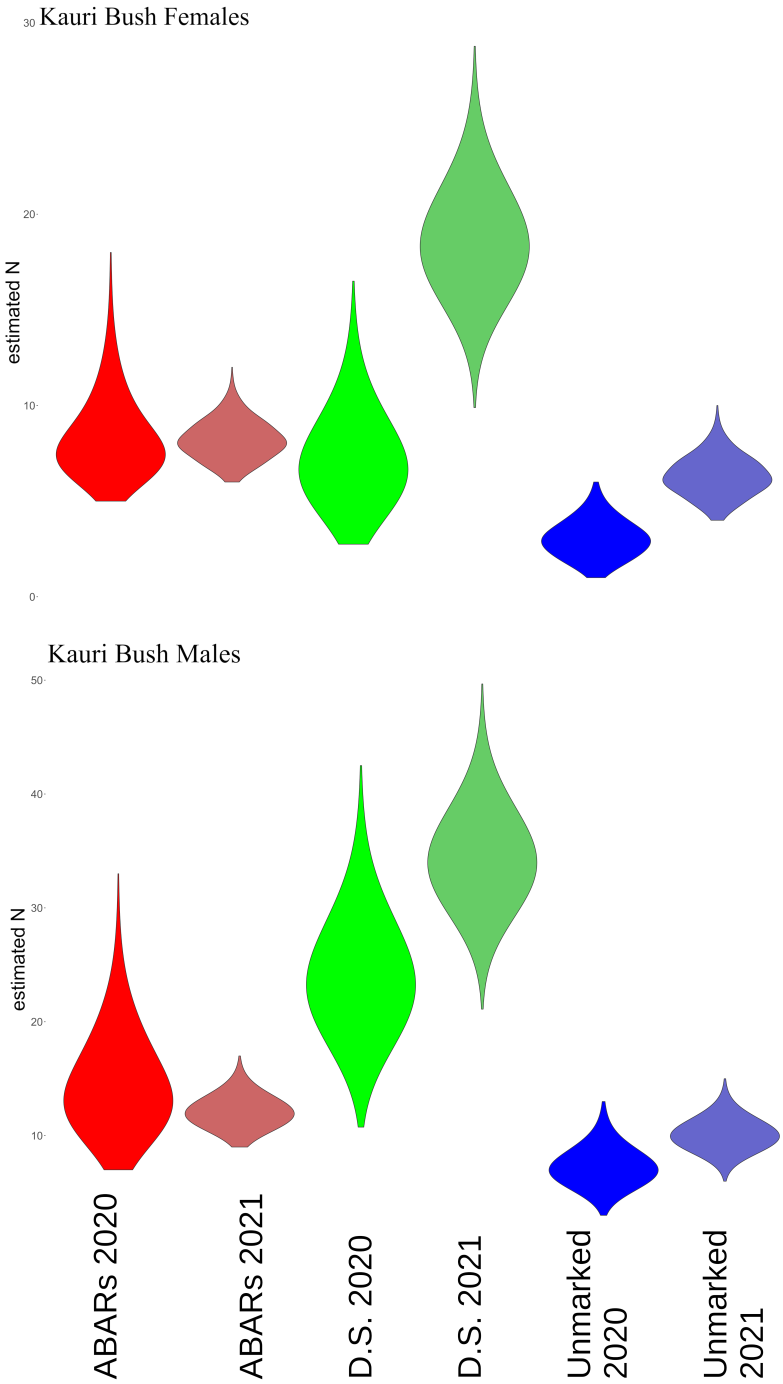


Figure 6.10: ABARs Estimated Females (top) and Males (bottom) NIBK in Kauri Bush over two sampling occasions with three methods: ABARs (Red), Distance Sampling (D.S. — Green), and Unmarked (Blue). The brighter variants (left of each pair) of each colour represent the 2020 May survey, whereas the dimmer ones (right of each pair) represent the 2021 March survey.

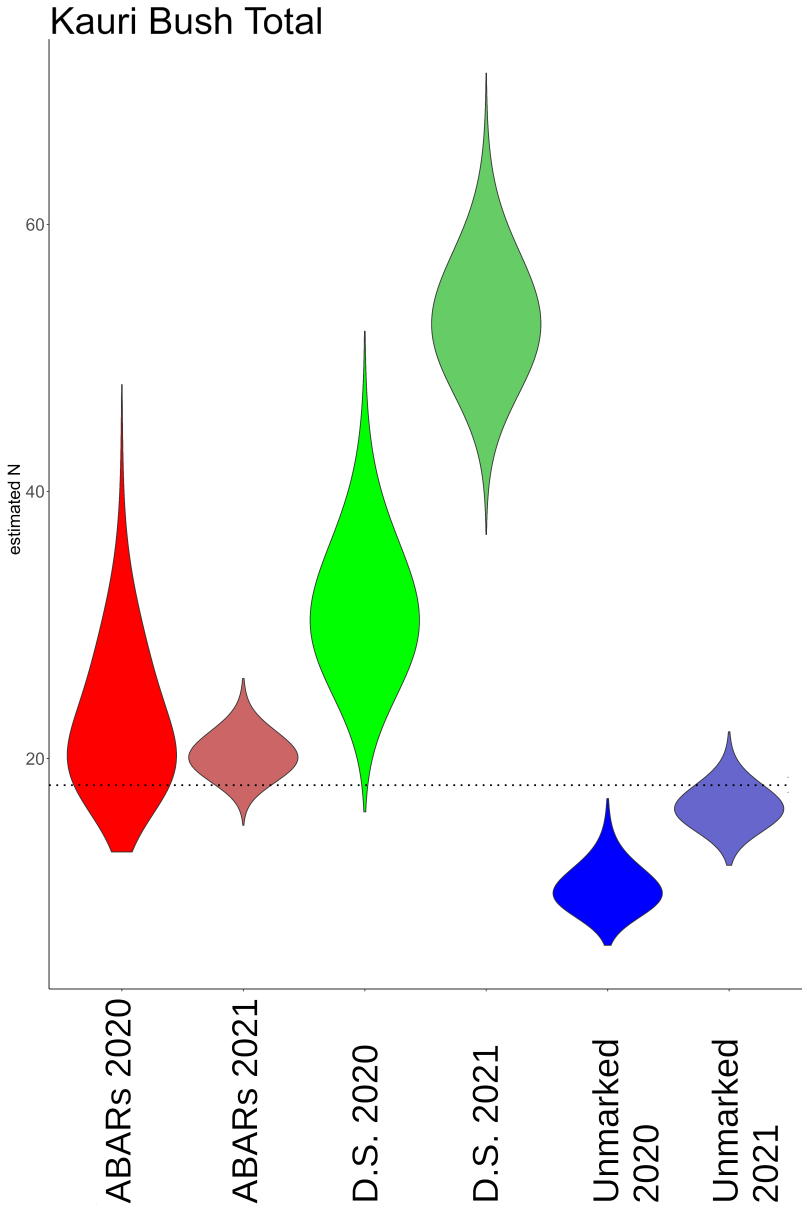


Figure 6.11: ABARs Estimated total NIBK in Kauri Bush over two sampling occasions with three methods: ABARs (Red), Distance Sampling (D.S. — Green), and Unmarked (Blue). The brighter variants (left of each pair) of each colour represent the 2020 May survey, whereas the dimmer ones (right of each pair) represent the 2021 March survey. The dotted line represents the number of kiwi found during the dog sweep survey on 5th June, 2020.

Table 6.2: Number of vocalisations detected by the ABARs per occasion per sex per night.

Occasion	Sex	Night						
		1	2	3	4	5	6	7
Kauri 2020	F	0	0	0	0			
		0	0	0	1			
		0	0	1	0			
		1	1	0	1			
	M	0	1	1	1			
		0	0	0	1			
		0	0	0	0			
		1	0	0	1			
Kauri 2021	F	2	2	2	2	0	2	1
		2	2	0	0	0	1	1
	M	2	1	1	1	1	0	1
		0	0	0	0	0	0	0
RSHG	F	0	0	1	1			
		0	0	0	0			
	M	1	1	0	1			
		0	0	1	0			
		0	1	0	1			
		1	1	1	0			
The Landing	F	0	0	0	0	0	0	0
		0	0	0	0	0	0	0
	M	1	0	0	0	0	0	0
		0	0	0	0	0	0	0
		0	0	0	0	0	0	0
		0	0	0	0	0	0	0

an intermediate level between the those resulting from the Unmarked and ABARs model from the same population (figures 6.12 and 6.13 and table 6.5). The posterior distributions of the covariates from these models were more similar to those from the fully ABARs–informed models (Figure 6.14). Tables with all the model–specific covariates posterior high density intervals are available as supplementary material.

Figure 6.12: Estimated Females (top) and Males (bottom) NIBK individuals in RSHG with differently informed models. The leftmost estimate (red) represents the Unmarked models (with no ABARs information), the second from the left (green) is informed with individuals from Kauri Bush 2020, the third from the left (cyan) is informed with individuals from Kauri 2020 and Kauri 2021, and the rightmost one is informed with ABARs from the respective target population.

Figure 6.13: Estimated Females (top) and Males (bottom) NIBK individuals in The Landing with differently informed models. The leftmost estimate (red) represents the Unmarked models (with no ABARs information), the second from the left (green) is informed with individuals from Kauri Bush 2020, the third from the left (cyan) is informed with individuals from Kauri 2020 and Kauri 2021, and the rightmost one is informed with ABARs from the respective target population.

Table 6.3: Estimating populations with unmarked populations and ABAR–informed models.

Sampling Occasion	Sex	Estimated Abundance					
		Unmarked			ABAR–informed		
		1st Quantile	Median	3rd Quantile	1st Quantile	Median	3rd Quantile
Kauri 2020	F	2	3	5	4	8	8
	M	5	7	11	9	14	25
Kauri 2021	F	2	5	6	6	8	10
	M	8	10	13	10	12	15
RSHG	F	7	11	16	9	13	17
	M	13	17	23	17	21	27
The Landing	F	0	34	97	4	50	100
	M	1	1	83	4	26	97

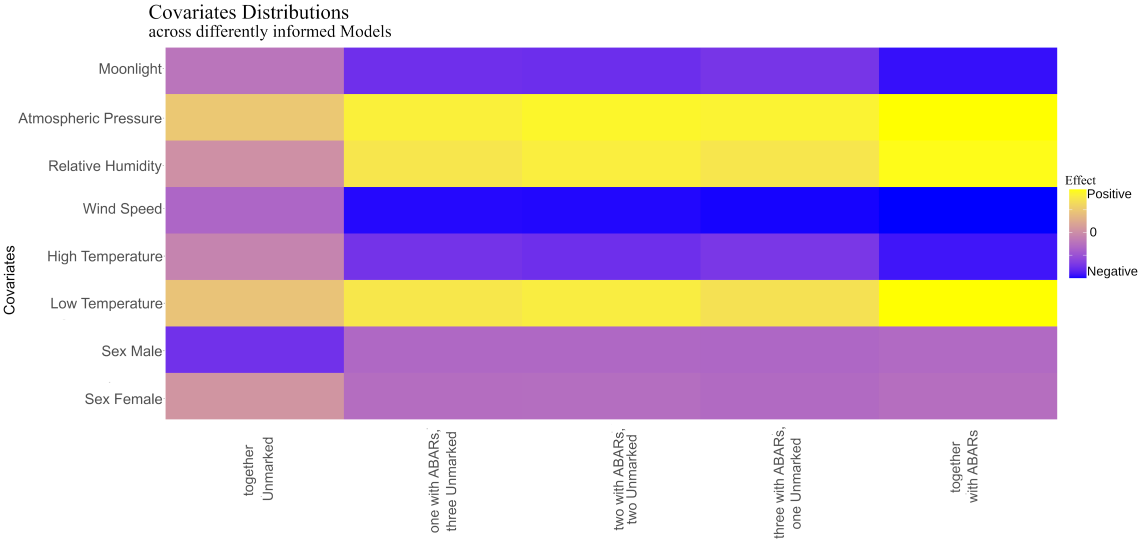


Figure 6.14: Effect of the covariates on the number of individual vocalisations recorded by ABARs . Posterior distributions range from negative (blue) to positive (yellow). All intermediate colours indicate posterior distributions mostly centered around zero (no effect of the covariate on the number of vocalisations). The brighter the yellow, the higher the proportion of the covariate distribution that fell above zero showing a positive effect on the number of vocalisations. Conversely the brighter the blue the higher the proportion of the distribution that fell below zero for that covariate and therefore showing a negative effect on the number of vocalisations. ABARs = Animal-Borne Acoustic Recorders

Table 6.4: Estimating Kauri Bush Population with three methods over two sampling occasions.

Method	Year	Females				Males			
		1st Quantile	Median	Mean	3rd Quantile	1st Quantile	Median	Mean	3rd Quantile
ABARs	2020	7.00	8.00	8.48	9.00	12.00	14.00	14.78	17.00
	2021	8.00	8.00	8.17	9.00	11.00	12.00	12.08	13.00
Distance Sampling	2020	5.50	7.00	7.25	8.50	20.75	23.50	24.03	27.00
	2021	16.78	18.44	18.62	20.22	31.67	34.11	34.33	36.78
Unmarked	2020	2.00	3.00	3.04	4.00	6.00	7.00	7.30	8.00
	2021	6.00	6.00	6.29	7.00	9.0	10.0	10.1	11.0

Table 6.5: Results from estimating the number of NIBK in RSHG and The Landing with progressively more informed models.ABARs = Animal-Borne Acoustic Recorders, NIBK = North Island Brown Kiwi, pop./pops. = population/populations, Q. = quantile, RSHG = Red Stony Hill Gully

Model	RSHG								The Landing							
	Females				Males				Females				Males			
	1st Q.	Median	Mean	3rd Q.	1st Q.	Median	Mean	3rd Q.	1st Q.	Median	Mean	3rd Q.	1st Q.	Median	Mean	3rd Q.
Unmarked	9	10	11	12	15	17	17	18	0	22	32	60	1	1	6	1
One pop.	10	12	12	13	17	19	19	21	0	15	26	67	2	49	46	77
Two pops.	10	12	12	14	17	19	20	21	1	16	26	44	6	51	47	78
ABARs	13	14	15	16	20	22	23	24	9	23	31	47	34	55	56	79

6.4 Discussion and Conclusions

PAM has emerged as a cost-effective and objective way of sampling for animal vocal cues (Pérez-Granados and Traba, 2021; Sugai et al., 2019). Nevertheless, more often than not we lack consistent means to relate detected vocal activities to number of animals in the area of detection of one's devices (Thomas and Marques, 2012). The use of SCR models together with ABARs can potentially accomodate a wide variety of variables and has the great advantage of being naturally scalable to multiple concurrent detectors in a given area (Royle et al., 2013).

The data from our four sampling occasions shows quite distinct numbers of vocalisations in the considered time window, both within (Kauri Bush 2020 versus 2021) and between populations. Interestingly, when looking at the distributions of the vocalisations in Kauri 2020 and RSHG, these look quite different at the community level, yet very similar at the individual level. This is quite possibly a result of these two gullies being neigbouring communities that were sampled only at a week apart. Further considering that birds from these two communities have been previously shown to display similar habits in terms of habitat utilisation (Dixon, 2015) reinforces the idea that the RSHG population comprises more birds than the population at Kauri Bush, at least within the detection of our ARUs.

The fact that the results from our first stage of analyses show the ABAR–informed models always estimating higher expected number of birds when compared to the Unmarked populations models is not surprising. As mentioned above one of the assumptions overcome by the ABAR–informed models is the detection of all the individuals in the area as this adds information about the silent animals. Unfortunately, the birds at The Landing were silent during the sampling window and although the ABARs to some extent compensated for the lack of vocalisations, all estimates were quite uncertain, as shown by the width of all this population's high density intervals (figure 6.4). This could be a location–specific factor, but critically points to the need to iterate the use of ABARs in managed populations, thus obtaining multiple samples over different years and seasons to better interpret the results in unmanaged populations.

The results from the second stage of our analyses shed some more light on the usefulness of the ABARs. Similarly to our previous consideration regarding the assumption of detecting all animals with the Unmarked populations model, here we see how the ABAR–informed models provide more consistent inferences than the other considered methods. The Unmarked populations model still returned consistent estimates over the two years, however underestimating the number of overall individuals, especially females. This is confirmed by the ascertained lower bound of eighteen individuals in Kauri Bush with the help of a Kiwi–finding dog shortly after the 2020 sampling, as this number sits outside the combined high density intervals for the Unmarked populations model (table 6.4). The distance sampling model returned the least consistent estimates, possibly because this assumed each vocalisation to come from a different individual, but also because of seasonal changes in behaviour (Nijman, 2007; Zuberogoitia et al., 2019).

Finally, the models including one and two *informing* populations with ABARs to inform abundance estimates on unmarked populations delivered some compelling results. Using data from informing populations to inform estimates in target populations resulted in these being closer to those resulting from ABAR–informed models for the target populations. Interestingly, median values for all but females at The Landing were intermediate between those from the Unmarked populations models and ABAR–informed models. This is very likely because of the total absence of female vocalisations from The Landing both at the community and individual level. Consequently it is understandable how the models informed with data from populations where females did vocalise, may tend to lower the expected values on populations with no recorded vocalisations. Additionally, it is quite interesting to see how the posterior distributions for most covariates of all these partially informed models markedly resemble more distribution of the full ABAR–informed model than the full Unmarked populations model. Although these covariates convey quite similar results to recent experiments (Chapters 4 and 5), they do not have real explanatory value on NIBK behaviour, as they are based on the projection of the behaviour of a few individuals on populations of estimated sizes. Nevertheless, their distribution is reasonably promising in regard of the number of sampling occasions needed to reach satisfactory results with these models (Pérez-Granados et al., 2019c).

According to the protocol proposed by Pérez-Granados et al. (2019b), and having previously evaluated the detection distance of our ARUs for the species of interest (300 m, see page 51), we have here approached the derivation of vocal activity rate for a cryptic and nocturnal species in the time window suggested by Colbourne and Digby (2016) as the optimal listening time–frame for NIBK acoustic monitoring. Thanks to the manual revision of batch–processed recordings in AviaNZ (Marsland et al., 2019) we know we have no false positives and would thus be in the fourth step of the protocol to derive consistent values to generalise the use of PAM for robust estimates of NIBK abundance Pérez-Granados et al. (2019b). However, the models are currently limited by the number and diversity of sites sampled, as shown by the high uncertainty in the estimates for the Northland site (The Landing). The number of further sampling occasions to reach reliable NIBK abundance estimates with this method (Pérez-Granados et al., 2019b, suggest at leat 20-30 sites to derive reliable VAR, and ABARs play an analogous role) would depend on the performance of the models and would ideally be identified when a difference in median estimated abundance between partially informed and ABAR–informed models smaller than a small number ϵ, should be reached through the iterative use of ABARs in managed populations.

Here we showed how employing ABARs in conjunction with ARUs and SECR can lead to more realistic and consistent estimates of abundance. However, using the ABARs in their current state, managers would see an an initial increase in the invasiveness of the procedures, as the animals would need to be handled more often, which would entail careful ethical considerations depending on the target species. This would also involve including a number of different populations, conditional to the heterogeneity of the target species' community, to allow for confident inferences on any other population. As such, one final consideration needs to address possible animal welfare issues. Although we know from a recent study (Chapter 5) that we have found no detectable difference between the vocal behaviour of a NIBK community that had been handled and one that is to date naïve to investigator's disturbance, there is no knowledge of how this could affect individual birds. In order to obtain such information it would be necessary to equip some animals from two different populations with ABARs programmed to start recording at a later date for a given number of days (e.g. eight); after half of the recording days would have passed, researchers would need to re–capture half of the animal and realease them shortly thereafter. The resulting comparison of vocal activity would give a quite strong indication on whether handling would represent a tolerable source of stress for the animals.

Alternatively and possibly more desirably, further tests and collaborations with transmitter–building companies could lead to replacing normal VHF transmitters with ABARs, programmed to record during the national KCC scheme and deploy ARUs programmed accordingly. This option would see no changes in the operations conducted by Kiwi practitioners tending to managed populations, although it would bring a wealth of information with the real possiblity of accurate abundance estimates based on passive acoustics.

To conclude, ABARs can be a reliable source of information to derive consistent abundance estimates for vocal animals that are able to carry them, nonetherless we would advise serious consideration of ethical implications appropriate to the target species. We are confident that future technical and technological developments will allow for more accurate and cost-efficient applications of these devices to inform more effective conservation efforts.

Chapter 7

Conclusions and Recommendations

Figure 7.1: Stephen in the field with Virginia and Kayla. Picture by Lucas Mugnier.

7.1 Conformity with the objectives of the thesis

Understanding the magnitude of biodiversity changes is crucial for policy makers (Brondízio et al., 2019). Reliable animal population monitoring practices are critical to enable conservation decision makers to properly allocate resources (IUCN, 2020). It is important to improve our wildlife monitoring techniques employing modern technologies and data analysis techniques (IUCN, 2021) to provide more robust yet cost–efficient species density estimates.
In Aotearoa New Zealand, the New Zealand Department of Conservation Te Papa Atawhai (DOC) defines the development of an objective and efficient monitoring framework for the iconic North Island Brown Kiwi (NIBK — *Apteryx mantelli*, Bartlett 1851) as critical in their 2018–2028 recovery plan (Germano et al., 2018). NIBK are currently monitored through specialised aural surveys, known as *Kiwi Call Counts* (KCC), where observers listen for NIBK vocalisations over two hours and annotate estimated distance, direction, and sex of all the detected vocalisations (Robertson et al., 2017). Besides the inconsistencies specifically identified by DOC in the numbers of vocalisations from KCC identified by DOC (Germano et al., 2018), other studies have found a high degree of variability in the data delivered by observers–reliant aural surveys, mostly depending on experience and hearing acuity (see chapter 1, section 1.1.2). The first objective of this thesis was to identify an inexpensive way of localising vocalising animals using ARUs, to minimise both bias (e.g. Alldredge et al., 2007) and cost (e.g. Williams et al., 2018). Using playback (chapter 4), we found that it is possible to use single microphone ARUs to localise vocalising animals with a reasonable degree of uncertainty depending on the environment. In chapter 6 we further applied this to inform a distance sampling model for the birds in Kauri bush (on Ponui Island). In accordance to Darras et al. (2018a), as ARUs and human observers have different detection ranges, it would be necessary to standardise the detection distance between ARUs–based and human–reliant KCC to integrate the results from these two sources. For historical data, this could be done by filtering out all the human detections above the ARUs detection range threshold (300 m in the case of the DOC AR4), whereas future applications could also deploy a grid of ARUs covering the presumed human detection radius around KCC listening stations.

Relating acoustic monitoring data to animal abundance can be particularly challenging because animals do not have a fixed call rate, and the playback technique has been reported as a possible tool to enhance or standardise detections (Evans et al., 2007; Okahisa et al., 2016) This technique is also familiar to Kiwi practitioners as it has been reported as a tool to enhance responses of NIBK during aural surveys (McLennan, 1992). The second objective of this thesis was to assess the suitability of the playback technique as a tool in NIBK acoustic monitoring. Over two different seasons, we conducted a formal evaluation of the degree to which NIBK in a high density population would respond to playback (chapter 4). We found that NIBK did not consistently respond to playback more than to control sounds. Although this may be different with populations of different density, our results discourage the use of playback within an acoustic monitoring framework with NIBK.

Recent technological advances in miniaturisation of electronic devices enabled animal–borne loggers have become to became a superlative source of data, as this is acquired directly from individual animals (Wilmers et al., 2015). These include animal-borne acoustic recorders (ABARs), which have quite recently entered the terrestrial scene (Gill et al., 2016; Stowell et al., 2017), and even more so employed for conservation–related purposes (Wijers et al., 2020). However, employing animal–borne devices entails repeated handling, which poses questions both for animal welfare issues (Curzer et al., 2013; Soulsbury et al., 2020), and for data reliability due to possible post–handling behavioural changes (Toy and Toy, 2021). The third objective of this thesis was to investigate whether NIBK communities with individuals that have been recently handled displayed a different vocal activity than communities of birds naïve to investigator's disturbance. Over a period of three weeks spanning before and after the annual transmitter change of the birds on Ponui Island, we deployed ARUs in both a gully inhabited by a large contingent of transmitterised birds, and one where birds are known to live but researchers have never worked (chapter 5). We found no detectable differences in the vocal activity of the two communities before or after the annual transmitter change. Although this is reassuring in terms of post–handling acoustic data reliability, it cannot be conclusive in regards of possible individual changes and in chapter 6 we propose an elucidative experiment using ABARs (see page 136).

Decision makers need robust estimates of animal density in order to support sensible conservation actions. PAM is cost–effective, uninvasive, and impervious to observer's bias, and also produces potentially permanent data records (Sugai et al., 2019). Nevertheless, density estimates based on PAM require a way to translate vocalisations' densities into animals' densities. The final objective of this thesis was to evaluate the use of ABARs concurrently to PAM surveys to estimate NIBK abundance, hence density. Over two years and four total sampling occasions, we acoustically sampled NIBK communities and estimated animal abundance using different acoustic methods (chapter 6). Abundance estimates using custom spatial mark–resight models informed with ABARs data were more accurate and robust over time than unmarked population models and distance sampling models. Furthermore, estimates of unmarked populations informed by ABARs estimates from other populations delivered encouraging results for eventually reaching fully uninvasive PAM abundance estimates.

7.2 Integrating Results from the Research Chapters

7.2.1 On Kiwi Vocal Behaviour and Activity

The North Island Brown Kiwi (*Apteryx mantelli*, Bartlett 1851) is an iconic and very charismatic species endemic to Aotearoa New Zealand. Nevertheless, this species is considered globally "Vulnerable" (BirdLife International, 2017) and nationally "At–Risk, Declining", being threatened by introduced mammals and predation by dogs (Department of Conservation Te Papa Atawhai, 2021). Moreover, its distribution is fragmented across the North Island of Aotearoa New Zealand, with potentially reduced gene flow, hence heightened risk of inbreeding. Despite decades of genetic research on this species, recent reviews have identified how the produced knowledge would not allow for predictable conservation outcomes (Undin et al., 2021). Additionally, only a few NIBK translocations could thus far be considered successful (Jahn et al., 2022). It is crucial to be able to properly assess animal abundance consistently, especially after translocations. Acoustic monitoring is the second most common method of post–release monitoring (Jahn et al., 2022), and interpreting this data can be quite daunting especially considering the surprising lack of knowledge regarding NIBK vocal behaviour. Formal descriptions of NIBK vocalisations are relatively recent (Corfield et al., 2008), as are reports of differences in vocal activity depending on external factors such as weather variables and lunar phase (Colbourne and Digby, 2016). Additionally, according to the latest Kiwi Recovery Plan, there is a critical need for more objective and effective framework for NIBK acoustic monitoring (see page 7), and a need for more knowledge about the behaviour of all Kiwi species (action 19.4, Germano et al., 2018).

This thesis delivers updated and novel information about NIBK vocal behaviour intertwined within the PAM–oriented experiments (chapters 4, 5, and 6). Over different seasons and time windows, these show how NIBK in the considered populations had a tendency to be more vocal at higher temperatures and relative humidities (chapter 4), and less so with higher lunar illumination (chapters 4, 5) and higher wind speed (chapter 4). Atmospheric pressure however resulted in conflicting results: NIBK were found to be both less vocally active at higher atmospheric pressure levels (Chapter 4), and less vocally active (chapter 5). The likely explanation for this resides in that moonlight and atmospheric pressure effects disappeared when a `day` covariate was introduced (chapter 5), indicating that they probably approximated time (i.e., days), rather than actual weather effects. However, in chapter 4 the posterior distributions of both moonlight and atmospheric pressure covariates were consistently associated with the response variable, notwithstanding the presence of covariates explicitly representing time (`day` and `data collection period` covariates). This would also hold for the covariates from chapter 6, which considered short sampling windows over multiple seasons. However the covariates from the models employed in this chapter relate to virtual/recruited individuals derived from the projection of the vocal behaviour of the

ABAR–equipped individuals; for this reason we are not considering them as informative of actual NIBK behaviour. It is still reassuring to see how they mostly align with the other chapters. Overall, across all these chapters male NIBK are more vocal than females, except that ABAR–equipped individuals from Kauri Bush of both sexes were similarly active. NIBK tended to be less vocal outside the breeding season as seen in chapters 4 (see figure 4.6) and 6 (figure 6.4). Finally, despite similar densities between NIBK communities at The Landing and on Ponui Island, both individual and community vocal activities resulted remarkably different between these two locations. Further studies would be necessary to investigate whether this might be an actual difference in vocal behaviour between island and mainland populations, more related to habitat type, or possibly latitude (although the time window of acoustic sampling were all adjusted for differences in sunset time).

By performing playback tests for assessing whether this technique could be of help in acoustic monitoring, we acquired information on the possibly territorial nature of NIBK (Corfield et al., 2008; Taborsky and Taborsky, 1992, 1995). According to the literature regarding territorial birds, we would expect these to not only respond to playback, but to do so more intensely when presented with strangers' vocalisations (e.g. Moskát et al., 2017). Additionally, in the case of NIBK, as duets have been reported to possibly have territorial defence functions (Corfield et al., 2008), we would expect more aggressive responses when broadcasting duets of strangers. However, animals from the high density community of Kauri Bush, Ponui Island did not vocalise more when presented with NIBK vocalisations than control sounds (either noise or vocalisations of Ruru, *Ninox novaeseelandiae*, Gmelin, 1788). We need to consider that this experimental test of playback (chapter 4) aimed at evaluating the potential of using playback for standardising aural surveys, and as such we did not broadcast from the depths of Kauri Bush, but from the pasture overlooking it (figure 4.3). However, the distance from vocalising birds was fairly low, with some birds vocalising just tens of metres from the broadcasting stations. It could be that what a NIBK considers territory may be a narrow area close to their roosting area, ideally this could be investigated with further playback experiments. However, the idea of a fluid concept of territoriality (Dixon, 2015; Potter, 1989) seems more appropriate for this species.

7.2.2 Relations to population densities

In the pursuit of non–invasive monitoring practices, identifying how numbers of vocalisations relate to numbers of individuals is essential (Pérez-Granados and Traba, 2021). In the absence of individual identification (Takagi, 2018; Terry et al., 2005), this is the only way to effectively estimate abundance from PAM. In chapter 6 we showed how models integrating information on individual vocal behaviour relay more realistic and consistent estimates (figures 6.6 and 6.10). We additionally show how models integrating estimates from populations with individually marked birds to estimate unmarked populations can deliver estimates closer to those obtained if these also had marked individuals.

These estimates are however still not close enough to those informed with ABARs in the target populations and, as illustrated in chapter 6, the use of these devices would need to be repeated over multiple seasons and populations in order to reach reliable estimates based only on PAM. It is however remarkable how our customised Spatial Mark Resight models for Unmarked population (see Royle et al., 2013, Chapter 18, and Appendix VI.i of this thesis) delivered estimates with posterior distributions always greatly overlapping the ones with ABARsinformation. Given the broad width of the distribution of estimated individuals for the population of The Landing, further investigations on mainland managed populations of both high and low presumed density would be necessary to reduce the degree of uncertainty. The fact that the Unmarked population models consistently delivered lower population estimates than the ABAR–informed models is likely because the added information from the ABARs provided knowledge of which proportion of animals did not vocalise during the sampling windows, thus allowing the ABARs–informed models to project this behaviour at the community level. In order to better estimate the Northland population it would be advisable to extend the sampling time window to the whole night, or perform multiple sampling over several widespread stations in the area.

7.2.3 Applications to conservation

This study provides a series of tools that can be immediately applied to NIBK monitoring nationwide. Firstly, we have tested and trialed a PAM system to localise sounds using ARUs produced by DOC with reasonable (approximately 15 m, figure 3.6) uncertainty, together with an assessment of their detection range with NIBK vocalisations (approximately 300 m, chapter 3). This allows for both paired sampling (as in Castro et al., 2019) and direct replacement of human observers with ARUs, granted that the appropriate detection range adjustments are adopted, such as deploying a grid of enough ARUs around listening stations as to cover the same range presumed for human observers. This would greatly reduce costs (Williams et al., 2018) while producing potentially permanent data records.

Secondly, having performed the first formal evaluation of using acoustic playback in NIBK acoustic surveys (chapter 4), this study provides support for the current Kiwi Best Practice Manual in discouraging its use, adding a behavioural component to the current reasoning based on changing technologies. In addition, the reporting of how a NIBK community naïve to investigator's disturbance and a regularly handled one could not be distinguished based on their vocal activity alone encourages the performing of aural surveys concurrently with transmitter change field trips, again reducing costs. Finally, having for the first time tested how animal-borne recorders used in conjunction to ARUs can be used to provide robust estimates of animal populations, this study paves the way to the development of an eventually completely PAM–reliant abundance estimating framework for NIBK.

7.3 Conclusion and future prospects

While this study has established some important building blocks for the development of sound PAM NIBK abundance estimates, it has opened many more doors than it has closed. This is to say that whereas this study was successful in addressing some important points in NIBK behaviour and conservation practices, it also set the stage for several important issues to be investigated. One of the most enticing would definitely be about whether there is such a thing as NIBK vocal individuality, and whether it can be of use in conservation, that is, whether we can recognise how many different individuals vocalise in an area. Such an exciting project would now be possible thanks to the ABARs data produced during this study, and that which can be produced in the future thanks to our trials and tests.
Moreover, a whole set of questions regarding territoriality and vocal behaviour in relationship to population density can now be tackled, protentially even without excessive ethical issues. Currently, the ABARs can only be left on the animals for a few weeks, thus requiring one extra handling event per year per bird, but further collaborations with transmitter–building technicianse integrate ABARs with regular VHF transmitters. If programmed to work for only a small amount of time, say one or two hours, these ABARs could likely be adapted to be left on the birds for longer periods of time, and with further testing possibly the whole year, thus averting the extra handling.

This would not only allow for more informed abundance estimates, but also for the compilation of the NIBK (and potentially other Kiwi species) repertoire with data from managed populations across Aotearoa New Zealand. Additionally, as it is not clear to what extent NIBK recognises other sounds as NIBK, for instance males are known to respond to impressions of their calls made with shepherd whistles (e.g. McLennan et al., 1987), hence experiments with artificially sythesised vocalisations of different levels of resemblance to natural vocalisations would deliver unprecedented insight on this aspect of NIBK/NIBK recognition.

In light of our findings, the next steps for kaitiaki to achieve reliable PAM–based abundance estimates are:

1. Collaborate with transmitter–building technicians to refine the ABARs. Ideally, the resulting devices would work for a whole year as regular VHF transmitters, and wake up to record audio for two hours during the National KCC scheme week. By retaining the original VHF transmitters build, such transmitters would also provide data on animals' activity through the already in–place `chick timers`, which would add a further layer of detail in describing NIBK vocal behaviour.

2. Identify a number of populations to equip with ABARs. These would ideally be already–managed populations, including animals from as many different environments, regiorns, and population densities as possible to encompass as much variability as possible (Pérez-Granados et al., 2019a, suggest between 20 and 30 sites).

3. In these populations, replace the regular VHF transmitters with ABARs on a representative number of individuals and, during National KCC scheme week, deploy ARUs in the interested areas. Ideally deploy VHF proximity data loggers to also objectively record the presence of non–vocal marked animals in the area. Some additional paired sampling with human observers performing KCC would help to standardise estimates in locations where it would be impractical to deploy ARUs. Also, performing some dog sweep surveys (where feasible) would provide further ground truth estimates.

4. Iterate comparing abundance estimates with ABARs–informed models and informing–populations models (see page 129) until the difference between these two estimates either reaches negligible size or stabilises.

Once the difference between estimates obtained by ABARs–informed models and informing–population models is satisfactory, the latter can be used to infer abundances without the need equipping animals with ABARs, effectively reaching reliable non invasive PAM–reliant abundance estimates in NIBK.

Finally, the methods described herein can be quickly adapted to be used with any terrestrial vocal species capable to carry ABARs, which would help standardise monitoring practices, provide more robust estimates, ultimately allowing conservation decision makers to more consciously allocate resources.

Bibliography

Abbey–Lee, R. N., Y. G. Araya–Ajoy, A. Mouchet, M. Moiron, E. F. Stuber, B. Kempenaers, and N. J. Dingemanse, 2018. Does perceived predation risk affect patterns of extra–pair paternity? A field experiment in a passerine bird. Functional Ecology 32:1001–1010.

Adams, D. C., 2000. Corticosterone responses of captive and wild northern brown kiwi (*Apteryx mantelli*): a thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Zoology at Massey University, Palmerston North, New Zealand. Thesis, Massey University.

Akçay, Ç., M. E. Tom, D. Holmes, S. E. Campbell, and M. D. Beecher, 2011. Sing softly and carry a big stick: signals of aggressive intent in the song sparrow. Animal Behaviour 82:377–382.

Alldredge, M. W., K. Pacifici, T. R. Simons, and K. H. Pollock, 2008. A novel field evaluation of the effectiveness of distance and independent observer sampling to estimate aural avian detection probabilities. Journal of Applied Ecology 45:1349–1356.

Alldredge, M. W., T. R. Simons, and K. H. Pollock, 2007. A field evaluation of distance measurement error in auditory avian point count surveys. The Journal of Wildlife Management 71:2759–2766.

Allen, M., L. Girod, R. Newton, S. Madden, D. T. Blumstein, and D. Estrin, 2008. Voxnet: An interactive, rapidly-deployable acoustic monitoring platform. *in* 2008 International Conference on Information Processing in Sensor Networks (ipsn 2008), Pp. 371–382. IEEE.

Alquezar, R. D. and R. B. Machado, 2015. Comparisons between autonomous acoustic recordings and avian point counts in open woodland savanna. The Wilson Journal of Ornithology 127:712–723.

Atkinson, I., 1959. Forest vegetation of the inner islands of the hauraki gulf. *in* Proceedings of the New Zealand Ecological Society, 7, Pp. 29–33.

Audacity Team, 2018. Audacity ®: Free audio editor and recorder [computer application]. version 2.2. 2.

Baker, M., 2016. 1,500 scientists lift the lid on reproducibility. Nature 533:452–455.

Bansal, N., 2020. Exploring the host-parasite relationship between brown kiwi (*Apteryx mantelli*), kiwi ticks (*Ixodes anatis*) and kiwi tick-borne haemoparasites: a thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Conservation Biology at Massey University, Manawatu, New Zealand. Thesis, Massey University.

Barker, N. K., 2008. Bird Song Structure and Transmission in the Neotropics: Trends, Methods and Future Directions. Ornitologia Neotropical 19:175–199.

Basse, B., J. McLennan, and G. Wake, 1999. Analysis of the impact of stoats, *Mustela erminea*, on northern brown kiwi, *Apteryx mantelli*, in New Zealand. Wildlife Research 26:227–237.

Bateman, J. and A. Uzal, 2021. The relationship between the Acoustic Complexity Index and avian species richness and diversity: a review. Bioacoustics 0:1–14.

Beason, R. C., 2004. What Can Birds Hear? Proceedings 21st Vertebrate Pest Conference Pp. 92–96.

Bedoya, C. and L. Molles, 2021. Acoustic censusing and individual identification of birds in the wild. bioRχiv `https://doi.org/10.1101/2021.10.29.466450`.

Beecher, M. D., M. B. Medvin, P. K. Stoddard, and P. Loesche, 1986. Acoustic adaptations for parent–offspring recognition in swallows. Experimental biology 45:179–193.

van Bemmelen, R., 2014. XC198866 – Kea – *Nestor notabilis*. *in* https://www.xeno-canto.org/. The XC team.

BirdLife International, 2017. *Apteryx mantelli*. The IUCN Red List of Threatened Species 2017. `https://dx.doi.org/10.2305/IUCN.UK.2017-3.RLTS.T45353580A119177586.en`. [Downloaded on 05 February 2020].

———, 2018. *Philesturnus rufusater*. The IUCN Red List of Threatened Species 2018. `https://dx.doi.org/10.2305/IUCN.UK.2018-2.RLTS.T103730503A132191453.en`. [Downloaded on 05 February 2020].

Blumstein, D. T., D. J. Mennill, P. Clemins, L. Girod, K. Yao, G. Patricelli, J. L. Deppe, A. H. Krakauer, C. Clark, K. A. Cortopassi, S. F. Hanser, B. McCowan, A. M. Ali, and A. N. G. Kirschel, 2011. Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus. Journal of Applied Ecology 48:758–767.

Bombaci, S. P. and L. Pejchar, 2019. Using paired acoustic sampling to enhance population monitoring of new zealand's forest birds. New Zealand Journal of Ecology 43:1–11.

Boulton, R. L., Y. Richard, and D. P. Armstrong, 2008. Influence of food availability, predator density and forest fragmentation on nest survival of New Zealand robins. Biological Conservation 141:580–589.

Bradbury, J. W. and S. L. Vehrencamp, 2011. Principles of animal communication. 2nd ed. Sinauer. ISBN: 978-0878930456.

Brand, A. R., 1938. Vibration frequencies of passerine bird song. The Auk 55:263–268.

Brandes, T. S., 2008. Automated sound recording and analysis techniques for bird surveys and conservation. Bird Conservation International 18:S163–S173.

Bridge, E. S., K. Thorup, M. S. Bowlin, P. B. Chilson, R. H. Diehl, R. W. Fléron, P. Hartl, R. Kays, J. F. Kelly, W. D. Robinson, et al., 2011. Technology on the move: recent and forthcoming innovations for tracking migratory birds. BioScience 61:689–698.

Brondízio, E. S., J. Settele, S. Díaz, and H. T. Ngo (eds.) 2019. Global assessment report of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. IPBES secretariat, Bonn, Germany. ISBN: 978-3-947851-20-1.

Browning, E., R. Gibb, P. Glover-Kapfer, and K. E. Jones, 2017. Passive acoustic monitoring in ecology and conservation. Tech. Rep. 2, WWF Conservation Technology Series.

Buckland, S. T., D. R. Anderson, K. P. Burnham, and J. L. Laake, 2012. Distance sampling: estimating abundance of biological populations. Springer Science & Business Media. ISBN: 978-0412426605.

Budka, M., P. Matyjasiak, J. Typiak, M. Okołowski, and M. Zagalska-Neubauer, 2019. Experienced males modify their behaviour during playback: the case of the chaffinch. Journal of Ornithology 160:673–684.

Cai, S., T. Collier, L. Girod, J.-Y. Lee, R. E. Hudson, K. Yao, C. E. Taylor, M. Bao, and Z. Wang, 2013. New wireless acoustic array node for localization, beamforming and source separation for bio-complexity bird data collection and study. *in* 2013 IEEE China Summit and International Conference on Signal and Information Processing, Pp. 210–214. IEEE.

Cameron, E. K. and P. J. de Lange, 2006. Vegetation and vascular flora of southern Ponui Island, Hauraki Gulf - a return visit. 16/10/05. Auckland Botanical Society 61:3–14.

Carney, K. M. and W. J. Sydeman, 1999. A review of human disturbance effects on nesting colonial waterbirds. Waterbirds 22:68–79.

Carpenter, B., A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell, 2017. Stan: A probabilistic programming language. Journal of Statistical Software 76:1–32.

Castro, I., A. De Rosa, N. Priyadarshani, L. Bradbury, and S. Marsland, 2019. Experimental test of birdcall detection by autonomous recorder units and by human observers using broadcast. Ecology and Evolution 9:2376–2397.

Catchpole, C. K. and P. J. Slater, 2003. Bird song: biological themes and variations. Cambridge University Press. ISBN: 978-1108469227.

Celis-Murillo, A., J. L. Deppe, and M. P. Ward, 2012. Effectiveness and utility of acoustic recordings for surveying tropical birds. Journal of Field Ornithology 83:166–179.

Challands, K. I., 1976. A preliminary investigation of the Pohangina Valley river terraces: a dissertation... for the degree of BA (Hons.) in Geography at Massey University. Thesis, Massey University.

Cobos, M., F. Antonacci, A. Alexandridis, A. Mouchtaris, and B. Lee, 2017. A survey of sound source localization methods in wireless acoustic sensor networks. Wireless Communications and Mobile Computing 2017.

Colbourne, R., 2005. Kiwi (*Apteryx* spp.) on offshore New Zealand islands. Tech. Rep. 208, Populations, translocations and identification of potential release sites. DOC Research & Development series.

Colbourne, R. and A. P. Digby, 2016. Call rate behabviour of brown kiwi (*Apteryx mantelli*) and great spotted kiwi (*A. haastii*) in relation to temporal and environmental paramenters. Tech. Rep. 348, Department of Conservation. ISBN: 978–0–478–15080–3.

Colbourne, R. and R. Kleinpaste, 1984. North island brown kiwi vocalisations and their use in censusing populations. Tech. Rep. 3, New Zealand Wildlife Service.

Corfield, J., L. Gillman, and S. Parsons, 2008. Vocalizations of the north island brown kiwi (*Apteryx mantelli*). The Auk 125:326–335.

Corfield, J., M. F. Kubke, S. Parsons, J. M. Wild, and C. Köppl, 2011. Evidence for an auditory fovea in the New Zealand kiwi (*Apteryx mantelli*). PLoS One 6:e23771.

Corfield, J. R., M. F. Kubke, S. Parsons, and C. Köppl, 2012. Inner-ear morphology of the New Zealand kiwi (*Apteryx mantelli*) suggests high-frequency specialization. Journal of the Association for Research in Otolaryngology 13:629–639.

Cornec, C., Y. Hingrat, T. Aubin, and F. Rybak, 2017. Booming far: the long-range vocal strategy of a lekking bird. Royal Society Open Science 4:170594.

Cragg, J. L., A. E. Burger, and J. F. Piatt, 2015. Testing the effectiveness of automated acoustic sensors for monitoring vocal activity of marbled murrelets *Brachyramphus marmoratus*. Marine Ornithology 43:151–160.

Craig, E., 2018. Call Count Monitoring of Northland Brown Kiwi 2017. Tech. rep., Department of Conservation, Whangarei.

———, 2021. Call count monitoring of northland brown kiwi 2020. Tech. rep., Department of Conservation, Whangarei.

Crunchant, A.-S., C. D. Nakka, J. T. Isaacs, and A. K. Piel, 2021. Acoustic sensors. P. 53, *in* S. A. Wich and A. K. Piel, eds. Conservation Technology. Oxford University Press. ISBN: 0198850247.

Cunningham, S., I. Castro, and M. Alley, 2007. A new prey-detection mechanism for kiwi (*Apteryx* spp.) suggests convergent evolution between paleognathous and neognathous birds. Journal of Anatomy 211:493–502.

Cunningham, S. J. and I. Castro, 2011. The secret life of wild brown kiwi: studying behaviour of a cryptic species by direct observation. New Zealand Journal of Ecology 35:209–219.

Curzer, H. J., M. C. Wallace, G. Perry, P. J. Muhlberger, and D. Perry, 2013. The ethics of wildlife research: A nine R theory. ILAR Journal 54:52–57.

Cynx, J., H. Williams, and F. Nottebohm, 1990. Timbre discrimination in Zebra Finch (*Taeniopygia guttata*) song syllables. Journal of Comparative Psychology 104:303–308.

Dabelsteen, T., 1981. The Sound Pressure Level in the Dawn Song of the Blackbird *Turdus merula* and a Method for Adjusting the Level in Experimental Song to the Level in Natural Song. Zeitschrift für Tierpsychologie 56:137–149.

Dadwal, N. and D. Bhatt, 2017. Response of male Pied Bushchats *Saxicola caprata* to playback of the songs of neighbours and strangers. Ornithological Science 16:141–146.

Darras, K., P. Batáry, B. Furnas, A. Celis–Murillo, S. L. Van Wilgenburg, Y. A. Mulyani, and T. Tscharntke, 2018a. Comparing the sampling performance of sound recorders versus point counts in bird surveys: A meta-analysis. Journal of Applied Ecology 55:2575–2586.

Darras, K., B. Furnas, I. Fitriawan, Y. Mulyani, and T. Tscharntke, 2018b. Estimating bird detection distances in sound recordings for standardizing detection ranges and distance sampling. Methods in Ecology and Evolution 9:1928–1938.

Deecke, V. B., 2006. Studying marine mammal cognition in the wild: a review of four decades of playback experiments. Aquatic mammals 32:461–482.

Dent, M. L. and R. J. Dooling, 2003. Investigations of the precedence effect in budgerigars: effects of stimulus type, intensity, duration, and location. The Journal of the Acoustical Society of America 113:2146–2158.

Department of Conservation Te Papa Atawhai, 2021. Brown Kiwi. `https://www.doc.govt.nz/nature/native-animals/birds/birds-a-z/kiwi/brown-kiwi/`. [Accessed on 14 November 2021].

Derégnaucourt, S., 2011. Birdsong learning in the laboratory, with especial reference to the song of the zebra finch (*Taeniopygia guttata*). Interaction Studies 12:324–350.

Derégnaucourt, S. and D. Bovet, 2016. The perception of self in birds. Neuroscience & Biobehavioral Reviews 69:1–14.

Digby, A., 2013. Whistling in the dark: an acoustic study of Little Spotted Kiwi. Thesis, Victoria University of Wellington.

Digby, A., B. D. Bell, and P. D. Teal, 2014a. Non-linear phenomena in little spotted kiwi calls. Bioacoustics 23:113–128.

Digby, A., M. Towsey, B. D. Bell, and P. D. Teal, 2013. A practical comparison of manual and autonomous methods for acoustic monitoring. Methods in Ecology and Evolution 4:675–683.

———, 2014b. Temporal and environmental influences on the vocal behaviour of a nocturnal bird. Journal of Avian Biology 45:591–599.

Dixon, T., 2015. What they do in the shadows: habitat utilisation and diet of brown kiwi (*Apteryx mantelli*) adults within a high-density island population: a thesis submitted in partial fulfilment of the requirements for the degree of Master of Science in Ecology, Massey University, Palmerston North, New Zealand. Thesis, Massey University.

Dooling, R. J., 2004. Audition: can birds hear everything they sing? Pp. 206–225, *in* Nature's Music: The Science of Birdsong. Elsevier. ISBN: 978-0124730700.

Dooling, R. J., B. Lohr, and M. L. Dent, 2000. Hearing in Birds and Reptiles. Pp. 308–359, *in* Comparative hearing: Birds and Reptiles. Springer. ISBN: 978-0387946849.

Dooling, R. J. and N. H. Prior, 2017. Do we hear what birds hear in birdsong? Animal Behaviour 124:283–289.

Dooling, R. J. and M. H. Searcy, 1981. Amplitude Modulation Thresholds for the Parakeet (*Melopsittacus undulatus*). Journal of Comparative Physiology. 143:383–388.

Douglas, S. B. and D. J. Mennill, 2010. A review of acoustic playback techniques for studying avian vocal duets. J. Field Orn. 81:115–129.

Dufty, A. M. and M. B. Crandall, 2005. Corticosterone secretion in response to adult alarm calls in american kestrels. Journal of Field Ornithology 76:319–325.

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

Efford, M. G. and J. Boulanger, 2019. Fast evaluation of study designs for spatially explicit capture–recapture. Methods in Ecology and Evolution 10:1529–1535.

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

Emlen, J. T., 1971. Population densities of birds derived from transect counts. The Auk 88:323–342.

Evans, S. A., S. M. Redpath, F. Leckie, and F. Mougeot, 2007. Alternative methods for estimating density in an upland game bird: the red grouse lagopus lagopus scoticus. Wildlife Biology 13:130–139.

Fan, J., Q. Luo, and D. Ma, 2010. Localization estimation of sound source by microphones array. Procedia Engineering 7:312–317.

Farina, A., 2013. Soundscape ecology: principles, patterns, methods and applications. Springer. ISBN: 978-9400773738.

Farmer, R. G., M. L. Leonard, and A. G. Horn, 2012. Observer effects and avian-call-count survey quality: rare-species biases and overconfidence. The Auk 129:76–86.

Ferreira, L. S., J. S. Damo, V. Sábato, J. E. Baumgarten, F. H. G. Rodrigues, and R. Sousa-Lima, 2021. Using playbacks to monitor and investigate the behaviour of wild maned wolves. Bioacoustics 30:74–92.

Fischer, J., R. Noser, and K. Hammerschmidt, 2013. Bioacoustic field research: a primer to acoustic analyses and playback experiments with primates. American Journal of Primatology 75:643–663.

Fotheringham, J. R., P. R. Martin, and L. Ratcliffe, 1997. Song transmission and auditory perception of distance in Wood Warblers (Parulinae). Animal Behaviour 53:1271–1285.

Freeman, B. G., G. A. Montgomery, and D. Schluter, 2017. Evolution and plasticity: Divergence of song discrimination is faster in birds with innate song than in song learners in Neotropical passerine birds. Evolution 71:2230–2242.

Fulford, T., 2016a. XC348754 – North Island Robin – *Petroica longipes*. *in* https://www.xeno-canto.org/. The XC team.

———, 2016b. XC348779 – North Island Saddleback – *Philesturnus rufusater*. *in* https://www.xeno-canto.org/. The XC team.

Gager, Y., 2019. Information transfer about food as a reason for sociality in bats. Mammal Review 49:113–120.

Gates, C. E. and W. B. Smith, 1972. Estimation of density of mourning doves from aural information. Biometrics Pp. 345–359.

Gentry, K. E., R. N. Lewis, H. Glanz, P. I. Simões, Á. S. Nyári, and M. S. Reichert, 2020. Bioacoustics in cognitive research: Applications, considerations, and recommendations. Wiley Interdisciplinary Reviews: Cognitive Science 11:e1538.

Germano, J., S. Barlow, I. Castro, R. Colbourne, M. Cox, C. Gillies, K. Hackwell, J. Harawira, M. Impey, A. Reuben, et al., 2018. Kiwi Recovery Plan 2018–2028/Mahere Whakaora Kiwi 2018–2028. Tech. Rep. 64, Department of Conservation Te Papa Atawhai — Threatened Species Recovery Plan. ISBN: 978–1–98–851473–4.

Gibbs, J. P. and S. M. Melvin, 1997. Power to detect trends in waterbird abundance with call-response surveys. The Journal of wildlife management 61:1262–1267.

Gibbs, J. P. and D. G. Wenny, 1993. Song Output as a Population Estimator: Effect of Male Pairing Status (El Canto Utilizado para Estimar el Tamaño de Poblaciones: El Efecto de Machos Apareados y No-apareados). Journal of Field Ornithology 64:316–322.

Gill, L. F., P. B. D'Amelio, N. M. Adreani, H. Sagunsky, M. C. Gahr, and A. Ter Maat, 2016. A minimum-impact, flexible tool to study vocal communication of small animals with precise individual-level resolution. Methods in Ecology and Evolution 7:1349–1358.

Graff, J., 2013. XC165265 – Australasian Bittern – *Botaurus poiciloptilus*. *in* https://www.xeno-canto.org/. The XC team.

Greif, S. and Y. Yovel, 2019. Using on-board sound recordings to infer behaviour of free-moving wild animals. Journal of Experimental Biology 222:jeb184689.

Haddaway, N. R., B. Macura, P. Whaley, and A. S. Pullin, 2018. ROSES RepOrting standards for Systematic Evidence Syntheses: pro forma, flow–diagram and descriptive summary of the plan and conduct of environmental systematic reviews and systematic maps. Environmental Evidence 7:7.

Hagenow, C., 1934. The equal tempered musical scale. American Journal of Physics 2:81–84.

Hart, J. D., 2020. Acoustic surveys of Water Rail *Rallus aquaticus* over-wintering on Alderney. Bird Study 67:245–250.

Hart, J. K. and K. Martinez, 2006. Environmental sensor networks: A revolution in the earth system science? Earth-Science Reviews 78:177–191.

Haselmayer, J. and J. S. Quinn, 2000. A comparison of point counts and sound recording as bird survey methods in amazonian southeast peru. The Condor 102:887–893.

Hassler, M., 2020a. World Ferns: Checklist of Ferns and Lycophytes of the World (version Nov 2018). *in* Species 2000 & ITIS Catalogue of Life, [2020-01-10] Beta. Species 2000: Naturalis, Leiden, the Netherlands. ISSN 2405-8858. [Digital resource at `www.catalogueoflife.org/col` accesed on 07 February 2020].

———, 2020b. World Plants: World Plants: Synonymic Checklists of the Vascular Plants of the World (version Nov 2018). *in* Species 2000 & ITIS Catalogue of Life, [2020-01-10] Beta. Species 2000: Naturalis, Leiden, the Netherlands. ISSN 2405-8858. [Digital resource at `www.catalogueoflife.org/col` accesed on 07 February 2020].

Hedley, R. W., K. K. Denton, and R. E. Weiss, 2017. Accounting for syntax in analyses of countersinging reveals hidden vocal dynamics in a songbird with a large repertoire. Animal Behaviour 131:23–32.

van Heezik, Y. and P. J. Seddon, 2017. Counting birds in urban areas: a review of methods for the estimation of abundance. Pp. 185–207, *in* Ecology and Conservation of Birds in Urban Environments. Springer.

Holland, J., T. Dabelsteen, S. B. Pedersen, and A. L. Paris, 2001. Potential ranging cues contained within the energetic pauses of transmitted Wren song. Bioacoustics 12:3–20.

Hurlbert, S. H., 1984. Pseudoreplication and the design of ecological field experiments. Ecological monographs 54:187–211.

Hynes, P., 1955. Ponui Island. Auckland Botanical Society 12:2.

Ichikawa, K., T. Akamatsu, T. Shinke, K. Adulyanukosol, and N. Arai, 2011. Callback response of dugongs to conspecific chirp playbacks. The Journal of the Acoustical Society of America 129:3623–3629.

Irwin, G., 2020. The archaeology of maori settlement and pa on ponui island, inner hauraki gulf, ad 1400-1800. Journal of the Polynesian Society, The 129:29–58.

IUCN, 2020. IUCN 2019 ANNUAL REPORT. `https://portals.iucn.org/library/sites/library/files/documents/2020-009-En.pdf`.

———, 2021. Addendum to the IUCN Programme 2021–2024. `https://portals.iucn.org/library/node/49760`.

Jahn, P., E. F. Cagua, L. E. Molles, J. G. Ross, and J. M. Germano, 2022. Kiwi translocation review: are we releasing enough birds and to the right places? New Zealand Journal of Ecology 46:3454.

Jamieson, S. E., I. Castro, T. Jensen, K. W. Morrison, and B. Durrant, 2016. Roosting preferences of north island brown kiwis (apteryx mantelli). The Wilson Journal of Ornithology 128:857–866.

Janik, V. M. and L. S. Sayigh, 2013. Communication in bottlenose dolphins: 50 years of signature whistle research. Journal of Comparative Physiology A 199:479–489.

Johnson, R. R. and J. J. Dinsmore, 1986. The use of tape-recorded calls to count virginia rails and soras. The Wilson Bulletin 98:303–306.

Jones, G. and B. M. Siemers, 2011. The communicative potential of bat echolocation pulses. Journal of Comparative Physiology A 197:447–457.

Juodakis, J., I. Castro, and S. Marsland, 2021a. Precision as a metric for acoustic survey design using occupancy or spatial capture-recapture. Environmental and Ecological Statistics 28:587–608.

Juodakis, J., S. Marsland, and N. Priyadarshani, 2021b. A changepoint prefilter for sound event detection in long-term bioacoustic recordings. The Journal of the Acoustical Society of America 150:2469–2478.

Keast, J., T. Kelly, H. Moorhouse, J. Tan, and S.-Y. Wei, 2010. Conservation of the North Island Brown Kiwi (*Apteryx mantelli*): current approaches, the successes and limitations, and proposals to ensure long term continuity.

Kendeigh, S. C., 1944. Measurement of bird populations. Ecological Monographs 14:67–106.

King, S. L., 2015. You talkin' to me? Interactive playback is a powerful yet underused tool in animal communication research. Biology Letters 11.

Knapp, C. and G. Carter, 1976. The generalized correlation method for estimation of time delay. IEEE Trans. Acoustics, Speech, and Signal Processing 24:320–327.

Kooyman, G., 2007. Animal-borne instrumentation systems and the animals that bear them: then (1939) and now (2007). Marine Technology Society Journal 41:6–8.

Krebs, C., 1998. Estimating abundance in animal and plant populations. Pp. 24–77, *in* Ecological methodology, 2nd ed. Benjamin Cummings California.

Krieg, C. A. and T. Getty, 2016. Not just for males: females use song against male and female rivals in a temperate zone songbird. Animal Behaviour 113:39–47.

Kroodsma, D. E., 1989a. Inappropriate experimental designs impede progress in bioacoustic research: a reply. Animal Behaviour 38:717–719.

———, 1989b. Suggested experimental designs for song playbacks. Animal Behaviour 37:600–609.

———, 1990. Using appropriate experimental designs for intended hypotheses in 'song' playbacks, with examples for testing effects of song repertoire sizes. Animal Behaviour 40:1138–1150.

———, 1992. Much ado creates flaws. Animal Behaviour 44:580–582.

———, 2017. Birdsong 'performance' studies: A sad commentary. Animal Behaviour 133:209–210.

Lambrechts, M. M. and A. A. Dhondt, 1995. Individual voice discrimination in birds. Pp. 115–139, *in* Current Ornithology, vol. 12. Springer.

Larsen, O. N., R. J. Dooling, and A. Michelsen, 2006. The role of pressure difference reception in the directional hearing of Budgerigars (*Melopsittacus undulatus*). Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology 192:1063–1072.

Leedale, A. E., S. A. Collins, and S. R. de Kort, 2015. Blackcaps (*Sylvia atricapilla*) increase the whistle part of their song in response to simulated territorial intrusion. Ethology 121:403–409.

Leite, G., 2018. XC348754 – North Island Robin – *Petroica longipes*. *in* https://www.xeno-canto.org/. The XC team.

Li, D. and Y. H. Hu, 2003. Energy-based collaborative source localization using acoustic microsensor array. EURASIP Journal on Advances in Signal Processing 2003:1–17.

Li, J., Y. Zhang, and Z. Zhang, 2011. High frequency components in avian vocalizations. Chinese Birds 2:125–131.

Linhart, P., R. Fuchs, S. Poláková, and H. Slabbekoorn, 2012. Once bitten twice shy: long-term behavioural changes caused by trapping experience in willow warblers *Phylloscopus trochilus*. Journal of Avian Biology 43:186–192.

Lipshutz, S. E., I. A. Overcast, M. J. Hickerson, R. T. Brumfield, and E. P. Derryberry, 2017. Behavioural response to song and genetic divergence in two subspecies of White–crowned Sparrows (*Zonotrichia leucophrys*). Molecular ecology 26:3011–3027.

Lohr, B., T. F. Wright, and R. J. Dooling, 1998. Detection of changes in timbre and harmonicity in complex sounds by Zebra Finches (*Taeniopygia guttata*) and budgerigars (*Melopsittacus undulatus*). Journal of Comparative Psychology (Washington, D.C. : 1983) 112:36–47.

———, 2003. Detection and discrimination of natural calls in masking noise by birds: estimating the active space of a signal. Animal Behaviour 65:763–777.

Lusk, C. H., 1984. Structure and dynamics of alluvial forest in the Pohangina Valley: a thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Botany. Thesis, Massey University.

Luther, D. A., R. Danner, J. Danner, K. Gentry, and E. P. Derryberry, 2017. The relative response of songbirds to shifts in song amplitude and song minimum frequency. Behavioral Ecology 28:391–397.

Luther, D. A. and J. Magnotti, 2014. Can animals detect differences in vocalizations adjusted for anthropogenic noise? Animal Behaviour 92:111–116.

Magrath, R. D., T. M. Haff, P. M. Fallow, and A. N. Radford, 2015. Eavesdropping on heterospecific alarm calls: from mechanisms to consequences. Biological Reviews 90:560–586.

Mankin, R. W. et al., 2012. Applications of acoustics in insect pest management. CAB Rev 7:1–7.

Marion, W. R., T. E. O'Meara, and D. S. Maehr, 1981. Use of playback recording in sampling elusive or secretive birds. Studies in Avian Biology 6:81–85.

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

Marsland, S., N. Priyadarshani, J. Juodakis, and I. Castro, 2019. AviaNZ: A future-proofed program for annotation and recognition of animal sounds in long-time field recordings. Methods in Ecology and Evolution 10:1189–1195.

Martin, G. R., 2017. The sensory ecology of birds. Oxford University Press. ISBN: 978-0199694532.

Martin, P. R., J. R. Fotheringham, and R. J. Robertson, 1995. A Prairie Warbler with a conspecific and heterospecific song repertoire. The Auk 112:770–774.

McGregor, P. K., 1994. Sound cues to distance: the perception of range. Pp. 74–94, *in* M. N. O. Davies and P. Green, eds. Perception and motor control in birds. Springer. ISBN: 978-0387528557.

———, 2000. Playback experiments: design and analysis. Acta ethologica 3:3–8.

McGregor, P. K., C. K. Catchpole, T. Dabelsteen, J. B. Falls, L. Fusani, H. C. Gerhardt, F. Gilbert, A. G. Horn, G. M. Klump, D. E. Kroodsma, et al., 1992. Design of playback experiments: the Thornbridge Hall NATO ARW consensus. Pp. 1–9, *in* Playback and studies of animal communication. Springer. ISBN: 978-0306442056.

McGregor, P. K. and J. R. Krebs, 1984. Sound degradation as a distance cue in great tit (*Parus major*) song. Behavioral Ecology and Sociobiology 16:49–56.

McGregor, P. K., J. R. Krebs, and L. M. Ratcliffe, 1983. The reaction of great tits (*Parus major*) to playback of degraded and undegraded songs: the effect of familiarity with the stimulus song type. The Auk Pp. 898–906.

McLennan, J., 1992. Nationwide monitoring of kiwi populations. DSIR Land Resources. Contract Report No. 92/21.

McLennan, J. and M. Potter, 1992. Distribution, population changes and management of brown kiwi in hawke's bay. New Zealand Journal of Ecology 16:91–102.

McLennan, J., M. Potter, H. Robertson, G. Wake, R. Colbourne, L. Dew, L. Joyce, A. McCann, J. Miles, P. Miller, et al., 1996. Role of predation in the decline of kiwi, *Apteryx* spp., in New Zealand. New Zealand Journal of Ecology 20:27–35.

McLennan, J., M. Rudge, and M. Potter, 1987. Range size and denning behaviour of Brown Kiwi, *Apteryx australis mantelli*, in Hawke's Bay, New Zealand. New Zealand Journal of Ecology 10:97–107.

Mellinger, D. K., K. M. Stafford, S. E. Moore, R. P. Dziak, and H. Matsumoto, 2007. An overview of fixed passive acoustic observation methods for cetaceans. Oceanography 20:36–45.

Miles, J. and I. Castro, 2000. Survey of Northern brown kiwi (*Apteryx mantelli*) on Ponui Island, Hauraki Gulf–1999. Tech. Rep. 21, Auckland Conservancy Technical Report Series. ISSN: 0113-3748 — ISBN: 0-478-21972-5.

Miles, J. R. G., 1995. Comparative ecology of northern brown kiwi (*Apteryx australis mantelli*) in Tongariro National Park and Tongariro Forest Park, central North Island: a thesis presented in partial fulfilment of the requirements for the degree of Master of Science at Massey University. Thesis, Massey University.

Miles, J. R. G., M. A. Potter, and R. A. Fordham, 1997. Northern brown Kiwi (Apteryx australis mantelli) in Tongariro National Park and Tongariro Forest: ecology and threats. Department of Conservation Wellington, New Zealand. ISBN: 0-478-01899-1.

Moon, L., G. Moon, J. Kendrick, and K. Baird, 2012. CD Track #1 of 60. *in* New Zealand Bird Calls. New Holland Publishers New Zealand. ISBN: 9781869663100.

Mortensen, R. M. and F. Rosell, 2020. Long-term capture and handling effects on body condition, reproduction and survival in a semi-aquatic mammal. Scientific Reports 10:1–16.

Morton, E. S., 1983. Grading, discreteness, redundancy, and motivation–structural rules. Pp. 183–212, *in* Acoustic communication in birds, 1st ed. Academic Press. ISBN: 978-0124268012.

Moskát, C., Z. Elek, M. Bán, N. Geltsch, and M. E. Hauber, 2017. Can Common Cuckoos discriminate between neighbours and strangers by their calls? Animal Behaviour 126:253–260.

Moskát, C., M. E. Hauber, M. Bán, A. Fülöp, N. Geltsch, A. Marton, and Z. Elek, 2018. Are both notes of the Common Cuckoo's call necessary for familiarity recognition? Behavioural processes .

Mossman, M. J. and L. A. Weir, 2005. North American Amphibian Monitoring Program (NAAMP). Pp. 307–313, *in* M. Lannoo, ed. Amphibian declines. University of California Press.

Mougeot, F. and V. Bretagnolle, 2000. Predation as a cost of sexual communication in nocturnal seabirds: an experimental approach using acoustic signals. Animal Behaviour 60:647–656.

Mulcahy, J., 2004. REW: Room Eq Wizard. Version 5.19.

Naguib, M., 1995. Auditory distance assessment of singing conspecifies in Carolina Wrens: The role of reverberation and frequency–dependent attenuation. Animal Behaviour 50:1297–1307.

———, 1996. Auditory distance estimation in song birds: Implications, methodologies and perspectives. Behavioural Processes 38:163–168.

———, 1998. Perception of degradation in acoustic signals and its implications for ranging. Behavioral Ecology and Sociobiology 42:139–142.

Naguib, M., G. M. Klump, E. Hillmann, B. Grießmann, and T. Teige, 2000. Assessment of auditory distance in a territorial songbird: Accurate feat or rule of thumb? Animal Behaviour 59:715–721.

Naguib, M. and R. H. Wiley, 2001. Review: Estimating the distance to a source of sound: Mechanisms and adaptations for long–range communication. Animal Behaviour 62:825–837.

Narins, P. M., 2018. Amphibian bioacoustics: From arch to zelick. The Journal of the Acoustical Society of America 143:1767–1767.

Narins, P. M., A. S. Feng, W. Lin, H. Schnitzler, A. Denzinger, R. A. Suthers, and C. Xu, 2004. Old world frog and bird vocalizations contain prominent ultrasonic harmonics. The Journal of the Acoustical Society of America 115:910–913.

Nelson, B. S., 2000. Avian dependence on sound pressure level as an auditory distance cue. Animal Behaviour 59:57–67.

Nelson, D. A., 1989. Song frequency as a cue for recognition of species and individuals in the field sparrow (*Spizella pusilla*). Journal of Comparative Psychology 103:171.

Niederhauser, J. M., A. L. DuBois, W. A. Searcy, S. Nowicki, and R. C. Anderson, 2018. A test of the eavesdropping avoidance hypothesis as an explanation for the structure of low–amplitude aggressive signals in the Song Sparrow. Behavioral Ecology and Sociobiology 72:47.

Nijman, V., 2007. Effects of vocal behaviour on abundance estimates of rainforest galliforms. Acta Ornithologica 42:186–190.

Okahisa, Y., A. Legault, B. Angin, V. Chartendrault, N. Barré, C. de Franceschi, T. Duval, F. Tron, R. Gula, and J. Theuerkauf, 2016. Using playback and habitat models to estimate the distribution and population size of the critically endangered Crow Honeyeater, *Gymnomyza aubryana*, in New Caledonia. Emu-Austral Ornithology 116:41–47.

Parlato, E. H., D. P. Armstrong, and J. G. Innes, 2015. Traits influencing range contraction in new zealand's endemic forest birds. Oecologia 179:319–328.

Patricelli, G. L., M. S. Dantzker, and J. W. Bradbury, 2008. Acoustic directionality of Red–winged Blackbird (*Agelaius phoeniceus*) song relates to amplitude and singing behaviours. Animal Behaviour 76:1389–1401.

Pérez-Granados, C., G. Bota, D. Giralt, J. Albarracín, and J. Traba, 2019a. Cost-effectiveness assessment of five audio recording systems for wildlife monitoring: differences between recording distances and singing direction. Ardeola 66:311–325.

Pérez-Granados, C., G. Bota, D. Giralt, A. Barrero, J. Gómez-Catasús, D. Bustillo-De La Rosa, and J. Traba, 2019b. Vocal activity rate index: a useful method to infer terrestrial bird abundance with acoustic monitoring. Ibis 161:901–907.

Pérez-Granados, C., J. Gómez-Catasús, D. Bustillo-de la Rosa, A. Barrero, M. Reverter, and J. Traba, 2019c. Effort needed to accurately estimate vocal activity rate index using acoustic monitoring: a case study with a dawn-time singing passerine. Ecological Indicators 107:105608.

Pérez-Granados, C. and J. Traba, 2021. Estimating bird density using passive acoustic monitoring: a review of methods and suggestions for further research. Ibis 163:765–783.

Perkins, A. J., A. Douse, G. Morgan, A. Cooper, and M. Bolton, 2017. Using dual–sex calls improves the playback census method for a nocturnal burrow–nesting seabird, the Manx Shearwater *Puffinus puffinus*. Bird Study 64:146–158.

Petersen, C. G. J., 1896. The yearly immigration of young plaice into the limfjord from the german sea, etc. Report of the Danish Biological Station (1895) 6:5–84.

Petrusková, T., H. Kahounová, I. Pišvejcová, A. Čermáková, T. Brinke, N. H. Burton, and A. Petrusek, 2021. Evaluating the potential effects of capturing and handling on subsequent observations of a migratory passerine through individual acoustic monitoring. Journal of Avian Biology 52.

Pierce, R. and I. Westbrooke, 2003. Call count responses of north island brown kiwi to different levels of predator control in northland, new zealand. Biological Conservation 109:175–180.

Plummer, M. et al., 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *in* Proceedings of the 3rd international workshop on distributed statistical computing, vol. 124, Pp. 1–10. Vienna, Austria.

Podolsky, R. H. and S. W. Kress, 1989. Factors affecting colony formation in Leach's Storm–petrel. The Auk 106:332–336.

Pope, A., 2008. Sine Wave Frequency Sweep with Audacity. `https://stormconsultancy.co.uk/blog/storm-news/sine-wave-frequency-sweep-with-audacity/`. [Downloaded on 05 February 2020].

Potamitis, I., S. Ntalampiras, O. Jahn, and K. Riede, 2014. Automatic bird sound detection in long real-field recordings: Applications and tools. Applied Acoustics 80:1–9.

Potter, M., 1990. Movement of North Island brown kiwi (*Apteryx australis mantelli*) between forest remnants. New Zealand Journal of Ecology 14:17–24.

Potter, M. A., 1989. Ecology and reproductive biology of the North Island brown kiwi (*Apteryx australis mantelli*): a thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Zoology at Massey University, Palmerston North. Thesis, Massey University.

Powell, R. A. and G. Proulx, 2003. Trapping and marking terrestrial mammals for research: integrating ethics, performance criteria, techniques, and common sense. ILAR journal 44:259–276.

Priyadarshani, N., I. Castro, and S. Marsland, 2018a. The impact of environmental factors in birdsong acquisition using automated recorders. Ecology and Evolution 8:5016–5033.

Priyadarshani, N., S. Marsland, and I. Castro, 2018b. Automated birdsong recognition in complex acoustic environments: a review. Journal of Avian Biology 49:jav–01447.

Priyadarshani, N., S. Marsland, J. Juodakis, I. Castro, and V. Listanti, 2020. Wavelet filters for automated recognition of birdsong in long-time field recordings. Methods in Ecology and Evolution 11:403–417.

Provost, K. L., W. M. Mauck III, and B. T. Smith, 2018. Genomic divergence in allopatric Northern Cardinals of the North American warm deserts is associated with behavioral differentiation. Ecology and Evolution 8:12456–12478.

Putman, B. J. and D. T. Blumstein, 2019. What is the effectiveness of using conspecific or heterospecific acoustic playbacks for the attraction of animals for wildlife management? a systematic review protocol. Environmental evidence 8:1–9.

Ralph, C. J., J. R. Sauer, and S. Droege, 1995. Monitoring bird populations by point counts. Gen. Tech. Rep. PSW-GTR-149. Albany, CA: US Department of Agriculture, Forest Service, Pacific Southwest Research Station. 187 p 149.

Ratcliffe, L. and R. G. Weisman, 1985. Frequency shift in the fee bee song of the black–capped chickadee. Condor 87:555–556.

———, 1986. Song sequence discrimination in the Black–capped Chickadee (*Parus atricapillus*). Journal of comparative psychology 100:361–367.

`R` Core Team, 2013. `R`: A language and environment for statistical computing.

Reid, D. S., C. M. Wood, S. A. Whitmore, W. J. Berigan, J. J. Keane, S. C. Sawyer, P. A. Shaklee, H. A. Kramer, K. G. Kelly, A. Reiss, et al., 2021. Noisy neighbors and reticent residents: Distinguishing resident from non-resident individuals to improve passive acoustic monitoring. Global Ecology and Conservation P. e01710.

Rhinehart, T. A., L. M. Chronister, T. Devlin, and J. Kitzes, 2020. Acoustic localization of terrestrial wildlife: Current practices and future opportunities. Ecology and Evolution 10:6794–6818.

Richards, D. G., 1981. Estimation of distance of singing conspecifics by the Carolina Wren. The Auk 98:127–133.

Ríos–Chelén, A. A., A. N. McDonald, A. Berger, A. C. Perry, A. H. Krakauer, and G. L. Patricelli, 2017. Do birds vocalize at higher pitch in noise, or is it a matter of measurement? Behavioral Ecology and Sociobiology 71:1–12.

Ritschard, M., K. Riebel, and H. Brumm, 2010. Female Zebra Finches prefer high–amplitude song. Animal Behaviour 79:877–883.

Robertson, H. A., R. Colbourne, and J. McLennan, 2017. Kiwi Best Practice Manual. Department of Conservation. ISBN: 978–1–98–851428–4.

Rosenstock, S. S., D. R. Anderson, K. M. Giesen, T. Leukering, and M. F. Carter, 2002. Landbird counting techniques: current practices and an alternative. The Auk 119:46–53.

Royle, J. A., R. B. Chandler, R. Sollmann, and B. Gardner, 2013. Spatial capture-recapture. Academic Press. ISBN: 978-0124059399.

Saito, K., K. Nakamura, M. Ueta, R. Kurosawa, A. Fujiwara, H. H. Kobayashi, M. Nakayama, A. Toko, and K. Nagahama, 2015. Utilizing the cyberforest live sound system with social media to remotely conduct woodland bird censuses in central japan. Ambio 44:572–583.

Saunders, J. C., P. V. Else, and G. R. Bock, 1978. Frequency selectivity in the Parakeet (*Melopsittacus undulatus*) studied with psychophysical tuning curves. Journal of Comparative and Physiological Psychology 92:406–415.

Scharf, B., 1970. Critical bands. chap. 5, *in* Foundations of Modern Auditory Theory, vol. 1. Academic Press, New York. ISBN: 978-0126919011.

Schlaepfer, M. A., D. Caldwell, J. C. Rorabaugh, M. J. Ryan, S. Stump, and M. J. Sredl, 2021. The use of evoked vocal responses to detect cryptic, low-density frogs in the field. Journal of Herpetology 55:174–180.

Schlicht, E. and B. Kempenaers, 2015. Immediate effects of capture on nest visits of breeding blue tits, cyanistes caeruleus, are substantial. Animal Behaviour 105:63–78.

Schnyder, H. A., D. Verelst, S. Bartenstein, U. Firzlaff, and H. Luksch, 2014. The avian head induces cues for sound localization in elevation. PLoS ONE 9:1–8.

Schwartzkopff, J., 1955. On the hearing of birds. The Auk 72:340–347.

Searcy, W. A., 1989. Pseudoreplication, external validity and the design of playback experiments. Animal Behaviour 38:715–717.

Sebastián-González, E., R. J. Camp, A. M. Tanimoto, P. M. de Oliveira, B. B. Lima, T. A. Marques, and P. J. Hart, 2018. Density estimation of sound-producing terrestrial animals using single automatic acoustic recorders and distance sampling. Avian Conservation and Ecology .

Sedláček, O., J. Vokurková, M. Ferenc, E. N. Djomo, T. Albrecht, and D. Hořák, 2015. A comparison of point counts with a new acoustic sampling method: a case study of a bird community from the montane forests of mount cameroon. Ostrich 86:213–220.

Seress, G., E. Vincze, I. Pipoly, T. Hammer, S. Papp, B. Preiszner, V. Bókony, and A. Liker, 2017. Effects of capture and video-recording on the behavior and breeding success of great tits in urban and forest habitats. Journal of Field Ornithology 88:299–312.

Simons, T. R., M. W. Alldredge, K. H. Pollock, and J. M. Wettroth, 2007. Experimental analysis of the auditory detection process on avian point counts. The Auk 124:986–999.

Slabbekoorn, H., 2004. Singing in the wild: the ecology of birdsong. Pp. 178–205, *in* Nature's music: The Science of Birdsong. Elsevier. ISBN: 978-0124730700.

Slaymaker, M., 2018. XC436719 – Stitchbird – *Notiomystis cincta*. *in* https://www.xeno-canto.org/. The XC team.

Soulsbury, C. D., H. E. Gray, L. M. Smith, V. Braithwaite, S. C. Cotter, R. W. Elwood, A. Wilkinson, and L. M. Collins, 2020. The welfare and ethics of research involving wild animals: A primer. Methods in Ecology and Evolution 11:1164–1181.

Sousa-Lima, R. S., T. F. Norris, J. N. Oswald, and D. P. Fernandes, 2013. A review and inventory of fixed autonomous recorders for passive acoustic monitoring of marine mammals. Aquatic Mammals 39.

Spiesberger, J. L. and K. M. Fristrup, 1990. Passive localization of calling animals and sensing of their acoustic environment using acoustic tomography. The american naturalist 135:107–153.

Stan Development Team et al., 2016. `RStan`: the `R` interface to `Stan`. R package version 2.

Stebbins, W. C., 1983. The Acoustic Sense of Animals. Harvard University Press. ISBN: 978-0674003262.

Stepanian, P. M., K. G. Horton, D. C. Hille, C. E. Wainwright, P. B. Chilson, and J. F. Kelly, 2016. Extending bioacoustic monitoring of birds aloft through flight call localization with a three-dimensional microphone array. Ecology and Evolution 6:7039–7046.

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

Stowell, D., E. Benetos, and L. F. Gill, 2017. On-bird sound recordings: automatic acoustic recognition of activities and contexts. IEEE/ACM Transactions on Audio, Speech, and Language Processing 25:1193–1206.

Strang, K. E., 2018. The ecology of feral cats (*Felis catus*) on a New Zealand offshore island: considerations for management: a thesis presented in partial fulfilment of the requirements for the degree of Doctor of Pholosophy in Zoology at Massey University, Manawatu, New Zealand. Thesis, Massey University.

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

Suits, B., 1998. Frequencies for equal-tempered scale. Physics of Music—Notes `https://pages.mtu.edu/~suits/notefreqs.html`.

Suzuki, R., S. Matsubayashi, R. W. Hedley, K. Nakadai, and H. G. Okuno, 2017. Harkbird: Exploring acoustic interactions in bird communities using a microphone array. Journal of Robotics and Mechatronics 29:213–223.

Taborsky, B. and M. Taborsky, 1991. Social organization of north island brown kiwi: Long-term pairs and three types of male spacing behaviour. Ethology 89:47–62.

———, 1992. Spatial organization of the North Island Brown Kiwi *Apteryx australis mantelli*: sex, pairing status and territoriality. Ibis 134:1–10.

———, 1995. Habitat Use and Selectivity by the Brown Kiwi (*Apteryx australis mantelli*) in a Patchy Environment. The Auk 112:680–689.

———, 1999. The mating system and stability of pairs in kiwi *Apteryx* spp. Journal of Avian Biology 30:143–151.

Takagi, M., 2018. Vocalizations of the ryukyu scops owl *Otus elegans*: individually recognizable and stable. Bioacoustics 29:1–17.

Tanimoto, A. M., P. J. Hart, A. A. Pack, and R. Switzer, 2017. Vocal repertoire and signal characteristics of 'Alalā, the Hawaiian Crow (*Corvus hawaiiensis*). The Wilson J. Ornithol. 129:25–35.

Taylor, C. E., J. T. Brumley, R. W. Hedley, and M. L. Cody, 2017. Sensitivity of California Thrashers (*Toxostoma redivivum*) to song syntax. Bioacoustics 26:259–270.

Taylor, C. E., Y. Huang, and K. Yao, 2016. Distributed sensor swarms for monitoring bird behavior: an integrated system using wildlife acoustics recorders. Artificial Life and Robotics 21:268–273.

Temple, S. A. and J. A. Wiens, 1989. Bird populations and environmental changes: can birds be bio-indicators. American Birds 43:260–270.

Tennyson, A. and P. Martinson, 2006. The extinct birds of new zealand. te papa press: Wellington. 180 p Isbn: 978-0-909010-21-8.

Terry, A. M. R., T. M. Peake, and P. K. McGregor, 2005. The role of vocal individuality in conservation. Frontiers in Zoology 2.

Thomas, L. and T. A. Marques, 2012. Passive acoustic monitoring for estimating animal density. Acoustics Today 8:35–44.

Thorpe, W. and D. Griffin, 1962. Ultrasonic frequencies in bird song. Ibis 104:220–227.

Towsey, M., L. Zhang, M. Cottman-Fields, J. Wimmer, J. Zhang, and P. Roe, 2014. Visualization of long–duration acoustic recordings of the environment. Procedia Computer Science 29:703–712.

Toy, R. and S. Toy, 2021. Changes in behaviour of great spotted kiwi (*Apteryx haastii*) following handling. Notornis 68:173–176.

Tyack, P., 2009. Acoustic playback experiments to study behavioral responses of free-ranging marine animals to anthropogenic sound. Marine Ecology Progress Series 395:187–200.

Undin, M., 2021. Studies of how to improve translocation outcomes of *Apteryx mantelli* focusing on breeding, hybrids, diversity, and telomeres: a thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Ecology at Massey University-Te Kunenga Ki Purehuroa, Palmerston North, New Zealand. Thesis, Massey University.

Undin, M., P. J. Lockhart, S. F. K. Hills, and I. Castro, 2021. Genetic rescue and the plight of ponui hybrids. Frontiers in Conservation Science 1:8. ISSN: 2673-611X.

Van Wilgenburg, S., P. Sólymos, K. Kardynal, and M. Frey, 2017. Paired sampling standardizes point count data from humans and acoustic recorders. Avian Conservation and Ecology 12:13.

Vargas–Castro, L. E., L. Sandoval, and W. A. Searcy, 2017. Eavesdropping avoidance and sound propagation: the acoustic structure of soft song. Animal Behaviour 134:113–121.

Vattiato, G., 2021. Measurement of wild north island brown kiwi (*Apteryx mantelli*) personalities. *in* Modelling Individual Heterogeneity In Behaviour For Wildlife Management And Conservation. University of Canterbury. PhD Thesis.

Venier, L. A., S. B. Holmes, G. W. Holborn, K. A. Mcilwrick, and G. Brown, 2012. Evaluation of an automated recording device for monitoring forest birds. Wildlife Society Bulletin 36:30–39.

Vieco Gálvez, D., 2019. The egg and the nest: obtaining information about the reproductive biology of Apteryx spp.(family: Apterygidae), a cryptic avian taxon, through eggshells: a thesis in partial fulfilment of the requirement for the degree of Doctor of Philosophy in Ecology at Massey University, Palmerston North, New Zealand. Thesis, Massey University.

Volodin, I. A., E. V. Volodina, A. V. Klenova, and V. A. Matrosova, 2015. Gender identification using acoustic analysis in birds without external sexual dimorphism. Avian Research 6:20.

Wagner, G., 2012. XC114338 – Morepork – *Ninox novaeseelandiae novaeseelandiae*. *in* https://www.xeno-canto.org/. The XC team.

Weary, D. M. and J. R. Krebs, 1992. Great tits classify songs by individual voice characteristics. Animal Behaviour 43:283–287.

Weilgart, L. S., 2007. A brief review of known effects of noise on marine mammals. International Journal of Comparative Psychology 20:159–168.

Weiser, E. L., R. B. Lanctot, S. C. Brown, J. A. Alves, P. F. Battley, R. Bentzen, J. Bêty, M. A. Bishop, M. Boldenow, L. Bollache, et al., 2016. Effects of geolocators on hatching success, return rates, breeding movements, and change in body mass in 16 species of arctic-breeding shorebirds. Movement Ecology 4:1–19.

Wijers, M., A. Loveridge, D. W. Macdonald, and A. Markham, 2021. CARACAL: a versatile passive acoustic monitoring tool for wildlife research and conservation. Bioacoustics 30:41–57.

Wijers, M., P. Trethowan, B. Du Preez, S. Chamaillé-Jammes, A. J. Loveridge, D. W. Macdonald, and A. Markham, 2020. Vocal discrimination of African lions and its potential for collar–free tracking. Bioacoustics 30:575–593.

Wijers, M., P. Trethowan, A. Markham, B. Du Preez, S. Chamaillé-Jammes, A. Loveridge, and D. Macdonald, 2018. Listening to lions: animal-borne acoustic sensors improve bio-logger calibration and behaviour classification performance. Frontiers in Ecology and Evolution 6:171.

Wiley, R. H., 1998. Ranging reconsidered. Behavioral Ecology and Sociobiology 42:143–146+147.

———, 2003. Is there an ideal behavioural experiment? Animal Behaviour 66:585–588.

Williams, E. M., C. F. J. O'Donnell, and D. P. Armstrong, 2018. Cost-benefit analysis of acoustic recorders as a solution to sampling challenges experienced monitoring cryptic species. Ecology and evolution 8:6839–6848.

Wilmers, C. C., B. Nickel, C. M. Bryce, J. A. Smith, R. E. Wheat, and V. Yovovich, 2015. The golden age of bio–logging: how animal–borne sensors are advancing the frontiers of ecology. Ecology 96:1741–1753.

Wilson, D. M. and J. Bart, 1985. Reliability of singing bird surveys: effects of song phenology during the breeding season. The Condor 87:69–73.

Wilson, D. R. and J. F. Hare, 2004. Animal communication: ground squirrel uses ultrasonic alarms. Nature 430:523.

Wöhr, M., 2018. Ultrasonic communication in rats: appetitive 50-khz ultrasonic vocalizations as social contact calls. Behavioral Ecology and Sociobiology 72:14.

Wong, A. and M. D. Gall, 2015. Frequency sensitivity in the auditory periphery of male and female Black–capped Chickadees (*Poecile atricapillus*). Zoology 118:357–363.

Wynia, A. L., V. Rolland, and J. E. Jiménez, 2019. Improved campephiline detection: An experiment conducted with the magellanic woodpecker. Ecology and Evolution 9:11724–11733.

Yip, D., L. Leston, E. Bayne, P. Sólymos, and A. Grover, 2017. Experimentally derived detection distances from audio recordings and human observers enable integrated analysis of point count data. Avian Conservation and Ecology 12.

Zahorik, P., 2002. Direct–to–reverberant energy ratio sensitivity. The Journal of the Acoustical Society of America 112:2110–2117.

Ziesemann, B., 2011. The social organisation and mating system of the Brown Kiwi (*Apteryx mantelli*). Thesis, Massey University.

Zuberogoitia, I., G. Burgos, J. A. González-Oreja, J. Morant, J. E. Martínez, and J. Zabala Albizua, 2019. Factors affecting spontaneous vocal activity of tawny owls *Strix aluco* and implications for surveying large areas. Ibis 161:495–503.

Appendices

I Appendix to Chapter 2

I.i Sample report paragraph #1 (Third party tracks)

“We obtained half of the original tracks from the Macaulay Library at the Cornell Lab of Ornithology and the rest from `xeno-canto.org` (ML1234, ML5678, ML9101, ML1121, ML3141, ML5161, XC7181, XC9202, XC1222, XC3242, XC5262, XC7282). All the original tracks were recorded in locations $>$ 800 km away from our experimental ones (see on-line repositories for details) and, given the sedentary behaviour reported for the target population (Otor, 2009), we do not see any reason to think that any of them were of individuals known by our targets. In `[Audio Editing Software Name and Version Here]` we joined the original tracks two by two (in the given order, for a total of six tracks), separated by one second of generated silence. We then applied a High Pass Filter at 1000 Hz, added 5 seconds of white noise at the beginning of each track (amplitude 0.4), and normalised each track to 0 dB amplitude. We did not apply any further algorithm; the finalised tracks are available at our research group website (`http://thisisour.url/research/media`) . Playback took place in three different locations within the plains of *Flat* natural reserve (#1: 12° 34.56 N, 78° 91.01 W; #2 11° 21.31 N 41° 51.61 W; #3 71° 81.92 N, 02° 12.22 W), on the Xth, Yth and Zth of Month Year respectively. All trials began at 21:00h and lasted 30 minutes each, consisting of 10 minutes of playback, 10 minutes of silence, repeat of the 10 minutes of playback. We broadcast a different pair of stimuli in each location to avoid pseudoreplication. The weather was generally stable throughout the three trials with $14^\circ C < T < 17^\circ C$, 65% $< H <$ 74%, and 3.4 Km/h<Wsp< 5.7 Km/h from 244° <Bearing< 275°. We recorded weather data using a `Brand` 1.23 portable weather station, that we set to log data every 15 seconds; the weather station logged data from 15 minutes before to 15 minutes after each trial and was placed on a tripod at a height of 1.5 m. Raw weather data is included as supplementary material to this paper and on our website. In all but one location (# 2) we did not record any activity/presence of individuals prior to the beginning of the playback. In location # 2, one male chased away a younger male around five minutes prior to the beginning of the experiment; both males were colour banded and we did not see them during the experiment. All three locations were sparsely vegetated, with only a few shrubs of Plant (*Genus speciensis*, L) within a radius of ~ 40 m from the loudspeaker. The loudspeaker (`Sp.Ker MAX`, frequency response $140 - 14,000$ Hz (chart included as Fig. S1), $\frac{\text{signal}}{\text{noise}} = 89 \pm 4$ dB) was set at 1.5 m height on a tripod, parallel to the ground. We used a `Sp.Lev M3734` Sound Pressure Level Meter (fast response, C setting – frequency response chart included as Fig. S2) to set broadcasting amplitude to 66 dB at 1 m distance, as this value has previously been reported to be the natural average for this species (Author *et al.*, 2011). We video-recorded all trials using a `O/Brand` 4.56 Action Camera fixed on top of the loudspeaker. During the trials, two observers sat in a previously (one week) assembled hide, approximately 30 m away, yet in sight, of the remotely activable loudspeaker. All the videos are available at our research group website.”

I.ii Sample report paragraph #2 (Self-recorded tracks)

"We recorded the original tracks at a sampling rate of 48 kHz and audio bit depth of 16, using a `MIC #1234` microphone (frequency response 40 to 20,000 Hz – chart included as Fig. S1 –, sensitivity -50 dB) connected to a `Sr.  EC X12345` digital recorder. The recording was carried out at the forest of *Trees* $(12\,^\circ 34.56N, 78\,^\circ 91.01W)$, in the city of *Town*, *Country*, during the breeding season, on Month Xth Year morning from hh:mm to hh:mm h. At the time of recording, temperature averaged 9° C, humidity 67%, and the wind was blowing from 256° at an average speed of 12 km/h (`Brand` X12 portable weather station, logging every 30 seconds throughout the recording session – raw log available as supplementary material). We recorded four different males perching on ~ 2 m high branches 5±2 m away from our hide; no obstacles were between us and the recorders at any time. No individual was solicited to vocalise. The measured SPLs were 62.3, 63.1, 66.0, and 65.4 dB measured at 5 ± 2 m (`Sp.Lev M3734` Sound Pressure Level Meter, fast response, A setting – chart included as Fig. S2); details for individual vocalisations are provided as comments with the tracks uploaded on-line (identifiers: ML7181, ML9202, XC1222 and XC3242). At the time of recording, base background noise ranged from 48 to 53 dB as measured with the same device and settings as the birds' vocalisations. We selected only tracks recorded whilst the recordee was facing the recorder. We used `[Audio Editing Software Name and Version Here]` to divide the five-hours long recording into individual vocalisations tracks (that we then uploaded on-line with the aforegiven identifiers). We did not apply any filter nor denoising algorithm. The playback experiment took place in the Forestpark Park, near the town of *City,Othercountry*, where we used an omni-directional (`Sp.Ker ∅`) loudspeaker, with frequency response between 40 Hz to 18,000 Hz (Fig. S3) and signal to noise ratio of 91 ± 2 dB; we placed the loudspeaker on the provided support (`Sp.047 tripod`), 1 m high at the presumed centres of four different territories. The presumed owners of territories were sighted and recognised (by their unique colour bands combinations) on-site prior to the beginning of the trials. Trials were initiated only if the territory owner had not been heard vocalising for $t \geq 10$ minutes. Vegetation was dense, with $0.3376 \pm 0.16\,{}^{\text{trees}}/_{m^2}$ in three out of four locations (coordinates, pictures and point-centred quarter method tables in the results section); location #3 was less dense, with $0.1875\,{}^{\text{trees}}/_{m^2}$. Dominant species were Onetree (*Primigenus specinæ*, L) in locations 1 (73%) and 2 (54%), Anothertree (*Aliumgenus speciferus*, L) in location 3 (80%) and Lastree (*Ultimugenus specii*, L) in location 4 (64%). We used a different track for each location to avoid pseudoreplication, and we set the loudspeaker to broadcast each track at its originally measured amplitude to avoid unnatural broadcast (5 m away from the speaker using the previously mentioned settings on hte `Sp.Lev M3734` Sound Pressure Level Meter)."

I.iii

Table I.i: Number of articles found and included from the different on-line databases

Source	Found	Included
PubMed Central	73	38
Scopus	177	115
Web of Science	601	406
Total	625	419

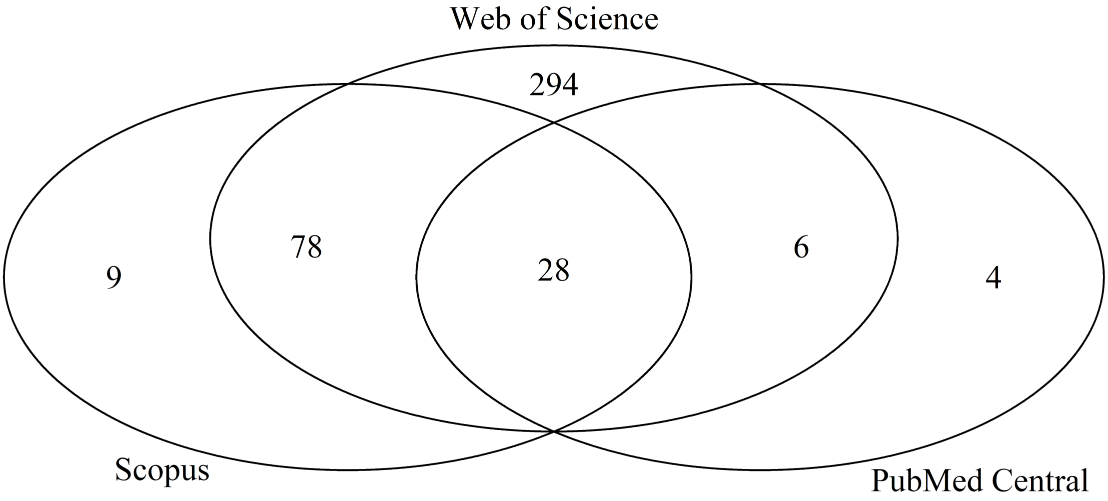
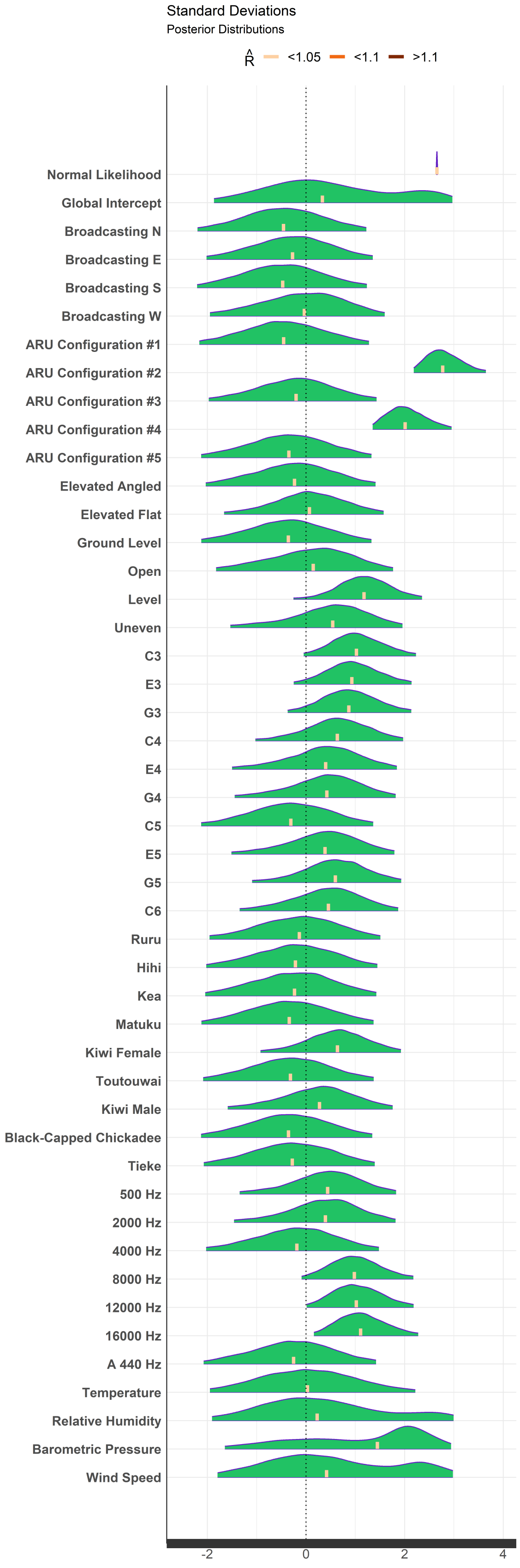


Figure I.i: Depiction of which on-line resources found the articles included in this study. Numbers within each section represents the number of articles found by each resource; i.e., we found 294 of the included articles only through Web of Science, whereas 6 articles were found by both PubMed Central and Web of Science.

II Appendix to Chapter 3

II.i Figures and tables



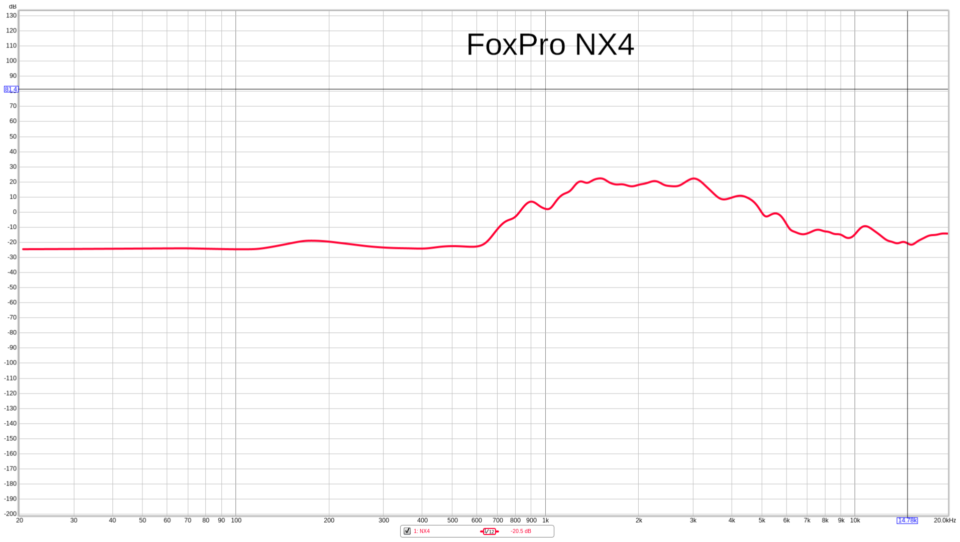


Figure II.ii: Frequency reponse curve of our FoxPro NX4 speker at 1/12 smoothing.

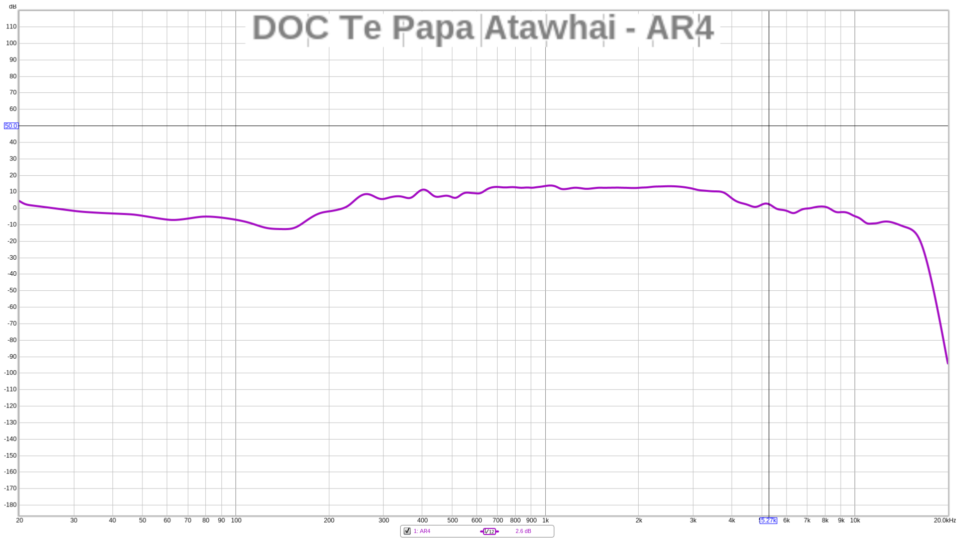


Figure II.iii: Frequency reponse curve of a DOC Te Papa Atawhai AR4 ARU at 1/12 smoothing.The fall above 16 KHz is due to the fact that the AR4 samples at 32 KHz, effectively only recording up to 16 KHz

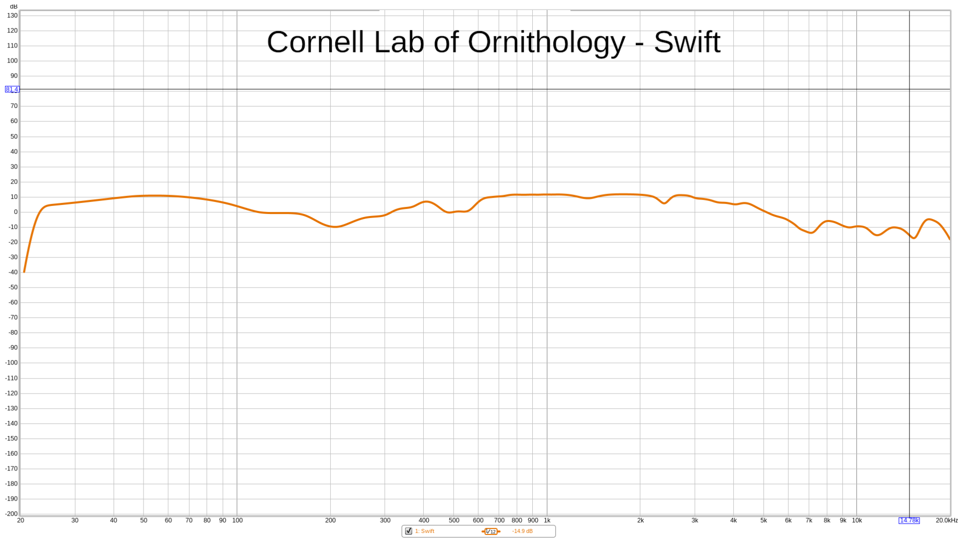


Figure II.iv: Frequency reponse curve of one of our Cornell Lab of Ornithology Swift ARUs at 1/12 smoothing.

Table II.i: Summary of the distribution of the Euclidian distances (in metres) between the first TDOA and energy estimates.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
energy	16.44	20.87	21.74	21.49	22.29	27.83
TDOA	0.00	3.16	8.00	6.65	9.00	15.52

Table II.ii: Summary of the distribution of the Euclidian distances (in metres) between the TDOA method and our best energy–based ARUs configuration (#4).

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
energy	2.68	9.62	11.31	13.29	16.21	39.95
TDOA	0.00	3.16	8.00	6.65	9.00	15.52

II.ii Generalised cross-correlation script

```
import os
import yaml
import csv
import wave
import numpy as np
import scipy.fft as fft
import matplotlib.pyplot as plt
from yaml import CLoader as Loader

class MicrophoneArray:
    def __init__(self, coordinates, centre, c_sound=343):
        self.coordinates = coordinates
        self.centre = centre
        self.c_sound = c_sound

    def load_data(self, dir, start, length, prefix, num_channels):
        data = []
        self.U = []
        for k in range(num_channels):
            fname = os.path.join(dir, '{:s}{:d}.wav'.format(prefix, k+1))
            fd = wave.open(fname)
            if k == 0:
                self.fs = fd.getframerate()
            width = fd.getsampwidth()

            if width == 2:
                datatype = np.int16
                maxval = 2**15
            elif width == 4:
                datatype = np.int32
                maxval = 2**31
            else:
                raise Exception("Can only handle 16 bit and 32 bit integer data")

            #Fast-forward to start value
            fd.setpos(round(start*self.fs))

            frames = fd.readframes(round(length*self.fs))
            x = np.frombuffer(frames, dtype=datatype)/maxval

            # Process data
            # Remove mean
            x -= np.mean(x)

            # Zero pad
            x = np.concatenate((x, np.zeros(100)))

            data.append(x)
            self.U.append(fft.rfft(x))
        self.data = np.array(data).T
        self.n_frames = len(x)

    def delay_time(self, theta, pos):

        x = [pos[0] - self.centre[0], pos[1] - self.centre[1]]
        return -(x[0]*np.cos(theta) + x[1]*np.sin(theta))/self.c_sound

    def delay_sample(self, theta, pos):
        return self.fs*self.delay_time(theta, pos)

    def gcc_phat(self, pairs, pos):
        # Calculate angles of arrival
        x = pos
        theta = np.arctan2(x[1] - self.centre[1], x[0] - self.centre[0])

        npairs = len(pairs)
        cc = []
        for k in range(npairs):
            i1, i2 = pairs[k]
            R = self.U[i1]*np.conj(self.U[i2])
            R = R/np.abs(R)
            foo = fft.irfft(R)
            d1 = self.delay_sample(theta, self.coordinates[i1])
            d2 = -self.delay_sample(theta, self.coordinates[i2])
            idx = d1 + d2
            cc.append(foo[(idx.round()).astype(int)])
        return np.array(cc).T

def load_experiment_details(config):
    """Load the relevant experiment details and offset from file"""
    with open(config['experiment_file']) as csvfile:
```

```
        reader = csv.DictReader(csvfile)
        rows = []
        for row in reader:
            rows.append(row)

    with open(config['sound_offsets_file']) as csvfile:
        reader = csv.reader(csvfile)
        offsets = []
        # Strip off the header row
        header = next(reader)
        for row in reader:
            offsets.append(float(row[0]))
        offset = offsets[config['sound_id']]

    return rows[config['experiment_id']], offset

def construct_mic_arrays(config, details):
    """Construct microphone arrays from provided details"""
    arr = {}
    x = np.array(config['coordinates'])
    for key in config['keys']:
        theta = config['array_orientations'][key]*np.pi/180
        c = np.cos(theta)
        s = np.sin(theta)

        # Fix the orientation
        R = np.array([[c, -s], [s, c]])
        centre = np.array(config['array_positions'][key])
        pos = np.dot(x, R) + centre # Using the inverse of R, but multiplying on right, means R is what we have
        arr[key] = MicrophoneArray(pos, centre, c_sound=config['speed_of_sound'])
    return arr



# ------------------- MAIN SCRIPT STARTS HERE ---------------------#
# Start by reading in our configuration file.
with open('./confi.yml') as configfile:
    config = yaml.load(configfile, Loader = Loader)

# Load the microphone arrays
details, offset = load_experiment_details(config)
# Or, just load offset directly
offset = -config['sound_offset']
print("offset=")
print(offset)
# Now construct the arrays
mic_arrays = construct_mic_arrays(config, details)

num_channels = config['num_channels']
length = config['window']
keys = config['keys']
dirs = {}
start = {}

aupfiles = {'c': 'path C', 'e': 'path E', 'w': 'path W'}
times = {'c': 'time C', 'e': 'time E', 'w': 'time W'}
for key in keys: # keys is a subset of ['w', 'c', 'e']
    aupfile = os.path.join(config['data_dir'], details[aupfiles[key]])
    base, fname = os.path.split(aupfile)
    fname_base, ext = os.path.splitext(fname)
    dirs[key] = os.path.join(base, config['wav_dir'], fname_base)
    start[key] = float(details[times[key]]) - offset

plt.figure(1)
i_subplot = 1
for key in keys:
    m = mic_arrays[key]
    m.load_data(dirs[key], start[key], length, config['channel_prefix'], config['num_channels'])
    t = 1/m.fs * np.arange(m.n_frames)
    # Plot waveform from centre channel
    plt.subplot(1,len(keys),i_subplot)
    plt.plot(t, mic_arrays[key].data)
    i_subplot += 1

x = np.arange(-75, 76, 1)
y = np.arange(-50, 51, 1)
X, Y = np.meshgrid(x, y)
cc = {}
for key in keys:
    m = mic_arrays[key]
    # Calculate GCC-PHAT values for each pair from current mic array. Don't
    # combine them yet
    cc[key] = m.gcc_phat(config['pairs'], [X, Y])
```

```
cc = np.concatenate([cc[key] for key in keys], axis=2)
img = (1/np.sum(1/np.abs(cc), 2)).T
plt.figure(2)
plt.imshow(img, origin=’lower’, extent=[-75.5, 75.5, -50.5, 50.5])

# Find the maximum
idx = np.unravel_index(img.argmax(), img.shape)
print(idx)
plt.plot(x[idx[1]], y[idx[0]], ’ro’)


print(details)
print("praa")
print(y[idx[0]])
plt.show()
```

III Energy/TOA comparison model

Parameter	$\hat{R}$	n_{eff}	mean	sd	2.5%	50%	97.5%
Intercept	1.000	1232	9.318	1.098	7.207	9.331	11.463
Energy Based	1.000	1238	12.111	1.100	10.001	12.110	14.212
TOA	1.000	1253	-2.651	1.104	-4.778	-2.672	-0.514

Table III.iii: Summary table of a simple model comparing estimates between using energy–based or time of arrival localisations. The more positive numbers showed for the energy based method mean bigger errors involved with this method.

IV Appendix to Chapter 4

Parameter	Rhat	n_eff	mean	sd	2.5%	50%	97.5%
bs[1]	1.0000	27172	-0.6466	1.0728	-2.7585	-0.6452	1.4394
bs[2]	1.0000	27150	0.1535	1.0721	-1.9618	0.1505	2.2268
bs[3]	1.0000	27158	0.5250	1.0721	-1.5885	0.5238	2.6010
ba[1]	0.9999	29043	-0.6730	1.0714	-2.7858	-0.6771	1.4391
ba[2]	0.9999	29114	0.1776	1.0709	-1.9263	0.1707	2.2857
ba[3]	0.9999	29084	0.5342	1.0705	-1.5746	0.5276	2.6373
bd[1]	1.0000	29285	-0.0070	1.9891	-3.9549	-0.0041	3.8945
bd[2]	0.9999	29599	0.0169	1.9989	-3.9180	0.0196	3.9213
bd[3]	1.0000	31097	0.0055	2.0024	-3.8802	0.0116	3.9438
bd[4]	1.0000	30773	-0.0142	1.9867	-3.8741	-0.0227	3.9247
bt[1]	1.0001	29236	-0.9130	1.2932	-3.4592	-0.9115	1.6091
bt[2]	1.0001	29532	0.9160	1.2829	-1.5955	0.9064	3.4438
bst[1]	0.9999	25510	0.4890	0.8615	-1.2057	0.4952	2.1905
bst[2]	0.9999	25149	0.0748	0.8607	-1.6052	0.0742	1.7817
bst[3]	0.9999	25143	-0.1533	0.8610	-1.8430	-0.1527	1.5526
bst[4]	0.9999	25454	-0.0651	0.8606	-1.7488	-0.0636	1.6338
bst[5]	0.9999	25323	-0.3338	0.8619	-2.0152	-0.3298	1.3659
bw[1]	1.0001	28740	-0.5947	0.1118	-0.8162	-0.5935	-0.3774
bw[2]	1.0000	29870	0.6599	0.1319	0.4024	0.6593	0.9189
bw[3]	1.0001	29236	-0.3677	0.0988	-0.5635	-0.3668	-0.1757
bw[4]	1.0000	28892	0.1000	0.0638	-0.0259	0.1002	0.2258
bw[5]	1.0001	28903	0.4708	0.1663	0.1455	0.4701	0.7998
bw[6]	1.0000	30506	0.1240	0.1477	-0.1641	0.1242	0.4142
bz	1.0001	29459	0.0317	1.5302	-2.9613	0.0398	3.0533
log-posterior	1.0000	25468	-16.9808	3.4442	-24.6607	-16.6490	-11.2094

Table IV.i: Poisson Model Summary Table

Parameter	Rhat	n_eff	mean	sd	2.5%	50%	97.5%
bs[1]	1.0000	29469	-0.9186	1.2892	-3.4345	-0.9145	1.6116
bs[2]	1.0000	29527	-0.9449	1.2887	-3.4488	-0.9457	1.5737
bt[1]	1.0000	30998	-0.9745	1.7376	-4.3744	-0.9666	2.4224
bt[2]	1.0001	30280	-0.8941	1.6230	-4.0817	-0.9057	2.2737
bd[1]	1.0000	30415	0.4040	1.2962	-2.1059	0.4052	2.9599
bd[2]	1.0000	30371	-0.0304	1.0347	-2.0427	-0.0292	2.0045
bd[3]	1.0000	30083	-0.8436	1.0525	-2.8996	-0.8443	1.2021
bd[4]	0.9999	29963	-1.4191	1.2799	-3.9213	-1.4289	1.0977
ba[1]	0.9999	28426	-0.8154	1.3018	-3.3597	-0.8209	1.7464
ba[2]	0.9999	28522	-1.0378	1.3013	-3.5766	-1.0415	1.4993
bst[1]	1.0000	26749	-0.3453	0.8867	-2.0975	-0.3436	1.3921
bst[2]	1.0000	26733	-0.3086	0.8812	-2.0427	-0.3078	1.4104
bst[3]	1.0000	26658	-0.3684	0.8861	-2.1205	-0.3676	1.3521
bst[4]	1.0000	26676	-0.3197	0.8857	-2.0671	-0.3220	1.4138
bst[5]	1.0000	26603	-0.5136	0.8902	-2.2836	-0.5078	1.2152
bw[1]	1.0000	28939	0.4590	0.3814	-0.3016	0.4640	1.1913
bw[2]	1.0000	29803	0.3218	0.6360	-0.8964	0.3146	1.5781
bw[3]	1.0000	29987	-0.1016	0.1892	-0.4763	-0.0988	0.2638
bw[4]	0.9999	30578	-0.0408	0.1301	-0.2983	-0.0410	0.2144
bw[5]	1.0000	29484	-0.7080	0.6249	-1.9217	-0.7098	0.5135
bw[6]	1.0000	30139	-0.3091	0.6536	-1.5835	-0.3064	0.9658
bz	1.0000	29367	-1.8549	1.6357	-5.0512	-1.8688	1.3632
log-posterior	1.0000	24741	-1152.0520	3.3330	-1159.4222	-1151.7187	-1146.4890

Table IV.ii: Exponential Model Summary Table

Parameter	Rhat	n_eff	mean	sd	2.5%	50%	97.5%
bs[1]	1.0001	27521	-0.3131	1.0882	-2.4404	-0.3080	1.8085
bs[2]	1.0001	27939	-0.2176	1.0826	-2.3487	-0.2133	1.9004
bs[3]	1.0000	27855	-0.2447	1.0805	-2.3562	-0.2420	1.8448
ba[1]	1.0000	27837	0.1302	1.0814	-1.9863	0.1321	2.2592
ba[2]	1.0000	27968	-0.5658	1.0782	-2.6797	-0.5619	1.5444
ba[3]	0.9999	27959	-0.3233	1.0772	-2.4347	-0.3216	1.7945
bd[1]	0.9999	31210	-0.0183	1.9980	-3.9369	-0.0086	3.8630
bd[2]	1.0001	30400	0.0104	2.0057	-3.9338	0.0078	3.9243
bd[3]	1.0000	30623	-0.0049	1.9953	-3.8927	-0.0048	3.9250
bd[4]	1.0000	30329	0.0031	1.9898	-3.9114	0.0053	3.8890
bt[1]	1.0000	29344	-1.3203	1.5296	-4.3033	-1.3111	1.6799
bt[2]	1.0002	29119	0.5645	1.4095	-2.1818	0.5806	3.3234
bst[1]	1.0001	25175	0.8421	0.8818	-0.8637	0.8429	2.5735
bst[2]	1.0002	25148	-0.3177	0.8763	-2.0097	-0.3131	1.3975
bst[3]	1.0001	25450	-0.1801	0.8879	-1.9134	-0.1802	1.5617
bst[4]	1.0001	25188	-0.2964	0.8818	-2.0120	-0.2955	1.4534
bst[5]	1.0002	25523	-0.8601	0.8910	-2.5969	-0.8648	0.8799
bw[1]	1.0002	29241	-0.1175	0.2595	-0.6373	-0.1158	0.3884
bw[2]	1.0000	28566	0.1100	0.3768	-0.6273	0.1085	0.8533
bw[3]	1.0000	30676	-0.2352	0.3062	-0.8421	-0.2315	0.3584
bw[4]	1.0000	29971	-0.1622	0.1628	-0.4846	-0.1603	0.1529
bw[5]	1.0003	29406	-0.1991	0.4196	-1.0118	-0.2010	0.6333
bw[6]	1.0001	29559	0.3167	0.4574	-0.5967	0.3173	1.2066
bz	1.0002	28096	-0.7633	1.5252	-3.7598	-0.7680	2.2352
log-posterior	1.0000	24519	-375.8899	3.5014	-383.6861	-375.5508	-370.0794

Table IV.iii: Binomial Model Summary Table

V Appendix to Chapter 5

V.i Exponential model

Parameter	Rhat	n_eff	mean	sd	2.5%	50%	97.5%
bs[1]	1.000	29282	0.112	0.747	-1.359	0.116	1.572
bs[2]	1.000	29950	-0.106	0.745	-1.563	-0.107	1.351
bsp[1]	1.000	29845	-0.030	0.348	-0.717	-0.031	0.642
bsp[2]	1.000	29808	-0.000	0.349	-0.690	-0.000	0.672
bsp[3]	1.000	30403	-0.066	0.371	-0.783	-0.064	0.658
bsp[4]	1.000	29974	0.088	0.373	-0.633	0.087	0.819
bg[1]	1.000	29756	-0.014	0.712	-1.394	-0.013	1.384
bg[2]	1.000	29840	0.004	0.712	-1.370	0.005	1.407
bo[1]	1.000	26885	0.043	0.501	-0.940	0.045	1.027
bo[2]	1.000	26880	-0.015	0.501	-0.997	-0.015	0.957
bo[3]	1.000	26952	-0.032	0.502	-1.021	-0.032	0.946
bo[4]	1.000	26927	-0.021	0.502	-1.004	-0.022	0.957
bh[1]	1.000	29732	-0.002	0.707	-1.380	-0.001	1.393
bh[2]	1.000	29807	0.018	0.708	-1.367	0.018	1.418
bw[1]	1.000	30351	-0.037	0.050	-0.137	-0.036	0.059
bw[2]	1.000	30612	-0.009	0.055	-0.116	-0.009	0.098
bw[3]	1.000	29447	-0.004	0.041	-0.085	-0.003	0.075
bw[4]	1.000	31251	-0.006	0.039	-0.083	-0.006	0.071
bw[5]	1.000	30811	0.001	0.039	-0.074	0.001	0.078
bw[6]	1.000	29726	-0.045	0.050	-0.143	-0.045	0.054
bz	1.000	28573	-3.068	1.352	-5.698	-3.070	-0.423
log-posterior	1.000	24435	-4623.354	3.254	-4630.634	-4623.007	-4617.957

Table V.i: Exponential Model Summary Table

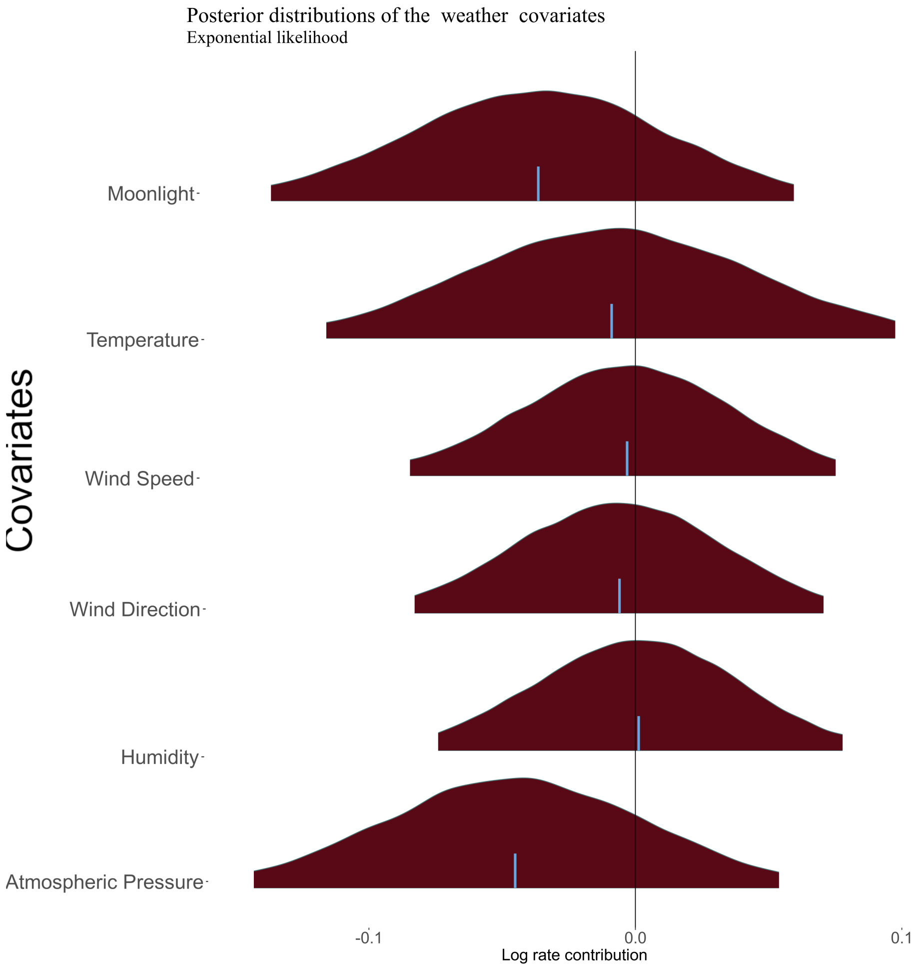


Figure V.i: Weather covariates for the Exponential model. Covariates with posterior distributions overlapping zero (vertical line) are not consistently associated with the response variable.

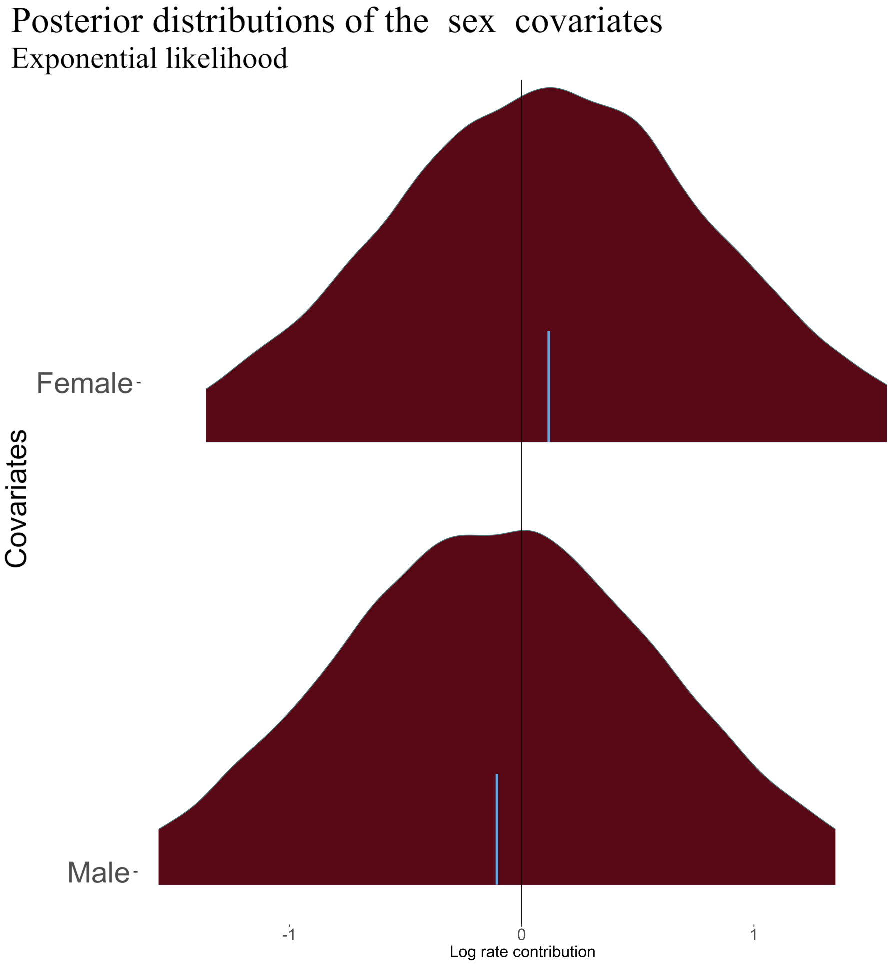


Figure V.ii: Sex covariate for the Exponential model. Covariates with posterior distributions overlapping zero (vertical line) are not consistently associated with the response variable.

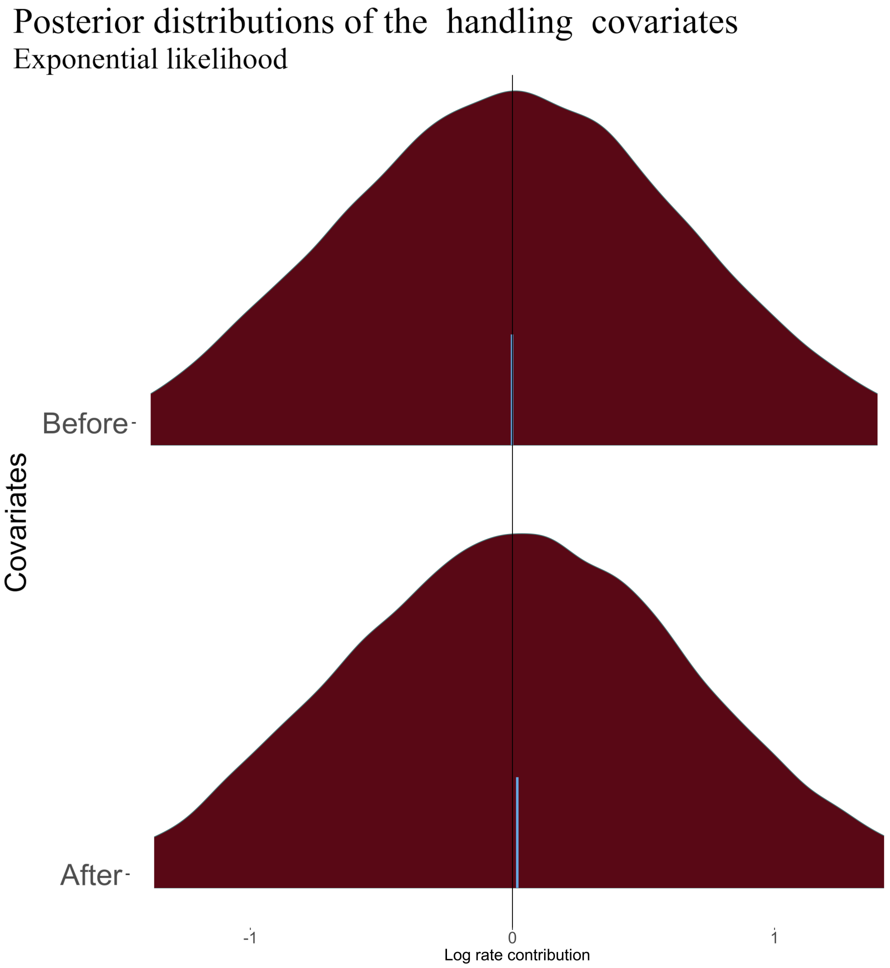


Figure V.iii: Handling covariate for the Exponential model. Covariates with posterior distributions overlapping zero (vertical line) are not consistently associated with the response variable.

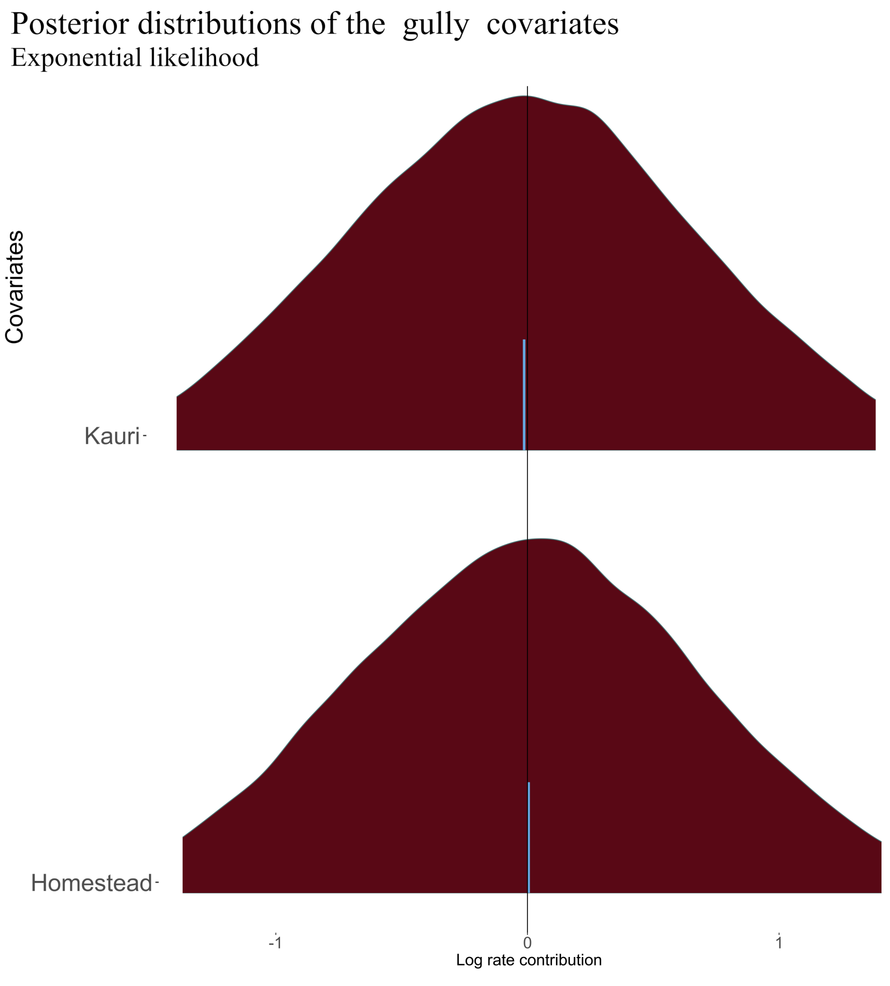


Figure V.iv: Gully covariate for the Exponential model. Covariates with posterior distributions overlapping zero (vertical line) are not consistently associated with the response variable.

VI Appendix to Chapter 6

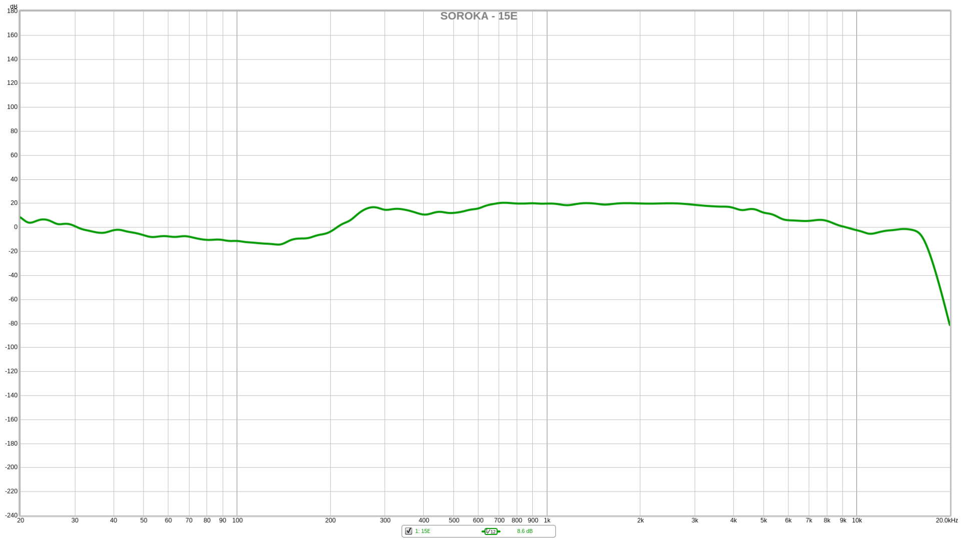


Figure VI.i: ABARs frequency response chart; similarly to the AR4, these devices only sample up to 32KHz, hence the fall in response above 16KHz.

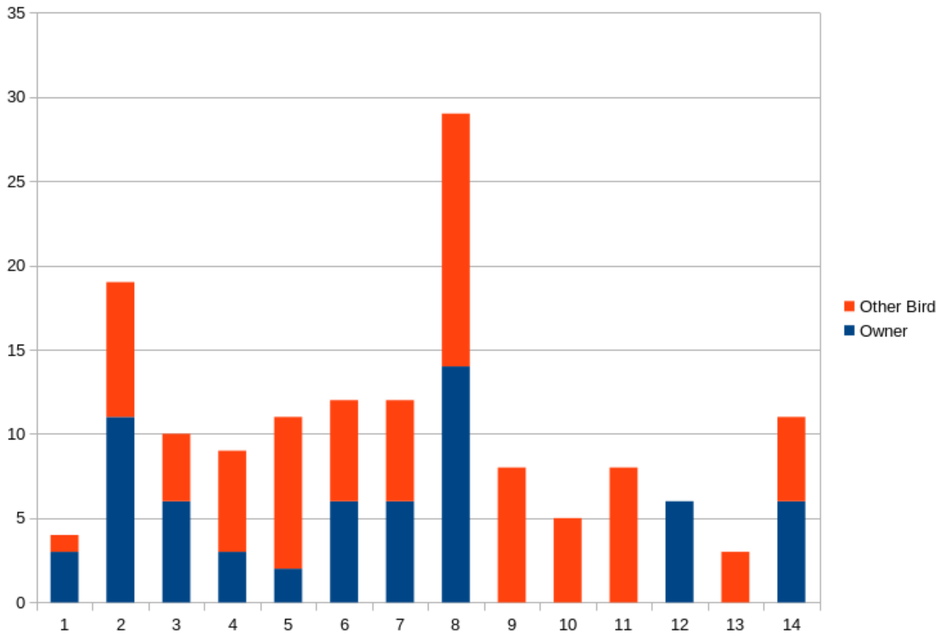


Figure VI.ii: Number of vocalisations recorded by each of the fourteen ABARs in 2020. The blue fraction of each bar represents the vocalisations of the animals wearing the device whilst the orange fraction represents the vocalisations of other NIBK.

VI.i Unmarked model

```
################ data
### alpha = global offset / intercept
### beta_sex = 2 levels, male female
### sex = two dimensions array with 1 as female and 2 as male {iota x max(nu[i])}
```

```
### iota = integer, number of populations // sampling events | input
### nu = one dimension array of integers of length iota, number of known individuals per population | input
### kappa = one dimension array of integers of length iota, number of sampling events per population | input
### lambda = three dimensions array {iota x nu[i] x kappa[i]}, rate of vocalisations per sampling event per individual per population
### eta = three dimensions array {iota x nu[i] x kappa[i]}, number of vocalisations per sampling event per individual per population | input
### rho = two dimensions array prob of recruitment of individual in population i of sex s
### gamma = recruitment bernoullian variable based on rho
### psi = two dimensions array, {iota x 2 (sex)} contains the number of individuals in population i of sex s
### mu = rate of vocalisation for recuited individuals // counterpart of lambda
### theta = number of vocalisation for recuited individuals // counterpart of eta
### omicron = real number of vocalisations per population, per sex per sampling event | input
### zeta = estimated number of individuals per population per sex

model{
########### covariates

       alpha~dnorm(0,1e-3)
       for(s in 1:2)
       {
              beta_sex[s]~dnorm(0,1e-3)
              #beta_toy[s]~dnorm(0,1e-3)
       }

       for(i in 1:iota)
       {
              beta_occ[i]~dnorm(0,1e-3)
       }
       for(i in 1:3)
       {
              beta_pop[i]~dnorm(0,1e-3)
       }

       for(q in 1:6)
       {
              beta_weather[q]~dnorm(0,1e-3)
       }

       for(i in 1:iota)
       {
              for(s in 1:2)
              {
                     rho[i,s]~dbeta(1,1)
                     for(w in 1:omega)
                     {
                            xi[i,s,w]~dbern(rho[i,s])
                            for(k in 1:kappa[i])
                            {
                                   logit(mu[i,w,k,s])<-alpha+beta_sex[s]+beta_weather[1]*lot[i,k]+beta_pop[qp[i]]+beta_occ[i]+beta_weather[2]*hit[i,k]+beta_weather[3]*ws[i,k]+beta_weather[4]*rh[i,k]+beta_weather[5]*ap[i,k]+beta_weather[6]*ml[i,k]
                                   theta[i,w,k,s]~dpois(mu[i,w,k,s]*xi[i,s,w])

                            }
                     }
              }
       }

       for(i in 1:iota)
       {
              for(s in 1:2)
              {
                     for(k in 1:kappa[i])
                     {
              omicron[i,s,k]~dsum(theta[i,1,k,s],theta[i,2,k,s],theta[i,3,k,s],theta[i,4,k,s],theta[i,5,k,s],theta[i,6,k,s],theta[i,7,k,s],theta[i,8,k,s],theta[i,9,k,s],theta[i,10,k,s],theta[i,11,k,s],theta[i,12,k,s],theta[i,13,k,s],theta[i,14,k,s],theta[i,15,k,s],theta[i,16,k,s],theta[i,17,k,s],theta[i,18,k,s],theta[i,19,k,s],theta[i,20,k,s],theta[i,21,k,s],theta[i,22,k,s],theta[i,23,k,s],theta[i,24,k,s],theta[i,25,k,s],theta[i,26,k,s],theta[i,27,k,s],theta[i,28,k,s],theta[i,29,k,s],theta[i,30,k,s],theta[i,31,k,s],theta[i,32,k,s],theta[i,33,k,s],theta[i,34,k,s],theta[i,35,k,s],theta[i,36,k,s],theta[i,37,k,s],theta[i,38,k,s],theta[i,39,k,s],theta[i,40,k,s],theta[i,41,k,s],theta[i,42,k,s],theta[i,43,k,s],theta[i,44,k,s],theta[i,45,k,s],theta[i,46,k,s],theta[i,47,k,s],theta[i,48,k,s],theta[i,49,k,s],theta[i,50,k,s],theta[i,51,k,s],theta[i,52,k,s],theta[i,53,k,s],theta[i,54,k,s],theta[i,55,k,s],theta[i,56,k,s],theta[i,57,k,s],theta[i,58,k,s],theta[i,59,k,s],theta[i,60,k,s],theta[i,61,k,s],theta[i,62,k,s],theta[i,63,k,s],theta[i,64,k,s],theta[i,65,k,s],theta[i,66,k,s],theta[i,67,k,s],theta[i,68,k,s],theta[i,69,k,s],theta[i,70,k,s],theta[i,71,k,s],theta[i,72,k,s],theta[i,73,k,s],theta[i,74,k,s],theta[i,75,k,s],theta[i,76,k,s],theta[i,77,k,s],theta[i,78,k,s],theta[i,79,k,s],theta[i,80,k,s],theta[i,81,k,s],theta[i,82,k,s],theta[i,83,k,s],theta[i,84,k,s],theta[i,85,k,s],theta[i,86,k,s],theta[i,87,k,s],theta[i,88,k,s],theta[i,89,k,s],theta[i,90,k,s],theta[i,91,k,s],theta[i,92,k,s],theta[i,93,k,s],theta[i,94,k,s],theta[i,95,k,s],theta[i,96,k,s],theta[i,97,k,s],theta[i,98,k,s],theta[i,99,k,s],theta[i,100,k,s])
```

```
                                }
                        }
                }

                for(i in 1:iota)
                {
                        for(s in 1:2)
                        {
                        zeta[i,s]<-sum(xi[i,s,])
                        }
                }
}
```

VI.ii ABARs model

```
################ data
### alpha = global offset / intercept
### beta_sex = 2 levels, male female
### sex = two dimensions array with 1 as female and 2 as male {iota x max(nu[i])}
### iota = integer, number of populations // sampling events | input
### nu = one dimension array of integers of length iota, number of known individuals per population | input
### kappa = one dimension array of integers of length iota, number of sampling events per population | input
### lambda = three dimensions array {iota x nu[i] x kappa[i]}, rate of vocalisations per sampling event per
     individual per population
### eta = three dimensions array {iota x nu[i] x kappa[i]}, number of vocalisations per sampling event per
     individual per population | input
### rho = two dimensions array prob of recruitment of individual in population i of sex s
### gamma = recruitment bernoullian variable based on rho
### psi = two dimensions array, {iota x 2 (sex)} contains the number of individuals in population i of sex s
### mu = rate of vocalisation for recuited individuals // counterpart of lambda
### theta = number of vocalisation for recuited individuals // counterpart of eta
### omicron = real number of vocalisations per population, per sex per sampling event | input
### zeta = estimated number of individuals per population per sex

model{
########### covariates

        alpha~dnorm(0,1e-3)
        for(s in 1:2)
        {
                beta_sex[s]~dnorm(0,1e-3)
                beta_toy[s]~dnorm(0,1e-3)
        }

        for(i in 1:iota)
        {
                beta_occ[i]~dnorm(0,1e-3)
        }
        for(i in 1:3)
        {
                beta_pop[i]~dnorm(0,1e-3)
        }

        for(q in 1:6)
        {
                beta_weather[q]~dnorm(0,1e-3)
        }

        for(i in 1:iota)
        {
                for(n in 1:nu[i])
                {
                        for(k in 1:kappa[i])
                        {
                                logit(lambda[i,n,k])<-alpha+beta_sex[sex[i,n]]+beta_toy[toy[i]]+beta_weather[1]*lot[i
                                     ,k]+beta_pop[qp[i]]+beta_occ[i]+beta_weather[2]*hit[i,k]+beta_weather[3]*ws[i,k
                                     ]+beta_weather[4]*rh[i,k]+beta_weather[5]*ap[i,k]+beta_weather[6]*ml[i,k]
                                eta[i,n,k]~dpois(lambda[i,n,k])
                        }
                }
        }

        for(i in 1:iota)
        {
                for(s in 1:2)
                {
                        rho[i,s]~dbeta(1,1)
                        for(w in 1:omega)
                        {
                                xi[i,s,w]~dbern(rho[i,s])
```

```
                                for(k in 1:kappa[i])
                                {
                                        logit(mu[i,w,k,s])<-alpha+beta_sex[s]+beta_toy[toy[i]]+beta_weather[1]*lot[i,k
                                             ]+beta_pop[qp[i]]+beta_occ[i]+beta_weather[2]*hit[i,k]+beta_weather[3]*ws
                                             [i,k]+beta_weather[4]*rh[i,k]+beta_weather[5]*ap[i,k]+beta_weather[6]*ml[
                                             i,k]
                                        theta[i,w,k,s]~dpois(mu[i,w,k,s]*xi[i,s,w])

                                }
                        }
                }
        }

        for(i in 1:iota)
        {
                for(s in 1:2)
                {
                        for(k in 1:kappa[i])
                        {
                omicron[i,s,k]~dsum(theta[i,1,k,s],theta[i,2,k,s],theta[i,3,k,s],theta[i,4,k,s],theta[i,5,k,s],theta
                     [i,6,k,s],theta[i,7,k,s],theta[i,8,k,s],theta[i,9,k,s],theta[i,10,k,s],theta[i,11,k,s],theta[i
                     ,12,k,s],theta[i,13,k,s],theta[i,14,k,s],theta[i,15,k,s],theta[i,16,k,s],theta[i,17,k,s],theta
                     [i,18,k,s],theta[i,19,k,s],theta[i,20,k,s],theta[i,21,k,s],theta[i,22,k,s],theta[i,23,k,s],
                     theta[i,24,k,s],theta[i,25,k,s],theta[i,26,k,s],theta[i,27,k,s],theta[i,28,k,s],theta[i,29,k,s
                     ],theta[i,30,k,s],theta[i,31,k,s],theta[i,32,k,s],theta[i,33,k,s],theta[i,34,k,s],theta[i,35,k
                     ,s],theta[i,36,k,s],theta[i,37,k,s],theta[i,38,k,s],theta[i,39,k,s],theta[i,40,k,s],theta[i
                     ,41,k,s],theta[i,42,k,s],theta[i,43,k,s],theta[i,44,k,s],theta[i,45,k,s],theta[i,46,k,s],theta
                     [i,47,k,s],theta[i,48,k,s],theta[i,49,k,s],theta[i,50,k,s],theta[i,51,k,s],theta[i,52,k,s],
                     theta[i,53,k,s],theta[i,54,k,s],theta[i,55,k,s],theta[i,56,k,s],theta[i,57,k,s],theta[i,58,k,s
                     ],theta[i,59,k,s],theta[i,60,k,s],theta[i,61,k,s],theta[i,62,k,s],theta[i,63,k,s],theta[i,64,k
                     ,s],theta[i,65,k,s],theta[i,66,k,s],theta[i,67,k,s],theta[i,68,k,s],theta[i,69,k,s],theta[i
                     ,70,k,s],theta[i,71,k,s],theta[i,72,k,s],theta[i,73,k,s],theta[i,74,k,s],theta[i,75,k,s],theta
                     [i,76,k,s],theta[i,77,k,s],theta[i,78,k,s],theta[i,79,k,s],theta[i,80,k,s],theta[i,81,k,s],
                     theta[i,82,k,s],theta[i,83,k,s],theta[i,84,k,s],theta[i,85,k,s],theta[i,86,k,s],theta[i,87,k,s
                     ],theta[i,88,k,s],theta[i,89,k,s],theta[i,90,k,s],theta[i,91,k,s],theta[i,92,k,s],theta[i,93,k
                     ,s],theta[i,94,k,s],theta[i,95,k,s],theta[i,96,k,s],theta[i,97,k,s],theta[i,98,k,s],theta[i
                     ,99,k,s],theta[i,100,k,s])
                        }
                }
        }

        for(i in 1:iota)
        {
                for(s in 1:2)
                {
                zeta[i,s]<-sum(xi[i,s,])+psi[i,s]
                }
        }
}
```

VI.iii ABARs–informed model for unmarked populations

```
################ data
### alpha = global offset / intercept
### beta_sex = 2 levels, male female
### sex = two dimensions array with 1 as female and 2 as male {iota x max(nu[i])}
### iota = integer, number of populations // sampling events | input
### nu = one dimension array of integers of length iota, number of known individuals per population | input
### kappa = one dimension array of integers of length iota, number of sampling events per population | input
### lambda = three dimensions array {iota x nu[i] x kappa[i]}, rate of vocalisations per sampling event per
     individual per population
### eta = three dimensions array {iota x nu[i] x kappa[i]}, number of vocalisations per sampling event per
     individual per population | input
### rho = two dimensions array prob of recruitment of individual in population i of sex s
### gamma = recruitment bernoullian variable based on rho
### psi = two dimensions array, {iota x 2 (sex)} contains the number of individuals in population i of sex s
### mu = rate of vocalisation for recuited individuals // counterpart of lambda
### theta = number of vocalisation for recuited individuals // counterpart of eta
### omicron = real number of vocalisations per population, per sex per sampling event | input
### zeta = estimated number of individuals per population per sex

model{
########### covariates

        alpha~dnorm(0,1e-3)
        for(s in 1:2)
        {
                beta_sex[s]~dnorm(0,1e-3)
        }
```

```
        for(i in 1:(delta+iota))
        {
                beta_occ[i]~dnorm(0,1e-3)
        }
        for(q in 1:6)
        {
                beta_weather[q]~dnorm(0,1e-3)
        }

        for(i in 1:iota)
        {
                for(n in 1:nu[i])
                {
                        for(k in 1:kappa[i])
                        {
                                logit(lambda[i,n,k])<-alpha+beta_sex[sex[i,n]]+beta_weather[1]*lot[i,k]+beta_occ[i]+
                                      beta_weather[2]*hit[i,k]+beta_weather[3]*ws[i,k]+beta_weather[4]*rh[i,k]+
                                      beta_weather[5]*ap[i,k]+beta_weather[6]*ml[i,k]
                                eta[i,n,k]~dpois(lambda[i,n,k])
                        }
                }
        }

        for(i in 1:(iota+delta))
        {
                for(s in 1:2)
                {
                        rho[i,s]~dbeta(1,1)
                        for(w in 1:omega)
                        {
                                xi[i,s,w]~dbern(rho[i,s])
                                for(k in 1:kappa[i])
                                {
                                        logit(mu[i,w,k,s])<-alpha+beta_sex[s]+beta_weather[1]*lot[i,k]+beta_occ[i]+
                                              beta_weather[2]*hit[i,k]+beta_weather[3]*ws[i,k]+beta_weather[4]*rh[i,k]+
                                              beta_weather[5]*ap[i,k]+beta_weather[6]*ml[i,k]
                                        theta[i,w,k,s]~dpois(mu[i,w,k,s]*xi[i,s,w])

                                }
                        }
                }
        }

        for(i in 1:(iota+delta))
        {
                for(s in 1:2)
                {
                        for(k in 1:kappa[i])
                        {
                omicron[i,s,k]~dsum(theta[i,1,k,s],theta[i,2,k,s],theta[i,3,k,s],theta[i,4,k,s],theta[i,5,k,s],theta
                      [i,6,k,s],theta[i,7,k,s],theta[i,8,k,s],theta[i,9,k,s],theta[i,10,k,s],theta[i,11,k,s],theta[i
                      ,12,k,s],theta[i,13,k,s],theta[i,14,k,s],theta[i,15,k,s],theta[i,16,k,s],theta[i,17,k,s],theta
                      [i,18,k,s],theta[i,19,k,s],theta[i,20,k,s],theta[i,21,k,s],theta[i,22,k,s],theta[i,23,k,s],
                      theta[i,24,k,s],theta[i,25,k,s],theta[i,26,k,s],theta[i,27,k,s],theta[i,28,k,s],theta[i,29,k,s
                      ],theta[i,30,k,s],theta[i,31,k,s],theta[i,32,k,s],theta[i,33,k,s],theta[i,34,k,s],theta[i,35,k
                      ,s],theta[i,36,k,s],theta[i,37,k,s],theta[i,38,k,s],theta[i,39,k,s],theta[i,40,k,s],theta[i
                      ,41,k,s],theta[i,42,k,s],theta[i,43,k,s],theta[i,44,k,s],theta[i,45,k,s],theta[i,46,k,s],theta
                      [i,47,k,s],theta[i,48,k,s],theta[i,49,k,s],theta[i,50,k,s],theta[i,51,k,s],theta[i,52,k,s],
                      theta[i,53,k,s],theta[i,54,k,s],theta[i,55,k,s],theta[i,56,k,s],theta[i,57,k,s],theta[i,58,k,s
                      ],theta[i,59,k,s],theta[i,60,k,s],theta[i,61,k,s],theta[i,62,k,s],theta[i,63,k,s],theta[i,64,k
                      ,s],theta[i,65,k,s],theta[i,66,k,s],theta[i,67,k,s],theta[i,68,k,s],theta[i,69,k,s],theta[i
                      ,70,k,s],theta[i,71,k,s],theta[i,72,k,s],theta[i,73,k,s],theta[i,74,k,s],theta[i,75,k,s],theta
                      [i,76,k,s],theta[i,77,k,s],theta[i,78,k,s],theta[i,79,k,s],theta[i,80,k,s],theta[i,81,k,s],
                      theta[i,82,k,s],theta[i,83,k,s],theta[i,84,k,s],theta[i,85,k,s],theta[i,86,k,s],theta[i,87,k,s
                      ],theta[i,88,k,s],theta[i,89,k,s],theta[i,90,k,s],theta[i,91,k,s],theta[i,92,k,s],theta[i,93,k
                      ,s],theta[i,94,k,s],theta[i,95,k,s],theta[i,96,k,s],theta[i,97,k,s],theta[i,98,k,s],theta[i
                      ,99,k,s],theta[i,100,k,s])
                        }
                }
        }

        for(i in 1:(iota+delta))
        {
                for(s in 1:2)
                {
                zeta[i,s]<-sum(xi[i,s,])+psi[i,s]
                }
        }
}
```

VII Glossary

Term	Definition
0 — full — amplitude	maximum amplitude allowed
440 Hz A	standard reference note for instrument tuning
ABAR	Animal–Borne Acoustic Recorder
AP	Atmospheric Pressure
ARU	Acoustic Recording Unit
ARW	Advanced Research Workshop
C Major arpeggio	cyclic sequence of notes including C, E, and G
C3	the third C on a standard piano
C6	the sixth C on a standard piano
dB	deciBel
dBA/dBC/dBZ	different deciBel weighting scales
DOC	Department of Conservation Te Papa Atawhai
DVAR	Detected Vocal Activity Rate
FLAC	Free Lossless Audio Codec
GCC	Generalised Cross–Correlation
IUCN	International Union for the Conservation of Nature
KCC	Kiwi Call Count
MCMC	Markov Chain Monte Carlo
NATO	North Atlantic Treaty Organization
NIBK	North Island Brown Kiwi
PAM	Passive Acoustic Monitoring
RH	Relative Humidity
ROSES	RepOrting standads for Systematic Evidence Syntheses
SECR	Spatially Explicit Capture-Recapture
SPL	Sound Pressure Level
T	Temperature
TDOA	Time Difference Of Arrival
VHF	Very High Frequency
WAV	Waveform audio file
WS	Wind Speed

Table VII.i: Glossary table.