The effects of climate change on the global migration of birds

O’MAHONY, NAIARA, KAREN

How to cite:
O’MAHONY, NAIARA, KAREN (2016) The effects of climate change on the global migration of birds, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/11543/

Use policy
The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full Durham E-Theses policy for further details.
The effects of climate change on the global migration of birds

Willow warbler, Phylloscopus trochilus, by Ian Fisher

By Naiara O’Mahony

School of Biological and Biomedical Sciences
Durham University
November 2015

Submitted for the degree of Doctor of Philosophy
Declaration

The material contained within this thesis has not previously been submitted for a degree at Durham University or any other university. The research reported within this thesis has been conducted by the author unless indicated otherwise.

Naiara O’Mahony

November 2015

© The copyright of this thesis rests with the author. No quotation from it should be published without the author’s prior written consent and information derived from it should be acknowledged.
Acknowledgements

I would like to thank my supervisors, Dr. Steve Willis and Dr. Phil Stephens, for their guidance, insight, patience and continued support throughout this project. I also owe particular thanks to Dr. David Baker, for all his help and encouragement. Thanks also to the Durham Conservation Ecology group, for support, constructive criticism and understanding throughout my PhD.

I owe thanks to Dr. James Pearce-Higgins for his enthusiasm and helpful comments. I would like to thank Dr. Robert Bagchi for his help and advice and especially for his guidance on statistical queries. I am grateful to BirdLife International and NatureServe, and all the volunteers who collected and compiled the data on the distributions of bird species. I am grateful for the Durham University Doctoral Scholarship which funded my PhD.

I thank the Northumbria ringing group (especially my trainer, Mr. Richard Barnes) for their encouragement and for inspiring me to learn more about birds. A huge thanks to Mr. Ian Fisher from the Northumbria ringing group, who kindly provided all photographs of birds provided in this thesis.

Finally, thanks to my friends for their unyielding support and patience, particularly those who have gone through this process with me at Durham (you know who you are!). Last but not least, to my loving family (Arantza, Cyril, Sinead, Niall, Wanda and my boys), as well as Jamie, Georgina and Rachel; I could not have finished this thesis without you.
Abstract

Migration is one of the most spectacular wonders of the natural world but is increasingly threatened by large-scale environmental change. Migrants are key components of biological systems at higher latitudes, where the magnitude and velocity of climate change is most pronounced. In this thesis, I consider how the distribution of migratory species may change in the future, across the globe, in response to environmental change.

Using global data on avian species distributions, the environmental and biological factors that affect migratory species richness were investigated. The cost of migration distance between breeding and non-breeding areas, and resident species richness were two important determinants of migratory species richness in a given location. Species distribution models (SDMs) were used to relate migratory species occurrence in breeding and non-breeding seasons to contemporary climatic variables. The importance of climatic variables for predicting migrant species’ occupancy depended on season (breeding or non-breeding). Whilst most SDMs focus on individual species breeding ranges, and model within the occupied realm, this can be problematic for migratory species that frequently use multiple realms, especially if trying to project newly suitable areas in future. Here, I developed a novel method to assess climatic suitability for migratory species globally, within species migration flyways. Selecting absence data to condition SDMs for species occurring across multiple realms can be problematic, as regions distant from the range may be suitable but unoccupied. To minimise this issue, I developed a modelling approach that selected absences preferentially from closer to a species distribution, using a distance weighting function; compared to other methods, this improved model performance on withheld test data. Using this approach, I projected the potential future breeding and non-breeding ranges of all global terrestrial long-distance migrants under climate change. Migrant breeding ranges were projected to shift poleward, but non-breeding range shift projections are less consistent in direction.

To date, predictions of migratory changes under future climate scenarios have usually estimated a single distance between breeding and non-breeding range centroids. However, this approach ignores the variation in migratory movements within species. I developed a method to estimate the range of potential migration distances for individual populations, both now and in the future, which agreed well with recovery data for a well-recorded European trans-Saharan migrant. This approach projected longer median migratory distances in the future for many species. This thesis highlights the importance of climate for migratory species, and suggests that the observed general decline in many long-distance migratory species may be exacerbated by ongoing climate change.
Table of Contents

Declaration .................................................................................................................. i
Acknowledgements ..................................................................................................... ii
Abstract ..................................................................................................................... iii
Table of Contents ....................................................................................................... iv

1 General Introduction ............................................................................................... 1
   1.1 Introduction ........................................................................................................ 2
   1.2 Climate change ................................................................................................... 2
   1.3 Species responses to climate change ................................................................. 3
      1.3.2 Adaptation and phenotypic plasticity ......................................................... 3
      1.3.3 Range shifts ............................................................................................... 3
      1.3.4 Biotic interactions ..................................................................................... 4
   1.4 The value of presence-absence data ................................................................. 5
   1.5 The usefulness of large scale climate data ...................................................... 5
   1.6 Species distribution models ............................................................................ 7
   1.7 The study of birds ............................................................................................ 8
   1.8 Migration ........................................................................................................... 8
   1.9 The effects of climate change on avian migratory species ............................. 10
      1.9.1 Timing of migration ................................................................................. 10
      1.9.2 Range shifts and migration distance ....................................................... 11
   1.10 Current limitations in understanding migrant bird responses to climate change .... 13
   1.11 Project aims .................................................................................................... 14
      1. The determinants of species richness across the globe .................................. 14
      2. How to select absence data for species distribution models ..................... 14
      3. What climatic factors are important for breeding/non-breeding migrants across the globe .. 14
      4. How is migration likely to change given future climate change projections .... 14

2. The determinants of migrant species richness across the globe .......................... 15
   2.1 Abstract ............................................................................................................ 16
   2.2 Introduction ...................................................................................................... 17
   2.3 Methods .......................................................................................................... 20
2.3.1 Species distribution data ................................................................. 20
2.3.2 Bioclimatic variables ................................................................. 20
2.3.3 Productivity and seasonality ............................................................. 21
2.3.4 Habitat diversity .............................................................................. 22
2.3.5 Distance .......................................................................................... 22
2.3.6 Modeling approach ......................................................................... 22
2.4 Results ............................................................................................... 25
2.5 Discussion .......................................................................................... 32
2.5.1 Global patterns of migrant species richness ........................................... 32
2.5.2 Potential drivers of migrant species richness .......................................... 32

3. How to select absence data for species distribution models ......................... 36

3.1 Abstract ............................................................................................ 37
3.2 Introduction ........................................................................................ 38
3.3 Methods ............................................................................................. 41
3.3.1 Creating virtual species distributions .................................................... 41
3.3.2 Absence selection ............................................................................. 42
3.3.3 Species Distribution Models ............................................................... 43
3.3.4 Absence selection method comparison ................................................ 45
3.3.5 Prevalence ........................................................................................ 46
3.4 Results ............................................................................................... 47
3.4.1 Comparison of absence selection methods ............................................ 47
3.4.2 Identifying the underlying climatic suitability ........................................ 47
3.4.3 Prevalence ........................................................................................ 51
3.5 Discussion .......................................................................................... 52
3.5.1 Is it necessary to define a study extent? .................................................. 52
3.5.2 Choosing an absence selection method ................................................ 53
3.5.3 Virtual species and the assessment of model performance ...................... 53
3.5.4 Conclusion ....................................................................................... 54

4. Seasonal niches in migratory birds: a biogeographic perspective .................... 55

4.1 Abstract ............................................................................................ 56
4.2 Introduction ........................................................................................ 57
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.3.6</td>
<td>Species richness</td>
<td>100</td>
</tr>
<tr>
<td>6.3.7</td>
<td>Species turnover</td>
<td>101</td>
</tr>
<tr>
<td>6.3.8</td>
<td>Migration distance</td>
<td>101</td>
</tr>
<tr>
<td>6.3.9</td>
<td>Change in migration distance</td>
<td>102</td>
</tr>
<tr>
<td>6.4</td>
<td>Results</td>
<td>104</td>
</tr>
<tr>
<td>6.4.1</td>
<td>Change in climatic suitability</td>
<td>104</td>
</tr>
<tr>
<td>6.4.2</td>
<td>Change in species richness</td>
<td>107</td>
</tr>
<tr>
<td>6.4.3</td>
<td>Change in species turnover</td>
<td>110</td>
</tr>
<tr>
<td>6.4.4</td>
<td>Change in migration distance</td>
<td>112</td>
</tr>
<tr>
<td>6.5</td>
<td>Discussion</td>
<td>115</td>
</tr>
<tr>
<td>6.5.1</td>
<td>Climatic suitability</td>
<td>115</td>
</tr>
<tr>
<td>6.5.2</td>
<td>Species richness and turnover</td>
<td>116</td>
</tr>
<tr>
<td>6.5.3</td>
<td>Migration distance</td>
<td>117</td>
</tr>
<tr>
<td>6.5.4</td>
<td>Conclusion</td>
<td>119</td>
</tr>
<tr>
<td>7.1</td>
<td>Synthesis</td>
<td>121</td>
</tr>
<tr>
<td>7.2</td>
<td>Thesis discussion</td>
<td>121</td>
</tr>
<tr>
<td>7.3</td>
<td>Assumptions and limitations</td>
<td>126</td>
</tr>
<tr>
<td>7.4</td>
<td>Future directions of further research</td>
<td>127</td>
</tr>
<tr>
<td>7.5</td>
<td>Conclusion</td>
<td>130</td>
</tr>
</tbody>
</table>

References: 131
Appendices: 155
Chapter 1

General Introduction

Redstart, *Phoenicurus phoenicurus*, by Ian Fisher
1.1 Introduction

In this thesis, I focus on migratory bird distributions across the globe, on the breeding and non-breeding seasons, and how they are related to climate. This focus is warranted, because migrants, which are key components of biological systems at higher latitudes, where the magnitude and velocity of climate change is most pronounced, are in decline worldwide. In order to manage populations of these important species effectively, there is a need to quantify how the distribution of migratory species may change in the future and to understand the potential consequences of these changes for biodiversity. To elaborate this focus, it is necessary to provide some brief background on several key issues. Here, I begin by exploring climate change and its impact on biological systems across a range of species, in particular range shifts. I then explain the utility of species distribution models for assessing the impacts of climate change on species. In light of that, I go on to explore the likely effects of climate change on migratory species.

1.2 Climate change

Global mean surface temperature has increased by 0.78°C (0.72 to 0.85°C) since the late 19th Century, a change categorised as “certain” in the 5th assessment report by the International Panel on Climate change (IPCC 2013). Oceans are warming, and sea levels rose by 0.19m (0.17 to 0.21m) between 1901 and 2010; furthermore, polar ice sheets have decreased in mass in recent decades (average annual sea ice extent has decreased by 2.7 ± 0.6% per decade, IPCC 2013). Change in precipitation patterns and an increase in the frequency of occurrence of extreme events are evident in recent decades (IPCC 2013). Moreover, climate change is spatially heterogeneous, with some locations, such as the Arctic, experiencing changes much larger than the global mean (IPCC 2007).

Anthropogenic climate change and other ongoing human-induced threats, such as habitat loss, present major threats to global biodiversity (Walther et al. 2002; Thomas et al. 2004; Brook, Sodhi & Bradshaw 2008; Pereira et al. 2010). Climate change is not occurring in isolation, and threats to biodiversity might intensify as climate change impacts interact with other factors, such as land-use change, in the future (Sala et al. 2000; Jetz, Wilcove & Dobson 2007; Mantyka-Pringle, Martin & Rhodes 2012; Mantyka-Pringle et al. 2015). Climate change is likely to impact at the genetic, species, community and ecosystem levels (Thomas, Franco & Hill 2006; Foden et al. 2013; Pacifici et al. 2015). Understanding species
responses to climate change is one of the most pressing challenges facing scientists today (Garcia et al. 2014).

1.3 Species responses to climate change

Evidence is accumulating of the impacts of changes in climate on ecological systems. An array of profound effects, across a broad range of organisms, have already been observed. Examples include changes in phenology (e.g. Menzel et al. 2006), range shifts (e.g. Root et al. 2003) and biotic interactions (e.g. Ockendon et al. 2014). I will now elaborate on each of these in turn.

1.3.2 Adaptation and phenotypic plasticity

Rapid environmental changes are occurring, and organisms must endure, acclimatise to, or escape from, adverse conditions (such as rising global temperatures) (Valladares, Gianoli & Gomez 2007). Populations can track environmental change through microevolution, a change in genotypes across generations in response to selection on a trait (Reale et al. 2003; Nussey et al. 2005). Populations that are genetically adapted to local climatic conditions should be more vulnerable to climate change than those comprising phenologically plastic populations (Chevin, Lande & Mace 2010; Phillimore et al. 2010). Phenology is the seasonal timing of life history events of plants and animals (Walther et al. 2002). Phenology, including the timing of migration, flowering, budburst, breeding and egg-laying of species has advanced in many species, distributed widely across taxonomic groups, and these shifts are recognised to be a response to climate change (Crick & Sparks 1999; Post et al. 2001; Visser & Both 2005; Menzel et al. 2006). Observed shifts in phenology can arise due to phenotypic plasticity (the ability of a genotype to exhibit a range of phenotypes (morphological, physiological and behavioural) in response to environmental variation (Fordyce 2006). The strategy chosen, 1) local adaptation, or 2) phenotypic plasticity, will determine the ability of a population to survive climate change (Jump & Penuelas 2005; Gimeno et al. 2009). The third strategy that a species can take is probably the most well studied one, which is to shift its distribution in response to climate change.

1.3.3 Range shifts

If species are able to track climatic conditions spatially, the implications of climatic change are a shift in species distributions (Barbet-Massin, Thuiller & Jiguet 2011). Studies across a wide range of taxa have detected a globally consistent trend in range shifts in the direction expected based on species physiological constraints, evidenced by poleward shifts...
in species distributions (Thomas & Lennon 1999; Parmesan & Yohe 2003; Root et al. 2003; Hickling et al. 2006; Devictor et al. 2008). A meta-analysis by Chen et al (2011) reported that species distributions have shifted to higher latitudes at a rate of 16.9 km per decade, three times faster than previously documented (Parmesan & Yohe 2003). Spatial shifts in species ranges have been the focus of recent research, and this is likely as a result of the widespread nature of these shifts, and their relevance to conservation and reserve selection (Araújo et al. 2004; Guisan et al. 2013).

Despite widespread spatial shifts, evidence suggests that species responses to climate change are currently “lagging” behind the rapid change in climate (Devictor et al. 2008; La Sorte & Jetz 2012). For species that have shifting climate niches, several factors will affect their ability to keep pace with this shift. Species-specific dispersal ability is particularly important (Schloss, Nuñez & Lawler 2012b), as is the environmental landscape through which a species must disperse (Bateman et al. 2014). Plant species may be particularly vulnerable as their sessile life-style limits their ability to keep pace with the rate of anthropogenic climate change (Neilson et al. 2005). The distributions of amphibians and reptiles (which play critical roles in food webs, often linking terrestrial and aquatic ecosystems), closely reflect rainfall and temperature patterns (Bickford et al. 2010). Amphibians have the highest threat of extinction of all terrestrial vertebrates (IUCN 2008). Although studies have shown a trend towards increased availability of climate space for amphibians and reptiles across Europe, these taxa are generally poor dispersers, and are therefore projected to lose range by 2050 (Araújo, Thuiller & Pearson 2006). Many such species will not be able to track future climate change and, without human intervention, are likely to become extinct (Devictor et al. 2008; Warren et al. 2013; Howard & Bickford 2014).

### 1.3.4 Biotic interactions

Responses to recent climate change vary greatly among species and locations. Responses include a diversity of trends in abundances, geographic range sizes, and directions of range shifts (Parmesan & Yohe 2003; Pearson & Dawson 2003; Hickling et al. 2006; Mair et al. 2012; Rapacciuolo et al. 2014). High interspecific variation in abundance trends in a group of 115 Lepidoptera species was explained by species-specific exposure and sensitivity to climate change in the past four decades (Palmer et al. 2015). If species are responding differently to climate change, then this could result in disrupted interactions between, and, indeed within, trophic levels in response to climate change (recently reviewed by Walther 2010).
Biotic interactions can strongly influence species spatial patterns, for example through predation, competition, resource-consumer interactions, and host-parasite interactions, and these interactions can influence how climate change affects species (Forchhammer et al. 2005; Suttle, Thomsen & Power 2007; Gilman et al. 2010; Wisz et al. 2012). Gilman et al. (2010) argue that failure to account for species interactions limits our ability to predict species responses to future climate change. It is essential to obtain a good understanding of species responses to climate, and develop the appropriate tools for understanding these responses, before we add the complexity of accounting for species interactions.

1.4 The value of presence-absence data

Species polygon data give geographic information about where a species is present or absent at a given location. These polygons are coarse generalisations of species distributions since they can include relatively large areas from which the species is absent, potentially overestimating the species’ true area of occupancy (Hoffmann et al. 2010; Somveille et al. 2013). Additionally, species occupancy maps for species occurring in areas such as Europe and North America are likely to be more accurate than those in other locations because of differences in investment in science and large scale biodiversity monitoring (Amano & Sutherland 2013; Somveille et al. 2013; Barnes et al. 2014, Sutherland et al. 2015). Information deficiencies are therefore likely to be greater on the non-breeding grounds than the breeding grounds (Faaborg et al. 2010). However, such limitations are not expected to significantly affect patterns at large scales and at coarse spatial resolutions (Somveille et al. 2013). Despite limitations, presence-absence data can make a considerable contribution towards ecology, such as the ability to estimate the biological effects of climate change, habitat loss and invasive species (Warren 2012). Although the aforementioned problems may limit the reliability of niche models, in many cases, they are still the best available niche estimates. The main virtue of species distribution modelling using presence-absence data is that they contain useful information, and are easily understood by policy makers and the general public (Guisan et al. 2013).

1.5 The usefulness of large scale climate data

Historic and future climate projections are freely available (e.g. WorldClim, Hijmans et al. 2005), which has enabled researchers to study ecological responses to climate change globally. Such studies include: predicting past distributions of species climatic niches using historic data to reconstruct past species distributions (hindcasting) to answer questions
about extinctions and locations of past refugia (Martinez-Meyer, Townsend Peterson & Hargrove 2004; Peterson, Martinez-Meyer & Gonzalez-Salazar 2004; Nogues-Bravo et al. 2008; Pearman et al. 2008; Nogués-Bravo 2009); using climate to assess drivers of population change (Warren et al. 2013) and using future climate projections to determine the impacts of climate change on biodiversity (recently reviewed in Bellard et al. 2012).

Climate data from historic climate observations comprise the baseline period used to infer current climate-species occupancy relationships, and are central to assessments of climate change (Elith & Leathwick 2009). Uncertainty in these baseline data can be large, particularly where climate dynamics are complex, such as in mountainous or coastal regions (Baker et al. 2016). Despite this, baseline data uncertainties are almost completely overlooked when assessing the potential responses of species to climate change, yet a recent study showed that species responses are sensitive to baseline uncertainty (Baker et al. 2016).

Uncertainty in future climate change projections is well known, and attempts are often made to incorporate the component of this uncertainty into climate change assessments (Hole et al. 2009; Bagchi et al. 2013; Warren et al. 2013; Baker et al. 2015). Although assessments of general circulation models (GCMs) show high confidence in reproducing observed mean surface temperature patterns at large spatial scales (IPCC 2013), there are large errors in mountainous and coastal regions (IPCC 2013). Furthermore, precipitation patterns can be strongly influenced by smaller scale topography and vegetation (Moore et al. 2015), which occur below the large spatial scale of GCMs, making regional simulations difficult (Leung & Qian 2003). Regional regimes such as monsoons are poorly captured in GCMs (Kang et al. 2002). Additionally, annual mean precipitaion has known biases in tropical areas (Brown et al. 2013). Finally, climate features that are highly influential such as cloud cover are poorly simulated within GCMs (Groisman, Bradley & Sun 2000; Trenberth & Fasullo 2010; Webb, Lambert & Gregory 2013), resulting in poor estimates in biologically important metrics such as evapo-transpiration (Kingston et al. 2009). The availability of climate data means ecologists can use these data in their research without understanding the weaknesses of these datasets. Attempts to account for uncertainty are incorporated by repeating analyses using projections from different GCMs (Bagchi et al. 2012; Garcia et al. 2012; Baker et al. 2015).
1.6 Species distribution models

One way to ameliorate the effects of climate change is to determine where species are likely to be located in the future. Shifts in species distributions in relation to climate change could mean that areas that are currently of conservation importance for priority species may not retain such species in the future (Hole et al. 2009; Bagchi et al. 2013; Baker et al. 2015). Climate change scenarios, from general circulation model (GCM) projections, can be used in species distribution models (SDMs) to forecast the location of suitable climate space for species in the future. Correlative SDMs are the approach most commonly used to describe the relationship between patterns in biodiversity and species distributions, and the environment (Dormann et al. 2012b). SDMs are used to predict the potential distributions of species in space and time by relating species occurrences (or abundances) to environmental variables (Guisan & Thuiller 2005).

SDMs have been used extensively to evaluate the potential impacts of climate and land-use change on species distributions (Guisan et al. 2013) and abundances (Howard et al. 2014), to inform protected areas selection and management (Araujo et al. 2011; Bagchi et al. 2013), to identify historical refugia for biodiversity (Carnaval et al. 2009; Graham et al. 2010), and to evaluate the potential for the spread and establishment of invasive species (Broennimann & Guisan 2008; Roura-Pascual et al. 2009). Although SDMs are widely used in ecology, limitations on their accuracy, predictive power and transferability are widely recognised (Araújo & Peterson 2012; Dormann et al. 2012b; Heikkinen, Marmion & Luoto 2012). These limitations often arise from unrealistic model assumptions (Guisan & Thuiller 2005; Araújo & Peterson 2012). In addition, limitations can also arise as a result of the data used to build the models (Wiens et al. 2009; Jiguet et al. 2010; Rocchini et al. 2011). Among these limitations, a poorly recognised source of model uncertainty arises as a consequence of the spatial extent from which absence data are drawn for model building (Thuiller et al. 2004; VanDerWal et al. 2009). However, delineating the extent of the study region is not straightforward, as it depends on knowledge of the natural history and the dispersal ability of a species, as well as information about the landscape of interest, including the configuration of suitable habitats and barriers to dispersal (Soberon & Peterson 2005; Barve et al. 2011). Defining this is particularly difficult for migrant species, that inhabit spatially and temporally distinct areas during their annual migratory cycle, and that have the ability to disperse large distances.
1.7 The study of birds

Birds are widespread, and have long been studied by both amateur naturalists and professional scientists; they are therefore well understood. However, 13% of global bird species are globally threatened with extinction (BirdLife International 2015). Although declines in birds are also attributed to habitat loss and degradation, and the impacts of invasive species (Szabo et al. 2012; Bellard et al. 2013) climate change is regarded as one of the greatest threats to birds during this century (Thomas et al. 2004; Warren et al. 2013). Furthermore, climate change is likely to exacerbate the negative impacts of threats such as habitat loss and degradation (Mac Nally et al. 2009).

Data on movements of birds exist from bird ringing studies, as well as from direct observation of arriving and departing individuals (Pearce-Higgins & Green 2014). Quantitative counting and mapping techniques have provided over 50 years of population and distribution data collection of birds (Moller & Fiedler 2010). Historic data exist from museum specimens, on birds’ historical distributions and on the timing and success of breeding from nest records that span many decades. Recent technological advances allow the real-time tracking of birds in time and space (Green et al. 2002; Visser et al. 2009; Egevang et al. 2010; Robinson et al. 2010). These data provide the opportunity to study the relationships between birds and their environment now, and to predict the effect that climate change will have on their distributions and populations. Broad-scale spatial patterns in biodiversity, from continental to global scales, have been investigated using bird data, such as: the drivers of species richness patterns (Storch et al. 2006; Rahbek et al. 2007a) and the global distributions of range (Orme et al. 2006) and body sizes (Olson et al. 2009). An aspect that is less well understood is that of avian migration, due to the difficulties in studying species which bi-annually occur in geographically disparate regions. I go on to discuss the importance of migratory species globally, and in light of climate change.

1.8 Migration

Migrations are persistent, predictable, directional movements from one location to another (Dingle 1996). Billions of migratory animals move across the globe each year in pursuit of increased foraging opportunities, greater safety, and improved reproductive opportunities (Alerstam & Lindstrom 1990; Alerstam, Hedenstrom & Akesson 2003). The migration phenomenon has long fascinated researchers, due to the astonishing distances covered by species, as well as the vast numbers of individuals involved (Bauer & Hoye 2014).
For example, more than two billion passerines are estimated to migrate to sub-Saharan Africa each year (Hahn, Bauer & Liechti 2009).

Highly predictable, seasonal movements, in combination with the large spatial scales and considerable numbers of individuals involved, render migration a unique biological phenomenon. Further, it is an under-appreciated aspect of biodiversity, which is intimately linked with that of resident communities (Bauer & Hoye 2014). Movements of migrants, such as salmon, to their natal breeding grounds transfer vast amount of nutrients and energy from the ocean to freshwater systems (Holtgrieve & Schindler 2011), as well as to surrounding terrestrial habitats (Chen et al. 2011a). Migratory seabirds also transport large quantities of nutrients to terrestrial breeding colonies (Michelutti et al. 2009). Furthermore, migratory movements represent a unique mechanism by which seeds, spores, as well as parasites, can cross biogeographic barriers (Nathan et al. 2008; Altizer, Bartel & Han 2011).

Grazing by migrants can alter nutrient cycling, primary productivity, plant biomass, and consequently the composition and long term persistence of plant communities (Zacheis, Hupp & Ruess 2001; Holdo et al. 2007). Therefore, the foraging of migrants results in consumer-resource interactions across several locations. Further, migratory predators such as seabirds, raptors, marine mammals and fish potentially exert top-down regulation on prey populations and resident communities (Bauer & Hoye 2014). Migratory species are also preyed upon during their journey, for instance, wildebeest migration is a critical resources for lions in the Serengeti, who time their reproduction with the arrival of the migrant wildebeest (Holdo et al. 2011). Migrants therefore have the potential to alter community structure and ecosystem functions along their routes (Bauer & Hoye 2014).

As all these examples show, migration is not simply the movement of animals, but the coupling of ecological communities globally, and the mediation of ecological community diversity and stability. Given the potential of migration to affect ecological networks across the world, the ramifications of the potential decline of migratory species are considerable. Migration is an increasingly threatened global phenomenon, as a result of climate change, habitat destruction, changes in land use practises, and overexploitation (Wilcove & Wikelski 2008). The loss of migratory behaviour may result in a loss of ecosystem services, as outlined above.
1.9 The effects of climate change on avian migratory species

Migratory species may be particularly vulnerable to climate change, with the potential for intricately timed and long-established relationships of migratory species with their environment to become disrupted. I will discuss two responses of migratory species to environmental change: changes in time, and changes in space.

1.9.1 Timing of migration

In response to seasonal changes, migrants often travel long distances during migration between a breeding site and an overwintering area (Aidley 1981). To achieve this, birds must gain sufficient energy (in the form of fuel) to cover the flight distance (Schaub & Jenni 2001). Migratory birds, particularly long distance migrants, are constrained by time and energy during migration (Wikelski et al. 2003). Arctic breeders have been shown to have reduced reproductive success or complete breeding failure if they arrive at their breeding grounds outside a very narrow time window, because of competition for high-quality breeding sites and time constraints to raise young (Alerstam & Lindstrom 1990; Bauer, Gienapp & Madsen 2008). Both the timing of arrival and arrival body condition are dependent on migration speed, as well as the conditions at intermediate stop-over sites where fuel is accumulated for the next migratory leg (Hedenstrom 1997; Schaub & Jenni 2001; Holdo, Holt & Fryxell 2009).

Natural selection should favour migratory birds that can synchronise migration with peaks of resource availability, particularly as migrating birds use chains of stopover sites en route which often differ in seasonality (Bauer, Gienapp & Madsen 2008). Migratory animals tightly follow the seasonal development of resources along their migration routes (the “green wave” hypothesis), in order to match timing of stopovers with local peaks in food availability (van der Graaf et al. 2006; Tottrup et al. 2008; Schindler et al. 2013).

The timing of the onset migration is under endogenous control (Berthold 2001) but triggered by photoperiod (Gwinner 1996). In light of climate change, if birds are limited by their circannual rhythms, they may face a mismatch between migration timing and resource availability in the wintering, breeding and stopover areas (Gordo et al. 2005; McNamara et al. 2011). For arctic breeding species such as Barnacle geese (Branta leucopsis), migrating along a seasonal gradient of plant profitability toward their breeding grounds allows them to follow highly profitable food in order to replenish resources and arrive in optimal conditions to breed (Kölzsch et al. 2015). The timing of arrival of Barnacle geese was uninfluenced by...
ecological barriers, indicative of partly fixed migration schedules (Kölzsch et al. 2015). These migration schedules might become non-adaptive in light of future climate change predictions, where onset of spring is predicted to advance in northerly sites (IPCC 2007).

Phenological mismatches among species and across trophic levels have been documented, for example in the oak-winter moth-tit food chain (Buse et al. 1999; Visser & Holleman 2001; Both et al. 2009). To determine the effects of global climate change on bird ecology, it is of vital importance to understand the causes of changes in bird migration timing, and whether these changes might result in phenological mismatches, with potentially severe consequences for migratory birds.

Long-distance migrants are particularly vulnerable to mismatches because climate in their breeding and wintering areas may be changing at different speeds, hampering appropriate adaptation (Both & Visser 2001). Further, short distance migrants spend the winter closer to the breeding areas in comparison with long-distance migrants, which may allow such species to fine-tune the timing of spring migration in response to climate (Rubolini et al. 2007; Moller, Rubolini & Lehikoinen 2008). Indeed, the breeding populations of long-distance migrant birds are already in decline (Both et al. 2006; Sanderson et al. 2006; Heldbjerg & Fox 2008; Moller, Rubolini & Lehikoinen 2008).

1.9.2 Range shifts and migration distance

Global climate change has led to increased winter temperatures in North-West Europe, and studies have shown that migration distances have decreased for short distance migrants (Visser et al. 2009). These changes in migratory distance are in response to non-breeding ranges shifting northward because of improved wintering conditions in these areas (Visser et al. 2009; Pulido & Berthold 2010; Kullberg et al. 2015), therefore migrants are wintering closer to breeding areas. The resulting shortened migration distance is likely to modify the timing of spring arrival of migrants to the breeding grounds (Coppack & Pulido 2004; Pulido 2007; Pulido & Berthold 2010), which means these species can better time their arrival to breeding areas to the timing of food availability (Coppack & Both 2002; Walther, Berger & Sykes 2005). Current climate change is favouring birds wintering closer to the breeding grounds as it reduces migration costs and facilitates the rapid adaptation to shifts in the timing of food emergence on the breeding grounds (Visser & Both 2005; Bradshaw & Holzapfel 2008). In contrast with these findings, longer migration distances are expected given climate change, as migratory species breeding ranges (often located at higher
latitudes) respond strongly to latitudinal climatic gradients, but non-breeding ranges (often located in tropical regions) do not. Therefore, breeding and non-breeding ranges may shift in opposing directions, thus increasing the likely migration distances of species in the future (Huntley et al. 2006; Doswald et al. 2009).

The degree to which changes to the migratory landscape will affect animals depends on the ability of migratory animals to alter the timing, direction and destinations of their journeys (Bowlin et al. 2010). Under climate change, long-distance migrants are likely to become resident as warmer climates allow for species to overwinter on the breeding grounds (Berthold 2001). Blackcaps (Sylvia atricapilla) in central Europe, which used to migrate almost exclusively to Mediterranean and African wintering grounds have altered their migration route over the past few decades, and now winter in the British Isles, where a combination of warmer temperatures and bird feeders increases survival (Berthold 2001; Bearhop et al. 2005). Furthermore, an artificial selection experiment in a population of blackcaps showed that residency will rapidly evolve in migratory populations if selection for shorted migration persists (Pulido & Berthold 2010).

There are a limited number of studies which have assessed the potential impacts of climate change on the non-breeding grounds of migrants (see Austin & Rehfisch 2005; Studds & Marra 2007; Barbet-Massin et al. 2009), and few have assessed the impact of climate change on migratory species on both their breeding and non-breeding ranges (Bohning-Gaese & Lemoine 2004; Doswald et al. 2009; Wilson et al. 2011). Doswald et al. (2009) showed that, although potential range extent varied among a group of trans-Saharan migrants, Sylvia warblers, in general, both breeding and non-breeding ranges were predicted to increase in extent given climate change projections for the end of the 21st Century. Migration distances were projected to increase, and in many cases novel potential future non-breeding areas were simulated, which suggested that new migration routes may need to develop in response to climate change. Certainly, as a result of such changes, birds may need to develop new migration strategies, for instance, by increasing the number or duration of stopovers used, in order to obtain sufficient energy for longer journeys (Schaub & Jenni 2001).
1.10 Current limitations in understanding migrant bird responses to climate change

The velocity of climate change has triggered urgent calls for proactive conservation action to address the potential ecosystem-wide impacts of climate change (Strange et al. 2011). The prioritisation of future conservation should explicitly account for dynamic (and uncertain) impacts of climate change on species, including shifts in species distributions (Colwell et al. 2008; Watson et al. 2009; Carroll, Dunk & Moilanen 2010). Therefore, conservation planning will depend in part on our ability to identify where species are likely to occur in the future, given climate change projections (Strange et al. 2011).

The study of migration patterns and the impacts of climate change at all stages of the migratory cycle (breeding, non-breeding, and staging areas) is needed to understand the alarming declines of long-distant migrant birds (Sanderson et al. 2006; Tøttrup et al. 2012; Small-Lorenz et al. 2013). Our ability to predict climate change impacts on migratory species is limited by a lack of clear understanding of how climate determines migratory species distributions and richness at present. Furthermore, the majority of studies on migratory species are focussed on the breeding season and are largely North America or European focussed. Therefore, there is a need for a global assessment of the determinants of migratory species occurrence and richness on both the breeding and non-breeding areas.

Even with this understanding, future predictions of migratory species distributions are limited by our ability to project ranges of highly mobile species. Reasons include the fact that delineating an appropriate region for model projections for migrants is not straightforward; as they inhabit spatially and temporally distinct areas during their annual migratory cycle, and have the ability to disperse large distances. Additionally, as previously mentioned, delineating the extent of a study region for modeling purposes is not straightforward.

Therefore, in this thesis, I aim to understand the drivers of migratory species occupancy and richness at a global extent, under current climatic conditions. I aim to create and develop species distribution models that overcome the limitations that emerge when modeling highly mobile species. I aim to use these improved models to determine the locations of climatically suitable space for migratory species on the breeding and non-breeding ranges given future, climate change projections. More detailed project aims are outlined in section 1.8.
1.11 Project aims

1. The determinants of species richness across the globe

Firstly, I will describe geographic variation in avian migratory species richness, and identify what environmental factors drive the global distribution of migratory birds. I will ask whether the drivers of migrant species richness are the same for both the breeding and the non-breeding areas. I will consider whether migrant species richness is related to environmental variates consistently across the globe, or whether they relate in different ways, or to differing degrees, between biogeographically distinct areas.

2. How to select absence data for species distribution models

Building SDMs requires data on species presence, as well as either true- or pseudo-absence data, but there are few guidelines regarding the spatial area from which absence data should be drawn. I will compare approaches to absence selection that differ in the spatial areas from which absences are drawn. Specifically, I will develop SDMs for virtual species, conditioned on absence data selected by different methods.

3. What climatic factors are important for breeding/non-breeding migrants across the globe

Next, I will create species distribution models that relate migrant distributions globally to climatic predictors, using the absence selection approach developed in Chapter 3. These analyses will reveal what climatic factors make a given location a desirable place for a migrating bird to breed or spend the non-breeding season. I will also assess whether the factors that predict migrant occurrence on the breeding areas are the same as those that predict migrant occupancy in the non-breeding regions.

4. How is migration likely to change given future climate change projections

Finally, I will use current species projections of breeding and non-breeding ranges to develop a method to estimate current migration distances. I will then predict future migrant species distributions using global climate change projections. I will apply the previously developed distance estimation method to future species distributions to investigate species specific changes in migratory distance. These analyses will reveal the influence of future climate change on migratory species in terms of: changes in available climatically suitable space, changes in species richness and changes in species turnover.
Chapter 2

The determinants of migrant species richness across the globe

Whinchat, *Saxicola rubetra*, by Ian Fisher
2.1 Abstract

Broad scale geographic patterns in species richness have been studied extensively, but these studies have largely overlooked migratory species, which biannually redistribute themselves across large ecological scales. I describe geographic variation in avian migratory species richness, and identify what environmental correlates drive the global distribution of migratory birds. Using comprehensive global distribution data for all extant birds, I calculated the number of breeding migrants, the number of non-breeding migrants, as well as the number of non-migratory species across the globe. I built statistical models with ecologically meaningful predictors related to hypothesised drivers of migratory species richness. Analyses were conducted at a global scale and also for individual biogeographic realms. Explanatory variables of species richness included: landscape productivity, habitat diversity, seasonality of productivity, consistency of seasonality, resident species richness, the distance of migration, and climatic predictors (Actual to potential evapotranspiration, and mean temperature of the warmest and coldest month). Geographic variation in avian migratory species richness was mapped and areas of high and low migrant richness were identified. Drivers of migratory species richness differed between the breeding and non-breeding areas, with mean temperature of the coldest month being an important predictor of migrants on the breeding grounds, and mean temperature of the warmest month being more important for migrants on the non-breeding grounds. The results also suggest that migratory species richness is correlated with different environmental variables in different parts of the world, for example mean temperature of the coldest month was important for non-breeding migrants in the northern hemisphere (Nearctic and Palearctic biogeographic realms), but not in the southern hemisphere (Afrotropical, Australasian, IndoMalayan, and Neotropical realms). The strong predictive power of my models suggests that the main mechanisms that describe global patterns of migratory species richness have been captured. Furthermore, the putative driver of species richness show pronounced seasonal and regional variation.
2.2 Introduction

Global biodiversity loss is a major driver of ecosystem change (Hooper et al. 2012). Understanding the processes that shape global species richness is fundamental for conservation priority setting under such biodiversity loss (Kier et al. 2009). Many explanatory hypotheses (outlined below) have been proposed to explain the unequal distribution of species globally, yet little consensus has been reached (Palmer 1994; Rahbek & Graves 2001; Jetz & Rahbek 2002; Orme et al. 2005; Rahbek et al. 2007a; Gotelli et al. 2009). Curiously, unravelling the determinants of spatial patterns in global species richness has largely lacked consideration of the fact that many species are migratory, a major ecological phenomenon that leads to substantial seasonal shifts in the distribution of species richness (But see Somveille, Rodrigues & Manica 2015).

Avian migration leads to a biannual redistribution of billions of birds across the globe from their breeding to their non-breeding regions. These, often cross-continental, movements can lead to seasonal increases of avian species richness in areas that would otherwise be relatively species poor (Wisz, Walther & Rahbek 2007b). Further scrutiny of this phenomenon is warranted by the ecological scale of migration. Additionally, an enhanced understanding of migrant assembly rules might cast light on macroecological studies that aim to understand global patterns of species richness. Indeed, migration gives the opportunity to further understand these patterns, by testing hypotheses of the drivers of richness for the same group of species on both their breeding and non-breeding ranges.

Evaluating migratory species richness patterns poses unique challenges to ecologists, as migrants spend part of their annual cycle in different habitats, at different latitudes, and often on different continents. Investigations that have examined the drivers of migrant richness have focused mostly on narrow geographical regions, primarily Europe or North America, and on one part of the migration cycle (usually the breeding season) (Hurlbert & Haskell 2003; Lemoine & Bohning-Gaese 2003; Monkkonen & Forsman 2005; Wisz, Walther & Rahbek 2007b; Honkanen et al. 2010; Morrison et al. 2013; Blackburn & Cresswell 2015). To obtain a more complete understanding of the determinants of migrant species richness, spatial and temporal variation of biologically meaningful measures must be incorporated into analyses at a global scale.

Seasonality in climate and resources has frequently been proposed as a key driver of migrant species richness (MacArthur 1959; Somveille, Rodrigues & Manica 2015). Migration allows birds to take advantage of spatial variation in the seasonal fluctuation of resources (Alerstam, Hedenstrom & Akesson 2003) and, consequently, migratory species are often better able to exploit seasonal resource availability than resident species (Evans et al. 2006). Resource availability can be measured
using the remote sensing measure of greenness, the normalised difference vegetation index (NDVI). Studies have shown that seasonality in NDVI explained the distribution of migrant (and resident) species richness throughout the year (H-Acevedo & Currie 2003). In addition, seasonality in rainfall is thought to trigger movement of Palaeartic migrants within Africa (Moreau 1972), because rainfall increases plant productivity and consequently insect abundance (Jones 1995). Furthermore the consistency of seasonality in a given location is likely to be an important driver of migrant species richness as areas that are consistently seasonal across years are more predictable from year to year. Hence, non-breeding migratory species richness was shown to be negatively related to inter-annual variability in NDVI (Wisz, Walther & Rahbek 2007a).

Productivity is considered to be a determinant of non-migrant avian species richness (Jetz & Rahbek 2002; Hawkins, Porter & Felizola Diniz-Filho 2003) but production at a given site that is not utilised by residents should be available to migrant species; therefore, I expect migrant species richness to be influenced by the magnitude of the seasonal pulse of productivity (Nieto, Flombaum & Garbulsky 2015). Another putative driver of non-migratory species richness, which may also apply to migrants, is habitat heterogeneity. Habitat heterogeneity provides niche diversity, therefore facilitating a greater number of species in an area (Jetz & Rahbek 2002; Hurlbert & Haskell 2003; Rahbek et al. 2007b; Honkanen et al. 2010; Allouche et al. 2012).

A neglected area of research is the potential importance of interspecific competition between migrants and residents as a determinant of migrant species richness in a given location (Leisler 1992). High resident richness could mean more competition for returning migrants for nest sites and resources, which could be detrimental to migrant species (Berthold et al. 1998; Lemoine & Bohning-Gaese 2003; Ahola et al. 2007). Alternatively, migrant birds may use residents as cues for profitable breeding sites and resource availability under limited time for site exploration (Monkkonen & Forsman 2002; Monkkonen & Forsman 2005).

Temperature is also thought to influence species richness due to physiological constraints (Currie 1991), and is also likely indirectly to affect food and shelter resources (Wright 1983; Currie 1991; Hurlbert & Haskell 2003; Coops et al. 2009). Areas with milder winters have been shown to receive more species during the non-breeding season (Wisz, Walther & Rahbek 2007a; Dalby et al. 2014). Increased numbers of breeding migrants have been correlated with winter harshness, and this is likely to be due to reduced competition with resident species (Lemoine & Bohning-Gaese 2003; Carnicer & Díaz-Delgado 2008; Schaefer, Jetz & Böhning-Gaese 2008). Energy alone, or in combination with precipitation, has been linked to large scale variation in species richness. Hawkins et al. (2003) found that annual actual evapotranspiration (AET) alone accounted for 70% of the
variation in global avian species richness. Changes in the timing of spring migration are amongst the most commonly reported phenological responses to climate change (Walther et al. 2002; Jonzen et al. 2006; Knudsen et al. 2011). It is therefore likely that climate influences migratory species richness.

Successful migration often requires birds to travel considerable distances, sometimes thousands of kilometres, between breeding and non-breeding grounds, within a limited time span (Liechti 2006). Migrants that return to the breeding grounds early have been shown to benefit from: acquiring better territories; higher survival rates for chicks; and the opportunity for extra matings and clutches (Møller 1994; Bearhop et al. 2005; Møller, Rubolini & Lehikoinen 2008; Newton 2008). Moreover, long distance migrants are often considered to be under selective pressure to minimise time spent migrating (Lindstrom & Alerstam 1992; Weber & Houston 1997). It follows that, given two potential migratory end-points of equal suitability, migrants should utilise the area that is closer, in order to minimise migratory distances. For example, the number of Palaearctic migratory species progressively declines in a north-south gradient within sub-Saharan Africa (Hockey 2000), perhaps because European migratory species migrate no further south than is necessary to find suitable conditions (Wisz, Walther & Rahbek 2007a; Newton 2008).

Here, I use species distribution data for migrant birds globally to address three issues related to migrant species richness. First, I aim to understand what determines migrant species richness across the globe, and whether these differ from those previously shown to predict avian richness (without differentiating between resident and migratory species). Second, I ask whether the drivers of migrant species richness are the same for both the breeding and the non-breeding areas. The spatial extent at which data are sampled can have a large impact on the results of studying the relative importance of environmental predictors (Davies et al. 2007; Rahbek et al. 2007a; Bradter et al. 2011). Therefore, thirdly, I consider whether migrant species richness is related to environmental variates consistently across the globe, or whether they relate in different ways, or to differing degrees, between biogeographically distinct realms.
2.3 Methods

2.3.1 Species distribution data

Global avian species distribution data of > 10,000 extant avian species were obtained from BirdLife International and NatureServe (Birdlife International & NatureServe 2014), with species defined according to BirdLife’s avian taxonomy (BirdLife International 2014). Primarily marine species were excluded from analyses, as the vast majority of their ranges are oceanic and therefore of different character to, and not comparable with, terrestrial species. Coastal species were also removed as many such species also forage in the marine and inter-tidal zones, such that their dependency upon terrestrial ecosystems is variable. The remaining terrestrial species distributions were gridded at a 0.5° X 0.5° (~60 x 60 km at the equator) equal-area grid in a Behrman equal area projection. Species were considered to be present in a grid-cell if their ranges intersected at least 50% of that cell.

Migrants were defined (in accordance with BirdLife International) as those that make regular cyclical movements beyond the breeding range, with predictable timing and destinations. This definition includes species that are only migratory in part of their range or population, and short distance migrants, but does not include truly nomadic species or altitudinal migrants. Breeding migrant richness was calculated as the sum of all migratory species that were present in a terrestrial cell during the breeding season. Non-breeding migrant richness was the sum of migrants that were present in a cell in the non-breeding season. Similarly, resident avian richness was defined as the sum of non-migrant avian species that were present in a cell. Resident species richness was used in our models as a predictor of migratory species richness, as a proxy for potential competition with migrants.

2.3.2 Bioclimatic variables

Mean monthly temperature and precipitation data were obtained from Worldclim (Hijmans et al. 2005, http://www.worldclim.org/) on a 0.04° resolution across the globe, from the period 1951-2000 (a period corresponding to the data used to produce the species’ range maps). These data were aggregated to a 0.5° resolution by ranking the 0.04° grid cells within each 0.5° cell by elevation and calculating the mean value for the cells that lay between the 25% and 75% quartiles of the elevation range for the 0.5° grid cell (Hole et al. 2009). This limits distortion of the mean climate by atypical values found at extremely high or low elevations. Bioclimatic explanatory variables which have been shown to be useful for modelling avian species distributions in biomes ranging from the arctic to the tropics were included in our analyses (Huntley et al. 2006; Hole et al. 2009; Araújo et al.
2011). The variables were: mean temperature of the warmest month (MTWM); seasonality and productivity metrics (see below); mean temperature of the coldest month (MTCO); and an estimate of the ratio of actual to potential evapotranspiration (APET). APET was estimated for each grid square using the model described in Prentice et al. (1992).

2.3.3 Productivity and seasonality

To investigate the effect of vegetation productivity and seasonality on migrant richness, NDVI (normalised difference vegetation index) data were used, processed by the Global Inventory Modelling and Mapping Studies group (GIMMS; Pettorelli et al. 2005; Tucker et al. 2005). NDVI correlates directly with vegetation productivity and provides information about net primary productivity (NPP), and vegetation biomass and quality (Reed et al. 1994; Pettorelli et al. 2005). These data were global at a 0.07 degree resolution (8km by 8km) and were available at bimonthly intervals from 1982 to 2006. For the purpose of this investigation, data were obtained for a period that maximised the overlap between the NDVI data and the period corresponding to when the BirdLife species distribution data were collected (from the earliest possible NDVI record (1982) to when the BirdLife polygons were finalised (2000)). The data were aggregated to a 0.5° resolution by estimating the mean NDVI from all pixels that were contained within each 0.5° cell.

Three NDVI metrics that I considered might influence migrant species richness were estimated as follows:

(1) The mean total growing season NDVI (INDVI) was calculated as the sum of positive NDVI values over the growing season across years, for each 0.5 degree cell across the globe. This value is a proxy for the productivity in a given cell.

(2) The seasonal amplitude of a cell’s productivity (NDVI\text{seasonality}) was estimated as the difference between the maximum and minimum non-negative annual NDVI values for a given cell in any one year. The mean of these annual amplitudes across the 19 year period (1982-2000) was obtained. This value is a proxy for the seasonality in a given cell.

(3) The inter-annual variability in seasonality (of primary productivity) is a measure of how consistently seasonal a cell is (NDVI\text{consistency}). A location that is consistently seasonal across years might be more attractive to migrants. To estimate NDVI\text{consistency}, I calculated the mean coefficient of variation of bimonthly NDVI estimates over the 19 year period (1982-2000). Specifically, the variance of NDVI for each of the 24 annual bimonthly time periods across the 19 years was divided by the
mean NDVI estimate for that time same period across the 19 years. This value is a proxy for the consistency in seasonality for a given cell.

### 2.3.4 Habitat diversity

Global land cover data were downloaded from the USGS Global Land Cover Characteristics Database (version 2, [http://edc2.usgs.gov/glcc/globe_int.php](http://edc2.usgs.gov/glcc/globe_int.php)). This source classifies 22 habitat classes using 1-km Advanced Very High Resolution Radiometer (AVHRR) data spanning a twelve month period (April 1992-March 1993). The USGS database was chosen due to its global coverage and because the time period of data collection overlapped with the species and NDVI data. Percentage coverage of each habitat class at 0.5° resolution was used to estimate Shannon habitat diversity (Shannon 1948), \( H \), for each cell.

### 2.3.5 Distance

Given the cost of migration, the number of non-breeding migrants in a cell is likely to be affected by proximity to the breeding grounds (and vice-versa). This “distance to the breeding grounds” metric was estimated, for each terrestrial cell globally, as the minimum great circle distance from that cell to the closest edge of each migrant species’ breeding distribution, calculated using the sp package in R (Pebesma & Bivand 2005). Each cell across the globe was then assigned a distance metric, which is the mean distance from that cell to nearest breeding cells for all migrant species. If a cell overlapped with a species’ breeding range the cell was assigned a distance of zero for that species. Similarly, the probability that a migrant bird will breed in a given cell is likely to be reflective, at least in part, of distance from the non-breeding grounds to that cell. “Distance to the non-breeding grounds” (to predict migrant richness on the breeding grounds) was calculated in a similar manner; but here the mean of minimum great circle distances between a given cell and the closest edge of all of the migrant non-breeding ranges was estimated. Spatial patterns of these distance predictors, as well as all other predictors, are displayed globally (Figure S 1).

### 2.3.6 Modeling approach

A multi-stage modelling approach was used to explore the potential determinants of migrant species richness that removed highly correlated predictor variables and minimised the potential confounding effects of spatial autocorrelation (as outlined in Voskamp et al. 2015). The initial pool of variables selected as potential determinants of migrant species richness on both the breeding and non-breeding grounds were: three bioclimatic variables (APET, MTWM, MTCO); seasonality (NDVI\(_{\text{seasonality}}\)); seasonal consistency (NDVI\(_{\text{consistency}}\)); habitat diversity (H); annual vegetation...
productivity (INDVI); resident avian species richness and distance to the breeding/non-breeding grounds. Linear and quadratic terms for these variables were used in modelling.

Spatial autocorrelation, where observations of nearby locations are more similar than expected at random, is a frequent phenomenon when analysing spatial data (Legendre 1993). Species richness values were highly spatially correlated (Moran’s I = 0.99). The effect of spatial autocorrelation was minimised when modelling by using a blocking approach where the transferability of fitted models to spatially segregated test data, was assessed. The globe was divided into 36 blocks, and these blocks were separated into two groups based on a checkerboard pattern. Models were created using one half of the checkerboard, sampling 10% of the data at random from within one set of blocks, and testing the resultant model on random samples of 10% of the cells from the left out set of blocks. This approach substantially reduced the spatial autocorrelation (mean Moran’s I across random subsets = 0.15, Table 2.1).

Species richness was modelled at two spatial extents, firstly at a global extent and secondly at the scale of individual biogeographic realms. To test whether migrant species richness is related to environmental variates in different ways, or to differing degrees, in individual biological realms, I produced models similar to the global models described above, but fitted instead to individual realms. Biogeographic realms were delimited following Olson et al. (2001) which defines eight major realms: Afrotropical, Australasian, Indo-Malayan, Nearctic, Neotropical, Palearctic, Antarctic and Oceanic. From these, the latter two realms were excluded from our analyses as both host few terrestrial migrants. Separate models were produced for the remaining six realms. More recently, Holt et al. (2013) proposed a more comprehensive series of biogeographic realms. However, the realms delimited by Olson et al. (2001) were used as they were more comparable in terms of individual realm extents, which minimised potential issues when comparing realm models.

Firstly highly correlated variables were removed (Pearsons’ r > 0.7, following Dormann et al. (2013)) from all models, i.e., for models of breeding and non-breeding areas, at both the global and realm level, as follows: All combinations of variables from our initial variable set (that included no variable pairs with correlations greater than r = 0.7), were selected i.e. a saturated model. For each season/realm combination, generalized linear models (GLMs; Poisson error distribution) of migratory species richness were fitted, that included all non-correlated combinations of covariates and their quadratic terms, to 1000 random subsets of the data using the blocking approach described above. Each time, the Akaike Information Criterion (AIC) was recorded, and the mean AIC from across the 1000 models was used to compare the among the candidate predictor
combinations. The combination of predictors with the lowest mean AIC was selected as the best model for that particular season/realm.

Secondly, considering all variables included in the best saturated model for a season/realm, the subset of these variables that produced the most parsimonious model was evaluated. This was assessed using model selection with AIC, where models with a ΔAIC of ≤6 and lower than the ΔAIC of all simpler models were considered (Richards 2008). For model selection, the MuMIn package in R (Barton 2015) was used, fitting 1000 models for each variable combination, using the sampling procedure described previously. This subsampling process resulted in subtly different variable selection among the 1000 models. To ensure that only those parameters with robust and repeatable influences on migratory species richness were identified, the frequency with which individual environmental covariates were selected in the top models across the 1000 subsets was recorded. I report the parameter values for those covariates that were included in at least 90% of the top models (example Figure S 2). Parameters defined as robust by this method were used to predict species richness for all terrestrial cells globally. Models fitted separately for each geographic realm used a similar blocking approach, with each realm split, checkerboard-fashion, into 10 equal-sized blocks, which were sub-divided into five training blocks and five test blocks. Following the same approach described for the global level models, the best model for each realm was identified.

Global and realm level models (each built on a subset of data as described above) were assessed in terms of how well they predicted migratory species richness using McFaddens $R^2$ (Beaujean 2012). Models that were calibrated at a global scale were also used to predict breeding and non-breeding migrant species richness at the individual realm level. As described above, models were built on subsets of the entire dataset in each case but were applied to the entire region of interest. A mean $R^2$ was calculated from across the 1000 iterations of the best model for a season/realm.

To test whether realm based models predict global species richness better than global models, predictions that were made from realm models fitted to individual realms were combined to predict migratory species richness on both the breeding and non-breeding ranges across the globe (excluding Antarctica and Oceania). AIC (based on the least squares case from Burnham and Anderson (2002)) was used to compare the global model to the amalgamated realm models, penalising the increased model complexity of the realm level models.
2.4 Results

My results show that global models of non-breeding migrant richness performed less well than amalgamated realm models, which can be seen clearly in Figure 2.1 where non-breeding migrant richness was vastly under-predicted when using global models. This improvement in models created using amalgamated realm models rather than global models was evident, though not as pronounced, for breeding migrant richness (Figure 2.1).

Realm level models, which are built using varying predictor combinations, are superior at predicting migratory species richness than global models (Table 2.1) in particular at extremes of high and low species richness (Figure 2.3). Furthermore, models are generally able to predict migratory species richness on the breeding grounds better than on the non-breeding grounds (Table 2.1). The majority of models perform well ($R^2 > 0.7$) at predicting migratory species richness, but some, such as the Neotropical realm models, performed poorly ($R^2 < 0.6$) in comparison with those from other realms (Table 2.1).

When global models, created using variables that perform well at predicting migrant richness (on both the breeding and non-breeding grounds) across the globe, were projected to individual realms, models did not recover species richness well in the majority of realms ($R^2 = 0.11$-$0.62$, Table 2.1). Nevertheless, these global models did predict predicted non-breeding migrant richness in the Nearctic (non-breeding $R^2=0.83$) and breeding migrant richness Palearctic realms (breeding $R^2 =0.78$) well.
Table 2.1. Model accuracy (evaluated using mean McFadden’s $R^2$) at predicting migrant species richness on the breeding and non-breeding grounds, using models fitted at global and biogeographic realm extents. Mean autocorrelation values of model residuals after subsampling (given as Moran’s I). $R^2$ (globe to realm) is the $R^2$ value of models created at the global scale, and applied to each individual realm. The number of grid cells underlying each model extent is also provided.

<table>
<thead>
<tr>
<th>Extent</th>
<th>Season</th>
<th>$R^2$</th>
<th>$R^2$ (globe to realm)</th>
<th>Moran’s I</th>
<th>Number of cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Breeding</td>
<td>0.73</td>
<td>-----</td>
<td>0.17</td>
<td>55,974</td>
</tr>
<tr>
<td></td>
<td>Non-breeding</td>
<td>0.61</td>
<td>-----</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>Afrotropical</td>
<td>Breeding</td>
<td>0.80</td>
<td>0.11</td>
<td>0.15</td>
<td>9,051</td>
</tr>
<tr>
<td></td>
<td>Non-breeding</td>
<td>0.70</td>
<td>0.45</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>Australasian</td>
<td>Breeding</td>
<td>0.74</td>
<td>0.34</td>
<td>0.13</td>
<td>3,926</td>
</tr>
<tr>
<td></td>
<td>Non-breeding</td>
<td>0.81</td>
<td>0.41</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>Indo-Malayan</td>
<td>Breeding</td>
<td>0.72</td>
<td>0.62</td>
<td>0.13</td>
<td>3,375</td>
</tr>
<tr>
<td></td>
<td>Non-breeding</td>
<td>0.71</td>
<td>0.27</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>Neotropical</td>
<td>Breeding</td>
<td>0.58</td>
<td>0.23</td>
<td>0.15</td>
<td>8,334</td>
</tr>
<tr>
<td></td>
<td>Non-breeding</td>
<td>0.41</td>
<td>0.27</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>Nearctic</td>
<td>Breeding</td>
<td>0.82</td>
<td>0.58</td>
<td>0.13</td>
<td>8,532</td>
</tr>
<tr>
<td></td>
<td>Non-breeding</td>
<td>0.95</td>
<td>0.83</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>Palaearctic</td>
<td>Breeding</td>
<td>0.88</td>
<td>0.78</td>
<td>0.16</td>
<td>22,765</td>
</tr>
<tr>
<td></td>
<td>Non-breeding</td>
<td>0.77</td>
<td>0.57</td>
<td>0.13</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2.1 Observed versus predicted migrant species richness from models generated using global (a,b) and combined realm extent (c, d) models for both the breeding (a,c) and non-breeding (b, d) areas. R squared values for the breeding and non-breeding richness predictions from global models are 0.85 and 0.71 respectively (a and b). R squared values for both the breeding and non-breeding richness predictions from combined realm models are 0.89 and 0.86 respectively (c and d). Red line corresponds to the 1:1 line.

The realm level models, when amalgamated, described migrant species richness better than the global level model (Breeding: Realm model AIC: 13403, Global model AIC: 152002 and non-breeding: Realm model AIC: 112952, Global model AIC: 130636). Amalgamated realm level models captured peaks in migrant species richness better than global models (Figure 2.2). In particular, global models strongly underestimated Palaearctic breeding diversity (Figure 2.2 (c)) and non-breeding diversity in the Indo-Malayan region (Figure 2.2 (d)).
Figure 2.2 Observed migrant breeding (a) and non-breeding (b) species richness. Breeding (c) and non-breeding (d) migrant species richness predictions based on the global model ($R^2=0.73$ and 0.61 respectively). Breeding (e) and non-breeding (f) migrant species richness predictions which are based on combined realm models ($R^2=0.85$ and 0.71 respectively).
Distance to the corresponding breeding or non-breeding grounds, productivity, resident richness and habitat heterogeneity were important for predicting migrant richness on both the breeding and non-breeding grounds when using global models (Figure 2.3, Figure 2.4). Mean temperature of the coldest month was the most important predictor of global breeding migrant richness, with the majority of breeding migrant species richness occurring at intermediate temperatures. Increased mean temperature of the warmest month and intermediate levels of resident richness were important predictors of non-breeding migrant richness globally (Figure 2.4).

Predictor importance varied among realms. Overall, resident richness was a consistently important predictor of both breeding and non-breeding migrant species richness. For areas of high breeding migrant occurrence (Nearctic and Palearctic realms) low levels of resident species richness determined migrant richness. The importance of predictors differs between these two realms, where Nearctic migrant richness is predicted by high levels of APET, habitat diversity (H), seasonality, productivity and MTWM, and Palearctic richness is predicted by low MTCO, consistency of seasonality and habitat diversity. Distance to the non-breeding grounds is an important predictor of breeding migrant richness in both realms.
Predictor importance for non-breeding migrant richness also varied among realms (Figure 2.4). For areas of high non-breeding migrant richness (Afrotropical, IndoMalayan and Neotropical realms) intermediate levels of resident species richness was an important determinant of migrant richness. Migratory richness decreased with increasing distance to the breeding grounds in these realms. Productivity and habitat diversity were important predictors of non-breeding migrant richness across these realms, however, relationships with richness varied among realms (Figure 2.4).
Figure 2.4 Predictors of non-breeding migrant species richness (APET (actual to potential evapotranspiration); CS (consistency); Dist (Distance); INDVI (productivity); MTCO (mean temperature of the coldest month); MTWM (mean temperature of the warmest month); RR (Resident species richness); Seas (Seasonality) and H (habitat heterogeneity)) that were not consistently included in the top models are blocked out in grey. Locations are global (GL), and realm: Afrotropical (AF); Australasia (AUS); IndoMalayan (IM); Neotropical (NEO); Nearctic (NE) and Palearctic (PAL). X-axis is the total range of values for each variable across the globe; y-axis is the maximum species richness globally (GL) and the maximum species richness within each individual realm (AF-PAL). These relationships are illustrative, and based on models produced on one half of the blocked data. Boxes highlight the principal regions of high migrant species richness on the non-breeding grounds.
2.5 Discussion

2.5.1 Global patterns of migrant species richness

The approach used in this study highlighted global patterns of migratory avian species richness, as well as the importance of underpinning predictors. Furthermore, it allowed the assessment of previously proposed drivers of species richness patterns, which has not been considered simultaneously for migratory species at this global spatial extent.

I show that migrant richness, in both the breeding and non-breeding season can be predicted accurately using a limited number of explanatory variables. Explanatory models of migrant richness performed well both at the global and at the biogeographic realm extent (Table 2.1). Models predicted migrant richness more accurately when conditioned and applied at a realm extent (Figure 2.1, Figure 2.2) than at a global extent. Models are expected to perform better when analysed at the realm level, because migrants are likely to respond differently to diverse environmental or biological drivers in different areas. In fact, realm level models captured the extremes (areas of high and low) migrant richness more accurately (Figure 2.2).

Models did not perform constantly well across realms, in particular, models of both breeding and non-breeding migrant richness in the Neotropical realm performed poorly ($R^2$ =0.58, $R^2$= 0.41 respectively). Other ecological processes not included in our analyses may influence the species richness of migrants in this region. For instance, the altitudinal range of the Andes was not taken into account in our analyses, which could act as a centre of recent speciation in the Neotropical area (Rahbek & Graves 2001; Kattan & Franco 2004). Moreover, biogeographical realms, some of which are large, can span multiple ecosystems, and variables which explain migrant species richness may not operate in the same way across the entire area. A finer scale than the realm level for conducting analyses could highlight the importance of locally important predictors which are overlooked in the realm-level analyses.

2.5.2 Potential drivers of migrant species richness

Resident richness was consistently selected as an important predictor of breeding migrant species richness at both the global and realm level (Figure 2.3, Figure 2.4). Migrants may use residents as cues for profitable sites (Monkkonen & Forsman 2005), where direct assessment of available breeding/non-breeding patches may not be possible. Further, the correlation of resident species richness with migratory richness could be attributed to underlying processes that drive resident species richness to be high in certain areas, such as topographic variability (Davies et al. 2007). The overall relationship between migrant and resident species richness at the realm level, is a
hump-shaped one, which is to be expected if migrants use residents as cues for profitable patches, but high resident richness results in competition for returning migrants for nest sites and resources, which could be detrimental to migrant species (Berthold et al. 1998; Lemoine & Bohning-Gaese 2003; Ahola et al. 2007).

Migration distance was a key predictor of non-breeding migrant richness in the areas of high non-breeding migrant richness (the Afrotropical, Neotropical, Australasian and Indo-Malayan realms, Figure 2.4). This supports the idea that migrant birds should not migrate further than is necessary to find suitable non-breeding habitat (Newton 2008). The southward gradient of declining species richness of Palaearctic migrants in the Afrotropical non-breeding grounds (similarly of Nearctic migrants in the Neotropical non-breeding grounds) is likely to be as a result of this concept. Further, non-breeding migrants are concentrated in the southern regions of the northern latitudes, rather than being distributed in areas of suitable climate in the southern hemisphere (Figure 2.2). Distance to the breeding grounds (Nearctic and Palearctic realms) was an important predictor of breeding migrant species richness, with intermediate migratory distances predicting breeding migrant species richness in the Nearctic realm. The relationship is not clear for the Palearctic realm, which is likely to be an artefact of the large size of the Palearctic realm.

Temperature was also an important predictor of migrant richness, but there were clear differences between realms. Mean temperature of the coldest month was an important predictor of migrant diversity (both breeding and non-breeding) in the Palearctic realm, and for non-breeding migrants in the Nearctic realm, both of which are in the northern hemisphere where thermal limitations are likely to be associated with the cold (Hickling et al. 2006; Pearce-Higgins & Green 2014). Mean temperature of the warmest month was more important for the southern hemisphere realms (Afrotropical, Neotropical, Indo-Malayan and Australasian), where drought and hot climates are likely to be more limiting to species richness (Newton 2004).

INDVI, a surrogate for net primary productivity, was selected across the realm level models as important for non-breeding migrant richness, but was only important for breeding migrants in the Nearctic realm, as well as for breeding richness from global models (perhaps influenced by the importance of this variable for breeding migrants in this realm, which hosts a large proportion of migrant diversity in the breeding season). This supports the idea that environmental productivity is a major determinant of broad scale taxonomic richness (Jetz & Rahbek 2002; Hawkins, Porter & Felizola Diniz-Filho 2003).
In order for migration to be a winning strategy, seasonal fluctuations in primary production must be predictable (Hurlbert & Haskell 2003). Surprisingly, seasonality and inter-annual seasonal variability (consistency) were not important predictors of breeding migratory richness at the global scale. The importance of these variables differed between seasons and among realms, with consistency predicting migrant diversity in the Palearctic and Afrotropical realms. Seasonality was only found to predict non-breeding richness in the Indo-Malayan region (Figure 2.4). Seasonality was an important predictor of breeding migrant richness in the Nearctic realm which is consistent with previous studies (Figure 2.3) (Dalby et al. 2014). Surprisingly, seasonality was a poor predictor of breeding migrant richness in the Palearctic region, where previous studies have shown seasonality to be a major driver of richness where incoming migrants benefit from a surplus of resources in areas of high seasonality (Somveille, Rodrigues & Manica 2015). Our analyses are limited by the fact that annual seasonality might be uninformative about how seasonal an area is at the period when the migrants are present in an area.

In agreement with previous studies of the determinants of avian diversity, habitat heterogeneity was an important predictor of migrant species richness at the global level (Jetz & Rahbek 2002; Davies et al. 2007). This is consistent with the habitat heterogeneity theory, where increased habitat diversity supports increased species numbers (Jetz & Rahbek 2002; Hurlbert & Haskell 2003; Rahbek et al. 2007b; Honkanen et al. 2010; Allouche et al. 2012). APET, a measure of moisture availability, is known to affect the timing of spring migration (Studds & Marra 2011), was an important predictor of migrant species richness in only the Nearctic, Australasian and Palearctic realms (Figure 2.3, Figure 2.4). This contrasts with Hawkins et al. (2003), who found that annual actual evapotranspiration (AET) accounted for 70% of the variation in global avian species richness. Therefore, the drivers of species richness of migrant birds may differ to those of non-migrants.

Diversity and distribution patterns are not solely related to climate and habitat factors such as those considered here. Other factors that could affect the species richness patterns of migrants include historical biogeography and small scale niche driven assembly patterns (Rahbek et al. 2007a). For example, the considerable species richness of Andean regions is likely to be driven by high levels of speciation as a result of the highly heterogeneous topography (Rahbek & Graves 2001). Further, migratory species may be displaced due urbanisation of landscapes, as well as human persecution (McCulloch, Tucker & Baillie 1992; Rodewald & Shustack 2008). These factors which have not been explicitly assessed may impact on migrant distributions.

In summary, previous macroecological studies have highlighted the importance of plausible drivers that explain species diversity, but few have analysed the drivers of migratory species at a
global scale. Often ignored in macroecological studies, migratory birds comprise over 80% of biodiversity of temperate regions across the globe (Rappole 1995). Recent research has highlighted the necessity to include migratory species in conservation planning (Runge et al. 2014), as rapid declines in migrant birds have been observed (Kirby et al. 2008). Studies that have analysed migrant species richness, have often focused on the breeding ranges of the species, overlooking the patterns of diversity of non-breeding migrant richness and the underlying drivers of these patterns, and how the two seasons might be related. Given the rapid biodiversity loss experienced globally, understanding drivers of species richness at a global scale is essential (Sala et al. 2000). Here I show that migratory species richness on both the breeding and non-breeding ranges can be predicted accurately using a small number of ecologically meaningful variables. Our study highlights the variables that drive migratory avian richness patterns on a large scale in different ecoregions of the world. Finally, global models fail to capture all the detail and intricacies that the narrower biogeographic realm extent models captured. Therefore I recommend future studies to be conducted at the global scale in order to capture the full annual cycle of dynamic migratory species, but that models should be calibrated at finer geographical scales, such as the biogeographic realm, to capture regional differences in relationships with these species.
Chapter 3

How to select absence data for species distribution models

Cuckoo, *Cuculus canorus*, by Ian Fisher
3.1 Abstract

Species Distribution Models (SDMs) are used widely in ecology for identifying species’ habitat preferences, and for predicting how the suitability of habitats might change in space or time. SDMs have found particular utility in predicting species’ potential responses to climate change. Building SDMs requires data on species’ presence, as well as either true- or pseudo-absence data, but there are few guidelines regarding the spatial extent from which absence data should be drawn. In fact, there is likely to be a trade-off in selecting the area from which absence data are drawn: too narrow an extent will likely truncate the full environmental response curve of a species, limiting the transferability of an SDM; too large an extent may incorporate suitable but unoccupied cells due to geographical constraints, potentially weakening underlying relationships. Selecting an extent over which to create an SDM is, thus, crucial. In spite of its direct effect on model performance, this remains a surprisingly overlooked step in the SDM process. Here, I compare approaches to absence selection that differ in the spatial extents from which absences are drawn. Specifically, I developed SDMs for virtual species, conditioned on absence data selected by different methods. I assessed model ability to predict species’ occurrence but also, importantly, their ability to detect the underlying climatic relationships for a species. Five absence selection approaches were used, including random selection, selecting concentric cells around a species’ range, and several intermediate, distance-weighted, selection methods. Models conditioned on absences drawn from extremely wide and extremely narrow extents provided poor descriptions of species’ underlying climatic niches. Overall, I found that SDMs using distance-weighted approaches to selecting absences consistently performed best in recreating the underlying climatic niche. The results suggest that existing, often arbitrary methods for absence selection might simulate current species’ distributions well but fail to identify underlying climatic relationships accurately. Using a weighted absence selection, whereby absences are preferentially selected close to the boundary of a species’ distribution, but with some absences drawn from further afield, overcomes the necessity to delimit an extent based on arbitrary decisions and outperforms other methods across climatic parameter space.
3.2 Introduction

Current rates of environmental change, including changes in climate, land-use and human exploitation of natural resources, are unprecedented in recent history (Sala et al. 2000; Pereira et al. 2010; Mantyka-Pringle, Martin & Rhodes 2012). Species worldwide are responding to climatic changes with changes in phenology, range and abundance changes, and local evolutionary responses (Parmesan & Yohe 2003; Chen et al. 2011b). Predicting species’ potential responses to climate change is vital for conservation planning and mitigation (Guisan et al. 2013), and to ensure the continuity of ecosystem service provision (Naidoo et al. 2008). Species Distribution Models (SDMs) correlate species’ spatial distributions with environmental conditions (Guisan & Zimmermann 2000b). SDMs have been used extensively to evaluate the potential impacts of climate and land-use change on species’ distributions (Guisan et al. 2013) and abundances (Howard et al. 2014), to inform protected areas selection and management (Araujo et al. 2011; Bagchi et al. 2013), to identify historical refugia for biodiversity (Carnaval et al. 2009; Graham et al. 2010) and to evaluate the potential for the spread and establishment of invasive species (Broennimann & Guisan 2008; Roura-Pascual et al. 2009).

Although SDMs are widely used in ecology, limitations on their accuracy, predictive power and transferability are widely recognised (Araújo & Peterson 2012; Dormann et al. 2012b; Heikkinen, Marmion & Luoto 2012). These limitations often arise from unrealistic model assumptions (Guisan & Thuiller 2005; Araújo & Peterson 2012). In addition, however, limitations can also arise as a result of the data used to build the models (Wiens et al. 2009; Jiguet et al. 2010; Rocchini et al. 2011). Among these limitations, a poorly recognised source of model uncertainty arises as a consequence of the spatial extent from which absence data are drawn for model building (Thuiller et al. 2004; VanDerWal et al. 2009).

Absence points provide a contrast to conditions where a species occurs. However, the study extent or “arena” (i.e., area from which absence data are drawn) is often defined arbitrarily, based on the region of data collection, a biogeographical realm, or a political or other convenient boundary around an area of interest. Even when only presence data are available, SDMs rely on artificial absence (often termed pseudo-absence) data drawn from the dataset provided (excluding presences). This, itself, is often arbitrarily defined (VanDerWal et al. 2009). In both presence-only and presence-absence modelling, arbitrarily defined arenas can be particularly problematic if a species occurs beyond or abutting the defined extent, potentially leading to the truncation of climatic niches (Thuiller et al. 2004; Barbet-Massin, Thuiller & Jiguet 2010).
Even when a species’ occupied range occurs entirely within the modelled extent, decisions about the source of absence data remain critical. This is because a fundamental trade-off exists in choosing the extent of the modelling arena. Specifically, for SDMs designed to link patterns of occurrence to climatic conditions, drawing absence data from a narrow extent: (1) minimises the risk of including cells that are climatically suitable but unoccupied due to historical processes or limits on dispersal (which would dilute the true climatic signal; Chefaoui & Lobo 2008; Anderson & Raza 2010); (2) avoids swamping the model with geographically distant data that are low on useful information (and which provide the appearance of high discrimination without improving model performance in areas of borderline suitability; Anderson & Raza 2010; Acevedo et al. 2012); but (3) limits the climatic space sampled (potentially constraining the predictive capacity of models; Thuiller et al. 2004; Barbet-Massin, Thuiller & Jiguet 2010). By contrast, drawing absence data from a very wide extent ensures that climate space is sampled more fully but risks dilution of the climatic signal, resulting in poor performance in areas of borderline suitability.

Several sampling strategies for selecting absence and pseudo-absence points have been suggested. Strategies include: selecting absences randomly from a predefined background area (Wisz & Guisan 2009); selecting absences within (or outside) a certain geographic distance from presences (Barbet-Massin et al. 2012; Vale, Tarroso & Brito 2014); and selecting absence points from outside a predefined region based on a preliminary model (Ward & Morgan 2014). Performance is usually judged based on one or more of a range of metrics, including the area under the curve (AUC) of the receiver operating characteristic, True Skill Statistic (TSS), Kappa, sensitivity and specificity. A major drawback of these metrics is that they are all calculated with reference to the model’s ability to recreate presence-absence distributions; models that score well by these metrics might, nevertheless, be poorly suited to predicting to novel situations. This is because presence/absence is a coarse metric that does not necessarily represent the underlying climatic suitability but is dependent on confounding processes that cause a species to occur in unsuitable areas (e.g. buffer effects; source-sink dynamics) or to be absent from suitable areas (e.g. spatial or dispersal constraints; historical biogeography). These processes could mean that presences and absences are weakly associated with relative suitability, especially at intermediate suitabilities (Beale, Lennon & Gimona 2008). Where one seeks to identify the impact of climate on suitability, a good SDM should be able to take coarse presence-absence data and still return a good representation of the underlying relationship between climate and probability of occupancy. An alternative to metrics based on presence-absence data, therefore, which holds more promise for predicting to novel conditions, would be to identify models that perform well in identifying the true relationship between climate and probability of occupancy.
Understanding how climate dictates the probability of occurrence of a species is challenging, as the real relationship between climate and the probability of occurrence is unknown. Virtual species distributions can overcome some of these limitations and can be useful in understanding and improving modelling approaches (Hirzel, Helfer & Metral 2001; Meynard & Quinn 2007; Meynard & Kaplan 2013). A virtual species’ distribution can be created using pre-determined relationships between environmental gradients and the probability of occurrence of the species. Crucially, virtual distributions are characterised by a known relationship between climate and probability of occupancy and, thus, allow SDMs to be assessed in relation to their ability to recover that true relationship.

Here, I use five approaches to selecting absence data for use in SDMs for virtual species distributions (including random selection, selecting concentric cells around a species’ range, and several intermediate, distance-weighted, selection methods), and compare their abilities to produce models that can identify the underlying relationship between climate and probability of occupancy (which I refer to as “suitability”). The approaches vary in the average spatial extent from which absences are drawn. I determine the effect of this variation on SDM accuracy and predictive performance. To quantify the importance of absence selection approaches, I focus on answering three questions: (1) how does absence sampling using differing methods affect the predictive ability of models; (2) what is the best absence sampling strategy in terms of recreating the underlying climatic suitability for a species; and (3) once the best absence sampling method is chosen, does changing the number of absences affect the predictive ability of SDMs?
3.3 Methods

3.3.1 Creating virtual species distributions

I generated global distribution patterns at a half degree (approx. 50km$^2$ at the equator) resolution for 50 virtual species that differed in their response to three bioclimatic variables: mean temperature of the coldest month (MTCO), mean temperature of the warmest month (MTWM) and seasonality of precipitation (SOP). The variables chosen reflect the type of variables used in SDMs of taxa from different global biomes and represent gradients and extremes that limit species in the real world. Global climate data were obtained from Worldclim (Hijmans et al. 2005, http://www.worldclim.org/) at a 2.5′ resolution and were aggregated to a 0.5° resolution.

For each individual climate variable, climatic suitability for a pseudo species was defined by a Gaussian relationship with a randomly selected mean and standard deviation, taken from within the plausible global ranges (MTCO (-50.1 - 28.3), MTWM (-7.4 - 39.1) and SOP (0 - 236)). The climatic suitability ($C_j$) of cell $j$ was calculated by multiplying the probabilities linked to each variable:

$$C_j = f(MTCO) \times f(MTWM) \times f(SOP) + \varepsilon$$

where $f$ represents a Gaussian probability function that determines the probability of occurrence of a species for each environmental variable and $\varepsilon$ is a normally distributed random error. To ensure that each climatic factor had the same weight of influence on species’ occurrence, each variable was normalised to fall between 0 and 1.

Species seldom occupy all suitable climate space (owing to constraints on the colonisation probability for suitable but distant climate space). Consequently, virtual species distributions did not necessarily occupy all areas of suitable climate. Instead, a virtual distribution comprised a core range of proximate cells where the likelihood of occupancy declined with distance from the range core. The probability of occurrence ($P_j$) of a virtual species in cell $j$ was the product of two processes: climatic suitability and spatial constraints: $P_j = C_j \times S_j$. $P_j$ was normalised to fall between 0 and 1. A focal cell from among all high suitability cells (with $C_j > 0.7$) was selected at random, where only cells that were neighboured on all sides by high suitability cells were eligible to be chosen. This cell was denoted the range centre cell. A spatial occupancy function, $S_j$, was then calculated as:

$$S_j = \frac{1}{1 + \varepsilon^z(D_j - D_{half})}$$

where $D_j$ is the distance from cell $j$ to the range centre cell and $D_{half}$ is the distance from the centre at which $S_j = 0.5$. $z$ determines the steepness of the decline in occupancy. Varying $D_{half}$ and $z$ alters...
the number of cells defined as presences for a species. The combination of stochasticity in climate suitability relationships and stochasticity in selecting the range core lead to wide variation in range sizes, without the need to vary \( z \) and \( D_{\text{half}} \). Consequently, parameter values (\( z = 0.001, D_{\text{half}} = 1000 \)) that lead to a plausible and varied selection of range sizes were used, yielding a median range size that fell well within the distribution of known avian range sizes obtained from BirdLife International (2011, http://www.birdlife.org/datazone/info/spcdownload, see Figure S3).

Finally, to generate species’ presences from the probability of occurrence, whilst also allowing for some noise in occupancy (to mimic real distributions), each cell was allocated a random number \( Q_j \) drawn from a uniform distribution. Occupancy was assigned to cells for which \( P_j > Q_j \). This approach simulated absences that might occur due to factors not considered in the model such as biotic interactions or local extinctions. This process was repeated until 50 virtual species distributions were generated.

### 3.3.2 Absence selection

To investigate the effect of the absence selection method on model accuracy, pseudo-absences were drawn using five different methods (Figure 3.1). These methods differed in the rapidity with which the probability of choosing a random unoccupied cell as an absence decayed with increasing distance (\( D_e \)) from the nearest edge of a given virtual species’ distribution. Method (1), with a highly restricted spatial extent, involved selecting the closest absences from concentric shells (CS) around the species’ occupied range. In Methods (2) to (4) absences were selected randomly but with a probability that declined with distance from the range edge. Specifically, probability of selection was proportional to: \( \frac{1}{D_e^3} \) (Method 2); \( \frac{1}{D_e^2} \) (Method 3); and \( \frac{1}{D_e} \) (Method 4). In each case, proximate unoccupied cells were preferentially selected as absences. However, distance from the species’ range edge from which absences were selected increased from Method (2) through to Method (4), increasing the number of more distant absences selected. In Method (5) (AR), absences were selected randomly (without weighting) from across the world. To account for the variability arising from the random selection of absences in the last four cases, models were fitted with ten different sets of randomly selected absences for each virtual species. Initially, models were conditioned on an equal number of absences and presences.
Figure 3.1 Illustration of methods used to select absences for modeling. a) Absences were selected randomly. (b-d) Absences were selected randomly but with a decreasing number of distant absences selected. Specifically, probability of selection was proportional to b) \( \frac{1}{D^1} \); c) \( \frac{1}{D^2} \); and d) \( \frac{1}{D^3} \) across the globe; and e) where the closest absences were selected in concentric shells around the species’ occupied range.

3.3.3 Species Distribution Models

The relationship between each virtual species’ distribution and the three climate variables was modelled using two machine-learning techniques (random forests, RFs, and gradient boosted models, GBMs) and two regression methods (generalised linear models, GLMs, and generalised additive models, GAMs). Further details of the modelling and cross-validation approach are provided below. Models were conditioned on data including absences drawn from each of the five absence selection approaches.
The effect of spatial autocorrelation in the SDMs was minimised using a blocking approach, in which the transferability of fitted models to spatially segregated test data was assessed. These methods follow those of Bagchi et al. (2013). Global climate data were split into five spatially disaggregated blocks, so that each block sampled the full range of covariate parameter space, but the mean of the climatic predictors differed little between blocks. Sampling units consisted of global ecoregions (http://www.worldwildlife.org/science/data), or parts of ecoregions if the ecoregions were very large. Large ecoregions (greater than 250,000 km²) were split into smaller sampling units by intersecting them with a 2.5° by 2.5° grid to create smaller subunits of a comparable size to smaller ecoregions. Areas separated geographically can comprise the same ecoregion, and ecoregions that were not neighbouring were considered separate sampling units. The five blocks were created using the “blockTools” package in R (Moore 2014). For SDMs, a jack-knife approach was employed where each block was left out in turn (test data) and models were fitted to the remaining four blocks (training data) (Pearson et al. 2007). In this way, predictor variables and block (or geographic location) were independent (minimising the effect of spatial autocorrelation) but the range of predictor variables were similar in the training and testing data. The resultant model fit to the test data was evaluated by calculating the threshold independent area under a receiver operating characteristic (ROC) curve (AUC) (Fielding & Bell 1997) by taking suitabilities from all left out blocks and comparing those to the presence absence data.

SDM methods also followed the approach of Bagchi et al. (2013). In brief, GLMs with a binomial distribution and a logit-link were used to fit polynomial relationships between each species’ occurrence and three climatic variables: mean temperature of the coldest month (MTCO), mean temperature of the warmest month (MTWM) and seasonality of precipitation (SOP). These data were obtained from WorldClim (Hijmans et al. (2005), http://www.worldclim.org/). All combinations of first to fourth order polynomials were fitted (total of 81 combinations) leaving one block out at a time, and testing model accuracy on the withheld block using AUC. The combination of polynomials that resulted in the highest mean AUC across the five blocks was used to fit the final five models. GAMs with a Bernoulli response and a logit link were fitted using thin-plate regression spines (“mgcv” package R, Wood 2006) to species’ occurrence data excluding one block at a time, and smoothness was established by generalised cross-validation (Wood 2006). GBM model performance was optimised by fitting an initial model with 5000 trees, with tree complexities between 1 and 4, and a learning rate of 0.001 (“gbm” package R, Ridgeway 2013). Model fit was tested on withheld blocks for all tree sizes (tree size ranged between 999 and 5000). The combination tree complexity and tree number with the lowest summed error on average from across the five blocks was used to fit the final five models. For RF models, cross-validation was used to select both the number of
variables used to build each tree (mtry) and the optimal number of trees (ntree) (package “randomForest” in R, Liaw & Wiener 2002). An initial forest was built with 1000 trees for each value of mtry between 1 and 3 and the AUC was calculated using the withheld block. Another forest was then grown with additional 500 trees and the model accuracy assessed. This process was continued until the AUC did not improve by more than 1% when an extra 500 trees were added. The mtry value and the number of trees that maximised the AUC across the five blocks were used to assess model fit and to fit the final models.

3.3.4 Absence selection method comparison

a) Predicting occurrences of virtual species

Models based on the five absence selection methods were compared to determine which approach best predicted virtual species distributions (presence-absence) using a jack-knife approach. The resultant model fit to the test data was evaluated by calculating the threshold independent area under a receiver operating characteristic (ROC) curve (AUC) (Fielding & Bell 1997) by taking suitabilities from all left out blocks and comparing those to the presence absence data. AUC is known to be affected by prevalence (Lobo, Jiménez-Valverde & Real 2008) but that concern does not apply for scenarios where the prevalence was held constant, with number of absences equal to number of presences (but see “prevalence” section below). For the four absence-selection approaches that randomly select multiple absence datasets (i.e. all except the CS approach), variation in AUC across replicates was minimal (see Table 3.1), so the mean AUC value across the ten replicates was used in subsequent analyses. Linear mixed effects models (R package “lme4”, Bates et al. 2014) were used to describe AUC as a function of the absence selection method (fixed effect), with species included as a random effect. The impact of absence selection method on AUC was tested using likelihood ratio tests to assess the fit of models from which the fixed effect was omitted relative to those which included the fixed effect.

b) Predicting the underlying climatic suitability for virtual species

SDMs using absence data from each of the five absence selection approaches were projected to the globe to assess which approach best predicted the original climatic suitability for each species. Root Mean Squared Error (RMSE) was used to describe the extent of departures from the underlying suitability (Cj).

I was principally interested in how well a model reproduces suitable or marginally suitable areas that could become suitable in future, and wished to minimise any bias in model fit due to the predominance of very low suitability cells for most species globally. For this reason, I divided cells
into two categories to minimise zero-inflation, which could otherwise inflate apparent model performance. The categories were: (1) high suitability cells, classified as the 2n most suitable cells, where n is the number of cells occupied by the virtual species, and (2) all other cells (hereafter, “lower suitability” cells), the majority of which had suitability close to, or actually, zero. I do not completely disregard low suitability cells as I am also interested in how models perform at low suitability sites, because poor fitting models could under- or over-predict species distributions in these areas. For the two suitability categories, linear mixed effects models were used to determine the effect of method for selecting absences on the RMSE, including species as a random effect.

3.3.5 Prevalence

An unbalanced design (where numbers of absence points and presence points differ) can introduce bias and affect model performance (McPherson, Jetz & Rogers 2004). Analyses described above conform to the recommendation of equal numbers of absences and presences (Senay, Worner & Ikeda 2013). In addition, I also tested whether increasing the number of absences improved the ability of models to predict climate suitability, as has been suggested by others (Barbet-Massin et al. 2012). Specifically, I used the best performing absence selection method ($\frac{1}{D_e^2}$) to evaluate the performance of models where the number of absences was one, two, five and ten times the number of presences. In each case, the RMSE between each species’ underlying climatic suitability and predicted suitability was estimated for all model permutations. The importance of the ratio of absences:presences in affecting RMSE was evaluated using linear mixed effects models, with species included as a random effect.
3.4 Results

3.4.1 Comparison of absence selection methods

Fitted SDMs showed discriminative power ranging from excellent to poor predictions (Table 3.1). Absence selection method had a significant effect on AUC ($\chi^2 = 636.2$, $p < 0.0001$), with the AR approach for selecting absences performing best in terms of AUC in comparison with the other absence selection methods (Figure 3.2).

Table 3.1 Model performance at predicting the distribution of virtual species within each left out test block was evaluated using the area under the curve (AUC) of the receiver operating curve. Absences were those used to condition the model in each of the absence selection methods. Mean AUCs across 50 species x 10 repeats (once for concentric shells) varied by absence selection method and by SDM approach (GAM, GBM, GLM, RF).

<table>
<thead>
<tr>
<th>Method for selecting absences</th>
<th>GAM</th>
<th>GBM</th>
<th>GLM</th>
<th>RF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absences random</td>
<td>0.98 (0.02)</td>
<td>0.98 (0.02)</td>
<td>0.99 (0.01)</td>
<td>0.99 (0.002)</td>
</tr>
<tr>
<td>One over distance</td>
<td>0.83 (0.10)</td>
<td>0.83 (0.10)</td>
<td>0.83 (0.10)</td>
<td>0.98 (0.01)</td>
</tr>
<tr>
<td>One over distance squared</td>
<td>0.74 (0.09)</td>
<td>0.73 (0.08)</td>
<td>0.74 (0.09)</td>
<td>0.97 (0.01)</td>
</tr>
<tr>
<td>One over distance cubed</td>
<td>0.71 (0.08)</td>
<td>0.70 (0.08)</td>
<td>0.71 (0.08)</td>
<td>0.97 (0.02)</td>
</tr>
<tr>
<td>Concentric shells</td>
<td>0.67 (0.07)</td>
<td>0.66 (0.07)</td>
<td>0.67 (0.07)</td>
<td>0.97 (0.02)</td>
</tr>
</tbody>
</table>

3.4.2 Identifying the underlying climatic suitability

Absence selection methods had a significant effect on model ability to recreate the underlying climatic suitability; RMSE differed according to absence selection method for both the lower suitability ($\chi^2 = 78.7$, $p < 0.0001$) and high suitability ($\chi^2 = 118.6$, $p < 0.0001$) cells. CS was relatively poor at identifying suitability within the areas of lower suitability, whilst AR was relatively poor at identifying suitability within areas of high suitability (Figure 3.3; see, also, Figure 3.4 for an individual species example). The CS approach often overestimated the suitability of unsuitable areas (Figure 3.4g) and the AR approach overestimated the suitability of marginally suitable areas (Figure 3.4c). Intermediate absence selection methods had lower RMSE values in both high and low suitability cells, with the ($\frac{1}{D^2}$) having the lowest RMSE in both categories (Figure 3.3).
Figure 3.2 Predictive performance (median area under the receiver operating characteristic curve (AUC)) of the 50 virtual species occupied ranges for each method for selecting absences (AR, \( \frac{1}{D_e} \), \( \frac{1}{D_e^2} \), \( \frac{1}{D_e^3} \), and CS) across all four SDM methods (GAM, GBM, GLM and RF).
Figure 3.3 The root mean squared error (RMSE) between the original climatic suitability and the predicted suitability for (a) high suitability cells and (b) lower suitability cells. Median RMSE is shown across all virtual species for each of the five absence selection approaches: absence random (AR); one over distance (\(\frac{1}{D_e}\)); one over distance squared (\(\frac{1}{D_e^2}\)); one over distance cubed (\(\frac{1}{D_e^3}\)) and concentric shells (CS).
Figure 3.4 Example of a virtual species’ a) climatic suitability and b) resultant derived species distribution. Ensemble climatic suitability for the species from models created using five absence selection methods: c) AR, d) $\frac{1}{D_e}$, e) $\frac{1}{D_e^2}$, f) $\frac{1}{D_e^3}$ and g) CS.
3.4.3 Prevalence

Increasing the number of absence points by two-, five- and ten-times the number of presences decreased the ability of models to identify the actual climatic suitability (Figure 3.5). This reduced ability was significant for both the high ($\chi^2_3 = 88$, $p < 0.0001$) and lower suitability ($\chi^2_3 = 110.4$, $p < 0.0001$) categories.

Figure 3.5 The root mean squared error (RMSE) between the observed and predicted climatic suitability when models were conditioned with absence records equalling the number of presences (P=A), and when absences are twice (x2), five (x5) and ten times (x10) the number of presences. Presented results are median RMSE for 50 virtual species when data are split into two categories (a) high suitability and (b) lower suitability.
3.5 Discussion

The ability of models to identify species’ climatic suitabilities is strongly affected by the method used to select absences. This also impacts on their ability to discriminate between species’ presences and absences. These findings have important implications for species distribution modelling. In particular, they suggest that current approaches to selecting absences, using arbitrary boundaries such as biological realms or country boundaries, will often fail to provide a good description of a species’ climatic niche. I discuss these findings in the context of three issues: (1) the benefits of not delimiting a study extent; (2) the choice of absence selection method; and (3) the utility of the virtual species approach and its implications for assessing model performance.

3.5.1 Is it necessary to define a study extent?

Ideally, the area of a study should encompass the complete geographic distribution of a species, as well as, crucially, areas that are accessible to the species over a relevant time period (Soberon & Peterson 2005; Barve et al. 2011; Guisan et al. 2014). Further, ensuring that the available niche space is sampled adequately is important for model transferability, particularly if models are projected in space or time (Randin et al. 2006). However, delineating the study region is not straightforward, as it depends on knowledge of the natural history and the dispersal ability of a species, as well as information about the landscape of interest, including the configuration of suitable habitats and barriers to dispersal (Soberon & Peterson 2005; Barve et al. 2011). Defining this is particularly difficult for migrant species (that inhabit spatially and temporally distinct areas during their annual migratory cycle, and that have the ability to disperse large distances) and for invasive species, whose potential area for expansion is unknown.

My results show that distance weighting can provide a robust means of selecting absences. This approach renders an a priori definition of a geographic extent of study area unnecessary, and minimizes the loss of information imposed by predefined boundaries. Using distance weighted absence selection, absences are preferentially selected close to the species’ occupied distribution (as recommended by Jiménez-Valverde, Lobo and Hortal (2008) and Lobo, Jiménez-Valverde and Hortal (2010)), but are also sampled from further afield, to minimise problems of environmental truncation (Raes 2012). In multi-species modelling, distance weighting also reduces any bias towards better quality models for species occurring towards the centre of a study region (Jiguet et al. 2006; Jiguet et al. 2007).
Choosing an absence selection method

When models were assessed on their ability to define areas of lower suitability for species, those trained using CS performed worst. This is because using a restricted area for absence selection reduces the available environmental parameter space under which models are calibrated, in turn reducing their predictive power (Thuiller et al. 2004). This is particularly problematic in situations in which the aim is to project potential occupancy into novel areas. For example, where the aim is to predict the establishment of invasive species (Guisan et al. 2014) or to predict shifts in migratory destinations (Doswald et al. 2009), models trained on restricted environmental data are unlikely to be informative. Truncating the climatic limits of a species can also lead to the incorrect assumption that species’ ranges are not constrained by climate (Beale, Lennon & Gimona 2008), a conclusion that is greatly dependent on study extent (Jiménez-Valverde et al. 2011).

When models were assessed on their ability to define areas of high suitability for species, those trained using AR performed worst. Moreover, AR over-predicts the underlying climatic suitability in regions of intermediate suitability (Figure 3.4c). When absences are selected at random, they include a greater proportion of very distant cells. Such absences can include extremely environmentally distinct absence localities. Models trained on data covering very broad environmental ranges can successfully distinguish high and low suitability areas but fail to capture the nuances of suitability in important areas proximate to the occupied range (Lobo, Jiménez-Valverde & Hortal 2010).

In contrast to the extremes of AR and CS, distance weighted absence selection approaches provided a better balance of performance across the full range of suitabilities (Figure 2.2). Overall, I recommend the $\frac{1}{D_e}$ weighting metric, which performed best at identifying climatic suitability in both high and lower suitability areas (see Figure 3.3).

Virtual species and the assessment of model performance

Species occupancy does not necessarily equate to climatic suitability. The advantage of using virtual species is that it is possible to assess model performance by comparing the known and modelled climatic suitabilities for each species, regardless of occupancy. This is important because models with apparently high discrimination between presences and absences (assessed using AUC) failed to capture the underlying climatic relationships for species (see also, Anderson & Raza 2010; Acevedo et al. 2012). This limits the utility of models to predict potentially suitable areas beyond the current range, for example to facilitate assisted colonisation (Hoegh-Guldberg et al. 2008). Using
virtual species in this way further underlines the inadequacy of AUC to identify models suitable for all applications (Lobo, Jiménez-Valverde & Real 2008; Jiménez-Valverde 2012).

The ability of models to determine the original climatic suitability of each virtual species was not improved when the number of absences used for modelling increased. Previous studies have shown that an unbalanced design, where there are more pseudo-absences than presences affects SDMs differently; for example, increasing the number of pseudo-absences improved regression models, but not for models based on classification techniques (Barbet-Massin et al. 2012). To avoid inducing bias, and because increasing the number of absences did not significantly improve model prediction ability, I advocate using a balanced number of presences and absences.

Of course, species are not only limited by climate. Species’ niches are also constrained by biotic interactions, for example, and the relationships between these can be complex (Soberón 2007). More realistic SDMs will require a better understanding of these complex interactions. Nonetheless, correlative SDMs should be viewed as a first approximation of the direction and impacts of future climate change and, thus, methodological issues must be addressed.

3.5.4 Conclusion

Species distribution models are widely used to predict spatial patterns of biodiversity, and many studies have focused on methodological techniques to enhance performance (Elith et al. 2006; Pearson et al. 2006; Barbet-Massin & Jetz 2014). Despite this, the selection of an appropriate study area over which to select absences can be problematic and is a frequently overlooked aspect of methodology. Here, I propose the use of a distance weighted absence selection approach that avoids the need to use a predefined study extent. In particular, I recommend the \( \frac{1}{D^2} \) weighting metric that balanced the reproduction of climatic suitability across a range of suitabilities, and performed well at reproducing the distributions of virtual species.
Chapter 4

Seasonal niches in migratory birds: a biogeographic perspective

Wheatear, *Oenathe oenathe*, by Ian Fisher
4.1 Abstract

Climate change is a major driver of species distribution shifts, and understanding the relative importance of climatic variables on species distributions is necessary to understand how future climate change might impact species. Yet, to date, few studies have assessed the role of climate in determining migratory bird distributions on both their breeding and non-breeding ranges.

Here, I quantify the relative importance of four climatic variables in explaining the occurrence of migratory birds on their breeding and non-breeding ranges. I modelled the distribution (presence-absence) of 430 migratory species on their seasonal breeding and non-breeding grounds, and compared the climatic variables that are important determinants of occupancy in different parts of the world, and for breeding versus non-breeding ranges. The importance of climatic predictors for predicting migrant occupancy differed depending on season, and between migration flyways. Temperature generally tended to be a better predictor of occupancy than precipitation, although the importance of precipitation increased at lower latitudes. The importance of climatic variables for migrant distributions varied around the world, and by season, which implies that relationships of migrants with climate at a given location and time cannot be extrapolated globally.
4.2 Introduction

Recent climate change has altered the geographical ranges of many species (Parmesan & Yohe 2003; Root et al. 2003), and has caused biodiversity loss globally (Sala et al. 2000). Species distribution models (SDMs) are often used to project the impacts of climate change on species distributions, and consequently to inform conservation planning (Elith & Leathwick 2009; Wiens et al. 2009; Underwood, D’Agrosa & Gerber 2010; Carvalho et al. 2011; Araújo & Peterson 2012). Climate change has not affected all regions equally to date, with the greatest changes in temperature occurring towards the poles (IPCC 2013). Poleward shifts have been documented in a variety of species (Thomas & Lennon 1999; Root et al. 2003; Hickling et al. 2006), and the velocity of these shifts is rapid (Chen et al. 2011b). With regional differences in climate change expected in the future, understanding the distribution of species relative to current climatic conditions will help better understand how species distributions will likely respond to future climate change.

Avian migration spans the globe, with billions of individuals making predictable movements across large spatial scales in pursuit of improved foraging conditions and reproductive opportunities (Alerstam, Hedenstrom & Akesson 2003). This redistribution of species radically changes the communities and ecosystems they encounter (Somveille et al. 2013; Bauer & Hoye 2014). Many migratory species have declined in recent decades (Kirby et al. 2008), potentially in part because such species are difficult to conserve (Runge et al. 2014). Migratory species have responded to recent climatic change with both range and phenological shifts, the latter including temporal advances in their arrival at breeding grounds (Jonzén et al. 2006; Balbontin et al. 2009; Lehikoinen & Sparks 2010). Few studies have explicitly assessed the effect of climate change on migratory species on both the breeding and non-breeding ranges (but see Doswald et al. 2009). In order to understand how climate change could affect migratory species in the future, it is important to understand the current determinants of the occurrence of migrants.

Migrant birds track seasonal fluctuations in resources, and seasonality is an important determinant of migrant species richness (Hurlbert & Haskell 2003; Wisz, Walther & Rahbek 2007b). Seasonality in rainfall is thought to trigger the movement of Palaearctic migrants within Africa (Moreau 1972), as rainfall increases plant productivity and consequently insect abundance. Ambient energy, usually in the form of temperature, influences metabolic rates and the thermoregulatory needs of species (Currie 1991; Currie et al. 2004; Lemoine, Schaefer & Böhning-Gaese 2007; Field et al. 2009). On the European breeding grounds the percentage and number of migrants increases in areas of high winter harshness (Newton & Dale 1996; Lemoine & Böhning-Gaese 2003), with declines in the proportion of European migrants, in response to increasing winter temperatures.
(Schaefer, Jetz & Böhning-Gaese 2008). It is therefore likely that migratory species respond to changes in both precipitation and temperature.

A question which has received little attention is whether migratory species climatic preferences differ depending on whether they are breeding or not. Seasonal differences in habitat use are known, with migrants having broader habitat preferences in the non-breeding grounds (Rappole 2013; Blackburn & Cresswell 2015), yet the preference of consistent climatic regimes is less well studied. The Swainson’s Flycatcher (*Myiarchus swainsoni*), an austral migrant, has been shown to track a consistent temperature regime in its seasonal movements throughout the year (Joseph and Stockwell (2000)). This suggests seasonal movements of some species could be predictable based on their tracking a climatic niche. Similarly, a number of Nearctic-Neotropical migrants have been shown to track the same climate on their breeding and non-breeding ground, though others switched their climatic preferences between seasons (Nakazawa *et al.* (2004)).

The majority of studies that have assessed the relationship between climatic variables and migrant distributions have focused on the breeding ranges of species, mainly North America and Europe (Hurlbert & Haskell 2003; Lemoine & Böhning-Gaese 2003; Monkkonen & Forsman 2005; Wisz, Walther & Rahbek 2007b; Honkanen *et al.* 2010; Morrison *et al.* 2013; Blackburn & Cresswell 2015). Rapid declines in migrant birds have been observed (Kirby *et al.* 2008), therefore understanding current drivers of migratory species occupancy at a global scale is essential if we are to understand these declines, and to assess the potential for effective climate change adaptation (Pearce-Higgins & Gill 2010). Here, I assess if, and how, climatic variables vary in their importance for migrants across the world. I test whether species differ in their climatic requirements across geographical ranges and seasons, and also explore whether the important determinants of migratory species distributions vary amongst migration flyways.
4.3 Methods

4.3.1 Species distribution data

A global dataset of the distribution of terrestrial avian species (excluding coastal and marine species) was derived from BirdLife International (Birdlife International & NatureServe 2011). I extracted data for all fully migratory species (those with completely distinct breeding and non-breeding ranges) globally. The breeding and non-breeding ranges were overlaid to a grid of 0.5 degree cells (approx. 50km x 50km at the equator). A species was considered to be present in a cell if the species polygon intersected 10% of a cell. Species whose range occupied fewer than 30 cells on their breeding or non-breeding grounds were excluded from further analyses, which resulted in 430 species used in subsequent analyses. The globe was divided into three major flyways (based on BirdLife International (http://www.birdlife.org/flyways/)) as follows: The Americas flyway (120° to 30° W), the African-Eurasian flyway (30° W to 60°E) and the Asian-Australasian Flyway (60°E to 120°E) (see Figure S 4).

4.3.2 Bioclimatic data

Bioclimatic variables at a resolution of 2.5 min were obtained from WORLDCLIM v1.4 for the 50 year interval 1950-2000 (Hijmans et al. 2005, http://www.worldclim.org/). These data were aggregated to a 30’ (0.5° ≈ 50km x 50km) resolution to match the species data. I extracted data for four uncorrelated bioclimatic variables, which have previously been shown to successfully describe the range extents of birds: mean temperature of the warmest quarter (MTWQ); annual precipitation (AP); seasonality of precipitation (SP) and seasonality of temperature (ST) (Bagchi et al. 2013; Dalby et al. 2014; Somveille, Rodrigues & Manica 2015).

4.3.3 Species distribution models (SDMs)

The relationship between each species’ distribution and the four climate variables was modelled using random forests (RFs) (Breiman 2001; Prasad, Iverson & Liaw 2006a; Cutler et al. 2007). RFs are a machine learning approach that builds many regression trees from bootstrapped data subsets and random subsets of predictors. These regressions are then aggregated to provide an average prediction (Breiman 2001; Liaw & Wiener 2002). Models were conditioned on data including absences drawn using a distance weighted selection approach, as outlined in Chapter 2. Briefly, the probability of choosing a random unoccupied cell as in absence cell decayed with distance where the weighting was $w = 1$/distance$^2$, so that absences closer to the range were preferentially chosen.
The effect of spatial autocorrelation in the SDMs was minimised using a blocking approach (following Bagchi et al. 2013), in which the transferability of fitted models to spatially segregated test data was assessed. Global climate data were split into five spatially disaggregated blocks, so that each block sampled the full range of covariate parameter space, but the mean of the climatic predictors differed little between blocks (Bagchi et al. 2013). Blocks comprised multiple sampling units, where each unit was a global ecoregion (http://www.worldwildlife.org/science/data), or part of an ecoregion if an ecoregion was very large. Large ecoregions (greater than 250,000 km²) were split into smaller sampling units by intersecting them with a 2.5° x 2.5° grid to create smaller subunits of a comparable size to smaller ecoregions. Areas separated geographically can comprise the same ecoregion, but areas of the same ecoregion that were not neighbouring were considered separate sampling units for blocking. The five blocks were created using the “blockTools” package in R (Moore 2014). For SDM modelling, each block was left out in turn (test data) and models were fitted to the remaining four blocks (training data). In this way, predictor variables and block (or geographic location) were independent (therefore minimising the effect of spatial autocorrelation) but the range of predictor variables were similar in the training and testing data.

Cross-validation was used in the RF models, to select both the number of variables used to build each tree (mtry) and the optimal number of trees (ntree) (package “randomForest” in R; Liaw and Wiener (2002)). An initial forest was built with 1000 trees for each value of mtry between 1 and 3 and the AUC was calculated using the withheld block. Another forest was then grown with additional 500 trees and the model accuracy was assessed. This process was continued until the AUC did not improve by more than 1% when an extra 500 trees were added. The mtry value and the number of trees that maximised the AUC across the five blocks were used to assess model fit and to fit the final models. The median AUC from across the five blocks was used to assess model accuracy, and cross-validation was used to optimise model predictive performance before fitting a final (optimal) set of models for each species.

4.3.4 Variable importance

Random forest models randomly selected 1-3 (mtry) variables at each node within a tree to determine the best binary split to explain variation in migrant presence within a cell (Prasad, Iverson & Liaw 2006b). The importance of each predictor in determining a species’ distribution on both the breeding and non-breeding grounds was assessed by estimating the decrease in performance of the model when each predictor was randomly permuted (Liaw & Wiener 2002). Each block was left out in turn and random forest models were fit to the remaining blocks, and mean variable importance across all five blocks was calculated. Mean relative variable importance for each of the four
bioclimatic predictors was extracted for each species and mapped. Cells with fewer than five species were excluded from maps. Relative importance for each species was aggregated for the two temperature (MTWQ and seasonality of temperature) and two precipitation (annual and seasonal precipitation) variables. To enable comparisons among species, relative variable importance was calculated by dividing the importance of each individual variable by the summed importance across all variables for each species (Howard et al. 2015).

4.3.5 Statistical analyses

Paired-sample t-tests were conducted to compare the relative importance of amalgamated precipitation and temperature variables for migrants on the breeding grounds, with migrants on the non-breeding grounds. Within subject (to control for species) one way ANOVAs were conducted to compare the relative importance for predicting species occupancy of all four climate variables among the breeding and the non-breeding ranges, respectively. Paired-sample t-tests were conducted to determine whether each of the four climatic variables differed in relative importance for determining migrant occupancy on the breeding versus the non-breeding ranges. Four separate two-way ANOVA were conducted for each climate variable. These were used to determine whether the importance for predicting migrant occurrence for each climate variable differed among season (breeding and non-breeding), migration flyway, and whether there was an interaction effect of season and migration flyway on importance.

Differences in relative variable importance of climatic variables on species occupancy across latitudes were tested using the interaction between latitude and climatic variable-type. Latitude was the median latitude of each migratory species range (in degrees north or south of the equator). This analysis was repeated separately for each climatic variable, but including the additional effect of the interaction between season (breeding or non-breeding) and latitude.
4.4 Results

Aggregated temperature variables were significantly more important for predicting migrant occurrence than aggregated precipitation values on both the breeding (Figure 4.1; paired t-test, $t_{863}=219$, $P<0.001$) and non-breeding (Figure 4.1; paired t-test, $t_{863}=254$, $P<0.001$) areas.

**Figure 4.1** Mean relative importance (±SE) of precipitation and temperature variables in determining species occurrence for 430 species for aggregated variables on the breeding (a) and non-breeding (b) grounds. Notches are the 95% confidence intervals of the median. Non-overlapping notches indicate a significant difference at the 5% level.
Figure 4.2 Mean relative importance (±SE) of individual variables for predicting occurrence (mean temperature of the warmest quarter (MTWQ); annual precipitation (AP); seasonality of precipitation (SP) and seasonality of temperature (ST)) for 430 migrant species on the breeding (a) and non-breeding (b) grounds.

The relative importance of each of the four climatic predictors (MTWQ, AP, SP, ST) for predicting migrant occurrence for 430 species differed significantly on both the breeding (ANOVA: \( F_{(3,1724)} = 61.21, P<0.0001, \) Figure 4.2) and non-breeding grounds (ANOVA: \( F_{(3,1293)} = 90.51, P<0.0001, \) Figure 4.2). Both temperature variables (MTWQ and ST) had a greater impact on breeding migrant occurrence than did precipitation variables (AP and SP) (Tukey’s post-hoc analyses, \( P < 0.001 \) for both). For the non-breeding grounds, seasonality of temperature was the most important predictor of migrant occurrence, while annual precipitation was more important than MTWQ at predicting migrant occurrence (Tukey’s post hoc test, \( P < 0.001 \), for both).

The relative importance of climatic variables for predicting migrant occupancy differed significantly depending on whether migratory species are on the breeding or non-breeding grounds.
(ANOVA: $F_{(3,1257)} = 49.8, P < 0.001, \text{Figure 4.3}$). MTWQ was more important for predicting the occupancy of breeding migrants than non-breeding migrants, whereas seasonality of temperature was significantly more important for predicting the occupancy of non-breeding migrants. Annual precipitation was significantly more important for non-breeding migrants than breeding migrants, and there was no significant difference in terms of the importance of precipitation seasonality for breeding and non-breeding migrants (Tukey’s post hoc test, $P < 0.001$, for all, Figure 4.3).

**Figure 4.3** Comparison of the relative importance for (a) mean temperature of the warmest quarter, b) annual precipitation c) seasonality of temperature, and d) seasonality of precipitation for predicting migrant species occupancy on the breeding and non-breeding areas.
4.4.1 Spatial patterns in the importance of climate variables

There are distinct spatial patterns of the relative importance of the four climatic variables across the globe for predicting breeding (Figure 4.4) and non-breeding (Figure 4.5) migrant occurrence. The relative importance of MTWQ in predicting migrant occurrence differed significantly among season (breeding and non-breeding) and flyway (Figure 4.6, Table 4.1). The relative importance of AP in predicting migrant occurrence differed significantly among seasons but not across flyways (Figure 4.6, Table 4.1). There was a significant interaction effect between season and flyway in terms of the relative importance of SP for predicting migrant occurrence (Figure 4.6, Table 4.1), specifically SP is important for breeding migrants in the Asian-Australasian flyway (Figure 4.4) and for non-breeding migrants in the Americas flyway (Figure 4.5 Figure 4.6). Finally, the importance of ST differed significantly among seasons, but was not significantly different among flyways (Figure 4.6, Table 4.1).
**Figure 4.4** Relative importance of each ecological predictor on migrant species occurrence on the breeding grounds as measured by a drop in regression accuracy after predictor removal from random forest models: a) mean temperature of the warmest quarter (MTWQ), b) annual precipitation (AP), c) seasonality of temperature (ST), d) seasonality of precipitation (SP).
Figure 4.5 Relative importance of each ecological predictor on migrant species occurrence on the non-breeding grounds as measured by a drop in regression accuracy after predictor removal from random forest models a) mean temperature of the warmest quarter (MTWQ), b) annual precipitation (AP), c) seasonality of temperature (ST), d) seasonality of precipitation (SP).
Table 4.1 Summaries of the results of four two-way analyses of variance (ANOVA). Four individual two-way ANOVAs were used to assess the drivers of spatial patterns in the importance of MTWQ, AP, SP and ST in determining the occurrence of migratory species within each grid cell. P-Values significant at the 5% level are shown in bold. Season is a two level factor (Breeding or non-breeding) and flyway is a three level factor (African-Eurasian, Americas, Asian-Australasian).

<table>
<thead>
<tr>
<th>Response Variable</th>
<th>Explanatory Variables</th>
<th>Degrees of freedom</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTWQ</td>
<td>Season</td>
<td>1</td>
<td>371.9</td>
<td>371.9</td>
<td>47.4</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>Flyway</td>
<td>2</td>
<td>1.8</td>
<td>0.9</td>
<td>0.112</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>Season x Flyway</td>
<td>2</td>
<td>73.5</td>
<td>36.8</td>
<td>1.706</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>Residuals</td>
<td>239</td>
<td>1875</td>
<td>7.8</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>AP</td>
<td>Season</td>
<td>1</td>
<td>46.0</td>
<td>46.5</td>
<td>9.41</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>Flyway</td>
<td>2</td>
<td>2.8</td>
<td>1.40</td>
<td>0.29</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>Season x Flyway</td>
<td>2</td>
<td>25.8</td>
<td>12.90</td>
<td>2.64</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>Residuals</td>
<td>239</td>
<td>1169.4</td>
<td>4.89</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SP</td>
<td>Season</td>
<td>1</td>
<td>9.4</td>
<td>9.39</td>
<td>1.964</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>Flyway</td>
<td>2</td>
<td>5.9</td>
<td>2.97</td>
<td>0.62</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>Season x Flyway</td>
<td>2</td>
<td>88.7</td>
<td>44.33</td>
<td>9.27</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>Residuals</td>
<td>239</td>
<td>1142.9</td>
<td>4.78</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>ST</td>
<td>Season</td>
<td>1</td>
<td>89.0</td>
<td>88.98</td>
<td>17.74</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td></td>
<td>Flyway</td>
<td>2</td>
<td>0.0</td>
<td>0.02</td>
<td>0.004</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>Season x Flyway</td>
<td>2</td>
<td>29.5</td>
<td>14.75</td>
<td>2.94</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Residuals</td>
<td>239</td>
<td>1198.5</td>
<td>5.01</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>
Figure 4.6 Comparison of the relative importance for mean temperature of the warmest quarter (MTWQ), annual precipitation (AP), seasonality of precipitation (SP) and seasonality of temperature (ST) for predicting migrant species occupancy on the breeding and non-breeding areas, compared across three flyways (a) Americas, (b) African-Eurasian and (c) Asian-Australasian.
The relative importance of MTWQ across latitude differed significantly between breeding and non-breeding migrants ($F_{(1,851)}=94.8, P<0.001$), with the importance of MTWQ increasing with latitude for predicting breeding migrants. MTWQ was important for predicting occupancy of non-breeding migrants in the temperate regions, but not in the tropical and boreal regions (Figure 4.7a). The importance of AP for predicting migrant occurrence differs marginally between the breeding and non-breeding seasons, with the relative importance of AP being high for predicting breeding migrant occupancy in the tropics ($F_{(1,851)}=5.39, P=0.02$, Figure 4.7b), but this importance diminishes in the temperate and boreal regions. AP was not an important predictor for non-breeding migrants and this did not vary with latitude (Figure 4.7b). The relative importance of ST for predicting migrant occupancy differed significantly across latitude for both breeding and non-breeding migrants ($F_{(1,851)}=63.86, P<0.01$) where importance of ST was more important for breeding migrants at high latitudes, but was most important for non-breeding migrants in the tropics and temperate areas (Figure 4.7c). Finally, there was no difference in the relative importance of SP for predicting occupancy of breeding and non-breeding migrants across latitude, with importance for migrants declining with latitude for both seasons in temperate and tropical areas ($F_{(1,851)}=0.32, p=0.5$, Figure 4.7d).
Figure 4.7 Latitudinal gradients (both hemispheres combined) of the relative importance of a) mean temperature of the warmest quarter (MTWQ), b) annual precipitation (AP), c) seasonality of temperature (ST) and d) seasonality of precipitation (SP) on migrant breeding (red) and non-breeding (blue) occurrence for 430 species. Symbols are mean importance values for species when binned into 10° categories (± standard error). The latitudes are categorised by the overall biome that they encompass: tropical, temperate and boreal.
4.5 Discussion

I have identified that the distributions of migratory species across the globe are highly variable in space in response to climatic variables. In general, temperature was more important than precipitation for predicting both breeding and non-breeding migrant distributions globally. Species responses to climate change are often characterised solely on temperature change (e.g. Parmesan & Yohe 2003; Thomas 2010), perhaps because biological responses to temperature are more easily understood than they are for precipitation, and therefore precipitation is often excluded from analyses that investigate the poleward shift in species distributions in response to temperature increases (Root et al. 2003). However, when considered simultaneously, evidence points to a range-limiting role for both moisture and temperature (Smith 2013; VanDerWal et al. 2013) which may result in multi-directional distribution shifts globally (VanDerWal et al. 2013).

The relative importance of climatic variables on migrant species occupancy varied with latitude, with precipitation being more influential at low latitudes, where both the timing and the absolute amount of rainfall in tropical wintering areas can have major impacts on non-breeding season performance of migratory birds (Studds & Marra 2007). The inter-annual variability of rainfall in the non-breeding regions often has cascading effects on plant productivity, arthropod abundance and therefore the condition and survival of birds (Szep et al. 2006; Studds & Marra 2007). Conversely, temperature was a more important driver of migrant occupancy at higher latitudes. This agrees with previous studies that have shown that migratory species are more likely to be temperature limited than water limited (Hawkins et al. 2003; Hickling et al. 2006; Huntley et al. 2007; Pearce-Higgins & Green 2014).

The high relative importance of temperature at higher latitudes and of precipitation at lower latitudes is consistent with the water-energy hypothesis, where the key drivers for species richness across the globe switch from moisture availability toward the equator to energy-related towards the poles (Hawkins et al. 2003; Whittaker, Nogués-Bravo & Araújo 2007). Assuming that the combination of individual species occupancy predictions allows for the prediction of species richness in an area (stacked SDMs Dubuis et al. 2011), climatic drivers of species occupancy are likely to be similar to those of species richness. This water-energy latitudinal pattern has been shown in a European-wide study of the relative importance of climate and land-use for birds (Howard et al. 2015), and matches latitudinal gradients in bird population responses to both temperature and precipitation (Pearce-Higgins & Green 2014; Pearce-Higgins et al. 2015).

The majority of studies of climate change impacts on biodiversity focus on temperature, and are generally conducted at higher latitudes (Both et al. 2006; Hickling et al. 2006; Sherry et al. 2007;
Devictor et al. 2008). Given that temperature is a more important predictor of migrant species occurrence than precipitation at higher latitudes (in particular for the breeding areas), the findings from these studies may not necessarily be applicable at lower latitudes, where precipitation is a more important driver of migratory species occurrence (Pearce-Higgins et al. 2015). By failing to account for the effects of precipitation at lower latitudes, a focus on temperature for finding a fingerprint of climate change, may underestimate the sensitivity of such species to climate change (VanDerWal et al. 2013). Crucially, that temperature is a more important predictor of migratory species distributions than precipitation (Tayleur et al. 2015) is important for future climate change predictions, given that the highest uncertainty for future climates is in forecasting precipitation (McSweeney et al. 2015).

The importance of seasonality of temperature and annual precipitation for migratory species occurrence did not differ among the three migration flyways. Mean temperature of the warmest quarter was more important for Nearctic-Neotropical migrants than Afro-European or Asian-Australasian migrants. Since temperature change is resulting in shifts in species distributions (Root et al. 2003) these migrants may be more affected by climate change than their Asian-Australasian or African-Eurasian counterparts. Seasonality of precipitation was more important for predicting occupancy of Asian-Australasian migrants than Americas or African-Eurasian migrants. Seasonality of precipitation has been shown to have a greater influence than temperature on the timing of migration for Australian migrants (Chambers 2008). My results show that species differ geographically in their climatic preferences, and that the relationships between climate and occupancy of migrants in one region cannot necessarily be extrapolated to other geographic regions.

Some species show overlaps between breeding and non-breeding climatic preferences (Martínez-Meyer, Peterson & Navarro-Sigüenza 2004; Nakazawa et al. 2004) but not others (Nakazawa et al. 2004; Laube, Graham & Böhning-Gaese 2015). Our results suggest that migratory species are (in general) not closely tracking their preferred climatic niche, as the relative importance of variables differed for migrants between the breeding and non-breeding areas. Migrants may move between environments that are optimal at different stages in their life cycle. The advantage of migrating may not be to follow optimal climatic conditions, but instead due to higher reproductive success and lower nest predation at higher latitudes during the breeding season, and lower winter mortality at lower latitudes during the non-breeding season (Bohning-Gaese et al. 2000; McKinnon et al. 2010).

A limitation of this study is that quantifying niche space at the species level fails to capture the finer scale relationships that individuals or populations may have with the environment (Bolnick et
Non-uniform distributions of individuals within a species’ range means that species may not respond to climate in the same way across its range. Narrower geographic extents than the broad flyway extent employed here might explain importance of climatic factors for migratory species at a finer scale, as it is obvious that the importance of variables for migrant occupancy varied geographically within some flyways, e.g. MTWQ within the Americas flyway. Moreover, there are other factors than climate that limit migrant distributions such as land-use, biotic interactions, genetic constraints on migration routes, geographic barriers, habitat, productivity, all which may have smaller scale impacts on migrant distributions which have not been explicitly assessed here (Alerstam, Hedenstrom & Akesson 2003; Jetz, Wilcove & Dobson 2007; McKinnon et al. 2010; Ockendon et al. 2014; Chudzińska et al. 2015).

Our results suggest that migrants may not favour the same climatic conditions on the breeding areas as the non-breeding areas, as has been shown by other studies (Doswald et al. 2009; Laube, Graham & Böhning-Gaese 2015). Migratory species may be responding to differing climatic conditions depending on season (whether they are breeding or not), therefore a new suite of SDMs that reflect temporal changes in niche requirements, instead of annual means are needed to improve our understanding of the importance of climatic variables for migratory species occurrence (Heikkinen, Luoto & Virkkala 2006). Additionally, if species’ niche availabilities are altered through climate change, then existing migratory strategies may no longer be available (Laube, Graham & Böhning-Gaese 2015). Exploring how migratory species will respond to the emergence of novel climates, or suitable climate space in novel locations, will be crucial for predicting how migratory species will respond to future climate change (Williams & Jackson 2007; Reside, VanDerWal & Kutt 2012). Our results highlight that migratory species respond to differing variables on their breeding and non-breeding ranges, and therefore assuming simple niche tracking when predicting future climate change shifts may be over simplistic.
Chapter 5

A novel approach for estimating avian migration distances

Swallow, *Hirundo rustica*, by Ian Fisher
5.1 Abstract

Migratory species may be particularly vulnerable to climate change, with the potential for intricately timed and long-established relationships of migratory species' with their environment to become disrupted. Breeding and non-breeding ranges of migratory species may shift in response to climate change, and in some cases may move apart, potentially leading to future declines. Conversely, some migrants might benefit from shorter migrations, and such shifts are already being observed. Migration distance is also likely to affect the arrival time of migrant birds on the breeding grounds, which has been linked to survival and reproductive success. Estimating migratory distances is therefore important to determine any costs or benefits to migration in the future.

To date, predictions of migratory changes under future climate scenarios have usually estimated the distance between breeding and non-breeding range centroids. However, this approach ignores the variation in migratory movements within species. Here, I develop a method to estimate the range of potential migration distances for species. Using correlative species distribution models (SDMs), I predict current landscape suitability using contemporary climate and species distribution data on their breeding and non-breeding ranges. I predict the range of migratory distances a species might undertake by combining the locality of high quality migratory end-points with travel-distance data, and I compare the predictions to recorded distances between migratory start and end points for an example species. I find that the method I developed estimated shorter migration distances than the centroid approach, and that these distances fitted observed distance estimations well for the example species.
5.2 Introduction

The persistence of species in the face of climate change depends on the ability of populations to keep pace with shifting climates, or adapt to changes in situ (Burrows et al. 2011). Shifts in latitude or elevation of the distributions of species in response to climate change have been widely reported in recent decades (Thomas & Lennon 1999; Parmesan & Yohe 2003; Root et al. 2003; Thomas 2010). Concurrently, breeding bird populations have been declining across Europe, with long distance migrants being particularly affected (Sanderson et al. 2006). Given that approximately 20% of the 9856 extant avian species migrate seasonally (BirdLife International 2008a; Kirby et al. 2008), surprisingly few studies have investigated the impact of changing climate on migratory species across their migration routes and at stopovers (Tottrup et al. 2008). The majority of climate change studies have focused on the breeding distributions of migratory birds, finding that some migrants have advanced their arrival dates to the breeding grounds in a pattern consistent with climate change (Jonzén et al. 2006; Balbontin et al. 2009; Lehikoinen & Sparks 2010). Few studies have assessed the potential impacts of climate change on the non-breeding grounds of migrants (but see Austin & Rehfisch 2005; Studds & Marra 2007; Barbet-Massin et al. 2009), and fewer still have assessed the impact of climate change on migratory species on both their breeding and non-breeding ranges (Bohning-Gaese & Lemoine 2004; Doswald et al. 2009; Wilson et al. 2011).

Doswald et al. (2009) demonstrated that, although potential range extent varied among a group of trans-Saharan migrants, Sylvia warblers, in general, both breeding and non-breeding ranges were predicted to increase in extent given climate change projections for the end of the 21st Century. Migration distances were projected to increase, and in many cases novel potential future non-breeding areas were simulated. Indeed, as a result of such changes, birds may need to develop new migration strategies, for instance, by increasing the number or duration of stopovers used, in order to obtain sufficient energy for longer journeys (Schaub & Jenni 2001).

The flight route chosen by migrants determines the total distance travelled, as well as the potential for encountering favourable stopover sites all which influences the overall energy and time needed for migration (Liechti 2006). Global climate change has led to warmer winters in North-West Europe, and studies have shown that migration distances have decreased for short distance migrants (Visser et al. 2009). These changes in migratory distance are consistent with predictions from climate change, whereby non-breeding ranges are shifting northward (Visser et al. 2009; Pulido & Berthold 2010; Kullberg et al. 2015), therefore migrants winter closer to breeding ranges. The resulting shortened migration distance is likely to modify the timing of spring arrival of migrants to the breeding grounds (Coppack & Pulido 2004; Pulido 2007; Pulido & Berthold 2010), which means these
species can better time their arrival to breeding areas to the timing of food availability (Coppack & Both 2002; Walther, Berger & Sykes 2005). Current climate change is favouring birds wintering closer to the breeding grounds as it reduces migration costs and facilitates the rapid adaptation to shifts in the timing of food emergence on the breeding grounds (Visser & Both 2005; Bradshaw & Holzapfel 2008). Furthermore, an artificial selection experiment in a population of blackcaps (Sylvia atricapilla) showed that residency will rapidly evolve in migratory populations if selection for shorter migration persists (Pulido & Berthold 2010).

In order to understand bird movements it is necessary to acquire data of where they are traveling. Migration distances can be calculated using remote sensing techniques (such as by fitting satellite tags or geolocators to birds) and ring recovery data (Green et al. 2002; Visser et al. 2009; Egevang et al. 2010; Robinson et al. 2010). Bird ringing means birds can be marked individually and reported later, providing evidence of their movements. Ringing data present a challenge for analysing migration because the probability of finding and recording a ringed bird varies geographically, resulting in non-random sampling (Korner-Nievergelt, Liechti & Thorup 2014) and therefore these data will be spatially biased. Existing tracking techniques are subject to a trade-off between weight and precision, where heavier but more precise geolocators can only be applied to birds of larger body masses, and lighter geolocators which can be applied to smaller birds but can have large measurement errors (Lisovski et al. 2012). New technology offers solutions to the spatial bias from bird ringing, but has limited application (with regards to species) and is currently non extensive in terms of tracking birds from the entirety of their range (and therefore still has the issue of spatial biases). Both satellite tagging and recovery data are essential for the study of migration, yet acquiring these data can be difficult, and these data can be spatially biased.

Species distribution models (SDMs) are widely used in ecology for identifying species’ habitat preferences, and for predicting how the suitability of habitats might change in space or time. SDMs have found particular utility in predicting species’ potential responses to climate change (Huntley et al. 2008; Doswald et al. 2009; Barbet-Massin, Thuiller & Jiguet 2012). Migration distances have been estimated from the distance between centroids of predicted breeding and non-breeding ranges (Doswald et al. 2009). These distance estimations often necessitate that the probabilities of occurrence estimates from SDMs are transformed to a binary presence/absence form, which means losing detailed information of climatic suitability for a species. Additionally, the choice of thresholding method is a source of uncertainty in the SDM process, which has been shown to drastically alter estimates of range shifts (Liu et al. 2005a; Nenzén & Araújo 2011). Finally, and perhaps most importantly, centroid-centroid distance estimates convey little nuanced detail of
migration at present, or how migrations might change in the future. While the usefulness of
centroids for analysing shifts in distances of species range shifts is evident (e.g. Lyons, Wagner &
Dzikiewicz 2010), centroids do not account for the fact that species ranges are variable in shape, and
may therefore not detect shifts that occur at range margins. Moreover, migrant species distributions
are not necessarily continuous, with the same species breeding on disparate continents or countries,
often making centroid estimates nonsensical.

The centroid-centroid approach for estimating migratory distance could be improved in several
ways: First, an approach that does not threshold modelled suitabilities avoids the uncertainty in
threshold selection and, importantly can account for the fact that climatic suitability varies across
the landscape. Second, a method that considers populations of individuals from across a species
breeding/non-breeding range. Third, a method that is biologically meaningful, whereby the
estimation of migration distances occurred between cells that are climatically suitable for a species,
and which takes into account the cost of migration distance (Somveille, Rodrigues & Manica 2015)
would be more appropriate than the centroid-centroid approach, which gives little information of
the underlying processes which might alter migration distance.

Several other factors affect the distribution of species, such as species-specific dispersal
ability, competition, resource availability and learned behaviour (Guisan & Thuiller 2005; Soberón
2007; Schloss, Nuñez & Lawler 2012a; Early & Sax 2014). Failing to account for such effects can result
in the overestimation of species range extents. When these factors are not explicitly considered,
projected species distributions have been constrained using biogeographic realms (Pigot, Owens &
Orme 2010), country boundaries (Acevedo et al. 2012) and distance buffers (Young et al. 2009). A
migrant bird that breeds in North America may have suitable climate space available in other parts
of the world (such as Europe), but given that a lot of migratory behaviour is under genetic control, it
is unlikely that migrants will shift their breeding/non-breeding areas to such an extent that they
switch continents altogether. As climate change will mean long distance migrants (that mostly
originate in Nearctic, Palaearctic) will be shifting their breeding ranges poleward (Hickling et al.
2006), they are unlikely to begin breeding in a new realm. Conversely, they could quite easily
become more prone to residency or migrate shorter distances if climatic conditions improve closer
to the breeding ranges, and switch realm with respect to the non-breeding range (as have Blackcaps
recently, Pulido and Berthold (2010)). Delineating an appropriate region for model projections for
migrants is not straightforward; as they inhabit spatially and temporally distinct areas during their
annual migratory cycle, and have the ability to disperse large distances.
Here I first develop SDMs that relate migratory breeding and non-breeding ranges to current climatic conditions. I then evaluate two methods for estimating the migration distance a typical individual of a species might undertake. The first is a commonly applied approach using SDMs to describe the distribution on both the breeding and non-breeding range, and calculating the centroid to centroid distance. The second is a distance weighted estimation approach whereby distance is calculated between climatically suitable cells on the breeding and non-breeding areas for a given species, favouring distances that are shorter.

5.3 Methods

5.3.1 Species distribution data

Breeding and non-breeding range polygons for 440 fully migratory bird species (those with fully non-overlapping breeding and non-breeding ranges), were obtained from BirdLife International (Birdlife International & NatureServe 2011). These polygons were overlaid to a 0.5 degree (approx. 56km x 56km at the equator) grid. A species was considered to be present in a grid-cell if the species’ polygon intersected the cell by 10% or more. From the initial 440 species, I excluded species that spent a significant portion of their time at sea because their occurrence is unlikely to be linked to terrestrial climate. Due to model building limitations, particularly when using data splitting for model validation, species that occupied fewer than 30 cells on their breeding or non-breeding grounds were excluded from further analyses. After this process, 340 species of long-distance migrant remained and were used for subsequent analyses.

5.3.2 Climate data

Bioclimatic variables for a 50-year interval (1950-2000) were obtained from WORLDCLIM v1.4 (Hijmans et al. 2005, http://www.worldclim.org/) at a 0.04° resolution. These data were aggregated to a 0.5° resolution to match the species data by calculating the mean value of all the 0.04° cells that fall within each 0.5 degree cell. Four uncorrelated bioclimatic variables were chosen \textit{a priori} for analyses: mean temperature of the warmest quarter (MTWQ); annual precipitation; seasonality of precipitation and seasonality of temperature. These variables can limit the distribution of species through both direct and indirect effects on vegetation, predation, and competition, and have been used successfully to describe the species distribution patterns of birds globally (Hurlbert & Haskell 2003; Huntley \textit{et al.} 2006; Abolafya \textit{et al.} 2013; Howard \textit{et al.} 2015; Somveille, Rodrigues & Manica 2015).
5.3.3 Species distribution models

The relationship between each species’ distribution and the four bioclimatic variables was modelled separately for the breeding and non-breeding ranges, using two modelling approaches: a machine-learning technique (random forests, RFs) and a regression method (generalised additive models, GAMs). Models were conditioned on presence-absence data, with presences coming from the gridded Birdlife polygons and absences drawn using the distance weighted absence selection approach, outlined in Chapter 2. The effect of spatial autocorrelation in SDMs was minimised using a blocking approach (following Bagchi et al. 2013), in which the transferability of fitted models to spatially segregated test data was assessed. Global climate data were split into five spatially disaggregated blocks, such that each block sampled the full range of covariate parameter space, but the mean of the climatic predictors differed little between blocks (Bagchi et al. 2013). Sampling units consisted of global ecoregions (http://www.worldwildlife.org/science/data), or parts of ecoregions if the ecoregions were very large. Large ecoregions (greater than 250,000 km²) were split into smaller sampling units by intersecting them with a 2.5° by 2.5° grid to create smaller subunits of a comparable size to smaller ecoregions. Areas separated geographically can comprise the same ecoregion, and ecoregions that were not neighbouring were considered separate sampling units. The five blocks were created using the “blockTools” package in R (Moore 2014). For SDMs, each block was left out in turn (test data) and models were fitted to the remaining four blocks (training data). In this way, predictor variables and block (or geographic location) were independent (therefore minimising the effect of spatial autocorrelation) but the range of predictor variables were similar in the training and testing data.

GAMs with a Bernoulli response and a logit link were fitted using thin-plate regression spines (“mgcv” package R, Wood (2006)) to species occurrence data excluding one block at a time, and smoothness was established by generalised cross-validation (Wood (2006)). For RF models, cross-validation was used to select both the number of variables used to build each tree (mtry) and the optimal number of trees (ntree) (package “randomForest” in R, Liaw and Wiener (2002)). An initial forest was built with 1000 trees for each value of mtry between 1 and 3 and the AUC was calculated using the withheld block. Another forest was then grown with additional 500 trees and the model accuracy assessed. This process was continued until the AUC did not improve by more than 1% when an extra 500 trees were added. The mtry value and the number of trees that maximised the AUC across the five blocks were used to assess model fit and to fit the final models. For both modelling approaches, the median AUC from across the five blocks was used to assess model accuracy, and cross-validation was used to optimise model predictive performance before fitting a final (optimal) set of models for each species.
5.3.4 Model projection

Areas suitable for occupancy during both the breeding and non-breeding season were determined by applying the SDMs for the breeding/non-breeding season across a species’ entire current migratory flyway. This allows areas that become newly suitable for spending the non-breeding season, but which occur between the breeding and non-breeding range, to be used in future, i.e. to allow for the development of new migratory strategies. The world was divided into three major migratory flyways (based on BirdLife international (http://www.birdlife.org/flyways/)) as follows: The Americas flyway (120°W to 30°W), the African-Eurasian flyway (30°W to 60°E) and the Asian-Australasian Flyway (60°E to 120°E), which encompass the migratory routes of the majority of terrestrial migrants. For wide ranging species that utilise several flyways, species distributions were modelled and projected separately to each of the flyways in which they occurred. For a small number of species (76 species) whose migration did not follow these general flyways, individual flyways (and hence combinations of regions to which models were applied) were established. For example, the Amur falcon (*Falco amurensis*) breeds in eastern Asia but spends the winter-in sub-Saharan Africa. For this species, the potential future breeding and non-breeding range was evaluated across Africa and eastern Asia.

5.3.5 Estimating migratory distances

Centroid-to-centroid migration distance

Continuous suitability data (from both the breeding and non-breeding range models) were converted to presence-absence using a thresholding approach that maximised the Kappa statistic. The Kappa statistic measures model accuracy while correcting for accuracy expected to occur by chance (Cohen 1960). Vincenty’s ellipsoid great circle distances (Vincenty 1975) were used to estimate the distance between the centroids of the breeding and non-breeding range for each species. Centroids were derived by taking the mean latitude and longitude of each migratory species’ breeding and non-breeding distribution.

Distance-weighted migratory distances

As an alternative to the centroid-centroid approach, I used an approach that estimated migration distances from localities drawn from across the breeding range. First, a cell was randomly selected from the modelled breeding range of a species (i.e. the range after applying a threshold, as described above), and 100 cells were also randomly selected from within the non-breeding areas. I then used a ‘least-cost’ approach to find the most suitable non-breeding cell to the focal breeding cell, from among the 100 candidate non-breeding locations. The non-breeding cell was selected
based on $P_{occ}$ (modelled suitability from the SDMS) weighted by the great circle distance between the two cells ($D_{min}$, where shorter distances are favoured), as follows:

$$C = \max_{i=1}^{100} \left( \frac{P_{occ,i}}{D_{min,i}} \right) \quad \text{[1]}$$

where $C$ is the cell chosen. The distance between the chosen cell and the breeding range cell recorded. This process was repeated 1000 times for each migratory species. The median of these 1000 distance estimations was recorded, to account for rare long distances that could skew the mean. Subsampling a series of non-breeding cells (100 cells per iteration) was necessary, as to estimate distances between each breeding cell and all non-breeding cells for all iterations and species, would be computationally prohibitive. For widespread migrants that use multiple flyways, a median migratory distance was recorded for each flyway. This process was conducted for all migrants. Using this approach, a range of migration distances were selected for each species, as well as the median distance.

This method for estimating migration distance was evaluated using an example species, the barn swallow ($Hirundo rustica$). This wide ranging species was selected as ringing recovery data were available, in particular for sub-Saharan Africa, in comparison to other trans-Saharan migrants. Ringing recovery data were obtained from three ringing and migration atlases (Britain and Ireland, Norway and Finnish (Wernham et al. 2002; Bakken, Runde & Tjorve 2006; Valkama et al. 2014). The ringing recovery data were African-European in extent, therefore the polygon data were clipped to the African-European flyway (Error! Reference source not found.).

The centroid-centroid and the distance weighted distribution approaches for estimating migration distance were compared, for all migratory species using paired t-tests. Both methods of estimating migration distance were also compared with ring recovery distance estimates for the example species.
Figure 5.1 BirdLife polygon data for the Barn swallow (*Hirundo rustica*), on the breeding range (red) and the non-breeding range (blue). The polygon data are constrained to the African-Eurasian migration flyway. Black points are the ringing and recovery data for Barn swallow obtained from the three European migration atlases (see text for details).
5.4 Results

The red-breasted flycatcher (*F. parva*) example illustrates that using centroid-centroid methods to estimate migration distances is flawed. Despite the non-breeding grounds of *F. parva* occurring in Pakistan and India (Figure S 6), this species is projected to have suitable climate in two disparate regions on the non-breeding range (Figure 5.2). This novel area of climatic suitability is an area where the species is categorised as “origin uncertain” by BirdLife International, that is, the species provenance in an area is not known (it may be native, reintroduced or introduced) (Figure S 6). The centroid-centroid approach for estimating migration distances is obsolete in this scenario, as it is unable to adapt to this, and instead choses a point in between the two non-breeding areas as its destination point, when in reality, the species is not projected to occupy this area in the non-breeding season. In contrast, the distance weighted method for estimating migratory distances shows a wide range of distances estimated for this species, to a number of disparate, climatically suitable locations (Figure 5.2b).
Figure 5.2 Predicted baseline breeding (red) and non-breeding (blue) distribution of the red-breasted flycatcher (*F. parva*). a) Arrows indicate the locations chosen to estimate distance using the centroid-centroid method. b) Frequency histogram of distances estimated using the distance weighted estimation method in blue. Dashed line indicates the median distance for this species (3980 km) and dotted line is the distance estimated using centroid to centroid distance (3383 km).
5.4.1 *Hirundo rustica* example

Median migration distances for the Europe-African migratory barn swallows were estimated to be 6131 km when the centroid-centroid approach was used, and 5740 km when the distance-weighted estimation approach was used.

Histograms of migratory distances (km) estimated using the distance weighted distribution approach were overlaid with histograms of distances estimated from ring recovery data (Figure 5.1). Ring recovery distance estimates fell within the distance estimates from polygon data using the distance weighted estimation approach. Distances calculated using the distance weighted method were slightly left skewed (Figure 5.3).

![Frequency histogram of the distances (in km) estimated using the distance weighted distribution approach (grey) the ring recovery data (blue). The ring recovery distance estimates exclude South Africa as it contained the majority of ringing recoveries, and therefore distances became right-skewed.](image)

**Figure 5.3** Frequency histogram of the distances (in km) estimated using the distance weighted distribution approach (grey) the ring recovery data (blue). The ring recovery distance estimates exclude South Africa as it contained the majority of ringing recoveries, and therefore distances became right-skewed.
5.4.2 Distance estimates across all migratory species

Centroid-centroid estimates of migratory distance are significantly larger than estimates from the distance weighted distribution approach (paired $t_{333} = -7.58$, $p < 0.0001$) for 334 migratory species (Figure 5.4).

![Box plot of migration distances](image)

**Figure 5.4** Migration distances (±SE) between the breeding and non-breeding ranges for 334 migratory species calculated using the centroid approach and the distance weighted approach. Notches indicate the 95% confidence intervals of the median, with a lack of overlap indicating a significant difference at the 5% level.

Migration distances were significantly larger for migrants that breed at higher latitudes (centroid-centroid: $F=76.9_{(1,318)}, P < 0.0001$ and distance weighted: $F=158.5_{(1,318)}, P < 0.0001$). Distances estimated using distance weighting were significantly larger for African-Eurasian migrants than for migrants of the Americas and Asian-Australasian flyways ($F_{(2,317)}=4.62$, $P = 0.01$). Centroid-centroid distance estimates were significantly larger for migrants of the Americas flyway than for African-Eurasian and Asian-Australasian migrants ($F_{(3,316)}=47.56$, $P < 0.001$). Mapped mean migration distance for migrants shows that distance is larger for more northern species for both methods for estimating distance (Figure 5.5a and b), but distance estimates were larger for North American migrants when the centroid-centroid approach is used than when the distance weighted distance estimate approach was employed.
Figure 5.5 Mean migration distance (km) for all breeding migrants based on a) the distance weighted estimation approach and b) the centroid-centroid approach for estimating distance.
5.5 Discussion

In order to estimate migration distances, an approach that incorporated both climatic suitability data for a species, as well as a cost of migration distance, was developed. Migration distances between breeding and non-breeding ranges were estimated more frequently between cells that were both climatically suitable and closer to one another. This approach also returned a range of likely migration distances that species would face rather than single estimate of mean migration distance. Finally, the advantage of this method over the centroid-centroid approach was that this approach was less rigid than the centroid approach, and allowed for the estimation of distances to novel climatically suitable locations that were predicted closer to the breeding areas of a given species.

The centroid-centroid method of estimating migration distances is a simple approach that has been used to estimate migratory distances for birds (doswald et al. 2009; la sorte et al. 2013). However, the centroid approach for estimating migration distance only estimates distance between the means of a species breeding/non-breeding range, and does not consider distances faced by individuals from across the range. This is particularly important as studies are reporting that species are shifting their range margins more than the centre of a range (mason et al. 2015), and therefore changes in migration distances may be underestimated. Additionally, the centroid-centroid approach is inflexible and can produce illogical migration distances when species’ have multiple breeding and non-breeding locations, as was illustrated by the example species F. parva. A further issue with this method is that in order to estimate distance, it is necessary to threshold the probabilistic climate suitability data to convert them to a binary presence-absence format. Thresholding has been criticized for being dependent on prevalence, and, importantly, for the arbitrary choice of threshold value which has been shown to drastically alter projections of species distributions (allouche, tsoar & kadmon 2006; nenzen & Araújo 2011).

The distance-weighted estimation approach provided a range of potential migratory distances for a species, as opposed to the centroid approach which is one value of mean migration distance. The distance weighted method is a more nuanced approach and was more flexible than the centroid-centroid method of estimating distances, as it allowed for the selection of novel climatically suitable non-breeding areas, if they were closer to the breeding range than the observed non-breeding areas. Distances were calculated in a more biologically meaningful way, whereby the majority of distances were estimated between climatically suitable cells in the breeding grounds and the non-breeding grounds, where I assume species are more likely to occur. Additionally, the cost of distance was incorporated to this method of estimating distance, whereby closer, climatically
suitable non-breeding areas were preferentially selected for calculating migration distance. Finally, the distance weighted approach maximised the use of the data by avoiding conversion of the data to a binary presence-absence format.

Across all species, migration distances were estimated to be larger when using the centroid-centroid approach than the distance-weighted estimation approach to calculate migration distance. This is likely to be as a result of the latter approach allowing for distances to be calculated to closer, climatically suitable locations. Migration distances were larger for migrants that breed at higher latitudes. In general, species that breed at the highest latitudes have the longest migration distances between the breeding and non-breeding ranges (Alerstam, Hedenstrom & Akesson 2003). The most extreme example of this is the Arctic tern (Sterna paradisaea) which breeds in the Arctic and spends the non-breeding season in the Antarctic (Rappole 2013). Migration distances were also larger for Afro-European migrants than for migrants of the Americas or Asian-Australasian flyway. This may be because of the Saharan desert barrier which requires a long distance crossing of this for migrants to reach the productive areas south of the Sahel (Moreau 1972).

When the distance weighted approach for estimating distances from polygon data was compared with ring recovery data estimates, it was apparent that the ring recovery data for H. rustica was biased in space. The majority of the recovery data for barn swallows in Africa were from South Africa where there is a ringing scheme, “The South African Bird Ringing Scheme”. This highlights the issue of non-random samples of ringing recovery data, where the probability of finding and recording a ringing recovery varies geographically (Korner-Nievergelt, Liechti & Thorup 2014).

Migratory species occupancy at a given site is dependent on the climatic suitability of the area (Guisan & Zimmermann 2000a), as well as the distance between the non-breeding and breeding grounds (Duijns et al. 2012). The suitability of the habitat is also related to the number of conspecifics and competitors, as well as the productivity and quality of the site (Ramos et al. 2015). Therefore, migratory species make complex decisions when selecting their habitat on both the breeding and non-breeding grounds. These additional considerations have not been explicitly tested here, but incorporating these effects would improve the realism of estimated migration distances.

An issue with the distance weighted distribution approach for estimating migration distance is that the density of individuals in a given cell was not taken into account. A cell is chosen for estimating distance based on the climatic suitability for a species and distance to the breeding range, but ignores the density of potential conspecific or competitors already in a given cell. Therefore, a cell may be repeatedly selected to estimate distance that would be unsuitable for a migrant (despite the suitable climate) if it was already at full of competitors and conspecifics. Finally, for some species,
suitable climate space was available for a migratory species closer to the breeding grounds that are currently unoccupied. This could be because despite the suitable climate space, the area is unsuitable for reasons not explicitly tested in this study (habitat homogeneity, competition, predators, inter-annual climate variability (Jetz & Rahbek 2002; Alerstam, Hedenstrom & Akesson 2003; Ahola et al. 2007; Allouche et al. 2012)). Further, species that make multiple stops on the non-breeding grounds might be simulated to winter too close to the breeding areas using the distance-weighted estimation approach, as they would be modelled to go to the nearest suitable sites, and not to move on when they become unsuitable.

The selection of suitable, closer, yet currently unoccupied non-breeding ranges could be of interest, as it could highlight novel areas for future overwintering, as migration is costly and shorter migration distances are likely to be selected for (Pulido & Berthold 2010). Furthermore, it could highlight areas where species are expected to occur but do not for other reasons such as competitive exclusion (Leathwick & Austin 2001; Meier et al. 2011), or a lack of necessary habitat (Torres et al. 2015) for the species. These additional contributors to the distribution of a species must be considered when using SDMs for conservation planning, because species distributions are not solely constrained by climate (Guisan & Thuiller 2005).

5.4.1 Conclusion

The distance-weighted estimation approach was chosen to estimate migration distances of species under future climate change (Chapter 6). This approach negates the need to threshold data and incorporated suitability estimates, whereby suitable cell-cell distances were more likely to be estimated. Moreover, a cost of migration distance was incorporated, whereby further distances were less likely to be selected for the distance estimation process. Finally, this approach was more flexible and allowed for the estimation of distances to novel, closer, yet climatically suitable areas that may become inhabited by migrants in the future.
Chapter 6

Long distance migrants face increased migration distances in response to climate change

Willow warbler, Phylloscopus trochilus, by Ian Fisher
6.1 Abstract

Migratory species are particularly vulnerable to climate change, with the potential for intricately timed and long-established relationships of migratory species' with their environment to become disrupted. Breeding and non-breeding ranges of migratory species may shift in response to climate change, and in some cases may move apart, potentially leading to future declines. Conversely, some migrants might benefit from shorter migrations, and such shifts are already being observed. Migration distance is also likely to affect the arrival time of migrant birds on the breeding grounds, which has been linked to survival and reproductive success. Projecting potential changes to migratory strategies in the future is important to understand additional costs or benefits to migratory species as a consequence of climatic change.

Correlative species distribution models were used to relate migratory species distributions across the globe, on breeding and non-breeding areas, to contemporary climates. Future potential distributions of migratory species on their breeding and non-breeding grounds were projected using an ensemble of future climate change scenarios. Distances between simulated potential breeding and non-breeding ranges were compared for individual species, using current and future climate projections, and changes in potential migratory distances in the future were evaluated.

In general, distances between the breeding and non-breeding areas are projected to increase for migrants in the future. This is largely a consequence of breeding ranges shifting poleward, while non-breeding ranges show little latitudinal shift in a given direction. Some novel potential non-breeding locations are simulated closer to the breeding areas, highlighting possible novel non-breeding sites for migrants in the future. Breeding species turnover was highest at higher latitudes, as was the change in breeding migrant species richness between now and the future. Important breeding sites for migrants are projected to become even more northerly in the future. Although there was variation in projected migratory patterns due to variable climate change projections, climatic suitability consistently improved for migrants on the breeding areas across all three global migration flyways.
6.2 Introduction

Anthropogenic climate change and ongoing human-induced threats, such as habitat loss, pose major threats to global biodiversity (Walther et al. 2002; Thomas et al. 2004; Brook, Sodhi & Bradshaw 2008; Pereira et al. 2010). Understanding species’ responses to climate change is one of the most pressing scientific challenges. Correlative species distribution models (SDMs) have been developed to assess the potential impacts of climate change on biodiversity (Peterson et al. 2011; Guisan et al. 2013). Climatic change is driving poleward shifts in species’ ranges for many taxa (Parmesan & Yohe 2003; Hickling et al. 2006; La Sorte & Thompson 2007) but, more recently, attention has shifted to the variation in responses among species (Eglington & Pearce-Higgins 2012; Gillings, Balmer & Fuller 2015; Palmer et al. 2015).

Migratory species may be particularly vulnerable to climate change, with the potential for intricately timed and long-established relationships between migratory species’ and their environments to become disrupted. Breeding bird populations have been declining across Europe, in particular long-distance migrants (Sanderson et al. 2006). Recent studies have reported changes in the spring migration of birds (Knudsen et al. 2011), with the advancement of arrival dates to the breeding grounds in a pattern consistent with climate change (Sparks 1999; Cotton 2003; Thorup, Tøttrup & Rahbek 2007; Saino et al. 2011). A change in climate may lead to a mismatch between timing of migration and resource availability on both the breeding and non-breeding areas (Both & Visser 2001; Both et al. 2006; Gordo 2007). Arctic breeders have been shown to have reductions in reproductive success or complete breeding failure if they arrive at their breeding grounds outside a very narrow time window (Alerstam & Lindstrom 1990; Bauer, Gienapp & Madsen 2008). Natural selection should favour birds that can synchronise migration with peaks of resource availability, particularly as migrating birds use chains of stopover sites en route, which often differ in seasonality (Bauer, Gienapp & Madsen 2008). Long-distance migrants are particularly vulnerable to mismatches because climate in their breeding and wintering areas may be changing at different speeds, hampering appropriate adaptation (Both & Visser 2001). Further, short distance migrants spend the winter closer to the breeding areas in comparison with long-distance migrants, which may allow such species to fine-tune the timing of spring migration in response to climate (Rubolini et al. 2007; Moller, Rubolini & Lehikoinen 2008). Indeed, the breeding populations of long-distance migrant birds are already in decline (Both et al. 2006; Sanderson et al. 2006; Heldbjerg & Fox 2008; Moller, Rubolini & Lehikoinen 2008).

Arrival date can be advanced in several ways: by increasing speed of migration, by departing the wintering grounds earlier, and by shortening the migration distance (Coppack & Both 2002). Global
climate change has led to increased winter temperatures in North-West Europe, and studies have shown that migration distances have decreased for short distance migrants (Visser et al. 2009). These changes in migratory distance are as a result of non-breeding ranges shifting northward (Visser et al. 2009; Pulido & Berthold 2010; Kullberg et al. 2015), resulting in migrants wintering closer to breeding areas. The resulting shortened migration distance is likely to modify the timing of spring arrival of migrants to the breeding grounds (Coppack & Pulido 2004; Pulido 2007; Pulido & Berthold 2010), which means these species can better time their arrival to breeding areas to the timing of food availability (Coppack & Both 2002; Walther, Berger & Sykes 2005). Current climate change is favouring birds wintering closer to the breeding grounds as it reduces migration costs and facilitates the rapid adaptation to shifts in the timing of food emergence on the breeding grounds (Visser & Both 2005; Bradshaw & Holzapfel 2008). Under climate change, migrants may become more sedentary as warmer climates allow for species to overwinter on the breeding grounds (Berthold 2001). An artificial selection experiment in a population of blackcaps (Sylvia atricapilla) suggested that residency could rapidly evolve in migratory populations if selection for shortened migration persists (Pulido & Berthold 2010). However, the observed shorter migration distances are inconsistent with what is expected from climate change, where migration distances are expected to increase because breeding ranges (often located at higher latitudes) respond strongly to latitudinal climatic gradients, but non-breeding ranges (often located in tropical regions) do not (Huntley et al. 2006; Doswald et al. 2009).

SDMs are used widely in ecology for identifying species’ habitat preferences, and for predicting how the suitability of habitats might change in space or time. SDMs have found particular utility in predicting species’ potential responses to climate change (Huntley et al. 2008; Doswald et al. 2009; Barbet-Massin, Thuiller & Jiguet 2012). Migration distances have been estimated from the distance between centroids of predicted breeding and non-breeding ranges (Doswald et al. 2009). Additionally, species range shifts are often calculated by estimating the distance between the centroid of the current species range with the centroid of the projected future or past species range (Lyons, Wagner & Dzikiewicz 2010; Gillings, Balmer & Fuller 2015). Importantly, centroid-centroid distance estimates convey little nuanced detail of how migrations might change in the future. Centroid shifts do take into consideration the shape of a species range, and therefore shifts may not be detected if they occur at range margins. Moreover, migrant species distributions are not necessarily continuous, with the same species breeding on disparate continents or countries, often making centroid estimates nonsensical.
Few studies have assessed the potential impacts of climate change on the non-breeding grounds of migrants (but see Austin & Rehfisch 2005; Studds & Marra 2007; Barbet-Massin et al. 2009), and fewer still have assessed the impact of climate change on migratory species on both their breeding and non-breeding ranges (Bohning-Gaese & Lemoine 2004; Doswald et al. 2009; Wilson et al. 2011). Doswald et al. (2009) showed that, although potential range extent varied among Sylvia warblers, a group of trans-Saharan migrants, in general, both breeding and non-breeding ranges were projected to increase in extent by the end of the 21st Century. Migration distances were projected to increase, and in many cases novel potential future non-breeding areas were simulated, which suggested that new migration routes may need to develop in response to climate change. Certainly, as a result of such changes, birds may need to develop new migration strategies, for instance, by increasing the number or duration of stopovers used, in order to obtain sufficient energy for longer journeys (Schaub & Jenni 2001).

Most studies of the effects of climate change migrant distributions have focused on the breeding ranges of species, and are mainly North America and Europe focused (Hurlbert & Haskell 2003; Lemoine & Bohning-Gaese 2003; Monkkonen & Forsman 2005; Wisz, Walther & Rahbek 2007b; Honkanen et al. 2010; Morrison et al. 2013; Blackburn & Cresswell 2015). However, it is clear that there are important links between the periods that migrants spend on the often widely separated breeding and non-breeding locations (Marra, Hobson & Holmes 1998; Gill et al. 2001; Gordo et al. 2005; Morrison et al. 2013). Without a solid understanding of the year-round geographic distributions of migrant species, long term conservation plans become difficult (Webster et al. 2002; Small-Lorenz et al. 2013). Despite this, there has been no (to my knowledge) previous assessment of the impacts of projected climate change on migratory species, on both the breeding and non-breeding seasons, at a global scale.

Here I assess the potential impacts of climate change on long-distance migratory birds across the globe using an ensemble of future climate change projections. I model the relationship between contemporary species distributions and climate, using correlative SDMs. From this the spatial and temporal patterns of projected climate change impacts for migratory species are assessed globally. Changes in migration distance are evaluated, allowing for species to develop new migratory strategies should novel climatically suitable areas become available closer to the breeding grounds in the future.
6.3 Methods

6.3.1 Species distribution data

Breeding and non-breeding range polygons for 443 fully migratory bird species (those with fully non-overlapping breeding and non-breeding ranges), were obtained from BirdLife International (Birdlife International & NatureServe 2011). These polygons were overlaid to a 0.5 degree (approx. 56km x 56km at the equator) grid. A species was considered to be present in a grid-cell if the species’ polygon intersected the cell by 10% or more. From the initial 443 species, I excluded species that spent a significant portion of their time at sea because their occurrence is unlikely to be linked to terrestrial climate. Due to model building limitations, particularly when using data splitting for model validation, species that occupied fewer than 30 cells on their breeding or non-breeding grounds were excluded from further analyses. After this process, 340 species of long-distance migrant remained, and were used for subsequent analyses.

6.3.2 Contemporary climatic data

Bioclimatic variables for a 50-year interval (1950-2000), were obtained from WorlClim v1.4 (Hijmans et al. 2005, http://www.worldclim.org/) at a 0.04° resolution. These data were aggregated to a 0.5° resolution to match the species data by calculating the mean value of all the 0.04° cells that fell within each 0.5 degree cell. Four non-correlated bioclimatic variables were chosen a priori for analyses as described in chapter 4: mean temperature of the warmest quarter; annual precipitation; seasonality of precipitation and seasonality of temperature.

6.3.3 Future climate projections

Future projections of bioclimatic variables were obtained from WorldClim (Hijmans et al. 2005) for two time periods: 2050 (average for 2041-2060) and 2070 (average for 2061-2080). Three generalised circulation models from the IPCC’S CMIP5 project (GCMs: CCSM4, HadGEM2-ES, MIROC-ESM-CHEM) for four representative concentration pathways (RCPs) were used. These GCMs were selected as they contained all four available RCP scenarios for both time periods. These data were obtained at a 0.04° and aggregated to a 0.5° resolution to match the species data, as was done for the baseline climate data. The four RCP scenarios (RCP2.6 [sometimes referred to as RCP3PD], RCP4.5, RCP6 and RCP8.5) were developed for the IPCC fifth assessment report (IPCC 2013) based on the fifth phase of the Coupled Model Inter-comparison Project5 (CMIP5, http://cmip-pcmdi.llnl.gov/cmip5/). These RCP scenarios incorporate the complex interactions of climate systems, ecosystems (such as land use and land cover change), and human activities (such as the emergence of new technologies, and socioeconomic development) to provide plausible descriptions
of how the future might unfold (Moss et al. 2010; van Vuuren et al. 2011; Rogelj, Meinshausen & Knutti 2012). The four RCPs (RCP2.6, RCP4.5, RCP6, and RCP8.5) are named after a possible range of radiative forcing values in 2100 relative to pre-industrial values (+2.6, +4.5, +6.0 and +8.5 W/m², respectively). RCP8.5 can be interpreted as a high emissions scenario, and RCP2.6 as the lowest emissions scenario (van Vuuren et al. 2011).

6.3.4 Species distribution models

The relationship between each species’ current distribution and the four bioclimatic variables was modelled separately for the breeding and non-breeding ranges, using two modelling approaches: a machine-learning technique (random forests, RFs) and a regression method (generalised additive models, GAMs) as explained in chapter 5. Models were conditioned on presence-absence data, with presences coming from the gridded BirdLife polygons (Birdlife International & NatureServe 2011) and absences drawn using the distance weighted absence selection approach, outlined in Chapter 2. The effect of spatial autocorrelation in SDMs was minimised using a blocking approach (following Bagchi et al. 2013), in which the transferability of fitted models to spatially segregated test data was assessed, as described in chapters 4 and 5. In total, there were 100 models for each species (2 SDMs X 10 pseudo-absence repetitions X 5 jackknife iterations).

Models calibrated using baseline data were then projected to the future using future climate change projections for the two time periods (2050 and 2070) for each of the four RCP scenarios using climate change predictions from 3 GCMs. To avoid projecting migratory species distributions to unrealistically distant locations, the SDMs were applied to the migration flyway that each species currently occupies. The world was divided into three major migratory flyways (as defined by BirdLife International: http://www.birdlife.org/flyways/) as follows: The Americas flyway (120°W to 30°W), the African-Eurasian flyway (30°W to 60°E) and the Asian-Australasian Flyway (60°E to 120°E), which encompass the migratory routes of the majority of terrestrial migrants (Figure S 4). This approach retained as suitable any areas that became climatically suitable but which occur between the breeding and non-breeding range within a migration flyway. For wide-ranging species that utilise several flyways, species distributions were modelled and projected separately to each of the flyways in which they occurred. A small number of species (76 species) follow migratory routes that do not match with the three major flyways. For such species, individual flyways (and hence combinations of regions to which models were applied) were established. For example, the Amur falcon (Falco amurensis) breeds in eastern Asia but spends the winter-in sub-Saharan Africa. For this species, the
potential future breeding and non-breeding range was evaluated across the African-Eurasian and part of the Asian-Australasian flyways.

6.3.5 Estimating climatic suitability for migrants globally

Models were used to predict the probability that a given cell would contain suitable climate for each migratory species using present climate and 24 (4 RCP scenarios X 3 GCMs X 2 time periods) future climate change projections. For each species, change in the climatic suitability was the summed climatic suitability across the migration flyway for the future period, divided by the summed current suitability for that species across the same flyway. This was calculated separately for each species on the breeding and non-breeding grounds. This projected change in suitability was calculated for each of the future climate change projections, and the 95% quantiles were used to assess uncertainty. For each cell across the globe, where the change in climatic suitability was >1, a species was projected to gain suitability, and where it was <1, a species was projected to lose climatic suitability. When the 95% quantiles for each species change in climatic suitability estimates overlapped with 0, these projections are less certain than projections which showed directional consensus in suitability change across the different GCMs, SDM methods and jack-knife iterations.

6.3.6 Species richness

The projected species richness for migrants was estimated (separately for breeding and non-breeding richness) by summing the probabilities that each of the species modelled would find suitable climate within a given cell in the migration flyway within which it occurs. Projected change in species richness between the present and future periods was calculated as a proportional change relative to current projected richness.

To test whether change in species richness was related to latitude, the world was divided into 6 longitudinal and 12 latitudinal bands to minimise the effect of spatial autocorrelation. The mean change in species richness was determined for each of the 108 blocks. An ANOVA was used to determine if the estimated mean species richness change for each block (breeding and non-breeding separately) differed among the longitudinal bands (Figure S 7). The longitudinal bands were: band 1= [70 to 80° latitude]; band 2= [50 to 70° latitude]; band 3= [30 to 50° latitude]; band 4= [0 to 30° latitude]; band 5= [-30 to 0° latitude]; and band 6= [-50 to -30° latitude].
6.3.7 Species turnover

Species turnover for each cell between the current and future time periods (2050, 2070) was calculated using the Bray-Curtis index of dissimilarity between two communities (Bray & Curtis 1957), as:

\[
T_j[t] = \frac{\sum_{k=1}^{s} |P_{jk}[t_2] - P_{jk}[t_1]|}{\sum_{k=1}^{s} |P_{jk}[t_1]| + \sum_{k=1}^{s} P_{jk}[t_2]}
\]

where \(t_1=\) baseline, \(t_2=\) future, \(j=\) species turnover for each cell between \(t_1\) and \(t_2\), and \(P_{jk}\) is the mean suitability of species \(k\) within a cell across the GCM projections. Species turnover was estimated separately for the breeding and non-breeding areas, and for the four RCP scenarios.

6.3.8 Migration distance

To estimate typical potential migration distances, the distance-weighted suitability approach outlined in Chapter 4 was used. First, continuous suitability data (from both the breeding and non-breeding range models) were converted to presence-absence projections, using a thresholding approach that maximised the Kappa statistic (Cohen 1960). The Kappa statistic measures model accuracy while correcting for accuracy expected to occur by chance (Cohen 1960). Second, a cell was randomly selected from the modelled breeding range of a species (i.e. the range after applying a threshold, as described above), weighted by the suitability of a cell, and 100 cells were randomly selected from within the modelled non-breeding areas. A ‘least-cost’ approach was used to find the most suitable non-breeding cell to the focal breeding cell, from among the 100 candidate non-breeding locations. The non-breeding cell was selected based on \(P_{OCC}\) (modelled climatic suitability from the SDMs) weighted by the great circle distance (km) between the two selected cells \(D_{min}\), where shorter distances are favoured, as follows:

\[
C = \max_{i=1}^{100} \left( \frac{P_{OCC,i}}{D_{min,i}} \right)
\]

This distance between the chosen non-breeding cell and the breeding range cell was recorded. This process was repeated 1000 times for each migratory species, the range of migratory distances recorded, and the median distance calculated. Subsampling a series of non-breeding cells (100 cells per iteration) was undertaken because to estimate distances between each breeding cell and all non-breeding cells for all iterations and species, would be computationally prohibitive. Secondly, this approach introduced stochasticity, which prevented a single suitable non-breeding
site close to the breeding grounds being selected in all iterations. For widespread migrants that use multiple flyways, a median migratory distance was recorded for each flyway.

6.3.9 Change in migration distance

Potential changes in median migration distance were expressed as the percentage change in distance between present and potential future median migration distances. Should novel, climatically suitable locations emerge for migrants in the future, then the range of potential migration distances could alter. Therefore, potential changes in the range of migration distances between present and potential future migration ranges were also estimated. Median migration distances across the three time periods (baseline, 2050, 2070) were compared using a repeated measure ANOVA (to account for the fact that the same species were found in each time period). The percentage change in median migration distance and the percentage change in the range of migratory distances between 2050 and 2070 were compared using paired t-tests (to control for species).

To assess the potential for migrants to select novel breeding or non-breeding destinations in the future, histograms of the frequency for which a given migration distance was selected across the 1000 replicates, were produced for each species. I test for a change in the modality of number of core non-breeding localities, and their distances from the breeding range. For instance, the shape of the histogram could change from a unimodal shape to a bi-modal distribution.

Using these distance frequency distributions (see, for example, Figure 6.1), parametric bootstraps were used to sequentially test the number of components (or modes) in a mixture model framework for each species (MixTools package, R, Benaglia et al. 2009). A mixture model is a probabilistic model that is used to represent the presence of sub-populations within an overall population, when the subpopulations have not been specifically identified in the data set. Gaussian mixture models were used to estimate the species-specific number of predicted migratory modes based on distance estimations. Parametric bootstraps were used to test (using the likelihood ratio statistic) whether a k-component fit, versus the alternative hypothesis of a (k+1)-component fits the distribution best. Testing ended once the p-value was above the significance level of P=0.05 (Young 2007). The overall change in the number of migration modes for all species across the three migration flyways were compared using a Chi-square test.
Figure 6.1 Simulated data were used to illustrate of the type of data used to assess number of migration modes. Frequency histogram shows the distances (1000 replicates) estimated between the chosen climatically suitable cell on the breeding ground and the climatically suitable, closer, cell chosen on the non-breeding range. For assessing the number of modes, the test of $k=1$ versus $k=2$ components showed that two components were significantly better than 1 ($\alpha=0.05$). However, when comparing $k=2$ versus $k=3$ components, $P=0.19$ indicating that there was no statistically significant evidence for selecting $k>2$. Given this, I assume that for this example, there are two migration modes.
6.4 Results

Species distribution models for predicted occurrence of 340 migratory species across the breeding and non-breeding areas showed good discriminatory power (breeding AUC: median = 0.94, min= 0.77, max=1.00) and non-breeding AUC: median=0.94, min=0.79, max=1.00). AUC values >0.9 reflect very good discrimination and those >0.7 show useful discrimination (Swets 1988).

Presented results are projections using the RCP8.5 projection unless otherwise stated. This is the most extreme climate change projection, but overall patterns did not differ greatly among RCP projections. I chose to present the most extreme projection to illustrate the worst-case scenario for migrants in the future.

6.4.1 Change in climatic suitability

Projected impacts of climate change on the representation of suitable climate for 334 long-distance migrants across the globe showed that migrants were expected to experience a gain in climatically suitable space of 56.3%, (95% CI [14.6, 90.5]) on their breeding areas by 2070, when using the RCP8.5 projection. There was considerable variation among migration flyways, with a similar proportion of migrants on the African-Eurasian flyway losing (Figure 6.2, median= 49.5%, (95% CI [10.2, 83.2])) and gaining (median=50.4%, (95%[16.8, 89.7])) suitable climate space on both breeding and non-breeding areas. In contrast, Asian-Australasian migrants were projected to lose 63.5%, 95% CI [2, 94.8] of their suitable climatic space in the non-breeding areas by the end of the 21st Century, while losing 44.8%, (95% CI [7, 91.7]) of suitable climate space on the breeding areas. Migrants in the Americas were projected to experience increased availability of suitable climatic space on both the breeding (median=62.4%, (95% CI [16.3, 92.4])) and non-breeding (median=65.6%, (95% CI [23, 95])) areas. By 2070 there was only consensus on whether species were projected to experience increased or decreased climatic suitability for fewer than 25% of species across the breeding range and 20% of species across non-breeding areas. A greater proportion of these species are likely to experience increased climatic suitability across the breeding and non-breeding ranges (Table 6.1). Non-breeding climatic suitability is more consistently diminished for African-Eurasian migrants than for migrants on the other two flyways (Table 6.1).
Figure 6.2 Number of migrant species on the breeding and the non-breeding areas which were projected (using the RCP 8.5 scenario) to experience increasing or decreasing climatic suitability across the three migration flyways, by the end of the 21st Century. Light shading displays the number of species for which climatic suitability was projected to increase/decrease based on the median climatic suitability (calculated across the GCMs, SDMs and jackknife iterations). For both seasons, dark colour shows the number of species for which there was a consistent trend in projected change estimates (where 95% quantiles did not overlap with 0).
Table 6.1 Change in climatic suitability for all species across the globe between the baseline period and the future time periods (2050, 2070). Reported are the percentages (numbers) of species likely (where the 95% quantiles show directional consensus) to experience increasing or decreasing climatic suitability in each time period across breeding and non-breeding regions for each migratory flyway (AF= African-Eurasian, AM= Americas, AA= Asia-Australasian). Decreasing climatic suitability values excluded as they are simply 100 minus the values present for increasing climatic suitability.

<table>
<thead>
<tr>
<th>Period</th>
<th>Flyway</th>
<th>Breeding Increasing CS</th>
<th>Non-breeding Increasing CS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>RCP 2.6</td>
<td>RCP 4.5</td>
</tr>
<tr>
<td>2050</td>
<td>AF</td>
<td>93.1% (27)</td>
<td>87.9% (29)</td>
</tr>
<tr>
<td></td>
<td>AM</td>
<td>77.4% (24)</td>
<td>79.4% (27)</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>88.2% (15)</td>
<td>77.8% (14)</td>
</tr>
<tr>
<td>2070</td>
<td>AF</td>
<td>92% (23)</td>
<td>83.3% (25)</td>
</tr>
<tr>
<td></td>
<td>AM</td>
<td>79% (23)</td>
<td>71.4% (20)</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>87.5% (14)</td>
<td>82.4% (14)</td>
</tr>
</tbody>
</table>
6.4.2 Change in species richness

There are clear spatial patterns in projected change in species richness for 2070 based on RCP 8.5 projections on both the breeding and non-breeding areas (Figure 6.3). Breeding migrant species richness differs significantly across latitude (ANOVA: $F_{(5,61)}=10.64$, $P < 0.0001$), where richness increases significantly at higher latitudes (50 to 80° latitude) and decreases significantly at latitudes of between 30° and 50° (Figure 6.4, Post Hoc Tukey $p < 0.01$ for both). Migrant species richness on the breeding grounds is projected to significantly decrease by 6.9 ($\pm 2$) species in latitudes of 50° to 50°, and increase by 2.6 ($\pm 4$) species at latitudes of 50 to 80°. The lowest change in migratory richness is in lower latitudes (30° to 50°), where there is currently low breeding migrant richness (Figure 6.3a). Overall, migratory species richness in the non-breeding areas is projected to decrease on average by -2.9 $\pm$ 1.7 species (Figure 6.3, Figure 6.4, ANOVA: $F_{(5,61)}=5.033$, $P < 0.001$). Migratory species richness in non-breeding areas was significantly decreased in latitudes of 0-30° (Figure 6.4, Post Hoc Tukey $p < 0.001$).

Areas of higher latitude that are currently relatively species poor (Figure 6.3a) are projected to experience increased species richness of breeding migrants for mid- and late-century climate change projections (presented RCP8.5 2070 projections, Figure 6.3b,c,e and f). Areas that currently contain high levels of breeding migrant species richness (Figure 6.3a) are projected to lose species richness in the future by over 30 species per cell (Figure 6.3b). For non-breeding migrants, areas in Brazil are projected to have increased species richness in mid-and late-century climate change projections, with more southern areas projected to experience loss in non-breeding species richness. Areas in Africa, India and Asia are projected to generally experience losses in non-breeding migrant species richness (Figure 6.3e,f).
Figure 6.3 Projected global impacts of climate change (RCP scenario 8.5) on species richness. Maps show projected changes in the number of species for which there is suitable climate between present and 2070. a) Is baseline breeding migrant species richness; b) Is the projected change in the number of species of migrants breeding in these areas. Future climates are likely to be suitable for a greater (red) or a fewer number of species (blue). Colour intensity indicates the magnitude of change. Grey area is low change, from -3 to + 3 species (includes no change). c) Shows areas that are projected to show the greatest gain (red) and loss (blue) in breeding migrant species richness in these areas (the lower and upper bounds of the 95% CI’s of predicted change in richness). d) Depicts baseline non-breeding migrant species richness; e) Shows the projected change in the number of migrants on the non-breeding grounds. Future climates are likely to be suitable for a greater number (red) or a fewer number (blue) of species. Colour intensity indicates the magnitude of change. f) regional projections of model residuals highlight areas that are projected to have the greatest gain (red) and loss (blue) in the number of non-breeding species (lower and higher bounds of the 95% CI’s of predicted change in suitability).
Figure 6.4 Mean change in species richness on the a) breeding and b) non-breeding areas for four latitudinal bands (from high latitude to low latitude). The longitudinal bands were: band 1= [70 to 80° latitude]; band 2= [50 to 70° latitude]; band 3= [30 to 50° latitude]; band 4= [0 to 30° latitude]; band 5= [-30 to 0° latitude]; and band 6= [-50 to -30° latitude]. Blue dashed line indicates no change in species richness.
6.4.3 Change in species turnover

Overall, median turnover in migratory species increased from 18% (95% confidence range=16.7%-21.8%) in 2050 to 21.3% (16.7-26.9%) in 2070 across their breeding ranges for all RCP scenarios (Figure 6.5). Median turnover in migrants on the non-breeding areas increased from 17.6% (16.1-21%) to 20.5% (16.1-26.7%) for the same periods (Figure 6.5).

![Graph showing species turnover](image)

**Figure 6.5** Median (95% quantiles) projected species turnover for each migration flyway (red= African-Eurasian; green= Americas; and blue=Asian-Australasian) calculated across the globe in two time periods (2050 and 2070) across for migrants on the breeding and non-breeding ranges.

Species turnover was greater across the breeding ranges than across the non-breeding ranges (Figure 6.6). Further, species turnover was projected to be higher for 2070 than for 2050 for both breeding and non-breeding migrants (Figure 6.5, Figure 6.6). Areas in northern latitudes were projected to experience the greatest amounts of species turnover (Figure 6.6).
Figure 6.6 Projected percentage turnover in species composition for mid-Century (a and b) and the end of the 21st Century (c and d) based on the RCP 8.5 scenario climate projections. Percentage species turnover of breeding migrants are depicted in a and c, and percentage species turnover of non-breeding migrants is shown in b and d.
6.4.4 Change in migration distance

Median migration distances (km) were projected to increase significantly across the three time periods (Repeated measures ANOVA: Figure 6.7: F=72.64_{(2,644)} < 0.001).

Figure 6.7 Difference in median (95% quantiles) projected migration distance (km) across all species for three time periods (baseline, 2050 and 2070).

Across all migrants, 22% of species are projected to have shorter median migration distances in the future, while 78% (252 species) are projected to have longer migration distances by 2070. Migration distance was significantly larger for migrants in 2070 than 2050 (Figure 6.8, Paired t-test: t=-3.29, df =319, P<0.001).
Figure 6.8 Percentage change in median migration distances for migrants between baseline and the two time periods (2050, 2070).

The range (standard deviation of migratory distances estimated across the 1000 iterations for each species) of migration distances that species are likely to experience in the future are projected to increase for 74% of species and decrease for 26% of species by 2070. The range of migration distances that species are likely to face are 4% larger for 2070 than for 2050 (Figure 6.9, Paired t-test: t=−2.1, df=319, p<0.05).

Figure 6.9 Percentage change in the range of migration distances that species are likely to travel for 2050 and 2070.
Figure 6.10 a) Observed breeding and non-breeding ranges for Thrush Nightingale (*Luscinia luscinia*) from BirdLife (Birdlife International & NatureServe 2011); b) frequency histogram of the baseline predicted migration distances (km) estimated using the distance weighted distribution approach. Dotted line is the median migration distance for this species (8132 km); c) projected (RCP 8.5) breeding and non-breeding distributions of *L. luscinia* for 2070 and c) frequency histogram of the future predicted migration distances (in km) estimated using the distance weighted distribution approach. The dotted line is the median migration distance estimated for this species (6105 km). Image of Thrush Nightingale from C. Bobzin, http://www.christofbobzin.de licenced under the creative commons attribution share alike 3.0, Germany.
Numbers of migration modes across all species were projected to increase by 2070 for 31% of migrants, decrease for 27% of migrants and remain the same for 42% of migrants. Whether increasing or decreasing, the predicted number of migration modes for species did not differ significantly between flyways ($\chi^2 = 8.9$, df= 4, P=0.06). The projected increase in the number of migration modes by 2070 was in response to the emergence of novel, climatically suitable locations that were less distant than current non-breeding locations for these species. An example species for which the number of flight modes was found to increase using the in a mixture model framework (described in methods) is depicted in Figure 6.10. For this species, the number of modes was predicted to increase from 2 to 3, as novel, climatically suitable areas that were closer to the breeding range for the species emerged in the region surrounding Bangladesh.

6.5 Discussion

This study provides the first assessment of climate change impacts on future migrant species at a global scale. My results suggest that climate change has the potential to dramatically impact the distributions of migratory species on both the breeding and non-breeding areas, across the globe. This assessment should highlight the potential threats that climate change poses for migrants in different locations globally, to aid the targeting of conservation and monitoring efforts. Here I discuss my findings in relation to change in climatic suitability, species richness, species turnover and migration distances.

6.5.1 Climatic suitability

There is considerable spatial heterogeneity in projected impacts of climate change on climatic suitability for migrants across the breeding and non-breeding areas, with the impacts of climate change increasing towards the end of the 21st century. Overall, climatic suitability was projected to increase globally, but there was considerable variability among flyways. Migrants of the Americas flyway were projected to have increased climatic suitability for both breeding and non-breeding areas, whereas Asian-Australasian migrants were projected to experience decreasing climatic suitability on the non-breeding areas. African-Eurasian migrants were projected to experience decreasing climatic suitability on both breeding and non-breeding areas. These findings are in accordance with a recent study (Stephens et al. 2016) whereby population trends for long-distance migrants in North America are positive for those species that are expected to respond positively to climate in the future. In contrast, long distance European migrants showed negative population trends for those that were expected to respond positively as well as those expected to respond negatively under future climate change. Given the findings by Stephens et al. (2016) and the predicted increase in climatic suitability for long-distance migrants in the Americas from my study,
migrants in the Americas flyway may fare better in the future than their African-Eurasian and Asian-Australasian counterparts.

6.5.2 Species richness and turnover

Species richness of breeding migrants was projected to increase at high latitudes, but decrease where there are currently large numbers of breeding migrants present. This is in agreement with studies which show that species are shifting their ranges poleward in a pattern that is consistent with climate change (Thomas & Lennon 1999; Hickling et al. 2006; Brommer, Lehikoinen & Valkama 2012). In contrast, the change in species richness of migrants on the non-breeding areas showed no consistent change with latitude. African-Eurasian and Asian-Australasian non-breeding areas were projected to experience declines in richness because of declines in climatic suitability in these areas, which contrasts with migrants in South America, where species richness is projected to increase in currently low diversity regions for non-breeding migrants as climatic suitability in this region increases. Climate change is therefore likely to affect breeding and non-breeding ranges to different extent, as has been proposed previously (Doswald et al. 2009). Note, however, that the quality of species occurrence data on the non-breeding grounds is poorer than for the breeding grounds for most species.

Forecasting the effects of future climate change on species distributions using SDMs can be fraught with uncertainties (Bagchi et al. 2013), but measuring the uncertainty associated with predictions makes interpreting these results more robust (Garcia et al. 2012; Bagchi et al. 2013; Baker et al. 2015). Although uncertainty was high, models that showed congruence projected migrants to experience increased climatic suitability across the three migration flyways for both breeding and non-breeding seasons (Table 6.1).

Turnover was projected to be highest between present and 2070, with areas in the northern latitudes experiencing higher species turnover than areas of lower latitudes. These climate change impacts on turnover could result in altered community composition and create non-analogue communities in the future (Urban, Tewksbury & Sheldon 2012). Species interactions such as competition and predation have not been assessed in this study, but have been shown to alter species responses to climate change (Mason et al. 2014), and is a further effect that must be taken into consideration. Areas of high turnover will have modified community structure, which may lead to ecosystem disruption (Peterson et al. 2002).
6.5.3 Migration distance

Migration distances were projected to increase for most long-distance migrants (79%), and decrease for 22%. This is despite the fact that for many species (31%), in future, some novel non-breeding areas were predicted to arise closer to the breeding areas than currently occur. This is likely due to the poleward shift of breeding ranges, while non-breeding migrant occurrence did not show consistent poleward latitudinal shifts. Non-breeding ranges located in tropical areas are not projected to move in consistent directions because of a lack of latitudinal climatic gradients (Huntley et al. 2006). Increasing migration distances will exacerbate the pressures upon long-distance migrants. Already migrants declines are occurring, attributed, in part, to likely mismatches with peaks of food availability on the breeding grounds (Both et al. 2006; Jones & Cresswell 2010; Saino et al. 2011). Migrants thus need to time their migration in accordance with resource dynamics along the migration route as well as at breeding/wintering sites in order to gain greater fuel loads for longer migrations (Saino et al. 2004; Gordo & Sanz 2008).

In contrast with long-distance migrants, short-distance migrants are predicted to decrease their migration distances, largely as a consequence of non-breeding areas shifting closer to the breeding areas (La Sorte & Thompson 2007; Visser et al. 2009). Short-distance migrants are therefore better able to predict conditions on the breeding grounds better as they are responding to cues that are closely linked to conditions indicating optimal arrival time, such as temperature (Cotton 2003; Gordo et al. 2005). In contrast, long-distance migrants may respond to cues that are entirely uncorrelated with the timing of spring phenology (Jones & Cresswell 2010). Short distance migrants should therefore not suffer the consequences of mistiming arrival or breeding as long-distance migrants do (Both et al. 2006). Indeed, if short-distance migrants can better adapt to changing climatic conditions than long-distance migrants, the latter may suffer a competitive disadvantage, leading to further declines in these species (Visser et al. 2009).

In many cases, potential future non-breeding ranges were simulated in regions far from the current non-breeding grounds, suggesting that in order for migrants to utilise these areas in response to climate change, they would need to develop new migration routes. This may not be possible, as although migrant species are highly mobile, migratory routes of birds are often genetically determined, thus migration routes may be conserved despite substantial climatic changes (Ruegg & Smith 2002; Ruegg, Hijmans & Moritz 2006). Additionally, factors that have not been considered in this study may limit species distributions in conjunction to climate. These include the availability of suitable habitats, change in land-use practices, as well as species interactions which will contribute to complexities in predicting future occurrences of migratory species (Hill,
In order to understand bird movements it is necessary to acquire data of where they are traveling. An important, but not yet understood, issue in the study of migrating birds, is understanding the extent to which breeding and non-breeding populations are connected (or migration connectivity) (Webster et al. 2002). It is clear that the periods that migrants spend in widely separated and disparate ecological locations in different periods of their annual cycle are inextricably linked (Marra, Hobson & Holmes 1998; Gill et al. 2001; Gordo et al. 2005; Morrison et al. 2013). Migrant species are often intensively studied during one season (usually breeding) and not the other (Salewski & Jones 2006). Without a solid understanding of the year-round geographic distributions and habitat requirements of migrant species, long term conservation plans become difficult (Webster et al. 2002; Small-Lorenz et al. 2013). The need to understand the alarming population declines of long-distance migrants (Sanderson et al. 2006) requires the study of migration patterns and the impacts of climate and land-use change on these patterns across all stages of the migratory cycle (breeding, non-breeding and staging areas) (Tøttrup et al. 2012; Small-Lorenz et al. 2013). Conservation of migratory species in the future will depend on obtaining information on migratory connectivity to inform any decisions and meeting the challenge of ensuring conservation strategies cover areas across the full migration route for species.

Here, I have shown that migrant distributions on the breeding and non-breeding areas can be accurately predicted using SDMs, and that predictions can be made of the location of suitable climate for migrants given future climate change projections. These geographic projections of climatically suitable locations for migrants could therefore be used to inform conservation management decisions (Guisan et al. 2013). Further, I have shown that as species distributions shift in response to projected changes in climate, concordant changes are expected to occur within communities. Correlative models currently ignore the biotic interactions species may face (from predators, competitors, prey) which can lead to misinformed predictions on the constraints species may face in terms of the reassembly of communities in the future (Schaefer, Jetz & Böhning-Gaese 2008). Understanding the long term implications of these trends represent an important challenge for scientists.
6.5.4 Conclusion

This study highlights that while there is considerable geographic variability associated with projecting the impacts of future climate change on migratory species across the globe, conclusions can be made about what challenges migrants are likely to face given these changes. In particular, I showed that climate change will substantially alter the breeding distributions of many migratory species; species distributions on the non-breeding areas are not projected to show such obvious directional shifts. Important breeding sites for migrants are projected to become even more northerly in the future. As a consequence of these divergent shifts between the breeding and non-breeding future ranges for species in the future, migration distances were projected to increase. Although there was variation in projected migratory patterns because of variable climate change projections, climatic suitability consistently improved for migrants on the breeding areas across all three global migration flyways. These results are a first step to determining the global consequences of climate change on migratory birds, and could be used to inform conservation planners in securing the persistence and future of currently declining long-distance migrants.
Chapter 7

General discussion

Photographs by Ian Fisher

Nightingale, *Luscinia megarhynchos*
7.1 Synthesis

In this thesis I investigated the relationship between environmental predictors and migratory species occurrence and richness, at a global scale, with the aim of using these relationships to infer possible effects of climate change on migratory species. This work has illustrated the myriad of influences that climate has on migratory species distribution patterns, highlighting that future climate change impacts on these important species are likely to be varied and complex. This analysis focuses on migratory bird distributions globally to assess the climatic determinants of their occurrence in order to predict climatically suitable locations for these species in the future. Furthermore, this project illustrates the importance of considering the impact of biotic factors, such as the cost of migration, for assessing potential shortfalls or benefits that range shifts might incur upon this group of mobile species. Here, I discuss the findings and novelty of this research, including recommendations for future research.

In the following section I set out the core results of each chapter and discuss these results as a whole within the wider framework of our current understanding, before outlining limitations and assumptions of the work not previously addressed. Finally, I present gaps and questions raised by this research, and suggest avenues for further research.

7.2 Thesis discussion

Rapid loss of biodiversity has occurred across the globe in recent decades, driven primarily by human modification of the environment, and is a major driver of ecosystem change (Sala et al. 2000; Butchart et al. 2010; Cardinale 2012; Hooper et al. 2012). Preventing the loss of biodiversity is a global priority (Rands et al. 2010), as the loss of biodiversity could have profound effects on ecosystem functioning (Loreau et al. 2001), with consequences for the resilience of ecosystems to environmental change (Mori, Furukawa & Sasaki 2013). Given that rapid declines in migrant birds have been observed (Sanderson et al. 2006; Kirby et al. 2008), understanding the drivers of migratory species richness across the globe, in both breeding and non-breeding seasons, is an essential first step toward mitigating declines in these species, and is central to conservation planning (Jiguet et al. 2005).

In Chapter 2, I investigate the importance of plausible drivers that explain migratory species diversity globally. Studies that have analysed migrant species richness have often focused on the breeding ranges of migrant species (Lemoine & Bohning-Gaese 2003; Barcena et al. 2004), overlooking the patterns and drivers of diversity of migrants in non-breeding areas, and how the two might be related. I show that migratory species richness on both the breeding and non-breeding
ranges can be predicted using a set of ecologically meaningful variables. Resident species richness was found to be an important predictor of migratory species richness globally, which supports the idea that migrants may be using areas of high resident richness as cues for profitable site selection (Monkkonen & Forsman 2002; Thomson, Forsman & Monkkonen 2003). Further, diversity was driven by the location relative to the breeding and non-breeding grounds, thus supporting the idea that the cost of migration distance affects the distance birds are willing to travel between the two sites (Wikelski et al. 2003; Newton 2008; Somveille, Rodrigues & Manica 2015).

By mapping spatial trends in migrant species richness globally, I present the locations of biodiversity hotspots for these vulnerable species, which can be used for assessing conservation priorities (Jiguet et al. 2005). However, species richness metrics can be dominated by common and widespread species (Jetz & Rahbek 2002; Rahbek et al. 2007b), which may conceal areas which are rich in small ranged species, often the focus of conservation efforts (Rahbek et al. 2007b). Another metric of biodiversity, phylogenetic diversity, has been proposed as an indicator of functional diversity and evolutionary potential (Lankau et al. 2011; Winter, Devictor & Schweiger 2012). The potential loss of evolutionary information has been realised (Heard & Mooers 2000; Purvis et al. 2000), and recent studies have shown that for birds, the locations of high species richness and high phylogenetic diversity do not necessarily overlap (Voskamp et al. 2015, in review).

Correlative models are the most widely used approach for modelling current and potential future species distributions (Beerling, Huntley & Bailey 1995; Araujo & Guisan 2006; Thuiller et al. 2006). Many studies have focused on methodological techniques to enhance performance of species distribution models (SDMs) (Elith et al. 2006; Pearson et al. 2006; Barbet-Massin & Jetz 2014). Despite this, the selection of an appropriate study area over which to select absences can be problematic and is a frequently overlooked aspect of SDM methodology. However, delineating the study region is not straightforward, as it depends on knowledge of the natural history and the dispersal ability of a species, as well as information about the landscape of interest, including the configuration of suitable habitats and barriers to dispersal (Soberon & Peterson 2005; Barve et al. 2011). Defining this is particularly difficult for migrant species (that inhabit spatially and temporally distinct areas during their annual migratory cycle, and that have the ability to disperse large distances) and for invasive species, whose potential area for expansion is unknown (Vaclavik & Meentemeyer 2009). In chapter 3, I propose the use of a distance weighted absence selection approach that avoids the need to use a predefined study extent. I found that the $\frac{1}{D_x^2}$ ($D_x$ is the distance from the nearest edge of a given virtual species’ distribution) weighting metric that balanced the reproduction of climatic suitability across a range of suitabilities, and performed well at
reproducing the distributions of virtual species was the best method for selecting absences. This is a novel and important solution to a common methodological issue.

In chapter 4, I use the absence selection approach created in Chapter 3 to develop SDMs that relate migrant distributions, globally, to climatic predictors at present. These analyses revealed that the importance of climatic variables for migrant distributions varied across the globe, and by season, which implies that relationships of migrants with climate in one migration flyway may not be the same as in others. Species polygon data, which give geographic information about where a species occurs and where it does not (presence-absence), was used for these analyses. The use of presence-absence data for SDMs has been criticised because they do not account for species abundances within a species range extent (Jiménez-Valverde, Lobo & Hortal 2008; Jimenez-Valverde et al. 2009). Furthermore, the role of presence-absence data in ecological research has been questioned, as heterogeneity among individuals or among populations cannot be accounted for at the species level (Davis et al. 1998; Dormann 2007; Kearney, Wintle & Porter 2010). However, here I have demonstrated that presence-absence data can contribute considerably to ecology, for instance revealing the influence of climate on species distributions at present, which is a valuable first step for understanding the broad effect of climate on the distributions of migratory species.

Accurate estimates of migratory distance are important to determine any costs or benefits to migration in the future should migratory distances change. I establish a method of estimating migration distance (Chapter 5) which improved on the simple centroid-centroid distance estimation approach used in previous studies (Helm, Hanski & Pärtel 2006; Huntley et al. 2008; Doswald et al. 2009). The approach developed was more biologically meaningful than centroid-centroid distances, as it also considered migration distance as a cost (Somveille, Rodrigues & Manica 2015). This is more appropriate than the centroid-centroid approach, which gives little information of the underlying processes which might alter migration distance. Although this method gave more nuanced estimates of migration distances for species, the method could be improved upon. For example, if abundance data were to be incorporated into the selection of climatically suitable cells on the non-breeding ranges, then the effects of conspecific competition (Salomonsen 1955), characterised by leap-frog migration (Swarth 1920; Stanley et al. 2015) which are currently not considered in this study could be incorporated in this distance estimation approach.

I investigate potential changes in migratory distributions and in species specific migratory distances, and discuss the implications of projected changes in these (Chapter 6). These analyses show the influence of future climate change on migratory species in terms of: changes in available climatically suitable space, changes in species richness and changes in species turnover. A number of
findings from this chapter raise matters that could have significant implications for the conservation of migratory species. First, migrant distributions on the breeding and non-breeding areas were accurately predicted using SDMs, thus future projections of the location of suitable climate for migrants were made which could inform conservation management decisions (Guisan et al. 2013). Second, higher latitudes were projected to experience larger species turnover than areas at lower latitudes. Climate change impacts on turnover could result in altered community composition and create non-analogue communities in the future (Urban, Tewksbury & Sheldon 2012), which may lead to ecosystem disruption (Peterson et al. 2002). Species interactions such as competition and predation were not assessed here, but have been shown to alter species responses to climate change (Mason et al. 2014), and is an effect that must be taken into consideration.

Third, areas that are climatically suitable for migrants were projected to increase globally, but there was considerable variability across flyways. Migrants of the Americas flyway were projected to have increased climatic suitability for both breeding and non-breeding areas, whereas Asian-Australasian migrants were projected to experience decreasing climatic suitability on the non-breeding areas, and African-Eurasian migrants were projected to experience decreasing climatic suitability on both breeding and non-breeding areas. Given these results, migrants in the Americas flyway may fare better in the future than their African-Eurasian and Asian-Australasian counterparts. Overwhelming evidence exists that long-distance migratory species are in serious decline in the Palearctic-African migration system (Bohning-Gaese & Bauer 1996; Sanderson et al. 2006; Ockendon et al. 2012), as well as migrants following the Palaearctic-Asian migration patterns (Laaksonen & Lehikoinen 2013). Long-term data sets in North America show over half of all Neartic-Neotropical migrants have experienced substantial declines over the past 40 years (BirdLife International 2008b). Despite predicted increases in climatic suitability for migratory species within the Americas flyway, important habitats for these migrants are under threat from anthropogenic impacts such as forest fragmentation, tropical deforestation and agricultural expansion which may contribute to declines of migrants in the Americas flyway (Murphy 2003; Kirby et al. 2008; Newton 2008). Protected area networks are a key component to protecting biodiversity from anthropogenic threats (Baker et al. 2015). For some migratory species, existing protected areas will be adequate to allow these species to adapt to change, but others (such as those of the Americas flyway) will benefit from conservation actions which improve the extent of suitable habitat (Hole et al. 2011). The need to consider shifts in the distribution of migratory species are recognised within international conservation initiatives (Trouwborst 2012). The identification of locations that will maintain present and future migratory species distributions is important for conservation investment, and the projections of future
migratory species distributions in this thesis are an important step for incorporating shifts into conservation planning (Carvalho et al. 2011; Hole et al. 2011)

Finally, migration distances were projected to increase for the majority of long-distance migrants. Increasing migration distances will exacerbate the pressures upon long-distance migrants. Already declines are occurring in migratory species, attributed, in part, to likely mismatches with peaks of food availability on the breeding grounds (Both et al. 2006; Jones & Cresswell 2010; Saino et al. 2011). In seasonal environments, an organism’s fitness is determined by the timing of life history events (Miller-Rushing et al. 2010). Natural selection should favour migratory birds that can synchronise migration with peaks of resource availability, particularly as migrating birds use chains of stopover sites en route which often differ in seasonality (Bauer, Gienapp & Madsen 2008). Timing of migration is of particular interest, as migratory birds, particularly long distance migrants, are constrained by time and energy during migration (Wikelski et al. 2003). Arctic breeders have been shown to have reductions in reproductive success or complete breeding failure if they arrive at their breeding grounds outside a very narrow time window (Alerstam & Lindstrom 1990; Bauer, Gienapp & Madsen 2008). Declines of the pied flycatcher in the Netherlands have been attributed to a disruption of the synchronicity between the flycatcher’s migration and the emergence of caterpillars caused by climate change (Both et al. 2006). Thus migrants need to time their migration in accordance with resource dynamics along the migration route as well as at breeding/wintering sites in order to gain greater fuel loads for longer migrations (Saino et al. 2004; Gordo & Sanz 2008). Although stopover habitats are used by many species as refuelling stations during migration, which can be critical for survival and successful reproduction, stop-overs are rarely incorporated into conservation strategies (Sheehy, Taylor & Norris 2011). For instance, the US National fish, wildlife and plants climate adaptation strategy does not address the needs of migratory species in their plan (Runge et al. 2014).

Observed declines in Afro-Palearctic migrants have been attributed to a myriad of factors, including climate change (Lemoine & Bohning-Gaese 2003), habitat loss or deterioration on breeding, non-breeding and staging locations (Berthold 2001; Schaub, Kania & Köppen 2005) and hunting pressure (McCulloch, Tucker & Baillie 1992). Well-loved migrants in the UK such as the spotted flycatcher (Muscicapa striata), the turtle dove (Streptopelia turtur) and the nightingale (Luscinia megarhynchos), have shown dramatic declines in recent decades (SUKB 2014). In Chapter 6 I project decreasing availability of suitable climatic space on both the breeding and non-breeding locations of African-European migrants. Conservation along the migration route (across breeding,
non-breeding and staging areas) will be of crucial importance to conserve these summer visitors in the face of projected climate change.

The aim of researching migration is to understand when, where, why and how animals migrate (Alerstam, Hedenstrom & Akesson 2003; Wilcove & Wikelski 2008). Migratory movements may have wide reaching ramifications for other animals, including humans, as migratory species have the potential to spread emerging diseases across geographic barriers (Liu et al. 2005b; Olsen et al. 2006). Furthermore, migration is not simply the movement of animals, but the coupling of ecological communities globally, and the mediation of ecological community diversity and stability (recently reviewed in Bauer & Hoye 2014). Given the potential of migration to affect ecological networks across the world, the ramifications of the potential decline of migratory species are vast.

The methods developed in this thesis need not apply solely to migratory birds, but to a range of migratory taxa. For instance, species distribution models have been applied to marine ecosystems (Dambach & Rödder 2011) and used to design reserve networks for riverine fish (Esselman & Allan 2011). Modeling migratory species in these taxa poses similar challenges to ecologists as modeling migratory birds does. For example delineating a boundary or extent to the study region for these species is a challenge, particularly for marine species, where simple boundaries (such as a country boundary) cannot be applied. The methods developed in chapter 3 overcome the necessity to delimit an extent based on arbitrary decisions, and could be used for the study of wide-ranging marine species. The conservation of terrestrial migrants is challenging, as these species not only face climate change impacts, but are increasingly affected by human modification of the landscape through which they migrate (Serneels & Lambin 2001). Terrestrial migration, such as for that of the endangered Mongolian saiga antelope (*Saiga tatarica mongolica*) is primarily challenged by anthropogenic forces such as degradation of pastures by livestock grazing as well as over-harvesting by poaching (Clark & Javzansuren 2006). Saiga numbers have plummeted by more than 95% in under two decades (Milner-Gulland et al. 2001). The conservation of terrestrial species thus needs to incorporate climate and well as land-use changes into conservation frameworks.

### 7.3 Assumptions and limitations

Climatic means provide useful metrics of relative change in climatic conditions over time, and are likely to be relevant in environments that respond slowly to climatic variability (Etterson & Shaw 2001), or where life histories allow species to buffer this variability (Post & Stenseth 1999; Morris et al. 2008). However, without explicitly accounting for variability in climatic conditions mean values alone are unlikely to be a good predictor of species’ sensitivities and adaptive capacity to
climate (Deutsch et al. 2008). While using 30-year averages revealed broad associations between species and climatic conditions, variables that correspond to the time (or indeed, the period before arrival) in which migrants are located on the breeding or non-breeding areas, are likely to vastly improve the robustness of these models (Huntley 1995; Heikkinen, Luoto & Virkkala 2006). For instance, rain in the Sahel in the preceding winter can advance migratory species spring arrival across Western Europe (Both 2010).

Both Chapters 4 and 6 draw on distributions derived from species distribution models (SDMs). Despite widespread use of these models to derive spatial predictions of the environmental suitability of species (Guisan & Thuiller 2005; Elith & Leathwick 2009), limitations on their accuracy and predictive power are widely recognised (Araújo & Peterson 2012; Dormann et al. 2012b; Heikkinen, Marmion & Luoto 2012). These models often rely solely on climatic descriptors of a species distribution, ignoring historical factors (such as geographic barriers), evolutionary processes and biotic interactions; all of which also influence distributions (Guisan & Thuiller 2005; Austin 2007; Jiménez-Valverde, Lobo & Hortal 2008). Improvements of SDMs could be made, both in terms of their biological realism and their transferability to novel environments, by the use of process-based distribution models (or mechanistic models) (Dormann et al. 2012a) which explicitly incorporate physiological processes (Kearney & Porter 2009), and can include processes such as dispersal (Barbet-Massin, Thuiller & Jiguet 2012) and biotic interactions (Kissling et al. 2011; Wisz et al. 2012).

### 7.4 Future directions of further research

In order to understand bird movements it is necessary to acquire data of where they are travelling. An important, but not yet understood, issue in the study of migrating birds, is migratory connectivity, the extent to which individuals from the same breeding area overwinter in the same areas (Webster et al. 2002). It is clear that the periods that migrants spend in widely separated and disparate ecological locations, in different periods of their annual cycle, are inextricably linked (Marra, Hobson & Holmes 1998; Gill et al. 2001; Gordo et al. 2005; Morrison et al. 2013). Migrant species are often intensively studied during one season (usually breeding) and not the other (Salewski & Jones 2006). Without a solid understanding of the year-round geographic distributions and habitat requirements of migrant species, long term conservation plans become difficult (Webster et al. 2002; Small-Lorenz et al. 2013).

Massive mark-recapture efforts of bird ringing (banding) means birds can be individually marked and later recaptured, providing evidence of their movements. In spite of such efforts, we still do not know where many individual birds ringed (usually on the breeding grounds) spend their winter because the probability of finding and recording a ringed bird varies geographically, and is
particularly low in some regions, such as on sub-Saharan non-breeding grounds (Korner-Nievergelt, Liechti & Thorup 2014). Exciting new advances in remote sensing techniques, as well as the analyses of genetic markers and chemical isotopes offer a level of insight that was previously unavailable (Webster et al. 2002; Rundel et al. 2013). Remote sensing techniques (such as fitting satellite tags or geolocators to birds) has revolutionised our understanding of avian migratory pathways as well as the locations of non-breeding areas and stop-overs for some species (Delmore, Fox & Irwin 2012; Stach et al. 2012).

Technological advances in measuring individual movements hold much promise for the future study of migratory species, such as reconstructing detailed pathways of individual birds (Delmore, Fox & Irwin 2012). For example, in order to better conserve the declining populations of cuckoo (Cuculus canorus) in Britain, the British Trust for Ornithology (BTO) have satellite tagged approximately 50 individuals to understand the circumstances that are contributing to their mortality (http://www.bto.org/science/migration/tracking-studies/cuckoo-tracking). Understanding the factors operating on migrating individuals throughout the annual cycle is essential for producing robust species distribution models in order to predict the ecological consequences of changes in climate and habitat in various locations throughout the year. In this thesis I have assessed species-specific responses to climate, but it is clear that responses of individuals to climate add further complexity to predicting ecological responses to climate change (Deutsch et al. 2008). Satellite tracking, as well as offering important information on poorly studied non-breeding localities for species, can provide vital information on variation in phenology and distribution among individuals within a species (Liechti et al. 2015). Conservation of migratory species in the future will depend on obtaining information on migratory connectivity to inform any decisions and meeting the challenge of ensuring conservation strategies cover areas across the full migration route for species.

An understanding of flight is of fundamental importance for studying bird migration (Alerstam & Hedenstrom 1998). In pioneering studies, theoretical concepts to analyse flight during migration were introduced to the field of optimisation theory (Tucker 1973; Pennycuick 1977; Pennycuick 1978; Weber & Houston 1997; Houston 1998; Weber, Ens & Houston 1998). Given the effects of climate change on migration in terms of phenological and range shifts, it is surprising that there have been no studies (to my knowledge) that analyse the mechanistic processes of migration, while incorporating the effects of these predicted shifts. Biomechanical and physiological studies of flight could aid conservation efforts (Denny & Helmuth 2009). Combining these mechanistic effects with migration theory may enable more accurate predictions of the potential effects of climate change on migratory birds. Bayly et al. (2011) used body mass of Grasshopper Warblers (Sylvia borin) as well as
distance of migration these birds travel from their wintering to their breeding sites, to calculate how much body mass a Grasshopper warbler would need in order to cover the distances required with a

i) non-stop flight method, or ii) using stopover sites. Body masses of recaptured birds, as well as flight range calculations suggested that less than 10% of birds in Portugal could have reached sub-Saharan Africa without pausing at stopover sites for re-fuelling. Flight range equations (Pennycuick 1977; Pennycuick 1978; Pennycuick 2008; Bayly, Rumsey & Clark 2011) could be used to describe the migration and fuelling strategies of migrants in the future. This could highlight potential threats for migrant species, should their migration distances increase (as was predicted in chapter 6), as well as whether or not, or how many, stop over locations need to be protected for species to complete migration (Bayly, Gómez & Hobson 2013). Provision of high-quality stopover sites that allow migrants to replace depleted nutrient and energy stores could aid conservation efforts of migratory species.

I have shown that species distributions can be accurately predicted using presence-absence data and climatic variables (chapters 4 & 5). Where available, abundance data can improve the accuracy of SDMs (Howard et al. 2014). Furthermore, biotic interactions such as competition and predation have not been assessed in this study, but have been shown to alter species responses to climate change (Wisz et al. 2012; Mason et al. 2014) and is a further effect that must be taken into consideration. For example, diurnal altitudinal migration for a mountain ungulate, the alpine chamois (*Rupicapra rupicapra*), was influenced by temperature, but disturbance from domesticated sheep dwarfed this effect, shifting the altitudinal range of chamois dramatically upslope (Mason et al. 2014). Boreal owls interact positively with woodpeckers, as woodpeckers excavate cavities in trees, which provide nesting sites for secondary cavity nesters such as owls (Virkkala 2006). Predictions of four owl species distributions at large scales were significantly improved by incorporating biotic interactions (with six woodpecker species) to models (Heikkinen et al. 2007). A lack of understanding of the factors that limit species’ distributions, which may result in species under-filling their climatic niche (Sunday, Bates & Dulvy 2012) could lead to inaccurate projections of the impacts of climate change on species (Mason et al. 2014).
7.5 Conclusion

This thesis represents an advance in our understanding of the issues of modeling migratory species, and highlights the need to improve these models to accurately predict the distributions of migratory species under future climate change projections. The challenges faced by migrants due to the rapidity of contemporary environmental change are unprecedented. The cumulative effects of climate and land-use change may seriously disrupt long established, intricately timed, relationships of migratory species to their environment (Small-Lorenz et al. 2013). The results from this thesis highlight that only considering one part of the annual migration cycle will seriously underestimate the effects of climate change on these species (Tøttrup et al. 2012; Small-Lorenz et al. 2013). Incorporating migratory movements in all stages of the migratory cycle into conservation planning frameworks is crucial in light of alarming population declines of long-distant migrants (Sanderson et al. 2006), and this thesis presents a first step towards better informed conservation planning for mobile species given future climate change projections.
References


SUKB (2014) The state of the UK’s birds. *The State of the UK’s Birds report is published by a partnership of eight organisations: RSPB; British Trust for Ornithology; Wildfowl & Wetlands Trust; Natural Resources Wales; Natural England; Northern Ireland Environment Agency; Scottish Natural Heritage; and the Joint Nature Conservation Committee.*


Appendices

Figure S 1 Global maps of variables used to predict migrant species richness a) Distance to the breeding range b) Distance to the non-breeding range c) Mean temperature of the coldest month d) Mean temperature of the warmest month e) Actual to potential evapotranspiration f) Habitat heterogeneity g) Productivity h) Seasonality i) Consistency of seasonality and j) Resident species richness.
Figure S 2 Mean standardised coefficient values (± 95% confidence intervals) of the full model of breeding migrant species richness for the breeding Australasian realm. Numbers indicate the frequency with which these variables were chosen in the best model across the 1000 randomly selected subsets of data. (1/2) is the predictor variable and (2/2) is the quadratic term of the predictor variable. Variables include APET (actual to potential evapotranspiration); Consistency of seasonality; INVDI (productivity); MTWM (mean temperature of the warmest month); non_br_distance (distance to the non-breeding grounds); Res.sprich (resident species richness) and SW (habitat diversity). Predictors that were included in the best models fewer than 90% of the time were excluded, in this case: SW, Consistency and INVDI were excluded from final predictions.
Figure S 3 Frequency histogram of the median area (km$^2$) of occupancy for a random sample of real avian species distributions obtained from BirdLife International. 50 species were randomly sampled from the avian BirdLife polygon data 100 times, and the median of each sample was derived. Median area of virtual species range sizes created (in blue) falls within the distribution of real species.
Table S1. For methods that involved selecting absences at random, the process was replicated ten times to account for variability among the potential absences chosen. The mean across virtual species of the coefficient of variation (CV) in AUC among the ten replicates for each absence selection approach and SDM was low, suggesting that AUC was reasonably robust to the selection of data.

<table>
<thead>
<tr>
<th>Method for selecting absences</th>
<th>GAM</th>
<th>GBM</th>
<th>GLM</th>
<th>RF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absences random</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>One over distance</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.01</td>
</tr>
<tr>
<td>One over distance squared</td>
<td>0.09</td>
<td>0.10</td>
<td>0.10</td>
<td>0.02</td>
</tr>
<tr>
<td>One over distance cubed</td>
<td>0.10</td>
<td>0.08</td>
<td>0.10</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Table S2. Model performance at predicting the distribution of virtual species within each left out test block was evaluated using AUC. Absences were those used to condition the model in each of the absence selection methods. Mean AUCs across 50 species x 10 repeats (once for CS) varied by absence selection method and by SDM approach.

<table>
<thead>
<tr>
<th>Method for selecting absences</th>
<th>Mean AUC (±SD) by species distribution model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GAM</td>
</tr>
<tr>
<td>Absences random</td>
<td>0.98 (0.02)</td>
</tr>
<tr>
<td>One over distance</td>
<td>0.83 (0.10)</td>
</tr>
<tr>
<td>One over distance squared</td>
<td>0.74 (0.09)</td>
</tr>
<tr>
<td>One over distance cubed</td>
<td>0.71 (0.08)</td>
</tr>
<tr>
<td>Concentric shells</td>
<td>0.67 (0.07)</td>
</tr>
</tbody>
</table>
Figure S 4 Three major global flyways (from BirdLife International - [http://www.birdlife.org/flyways/](http://www.birdlife.org/flyways/)), the Americas flyway (green), the African-Eurasian flyway (red), and the Asian-Australasian flyway (blue).
Figure S 5 The relative importance of each climatic variable (MTWQ, AP, SP, ST) for predicting migrant occurrence of both breeding and non-breeding migrants on each of the six biogeographic realms (Afrotropical, Australasian, Indo-Malayan, Nearctic, Neotropical and Palearctic).
Figure S 6 Figure of *Ficedula parva* distribution adapted from BirdLife International (2016). Downloaded from http://www.birdlife.org on 03/05/2016. The term origin uncertain means the species is/was present, but it is not known if it is present during part or all of the year.
The world was split into 9 longitudinal bands and 12 latitudinal bands (a) for analyses. These longitudinal bands were subsequently grouped into four bands (b), the fourth band is comprised of the 4\textsuperscript{th}-9\textsuperscript{th} latitudinal bands.

**Figure S 7** The world was split into 9 longitudinal bands and 12 latitudinal bands (a) for analyses. These longitudinal bands were subsequently grouped into four bands (b), the fourth band is comprised of the 4\textsuperscript{th}-9\textsuperscript{th} latitudinal bands.
Figure S 8 Projected impacts of climate change (RCP scenario 8.5) on species richness across the globe. The maps show projected changes in the number of species for which there is suitable climate between present and the mid-21st Century (2050). A) Shows the projected change in the number of species of breeding migrants. Future climates are likely to be suitable for a greater number of species (red) or a fewer number of species (blue). Colour intensity indicates the magnitude of change. B) Shows areas that are projected to show the greatest gain (red) and loss (blue) in the number of breeding migrant species in these areas (lower and higher bounds of the 95% CI’s). C) Shows the projected change in the number of species of non-breeding migrants. Future climates are likely to be suitable for a greater number of species (red) or a fewer number of species (blue). Colour intensity indicates the magnitude of change. and D) Shows areas that are projected to show the greatest gain (red) and loss (blue) in the number of non-breeding species in these areas.