WHISTLING IN THE DARK

An acoustic study of little spotted kiwi

ANDREW DIGBY

A thesis submitted to the Victoria University of Wellington in fulfilment of the requirements for the degree of Doctor of Philosophy.

Victoria University of Wellington

2013

Whistling in the dark: An acoustic study of little spotted kiwi. Ph.D. © Andrew Digby 2013

SUPERVISORS:
Dr. Ben D. Bell
Dr. Paul D. Teal

LOCATION:
Victoria University of Wellington, Wellington, New Zealand

GREY KIWI.
APTERYX OWENI.
(TWO-FIFTHS NATURAL SIZE)

To whistle in the dark:

1. To make a show of bravery despite one's fears; to put on a brave front.

2. To speak of something despite having little knowledge of it.

      – Wiktionary, http://en.wiktionary.org/wiki/whistle_in_the_dark

ACKNOWLEDGMENTS

This thesis was written using the `classicthesis` Latex template. Most plots were made using `ggplot2` in R. Artwork for the frontspiece and chapters 1–3 was provided courtesy of the Alexander Turnbull Library, Wellington, and the Museum of New Zealand Te Papa Tongarewa.

CONTENTS

ABSTRACT

Song function and evolution are two central topics of avian bioacoustic research. Discerning why birds sing and why they have such diversity of song can yield rich information on behaviour and speciation, and can provide important tools for conservation. Knowledge of vocal behaviour of a wide range of bird groups is necessary, yet avian bioacoustic studies have been hampered by a bias towards male song birds, with many species and even whole groups relatively unstudied.

The New Zealand kiwi (Apterygidae) are vocal and threatened taxa, and calls play an important part in their conservation. Yet kiwi acoustic behaviour is poorly understood, and although the five species differ significantly in their ecology, only one has been subject to detailed acoustic study. This thesis addresses this gap in knowledge of kiwi ecology with the first acoustic study of the little spotted kiwi (LSK). The principal aims are to improve understanding of kiwi calling behaviour, and to provide further acoustic tools for kiwi conservation. On a broader scale, enhanced knowledge of the acoustic ecology of the taxonomically and ecologically distinct kiwi will provide insight into song function and signal evolution in all birds.

Sexual call dimorphism in LSK is shown to be unrelated to size differences, and instead has likely functional significance, with male calls more suited for territorial defence. There is striking 'acoustic cooperation' between the sexes which constrains the function of duets in LSK. Analysis of complex vocal features, the first in any ratite, reveals that two-voicing is unexpectedly rare in this species, but that non-linear phenomena are common. Their association with territorial calls and high frequencies indicates that these features provide acoustic emphasis to enhance resource defence or convey aggression or fitness information. Non-linear phenomena are very common in nestling LSK calls, in accordance with the hypothesis that they add unpredictability to prevent habituation. LSK have surprisingly low inter-individual call variability, suggesting that this species may not use calls for individual identification. This lack of variability may be a result of the low genetic diversity in this species.

A long-term dataset reveals significant fluctuations in calling rates with temporal and environmental factors. These trends indicate that calls serve an important function for reproduction and pair contact, and that calls may reflect foraging activity.

They also provide evidence that kiwi are adversely affected by light pollution. A comparison of automated acoustic methods with manual call counts shows that while they have a different biases, autonomous techniques are highly effective for kiwi conservation monitoring. Microphone array methods have great potential for enhancing conservation and behavioural information through spatial monitoring of kiwi, but are demonstrated not to be suitable with currently available equipment.

INTRODUCTION

Credit: Gould, John, 1804-1881. Gould, John, 1804-1881: Apteryx owenii [Little spotted kiwi]. Gould. J. Gould delt, C. Hullmandel imp. [London, Gould, 1848]. Ref: PUBL-0026-6-03. Alexander Turnbull Library, Wellington, New Zealand. http://natlib.govt.nz/records/23172048. Permission of the Alexander Turnbull Library, Wellington, New Zealand, must be obtained before any reuse of this image.

Bioacoustics is the interface between biology and acoustics. It is the study of the production and reception of sound in animals. Dominant topics in this field are the behavioural significance of animal vocalisations, the evolution of these signals, and their use in monitoring species (Bradbury & Vehrencamp 1998). These are the underlying themes of this thesis, which analyses kiwi bioacoustics with two primary aims:

1. **Call function and evolution.** To investigate kiwi acoustic ecology to contribute to understanding of avian call functions and vocal evolution (§ 1.1).

2. **Conservation bioacoustics.** To develop the use of acoustic methods for kiwi conservation (§ 1.4).

1.1   BIRD SONG FUNCTION AND EVOLUTION

1.1.1   *The functions of bird songs and calls*

Bird song is ubiquitous, yet why birds sing is not clearly understood. Determining the functions of their songs and calls is a principal focus of avian bioacoustics research, and provides insight into many other areas of avian ecology. Calls and songs can be used to discriminate between species, sex and individuals (Catchpole & Slater 2008); to assess territory dynamics (Kirschel et al. 2011) and territory quality (Galeotti 1998); to infer the fitness, age, social status, reproductive success (Janicke et al. 2008, Ballentine 2009, de Kort et al. 2009) or even personality (Guillette & Sturdy 2011) of individuals. Such information, obtained passively, has important applications for research and conservation management. Understanding *why* these vocalisations are produced is necessary for interpreting and using this information.

1.1.1.1   *Calls or songs?*

There is no standard definition distinguishing bird songs and calls. A common distinction is that songs tend to be longer and more complex than calls (Langmore 1998). Another regular contrast is that calls are more spontaneous than song, often produced in a certain contexts such as for alarm and contact, whereas songs are used for territory defence and mate attraction (Marler 2004). Some authors further define songs as usually being produced by males during the breeding season, and calls by both sexes year-round (Catchpole & Slater 2008). However, such definitions are suited towards song birds in temperate regions, in which males are far more vocal and song output is seasonal (Langmore 1998). Indeed, songs have traditionally been associated with the song birds, or oscines, and not with other bird groups (Catchpole & Slater 2008). But some of the definitions of calls and songs suggest that non-oscines may be considered to sing too. Kiwi vocalisations, for example, fall into both categories: they can be classed as calls from their simple structure and production by both sexes, but as songs from their length and association with territory defence and reproduction. Determining whether kiwi sing or call, or both, is an underlying theme of this thesis. Throughout this manuscript kiwi vocalisations are referred to as 'calls', but this should not be interpreted as distinguishing 'call' from 'song'.

1.1.1.2 *Sexual selection and territory defence*

In songbirds, or oscines, bird song is mostly produced by males. Its primary role is widely believed to be for intra– and inter–sexual selection: to repel rivals and attract females (Collins 2004, Podos et al. 2004). This also applies to non-oscines, with calls reflecting male condition and reproductive success in species such as owls (Appleby & Redpath 1997a, Galeotti 1998, Hardouin et al. 2007), penguins (Marks et al. 2010) and seabirds (Genevois & Bretagnolle 1994, Janicke et al. 2008). Territory defence is a related function of male oscine song (Catchpole & Slater 2008), and is employed by both sexes in non-oscines (Fedy & Stutchbury 2005, Geberzahn et al. 2009). Vocalisations also contribute to sexual selection and reproduction in many bird taxa through duets, which provide functions of mate-guarding, signalling of commitment between mates, and ensuring reproductive synchrony (Benedict 2008, Hall 2009). Solo male songs can also be important for reproduction, having been shown to induce nest–building and ovarian development in female oscines (Brockway 1965, Bentley et al. 2000).

1.1.1.3 *Identification*

Bird calls are also central to identification at many levels: of species, sex, kin, mate, neighbours, parents, offspring, and individuals. Calls may be important in preventing hybridisation and inter-species competition, especially between sympatric and monomorphic species such as seabirds (Bolton 2007, McKown 2008, Petrusková et al. 2008, Curé et al. 2012). Similarly, many species exhibit neighbour recognition (the 'dear enemy effect') to avoid unnecessary competition (Speirs & Davis 1991, McGregor 1993, Clark et al. 2006, Catchpole & Slater 2008). Acoustic mate recognition is common in colonial nesting species such as penguins and seabirds (Speirs & Davis 1991, Clark et al. 2006, Curé et al. 2009). Vocal discrimination between parents and offspring reduces fitness costs to both (Knudsen & Evans 1986, Beecher 1991, Pierotti 1991, Ramos 2003). At a finer level, individually distinctive calls occur frequently in species with reduced reliance on visual cues, such as colonial nesters (Beecher et al. 1981, Stoddard & Beecher 1983, Robisson et al. 1993), nocturnal birds (Galeotti & Pavan 1991, Jones & Smith 1997, Otter 1996, Peake et al. 1998) and cryptic birds living in dense habitat (Gilbert et al. 1994, Kirschel et al. 2009). Whether these vocal differences are apparent to the birds themselves is less clear, although recognition of individuals from vocalisations has been demonstrated by playback experiments in some species (e. g. Galeotti & Pavan 1993, Aubin et al.

2000). Individual identification may not be necessary in some species, with classification such as mate or non-mate being sufficient. However, individual identification could provide survival benefits, such as in providing signaller reliability when assessing threat from alarm calls (Blumstein et al. 2004, Yorzinski et al. 2006, McDonald 2012).

1.1.1.4 *Other functions*

Social birds tend to have a more diverse vocal repertoire than solitary birds, suggesting that vocalisations play a role in social cohesion (Marler 2004). Passerines have the largest vocal repertoires, as a result of their ability to learn songs (Hultsch & Todt 2004, Kroodsma 2004, Catchpole & Slater 2008). Yet the functions of *songs* are largely restricted to territory defence and mate attraction, and even in oscines *calls* tend to have more uses. Many birds produce alarm calls, which occur in many different forms (mobbing calls, predator alert calls, distress calls), and have common acoustic characteristics between species (Marler 2004, Martin et al. 2011). Calls are also used to signal aggression (Morton 1977, Galeotti 1998, Marler 2004), maintain contact between members of a pair or flock (Marler 2004), and signal food availability to conspecifics (Wilson & Mennill 2010). Begging calls are employed to solicit food, most commonly from parents by offspring, but also by adult females in courtship behaviour (Lack 1953, Marler 2004).

Songs and calls serve a wide range of functions, and are the most important form of communication for many bird species. There is also huge diversity in the vocalisations between species; more so than in most other taxa. Yet in many species and even in whole bird groups, particularly in non-oscines, call functions remain unexplored, and even basic call descriptions do not exist for many species (Maurer et al. 2008).

1.1.2 *The evolution of bird song*

Understanding how avian communication became so diverse and complex is a central theme of bioacoustic research that is related to song function. Several hypotheses explain how acoustic signal divergence can occur in birds (Podos et al. 2004, ten Cate 2004):

1. *Acoustic adaptation:* sound transmission varies with habitat type (Morton 1975), so selection for efficient signalling can lead to acoustic differences between

closely-related species occupying different habitats (Ey & Fischer 2009, Tobias et al. 2010). A variant of this adaptation occurs in response to anthropogenic noise, with many species known to change song structure and frequency to facilitate more effective communication in noisy environments (Slabbekoorn & Peet 2003, Luther & Derryberry 2012).

2. *Morphological adaptation:* variation in habitat features such as altitude and food availability can also influence body size, which can produce song differences as a result of morphological constraints (Podos 2001, Derryberry et al. 2012)

3. *Species recognition:* calls may diverge due to selection for avoiding hybridisation. Hybrids, which often have lower fitness, frequently have intermediate vocalisations (den Hartog et al. 2007, Brown et al. 2010). Evidence for this effect is seen in associations between call structure and phylogeny (Farnsworth & Lovette 2008), and in songs of closely-related sympatric species being more divergent than allopatric species (Seddon & Sorenson 2005).

4. *Acoustic competition:* calls may diverge to avoid acoustic competition with conspecifics, or with other related or unrelated species (Doutrelant & Lambrechts 2001, Grant & Grant 2010).

5. *Sexual selection:* selection for certain song types or patterns due to inter- or intra-sexual selection is a key factor in song divergence (Searcy & Andersson 1986, Slabbekoorn & Smith 2002, Podos et al. 2004).

6. *Cultural transmission:* songs and calls may diverge as a result of copying errors and vocal matching in song-learning species (Baker et al. 2006, Parker et al. 2010).

Despite these divergent pressures, birds retain the general vocal characteristics of their ancestors (ten Cate 2004). As a result, vocalisations can be used to infer genetic information. Correlations between phylogeny and acoustic features have been found in several bird families (McCracken & Sheldon 1997, Seddon & Sorenson 2005, Farnsworth & Lovette 2008, Miller & Baker 2009). Calls have been used to help define new species (Kennedy et al. 2001), revise phylogenies (Whitney et al. 2000) or indicate hybridisation (Brown et al. 2010).

The extent to which acoustic divergence drives speciation is uncertain and somewhat controversial (Slabbekoorn & Smith 2002, Ritchie 2007, Catchpole & Slater 2008). This is largely due to the inability to observe such speciation occurring,

but is contributed to by a bias towards bioacoustic study of passerines, in which song-learning obscures evolutionary mechanisms (ten Cate 2004). Analysing vocal evolution in non-learners removes this complication, providing stronger evidence for the selective and phylogenetic influences on signal divergence (Seddon & Sorenson 2005, Seddon & Tobias 2007, Tobias et al. 2010).

1.1.3   *Understanding bird song: acoustic ecology*

Investigation into the function and evolution of bird song first requires knowledge of the calls and calling behaviour — the *acoustic ecology* — of individual species. Only then can vocalisations of different species be compared to elucidate patterns with other aspects of ecology. Yet a quantitative description of calls and song is lacking for many birds, particularly the non-oscines. This thesis forms an investigation of the acoustic ecology of one such bird, the kiwi (Apterygidae), with particular focus on the little spotted kiwi (*Apteryx owenii*).

1.2   LITTLE SPOTTED KIWI

1.2.1   *Basic ecology*

Little spotted kiwi (LSK) were common on the New Zealand mainland prior to European settlement (Heather & Robertson 2005), but declined rapidly during the 20th century, mostly due to predation (McLennan et al. 2004). The species is now thought to be extinct on the mainland (Heather & Robertson 2005), with the last population remnants found in the early 1980s (Jolly & Colbourne 1991, Shepherd et al. 2012). The entire species descends from at most five birds translocated to Kapiti island in 1912 (Colbourne & Robertson 1997, Ramstad et al. 2013), and the current population numbers around 1600 in eight island populations (Figure 1.1; Holzapfel et al. 2008, H. Robertson, per. comm.). It is currently the second rarest kiwi, classified as near threatened by the IUCN (IUCN 2011) or recovering (conservation dependent, range restricted) in the New Zealand Threat Classification System (Miskelly et al. 2008).

The LSK is the smallest of the five recognised species (Burbidge et al. 2003), with mean weights of 1350 g and 1125 g for females and males respectively (Heather & Robertson 2005). It is strictly nocturnal, and thought to share the unusual sensory
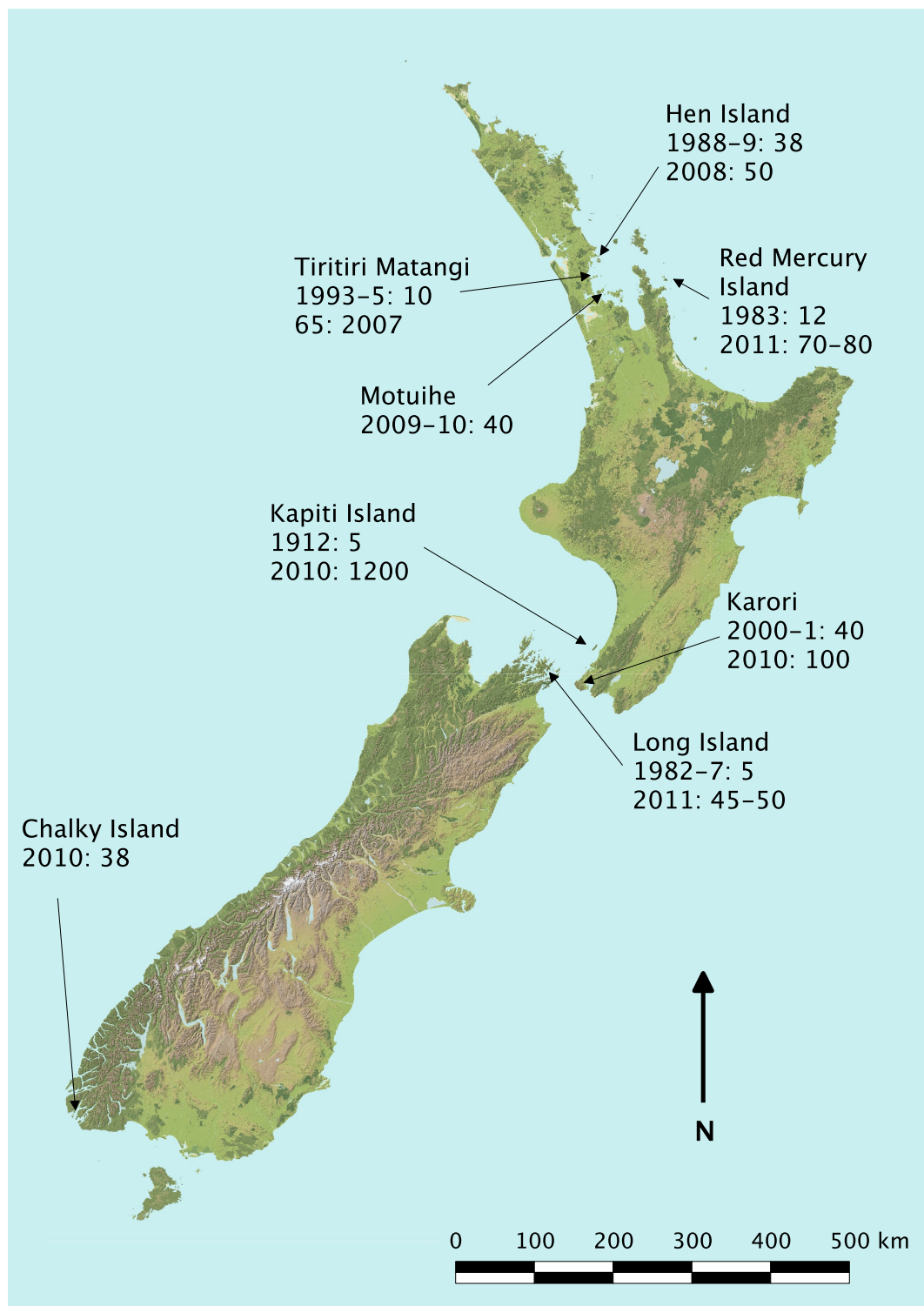
Figure 1.1: Little spotted kiwi populations in New Zealand. All are translocated populations: first dates and number provide the dates and number of kiwi in the initial translocation. Additional dates and numbers give latest population estimates (Holzapfel et al. 2008, H. Robertson, New Zealand Department of Conservation, pers. comm., 2011). Base map credit: www.Geographx.co.nz.

repertoire of brown kiwi (*Apteryx mantelli*; formerly North Island brown kiwi), with poor vision (Martin et al. 2007, Corfield 2009), a highly-developed olfactory system (Bang 1971, Corfield 2009, Cunningham et al. 2009, Castro et al. 2010), acute hearing most sensitive to high frequencies (Corfield 2009, Corfield et al. 2011), and unusual tactile ability ('remote touch') through a bill-tip organ (Cunningham et al. 2007, 2009).

LSK are highly territorial, with a territory size of 2–3 ha on Kapiti Island (Colbourne & Robertson 1997, Holzapfel et al. 2008). However territoriality and territory size may vary between populations, as in brown kiwi (*Apteryx mantelli*) (McLennan et al. 1987, Taborsky & Taborsky 1992). Like brown kiwi (Colbourne & Kleinpaste 1983, McLennan 1988, Taborsky & Taborsky 1999), LSK are thought to be largely monogamous. However, this has not been assessed, and extra-pair paternity and polygamy are known or likely in some populations of brown kiwi (Ziesemann et al. 2011) and tokoeka (*Apteryx australis*; Falla 1979 in Colbourne 1991, Taborsky & Taborsky 1999). LSK are thought to lay 1–2 eggs per year (Jolly 1989), although recent research shows that some individuals may have a higher reproductive output (H. Taylor, Victoria University of Wellington, pers. comm., 2013). Incubation takes 65–75 days (Jolly 1989, Heather & Robertson 2005) and is solely performed by males, with only one known exception of shared incubation (Potts 1872, in Colbourne 2002). The young are tolerated in the adult territory for about a year before dispersal (Colbourne 1992). LSK are solitary, and do not form social groups, unlike tokoeka, rowi (*Apteryx rowi*) and great spotted kiwi (*Apteryx haastii*; Colbourne 1991, 1992).

1.2.2   *Acoustic behaviour*

Kiwi calling behaviour and call functions are relatively poorly understood. Although several qualitative accounts exist (e. g. Buller 1888, Clark 1952), much of the knowledge of calling behaviour comes from call surveys of just one species, the brown kiwi (Colbourne & Kleinpaste 1984, Miles et al. 1997, Colbourne 2006). Quantitative spectral analyses have only been undertaken in two studies, which investigated spectral composition, duetting behaviour, seasonal variation, individual variation and dialects in calls of two brown kiwi populations (Corfield 2005, Hojem 2006, Corfield et al. 2008a). No spectral analyses have been published for the other four kiwi species.

Censuses reveal that while both sexes call, males do so much more frequently in all species except for great spotted kiwi (Colbourne 2006). Call rates typically peak at dusk and dawn (Colbourne & Kleinpaste 1984, Miles et al. 1997, Heather & Robertson 2005), but there is significant nightly variation (Colbourne & Kleinpaste 1984). Weather appears to impact calling behaviour, but there is conflicting evidence for this (Miles & McLennan 1997, Pierce & Westbrooke 2003), and weather also affects detection distance (Pierce & Westbrooke 2003). Similarly, there are significant variations of the effect of light levels on call rates: some studies report a reduction with moon phase (Buller 1888, Clark 1952, Colbourne & Kleinpaste 1983, 1984), whereas others suggest no effect (Miles et al. 1997, Pierce & Westbrooke 2003). There are substantial seasonal variations in calling activity, with peaks before and after the main breeding period in brown kiwi (Colbourne & Kleinpaste 1984, Miles et al. 1997, Scrimgeour & Pickett 2011). However, seasonal differences are less pronounced in great spotted kiwi and Stewart Island tokoeka (Colbourne 2006).

Spatial as well as temporal variation exists in brown kiwi, with dialects found in two Northland populations (Hojem 2006). No other quantitative dialect studies exist, although great spotted kiwi show no strong geographic call variation (McLennan & McCann 2002).

A large proportion of brown kiwi calls occur as part of a duet, particularly for females (Corfield 2005). Duet function has not been investigated in kiwi, but as for other species is likely to serve multiple roles (Hall 2009). Duets occur most often at the start of the night when pairs are separated, suggesting pair contact is a function of duets (Colbourne & Kleinpaste 1984, Hojem 2006). Most duets take place between mates (Taborsky & Taborsky 1992, Miles & McLennan 1997, Corfield et al. 2008a), can occur when pairs are close together (Colbourne & Kleinpaste 1984), and are more frequent in the breeding season (Taborsky & Taborsky 1992, Corfield 2005, Hojem 2006), suggesting that pair bonding and reproductive synchrony are also drivers. There is evidence that duets contribute to joint territory defence, since kiwi often duet in response to nearby conspecifics (Colbourne & Kleinpaste 1984, Corfield 2005), when being chased (Colbourne & Kleinpaste 1984), and following simulated territorial intrusion (Colbourne & Kleinpaste 1984, Miller & Pierce 1995, Miles et al. 1997). However, some of these responses may be for mate-guarding, which is further supported by inter-sexual difference in contribution to duets (Corfield 2005, Corfield et al. 2008a) and in sex-specific responses to playback (Miles et al. 1997).

In common with other nocturnal species (e. g. May 1994, Galeotti & Pavan 1991), spectral analyses have shown that brown kiwi have individually distinctive calls (Corfield 2005, Hojem 2006). This is supported by anecdotal evidence that brown kiwi often do not respond to a neighbour's call (McLennan et al. 1987, Taborsky & Taborsky 1991, 1992), and respond more frequently to a strange call during playback (Miles et al. 1997). Female brown kiwi calls contain formants (Corfield et al. 2008a), and males may contain signatures of two-voicing[1] (Corfield 2005); both information-bearing elements that have been linked with individual identification in other species (Suthers 1994, Aubin et al. 2000).

Although kiwi calls have been used extensively for population censuses over several decades (Colbourne 2006), understanding of their call functions is still limited. Much of the knowledge of kiwi calls presented here comes from qualitative or circumstantial evidence, and nearly all from a single species, the brown kiwi. Calls of the other four species remain unstudied, and there have been no comparisons between the species, despite their substantial differences in other aspects of their ecology (Heather & Robertson 2005, Castro 2011). However, beyond improving knowledge of kiwi ecology, what is the wider relevance of understanding kiwi call function?

1.3   WHY STUDY KIWI ACOUSTICS?

Kiwi are ecologically and taxonomically very different from other birds, yet are acoustically similar, sharing common vocal features such as duets and use of calls for territory defence. Kiwi have very different ecology from their palaeognath[2] relatives such as emu and cassowary (Taylor et al. 2000, Mack & Jones 2003), and the kiwi species differ in their physiology and ecology (Heather & Robertson 2005, Castro 2011). As non-oscines, they are expected not to learn their songs (Jarvis 2004, Kroodsma 2005), and as ground-dwellers, analyses are simplified by uniformity of calling height (Seddon & Sorenson 2005). These features make kiwi subject species that are well-suited for exploring theories of avian vocal evolution.

The expected lack of song learning also considerably aids interpretation of call function; for example, vocal plasticity in songbirds can mask the functional significance of dialects or individually distinctive vocalisations (Kroodsma 2004, Seddon &

---

1 See § 3.1 for a definition

2 Palaeognathae ("old jaws") are the clade of birds containing the ratites and tinamous, and are distinct from the Neognathae ("new jaws"), the clade containing all other extant bird species.

Sorenson 2005, Fox 2008a). Yet bioacoustic studies of non-oscines are in the minority, and very few palaeognaths have been acoustically investigated (Beaver 1978, Mack & Jones 2003, Corfield et al. 2008a). Avian bioacoustic studies are also heavily biased towards birds from temperate regions (Slater & Mann 2004). Although kiwi inhabit a temperate climate, they have 'tropical' characteristics to their behavioural and acoustic ecology, such as year-round territory defence, duetting, and frequent female song (Colbourne & Kleinpaste 1984, Corfield 2005). Female vocalisations have typically received little attention because most females in temperate zones do not sing, but female calls and inter-sexual vocal differences are central to understanding acoustic behaviour (Langmore 1998, Collins 2004, Geberzahn et al. 2009).

The kiwi's unique ecological niche also merits attention with respect to bioacoustics. Kiwi are a rare example of a ground-dwelling, nocturnal bird. They are part of a nocturnal guild of communicators such as owls, seabirds, insects and frogs that rely heavily on sound for communication. Yet they have an unique sensory repertoire, and with poor vision (Martin et al. 2007) are more reliant on acoustic cues than other nocturnal birds, which usually have excellent vision (Martin et al. 2004, Corfield 2009). Furthermore, kiwi are habitat generalists, occupying a wide range of habitats both within and across species (Heather & Robertson 2005, Castro 2011). These also can have accompanying physiological differences: for example, upland and lowland great spotted kiwi have size differences (McLennan & McCann 2002).

Although kiwi are ideal taxa for investigating avian call function and vocal evolution, they have received relatively little bioacoustic attention. This thesis intends to redress this with the first study of acoustic ecology of the little spotted kiwi. By advancing understanding of the acoustic behaviour of this species, the aim is to establish a basis for further comparative studies within kiwi and among palaeognaths, and to contribute to the wider themes of avian call function and evolution. To facilitate such comparisons, this thesis provides a description of LSK calls, and addresses a series of questions regarding their vocal behaviour (Box 1.1).

1.4   CONSERVATION BIOACOUSTICS

1.4.1   *Benefits of bioacoustic methods*

Throughout the history of conservation, sound has been used to measure species abundance (Kendeigh 1944, Dawson & Bull 1975). Acoustic methods offer a number

<div style="border:1px solid #000; padding:1em;">

LSK acoustic behaviour questions

1. *Why do LSK have sexually dimorphic calls?*

2. *Do LSK calls contain information-bearing spectral features such as two-voicing* and non-linear phenomena*, and why?*

3. *Do LSK have individually distinctive calls?*

4. *How does LSK calling activity vary temporally, and how can this inform call function theories?*

</div>

Box 1.1: Questions addressed by this thesis regarding acoustic behaviour of LSK. (*See § 3.1 for a definition of these terms.)

of advantages over other monitoring techniques. They can provide more ecologically important information than visual survey methods (Kirschel et al. 2011, Mennill 2011), and are also very efficient, yielding inexpensive and rapid assessment of biodiversity and species abundance (Parker 1991, Riede 1993, 1998, Blumstein et al. 2011). Passive and non-invasive, acoustic methods are less likely to affect animal behaviour, making them ideal for sensitive and easily-disturbed species (Bridges & Dorcas 2000, Mennill 2011). Similarly, they are well-suited to animals that are difficult to observe visually. These include cryptic or sparsely-distributed animals such as bitterns (Gilbert et al. 1994, Eriksen 1999, Wahlberg & Møhl 2003) or elephants (Payne & Thompson 2003), nocturnal-calling species such as corncrake (May 1998, Peake & McGregor 2001) and owls (Redpath et al. 2000, Nagy & Rockwell 2012), and those occupying dense (Jiménez et al. 2003, Mennill & Vehrencamp 2008b) or inaccessible habitat (Hutto & Stutzman 2009). Using autonomous acoustic recorders also avoids temporal sampling bias that can arise from human-based surveys (Peterson & Dorcas 1994, Bridges & Dorcas 2000), and provides a permanent data record. This reduces survey cost by minimising time spent by experts identifying species, allows for the use of automatic classification methods, and increases accuracy by enabling re-analysis (Hobson et al. 2002, Rempel et al. 2005, Hutto & Stutzman 2009, Swiston & Mennill 2009).

1.4.2 *Examples of conservation bioacoustics*

With many benefits of bioacoustic monitoring, its applications to conservation are diverse (Gaunt & McCallum 2004). Acoustic methods can be used to assess ecosystem health and biodiversity (Riede 1993, 1998), or to measure the impacts of anthropogenic noise (Blumstein et al. 2011, Chan & Blumstein 2012) and other disturbances (Farnsworth & Russell 2007, Hüppop et al. 2006, Hüppop & Hilgerloh 2012). Acoustic monitoring is frequently used to determine species presence or absence (Swiston & Mennill 2009), abundance (Alldredge et al. 2007, Diefenbach et al. 2007), population trends (Miller & Pierce 1995, Pierce & Westbrooke 2003), variation in temporal patterns (Bridges & Dorcas 2000) and migration (Farnsworth & Russell 2007). Call surveys are also used in conservation to assess efficacy and impacts of management methods such as predator control (Robertson & Colbourne 2001, Robertson et al. 1999, Pierce & Westbrooke 2003).

Geographical variations in calls can provide evidence for founder effects, reproductive isolation and bottlenecks (Baker & Jenkins 1987, Baker et al. 2006, Seddon & Tobias 2007), which can have strong implications for the translocation and management of anthropogenically fragmented species (Rowe & Bell 2007, Parker et al. 2010, Robin et al. 2011). Acoustic methods are also important in assessing taxonomic status of morphologically cryptic species (Bolton 2007, Bretagnolle & Genevois 1997), and have been used to detect and define species (Isler et al. 1997, Robbins & Stiles 1999, Gaunt & McCallum 2004). Calls are also used to determine sex in monomorphic species such as ducks and seabirds (Volodin et al. 2005, 2009, Bolton 2007, Bourgeois et al. 2007). Identifying individuals from calls is possible in many bird species and has many important applications in conservation (Terry et al. 2005), such as measuring sampling bias and improving estimates of population size (Saunders & Wooller 1988, Jones & Smith 1997, Eriksen 1999, Peake & McGregor 2001, Terry & McGregor 2002), and monitoring individual survival (Galeotti & Sacchi 2001, Gilbert et al. 2002). Spatial information from microphone arrays (Blumstein et al. 2011) can be used to monitor individuals (Payne & Thompson 2003) or measure territory use and dynamics (Mennill et al. 2006, Fitzsimmons et al. 2008c,b, Koloff & Mennill 2011, Kirschel et al. 2011, Mennill 2011).

1.4.3 *Kiwi conservation bioacoustics*

Bioacoustic methods make an important contribution to kiwi conservation through call count surveys. These were formalised with the introduction of a New Zealand

Kiwi conservation acoustics questions

1. *Are autonomous recorders and automatic detection methods suitable for call point counts?*

2. *How can temporal variability in kiwi calls assessed by autonomous recorders inform monitoring programmes?*

3. *Are microphone arrays suitable for monitoring kiwi?*

Box 1.2: Questions assessing suitability of acoustic methods for kiwi conservation.

monitoring scheme in the 1980s (McLennan 1992, Robertson & Colbourne 2003). Results are used to determine kiwi distribution (Miller & Pierce 1995), relate density to location and habitat (Colbourne 2006), and assess population trends resulting from conservation management practices such as predator control and advocacy (Pierce & Westbrooke 2003). Playback of calls also is a key tool in catching kiwi for research and conservation management (Robertson & Colbourne 2003). However, acoustic methods are under-utilised in kiwi conservation. Beyond call counts, most of which are performed with human observers, there is little widespread use of other bioacoustic methods, despite their suitability to kiwi. This thesis aims to improve this situation by assessing further methods of acoustic monitoring for kiwi conservation (Box 1.2).

1.5   THESIS STRUCTURE AND PUBLICATIONS

This thesis uses analysis of LSK bioacoustics to address eight main hypotheses and questions regarding kiwi call function and acoustic conservation methods. These contribute towards the wider topics of the function and evolution of avian vocalisations (Figure 1.2).

The thesis is written as a collection of six independent papers, five of which have either been submitted for publication prior to completion of the thesis, or will be soon afterwards (Table 1.1). I am the lead author on each of these papers, with my supervisors Dr Ben Bell and Dr Paul Teal as co-authors. Dr Michael Towsey, of the University of Queensland, Brisbane, Australia, is a co-author for the papers from Chapters 5 & 6. Dr Towsey provided the call detection software in these chapters,
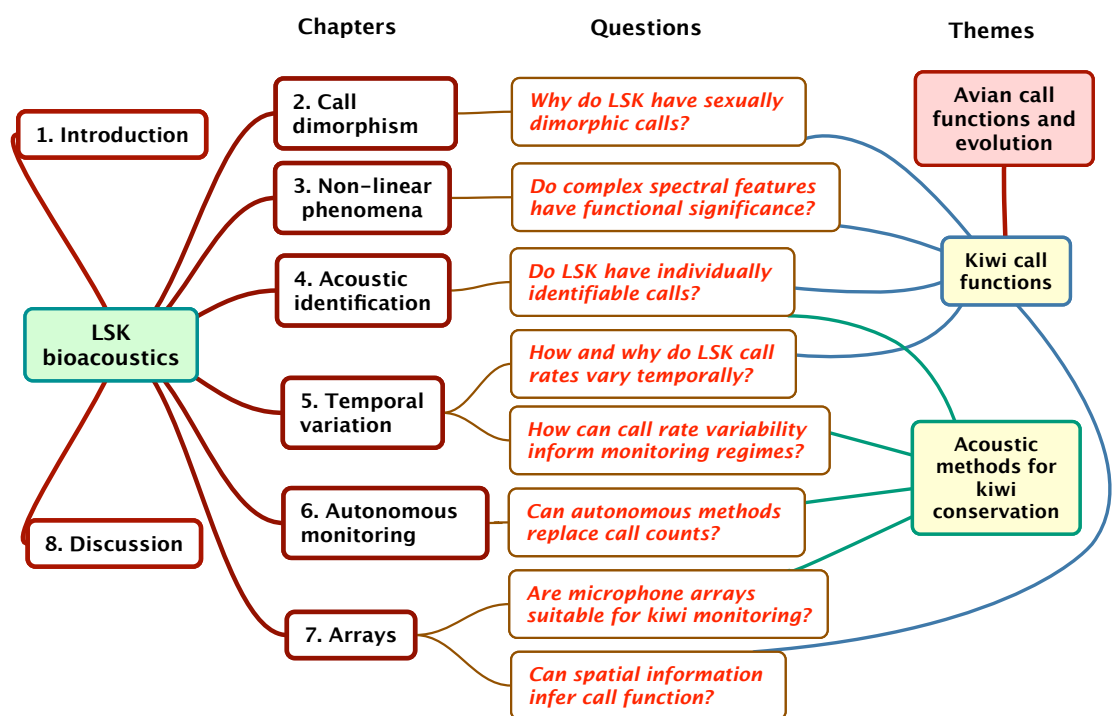
Figure 1.2: The questions and themes addressed by this thesis.

which was then adapted to little spotted kiwi. Chapters submitted for publication are reproduced in published or submitted form in the style of the journal, but with minor modifications to the formatting, and with added cross-references between chapters. As a collection of papers, there is unavoidably some repetition between chapters, particularly in the introductions. Any results which could not be included in manuscripts due to constraints of time or space are included in appendices at the end of the relevant chapter. Scientific names of species are provided upon the first use in each chapter.

Table 1.1: Manuscripts submitted or in preparation for submission from material in this thesis.

Chapter	Title	Journal	Status
2	*Vocal cooperation between the sexes in little spotted kiwi*	Ibis	Published, Apr 2013
3	*The behavioural significance of non-linear phenomena in kiwi calls*	Bioacoustics	Submitted, Nov 2012
4	*Vocal individuality in little spotted kiwi calls*	To be decided	In prep.
5	*Temporal variation in little spotted kiwi calls*	To be decided	In prep.
6	*A practical comparison of manual and autonomous methods for acoustic monitoring*	Methods in Ecology & Evolution	Published, May 2013

VOCAL COOPERATION BETWEEN THE SEXES IN LITTLE SPOTTED KIWI[1]

Little Spotted Kiwi by Keulemans, J. G., 1872, New Zealand. Reproduced courtesy of Museum of New Zealand Te Papa Tongarewa. (O.009829).

ABSTRACT

Sexual call dimorphism in birds is usually associated with sexual size dimorphism. Departures from this relationship can be used to infer call function, but research into inter-sexual call differences, as with song function in general, has been restricted by a bias towards male passerines. The nocturnal and flightless New Zealand kiwi

---

(family: Apterygidae) are acoustically similar but taxonomically and ecologically very different from other birds, so provide a contrast in exploring avian call function and evolution. Yet kiwi acoustic ecology is poorly understood, with the calls of only one of the five kiwi species spectrally described, and acoustic differences between the sexes virtually unexplored. We have conducted the first bioacoustic study of the little spotted kiwi *Apteryx owenii*, and assessed sexual call dimorphism in this species. We found significant inter-sexual differences in call temporal and frequency characteristics that were not related to size dimorphism. Contribution to duets and variation in temporal structure with call context also differed between the sexes. We suggest that these differences indicate divergent call function, with male calls more suited for territory defence, and female calls for pair contact. There was a striking lack of overlap in the frequency spectrum distributions of male and female calls, which was also unrelated to size and was further emphasised by the presence of formants in female calls. We propose that this provides evidence for inter-sexual acoustic cooperation in call frequency, of a type which to our knowledge has not previously been described in birds. This may result from selection for enhanced joint resource defence in these kiwi.
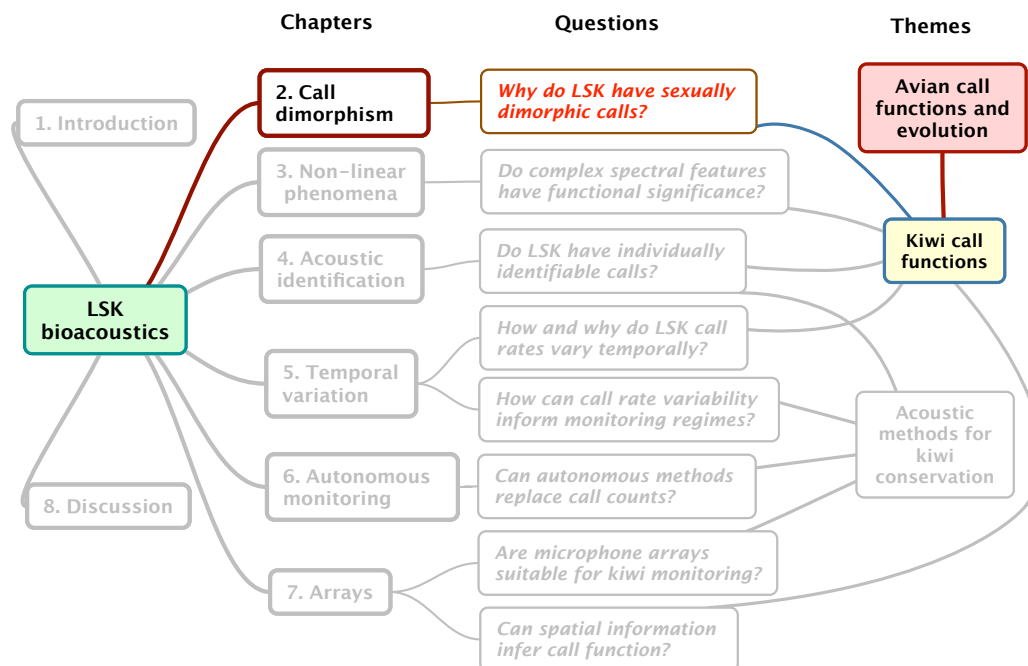
Topics covered by this chapter.

2.1 INTRODUCTION

An inverse relationship between vocalisation frequency and body size is found in many taxa (Ryan & Brenowitz 1985), including diverse bird groups (e.g. Bertelli & Tubaro 2002, Martin et al. 2011). This is because larger individuals usually have larger vocal structures, which produce sounds with a lower fundamental frequency (Wallschläger 1980, Ryan & Brenowitz 1985, Seddon & Sorenson 2005). In sexually dimorphic bird species, the larger sex usually has lower frequency calls (Maurer et al. 2008, Geberzahn et al. 2009). However, there are exceptions. In some sexually dimorphic species the larger sex has higher-pitched calls (Herting & Belthoff 2001), whereas in others there are no inter-sexual call differences (Janicke et al. 2007). In contrast, some sexually monomorphic species have heterogeneous calls (Bourgeois et al. 2007, Volodin et al. 2009). Furthermore, the slope of the size-frequency regression varies with avian sub-taxa (oscines, sub-oscines and non-passerines; Ryan & Brenowitz 1985). Call frequency can also be influenced by other factors, such as habitat type (Bertelli & Tubaro 2002, Nicholls & Goldizen 2006), ambient noise (Slabbekoorn & Peet 2003, Slabbekoorn & den Boer-Visser 2006, Luther & Derryberry 2012), bill morphology (Grant & Grant 2010) or even the context of calls (Geberzahn et al. 2009).

Departures from the expected frequency-size relation make sexual call dimorphism an important mechanism for determining call function. Vocal similarities between the sexes can imply convergence in call function (Wright & Dahlin 2007, Kirschel et al. 2011). Thus differences can indicate functional divergence. However, a bias towards male song has hampered research into song function, with female song and inter-sexual call dimorphism unstudied in the majority of bird species (Collins 2004, Marler 2004). Investigating call differences between the sexes in a wider range of species, particularly those with unusual ecology or taxonomy, is likely to provide valuable information on avian call function.

The New Zealand endemic kiwi (Apterygidae) are ideal taxa for investigating avian call function. All five species (Burbidge et al. 2003) are vocal and sexually dimorphic in size and call, with females being on average larger and having lower frequency calls (Heather & Robertson 2005). Studies of brown kiwi (*Apteryx mantelli*) have shown that it is acoustically similar to other birds, exhibiting features such as dialects (Hojem 2006) and duets (Corfield 2005), and has likely call functions of territory defence, mate attraction and pair contact (Colbourne & Kleinpaste 1984, Corfield et al. 2008a).

While vocally comparable, kiwi are ecologically unique among birds. They are nocturnal and flightless, and have an unusual sensory repertoire: poor vision (Martin et al. 2007), a highly-developed olfactory capability (Cunningham et al. 2009, Castro et al. 2010), high-frequency sensitivity (Corfield et al. 2011) and vibrotactile ability from a bill-tip organ (Cunningham et al. 2007). These ecological differences and common acoustic characteristics make kiwi a useful counterpoint to the passerine and suboscine species upon which the vast majority of bioacoustic studies have focussed (Catchpole & Slater 2008). The five species of kiwi also vary in their morphology and behaviour (Castro 2011). Comparing kiwi calling behaviour between the sexes and among the five species, as well as with other birds, can therefore potentially offer powerful new insights into avian call function.

As palaeognaths, kiwi are also taxonomically distinct, and are not expected to learn their calls (Jarvis 2004). Acoustic contrasts among the kiwi species and other palaeognaths could thus provide important tests of avian signal evolution. Improving understanding of kiwi calling behaviour is also important for their conservation. Calls are the main census tool used for kiwi (Colbourne 2006, Holzapfel et al. 2008), with all five species threatened or near-threatened (IUCN 2011).

Despite these motivations, only one kiwi species, the brown kiwi, has been subject to detailed acoustic study. The assumption that call dimorphism is a result of size dimorphism (Corfield et al. 2008a) is untested in kiwi. Formants — amplification of certain frequencies due to vocal tract filtering — are evident in female brown kiwi calls, and have been postulated to serve as an information-bearing element (Corfield 2005, Corfield et al. 2008a). However, the function of such spectral features has not been investigated, and the spectral characteristics of other kiwi species have not been examined. Duetting has been quantified in brown kiwi (Corfield 2005), but not in other kiwi species. Whether kiwi duets serve similar functions as in other birds, such as territory defence, pair contact and pair bonding (Benedict 2008, Hall 2009), remains unclear. Kiwi calling behaviour, particularly female call function and intersexual call differences, is therefore largely unexplored. Improved understanding of kiwi acoustic ecology will aid in their conservation and contribute to developing theories of avian call function and evolution.

2.1.1   *This study*

In this paper we advance knowledge of kiwi calling behaviour with the first acoustic study of the little spotted kiwi (*Apteryx owenii; LSK*), the smallest and second

most rare of the kiwi species (Heather & Robertson 2005, Holzapfel et al. 2008). The specific aims of this study were to (1) quantitatively describe and compare the calls of each sex, (2) use morphological measures to determine whether inter-sexual call dimorphism is a result of size dimorphism, and (3) use these comparisons to investigate possible inter-sexual differences in call function in LSK. Such a combined acoustic and morphometric analysis has not previously been undertaken in any kiwi species, and will improve understanding of their acoustic ecology. This will assist in kiwi conservation and contribute to developing theories of call function in other bird species, particularly those with sexual dimorphism. Furthermore, the first spectral description of this palaeognath will provide a new reference point for assessing call divergence in birds.

2.2 METHODS

2.2.1 *Study site*

Calls of LSK were recorded in Zealandia (formerly Karori Sanctuary), a 225 ha fenced reserve free of mammalian predators in Wellington, New Zealand. The reserve encompasses a valley with steep-sided gullies and primarily native lowland podocarp forest. Forty LSK were introduced to the reserve in 2000–2001 from Kapiti Island, 50 km to the north, and by 2010 the population had grown to an estimated 100 individuals (H. Robertson, New Zealand Department of Conservation, pers. comm., 2010).

2.2.2 *Call recordings*

Calls were recorded at 48 kHz, 24-bit precision, with hand-held equipment (Sound Devices 722 with a Telinga Twin Science microphone and a Telinga 53 cm foldable parabola), on 140 nights from May 2009 to April 2012. Calls were sampled in the first three hours after civil twilight on nights with no rain and low wind noise. Recordings were made from high points or close to known kiwi. Prior to April 2010, all calls were recorded serendipitously, but from April–August 2010, 52 kiwi in the reserve were fitted with radio transmitters for a population census. The transmitters were removed from many of these kiwi in late 2010, but a further 11 kiwi were tagged in 2011 for a separate study, and transmitters were retained on 15 males until April 2012. The transmitters facilitated the tracking of individuals and independent

identification of calling birds, and enabled calls to be recorded at close range. Birds were also fitted with leg bands with reflective colour tape, so that they could be identified at night in the absence of transmitter information. After April 2010, playback was also used to elicit calls from individuals. To avoid habituation and undue disturbance, this was used less than once per week for each individual, and outside of the September–January breeding season.

2.2.3 *Call samples*

A consequence of recording wild kiwi is that the distance to calls will be non-uniform. Call distance also does not reliably infer call attenuation or degradation. For example, a kiwi calling across a valley could be recorded with fewer intermediate obstacles than one at a much closer distance on flat terrain. So call spectrographic quality was used as a selection criterion rather than distance. Estimating distances to a sound source is difficult (Naguib & Wiley 2001), especially at night. However, confidence in distance measures was provided by the close range of most calls, transmitter signals and knowledge of site topography.

A total of 775 LSK calls were recorded. Each was assigned a subjective five-level quality score, based on amplitude and interference from other sound sources. Only the 212 calls with a quality score in the highest two levels (1–2) were used in analyses (Table 2.1). Each LSK call comprised a series of repeated syllables. Some calls were recorded at very close range, and so contained clipped syllables that saturated the recorder. Since they were of very high signal-to-noise, these calls were retained to assess the presence of spectral features and for duet analysis, but with the clipped syllables removed (Sample A, Table 2.1). However, the entire calls were discarded from the quantitative temporal and frequency analyses (Sample B). Long pauses were also removed from this sample by discarding the first syllable occurring after a gap of more than 2 seconds.

For morphometric-call analyses, calls were only used if kiwi identity could be verified from telemetry or visual confirmation of coloured leg bands. Only adults were included, defined as those translocated to Zealandia in 2000–2001, or adult-sized ($\geq 1\,$kg) and found sharing a burrow with another known adult (Robertson & Colbourne 2003). Calls were trimmed to avoid sudden frequency or temporal variation due to priming at the call start (Jones & Smith 1997), and similar deviations at the call end. Inspection of all calls showed that these fluctuations only occurred before the 15th syllable and after the 35th syllable in each call, so only syllables within

this range were retained. Calls containing fewer than ten syllables in this region were discarded, to ensure sufficient sample size for calculating syllable means. This resulted in 81 calls from 15 individuals (nine male, six female; Sample C).

For comparing power spectra between the sexes, the highest quality calls with no clipping were used, without the identification requirement (Sample D).

Table 2.1: The call samples used for analyses. $^*f_0$ = fundamental frequency.

	Calls			Syllables			
Sample	M	F	Total	M	F	Total	Use
A: Quality=1-2	158	54	212	4539	1545	6084	Spectral features and duets
B: A + no clip or pauses	148	51	199	4066	1386	5452	Comparison of f_0^* distributions
C: B + ID + $\geq$ 25 syllables	66	15	81	1176	263	1439	Morphology-call models
D: Quality=1 & no clip	93	29	122	2756	771	3527	Comparison of power spectra

2.2.4 *Call analyses*

Call and syllable parameters were measured from spectrograms using the software Raven Pro 1.4 (Charif et al. 2010), with: 1024-sample Hann window, 50% overlap, and 23.4 Hz frequency resolution. The temporal and frequency extents of the fundamental frequency (f_0) of each syllable were determined from the energy distribution, which provides robustness against variations in background noise (Cortopassi 2006). These measures were made at the peak energy and at the 5th, 50th (central), and 95th percentiles of the energy distribution. The central 90% of energy was used to measure syllable length and bandwidth. Call duration was calculated as the sum of the syllable lengths and gaps. Syllable rate was defined as the call duration divided by the number of syllables, and call duty cycle as the sum of the syllable lengths divided by the call duration. The waveform of each call was also checked for detector saturation.

Power spectra were measured in Adobe Audition 4.0 using a 1024-sample Hann window. The spectrum was measured across the whole call, except for initial syl-

lables showing large frequency variation due to priming. A separate background noise spectrum was measured from a silent period adjacent to the call. Spectra were then background-subtracted, and normalised to unit area under the curve.

All statistical analyses were conducted in R (2.1.15; R Development Core Team 2011). Call parameter distributions were asserted to be significantly different between the sexes if the 95% confidence intervals of the medians did not overlap.

2.2.5 *Morphology measures*

Capture must be minimised in wild kiwi as it causes stress and can disrupt breeding (Robertson & Colbourne 2003). Morphometric measures were therefore taken only when birds were captured for other purposes. This was unavoidably infrequent, often several weeks or even months apart from call recordings. The standard kiwi measurements of weight, bill length and tarsus length, tarsus width and tarsus depth were taken (Robertson & Colbourne 2003). Tarsus width and depth are the minimum lateral and minimum longitudinal dimensions of the tarsometatarsus, respectively. For many recaptures only weight and bill-length were measured. Birds were sexed according to bill length, which is sexually dimorphic in adult LSK (M$\leq$ 75 mm; F$>$ 75 mm; Colbourne & Robertson 1997).

Variation in morphological measures was assessed by a coefficient of variation ratio, calculated for the nine males and six females in call sample C (Table 2.1). This ratio was defined as the coefficient of variation of all M measures from K kiwi ($CV_a = \sigma_M/\mu_M$) divided by the mean coefficient of variation within individuals, calculated for cases where each individual had more than one measure ($\overline{CV_w}$; Equation 2.1). The ratio was calculated separately for each variable and each sex. Ratio values greater than one indicated that variation was greater among than within individuals, and that the metric provided some discrimination among individuals (e.g. Fitzsimmons et al. 2008a).

$$CV_r = \frac{CV_a}{CV_w} = \frac{\sigma_M/\mu_M}{\frac{1}{K}\sum_{j=1}^{K}\sigma_j/\mu_j} \tag{2.1}$$

Bill length and tarsus dimensions undergo slow change in adult kiwi (Robertson & Colbourne 2003). This was supported by our repeated measures, which also showed no coherence between the direction of changes in these metrics and changes in weight. Any variation was therefore assumed to be due to measurement error, and so the means of tarsus dimensions and bill length were used for each kiwi. Since

the variation within individuals was low compared to that among kiwi (Table 2.2), means would retain size difference information.

Table 2.2: Number of repeated measures (n) and the coefficient of variation ratio (CV_r) for morphological parameters of the nine males and six females in the morphology-call sample. Coefficient of variation ratios greater than one indicate that variation among individuals was greater than variation within, so the metric provided some discrimination among individuals in that sex.

	Male		Female	
Parameter	n	CV_r	n	CV_r
Bill length	1-6	2.3	1-4	2.1
Tarsus depth	1-3	1.8	1-2	4.4
Tarsus length	1-3	1.4	1-2	6.9
Tarsus width	1-3	1.6	1-2	1.5
Weight	2-8	1.1	1-4	0.9

Kiwi weights fluctuate seasonally, with males in sample C undergoing a median weight increase of 16% between the breeding (September–February) and non-breeding periods (April–August). No females were measured close to breeding, and the infrequent captures of both sexes led to large separations between calls and nearest weight measures (mean = 63 days, range = 0–418 days). This meant that any changes in calls due to weight variation could not be reliably modelled. However, seasonal weight changes were likely due to variation in fat deposits, which should not have affected overall body size and hence dimensions of the vocal structures. That weight fluctuations were large compared to variation in tarsus measures (Table 2.2), and not correlated with them, supports this assumption that seasonal weight changes were not accompanied by corresponding body size changes. We therefore used the mean weight of each kiwi, and so assessed the call-size relationship among kiwi, rather than attempting to model weight-call variations within individuals.

2.2.5.1 *Morphology PCA*

Mean tarsus length and mean weight measures were highly correlated in the 15 kiwi in the morphology-call sample (Figure 2.1). These metrics were therefore combined in a principal component analysis (PCA). This was applied to both sexes, since males

and females had similar tarsus-weight correlations. The first principal component provided a size measure common to both sexes, enabling sex to be included as an explanatory variable in call-morphology models. Bill length was not highly correlated with tarsus dimensions and weight within each sex (Figure 2.1), so it was not included in the PCA. Bill length could therefore be included as a separate variable in the model to determine whether it influenced calls independently of body size.

Figure 2.1: Mean morphological parameters of the nine males and six females used for call-morphology analysis (sample C). Size PC1 is the first component of the principal component analysis applied to mean weight and mean tarsus. Numbers in the upper panels show correlations for each sex.

2.2.6 *Call parameter-morphology modelling*

Two mixed regression models were used to determine the relationship between call parameters and kiwi size, using calls from sample C (Table 2.1). The response variables chosen to represent call frequency and temporal structure were (1) mean of the centre syllable frequency and (2) syllable rate. Centre frequency was the frequency parameter most closely correlated with all others, and syllable rate was the temporal variable showing most disparity between the sexes. The explanatory variables for both models were the kiwi parameters sex, mean bill length, and mean size (first component of the tarsus-weight PCA), and the call parameters of distance and

whether the call was elicited with playback. Since the modelling was intended to assess sexual call dimorphism, interactions between sex and all other variables were included in the model. Interactions between kiwi morphology (size, bill length) and call parameters (distance and playback) were not included since they appeared unlikely from inspection of the data.

All continuous variables were normalised to facilitate comparison of coefficients. Data were inspected for outliers, and distance was log-transformed to reduce residuals from distant calls. The collinearity of the normalised dependent variables was confirmed to be low.

2.2.6.1   *Random component*

Multiple calls were recorded from each kiwi, so the necessity of a mixed model was tested by comparing the full model with and without a random intercept term for kiwi identity, using Akaike's Information Criteria corrected for small sample size (AICc; Hurvich & Tsai 1989) and a likelihood ratio test. Any heterogeneity of residuals was corrected by adding combinations of variance covariates to the full model, and then using residual–covariate plots and AICc to select the optimal covariance component (Zuur et al. 2009). Models were calculated in the R package nlme (Pinheiro et al. 2012).

2.2.6.2   *Fixed component*

Selection of the optimal fixed structure followed an information-theoretic approach (Anderson et al. 2000, Anderson & Burnham 2002). A suite of 27 models was constructed, each with the same random component but a different combination of explanatory variables. These models ranged from the null model (no explanatory variables) to the full model containing all explanatory variables and relevant interactions. Candidate models were selected based on ecological knowledge and inspection of covariate plots. A 95% confidence set of models was obtained by ranking by AICc, and then including all models with cumulative Akaike weight up to 0.95. An average model and the relative importance of the dependent variables were calculated from this 95% confidence set. Model adequacy was checked using the full model.

2.3   RESULTS

2.3.1   *Vocalisation types*

The most prominent of adult LSK vocalisations is the loud whistle call, audible over long distances. Low-frequency grunts were also heard, between pairs in the burrow during the day, during courtship, and frequently after (simulated) territorial intrusion by another kiwi. This suggests a pair-bond reinforcement function for this short range communication. Other sounds produced were loud sniffs and snorts, usually accompanied by heavy footfall and litter-rustling following playback; and occasional squeals and snarls during fights between birds, or when members of a pair appeared to surprise each other. All results in this paper refer to the whistle call (hereafter 'call').

2.3.2   *Syllable types*

LSK calls exhibited little structural variation within each sex, predominantly consisting of a single syllable type of simple frequency-modulated upward sweeps (Figure 2.2). Each syllable comprised a 'trill' of fine-scale frequency modulations (Figure 2.3). A descending frequency component ('hook') often occurred at the syllable end (Figure 2.3A). This was more common in male than female calls (M = 74% of calls, F = 26% of calls; equality of proportions test, $\chi^2 = 36.7, P < 0.001$).

We used a generalised linear model to assess the factors influencing this 'hook'. Its presence or absence was modelled as a function of call distance, whether the call was elicited by playback, and the interaction of these effects. The modelling procedure followed that used for the call-morphology models. Calls were used from the highest two quality levels with no clipping and a positive identification of kiwi. Only male calls were tested, to ensure sufficient sample size for the effects. Adding a random effect for kiwi identity decreased the model fit ($\Delta AICc = 2.2$), so a mixed model was not necessary. From a model suite of all possible combinations of effects, the 95% confidence set contained the two models without interactions (Table 2.A.1). Distance was the primary influence, with hooks being significantly more likely to be detected in calls at closer range (Figure 2.A.1). Playback showed some positive association with hook presence, but this effect was not significantly different from zero.
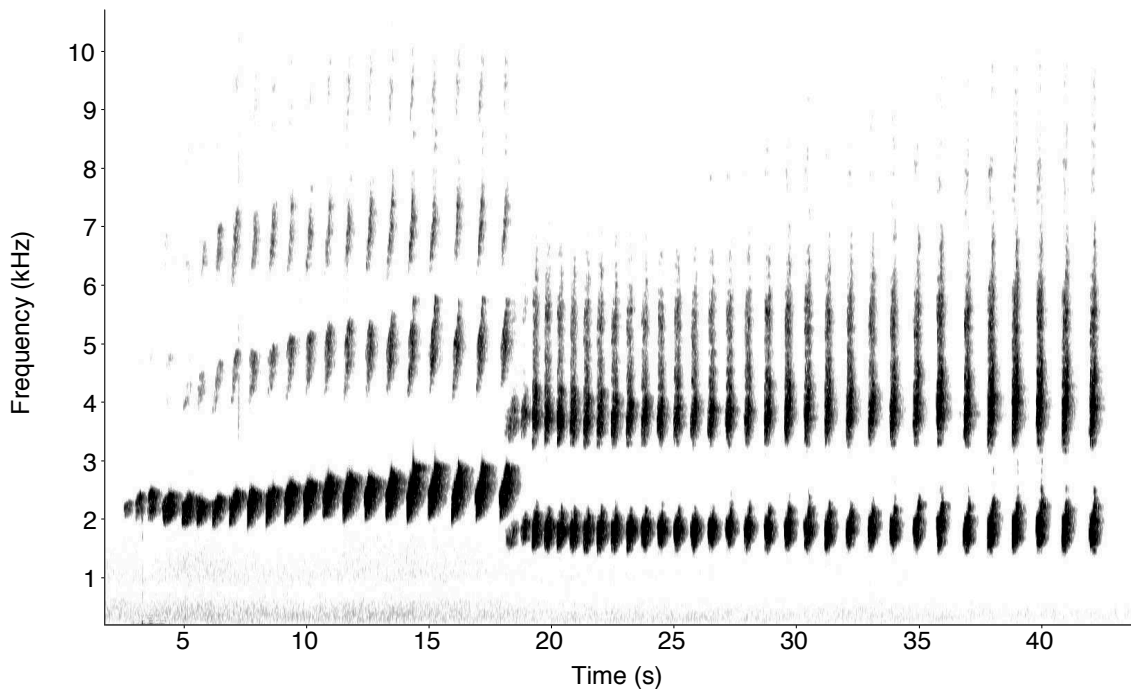
Figure 2.2: Spectrogram of typical male (left) and female (right) little spotted kiwi whistle calls. This is a non-overlapping pair duet; duets in which the calls of each sex overlapped antiphonally and synchronously were also recorded.

2.3.3 *Harmonic structure and spectral energy distribution*

Male calls showed clear harmonic structure, with up to nine harmonics to the sampling limit of 24 kHz. Female calls contained up to 12 harmonics, but the gaps between harmonics were less distinct and energy was more broadband above the fundamental frequency than in male calls (Figure 2.2). Female vocalisations also commonly displayed formant structure. These appeared as spectral shaping above f_0, and associated suppression of harmonic strength at certain ranges, most noticeably in the range ~ 2.25–3.25 kHz (Figure 2.2 and Figure 2.3). Formants were often difficult to discern, but occurred in at least 63% of female calls. They were not found in any male calls.

The frequency power spectra of male and female calls were strikingly disjoint, with little or no overlap between the sexes in the fundamental and first harmonics (Figure 2.4). The suppression of the lower half of the median female first harmonic at ~ 3 kHz was strikingly coincident with the median dominant energy of male calls. In addition, formants in female calls provided amplification at slightly higher frequencies (~ 4 kHz). This energy suppression was verified not to be an artefact of the recording equipment.
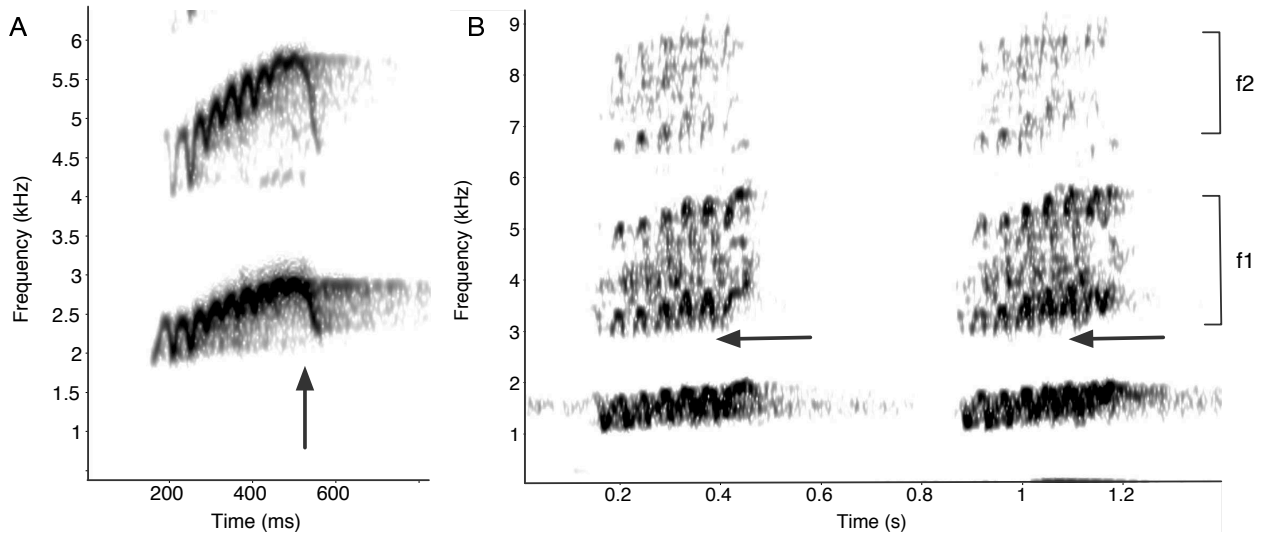
Figure 2.3: (A) Male little spotted kiwi syllable, showing rapid trill and 'hook' at the end (arrow). 768-sample Hann window, 75% overlap, 46.9 Hz resolution. (B) Formants in a female LSK call. Frequencies in the range $3.0 - 5.5$ kHz (f1) and $6.5 - 9.0$ kHz (f2) are amplified, whereas those at ~ 2.5 kHz and 6.0 kHz are suppressed. Suppression of the lower half of the first harmonic at $2.5 - 3.0$ kHz is evident in the missing troughs of the trill (arrows). 512-sample Hann, 75%, 46.9 Hz.

2.3.4 *Call spectral parameters*

Male and female syllables differed significantly in most frequency and temporal parameters (Figure 2.5), as measured from the fundamental frequency of calls in sample B (Table 2.1). Only call duty cycle and call length, measured by time or number of syllables, were equivalent between the sexes.

2.3.5 *Morphology PCA*

The first component of the PCA explained 84% of the variance in mean tarsus and mean weight for kiwi in the morphology-call sample (Table 2.3). It therefore provided a suitable indicator of kiwi size, demonstrated by its close correlation with mean weight and tarsus in both sexes (Figure 2.1).

2.3.6 *Sexual size dimorphism*

The 15 kiwi with calls in sample C exhibited overlap between the sexes in mean tarsus, weight and size, but not bill length (Figure 2.1). The sexes were significantly
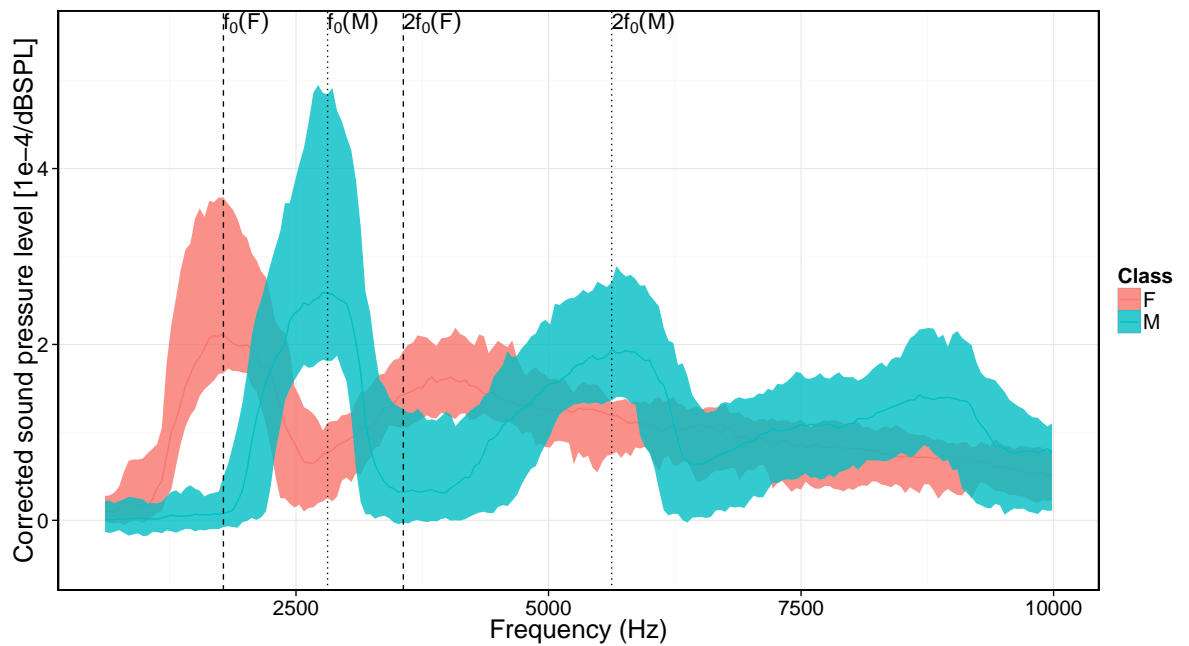
Figure 2.4: Average spectral energy distributions of male and female little spotted kiwi calls in sample D. The sexes had non-overlapping energy distributions, and the second energy peak in the female calls appeared at a higher frequency than that expected by harmonic structure ($2f_0$(F)). Lines and shading show medians and the 90% range of call power spectra (background-subtracted and normalised). Vertical lines (dashed = female; dotted = male) show the fundamental frequency (f_0) calculated as the energy peak in the median, and the expected frequency of the first harmonic ($2f_0$), based on the fundamental frequency.

Table 2.3: Importance and loadings of components in the principal component analysis of mean weight and mean tarsus measures, calculated for the 15 kiwi (9 males and 6 females) with calls in sample C.

Importance	PC1	PC2
Standard deviation	1.838	0.600
Proportion of variance	0.844	0.090
Loadings		
Mean weight	0.518	-0.114
Mean tarsus depth	0.487	-0.583
Mean tarsus width	0.524	-0.064
Mean tarsus length	0.469	0.802
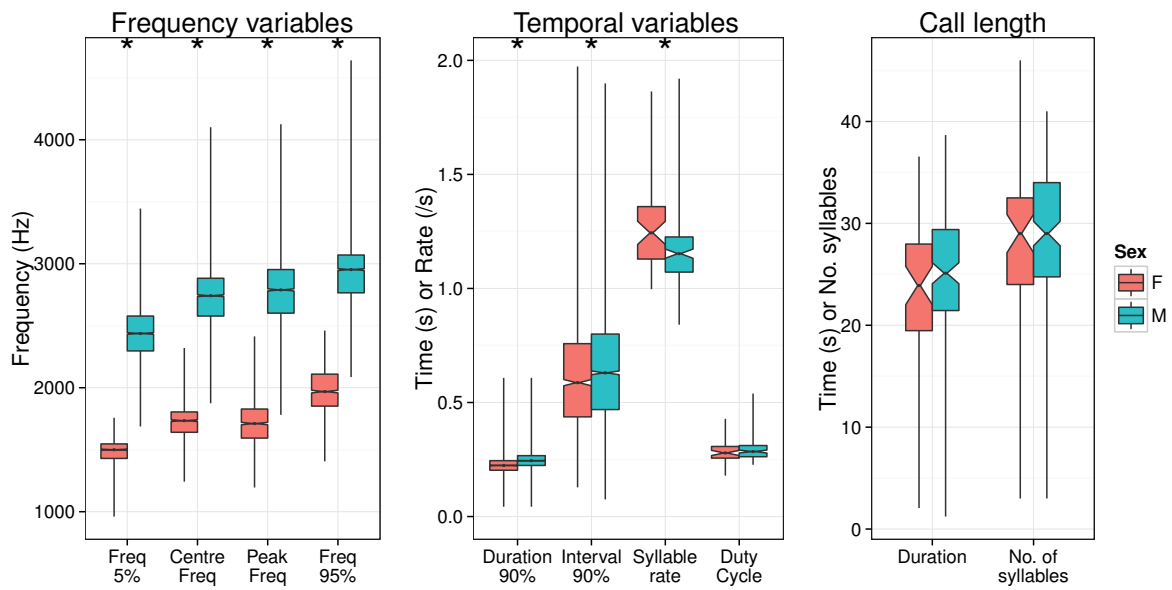
Figure 2.5: Little spotted kiwi calls showed significant inter-sexual differences in all parameters, with the exception of call duty cycle and length. Calls are from sample B. Boxes, lines and whiskers depict the central 50th percentile, median, and range, and notches depict 95% confidence intervals of medians. Asterisks denote parameters that are significantly different between the sexes, as defined by non-overlapping 95% confidence intervals of the median.

different in all morphological variables (Table 2.4), confirming size dimorphism in this sample.

Table 2.4: Sexual size dimorphism in the kiwi in the morphology-call sample. All mean morphological variables were significantly different between the sexes (Mann-Whitney U, 5%). n is the sample size, and s.e. is the standard error of the mean.

	Males			Females			Difference	
Variable	n	mean	s.e.	n	mean	s.e.	U	P
mean weight	9	1198.6	21.5	7	1407.7	27.0	0	< 0.001
mean tarsus depth	9	13.9	0.3	7	15.0	0.3	9	0.02
mean tarsus width	9	9.9	0.1	7	10.7	0.1	0	< 0.001
mean tarsus length	9	68.1	0.7	7	73.6	1.1	0	0.001
mean bill length	9	67.5	0.4	7	80.6	0.6	0	< 0.001
size PC1	9	-1.3	0.4	7	1.7	0.4	0	< 0.001

2.3.7 *Frequency-morphology model*

The optimal random component of the frequency-morphology model had an exponential variance structure with a different kiwi size covariate for each sex. The kiwi random effect was highly significant ($\Delta AICc = 11.4$, $L = 14.35$, df$= 1$, $P < 0.001$) and the intra-class correlation coefficient for calls within kiwi was 0.41. The 95% confidence set of models contained 11 models from the full set of 27 (Table 2.A.2). The model containing sex as the sole explanatory variable was the optimal model, more than twice as likely to be the correct model than the next highest-ranked (evidence ratio $= 0.34/0.16 = 2.2$).

All models showed that sex was by far the dominant factor influencing call frequency: it had a high relative importance, and was the only variable with a 95% confidence interval that did not intersect with zero in all models, including the average (Figure 2.6). In contrast, size, bill length, playback and distance and all interactions had very little effect on frequency: these variables had low relative importance and were included in few models, and mostly those with low rank.
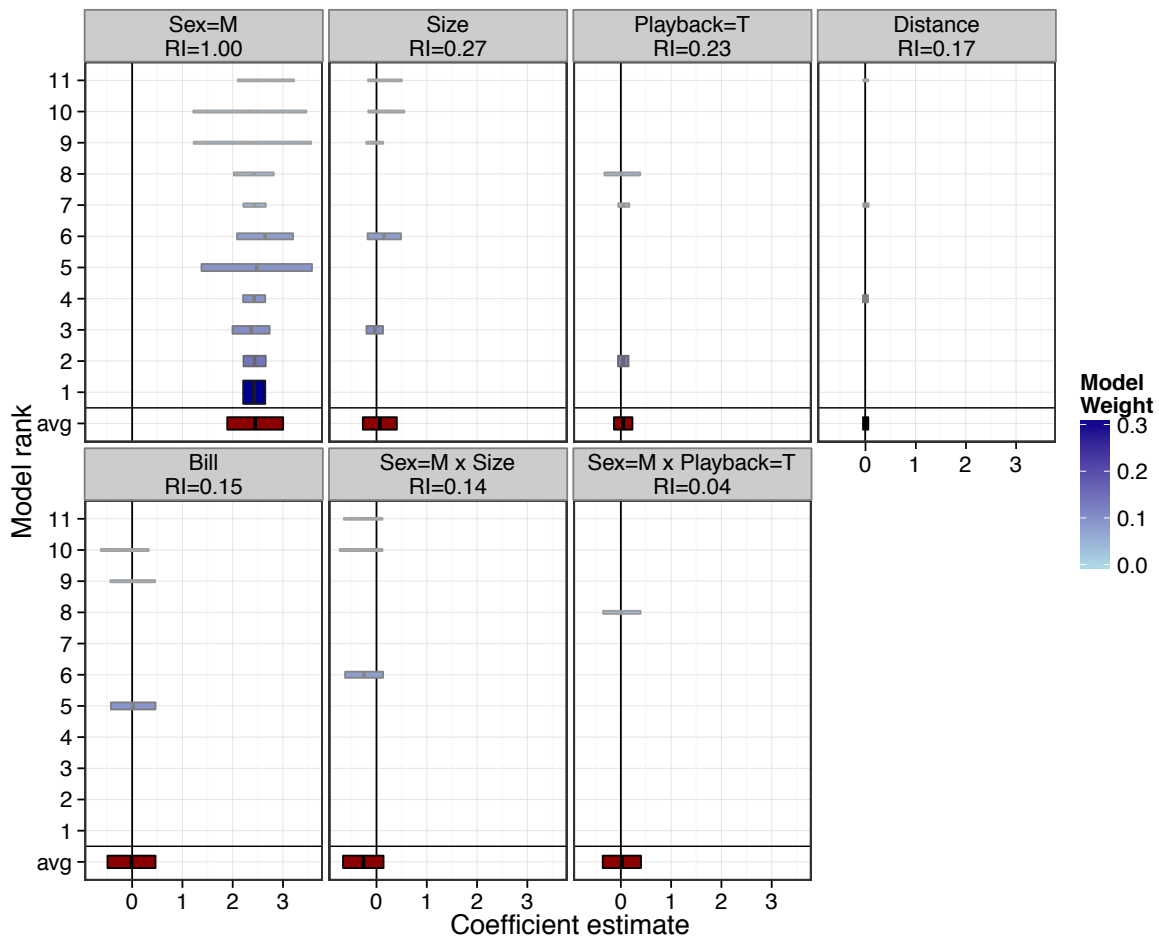
Figure 2.6: Fixed effect coefficients for the 95% confidence set of frequency–morphology models, and their weighted average. The response variable is mean call centre frequency. Sex was the most dominant and only significant effect on call frequency, with high relative importance (RI) and the only parameter with 95% confidence intervals (shaded bars) that did not intersect with zero. Models are ranked by AICc, with the optimal models and the weighted average over all models at the bottom of each panel. Bar thickness and shading reflect model Akaike weight.

2.3.8 *Temporal-morphology model*

The temporal model had an optimal random component with playback as a variance covariate. The kiwi random effect slightly improved the model ($\Delta AICc = 0.1, L = 2.845$, df$= 1, P = 0.046$) and so was included. Intra-class correlation for calls within kiwi was 0.18. There were just three models in the 95% confidence set selected from twenty-seven. This indicated low model uncertainty, which was confirmed by the two highest-ranked models having much higher evidence ratios than the third (11.9 and 3.5 respectively; Table 2.A.3). Sex, playback and their interaction were the dominant effects on syllable rate (Figure 2.7). Syllable rates were lower for males than females. Playback affected syllable rate, but differently for each sex: female rates were reduced and male rates increased for calls induced by playback.
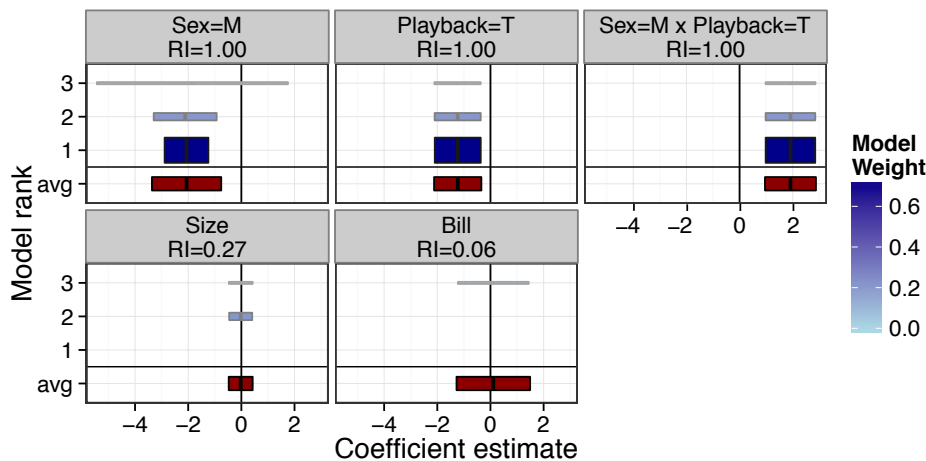
Figure 2.7: Effect coefficients for the 95% confidence set of temporal–morphology models with a response variable of call syllable rate. Sex and playback, and their interaction, were the dominant effects on syllable rate, whereas size and bill length had negligible impact. Males had a lower syllable rate, and playback reduced syllable rate in female calls, but increased it in male calls. Abbreviations and symbols follow Figure 2.6.

2.3.9 *Duetting*

A call was considered to be part of a duet if a nearby kiwi called immediately before, during or afterwards. From the 212 calls in sample A, 73 (34%) were part of a duet, and in 13 of these the calls overlapped. In all cases the duet was between birds of opposite sex, and was often confirmed to be between members of a pair. Females were significantly more likely to call as part of a duet than males (61% of female

calls were in duets, compared to 25% of males; equality of proportions test, $\chi^2 = 21.28, P < 0.001$). Females called first in fewer of their duets than males, although the difference was not significant (females initiated 39% of duets they were involved in, males 55%; $\chi^2 = 1.195, P = 0.274$). However, when only considering call-and-reply duets (not overlapping), females were significantly less likely to call first (females initiated 30% of call-reply duets, males 64%; $\chi^2 = 5.58, P = 0.018$). This contrasted with overlapping duets, in which females initiated 5/6 they were involved in, and males just 1/7 (sample too small for significance test). So in non-overlapping duets females were more likely to create the duet by replying to a partner's call, whereas in overlapping duets males tended to create the duet by joining the female call.

2.4 DISCUSSION

2.4.1 *Little spotted kiwi call characteristics and functional significance*

Kiwi have evolved physiological characteristics to suit their nocturnal, ground-dwelling, forest habitat niche (Martin et al. 2007, Corfield et al. 2008b, 2011). Their calls appear to have adapted similarly, differing substantially from their palaeognath relatives (e.g. Beaver 1978, Mack & Jones 2003, Corfield et al. 2008a). The peak energy of LSK calls is in the 1–3 kHz frequency 'window' that provides the best transmission of sound at ground level, both in forests and open habitat (Morton 1975, Marten et al. 1977, Marten & Marler 1977). The frequency modulation, high amplitude, repeated syllables and long duration of the call are characteristics that facilitate determination of distance (ranging) and direction (Naguib & Wiley 2001, Slabbekoorn et al. 2002, Catchpole & Slater 2008). LSK calls, both male and female, are clearly selected for locatability as well as long-range communication.

Spectral features such as 'hooks' may have functional significance. They are more likely to be detected in closer calls, suggesting that any function occurs over relatively short range. Hooks have slightly greater prevalence in calls elicited by playback, suggesting that they may signal aggressive intent to an intruder, and so would only require detection at closer distances. This playback effect is weak, but may be masked by the assumption that all calls recorded without playback were non-territorial. This certainly underestimates the proportion of aggressive calls, since some of those classified as non-aggressive (no playback) are likely to have been elicited by other kiwi.

The overall uniformity of syllable types in LSK calls contrasts strongly with the greater structural variation of brown kiwi calls (Corfield et al. 2008a; Chapter 4). It is possible that this may be related to the much lower genetic diversity in LSK (Ramstad et al. 2010), although differences in ecological factors such as social structure, habitat and territory size may also contribute.

2.4.2 *Call variation with morphology*

Inferences of call variation with body size rely on morphometrics that accurately represent size. With only vestigial wings and no tail, size measures of kiwi are limited compared to most other birds. However, the strong correlation between size and mean weight, and the degree of size dimorphism of individuals in our sample, demonstrates that our data are suitable for making these inferences.

Our results show that differences in LSK call frequency and temporal structure are not associated with size differences, both within and between the sexes. That inter-sexual call frequency differences are not a result of size dimorphism is in contrast to many other bird species (Eda-Fujiwara et al. 2004, Seddon & Sorenson 2005, Martin et al. 2011). It suggests that the kiwi vocal system does not scale allometrically, a result in common with mammals such as baboons and humans (Rendall et al. 2005).

Bill length does not influence call frequency or temporal characteristics, so LSK calls do not appear to be subject to constraints imposed by bill morphology found in many songbirds (e.g. Podos 1997, Hoese et al. 2000). The calling posture of LSK, with extended neck, bill pointing vertically upwards, and narrow gape (A. Digby, pers. obs.), suggests that the vocal tract is a more important acoustic filter than the bill (Fitch 1999, Riede et al. 2006)

Since inter-sexual call differences do not depend on size, the physiological nature and drivers of call dimorphism in LSK are unclear. In songbirds, greater mass and control of syringeal muscles enables males to produce a greater frequency range and more diverse song than females, which may be an important factor in the evolution of vocal complexity (Riede et al. 2010). However, in sub-oscines call frequency is related to airsac pressure rather than syringeal muscle control (Amador et al. 2008). This suggests that in the non-oscine LSK, vocal tract morphology, rather than the syringes, may be under selective pressures that differ between the sexes. We propose two non-exclusive selective forces: (1) inter-sexual differences in call function, and (2) inter-sexual acoustic cooperation.

2.4.2.1 *Inter-sexual dimorphism in call function*

We suggest that inter-sexual selective pressures on calls are related to territoriality. Male brown kiwi play a more dominant role than females in territory defence (Colbourne & Kleinpaste 1983, 1984, Robertson & Fraser 2009), which serves a likely reproductive function, rather than to protect food sources or shelter sites (Taborsky & Taborsky 1991, 1992, 1999). Male LSK appear to be similarly dominant (A. Digby pers. obs.). It is possible that this male dominance is related to the breeding system. Incubation is very long (65–80 days) and male-only in brown kiwi and LSK (Castro 2011), making sequential polyandry likely (Maurer et al. 2011). While social monogamy has previously been reported as the norm for all kiwi species (McLennan 1988, Heather & Robertson 2005), it is certainly not exclusive. An example is a cooperative and flexible mating system in a high density population of brown kiwi which do not defend territories (Ziesemann et al. 2011, I. Castro, Massey University, New Zealand, pers. comm., 2011).

That similar systems have not been found in LSK may reflect a relative lack of study effort, but it is feasible that social and sexual monogamy is a more effective breeding strategy where it is possible. This is particularly the case for male LSK, which would be likely to increase fitness less than females in a non-monogamous mating system. As sole incubators and with a very large egg to body size ratio (Reid 1971), the number of offspring males can produce each breeding season is limited to the one or two eggs they can incubate (Jolly 1989, Heather & Robertson 2005). So males may have a greater requirement to guard paternity, and achieve this through maintaining territory. Males would therefore have stronger pressure to use calls more often for territory defence, and to produce a call which is more effective for doing so, since natural selection favours calls that most effectively reach their targets (Endler 1992).

Male and female LSK calls differ substantially in their transmission properties. At ground level in forested habitat, sounds below 2 kHz are excessively attenuated compared to those in the 2–3 kHz range (Marten & Marler 1977, Marten et al. 1977). Lower frequency sounds are also more likely to be masked by wind noise (Ryan & Brenowitz 1985, Morgan & Raspet 1992, Slabbekoorn 2004). Tonal, narrowband calls are more easily detected by terrestrial vertebrate hearing apparatus than broadband ones (Wiley & Richards 1978). Therefore male LSK calls, with their higher frequency and more tonal properties, will transmit and be received more efficiently in their forest habitat than the lower frequency, broadband female calls. This is

borne out empirically: male LSK calls can be detected over greater distances than females (A.P. Digby unpubl. data).

Male LSK may have developed a higher frequency and tonal call not as a result of their smaller body size, but by virtue of stronger selection for acoustic adaptation to produce a call that is more efficient in resource defence. Female calls, transmitting less efficiently, would be more suited for pair contact, or to communicate aggression over shorter distances (Morton 1977).

Males are more vocal than females in brown kiwi and LSK, with higher calling rates in both species (Colbourne & Kleinpaste 1984, Corfield 2005; Chapter 5). Although we found that call length did not vary significantly between the sexes in LSK, males had a higher median call length and significantly longer syllable duration and interval, features that would aid detection.

The disparity in duet contribution further suggests functional differences between the sexes. Brown kiwi and LSK duet frequently (Colbourne & Kleinpaste 1984, Corfield et al. 2008a, this study), and following simulated territorial intrusion often join in overlapping duets (A. Digby, pers. obs.). These duets may serve territory defence, mate guarding or pair bonding functions (Hall 2009). Females have a higher proportion of their calls as part of duets in both brown kiwi and LSK, yet are less likely to call first (Corfield et al. 2008a, this study). This suggests that female calls may be used primarily for pair contact, by responding to their mate's call. It could alternatively imply that females respond to their partner to mate-guard, but this is at odds with overlapping duets nearly always being formed by males joining the female in both species (Corfield et al. 2008a, this study). This indicates that males join duets either to mate-guard, or to enhance territory defence. Below we propose the latter.

Variation in spectral features between male and female calls also implies intersexual functional differences. Male calls contained a higher proportion of 'hooks', which were loosely associated with calls stimulated by territorial intrusion. That female calls were more broadband, and commonly contained formants which are absent in male calls, also suggests differences in function.

Our finding that males increase their call rates when faced with territorial intrusion, whereas females adopt a slower call, may also indicate functional dimorphism. These responses may also depend on the sex of the intruding kiwi, which was not tested in the current study. Playback experiments would help determine the reasons for these sex-specific responses.

2.4.2.2 *Inter-sexual acoustic cooperation*

We propose that male and female LSK may have developed different calls to avoid
inter-sexual acoustic competition. This is supported by two pieces of evidence: (1)
the sexes had non-overlapping spectral energy distributions that were not a result
of their size differences, and (2) female calls had amplitude suppression of the first
harmonic (f_1) in the exact frequency range most utilised by males, and amplification
at slightly higher frequencies (Figure 2.4). This results in complementary frequency
distributions between the sexes that are unrelated to size dimorphism. We believe
this is the first time such inter-sexual vocal cooperation in frequency spectral profile
has been described in birds.

Female LSK appear to use vocal tract filtering to achieve this spectral shaping,
suggesting that avoiding acoustic conflict with males may be a driver for formant
production in kiwi. This is an alternative to the hypothesis that formants might be
used by kiwi to convey body size or identity (Corfield et al. 2008a), in a similar man-
ner to other birds that are reliant on non-visual cues, such as oilbirds (Suthers 1994).
Among kiwi, formants have only been found in female calls of the LSK and brown
kiwi (Corfield 2005, Corfield et al. 2008a). This sex-specificity suggests that form-
ants more likely function for 'acoustic cooperation' than to impart acoustic identity
or size information. However, males may not have a requirement for vocal identifica-
tion, or may use alternative methods. One candidate is the use of non-linear acoustic
phenomena such as subharmonics, which have been postulated to provide caller in-
formation in other species (Wilden et al. 1998, Riede et al. 2004, 2007, Zollinger
et al. 2008), and which are more commonly found in male than female LSK calls
(Chapter 3).

Modification of the female frequency distribution may serve to reduce formant
spacing, which is usually negatively correlated with trachea length and hence body
size (Fitch 1997). Several bird species with similar ecological traits to kiwi (dense
habitat, monogamy, sexual size dimorphism) have elongated tracheae, which may
be used to convey a false impression of size for the purposes of mate attraction and
territory defence (Fitch 1999, Charlton et al. 2007). Kiwi do not possess unusually
long tracheae (Forbes 1881; M. Alley and S. Jamieson, Massey University, New Zeal-
and, pers. comm., 2012), but female LSK may be able to manipulate their vocal tract
to decrease formant spacing and exaggerate size in a similar way to other animals
(Charlton et al. 2011).

In assessing any functional significance of the complementary spectral energy distributions, it is important to consider kiwi auditory ability, especially since the formant information in female calls is at high frequencies for birds ($\geq 4\,\text{kHz}$). This should be audible to LSK if they have similar hearing abilities to brown kiwi, which have unusually high frequency responses, with greatest sensitivity in the 4–6 kHz range (Corfield et al. 2011). However, the spectral resolution of kiwi hearing, and their ability to discern shifts in formant frequencies, is unknown.

The apparent inter-sexual frequency cooperation presented in LSK has clear implications for the function of overlapping duets. With disparate call frequencies, LSK have limited ability to 'jam' transmission of a mate's signal in paternity guarding (Hall 2009). Rather, their complementary frequency spectra mean that both sexes are audible when calling together, so either joint resource defence or reinforcement of the pair bond are more likely functions of these duets in LSK. That males nearly always create overlapping duets in both LSK and brown kiwi (Corfield et al. 2008a, this study) suggests that they have greater pressure to maintain territory and/or the pair bond. However, these results rely on calls from a small number of individuals, and there is inter-pair variation in duetting behaviour in brown kiwi (Corfield 2005). Playback experiments from a larger sample of individuals are necessary to make stronger inferences about duet function in kiwi.

We caution that the lack of frequency overlap between the sexes in LSK may be exaggerated by the low inter-individual call variation relative to intra-individual variation in this species (Chapter 4). The selective relevance may also be obscured by the fact that LSK suffered an extreme genetic bottleneck about 100 years ago (Ramstad et al. 2010). These implications can be further resolved by comparing calls from LSK populations with different genetic diversity, and assessing whether other kiwi species exhibit similar inter-sexual acoustic cooperation.

Disparities in contribution to resource defence are a likely driver for the acoustic differences between male and female LSK. We suggest that males use calls primarily for territory defence, whereas female calls are more suited for shorter-range functions such as pair contact. Functions such as mate attraction should also be considered by comparing male and female calling rates and territorial behaviour (Mennill 2011). Regardless, these inter-sexual vocal differences that are unrelated to size imply a divergence of call function that is in contrast with other species (Kirschel et al. 2011). This challenges the common assumption that male and female calls serve similar functions in birds (Collins 2004). Departures from this paradigm (e.g. Fedy & Stutchbury 2005, Maurer et al. 2011, this study) highlight the benefits of

bioacoustic analyses of a wide range of species for advancing understanding of avian call function.

2.A APPENDIX A[2]

Table 2.A.1: The 95% confidence set of the generalised linear models relating hook presence to playback and distance in male calls. Bold type indicates effects with 95% confidence intervals that do not intersect zero. $\Delta AICc$ = difference in AICc from optimal model; k = degrees of freedom; w = Akaike weight, which is the estimate that model is the true model, assuming the true model is in the set; L = log likelihood.

Rank	Effects	$\Delta AICc$	k	w	L	Resid. deviance
1	**Distance**	0.00	2	0.61	-29.13	58.26
2	**Distance**, Playback	0.91	3	0.39	-28.51	57.01
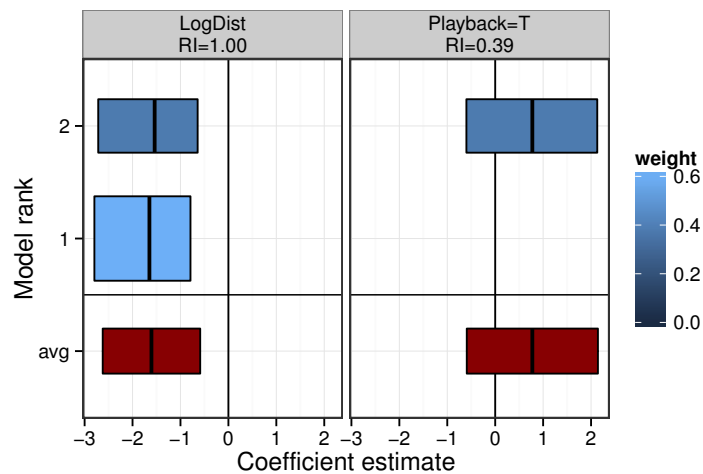
Figure 2.A.1: Effect coefficients for the 95% confidence set of models relating hook presence to call distance and playback, and their weighted average (avg). Closer calls were more likely to contain hooks, as were those induced by playback, although this effect was weak. Bars show coefficient estimates and their 95% confidence intervals, with thickness and shading reflecting model Akaike weight. RI is the relative importance of each effect.

---

2 This appendix appears as Supplementary Material in the Ibis manuscript.

Table 2.A.2: The 95% confidence set of the frequency–morphology models, selected by AICc rank. The response variable is the mean centre syllable frequency for each call. Sex is the only non-zero effect (bold type). Numbers in parenthesis are 95% confidence intervals for the random intercept and the residual standard deviations (s.d.). PB = playback; $\Delta AICc$ = difference in AICc from optimal model; k = degrees of freedom; w = Akaike weight, which is estimate that model is the true model, assuming the true model is in the set; L = log likelihood.

Rank	Effects	$\Delta AICc$	k	w	L	Random effect s.d.	Residual s.d.
1	**Sex**	0.00	6	0.31	5.70	0.147 (0.091,0.240)	0.178 (0.145,0.219)
2	**Sex**,PB	1.58	7	0.14	6.11	0.148 (0.091,0.241)	0.177 (0.144,0.217)
3	**Sex**,Size	2.20	7	0.10	5.80	0.144 (0.087,0.238)	0.178 (0.145,0.219)
4	**Sex**,Distance	2.36	7	0.10	5.72	0.147 (0.090,0.240)	0.178 (0.144,0.219)
5	**Sex**,Bill	2.39	7	0.09	5.70	0.147 (0.091,0.240)	0.178 (0.145,0.219)
6	**Sex**,Size,Sex×Size	2.67	8	0.08	6.80	0.135 (0.082,0.224)	0.178 (0.145,0.219)
7	**Sex**,PB,Distance	3.72	8	0.05	6.27	0.149 (0.091,0.243)	0.175 (0.142,0.215)
8	**Sex**,PB,Sex×PB	4.03	8	0.04	6.12	0.148 (0.090,0.241)	0.177 (0.144,0.217)
9	**Sex**,Size,Bill	4.66	8	0.03	5.80	0.144 (0.087,0.239)	0.178 (0.145,0.219)
10	**Sex**,Size,Bill,Sex×Size	4.73	9	0.03	7.03	0.132 (0.080,0.219)	0.178 (0.145,0.219)
11	**Sex**,Size,Distance,Sex×Size	4.96	9	0.03	6.92	0.134 (0.080,0.223)	0.177 (0.144,0.217)

Table 2.A.3: The 95% confidence set of temporal–morphology models. The response variable is the call syllable rate. Conventions and abbreviations as Table 2.A.2

Rank	Effects	$\Delta AICc$	k	w	L	Random effect s.d.	Residual s.d.
1	**Sex,PB,Sex×PB**	0.00	7	0.73	-96.16	0.335 (0.167,0.670)	0.921 (0.760,1.116)
2	**Sex,PB**,Size,**Sex×PB**	2.45	8	0.21	-96.16	0.336 (0.168,0.673)	0.919 (0.758,1.115)
3	Sex,**PB**,Size,Bill,**Sex×PB**	4.96	9	0.06	-96.14	0.341 (0.169,0.687)	0.917 (0.755,1.114)

2.B   APPENDIX B[3]

2.B.1   *Variation during calls*

The frequency and bandwidth of the syllables in LSK calls increases as the calls pro-
gresses (Figure 2.B.1). Male calls exhibit a greater frequency increase than females.
For both sexes the gap between syllables rises during the call, whereas the syllable
duration remains remarkably constant (Figure 2.B.2). The call sample for these plots
is Sample B (Table 2.1).
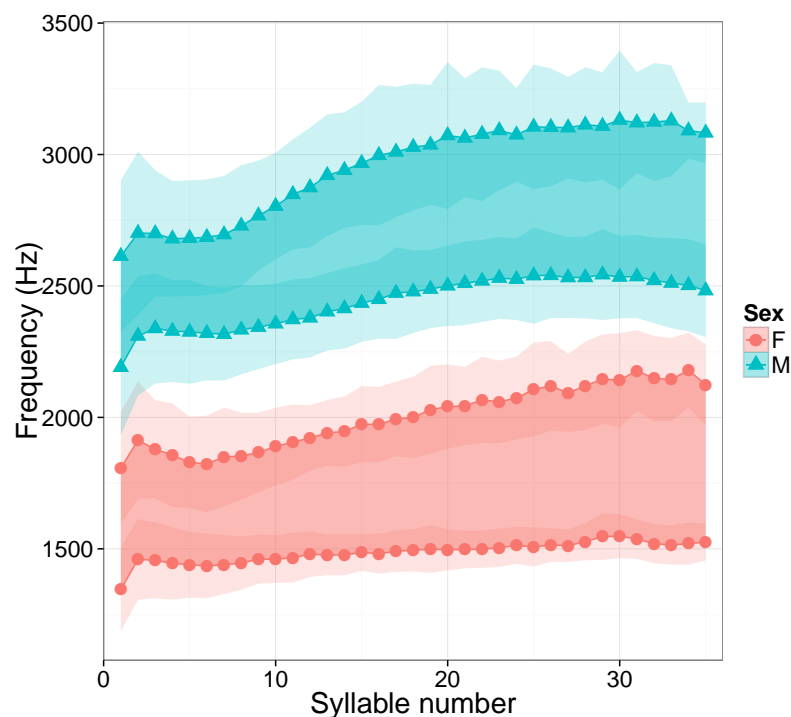
Figure 2.B.1: Frequency variations during the call. Upper lines for each sex are 95% fre-
quency; lower lines are 5% frequency, and 90% bandwidth is the gap between
these. Lines and shading depict means and one standard deviation about the
mean for each syllable number and sex.

2.B.2   *Syllable variations*

LSK calls showed high stereotypy, with most calls dominated by a single syllable
type. However, variations of this syllable structure occurred, in the form of a sudden

---

3 This appendix contains additional information describing LSK calls that was not included in the ma-
nuscript accepted in Ibis due to space restrictions.
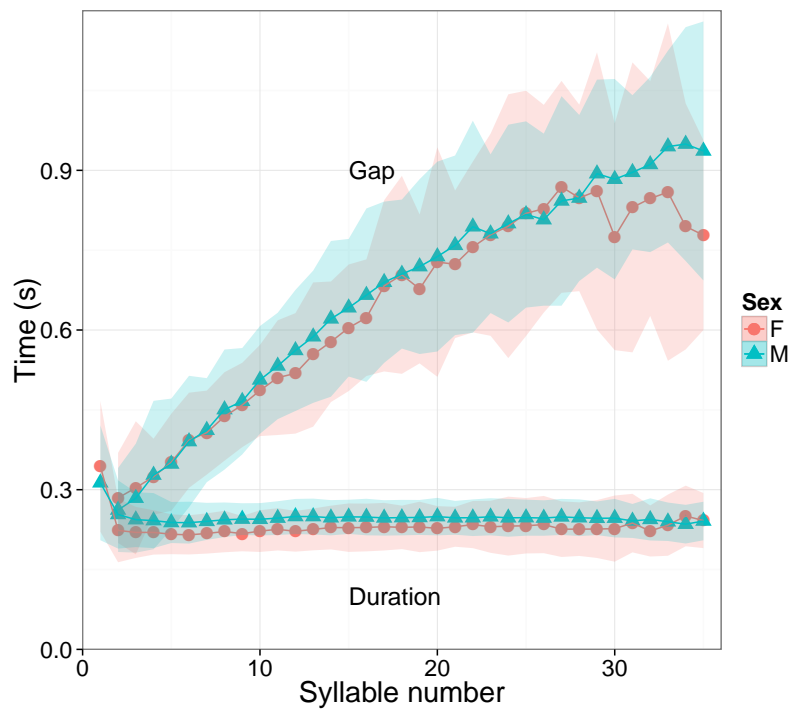
Figure 2.B.2: Temporal variations during the call, showing the syllable duration and gap between syllables. Lines and shading depict means and one standard deviation about the mean for each syllable number and sex.

transition during the syllable to either a higher frequency, or to a higher bandwidth of the syllable 'trill' (Figure 2.B.3). These variants were common in male calls, with 40% containing at least one syllable with a higher frequency component, and 49% containing more than one syllable with a high bandwidth section. The proportions in female calls were 31% and 30% respectively. These features nearly always occurred at the end of the call, but rarely featured in more than three syllables. Their cause is unclear, but their rarity and occurrence at the end of calls may suggest that they signify a loss of vocal control.
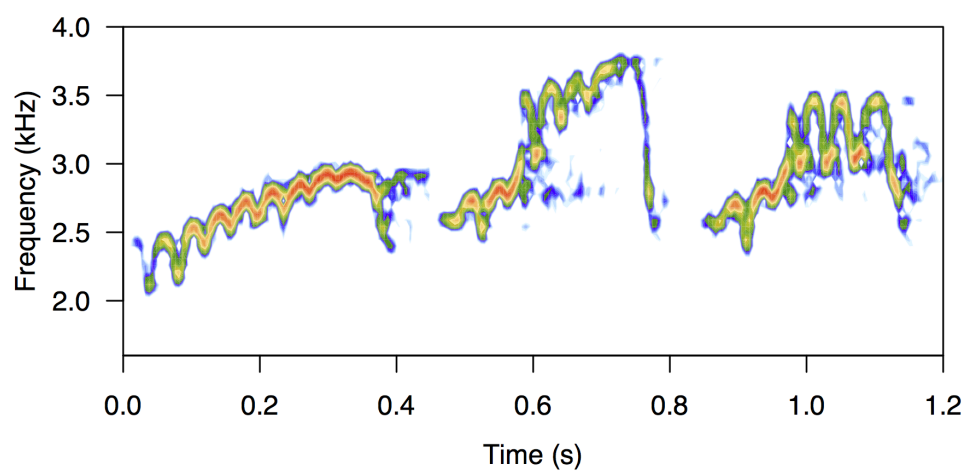
Figure 2.B.3: Variations of male LSK syllable types. A typical syllable (left), a syllable with a high frequency component (middle), and with a high bandwidth section (right). The latter two variants were usually found at the end of calls. These examples were taken from a single call, but only the last two occurred consecutively, and syllable gaps have been shorted to better illustrate syllable structure. These syllable variants also occurred in female LSK calls, but were less clearly defined due to the more broadband energy distribution.

TWO-VOICING AND NON-LINEAR PHENOMENA[1]

Cover of Manual of birds (Little spotted kiwi) by Buchanan, J., 1865-1882, after Keulemans, J. G., 1872-1882. Reproduced courtesy of Museum of New Zealand Te Papa Tongarewa. (1992-0035-2279/35).

---

1 This chapter appears as submitted to *Bioacoustics* in November 2012, except for minor textual amendments and the addition of an Appendix and colour figures.

ABSTRACT

Complex features in vocalisations can provide important behavioural information. Examples are two-voicing, which provides an identification mechanism in several birds species, and non-linear phenomena (NLP), which may serve to increase acoustic impact or convey fitness or identity information. Despite the important contribution they can make to understanding behaviour and call function, the use of these spectral features is unexplored in most bird species and many genera. This applies to the New Zealand kiwi (family: Apterygidae), with two-voicing and NLP-like features indicated in one species, but not explored further. Yet as nocturnal and cryptic birds heavily reliant on vocal communication, kiwi are ideal species in which to assess the function of complex vocal characteristics. Furthermore, kiwi are acoustically typical but taxonomically and ecologically unique among birds, so can provide an important reference for determining the behavioural relevance of these features in other bird species. We have assessed the occurrence of two-voicing and NLP in little spotted kiwi (*Apteryx owenii*; LSK), in the first study of such features in a ratite. Two-voicing is unexpectedly rare in this species, but NLP in the form of subharmonics are common. We tested the functional significance of these features by comparing their occurrence with call spectral parameters and conditions. Subharmonic use increased with frequency and during territorial calls, indicating that this feature may provide acoustic impact, possibly to convey aggression or fitness information. Counter to expectations, our results suggest that NLP and two-voicing are unlikely to provide identifying information in this kiwi species.

3.1 INTRODUCTION

Bioacoustic studies tend to characterise species using common call or syllable types, and typically disregard complex spectral features. However, these irregular elements may represent an opportunity rather than a barrier in understanding call function and acoustic behaviour (Fitch et al. 2002, Blumstein & Récapet 2009). They may impart important behavioural information to conspecifics (e.g. Townsend & Manser 2011), or enhance vocal diversity (Zollinger et al. 2008). The complex feature that has received most attention in avian communication is two-voicing. Non-linear phenomena (NLP) are another set of acoustic traits which have been widely documented in mammals (e.g. Wilden et al. 1998, Schneider & Anderson 2011), but until recently

Chapters   **Questions**   **Themes**
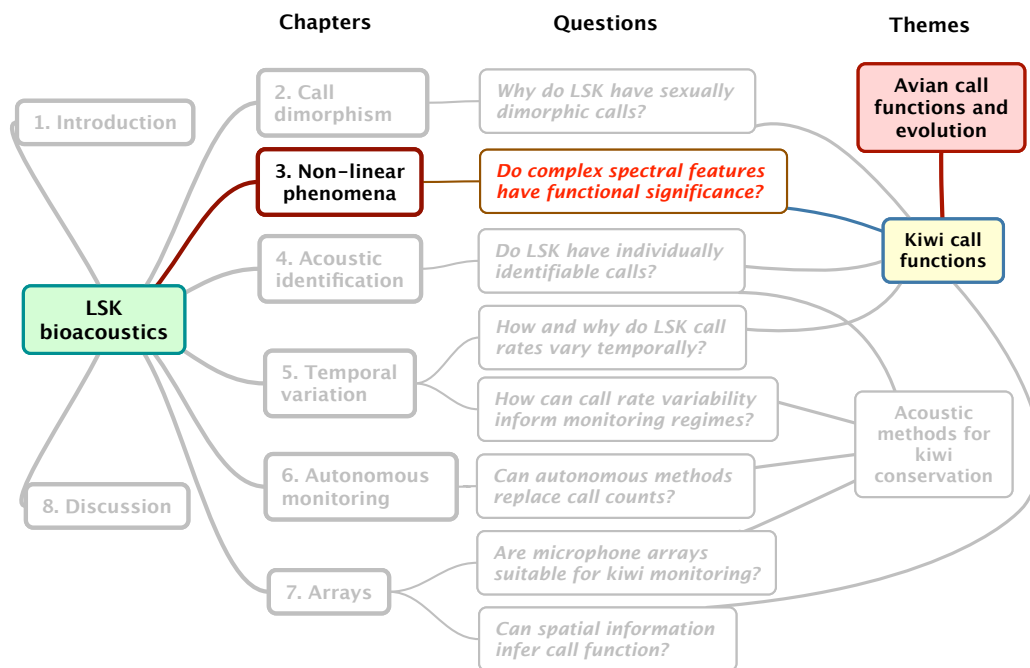
Topics covered by this chapter.

have been less studied in birds (Fee et al. 1998, Zollinger et al. 2008, Suthers et al. 2010).

3.1.1 *Two-voicing*

The vocal organ of birds, the syrinx, produces two sound sources which can be independently controlled (Suthers 1990). This provides an ability to produce two simultaneous sounds with non-harmonically related frequencies, referred to as 'two-voicing' (Greenewalt 1968). This feature is thought to serve to increase song complexity in songbirds (Zollinger et al. 2008, Suthers et al. 2010). Two-voicing has also been found in a number of non-oscines. Many of these are colonial nesting species (Robisson et al. 1993, Krakauer et al. 2009, Krull & Hauber 2010), and/or nocturnal birds (Suthers 1994) which cannot rely on visual cues to identify mates or offspring. This suggests an identification function, as clearly demonstrated in penguins (Aubin et al. 2000), in which use of two-voicing varies according to breeding ecology (Jouventin & Aubin 2002).

3.1.2 *Non-linear phenomena*

NLP arise from the intrinsic non-linear vibrational properties of the vocal apparatus. Four types are commonly documented (Wilden et al. 1998, Fitch et al. 2002, Zollinger et al. 2008):

- *Biphonation* is the appearance of simultaneous but independent fundamental frequencies (f_0, g_0) generated by a single sound source of two weakly-coupled oscillators (e.g. one side of the syrinx). This contrasts with two-voicing, in which the independent frequencies each arise from a separate source. Two principal forms are identified: (1) two independent frequencies that show no obvious interaction, and (2) sidebands above and below the fundamental due to its modulation by a lower frequency signal.

- *Subharmonics* occur when two coupled oscillators vibrate at different frequencies with an integer ratio. The oscillation period of the combined system is decreased by this ratio, so that intermediate harmonics appear at integer fractions of the fundamental frequency, f_0 (e.g. $f_0/2, f_0/3$).

- *Frequency jumps* are sudden, discontinuous changes in f_0, separated by a very short silent period. They are caused by gradual changes in vocal tension or pressure causing a sudden transition between vibratory regimes.

- *Deterministic chaos* is the sudden transition to broadband energy caused by de-synchronisation of coupled oscillators. It is distinguished from random ('white') noise by residual periodicity resulting in intervening harmonic 'windows', preceding subharmonics, or formants within the noisy segment.

Features spectrally resembling frequency jumps, subharmonics and biphonation can also result from two-voicing rather than non-linear processes (Mann et al. 2006, Zollinger et al. 2008, Suthers et al. 2010). Additionally, subharmonics and biphonation can arise from non-linear interaction of sounds from both sides of the syrinx with each other and the vocal tract (Nowicki & Capranica 1986, Laje & Mindlin 2005, Laje et al. 2008). It is therefore not always possible to determine from spectrograms alone whether apparent NLP are due to two independent voices, or whether they are unilaterally or bilaterally produced. This is only achievable by measuring syringeal airflow (Zollinger et al. 2008, Suthers et al. 2010), or by altering air density and sound speed through atmospheric manipulation (Laje & Mindlin 2005). However, the purpose of the current investigation is not to determine the origin of NLP, but their frequency of occurrence and possible functional significance.

3.1.2.1 *NLP function*

Although common in mammals (Wilden et al. 1998), and evident in diverse bird species (Fee et al. 1998, Beckers & ten Cate 2006, Suthers et al. 2010), the functional significance of NLP is unclear. By increasing vocal diversity, these features may be important in sexual selection in songbirds (Fee et al. 1998, Zollinger et al. 2008, Suthers et al. 2010). NLP occur at the limits of acoustic stability of the vocal tract and are more common in vocalisations of infants, subordinates (Wilden et al. 1998, Fitch et al. 2002) and unhealthy individuals (Wilden et al. 1998, Riede 2000). They therefore may provide an indicator of age, status or fitness (Riede et al. 2007). Another potential function is to provide an honest or dishonest indicator of size: by revealing formant structure, chaos may convey true size (Fitch et al. 2002), or subharmonics might provide a false impression of size by lowering perceived frequency (Wilden et al. 1998, Fitch et al. 2002).

By reflecting asymmetries in the vocal tract, NLP provide information which may enable discrimination between individuals (Fitch et al. 2002). This is supported by evidence from canids (Wilden et al. 1998, Mitchell et al. 2006, Volodina et al. 2006, Schneider & Anderson 2011), and may provide a similar function in juvenile birds (Suthers et al. 2010).

A commonly-proposed function of NLP is to yield variable calls that are harder to ignore (Fitch et al. 2002). This 'unpredictability hypothesis' is supported by the use of NLP in mammal alarm calls (Blumstein & Récapet 2009, Townsend & Manser 2011) and social calls (Schneider & Anderson 2011). It can also explain the more frequent use of NLP in calls of animals that are more likely to suffer being ignored, such as begging chicks (Suthers et al. 2010), young or subordinate animals (Wilden et al. 1998, Fitch et al. 2002) and animals under stress (Wilden et al. 1998). However, not all species exhibit NLP use in accordance with this auditory impact hypothesis (Riede et al. 2007).

NLP may provide other functions, such as directional movement or orientation information in cetaceans (Miller 2002, Filatova et al. 2012). It is likely that NLP serve multiple functions that may vary between species. In order to further understand their function, it is therefore necessary to investigate NLP occurrence in a diverse range of species.

3.1.3 *Spectral features in kiwi calls*

The endemic New Zealand kiwi (family: Apterygidae) are almost exclusively nocturnal (Heather & Robertson 2005, Castro 2011) and have poor vision (Martin et al. 2007), so are expected to rely heavily on vocalisations for communication. In all five species (Burbidge et al. 2003), both sexes vocalise with a loud whistle call (Heather & Robertson 2005). In brown kiwi (*Apteryx mantelli*) this is thought to be used primarily for territory defence and pair contact (Colbourne & Kleinpaste 1984, Corfield et al. 2008a). Yet kiwi call function is poorly understood, with the brown kiwi the only species acoustically analysed prior to this study (Colbourne & Kleinpaste 1984, Corfield 2005, Hojem 2006, Corfield et al. 2008a). It has been anticipated that kiwi use two-voicing for conspecific identification, in common with other species reliant on vocal cues (Corfield 2005). Two-voicing in brown kiwi was indicated in Corfield (2005), but only from a single male. In later results from the same study, Corfield et al. (2008a) drew attention to "unexpected structural features that call for further investigation" intermediate to the main harmonic series. These features superficially resemble subharmonics (Corfield et al. 2008a, Fig.1), but NLP have not been explored in any kiwi species.

In this study we investigate the calls of little spotted kiwi (*Apteryx owenii*; LSK) for two-voicing and NLP. LSK are the second-rarest kiwi species, extirpated from the mainland of New Zealand in the 1980s, with a current population of ∼1600 individuals across eight predator-free island sanctuaries (Holzapfel et al. 2008, H. Robertson, per comm, 2012). In a companion study (Digby et al. 2013), we showed that inter-sexual call differences are not a consequence of size dimorphism in LSK, and may have functional significance. The aim of this paper is to (1) assess the occurrence of two-voicing and NLP in LSK calls, and to (2) measure their variation with call circumstances and spectral parameters, to test the hypothesis that these features have functional significance. Understanding the use of two-voicing and NLP in the taxonomically and ecologically unique kiwi will enhance understanding of their function in other birds.

The aim of this chapter is to (1) assess the occurrence of two-voicing and NLP in LSK calls, and to (2) measure their variation with call circumstances and spectral parameters, to assess whether these features have functional significance. Three hypotheses are tested for NLP occurrence: that they (1) are used to add emphasis to calls during resource defence, (2) occur at the limits of vocal performance, and (3) are distance-limited. The respective predictions are that non-linear features will

occur significantly more often in (1) calls induced by playback, (2) higher frequency syllables, and (3) closer calls. Understanding the use of two-voicing and NLP in the taxonomically and ecologically unique kiwi will enhance understanding of their function in other birds.

3.2 METHODS

3.2.1 *Field recordings*

LSK calls were recorded over 140 nights between November 2009 and April 2012 at Zealandia, a 225 ha fenced reserve in Wellington, New Zealand. Forty little spotted kiwi were translocated to the reserve from nearby Kapiti Island in 2000–2001, and in 2010 the population was estimated at 100 individuals (H. Robertson, New Zealand Department of Conservation, pers. comm., 2010). During the study period 15–52 kiwi, the majority of which were male, were fitted with radio transmitters for separate studies. This enabled individuals to be located and calling kiwi identified. Calls were digitised at 48 kHz, 24-bit precision with hand-held recording equipment (Sound Devices 722 with Telinga Twin Science microphone and Telinga 53 cm foldable parabola). Calls were either recorded serendipitously from elevated listening positions, or by using conspecific playback to simulate territorial intrusion close to a known individual located by radio telemetry. During this process the calls of a male and then female LSK from the same population were broadcast using a digital recorder (Sony PCM-M10P) and small 8-watt loudspeaker (AJ Productions Ltd, Hamilton, New Zealand). Following playback the target individual and its mate would often approach the loudspeaker and call, enabling calls to be recorded at very close range. Since more males had transmitters attached, and because males are more vocal (Digby et al. 2013; Chapter 5) and more likely to approach simulated intrusion (Robertson & Fraser 2009), more male than female calls were recorded. Where possible the identity of the calling kiwi was determined using radio telemetry or coloured leg bands.

The distance to recorded calls was determined by estimating the call position in relation to geographic features, and marking the recording and estimated kiwi positions with Global Positioning System (GPS) waypoints. The distance between these waypoints was then measured using GPS software (Garmin BaseCamp 3.0). When calls occurred at such close range that the kiwi location was visible, the call distance was estimated in the field.

3.2.1.1 *Ethical notes*

Playback was used sparingly to avoid habituation and undue disturbance, and was not used in the main September–January breeding season (Heather & Robertson 2005).

The transmitters were attached for a Department of Conservation census (led by Dr. Hugh Robertson) and an unrelated Victoria University of Wellington study (led by Helen Taylor, Department of Conservation Permit WE/31248/FAU, and approved by Victoria University's Animal Ethics Committee). 'Kiwi chick' and 'juvenile kiwi' transmitters were used (Sirtrack Ltd and KiwiTrack Ltd, Havelock North, New Zealand), weighing 6 g and 11 g respectively. The weight of the largest transmitter was (mean and range) 0.9% (0.8%–1.2%) of male LSK body weight and 0.8% (0.5%–1.0%) of female weight. This was well below the 3% maximum recommended for kiwi (Robertson & Colbourne 2003). Kiwi capture, handling and transmitter attachment methods all followed established procedures described in Robertson & Colbourne (2003), and no detrimental effects of these were observed.

3.2.2 *Bioacoustic analyses*

LSK calls consist of a repeated series of around 30 short syllables (Digby et al. 2013; Chapter 2). Calls were assigned a five-level subjective quality score, based on amplitude and interference from other sources. Only calls from the two highest quality levels were used for analysis, provided they also contained at least ten syllables and no saturation ('clipping') of the detector. The resulting sample (A) contained 194 calls (145 male, 49 female) from 759 recorded, with a total of 5705 syllables (4246 male, 1459 female). A further subset (B) contained only calls in which the identity of the kiwi was independently identified, using either coloured leg bands or radio telemetry. Only adults were included in this sample, which was used to investigate factors affecting appearance of spectral features, with kiwi identity necessary to account for repeated measures. Subsample B contained 98 calls from 16 individuals (76 calls from ten males, 22 calls from seven females).

Spectrographic analysis was conducted in Raven Pro 1.4 software (Charif et al. 2010). The default spectrogram parameters were a 1024-sample Hann window, 50% overlap, and frequency resolution of 23.4 Hz. A shorter window and higher frequency resolution were used for inspecting each syllable for the presence of two-

voicing and NLP features. Syllable frequency was measured at the median of the energy distribution.

3.2.3 *Statistical analyses*

Factors influencing the appearance of spectral features were assessed using generalised linear mixed models on the call subset B. The sexes were modelled separately. The response variable was the presence or absence of two-voicing or NLP feature in each syllable. The explanatory variables were syllable frequency, call distance, whether the call was elicited by playback, and all interactions between these effects. Since repeated calls were recorded from each kiwi, a call identifier, nested in kiwi identity, was included as a random component in the model. Syllables in each call are temporally correlated, so an auto-regressive correlation structure was used, with syllable number within each call as a covariate. Models were calculated with penalised quasi-likelihood (PQL) estimation using `glmmPQL` (Venables & Ripley 2002) in R (2.1.15; R Development Core Team 2011). PQL estimation with a binary response variable can lead to underestimates in fixed and random effect sizes if cluster sizes are small (Rodríguez & Goldman 2001, Breslow 2003). However, in our data cluster sizes (syllables per call) were large (10–40), so such bias was minimal (Breslow 2003). Preferred informatic-theoretic model selection methods are not possible with quasi-likelihood estimation, so to avoid problems associated with model selection (Anderson et al. 2000), inferences were based on the full model with no term deletion (Zuur et al. 2012).

A generalised linear mixed model was also used to test whether NLP use differed between kiwi, controlling for playback and distance. The response variable was subharmonic presence in a call, and the explanatory variables were call distance, playback, and their interaction. The model was fitted to call sample B using R package `lme4` (Bates et al. 2012). Variation in individual NLP use was determined by comparing the AIC values of models with and without kiwi as a random effect.

3.3   RESULTS

3.3.1   *Two-voicing*

Only one example of a spectral signature of two-voicing was found from 5705 syllables in 194 calls (Figure 3.3.1). It is possible that this feature is actually a result of biphonation, since it is impossible to spectrographically distinguish between the two (Zollinger et al. 2008).
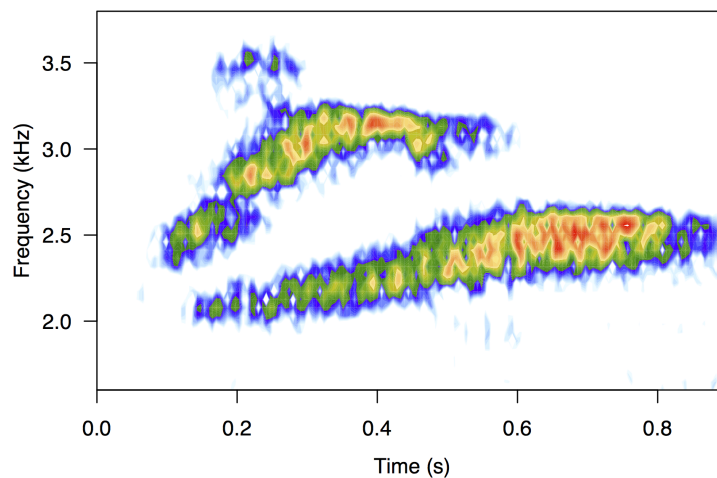
Figure 3.3.1: The only example of two-voicing found in the LSK call sample. This occurred in the first syllable of a male call, with two clearly independent signals from 200–600 ms.

3.3.2   *Non-linear phenomena*

Subharmonic features were commonly found in LSK calls (Table 3.3.1 and Figure 3.3.2). Most were period doubling ($f_0/2$), but higher order regimes were noted in both sexes (Figure 3.3.3). Subharmonics most commonly occurred throughout the whole call (51% of male calls and 63% of female calls containing subharmonics), or in the last half of calls for males (15% of subharmonic calls). The proportion of calls in which subharmonics were found did not differ significantly between the sexes (equality of proportions test: $\chi_1^2 = 2.22, P = 0.136$).

Signatures of **biphonation** were seen rarely in calls of either sex (Table 3.3.1). The occurrence was too low to test inter-sexual differences. Only side-band biphonation was found (Figure 3.3.4), with the exception of the two-voicing example noted

Table 3.3.1: Occurrence of signatures of NLP in LSK calls, from a total of 49 female and 145 male calls. Instances where NLP feature occurrence could not be clearly determined are excluded.

	Frequency (%)	
NLP feature	Male	Female
Subharmonics	46	33
Biphonation	4	8
Deterministic chaos	1	6
Frequency jumps	0	0
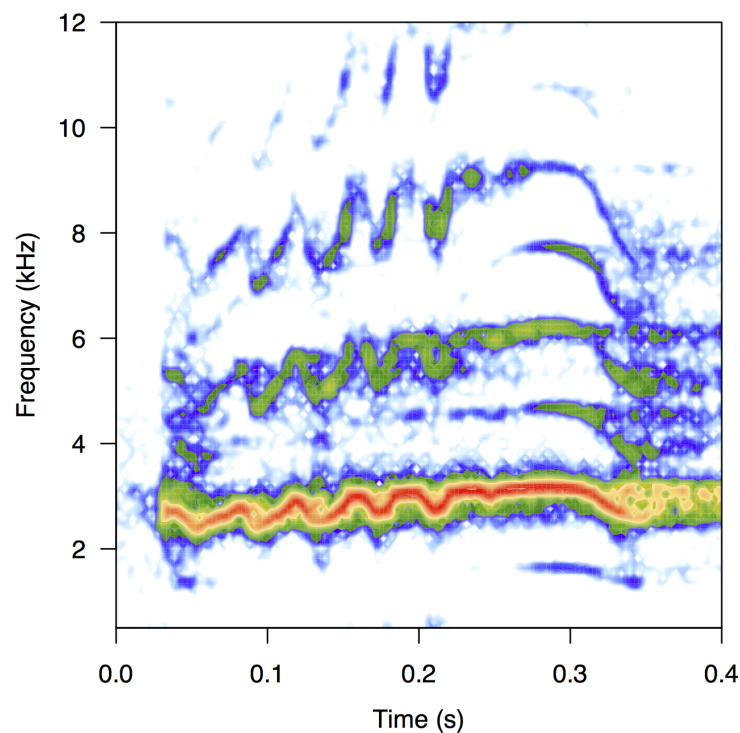
Figure 3.3.2: Subharmonics in a syllable of a male LSK call. The fundamental harmonic is at 2.5–3.0 kHz, with the first, second and third harmonics starting at ~5, 7.5 and 10 kHz. Intermediate to these is a subharmonic series with a period of $f_0/2$, clearly apparent at the end (300 ms) of the syllable.
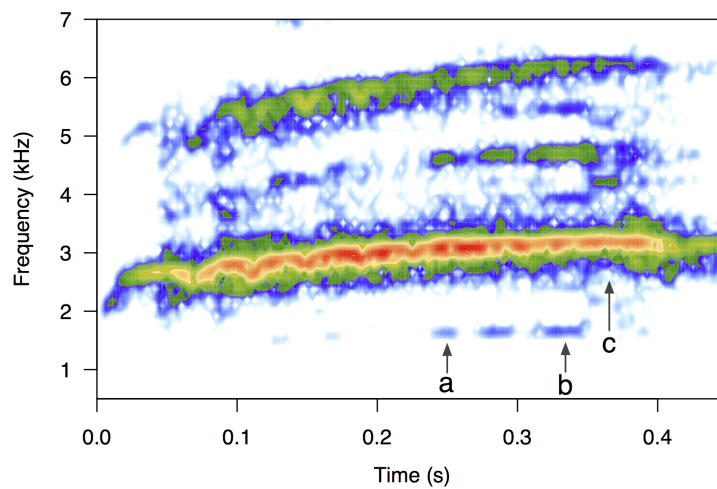
Figure 3.3.3: Different order subharmonics in a syllable of a male LSK call. Subharmonics appear first at $f_0/2$ (a), then $f_0/4$ (b) before a change to a $f_0/3$ regime (c).

above, which may represent biphonation with two independent fundamental frequencies. The side-band biphonation usually occupied only part of the syllable.

Frequency jumps were not detected in any call. **Deterministic chaos** was identified in both sexes, but was very rare (Table 3.3.1). It only occurred at the end of syllables and near the end of the call (Figure 3.3.5).

3.3.3   *Variation in NLP occurrence*

Modelling of factors influencing NLP presence was only possible for subharmonics in male calls; effect sizes were too small to test the appearance of subharmonics in female calls and other NLP types and two-voicing. The significant effects on subharmonic appearance in male calls were frequency, playback, and the interaction of frequency and distance (Table 3.3.2). Subharmonics were significantly more likely in playback calls and in higher frequency syllables. The significant interaction indicated that frequency dependence declined with distance. However, the model indicated that frequency would only have a negative impact on subharmonic appearance at distances of over 525 m, which was beyond the maximum distance at which subharmonics were detected (125 m). Distance itself did not directly affect subharmonic presence. The model showed a high correlation between syllables within each call ($\alpha = 0.73$, intra-class correlation = 0.88), but only weak correlation between syllables from different calls for each kiwi (intra-class correlation = 0.06). The model conformed to assumptions of homogeneity and normality, and temporal in-
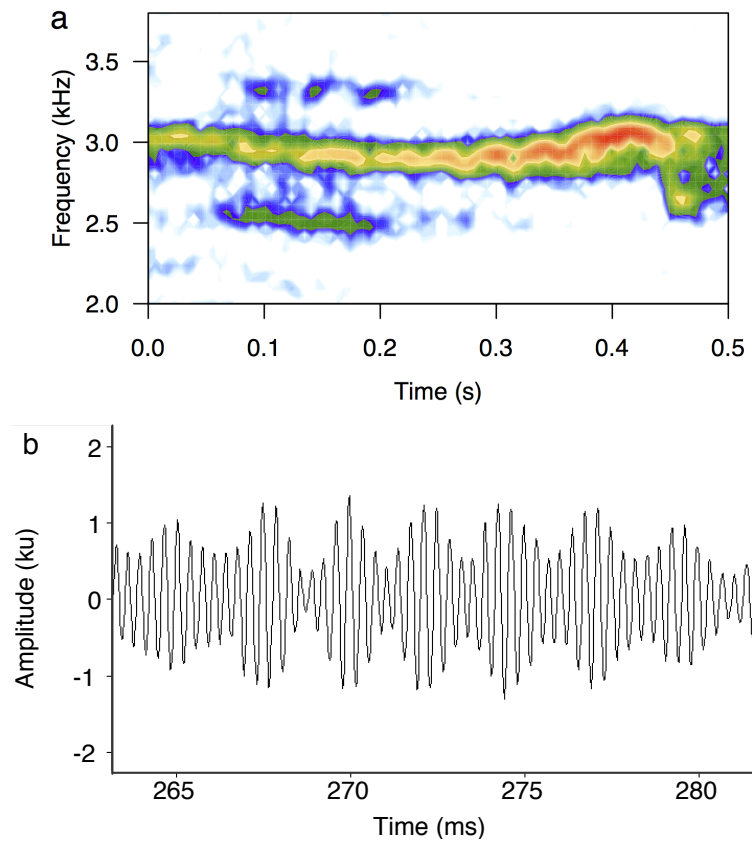
Figure 3.3.4: Sideband biphonation in a syllable of a male LSK call. Sidebands occur in the spectrogram (a) between 100–200 ms, at ±400 Hz relative to the fundamental frequency. The sideband at +400 Hz is ∼10 dB weaker and just visible in this scaling. The waveform (b) shows a periodicity of ∼2.5 ms for this section, corresponding to a 400 Hz modulating frequency which produces the sidebands.
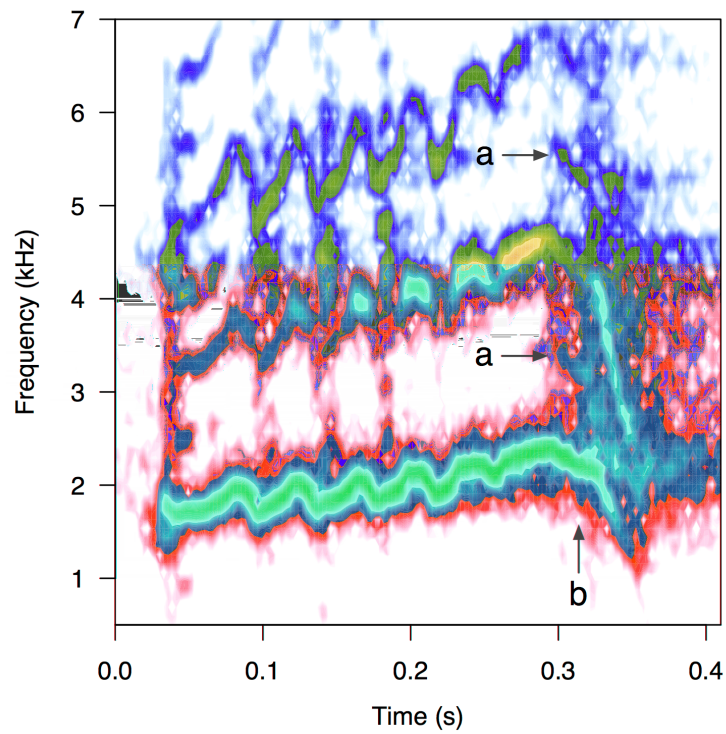
Figure 3.3.5: Deterministic chaos and subharmonics in a female LSK syllable. The syllable shows harmonic structure before chaos occurs at ~320 ms (b). Subharmonics are apparent at $3f_0/2$ and $5f_0/2$ (a) immediately prior to the chaotic section. Subharmonics are also evident at the first half of the syllable.

dependence of residuals confirmed the suitability of the auto-regressive correlation structure.

Table 3.3.2: Mixed model effect coefficients of subharmonic presence in 76 calls from ten male LSK. Subharmonic presence increases with frequency and playback, but decreases with the interaction of frequency and distance. Residual standard deviation was $\sigma = 0.704$, and the first order autoregressive correlation parameter between syllables in the same call was $\rho = 0.750$. Frequency and log(distance) are normalised. Asterisks and bold type denote effects that are significantly non-zero at the 5% level. PB = T signifies that playback was used to elicit a call.

Effect	Estimate	Std. Error	d.f.	t	P	
(Intercept)	-2.29	0.64	2264	-3.60	< 0.001	
frequency	1.32	0.30	2264	4.35	< 0.001	*
PB = T	1.69	0.68	63	2.49	0.015	*
log(distance)	-1.19	0.66	63	-1.81	0.075	
frequency × PB = T	-0.60	0.32	2264	-1.91	0.056	
frequency × log(distance)	-0.89	0.30	2264	-2.93	0.003	*
PB = T × log(distance)	0.21	0.72	63	0.29	0.769	
frequency × PB = T × log(distance)	0.55	0.32	2264	1.74	0.083	

The low intra-class correlation between syllables in different calls from the same kiwi indicated little consistency in individual subharmonic use. This was supported by the model relating subharmonic presence to distance and playback in male calls. Including kiwi as a random effect did not improve the model fit ($\Delta AIC = 1.51, P = 0.31$). The intra-class correlation between calls from the same kiwi in the random effect model was low ($\rho = 0.017$). There were therefore no strong differences in subharmonic use between individuals. Effect sizes were too small to test for the difference between individuals for subharmonics in female calls and other NLP types.

3.4 DISCUSSION

Two-voicing was rare in our sample of LSK calls, with only one example of independent fundamental frequencies. This is counter to the expectation that kiwi might use two-voicing for identification, in common with other bird species that are unable to rely on visual cues (Corfield 2005).

In contrast to two-voicing, LSK calls frequently contained signatures of NLP, with subharmonics common in calls from both sexes. Other NLP types were present, but rare. The scarcity of biphonation in LSK was similar to its rarity in other species (Mann et al. 2006, Zollinger et al. 2008, Suthers et al. 2010). Deterministic chaos was also rare, but this was not unexpected given the short duration of syllables in LSK calls.

Features spectrally resembling subharmonics may be due to two-voicing (Suthers et al. 2010). However, this was very unlikely in our sample, given the rarity of two-voicing with independent f_0. For the subharmonic features to have resulted from non-interacting voices would have meant that two-voicing only occurred with spectrally identical signals that had integer frequency ratios. The positive association between these features and syllable frequency also was in accordance with subharmonics in chimpanzees, which are more common with increasing f_0 (Riede et al. 2007).

Biphonation with independent frequencies cannot be resolved spectrographically from two-voicing (Zollinger et al. 2008). But all examples reported here are side-band biphonation, in which the modulation of the fundamental frequency by a lower frequency signal is clear (Figure 3.3.4), and which can only result from non-linear processes (Nowicki & Capranica 1986).

We could not determine whether the subharmonics and biphonation were unilaterally or bilaterally produced without measurement of syringeal airflow or atmospheric manipulation (Laje et al. 2008, Zollinger et al. 2008, Suthers et al. 2010). Such experiments were not possible or desirable in the near-threatened LSK (IUCN 2011). Nevertheless, these features, and the presence of deterministic chaos, which cannot be confused with any linear system, confirm the presence of NLP in LSK calls.

3.4.1   *NLP function*

The tests of subharmonic occurrence suggests that these features have functional significance in male LSK calls. The increase in subharmonic appearance in calls elicited by playback, when the kiwi were faced with apparent territorial intrusion, supports the acoustic impact hypothesis. Call temporal structure changes in these circumstances (Digby et al. 2013; Chapter 2), and subharmonics may provide further acoustic emphasis to convey aggression or urgency. The greater prevalence of subharmonics at higher frequencies supports the hypothesis that they occur at the acoustic limits of the vocal tract, suggesting an association with vocal stress. LSK

syllables increase in frequency as the call progresses (Digby et al. 2013; Chapter 2), reaching an asymptotic limit at the conclusion of the call. This suggests a performance limit is being reached, and the occurrence of subharmonics at this limit indicates a loss of vocal control when 'vocalising at the edge' (Riede et al. 2007). Loss of vocal control driving NLP production has been used to explain the predominance of chaotic structure in young animal vocalisations (Wilden et al. 1998). This is supported in LSK, with subharmonics, chaos, biphonation and frequency jumps common in chick calls (§ 3.A).

Factors influencing other NLP types could not be modelled due to their rarity. However, deterministic chaos also showed a relation with frequency, occurring only at the end of notes at the end of calls, when highest frequencies were reached. Chaos may therefore be associated with a loss of vocal control in LSK, as suggested for other species (Riede 2000, Suthers et al. 2006, Zollinger et al. 2008). Biphonation most frequently occurred at the start of calls in our adult sample, and had no clear association with higher frequencies. It appears that this form of NLP has a different causal mechanism and so may serve a different function.

NLP could provide an indicator of body size or fitness in LSK. Subharmonics may serve to lower the pitch of a call, and give the false impression of a larger body size during conflict (Wilden et al. 1998, Fitch et al. 2002). The prevalence of subharmonics in aggressive (playback) calls in LSK is in accordance with this hypothesis. Frequency and temporal call characteristics in LSK do not reflect body size information (Digby et al. 2013; Chapter 2), so such a mechanism for dishonest size indication may be advantageous. Chaotic structure in calls can emphasise formants, providing a means of conveying body size over long distance (Fitch et al. 2002, Reby et al. 2005). By altering formant structure, this can also be used to dishonestly signal size (Charlton et al. 2011). However, this function of deterministic chaos would appear to be less significant in LSK, since instances are rare, and then only occur for a very short period. Additionally, formants are apparent in the broadband female LSK calls without deterministic chaos (Digby et al. 2013; Chapter 2).

The occurrence of subharmonics in male calls did not vary greatly with individual, once playback and distance were accounted for. This suggests that factors affecting subharmonic production were equivalent across all males. If subharmonics were functional, then their use did not vary between individuals in the same call circumstances.

The effect of subharmonic presence increasing with frequency declined with distance in our model, suggesting that any function would operate only over shorter ranges. Although LSK calls transmit to distances of at least 800 m (A. Digby, unpublished data), we detected no NLP features in calls beyond 125 m. Territory areas at the study site are ~2–6 ha (A. Digby, unpubl. data), so NLP would be detectable only to neighbours, mates and intruders. This suggests that LSK calls may have distance–dependent functions, as suggested for other species (Barker 2008). For example, at large distances the call may impart species, sex and approximate location information, whereas at short range further cues such as stress, fitness, or aggressive intent may be apparent.

The rarity of two-voicing suggests that it is not used for individual identification in LSK, but NLP may provide an alternative mechanism for this by reflecting idiosyncrasies of the vocal tract (Fitch et al. 2002). Yet NLP were evident in fewer than half of all calls, so would not serve a consistent means of identification. Formants may provide an alternative, since they also reflect individual vocal tract morphology, and are frequent in female LSK and brown kiwi calls (Corfield et al. 2008a, Digby et al. 2013). But their occurrence in the calls of just one sex cautions against their interpretation as an identification mechanism. A functional relevance for formants might be to avoid frequency overlap between the sexes during duetting (Digby et al. 2013; Chapter 2).

With two-voicing very rare, and NLP and formants infrequent, do LSK use calls for recognition at all? It is feasible that as a socially monogamous bird defending year-round territories, these kiwi have little requirement for vocal individual identification. Every nearby calling intruder may be assessed as a potential threat, irrespective of identity (Bard et al. 2008). A lack of vocal discrimination in LSK is indicated by their low inter-individual call variation (A. Digby, unpublished data). This is in contrast with brown kiwi (Corfield 2005), but is supported by observations of an LSK attacking its mate that had just called (A. Digby, personal observations). In the absence of auditory recognition, LSK may distinguish between neighbours and intruders using alternative cues. One likely mechanism is scent, which appears to be an important mode of communication among brown kiwi (Castro et al. 2010).

Although NLP were indirectly indicated in brown kiwi in Corfield et al. (2008a), this is the first study into their occurrence in any ratite. Two-voicing is unexpectedly rare in LSK, but NLP in the form of subharmonics are relatively common. The variation in their appearance with frequency and playback in males suggests that they are functionally significant, possibly providing additional vocal impact to con-

vey aggression or body size or fitness. Analysis of these features in other kiwi species, and during simulated territorial intrusion in playback experiments, would further explore the functional significance identified here. Our confirmation of the occurrence and behavioural relevance of NLP in the ecologically and evolutionarily distinct kiwi provides a new perspective on the function of these features in avian communication.

3.A   NLP IN LSK CHICK VOCALISATIONS[2]

3.A.1   *Introduction*

Over two nesting seasons during the study period autonomous recorders were placed at LSK nests in Zealandia. This was initially for two main reasons. The first was to use a microphone inside the nest to determine whether 'pipping', made by the chick (Figure 3.A.1) as it breaks through the inner part of the egg in the days prior to hatch, could be acoustically detected within the nest of wild kiwi (Figure 3.A.2). This would provide advanced warning of hatch, enabling chick rescue in areas where they are prone to predation. This work was suggested by and carried out with Rogan Colbourne (Department of Conservation). The second motivation was to use recorders to provide nest exit and entrance times for monitoring incubating males, prior to the use of motion-activated infrared remote cameras. This was part of a pilot study into LSK reproductive success, led by Kristina Ramstad (Victoria University of Wellington; VUW).

Figure 3.A.1: One of the LSK chicks monitored.

Results from one nest with an internal microphone in the first season (2010-2011) indicated that the 'pipping' could not be easily detected, but that acoustic methods were suitable for monitoring incubation patterns. The recordings also provided many

2 this section was not included in the manuscript submitted to bioacoustics, and will be prepared as a separate paper.

calls from the chicks, and studying this vocal behaviour became the focus for the subsequent nesting season. The aims were to assess: (1) call ontogeny in hatchling LSK, (2) the function of chick calls, and (3) male-offspring interactions, with a view to determining the function of males remaining with chicks post-hatch. These were investigated by examining variation in chick calls with age, and changes in output and structure in relation to the presence of the adult male. Data analysis is ongoing at the time of writing, so full results cannot be included in this thesis, but will be published following submission. However, preliminary results and examples of chick vocalisations are presented here.

Figure 3.A.2: A little spotted kiwi nest.

3.A.2 *Methods*

All nests monitored were in Zealandia, and were located from tracking of males fitted with radio transmitters. These were attached for the Department of Conservation census led by Hugh Robertson in 2010, and then retained from 2011–2013 for the reproductive study led by VUW PhD student Helen Taylor (§ 3.2.1.1). An internal microphone was installed in one nest in the first breeding season (2010–2011), and in two nests in the second season (2011–2012) (Table 3.A.1). Microphones were installed prior to hatch dates, estimated from incubation start dates assuming a 65–75 day incubation period (Heather & Robertson 2005). The recorders and microphones were removed after the nests had been vacated by both male and chick.

Table 3.A.1: Nests monitored acoustically with internal microphones in Zealandia from 2010 – 2012.

Season	Male band	Recording start	Recording end	Notes
2010–2011	O-32668	13/Nov/2010	15/Jan/2011	
2011–2012	O-31403	17/Nov/2011	18/Jan/2011	Two eggs laid and hatched
2011–2012	O-31659	25/Dec/2011	28/Mar/2012	

Installation of recorders took place at night once the male had left the nest. Radio telemetry was used to confirm that the male was far enough away not to hear activity at the nest, and care was taken to listen for the female (usually not transmittered) returning. A Song Meter 2 autonomous recorder (Wildlife Acoustics Ltd, Concord, MA, USA) was placed close to the nest, with one microphone attached directly to the recorder to record vocalisations outside the nest, and the other placed inside the nest and attached with a 10 m extension cable. The internal microphone was inserted by hand-drilling a 20 mm hole into the nest chamber, ensuring that the hole was within 30 cm of the egg, but not directly above it (Figure 3.A.3). The hole was then lined with 40 cm length of 20 mm diameter PVC piping, and a clamp affixed to the outside on the surface to ensure that it did not fall into the nest. The microphone, attached to the extension cable, was inserted into the piping until the microphone protruded by ∼1 cm into the nest cavity.

Recordings were made at 16 kHz sampling rate and 16-bit precision, and were conducted in continuous three-hour blocks from sunset to sunrise each night. Although the male would be absent for much of this period at the incubation start, this covers the period when the male would be exiting and entering the nest, and would include time when the chick was with the male and on its own. LSK chicks typically spend the first 5–7 days after hatch without leaving the nest, and thereafter leave soon after sunset, remaining in the vicinity of the nest and returning frequently in the first few weeks (A. Digby, unpubl. data).

Spectrograms were analysed in Adobe Audition 5.0 (Adobe Systems, CA, USA) and Raven Pro (Version 1.4 & 1.5; Charif et al. 2010). Hann window size varied from 1024 samples to 256 samples, depending on the temporal extent being examined. Periods of vocal activity were located by viewing the full extent of each three hour file, and then examining each period in more detail. Each syllable of chick vocalisation was inspected for two-voicing and NLP, and annotated accordingly.
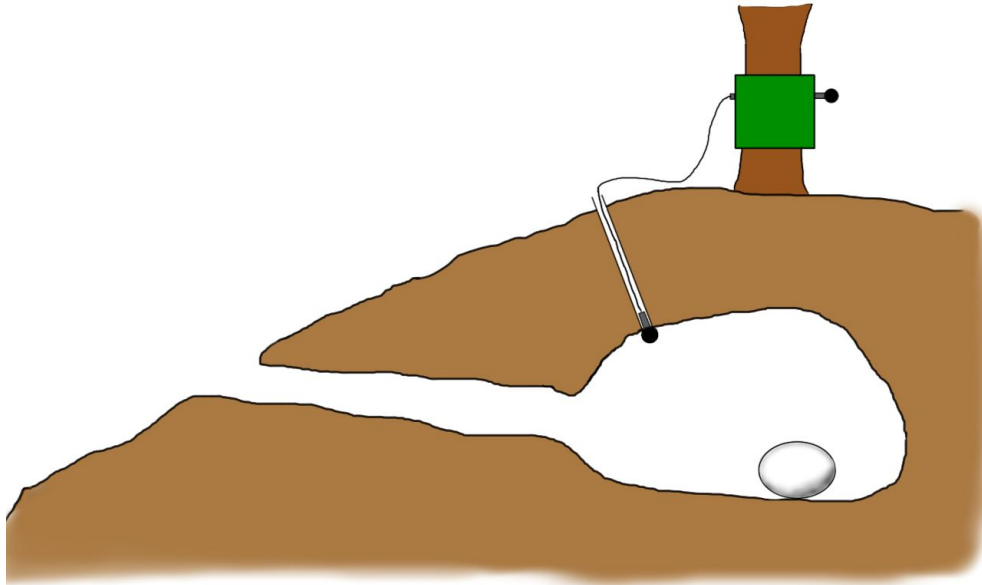
Figure 3.A.3: Nest acoustic monitoring setup. A Song Meter 2 was placed near to the nest, with the external microphone attached directly to the recorder, and an internal nest microphone placed inside the nest chamber and connected via an extension cable.

The resulting output contained the time of each syllable, and the presence and type of NLP or two-voicing in each.

The timing of the chick and male entering and exiting the nest was assessed acoustically in the first season. The internal and external microphones enabled the direction of movement (entry or exit) to be established. Whether the male or chick was entering or leaving was determined by the amplitude of nest litter rustling, whether the kiwi could be heard camouflaging the entrance upon exit, and by the vocalisations inside and outside the nest afterwards. In the subsequent two seasons motion-activated cameras were also used to monitor activity at the nest, and verified that the acoustic assessment was accurate.

3.A.3 *Results*

The four eggs in three nests all hatched successfully (Table 3.A.2). Microphone failures were common, including three in one nest within five weeks. This appeared to be due to water damage. The 'pipping' prior to hatching was not detected for any nest, but hatching was detected acoustically in two nests, and not in the third as a result of microphone failure. Hatching was indicated by either the sound of the chick struggling to emerge from the egg, or by faint calls from the chick. In the

second season hatch dates agreed with those determined by 'chick timer' telemetry tags (Taylor et al., in prep.). The first repeated calls occurred between two and five days after hatch.

Table 3.A.2: Hatch and first call dates of LSK chicks in nests acoustically monitored. The date of first call refers to the first repeated call from the nestling; occasional faint, single syllable calls were usually heard prior to this.

Season	Male band	Hatch date	First call
2010–2011	O-32668	29/Dec/2010	03/Jan/2011
2011–2012	O-31403	20/Nov/2011	25/Nov/2011
2011–2012	O-31659	01/Feb/2012	03/Feb/2011

Vocalisations were recorded from chicks in all nests. Nestlings were frequently vocal, with the vast majority of calls being simple frequently-modulated 'whistles', not dissimilar to male adult calls (Figure 3.A.4). Non-linear phenomena (NLP) were common in these syllables, with deterministic chaos, subharmonics, frequency jumps, and biphonation all observed (Figure 3.A.5 – Figure 3.A.9). Two biphonation types were recorded: 'side-band' biphonation (Figure 3.A.7), and biphonation with independent fundamental frequencies (Figure 3.A.6), which is indistinguishable from two-voicing. The nestlings made loud and rapid calls when the male returned to the nest, and these calls contained a high proportion of non-linear features. Qualitative inspection also indicated that NLP decreased in occurrence which chick age. A number of variations to the standard syllable type were found, including 'moans', 'squeaks' that were lower frequency with rich harmonic structure (Figure 3.A.10).

3.A.4 *Discussion*

Nestling LSK are more vocal and have a wider repertoire of calls than adults. Their vocalisations feature a main syllable type, but contain several variations of this. Non-linear phenomena are very common in LSK chick calls, with the four principal variants (§ 3.1.2) all evident. Frequency jumps and biphonation with independent f_0 (or two-voicing), which were not found in the calls of adults, occur frequently in those of nestlings. Deterministic chaos, rare in adults, is common in chick calls.

This prevalence of NLP in nestling calls is in accordance with their greater occurrence in young mammals (Wilden et al. 1998, Fitch et al. 2002). This may be due to the weaker vocal control in juveniles (Wilden et al. 1998), since NLP is thought
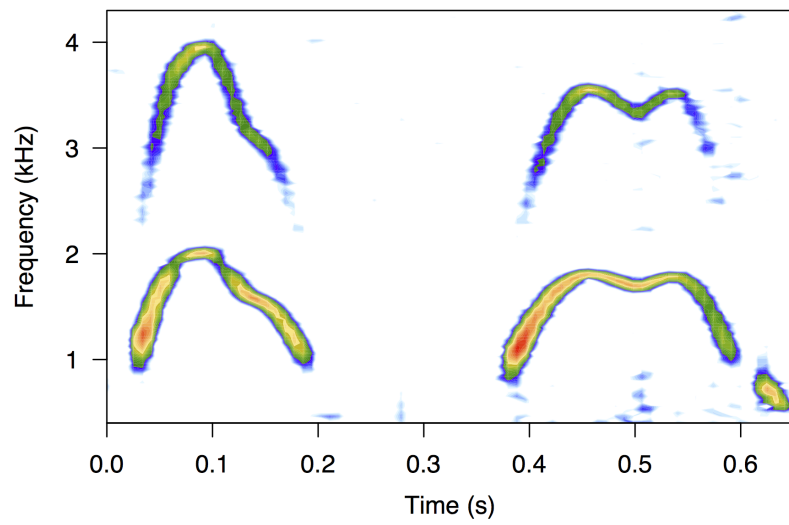
Figure 3.A.4: The typical LSK chick call structure: a series of repeated frequency-modulated syllables with harmonics.
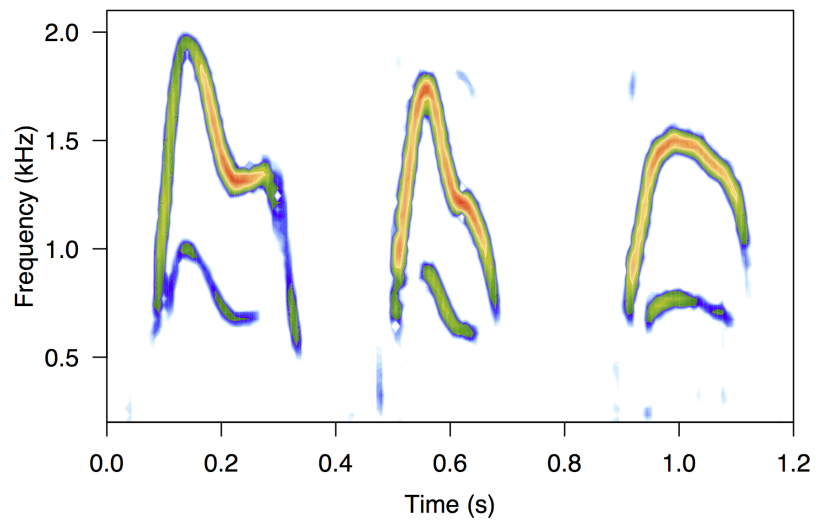
Figure 3.A.5: Subharmonics in an LSK chick vocalisation. Each syllable has a replicate at half the fundamental frequency.
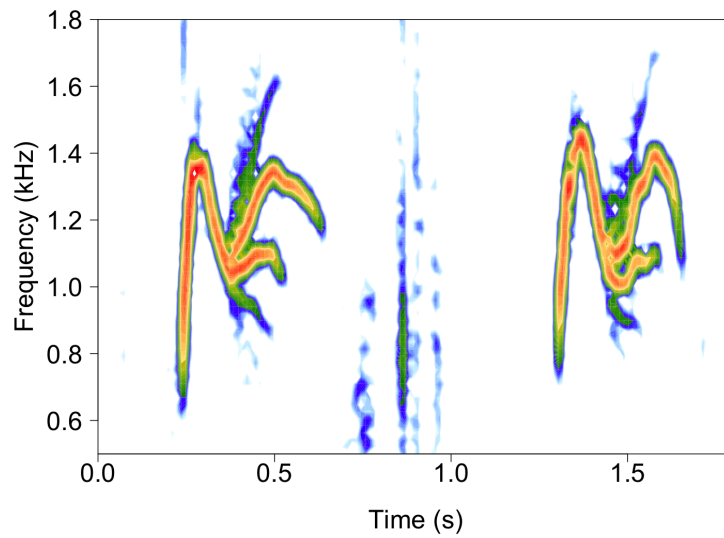
Figure 3.A.6: Biphonation or two-voicing in a chick call. A bifurcation during each syllable creates diverging independent fundamental frequencies. These could represent biphonation or two-voicing, which are indistinguishable in spectrograms.
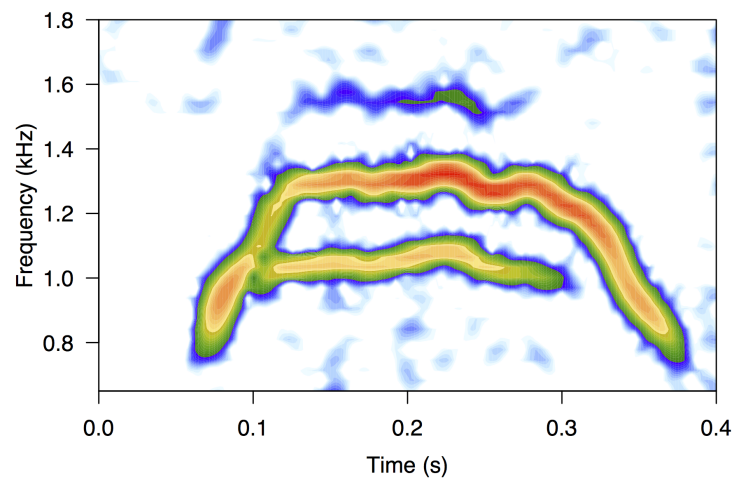
Figure 3.A.7: Biphonation with sidebands at ±300 Hz in the syllable of a nestling call.
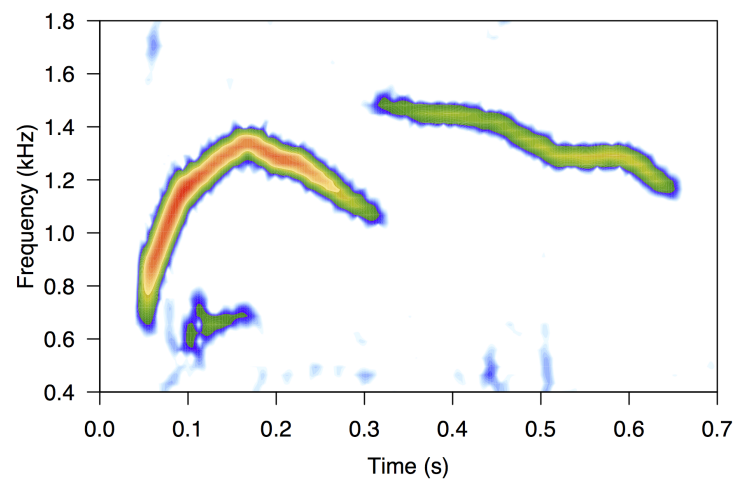
Figure 3.A.8: A frequency jump in a chick call. At 0.3 s the syllable frequency suddenly increases with no intermediate transition.
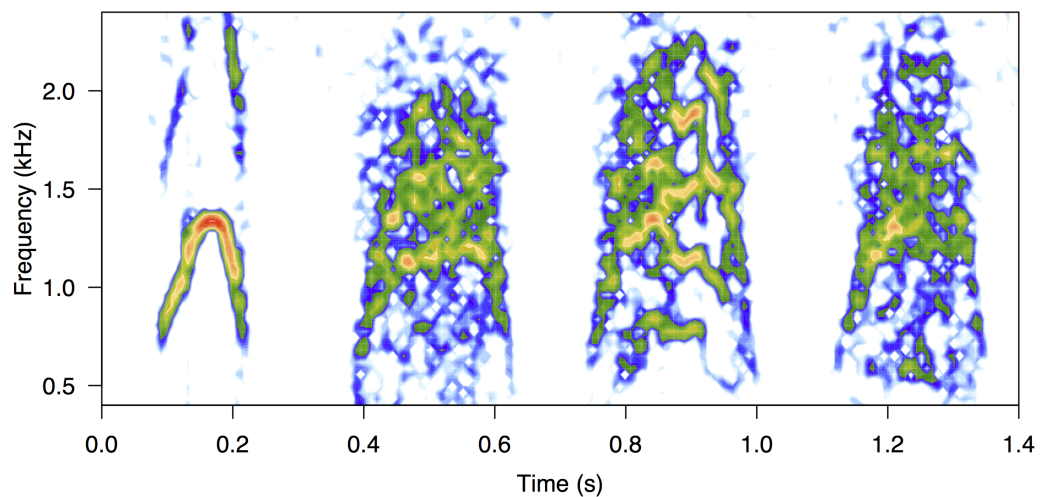
Figure 3.A.9: Deterministic chaos in a chick call. The first syllable is a typical LSK, but the following three all show signs of chaotic structure. Note the residual harmonic structure, which is particularly evident in the syllable from 0.8–1.0 s.
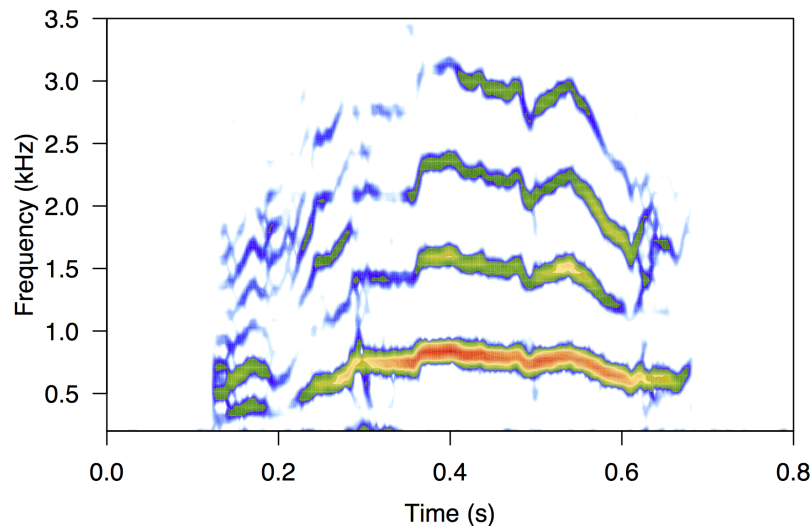
Figure 3.A.10: A low-frequency 'squeak' from an LSK nestling. The majority of syllables were as seen in Figure 3.A.4, but several different syllable types were found.

to occur in vocalisations at the physical limits of the vocal tract (Riede et al. 2007). The occurrence of NLP in LSK chicks is in agreement with the 'unpredictability hypothesis' for their function. Non-linearities are thought to increase the acoustic impact of vocalisations and prevent habituation (Blumstein & Récapet 2009, Townsend & Manser 2011). They are therefore expected to be used more often by animals in danger of being ignored, such as juveniles (Fitch et al. 2002).

The loss of vocal control and unpredictability hypotheses are supported by the large proportion of NLP in the sudden loud chick calls that occur when the adult male returns to the nest. The chick is clearly stimulated in these circumstances, but the functional significance of this signalling is unclear. Adults are not thought to feed their young in both LSK (Jolly 1989) and brown kiwi (McLennan et al. 2004), so the nestling is probably not begging for food. One possibility is that the chick's calls inform the adult that it is alive and healthy (although note that NLP can increase in pathological animals; Wilden et al. 1998, Riede 2000), so that the male remains in the nest. Male LSK share the nest burrow with the chick for up to three months after hatch (A. Digby, pers. obs.). Why this occurs is unclear, since the chick can feed itself, and nests commonly offer less shelter than typical burrows in brown kiwi and LSK (McLennan et al. 1987, Jolly 1989). This sharing may benefit the chick by providing warmth or protection from predators (McLennan et al. 1987, Colbourne 2002). There is clearly some advantage to this behaviour, which the chick may encourage with its repeated calls rich in non-linear features. Quantitative analysis of the chick calls recorded in this study is ongoing, and may yield further information on this relatively unknown aspect of LSK reproduction.

VOCAL IDENTIFICATION

ABSTRACT

Vocal discrimination occurs on a wide range levels in avian communication, and has important survival implications. The fitness of many birds depends upon the ability to distinguish mates, neighbours, offspring, parents or even individuals by call. This is particularly important when visual cues cannot be used, and so colonial-nesting, cryptic or nocturnal bird species often have individually distinctive calls. The ability to acoustically discriminate between individuals can provide important conservation benefits, such as in enabling correction for census biases and monitoring individual survival.

Acoustic distinction among individuals is likely to be important in the nocturnal and cryptic kiwi, yet has been assessed in just one species, the brown kiwi (*Apteryx mantelli*). This chapter tests the hypothesis that little spotted kiwi (*Apteryx owenii*; LSK) have individually distinctive calls. This was measured by the ability of three different classification methods (discriminant function analysis, support vector machines, and statistical classifiers) to assign identity using mean call or syllable temporal and frequency parameters.

The LSK in the study sample had low inter-individual call variation. Using call variables, only discriminant function analysis and support vector machines applied to males assigned identity better than that expected by chance, with successful classification achieved for at most 57% of male calls. Classification from syllable variables was far more reliable, with all methods identifying individuals better than chance for both sexes, and 71–79% of calls correctly classified. There was no significant difference in performance among the three classifiers. Classification accuracy decreased substantially with increasing numbers of kiwi.

These results suggest that LSK are likely to be able to distinguish individual conspecifics by call, but that their ability to do so is limited or relies upon more complex spectral features. The variation between the sexes and among individuals suggests that sexual selection pressures may shape vocal stereotypy, although it is feasible that LSK do not have a strong requirement for vocal discrimination. The variability in classifier performance restricts the potential use of vocal identification in conservation of these kiwi. However, the low call variation of LSK contrasts with that of brown kiwi (*Apteryx mantelli*), and may reflect the differences in genetic diversity between these two species. If confirmed, this could provide a powerful conservation

tool for rapidly assessing genetic diversity in kiwi, or for locating different genetic strains.
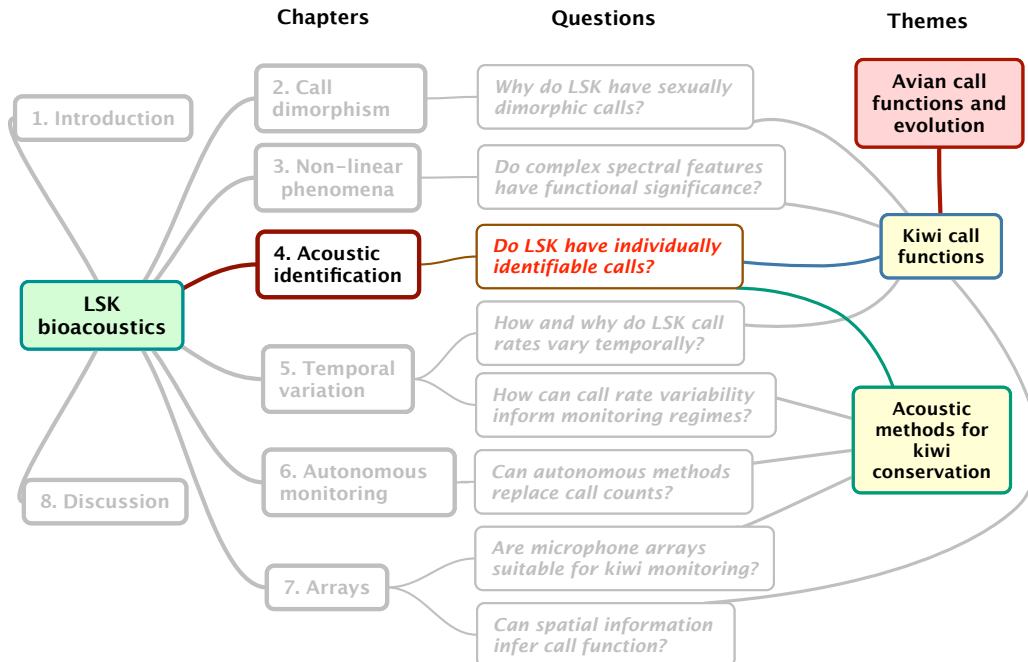
Topics covered by this chapter.

4.1 INTRODUCTION

4.1.1 *Function of vocal individuality*

Bird vocalisations enable discrimination at a wide range of levels. The species, sex, or exact identity of an individual can be discerned from calls, and kin, mates, neighbours, parents or offspring recognised acoustically. This discrimination is of vital importance to vocal species. Calls may help avoid hybridisation and inter-species competition in monomorphic species (Bolton 2007, McKown 2008, Curé et al. 2012), or enable mate recognition that is essential to successful reproduction (Aubin et al. 2000, Clark et al. 2006, Curé et al. 2009). Neighbour recognition prevents unnecessary competition in territorial birds (McGregor 1993, Clark et al. 2006). Vocal discrimination between parents and offspring ensures that parents feed only their young (Beecher 1991, Bradbury & Vehrencamp 1998), and that juveniles do not risk aggression by begging from other adults (Knudsen & Evans 1986, Pierotti 1991,

Ramos 2003). Acoustic identification of individuals may be used to assess reliability of signals such as alarm calls in social birds (Yorzinski et al. 2006, McDonald 2012).

The extent to which birds use vocalisations for discrimination purposes depends on their ecology. Parent-offspring recognition varies according to breeding behaviour, such as whether the species is a solitary or colonial breeder (Beecher et al. 1981, Stoddard & Beecher 1983, Beecher 1991) or whether parental care continues after fledging (Jones et al. 1987, Draganoiu et al. 2006, Klenova et al. 2009). Acoustic mate recognition occurs in passerines (Beletsky 1983, Wiley et al. 1991), but individually distinctive calls are most common in non-oscine colonial species which cannot rely on spatial cues for identification, such as penguins (Robisson et al. 1993, Aubin et al. 2000) and other seabirds (Jones et al. 1987, Mulard et al. 2009), bee eaters (Lessells et al. 1995) and swallows (Beecher et al. 1981, Stoddard & Beecher 1983). This dependence on breeding system is most clearly illustrated in penguins. Non-nesting penguins cannot use spatial information to find their mate, and so rely on vocal cues more than penguin species which have a fixed nest site. As a result, the non-nesting species have a more complex acoustic recognition system (Jouventin & Aubin 2002). Nocturnal birds are also less reliant on visual cues and so frequently can be individually identified by calls, as in owls (Galeotti & Pavan 1991, Otter 1996, Appleby & Redpath 1997b), frogmouths (Jones & Smith 1997) and oilbirds (which are also colonial nesters) (Suthers & Hector 1985, Suthers 1994). Similarly, cryptic birds living in dense habitat such as bitterns (Gilbert et al. 1994) and ant birds (Kirschel et al. 2009) also frequently display high intra-individual variation.

While many species exhibit individually distinctive vocalisations, the extent to which birds themselves discern and use these differences is often less clear. Experiments demonstrate the ability of birds to discriminate between individuals (Falls & Brooks 1975, McGregor 1993, Hyman & Hughes 2006), but it is often difficult to determine whether 'true' individual identification or classification on a broader scale (e. g. status or mate/non-mate) is taking place (Tibbetts & Dale 2007, McDonald 2012).

4.1.2 *Conservation applications*

Regardless of how it is used by animals themselves, vocal individuality can be an important tool in conservation. Acoustic methods are commonly used for assessing abundance (Dawson & Bull 1975, Rosenstock et al. 2002), and are particularly important for visually cryptic species (e. g. Redpath 1994, Luukkonen et al. 1997,

O'Donnell 2011). If the species has individually distinctive calls, then the ability to determine not just *how many* but *which* individuals are present can be used to quantify and correct for census biases (Saunders & Wooller 1988, Terry & McGregor 2002, Terry et al. 2005). Even if individual vocal variation is low, it can still be used to assess a minimum population size if combined with temporal or spatial caller information (Maurer et al. 2008).

The application of vocal individuality to censuses relies on *discrimination* – the ability to distinguish between individuals at a given point in time. If the vocal differences persist over time, then this enables *identification* (Terry et al. 2005). Vocal identification can provide life history or behavioural traits that are important in conservation management of threatened species, such as assessing mate retention or territory dynamics (Galeotti & Pavan 1991, Kirschel et al. 2011), or monitoring survival (Galeotti & Sacchi 2001, Gilbert et al. 2002).

4.1.3   *Methods of acoustic identification*

Individuals can be acoustically identified using qualitative spectrographic assessment (Terry et al. 2001), but the majority of analyses use quantitative spectrographic measures. Differences among individuals in frequency and temporal variables is often demonstrated using ANOVA (e.g. Jones et al. 1987, McKown 2008) or linear mixed models (e.g. Maurer et al. 2008). Classification of individuals based on these variables is commonly performed with discriminant analysis (Yorzinski et al. 2006, Baldo & Mennill 2011, Robin et al. 2011, Kirschel et al. 2011). Other available classification methods used in acoustic identification include artificial neural networks (Terry & McGregor 2002, McKown 2008), support vector machines (Cheng et al. 2012), self-organising maps (Kirschel et al. 2009), and hidden Markov Models (Kirschel et al. 2009, Cheng et al. 2012). These methods are also used for acoustic species discrimination (Acevedo et al. 2009).

Birds can also be identified individually through two-voicing (Chapter 3). This phenomenon is found in many species (Weisman et al. 1990, Zollinger et al. 2008, Krakauer et al. 2009), and is common in those that are nocturnal or colonial nesters (Suthers 1994, Krull & Hauber 2010). It serves an identification function in penguins, creating a stereotypic 'beat' that is unique to individuals and propagates effectively (Aubin et al. 2000, Aubin & Jouventin 2002).

4.1.3.1 *Call-type dependent vs independent methods*

The performance of many vocal identification methods varies with sample size. Errors usually increase with the number of individuals, and error rate can vary considerably depending on the combination of individuals included in the classification (McKown 2008). A more fundamental problem is that the methods described above are all call-type dependent, applicable to only a subset of calls of an individual's repertoire. These techniques require that all individuals in a species use similar calls (call sharing), and necessitate a different classifier for each call type, which is problematic for species such as song birds which have large vocal repertoires (Fox 2008a). Furthermore, methods such as discriminant analysis and support vector machines can only classify from a known 'library' of calls, and therefore are unable to distinguish unknown individuals (Terry et al. 2005). Call-type dependent methods are also reliant on vocal stability. Most studies of vocal identification are short term, and few have demonstrated stability over a number of years (Saunders & Wooller 1988, Gilbert et al. 2002, Klenova et al. 2009). However, animals frequently change vocalisation structure, as a result of season (Gilbert et al. 1994), body condition (Marks et al. 2010), social context (Geberzahn et al. 2009), developmental state (Klenova et al. 2009), or call learning (Kroodsma 2004).

More generic classifiers are able to distinguish individuals by their 'voice', irrespective of the signal content or structure. This appears to be how animals discriminate using vocalisations: humans identify each other by voice, regardless of the words spoken (Fox 2008a), and male great tits are able to distinguish other males based on unfamiliar songs (Weary et al. 1992). Averaged call characteristics can be used to discriminate between individuals (Weary et al. 1990), but such an approach is relatively inefficient. Short-term vocal features provide a more efficient approach. For example, formants in mammal and bird vocalisations reflect the shape of the vocal tract that is unique to each individual (Suthers 1994, Fitch et al. 2002, Fitch & Fritz 2006). In human speech recognition, vocal features such as mel-frequency cepstral coefficients (MFCCs) are extracted from the vocalisation, and then assigned to an individual with a trained classifier (Fox 2008a). The application of these methods to animals is in its infancy, but they have been successfully applied to birds (Fox 2008b, Cheng et al. 2010, 2012). Such call-type independent classification methods offer considerable benefits, since they are not species-specific, do not require many call examples, can classify previously-unknown individuals, do not require call type separation, and can utilise automatic feature extraction and classification methods (Fox 2008a).

Another method for distinguishing between individuals are non-linear phenomena (NLP; Chapter 3). These complex vocalisations reflect the structure of the vocal tract, and as such are call-type independent and unique to the individual. NLP have been shown to serve an identification function in canids (Tooze et al. 1990, Mitchell et al. 2006, Volodina et al. 2006). Analysis in birds is less developed, but chaos in magpie fledgling calls may provide identification information (Suthers et al. 2010).

4.1.4 *Vocal individuality in kiwi*

As nocturnal species likely to be strongly reliant on vocal communication, kiwi may be able to discriminate between individual conspecifics using calls (Corfield et al. 2008a). This would provide substantial benefits to kiwi conservation research and management, through the ability to observe individual life histories, and through reducing bias of call censuses, which are one of the main monitoring tools for kiwi (Colbourne 2006).

Two studies of different populations have reported individual differences in call spectral structure of brown kiwi (*Apteryx mantelli*; Corfield 2005, Hojem 2006). The calls of this species appear to be highly stereotyped, with Corfield (2005) correctly classifying > 85% of calls to individuals using discriminant analysis. However, acoustic individual identification of kiwi has been conducted in no other species. With marked differences in ecology (Castro 2011) and genetic diversity between species (Ramstad et al. 2010), vocal individuality needs to be assessed in each before it can be widely utilised as a conservation tool.

In this chapter the potential for vocal recognition is assessed in LSK. Although they have the lowest threat classification of all kiwi species (Miskelly et al. 2008, IUCN 2011), there is strong concern over the genetic diversity of LSK (Miller et al. 2011), with a few individuals thought to have dominated recruitment in some populations (Ramstad et al. 2013). The ability to identify these individuals by call would therefore be valuable in conservation management. Furthermore, brown kiwi exhibit strong individual call differences, but have higher genetic diversity than LSK, which have the lowest genetic diversity among kiwi (Ramstad et al. 2010). Call variety can reflect genetic diversity (Araya-Ajoy et al. 2009, Robin et al. 2011), so evaluating vocal variation in LSK will determine whether vocal individuality is likely in the remaining three kiwi species. By comparing relationships between vocal and genetic diversity among kiwi populations and between species, call variation may even prove useful as an initial measure of genetic diversity.

To assess individual vocal differences in LSK, a variety of call-type dependent classification methods are applied in this chapter to high-quality calls from a single population. Call-type dependent techniques are appropriate for LSK, since they only have one call type (Chapter 2), and as long-lived birds (Robertson & Colbourne 2004) that are unlikely to be call learners (Kroodsma 2005), their calls are expected to be temporally stable. The research hypothesis is that LSK will have individually distinctive calls, with the prediction that using frequency and temporal variables, the classifications methods will correctly assign identity better than expected by chance.

4.2   METHODS

4.2.1   *Recording methods and data*

Calls were sampled from a single LSK population in Zealandia, a 225 ha fenced reserve in Wellington, New Zealand. The population was founded from 40 individuals translocated from Kapiti Island in 2000–2001. Recordings were made on 140 nights between May 2009 and April 2012. Recording equipment was a Sound Devices 722 digital recorder and a Telinga Twin Science microphone with a Telinga 53 cm foldable parabola. Calls were digitised at 48 kHz sample rate and 16-bit precision. From April 2010 to the end of the study between 15 and 52 LSK were fitted with radio transmitters for a census (Hugh Robertson, Department of Conservation), and a separate study (Helen Taylor, Victoria University of Wellington). Radio telemetry using these transmitters or colour reflective leg bands were used to identify calling kiwi. Spontaneous calls were recorded, but playback was also used to elicit calls from known individuals. To avoid habituation or excessive disturbance, playback was only used at most once per week per kiwi, and only outside of the September – January breeding season (Jolly 1989).

A total of 775 calls were recorded, and assigned a five-level subjective quality score from spectrogram inspection, based on signal to noise and sound interference from other sources. Only calls in the two highest-quality levels from positively-identified adults (§ 2.2.3) were used for analysis, resulting in a sample of 100 calls from 16 individuals (ten male, six female).

4.2.2 *Spectral analyses*

Spectrographic analyses were conducted in Raven Pro 1.4 (Charif et al. 2010), with a 1024-sample Hann window, 50% overlap, and 23.4 Hz frequency resolution. A series of energy-based frequency and temporal measures (Table 4.2.1) were made for each syllable within each call. Analyses of these measures were conducted in R (version 2.15; R Development Core Team 2011).

Table 4.2.1: Energy-based spectral measures of syllables and calls in the individual identification dataset.

Syllable variable	Description
5% frequency	Frequency of the 5th energy percentile
95% frequency	Frequency of the 95th energy percentile
Q1 frequency	Frequency of 1st energy quartile
Q3 frequency	Frequency of 3rd energy quartile
centre frequency	Frequency of the 50th energy percentile
maximum frequency	Frequency of the maximum energy
90% bandwidth	Frequency span between 5th and 95th energy percentiles
IQR bandwidth	Frequency span of inter-quartile energy range
90% duration	Duration between the 5th and 95th energy percentiles
IQR duration	Duration of inter-quartile energy range
90% syllable gap	Gap between syllables, measured between consecutive 95th and 5th energy percentiles
IQR syllable gap	Duration of interval between consecutive Q3 and Q1

Call variable	Description
mean 5% frequency	Mean of 5% frequency.
	(Repeated for all syllable variables)
syllable rate	Number of syllables divided by call length
call duration	Time from 5th percentile of first note to 95th of last
duty cycle	Sum of syllable durations divided by sum of syllable durations and gaps

4.2.3   *Individual variation*

4.2.3.1   *Variation within calls*

Syllable frequency and temporal properties vary systematically during LSK calls (Chapter 2), and also exhibit random variations at the call start ('priming'), and occasionally at the call end. Both of these types of changes have the potential to affect discrimination (Jones & Smith 1997, Corfield 2005). Plots of each variable against syllable number from the start of the call were used to identify the portion of the call suitable for subsequent analyses, and to assess which variables showed most differences among individuals.

4.2.3.2   *Coefficients of variation*

Coefficients of variation were used to identify parameters that would provide most discrimination among individuals. The ratio of among-individual variation to the mean of within-individual variation was calculated for all syllable and call variables. Variables with a ratio of greater than one indicated a greater variability among individuals than within, and so could potentially be used to distinguish individuals (Fitzsimmons et al. 2008a).

4.2.3.3   *Principal component analysis*

Principal component analysis (PCA) was employed to provide an initial assessment of differences between individuals across multiple spectral variables. Since observations should be independent for PCA (Zuur et al. 2007), it was only applied to syllable means, rather than for each syllable. Individuals with more than two high-quality calls were included in the sample. To avoid random fluctuations at the call start and end, only syllables between the 15th and 35th in each call were used (§ 4.3.1). Calls were removed if they subsequently contained fewer than ten syllables. The resulting sample contained 71 calls from nine males, and 19 calls from five females. Pauses in calls were removed by discarding syllables with an inter-syllable gap of more than two seconds. The number of variables was reduced using collinearity and coefficient of variation criteria. In cases of high correlation ($r \geq 0.8$), the variable with the lower coefficient of variation ratio was removed, leaving six variables for males and four for females (Table 4.2.2). One male (band O-32663) produced two unusual calls with very high mean frequencies (95% frequency $>4\,$kHz) due to

a high proportion of syllables containing high frequency sections (§ 2.B.2). These calls were abnormal for that individual and all others, and such outliers can detrimentally affect PCA, so these were removed from the dataset. PCA was applied to each sex separately due to the large frequency difference between males and females (Chapter 2).

Table 4.2.2: Call variables used in the PCA.

Sex	Call variables
Male	mean centre frequency, mean 5% frequency , mean 95% frequency, mean bandwidth, mean duration, syllable rate
Female	mean maximum frequency, mean bandwidth, mean duration, syllable rate

PCA was applied to the correlation matrix due to the differently-scaled input variables (Zuur et al. 2007), and was calculated with R package `prcomp`. Distance biplots were used to compare principal component scores among individuals.

4.2.4 *Classifiers*

Principal component analysis provides indication of difference between groups, but does not allow discrimination between them. Three different multivariate classification methods were tested for their ability to discriminate between individual kiwi from call parameters. These were discriminant function analysis, statistical classifiers, and support vector machines.

4.2.4.1 *Discriminant function analysis*

Discriminant function analysis (DFA) requires independent samples (Zuur et al. 2007), so could only be applied to call means. To avoid collinearity of input variables, the principal components derived in § 4.2.3.3 were used instead of the original call variables. It is recommended that the number of observations per group is at least 4–5 times the number of variables for discriminant analysis (Zuur et al. 2007). However, the constraints of ecological observational data mean that this is often relaxed in many studies (e.g. Corfield 2005, Maurer et al. 2008, Fitzsimmons et al. 2008a, Kirschel et al. 2011). The minimum is two observations per group, or a minimum group size equal to the number of variables, although this too is often violated (Zuur

et al. 2007). Therefore females with at least four calls, and males with at least six calls, were used in the samples, since these were the numbers of call variables used for the PCA (Table 4.2.2). This resulted in a sample of 66 calls from seven males, and 15 calls from three females. The sexes were analysed separately. Homogeneity of variance between groups and normality was ensured with boxplots and histograms. Analyses were carried out using function lda in R package MASS (Venables & Ripley 2002).

4.2.4.2 *Statistical classifiers*

Multi-parameter individual discrimination was also assessed using statistical classifiers, with software C5.0 (version 2.09; RuleQuest Research, St Ives, NSW, Australia. http://www.rulequest.com/see5-info.html). This builds classifiers with either decision trees (hierarchical rules) or rulesets (unordered rules). Classification was performed on the same PCA dataset as used for discriminant function analysis. Machine learning algorithms such as decision trees and rulesets do not have the independence requirement of DFA, so classification was also performed using syllable measures (Table 4.2.1). All individuals in the high-quality sample were used for this, resulting in a sample of 1245 syllables from ten males, and 300 syllables from six females.

Table 4.2.3: Syllable variables use for both sexes in classification with C5.0 software.

Syllable variables	Syllable number, centre frequency, 5% frequency, 95% frequency, 90% bandwidth, 90% duration, 90% syllable gap

Classifiers were constructed using decision trees and rulesets, with and without softening for the decision tree. Softening converts 'sharp' rule thresholds in decision trees into 'fuzzy' boundaries, which can improve classification accuracy for some datasets. Adaptive boosting was applied to all models. This is an iterative process, generating a classifier at each step which uses the errors of the previous classifier to reduce the error rate.

Classifier accuracy usually varies with the number of classes (e.g. McKown 2008). This was assessed by applying the best classifier to a subset without replacement of n individual kiwi, where n ranged from two to the full number of individuals. This was applied 20 times to each value of n, and the mean proportion of calls correctly classified was then plotted against n.

4.2.4.3 *Support vector machines*

Support vector machines (SVMs) classify by using kernel techniques to transform the data into a higher dimensional space so class boundaries become linear Cortes & Vapnik 1995. They have been shown to work effectively for acoustic identification of birds (Acevedo et al. 2009). SVMs were applied to the principal component dataset used for DFA, and the same syllable variables used for C5.0. Function `ksvm` in R package `kernlab` (Karatzoglou et al. 2004) was used. A Gaussian radial base function was used, and the $C - SVM$ algorithm. The cost parameter C was optimised using package `caret` (Kuhn 2008).

4.2.4.4 *Comparing classifiers*

Since sample sizes were too small to split into separate training and testing datasets, the performance of each classifier was assessed with ten repeats of ten-fold cross-validation using `caret`. Comparisons were made using the unweighted Kappa statistic, which is a measure of accuracy corrected for the probability of successful classification by chance (Cohen 1960). This provides a significant advantage over the more usual measure of accuracy for unbalanced datasets (Carletta 1999, Sim & Wright 2005). There is concern over the use of a single statistic to quantify multiclass classifications (Kraemer et al. 2004), so the classifiers were also compared with the area under the receiver operating characteristic (ROC) curve (Kuhn 2008). The ROC curve is a plot of the true positive rate against the false positive rate, and was constructed using a one-versus-all method for each kiwi.

4.3 RESULTS

4.3.1 *Variation within calls*

Frequency variables showed substantial fluctuation at the beginning and end of calls in both sexes. From assessment of all high-quality calls, only syllables between the 15th and 35th were used for subsequent analyses (Figure 4.3.1). The variables 95% frequency and 90% bandwidth showed most promise for distinguishing among individuals when variation within the call was considered (Figure 4.3.2).
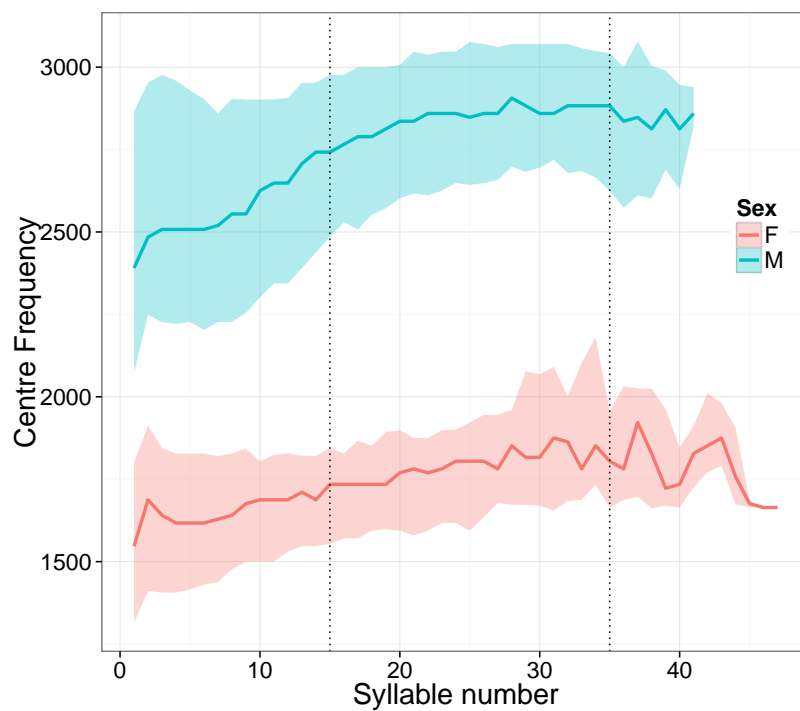
Figure 4.3.1: Change of frequency during male and female LSK calls. Solid lines show median syllable central frequency over all calls in the high-quality dataset. Shaded areas are 95% range. Frequency variations are less between the dashed lines, so only syllables within these limits were used for individual discrimination analyses.
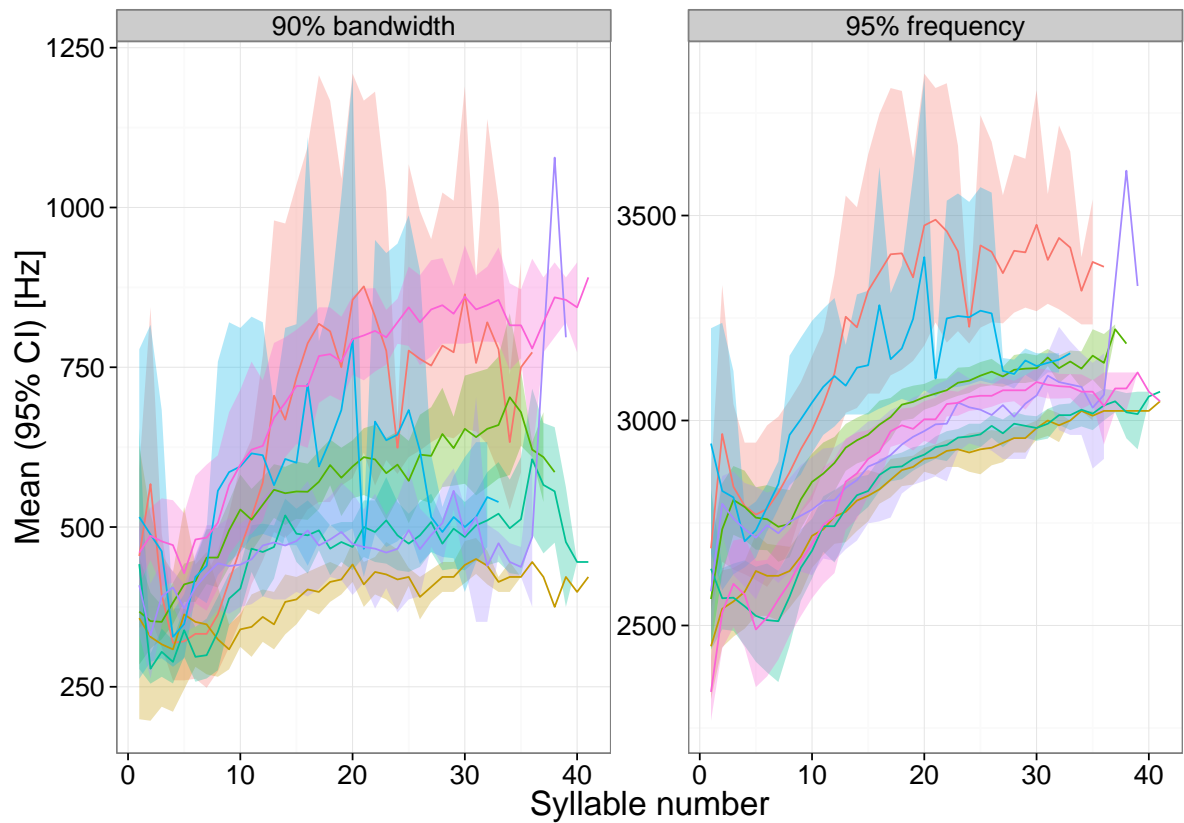
Figure 4.3.2: Change of frequency parameters during calls, for individual male LSK with at least five calls in the high-quality dataset. Lines show the mean syllable values for each kiwi, and shaded areas are 95% confidence intervals about this mean. Although two males show large fluctuations, the others have good consistency of frequency parameters between calls (after priming at the call start) and little overlap among individuals in the last half of the call.

4.3.2 *Coefficients of variation*

In male LSK, frequency variables had the highest coefficient of variation ratios, while the temporal variables ratios were approximately one. This indicated that individuals differed more in frequency than temporal call structure (Figure 4.3.3). Females showed a similar pattern, except that syllable duration also differed between individuals, whereas it did not in males. Measures based on the 5th and 95th percentiles tended to show more discrimination power than those using the first and third energy quartiles (Q1 and Q3), so the latter were dropped from all further analyses.
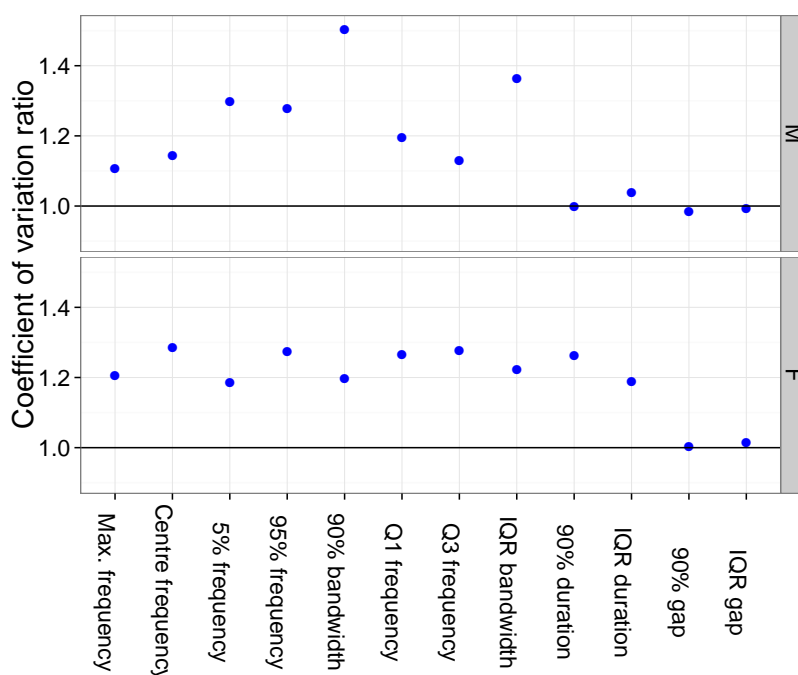
Figure 4.3.3: Coefficient of variation ratios for syllable variables in male and female calls. A ratio of greater than one indicates that there is greater variation among than within individuals, suggesting that the variable has the potential to discriminate individuals.

Call variables showed greater discrepancies between the sexes. In males all except mean syllable duration had a ratio of greater than one, whereas in females the reverse applied: only the ratio for mean syllable length was substantially greater than one (Figure 4.3.4).
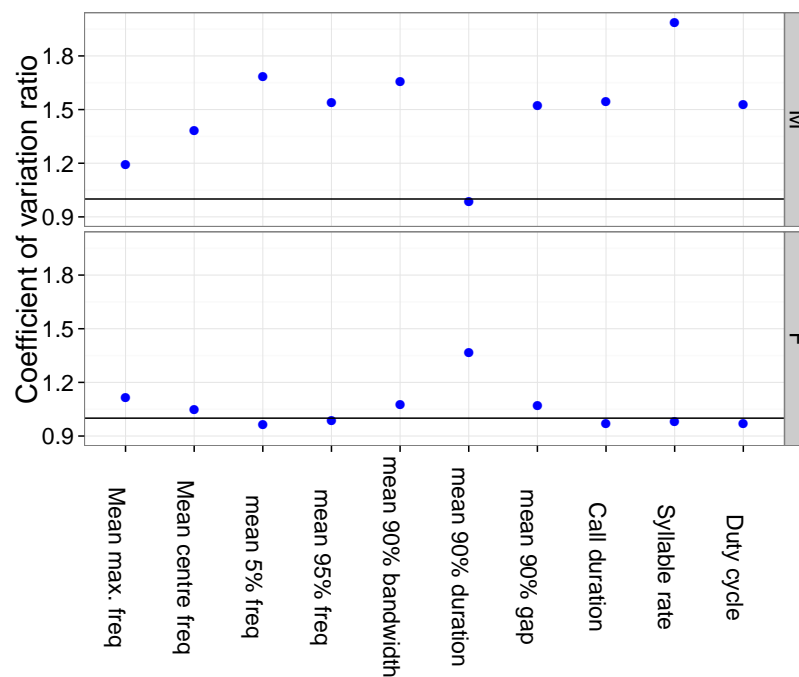
Figure 4.3.4: Coefficient of variation ratios for call variables. There is substantial difference in variable ratios between the sexes. Only one variable, mean syllable duration, shows potential for discrimination between females.

4.3.3 *Principal component analysis*

The first four principal components accounted for at least 90% of the variation for both males and females, and the first two components contributed more than 60% of the variance in each case (Table 4.3.1). There was substantial overlap between most individuals in principal component space (Figure 4.3.5 & Figure 4.3.6). However, some subsets of males occupied disparate regions of the biplot.

4.3.4 *Discriminant function analysis*

Discriminant function analysis was applied to 66 calls from seven males, and 15 calls from three females. The first three discriminant functions contributed more than 95% of the trace for males, and only two discriminants provided all trace variation in females. While the calls of each individual were mostly grouped together in discriminant function space, there was substantial overlap of these groups (Figure 4.3.7 & Figure 4.3.8).
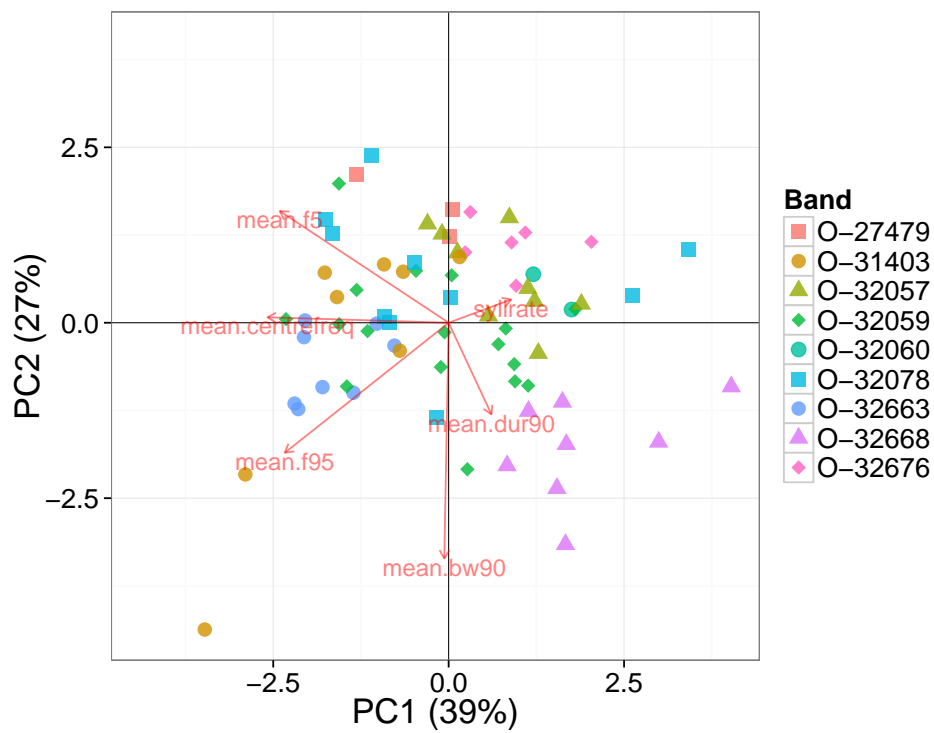
Figure 4.3.5: Distance biplot of the first two principal components for male calls. Individuals are identified by leg band numbers. Calls from each individual are broadly grouped, but there is substantial overlap between the locii of points for each. This indicates that low inter-individual variation in male LSK calls. Abbreviations: f5 = 5% frequency, syllrate = syllable rate, dur90 = 90% duration, bw90 = 90% bandwidth, f95 = 95% frequency.

Table 4.3.1: Principal components of call variables for male and female LSK in the high quality call sample.

Males	PC1	PC2	PC3	PC4	PC5
Standard deviation	1.528	1.262	0.982	0.950	0.453
Proportion of Variance	0.389	0.266	0.161	0.150	0.034
Cumulative Proportion	0.389	0.655	0.815	0.966	1.000

Females	PC1	PC2	PC3	PC4
Standard deviation	1.468	1.080	0.808	0.157
Proportion of Variance	0.539	0.291	0.163	0.006
Cumulative Proportion	0.539	0.831	0.994	1.000
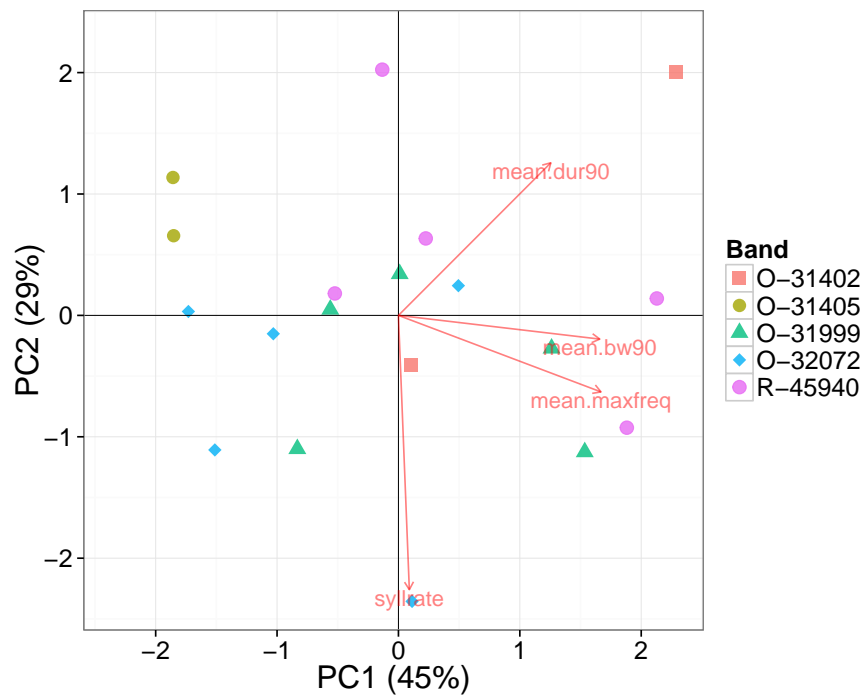
Figure 4.3.6: Distance biplot for female calls. There is only poor grouping and large overlap between individuals. Abbreviations as Figure 4.3.5.
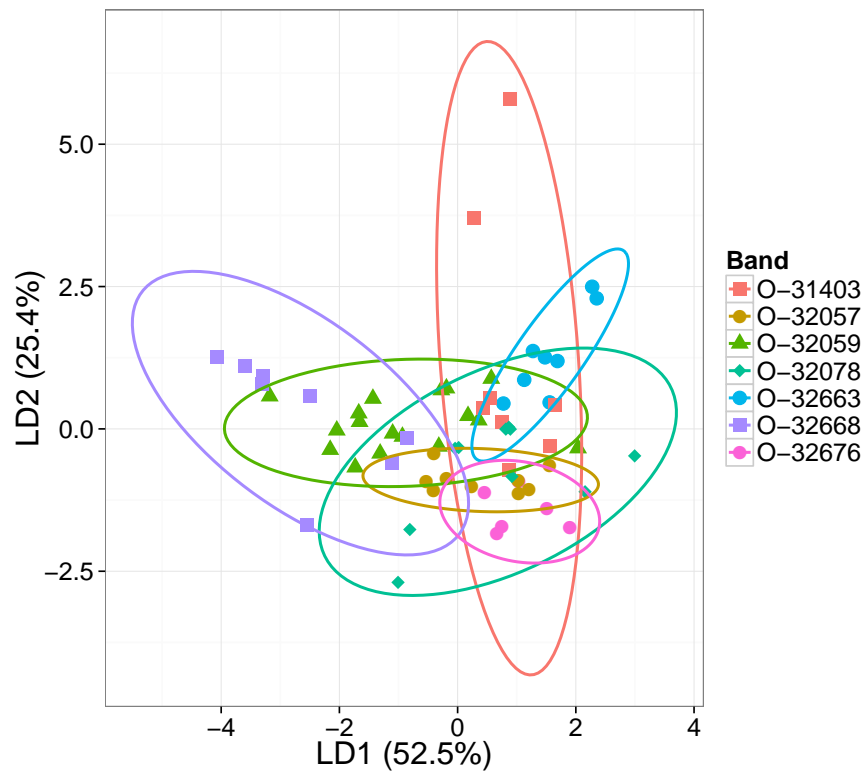
Figure 4.3.7: Scatterplots of discriminant functions for male LSK, derived from PCA on call means. Ellipses show 95% confidence intervals about the mean. Calls from each individual are mostly grouped together, but these locii overlap, and are large for some individuals. This suggests that individual identification accuracy from discriminant function analysis will be low overall, but variable between individuals.
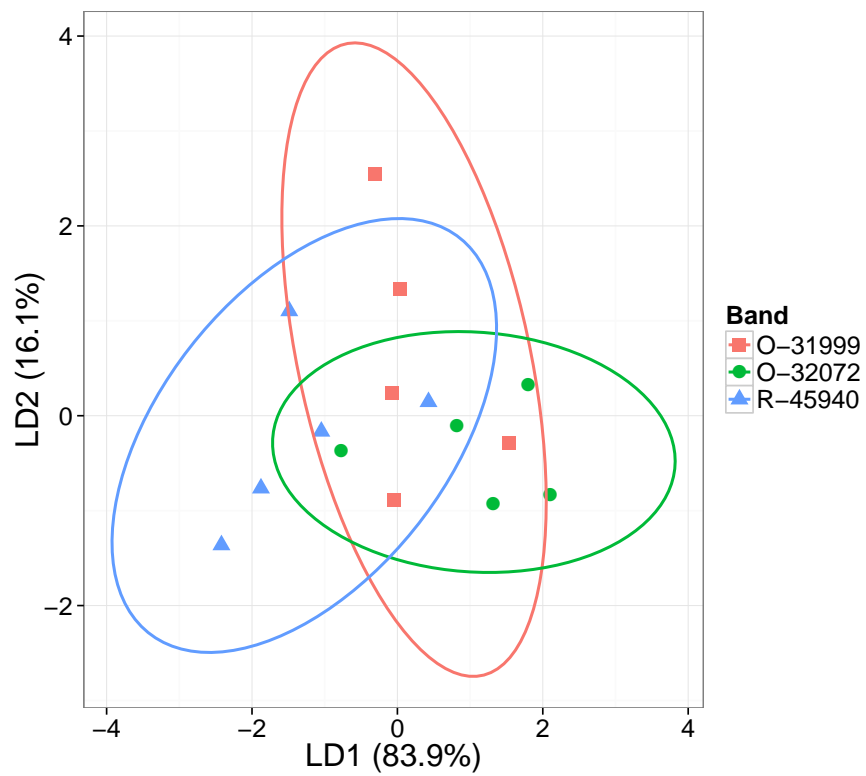
Figure 4.3.8: Scatterplots of discriminant functions for female LSK. Confidence intervals are large, and exhibit substantial overlap, even for a small number of individuals.

Cross-validation showed that discriminant function analysis accuracy varied substantially between individuals. While some individuals had relatively high precision and recall (Box 5.1, page 135), others had low values (Figure 4.3.9). The mean precision and recall were 52.9% and 56.8% for males, and 29.1% and 28.1% for females.

4.3.5 *Statistical classifiers*

There was very little difference in performance between decision trees with and without softening, so the values without softening are reported. For both sexes the syllable classifiers performed substantially better than the call classifiers (Table 4.3.2). Rulesets marginally outperformed decision trees. For the ruleset classifiers, precision and recall ranged substantially between individuals (Figure 4.3.10 & Figure 4.3.11).

Table 4.3.2: Performance measures for the statistical classifiers from C5.0 software, for both sexes, and applied to both call and syllable variables. Precision and recall values are means across all kiwi.

Input variables	Method	Sex	Kappa	Precision	Recall
Call	Tree	M	0.335	43.9	43.7
Call	Ruleset	M	0.351	44.5	45.9
Call	Tree	F	-0.011	20.6	18.9
Call	Ruleset	F	-0.033	19.2	19.2
Syllable	Tree	M	0.711	71.7	72.7
Syllable	Ruleset	M	0.717	72.8	72.2
Syllable	Tree	F	0.594	70.3	71.6
Syllable	Ruleset	F	0.610	72.9	71.4

The overall best performing classifier was a ruleset applied to syllables of male calls. As the number of males input to this classification method was increased, there was a strong decrease in classifier performance (Figure 4.3.12).

4.3.5.1 *Support vector machines*

SVM performance was highly variable among individuals, both for syllable and call classification (Figure 4.3.13 & Figure 4.3.14). As with C5.0, classification using
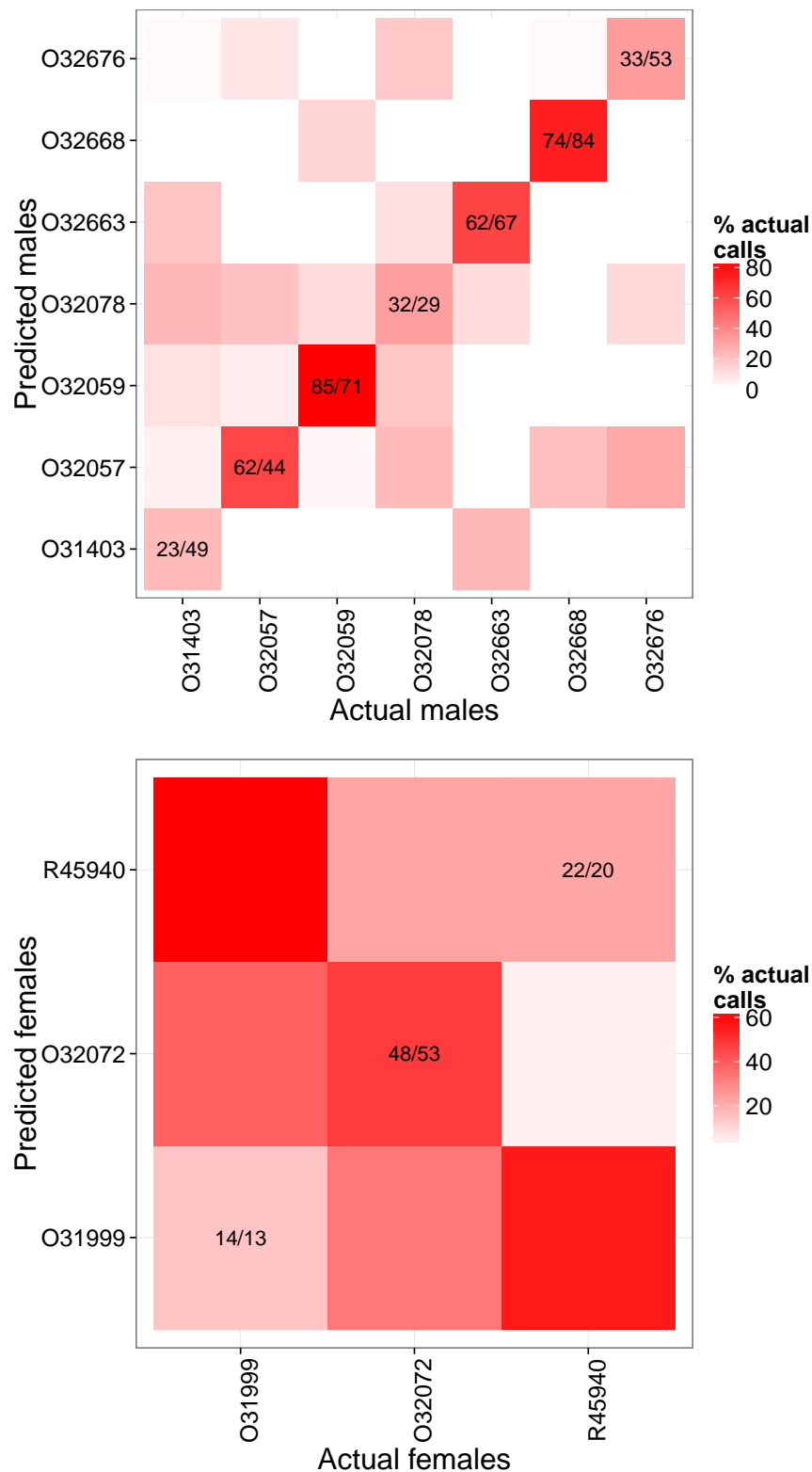
Figure 4.3.9: Confusion matrices for identification of male and female LSK by discriminant function analysis. Scores are evaluated by cross-validation, with numbers on the diagonal showing the precision/recall for each individual.
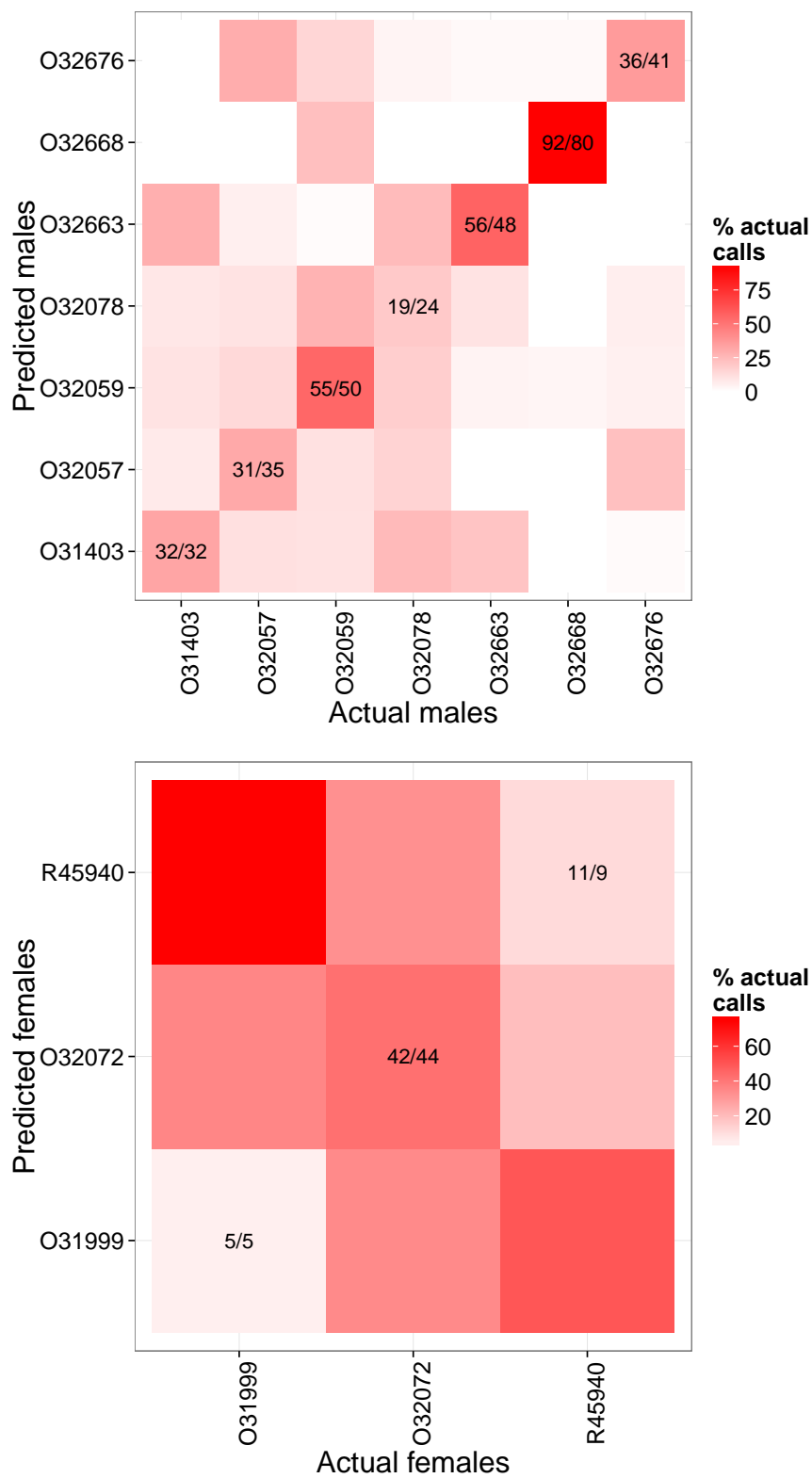
Figure 4.3.10: Confusion matrices for `C5.0` ruleset classifiers for call variables. Shading indicates the proportion of actual calls in each cell. Numbers on the diagonal show the precision/recall for each individual.
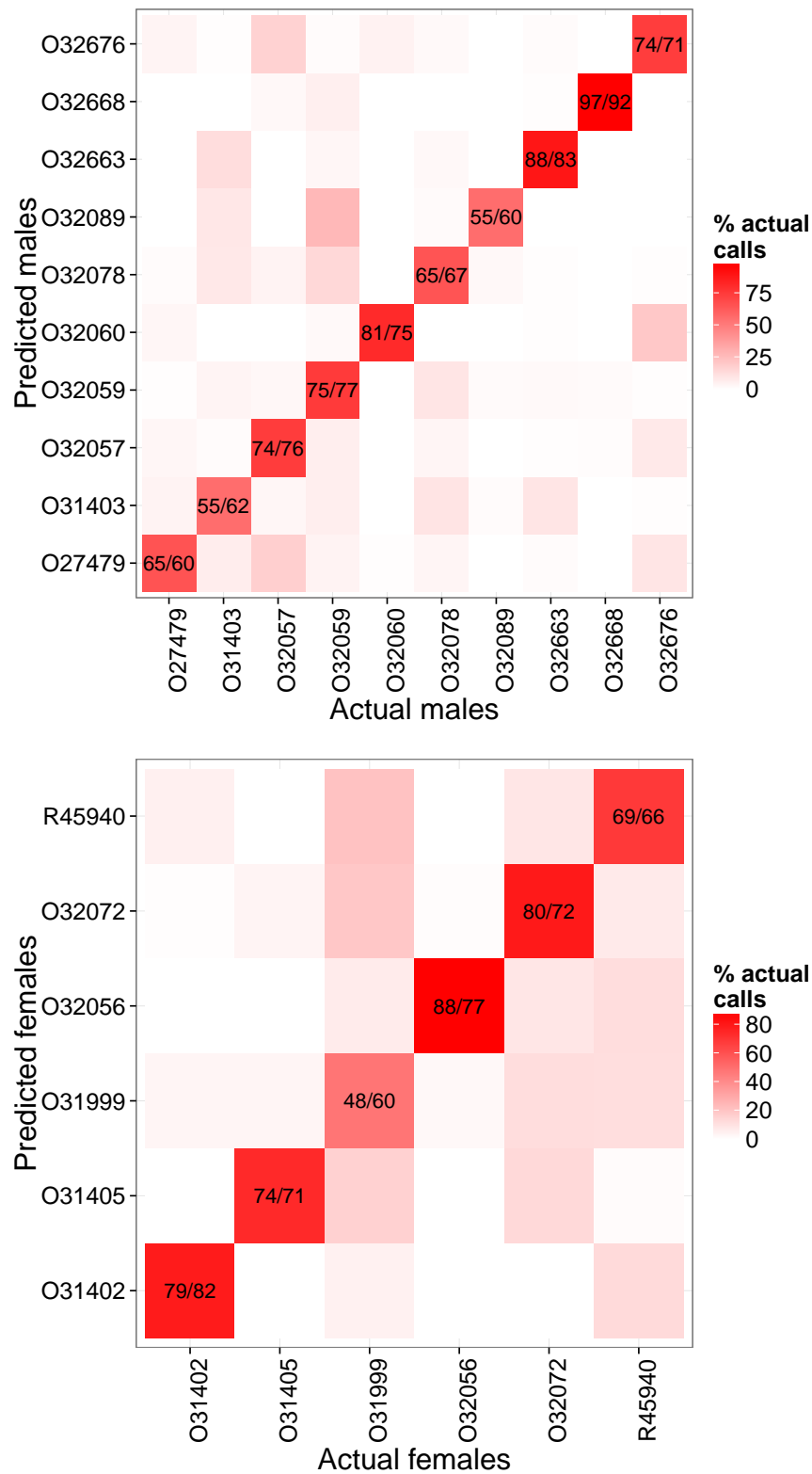
Figure 4.3.11: Confusion matrices for `C5.0` ruleset classifiers for syllable variables. Shading and annotations as Figure 4.3.10.
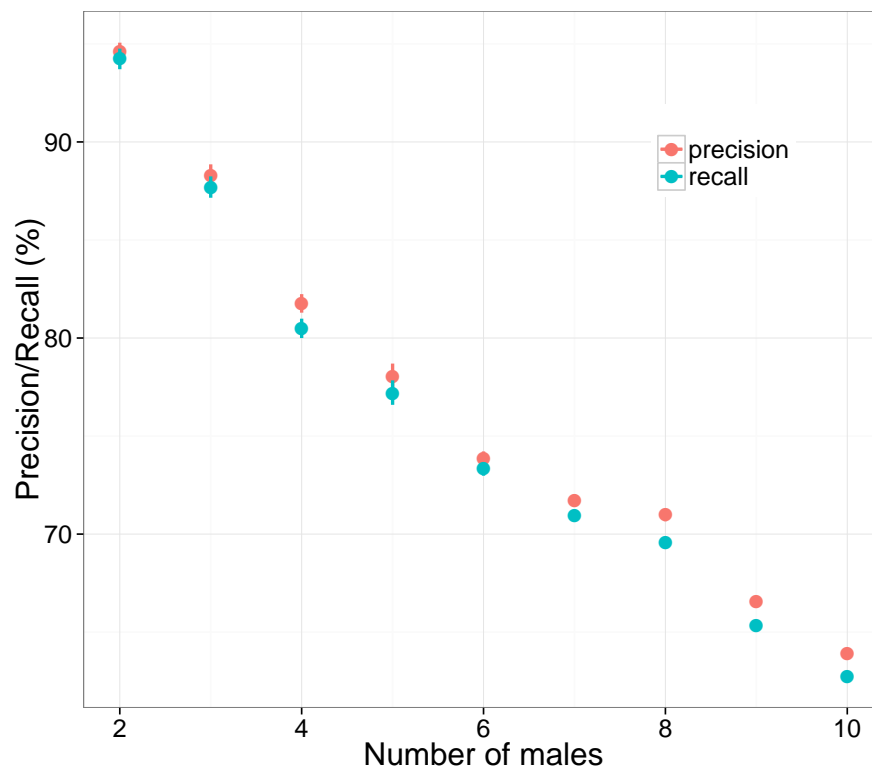
Figure 4.3.12: The change in classifier performance with sample size. The classifier is a `C5.0` ruleset applied to male syllables. Filled circles give the mean recall and precision against the number of individual kiwi in the sample. Error bars indicate 95% confidence intervals about the mean.

syllable variables was far better. For call variables, mean precision and recall were 48.6/47.3% for males, and 8.6/6.6% for females. For syllables, mean precision and recall were 65.8/70.5% for males, and 70.5/79.3% for females.

4.3.5.2 *Comparing classifiers*

The classifiers applied to syllable variables far out-performed those applied to call variable principal components, as measured by the Kappa and ROC statistics (Figure 4.3.15 & Figure 4.3.16). All methods were no better than chance at classifying females from call variables. Male classification from calls was more reliable than for females. There was no significant difference at the 5% level between each of the classification methods within the call and syllable variable datasets.

4.4   DISCUSSION

Principal component analysis demonstrated that individual LSK in the sample shared similar temporal and frequency call parameters. Calls from each individual tended to be grouped together in principal component space, although these groups mostly had large overlap between individuals. As a consequence, discrimination techniques were limited in their ability to identify individuals from calls. All methods were unable to discriminate females using mean call variables, and accuracies were low (Kappa < 0.5) for male call means. However, the `C5.0` and SVM classifiers achieved relatively high accuracies when applied to syllable variables. This was likely due to the ability to model the variation in frequency and temporal measures during the call (Chapter 2). There was no significant difference between the discriminant function analysis, support vector machine and statistical classifier (tree or ruleset) methods in classification performance. Accuracy varied substantially between individuals, with some much more distinct in the parameter space than others. As expected, the overall classification success decreased with the number of individuals in the sample.

These acoustic identification results can be considered from conservation or ecological perspectives. From a conservation point of view, the classifier performance compares poorly with results from other species. In this study, (unbalanced) accuracies were 57% for males and 28% for females using DFA. This contrasts with DFA accuracies of 77% for petrels (McKown 2008), 93% for antbirds (Kirschel et al. 2011), and 85% and 79% for male and female brown kiwi (Corfield 2005). However, the results for male LSK are comparable to those for other species, such as
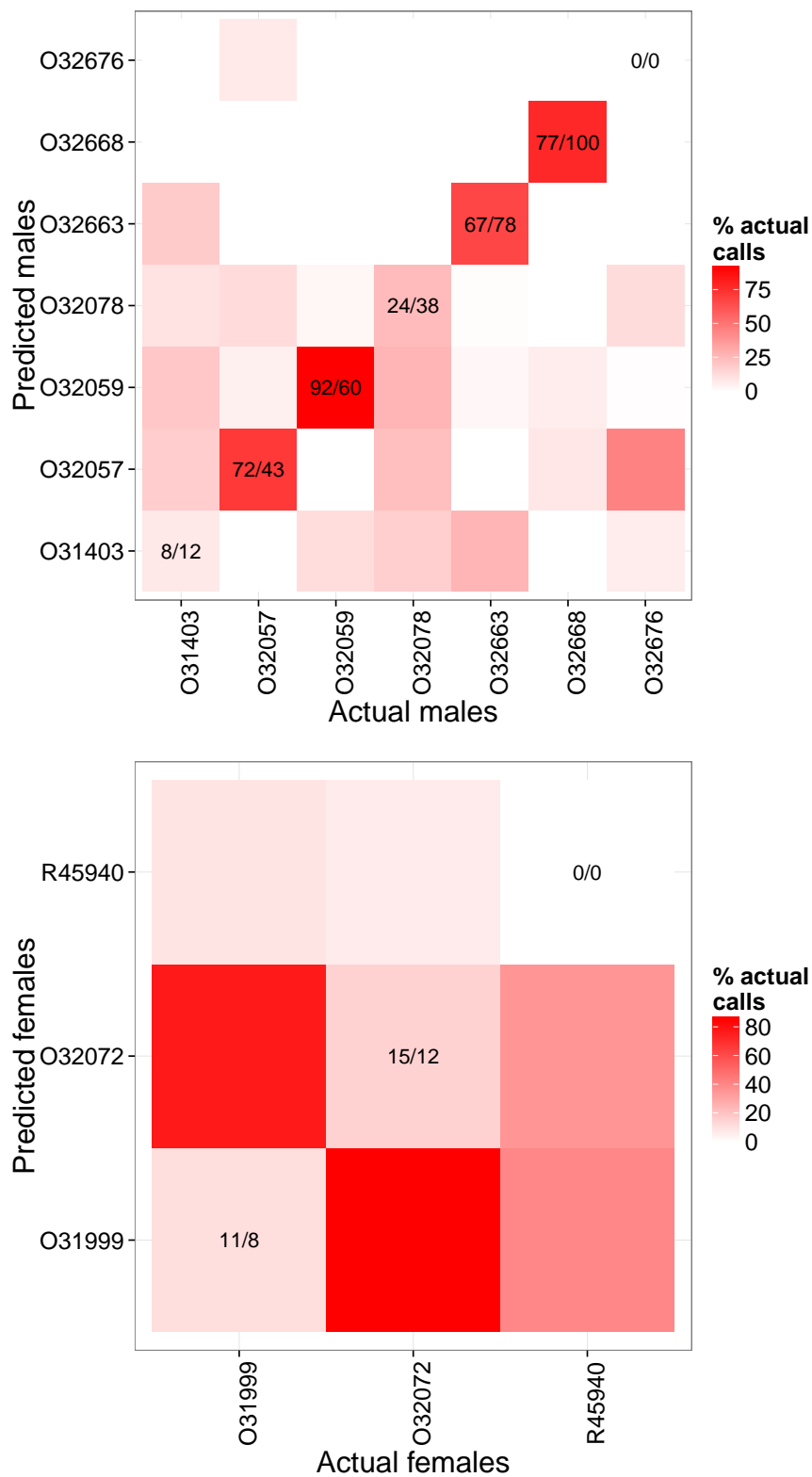
Figure 4.3.13: Cross-validation confusion matrices for support vector machine classifiers for call variables. Shading and annotations as Figure 4.3.10.
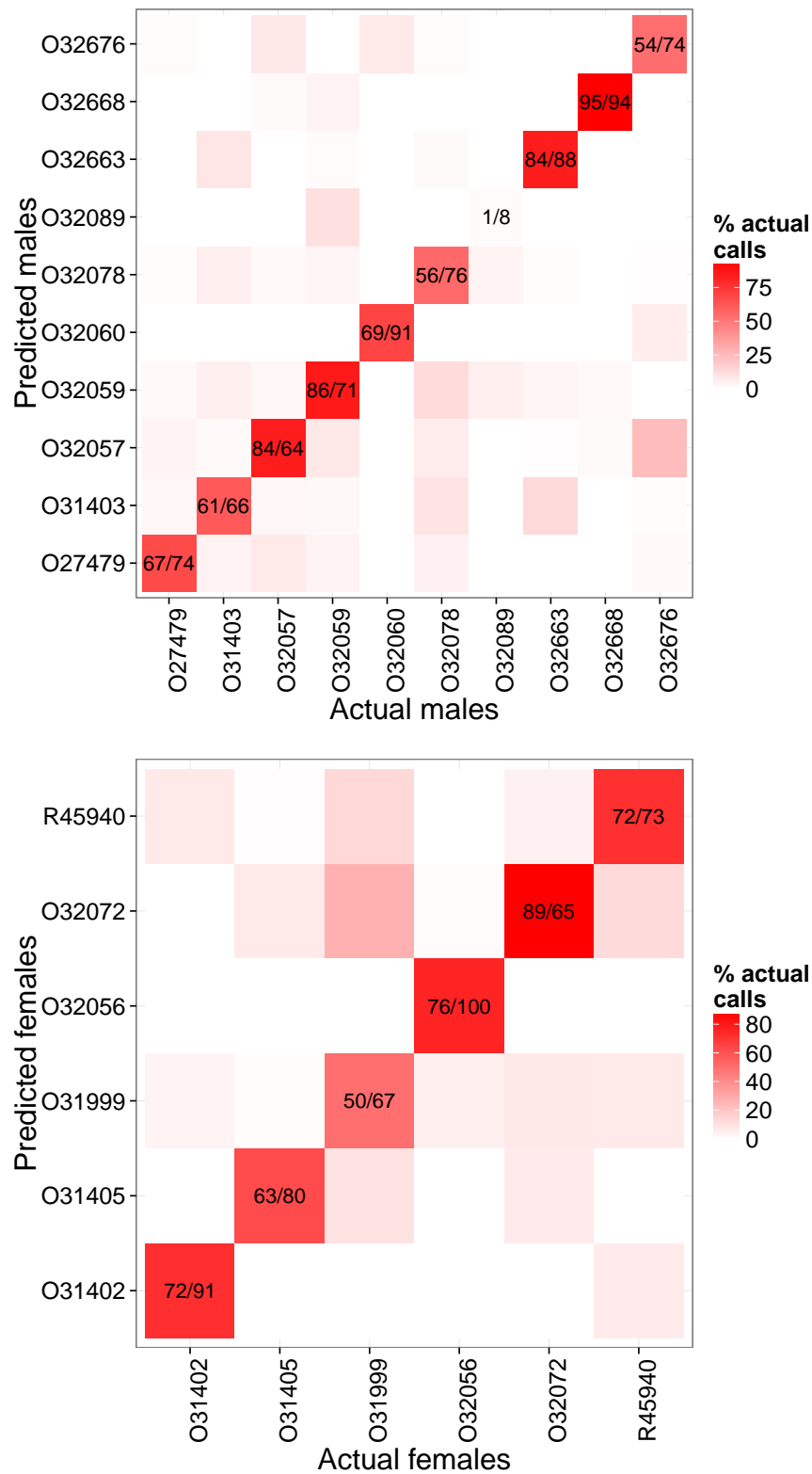
Figure 4.3.14: Cross-validation confusion matrices for support vector machine classifiers for syllable variables. Shading and annotations as Figure 4.3.10.
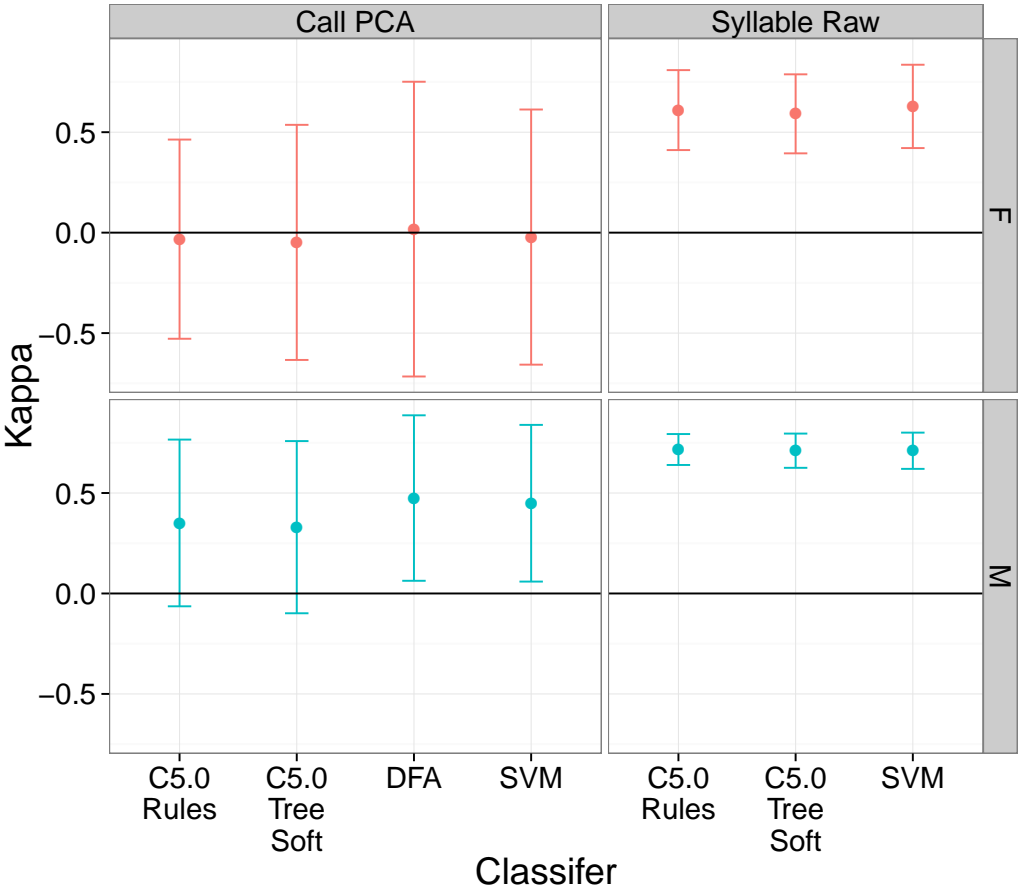
Figure 4.3.15: Performance comparison of each classification method for both call and syllable variables. The Kappa statistic measures classifier accuracy over all classes, adjusted for different class size. Values of one indicate perfect accuracy, whereas zero is the value expected from chance assignment. Error bars are ± 2 standard deviations.
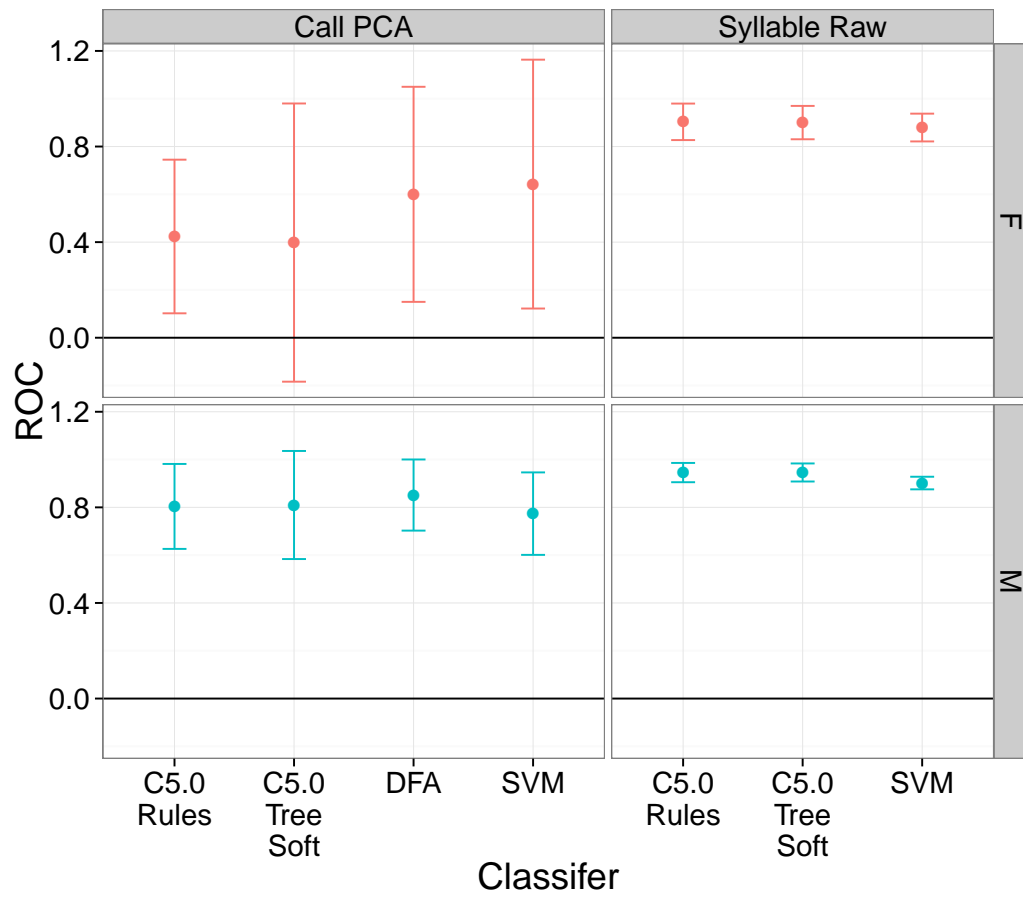
Figure 4.3.16: Area under the receiver operating characteristic (ROC) curve for each classifier. Larger values indicate better performance. Error bars are ±2 standard deviations.

great currasows (48%; Baldo & Mennill 2011) and American crows (65%; Yorzinski et al. 2006). When applied to syllables, the machine learning methods provided classification accuracies far better than expected by chance, indicating that acoustic identification is feasible.

For conservation applications a major limitation of these classifiers is that they can only identify known individuals, so would only be suitable for re-identification. They might, for example, be used to monitor the stability of territories or pair bonds with time. The large overlap in spectral parameter space suggests that these methods are only suitable for relatively small numbers of LSK. However, this is less problematic for kiwi than other species, since their territoriality means that discrimination of large numbers of individuals would be necessary. A substantial disadvantage of these methods is the effort required to collect sufficient training data, particularly for females which call less (§ 5.3). Obtaining enough high quality calls is difficult for kiwi, since they are cryptic and typically call just a few times each night. This restriction favours the machine learning classifiers, which can be applied to syllables rather than call means. Overall, these results suggest that call identification is possible for LSK for conservation purposes, and may have some limited applications.

From an ecological perspective, the discrimination results suggest that LSK are likely to be able to distinguish conspecifics by call to some extent. However, the close overlap of individuals indicates that identification may not always be easily accomplished. This is supported by an observation of a male LSK apparently attacking its mate which had just called (A. Digby, pers. obs.). It is possible that as territorial birds, LSK do not require strong vocal complexity and individual distinctiveness (Robisson 1992). The inter-sexual difference in call stereotypy is not in accordance with this, however. The classifiers indicated that females have less unique calls than males. This may have ecological relevance: perhaps male calls vary more due to sexual selection pressures such as advertisement of size or fitness. However, this may be an artefact of the temporal and frequency measures used, or the small sample size for females.

Assuming LSK do discriminate acoustically (but see discussion in Chapter 3), they are likely to use more sophisticated methods than simple temporal and frequency call features. Future attempts at identification should focus on short-term vocal characteristics such as formants and non-linear phenomena (NLP; Chapter 3). Such features of voice are more likely to be the characters used by the birds themselves to identify conspecifics (Weary et al. 1990, Fox 2008a), and would enable discrimination from fewer calls. However, the relatively infrequent appearance of NLP in LSK

calls, and the low inter-individual variation in their use (§ 3.3.3), suggest that they may not serve an identification function in these kiwi (§ 3.4.1).

The low call variation between LSK in this study is in contrast with the greater variation and higher classification success rate found in brown kiwi (Corfield 2005). LSK have a much lower genetic diversity than other kiwi (Ramstad et al. 2010), and this may affect their vocal diversity. As non-oscines, kiwi are not expected to learn their calls, and so vocalisations may reflect genetic differences (Kroodsma 2005). Genetic diversity is related to song diversity or structure at an individual level in a number of bird species (Marshall et al. 2003, Seddon et al. 2004, Pfaff et al. 2007, Araya-Ajoy et al. 2009). A similar effect has been found among populations, with more genetically diverse populations exhibiting greater song differences than more genetically similar populations (Robin et al. 2011). Although possibly confounded by ecological factors such as morphological and habitat differences, it is feasible that a relation between genetic and call diversity may also be present at the species level, and at least partially explain the vocal differences between little spotted and brown kiwi.

If genetic and call diversity are related in this way, then the stereotypic differences between LSK individuals, albeit limited, would suggest that acoustic individual identification is likely in the other kiwi species. However, in order to establish whether this relationship is a real effect, then a direct comparison is necessary of all kiwi species using identical methodology, such a multivariate dissimilarity techniques (Anderson et al. 2006, Anderson 2006). Contrasting call diversity between different LSK populations of different genetic diversity (Ramstad et al. 2013), would help assess links between call and genetic diversity.

If verified, the link between genetics and calls offers a potentially powerful tool for conservation management. Acoustic methods could enable rapid and inexpensive comparisons of genetic differences among populations, or be used to locate or identify different genetic strains. Although clearly inferior to conventional genetic techniques, this could prove a useful preliminary method for preparing translocations, managing translocated populations, or prioritising areas for conservation protection.

LITTLE SPOTTED KIWI VOCAL ACTIVITY

ABSTRACT

Variation in vocal activity can provide rich behavioural information and assessment of such changes is important in conservation management. This is particularly true of kiwi (Apterygidae), for which calls are used heavily in monitoring, but their use is inhibited by insufficient and conflicting evidence of call variation with temporal and environmental factors. Long-term analysis of kiwi vocal activity is necessary to guide the application of calls for conservation monitoring, and to enhance knowledge of kiwi behaviour. This chapter presents results from a nightly survey of little spotted kiwi (LSK; *Apteryx owenii*) calling patterns over three years, using autonomous

acoustic recording and automatic call detection methods. This has yielded the largest kiwi call dataset ever recorded. These data are combined with meteorological and ephemeris data in a generalised mixed additive model to test the hypothesis that call rates vary significantly with season, time of night, and environmental conditions.

Male LSK called three times more often than females, consistent with the hypothesis that males have a greater role in territory defence. Both sexes showed a peak in calling activity at the start and end of the night, which is also in accordance with a territorial call function. There was significant seasonal variation in calling, with both sexes most vocal at the beginning of the breeding period. An unexplained increase in female calls only during the autumn was also observed. LSK call rates were significantly affected by environmental conditions, aside from a decrease with increasing wind and rain which was a likely result of reduced detection probabilities. Calls decreased with temperature and moderate air pressures, but increased with humidity. It is unclear to what extent these effects are caused by variation in sound propagation, or reflect changes in calling rates. A sharp increase with ground moisture content was in accordance with results from brown kiwi (*Apteryx mantelli*), and can be explained by improved foraging or reduced likelihood of detection by predators or conspecifics in damp conditions. A strong decrease in calls occurred with increasing cloud cover, but there was no effect of moon phase. This suggests that light pollution may influence kiwi activity at the study site. These environmental impacts are used to issue guidance for LSK call monitoring, and demonstrate the effectiveness of acoustic methods in providing behavioural information not available to other monitoring methods.

5.1   INTRODUCTION

5.1.1   *The relevance of vocal activity*

Changes in vocal activity can provide behavioural information. The most obvious example is temporal variation in bird song, with variation in calling rates with time of day and season revealing behavioural patterns of feeding, territorial behaviour, and reproduction (Catchpole & Slater 2008). Vocal output association with environmental conditions can also be due to other influences, such as food availability and predator avoidance. For example, moon phase affects activity in nocturnal prey and predators (Penteriani et al. 2011). Birds may be less active during bright moon periods due to predator avoidance and a reduction in prey availability (Mougeot
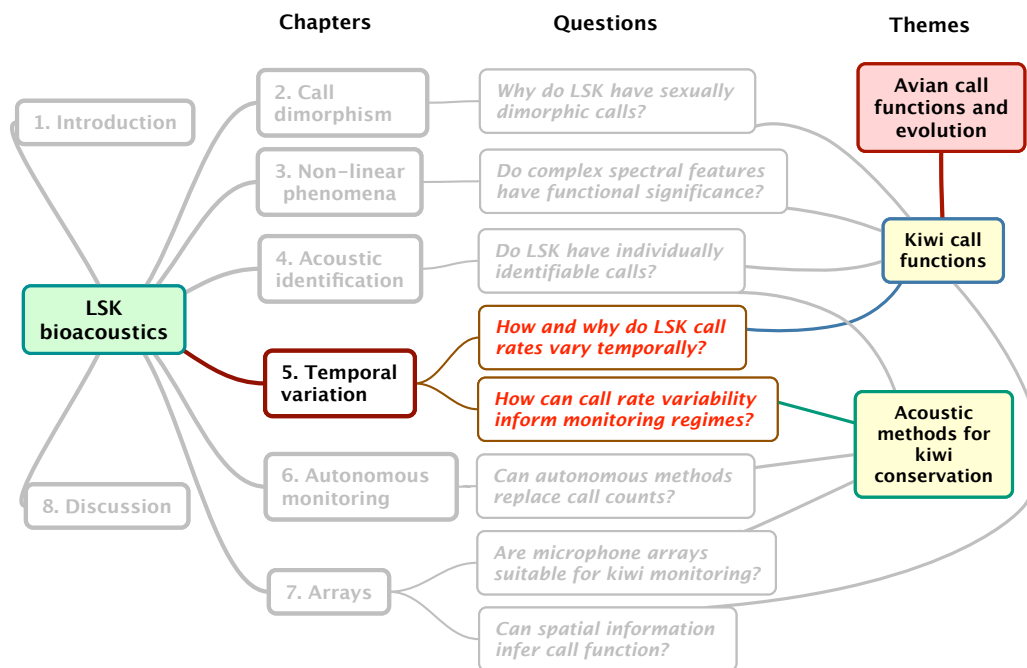
Topics covered by this chapter.

& Bretagnolle 2000). Yet some species are more vocal on moonlit nights, suggesting that there may be a trade-off between risks of predation and collision (McKown 2008). Activity during moonlight can infer ecological information such as breeding status: moonlight avoidance in owls (Penteriani et al. 2011) and petrels (Mougeot & Bretagnolle 2000) differs between breeding and non-breeding individuals. Vocal variation can provide insight into other modes of communication: for example, eagle owls (*Bubo bubo*) call more on moonlit nights since they utilise a white throat patch which is used for communication and is only visible during vocal displays (Penteriani et al. 2010).

Weather conditions also invoke behavioural changes and thus differences in vocal output in many species. For example, rain can restrict territorial displays, cloud can reduce foraging (Bas et al. 2008), and wind influences migration activity (Hüppop & Hilgerloh 2012). Weather conditions also influence sound transmission properties, which can strongly influence behaviour (Larom et al. 1997).

Measurement of call rates is also important for conservation monitoring and management. Call point counts are commonly used to survey for species abundance (Dawson & Bull 1975, Rosenstock et al. 2002, Thompson 2002). Yet for many species counts vary depending on the time of day or weather conditions under which

they were conducted (Bas et al. 2008). Temporal census bias can also lead to whole patterns of behaviour being missed, and subsequent underestimates of population size (Bridges & Dorcas 2000).

5.1.2   *Kiwi vocal behaviour*

5.1.2.1   *Kiwi call monitoring*

Call counts are particularly appropriate for monitoring of cryptic or nocturnal species which are difficult to monitor by other means. This applies to kiwi, for which call censuses are an important part of conservation. Call counts are used to assess kiwi distribution with location and habitat, monitor population trends, and locate remnant populations (Robertson & Colbourne 2003, Colbourne 2006). They also provide information on the efficacy of conservation management measures such as advocacy or predator control (Robertson et al. 1993, Pierce & Westbrooke 2003), and can identify predation events (Colbourne 2006).

A standardised national scheme of call monitoring, established in 1993, assesses long-term kiwi population trends across New Zealand (Robertson & Colbourne 2003). Twenty sites covering all kiwi populations and a range of management levels are monitored. Each site is visited for three consecutive years, at the same time of year, to establish a baseline, and then at 5-yearly intervals to monitor trends in call rates.This procedure was designed to be capable of detecting a 25% variation in mean call counts (McLennan 1992).

Kiwi call counts are conducted under similar conditions. Counts are short (1–2 hours), and take place in the first four hours of darkness (Robertson & Colbourne 2003). Observers are requested to listen on dark nights near the new moon, in fine weather conditions, and at certain times of year (Colbourne 2006). These temporal and environmental constraints are imposed in order to maximise detectability and maintain consistency, which is necessary when monitoring population trends.

5.1.2.2   *Variation in kiwi vocal activity*

There is substantial variation in kiwi calling activity. The most noticeable are seasonal variations, with brown kiwi (*Apteryx mantelli*) exhibiting high call rates prior to the main breeding season, usually during June – August (Colbourne & Kleinpaste 1984). Yet call rate peaks vary between populations (Miles et al. 1997), possibly

as a result of differences in the breeding period. Seasonal fluctuations also differ between species: brown kiwi calling has large fluctuations, but great spotted kiwi (*Apteryx haasti*) and Stewart Island tokoeka (*Apteryx australis*) show little seasonal variation (Colbourne 2006).

Moonlight also appears to influence brown kiwi calling activity, with many authors reporting reduced call rates during brighter nights and periods nearer to full moon (Buller 1888, Clark 1952, Colbourne & Kleinpaste 1984). It has been postulated that kiwi may even remain inside their burrows on the brightest nights (Colbourne & Kleinpaste 1984). Yet low brown kiwi call rates have been reported on very dark nights (Colbourne & Kleinpaste 1984), and in some brown kiwi populations there is no discernible effect of moon phase on calling activity (Miles et al. 1997, Pierce & Westbrooke 2003).

Weather conditions also affect calling behaviour. Fewer calls are recorded with increased wind and rain (Miles et al. 1997, Pierce & Westbrooke 2003), which is at least in part due to reduced detection probabilities. Increased ground moisture increases call activity (Pierce & Westbrooke 2003). In general weather effects are not well understood: for example, some studies report that temperature affects call rates (Miles et al. 1997), whereas others do not (Pierce & Westbrooke 2003). Call rates can also vary dramatically between consecutive nights with the same weather conditions (Colbourne & Kleinpaste 1984).

The contradictions in the effects of environmental conditions suggest that kiwi calling activity, and hence behaviour, is not fully understood. Yet if changes in call rates are to be used as a proxy for population trends, it is essential that factors affecting calling activity are known. Measurement of these variations is necessary to guide census efforts, such as determining the optimal time of night, season and count length, and the number of consecutive nights required to accurately sample population density (Miles et al. 1997). Some of these aspects have been investigated in brown kiwi (e.g. Colbourne & Kleinpaste 1984, Miles et al. 1997, Pierce & Westbrooke 2003), but studies are heavily biased towards that species, with vocal activity virtually unexplored in some species such as little spotted and great spotted kiwi. The different species have quite different ecology, which is likely to affect activity. For example, due to their size LSK are likely to have been the only kiwi preyed upon by the extinct laughing owl (*Sceloglaux albifacies*; Worthy 2001, Worthy & Holdaway 2002; Figure 5.1.1), and as such may exhibit different moonlight avoidance to the other species.

SCELOGLAUX ALBIFACIES.

Figure 5.1.1: The laughing owl (*Sceloglaux albifacies*), an extinct predator of little spotted kiwi. Credit: Keulemans, John Gerrard, 1842-1912. Keulemans, John Gerrard, 1842-1912 :Sceloglaux albifacies. J. G. Keulemans del. Mintern Bros Imp. [London, 1876]. Rowley, George Dawson :Ornithological miscellany. London, Trubner Co, 1876. Ref: PUBL-0035-1-035. Alexander Turnbull Library, Wellington, New Zealand. http://natlib.govt.nz/records/23241413.

Knowledge of kiwi calling activity can also aid conservation efforts by guiding when calls should be used for capture. Playback at night is a common method of catching kiwi (Robertson & Colbourne 2003), and is currently advised during dark moon periods, since kiwi appear to be less wary in lower light levels. However, it's possible that kiwi vocal activity, and hence response to calls, may be greater during other moon phases or ambient conditions, and that this may vary between species.

5.1.2.3 *Long-term kiwi monitoring*

A greater understanding of kiwi calling patterns is therefore important for optimising and minimising the bias in short-term monitoring programmes, and for improving knowledge of kiwi behaviour. However, few studies have assessed long-term trends in kiwi calling patterns, and even those that have conducted many counts have still monitored calling activity over relatively short periods (Colbourne & Kleinpaste 1984, Miles et al. 1997). This is due to the considerable time demands in carrying out call counts with observers *in situ*. Autonomous acoustic recorders now offer a method for monitoring of kiwi acoustic activity over long time frames and a wide range of ambient conditions. This can be used to guide, supplement or even replace existing call counts.

5.1.2.4 *Chapter aims*

In this chapter a long-term dataset of nightly LSK call recording is used to assess variation in call rates, to enhance understanding of kiwi behaviour and to assist in the use of call counts for kiwi conservation. A generalised model is used to test the hypothesis that LSK call rates vary significantly with temporal and environmental factors, as a result of calls being used for specific functions and variation in the detection ability of acoustic methods. This predictions made from this hypothesis are that LSK call rates will vary significantly with:

1. Time of night and season, due to the functional relevance of calls.

2. Moon phase and ground moisture, reflecting behavioural changes as found in brown kiwi.

3. Wind speed, wind direction and rainfall, through decreased detection probabilities.

5.2   METHODS

5.2.1   *Recordings*

The recording location was on a spur in a central region of Zealandia (41.29776 S, 174.74432 E; 190 m elevation), a fenced 225 ha mammalian predator-free reserve in Wellington, New Zealand (Figure 5.2.1). This area had high kiwi density, with at least 14 territories within a radius of 400 m from the monitoring site, estimated from call counts and radio telemetry studies (§ 3.2.1.1). A Song Meter 2 autonomous recorder (Wildlife Acoustics Inc., Concord, MA, USA) was attached to a tree at a height of 1.5 m. Recordings were made in stereo at 22.05 kHz and saved in three hour blocks as lossless compressed wave files ('wac0' format). The default hardware gain stage of $+24$ dB per channel was used, and no software gain. Recordings were scheduled from ten minutes after sunset to ten minutes after sunrise each night, and ran from 7th November 2009 to 30th November 2012. Recordings were made on 971 nights over the three year study period; 148 nights were lost due to recorder malfunctions or flat batteries. The recorder was resupplied with batteries and memory cards weekly.

5.2.2   *Call detection*

LSK calls within the recordings were detected automatically using a two-stage software process: (1) event detection and (2) event classification.

5.2.2.1   *Event detection*

Event detection software was written in C# by Michael Towsey (Queensland University of Technology, Brisbane, Australia), and modified (by Michael Towsey and Andrew Digby) for use with LSK calls. These consist of long series of repeated syllables (Digby et al. 2013), so an oscillation detector was developed based on a discrete cosine transformation (Towsey et al. 2012). The software used cross-correlation to detect potential LSK calls based on amplitude and periodicity, producing a list of 'events' for each sound file scanned. Since the sexes have largely non-overlapping call energies (Digby et al. 2013), separate searches were performed for male and female calls, in the frequency ranges 2.2–3.2 kHz and 1.3–2.0 kHz respectively. Each event was assigned eight scores, characterising its intensity, periodicity, change in
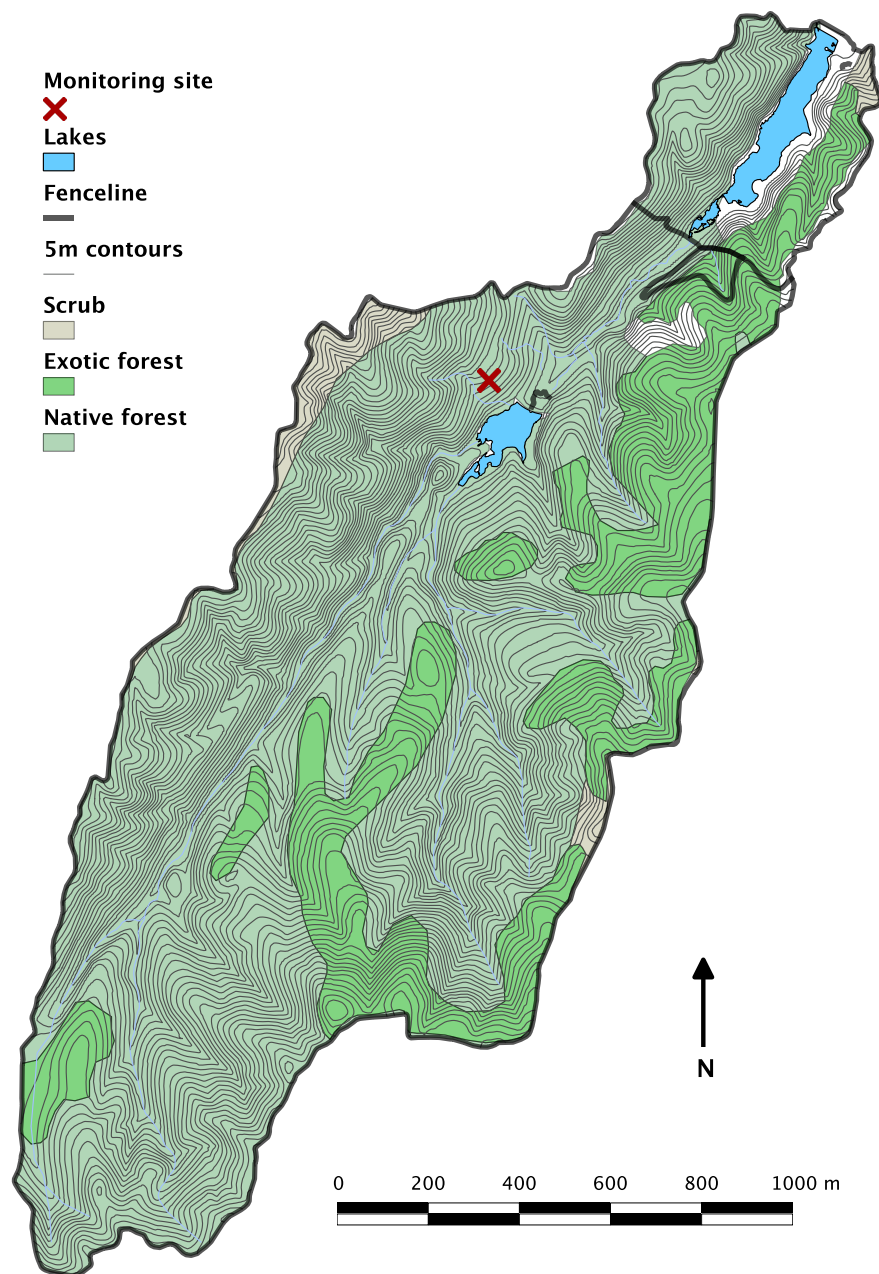
Figure 5.2.1: The recorder location within Zealandia.

periodicity, frequency modulation of each element, change in amplitude across elements, bandwidth, and signal-to-noise ratio. A weighted combination of these scores was used to produced a combined score, and events above a given threshold of this score are marked as candidate kiwi calls. The scores and their weighting were optimised to detect LSK calls, whilst minimising false positives from other crepuscular or nocturnal species with similar calls that share LSK habitat, such as kaka (*Nestor meridionalis*; a parrot) , weka (*Gallirallus australis*; a rail), and morepork (*Ninox novaeseelandiae*; an owl). A minimum length of 7 s was imposed for all events. To avoid duplicate detections from calls containing pauses, a Python script was used to merge events if they were within 5 s of another event in the same frequency band and with a similar combined score. Processing time on a Windows PC (4-core Xeon 2.4 GHz; 4 GB RAM) was approximately two minutes for each hour of recording.

5.2.2.2 *Event classification*

TRAINING THE CLASSIFIER   The list of potential LSK calls from each sound file were processed by a statistical classifier using C5.0 software (version 2.09; Rule-Quest Research, Australia), running on Mac OS X. A classifier was built using 22 three-hour files from the same site, randomly selected over a range of dates, and then filtered to include only nights with low background noise. LSK calls within this training set were identified by manual inspection of spectrograms in Raven Pro 1.4 (Charif et al. 2010). Each call was annotated with sex and a subjective quality score reflecting the signal-to-noise ratio of the call in the spectrogram, ranging from one (very clearly visible) to ten (only just visible). Calls with the lowest quality classification (ten) were removed from the training set, as were calls which overlapped with other biotic or abiotic sound sources. The subsequent training data contained 459 LSK calls longer than 7 s (368 male, 91 female).

The event detector returned 670 events for this dataset. For each event all eight detector scores and the true classification (kiwi or not kiwi) were input to the classifier. Different software settings for building the classifier were tested. These were decision tree (hierarchical rules) or ruleset (unordered rules), softening, and adaptive boosting. Boosting uses iterative classifier generation to incrementally improve the error rate, and softening changes 'sharp' decision thresholds into 'fuzzy' boundaries, which can improve accuracy. The method chosen was that which gave the lowest error rate and highest recall and precision values from ten repeats of ten-fold cross-validation (Box 5.1). For each event C5.0 returned a classification (kiwi or not

kiwi) and confidence score (0–1), then the sex of each call was assigned using the designation from the event detector.

Verification measures

- TP: True positive. When an event is correctly classed as a kiwi.

- FP: False positive. When an event is classed as a kiwi, but actually isn't.

- FN: False negative. A 'miss'. When a kiwi call is not detected or classified as such.

- TN: True negative. When a detection is correctly not classified as a kiwi.

- Recall:
$$R = \frac{TP}{TP + FN}$$

- Precision:
$$P = \frac{TP}{TP + FP}$$

- Accuracy:
$$A = \frac{TP + TN}{TP + FP + FN + TN}$$

Box 5.1: Verification measures used to assess detector performance

---

TESTING THE CLASSIFIER   With an optimal classifier chosen from cross-validation, the detection and classification software was then tested on a separate dataset not used for training. Twenty-two three-hour files from the same site were randomly chosen from dates across three years, on nights with low to moderate background noise. These recordings were inspected spectrographically, and all LSK calls annotated with sex and quality. This testing dataset contained 604 calls longer than 7 s with all quality levels (470 male, 134 female). The detection and classification software was run on these files, and each detection was verified by inspecting spectrograms in Raven Pro 1.4 . These verification results were used to characterise the performance of the software over all recordings.

RUNNING THE CLASSIFIER   Once the performance of the software was determined from a subset of recordings, it was run over all of the recorder sound files from the 3 years of monitoring. The event files for each recording were merged by month, and then input to the decision tree classifier, which produced a list of candidate kiwi calls for each month. These were not manually inspected: doing so would require a very large amount of time, and the high precision indicated from the classifier testing (see § 5.3) indicated that the detections would contain very few false positives.

5.2.3 *Meteorological and ephemeris data*

Meteorological data for the monitoring period were obtained from the New Zealand National Climate Database (http://cliflo.niwa.co.nz). Hourly observations from the closest automatic weather station (AWS) to the monitoring site were used, which was at Kelburn in Wellington (World Meteorological Organisation synoptic number 93437; 41.285 S, 174.768 E, 125 m elevation). This was 2.6 km from the study site. Cloud observations were not available at that station, so were taken from Wellington airport AWS (synoptic number 93436, 41.322 S, 174.804 E, 4 m), which was 6.4 km away. Cloud cover would not vary substantially over such distance, and the Kelburn site shares similar weather conditions with the monitoring area. This was supported by the high correlation ($r = 0.91$) between temperatures recorded by the Song Meter during recording with those from the AWS. Despite the availability of these on-site temperatures, those from the AWS were used to maintain consistency with other observations.

Moon phase was calculated at each hour from sunset to sunrise each night, using the package moonsun (Komsta 2010) in R (version 2.15, R Development Core Team 2011). Cloud cover is reported at multiple levels, given by cloud height, but only a single cloud cover variable was required for the model. A weighted cloud cover was therefore derived, with the cover scaled by inverse height, from a scaling factor of 1 at 0 m, to 0 at the maximum cloud height. Lower (thicker) cloud therefore received more weighting. Available meteorological data did not include a measure of ground moisture, which affects call rates in brown kiwi (Pierce & Westbrooke 2003). Ground moisture was therefore approximated from hourly rainfall, by summing precipitation in the twelve hours previous to each hourly time interval. This variable is hereafter referred to as 12-hour or cumulative rain.

Table 5.2.1: Meteorological data obtained from the New Zealand National Climate Database. Hourly observations are made on the hour (UTC), and are either instantaneous (Inst.) or apply to the previous hour. Station numbers are World Meteorological Organisation synoptic numbers.

Variable	Station location	Station number	Frequency	Period
2 m temperature	Kelburn	93437	Hourly	Inst.
Relative humidity	Kelburn	93437	Hourly	Inst.
Station pressure	Kelburn	93437	Hourly	Inst.
Surface wind speed	Kelburn	93437	Hourly	Hour
Surface wind direction	Kelburn	93437	Hourly	Inst.
Maximum gust speed	Kelburn	93437	Hourly	Hour
Maximum gust direction	Kelburn	93437	Hourly	Hour
Rainfall	Kelburn	93437	Hourly	Hour
Cloud cover	Airport	93436	Hourly	Inst.
Cloud height	Airport	93436	Hourly	Inst.

5.2.4 *Statistical analyses*

5.2.4.1 *Data aggregation*

Call numbers and meteorological and ephemeris data were summarised in hourly time bins for each night, with the first bin starting at sunset. All clock times were converted to UTC, and calculations were performed in package zoo (Zeileis & Grothendieck 2005) in R. The numbers of male and female LSK calls in each hourly interval from sunset were aggregated. Instantaneous hourly meteorological observations (Table 5.2.1) were linearly interpolated to the mid-point of each time bin, and stored against the bin end time. Summary and cumulative observations (previous hour maximum gust and rainfall) were also interpolated to the bin centres, as was moon phase. The actual recording time in each hourly time bin was calculated from sensor files logged by the recorder at five minute intervals. Bins did not contain a full hour of recording if sunset occurred or the recorder ran out of batteries during the hour. Bins with less than ten minutes of recording were removed from the dataset.

5.2.4.2 *Additive models*

It was decided on the basis of correlation matrices and variance inflation factors to drop maximum gust speed and direction from the model because of collinearity. Outlier inspection removed two observations containing hourly rain greater than 15 mm and a further nine with 12-hourly rain greater than 50 mm. Interactions were included for wind speed with direction, moon phase with cloud cover, and sex with month, time of night, wind speed and precipitation. Further interactions were not considered likely to affect call rates and so were excluded to avoid over-complicating the model. An offset term allowed for counts from variable time periods. The subsequent model terms were:

- Response variable: counts per hour for each sex

- Explanatory variables: sex, year, month, hours after sunset, moon phase, hourly rain, 12-hour rain, wind speed, wind direction, weighted cloud cover. All calculated for each hour of each night.

- Interactions: sex × month, hours after sunset, wind speed, hourly rain; wind speed × wind direction; moon phase × weight clouded cover.

- Offset: log(bin length).

Initial analysis showed non-linear relationships between counts and some explanatory variables, so generalised additive models (GAMs) were fitted using package `mgcv` (Wood 2011) in R. A negative binomial link function was used due to slight overdispersion ($\widehat{\phi} = 2.24$) in a Poisson model. Zero inflation was unlikely given the low dispersion parameter in the negative binomial model ($\widehat{\phi} = 1.15$), and was confirmed with histograms of counts per time bin (Figure 5.2.2)

An initial GAM showed residual heterogeneity for relative humidity, rain, and 12-hour rain. A variance structure is not possible in GAMs in R, so variable transforms were used, selected by improvements in residual homogeneity. Relative humidity was squared, and the two rain variables were transformed by $log(x + 1)$.

Smoothers to model non-linearities were fitted to all except categorical variables (sex and year). Further models were constructed with separate smoothers for each sex for month, time of night, wind speed and wind direction, and with a smoother for moon phase that depended on weighted cloud cover. Different smoothing classes were tested: thin-plate regression splines, cubic regression splines, and P-splines (Wood 2003, 2004). Cyclic splines were used for month and wind direction variables,
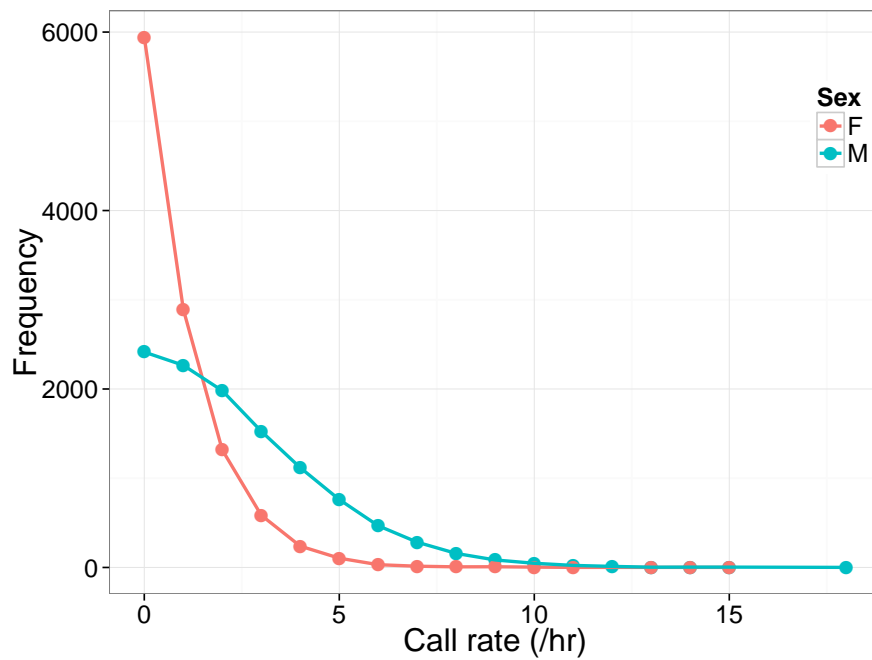
Figure 5.2.2: Numbers of counts in each time bin after sunset, for males and females. There is no strong evidence of zero-inflation within either sex.

to ensure continuity at end points (e. g. Dec–Jan and 0°). All models were compared using AIC and dispersion. The best model had cyclic P-spline smoothers for month and wind direction, a P-spline smoother for time of night, cubic regression smoothers for all other variables, and separate sex smoothers for month, time of night, wind speed and wind direction.

Auto-correlation functions (ACFs) indicated independence problems in the GAM, so a correlation structure was added to the model, using function gamm in mgcv. For such a large dataset the correlation structure must be applied to relatively small groups (Wood 2011), so two structures were tested: correlations among time bins in each night, and among counts within each week. Generalised additive mixed models (GAMMs) with correlation structures take many hours to run for large datasets, and do not yield likelihoods or AIC, so generalised linear squares (GLS) models (i. e. without smoothing terms; Zuur et al. 2009) were used to assess the approximate performance of these correlation groupings. A range of correlation structures were tested, including auto-regressive moving average (ARMA) models (Zuur et al. 2009), which provide a residual function from an auto-regressive parameter, p, and a moving average parameter, q. All combinations with $p = 1, 2, 3$ and $q = 0, 1, 2, 3$ were tested, and models compared using deviance and AIC.

$$
\begin{aligned}
Call\,Rate = {}& Sex + Year + s(Month, by = Sex, bs = `cp') \\
& + s(SShr, by = Sex, bs = `ps') + s(Moon, bs = `cr') \\
& + s(WindSpd, by = Sex, bs = `cr') + s(log(Rain + 1), by = Sex, bs = `cr') \\
& + s(Tair, bs = `cr') + s(Pressure, bs = `cr') \\
& + s((RH)^2, bs = `cr') + s(CldWt, bs = `cr') \\
& + s(log(CumRain + 1), bs = `cr') + s(WindDir, bs = `cc') \\
& + Moon \times CldWt + WindSpd \times WindDir \\
& + offset(log(bin\_length)
\end{aligned}
$$

Correlation structure: ARMA(p=2, q=2, form= 1|(Week × Sex)

Equation 5.1: The final generalised additive mixed model used to model call rates with sex, time, weather and moon variables. Smoother types are indicated by *bs*: '*cp*' = cyclic P-spline, '*ps*' = P-spline, '*cr*' = cubic regression spline. Other abbreviations: *RH* = relative humidity, *SShr* = hours after sunset, *CumRain* = cumulative rainfall, *CldWt* = weighted cloud cover, *bin_length* = time covered by each observation period.

The five best-performing correlation structures were applied to GAMMs containing smoothing terms for all explanatory variables, and compared using residual plots and ACFs. The ARMA(2,2) model with within-week grouping was assessed to be the best correlation structure. Informatic-Theoretic methods are preferred for model selection (Anderson et al. 2000, Anderson & Burnham 2002), particularly since generalised additive models yield approximate significance values (Zuur et al. 2009). However, since no AIC measures were available in the GAMM, a conservative approach of no model selection was adopted (Zuur et al. 2012). The final model used is given in Equation 5.1

5.3   RESULTS

5.3.1   *Recording data*

A total of 10,802 hours of recordings were made on 971 nights, and were aggregated into 11,386 time bins. After removing outliers (§ 5.2.4.2), and time bins less than ten minutes long (§ 5.2.4.1), the total monitoring time was 10,776 hours in 11,151 bins. The event detection software took about 15 days of computing time to process

these data. Event classification time using C5.0 was negligible: less than one minute for the entire dataset.

5.3.2 *Classifier results*

From cross-validation of the training data, the optimal C5.0 classifier used a decision tree with no softening and 10-classifier boosting. This classifier had a mean error rate of 0.9%, with 97.1% precision and 93.4% recall from the events passed to it, or 71.1% recall including calls missed by the event detector. Such cross-validation scores from training data are optimistic, and testing on the unseen dataset gave a more conservative precision and recall of 94% and 58% respectively. Excluding the very faintest calls (quality = 10) increased precision and recall to 93% and 80% (Figure 5.3.1). False positives were rare, and from a range of sources: occasionally song thrush (*Turdus philomelos*), tui (*Prosthemadera novaeseelandia*) or morepork calls were classified as male LSK, but noise fluctuations were also mistaken for kiwi calls. The software detected some calls that were not visible on the spectrogram, but were just audible. False negatives were mainly due to very faint calls, but also resulted from acoustic interference from other kiwi and other species. These occurred when bird song overlapped kiwi calls, or when two kiwi calls of the same sex were separated by a gap of less than 5 s, in which case they appeared as only one event due to merging criteria (§ 5.2.2.1).

A total of 638,720 events were detected by the event detector in the full three year dataset, of which 36,238 (26,962 male, 9,276 female) were classified as LSK calls and so were input to the model. The time of the first call each night, measured from sunset, was mode/median = 43/46 minutes for males, and 58/74 minutes for females.

5.3.3 *Model results*

The best GAMM showed that the number of calls detected depended strongly upon sex, month, time of night, wind speed, wind direction, cloud cover, hourly rain, and ground moisture (Table 5.3.1). There were weaker but still significant effects of air temperature, pressure, and the interaction of wind speed with direction. The effect of relative humidity was just non-significant. Year, moon phase and the interaction of moon phase with cloud cover all did not influence counts. The overall hypothesis that call rates are affected by temporal and environmental conditions is therefore
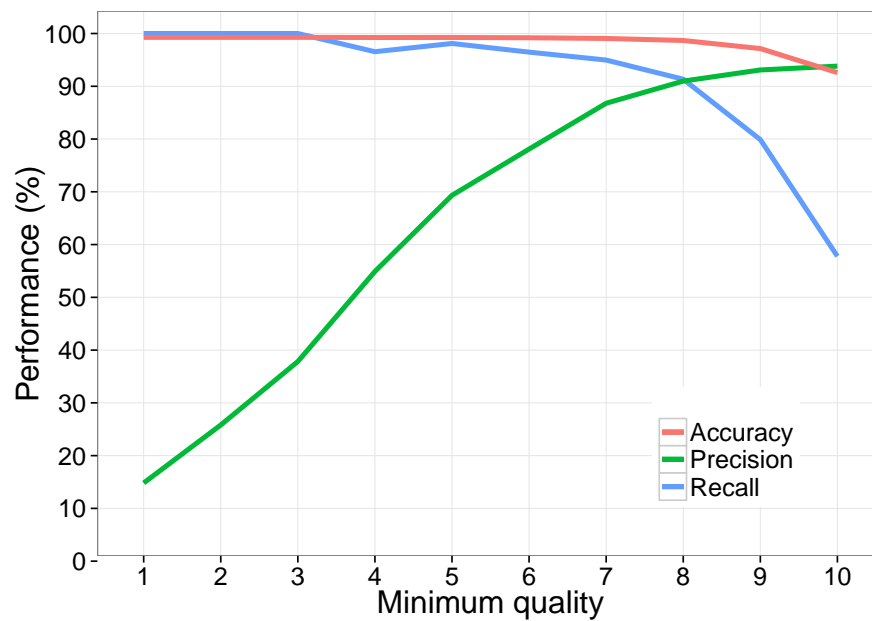
Figure 5.3.1: Performance of the best `C5.0` classifier from testing on an unseen dataset. Precision, recall and accuracy are given as a function of the minimum quality of calls in the test dataset. Increasing quality scores indicates decreasing signal-to-noise, such that calls with a quality score of ten are barely audible or visible on a spectrogram. Precision decreases with quality since the number of true positives decrease, but the false positives remain constant.

supported, but not all predictions are realised, as seen in the non-significance of moon phase, and the unexpected significance of temperature and pressure.

The sexes showed similar patterns of calling activity with time of night, peaking between the first and second hours after sunset, and then dropping throughout the night, with a slight increase at dawn (Figure 5.3.2). Although both sexes had highest calling rates during spring (September – October), males and females otherwise differed in their seasonal calling patterns. Females had a calling peak in autumn (∼April), but males did not, and male rates were lowest in mid-winter (June), whereas females were lowest in mid-summer (December – January).

Combined call rates for both sexes showed a decrease with increasing temperature and cloud cover. Calls decreased significantly with rain and wind speed, although there were differences in how the sexes were affected by these variables (Figure 5.3.3). There was a weak tendency for calls to increase with rising humidity, and for lower call rates at intermediate air pressures. Calls initially increased sharply from dry to moist ground conditions. Wind direction substantially affected the number of calls detected at the study site, with fewer calls during southerly winds, and more during northerlies.

Model validation showed that residuals conformed to assumptions of normality and homogeneity. Auto-correlation functions indicated very minor lack of independence at some lags (Figure 5.3.4), but no residual patterns as seen in the models without correlation structure. A temporal variogram confirmed that there were no significant independence problems (Figure 5.3.5).

5.4 DISCUSSION

5.4.1 *Automatic detector performance*

The high precision from the event detector and classifier was important for assessing call variability. Ideally all detections should have been inspected to remove all false positives, but this would have been very time-consuming for little gain. The event detector was optimised to minimise false positives, and there was accordingly a low (∼6%) false positive rate. Furthermore, the estimated false positives and true negatives are likely to be over-estimated, since the detector and classifier were trained and tested on the first three hours after sunset, when both types of error would be higher than at quieter times of the night. That relatively few false positives were due

Table 5.3.1: Results of the call rate GAMM, showing significance of linear model effects and smoothing terms. edf = estimated degrees of freedom for the spline, giving the non-linearity of the smoother (edf = 1 is a straight line). Significant codes (Sig.): *** < 0.001, ** < 0.01, * < 0.05. Abbreviations as Equation 5.1. Model adjusted R^2 = 0.36. Correlation structure = ARMA(2,2), with ϕ_1 = 1.47, ϕ_2 = -0.474, θ_1 = -1.39, θ_2 = 0.411. Dispersion: $\widehat{\phi}$ = 1.034.

Effect	Estimate	Std. Error	t	P	Sig.
(Intercept)	-31.00	35.11	-0.88	0.377	
Sex = F	-1.12	0.02	-73.09	<0.001	***
Year	0.01	0.02	0.79	0.432	
Moon:CldWt	0.00	0.00	1.01	0.311	
WindSpd×WindDir	0.00	0.00	2.16	0.031	*

Smoother	edf	Est. rank	F	P	
Month Sex=M	4.51	9.00	5.60	<0.001	***
Month Sex=F	5.46	9.00	4.18	<0.001	***
SShr Sex=M	8.45	8.45	47.64	<0.001	***
SShr Sex=F	8.51	8.51	52.81	<0.001	***
Moon	1.00	1.00	0.00	0.970	
WindSpd Sex=M	4.69	4.69	54.45	<0.001	***
WindSpd Sex=F	5.17	5.17	129.78	<0.001	***
log(Rain) Sex=M	3.42	3.42	169.14	<0.001	***
log(Rain) Sex=F	3.21	3.21	117.79	<0.001	***
Temperature	1.97	1.97	4.77	0.009	**
Pressure	2.64	2.64	3.01	0.036	*
RH^2	2.52	2.52	2.79	0.049	*
WindDir	5.96	8.00	5.63	<0.001	***
CldWt	1.00	1.00	37.91	<0.001	***
log(Cum. Rain)	4.58	4.58	11.65	<0.001	***
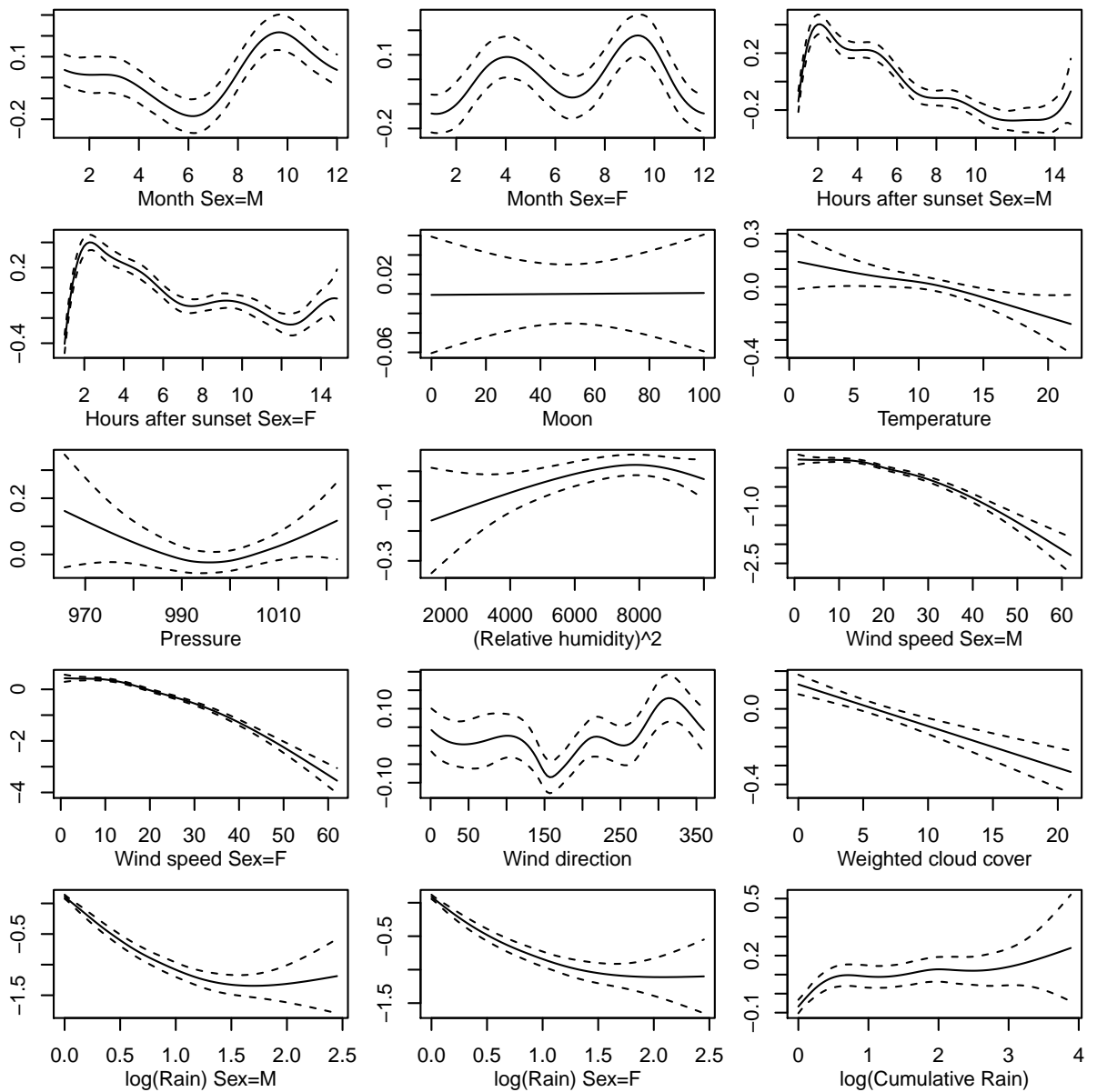
Figure 5.3.2: Variation in calls detected with temporal, weather and moon variables, from the generalised additive model. The vertical scale is the contribution to the fitted counts in the model (not actual counts), and has the same units for all plots. Month and time of night variables have a different smoother for males and females; all other variables are for the sexes combined. Dashed lines give approximate 95% confidence intervals.
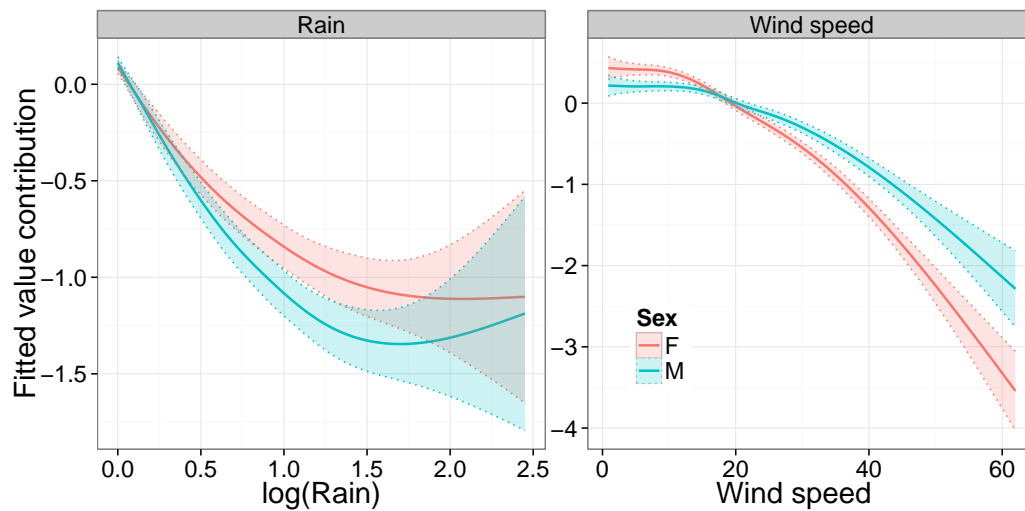
Figure 5.3.3: Variation in detected calls with wind and rain for each sex. Female calls decreased significantly more with wind than males, whereas male calls declined more with increasing rain than females.
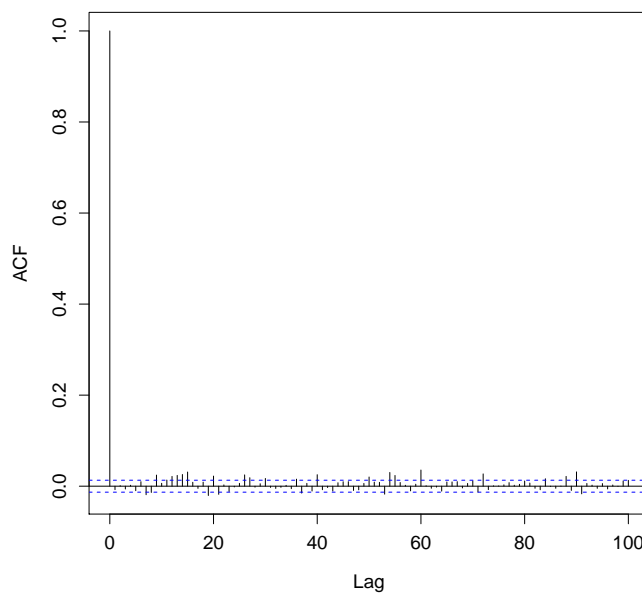
Figure 5.3.4: Auto-correlation function of normalised residuals for the final GAMM. Blue dotted lines show 95% confidence intervals for zero correlation. Although some lags have correlations that are significantly non-zero, there is no patterns to these, and correlation values are nevertheless small.
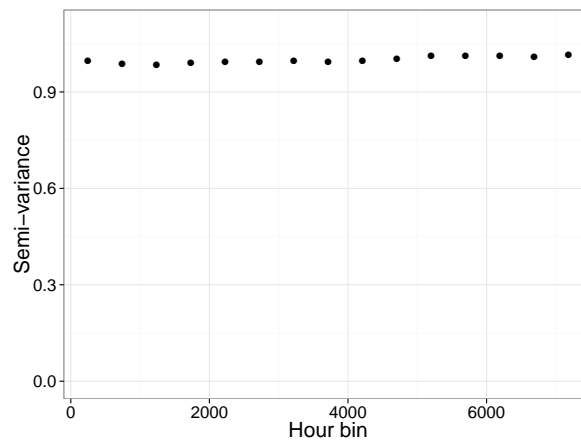
Figure 5.3.5: Variogram of GAMM normalised residuals with observation number (hour bin). The near-constant semi-variance indicates residual independence.

to other species confirms that they were not biased towards a particular time period (e. g. dusk or dawn). False positives due to noise indicated that these may have increased during high wind or rain. Since acoustic methods perform poorly in such conditions anyway, modifying this behaviour was not important. However, a future assessment of detector performance with environmental factors would be useful.

A cost of the high precision was that the recall was relatively low (58%), mainly due to the event detector missing very faint calls. Excluding the very faintest calls increases the recall to a respectable 80%, by effectively reducing the monitoring area. Including all quality calls, the recorder plus automatic detection software covered an area of at least 25 ha (Chapter 6). This encompasses at least ten kiwi territories at the study site, which provides an ample sample size for assessing call variability over an extended period.

Apart from very faint calls, the majority of the remaining false negatives were due to calls overlapping calls from other kiwi or species. This is an inherent limitation of acoustic and automatic detection methods, which are therefore less suited for areas with high call rates, or where other vocal species share call characteristics. However, preliminary results show that the detector used is able to reliably discriminate LSK calls from those of weka and morepork, even though these are frequently confused with kiwi by human listeners (Robertson & Colbourne 2003). Automatic call detection software is also less appropriate for species with more variable vocalisations, and may need to be retrained to cope with dialects. However, LSK calls are highly uniform between individuals, and between populations (unpubl. data), so are ideal candidates for such methods.

This study has produced the largest collection of kiwi calls ever recorded from a single site, with continuous observations over a long time baseline. Comparing the length of monitoring with earlier manual studies illustrates the benefits of autonomous acoustic methods. For example, Miles et al. (1997) observed 1079 calls in 478 hours of listening, and (Colbourne & Kleinpaste 1984) 1032 calls in over 27–63 nights. Automated techniques in this study yielded 36,238 candidate calls in 10,776 hours of monitoring, for considerably less effort.

5.4.2 *LSK calling behaviour*

The long time baseline of this study and accompanying environmental data enables behavioural inferences to be made from LSK call rate variability. However, call counts are the result of the number of calls made by a species, combined with the probability of detection (Bas et al. 2008). As such, the derived call variation measures the detection method as well as the intrinsic call rates. While in many cases the detector imposes no bias (e. g. temporal variations), this must be considered when assessing call variability.

5.4.2.1 *Sex differences*

There was a significant difference in detected call rates between the sexes, with a ratio of male to female calls of 2.9:1. This sex ratio derived here is similar to that of 2.5:1 reported from a sample of 129 LSK calls from call counts (Colbourne 2006). Male LSK clearly have a greater requirement to vocalise than females. The territoriality of LSK and their reliance on vocal communication, combined with inter-sexual call spectral and transmission differences (Chapter 2), suggests that this sex imbalance may be due to males being more territorial than females.

5.4.2.2 *Temporal variation*

MONTHLY   LSK call rates showed strong seasonal variation, with both sexes most vocal at the start of the September – February breeding season. This conforms to the relationship between breeding and vocalisations found in most bird species (Catchpole & Slater 2008). Male LSK show a similar seasonal calling pattern to that of brown kiwi (sexes combined) reported in Miles et al. (1997), with a minimum in mid-winter, and a peak around October. However, brown kiwi show substantial variations among populations in seasonal calling, with other populations peaking

in mid-winter (Colbourne & Kleinpaste 1984). LSK populations can vary in their breeding season (Hugh Robertson, New Zealand Department of Conservation and Helen Taylor, Victoria University, Wellington, New Zealand, pers. comm., 2012), so may also exhibit vocal activity variations among populations. The strong LSK female calling peak in autumn (April), without a corresponding male increase, has not been reported in other kiwi, and is unexplained.

ANNUAL   The non-significant effect of year on call rates is at odds with the expected LSK population increase estimated at the study site (Hugh Robertson, New Zealand Department of Conservation, pers. comm., 2010). However, it is in accordance with annual 'manual' LSK call counts conducted at the same site (Erin Daldry, Zealandia, Wellington, New Zealand, pers. comm., 2013). This could be due to the relatively high density of kiwi around the recorder location (A. Digby, unpubl. data). Recruited individuals may have difficulty establishing a territory in this area, and as 'floaters' are less likely to call until they become territory holders (Colbourne & Kleinpaste 1984, Taborsky & Taborsky 1991). In addition, call rates can reflect population stability, and have been observed not to change with population density in other kiwi species (Rogan Colbourne, Hugh Robertson, New Zealand Department of Conservation, pers. comm., 2012).

TIME OF NIGHT   Within each night, LSK show similar calling patterns to brown kiwi (Colbourne & Kleinpaste 1984, Miles et al. 1997). The greatest calling activity is in the first hours after dusk, with calls decreasing during the night, before a small increase prior to dawn. Timing calls from sunset means that the pre-dawn peak is diminished due to night length changes during the year. These calling peaks are comparable with the dusk and dawn chorus in diurnal birds, suggesting that they may have similar causal factors, such as optimal sound transmission (although see Brown & Handford 2002), poor feeding conditions, territory defence, and mate-guarding (Larom et al. 1997, Catchpole & Slater 2008). However transmission conditions are not optimal during the evening peak in kiwi calls, kiwi foraging is unaffected by light levels, and mate-guarding is an unlikely driver since similar patterns are observed all year in brown kiwi (Miles et al. 1997) and LSK (A. Digby, pers. obs.). Territory defence and pair contact are more likely causes of the post-dusk and pre-dawn peaks in kiwi vocal activity (Colbourne & Kleinpaste 1984). That the dawn peak is much weaker than that at dusk (even allowing for variable night length) is in accordance with a mate contact rather than territory defence function. Assuming that all indi-

viduals marking territory at the beginning of the night would do so at the end, then the dusk and dawn peaks should be comparable. However, if only individuals sharing a burrow with a mate call for contact just prior to dawn, then the calling rate would be expected to be much less, since pairs share burrows only a small proportion of the time: $\sim$ 8% for brown kiwi (McLennan et al. 1987), and $\sim$ 20% for LSK (Hugh Robertson, New Zealand Department of Conservation, pers. comm., 2011).

5.4.2.3 *Environmental conditions*

Although correlations between calling activity and environmental conditions are known for brown kiwi (Miles et al. 1997, Pierce & Westbrooke 2003), this is the first study to measure call variation in any other kiwi species. It also provides the first assessment of kiwi vocal activity from continuous sampling over an extended period.

WIND AND RAIN   The very strong effects of wind speed and rainfall on the observed counts were very likely a result of the acoustic detection methods, which are known to be adversely affected by strong wind and rain (Miles et al. 1997, Pierce & Westbrooke 2003, Bas et al. 2008). Detection rates showed a sex bias with wind and rain, with female calls masked more by wind, and males more by rain. The wind effect is expected: the lower frequencies of female calls are masked more by wind noise than the higher male call frequencies. Rain noise is more broadband, so affects the sexes more equally than wind, but why it should obscure male calls more is unclear. The slight upturn in LSK call rates during very heavy rain may be a result of increased false positives. No previous studies have quantified the difference in detection between the sexes with increasing wind and rain, which should be accounted for when analysing count data. The changes in calls detected with wind direction are likely a result of the study site topography. Wellington frequently receives southerly and northerly gales, and the recorder site was south-facing, protected from the northerly winds.

GROUND MOISTURE   The sharp increase in calls that occurs from dry to wet ground conditions mirrors the increase in call rates with ground moisture reported in brown kiwi (Pierce & Westbrooke 2003). This levels off in LSK calls, suggesting a dichotomy in call rates between wet and dry ground conditions. This difference may be related to food availability. LSK feed predominantly on invertebrates from the top layers of

soil, and seasonal changes in prey consumed suggests that soil drying affects feeding habits (Colbourne et al. 1990). Seasonal variation in foraging habitats in brown kiwi may also be related to soil moisture (Cunningham & Castro 2011). Wetter soil conditions would improve food availability for LSK, and so more foraging activity might prompt higher calling rates. Alternatively, kiwi may be harder to detect in winter compared to summer because damp leaf litter might muffle their movements (Cunningham & Castro 2011). LSK may therefore call more in wet ground conditions because after the call they are less likely to be detected when foraging by predators or aggressive conspecifics.

TEMPERATURE, HUMIDITY AND PRESSURE   The significant effect of increasing calls with increasing relative humidity may reflect the association with ground moisture, or result from improved sound transmission at higher humidities (Marten et al. 1977). The decrease for very high humidities is likely due to rain inhibiting detection. The decrease in calls with temperature is hard to explain from meteorological conditions. Cooler temperatures induce slower sound speeds and hence less efficient transmission, although atmospheric conditions associated with colder temperatures (such as inversion layers) also influence sound propagation. Any ecological explanation for the temperature relation is also obscure, since the trend is counter to the seasonal variation for males (which dominate the counts), in which call rates are lower during the cooler months. It is also opposite to the decrease in brown kiwi calls with decreasing temperature reported by Miles et al. (1997), although they combined rain and temperature conditions. The relationship between calls and pressure is similarly difficult to account for, and may relate to atmospheric conditions affecting sound transmission. Alternatively, lower pressure may be related to moist ground conditions and subsequent higher calling rates, but this does not explain the apparent increase in calls with higher pressure. Note also that all meteorological variables in the model show no strong cross-correlation (§ 5.2.4.2).

LIGHT LEVELS   LSK showed a strong decrease in vocal activity with increasing cloud cover. This may be a result of the location of the study site, which at 2 km from Wellington city centre receives significant light pollution. This can have a wide range of deleterious impacts (Longcore & Rich 2004), and reflected light levels increase with cloud cover (Kyba et al. 2011). Cloud cover may therefore indirectly influence kiwi activity. However, this response to light levels is at odds with moon phase having no effect on LSK calling rates. Some brown kiwi studies report a similar lack of moon

effect (Miles et al. 1997, Pierce & Westbrooke 2003), but many authors report that kiwi are less vocal on bright moonlit nights (e. g. Buller 1888, Clark 1952, Colbourne & Kleinpaste 1984). Furthermore, LSK are probably the only kiwi species that were susceptible to predation from a nocturnal volant species (the laughing owl; Worthy 2001, Worthy & Holdaway 2002), and so would be expected to avoid activity on moonlit nights, in common with other nocturnal birds (Brigham et al. 1999, Mougeot & Bretagnolle 2000). A likely explanation for this paradox is that light pollution is dominant over any effect of the moon. Cloud cover can increase sky luminance by a factor of ten within cities, and yield ambient light levels many times brighter that that caused by a bright moon (Kyba et al. 2011). These results therefore suggest that LSK behaviour at the study site may be affected by light pollution.

5.4.2.4   *Other factors affecting calling activity*

Although the model showed many significant environmental effects on call rates, it only explained about a third of the variation in the data. Brown kiwi are similarly known to exhibit large variations in calls under identical weather conditions, even on consecutive nights (Colbourne & Kleinpaste 1984). This is to be expected in a species for which calls are a dominant form of communication. Social factors such as mate contact and territory defence are likely to be far greater influences on vocal behaviour than environmental conditions. For example, kiwi calls often cascade in a 'domino effect', with one call setting off a chain of others (A. Digby, pers. obs.). Call rates will also be affected by stochastic processes, such as the foraging path taken by a neighbouring kiwi that may or may not prompt a territorial call. Indeed, it is somewhat surprising that weather conditions affect calling behaviour at all, and that cloud cover, for example, could have as much impact on calls as season or time of night.

5.4.3   *Conservation implications*

The environmental and temporal influences determined by this study can guide acoustic monitoring of LSK. In order to optimise detection conditions, counts should take place at the beginning of the breeding period (September – October), and include the second hour after sunset. The current practice of not monitoring in strong wind or rain is reinforced by the very strong effects on counts and the sex bias shown here. However, there appears to be no benefit in only counting during a dark moon phase in this species, at least at this study site. Counts would further be maxim-

ised by being conducted in conditions of cooler air temperatures and damp ground conditions. The variations with season and light conditions may be site-specific, so it would be prudent to repeat a similar experiment, over a shorter time frame, at another LSK population.

The significant variation in calling rates with most weather variables suggests that censuses would benefit from accurate meteorological records during call counts, rather than broad categorical conditions currently recorded (e. g. cold/mild/warm for temperature). Such information is freely available from nearby automatic weather stations, or can be easily obtained from inexpensive portable thermometers and hygrometers. The apparent sensitivity of LSK to light pollution suggests that a direct measurement of ambient light levels would also be beneficial.

Despite their limitations in adverse weather conditions, these results highlight the benefits of acoustic methods. The changes in LSK vocal activity with environmental conditions, assessed passively over a long time period, provide a unique behavioural insight not available from other monitoring methods.

AUTONOMOUS MONITORING METHODS FOR KIWI[1]

SUMMARY

1. Autonomous acoustic recorders are widely available and can provide a highly efficient method of species monitoring, especially when coupled with software to automate data processing. However, the adoption of these techniques is restricted by a lack of information quantifying how their use can change the results of existing manual surveys.

2. We assessed the performance of semi-automated and fully-automatic acoustic techniques by direct comparison with a manual survey, using commercially-available autonomous recorders and novel detection and classification software. We compared the detection capability, time requirements, areal coverage and weather condition bias of these three methods using an established call monitoring programme for a nocturnal bird, the little spotted kiwi (*Apteryx owenii*).

3. The automated methods had very high precision (>98%) and required less than 3% of the time taken by the counts with field observers. The areal coverage of the semi-automatic and fully-automatic methods was 85% and 42% of the manual survey. The autonomous methods were more adversely affected by wind, and did not show a positive association between ground moisture and call rates that was apparent from the manual counts. However, all methods produced the same results for the most important conservation information from the survey: the annual change in calling activity.

4. Autonomous monitoring techniques can produce different inferences if sampling is not adjusted accordingly. Nevertheless, the sensitivity, robustness and high accuracy of automated methods demonstrates that they offer a suitable and extremely efficient alternative to field observer point counts for species monitoring.

6.1 INTRODUCTION

6.1.1 *Call counts*

Point call counts with field observers are one of the most common methods of measuring abundance or presence or absence of birds (Rosenstock et al. 2002). Such counts are prone to several well-known biases. Observers vary greatly in their ability to detect and classify calls due to differences in age, experience and hearing (Emlen & Dejong 1992, Rosenstock et al. 2002, Hutto & Stutzman 2009), and errors become more prevalent with high call rates (Hutto & Stutzman 2009). The presence of observers can also affect vocal activity through disturbance of the study species. Counts usually take place during fine weather to maximise detections and minimise variability, but can still be biased by weather conditions (Bas et al. 2008). Since counts

| Chapters | Questions | Themes |
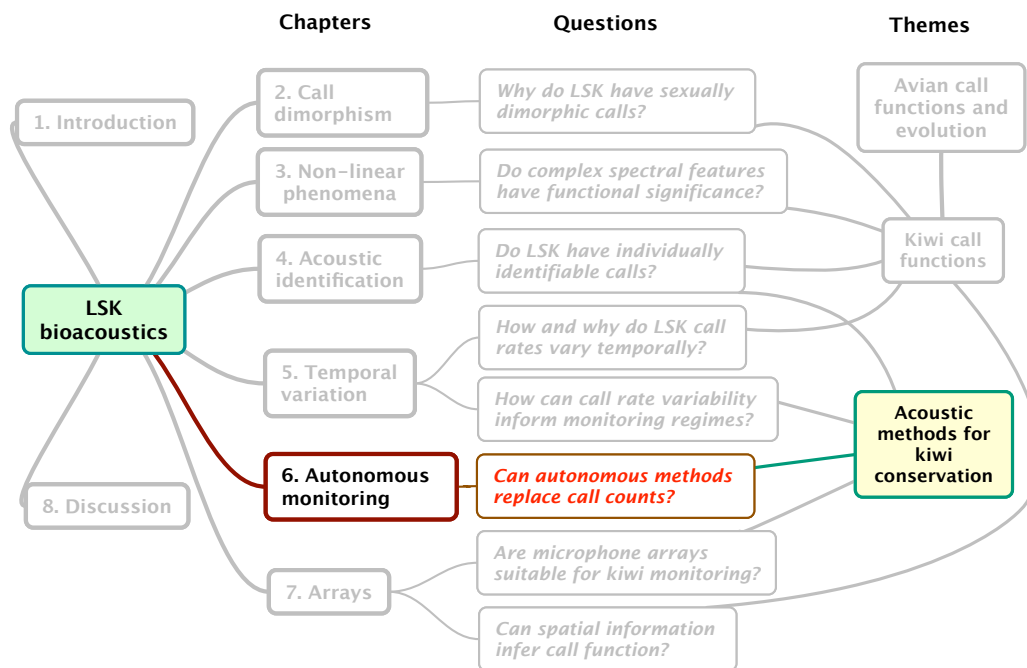
Topics covered by this chapter.

are usually short, they are prone to temporal biases arising from daily or seasonal activity variation (Diefenbach et al. 2007, Bas et al. 2008). This can lead to behavioural information being missed or population sizes significantly underestimated (Bridges & Dorcas 2000). In addition to these biases, manual counts require substantial observer effort, which can be expensive, particularly when trained observers are required to monitor in remote areas.

6.1.2 *Autonomous recorders*

Autonomous acoustic recorders offer an alternative not subject to many of the biases of manual call monitoring. By not requiring the presence of an observer they can be less expensive (Charif 2008), can be used in inaccessible or inhospitable habitats (Hutto & Stutzman 2009), avoid bias from subject disturbance (Alldredge et al. 2007), and minimise temporal bias through extended sampling. Acoustic recorders can provide reliable data more rapidly than human-based survey techniques (Parker 1991, Riede 1993, 1998), and can perform better at sampling bird communities than skilled observers using audio-visual field point counts (Celis-Murillo et al. 2009). Recorders provide a permanent data record for reanalysis by independent observers

(Swiston & Mennill 2009), and combined in microphone arrays (Mennill et al. 2012) can yield spatial and behavioural information that is not available to an observer in the field (Fitzsimmons et al. 2008b, Mennill & Vehrencamp 2008a, Kirschel et al. 2011).

A principal obstacle to the widespread use of autonomous acoustic recorders is that they require substantial expert time to process the recordings. Although this can be substantially less time than required for field point counts (Charif 2008), autonomous surveys sample over much longer time periods and so often yield very large data volumes.

Automatic species detection and classification software provide an efficient method of processing such large datasets. A variety of methods are available (Brandes 2008), with some accessible through dedicated software packages (Charif 2008). These methods present two principal issues. One is that classification methods such as Hidden Markov Models and neural networks rely on large training datasets, which can be restrictive for species with low call rates (Acevedo et al. 2009, Towsey et al. 2012). The second is the difficulty in minimising the number of false negative and false positive errors (Waddle et al. 2009, Bardeli et al. 2010).

A further barrier preventing wide-scale adoption of autonomous acoustic methods for conservation management is that recorders usually have a lower sensitivity than a human listener (Hutto & Stutzman 2009). This can be remedied by modified sampling, but necessitates that their efficacy be assessed against existing monitoring methods. This has not been undertaken for most species, and is rarely performed in the context of established monitoring programmes.

6.1.3 *Kiwi monitoring*

Of particular benefit for cryptic vocal species, point call counts are an established tool for monitoring kiwi (Apterygidae). Call surveys have been used in kiwi conservation management to detect previously unknown remnant populations at risk of extinction (Jolly & Colbourne 1991) and to measure response to predator control or advocacy (Robertson et al. 1993, Pierce & Westbrooke 2003, Robertson & Colbourne 2004). They are widely used to assess kiwi distribution and population trends throughout New Zealand (Robertson & Colbourne 2003, Robertson et al. 2005, Colbourne 2006). The vast majority of surveys use manual counts, with observers listening for 1–2 hours on up to eight nights per year at designated sites (Robertson & Colbourne 2003). As a result these surveys are prone to temporal sampling bias, with popula-

tion estimates dependent on the density of listening stations, the length of counts, and the number of consecutive nights on which surveys take place (Colbourne & Kleinpaste 1984, Miles et al. 1997). Call rates can also vary with weather conditions (Pierce & Westbrooke 2003) and moon phase (Miles et al. 1997). Autonomous recorders have recently been used for kiwi surveys (S. Cockburn, New Zealand Department of Conservation, pers. comm., 2011), but no published results exist of evaluation against existing count methods.

6.1.4 *This study*

In this paper we compare autonomous acoustic recorders and call detection software with manual methods, using an established kiwi call count programme as a test case. We assess method sensitivity and reliability from a conservation monitoring perspective, and measure variation in performance due to different ambient conditions. By application to a real monitoring situation, these results provide an applied and quantitative assessment of the efficacy of automated methods for conservation monitoring of vocal species.

6.2 MATERIALS AND METHODS

6.2.1 *Call counts*

Since 2002, kiwi call counts have been conducted annually between January and March at Zealandia, a 225 ha mammalian predator-free fenced wildlife reserve in Wellington, New Zealand. In 2000–2001, 40 little spotted kiwi (*Apteryx owenii*; LSK) were translocated to the sanctuary from nearby Kapiti Island, and a 2010 census estimated the population size at ~100 (H. Robertson, New Zealand Department of Conservation, pers. comm., 2010). Counts follow the protocol described in the national Kiwi Call Count Scheme (Robertson & Colbourne 2003). From two sites in the reserve, two or three observers record the sex, estimated distance and magnetic bearing of all kiwi calls heard for an hour from 30 minutes after sunset. Weather conditions are recorded (Table 6.2.1), and surveys only take place on nights with low or moderate background noise. Established procedure at Zealandia requires that the observers were varied as much as possible between counts.

Table 6.2.1: Environmental variables recorded by the observers during each call count

Variable	Values
Temperature	Cold, Mild, Warm
Rain	Nil, Light, Moderate
Wind	Calm, Light, Moderate, Strong
Cloud cover	Clear, Partly cloudy, Overcast
Light	Light, Dark, Black
Noise	None, Slight, Moderate
Ground condition	Dry, Damp, Wet

6.2.2 *Call recording*

On 53 nights from 17th January – 16th March 2011 and 18th January – 30th March 2012, an autonomous recorder (Song Meter 2, Wildlife Acoustics Inc., Concord, MA, USA) was positioned near a kiwi count site in Zealandia. This site was on top of a 10 m high viewing platform above the tree canopy, on a spur in the central part of the reserve (S41.29776 E174.74432). Due to security considerations and the likely impact of wind noise on the microphones, the recorder was placed next to the platform at 2 m above ground level. The recorder was therefore approximately 10 m below and 5 m away horizontally from the observers. To assess if this distance affected the likelihood of detecting kiwi calls, for 27 nights in 2012 a microphone connected to one channel was positioned in the tree canopy 2 m horizontally from the platform edge. By comparing calls detected by each channel, the suitability of the ground-level listening position for comparison with the canopy-level manual call counts could be assessed.

Recordings were conducted for 70 minutes each night, starting 5 minutes before and ending 5 minutes after each count. Recordings were made in stereo (microphone frequency response = 20–20000 Hz), and were digitised at 16kHz, 16-bit precision. A hardware gain of +48 dB was applied to each channel.

6.2.3 *Spectrographic analysis*

The recordings from the count were analysed spectrographically by an experienced observer (AD). Spectrograms were viewed in Raven Pro 1.4 (Charif et al. 2010)

with 512-sample Hann window, 31.3 Hz spectral resolution. Files were viewed in 15 minute segments, and to improve detectability of LSK calls against background noise, spectrograms were smoothed with ten-sample averaging. Each LSK call was marked and annotated with its sex and a subjective signal-to-noise quality score, ranging from 1 (very high signal-to-noise) to 10 (call barely visible or audible). Detection was primarily conducted from visual inspection of the spectrograms, but aural confirmation was used for very faint calls.

6.2.4   *Automatic call detection*

Calls within the recordings were detected using custom software optimised for detection of LSK. Although detectors for multiple species are available, customised recognisers usually offer better performance for autonomous monitoring in acoustically unconstrained environments (Bardeli et al. 2010, Towsey et al. 2012). The LSK software consisted of a two-stage process of event detection and event classification.

6.2.4.1   *Event detection*

LSK calls consist of a long series of repeated rising syllables or elements that persist for 20–40 seconds (Digby et al. 2013; Chapter 2).The sexes call in largely non-overlapping frequency ranges (Digby et al. 2013; Chapter 2), hence it was possible to search for male and female calls in separate bands (2.2–3.2 kHz and 1.3–2.2 kHz respectively) using the same algorithm. The first step was to locate potential LSK calls by calculating a cross-correlation coefficient to detect repeated syllables above a threshold of amplitude, periodicity and duration. Several other bird species have calls with a similar repetition of elements in the same frequency band. It was therefore necessary to extract additional features concerning change in periodicity over duration of call, frequency modulation of each call element, change in amplitude across elements and signal to noise ratio. A minimum event length was set at 7 s and, to avoid duplicate detections of calls containing pauses, events were merged if they were within 5 s of another event in the same frequency band. A total of seven features were calculated for each potential call event. An eighth feature was a weighted combination of the above normalised feature scores. Software was written in C#.

6.2.4.2   *Event classification*

Potential LSK calls were then further classified by a decision tree classifier, using `C5.0` software (version 2.09; RuleQuest Research, Australia). The decision tree was trained using 22 three-hour files with low background noise from the same site, randomly selected over a range of dates. LSK calls with the lowest quality score or which overlapped other biotic or abiotic sources were removed from the training set. This resulted in 459 training LSK calls (368 male, 91 female) longer than 7 s. For each event the true classification (kiwi or not kiwi) and all eight detector scores for each event were input to the classifier. Different methods of building the decision tree were evaluated from ten repeats of ten-fold cross-validation, and compared with error rate, recall and precision. Recall is defined as $TP/(TP + FN)$, and precision as $TP/(TP + FP)$, where TP = true positives, FN = false negatives, and FP = false positives. The final classifier used a decision tree with adaptive boosting and no softening. It had a mean error rate of 0.9%, and a cross-validated recall and precision of 93.4% and 97.1% respectively. This recall value on training data was higher than would be expected in operational conditions due to the phenomenon of 'over-learning' (learning specific features of the training data that are not actually relevant in the operational environment). In addition the training data contained only unambiguous calls (quality score of at least nine) to assist learning. This decision tree was used to classify test data containing previously 'unseen' events in call count files.

6.2.5   *Comparing counts*

The calls recorded by the call count observers (manual detections) were matched with detections from the manual spectrogram scanning (semi-automatic detections) and classification software (automatic detections). For each method, detections were marked either as true positives, false positives, or false negatives (Figure 6.2.1). The manual counts were taken to be the truth, except for sexing calls, and in cases when a call was recorded as close or medium distance (<250 m) by the observers, but no LSK call and a call from another species was apparent in the spectrogram at the time. Species at the study site that can be confused acoustically with LSK are ruru (*Ninox novaeseelandiae*), a native owl, and kaka (*Nestor meridionalis*), a native parrot. In these cases, and when the sex was recorded incorrectly, the manual detection was marked as a false positive. This approach unavoidably underestimates the number of

false positives and overestimates the true positives for the manual detections, since distant calls not visible on spectrograms could not be verified.
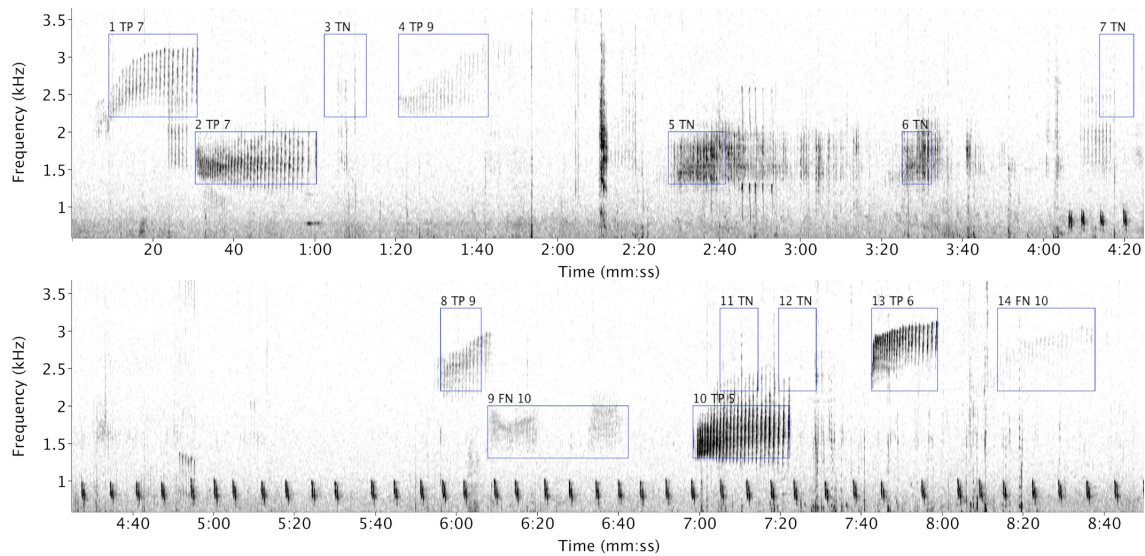
Figure 6.2.1: A section of a spectrogram from the autonomous recorder, with detections from the automatic software. The spectrogram is split into two contiguous lines. Annotations above each detection give an event number, a verification of the detection (TP = true positive, FN = false negative, TN = true negative), and the call quality (1–10) for kiwi calls. Events 1, 4, 8, 13 and 14 are male LSK calls, and events 2, 9 and 10 are female LSK calls. The faintest kiwi calls (9 and 14) are not detected by the automatic software, but were by the other methods. Events 3, 5 and 6 are duck calls. The repeated signals in the lower line at ~0.8 kHz are from an owl (*Ninox novaeseelandiae*).

6.2.6 *Recorder detection limits*

The distance limits at which the semi-automatic and automatic methods could detect LSK calls was determined by broadcasting kiwi calls at realistic amplitude at varying distances from the recorder. The appropriate speaker volume was determined by playing a LSK call recorded at known close range (<5 m) at different volume settings, and recording it from the same distance with the same equipment as used to make the initial recording (Sound Devices SD722 recorder with Telinga Twin Science microphone and Telinga 53 cm foldable parabola). The appropriate volume setting was that for which the re-recorded power spectra most closely matched that of the original call over the peak energy range. With appropriate volume settings determined for each sex, male and female LSK calls were played at 100 m intervals

at distances ranging from 100–1200 m from the Song Meter. Calls were broadcast using a Sony PCM-M10P recorder and a MIPRO MA-101 45-watt loudspeaker. Tests were conducted at night in low background noise conditions. The distance limits were determined by manually and automatically scanning the spectrograms from the autonomous recorder, and ascertaining the distance threshold beyond which the known calls were not detected.

6.2.7 *Count variation with ambient conditions*

Generalised linear models were used to assess variation in detection with ambient conditions, allowing for an annual trend. A different model was calculated for each detection method (manual, semi-automatic, and automatic). The response variable was the number of calls detected during each count, and explanatory variables were year (2011 or 2012) and environmental conditions recorded by the observers, except wind direction (Table 6.2.1). Poisson models with a log link were fitted using function glm in R (version 2.15; R Development Core Team 2011). Observations from 52 nights were used; one night with a count abandoned due to poor weather was removed. Full models were checked for overdispersion and adequacy (Zuur et al. 2009). Model selection followed an informatic-theoretic approach (Anderson et al. 2000), with models fitted for all possible combinations of explanatory variables without interactions. These were ranked by corrected Akaike Information Criteria (AICc; Hurvich & Tsai 1989), and a 95% confidence model set was selected based on cumulative Akaike weight. Inferences for the three detection methods were then based on the weighted average model and the relative variable importance calculated from this set.

6.3   RESULTS

6.3.1 *Time expenditure*

Spectrogram inspection took 2–5 minutes per hour of recording, depending on the call rate. The detection software took approximately two minutes per recording hour on a four-core 2.4 GHz computer. Replacement of recorder memory cards and batteries was required about every six weeks. Including these resupply times, the total time required for the semi-automatic and automatic detection methods was less than one person-hour per week. The manual counts required at least 28 person-

hours per week, with two people spending one hour listening and one hour walking to and from the site for each night. This is a conservative estimate, and does not include additional time in processing the hand-written count results and organising observers.

6.3.2 *Comparison between ground-level and canopy sites*

A comparison of the ground-level and canopy recording sites yielded 348 calls (291 M, 57 F) from the canopy microphone, and 345 (284 M, 61 F) from the microphone near ground level. The median difference in the number of calls detected at each site for each count was not significantly different from zero (Wilcoxon paired signed rank test, $V = 94.5$, $P = 0.702$). The ground level microphone site was therefore comparable with the observer listening position, and so recordings from this site were used for all analyses.

6.3.3 *Detection distance limit*

The simulated kiwi calls of both sexes were reliably detected from manual spectrogram inspection to at least 400 m. Some calls were discerned up to 600 m, but beyond 400 m were very faint on the spectrograms, and not evident for all repeats at that distance. The detection software consistently located calls of both sexes to 300 m, and some up to 400 m.

6.3.4 *Detection performance*

A total of 900 calls (721 male, 171 female, 8 sex unknown) were detected by all three methods over 52.2 hours on 53 evenings. Recall of this total ranged from 40% for the automatic detection, to 94% for the manual counters (Table 6.3.1). The semi-automatic and automatic methods detected 85% and 42% respectively of the calls heard by the human listeners. The automated methods detected proportionally more female calls: 93% and 47% of the female calls heard, compared to 84% and 42% of male calls. The observers missed 6% of the total calls, and misclassified 4% of the calls they detected. The automatic detector found one very faint call which was missed by both the other methods; otherwise false negatives from the semi-automatic detection were also false negatives for the automatic method.

Table 6.3.1: Detection scores of the three methods, with true positives (TP), false positives (FP), false negatives (FN), recall (R) and precision (P) shown for the total of 900 calls. Note that the manual false positives are a lower estimate, since for distant calls accuracy could not be verified from spectrograms.

Method	TP	FP	FN	R (%)	P (%)
Manual	849	36	51	94.3	95.9
Semi-automatic	723	0	177	80.3	100.0
Automatic	358	7	542	39.8	98.1

The detection distance limit of at least 400 m for manual spectrogram inspection equates to a minimum areal coverage of ∼50 ha for semi-automatic detection. Assuming a uniform density of calling kiwi, the recall values indicate that the manual and automatic methods therefore sampled calls from areas of ∼59 ha and ∼25 ha respectively. This latter figure agrees with the areal coverage of 28 ha calculated from the 300 m detection limit for automatic detection.

As expected, calls not detected by the semi-automatic method were those at further distances: 85% of those missed were marked as > 250 m by the observers. Similarly, the calls not detected by the automatic method, but visible in the spectrograms, tended to be faint calls with low quality scores (Figure 6.3.1), or those which overlapped with other signals at similar frequencies.

6.3.5 *Variation with year and ambient conditions*

The inferences from the average generalised linear models differed between detection methods. Ground moisture, cloud cover and background noise significantly affected calls detected by the field observers, whereas only wind strength affected semi-automatic detection, and noise and wind strength influenced automatic counts (Figure 6.3.2 and Table 6.A.1). However, the sign of the effects on calls detected were the same for all models, with counts reduced with increasing wind strength, noise and cloud cover, and dry ground. The most important variables influencing detection were ground conditions for manual counts, and wind strength for semi-automatic and automatic detection (Figure 6.3.3).
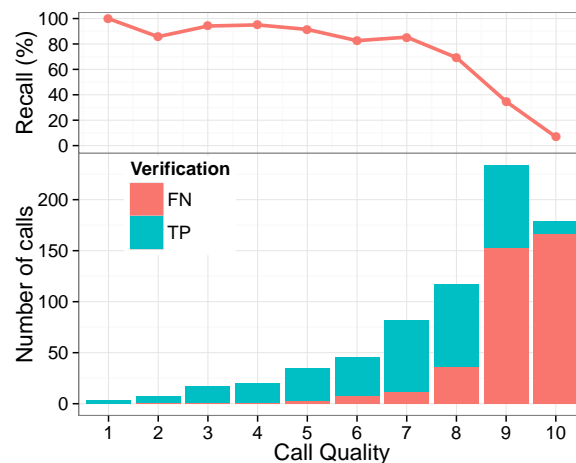
Figure 6.3.1: The recall and number of calls that were either detected (TP) or not detected (FN) by the automatic software, as a function of subjective spectrographic call quality (10 is low quality). Bars are stacked, so total bar height shows the number of calls in each quality bin. The majority of calls not detected are those in the two lowest quality bins (9 and 10).

6.4 DISCUSSION

As expected, kiwi call counts conducted by observers in the field were more sensitive than those from autonomous recorders. Misclassification and false negative rates for the manual counts were low, confirming that counts using trained volunteers, many with no previous experience of bird counts, can yield accurate results. While autonomous recorders were less sensitive, their areal coverage was comparable with manual surveys, and adequate for many applications. For example, assuming an LSK territory size of 3 ha for a dense population (Colbourne & Robertson 1997, Holzapfel et al. 2008), the semi-automatic survey was able to monitor approximately 16 territories, compared to 19 for the manual count. The fully automatic survey samples about half the number of territories, but this can be increased simply by using more recorders. The automatic method is precise and robust, with very few false positives that appear to be unaffected by noise conditions, which often cause problems for automatic signal detection and classification (Bardeli et al. 2010).

It should be noted that the performance of the manual counters is likely overestimated in this study, since distant detections could not be confirmed by the recorders. Conversely, the semi-automatic recall was a lower limit, since some very faint calls that were missed during spectrogram scanning were seen upon re-inspection, and may have been detected with more careful spectrogram inspection. However, these
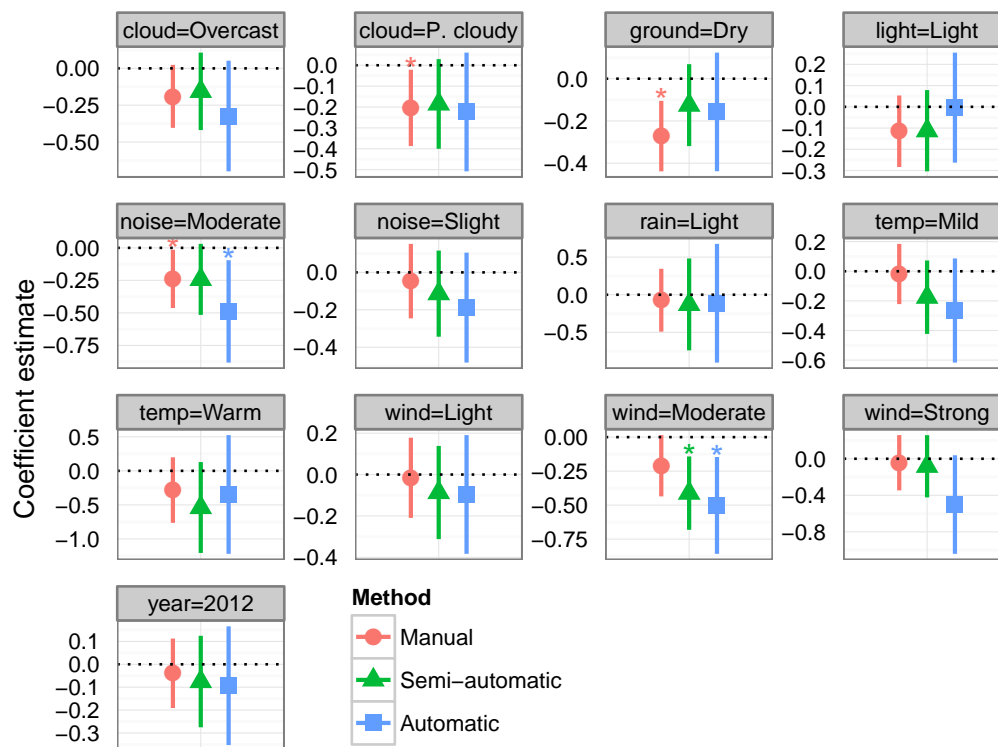
Figure 6.3.2: Effect coefficient estimates and 95% confidence intervals for the average models determining the influence of year and environmental conditions on counts from each detection method. Effects which significantly impact counts are those for which the coefficient confidence range do not intersect zero, and are marked by an asterisk.
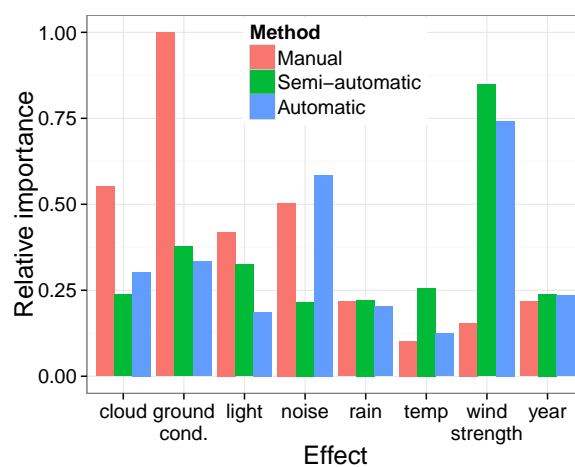
Figure 6.3.3: The relative importance of the year and environmental variables in the average model for each method.

caveats would make at most only minor adjustments to the performance of each method and do not alter the overall conclusions.

Automatic detection software is ideally suited to species such as LSK that have low call variability. For animals with a wider repertoire, such as song birds, or for detection of multiple species, a larger training set is necessary and detector performance is likely to be reduced (Bardeli et al. 2010). A further restriction is that the detection software in this study was often unable to resolve calls that overlapped with other sources, either conspecifics or other species. This would limit its usefulness in measuring abundance in areas of high population density, or where there are high numbers of acoustically competing species. However, manual counts are also prone to such bias (Hutto & Stutzman 2009), which is caused by either the insufficiency of human auditory ability in temporally or spectrally resolving simultaneous calls, or the requirement that observers record call details as they occur. The latter also affects the quality of observations made, since call parameters were occasionally omitted or inaccurate during periods of high calling activity. It is significant that the semi-automatic method tended to perform better than the observers during high call rates.

The principal advantage of using autonomous recorders is the gain in efficiency, with the semi-automated survey around 30 times more efficient than traditional counts. These substantial gains in productivity offered by recorders could be utilised by sampling over longer periods each night and at other times of year to reduce temporal biases. The only significant time input required for the fully automated method are in resupplying recorders and data management, and a nightly single point survey could be achieved with a time expenditure of only about an hour per month. For remote areas or in determining species presence or absence, this method is extremely attractive.

The automatic recorder surveys also offer a major benefit in reduction of temporal and observer bias that can affect manual surveys. The automatic detection method is fully repeatable as it requires no subjective input at all, and so is well-suited for assessing variation in counts between sites or time periods. Minimal disturbance of the target species, and a permanent data record, are further benefits of the autonomous methods.

The generalised linear models suggested that wind noise affected the autonomous detections more than the manual counts. This may have been due to the observers being slightly above the canopy, and so surrounded by fewer scattering surfaces

and sources of noise in windy conditions than the microphones, which were below canopy level. Also, the microphones were static, whereas the observers were able to move to minimise the effect of wind noise. This reflects an inherent restriction of acoustic recorders, which are adversely affected by wind noise on microphones, and should be placed in a sheltered position.

The significant impact of ground conditions on the manual counts, but not on the autonomous counts, is a more serious concern. Pierce & Westbrooke (2003) reported an increase in call rate with increasing ground moisture index for brown kiwi (*Apteryx mantelli*), suggesting that this is a real effect. The inability of the autonomous methods to detect a significant influence may be a result of their reduced sensitivity, and subsequent lower number of counts and wider effect confidence intervals. Additionally, higher call rates under damp or wet ground conditions may have reduced the recall of the fully-automated method. However, the ground moisture coefficient estimates for the autonomous methods are similar to the manual counts, albeit non-significant. A longer sampling period is necessary to sufficiently reduce confidence intervals for the semi-automatic and automatic methods. This would be easily done given their efficiency, and it should be noted that the short recording period of one hour per night is artificially low, used only to provide direct comparison with the manual method.

The year effect would be the one of main interest in most conservation applications, and shows no difference between the three methods. It is slightly surprising that no call increase is apparent between years, given that the LSK population in Zealandia is growing at ~10% per annum (H. Robertson, New Zealand Department of Conservation, pers. comm., 2010). However, counts taken over many years would be required to measure such a trend, and call rates do not always follow population changes (Robertson & Colbourne 2003).

6.4.1 *Recommendations*

This study has evaluated the use of autonomous recorders for point call counts, applied to a typical, volunteer-run monitoring programme, rather than one-off research experiments. These results demonstrate that while trained lay observers provide accurate surveys, autonomous recorders offer a viable alternative. Semi-autonomous surveys can yield comparable coverage to observers for considerably less effort, and fully automated methods offer robust alternatives that are well-suited to longer-term projects and those in remote or unmonitored areas. Our modelling results show that

autonomous methods can provide similar ecological inferences to manual counts, but that care must be taken to ensure sufficient temporal sampling to adjust for the different biases these are subject to.

Manual call counts can offer significant benefits for conservation advocacy and community engagement. We recommend that autonomous recorders do not replace these, but are utilised to increase the spatial and temporal coverage of existing call count regimes. Acoustic recorders should also be used during call counts to verify the accuracy of observations (Hutto & Stutzman 2009), particularly during periods of high call rates. Autonomous recorders would also be suitable for areas where existing surveys are not established, and are particularly appropriate for determining presence or absence of a species in remote areas where call counts are costly and inefficient.

6.A APPENDIX

Table 6.A.1: Effect coefficients for the averaged generalised linear models of calls detected for each method. The deviance and AIC values for the best-ranked model in each 95% confidence set are also shown. *P* values marked in bold are significant at the 5% level. Est. = estimate; s.e. = standard error; Dev. = deviance.

	Manual				Semi-automatic				Automatic			
	Est.	s.e.	*z*	*P*	Est.	s.e.	*z*	*P*	Est.	s.e.	*z*	*P*
Intercept	2.97	0.18	15.9	**<0.001**	2.63	0.22	12.0	**<0.001**	1.97	0.36	5.3	**<0.001**
cloud=Overcast	-0.19	0.11	1.7	0.082	-0.09	0.11	0.8	0.450	0.49	0.20	2.4	**0.015**
cloud=P. cloudy	-0.20	0.09	2.2	**0.028**	-0.41	0.13	3.0	**0.003**	0.30	0.16	1.8	0.076
ground=Dry	-0.27	0.08	3.2	**0.001**	-0.08	0.17	0.5	0.635	-0.10	0.14	0.7	0.511
light=Light	-0.12	0.08	1.3	0.181	-0.11	0.09	1.2	0.249	-0.50	0.18	2.8	**0.006**
noise=None	0.24	0.11	2.1	**0.037**	-0.13	0.10	1.3	0.205	-0.50	0.27	1.8	0.069
noise=Slight	0.19	0.10	1.9	0.055	-0.16	0.13	1.2	0.245	-0.32	0.19	1.7	0.092
year=2012	-0.04	0.08	0.5	0.612	-0.19	0.11	1.7	0.091	-0.22	0.14	1.5	0.123
rain=Nil	0.07	0.21	0.3	0.732	-0.18	0.12	1.4	0.166	-0.16	0.14	1.1	0.272
temp=Mild	-0.02	0.10	0.2	0.859	-0.54	0.33	1.6	0.114	-0.09	0.13	0.7	0.480
temp=Warm	-0.28	0.24	1.2	0.250	0.24	0.14	1.7	0.082	-0.00	0.13	0.0	0.977
windstr=Light	-0.02	0.10	0.2	0.874	0.13	0.12	1.0	0.302	0.11	0.39	0.3	0.779
windstr=Moderate	-0.21	0.11	1.8	0.067	-0.08	0.10	0.7	0.460	-0.27	0.17	1.5	0.139
windstr=Strong	-0.04	0.15	0.3	0.777	0.13	0.30	0.4	0.680	-0.35	0.43	0.8	0.435
	Dev.	AIC			Dev.	AIC			Dev.	AIC		
Best-ranked model	92.0	319.9			48.1	259.6			44.0	230.5		

MICROPHONE ARRAYS FOR KIWI MONITORING

ABSTRACT

Microphone arrays have traditionally been under-utilised in terrestrial ecology, but can provide rich population and behavioural information that is not available to other methods. Arrays are particularly well-suited for conservation of cryptic species, through the ability to monitor individuals, assess territory dynamics, and improve census estimates. Such spatial information is valuable in kiwi conservation, but, despite the reliance on acoustic methods for kiwi monitoring, arrays are not yet an established method. Reasons for this include the high cost, complexity and large amounts of equipment traditionally required. However, the advent of commercially-available time-synchronised autonomous recorders provides the potential for an af-

fordable and easily-deployed array. This chapter assesses the feasibility of using such equipment to spatially monitor little spotted kiwi (LSK; *Apteryx owenii*) in their natural habitat.

Accurate time delay determination is critical to array methods. Algorithms based on generalised cross-correlation were tested, and a modified phase-transform version was derived. Optimal microphone configurations were trialled, and the time-delay algorithms and equipment tested with triangulation in both open and forested habitats. A novel maximum-likelihood localisation method was developed for this application.[1]

The time delay algorithm worked well in open habitat, but was unable to reliably triangulate source locations in the forest. An alternative time-delay method using blind source separation was also tested, but performed poorly. As a result, localisation of kiwi calls in their typical habitat was not possible using this equipment, even with advanced time-delay calculation techniques. The main reason appeared to be the highly-reverberant environment of the forest environment, which was typical lowland podocarp New Zealand bush. Interference from other sound sources and the low call signal strength due to the necessarily wide array spacing were other likely sources of error. These difficulties indicate that array methods cannot be easily implemented for kiwi monitoring with standard equipment. However, kiwi calls are nevertheless well-suited to array methods, and the results from this study allow recommendations to be made to guide future investigations.

7.1   INTRODUCTION

7.1.1   *Array applications for ecology*

Microphone arrays have a long history of use in fields such as military and aviation, but have only relatively recently been applied to ecology (Blumstein et al. 2011). Using multiple microphones provides two main benefits. First, it enables discrimination of multiple sources, which can be used to improve estimates of population abundances and densities (e. g. Celis-Murillo et al. 2009, Dawson & Efford 2009). Secondly, knowledge of the exact position of each microphone, and the time delay in a source reaching each, allows determination of the source location. This latter

---

1 Paul Teal and Yusuke Hioka provided invaluable advice in devising the methods used in this chapter. The new maximum-likelihood localisation method was developed by Paul.
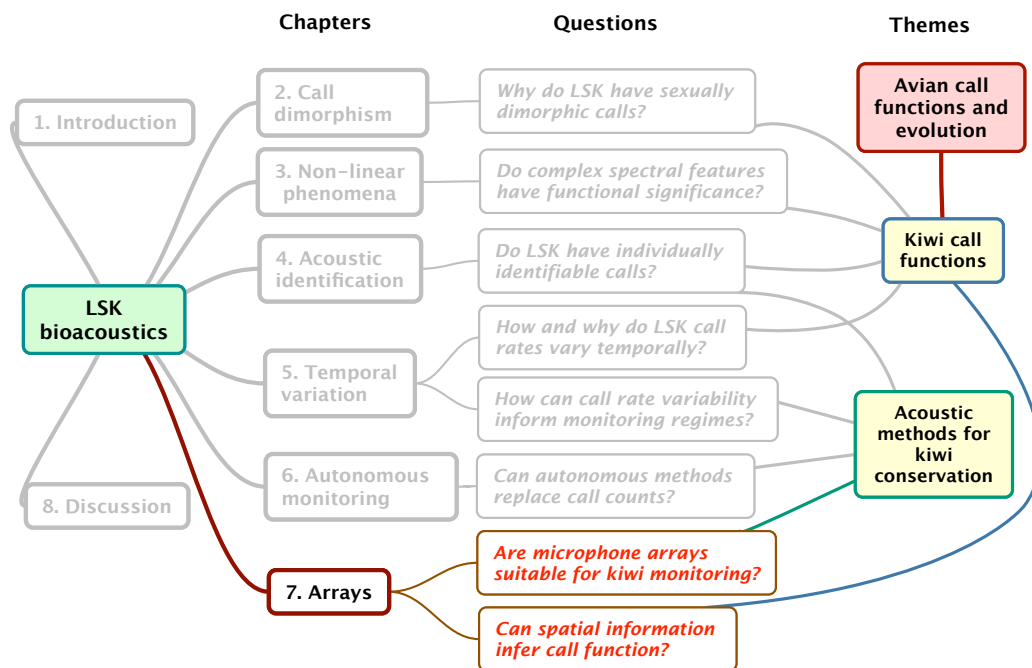
Topics covered by this chapter.

method in particular has many applications in ecology (Blumstein et al. 2011), but remains relatively under-used.

Arrays have had longest use in marine ecology, since they provide a means of monitoring cetaceans (Wahlberg et al. 2001, Muanke & Niezrecki 2007) even over large distances (Stafford et al. 1998). The ability to study unseen animals has clear applications in terrestrial ecology for monitoring species which are hard to monitor visually. Examples include nocturnal or cryptic species (Wahlberg & Møhl 2003), those in dense habitats (e. g. Mennill et al. 2006, Collier et al. 2010b) or which are sparsely distributed (Payne & Thompson 2003).

Acoustic array methods can provide more information about the ecology of a species than conventional, non-acoustic methods (Mennill 2011). They can yield insight into territory dynamics (Kirschel et al. 2011), or be used to measure abundances, densities or population structures (Payne & Thompson 2003, Dawson & Efford 2009). Spatial information from an array can assist identification of individuals in population studies (Payne & Thompson 2003, Bardeli et al. 2010).

Microphone arrays also offer many benefits to the field of behavioural ecology, and enable studies of animal communication not possible with any other method. For

example, arrays have been used to determine the dynamics of alarm calls in birds (Bower & Clark 2005) and mammals (Collier et al. 2010a). They are particularly useful for studies of finely synchronised antiphonal duets in tropical birds (Mennill et al. 2006). The ability to study the location and movement of calling individuals has enabled inferences to be made about duet functions (Mennill & Vehrencamp 2008a) and the use of counter-singing and eavesdropping (Fitzsimmons et al. 2008c,b). Directional information can provide insight into tactics and sexual selection in signalling behaviour, such as in courtship displays of lek species (Dantzker et al. 1999, Patricelli & Krakauer 2010). Arrays can also be beneficial in studies of signal structure: for example, Krakauer et al. (2009) used positional information from an array to determine the presence of two-voicing in sage grouse.

These insights into population and behavioural ecology provided by array methods have clear conservation applications. The ability to determine abundance, territory extents and dynamics, and to identify individuals, is central to monitoring and management of threatened species. In addition, arrays can be used to provide spatial information of invasive species (Hu et al. 2009). Arrays can also yield greater signal strength for acoustic monitoring: once a call is located, the array signals can be focussed on it to decrease noise and reverberations.

7.1.2 *Disadvantages and requirements*

Acoustic array methods are obviously only applicable to vocal species, and are less accurate for those with low signal-to-noise (SNR) or broadband calls (Møhl et al. 2001, Wahlberg & Møhl 2003). They are best suited to animals with loud, tonal, and frequency-modulated vocalisations (McGregor et al. 1997, Mennill et al. 2012).

There are several technical drawbacks to the use of microphone arrays in ecology. Arrays can require a large amount of equipment, particularly wired systems which can use many hundreds of metres or even kilometres of cable (Mennill et al. 2006, Fitzsimmons et al. 2008c). Wireless systems are more portable, but require very accurate time synchronisation between nodes (Chen et al. 2003). Multiple microphones result in the collection of large amounts of data, which can result in data management or signal detection issues. The calculations of time delays and source localisation can be complex, and there are several sources of error. The main uncertainties arise from inaccuracies in receiver position, both horizontal and vertical, and multi-path problems introduced by reverberations (Spiesberger 1999, Wahlberg et al. 2001, Bower & Clark 2005, Blumstein et al. 2011).

However, many of these issues can be overcome. Equipment for constructing microphone arrays is now relatively inexpensive and more widely available, with hand-held GPS units ubiquitous, and recorders offering time synchronisation commercially available (Mennill et al. 2012). Time of arrival and localisation algorithms have also improved, and offer solutions for dealing with multi-path errors (Spiesberger 1999, Mennill et al. 2006, Chen et al. 2006). With these advances, microphone arrays are becoming more commonly used in terrestrial ecology, and positional accuracies of under one metre are now possible, even in challenging conditions (Collier et al. 2010b).

7.1.3 *Arrays for kiwi*

Kiwi are nocturnal and cryptic, so locating individuals currently relies upon radio-telemetry methods, which are invasive, expensive, time-consuming, and can be prone to bias (Maloney 1996, Moorhouse & Macdonald 2005). With long, loud, repeated, tonal and frequency-modulated calls (Corfield et al. 2008a; Chapter 2), kiwi are ideal subjects for microphone array monitoring (McGregor et al. 1997, Mennill et al. 2012). Their flightlessness also assists, reducing the localisation problem to a two dimensional surface (Collier et al. 2010b).

Microphone arrays would benefit kiwi call censuses, which are the main method of large-scale kiwi monitoring (Colbourne 2006). Such surveys are conducted manually (Robertson & Colbourne 2003) or at most with single element recorders, and hence record multiple calls from individuals. Spatial information of calls would enable a more accurate translation between call rates and kiwi densities.

At least one kiwi species, the brown kiwi (*Apteryx mantelli*), has individually distinctive calls (Corfield 2005). As non-oscines, kiwi are not expected not to learn calls (Kroodsma 2005), and all species are long-lived (Heather & Robertson 2005), so their calls should remain stable over time. Microphone arrays could therefore be used to monitor individuals over extended periods. This would enable mapping of territories and study of territory dynamics in kiwi (e.g. Kirschel et al. 2011). Spatial information of duetting pairs would provide tests of duet function in kiwi (e. g. Mennill et al. 2006).

Despite the many benefits of using microphone arrays for kiwi research and conservation, it is not an established monitoring method. Only a few preliminary investigations have been conducted (Stuart Cockburn, New Zealand Department of Conservation, pers. comm., 2011). This is largely because arrays have previously

required custom-made expensive equipment and complex algorithms. Furthermore, kiwi monitoring often takes place in remote and rugged terrain, so equipment that is robust and easily deployed is required. However, in 2010 the introduction of wireless time synchronisation capabilities in a commercially-available autonomous recorder (model SM2-GPS, Wildlife Acoustics Ltd, Massachusetts, USA) provided potentially suitable equipment. This chapter describes attempts to use these recorders in an acoustic array for kiwi monitoring.

7.2   MICROPHONE ARRAY METHODS

7.2.1   *Time-of-arrival-difference methods*

Integral to all array methods is the ability to calculate the time difference of arrival (TDOA) of a signal at two microphones. There are many methods of achieving this (Chen et al. 2006, Benesty et al. 2008b), but the most common are correlation-based techniques, which are the focus here.

7.2.1.1   *Cross-correlation*

The simplest method of TDOA calculation is to maximise the cross-correlation function (CCF) between the two signals. The lag time that maximises this function is an estimate of the true delay. In practice, the CCF of a digital signal cannot be calculated, so it is usually replaced by its time-averaged estimate (Chen et al. 2006). Because the CCF often has high-frequency oscillations, the complex envelope is often used to identify the correlation peak. This standard cross-correlation method is widely used, but is prone to errors from reverberation (Benesty et al. 2008b)

7.2.1.2   *Generalised cross-correlation (GCC)*

In generalised cross-correlation methods a frequency-domain weight function is applied to the cross-correlation (Knapp & Carter 1976, Brandstein 1997):

$$R_{x_1 x_2}(\tau) = \int_\infty^\infty W(\omega) X_1(\omega) X_2^*(\omega) e^{i\omega\tau} d\omega \qquad (7.1)$$

The delay estimate is the τ which maximises the cross-correlation R between two signals X_1 and X_2 (* denotes the complex conjugate). Different weight functions (W)

can be used, and the case $W = 1$ is simple cross-correlation. Three other weighting systems introduced here are maximum-likelihood (ML), phase-transform (PHAT), and a pitch-based phase-transform method.

MAXIMUM LIKELIHOOD   This is a signal-to-noise (SNR) weighted version of GCC. The weight function is based on the signal SNR:

$$W_{ML}(\omega) = \frac{|X_1(\omega)||X_2(\omega)|}{|N_1(\omega)|^2|X_2(\omega)|^2 + |N_2(\omega)|^2|X_1(\omega)|^2} \tag{7.2}$$

where $|N_1(\omega)|^2$ and $|N_2(\omega)|^2$ are the noise power spectra estimated from the data (Brandstein et al. 1995, Brandstein 1997). This emphasises frequencies in the cross-correlation based on the SNR properties of the signals. It therefore performs best in noisy environments, but performance degenerates quickly in reverberant conditions (Brandstein 1997).

PHASE TRANSFORM   GCC-PHAT uses only the phase of the cross-spectrum, discarding amplitude (Knapp & Carter 1976, Brandstein 1997, Benesty et al. 2008a). The weighting function is the magnitude of the cross-spectrum:

$$W_{PHAT} = \frac{1}{|X_1(\omega)X_2^*(\omega)|} \tag{7.3}$$

GCC-PHAT places equal emphasis on each component of the phase, resulting in robustness to reverberant and relatively noisy environments (Brandstein 1997, Chen et al. 2006). As a consequence this method is widely used. However, it performs less well for low SNR signals (Brandstein 1997), and degrades in highly reverberant conditions, so variants that are less susceptible to noisy environments have been proposed (Benesty et al. 2008b).

PITCH-BASED WEIGHT FUNCTION   A pitch-based function attempts to combine the noise-insensitivity of ML with the reverberation-resistance of PHAT (Brandstein 1997). This is a frequency-dependent version of PHAT, applicable to signals containing periodicity, such as harmonics in human speech. The periodic sections of the signal should be less influenced by noise and reverberation, and so receive more weighting in the cross-correlation. The weighting function is:

$$W_P(\omega) = \frac{(1 - max(E_{m1}E_{m2})^\gamma}{|X_1(\omega)X_2^*(\omega)|}, \omega \in [a_m, b_m] \tag{7.4}$$

where E_m is the error associated with the mth harmonic of the fundamental, centred on the frequency region $[a_m, b_m]$. This method performs better than both ML and PHAT in noisy and reverberant environments. However, it is applicable only to signals with well-resolved harmonics, requiring a model of the frequency structure, and an accurate estimate ($< 1\,\mathrm{Hz}$) of the fundamental frequency.

7.2.2 *Localisation methods*

With a time delay of a signal arriving at two microphones, the direction or position of the source can be derived. As with time delay calculations, there are many possible methods for source localisation. The most common techniques are described here, and a novel method is introduced.

7.2.2.1 *Triangulation*

Triangulation is the simplest of the localisation methods, and returns a source direction given two signals from adjacent microphones. With a far-field source (Figure 7.2.1), a plane wavefront is incident at each sensor with a time delay τ that depends on the sensor spacing d, the incident angle θ and the sound speed c according to:

$$\theta = asin(\frac{c\tau}{d}) \tag{7.5}$$
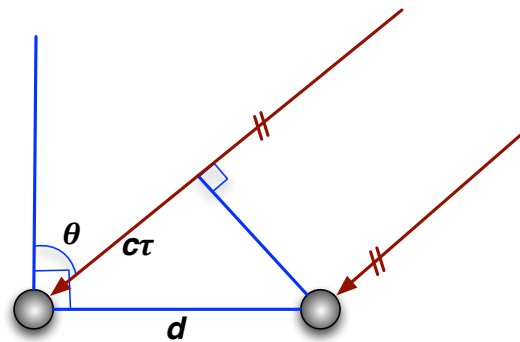
Figure 7.2.1: A sound wave incident on two microphones. The incident angle θ can be calculated from the time delay τ, the microphone separation d, and the sound speed c (Equation 7.5).

For a single close microphone pair, the source locus is a cone. By triangulating multiple pairs, the intersection of these cones enables the source position to be determined.

7.2.2.2 *Beam-forming*

Beamforming combines individual signals from multiple microphones so that the target signal interferes constructively, and noise interferes destructively. There are many variants, but the simplest is delay-and-sum beam-forming (Benesty et al. 2008a). This method sums the signals from each sensor, shifted by the expected time delay, given known sensor locations and an estimated candidate location. The position with maximum power of this reconstructed signal then gives the location estimate.

7.2.2.3 *Multilateration*

Whereas triangulation calculates source location based on angles, multilateration uses distances. The time delay between two microphones describes a half-hyperboloid. For multiple sensors, the hyperboloids corresponding to each possible pair intersect at a point locus (Figure 7.2.2). In three-dimensions, at least four sensors are required to produce a point estimate of the source location. More sensors than this improve the accuracy of the estimate.
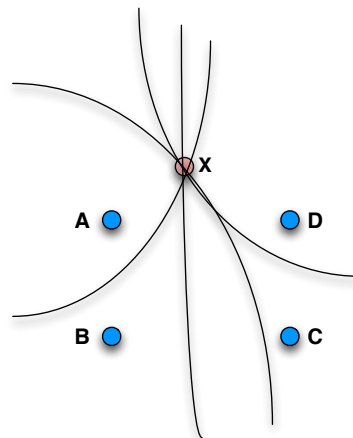
Figure 7.2.2: Schematic of multilateration. The time delay between each pair of microphones (A - D) describes a hyperbola of possible source locations. The intersection of the hyperbolae from multiple pairs produces an point estimate of the source location (X).

7.2.2.4 *Correlation sum*

Correlation sum localisation generates time delay functions that provide a likelihood measure at each possible delay value for a sensor pair (Collier et al. 2010b, Collier 2011). The source location is then estimated by finding the point which maximises the sum of these functions, at lags corresponding to the expected delay for that point.

7.2.2.5 *Novel maximum-likelihood algorithm*

An alternative maximum likelihood (ML) localisation method has been developed by Paul Teal. It assumes that for a series of sensor pairs the separation between pairs is much larger than that between the two microphones in each pair. Given an initial position estimate, an ML procedure minimises the sum of weighted temporal and spatial errors to provide an estimate of the source location. These errors, each having different magnitude, arise from:

- Time-delay measurement errors between the sensor pairs

- Time-delay measurement errors within the sensor pairs

- Positional measurement errors of sensor pairs, relative to the first pair

- Relative positional measurement errors of sensors within pairs

The initial position estimate is given by either triangulation or multilateration. For triangulation, the time delays within each sensor pair defines a half-cone. The intersection of these half-cones from all sensor pairs gives the location estimate. Multilateration does the same, but with the intersection of hyperboloids described by each sensor pair. Time delays for the P sensors are then used to construct $2P - 2$ equations to solve for the initial position estimate.

The procedure can be extended to estimate locations for multiple sources in an single optimisation, assuming that calls are not concurrent and sources are stationary during calls. Reduction of target location estimates can be achieved by calibrating the array using calls from known source locations, which can also be achieved in a single optimisation. Errors can be further minimised by assuming targets are at ground level, and incorporating known topography into the solution.

7.2.3 *Practical considerations*

7.2.3.1 *Time synchronisation*

In order to calculate time delays with sufficient accuracy, precise time synchronisation of the signals from multiple microphones is necessary. In wired arrays, the microphone signals can be passed to separate channels of a multi-channel recorder, so synchronisation depends solely upon the sample rate of the recording. The main disadvantages of wired arrays are that they require large amounts of expensive and heavy cable (Mennill et al. 2006, Fitzsimmons et al. 2008c), and are not easily portable as a result.

Wireless arrays offer a more portable and easily deployable solution, but at the expense of increased difficulty of time synchronisation. One solution is to use a local radio frequency network, correlating signals at each microphone with the transmitted signal (Allen et al. 2008, Blumstein et al. 2011). This has the advantage of not relying on a central time source, and can provide accuracies of $\sim 10\mu s$ (Collier et al. 2010b). An alternative method is to use time signals from Global Positioning System (GPS) satellites to synchronise each element in the array (e. g. Mennill et al. 2012). Such methods typically give micro-second precision, but rely on satellite reception, which is not always possible in habitat with dense vegetation or steep terrain. Although these methods synchronise recordings at a particular time, the clock speeds on each recorder will drift relative to each other, and so repeated synchronisation is necessary.

7.2.3.2 *Receiver location*

For both wired and wireless arrays, determination of microphone position is often a dominant source of error in localisation methods (Wahlberg et al. 2001, Mennill et al. 2006). GPS equipment is commonly used, and typically provides errors of over 1 m for handheld devices, and sub-metre accuracy for professional survey equipment (Blumstein et al. 2011). However, dense forest habitats can significantly degrade GPS accuracy, and necessitate repeated sampling (Mennill et al. 2006). An alternative that does not rely on GPS for microphone location is a self-localising system of nodes in a networked array (Girod et al. 2006). Each node emits a ranging acoustic 'chirp' which is detected by the other elements, and its relative position estimated by multilateration. By repeated sampling, this method can result in impressive node position accuracies ($\sim$5 cm) (Girod et al. 2006).

7.3   ARRAY METHODS FOR KIWI CALLS

7.3.1   *Time delay methods suitable for kiwi*

Generalised cross-correlation methods were selected as the method for estimating time delays from kiwi calls. The array was to be tested on little spotted kiwi (LSK) in Zealandia (§ 2.2.1). This site was acoustically reverberant due to steep-sided gullies, and also subject to moderate levels of background noise due to wind and streams. Maximum-likelihood and variants of Phase-Transform GCC methods were therefore tested. The pitch-based weight function was not employed because it was designed for human speech in room environments, and a required model of the frequency excitation spectrum does not exist for kiwi. Additionally, harmonics are often not distinct in female LSK calls, and harmonics in both sexes are usually only detectable in close calls.

7.3.2   *Recorders*

The objective was to determine whether a microphone array for monitoring of kiwi could be easily deployed using commercially-available recorders. The only such recorder at the time of the study was the Wildlife Acoustics Song Meter 2 (Table 7.3.1 & Figure 7.3.1). In August 2010, a GPS time-synchronisation option was made available for this two-channel, programmable recorder, providing capability for the recorders to be combined in an array. A GPS receiver (model GPS16x-LVS, Garmin, Kansas, USA) was plugged into each recorder, and provided time signals to synchronise the internal clock to an accuracy of ∼1 ms. Synchronisation was maintained throughout recording. Four Song Meters, already utilised as standalone autonomous recorders for kiwi, were upgraded for a total cost of US\$2,400.

7.3.3   *Recorder configuration*

Before the recorders could be used in an array, determination of the appropriate microphone configuration was necessary. Different arrangements were tested by playing a chirp sequence (Figure 7.3.2) at different angles to the recorder. These angles were measured clockwise from the normal to the recorder (Figure 7.3.3).

Table 7.3.1: Specifications of the Wildlife Acoustics Song Meter 2 recorders used in the study.

Specification	Value
Weight	0.91 kg (excl. batteries)
Price	US$700
Sample rate	4–48 kHz
Power supply	4 D-cell batteries, or 6-12V external
Power consumption	50-65 mA for stereo, ≤24 kHz
Media	Up to 4×32 GB SDHC
GPS upgrade cost	US$600 each
GPS receiver	Garmin GPS16x-LVS
Microphone response	20–20,000 Hz
Microphone SNR	> 62dB

Figure 7.3.1: The Wildlife Acoustics Song Meter 2. In this case the microphones are attached directly to the recorder.
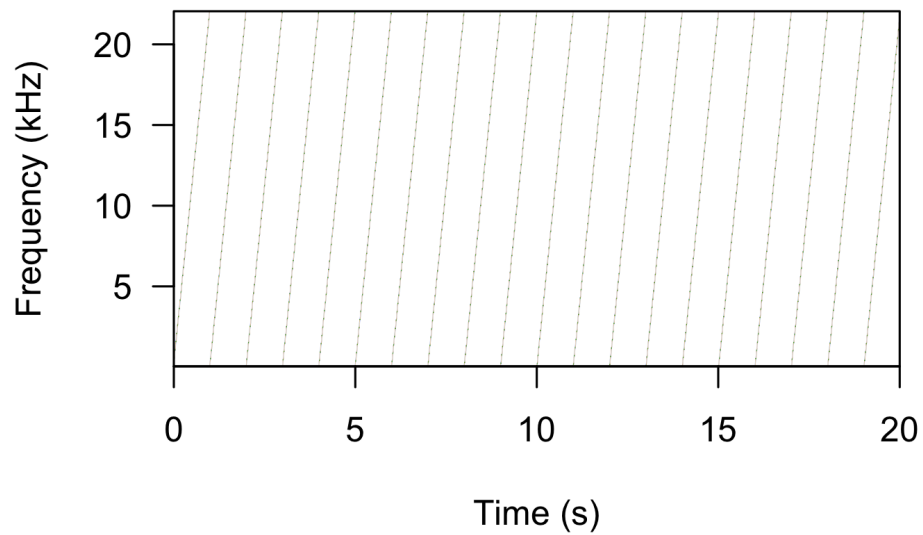
Figure 7.3.2: Spectrogram of the chirp sequence used to calculate microphone impulse responses, consisting of 20 one-second frequency modulations.
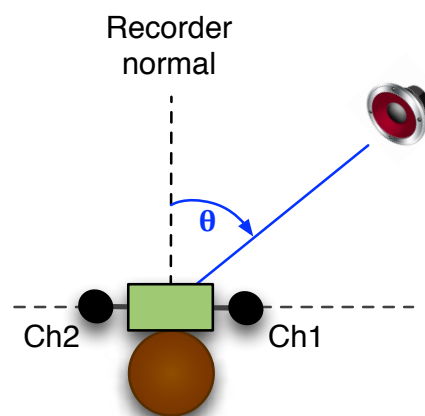
Figure 7.3.3: Test angles were measured clockwise from the normal to the face of the Song Meter 2.

7.3.3.1 *Microphone orientation*

METHOD   The microphones used for the trial (model SMX-II, Wildlife Acoustics Inc.) were specified as omni-directional. However, initial tests indicated that microphone responses differed between the two channels when the microphones were in the default locations on each side of the recorder (Figure 7.3.1). This would add considerably to the difficulty of time delay calculation. A series of tests were therefore conducted to determine the optimal microphone orientation. The chirp sequence was played at angles of −90°, 0°, and 90° (Figure 7.3.3). The microphones were then swapped and the test repeated, to distinguish any effects from differences between the microphones. The impulse responses of each microphone were then compared for different microphone orientations (Figure 7.3.4). Extension cables were used to place the microphones at any position relative to each other.
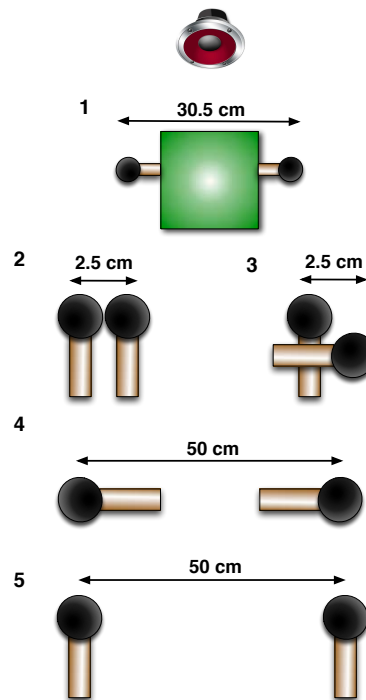
Figure 7.3.4: Microphone orientations trialled. The signal direction in this illustration is from the top of the diagram, at 0°. Directions of 90° and -90° were also tested.

RESULTS   With the microphones on the recorder, there was a large difference between the phase responses when the source was close to the microphone axis (Figure 7.3.5). This could have been due to reflections or diffraction from the recorder surfaces, or the microphone characteristics. Further tests in which the microphones were not parallel (tests 1, 3 and 4 in Figure 7.3.4) showed differences in the amplitude of

impulse responses. Swapping and rotating the microphones indicated that this was likely to be a result of directivity in the microphones. When the microphones were oriented parallel to each other, the responses showed no such differences, and only a delay that corresponded well with the separation between them (Figure 7.3.6). It was therefore deemed necessary to orient the microphones in the same direction. This was achieved by mounting the microphones on a block of wood, and connecting them to the recorder with extension cables.
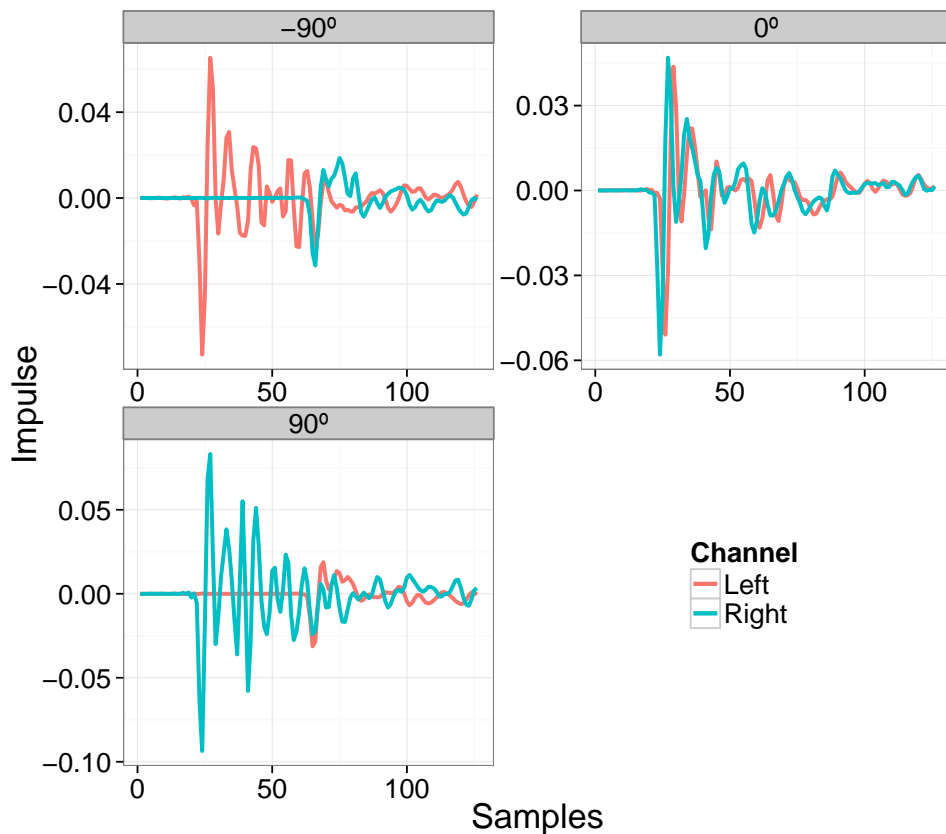
Figure 7.3.5: Impulse responses with the microphones mounted on the recorder (30.5 cm separation) from a source with incident angles of $-90°$, $0°$ and $90°$. The responses from each microphone were similar when the source is perpendicular to the microphone axis ($0°$), but differed in shape (in addition to translation from the time delay) when the source was parallel to the axis ($\pm 90°$).

7.3.3.2   *Microphone separation*

Time delay calculations can be affected by spatial aliasing, which causes multiple peaks in the cross-correlation function at higher frequencies (Benesty et al. 2008b). Multiple peaks within the physical time delay range makes determination of the cor-
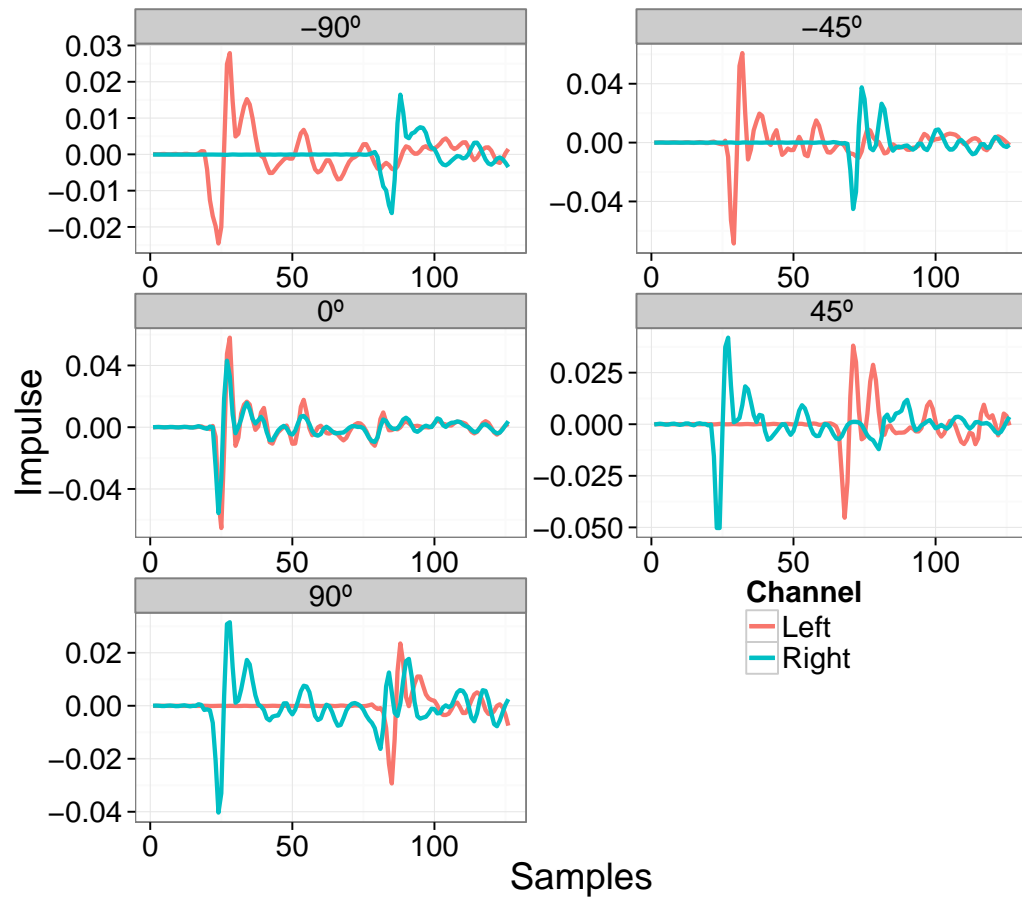
Figure 7.3.6: Impulse responses of microphones mounted in parallel, with 48 cm separation. There is good agreement in the shape of the impulse responses (allowing for the time delay) at all source angles.

rect peak difficult, and so a low-pass filter is used to prevent them. Frequencies at which aliasing occurs depends on the sound speed and microphone spacing, according to $f > c/2d$ (Benesty et al. 2008b). A large microphone separation (d) is desirable as it improves time delay accuracy, but increasing d decreases the frequency at which spatial aliasing occurs. So a trade-off occurs between increasing TDOA accuracy and limiting spatial aliasing. Male LSK calls contain fundamental frequency energies up to 3.5 kHz. With a sound speed of 320 m/s, this would limit microphone spacing to 4.5 cm to avoid any spatial aliasing. This is prohibitively small, requiring sub-millisecond timing accuracy, so tests were conducted to determine whether spatial aliasing was a problem when recording kiwi calls with the Song Meter 2.

One channel from a recorded kiwi call was shifted with respect to the other channel by a range of time delays. These corresponded to different separations of the microphones for each channel, assuming the signal arrived from the microphone axis ($\theta = 90°$). In this way simulated signals for different microphone separations were generated, ranging from 0 m to 10 m. The signals were bandpass filtered from 2–5 kHz, and cross-correlation of the channels examined for multiple peaks, which would indicate spatial aliasing within this frequency window.

The cross-correlation function for the whole 25 s call showed no multiple peaks, and just a single peak close to the expected delay (Figure 7.3.7). This indicated that there was no strong periodicity across the whole call with a period shorter than the time delay. The substantial changes in syllable structure, length and separation that occur over the call (Chapter 2) were the likely cause of this.

These tests showed that spatial aliasing was not a problem when cross-correlating calls of typical length (Chapter 2) in the frequency range of male LSK. Since females have lower call frequencies and aliasing decreases with frequency, this would apply to both sexes. Any microphone separation in the tested range of up to 10 m was therefore suitable. A value of 0.5 m was selected, being large enough to ensure reasonable time delays, but sufficiently small to allow accurate relative microphone orientation and practical deployment in a forest environment.

7.3.3.3 *Cross-talk*

With a multi-channel recorder, cross-talk — when the signal from one channel 'bleeds' into another — can cause obvious problems for cross-correlation. Initial cross-correlation tests showed a peak in the cross-correlation function during the interval between syl-
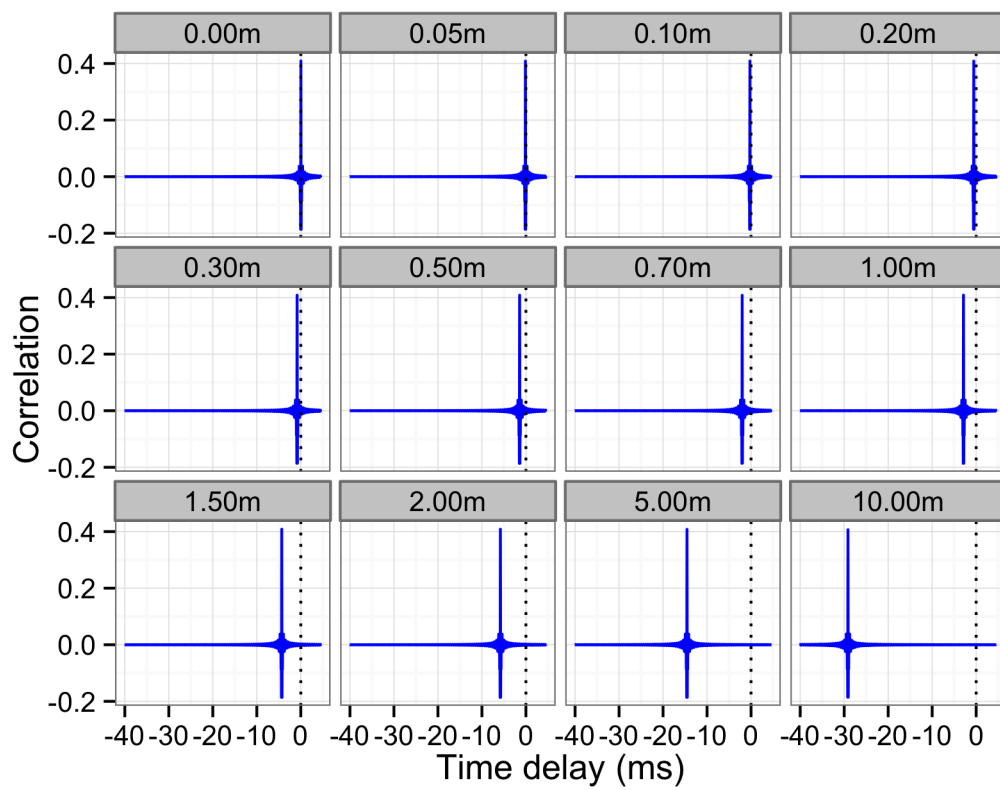
Figure 7.3.7: Cross-correlations at different simulated microphone separations, from 0–10 m, for a male LSK call aligned with the microphone axis. For each separation there is only a peak in the correlation function at the expected time delay, indicating that spatial aliasing would not influence time delay calculations for these kiwi calls.

lables in a LSK call. Cross-talk was a suspected cause of this. However, tests by the Song Meter manufacturer (Ian Agranat, Wildlife Acoustics Ltd, Massachusetts, USA, pers. comm., 2011) showed that cross-talk in the SM2 was $\sim$56 dBV between the channels. This was below the SMX-II microphone noise floor of -49 dBV at 1 Pa, so cross-talk below ~ 50 dB separation would not be detectable by the recorder. The observed correlation peak therefore remains unexplained.

7.3.4 *Time delay method*

To investigate the performance of different GCC algorithms, a male LSK call was played at known angles from a Song Meter 2 in an open field. The call was played at 10 m and 30 m from the recorder, with two volume settings at the further distance, to simulate close and distant calls (Figure 7.3.8). The angle calculated from the time delay between the two channels of the recorder was then compared with the known angle to assess the relative accuracy of the algorithm. Male calls were used for all tests, because they were the main signal to be surveyed, typically outnumbering female calls by $\sim 3 : 1$ (§ 5.3). Males also have a higher frequency than females (M: $f_0 = 2.5 - 3.5$ kHz; F:$f_0 = 1.5 - 2.5$ kHz; § 2.31), and since higher frequencies are more prone to spatial aliasing and lower GCC accuracy, any such issues would first become evident in male calls. Maximum-likelihood (GCC-ML) and phase-transform (GCC-PHAT) methods were tested.

7.3.4.1 *Maximum likelihood GCC*

Initial tests of this method using the full male call indicated that this method performed poorly, even in a low reverberant environment (Figure 7.3.9). Since the deployment environment at Zealandia was more reverberant, and ML is known to perform more poorly in such conditions (Brandstein 1997), this method was not pursued further.

7.3.4.2 *Phase-transform GCC*

The unsuitability of the other two GCC methods considered left GCC-PHAT as the remaining option. Initial tests showed promising results, but further adjustments were made to refine the algorithm.
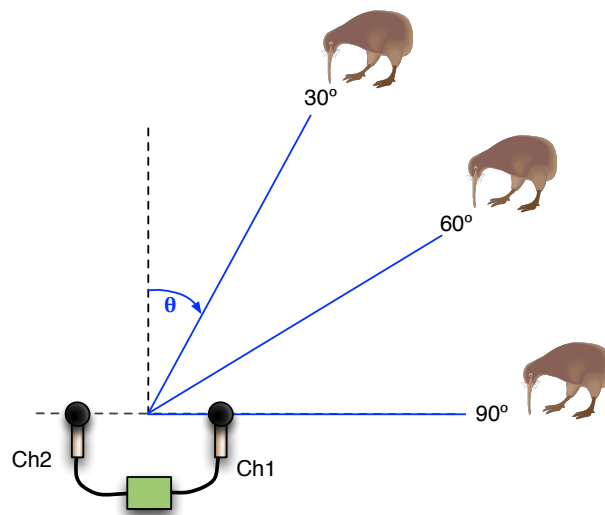
Figure 7.3.8: Test configuration to compare performance of GCC TDOA methods. A male LSK call was played at different angles from the recorder, measured clockwise from the recorder normal. Microphones were aligned parallel to each other, and attached to the recorder via extension cables. Calls were played at 10 m and 30 m from the recorder, and at two different volumes at 30 m to simulate calls of different distances.
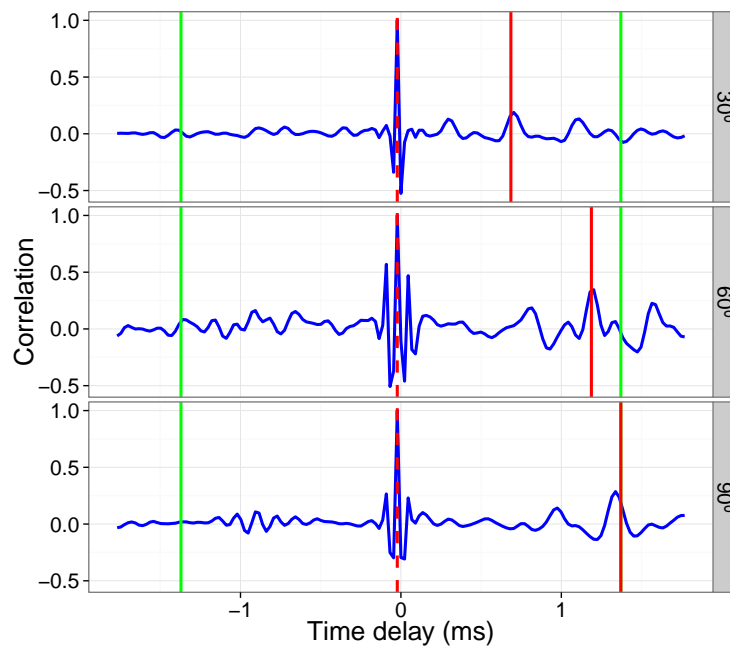
Figure 7.3.9: Maximum-likelihood generalised cross-correlations for recordings of a male LSK call played high volume 30 m from the recorder, at 30°, 60° and 90° to the microphone axis normal. The red dotted line depicts the time delay of the cross-correlation peak, and the solid red line the true delay. Green lines are the physical limits. The ML method performed poorly in these tests.

frequency filtering. The majority of energy in LSK calls is found in the fundamental harmonic, so low-pass filter of $2 - 5\,\text{kHz}$ was applied to the male calls to avoid spatial aliasing and minimise interference from other sound sources.

segmentation. LSK calls consist of a long series of repeated syllables, with intervening silent periods of $\sim 1\,\text{s}$. These gaps were removed from the call by applying a filter to retain only call segments with an amplitude above a given threshold. Noise spikes and unusually short syllables were discarded by a further filter. GCC-PHAT was then applied to each signal segment, and the resulting correlation matrices were trimmed to the length of the shortest, above a specified minimum signal length. The final correlation function was calculated by averaging the trimmed correlations over all signal segments in the call.

normalisation. In many tests the correlation function showed a dominant peak at zero delay, suggesting that the GCC-PHAT weight function was normalising the signal too much. The strength of the normalisation was adjusted by introduction of a weight index, n, such that the weight function in Equation 7.3 was: $W_{PHAT} = 1/|X_1(\omega)X_2^*(\omega)|^n$. This index was varied between values of 0 (basic GCC) and 1 (GCC-PHAT) to find the optimal weighting for most accurately recovering the true time delay. A value of $n = 0.5$ was chosen from these tests.

unbiased correlation. Cross-correlation has a bias towards zero lag, since at lower lags more values are being correlated, resulting in a higher correlation peak. To avoid dominant peaks at zero delay unbiased correlation was used, which weights the correlation function at lag m, for sample length N, by $R_{unb}(m) = 1/(N - |m|)R(m)$.

non-physical solutions. The microphone separation (d) and sound speed (c) impose a physical limit on the possible time delay solutions of $|t| \leq d/c$. Peaks in the cross-correlation function greater than this lag value were discarded.

circular convolution. To avoid the detrimental effect of circular convolution during cross-correlation, linear convolution was enforced by zero-padding the signals prior to correlation.

7.3.4.3 *TDOA results*

The final GCC-PHAT method produced a smoother correlation function with a maximum closer to the true value than conventional GCC-PHAT (Figure 7.3.10 & Fig-

ure 7.3.11). Converting the time delays to angles demonstrates the improvement in accuracy from the modified GCC-PHAT (Figure 7.3.12). The performance deteriorated at higher incident angles (closer to the microphone axis).
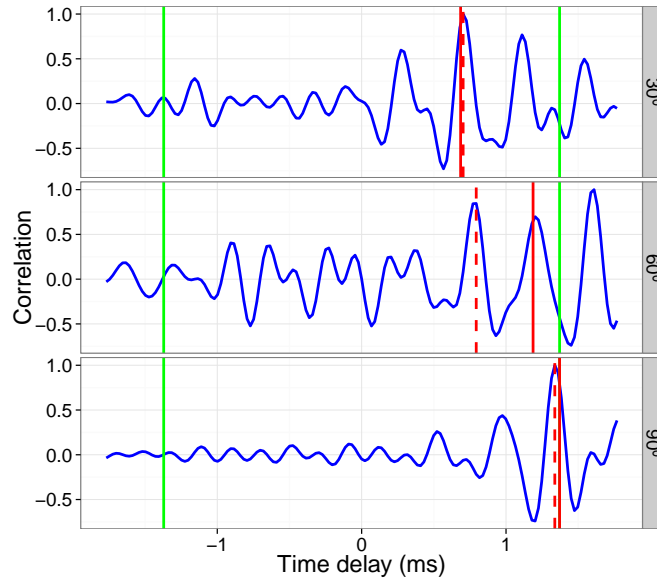
Figure 7.3.10: Cross-correlations from basic GCC-PHAT for the high volume 30 m test. Symbols as Figure 7.3.9. This method is more accurate than GCC-ML, but chooses the incorrect peak for the 60° test.

7.3.5   *Full field test*

7.3.5.1   *Test site*

With the TDOA algorithm working well in an low-reverberant environment, the next stage was to assess its performance in LSK habitat. An area was selected in Zealandia which was quite flat (Figure 7.3.13), but which was covered in young native lowland podocarp forest of typical density. This was a moderately reverberant environment, but typical of the situation in which kiwi monitoring takes place. In fact, the area encompassed a kiwi call count listening site that was established in 2001 (Raewyn Empson, Zealandia, Wellington, New Zealand, pers. comm., 2010).

7.3.5.2   *Recorder configuration*

Four SM2 recorders were placed in a square formation, approximately 50 m apart (Figure 7.3.13). The recorders were attached to trees at 2 m above ground level.
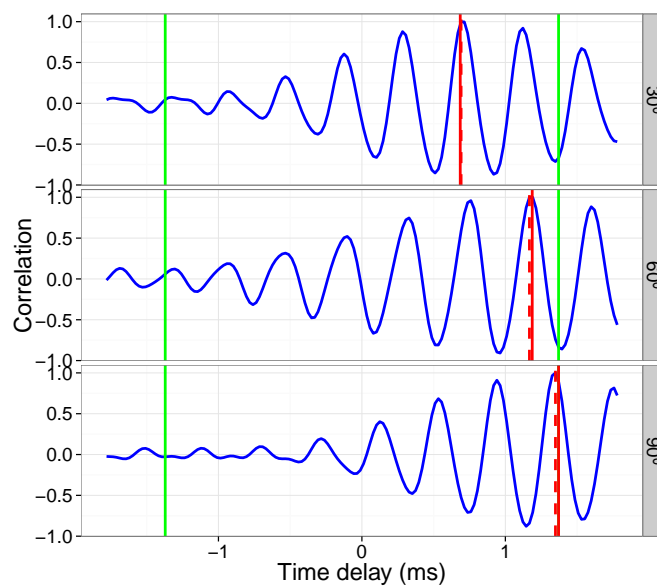
Figure 7.3.11: Cross-correlations from the final GCC-PHAT method for the high volume 30 m test. Symbols as Figure 7.3.9. This algorithm out-performs both conventional GCC-PHAT and GCC-ML, and closely recovers the true time delay for each test.

Two microphones were attached to each recorder with 1 m extension cables, and fixed at each end of a flat strip of wood, perpendicular to its surface, so that they were parallel (Figure 7.3.14). The microphones were directed approximately towards the centre of the square. One microphone pair was aligned vertically, to improve accuracy along that axis; the other three were placed parallel to the ground. Each Song Meter was time-synchronised using its GPS receiver. As well as providing time signals, the GPS information was used to provide an accurate estimate of the receiver position. Following the test, the recorders were left in position, sampling GPS coordinates at one minute every other minute for two days. This provided approximately 2000 samples. The 10% of samples furthest from the centroid were discarded, and the mean of the remaining samples provided the averaged location of the recorder (Daniel Mennill, University of Windsor, Windsor, Canada, pers. comm., 2011).

7.3.5.3 *Test sounds*

A 35 s male kiwi call was played at eight locations within and just outside the array (Figure 7.3.13), using a 45-watt speaker (MiPro MA-101) connected to a digital recorder/player (Sony PCM-M10P). The speaker volume was set to be close to the true
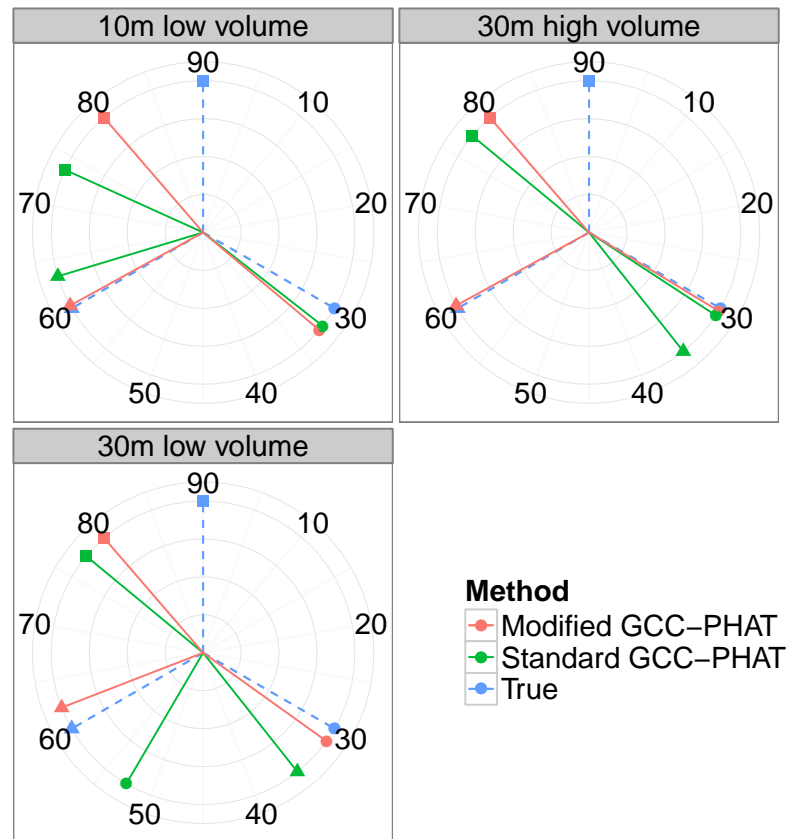
Figure 7.3.12: Comparison of the modified and standard GCC-PHAT time delay algorithms in recovering incident angles of a male LSK call played at different angles and distances from the recorder. The circular scale shows the angle from the normal to the microphone axis, measured clockwise. Filled symbols depict the corresponding true value; note the large inaccuracies from standard GCC-PHAT for the 30 m low volume test.
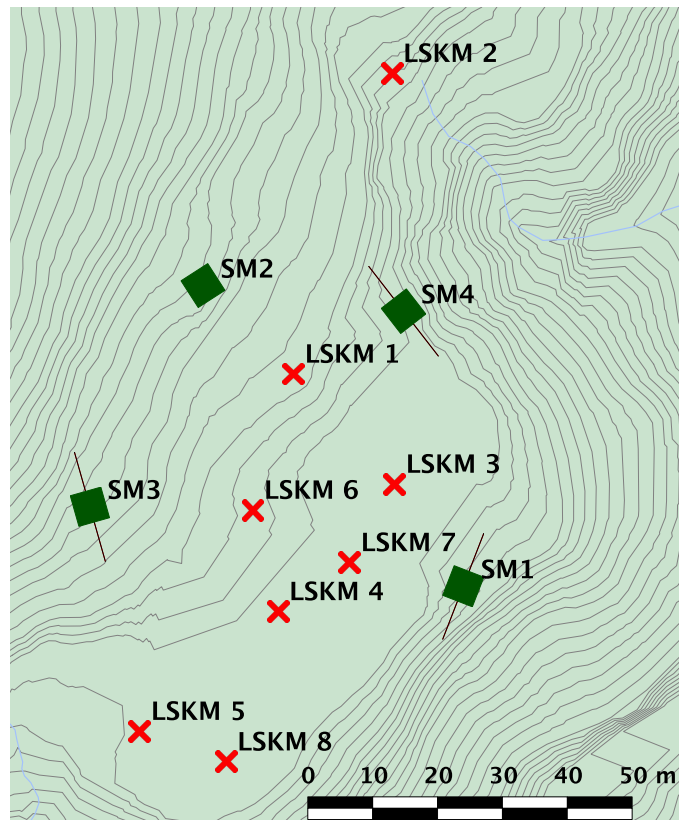
Figure 7.3.13: The array test site in Zealandia. Green squares mark the recorder locations, with microphone axes indicated by solid black lines. The recorder without a line has the microphone axis oriented vertically. Red crosses mark male LSK call test locations. Contours are at 1 m intervals, and green shading denotes native forest.

Figure 7.3.14: Recorder configuration in the full field test. The microphones are aligned parallel to each other at 50 cm separation, pointing towards the centre of the array. The GPS receiver (top left of picture) is plugged into the auxiliary port on top of the recorder.

amplitude of a kiwi call, calibrated from a call made at a known distance (Chapter 6). Before the first and after the last test, a chirp sequence was played at the location of the first test site (LSKM 1; Figure 7.3.13). This was to allow correction for any clock drift, and to check microphone impulse responses. The Song Meter sampling rates were 48,000 kHz. The location of test positions was determined with a hand-held GPS receiver (Garmin GPSMAP CSX60). The procedure followed that for determining the recorder locations, except that the GPS unit was left at each test location for 5–10 minutes, resulting in ∼100 samples per site.

7.3.5.4   *Triangulation*

The time delay between channels was calculated at each recorder for each test sound using the segmented GCC-PHAT algorithm described in § 7.2.1.2. The delays were converted to incident angles (Equation 7.5) for the three recorders with a horizontal microphone alignment, and triangulated to estimate source position. In only one test did all three angles from each recorder intersect and provide an accurate determination of the test location (Figure 7.3.15).

In order to investigate why the GCC-PHAT algorithm was unable to reliably recover time delays, impulse responses for the three horizontally-aligned recorders
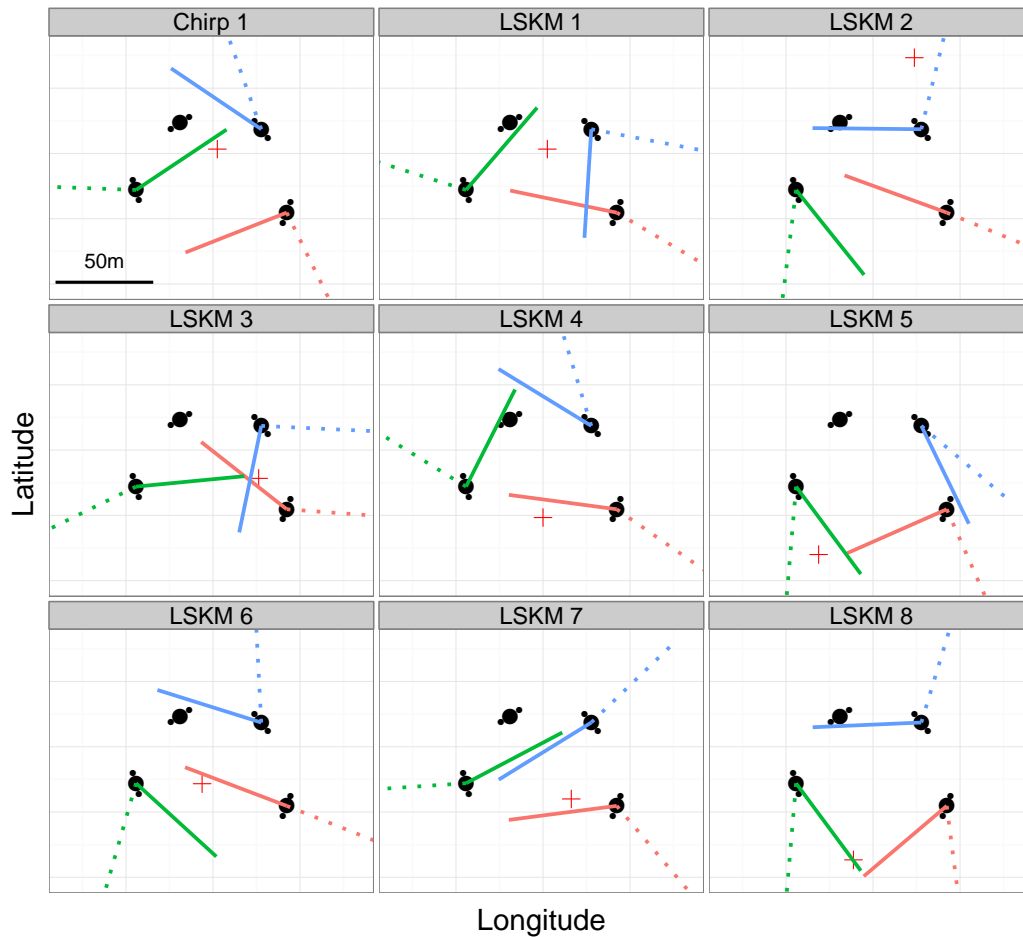
Figure 7.3.15: Triangulations of the eight male kiwi calls (LSKM) and the first chirp test. Circles denote recorder positions, with microphone separation exaggerated to show recorder orientation. The north-west recorder had microphones aligned vertically, so was not used in the triangulations. Crosses depict true source location, and lines are the triangulated angles calculated for each recorder using the modified GCC-PHAT algorithm. Each time delay gives two potential angles: the solid and dotted lines show the solution inside and outside the array respectively. Only one test (LSKM 3) yielded an accurate recovery of the source location.

were derived from the two chirp sequences. Only one recorder (SM3) provided time delays between the channels that were close to expected value. For another (SM1) the impulse responses could not be calculated since the signal was too weak, and the third (SM4) showed significant structural differences between the channels (Figure 7.3.16). This suggests that reverberations or low signal-to-noise were the principal cause of the inability of the correlation to accurately estimate time delays.
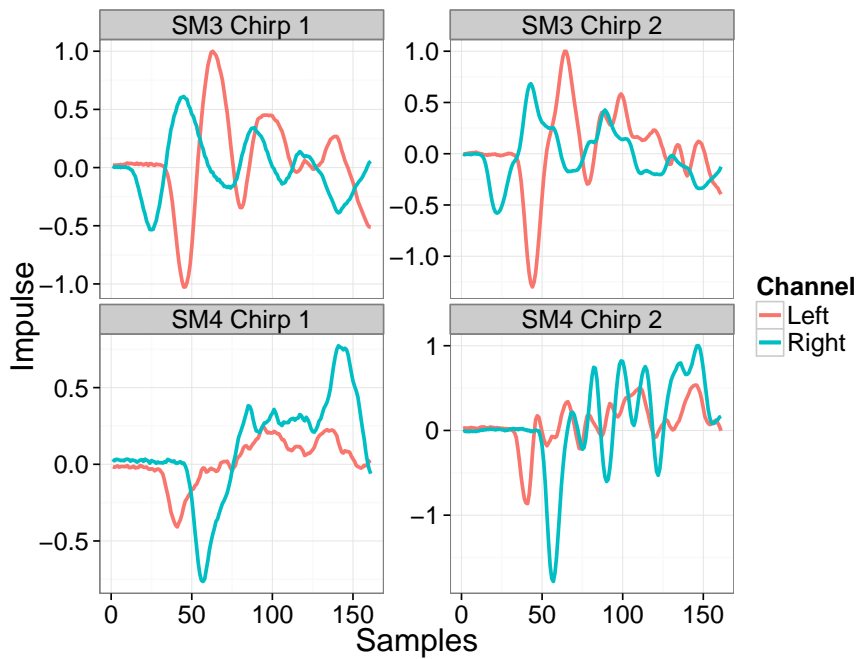
Figure 7.3.16: Impulse responses from two recorders in the field experiment, calculated from the chirp sounds. Recorder SM3 shows good agreement between the channels, whereas in the other there are substantial structural differences in the signals received at each microphone. This suggests that reverberations or low signal-to-noise are the main source of error in time delay calculations.

7.3.5.5 *Blind source separation*

Time delays for all recorders and tests were also calculated using a blind source separation (BSS) algorithm, provided by Yuanhang Zheng and Walter Kellermann of the University of Erlangen-Nürnberg, Germany. Blind source separation involves the recovery of signals from multiple mixed signals, and is robust against reverberations and background noise (Aichner et al. 2006). In this implementation a bandpass filter was first applied to the kiwi call and then a voice activity detector was used to isolate the syllables of the call. The resulting signal was passed to the BSS, full details of which are in Aichner et al. (2006), Lombard et al. (2009). The algorithm, run in

Matlab, returned a list of possible incident angles, of which the mode was used to estimate the source location.

The algorithm was applied to the full field test data to provide incident angles for each test call and recorder. However, this method was also unable to accurately triangulate to the true source location in all tests, and for six cases the signals were too weak to yield a finite time delay (Figure 7.3.17). Overall this technique performed worse than GCC-PHAT.
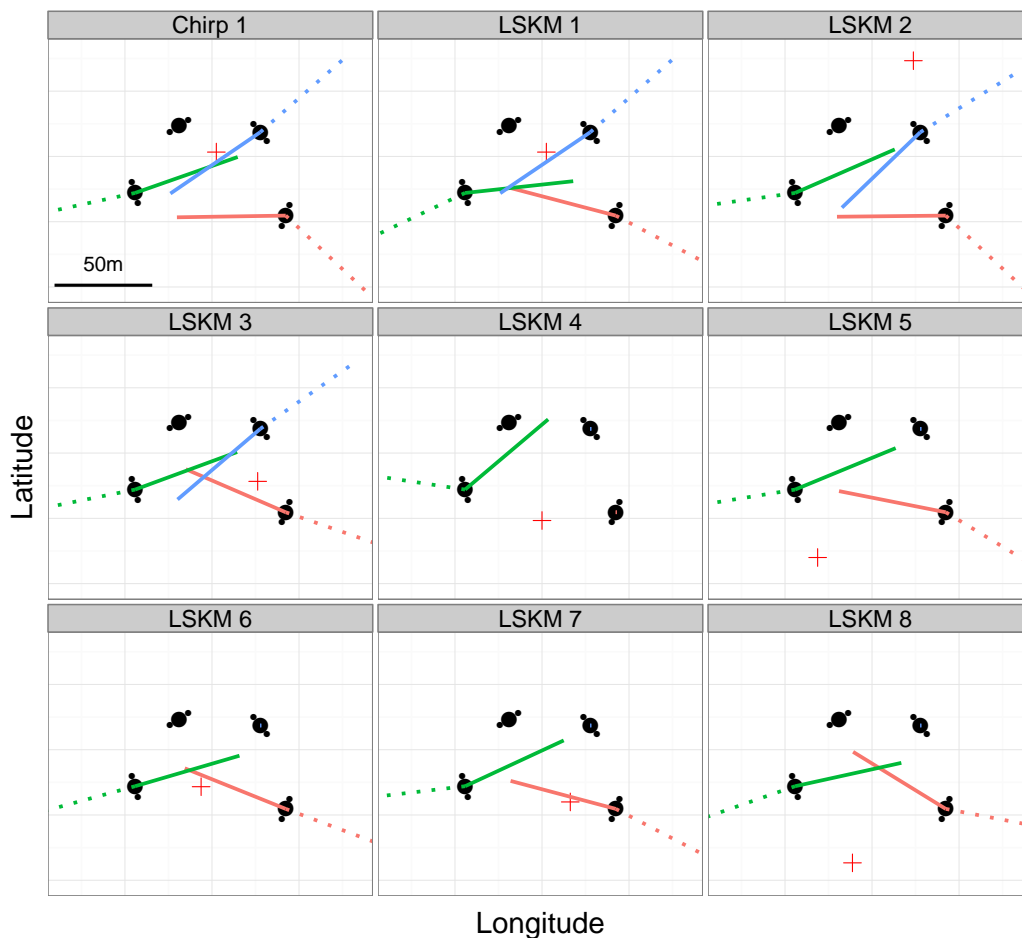
Figure 7.3.17: Triangulation results using blind source separation. Symbols as Figure 7.3.15. No source locii for a recorder means that the signal-to-noise ratio was too low to yield an estimate from the algorithm, except for the north-west recorder, which was vertically oriented and so not included in the triangulation.

Reliable time delay estimates were not possible even from the highly synchronised ($dt = 1/48kHz = 20\,\mu s$) channels in each recorder. Performance would be worse

between channels from different recorders that were synchronised to lower precision by GPS time signals ($dt \approx 1\,\text{ms}$). It was therefore decided not to continue with multilateration tests.

7.4 DISCUSSION

Both the modified GCC-PHAT time delay algorithm devised for this study and the blind source separation method did not work reliably in the forest environment in which the array was to be deployed. For this reason no localisation results were obtained, either from triangulation, multilateration, or the maximum likelihood method described in § 7.2.2.5.

The results presented here are in contrast to another study published after this analysis, in which the same recorders were used to localise a variety of animal calls in two habitat types (Mennill et al. 2012). In that study, four Song Meter 2 recorders each with two microphones directly attached and GPS time synchronisation were arranged in arrays with 25 m and 50 m spacing. Twenty-five simulated and animal signals were broadcast inside and outside the arrays in both open and forest sites. All eight channels were inspected, and suitable signal sections passed to a correlation sum localisation method, analogous to delay-and-sum beam-forming (see Mennill et al. 2006 for details). This provided a source location estimate with an associated quality index and surrounding 'probability cloud' of less likely location estimates, which were used to determine whether the location was 'reliable' or not. The overall accuracy of source location estimation was 1.87±0.13 m. In this section these impressive results are used to examine the deficiencies in the methods presented here, and to assess the suitability of the equipment for localising kiwi calls.

7.4.1 *Recording environment and TDOA algorithm*

The large differences in the impulse response received between microphones (Figure 7.3.16) indicated a highly reverberant acoustic environment. Although not densely vegetated by New Zealand standards (Figure 7.3.14), the evergreen foliage in the test environment provides many reflective surfaces. This is in contrast with the forest habitat of Mennill et al. (2012), which appears rather open in comparison, with very little understory, and small to moderately sized tree trunks providing the main reflective surfaces. That Mennill et al. (2012) reported no significant difference in location accuracy between the field and forest sites suggests that either the

acoustic environment between the two was similar, or that their method is far more robust to reverberations that those used here. The modified GCC-PHAT time delay method derived here worked well in an open field, but not in the forest, suggesting that reverberations in a challenging acoustic environment are likely to be the primary reasons for the failure in the forest habitat. This is despite the considerable attempts to devise an algorithm that is robust to reverberations. In contrast, Mennill et al. (2012) appear to have used simple cross-correlations to calculate time delays.

7.4.2   *Source signal type and strength*

Mennill et al. (2012) reported that location accuracy was greater for signals containing frequency modulation, in common with earlier recommendations (McGregor et al. 1997). Tonal sounds and rapidly-repeated signals led to poor performance at high frequencies, but not at low frequencies. However, all signal types produced useable location estimates, except for woodpecker drumming sounds. Their results suggest that LSK calls are suitable for array localisation methods. The barred antshrike call used in Mennill et al. (2012) is similar to that of a male LSK, and was located with an average accuracy of 1.1±0.2m.

However, the strength of the test sound is undoubtedly a factor. Mennill et al. (2012) played sounds at unnaturally loud amplitudes of 95 dB at 1 m. In contrast, the kiwi tests used calls at realistic volumes. This was to simulate the true conditions under which the array would have to operate, but undoubtedly contributed to a deterioration in accuracy. Many kiwi test calls had low signal-to-noise on several recorders, and six were too weak to yield time delay estimates with the BSS method. This could be resolved by increasing the call amplitude, but that would not provide a test of whether the methods work in practice.

7.4.3   *Signal selection*

Time delays were calculated for the entire kiwi call, with syllable gaps removed. There was no quality selection of signals to exclude biotic or abiotic background noise, such as that from other species (the test was conducted during the day), wind noise, or anthropogenic signals (e. g. aeroplanes and motor vehicles). Some of the syllables would have been affected by such interference, and so the resulting mean correlations would have been less likely to yield accurate time delays. Mennill et al. (2012) discarded 16% of their signal portions due to background noise interference,

and then further filtered the signals based on the reliability of the resulting location estimates. The overall reliability varied from 1–93%, depending on the source signal type, and at forest sites only 33% of sounds provided reliable estimates, after filtering for background noise. Overall, only 27% of annotated sounds were used. In contrast, all test sounds in the present study were used to produce time delay estimates.

This quality thresholding is a major difference between the methods. Quality filtering of signals was not applied in the kiwi tests because acoustic interference was low. This was confirmed by using a narrower 2–3.5 kHz frequency filter to exclude most diurnal bird songs, but which produced almost identical triangulation results. However, the results from Mennill et al. (2012) demonstrate that interference from other signals severely limits location accuracy, so a method of rejecting sections of calls that contain such disturbances is required. Rather than calculating a time delay from the mean of correlations of all call syllables, a delay (or location) estimate should be calculated for each separately, and filtered according to their precision.

Mennill et al. (2012) found that using short selections of less than 1 s produced better location estimates. Similar length sections were used in the current study, since the call segments were LSK syllables which are less than 1 s long (Chapter 2). However, correlating short signal lengths can lead to a bias towards zero time delay as a result of 'edge effects' (Yeredor 2012). It is therefore possible that the segmentation of the kiwi calls into short sections prior to correlation resulted in such a bias and contributed to the poor triangulation accuracy.

7.4.4 *Array configuration*

Location estimate accuracies deteriorate with increasing array spacing, as the distance between the source and the receiver increase (Bower & Clark 2005, McGregor et al. 1997). This in confirmed in Mennill et al. (2012), with 25 m element spacing performing better than 50 m separation. However, 25 m spacing is small for kiwi, even for LSK which have small territories (2–6 ha at the study site; A. Digby, unpubl. data). With such spacing most calls would be recorded outside of the array. The location accuracy of sources outside of the array is significantly worse than those inside (Bower & Clark 2005, McGregor et al. 1997), and the difference can be substantial (e.g. 10.2 ± 1.6 m compared to 1.87 ± 0.13 m in Mennill et al. 2012). The node spacing of $\sim$50 m used in the present study was more suitable for kiwi monitoring, although did result in weak signal to noise ratios for some test locations. The op-

timal spacing will depend on the habitat, but there is a trade-off between increased coverage and position accuracies in determining array size (Bower & Clark 2005).

Interestingly, Mennill et al. (2012) used microphones attached directly to the Song Meter recorders, so presumably experienced none of the diffraction or reflection problems this caused in the kiwi experiment. As noted earlier and also by Mennill et al. (2012), using extension cables to increase the microphone separation will yield more accurate time delay estimates.

7.4.5   *GPS error*

Even though Mennill et al. (2012) used survey-grade GPS equipment, their location inaccuracies were dominated by GPS errors. These errors were significantly higher at forest compared to field sites (2.23 ± 0.031 m and 0.006 ± 0.001 m respectively. When they used the same hand-held GPS (Garmin GPS60CSx) as employed in this study, their overall position error increased by 125% (from 1.87 ± 0.13 m to 4.22 ± 0.22 m. The poor location estimates from triangulation of the kiwi calls would therefore have been significantly affected by GPS errors, particularly in the source position, which were only averaged over several minutes. Future experiments should integrate GPS signals at test locations over a longer period, although this is logistically difficult at a public site. Using survey-grade GPS equipment would be desirable, but is not essential. The location accuracies achieved with hand-held GPS units by Mennill et al. (2012) are sufficient for kiwi monitoring, in which accuracies of $\sim$10 m should be sufficient for most applications.

7.4.6   *Recommendations*

The experiments described in this chapter demonstrate that implementing a microphone array for kiwi monitoring using current commercially-available recorders is challenging. However, the results of Mennill et al. (2012) show that such arrays are possible, at least in simpler acoustic environments. Comparison of that study with this one indicate that the highly reverberant environment, large node spacing, and lack of signal selection were the primary reasons for the inability of this method to yield reliable triangulation estimates. The signal selection is particularly significant: although in most cases the test location could not be recovered from the whole kiwi call, the experience of Mennill et al. (2012) suggests that this is to be expected. Signal portions with high background noise should therefore be rejected, and location

estimates and reliability measures derived for each kiwi syllable. Shortage of time restricted further testing, but it is likely that this technique would yield far more accurate results. Whereas the microphone setup and correlation calculations in this study were more sophisticated than those of Mennill et al. (2012), the lack of signal selection and reliability measures negated the benefits from these.

Future attempts to apply microphone arrays for kiwi should initially be tested in open areas, such as scrub and grassland habitat. A survey-grade GPS would improve accuracy, but is not essential on the scales required for kiwi monitoring. Array spacing is problematic for kiwi. Larger node separation results in larger errors, but closer spacing would result in substantial error increases due to most calls being outside of the array boundary. One solution would be to expand the array with additional microphones, which allows for a larger array for increased spatial coverage while still maintaining accuracy (e. g. Collier et al. 2010b; but see Mennill et al. 2006). The large spatial scales of kiwi territory dynamics, on the order of tens of hectares (McLennan et al. 1987), also means that the sub-5 m accuracies possible with small arrays (Collier et al. 2010b) would be unnecessary for most monitoring applications.

The careful signal selection required to detect calls and identify high signal-to-noise portions within them, means that the process would require lengthy manual input. However, automatic detection methods (Chapter 6) can return signal quality information that could be used to select suitable sections for localisation. It is therefore certainly feasible that the whole process could be fully automated.

In summary, microphone arrays cannot currently be easily implemented for spatial monitoring of LSK without specialist equipment and advanced algorithms. This is also likely to apply to other kiwi species since they have similar call structure (e. g. Corfield et al. 2008a). However, the rapid increase of array technology (Blumstein et al. 2011), coupled with increasing demand for spatial monitoring capabilities, will result in more commercially-available systems that can be applied by non-expert users. Given the intrinsic suitability of acoustic array methods for kiwi calls, further development of the techniques and equipment will prove advantageous for kiwi research and conservation.

CONCLUDING DISCUSSION

This thesis presents the most detailed study into kiwi acoustic ecology yet conducted, and the first investigation into the calls of the little spotted kiwi (LSK). The research had two primary motivations: to further understanding of kiwi call functions and to develop the use of acoustic methods for kiwi conservation. These objectives contribute to the wider themes of avian call function, signal evolution and conservation bioacoustics. The objectives were addressed with a series of questions, which are now reviewed to summarise the findings of this thesis.

8.1   KIWI CALL FUNCTIONS

8.1.1   *Call dimorphism*

Why do LSK have sexually dimorphic calls?

Call dimorphism within and between the sexes is unrelated to morphology in LSK (Chapter 2). This implies a functional significance for the inter-sexual vocal differences. The higher calling rates of males and the spectral properties of their calls suggest that males play a more active role in acoustic territory defence. Female calls, in contrast, are more suited for shorter-range functions such as pair contact.

The striking disparity in spectral distributions between the sexes is unrelated to body size and further accentuated by formants in female calls. This feature constrains likely duet functions in LSK, indicating that duets are used for resource defence or pair-bonding rather than mate-guarding.

8.1.2   *Spectral features*

Do LSK calls contain information-bearing spectral features such as two-voicing and non-linear phenomena, and why?

Non-linear phenomena (NLP) in the form of subharmonics are common in LSK calls (Chapter 3); the first confirmation that such features occur in palaeognaths. Their association with territorial calls and high frequencies suggests that they provide acoustic impact, possibly to signal aggressive intent or fitness. These features may indicate loss of vocal control or stress. 'Hooks' in male calls also occur more frequently with simulated territorial intrusion, suggesting that they too may convey aggression.

Non-linear features are common and diverse in nestling LSK. This is in accordance with NLP being more frequently observed in juveniles in other animals (Fitch et al. 2002), and supports the hypothesis that they add unpredictability to the calls of animals which are in danger of being ignored. These findings suggest that vocalisations provide an effective means for investigating the relationship between adults and nestlings in this species.

Two-voicing is very rare in adult LSK. This is surprising, given that kiwi have been considered likely to use such features for identification (Corfield 2005), in common with other birds unable to rely on visual cues (Aubin et al. 2000). If LSK do use calls for individual discrimination, then another mechanism must be responsible.

8.1.3 *Vocal individuality*

Do LSK have individually distinctive calls?

Vocal differences are low among LSK, particularly in females (Chapter 4). Males have more individually distinctive calls, but stereotypy varies among individuals. Why some individuals have more distinctive calls than others is unclear. It is not due to size differences, so may reflect other traits such as fitness. Call variation among male LSK may therefore be related to sexual selection.

The overall low variation in frequency and temporal call properties among LSK suggests that either other spectral features such as NLP or formants are used as identifiers, or that LSK do not rely heavily upon vocalisations for identification. The contrast with the strong frequency and temporal differences in brown kiwi (*Apteryx mantelli*) calls (Corfield 2005) suggests the latter. Alternative cues for identification include olfactory signals (Castro et al. 2010) or spatial information: the strong territoriality of LSK may reduce the requirement for individually distinctive calls (Robisson 1992). Yet these arguments also apply to other kiwi, and cannot explain the differences between the species. A more likely scenario is that the low call variation in LSK is related to their low genetic diversity.

8.1.4 *Vocal activity*

How does LSK calling activity vary temporally, and how can this inform call function theories?

The seasonal variation in LSK vocal activity shows that calls serve a clear role in reproduction (Chapter 5). The strong peak at the start of the breeding period suggests that mate attraction or pair bonding are important call functions. However, the peak in female calls in the autumn is unexplained.

The maximum in call rate at the beginning of the night indicates that calls are used for pair contact and territory defence. That weaker peak at the end of the night is more in accordance with mate contact rather than territory defence function.

The surprisingly strong effects of weather conditions on calling rates allow inferences to be made about LSK behaviour. The changes in calling activity with relative humidity and temperature are likely to be related to foraging conditions. The increase in calls with wet ground conditions may occur for the same reason, although it is unclear why kiwi should call more when feeding is easier. An alternative explanation is that wet leaf litter reduces the likelihood of detection by predators or conspecifics.

Predation risk is at odds with the lack of moon effect on LSK calling activity. It would be anticipated that with a volant predator (the laughing owl) that used both hearing and vision to detect prey, then LSK would be less active and vocal on moonlit nights. However, site-specific bias is likely responsible for this result, supported by the strong cloud-effect. The implication that kiwi are adversely affected by light pollution may have profound impacts on their conservation, amid the increasing urbanisation of New Zealand.

8.2   KIWI CONSERVATION BIOACOUSTICS

8.2.1   *Autonomous acoustic methods*

Are autonomous recorders and automatic detection methods suitable for call point counts?

This research has demonstrated the suitability of autonomous acoustic recorders and automatic detection software for kiwi monitoring. They compare well with manual call counts (Chapter 6), and the assessment of temporal and environmental effects on calling activity (Chapter 5) would not be feasible without automatic methods. Autonomous methods are highly efficient compared to human observers, and although they are less sensitive, provide sufficient areal coverage for most monit-

oring purposes. They are less effective for species with more variable calls, and in areas with higher calling rates and higher acoustic competition, but this is also true for manual counts. Due to their reduced sensitivity, autonomous methods require longer monitoring time frames to produce comparable results to human-based counts (Chapter 5 & Chapter 6). This is easily achieved, however, and results in reduced temporal bias.

8.2.2 *Call variation*

How can temporal variability in kiwi calls assessed by autonomous recorders inform monitoring programmes?

The long temporal study of calling activity provides guidance for the use of calls for kiwi monitoring (Chapter 5). Counts should not be conducted during rain or moderately windy conditions, not only because of the reduced detection probability, but also because of a consequent sex bias. The lack of moon effect indicated here would provide more flexibility for scheduling counts, but should be confirmed in other populations and species due to possible site-specific bias. Little spotted kiwi counts would be most effective at the beginning of the breeding period, and in wet ground conditions.

8.2.3 *Microphone arrays*

Are microphone arrays suitable for monitoring kiwi?

Microphone arrays would be useful in kiwi conservation, but are not easily deployed with currently-available equipment and algorithms (Chapter 7). Triangulation methods work well in open habitats, but appear to suffer from reverberation in forested environments. In addition, array size is problematic for kiwi, which often have large territories (McLennan et al. 1987). Other results suggest that arrays are currently feasible with available equipment (Mennill et al. 2012), but depend on unnaturally loud call volumes, custom software and relatively open habitats which limits the application of these results to kiwi in New Zealand. Therefore while array methods have great promise for kiwi monitoring, they are not a practical conservation tool at present.

8.2.4 *Individual identification*

The ability to identify individual animals acoustically has many applications in conservation (Terry et al. 2005). That LSK are to some extent vocally distinctive could be applied to their conservation, such as in assessing male survival. However, the overall low call variation of LSK (Chapter 4) means that acoustic identification cannot be widely utilised. Of more interest from a conservation perspective is the possibility that call variability is related to genetic diversity, which could enable inexpensive, passive and rapid genetic assessment of populations.

8.3 FUTURE WORK

A number of studies of kiwi vocalisations have been conducted (Colbourne & Kleinpaste 1984, Miles et al. 1997, Pierce & Westbrooke 2003, Hojem 2006), including recent spectrographic analyses (Corfield 2005, Corfield et al. 2008a). Yet kiwi acoustic research remains almost completely biased towards a single species, and the functional significance of kiwi calls remains unclear. This thesis has contributed to addressing that bias, and has furthered understanding of kiwi call functions. Each insight raises more questions, however, and so this study has highlighted a number of areas of research which would be fruitful to pursue in elucidating the function of kiwi calls and enhancing their use in conservation.

CALL VARIATION AND GENETIC DIVERSITY. Establishing a relation between call and genetic variation could have profound implications for kiwi conservation, by enabling approximate genetic diversity to be easily assessed. This could be achieved by measuring calls of different LSK populations for dialects, and correlating call and genetic diversity across the kiwi species.

DUET FUNCTION. Duets are a significant part of the vocal repertoire of kiwi, yet have received very little attention. A long temporal dataset from autonomous recorders could elucidate duet function from temporal variation. Playback experiments could also be used to determine whether duets have sex-specific functions (e. g. Koloff & Mennill 2011).

TEMPORAL VARIATION. This study has indicated that kiwi vocal activity may depend on previously unstudied environmental conditions such as humidity and pressure, and has raised doubts on the influence of other factors such as moonlight. Since such effects have conservation implications through guiding census

protocols, their impact should be measured in other populations and species using autonomous recorders.

OTHER KIWI SPECIES. Taxonomically and ecologically unique, ground-dwelling, and likely non-call learners, kiwi are an ideal bird group with which to test avian vocal evolution and call function theories. This could be achieved by comparing vocal features, ecology and phylogeny across the kiwi species and other palaeognaths. This study provides a basis for such comparative studies, yet still calls of only two of the five kiwi species have been quantitatively analysed. Future efforts at kiwi acoustic research should focus on the other three species, the calls of which remain virtually unstudied.

Acevedo, M. A., Corrada-Bravo, C. J., Corrada-Bravo, H., Villanueva-Rivera, L. J., & Aide, T. M. 2009. Automated classification of bird and amphibian calls using machine learning: a comparison of methods. *Ecological Informatics*, 4:206–214. (Cited on pages 95, 103, and 158.)

Aichner, R., Buchner, H., Yan, F., & Kellermann, W. 2006. A real-time blind source separation scheme and its application to reverberant and noisy acoustic environments. *Signal Processing*, 86(6):18–18. (Cited on page 201.)

Alldredge, M. W., Pollock, K., Simons, T., Collazo, J., Shriner, S., & Johnson, D. 2007. Time-of-detection method for estimating abundance from point-count surveys. *The Auk*, 124(2):653–664. (Cited on pages 27 and 157.)

Allen, M., Girod, L., Newton, R., Madden, S., Blumstein, D. T., & Estrin, D. 2008. VoxNet: an interactive, rapidly-deployable acoustic monitoring platform. In *7th International Conference on Information Processing in Sensor Networks*, pages 371–382. IEEE. (Cited on page 183.)

Amador, A., Goller, F., & Mindlin, G. 2008. Frequency modulation during song in a suboscine does not require vocal muscles. *Journal of Neurophysiology*, 99(5):2383–2389. (Cited on page 51.)

Anderson, D. R. & Burnham, K. P. 2002. Avoiding pitfalls when using information-theoretic methods. *The Journal of Wildlife Management*, 66(3):912–918. (Cited on pages 41 and 140.)

Anderson, D. R., Burnham, K. P., & Thompson, W. L. 2000. Null hypothesis testing: problems, prevalence, and an alternative. *The Journal of Wildlife Management*, 64(4):912–923. (Cited on pages 41, 71, 140, and 164.)

Anderson, M. J. 2006. Distance-based tests for homogeneity of multivariate dispersions. *Biometrics*, 62:245–253. (Cited on page 123.)

Anderson, M. J., Ellingsen, K. E., & McArdle, B. H. 2006. Multivariate dispersion as a measure of beta diversity. *Ecology Letters*, 9:683–693. (Cited on page 123.)

Appleby, B. M. & Redpath, S. M. 1997a. Indicators of male quality in the hoots of tawny owls (*Strix aluco*). *Journal of Raptor Research*, 31:65–70. (Cited on page 17.)

Appleby, B. M. & Redpath, S. M. 1997b. Variation in the male territorial hoot of the tawny owl *Strix aluco* in three English populations. *Ibis*, 139(1):152–158. (Cited on page 94.)

Araya-Ajoy, Y.-m., Chaves-Campos, J., Kalko, E. K. V., & DeWoody, J. A. 2009. High-pitched notes during vocal contests signal genetic diversity in ocellated antbirds. *PLoS ONE*, 4(12):e8137. (Cited on pages 97 and 123.)

Aubin, T. & Jouventin, P. 2002. How to vocally identify kin in a crowd: The penguin model. *Advances in the Study of Behavior*, 31:243–278. (Cited on page 95.)

Aubin, T., Jouventin, P., & Hildebrand, C. 2000. Penguins use the two-voice system to recognize each other. *Proceedings of the Royal Society B: Biological Sciences*, 267(1448):1081–1087. (Cited on pages 17, 24, 65, 93, 94, 95, and 211.)

Baker, A. & Jenkins, P. 1987. Founder effect and cultural evolution of songs in an isolated population of chaffinches, *Fringilla coelebs*, in the Chatham Islands. *Animal Behaviour*, 35(6):1793–1803. (Cited on page 27.)

Baker, M., Baker, M., & Tilghman, L. 2006. Differing effects of isolation on evolution of bird songs: examples from an island-mainland comparison of three species. *Biological Journal of the Linnean Society*, 89:331–342. (Cited on pages 19 and 27.)

Baldo, S. & Mennill, D. J. 2011. Vocal behavior of great curassows, a vulnerable Neotropical bird. *Journal of Field Ornithology*, 82(3):249–258. (Cited on pages 95 and 122.)

Ballentine, B. 2009. The ability to perform physically challenging songs predicts age and size in male swamp sparrows, *Melospiza georgiana*. *Animal Behaviour*. (Cited on page 16.)

Bang, B. 1971. Functional anatomy of the olfactory system in 23 orders of birds. *Acta Anatomica*, 79:1–76. (Cited on page 22.)

Bard, S. C., Hau, M., Wikelski, M., & Wingfield, J. C. 2008. Vocal distinctiveness and response to conspecific playback in the spotted antbird, a Neotropical suboscine. *Condor*, 104:387–394. (Cited on page 80.)

Bardeli, R., Wolff, D., Kurth, F., Koch, M., Tauchert, K. H., & Frommolt, K. H. 2010. Detecting bird sounds in a complex acoustic environment and application to bioacoustic monitoring. *Pattern Recognition Letters*, 31(12):1524–1534. (Cited on pages 158, 161, 167, 169, and 175.)

Barker, N. 2008. Bird song structure and transmission in the Neotropics: trends, methods and future directions. *Ornitologia Neotropical*, 19(2):175–199. (Cited on page 80.)

Bas, Y., Devictor, V., Moussus, J.-P., & Jiguet, F. 2008. Accounting for weather and time-of-day parameters when analysing count data from monitoring programs. *Biodiversity and Conservation*, 17(14):3403–3416. (Cited on pages 127, 128, 148, 150, 156, and 157.)

Bates, D., Maechler, M., & Bolker, B. 2012. lme4: Linear mixed-effects models using S4 classes. `cran.r-project.org/web/packages/lme4/index.html`. Downloaded 07 Jul 2013. (Cited on page 71.)

Beaver, P. 1978. Ontogeny of vocalization in the greater rhea. *The Auk*, 95(2):382–388. (Cited on pages 25 and 50.)

Beckers, G. & ten Cate, C. 2006. Nonlinear phenomena and song evolution in *Streptopelia* doves. *Acta Zoologica Sinica*, 52S:482–485. (Cited on page 67.)

Beecher, M. 1991. Signature versus perceptual adaptations for individual vocal recognition in swallows. *Behaviour*, 118(1/2):15–25. (Cited on pages 17, 93, and 94.)

Beecher, M., Beecher, I., & Lumpkin, S. 1981. Parent-offspring recognition in bank swallows (*Riparia riparia*): I. Natural history. *Animal Behaviour*, 29(1):86–94. (Cited on pages 17 and 94.)

Beletsky, L. D. 1983. Vocal mate recognition in male red-winged blackbirds, *Agelaius Phoeniceus*. *Behaviour*, 84:124–134. (Cited on page 94.)

Benedict, L. 2008. Occurrence and life history correlates of vocal duetting in North American passerines. *Journal of Avian Biology*, 39:57–65. (Cited on pages 17 and 34.)

Benesty, J., Chen, J., & Huang, Y. 2008a. Conventional beamforming techniques. In *Microphone array signal processing*. Springer, London. (Cited on pages 179 and 181.)

Benesty, J., Chen, J., & Huang, Y. 2008b. Direction-of-arrival and time-difference-of-arrival estimation. In *Microphone array signal processing*. Springer, London. (Cited on pages 178, 179, 188, and 190.)

Bentley, G. E., Wingfield, J. C., Morton, M. L., & Ball, G. F. 2000. Stimulatory effects on the reproductive axis in female songbirds by conspecific and heterospecific male song. *Hormones and behavior*, 37(3):179–189. (Cited on page 17.)

Bertelli, S. & Tubaro, P. 2002. Body mass and habitat correlates of song structure in a primitive group of birds. *Biological Journal of the Linnean Society*, 77:423–430. (Cited on page 33.)

Blumstein, D. T., Mennill, D. J., Clemins, P., Girod, L., Yao, K., Patricelli, G., Deppe, J. L., Krakauer, A. H., Clark, C., Cortopassi, K. A., Hanser, S. F., McCowan, B., Ali, A. M., & Kirschel, A. N. G. 2011. Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus. *Journal of Applied Ecology*, 48(3):758–767. (Cited on pages 26, 27, 174, 175, 176, 183, and 207.)

Blumstein, D. T. & Récapet, C. 2009. The sound of arousal: the addition of novel non-linearities increases responsiveness in marmot alarm calls. *Ethology*, 115(11):1074–1081. (Cited on pages 64, 67, and 90.)

Blumstein, D. T., Verneyre, L., & Daniel, J. C. 2004. Reliability and the adaptive utility of discrimination among alarm callers. *Proceedings of the Royal Society B: Biological Sciences*, 271:1851–1857. (Cited on page 18.)

Bolton, M. 2007. Playback experiments indicate absence of vocal recognition among temporally and geographically separated populations of Madeiran Storm-petrels *Oceanodroma castro*. *Ibis*, 149:255–263. (Cited on pages 17, 27, and 93.)

Bourgeois, K., Curé, C., Legrand, J., Gómez-Díaz, E., Vidal, E., Aubin, T., & Mathevon, N. 2007. Morphological versus acoustic analysis: what is the most efficient method for sexing yelkouan shearwaters *Puffinus yelkouan*? *Journal of Ornithology*, 148:261–269. (Cited on pages 27 and 33.)

Bower, J. & Clark, C. 2005. A field test of the accuracy of a passive acoustic location system. *Bioacoustics*, 15(1):1–14. (Cited on pages 176, 205, and 206.)

Bradbury, J. & Vehrencamp, S. 1998. *Principles of animal communication*. Sinauer Associates, Sunderland, MA, USA, 1st edition. (Cited on pages 15 and 93.)

Brandes, S. T. 2008. Automated sound recording and analysis techniques for bird surveys and conservation. *Bird Conservation International*, 18(Supplement):S163–S173. (Cited on page 158.)

Brandstein, M. 1997. A pitch-based approach to time-delay estimation of reverberant speech. In *1997 IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, USA. (Cited on pages 178, 179, and 192.)

Brandstein, M., Adcock, J., & Silverman, H. F. 1995. A practical time-delay estimator for localizing speech sources with a microphone array. *Computer Speech and Language*, 9:1–26. (Cited on page 179.)

Breslow, N. 2003. Whither PQL? *University of Washington Biostatistics Working Paper*, 192:1–24. (Cited on page 71.)

Bretagnolle, V. & Genevois, F. 1997. Geographic variation in the call of the blue petrel: effects of sex and geographical scale. *The Condor*, 99(4):985–989. (Cited on page 27.)

Bridges, A. & Dorcas, M. 2000. Temporal variation in anuran calling behavior: implications for surveys and monitoring programs. *Copeia*, 2000(2):587–592. (Cited on pages 26, 27, 128, and 157.)

Brigham, R., Gutsell, R., & Wiacek, R. 1999. Foraging behaviour in relation to the lunar cycle by Australian owlet-nightjars *Aegotheles cristatus*. *Emu*, 99:253–261. (Cited on page 152.)

Brockway, B. F. 1965. Stimulation of ovarian development and egg laying by male courtship vocalization in budgerigars (*Melopsittacus undulatus*). *Animal Behaviour*, 13(4):575–578. (Cited on page 17.)

Brown, R. M., Nichols, R. A., & Faulkes, C. G. 2010. Range expansion and hybridization in Round Island petrels (*Pterodroma* spp.): evidence from microsatellite genotypes. *Molecular Ecology*, 19(15):3157–3170. (Cited on page 19.)

Brown, T. J. & Handford, P. 2002. Why birds sing at dawn: the role of consistent song transmission. *Ibis*, 145(1):120–129. (Cited on page 149.)

Buller, W. L. 1888. A history of the birds of New Zealand. [With plates.]. *Sir Walter Lawry Buller*, London. (Cited on pages 22, 23, 129, and 152.)

Burbidge, M. L., Colbourne, R. M., Robertson, H. A., & Baker, A. J. 2003. Molecular and other biological evidence supports the recognition of at least three species of brown kiwi. *Conservation Genetics*, 4(2):167–177. (Cited on pages 20, 33, and 68.)

Carletta, J. 1999. Assessing agreement on classification tasks: the kappa statistic. *Computational Linguistics*, 22(2):1–6. (Cited on page 103.)

Castro, I. 2011. *Kiwi. A natural history*. New Holland, Auckland, New Zealand, illustrated edition. (Cited on pages 24, 25, 34, 52, 68, and 97.)

Castro, I., Cunningham, S. J., Gsell, A., Jaffe, K., Cabrera, A., & Liendo, C. 2010. Olfaction in birds: A closer look at the Kiwi (Apterygidae). *Journal of Avian Biology*, 41:213–218. (Cited on pages 22, 34, 80, and 211.)

Catchpole, C. & Slater, P. J. B. 2008. *Bird song: biological themes and variations*. Cambridge University Press, Cambridge, UK, 2nd edition. (Cited on pages 16, 17, 18, 19, 34, 50, 126, 148, and 149.)

Celis-Murillo, A., Deppe, J. L., & Allen, M. F. 2009. Using soundscape recordings to estimate bird species abundance, richness, and composition. *Journal of Field Ornithology*, 80(1):64–78. (Cited on pages 157 and 174.)

Chan, A. & Blumstein, D. 2012. Anthropogenic noise's first reverberation into community ecology. *Proceedings of the Royal Society B: Biological Sciences*, 279:2725–2726. (Cited on page 27.)

Charif, R. 2008. Automated detection of cerulean warbler songs using XBAT data template detector software. Technical report, The Cornell Lab of Ornithology, Ithaca, NY, USA. (Cited on pages 157 and 158.)

Charif, R., Strickmann, L., & Waack, A. 2010. *Raven Pro 1.4 user's manual*. The Cornell Lab of Ornithology, Ithaca, NY, USA. (Cited on pages 37, 70, 84, 99, 134, and 160.)

Charlton, B., Reby, D., & McComb, K. 2007. Female red deer prefer the roars of larger males. *Biology Letters*, 3(4):382–385. (Cited on page 54.)

Charlton, B. D., Ellis, W. A. H., McKinnon, A. J., Cowin, G. J., Brumm, J., Nilsson, K., & Fitch, W. T. 2011. Cues to body size in the formant spacing of male koala (*Phascolarctos cinereus*) bellows: honesty in an exaggerated trait. *Journal of Experimental Biology*, 214:3414–3422. (Cited on pages 54 and 79.)

Chen, J., Benesty, J., & Huang, Y. 2006. Time delay estimation in room acoustic environments: an overview. *EURASIP Journal on Applied Signal Processing*, 2006(26503):1–19. (Cited on pages 177, 178, and 179.)

Chen, J., Yip, L., Elson, J., Wang, H., & Maniezzo, D. 2003. Coherent acoustic array processing and localization on wireless sensor networks. *Proceedings of the IEEE*, 91(8):1154. (Cited on page 176.)

Cheng, J., Sun, Y., & Ji, L. 2010. A call-independent and automatic acoustic system for the individual recognition of animals: A novel model using four passerines. *Pattern Recognition*, 43(11):3846–3852. (Cited on page 96.)

Cheng, J., Xie, B., Lin, C., & Ji, L. 2012. A comparative study in birds: call-type-independent species and individual recognition using four machine-learning methods and two acoustic features. *Bioacoustics*, 21(2):157–171. (Cited on pages 95 and 96.)

Clark, J., Boersma, P., & Olmsted, D. 2006. Name that tune: call discrimination and individual recognition in Magellanic penguins. *Animal Behaviour*, 72:1141–1148. (Cited on pages 17 and 93.)

Clark, J. D. 1952. Random notes on the kiwi. *Notornis*, 4(8):211. (Cited on pages 22, 23, 129, and 152.)

Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46. (Cited on page 103.)

Colbourne, R. M. 1991. Further evidence of female incubation and family grouping in brown kiwis. *Notornis*, 38(3):248–249. (Cited on page 22.)

Colbourne, R. M. 1992. Little spotted kiwi (*Apteryx owenii*): recruitment and behaviour of juveniles on Kapiti Island, New Zealand. *Journal of the Royal Society of New Zealand*, 22(4):321–328. (Cited on page 22.)

Colbourne, R. M. 2002. Incubation behaviour and egg physiology of kiwi (*Apteryx* spp.) in natural habitats. *New Zealand Journal of Ecology*, 26(2):129–138. (Cited on pages 22 and 90.)

Colbourne, R. M. 2006. Kiwi call scheme. Technical Report S5020/129, Department of Conservation, Wellington, New Zealand. (Cited on pages 22, 23, 24, 28, 34, 97, 128, 129, 148, 158, and 177.)

Colbourne, R. M., Baird, K., & Jolly, J. 1990. Relationship between invertebrates eaten by little spotted kiwi, *Apteryx owenii*, and their availability on Kapiti Island, New Zealand. *New Zealand Journal of Zoology*, 17:533–542. (Cited on page 151.)

Colbourne, R. M. & Kleinpaste, R. 1983. A banding study of North Island brown kiwis in an exotic forest. *Notornis*, 30(2):109–124. (Cited on pages 22, 23, and 52.)

Colbourne, R. M. & Kleinpaste, R. 1984. North Island brown kiwi vocalisations and their use in censusing populations. *Notornis*, 31(3):191–201. (Cited on pages 22, 23, 25, 33, 52, 53, 68, 128, 129, 131, 148, 149, 152, 159, and 214.)

Colbourne, R. M. & Robertson, H. 1997. Successful translocations of little spotted kiwi (*Apteryx owenii*) between offshore islands of New Zealand. *Notornis*, 44:253–258. (Cited on pages 20, 22, 38, and 167.)

Collier, T. 2011. *Wireless sensor network-based acoustic localization for studying animal communication in terrestrial environments*. PhD thesis, University of California, Los Angeles. (Cited on page 182.)

Collier, T. C., Blumstein, D. T., Girod, L., & Taylor, C. E. 2010a. Is alarm calling risky? Marmots avoid calling from risky places. *Ethology*, 116(12):1171–1178. (Cited on page 176.)

Collier, T. C., Kirschel, A. N. G., & Taylor, C. E. 2010b. Acoustic localization of antbirds in a Mexican rainforest using a wireless sensor network. *The Journal of the Acoustical Society of America*, 128(1):182. (Cited on pages 175, 177, 182, 183, and 207.)

Collins, S. 2004. Vocal fighting and flirting: the functions of birdsong. In Marler, P. & Slabbekoorn, H., editors, *Nature's Music: The Science of Birdsong*, pages 39–79. Elsevier Academic Press, London, UK. (Cited on pages 17, 25, 33, and 55.)

Corfield, J. 2005. Description, dueting, seasonal variations, and individual identification of the vocalisations of the Brown Kiwi (*Apteryx mantelli*). Master's thesis, Auckland University of Technology, Auckland, New Zealand. (Cited on pages 22, 23, 24, 25, 33, 34, 53, 54, 55, 68, 77, 80, 97, 100, 101, 117, 123, 177, 211, and 214.)

Corfield, J., Gillman, L., & Parsons, S. 2008a. Vocalisations of the North Island brown kiwi (*Apteryx mantelli*). *The Auk*, 125(2):326–335. (Cited on pages 22, 23, 24, 25, 33, 34, 50, 51, 53, 54, 55, 68, 80, 97, 177, 207, and 214.)

Corfield, J., Kubke, M. F., Parsons, S., Wild, J. M., & Köppl, C. 2011. Evidence for an auditory fovea in the New Zealand kiwi (*Apteryx mantelli*). *PLoS ONE*, 6(8):e23771. (Cited on pages 22, 34, 50, and 55.)

Corfield, J. R. 2009. *Evolution of the brain and sensory systems of the kiwi*. PhD thesis, University of Auckland, Auckland, New Zealand. (Cited on pages 22 and 25.)

Corfield, J. R., Wild, J. M., Hauber, M. E., Parsons, S., & Kubke, M. F. 2008b. Evolution of brain size in the Palaeognath lineage, with an emphasis on New Zealand ratites. *Brain, Behaviour and Evolution*, 71(2):87–99. (Cited on page 50.)

Cortes, C. & Vapnik, V. 1995. Support-vector networks. *Machine Learning*, 20(3):273–297. (Cited on page 103.)

Cortopassi, K. A. 2006. Automated and robust measurement of signal features. Cornell Laboratory of Ornithology, Ithaca, NY, USA. www.birds.cornell.edu/brp/research/algorithm/automated-and-robust-measurement-of-signal-features. Accessed 17 May 2012. (Cited on page 37.)

Cunningham, S. J. & Castro, I. 2011. The secret life of wild brown kiwi: studying behaviour of a cryptic species by direct observation. *New Zealand Journal of Ecology*, 35(3):209–219. (Cited on page 151.)

Cunningham, S. J., Castro, I., & Alley, M. R. 2007. A new prey-detection mechanism for kiwi (*Apteryx* spp.) suggests convergent evolution between paleognathous and neognathous birds. *Journal of Anatomy*, 211(4):493–502. (Cited on pages 22 and 34.)

Cunningham, S. J., Castro, I., & Potter, M. 2009. The relative importance of olfaction and remote touch in prey detection by North Island brown kiwis. *Animal Behaviour*, 78(4):899–905. (Cited on pages 22 and 34.)

Curé, C., Aubin, T., & Mathevon, N. 2009. Acoustic convergence and divergence in two sympatric burrowing nocturnal seabirds. *Biological Journal of the Linnean Society*. (Cited on pages 17 and 93.)

Curé, C., Mathevon, N., Mundry, R., & Aubin, T. 2012. Acoustic cues used for species recognition can differ between sexes and sibling species: evidence in shearwaters. *Animal Behaviour*. (Cited on pages 17 and 93.)

Dantzker, M., Deane, G., & Bradbury, J. 1999. Directional acoustic radiation in the strut display of male sage grouse *Centrocercus urophasianus*. *Journal of Experimental Biology*, 202:2893–2909. (Cited on page 176.)

Dawson, D. & Bull, P. 1975. Counting birds in New Zealand forests. *Notornis*, 22(2):101–109. (Cited on pages 25, 94, and 127.)

Dawson, D. K. & Efford, M. G. 2009. Bird population density estimated from acoustic signals. *Journal of Applied Ecology*, 46:1201–1209. (Cited on pages 174 and 175.)

de Kort, S. R., Eldermire, E. R. B., Valderrama, S., Botero, C. A., & Vehrencamp, S. L. 2009. Trill consistency is an age-related assessment signal in banded wrens. *Proceedings of the Royal Society B: Biological Sciences*, 276(1665):2315–2321. (Cited on page 16.)

den Hartog, P., de Kort, S., & ten Cate, C. 2007. Hybrid vocalizations are effective within, but not outside, an avian hybrid zone. *Behavioral Ecology*, 18:608–614. (Cited on page 19.)

Derryberry, E. P., Seddon, N., Claramunt, S., Tobias, J. A., Baker, A., Aleixo, A., & Brumfield, R. T. 2012. Correlated evolution of beak morphology and song in the neotropical woodcreeper radiation. *Evolution*, 66(9):2784–2797. (Cited on page 19.)

Diefenbach, D., Marshall, M., & Mattice, J. 2007. Incorporating availability for detection in estimates of bird abundance. *The Auk*, 124(1):96–106. (Cited on pages 27 and 157.)

Digby, A., Bell, B. D., & Teal, P. D. 2013. Vocal cooperation between the sexes in Little Spotted Kiwi *Apteryx owenii*. *Ibis*, 155(2):229–245. (Cited on pages 68, 69, 70, 78, 79, 80, 132, and 161.)

Doutrelant, C. & Lambrechts, M. 2001. Macrogeographic variation in song – a test of competition and habitat effects in blue tits. *Ethology*, 107:533–544. (Cited on page 19.)

Draganoiu, T., Nagle, L., Musseau, R., & Kreutzer, M. 2006. In a songbird, the black redstart, parents use acoustic cues to discriminate between their different fledglings. *Animal Behaviour*, 71:1039–1046. (Cited on page 94.)

Eda-Fujiwara, H., Yamamoto, A., Sugita, H., Takahashi, Y., Kojima, Y., Sakashita, R., Ogawa, H., Miyamoto, T., & Kimura, T. 2004. Sexual dimorphism of acoustic signals in the oriental white stork: non-invasive identification of sex in birds. *Zoological Science*, 21(8):817–821. (Cited on page 51.)

Emlen, J. & Dejong, M. 1992. Counting birds: the problem of variable hearing abilities. *Journal of Field Ornithology*, 63:26–31. (Cited on page 156.)

Endler, J. A. 1992. Signals, signal conditions, and the direction of evolution. *The American Naturalist*, 139:S125–S153. (Cited on page 52.)

Eriksen, T. 1999. Booming bitterns *Botaurus stellaris*: ecology, behaviour and conservation implications. Master's thesis, Copenhagen University, Copenhagen, Denmark. (Cited on pages 26 and 27.)

Ey, E. & Fischer, J. 2009. The "acoustic adaptation hypothesis" - a review of the evidence from birds, anurans and mammals. *Bioacoustics*, 19(1-2):21–48. (Cited on page 19.)

Falla, R. A. 1979. Some aspects of New Zealand birdlife. *Sanderson Memorial Lecture, Forest and Bird*, 211:18–25. (Cited on page 22.)

Falls, J. B. & Brooks, R. J. 1975. Individual recognition by song in white-throated sparrows. II. Effects of location. *Canadian Journal of Zoology*, 59:2380–2385. (Cited on page 94.)

Farnsworth, A. & Lovette, I. J. 2008. Phylogenetic and ecological effects on interspecific variation in structurally simple avian vocalizations. *Biological Journal of the Linnean Society*, 94(1):155–173. (Cited on page 19.)

Farnsworth, A. & Russell, R. W. 2007. Monitoring flight calls of migrating birds from an oil platform in the northern Gulf of Mexico. *Journal of Field Ornithology*, 78(3):279–289. (Cited on page 27.)

Fedy, B. C. & Stutchbury, B. J. M. 2005. Territory defence in tropical birds: are females as aggressive as males? *Behavioral Ecology and Sociobiology*, 58:414–422. (Cited on pages 17 and 55.)

Fee, M. S., Shralman, B., Pesaran, B., & Mitra, P. P. 1998. The role of nonlinear dynamics of the syrinx in the vocalizations of a songbird. *Nature*, 395(3):67–71. (Cited on pages 65 and 67.)

Filatova, O. A., Deecke, V. B., Ford, J. K. B., Matkin, C. O., Barrett-Lennard, L. G., Guzeev, M. A., Burdin, A. M., & Hoyt, E. 2012. Call diversity in the North Pacific killer whale populations: implications for dialect evolution and population history. *Animal Behaviour*, 83(3):595–603. (Cited on page 67.)

Fitch, W. 1997. Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *The Journal of the Acoustical Society of America*, 102(2):1213–1222. (Cited on page 54.)

Fitch, W. & Fritz, J. 2006. Rhesus macaques spontaneously perceive formants in conspecific vocalizations. *The Journal of the Acoustical Society of America*, 120(4):2132–2141. (Cited on page 96.)

Fitch, W. T. 1999. Acoustic exaggeration of size in birds via tracheal elongation: comparative and theoretical analyses. *Journal of Zoology*, 248(01):31–48. (Cited on pages 51 and 54.)

Fitch, W. T., Neubauer, J., & Herzel, H. 2002. Calls out of chaos: the adaptive significance of nonlinear phenomena in mammalian vocal production. *Animal Behaviour*, 63(3):407–418. (Cited on pages 64, 66, 67, 79, 80, 86, 90, 96, and 211.)

Fitzsimmons, L. P., Barker, N. K., & Mennill, D. J. 2008a. Individual variation and lek-based vocal distinctiveness in songs of the screaming piha (*Lipaugus vociferans*), a suboscine songbird. *The Auk*, 125(4):908–914. (Cited on pages 38, 100, and 101.)

Fitzsimmons, L. P., Foote, J. R., & Ratcliffe, L. M. 2008b. Frequency matching, overlapping and movement behaviour in diurnal countersinging interactions of black-capped chickadees. *Animal Behaviour*, 75:1913–1920. (Cited on pages 27, 158, and 176.)

Fitzsimmons, L. P., Foote, J. R., Ratcliffe, L. M., & Mennill, D. J. 2008c. Eavesdropping and communication networks revealed through playback and an acoustic location system. *Behavioral Ecology*, 19(4):824–829. (Cited on pages 27, 176, and 183.)

Forbes, W. A. 1881. On the conformation of the thoracic end of the trachea in the "ratite" birds. *Proceedings of the Zoological Society of London*, 49(3):778–788. (Cited on page 54.)

Fox, E. J. S. 2008a. A new perspective on acoustic individual recognition in animals with limited call sharing or changing repertoires. *Animal Behaviour*, 75(3):1187–1194. (Cited on pages 25, 96, and 122.)

Fox, E. J. S. 2008b. *Call-independent identification in birds*. PhD thesis, University of Western Australia, Perth, Australia. (Cited on page 96.)

Galeotti, P. 1998. Correlates of hoot rate and structure in male tawny owls *Strix aluco*: implications for male rivalry and female mate choice. *Journal of Avian Biology*, pages 25–32. (Cited on pages 16, 17, and 18.)

Galeotti, P. & Pavan, G. 1991. Individual recognition of male tawny owls (*Strix aluco*) using spectrograms of their territorial calls. *Ethology Ecology & Evolution*, 3(2):113–126. (Cited on pages 17, 24, 94, and 95.)

Galeotti, P. & Pavan, G. 1993. Differential responses of territorial tawny owls *Strix aluco* to the hooting of neighbours and strangers. *Ibis*, 135(3):300–304. (Cited on page 17.)

Galeotti, P. & Sacchi, R. 2001. Turnover of territorial scops owls *Otus scops* as estimated by spectrographic analyses of male hoots. *Journal of Avian Biology*, 32(3):256–262. (Cited on pages 27 and 95.)

Gaunt, S. L. L. & McCallum, D. A. 2004. Birdsong and conservation. In Marler, P. & Slabbekoorn, H., editors, *Nature's Music: The Science of Birdsong*. Elsevier Academic Press, London. (Cited on page 27.)

Geberzahn, N., Goymann, W., Muck, C., & ten Cate, C. 2009. Females alter their song when challenged in a sex-role reversed bird species. *Behavioral Ecology and Sociobiology*, 64(2):193–204. (Cited on pages 17, 25, 33, and 96.)

Genevois, F. & Bretagnolle, V. 1994. Male blue petrels reveal their body mass when calling. *Ethology Ecology & Evolution*, 6:377–383. (Cited on page 17.)

Gilbert, G., McGregor, P. K., & Tyler, G. 1994. Vocal individuality as a census tool: practical considerations illustrated by a study of two rare species. *Journal of Field Ornithology*, 65(3):335–348. (Cited on pages 17, 26, 94, and 96.)

Gilbert, G., Tyler, G., & Smith, K. 2002. Local annual survival of booming male great bittern *Botaurus stellaris* in Britain, in the period 1990–1999. *Ibis*, 144(1):51–61. (Cited on pages 27, 95, and 96.)

Girod, L., Lukac, M., Trifa, V., & Estrin, D. 2006. The design and implementation of a self-calibrating distributed acoustic sensing platform. In *The 4th ACM Conference on Embedded Networked Sensor Systems*, pages 71–84, New York, NY, USA. ACM Press. (Cited on page 183.)

Grant, B. & Grant, P. 2010. Songs of Darwin's finches diverge when a new species enters the community. *Proceedings of the National Academy of Sciences of the United States of America*, 107(47):20156–20163. (Cited on pages 19 and 33.)

Greenewalt, C. H. 1968. *Bird song: acoustics and physiology*. Smithsonian Institution Press, Washington, USA. (Cited on page 65.)

Guillette, L. M. & Sturdy, C. B. 2011. Individual differences and repeatability in vocal production: stress-induced calling exposes a songbird's personality. *Naturwissenschaften*, 98(11):977–981. (Cited on page 16.)

Hall, M. E. 2009. A review of vocal duetting in birds. *Advances in the Study of Behavior*, 40:67–121. (Cited on pages 17, 23, 34, 53, and 55.)

Hardouin, L. A., Reby, D., Bavoux, C., Burneleau, G., & Bretagnolle, V. 2007. Communication of male quality in owl hoots. *The American Naturalist*, 169(4):552–562. (Cited on page 17.)

Heather, B. & Robertson, H. 2005. *The field guide to the birds of New Zealand*. Penguin Books, Auckland, New Zealand, 2nd edition. (Cited on pages 20, 22, 23, 24, 25, 33, 35, 52, 68, 70, 83, and 177.)

Herting, B. L. & Belthoff, J. R. 2001. Bounce and double trill songs of male and female western screech-owls: characterization and usefulness for classification of sex. *The Auk*, 118(4):1095–1101. (Cited on page 33.)

Hobson, K., Rempel, R., Greenwood, H., Turnbull, B., & van Wilgenburg, S. 2002. Acoustic surveys of birds using electronic recordings: new potential from an omnidirectional microphone system. *Wildlife Society Bulletin*, pages 709–720. (Cited on page 26.)

Hoese, W., Podos, J., Boetticher, N., & Nowicki, S. 2000. Vocal tract function in birdsong production: experimental manipulation of beak movements. *Journal of Experimental Biology*, 203:1845–1855. (Cited on page 51.)

Hojem, C. 2006. The real New Zealand national anthem: calling structure and calling rate of North Island brown kiwi (*Apteryx mantelli*) on Ponui Island and comparisons with Rarewarewa Reserve. Master's thesis, University of Auckland, Auckland, New Zealand. (Cited on pages 22, 23, 24, 33, 68, 97, and 214.)

Holzapfel, S. A., Robertson, H. A., McLennan, J. A., Sporle, W., Hackwell, K., & Impey, M. 2008. Kiwi (*Apteryx* spp.) recovery plan 2008–2018. Threatened Species

Recovery Plan 60. Department of Conservation, Wellington, New Zealand. (Cited on pages 20, 21, 22, 34, 35, 68, and 167.)

Hu, W., Bulusu, N., Chou, C. T., Jha, S., Taylor, A., & Tran, V. N. 2009. Design and evaluation of a hybrid sensor network for cane toad monitoring. *ACM Transactions on Sensor Networks (TOSN)*, 5(1):4:1–4:28. (Cited on page 176.)

Hultsch, H. & Todt, D. 2004. Learning to sing . In Marler, P. & Slabbekoorn, H., editors, *Nature's Music: The Science of Birdsong*. Elsevier, London. (Cited on page 18.)

Hüppop, O., Dierschke, J., Exo, K.-M., Fredrich, E., & Hill, R. 2006. Bird migration studies and potential collision risk with offshore wind turbines. *Ibis*, 148:90–109. (Cited on page 27.)

Hüppop, O. & Hilgerloh, G. 2012. Flight call rates of migrating thrushes: effects of wind conditions, humidity and time of day at an illuminated offshore platform. *Journal of Avian Biology*, 43(1):85–90. (Cited on pages 27 and 127.)

Hurvich, C. M. & Tsai, C.-L. 1989. Regression and time series model selection in small samples. *Biometrika*, 76(2):297–307. (Cited on pages 41 and 164.)

Hutto, R. L. & Stutzman, R. J. 2009. Humans versus autonomous recording units: a comparison of point-count results. *Journal of Field Ornithology*, 80(4):387–398. (Cited on pages 26, 156, 157, 158, 169, and 171.)

Hyman, J. & Hughes, M. 2006. Territory owners discriminate between aggressive and nonaggressive neighbours. *Animal Behaviour*, 72:209–215. (Cited on page 94.)

Isler, M., Isler, P., & Whitney, B. 1997. Biogeography and systematics of the *Thamnophilus punctatus* (Thamnophilidae) complex. *Ornithological Monographs*, 48:355–381. (Cited on page 27.)

IUCN 2011. IUCN Red List of threatened species. Version 2011.2. www.iucnredlist.org. Accessed 17 May 2012. (Cited on pages 20, 34, 78, and 97.)

Janicke, T., Hahn, S., Ritz, M., & Peter, H. 2008. Vocal performance reflects individual quality in a nonpasserine. *Animal Behaviour*, 75(1):91–98. (Cited on pages 16 and 17.)

Janicke, T., Ritz, M. S., Hahn, S., & Peter, H.-U. 2007. Sex recognition in brown skuas: do acoustic signals matter? *Journal of Ornithology*, 148:565–569. (Cited on page 33.)

Jarvis, E. 2004. Learned birdsong and the neurobiology of human language. *Annals of the New York Academy of Sciences*, 1016:749–777. (Cited on pages 24 and 34.)

Jiménez, I., Londoño, G. A., & Cadena, C. D. 2003. Efficiency, bias, and consistency of visual and aural surveys of curassows (Cracidae) in tropical forests. *Journal of Field Ornithology*, 74(3):210–216. (Cited on page 26.)

Jolly, J. 1989. A field study of the breeding biology of the little spotted kiwi (*Apteryx owenii*) with emphasis on the causes of nest failures. *Journal of the Royal Society of New Zealand*, 19:433–447. (Cited on pages 22, 52, 90, and 98.)

Jolly, J. & Colbourne, R. M. 1991. Translocations of the little spotted kiwi (*Apteryx owenii*) between offshore islands of New Zealand. *Journal of the Royal Society of New Zealand*, 21(2):143–149. (Cited on pages 20 and 158.)

Jones, D. & Smith, G. 1997. Vocalisations of the marbled frogmouth II: an assessment of vocal individuality as a potential census technique. *Emu*, 97:296–304. (Cited on pages 17, 27, 36, 94, and 100.)

Jones, I. L., Falls, J. B., & Gaston, A. J. 1987. Vocal recognition between parents and young of ancient murrelets, *Synthliboramphus antiquus* (Aves: Alcidae). *Animal Behaviour*, 35(5):1405–1415. (Cited on pages 94 and 95.)

Jouventin, P. & Aubin, T. 2002. Acoustic systems are adapted to breeding ecologies: individual recognition in nesting penguins. *Animal Behaviour*, 64(5):747–757. (Cited on pages 65 and 94.)

Karatzoglou, A., Smola, A., Hornik, K., & Zeileis, A. 2004. kernlab - An S4 package for kernel methods in R. *Journal of Statistical Software*, 11(9):1–20. (Cited on page 103.)

Kendeigh, S. C. 1944. Measurement of bird populations. *Ecological Monographs*, 14(1):67–106. (Cited on page 25.)

Kennedy, R. S., Fisher, T. H., Harrap, S., Diesmos, A. C., & Manamtam, A. S. 2001. A new species of woodcock (Aves: Scolopacidae) from the Philippines and a re-evaluation of other Asian/Papuasian woodcock. *Forktail*, 17:1–12. (Cited on page 19.)

Kirschel, A. N. G., Cody, M. L., Harlow, Z. T., Promponas, V. J., Vallejo, E. E., & Taylor, C. E. 2011. Territorial dynamics of Mexican ant-thrushes *Formicarius moniliger*

revealed by individual recognition of their songs. *Ibis*, 153(2):255–268. (Cited on pages 16, 26, 27, 33, 55, 95, 101, 117, 158, 175, and 177.)

Kirschel, A. N. G., Earl, D. A., Yao, Y., Escobar, I. A., Viches, E., Vallejo, E. E., & Taylor, C. E. 2009. Using songs to identify individual Mexican antthrush *Tormicarius moniliger*: comparison of four classification methods. *Bioacoustics*, 19:1–20. (Cited on pages 17, 94, and 95.)

Klenova, A. V., Volodin, I. A., & Volodina, E. V. 2009. The variation in reliability of individual vocal signature throughout ontogenesis in the red-crowned crane *Grus japonensis*. *Acta Ethologica*, 12(1):29–36. (Cited on pages 94 and 96.)

Knapp, C. & Carter, G. 1976. The generalized correlation method for estimation of time delay. *IEEE Transactions On Acoustics, Speech, And Signal Processing*, 24(4):320–327. (Cited on pages 178 and 179.)

Knudsen, B. & Evans, R. M. 1986. Parent-young recognition in herring gulls (*Larus argentatus*). *Animal Behaviour*, 34(1):77–80. (Cited on pages 17 and 93.)

Koloff, J. & Mennill, D. 2011. Aggressive responses to playback of solos and duets in a Neotropical antbird. *Animal Behaviour*, 82:587–593. (Cited on pages 27 and 214.)

Komsta, L. 2010. moonsun: Basic astronomical calculations with R. Version 0.1.2. Downloaded 21 Sep 2012. (Cited on page 136.)

Kraemer, H. C., Periyakoil, V. S., & Noda, A. 2004. *Kappa coefficients in medical research*, volume 1 of *Tutorials in biostatistics*. John Wiley & Sons, Hoboken, NJ, USA. (Cited on page 103.)

Krakauer, A., Tyrrell, M., Lehmann, K., Losin, N., Goller, F., & Patricelli, G. 2009. Vocal and anatomical evidence for two-voiced sound production in the greater sage-grouse *Centrocercus urophasianus*. *Journal of Experimental Biology*, 212:3719–3727. (Cited on pages 65, 95, and 176.)

Kroodsma, D. 2005. *The singing life of birds*. Houghton Mifflin, Boston, MA, USA. (Cited on pages 24, 98, 123, and 177.)

Kroodsma, D. E. 2004. The diversity and plasticity of birdsong. In Marler, P. & Slabbekoorn, H., editors, *Nature's Music: The Science of Birdsong*. Elsevier Academic Press, London. (Cited on pages 18, 24, and 96.)

Krull, C. R. & Hauber, M. E. 2010. Two voice system found in the Australasian gannet (*Morus serrator*). *Notornis*, 57(3):162–165. (Cited on pages 65 and 95.)

Kuhn, M. 2008. Building predictive models in R using the caret package. *Journal of Statistical Software*, 28(5):1–26. (Cited on page 103.)

Kyba, C. C. M., Ruhtz, T., Fischer, J., & Holker, F. 2011. Cloud coverage acts as an amplifier for ecological light pollution in urban ecosystems. *PLoS ONE*, 6(3):e17307. (Cited on pages 151 and 152.)

Lack, D. 1953. *The life of the robin*. Penguin Books, London, UK, 3rd edition. (Cited on page 18.)

Laje, R. & Mindlin, G. 2005. Modeling source-source and source-filter acoustic interaction in birdsong. *Physical Review E*, 72(3):036218. (Cited on page 66.)

Laje, R., Sciamarella, D., Zanella, J., & Mindlin, G. B. 2008. Bilateral source acoustic interaction in a syrinx model of an oscine bird. *Physical Review E*, 77(1):011912. (Cited on pages 66 and 78.)

Langmore, N. 1998. Functions of duet and solo songs of female birds. *Trends in Ecology & Evolution*, 13(4):136–140. (Cited on pages 16 and 25.)

Larom, D., Garstang, M., Payne, K., Raspet, R., & Lindeque, M. 1997. The influence of surface atmospheric conditions on the range and area reached by animal vocalizations. *Journal of Experimental Biology*, 200:421–431. (Cited on pages 127 and 149.)

Lessells, C., Rowe, C., & McGregor, P. K. 1995. Individual and sex differences in the provisioning calls of European bee-eaters. *Animal Behaviour*, 49:244–247. (Cited on page 94.)

Lombard, A., Rosenkranz, T., Buchner, H., & Kellermann, W. 2009. Multidimensional localization of multiple sound sources using averaged directivity patterns of blind source separation systems. In *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 233–236, Tapei, Taiwan. IEEE. (Cited on page 201.)

Longcore, T. & Rich, C. 2004. Ecological light pollution. *Frontiers in Ecology and the Environment*, 2(4):191–198. (Cited on page 151.)

Luther, D. A. & Derryberry, E. P. 2012. Birdsongs keep pace with city life: changes in song over time in an urban songbird affects communication. *Animal Behaviour*, 83(4):1059–1066. (Cited on pages 19 and 33.)

Luukkonen, D., Prince, H., & Mao, I. 1997. Evaluation of pheasant crowing rates as a population index. *The Journal of Wildlife Management*, 61(4):1338–1344. (Cited on page 94.)

Mack, A. L. & Jones, J. 2003. Low-frequency vocalizations by cassowaries (*Casuarius* spp.). *The Auk*, 120(4):1062–1068. (Cited on pages 24, 25, and 50.)

Maloney, M. 1996. Inaccuracy of a radio-tracking system for small mammals: The effect of electromagnetic interference. *Journal of Zoology*, 239(2):401–406. (Cited on page 177.)

Mann, D. A., O'Shea, T. J., & Nowacek, D. P. 2006. Nonlinear dynamics in manatee vocalizations. *Marine Mammal Science*, 22(3):548–555. (Cited on pages 66 and 78.)

Marks, E., Rodrigo, A., & Brunton, D. 2010. Ecstatic display calls of the Adelie penguin honestly predict male condition and breeding success. *Behaviour*, 147(2):165–184. (Cited on pages 17 and 96.)

Marler, P. 2004. Bird calls: a cornucopia for communication. In Marler, P. & Slabbekoorn, H., editors, *Nature's Music: The Science of Birdsong*, pages 132–177. Elsevier Academic Press, London, UK. (Cited on pages 16, 18, and 33.)

Marshall, R. C., Buchanan, K. L., & Catchpole, C. K. 2003. Sexual selection and individual genetic diversity in a songbird. *Biology Letters*, 270 Suppl 2:S248–50. (Cited on page 123.)

Marten, K. & Marler, P. 1977. Sound transmission and its significance for animal vocalization. *Behavioral Ecology and Sociobiology*, 2(3):271–290. (Cited on pages 50 and 52.)

Marten, K., Quine, D., & Marler, P. 1977. Sound transmission and its significance for animal vocalization. II. Tropical forest habitats. *Behavioral Ecology and Sociobiology*, 2:291–302. (Cited on pages 50, 52, and 151.)

Martin, G., Rojas, L., Ramírez, Y., & McNeil, R. 2004. The eyes of oilbirds (*Steatornis caripensis*): pushing at the limits of sensitivity. *Naturwissenschaften*, 91:26–29. (Cited on page 25.)

Martin, G. R., Wilson, K.-J., Martin Wild, J., Parsons, S., Kubke, M. F., & Corfield, J. 2007. Kiwi forego vision in the guidance of their nocturnal activities. *PLoS ONE*, 2(2):e198. (Cited on pages 22, 25, 34, 50, and 68.)

Martin, J. P., Doucet, S. M., Knox, R. C., & Mennill, D. J. 2011. Body size correlates negatively with the frequency of distress calls and songs of Neotropical birds. *Journal of Field Ornithology*, 82(3):259–268. (Cited on pages 18, 33, and 51.)

Maurer, G., Double, M. C., Milenkaya, O., Süsser, M., & Magrath, R. D. 2011. Breaking the rules: sex roles and genetic mating system of the pheasant coucal. *Oecologia*, 167(2):413–425. (Cited on pages 52 and 55.)

Maurer, G., Smith, C., Süsser, M., & Magrath, R. D. 2008. Solo and duet calling in the pheasant coucal: sex and individual call differences in a nesting cuckoo with reversed size dimorphism. *Australian Journal of Zoology*, 56(3):143. (Cited on pages 18, 33, 95, and 101.)

May, L. 1994. Individually distinctive corncrake *Crex rex* calls: a pilot study. *Bioacoustics*, 6:25–32. (Cited on page 24.)

May, L. 1998. Individually distinctive corncrake *Crex rex* calls: a further study. *Bioacoustics*, 9:135–148. (Cited on page 26.)

McCracken, K. & Sheldon, F. 1997. Avian vocalizations and phylogenetic signal. *Proceedings of the National Academy of Sciences of the United States of America*, 94(8):3833–3836. (Cited on page 19.)

McDonald, P. G. 2012. Cooperative bird differentiates between the calls of different individuals, even when vocalizations were from completely unfamiliar individuals. *Biology Letters*, 8(3):365–368. (Cited on pages 18 and 94.)

McGregor, P. K. 1993. Signalling in territorial systems: a context for individual identification, ranging and eavesdropping. *Philosophical Transactions: Biological Sciences*, 340(1292):237–244. (Cited on pages 17, 93, and 94.)

McGregor, P. K., Dabelsteen, T., Clark, C., Bower, J., & Holland, J. 1997. Accuracy of a passive acoustic location system: empirical studies in terrestrial habitats. *Ethology Ecology & Evolution*, 9(3):269–286. (Cited on pages 176, 177, 204, and 205.)

McKown, M. 2008. *Acoustic communication in colonial seabirds: individual, sexual, and species-specific variation in acoustic signals of Pterodroma petrels*. PhD thesis, University of North Carolina. (Cited on pages 17, 93, 95, 96, 102, 117, and 127.)

McLennan, J. 1988. Breeding of North Island brown kiwi, *Apteryx australis mantelli*, in Hawke's Bay, New Zealand. *New Zealand Journal of Ecology*, 11:89–97. (Cited on pages 22 and 52.)

McLennan, J. 1992. Nationwide monitoring of kiwi populations. Technical Report No. 92/21, DSIR Land Resources, Havelock North, New Zealand. (Cited on pages 28 and 128.)

McLennan, J., Dew, L., Miles, J., Gillingham, N., & Waiwai, R. 2004. Size matters: predation risk and juvenile growth in North Island brown kiwi (*Apteryx mantelli*). *New Zealand Journal of Ecology*, 28(2):241–250. (Cited on pages 20 and 90.)

McLennan, J. & McCann, T. J. 2002. Genetic variability, distribution and abundance of great spotted kiwi (*Apteryx haastii*). In Overmars, F., editor, *Some early 1990s studies in kiwi (*Apteryx *spp.) genetics and management*. Science and Research Internal Report 191, Department of Conservation, Wellington, New Zealand. (Cited on pages 23 and 25.)

McLennan, J., Rudge, M. R., & Potter, M. A. 1987. Range size and denning behavior of brown kiwi, *Apteryx australis mantelli*, in Hawkes Bay, New Zealand. *New Zealand Journal of Ecology*, 10:97–107. (Cited on pages 22, 24, 90, 150, 207, and 213.)

Mennill, D., Burt, J., Fristrup, K., & Vehrencamp, S. 2006. Accuracy of an acoustic location system for monitoring the position of duetting songbirds in tropical forest. *The Journal of the Acoustical Society of America*, 119:2832. (Cited on pages 27, 175, 176, 177, 183, 203, and 207.)

Mennill, D. & Vehrencamp, S. 2008a. Context-dependent functions of avian duets revealed by microphone-array recordings and multispeaker playback. *Current Biology*, 18(17):1314–1319. (Cited on pages 158 and 176.)

Mennill, D. & Vehrencamp, S. 2008b. Supplement to context-dependent functions of avian duets revealed by microphone-array recordings and multispeaker playback. *Current Biology*, 18:1–6. (Cited on page 26.)

Mennill, D. J. 2011. Individual distinctiveness in avian vocalizations and the spatial monitoring of behaviour. *Ibis*, 153(2):235–238. (Cited on pages 26, 27, 55, and 175.)

Mennill, D. J., Battiston, M., Wilson, D. R., Foote, J. R., & Doucet, S. M. 2012. Field test of an affordable, portable, wireless microphone array for spatial monitoring of

animal ecology and behaviour. *Methods in Ecology and Evolution*, 3(4):704–712. (Cited on pages 158, 176, 177, 183, 203, 204, 205, 206, 207, and 213.)

Miles, J. & McLennan, J. 1997. A new technique for radio-tagging immature kiwi (*Apteryx* spp.). *Notornis*, 4(5):44–47. (Cited on page 23.)

Miles, J., Potter, M., & Fordham, R. 1997. Northern brown kiwi (*Apteryx australis mantelli*) in Tongariro National Park and Tongariro Forest - ecology and threats. *Science for Conservation*, 51:1–23. (Cited on pages 22, 23, 24, 128, 129, 131, 148, 149, 150, 151, 152, 159, and 214.)

Miller, E. H. & Baker, A. J. 2009. Antiquity of shorebird acoustic displays. *The Auk*, 126(2):454–459. (Cited on page 19.)

Miller, H. C. H., Bowker-Wright, G. G., Kharkrang, M. M., & Ramstad, K. K. 2011. Characterisation of class II B MHC genes from a ratite bird, the little spotted kiwi (*Apteryx owenii*). *Immunogenetics*, 63(4):223–233. (Cited on page 97.)

Miller, P. & Pierce, R. 1995. Distribution and decline of the North Island brown kiwi (*Apteryx australis mantelli*) in Northland. *Notornis*, 42(3):203–211. (Cited on pages 23, 27, and 28.)

Miller, P. J. O. 2002. Mixed-directionality of killer whale stereotyped calls: a direction of movement cue? *Behavioral Ecology and Sociobiology*, 52(3):262–270. (Cited on page 67.)

Miskelly, C., Dowding, J., Elliott, G., Hitchmough, R., Powlesland, R., Robertson, H., Sagar, P., Scofield, R., & Taylor, G. 2008. Conservation status of New Zealand birds, 2008. *Notornis*, 55:117–135. (Cited on pages 20 and 97.)

Mitchell, B., Makagon, M., Jaeger, M., & Barrett, R. 2006. Information content of coyote barks and howls. *Bioacoustics*, 15(3):289. (Cited on pages 67 and 97.)

Møhl, B., Wahlberg, M., & Heerfordt, A. 2001. A large-aperture array of nonlinked receivers for acoustic positioning of biological sound sources. *The Journal of the Acoustical Society of America*, 109:434–437. (Cited on page 176.)

Moorhouse, T. & Macdonald, D. 2005. Indirect negative impacts of radio-collaring: sex ratio variation in water voles. *Ecology*, 42:91–98. (Cited on page 177.)

Morgan, S. & Raspet, R. 1992. Investigation of the mechanisms of low-frequency wind noise generation outdoors. *The Journal of the Acoustical Society of America*, 92:1180. (Cited on page 52.)

Morton, E. 1975. Ecological sources of selection on avian sounds. *The American Naturalist*, 109(965):17–34. (Cited on pages 18 and 50.)

Morton, E. 1977. On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *The American Naturalist*, 111:855–869. (Cited on pages 18 and 53.)

Mougeot, F. & Bretagnolle, V. 2000. Predation risk and moonlight avoidance in nocturnal seabirds. *Journal of Avian Biology*, 31:376–386. (Cited on pages 126, 127, and 152.)

Muanke, P. B. & Niezrecki, C. 2007. Manatee position estimation by passive acoustic localization. *The Journal of the Acoustical Society of America*, 121:2049–2059. (Cited on page 175.)

Mulard, H., Aubin, T., & White, J. 2009. Voice variance may signify ongoing divergence among black-legged kittiwake populations. *Biological Journal of the Linnean Society*, 97:289–297. (Cited on page 94.)

Naguib, M. & Wiley, R. H. 2001. Estimating the distance to a source of sound: mechanisms and adaptations for long-range communication. *Animal Behaviour*, 62(5):825–837. (Cited on pages 36 and 50.)

Nagy, C. M. & Rockwell, R. F. 2012. Identification of individual eastern screech-owls *Megascops asio* via vocalization analysis. *Bioacoustics*, 21(2):127–140. (Cited on page 26.)

Nicholls, J. & Goldizen, A. 2006. Habitat type and density influence vocal signal design in satin bowerbirds. *Journal of Animal Ecology*, 75:549–558. (Cited on page 33.)

Nowicki, S. & Capranica, R. R. 1986. Bilateral syringeal coupling during phonation of a songbird. *The Journal of Neuroscience*, 6(12):3595–3610. (Cited on pages 66 and 78.)

O'Donnell, C. 2011. Breeding of the Australasian bittern (*Botaurus poiciloptilus*) in New Zealand. *Emu*, 111:197–201. (Cited on page 95.)

Otter, K. 1996. Individual variation in the advertising call of male northern saw-whet owls. *Journal of Field Ornithology*, 67(3):398–405. (Cited on pages 17 and 94.)

Parker, K., Hauber, M., & Brunton, D. 2010. Contemporary cultural evolution of a conspecific recognition signal following serial translocations. *Evolution*, 64(8):2431–2441. (Cited on pages 19 and 27.)

Parker, T. A. 1991. On the use of tape recorders in avifaunal surveys. *The Auk*, 108(2):443–444. (Cited on pages 26 and 157.)

Patricelli, G. L. & Krakauer, A. H. 2010. Tactical allocation of effort among multiple signals in sage grouse: an experiment with a robotic female. *Behavioral Ecology*, 21(1):97–106. (Cited on page 176.)

Payne, K. & Thompson, M. 2003. Elephant calling patterns as indicators of group size and composition: the basis for an acoustic monitoring system. *African Journal of Ecology*, 41:99–107. (Cited on pages 26, 27, and 175.)

Peake, T., McGregor, P. K., Smith, K., & Tyler, G. 1998. Individuality in corncrake (*Crex crex*) vocalizations. *Ibis*, 140:120–127. (Cited on page 17.)

Peake, T. M. & McGregor, P. K. 2001. Corncrake (*Crex crex*) census estimates: a conservation application of vocal individuality. *Animal Biodiversity and Conservation*, 24:81–90. (Cited on pages 26 and 27.)

Penteriani, V., Delgado, M., & Campioni, L. 2010. Moonlight makes owls more chatty. *PLoS ONE*, 5(1). (Cited on page 127.)

Penteriani, V., Kuparinen, A., & Delgado, M. 2011. Individual status, foraging effort and need for conspicuousness shape behavioural responses of a predator to moon phases. *Animal Behaviour*, 82:413–420. (Cited on pages 126 and 127.)

Peterson, C. & Dorcas, M. 1994. Automated data acquistition. In Heyer, W. R., McDiarmid, R. W., Donnelly, M., & Hayek, L., editors, *Measuring and monitoring biological diversity - standard methods for amphibians*. Smithsonian Institution Press, Washington, D.C. (Cited on page 26.)

Petrusková, T., Osiejuk, T., Linhart, P., & Petrusek, A. 2008. Structure and complexity of perched and flight songs of the tree pipit (*Anthus trivialis*). *Annales Zoologici Fennici*, 45(2):135–148. (Cited on page 17.)

Pfaff, J. A., Zanette, L., MacDougall-Shackleton, S. A., & MacDougall-Shackleton, E. A. 2007. Song repertoire size varies with HVC volume and is indicative of male quality in song sparrows (*Melospiza melodia*). *Proceedings of the Royal Society B: Biological Sciences*, 274(1621):2035–2040. (Cited on page 123.)

Pierce, R. & Westbrooke, I. 2003. Call count responses of North Island brown kiwi to different levels of predator control in Northland, New Zealand. *Biological Conservation*, 109(2):175–180. (Cited on pages 23, 27, 28, 128, 129, 136, 150, 152, 158, 159, 170, and 214.)

Pierotti, R. 1991. Infanticide versus adoption: an intergenerational conflict. *The American Naturalist*, 138(5):1140–1158. (Cited on pages 17 and 93.)

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team 2012. nlme: linear and nonlinear mixed effects models. Version 3.1-104. http://cran.r-project.org/web/packages/nlme/index.html. Downloaded 03 Aug 2013. (Cited on page 41.)

Podos, J. 1997. A performance constraint on the evolution of trilled vocalizations in a songbird family (Passeriformes: Emberizidae). *Evolution*, 51(2):537–551. (Cited on page 51.)

Podos, J. 2001. Correlated evolution of morphology and vocal signal structure in Darwin's finches. *Nature*, 409(6817):185–188. (Cited on page 19.)

Podos, J., Huber, S., & Taft, B. 2004. Bird song: the interface of evolution and mechanism. *Annual Review of Ecology, Evolution, and Systematics*, pages 55–87. (Cited on pages 17, 18, and 19.)

Potts, T. H. 1872. On the birds of New Zealand. *Transaction & Proceedings of the New Zealand Institute*, 5:171–205. (Cited on page 22.)

R Development Core Team 2011. R: A language and environment for statistical computing. Version 2.15. R Foundation for Statistical Computing, Vienna. http://www.R-project.org. Downloaded 02 Apr 2012. (Cited on pages 38, 71, 99, 136, and 164.)

Ramos, J. 2003. Intraspecific aggression by roseate tern adults on chicks in a tropical colony. *Waterbirds*, 26(2):160–165. (Cited on pages 17 and 94.)

Ramstad, K., Pfunder, M., Robertson, H. A., Colbourne, R., Allendorf, F., & Daugherty, C. H. 2010. Fourteen microsatellite loci cross-amplify in all five kiwi species (*Apteryx* spp.) and reveal extremely low genetic variation in little spotted kiwi (*A. owenii*). *Conservation Genetics Resources*, 2:333–336. (Cited on pages 51, 55, 97, and 123.)

Ramstad, K. M., Colbourne, R. M., Robertson, H. A., Allendorf, F. W., & Daugherty, C. H. 2013. Genetic consequences of a century of protection: serial founder

events and survival of the little spotted kiwi (*Apteryx owenii*). *Proceedings of the Royal Society B: Biological Sciences*, 280(1762):20130576. (Cited on pages 20, 97, and 123.)

Reby, D., McComb, K., Cargnelutti, B., Darwin, C., Fitch, W. T., & Clutton-Brock, T. 2005. Red deer stags use formants as assessment cues during intrasexual agonistic interactions. *Proceedings of the Royal Society B: Biological Sciences*, 272(1566):941–947. (Cited on page 79.)

Redpath, S. M. 1994. Censusing tawny owls *Strix aluco* by the use of imitation calls. *Bird Study*, 41(3):192–198. (Cited on page 94.)

Redpath, S. M., Appleby, B. M., & Petty, S. 2000. Do male hoots betray parasite loads in tawny owls? *Journal of Avian Biology*, 31(4):457–462. (Cited on page 26.)

Reid, B. 1971. The weight of the kiwi and its egg. *Notornis*, 18(4):245–249. (Cited on page 52.)

Rempel, R., Hobson, K., Holborn, G., van Wilgenburg, S., & Elliott, J. 2005. Bioacoustic monitoring of forest songbirds: interpreter variability and effects of configuration and digital processing methods in the laboratory. *Journal of Field Ornithology*, 76:1–11. (Cited on page 26.)

Rendall, D., Kollias, S., Ney, C., & Lloyd, P. 2005. Pitch (F_0) and formant profiles of human vowels and vowel-like baboon grunts: The role of vocalizer body size and voice-acoustic allometry. *The Journal of the Acoustical Society of America*, 117(2):944–955. (Cited on page 51.)

Riede, K. 1993. Monitoring biodiversity: analysis of Amazonian rainforest sounds. *Ambio*, 22(8):546–548. (Cited on pages 26, 27, and 157.)

Riede, K. 1998. Acoustic monitoring of Orthoptera and its potential for conservation. *Journal of Insect Conservation*, 2(3):217–223. (Cited on pages 26, 27, and 157.)

Riede, T. 2000. *Vocal changes in animals during disorders*. PhD thesis, Humboldt-Universität, Berlin, Germany. (Cited on pages 67, 79, and 90.)

Riede, T., Arcadi, A. C., & Owren, M. J. 2007. Nonlinear acoustics in the pant hoots of common chimpanzees (*Pan troglodytes*): vocalizing at the edge. *The Journal of the Acoustical Society of America*, 121(3):1758–1767. (Cited on pages 54, 67, 78, 79, and 90.)

Riede, T., Fisher, J. H., & Goller, F. 2010. Sexual dimorphism of the zebra finch syrinx indicates adaptation for high fundamental frequencies in males. *PLoS ONE*, 5(6):e11368. (Cited on page 51.)

Riede, T., Owren, M. J., & Arcadi, A. C. 2004. Nonlinear acoustics in pant hoots of common chimpanzees (*Pan troglodytes*): frequency jumps, subharmonics, biphonation, and deterministic chaos. *American Journal of Primatology*, 64(3):277–291. (Cited on page 54.)

Riede, T., Suthers, R. A., Fletcher, N. H., & Blevins, W. E. 2006. Songbirds tune their vocal tract to the fundamental frequency of their song. *Proceedings of the National Academy of Sciences of the United States of America*, 103(14):5543–5548. (Cited on page 51.)

Ritchie, M. 2007. Sexual selection and speciation. *Annual Review of Ecology, Evolution, and Systematics*, 38:79–102. (Cited on page 19.)

Robbins, M. & Stiles, F. 1999. A new species of pygmy-owl (Strigidae: *Glaucidium*) from the Pacific slope of the northern Andes. *The Auk*, 116(2):305–315. (Cited on page 27.)

Robertson, H. & Colbourne, R. M. 2001. Survival of little spotted kiwi exposed to the rodenticide brodifacoum. *The Journal of Wildlife Management*, 65(1):29–34. (Cited on page 27.)

Robertson, H. & Colbourne, R. M. 2003. Kiwi (*Apteryx* spp.) best practice manual. Technical report, Department of Conservation, Wellington, New Zealand. (Cited on pages 28, 36, 38, 70, 128, 131, 147, 158, 159, 170, and 177.)

Robertson, H., Colbourne, R. M., & Graham, P. 1999. Survival of brown kiwi (*Apteryx mantelli*) exposed to brodifacoum poison in Northland, New Zealand. *New Zealand Journal of Ecology*, 23(2):225–231. (Cited on page 27.)

Robertson, H. A. & Colbourne, R. M. 2004. Survival of little spotted kiwi (*Apteryx owenii*) on Kapiti Island. *Notornis*, 51:161–163. (Cited on pages 98 and 158.)

Robertson, H. A., Colbourne, R. M., & Nieuwland, F. 1993. Survival of little spotted kiwi and other forest birds exposed to brodifacoum rat poison on Red Mercury Island. *Notornis*, 40(4):253–262. (Cited on pages 128 and 158.)

Robertson, H. A. & Fraser, J. R. 2009. Use of trained dogs to determine the age structure and conservation status of kiwi *Apteryx* spp. populations. *Bird Conservation International*, 19(02):121–129. (Cited on pages 52 and 69.)

Robertson, H. A., McLennan, J., Colbourne, R. M., & McCann, T. J. 2005. Population status of great spotted kiwi (*Apteryx haastii*) near Saxon Hut, Heaphy Track, New Zealand. *Notornis*, 52:27–33. (Cited on page 158.)

Robin, V. V., Katti, M., Purushotham, C., Sancheti, A., & Sinha, A. 2011. Singing in the sky: song variation in an endemic bird on the sky islands of southern India. *Animal Behaviour*, 82(3):513–520. (Cited on pages 27, 95, 97, and 123.)

Robisson, P. 1992. Vocalizations in *Aptenodytes* penguins: application of the two-voice theory. *The Auk*, 109(3):654–658. (Cited on pages 122 and 211.)

Robisson, P., Aubin, T., & Bremond, J. 1993. Individuality in the voice of the emperor penguin *Aptenodytes forsteri*: adapation to a noisy environment. *Ethology*, 94:279–279. (Cited on pages 17, 65, and 94.)

Rodríguez, G. & Goldman, N. 2001. Improved estimation procedures for multilevel models with binary response: a case-study. *Journal of the Royal Statistical Society: Series A*, 164(2):339–355. (Cited on page 71.)

Rosenstock, S. S., Anderson, D. R., Giesen, K. M., Leukering, T., Carter, M. F., & Thompson III, F. 2002. Landbird counting techniques: current practices and an alternative. *The Auk*, 119(1):46–53. (Cited on pages 94, 127, and 156.)

Rowe, S. J. & Bell, B. D. 2007. The influence of geographic variation in song dialect on post-translocation pair formation in North Island kokako (*Callaeas cinerea wilsoni*). *Notornis*, 54(1):28–37. (Cited on page 27.)

Ryan, M. J. & Brenowitz, E. A. 1985. The role of body size, phylogeny, and ambient noise in the evolution of bird song. *The American Naturalist*, 126(1):88–100. (Cited on pages 33 and 52.)

Saunders, D. & Wooller, R. 1988. Consistent individuality of voice in birds as a management tool. *Emu*, 88:25–32. (Cited on pages 27, 95, and 96.)

Schneider, J. N. & Anderson, R. E. 2011. Tonal vocalizations in the red wolf (*Canis rufus*): Potential functions of nonlinear sound production. *The Journal of the Acoustical Society of America*, 130(4):2275–2284. (Cited on pages 64 and 67.)

Scrimgeour, J. & Pickett, A. J. 2011. Taxon plan for western brown kiwi (*Apteryx mantelli*). Technical report, Department of Conservation, Wellington, New Zealand. (Cited on page 23.)

Searcy, W. & Andersson, M. 1986. Sexual selection and the evolution of song. *Annual Review of Ecology, Evolution, and Systematics*, 17(1):507–533. (Cited on page 19.)

Seddon, N., Amos, W., Mulder, R. A., & Tobias, J. A. 2004. Male heterozygosity predicts territory size, song structure and reproductive success in a cooperatively breeding bird. *Proceedings of the Royal Society B: Biological Sciences*, 271(1550):1823–1829. (Cited on page 123.)

Seddon, N. & Sorenson, M. 2005. Ecological adaptation and species recognition drives vocal evolution in Neotropical suboscine birds. *Evolution*, 59(1):200–215. (Cited on pages 19, 20, 24, 33, and 51.)

Seddon, N. & Tobias, J. 2007. Song divergence at the edge of Amazonia: an empirical test of the peripatric speciation model. *Biological Journal of the Linnean Society*, 90(1):173–188. (Cited on pages 20 and 27.)

Shepherd, L. D., Worthy, T. H., Tennyson, A. J. D., Scofield, R. P, Ramstad, K. M., & Lambert, D. M. 2012. Ancient DNA analyses reveal contrasting phylogeographic patterns amongst kiwi (*Apteryx* spp.) and a recently extinct lineage of spotted kiwi. *PLoS ONE*, 7(8):e42384. (Cited on page 20.)

Sim, J. & Wright, C. C. 2005. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical Therapy*, 85:257–268. (Cited on page 103.)

Slabbekoorn, H. 2004. Singing in the wild: the ecology of birdsong. In Marler, P. & Slabbekoorn, H., editors, *Nature's Music: the Science of Birdsong*, pages 178–205. Elsevier Academic Press, London, UK. (Cited on page 52.)

Slabbekoorn, H. & den Boer-Visser, A. 2006. Cities change the songs of birds. *Current Biology*, 16:2326–2331. (Cited on page 33.)

Slabbekoorn, H., Ellers, J., & Smith, T. 2002. Birdsong and sound transmission: the benefits of reverberations. *The Condor*, 104(3):564–573. (Cited on page 50.)

Slabbekoorn, H. & Peet, M. 2003. Birds sing at a higher pitch in urban noise. *Nature*, 424(6946):267. (Cited on pages 19 and 33.)

Slabbekoorn, H. & Smith, T. 2002. Bird song, ecology and speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 357:493–503. (Cited on page 19.)

Slater, P. & Mann, N. 2004. Why do the females of many bird species sing in the tropics? *Journal of Avian Biology*. (Cited on page 25.)

Speirs, E. & Davis, L. 1991. Discrimination by Adelie penguins *Pygoscelis adeliae*, between the loud mutual calls of mates, neighbors and strangers. *Animal Behaviour*, 41:937–944. (Cited on page 17.)

Spiesberger, J. 1999. Locating animals from their sounds and tomography of the atmosphere: experimental demonstration. *The Journal of the Acoustical Society of America*, 106:837. (Cited on pages 176 and 177.)

Stafford, K. M., Fox, C. G., & Clark, D. S. 1998. Long-range acoustic detection and localization of blue whale calls in the northeast Pacific Ocean. *The Journal of the Acoustical Society of America*, 104(6):3616–3625. (Cited on page 175.)

Stoddard, P. & Beecher, M. 1983. Parental recognition of offspring in the cliff swallow. *The Auk*, 100:795–799. (Cited on pages 17 and 94.)

Suthers, R. 1990. Contributions to birdsong from the left and right sides of the intact syrinx. *Nature*, 347:473–477. (Cited on page 65.)

Suthers, R. 1994. Variable asymmetry and resonance in the avian vocal tract: a structural basis for individually distinct vocalizations. *Journal of Comparative Physiology A*, 175:457–466. (Cited on pages 24, 54, 65, 94, 95, and 96.)

Suthers, R. & Hector, D. 1985. The physiology of vocalization by the echolocating oilbird, *Steatornis caripensis*. *Journal of Comparative Physiology A*, 156:243–266. (Cited on page 94.)

Suthers, R. A., Narins, P. M., Lin, W.-Y., Schnitzler, H.-U., Denzinger, A., Xu, C.-H., & Feng, A. S. 2006. Voices of the dead: complex nonlinear vocal signals from the larynx of an ultrasonic frog. *Journal of Experimental Biology*, 209(24):4984–4993. (Cited on page 79.)

Suthers, R. A., Wild, J. M., & Kaplan, G. 2010. Mechanisms of song production in the Australian magpie. *Journal of Comparative Physiology A*, 197(1):45–59. (Cited on pages 65, 66, 67, 78, and 97.)

Swiston, K. A. & Mennill, D. J. 2009. Comparison of manual and automated methods for identifying target sounds in audio recordings of pileated, pale-billed, and putative ivory-billed woodpeckers. *Journal of Field Ornithology*, 80(1):42–50. (Cited on pages 26, 27, and 158.)

Taborsky, B. & Taborsky, M. 1991. Social organization of North Island brown kiwi: long-term pairs and three types of male spacing behaviour. *Ethology*, 89(1):47–62. (Cited on pages 24, 52, and 149.)

Taborsky, B. & Taborsky, M. 1999. The mating system and stability of pairs in kiwi (*Apteryx* spp.). *Journal of Avian Biology*, 30(2):143–151. (Cited on pages 22 and 52.)

Taborsky, M. & Taborsky, B. 1992. Spatial organization of the North Island brown kiwi (*Apteryx australis mantelli*): sex, pairing status and territorality. *Ibis*, 134(1):1–10. (Cited on pages 22, 23, 24, and 52.)

Taylor, E. L., Blache, D., Groth, D., Wetherall, J. D., & Martin, G. B. 2000. Genetic evidence for mixed parentage in nests of the emu (*Dromaius novaehollandiae*). *Behavioral Ecology and Sociobiology*, 47(5):359–364. (Cited on page 24.)

ten Cate, C. 2004. Birdsong and evolution. In Marler, P. & Slabbekoorn, H., editors, *Nature's Music: the science of birdsong*. Elsevier Academic Press, London. (Cited on pages 18, 19, and 20.)

Terry, A., Peake, T., & McGregor, P. K. 2005. The role of vocal individuality in conservation. *Frontiers in Zoology*, 2(1):10. (Cited on pages 27, 95, 96, and 214.)

Terry, A. M. R. & McGregor, P. K. 2002. Census and monitoring based on individually identifiable vocalizations: the role of neural networks. *Animal Conservation*, 5:103. (Cited on pages 27 and 95.)

Terry, A. M. R., McGregor, P. K., & Peake, T. 2001. A Comparison of some techniques used to assess vocal individuality. *Bioacoustics*, 11:169–188. (Cited on page 95.)

Thompson, W. L. 2002. Towards reliable bird surveys: accounting for individuals present but not detected. *The Auk*, 119(1):18–25. (Cited on page 127.)

Tibbetts, E. A. & Dale, J. 2007. Individual recognition: it is good to be different. *Trends in Ecology & Evolution*, 22(10):529–537. (Cited on page 94.)

Tobias, J. A., Aben, J., Brumfield, R. T., Derryberry, E. P., Halfwerk, W., Slabbekoorn, H., & Seddon, N. 2010. Song divergence by sensory drive in Amazonian birds. *Evolution*, 64(10):2820–2839. (Cited on pages 19 and 20.)

Tooze, Z. J., Harrington, F. H., & Fentress, J. C. 1990. Individually distinct vocalizations in timber wolves, *Canis lupus*. *Animal Behaviour*, 40(4):723–730. (Cited on page 97.)

Townsend, S. W. & Manser, M. B. 2011. The function of nonlinear phenomena in meerkat alarm calls. *Biology Letters*, 7(1):47–49. (Cited on pages 64, 67, and 90.)

Towsey, M., Planitz, B., Nantes, A., Wimmer, J., & Roe, P. 2012. A toolbox for animal call recognition. *Bioacoustics*, 21(2):107–125. (Cited on pages 132, 158, and 161.)

Venables, W. N. & Ripley, B. D. 2002. *Modern applied statistics with S*. Springer Verlag, New York, USA, 4th edition. (Cited on pages 71 and 102.)

Volodin, I., Kaiser, M., Matrosova, V., Volodina, E., Klenova, A., Filatova, O., & Kholodova, M. 2009. The technique of noninvasive distant sexing for four monomorphic *Dendrocygna* whistling duck species by their loud whistles. *Bioacoustics*, 18:277–290. (Cited on pages 27 and 33.)

Volodin, I., Volodina, E., & Klenova, A. 2005. Individual and sexual differences in the calls of the monomorphic white-faced whistling duck *Dendrocygna viduata*. *Acta Ornithologica*, 40(1):43–52. (Cited on page 27.)

Volodina, E. V., Volodin, I. A., Isaeva, I. V., & Unck, C. 2006. Biphonation may function to enhance individual recognition in the dhole, *Cuon alpinus*. *Ethology*, 112:815–825. (Cited on pages 67 and 97.)

Waddle, J., Thigpen, T. F., & Glorioso, B. M. 2009. Efficacy of automatic vocalization recognition software for anuran monitoring. *Herpetological Conservation and Biology*, 4(3):384–388. (Cited on page 158.)

Wahlberg, M. & Møhl, B. 2003. Localising bitterns (*Botaurus stellaris*) with an array of non-linked microphones. *Bioacoustics*, 13(3):233–245. (Cited on pages 26, 175, and 176.)

Wahlberg, M., Møhl, B., & Madsen, P. T. 2001. Estimating source position accuracy of a large-aperture hydrophone array for bioacoustics. *The Journal of the Acoustical Society of America*, 109:397–406. (Cited on pages 175, 176, and 183.)

Wallschläger, D. 1980. Correlation of song frequency and body weight in passerine birds. *Cellular and Molecular Life Sciences*, 36(4):412–412. (Cited on page 33.)

Weary, D., Lemon, R., & Perreault, S. 1992. Song repertoires do not hinder neighbor-stranger discrimination. *Behavioral Ecology and Sociobiology*, 31:441–447. (Cited on page 96.)

Weary, D., Norris, K., & Falls, J. 1990. Song features birds use to identify individuals. *The Auk*, 107(3):623–625. (Cited on pages 96 and 122.)

Weisman, R., Ratcliffe, L., Johnsrude, I., & Hurly, T. 1990. Absolute and relative pitch production in the song of the black-capped chickadee. *Condor*, 92:118–124. (Cited on page 95.)

Whitney, B. M., Pacheco, J. F., Buzzetti, D. R. C., & Parrini, R. 2000. Systematic revision and biogeography of the *Herpsilochmus pileatus* complex, with description of a new species from northeastern Brazil. *The Auk*, 117(4):869–891. (Cited on page 19.)

Wilden, I., Herzel, H., Peters, G., & Tembrock, G. 1998. Subharmonics, biphonation, and deterministic chaos in mammal vocalization. *Bioacoustics*, 8:1–30. (Cited on pages 54, 64, 66, 67, 79, 86, and 90.)

Wiley, R., Hatchwell, B., & Davies, N. 1991. Recognition of individual males' songs by female dunnocks: a mechanism increasing the number of copulatory partners and reproductive success. *Ethology*, 88(2):145–153. (Cited on page 94.)

Wiley, R. & Richards, D. 1978. Physical constraints on acoustic communication in the atmosphere: implications for the evolution of animal vocalizations. *Behavioral Ecology and Sociobiology*, 3:69–94. (Cited on page 52.)

Wilson, D. R. & Mennill, D. J. 2010. Black-capped chickadees, *Poecile atricapillus*, can use individually distinctive songs to discriminate among conspecifics. *Animal Behaviour*, 79(6):1267–1275. (Cited on page 18.)

Wood, S. N. 2003. Thin plate regression splines. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 65(1):95–114. (Cited on page 138.)

Wood, S. N. 2004. Stable and efficient multiple smoothing parameter estimation for generalized additive models. *Journal of the American Statistical Association*, 99(467):673–686. (Cited on page 138.)

Wood, S. N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statisical Methodology)*, 73(1):3–36. (Cited on pages 138 and 139.)

Worthy, T. H. 2001. A fossil vertebrate fauna accumulated by laughing owls (*Sceloglaux albifacies*) on the Gouland Downs, northwest Nelson, South Island. *Notornis*, 48(4):225–233. (Cited on pages 129 and 152.)

Worthy, T. H. & Holdaway, R. N. 2002. *The lost world of the moa: prehistoric life of New Zealand*. Indiana Unversity Press, Bloomington, IN, USA, 1st edition. (Cited on pages 129 and 152.)

Wright, T. F. & Dahlin, C. R. 2007. Pair duets in the yellow-naped amazon (*Amazona auropalliata*): phonology and syntax. *Behaviour*, 144(2):207–228. (Cited on page 33.)

Yeredor, A. 2012. Analysis of the edge-effects in frequency-domain tdoa estimation. In *2012 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 3521–3524, Kyoto, Japan. (Cited on page 205.)

Yorzinski, J. L., Vehrencamp, S. L., McGowan, K. J., & Clark, A. B. 2006. The inflected alarm caw of the American crow: differences in acoustic structure among individuals and sexes. *Condor*, 108(3):518. (Cited on pages 18, 94, 95, and 122.)

Zeileis, A. & Grothendieck, G. 2005. zoo: S3 infrastructure for regular and irregular time series. *Journal of Statistical Software*, 14(6):1–27. (Cited on page 137.)

Ziesemann, B., Brunton, D. H., & Castro, I. C. 2011. Nesting success and breeding ecology in a high-density population of brown kiwi (*Apteryx mantelli*). *Emu*, 111(2):148–154. (Cited on pages 22 and 52.)

Zollinger, S. A., Riede, T., & Suthers, R. A. 2008. Two-voice complexity from a single side of the syrinx in northern mockingbird *Mimus polyglottos* vocalizations. *Journal of Experimental Biology*, 211(12):1978–1991. (Cited on pages 54, 64, 65, 66, 67, 72, 78, 79, and 95.)

Zuur, A. F., Ieno, E. N., & Smith, G. M. 2007. *Analysing ecological data*. Statistics for Biology and Health. Springer, New York, 1st edition. (Cited on pages 100 and 101.)

Zuur, A. F., Ieno, E. N., Walker, N., Saveliev, A. A., & Smith, G. M. 2009. *Mixed effects models and extensions in ecology with R*. Springer, New York, USA, 1st edition. (Cited on pages 41, 139, 140, and 164.)

Zuur, A. F., Saveliev, A. A., & Ieno, E. N. 2012. *Zero inflated models and generalized linear mixed models with R*. Highland Statistics Ltd, Newburgh, United Kingdom, 1st edition. (Cited on pages 71 and 140.)