Predicting the distribution and impacts of non-native birds in the Iberian Peninsula

Martin John Patrick Sullivan

PhD

2014

School of Environmental Sciences
University of East Anglia

Abstract

Increasing numbers of species are being transported beyond their natural range boundaries by humans. These non-native species can have severe negative impacts on native biodiversity. In order to guide management of these species it is important to be able predict where non-native species will spread to, and what impact they will have. This thesis aims to improve our understanding in both these areas, using the expansion of non-native birds in the Iberian Peninsula as a study system. The number of non-native passerines in the Iberian Peninsula has increased in the late 20th century, with the common waxbill *Estrilda astrild,* yellow-crowned bishop *Euplectes afer,* red avadavat *Amandava amandava* and black-headed weaver *Ploceus melanocephalus* all established as breeding species since 1960.

Methods to (1) account for dispersal limitation when modelling the distribution of spreading non-native species and (2) evaluate the likely transferability of native trained species distribution models were developed. The consistency of the species-environment relationship during expansion in the non-native range was also examined. The ability of vacant niches to facilitate the spread of non-native species was tested, and a framework for detecting the early impacts of non-native species was developed.

Species distribution models of the potential distribution of non-native species are improved by incorporating dispersal. Dispersal is an important constraint on the distribution of non-native species, and interacts with environmental suitability to alter the species-environment relationship between the range-margin and the range core, and over time. Despite accounting for dispersal limitation in their evaluation, the performance of native-trained species distribution models was poor when most environmental conditions that were analogous to the species native range were within the species niche.

Non-native birds in the Iberian Peninsula utilised similar resources to native seed-eating birds, but small differences in resource utilisation allowed them to exploit rice fields, where resources were under-exploited by native species. Non-native birds could also interact with native reedbed nesting passerines, and indeed aggression between black-headed weavers and native *Acrocephalus* warblers has been recorded. However, we did not find evidence for competition between these species at current population densities of black-headed weavers.

Further work on non-native species needs to extend the hybrid dispersal-species distribution models developed here, and also to conduct more assessments of the impacts of non-native species in the early stages of their invasion.

Contents

List of Figures

List of Tables

Acknowledgements

I would firstly like to thank my supervisors, Aldina Franco and Richard Davies, for their help, support, enthusiasm, guidance and ideas throughout my PhD. I would also thank all researchers in CEEC and especially Strangles for their advice and support.

This work was funded by a NERC PhD studentship, with an additional fieldwork grant from the British Ornithologists Union. The ICNF and SPEA provided digital atlas data for Portugal, and Birdlife International provided global range maps for non-native species. Valerie Olson, Richard Davies and Amy Romans gave me access to the Bowerbird database of bird traits, which provided a foundation for the more detailed trait data I collected. Thanks to Tommaso La Manita for providing me with information and literature about non-native birds in Italy, Colin Key for interesting discussions about non-native birds in Portugal and to British Birds for providing me literature on non-native birds in Portugal they had published free of charge.

A lot of people and organisations helped me during fieldwork. The ICNF provided logistical support, allowing me to stay at Sitó das Hortas and Ria Formosa and providing access to Ponta da Evra and Ria Formosa. In particular I would like to thank Vitor Encarnacão for going out of his way to help set me up with information, accommodation and contacts. Associação PATO kindly allowed me to stay at their field station for a nominal charge, and gave me access to Paul de Tornada. Thanks to all the staff at Paul de Tornada for being so helpful and welcoming. I would particularly like to thank Helder Cardoso for being so receptive to my research, and for many interesting conversations. Helder Cardoso and Vitor Encarnacão both provided me with ringing data, which while not appearing in this thesis was very interesting to look at. Paula Moniz and Fillipe Moniz helped arrange for me to stay at Ria Formosa and Castro Marim respectively. Gonçalo Cardoso and Ricardo Lopes kindly invited me to give a seminar at CIBIO, where I had many stimulating discussions. It is normal to thank your parents in a thesis acknowledgements, however, I would also like to thank someone else's. Ricardo Correia's parents went out of their way to be helpful to me, whether it was providing me with dinner, letting me print a boarding pass or getting discounts on new tyres. Ricardo helped me a lot on fieldwork, most notably helping me buy a boat and introducing me to the delights of chocolate salami. Less helpful however was his decision to regale me with fieldwork horror stories of being chased by bulls when I phoned him shaken after being threatened to be shot on my first day's fieldwork.

A number of field assistants helped me collect data, and I would like to thank them for that as well as the company and ideas they provided. In order of length of time helping me these are Supatchaya Tetachoochert, Hannah Mossman, Chris Panter and Chris Sullivan.

Several of my chapters have co-authors from outside my supervisory team. Luís Reino contributed data on the expansion of the common waxbill (as well as ideas and comments on the draft manuscript) to chapter 2, Hannah Mossman contributed to all aspects of chapter 6, while chapter 7 builds on the work of an MSc student, James Grundy, who I helped supervise.

Thanks to friends and family for support along the way. Finally, thanks to Hannah Mossman for intellectual input, and for everything else.

1 Introduction

Although some species are very widespread, no species is found ubiquitously throughout the globe, and some species have very restricted distributions. Species distributions can be defined as the geographical areas within a species fundamental niche (which itself is defined as the range of environmental conditions in which populations of a species have a positive growth rate (Holt 2009; Soberon & Nakamura 2009)) that the species has been able to disperse to and is not excluded through biotic interactions. Species distributions are emergent properties of species (Olalla-Tarraga *et al.* 2011), and result from the impact of abiotic and biotic factors on populations and individuals. Species range boundaries represent the transition from suitable to unsuitable biotic and abiotic conditions (McInnes, Purvis & Orme 2009). However, suitable conditions may occur beyond these range boundaries. Therefore at a global scale species are typically dispersal limited.

Patterns of speciation, dispersal and extinction have resulted in biogeographic realms with distinct evolutionary histories (Holt *et al.* 2013). Considerable interchange between these regions has occurred in the past, for example following the forming of the Panamanian isthmus (Tilman 2011). These events have allowed species to overcome previous dispersal constraints and spread to suitable areas, leading to lineages with disparate evolutionary histories mingling.

Currently, an unprecedented degree of interchange is occurring due to the transport of species around the world by humans. These non-native species provide a natural experiment and an applied problem. The study of non-native species allows fundamental ecological questions to be explored (Crawley 2005). For example, introducing species into new areas allows investigation of the role of biotic, abiotic and evolutionary conditions in limiting species distributions (Alexander & Edwards 2010), the ability of communities to accept new species (Briggs 2007) and the role of limiting similarity in controlling community membership (Strauss, Webb & Salamin 2006). Non-native species also provide applied problems, as they are a major cause of biodiversity loss (Clavero & García-Berthou 2005) and economic damage (Lodge *et al.* 2006).

1.1 Invasion biology

1.1.1 Definitions

A plethora of terminology and definitions exist in invasion biology (Blackburn, Lockwood & Cassey 2009). Species can be alien, exotic, non-native, naturalised, invasive or belong to the neobiota. Many of these terms are typically treated as synonymous, but some have more specific meanings (Blackburn *et al.* 2011). I will first consider the apparently synonymous words. Alien and exotic imply that the species they refer to have come from different places and potentially do not belong in the place they are now found, while neobiota implies species that are now found in a location but were not there in the past. There are several issues raised with these terms. Firstly, based on the definitions given above they could potentially be applied to naturally range shifting species (I am using the word naturally even though these species may be shifting their range in response to anthropogenic drivers (Parmesan & Yohe 2003)) as well as species transported to new areas by humans. Secondly, by implying that these species do not belong in the new location these terms have negative connotations. Misunderstandings based on non-neutral terminology can lead to criticisms of invasion biology (Simberloff 2011; Zisenis 2012). I will therefore use the neutral term non-native species. I define non-native species as species that have been transported beyond their natural range boundaries by humans. The emphasis of this definition on transportation by humans is important as this is the key conceptual difference between non-native species and naturally range expanding species.

Some terms are used to describe specific stages in the invasion framework (Blackburn *et al.* 2011). This framework describes the processes that a species must pass through in order to become an established non-native species. Species must be transported to a new location, overcoming the barrier of geography, before being introduced (including accidental escapes) to the new location. Once there, they need to be able to survive and reproduce to a sufficient extent for the population growth rate to be positive and the species to establish self-sustaining populations. They can then spread from the initial site of introduction to other areas with suitable environmental conditions for them (Blackburn *et al.* 2011). Species with self-sustaining populations are often described as established or naturalized, while widespread species are sometimes called invasive (Blackburn *et al.* 2011). However, the term invasive usually refers to species with negative impacts (Williamson 1996). An issue with this use of the word invasive is that it implies that species either have or do not have negative impacts, when impact is probably best thought of as a continuous variable (i.e. species differ in their invasiveness) (Lodge *et al.* 2006). In plant ecology, non-native species that have been established for many centuries are distinguished from more recently established non-native species with the terms archaeophyte and neophyte (Preston, Pearman & Hall 2004).

Although it is common practice to refer to species as non-native, being non-native is a property of populations rather than species. It is possible to have both native and non-native populations of a species within the same country. In the UK, for example, there are both native and non-native populations of welsh poppy *Meconopsis cambrica* (Valtuena, Preston & Kadereit 2011), mountain hare *Lepus timidus* (Harris & Yalden 2008) and greylag goose *Anser anser* (Mitchell, Hearn & Stroud 2012), with considerable mixing of populations of the latter (Mitchell, Hearn & Stroud 2012). Likewise, it is possible for some populations of a non-native species to be established while others are not (Santos, Clacell & Sol 2007).

1.1.2 Transport and introduction

Species can either be deliberately transported to new areas (e.g. plants in the horticultural trade) or they can be transported unintentional stowaways. Certain taxa are more likely to be transported than others (Blackburn, Gaston & Parnell 2010), and certain transport vectors will favour some taxa more than others (Hulme 2009). For instance, marine invertebrates are often unintentionally transported in ship ballast water (Briggs 2007), so transport risk is likely to be related to shipping patterns. Links between trade and transport risk mean that patterns of introduction partly result from socio-economic factors. For instance, former communist countries in Eastern Europe were relatively isolated from international trade compared to countries in Western Europe, and have fewer non-native species (Chiron, Shirley & Kark 2010).

In order to deliberately transport a species, wild individuals have to be taken into captivity. This is more likely for species that come into close contact with humans, or species that are highly in demand in the pet or horticultural industry (Carrete & Tella 2008; Blackburn, Gaston & Parnell 2010; Bradley *et al.* 2012). Pre-establishment process can pose strong selection pressures that result in transported populations being a non-random subset of the native population (Carrete *et al.* 2012). For example, individuals with low neophobia are more likely to be caught in traps and therefore transported (Carrete *et al.* 2012).

Once transported to a new area, individuals of a species need to get from captivity into the wild (Blackburn *et al.* 2011). Sometimes individuals are deliberately introduced into the wild (Green 1997); indeed this is common practice with plants where non-native species are routinely planted in gardens (Blackburn *et al.* 2011).

It is important to consider patterns of transport and introduction when analysing later stages in the invasion framework. For example, because transport and introduction are non-random (Blackburn, Gaston & Parnell 2010, Figure 1.1), analyses seeking to identify traits associated with establishment success should use introduced species rather than all species as the source pool (Cassey *et al.* 2004). Species are more likely to become established when large numbers of propagules are released (Green 1997; Blackburn & Duncan 2001; Colautti, Grigorovich & MacIsaac 2006). As well as affecting which species become established, uneven introduction effort can affect the spatial patterns of non-native species distributions (Muñoz & Real 2006).

1.1.3 Establishment

After accounting for propagule pressure, the ability of a species to establish is influenced by a combination of biotic and abiotic factors. Climatic niche matching is common in non-native species (Petitpierre *et al.* 2012; Strubbe *et al.* 2013), with species more likely to establish successfully if introduced in areas that are climatically similar to their native range (Blackburn, Lockwood & Cassey 2009). Niche matching has been used in non-native species risk assessments to predict if a species is likely to become established (Thuiller *et al.* 2005; Kumschick & Richardson 2013). However, niche shifts have been observed in some biological invasions (Broennimann *et al.* 2007; Alexander & Edwards 2010; Gallagher *et al.* 2010).

Properties of the native community can influence the establishment success of non-native species. This biotic resistance has attracted considerable interest (Levine, Adler & Yelenik 2004). Experimental work has shown that diverse communities are harder to invade (Naeem *et al.* 2000). This is due to both sampling effects, where diverse communities are more likely to contain strong competitors, and greater resource use by diverse communities (Fargione & Tilman 2005). However, this pattern is seldom observed in natural communities (Fridley *et al.* 2007). Diverse native communities are often found in resource rich environments, and in areas that receive a strong propagule supply of both native and non-native species (Levine 2000; Fridley *et al.* 2007). Biotic resistance can be detected in patterns of establishment success, with non-native species that are distantly related to native species less likely to become established (Strauss, Webb & Salamin 2006). These studies have been carried out on plants, where species compete for space as well as limiting resources. The mechanisms of biotic resistance in animal invasions are less well known. Morphological overdispersion of non-native birds in oceanic islands has been interpreted as evidence for the importance of competition in influencing establishment success (Lockwood, Moulton & Anderson 1993).

However, Duncan and Blackburn (2002) found morphological overdispersion in introduced gamebirds in New Zealand when similar species where unlikely to have interacted, indicating that factors other than competition can result in morphological overdispersion.

The success of some non-native species has been linked to the absence of specialist predators (Keane & Crawley 2002) as well as release from competition (Batalha, Ramos & Cardoso 2013). Non-native species can actively release themselves from natural enemies; the non-native glassy-winged sharpshooter *Homalodisca coagulate* produces chemicals that kill native spiders in French Polynesia (Suttle & Hoddle 2006).

Considerable effort has been invested in the search for traits associated with successful invaders. In birds, traits such as migration that influence the impact of Allee effects have been found to reduce establishment success (Blackburn, Cassey & Lockwood 2009), while successfully established species tend to have lager bodies (Blackburn, Lockwood & Cassey 2009), larger brains and broader niches than unsuccessful species (Vall-llosera & Sol 2009). Work in plants has revealed complex interactions between the traits of non-native species and the environmental conditions in the invaded location; resource pulses can aid the establishment of non-native species (Tilman 2004), but this is especially true for fast growing (often ruderal) species (Mata, Haddad & Holyoak 2013). In contrast, resource pulses can hinder establishment success of slow growing but competitively dominant species (Mata, Haddad & Holyoak 2013).

It is important to note that current patterns of establishment success represent a snapshot in time; some species that had self-sustaining populations in the past are now extinct in their non-native range (Monticelli 2008b; Zenni & Nuñez 2013), while some species that are currently established will not persist indefinitely.

1.1.4 Impact

Globally, non-native species are a leading cause of biodiversity loss (Clavero & García-Berthou 2005; Clavero *et al.* 2009). Non-native species can reduce populations of native species by preying on them, competing with them, transmitting diseases to them and changing the characteristics of the ecosystems in which they live (Kumschick & Nentwig 2010). Some of the most dramatic impacts are through predation, especially when a non-native predator is introduced to a previously predator free island (Blackburn *et al.* 2004; Hilton & Cuthbert 2010). Non-native species can also have negative economic impacts. This can result from damage to crops and infrastructure, as well as through disruption of ecosystem services

(Pimentel, Zuniga & Morrison 2005; Lodge *et al.* 2006; Cook *et al.* 2007). For example, in Great Britain the combined damage and control costs of non-native species has been estimated at over £2 billion (GB non-native species secretariat 2013).

Non-native plants can affect native communities by changing fire regimes (Rossiter *et al.* 2003), which can lead to declines in native species and changes in nutrient cycling (Mack, D'Antonio & Ley 2001). Non-native plants can also outcompete native species (Suttle, Thomsen & Power 2007), and this can lead to declines in native consumers (Sullivan et al. in prep.). However, while dominance by non-native species could be because they are the drivers of change (and actively have a negative impact on native species), they could also be the passengers of change, responding to environmental change that favours non-native species over native species (Didham *et al.* 2005; MacDougall & Turkington 2005; D'Antonio, Hughes & Tunison 2011).

Non-native species are frequently suspected of having a negative impact by competing with native species (Lever 2005). Competition has traditionally been thought of as important in structuring bird communities (Cody 1974), with examples such as character displacement in wader bill morphology supporting this view (Recher 1966). In contrast, Newton (1998) considered that “bird communities are far from stable, being composed of loose assemblages of species which, for the most part, fluctuate independently of one another.” While many non-native bird species are considered likely to compete with natives (Lever 2005), interspecific competition has rarely been demonstrated to have a negative effect on native bird species (Blackburn, Lockwood & Cassey 2009).

The impacts of non-native species are not always negative. Non-native species can provide beneficial ecosystem services and resources that can be exploited by native species (Mattingly, Orrock & Reif 2012). For example, the red swamp crayfish *Procambarus clarkii* has a positive impact on native predators despite having a negative impact on native species in lower trophic levels (Tablado *et al.* 2010). Non-native species can be popular with the public (Strubbe, Shwartz & Chiron 2011), so provide cultural ecosystem services (Sullivan 2012). Indeed, some birdwatchers will undertake international travel to see established populations of non-native species (Monticelli 2008). The transport of non-native species has the potential to increase ecological diversity, both by increasing the alpha diversity of locations by adding non-native species without native extinction, and through speciation in introduced populations of non-native species (Thomas 2013).

Identifying whether species are likely to have a negative impact is crucial for determining management decisions (Lodge *et al.* 2006). Risk assessments, based on species traits and documented impacts elsewhere, provide one tool for assessing whether a potential or newly established non-native species is likely to have a negative impact (Andersen *et al.* 2004; Shirley & Kark 2009; Leung *et al.* 2012; Kumschick, Bacher & Blackburn 2013; Kumschick & Richardson 2013). While risk assessments are a promising tool for identifying species that are likely to have a negative impact, some frameworks have been criticised for an over-reliance on anecdotal reports of impact (Strubbe, Shwartz & Chiron 2011). There is demand from policy makers to have an evidence base to justify management decisions (Defra 2003). It is therefore desirable to have a framework for detecting the early impacts of non-native species while control is still feasible (Sullivan, Grundy & Franco 2014).

1.2 Species distribution modelling

By comparing environmental conditions in areas where a species is present to conditions where it is absent it is possible to quantify its environmental associations (Guisan & Zimmermann 2000). Interest in the field grew in the late 1980s and early 1990s, partly due to the advent of geographic information systems (GIS) to handle spatial environmental datasets (Miller, Stuart & Howell 1989; Hill 1991; Walker & Cocks 1991). The potential use of species distribution models (SDMs) to predict the potential distribution of non-native species has also been recognised since the early 1990s (Yee & Mitchell 1991). Interest in species distribution modelling has grown exponentially (Figure 1.2), and numerous methods are now available to fit species distribution models to both presence-absence and presence-only data (Elith & Leathwick 2009).

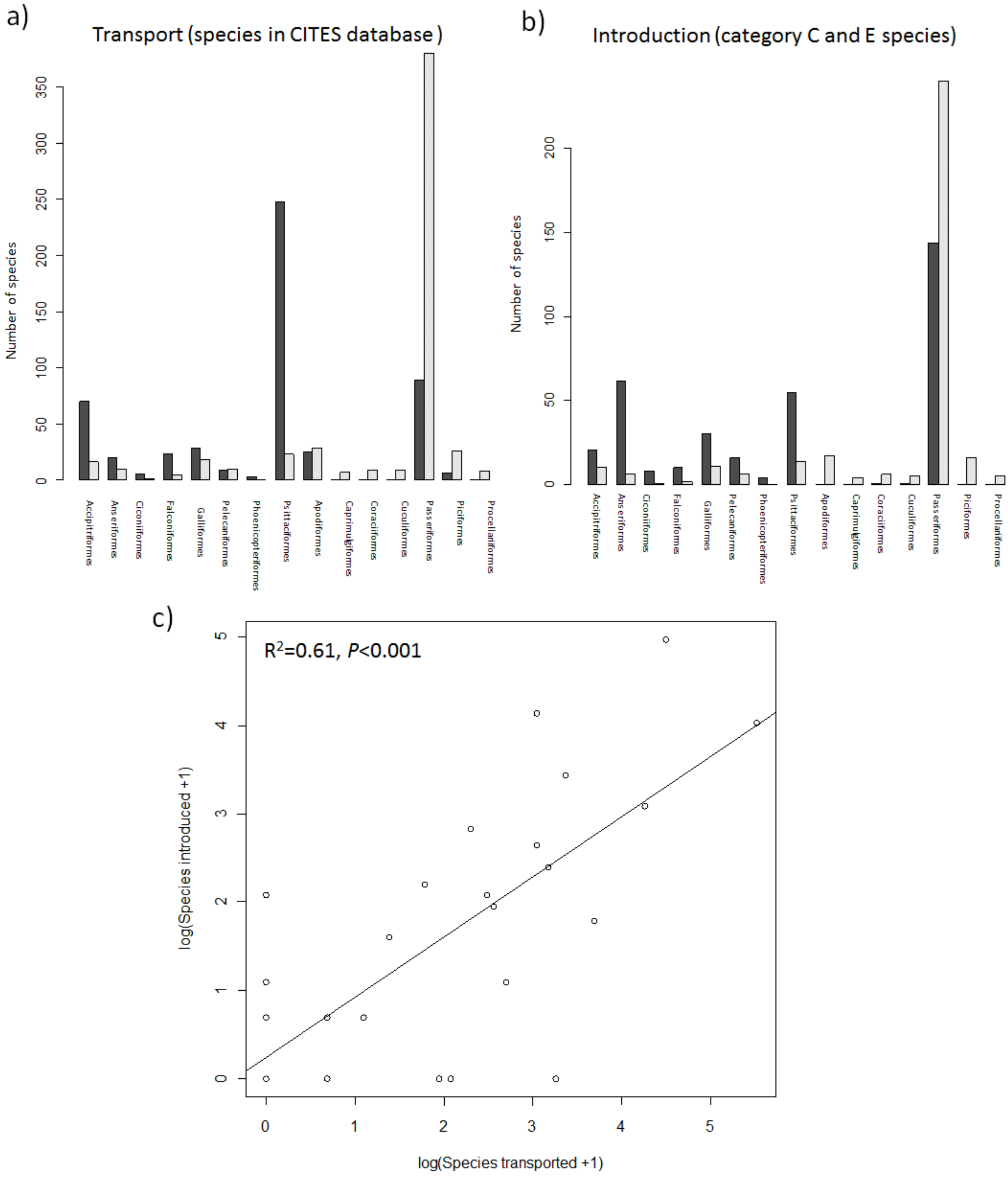


Figure 1.1. Non-random nature of introductions of non-native species to southern Europe. (A) Actual (dark bars) and expected (pale bars) number of species transported to southern Europe (defined as Portugal, Spain, France, Italy and Greece) in different bird families. Families in which there was a significant difference between observed and expected values have been shown. Expected values are median values from 10000 samples of all bird species. (B) As A, but showing numbers of introduced (i.e. category C and E) species in southern Europe. (C) Relationship between the logged number of species transported to southern

Europe in each family and the logged number of species introduced to southern Europe in each family. The positive relationship indicates that introduction is contingent on the preceding transport stage. Note that the CITES database is not a complete list of all species transported to a country, so there can be more species introduced than transported.

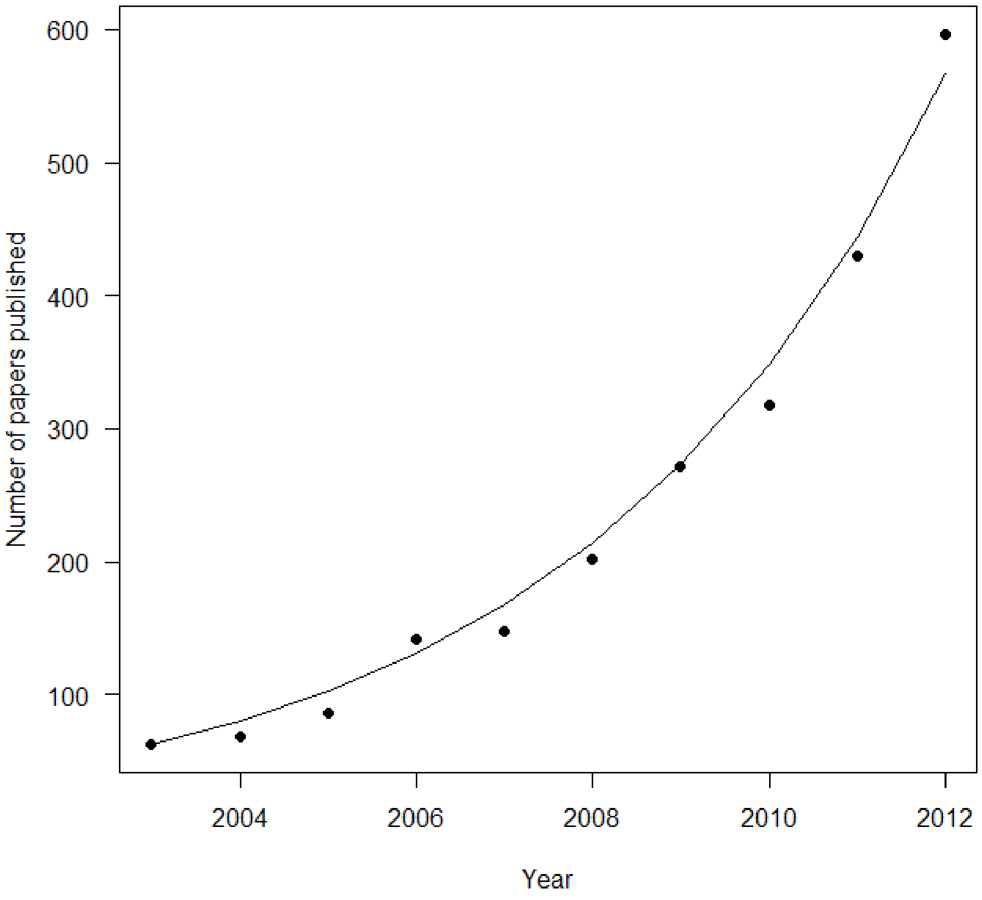


Figure 1.2. Number of publications returned by a web of science search with keywords ““species distribution” model” in each year between 2003 and 2012. The fitted line shows the number of publications predicted by an exponential model fitted to this data and parameterised by maximum likelihood (number of publications in year$_{t+1}$ = number of publications in year$_t$ x $e^{0.2\ 440284}$).

The distribution of a species can be linked to the environment using logistic generalised linear models (GLMs). These estimate the logit of the occurrence probability of a species as a function of linear relationships with input variables (Crawley 2007). More complex relationships can be modelled, for instance quadratic relationships can be modelled by including both the input variable and the squared input variable in the model. The complexity of the model has to be specified *a priori*, although fitted models can be simplified to remove polynomial terms (Crawley 2007). Generalised additive models (GAMs) provide greater flexibility for modelling non-linear relationships. These fit non-parametric smoothers to describe the relationship between the input variable and the logit occurrence probability, with the complexity of these smoothers determined by generalised cross-validation (Wood 2006). Other methods for fitting non-linear relationships have been developed. Multivariate adaptive regression splines (MARS) fit piecewise regressions, where the regression coefficient changes at different levels of the input variable (Friedman 1991). Gaussian random fields (GRaFs) use a Bayesian approach, using the observed species-environment relationship to update a prior expectation of the distribution of occurrence probabilities across the range of the input variable (Golding 2014).

Some SDM methods seek to determine the probability distribution of a species occurrence in environmental space. Support vector machines assume that occurrence data in environmental space come from an unknown distribution that potentially has high dimensionality. This unknown distribution is the probability of occurrence of a species in environmental space. Instead of estimating the probability density of this distribution, support vector machines provide a measure of support for the probability distribution being greater than zero at a given point in environmental space (Drake, Randin & Guisan 2006). Maximum entropy modelling (MAXENT) also takes observations of the occurrence of a species in environmental space (defined by pseudoabsence points) as a subset the incompletely observed true distribution of the species in environmental space. The MAXENT model is recursively updated to find the maximum entropy distribution of occurrence probabilities in environmental space, with the constraint that the expected values of environmental variables under this distribution matches the empirical average of that environmental variable (Phillips, Anderson & Schapire 2006). Interestingly, MAXENT has been shown to be mathematically equivalent with a Poisson generalised linear model (specifically a Poisson point-process model) (Renner & Warton 2013), so can be thought of as a specialised case of a generalised linear model.

A number of machine learning methods have been developed. These methods seek to develop rules to classify whether a species in present or absent in a location. Classification and regression trees (CART) recursively partitions data into smaller increasingly homogenous subsets based on splitting rules (Crawley 2007). This method has been further developed to use multiple CARTs. This new method is called random forests. Many simple CARTs are developed, and each CART uses input data to classify the occurrence of a species. These individual classifications are collated to produce the consensus classification of the random forest (Breiman 2001a). CART has also been extended with boosted regression trees (BRT). In this iterative process an initial CART is constructed, then additional CARTs are constructed to model residual variation (Li & Wang 2013). Genetic algorithm for rule set production (GARP) and artificial neural networks (ANN) both use complex interactive processes to construct models. In GARP, rules that determine the distribution of a species are randomly determined, the performance of the model is then tested on independent testing data, then rules in the models are shuffled with rules in good performing models more likely to be retained (Elith *et al.* 2006). In ANN, processing elements (neurons) are connected in potentially complicated networks. Each neuron receives one or more inputs (e.g. from other neurons), sums these inputs and after applying a function produces an output (potentially to another neuron) (Li & Wang 2013).

SDMs can be divided into groups based on their underlying philosophy and assumptions. The first group, stochastic data models, assume that observations result from a combination of an underlying model that can be parameterised and stochastic noise (Breiman 2001b). The philosophy of stochastic data modelling underlies most of the statistics used by ecologists (Crawley 2007), and models dominate the early species distribution modelling literature (Guisan & Zimmermann 2000). This group includes GLMs, GAMs, GRaFs, MARS and MAXENT. The second group, machine learning models, make no assumptions about the processes that generate observations, and instead seek to generate accurate predictions of the observations based on input variables (Breiman 2001b). These methods are popular for species distribution modelling (Elith *et al.* 2006), and have been employed in other fields in ecology, for example to predict extinction risk (Jones, Fielding & Sullivan 2006). This group includes CART, random forests, BRTs, GARPs and ANNs.

There has been considerable debate over which SDM method is best. Many empirical comparisons have been performed, often with conflicting results (Elith *et al.* 2006; Elith & Graham 2009; Rapacciuolo *et al.* 2012; Rodríguez-Rey, Jiménez-Valverde & Acevedo 2013). Models in the stochastic data model group tend to show similar performance, with the best model varying between comparisons. Random forests and BRTs can show different patterns in model performance to stochastic data models (Rapacciuolo *et al.* 2012), and sometimes are the best performing modelling methods (Elith *et al.* 2006). Models produced by machine learning techniques are often complex, although there is debate over whether random forests are simple (because each CART is simple) or complex (because there are multiple CARTs) (Li & Wang 2013). Simpler SDMs have been found to perform better when transferred to new conditions (Heikkinen, Marmion & Luoto 2012; Wenger & Olden 2012), with complicated machine learning methods potentially over-fitting to data. Averaging predictions from multiple modelling techniques has been suggested as a method of dealing with uncertainty resulting from choice of modelling method (Marmion *et al.* 2009).

SDMs that use presence-absence data assume that a species is absent from areas where it is not recorded, and that these absences arise because the environmental conditions in that area are unsuitable for the species. This assumption is frequently violated, as species are not always recorded when they are present (Lobo, Jiménez-Valverde & Hortal 2010; Kéry 2011), while dispersal limitation and biotic interactions can prevent species from occupying a location that would be suitable for it based on the modelled environmental variables (Guisan & Thuiller 2005; Gallien *et al.* 2010; Wisz *et al.* 2013). These absences will result in SDMs underestimating the prevalence of a species (Kéry 2011). These absences will have more serious consequences if they are correlated with environmental variables, for example causing SDMs to fail to characterise the shape of relationships properly (see Figure 1.3 for an example). Presence-only methods (e.g. MAXENT) do require absences, so may appear to avoid these problems (Phillips, Anderson & Schapire 2006). However, they are still vulnerable to bias in occurrence data caused by uneven recorder effort, dispersal limitation and biotic interactions (Phillips *et al.* 2009; Elith, Kearney & Phillips 2010). Methods that model these processes have been developed (Bierman *et al.* 2010; Hill 2011; Kéry 2011; Kissling *et al.* 2012; Sullivan *et al.* 2012). Hierarchical models offer one way of incorporating the multiple processes that shape species distributions. These involve two linked models; a process model links environmental variation to the species true occurrence status (which is a latent variable), and an observation model uses covariates of recording effort to link true occurrence status to the recorded occurrence status of a location (Kéry 2011). While hierarchical models show great promise, their complexity of implementation has probably hindered their uptake. There is therefore need to create simple flexible methods to account for the problems of uneven recorder effort, dispersal limitation and biotic interactions.

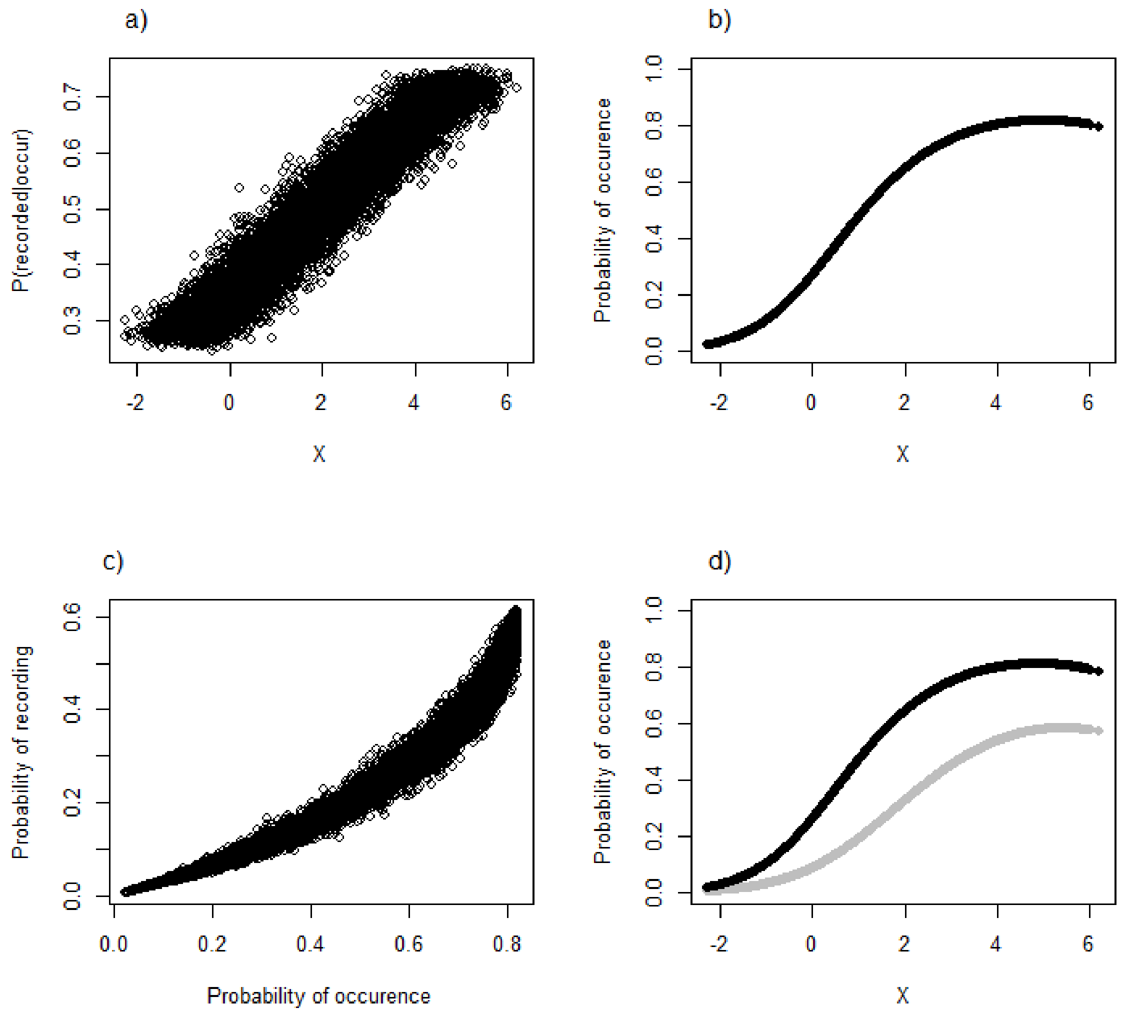


Figure 1.3. Effect of biased recording effort on the ability of SDMs to characterise response curves demonstrated using a simple simulation. (A) Recording effort (the conditional probability that a species is recorded given it occurs) was strongly (*r* = 0.95) correlated with the environmental variable X. (B) The occurrence of the simulated species was also positively related to that environmental variable. (C) This correlation meant that the relationship between the probability of occurrence of the species and the probability the species was recorded was noisy and non-linear. (D) SDMs (in this case a GLM) fitted to the recorded distribution of the species underestimated its prevalence compared to SDMs fitted to the true distribution, and did not estimate the shape of the response curve as accurately.

1.3 Climate change

Anthropogenic climate change is changing the abiotic conditions organisms experience (Parmesan & Yohe 2003). Biotic responses to climate change have broadly been consistent with climatic niche tracking, with pole-ward range expansion (Parmesan & Yohe 2003; Chen *et al.* 2011), changes in phenology (Fitter & Fitter 2002) and local extinctions of the trailing edges of species distributions (Franco *et al.* 2006; Thomas, Franco & Hill 2006). However, these patterns hide considerable variation between species (Chen *et al.* 2011), with the distributions of some species lagging behind their climatic niche (Menendez *et al.* 2006; Devictor *et al.* 2012). The effects of climate change can be complex (Walther 2010). Climate can alter the strength of species interactions (Memmott *et al.* 2007; Barton, Beckerman & Schmitz 2009), leading to cascading indirect effects (Barton & Schmitz 2009). Even if species interactions are unchanged, interactions with competitors or facilitators that respond strongly to climate can overturn species physiological responses to climate (Suttle, Thomsen & Power 2007). These indirect effects can result in different responses to ambient variation in climate and sustained directional climate change (Sullivan *et al.* in prep.). There is considerable variation in the effect of species interactions between communities (Adler, Dalgleish & Ellner 2012); in some communities they lead to idiosyncratic responses to climate (Suttle *et al.* in prep.), while other communities remain stable under climate change (Grime *et al.* 2008).

Although these complex responses to climate change have led to doubts being raised over the ability of correlative SDMs to predict species potential distributions under climate change (Schmitz *et al.* 2003; Higgins, O'Hara & Romermann 2012), they are widely used to assess how species will respond to climate change (Walther *et al.* 2009; Barbet-Massin, Thuiller & Jiguet 2010; Bellard *et al.* 2013). As well the challenges posed by complex responses of species to climate, which can sometimes be addressed using hybrid correlative and mechanistic models (Higgins, O'Hara & Romermann 2012; Higgins, Roemermann & O'Hara 2012; Kissling *et al.* 2012), climate change will create no analogue climates (Ohlemüller *et al.* 2006). Predicting the potential distribution of species under future climate change therefore sometimes requires extrapolating models beyond the range of training data (Wenger & Olden 2012). Interestingly, this problem is also encountered when constructing models of the potential distribution of non-native species, as climatic conditions in species non-native range are often non-analogous to those in their native range (Jimenez-Valverde *et al.* 2011).

Climate change may benefit many non-native species (Walther *et al.* 2009; Bradley *et al.* 2010; Bradley *et al.* 2012; Bellard *et al.* 2013), for example by removing abiotic constraints to establishment (Thomsen *et al.* 2006). As well as increasing the potential distribution of many non-native species (Bellard *et al.* 2013), resource pulses following extreme climatic events may promote community invasability (Tilman 2004; Diez *et al.* 2012). One potentially complex way climate could promote future invasions is by interacting with trade. In the USA, xerophytes are under-represented in the non-native flora, but demand for these species is rising following recent climate change (Bradley *et al.* 2012). This has resulted in the development of new horticultural trade partnerships, which is often followed by rapid transfer of non-native species (Levine 2003). These species have been selected because their climatic niche matches conditions in their non-native range, increasing the probability of establishment (Blackburn, Lockwood & Cassey 2009).

1.4 Non-native birds in the Iberian Peninsula

Although Mediterranean bird communities have been considered to be difficult to invade (Groves & Di Castro 1991) 23 species of non-native passerine bird have been recorded breeding in the Iberian Peninsula, of which six are considered to have self-sustaining populations (Santos, Clacell & Sol 2007; Monticelli 2008). The most widespread of these, the common waxbill *Estrilda astrild*, was introduced to the Lisbon area in 1964, and the Algarve by 1977, and has spread along the coast of Portugal, and inland into western Spain (Silva, Reino & Borralho 2002). This pattern of expansion is mirrored by several other less widespread non-native species which have become more abundant in recent years (Marti & de Moral 2003; Equipa Atlas 2008).

In addition to the common waxbill, substantial self-sustaining populations of yellow-crowned bishop *Euplectes afer,* red avadavat *Amandava amandava* and black-headed weaver *Ploceus melanocephalus* are also established in western Iberia. The common waxbill and yellow-crowned bishop are widespread in sub-Saharan Africa, while the black-headed weaver has a more restricted distribution across the central belt of Africa (Matias 2002). All these species are native to countries that were former Portuguese colonies, and are popular cage birds in Portugal. The red avadavat is native to the Indian subcontinent and south-east Asia, and shares a similar native distribution to a number of other munias and manikins that have been recorded as escapes in Portugal and Spain (Matias 2002).

Species distribution modelling has been carried on the common waxbill in Portugal, using a dataset of colonisation dates of 20km x 20 km UTM grid cells. Spatial factors were found to explain more variation than environmental factors (Silva, Reino & Borralho 2002; Reino 2005), even though this was modelled using nearest neighbour counts rather than applying a formal dispersal model. Common waxbills were more likely to occur in low lying areas with high humidity, high annual rainfall and warm temperatures (Reino 2005; Reino, Moya-Larano & Heitor 2009). Common waxbills may be expected to increase their range with climate change, as the risk of colonisation of a square increased by 47% for a 1°C increase in temperature (Reino, Moya-Larano & Heitor 2009). Land-use was not found to be significant in these studies, being removed in stepwise regression. However, only two land-use classes (river density and presence of wetland) were considered. A global species distribution model for the common waxbill using data from native and non-native ranges has been produced (Stiels *et al.* 2011), predicting substantial range expansion with climate change. However, occurrence data from the Iberian Peninsula was not used in this analysis.

In addition to the common waxbill, species distribution models predicting the potential future range have been produced for the monk parakeet *Myiopsittia monacus* (Muñoz & Real 2006) and red-billed leiothrix *Leiothix lutea* (Herrando *et al.* 2010). These studies found that the current distribution of species was closely linked to the location of potential release sites. The habitat associations of non-native birds have been studied in the Valencia area, where Estrilid finches where shown to be associated with marshland habitats while parrots (Psittacidae) were associated with urban parks (Murgui 2001; Murgui & Valentin 2003).

To date, little is known about the impact of non-native birds in the Iberian Peninsula on native bird species. Batalha *et al.* (2013) compared the ecological requirements and morphological traits of common waxbills and selected native species and concluded that the common waxbill occupied a marginal ecological niche, so was unlikely to have a negative impact on native species. There have been anecdotal reports of aggression between black-headed weavers *Ploceus melanocephalus* and native *Acrocephalus* warblers (Cardoso 2008), leading to suggestions that they could have a negative impact (Matias 2002).

1.5 Thesis outline

This thesis can be split to two sections. The first section (chapters two to five) is concerned with elucidating the environmental variables that constrain the distribution of non-native species, while the second section (chapters six and seven) investigates how non-native species interact with native species. The emphasis of this thesis is on developing methodology to predict the potential distribution and impacts of non-native species, although some chapters focus more strongly on understanding the ecology of the non-native birds in the Iberian Peninsula.

Modelling the potential distribution of non-native species is challenging, as non-native species are often dispersal limited, violating a key assumption of species distribution modelling (Gallien *et al.* 2010). Although this issue has been frequently acknowledged (Elith, Kearney & Phillips 2010; Václavík & Meentemeyer 2012), few methods have been proposed to address it (Dullinger *et al.* 2009; Gallien *et al.* 2012). In ***Chapter Two*** I develop a simple and flexible method, dispersal weighting, to incorporate dispersal information when fitting SDMs. The performance of dispersal weighting and existing methods that do not account for dispersal limitation are compared using simulations. I also compare the performance of the methods at modelling the distribution of the common waxbill.

The habitat associations of species whose distributions are not at equilibrium with their environment have been found to vary between range margins and the range core (Oliver *et al.* 2009). This could be due to dispersal or climate, but disentangling the two is complicated as they are often correlated. In ***Chapter Three*** I compare the habitat associations of the common waxbill along multiple expansion axes, thus ensuring the two are not correlated. I compare the effect of dispersal and climate on the species habitat associations.

SDMs of the potential distribution of non-native species are often constructed using distribution data from the species native range (Stiels *et al.* 2011b), for example because the species is insufficiently widespread in its non-native range to construct SDMs. Native trained SDMs often over-predict the potential distribution of non-native species (Jimenez-Valverde *et al.* 2011). In ***Chapter Four*** I develop methods to identify whether this apparent over-prediction is due to dispersal limitation, niche unfilling due to other limiting environmental variables or poor model specification. I also develop and test a framework for predicting when native trained SDMs will perform well at classifying which areas of a species non-native range are potentially suitable.

Considerable attention has been paid to identifying whether species exhibit niche shifts between their native and non-native distribution (Pearman *et al.* 2008). However, these studies have typically considered the non-native niche to be static. In ***Chapter Five***, I use data on the distribution of the common waxbill at multiple time points to investigate how its climatic niche has changed as it has spread through the Iberian Peninsula.

In the Iberian Peninsula, one particular land-use, rice fields, is particularly heavily invaded by non-native birds. Rice fields are a fairly recent land-use, and potentially offer resources that are underexploited by native species. In ***Chapter Six*** we test whether non-native species are better able to exploit the resources provided by rice fields, and whether this has facilitated their establishment.

Detecting the impacts of non-native species is easier when they are widespread, but by this point eradication is difficult (Lodge *et al.* 2006). There is therefore a trade-off between strength of evidence for there being an impact and feasibility of control. In ***Chapter Seven*** I develop a framework for detecting the early impacts of non-native species. I apply this framework to investigate whether the non-native black-headed weaver *Ploceus melanocephalus* has a negative impact through competition on two ecologically similar native species.

In ***Chapter Eight*** I bring together the findings of the other chapters and suggest future research areas that would enhance our ability to predict the potential distribution and impacts of non-native species.

References

Adler, P.B., Dalgleish, H.J. & Ellner, S.P. (2012) Forecasting plant community impacts of climate variability and change: when do competitive interactions matter? *Journal of Ecology,* **100,** 478-487.

Alexander, J.M. & Edwards, P.J. (2010) Limits to the niche and range margins of alien species. *Oikos,* **119,** 1377-1386.

Andersen, M.C., Adams, H., Hope, B. & Powell, M. (2004) Risk Assessment for Invasive Species. *Risk Analysis,* **24,** 787-793.

Barbet-Massin, M., Thuiller, W. & Jiguet, F. (2010) How much do we overestimate future local extinction rates when restricting the range of occurrence data in climate suitability models? *Ecography,* **33,** 878-886.

Barton, B.T., Beckerman, A.P. & Schmitz, O.J. (2009) Climate warming strengthens indirect interactions in an old-field food web. *Ecology,* **90,** 2346-2351.

Barton, B.T. & Schmitz, O.J. (2009) Experimental warming transforms multiple predator effects in a grassland food web. *Ecology Letters,* **12,** 1317-1325.

Batalha, H.R., Ramos, J.A. & Cardoso, G.C. (2013) A successful avian invasion occupies a marginal ecological niche. *Acta Oecologica,* **49,** 92-98.

Bellard, C., Thuiller, W., Leroy, B., Genovesi, P., Bakkenes, M. & Courchamp, F. (2013) Will climate change promote future invasions? *Glob Chang Biol*.

Bierman, S.M., Butler, A., Marion, G. & Kühn, I. (2010) Bayesian image restoration models for combining expert knowledge on recording activity with species distribution data. *Ecography,* **33,** 451-460.

Blackburn, T.M., Cassey, P., Duncan, R.P., Evans, K.L. & Gaston, K.J. (2004) Avian extinction and mammalian introductions on oceanic islands. *Science,* **305,** 1955-1958.

Blackburn, T.M., Cassey, P. & Lockwood, J.L. (2009) The role of species traits in the establishment success of exotic birds. *Global Change Biology,* **15,** 2852-2860.

Blackburn, T.M. & Duncan, R.P. (2001) Determinants of establishment success in introduced birds. *Nature,* **414,** 195-197.

Blackburn, T.M., Gaston, K.J. & Parnell, M. (2010) Changes in non-randomness in the expanding introduced avifauna of the world. *Ecography,* **33,** 168-174.

Blackburn, T.M., Lockwood, J.L. & Cassey, P. (2009) *Avian Invasions: The Ecology and Evolution on Exotic Birds*. Oxford University Press, Oxford.

Blackburn, T.M., Pysek, P., Bacher, S., Carlton, J.T., Duncan, R.P., Jarosik, V., Wilson, J.R. & Richardson, D.M. (2011) A proposed unified framework for biological invasions. *Trends Ecol Evol,* **26,** 333-339.

Bradley, B.A., Blumenthal, D.M., Early, R., Grosholz, E.D., Lawler, J.J., Miller, L.P., Sorte, C.J.B., D'Antonio, C.M., Diez, J.M., Dukes, J.S., Ibanez, I. & Olden, J.D. (2012) Global change, global trade, and the next wave of plant invasions. *Frontiers in Ecology and the Environment,* **10,** 20-28.

Bradley, B.A., Blumenthal, D.M., Wilcove, D.S. & Ziska, L.H. (2010) Predicting plant invasions in an era of global change. *Trends in Ecology & Evolution,* **25,** 310-318.

Breiman, L. (2001a) Random Forests. *Machine Learning,* **45,** 5-32.

Breiman, L. (2001b) Statistical Modeling: The Two Cultures. *Statistical Science,* **16,** 199-215.

Briggs, J.C. (2007) Marine biogeography and ecology: invasions and introductions. *Journal of Biogeography,* **34,** 193-198.

Broennimann, O., Treier, U.A., Muller-Scharer, H., Thuiller, W., Peterson, A.T. & Guisan, A. (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters,* **10,** 701-709.

Cardoso, H. (2008) Ploceus melanocephalus. *Atlas das Aves Nidificantes em Portugal* (ed. A. Equipa), pp. 540-541. ICNB, Lisbon.

Carrete, M., Edelaar, P., Blas, J., Serrano, D., Potti, J., Dingemanse, N.J. & Tella, J.L. (2012) Don't neglect pre-establishment individual selection in deliberate introductions. *Trends in Ecology & Evolution,* **27,** 67-68.

Carrete, M. & Tella, J.L. (2008) Wild-bird trade and exotic invasions: a new link of conservation concern? *Frontiers in Ecology and the Environment,* **6,** 207-211.

Cassey, P., Blackburn, T.M., Jones, K.E. & Lockwood, J.L. (2004) Mistakes in the analysis of exotic species establishment: source pool designation and correlates of introduction success among parrots (Aves: Psittaciformes) of the world. *Journal of Biogeography,* **31,** 277-284.

Chen, I.C., Hill, J.K., Ohlemüller, R., Roy, D.B. & Thomas, C.D. (2011) Rapid Range Shifts of Species Associated with High Levels of Climate Warming. *Science,* **333,** 1024-1026.

Chiron, F., Shirley, S.M. & Kark, S. (2010) Behind the Iron Curtain: Socio-economic and political factors shaped exotic bird introductions into Europe. *Biological Conservation,* **143,** 351-356.

Clavero, M., Brotons, L., Pons, P. & Sol, D. (2009) Prominent role of invasive species in avian biodiversity loss. *Biological Conservation,* **142,** 2043-2049.

Clavero, M. & García-Berthou, E. (2005) Invasive species are a leading cause of animal extinctions. *Trends in Ecology & Evolution,* **20,** 110.

Cody, M.I. (1974) *Competition and the Structure of Bird Communities*. Princeton University Press, Princeton.

Colautti, R.I., Grigorovich, I.A. & MacIsaac, H.J. (2006) Propagule pressure: A null model for biological invasions. *Biological Invasions,* **8,** 1023-1037.

Cook, D.C., Thomas, M.B., Cunningham, S.A., Anderson, D.L. & De Barro, P.J. (2007) Predicting the economic impact of an invasive species on an ecosystem service. *Ecological Applications,* **17,** 1832-1840.

Crawley, M.J. (2005) Ecology: Learning from the Aliens. *Science,* **310,** 623-624.

Crawley, M.J. (2007) *The R Book*. John Wiley & Sons, Chichester.

D'Antonio, C.M., Hughes, R.F. & Tunison, J.T. (2011) Long-term impacts of invasive grasses and subsequent fire in seasonally dry Hawaiian woodlands. *Ecological Applications,* **21,** 1617-1628.

Defra (2003) Review of non-native species policy: Report of the working group. Department for Environment, Food and Rural Affairs, London.

Devictor, V., van Swaay, C., Brereton, T., Brotons, L., Chamberlain, D., Heliölä, J., Herrando, S., Julliard, R., Kuussaari, M., Lindström, Å., Reif, J., Roy, D., Schweiger, O., Settele, J., Stefanescu, C., Van Strien, A., Van Turnhout, C., Vermouzek, Z., WallisDeVries, M., Wynhoff, I. & Jiguet, F. (2012) Differences in the climatic debts of birds and butterflies at a continental scale. *Nature Climate Change,* **2,** 121-124.

Didham, R.K., Tylianakis, J.M., Hutchison, M.A., Ewers, R.M. & Gemmell, N.J. (2005) Are invasive species the drivers of ecological change? *Trends in Ecology & Evolution,* **20,** 470-474.

Diez, J.M., D'Antonio, C.M., Dukes, J.S., Grosholz, E.D., Olden, J.D., Sorte, C.J.B., Blumenthal, D.M., Bradley, B.A., Early, R., Ibáñez, I., Jones, S.J., Lawler, J.J. & Miller, L.P. (2012) Will extreme climatic events facilitate biological invasions? *Frontiers in Ecology and the Environment,* **10,** 249-257.

Drake, J.M., Randin, C. & Guisan, A. (2006) Modelling ecological niches with support vector machines. *Journal of Applied Ecology,* **43,** 424-432.

Dullinger, S., Kleinbauer, I., Peterseil, J., Smolik, M. & Essl, F. (2009) Niche based distribution modelling of an invasive alien plant: effects of population status, propagule pressure and invasion history. *Biological Invasions,* **11,** 2401-2414.

Duncan, R.P. & Blackburn, T.M. (2002) Morphological over-dispersion in game birds (Aves : Galliformes) successfully introduced to New Zealand was not caused by interspecific competition. *Evolutionary Ecology Research,* **4,** 551-561.

Elith, J. & Graham, C.H. (2009) Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography,* **32,** 66-77.

Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberon, J., Williams, S., Wisz, M.S. & Zimmermann, N.E. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography,* **29,** 129-151.

Elith, J., Kearney, M. & Phillips, S. (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution,* **1,** 330-342.

Elith, J. & Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics,* **40,** 677-697.

Equipa Atlas (2008) *Atlas das Aves Nidificantes em Portugal (1999-2005)*. Instituto da Conservacao da Natureza e da Biodiversividade, Sociedade Portuguesa para o Estudo das Aves, Parque Natural da Madeira e Secretaria Regional do Ambiente e do Mar, Lisbon.

Fargione, J.E. & Tilman, D. (2005) Diversity decreases invasion via both sampling and complementarity effects. *Ecology Letters,* **8,** 604-611.

Fitter, A.H. & Fitter, R.S.R. (2002) Rapid changes in flowering time in British plants. *Science,* **296,** 1689-1691.

Franco, A.M.A., Hill, J.K., Kitschke, C., Collingham, Y.C., Roy, D.B., Fox, R., Huntley, B. & Thomas, C.D. (2006) Impacts of climate warming and habitat loss on extinctions at species' low-latitude range boundaries. *Global Change Biology,* **12,** 1545-1553.

Fridley, J.D., Stachowicz, J.J., Naeem, S., Sax, D.F., Seabloom, E.W., Smith, M.D., Stohlgren, T.J., Tilman, D. & Von Holle, B. (2007) The invasion paradox: Reconciling pattern and process in species invasions. *Ecology,* **88,** 3-17.

Friedman, J.H. (1991) Multivariate Adaptive Regression Splines. 1-67.

Gallagher, R.V., Beaumont, L.J., Hughes, L. & Leishman, M.R. (2010) Evidence for climatic niche and biome shifts between native and novel ranges in plant species introduced to Australia. *Journal of Ecology,* **98,** 790-799.

Gallien, L., Douzet, R., Pratte, S., Zimmermann, N.E. & Thuiller, W. (2012) Invasive species distribution models - how violating the equilibrium assumption can create new insights. *Global Ecology and Biogeography,* **21,** 1126-1136.

Gallien, L., Münkemüller, T., Albert, C.H., Boulangeat, I. & Thuiller, W. (2010) Predicting potential distributions of invasive species: where to go from here? *Diversity and Distributions,* **16,** 331-342.

GB non-native species secretariat (2013) http://www.nonnativespecies.org/home/index.cfm.

Golding, N. (2014) GRaF: Species distribution modelling using latent Gaussian random fields.

Green, R.E. (1997) The influence of numbers released on the outcome of attempts to introduce exotic bird species to New Zealand. *Journal of Animal Ecology,* **66,** 25-35.

Grime, J.P., Fridley, J.D., Askew, A.P., Thompson, K., Hodgson, J.G. & Bennett, C.R. (2008) Long-term resistance to simulated climate change in an infertile grassland. *Proceedings of the National Academy of Sciences of the United States of America,* **105,** 10028-10032.

Groves, R.H. & Di Castro, F. (1991) *Biogeography of Mediterranean Invasions*. Cambridge University Press, Cambridge.

Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters,* **8,** 993-1009.

Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling,* **135,** 147-186.

Harris, S. & Yalden, D. (2008) Mammals of the British Isles: Handbook. The Mammal Society, Southampton.

Heikkinen, R.K., Marmion, M. & Luoto, M. (2012) Does the interpolation accuracy of species distribution models come at the expense of transferability? *Ecography,* **35,** 276-288.

Herrando, S., Llimona, F., Brotons, L. & Quesada, J. (2010) A new exotic bird in Europe: recent spread and potential range of Red-billed Leiothrix Leiothrix lutea in Catalonia (northeast Iberian Peninsula). *Bird Study,* **57,** 226-235.

Higgins, S.I., O'Hara, R.B. & Romermann, C. (2012) A niche for biology in species distribution models. *Journal of Biogeography,* **39,** 2091-2095.

Higgins, S.I., Roemermann, C. & O'Hara, R.B. (2012) Special Issue: The ecological niche as a window to biodiversity. *Journal of Biogeography,* **39,** 2089-2089.

Hill, M.O. (1991) Patterns of species distribution in britain elucidated by canonical correspondence-analysis. *Journal of Biogeography,* **18,** 247-255.

Hill, M.O. (2011) Local frequency as a key to interpreting species occurrence data when recording effort is not known. *Methods in Ecology and Evolution*.

Hilton, G.M. & Cuthbert, R.J. (2010) The catastrophic impact of invasive mammalian predators on birds of the UK Overseas Territories: a review and synthesis. *Ibis,* **152,** 443-458.

Holt, B., Lessard, J.P., Borregaard, M.K., Fritz, S.A., Araujo, M.B., Dimitrov, D., Fabre, P.H., Graham, C.H., Graves, G.R., Jonsson, K.A., Nogues-Bravo, D., Wang, Z.H., Whittaker, R.J., Fjeldsa, J. & Rahbek, C. (2013) An update of Wallace's zoogeographic regions of the world. *Science,* **339,** 74-78.

Holt, R.D. (2009) Bringing the Hutchinsonian niche into the 21st century: Ecological and evolutionary perspectives. *Proceedings of the National Academy of Sciences,* **106,** 19659-19665.

Hulme, P.E. (2009) Trade, transport and trouble: managing invasive species pathways in an era of globalization. *Journal of Applied Ecology,* **46,** 10-18.

Jimenez-Valverde, A., Peterson, A.T., Soberon, J., Overton, J.M., Aragon, P. & Lobo, J.M. (2011) Use of niche models in invasive species risk assessments. *Biological Invasions,* **13,** 2785-2797.

Jones, M.J., Fielding, A. & Sullivan, M. (2006) Analysing extinction risk in parrots using decision trees. *Biodiversity and Conservation,* **15,** 1993-2007.

Keane, R.M. & Crawley, M.J. (2002) Exotic plant invasions and the enemy release hypothesis. *Trends in Ecology & Evolution,* **17,** 164-170.

Kissling, W.D., Dormann, C.F., Groeneveld, J., Hickler, T., Kuhn, I., McInerny, G.J., Montoya, J.M., Romermann, C., Schiffers, K., Schurr, F.M., Singer, A., Svenning, J.C., Zimmermann, N.E. & O'Hara, R.B. (2012) Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. *Journal of Biogeography,* **39,** 2163-2178.

Kumschick, S., Bacher, S. & Blackburn, T.M. (2013) What determines the impact of alien birds and mammals in Europe? *Biological Invasions,* **15,** 785-797.

Kumschick, S. & Nentwig, W. (2010) Some alien birds have as severe an impact as the most effectual alien mammals in Europe. *Biological Conservation,* **143,** 2757-2762.

Kumschick, S. & Richardson, D.M. (2013) Species-based risk assessments for biological invasions: advances and challenges. *Diversity and Distributions,* **19,** 1095-1105.

Kéry, M. (2011) Towards the modelling of true species distributions. *Journal of Biogeography,* **38,** 617-618.

Leung, B., Roura-Pascual, N., Bacher, S., Heikkila, J., Brotons, L., Burgman, M.A., Dehnen-Schmutz, K., Essl, F., Hulme, P.E., Richardson, D.M., Sol, D. & Vila, M. (2012) TEASIng apart alien species risk assessments: a framework for best practices. *Ecology Letters,* **15,** 1475-1493.

Lever, C. (2005) *Naturalised Birds of the World*. T & AD Poyser, London.

Levine, J.M. (2000) Species Diversity and Biological Invasions: Relating Local Process to Community Pattern. *Science,* **288,** 852-854.

Levine, J.M., Adler, P.B. & Yelenik, S.G. (2004) A meta-analysis of biotic resistance to exotic plant invasions. *Ecology Letters,* **7,** 975-989.

Levine, J.M.a.D.A.C.M. (2003) Forecasting Biological Invasions with Increasing International Trade. *Conservation Biology,* **17,** 322-326.

Li, X. & Wang, Y. (2013) Applying various algorithms for species distribution modelling. *Integrative Zoology,* **8,** 124-135.

Lobo, J.M., Jiménez-Valverde, A. & Hortal, J. (2010) The uncertain nature of absences and their importance in species distribution modelling. *Ecography,* **33,** 103-114.

Lockwood, J.L., Moulton, M.P. & Anderson, S.K. (1993) Morphological Assortment and the Assembly of Communities of Introduced Passeriforms on Oceanic Islands: Tahiti Versus Oahu. *The American Naturalist,* **141,** 398-408.

Lodge, D.M., Williams, S., MacIsaac, H.J., Hayes, K.R., Leung, B., Reichard, S., Mack, R.N., Moyle, P.B., Smith, M., Andow, D.A., Carlton, J.T. & McMichael, A. (2006) Biological invasions: Recommendations for US policy and management. *Ecological Applications,* **16,** 2035-2054.

MacDougall, A.S. & Turkington, R. (2005) Are invasive species the drivers or passengers of change in degraded ecosystems? *Ecology,* **86,** 42-55.

Mack, M.C., D'Antonio, C.M. & Ley, R.E. (2001) Alteration of ecosystem nitrogen dynamics by exotic plants: A case study of C-4 grasses in Hawaii. *Ecological Applications,* **11,** 1323-1335.

Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K. & Thuiller, W. (2009) Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions,* **15,** 59-69.

Marti, R. & de Moral, J.C. (2003) Atlas de las Aves Reproductoras de Espana. Direccion General de Conservacion de la Naturaleza-Sociedad Espanola de Ornithologia, Madrid.

Mata, T.M., Haddad, N.M. & Holyoak, M. (2013) How invader traits interact with resident communities and resource availability to determine invasion success. *Oikos,* **122,** 149-160.

Matias, R. (2002) *Aves exóticas que nidificam em Portugal continental*. ICNB, Lisbon.

Mattingly, W.B., Orrock, J.L. & Reif, N.T. (2012) Dendroecological analysis reveals long-term, positive effects of an introduced understory plant on canopy tree growth. *Biological Invasions,* **14,** 2639-2646.

McInnes, L., Purvis, A. & Orme, C.D.L. (2009) Where do species' geographic ranges stop and why? Landscape impermeability and the Afrotropical avifauna. *Proceedings of the Royal Society B: Biological Sciences,* **276,** 3063-3070.

Memmott, J., Craze, P.G., Waser, N.M. & Price, M.V. (2007) Global warming and the disruption of plant-pollinator interactions. *Ecol Lett,* **10,** 710-717.

Menendez, R., Megias, A.G., Hill, J.K., Braschler, B., Willis, S.G., Collingham, Y., Fox, R., Roy, D.B. & Thomas, C.D. (2006) Species richness changes lag behind climate change. *Proceedings of the Royal Society B-Biological Sciences,* **273,** 1465-1470.

Miller, R.I., Stuart, S.N. & Howell, K.M. (1989) A methodology for analyzing rare species distribution patterns utilizing GIS technology: The rare birds of Tanzania. *Landscape Ecology,* **2,** 173-189.

Mitchell, C., Hearn, R. & Stroud, D. (2012) The merging populations of Greylag Geese breeding in Britain. *British Birds,* **105**.

Monticelli, D. (2008) Finding introduced birds near Lisbon, Portugal. *Birdwing World,* **21,** 203-206.

Murgui, E. (2001) Factors influencing the distribution of exotic bird species in Comunidad Valenciana (Spain). *Ardeola,* **48,** 149-160.

Murgui, E. & Valentin, A. (2003) Relacion entre las caracteristicas del paisaje urbano y la xomunidad de aves introducidas en la Ciudad de Valencia (Espana) *Ardeola,* **50,** 201-214.

Muñoz, A.-R. & Real, R. (2006) Assessing the potential range expansion of the exotic monk parakeet in Spain. *Diversity and Distributions,* **12,** 656-665.

Naeem, S., Knops, J.M.H., Tilman, D., Howe, K.M., Kennedy, T. & Gale, S. (2000) Plant diversity increases resistance to invasion in the absence of covarying extrinsic factors. *Oikos,* **91,** 97-108.

Newton, I. (1998) *Population Limitation in Birds*. Academic Press, San Diego.

Ohlemüller, R., Gritti, E.S., Sykes, M.T. & Thomas, C.D. (2006) Towards European climate risk surfaces: the extent and distribution of analogous and non-analogous climates 1931–2100. *Global Ecology and Biogeography,* **15,** 395-405.

Olalla-Tarraga, M.A., McInnes, L., Bini, L.M., Diniz-Filho, J.A.F., Fritz, S.A., Hawkins, B.A., Hortal, J., Orme, C.D.L., Rahbek, C., Rodriguez, M.A. & Purvis, A. (2011) Climatic niche conservatism and the evolutionary dynamics in species range boundaries: global congruence across mammals and amphibians. *Journal of Biogeography,* **38,** 2237-2247.

Oliver, T., Hill, J.K., Thomas, C.D., Brereton, T. & Roy, D.B. (2009) Changes in habitat specificity of species at their climatic range boundaries. *Ecology Letters,* **12,** 1091-1102.

Parmesan, C. & Yohe, G. (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature,* **421,** 37-42.

Pearman, P.B., Guisan, A., Broennimann, O. & Randin, C.F. (2008) Niche dynamics in space and time. *Trends in Ecology & Evolution,* **23,** 149-158.

Petitpierre, B., Kueffer, C., Broennimann, O., Randin, C., Daehler, C. & Guisan, A. (2012) Climatic Niche Shifts Are Rare Among Terrestrial Plant Invaders. *Science,* **335,** 1344-1348.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling,* **190,** 231-259.

Phillips, S.J., Dudik, M., Elith, J., Graham, C.H., Lehmann, A., Leathwick, J. & Simon, F. (2009) Sample Selection Bias and Presence-Only Distribution Models: Implications for Background and Pseudo-Absence Data. *Ecological Applications,* **19,** 181-197.

Pimentel, D., Zuniga, R. & Morrison, D. (2005) Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics,* **52,** 273-288.

Preston, C., Pearman, D. & Hall, A. (2004) Archaeophytes in Britain. *Botanical Journal of the Linnean Society,* **145,** 257-294.

Rapacciuolo, G., Roy, D.B., Gillings, S., Fox, R., Walker, K. & Purvis, A. (2012) Climatic Associations of British Species Distributions Show Good Transferability in Time but Low Predictive Accuracy for Range Change. *PLoS ONE,* **7,** e40212.

Recher, H.F. (1966) Some Aspects of the Ecology of Migrant Shorebirds. *Ecology,* **47,** 393-407.

Reino, L. (2005) Variation partitioning for range expansion of an introduced species: the common waxbill Estrilda astrild in Portugal. *Journal of Ornithology,* **146,** 377-382.

Reino, L., Moya-Larano, J. & Heitor, A.C. (2009) Using survival regression to study patterns of expansion of invasive species: will the common waxbill expand with global warming? *Ecography,* **32,** 237-246.

Renner, I.W. & Warton, D.I. (2013) Equivalence of MAXENT and Poisson Point Process Models for Species Distribution Modeling in Ecology. *Biometrics,* **69,** 274-281.

Rodríguez-Rey, M., Jiménez-Valverde, A. & Acevedo, P. (2013) Species distribution models predict range expansion better than chance but not better than a simple dispersal model. *Ecological Modelling,* **256,** 1-5.

Rossiter, N.A., Setterfield, S.A., Douglas, M.M. & Hutley, L.B. (2003) Testing the grass-fire cycle: alien grass invasion in the tropical savannas of northern Australia. *Diversity and Distributions,* **9,** 169-176.

Santos, D.M., Clacell, J. & Sol, D. (2007) Lista europea de aves: categorías C y E. SEO.

Schmitz, O.J., Post, E., Burns, C.E. & Johnston, K.M. (2003) Ecosystem responses to global climate change: Moving beyond color mapping. *Bioscience,* **53,** 1199-1205.

Shirley, S.M. & Kark, S. (2009) The role of species traits and taxonomic patterns in alien bird impacts. *Global Ecology and Biogeography,* **18,** 450-459.

Silva, T., Reino, L.M. & Borralho, R. (2002) A model for range expansion of an introduced species: the common waxbill Estrilda astrild in Portugal. *Diversity and Distributions,* **8,** 319-326.

Simberloff, D. (2011) Non-natives: 141 scientists object. *Nature,* **475,** 36.

Soberon, J. & Nakamura, M. (2009) Niches and distributional areas: Concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences of the United States of America,* **106,** 19644-19650.

Stiels, D., Schidelko, K., Engler, J., van den Elzen, R. & Rödder, D. (2011) Predicting the potential distribution of the invasive Common Waxbill *Estrilda astrild* (Passeriformes: Estrildidae). *Journal of Ornithology***,** 1-12.

Strauss, S.Y., Webb, C.O. & Salamin, N. (2006) Exotic taxa less related to native species are more invasive. *Proceedings of the National Academy of Sciences,* **103,** 5841-5845.

Strubbe, D., Broennimann, O., Chiron, F. & Matthysen, E. (2013) Niche conservatism in non-native birds in Europe: niche unfilling rather than niche expansion. *Global Ecology and Biogeography***, 22**, 962–970.

Strubbe, D., Shwartz, A. & Chiron, F. (2011) Concerns regarding the scientific evidence informing impact risk assessment and management recommendations for invasive birds. *Biological Conservation,* **144,** 2112-2118.

Sullivan, M.J.P., Davies, R.G., Reino, L. & Franco, A.M.A. (2012) Using dispersal information to model the species–environment relationship of spreading non-native species. *Methods in Ecology and Evolution,* **3,** 870-879.

Sullivan, M.J.P., Grundy, J. & Franco, A.M.A. (2014) Assessing the impacts of the non-native Black-headed Weaver on native Acrocephalus warblers. *Ibis,* **156,** 231-232.

Sullivan, M.P. (2012) Ecosystem services: do we need birds? *Ibis,* **154,** 884-886.

Suttle, K.B. & Hoddle, M.S. (2006) Engineering enemy-free space: an invasive pest that kills its predators. *Biological Invasions,* **8,** 639-649.

Suttle, K.B., Thomsen, M.A. & Power, M.E. (2007) Species interactions reverse grassland responses to changing climate. *Science,* **315,** 640-642.

Tablado, Z., Tella, J.L., Sanchez-Zapata, J.A. & Hiraldo, F. (2010) The Paradox of the Long-Term Positive Effects of a North American Crayfish on a European Community of Predators. *Conservation Biology,* **24,** 1230-1238.

Thomas, C.D. (2013) The Anthropocene could raise biological diversity. *Nature,* **502,** 7-7.

Thomas, C.D., Franco, A.M.A. & Hill, J.K. (2006) Range retractions and extinction in the face of climate warming. *Trends in Ecology & Evolution,* **21,** 415-416.

Thomsen, M.A., D'Antonio, C.M., Suttle, K.B. & Sousa, W.P. (2006) Ecological resistance, seed density and their interactions determine patterns of invasion in a California coastal grassland. *Ecology Letters,* **9,** 160-170.

Thuiller, W., Richardson, D.M., Pysek, P., Midgley, G.F., Hughes, G.O. & Rouget, M. (2005) Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology,* **11,** 2234-2250.

Tilman, D. (2004) Niche tradeoffs, neutrality, and community structure: A stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences of the United States of America,* **101,** 10854-10861.

Tilman, D. (2011) Diversification, Biotic Interchange, and the Universal Trade-Off Hypothesis. *American Naturalist,* **178,** 355-371.

Vall-llosera, M. & Sol, D. (2009) A global risk assessment for the success of bird introductions. *Journal of Applied Ecology,* **46,** 787-795.

Valtuena, F.J., Preston, C.D. & Kadereit, J.W. (2011) Evolutionary significance of the invasion of introduced populations into the native range of Meconopsis cambrica. *Molecular Ecology,* **20,** 4318-4331.

Václavík, T. & Meentemeyer, R.K. (2012) Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. *Diversity and Distributions,* **18,** 73-83.

Walker, P.A. & Cocks, K.D. (1991) Habitat - a procedure for modeling a disjoint environmental envelope for a plant or animal species. *Global Ecology and Biogeography Letters,* **1,** 108-118.

Walther, G.-R. (2010) Community and ecosystem responses to recent climate change. *Philosophical Transactions of the Royal Society B-Biological Sciences,* **365,** 2019-2024.

Walther, G.-R., Roques, A., Hulme, P.E., Sykes, M.T., Pyšek, P., Kühn, I., Zobel, M., Bacher, S., Botta-Dukát, Z., Bugmann, H., Czúcz, B., Dauber, J., Hickler, T., Jarošík, V., Kenis, M., Klotz, S., Minchin, D., Moora, M., Nentwig, W., Ott, J., Panov, V.E., Reineking, B., Robinet, C., Semenchenko, V., Solarz, W., Thuiller, W., Vilà, M., Vohland, K. & Settele, J. (2009) Alien species in a warmer world: risks and opportunities. *Trends in Ecology and Evolution,* **24,** 686-693.

Wenger, S.J. & Olden, J.D. (2012) Assessing transferability of ecological models: an underappreciated aspect of statistical validation. *Methods in Ecology and Evolution,* **3,** 260-267.

Williamson, M.H. (1996) *Biological Invasions*. Springer.

Wisz, M.S., Pottier, J., Kissling, W.D., Pellissier, L., Lenoir, J., Damgaard, C.F., Dormann, C.F., Forchhammer, M.C., Grytnes, J.A., Guisan, A., Heikkinen, R.K., Hoye, T.T., Kuhn, I., Luoto, M., Maiorano, L., Nilsson, M.C., Normand, S., Ockinger, E., Schmidt, N.M., Termansen, M., Timmermann, A., Wardle, D.A., Aastrup, P. & Svenning, J.C. (2013) The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological Reviews,* **88,** 15-30.

Wood, S.N. (2006) *Generalized Additive Models: An Introduction with R*. Chapman and Hall, London.

Yee, T.W. & Mitchell, N.D. (1991) Generalized additive-models in plant ecology. *Journal of Vegetation Science,* **2,** 587-602.

Zenni, R.D. & Nuñez, M.A. (2013) The elephant in the room: the role of failed invasions in understanding invasion biology. *Oikos***, 122**, pages 801–815.

Zisenis, M. (2012) EU DAISIE Research Project: Wanted-death penalty to keep native species Competitive? *Journal of Agricultural & Environmental Ethics,* **25,** 597-606.

2 Using dispersal information to model the species-environment relationship of spreading non-native species

Published as Sullivan, M.J.P., Davies, R.G., Reino, L. & Franco, A.M.A. (2012) Using dispersal information to model the species–environment relationship of spreading non-native species. Methods in Ecology and Evolution, ***3,*** *870-879.*

Non-native species can be major drivers of biodiversity loss and cause economic damage. Predicting the potential distribution of a non-native species, and understanding the environmental factors that limit this distribution is useful for informing their potential management. This is often done using species distribution models (SDMs) that attempt to classify grid-cells as suitable or unsuitable for a species based on a set of environmental covariates. A key assumption of SDMs is that a species is in equilibrium with its environment. Spreading non-native species often violate this assumption due to dispersal limitation. We present a simple method for dealing with this problem; dispersal weighting (DW). This uses the probability that a species can disperse to a grid-cell to weight a SDM. We use simulations to compare the ability of DW and unweighted models at parameterising the true species-environment relationship (SER) of a simulated species, and to test their ability at predicting the future distribution of this species. We investigate how varying the degree of spatial autocorrelation in explanatory variables affects the performance of the methods. DW models outperformed unweighted models at parameterising the SER, and at predicting the future distribution of the species when dispersal probabilities were incorporated into the model predictions. Unweighted models had a stronger tendency than DW models to overestimate the magnitude of relationships with spatially autocorrelated explanatory variables, but underestimate the magnitude of relationships with randomly distributed variables. We then applied our method to a real case study, using it to model the distribution of the non-native common waxbill *Estrilda astrild* in the Iberian Peninsula as a function of climate and land-use variables. The relative performance of DW and unweighted models reflected the results of the simulation. We conclude that DW models perform better than unweighted models at modelling the true SER of non-native species, and recommend using DW whenever enough data exists to create a dispersal model.

2.1 *Introduction*

Human-assisted dispersal has allowed species to cross biogeographic barriers, introducing them to new environments where they interact with novel species assemblages. These non-native species can have negative impacts on native biodiversity (Williamson 1996), and can cause economic damage by becoming pests (Pimentel, Zuniga & Morrison 2005) or disrupting ecosystem services (Cook *et al.* 2007). In order to evaluate the potential impacts of these species, and devise management strategies to control them, it is useful to be able to predict their potential distribution and understand the environmental factors that limit this distribution. Species distribution models (SDMs) have often been employed to do this (Real *et al.* 2008; Strubbe & Matthysen 2009). Where presence-absence data are available, records of non-native species can be mapped onto a grid, and models use environmental covariates to discriminate between grid-cells that are occupied and unoccupied. SDMs assume that the species being modelled is at equilibrium with the environment (Guisan & Thuiller 2005), so unoccupied grid-cells are unsuitable for the species. This assumption is likely to be violated by spreading non-native species, which have yet to reach all environmentally suitable areas (Václavík & Meentemeyer 2012), and also by range-shifting species responding to environmental change (Elith, Kearney & Phillips 2010), as dispersal limitation may prevent them from keeping pace with the movement of suitable environmental conditions (Menéndez *et al.* 2006; Brooker *et al.* 2007). Spreading species can therefore be absent from a grid-cell due to low environmental suitability or dispersal limitation. The spatial structure of explanatory variables may interact with dispersal limitation to affect model inference (Václavík, Kupfer & Meentemeyer 2012); for example environmental variables that do not causally influence the distribution of a species may be erroneously identified as limiting the distribution if they occur on a gradient aligned to species' axis of dispersal.

The need to account for dispersal limitation when modelling the distribution of non-native and range-shifting species has been recognised (Peterson 2003; Guisan & Thuiller 2005; Gallien *et al.* 2010). Invasion dynamics have been simulated using dispersal models that incorporate environmental suitability (Smolik *et al.* 2010; Travis *et al.* 2011), and dispersal models have been used to produce realistic predictions of species' distributions under climate change scenarios (Engler & Guisan 2009). Despite this, there are few examples of dispersal models being used to influence the fitting of species distribution models. Several studies (e.g. Muñoz & Real 2006; Dullinger *et al.* 2009) have used covariates such as roads that might be related to the transport and introduction of non-native species as proxies for dispersal, while Václavík and Meentemeyer (2009) used propagule pressure calculated from a dispersal model as a covariate. The most direct approach to dealing with the problem of absences due to dispersal limitation was by Elith et al. (2010) who estimated the maximum area a non-native species could have spread to, and restricted pseudoabsence background points to that area. Despite these techniques it is still not the state of practice to incorporate dispersal limitation into models of the distribution of spreading species (e.g. Heidy Kikillus, Hare & Hartley 2010; Gormley *et al.* 2011).

We present a simple new method that accounts for dispersal limitation in the fitting of a SDM. We first construct a dispersal model, and then use this to weight a SDM of the species-environment relationship (SER). In this way the importance of absences due to dispersal limitation is reduced, so the model fitting procedure is closer to the desired situation where the model discriminates between presences and absences due to suitable and unsuitable environmental conditions respectively.

We compare the ability of this method with models that do not account for dispersal limitation at parameterising the SER and predicting the future distribution of a simulated non-native species. We explore how both modelling techniques perform when the spatial structure of explanatory variables is varied. Both techniques are then applied to model the distribution of a non-native bird, the common waxbill *Estrilda astrild*, in the Iberian Peninsula.

2.2 Methods

2.2.1 General model framework

We used a dispersal model, in this case a cellular automaton (Carey 1996), to calculate the probability that a species could disperse to a given grid-cell. These probabilities were used to weight a linear model, so it was fitted more closely to data points where the species was likely to have been able to disperse to and the assumption of equilibrium likely to hold. We refer to this as dispersal weighting (DW).

DW is most easily understood by considering model fitting by least squares. In ordinary least squares $\sum d^2$ is minimised in model fitting, where d is the difference between the response variable and fitted values predicted by the model, while in DW least squares $\sum p_{disp} \times d^2$ is minimised, where p_{disp} is the probability of the grid-cell being dispersed to. DW can also be applied to generalised linear models (GLM), where the vector of dispersal probabilities is supplied as prior weights to the iteratively reweighted least squares algorithm used in model fitting. DW can be easily implemented in R (R Development Core Team 2010) by supplying a vector of dispersal probabilities to the weights argument of model fitting functions such as lm.

While dispersal information is used in the fitting of DW models, predictions from the fitted model object will only relate to how environmentally suitable grid-cells are, and will not be influenced by the probability that the species could disperse to each grid-cell. We call this unweighted prediction (UP). However, for a grid-cell to be occupied it has to be both suitable and dispersed to. If these events are assumed to be independent (this assumption is only likely to hold if a globally derived dispersal kernel is used; see discussion), then, using the multiplication rule for independent events, the probability that a grid-cell is occupied is the product of the probability that it is dispersed to (calculated from the dispersal model) and the probability that it is environmentally suitable. We refer to this as weighted prediction (WP). The use of dispersal information in model fitting and prediction is summarised in Table 2.1.

Table 2.1. Use of dispersal information in the different models used in this paper

Model	Dispersal information used in model fitting?	Dispersal information used in prediction?
GLM UP	N	N
DWGLM UP	Y	N
GLM WP	N	Y
DWGLM WP	Y	Y

2.2.2 Simulation

We used a simulation to compare the performance of DW and unweighted GLMs (hereafter DWGLM and GLM respectively) at parameterising the SER of a simulated non-native species. Each cell in a 50×50 grid was assigned a probability of being suitable based on known relationships with three environmental variables. The simulated species was 'introduced' to a grid-cell (coordinates 48, 28), and was allowed to spread to suitable grid-cells based on known dispersal rules (Appendix 2-1, see Appendix 2-2 for examination of the influence of introduction location). The dispersal rules were also used to provide weights for the DWGLM. To investigate whether the relative performance of GLMs and DWGLMs change as a species spreads and occupies a larger portion of suitable grid-cells, we ran simulations for three, five and 10 generations. To investigate if spatial autocorrelation in an explanatory variable influenced model performance, we ran simulations where one variable, *a*, was randomly distributed (non-spatial scenario, mean correlation between *a* and the X axis <0.001 ± 0.020 standard deviation), and where *a* was strongly correlated with the X axis (spatial scenario, mean correlation between *a* and the X axis 0.993 ± <0.001 standard deviation). The other two environmental variables were randomly distributed in all simulation scenarios. Each simulation scenario was run 1000 times, and the occurrence of the simulated species was modelled as a function of the three environmental variables using logistic DWGLMs and GLMs. The environmental variables were reset for each iteration and chosen according to the previously described rules.

We also assessed the ability of the models to predict the simulated species' future distribution. Models were fitted using the distribution after five generations, and their performance was assessed against the distribution after 10 and 20 generations by calculating the area under the receiver operating characteristic curve (AUC). AUC has been criticised as values are dependent on the ratio between the extent of occurrence of a species and the extent of the study area (Lobo, Jiménez-Valverde & Real 2008). That criticism is not applicable to our use of AUC as we used it to compare GLMs and DWGLMs using the same distribution data. Simulations were run 50 times in this assessment, and we assessed model performance using both unweighted and WP (see general model framework); in the latter case the relative dispersal pressure at the end of the simulation was used to weight predictions. By using the same dispersal rules to run the simulation and provide the dispersal weights we were in effect using a perfect dispersal model. As dispersal models constructed with real data are almost certainly imperfect descriptions of the true situation, we tested the sensitivity of DWGLMs to errors in the dispersal model. Stochastic errors were introduced to the dispersal model predictions by adding or subtracting a random number drawn from a uniform distribution up to a maximum error value for each grid-cell. We did this for errors of up to ±0.05, ±0.1, then at increasing 0.1 increments up to ±0.9. These errors were added to the dispersal probability predictions used to fit DWGLMs after five simulation generations, and also to the probability of dispersal used to calculate WPs when the models were tested after 10 generations.

2.2.3 Modelling the distribution of the common waxbill

To compare the performance of DWGLMs and GLMs when real data were used we applied the modelling techniques to model the occurrence of the common waxbill in the Iberian Peninsula as a function of climate and land-use variables. The common waxbill is a largely granivorous estrildid finch species (Passeriformes: Estrildidae) native to sub-Saharan Africa, where it is often associated with mesic habitats (Payne 2010). It has been introduced to South America, the Iberian Peninsula and several oceanic islands (Lever 2005). In the Iberian Peninsula it was first recorded in western Portugal in 1964 (Reino & Silva 1998) and is now the most widespread non-native bird species in the Iberian Peninsula (Silva, Reino & Borralho 2002).

2.2.3.1 Calculating dispersal probabilities

We obtained data on the expansion of the common waxbill between 1964 and 1999 from Reino, Moya-Larano & Heitor (2009), supplemented with additional records from Spain (Figure 2.1). Occurrences were mapped in 1964, 1974, 1984, 1994 and 1999 on a UTM grid of 10 × 10km cells covering continental Portugal and Spain. We used a coarser time scale than previous studies of the expansion of the common waxbill to try and mitigate the effects of temporary high spatial heterogeneity in recorder effort caused by local bird atlas projects (e.g. Elias & Reino 1994).

We used this dataset to inspect the shape and parameterise the dispersal kernel that best described the expansion of the common waxbill over 10 year periods (Appendix 2-3). The dispersal kernel was run in cellular automata dispersal models (see Carey 1996 for an example) starting from 1964, 1974, 1984, and 1994, and using real occurrence data for each starting year, to calculate the probability of each grid-cell being dispersed to by the following time period (Appendix 2-3). The addition rule for non-mutually exclusive events was used to calculate the overall probability of each grid-cell being dispersed to by 2004 from these, giving a single probability between zero and one that the cell had been dispersed to (Figure 2.1).

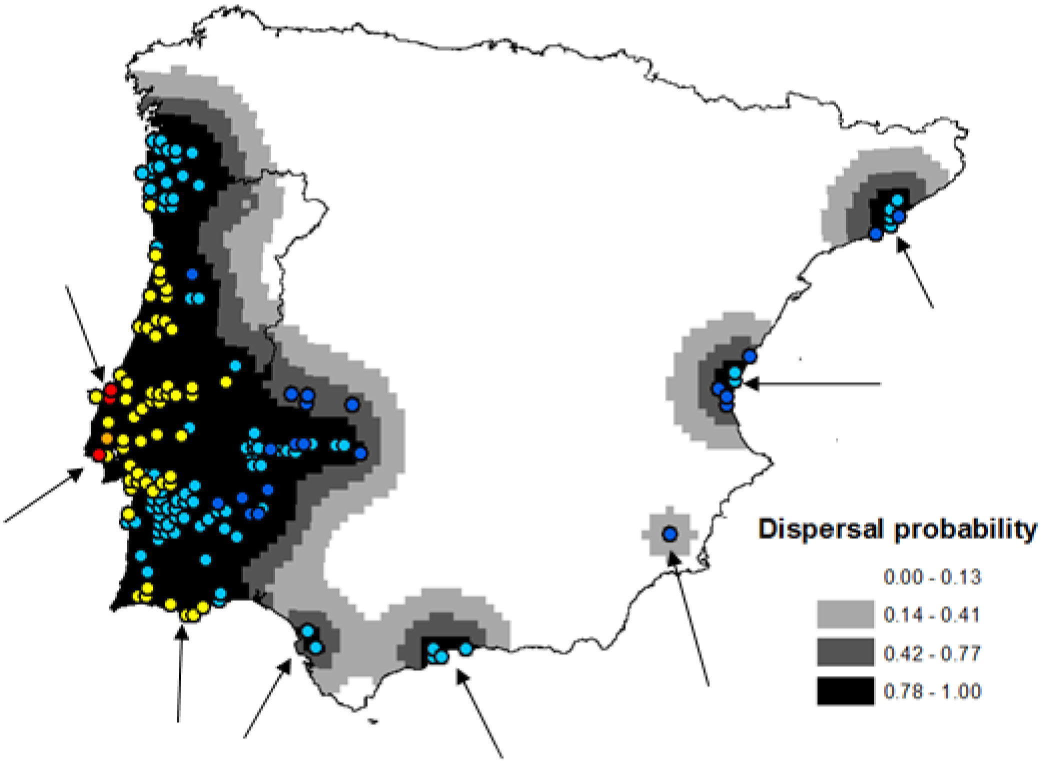


Figure 2.1. Probability of grid-cells in the Iberian Peninsula being dispersed to by the common waxbill by 2004. Darker shades indicate higher probabilities. Coloured circles show the colonisation date of grid-cells used to construct the dispersal model; red = by 1964, orange = by 1974, yellow = by 1984, light blue = by 1994 and dark blue = by 1999. Arrows show the location of presumed separate introduction events (see Appendix 2-3 for justification).

2.2.3.2 Explanatory variables

We modelled the occurrence of the common waxbill as a function of both climate and land-use variables. Mean precipitation, temperature and daily temperature range were obtained for 10' grid-cells for each month between 1991 and 2000 from the CRU TS1.2 (Mitchell *et al.* 2004) and interpolated to a 1km^2 resolution where appropriate (Appendix 2-4). Land-cover variables were obtained from the Corine land-cover classes (Appendix table 2.2), and the area of each class in each 10km UTM grid-cell was extracted from Corine 2000 vector layers for Portugal (Caetano, Nunes & Nunes 2009) and Spain (Instituto Geográfico Nacional 2011) in Arc GIS 9.3 (ESRI 2008). To allow us to compare how the performance of DWGLMs and GLMs with different covariate sets related to their performance in spatial and non-spatial simulation scenarios, we assessed the degree of spatial autocorrelation in explanatory variables by calculating Moran's I in the first distance class in SAM (Rangel, Diniz-Filho & Bini 2010).

2.2.3.3 Distribution data

As our primary interest was comparing the performance of GLM and DWGLM we only used distribution data from the Iberian Peninsula, a part of the invaded range where sufficient data were available to construct a dispersal model. We obtained data on the occurrence of common waxbills in 10 × 10 UTM grid-cells in the Iberian Peninsula from the most recent Portuguese (Equipa Atlas 2008) and Spanish (Marti & de Moral 2003) atlases of breeding birds. The survey periods for both atlases overlapped considerably (the Portuguese atlas ran from 1999 to 2005, while the Spanish atlas ran from 1998 to 2002). Where grid-cells straddled the national border they were considered occupied if common waxbills were recorded there in either national atlas.

2.2.4 Data analysis

We constructed logistic DWGLMs and GLMs of the occurrence of the common waxbill as a function of climate and land-use variables, using the dispersal probability for each grid-cell to weight the DWGLMs. To aid comparison of the different modelling techniques the same explanatory variables were used in the global models for each method. Following preliminary analysis (Appendix 2-4) three climate and five land-use variables were selected, as well as appropriate quadratic terms and interactions (Appendix table 2.3) and a proxy for recorder effort (Appendix 2-4). We used multi-model inference (MMI, Burnham & Anderson 2002) to fit all valid simplifications of the global climate and land-use models and identify the 95% confidence set of models with the most support (Appendix 2-4, results presented in Appendix table 2.3).. Model performance was assessed by cross-validation, with data split into mutually exclusive training (75%) and testing (25%) sets. The MMI procedure described above was performed on the training set, and the accuracy of the model averaged predictions of the resulting 95% confidence set of models was tested on the testing set. Model performance was assessed by calculating the AUC for WPs and UPs. This procedure was repeated for 500 iterations for each model set. Unless otherwise stated, all analyses were performed in R version 2.12 (R Development Core Team 2010).

2.3 Results

2.3.1 Simulation

The simulated species colonised more of the available grid-cells when explanatory variables were spatially structured than when they were randomly distributed. In the spatial scenario, the median area colonised after five generations was 2.4% of the grid, with 45.8% colonised after 20 generations. In the non-spatial scenario 1.2% of the grid was colonised after five generations, with 28.3% colonised after 20 generations.

DWGLMs performed better than GLMs at parameterising the SER of the simulated species (Figure 2.2). Compared to DWGLMs, GLMs tended to underestimate the magnitude of relationships with randomly distributed variables, but overestimate the magnitude of relationships with strongly spatially correlated variables, indicating that while the spatial structure of explanatory variables had a strong effect on how GLMs parameterised the SER for dispersal limited species, the effect was less pronounced for DWGLMs. GLM parameter estimates for the strongly spatially correlated variable *a* improved when the simulation was run for more generations (Figure 2.2). Despite this, the proportion of simulation runs in which DWGLMs produced closer estimates of the true parameter value increased with the number of generations (e.g. in the non-spatial scenario this happened in 65.7% of runs after three generations and in 84.9% of runs after 10 generations), and was also higher in the spatial than non-spatial scenario (after 10 generations DWGLMs produced closer estimates of the true parameter value in 99.2% of runs in the spatial scenario compared to 84.9% in the non-spatial scenario). GLMs produced better parameter estimates for one randomly distributed variable, *b*, in all but one simulation scenario; however there was considerable overlap between parameter estimates derived by both methods, and the median value from DWGLMs was closest to the true value of 0.2 (range of median parameter estimates from all simulation scenarios: GLMs = 0.04-0.076, DWGLMs = 0.124-0.154). These results indicate that the superior performance of DWGLMs compared to GLMs was most pronounced when variables had larger true parameter values, were spatially autocorrelated, and when models were fitted after a number of generations.

When WP (multiplying the predicted suitability of a grid-cell with the predicted probability the grid-cell was dispersed to, see general framework in methods) was used, DWGLMs performed better at predicting the future distribution of the simulated species in all simulation scenarios (Figure 2.3; range of median AUC values from all simulation scenarios:

GLMs=0.645-0.897, DWGLMs=0.979-0.993). This indicates that DWGLMs were better at classifying the suitability of grid-cells in areas where the species was able to disperse to. When UP was used GLMs and DWGLMs showed similar performance when environmental variables were randomly distributed (median AUC values after 10 and 20 generations: GLMs=0.61 and 0.586, DWGLMs=0.631 and 0.612). When one environmental variable was strongly spatially correlated GLM performed better than DWGLMs after 10 generations (median AUC values: GLMs=0.886, DWGLMs=0.78) when the simulated species occupied 11.1% of the grid-cells, but after 20 generations when 45.8% of grid-cells were occupied both methods showed similar performance (median AUC values: GLMs=0.782, DWGLMs=0.823), indicating better classification of grid-cell occupancy by GLMs only for early stages of invasion.

The performance of WPs from both models declined when errors were introduced into the dispersal model (Figure 2.4). This was especially pronounced when explanatory variables were randomly distributed. The decline in performance was steeper for DWGLMs, but they still outperformed GLMs when the maximum introduced error in dispersal probability was less than 0.6, indicating that DWGLMs are fairly robust to errors in the dispersal model.

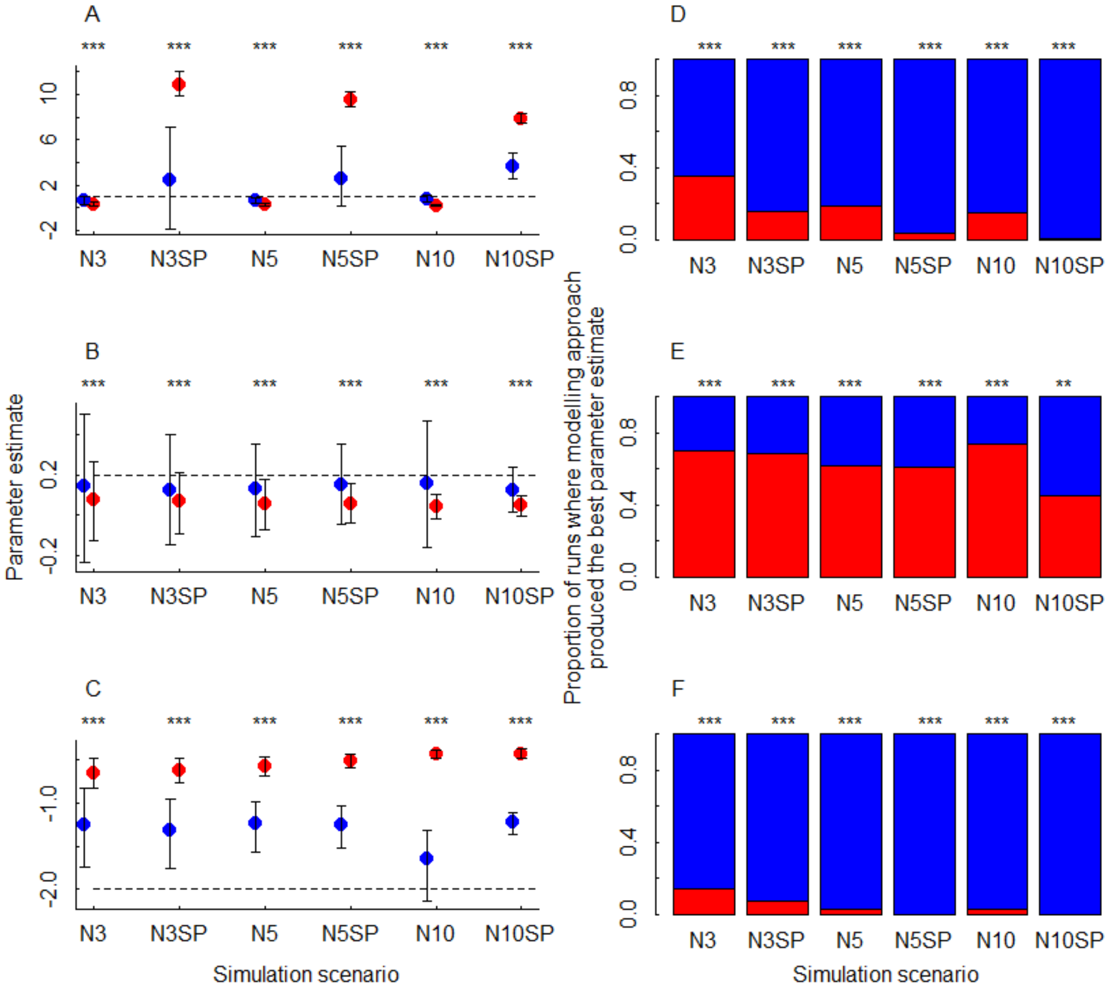


Figure 2.2. Performance of DWGLMs and GLMs at parameterising the SER of a simulated invasive species. The number of generations the simulation was run for is shown by the number after N (e.g. N3 = 3 generations). Spatial simulation scenarios are denoted by SP. (A-C) Estimates of the slope of the relationship between the species' occurrence and environmental parameters *a* (A), *b* (B) and *c* (C) produced by DWGLMs (blue) and GLMs (red). Points show the median and error bars the inter-quartile range of estimates from 1000 runs. The dashed line shows the true parameter value. The difference between DWGLM and GLM parameter estimates was tested with Mann-Whitney *U* tests; Bonferoni adjusted *P* values are displayed, *** *P*<0.001. (D-F) Proportion of runs in which DWGLMs (blue bars) and GLMs (red bars) produced the closest parameter estimate to the true value for parameters *a* (D), *b* (E) and *c* (F). Binomial tests were used to test whether the proportions were significantly different from the null expectation of 0.5; *** *P*<0.001, ** *P*<0.01.

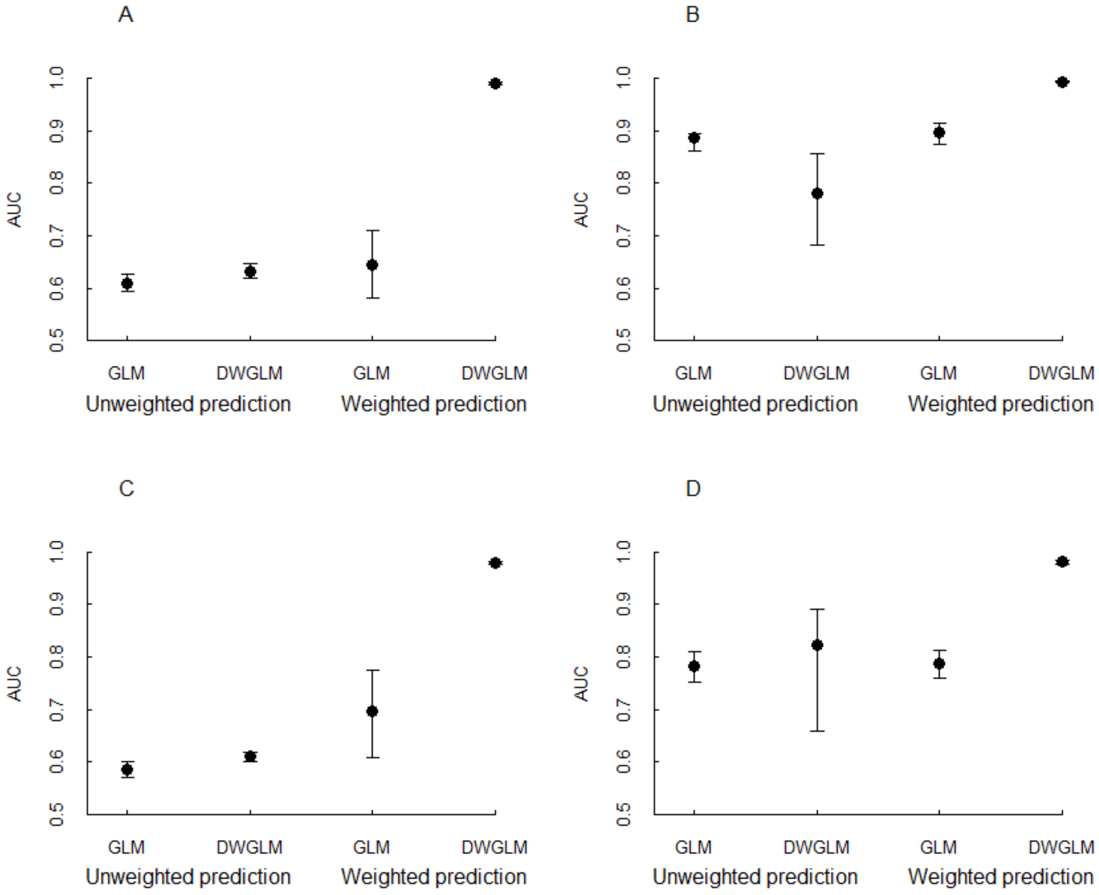


Figure 2.3. Performance of GLMs and DWGLMs at classifying the suitability of grid-cells for a simulated species. Models were constructed after five generations, and tested on the distribution after 10 (A and B) and 20 (C and D) generations. See general model framework for description of weighted and unweighted validation methods. Models were constructed for (A and C) randomly distributed explanatory variables and (B and D) where one variable was strongly spatially autocorrelated. Median and interquartile range AUC values from 50 simulation runs are shown.

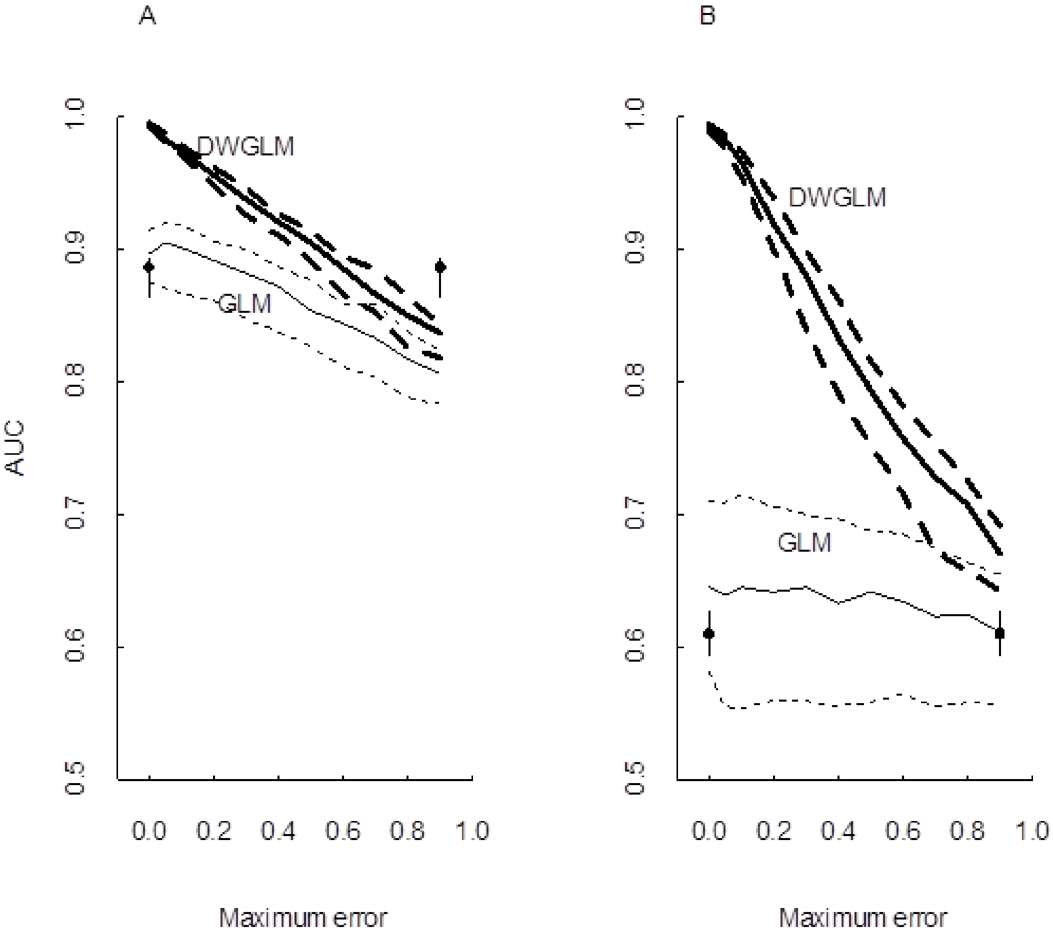


Figure 2.4. Effect of errors in the dispersal model on the performance of DWGLMs and GLMs in (A) spatial and (B) non-spatial simulation scenarios. Errors were drawn from a uniform distribution up to a maximum value and introduced to the dispersal probabilities for each grid-cell. AUC values were calculated by testing the ability of models constructed after five generations of a simulated non-native species to classify grid-cells as suitable for that species after 10 generations. Median AUC values for DWGLM WPs are shown by the bold line, with the other line showing AUC values for GLM WPs. Dashed lines delimit the interquartile range from 50 simulation runs. For comparison median and interquartile range AUC values for UPs of GLMs have been shown by points and error bars at both ends of the x axis.

2.3.2 Common waxbill model

Dispersal was important in structuring the common waxbill distribution, with the dispersal model explaining 19.2% of variation in the occurrence data (Figure 2.1). The majority of absences were due to dispersal limitation; 70.0% of absences had a probability of being dispersed to of <0.1, compared to 0.006% of presences. Despite this absence data were still available for model fitting, with 770 absences and 594 presences having a probability of being dispersed to of >0.5.

The main differences between GLMs and DWGLMs of common waxbill occurrence were the magnitude of relationships with explanatory variables and the importance given to interactions (Appendix table 2.3). These differences resulted in DWGLMs having fewer omission errors and predicting a larger potential distribution (Figure 2.5).

The relative performance of GLMs and DWGLMs of common waxbill occurrence was similar to the results of the simulations. Climate variables showed stronger spatial autocorrelation than land-use variables (Table 2.2). This was reflected in the performance of climate and land-use based models of common waxbill occurrence when assessed by cross-validation (Figure 2.6). The relative performance of GLMs and DWGLMs using climate covariates was similar to the situation in the spatial simulation scenarios; GLMs performed better than DWGLMs when UPs were assessed (median AUC values 0.937 and 0.919 respectively), while DWGLMs performed better than GLMs when WPs were assessed (median AUC values 0.962 and 0.951 respectively). The performance of land-use based models was more similar to the non-spatial simulation scenario. GLMs and DWGLMs performed similarly when using UPs (median AUC values 0.867 and 0.861 respectively), but DWGLMs performed better than GLMs models when using WPs (median AUC values 0.966 and 0.948 respectively). This indicates that for both sets of covariates DWGLMs models were better at classifying the suitability of grid-cells for the common waxbill when dispersal limitation was corrected for, confirming similar results for the simulations compared to those observed with real data.

Table 2.2. Spatial autocorrelation of explanatory variables used to model the occurrence of the common waxbill.

Explanatory variable	Moran's I
MTCM	0.650
MDTR	0.697
MAP	0.693
Rice	0.180
Irrigated agriculture	0.167
Parks and gardens	0.137
Built	0.336
Woody agriculture	0.409
Recorder effort	0.099

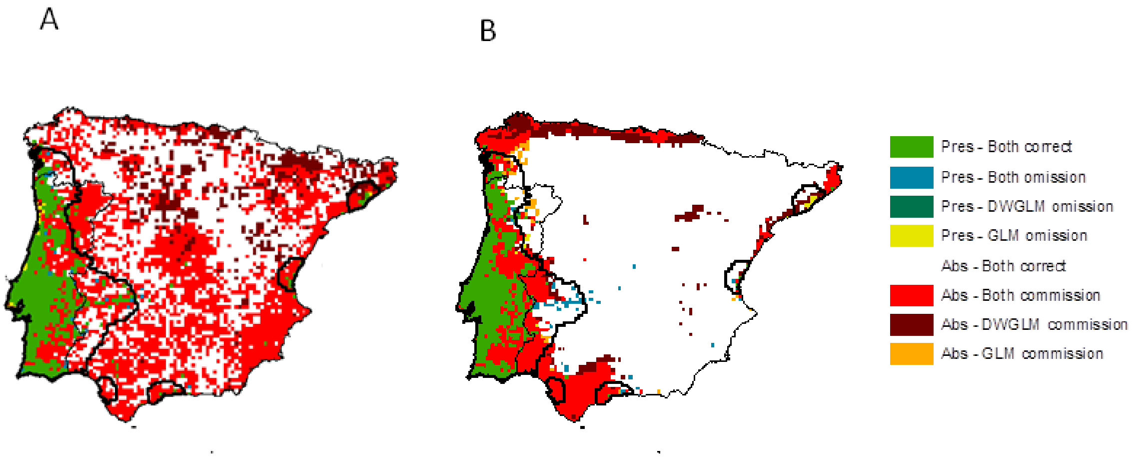


Figure 2.5. Potential distribution of the common waxbill using (A) land-use and (B) climate variables. A threshold that minimised the difference between omissions and commissions (Jiménez-Valverde & Lobo 2007) was used to convert continuous suitability values to a binary classification. This threshold was lower for GLMs. Areas within the thick black lines have a dispersal probability of >0.5 in 2005.

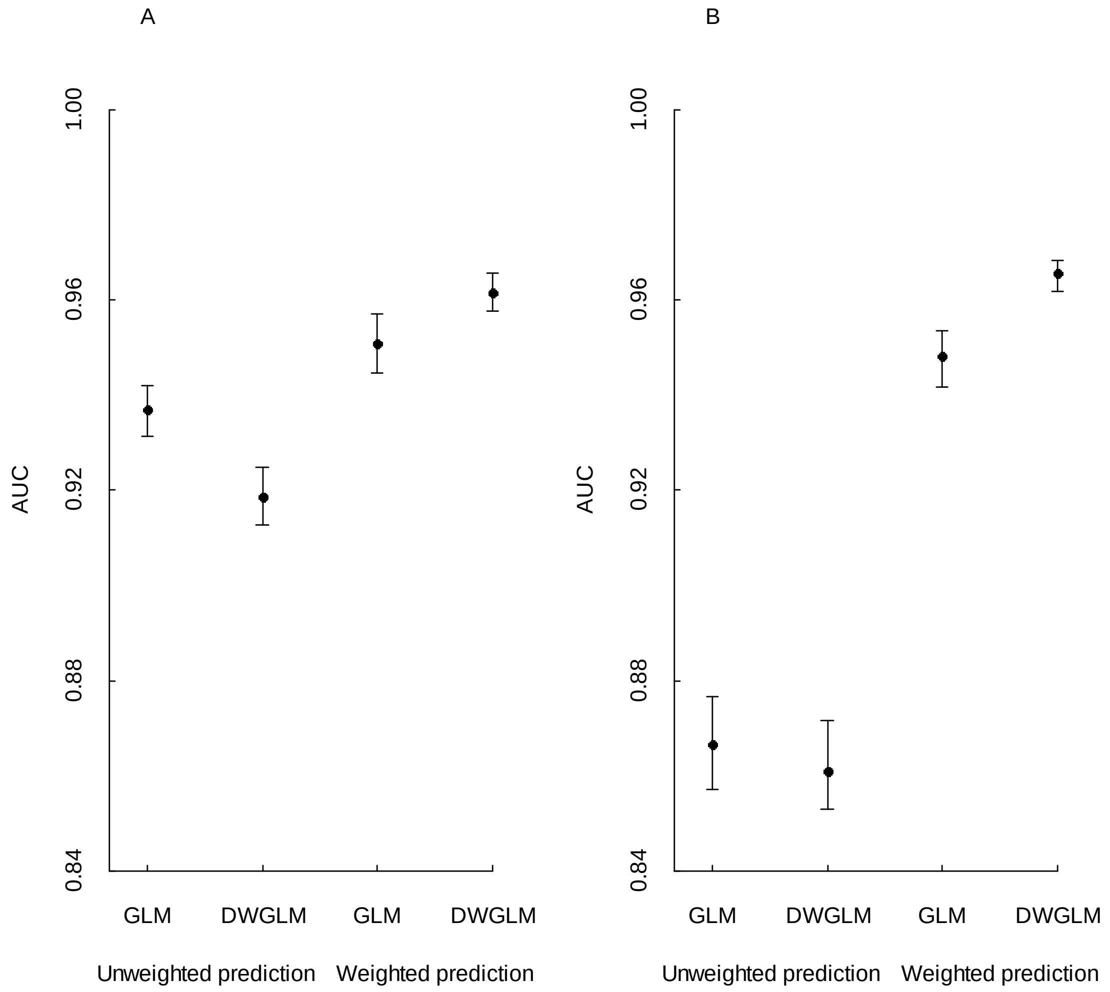


Figure 2.6. Performance of GLMs and DWGLMs at classifying the suitability of grid-cells for the common waxbill. Models were constructed using climate (A) and land-use (B) explanatory variables. See general model framework for description of prediction methods. Median and interquartile range AUC values from 500 cross-validation runs are shown.

2.4 Discussion

2.4.1 Comparison of GLMs and DWGLMs

Logistic GLMs have been frequently used to model the distribution of spreading non-native species (Reino 2005; Real *et al.* 2008). These models have proved useful at distinguishing between areas that are occupied and unoccupied by a species based on sets of environmental covariates, but the performance of such correlative models may be affected by not considering dispersal limitation (Beale, Lennon & Gimona 2008; Gallien *et al.* 2010). We proposed a new method, DW, that downweights the importance of grid-cells where a species is likely to be absent due to dispersal limitation, and tested its performance against GLMs. These analyses demonstrated that DWGLMs performed better than GLMs at parameterising the true SER and at classifying the suitability of grid-cells in areas where the modelled species was likely to have dispersed to. However, when explanatory variables were distributed along a spatial gradient, GLMs performed better than DWGLMs at classifying areas as occupied or unoccupied across the whole study area. The differences in model performance can be understood with reference to the pool of presences and absences that models are fitted to classify between. In DWGLMs, absences due to dispersal limitation are downweighted, so the absence pool largely contains absences due to unsuitable environmental conditions. In contrast, in GLMs the absence pool contains absences due to both unsuitable environmental conditions and dispersal limitation.

The degree of spatial autocorrelation in explanatory variables affected the performance of the different methods. When environmental variables were randomly distributed, dispersal limitation of non-native species led to explanatory variables occurring with favourable values in the absence pool. Downweighting the importance of those dispersal limited grid-cells reduces their frequency in the absence pool, so the distribution of environmental variables in the presence and absence pool will better reflect the environmental preferences of the non-native species. UPs from DWGLMs and GLMs performed poorly at classifying the potential distribution of the simulated species across the whole study area. This is likely to be because both models were penalised for correctly predicting suitable sites that were not yet occupied due to dispersal limitation. When WPs were used, the performance of both methods improved, but DWGLMs performed considerably better than GLMs as they were better at parameterising the SER.

GLMs were more prone than DWGLMs to overparameterising the relationship with spatially autocorrelated explanatory variables. If explanatory variables are distributed so that they have increasingly favourable values near the site of introduction of a non-native species, then they will occur with favourable values in the presence pool and unfavourable values in the absence pool due to dispersal limitation alone, accentuating the pattern that might be observed due to the SER. Because of this, in an invasion's early stage these spatially autocorrelated variables make good predictors of the non-native species' distribution. By overparameterising relationships with spatially autocorrelated variables GLMs exploit the spatial information contained by them, so perform better than DWGLMs. As the non-native species spreads further the spatial information contained in spatially autocorrelated variables becomes less useful; this was demonstrated by the reduction of the relative performance of GLMs as the simulation was run for more generations. As with the case where explanatory variables were randomly distributed, DWGLMs performed better than GLMs when using WPs. Our results support previous studies that show that associations between the spatial structure of explanatory variables and the dispersal potential of a species can lead to models that do not account for dispersal limitation identifying statistical relationships where little or no causal relationship exists (Bahn & McGill 2007; Beale, Lennon & Gimona 2008).

AUC values of WPs in the simulations were almost certainly higher than would be achieved with real data, as a perfect dispersal model was used (i.e. the dispersal model used to construct the simulations was used to provide dispersal weights), and the simulated species distribution depended on only three environmental variables, all of which were modelled. We are confident that this does not affect the conclusions drawn above for three reasons. Firstly, both DWGLM and GLM WPs benefited from a perfect dispersal model, so this would not affect comparisons between them. Secondly, when stochastic errors were introduced into the dispersal model, DWGLMs still outperformed GLMs when WPs were assessed. The faster decline in performance of DWGLM compared to GLM WPs was probably because dispersal information was used twice (fitting and prediction) in DWGLM, compared to once (prediction) in GLM. Thirdly, and most importantly, the similarity between the results of the simulations and the application with the common waxbill support generalisation of the simulation results to real scenarios.

The relative performance of GLMs and DWGLMs at modelling the distribution of the common waxbill in the Iberian Peninsula were similar to the results of the simulations. This extended to observations of the effect of spatial autocorrelation in explanatory variables; climate variables showed stronger spatial autocorrelation than land-use variables, and the relative performance of the modelling methods mirrored the spatial and non-spatial simulation scenarios respectively. This is not perfect as the difference in spatial autocorrelation was not as extreme as in the simulation scenarios.

Many grid-cells had suitable land-use but low dispersal probabilities; therefore the increase in performance when WP was used compared to when UP was used was greater for land-use models. Additionally, many dispersal limited grid-cells were also climatically unsuitable.

DWGLMs performed worse in UP as they predicted a larger potential distribution, and were penalised for classifying dispersal limited grid-cells as suitable.

Dispersal was an important constraint on the common waxbill distribution, and the majority of absences available to the GLM were strongly downweighted by DWGLM. A small number of presences were also strongly downweighted. As these only represented a small fraction of the total set of presences, it is unlikely that this would have had a major effect on model performance. One of the reasons for errors in the dispersal model was that the dispersal kernel used did not vary spatially, while in reality the common waxbill expanded faster along the northward expansion axes than in other directions (Silva, Reino & Borralho 2002). Using a global dispersal model has a useful property in that the probabilities that a grid-cell has been dispersed to and is environmentally suitable are independent, so can be multiplied to provide the probability that a grid-cell will be occupied (as in WP). This is unrealistic as dispersal will interact with environmental suitability (Marushia & Holt 2006); unsuitable areas will slow dispersal (McRae 2006), while corridors or stepping stones of suitable habitat can assist population spread (Sondgerath & Schroder 2002). In this paper our primary aim was to understand the SER so we first calculated dispersal probabilities then parameterised environmental suitability given these dispersal probabilities. This allowed us to account for dispersal limitation when modelling the SER. Smolik et al. (2010) did the opposite, estimating dispersal probabilities given environmental suitability. Their approach does not mitigate the impact of dispersal limitation on fitting environmental suitability models, but could be useful for simulating the future distribution of spreading species. Ideally the parameters of the dispersal and environmental suitability models would be pararamatised simultaneously. Bayesian hierarchical models show promise as a means of achieving this; for example Stanaway et al. (2011) used them to simultaneously estimate parameters of two dispersal models acting at different spatial scales.

2.4.2 Choice of dispersal model

Constructing a dispersal model is the most challenging part of using this method (see Appendix 2-3 for discussion). The model used will depend on the data available. When, as in our case with the common waxbill, distribution data are available at different time periods, the parameter(s) of a dispersal kernel can be estimated using numerical optimisation by running cellular automata with different parameter values (Carey 1996; Smolik *et al.* 2010) and assessing the fit of the resulting models to observed data. Dispersal kernels could also be estimated from recaptures of marked organisms (Paradis *et al.* 1998) or movements of radio tagged individuals (Driezen *et al.* 2007). Cellular automata can be constructed with limited data, but they do not take variation in dispersal probabilities due to demographic factors into account (Carey 1996), so, where more data are available, it may be preferable to use other spread models (see Hengeveld 1989 for examples). The only requirement for a dispersal model to be used in DW is that it can provide the probability that each grid-cell has been dispersed to.

2.4.3 Use with other modelling techniques

The method of weighting grid-cells by dispersal probabilities can be applied to any other presence-absence modelling technique that accepts case weights, such as generalised additive models. The technique could potentially be applied to presence-pseudoabsence techniques. Pseudoabsence points can be restricted to buffer zones around presence points, with the aim of accounting for dispersal limitation (Elith, Kearney & Phillips 2010). This method could be extended so that, rather than simply restricting pseudoabsences to an area of potential dispersal, the probability of drawing a grid-cell as a pseudoabsence point is proportional to the dispersal probability of that grid-cell.

2.5 Conclusion

Dispersal is an important constraint on the distribution of spreading species such as the common waxbill, with species absent from many areas due to dispersal limitation alone. We demonstrated that models which downweighted absences due to dispersal limitation performed better than unweighted models at parameterising the SER and classifying the suitability of grid-cells for the modelled species. A number of other issues not addressed here contribute to the challenge of modelling range-shifting species distributions (Elith, Kearney & Phillips 2010), however, by using DW to help tackle the problem of dispersal limitation we can start to increase confidence in our ability to model the distributions of these species.

Acknowledgments

We thank an anonymous reviewer for their comments on the manuscript. The ICNB provided digital atlas data for Portugal. This research was carried out on the High Performance Computing Cluster supported by the Research and Specialist Computing Support service at the University of East Anglia. MJPS was funded by a NERC PhD studentship.

References

Bahn, V. & McGill, B.J. (2007) Can niche-based distribution models outperform spatial interpolation? *Global Ecology and Biogeography,* **16,** 733-742.

Beale, C.M., Lennon, J.J. & Gimona, A. (2008) Opening the climate envelope reveals no macroscale associations with climate in European birds. *Proceedings of the National Academy of Sciences,* **105,** 14908-14912.

Brooker, R.W., Travis, J.M.J., Clark, E.J. & Dytham, C. (2007) Modelling species' range shifts in a changing climate: The impacts of biotic interactions, dispersal distance and the rate of climate change. *Journal of Theoretical Biology,* **245,** 59-65.

Burnham, K.P. & Anderson, D.R. (2002) *Model selection and multi-model inference, a practical information-theoretic approach*. Springer, New York.

Caetano, M., Nunes, V. & Nunes, A. (2009) CORINE Land Cover 2006 for Continental Portugal. *Relatório técnico*. Instituto Geográfico Português

Carey, P.D. (1996) DISPERSE: A Cellular Automaton for Predicting the Distribution of Species in a Changed Climate. *Global Ecology and Biogeography Letters,* **5,** 217-226.

Cook, D.C., Thomas, M.B., Cunningham, S.A., Anderson, D.L. & De Barro, P.J. (2007) Predicting the economic impact of an invasive species on an ecosystem service. *Ecological Applications,* **17,** 1832-1840.

Driezen, K., Adriaensen, F., Rondinini, C., Doncaster, C.P. & Matthysen, E. (2007) Evaluating least-cost model predictions with empirical dispersal data: A case-study using radiotracking data of hedgehogs (Erinaceus europaeus). *Ecological Modelling,* **209,** 314-322.

Dullinger, S., Kleinbauer, I., Peterseil, J., Smolik, M. & Essl, F. (2009) Niche based distribution modelling of an invasive alien plant: effects of population status, propagule pressure and invasion history. *Biological Invasions,* **11,** 2401-2414.

Elias, G.L. & Reino, L. (1994) Atlas of wintering birds of Baixo Altentejo: 1992-1995. *Bird Census News,* **7,** 57-62.

Elith, J., Kearney, M. & Phillips, S. (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution,* **1,** 330-342.

Engler, R. & Guisan, A. (2009) MIGCLIM: Predicting plant distribution and dispersal in a changing climate. *Diversity and Distributions,* **15,** 590-601.

Equipa Atlas (2008) *Atlas das Aves Nidificantes em Portugal (1999-2005)*. Instituto da Conservacao da Natureza e da Biodiversividade, Sociedade Portuguesa para o Estudo das Aves, Parque Natural da Madeira e Secretaria Regional do Ambiente e do Mar, Lisbon.

ESRI (2008) ArcGIS Desktop. Environmental Systems Research Institute.

Gallien, L., Münkemüller, T., Albert, C.H., Boulangeat, I. & Thuiller, W. (2010) Predicting potential distributions of invasive species: where to go from here? *Diversity and Distributions,* **16,** 331-342.

Gormley, A.M., Forsyth, D.M., Griffioen, P., Lindeman, M., Ramsey, D.S.L., Scroggie, M.P. & Woodford, L. (2011) Using presence-only and presence–absence data to estimate the current and potential distributions of established invasive species. *Journal of Applied Ecology,* **48,** 25-34.

Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters,* **8,** 993-1009.

Heidy Kikillus, K., Hare, K.M. & Hartley, S. (2010) Minimizing false-negatives when predicting the potential distribution of an invasive species: a bioclimatic envelope for the red-eared slider at global and regional scales. *Animal Conservation,* **13,** 5-15.

Hengeveld, R. (1989) *Dynamics of biological invasions*. Chapman and Hall, London.

Instituto Geográfico Nacional (2011) Corine Land Cover 2000 España. http://www.fomento.gob.es/MFOM/LANG_CASTELLANO/DIRECCIONES_GENERALES/INSTITUTO_GEOGRAFICO/Teledeteccion/corine/clc2000/. Accessed19/12/2011.

Jiménez-Valverde, A. & Lobo, J.M. (2007) Threshold criteria for conversion of probability of species presence to either–or presence–absence. *Acta Oecologica,* **31,** 361-369.

Lever, C. (2005) *Naturalised Birds of the World*. T & AD Poyser, London.

Lobo, J.M., Jiménez-Valverde, A. & Real, R. (2008) AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography,* **17,** 145-151.

Marti, R. & de Moral, J.C. (2003) Atlas de las Aves Reproductoras de Espana. Dirección General de Conservación de la Naturaleza-Sociedad Española de Ornithología, Madrid.

Marushia, R.G. & Holt, J.S. (2006) The effects of habitat on dispersal patterns of an invasive thistle, Cynara cardunculus. *Biological Invasions,* **8,** 577-593.

McRae, B.H. (2006) Isolation by resistance. *Evolution,* **60,** 1551-1561.

Menendez, R., Megias, A.G., Hill, J.K., Braschler, B., Willis, S.G., Collingham, Y., Fox, R., Roy, D.B. & Thomas, C.D. (2006) Species richness changes lag behind climate change. *Proceedings of the Royal Society B-Biological Sciences,* **273,** 1465-1470.

Mitchell, T.D., Carter, T.R., Jones, P.D., Hulme, M. & New, M. (2004) A comprehensive set of high-resoultion grids of monthly climate for Europe and the globe: the observed record (1901-2000) and 16 scenarios (2001-2100). *Working Paper*. Tyndall Centre for Climate Change Research, Norwich, UK.

Muñoz, A.-R. & Real, R. (2006) Assessing the potential range expansion of the exotic monk parakeet in Spain. *Diversity and Distributions,* **12,** 656-665.

Paradis, E., Baillie, S.R., Sutherland, W.J. & Gregory, R.D. (1998) Patterns of natal and breeding dispersal in birds. *Journal of Animal Ecology,* **67,** 518-536.

Payne, R.B. (2010) Family Estrildidae (Waxbills). *Handbook of the Birds of the World* (eds J. del Hoyo, A. Elliott & D.A. Christie). Lynx Edicions, Barcelona.

Peterson, A.T. (2003) Predicting the geography of species' invasions via ecological niche modelling. *Quarterly Review of Biology,* **78,** 419-433.

Pimentel, D., Zuniga, R. & Morrison, D. (2005) Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics,* **52,** 273-288.

R Development Core Team (2010) R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Rangel, T.F., Diniz-Filho, J.A.F. & Bini, L.M. (2010) SAM: a comprehensive application for Spatial Analysis in Macroecology. *Ecography,* **33,** 46-50.

Real, R., Márquez, A.L., Estrada, A., Muñoz, A.R. & Vargas, J.M. (2008) Modelling chorotypes of invasive vertebrates in mainland Spain. *Diversity and Distributions,* **14,** 364-373.

Reino, L. (2005) Variation partitioning for range expansion of an introduced species: the common waxbill Estrilda astrild in Portugal. *Journal of Ornithology,* **146,** 377-382.

Reino, L., Moya-Larano, J. & Heitor, A.C. (2009) Using survival regression to study patterns of expansion of invasive species: will the common waxbill expand with global warming? *Ecography,* **32,** 237-246.

Reino, L. & Silva, T. (1998) The distribution and expansion of the Common Waxbill (*Estrilda astrild*) in the Iberian Peninsula. *Biol Cons Fauna,* **102,** 163-167.

Silva, T., Reino, L.M. & Borralho, R. (2002) A model for range expansion of an introduced species: the common waxbill Estrilda astrild in Portugal. *Diversity and Distributions,* **8,** 319-326.

Smolik, M.G., Dullinger, S., Essl, F., Kleinbauer, I., Leitner, M., Peterseil, J., Stadler, L.M. & Vogl, G. (2010) Integrating species distribution models and interacting particle systems to predict the spread of an invasive alien plant. *Journal of Biogeography,* **37,** 411-422.

Sondgerath, D. & Schroder, B. (2002) Population dynamics and habitat connectivity affecting the spatial spread of populations - a simulation study. *Landscape Ecology,* **17,** 57-70.

Stanaway, M.A., Reeves, R. & Mengersen, K.L. (2011) Hierarchical Bayesian modelling of plant pest invasions with human-mediated dispersal. *Ecological Modelling,* **222,** 3531-3540.

Strubbe, D. & Matthysen, E. (2009) Predicting the potential distribution of invasive ring-necked parakeets *Psittacula krameri* in northern Belgium using an ecological niche modelling approach. *Biological Invasions,* **11,** 497-513.

Travis, J.M.J., Harris, C.M., Park, K.J. & Bullock, J.M. (2011) Improving prediction and management of range expansions by combining analytical and individual-based modelling approaches. *Methods in Ecology and Evolution,* **2,** 477-488.

Václavík, T., Kupfer, J.A. & Meentemeyer, R.K. (2012) Accounting for multi-scale spatial autocorrelation improves performance of invasive species distribution modelling (iSDM). *Journal of Biogeography,* **39,** 42-55.

Václavík, T. & Meentemeyer, R.K. (2009) Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? *Ecological Modelling,* **220,** 3248-3258.

Václavík, T. & Meentemeyer, R.K. (2012) Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. *Diversity and Distributions,* **18,** 73-83.

Williamson, M.H. (1996) *Biological Invasions*. Springer, New York.

2.6 Appendix 2-1: Simulation construction

Each cell in a 50×50 grid was assigned a probability of being suitable (P_{hab}) based on the equation logit(P_{hab}) = $1a+0.2b-2c-3$, where a, b and c were environmental variables with randomly generated values taken from a normal distribution with a mean of zero and a standard deviation of one. The simulated species was 'introduced' to a grid cell (coordinates 48, 28), and was allowed to spread according to the following dispersal rules for n generations; for each generation the species could disperse from its current cell to the eight nearest neighbour cells with a probability of one, and to the next nearest 16 cells with a probability of 0.5. For a cell to be colonised it had to be suitable and be dispersed to; as these were independent events the probability of a cell being colonised was the product of the probabilities of the cell being suitable and being dispersed to. Rather than use a single threshold value to determine whether a grid cell was suitable we ensured that each grid cell would be colonised with the calculated probability of colonisation by drawing a number from a uniform distribution between zero and one for each potential colonisation event. If the probability of colonisation was greater or equal to this number then the grid cell was colonised.

The dispersal rules used in the simulations were used to calculate the relative dispersal pressure for each grid cell. To do this the summed dispersal pressure of each grid cell was divided by the maximum summed dispersal pressure of any grid cell to give a relative dispersal pressure between zero and one. This relative dispersal pressure was used to weight the DWGLM. The occurrence of the simulated non-native species was modelled as a function of the three environmental variables using DWGLMs and GLMs with the structure $PA_{\textit{simulated species}} \sim a+b+c$, a binomial error distribution and a logit link function.

2.7 Appendix 2-2: Effect of introduction site on simulation results

We introduced the simulated species to a grid cell towards the right-hand margin of the grid because we wanted to examine the performance of the modeling methods at parameterising the SER of a species that was spreading along an environmental gradient into less suitable conditions (in our simulations, the relationship with variable a in the spatial simulation scenario). If we had introduced the simulated species to the left-hand side of the grid, then, in the spatial simulation scenario, dispersal limitation would lead it to initially occupying grid cells where the environmental variable had less suitable values. Given the pool of presences and absences available for model fitting, we would expect GLMs to parameterise the relationship with the environmental variable as negative. We tested this by running 50 simulations for five generations where the simulated species was introduced to the left-hand side of the grid (coordinates 10,25). GLMs estimated the relationship with *a* as being negative (median slope = -5.217, Q1 = -5.615, Q3 = -4.834). Most DWGLMs also estimated the relationship as negative (median slope = -1.252, Q1=-3.659, Q3=1.181), but the parameter estimates were closer to the true value of 1 (Wilcoxon singed-rank test, V=1160, $P<0.001$), and the inter-quartile range overlapped the true value of *a*.

To test if the site of introduction had an important effect on parameter estimates in the non-spatial simulation scenario we carried out simulations were the species spread for five generations from introduction locations with coordinates 10,25, 20,25, 30,25 and 40,25. 50 simulation runs were performed from each introduction location. The location of introduction did not have a major effect on parameter estimates, with DWGLMs consistently producing higher estimates of *a* than GLMs (median DWGLM estimates = 0.556-0.659, median GLM estimates = 0.198-0.332).

2.8 Appendix 2-3: Constructing the dispersal model

Constructing the dispersal kernel

We used data on the expansion of the common waxbill (Reino, Moya-Larano & Heitor 2009, with additional data from Spain) to estimate the shape and parameters of a dispersal kernel. We decided to use a negative exponential function for the dispersal kernel after inspecting the shape of a histogram of the distance between each newly colonised grid cell and the nearest occupied grid cell in the previous time step. In a negative exponential dispersal kernel, the decline in dispersal probability P with distance is given by $P = e^{-bx}$, with the parameter b determining the rate of decline, and x denoting distance. It is possible to estimate b using maximum likelihood. In order to do this, the dispersal kernel first has to be re-written into a logit scale,

$\text{logit}(P) = \log(P/1\text{-}P) = \log(e^{-bx}/(1\text{-} e^{-bx}))$.

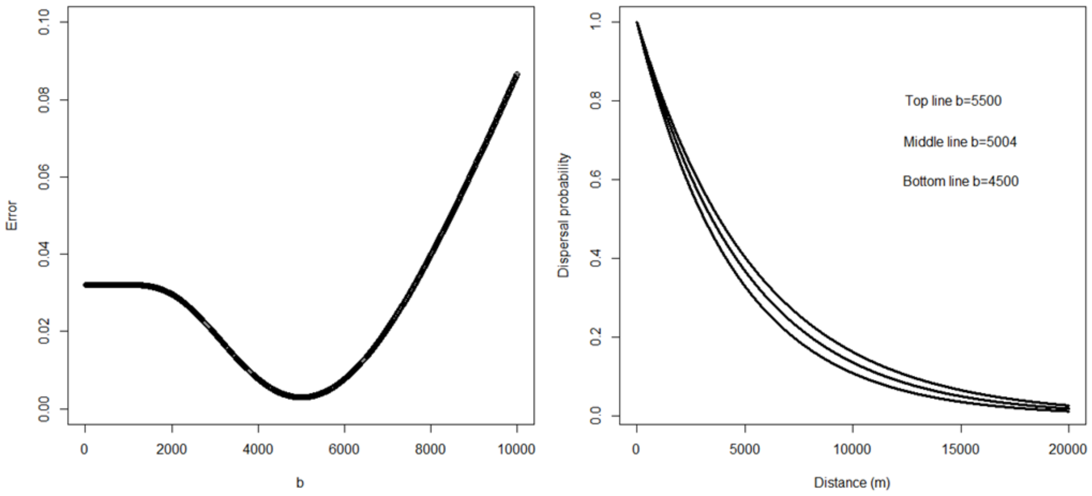
This can then be substituted into a binomial likelihood function,

$$likelihood = \sum -y \times \log(1+e^{P}) - ((1-y) \times \log(1+e^{P})),$$

where *P* is the dispersal probability calculated from the dispersal kernel and *y* is the occupancy status of a given grid-cell. However, there are a number of issues with this approach that required a more complex parameterisation procedure. Firstly, this only considers the probability of a grid-cell being colonised from a single nearest neighbour, when multiple grid-cells colonised in the previous time step may contribute to a grid-cell's colonisation probability. Secondly, absences inside the range margin did not represent errors in the dispersal model as they were not due to dispersal limitation. Thirdly, we were aware that some of the apparent absences in the distribution data may have been due to low recording effort rather than be true absences, and that if this was not corrected for the dispersal model would be wrongly penalised for predicting false presences. Modifying the basic likelihood function given above to incorporate these is non-trivial, therefore, we took an algorithmic approach to dispersal kernel parameterisation. We acknowledge that this involves a trade-off between biological realism and statistical interpretability.

We used the addition rule of probability for non-mutually exclusive events to calculate, given the probability of dispersal from each individual occupied grid cell, the overall probability of a grid-cell being unoccupied (see running the dispersal model for more details). In order to avoid penalising models for absences within the range margin we assessed model predictions by comparing the intersection of minimum convex polygons encompassing the species' actual and predicted distribution. To compensate for lower recorder effort we divided the area incorrectly predicted to be occupied by a correction factor; this was calculated by dividing the area occupied during the Portuguese and Spanish bird atlases (in 2004) by the area predicted to be occupied in 2004 based on the expansion rate observed between 1964 and 1999. The dispersal kernel was thus parameterised by minimising the sum at each 10 year time step of the area where common waxbills were falsely predicted to be absent and the area where common waxbills were falsely predicted to be present, with the latter dived by the correction factor described above to compensate for low recorder effort.

The *b* parameter was estimated as 5004. Uncertanty around this estimate, and the consequences for model fit, are shown in Appendix figure 2.1.



Appendix figure 2.1. Profile of dispersal kernel fit with different values of *b*, and consequences of uncertainty of fit (i.e. choosing parameter estimates with similar degrees of support) for overall dispersal kernel.

Isolated populations of common waxbills have been recorded at large (>100km) distances from the nearest grid cell occupied in the previous time step (Fig. 1). These populations could have arisen from separate introductions or by long-distance dispersal (Silva, Reino & Borralho 2002). The origin of these populations has implications for parametrising the dispersal kernel; if they originate from separate introductions then they should not be counted as omission errors. Previous authors have suggested that common waxbill populations in the Algarve in Portugal and along the Mediterranean coast of Spain originate from separate introductions (Ferreira 1984; Reino & Silva 1998). This is supported by the large number of exotic birds recorded in these areas (Costa, Elias & Farinha 1997; Lever 2005) which could indicate that these are major sites of introduction for non-native birds, and the large numbers of common waxbills traded in the Iberian Peninsula (WCMC-CITES 2012). We therefore excluded records of common waxbills in the Algarve in 1984 and along the Mediterranean coast of Spain in 1994 when calculating the area where common waxbills were falsely predicted to be absent. This meant that the resulting dispersal kernel was smaller than it would have been if these isolated populations had been included.

Running the dispersal model

The parameterised dispersal kernel was run in cellular automata starting from 1964, 1974, 1984 and 1994. For each time step the observed distribution data was input into the cellular automaton dispersal model. The probability of a grid cell being colonised by the next ten-year time step was calculated as a function of distance to other occupied grid cells using the addition rule of probability for non-mutually exclusive events. Due to computational limitations, calculations only included the nearest 10 occupied cells, but as the probability of colonisation from more distant grid cells was low this is unlikely to have influenced the results. This gave four probabilities of colonisation for each grid cell. The overall probability of each grid cell being colonised by 2004 could be calculated from these using the addition rule for non-mutually exclusive events.

Discussion on methods for constructing the dispersal model

The dispersal kernel we estimated might differ from one constructed using individual movement data for two reasons. Firstly, it describes the movement of a population, not individuals. Secondly, the distribution data we had available to construct the dispersal kernel were at a coarser time scale than the generation time of the study species; we had distribution data at ten-year time steps, while the common waxbill can breed one or more times a year (Payne 2010).

The rate of expansion of the distribution of the common waxbill was slower in the first decade following introduction than in later time periods (Silva, Reino & Borralho 2002). This slow initial expansion is commonly observed in biological invasions (Hengeveld 1989). Dispersal kernels constructed using only data from the early stages of an expansion are likely to underestimate the subsiquent rate of expansion. This will not cause a problem for using dispersal weighting if the dispersal model describes the expansion up to the distribution used to fit the SDM. For example, a dispersal model constructed using the expansion of the common waxbill between 1964 and 1974 could be used to weight a SDM fitted to data from 1974, but the subsiqent increase in expansion rate would mean the dispersal model would overestimate dispersal limitation if it was used to weight a SDM fitted to distribution data from later time periods.

There were several advantages to using cellular automata given the data available to us. We had data in discrete spatial (UTM grid cells) and temporal (ten-year time periods) units, which matches the operation of cellular automata over discrete spatial and temporal units. Additionally, cellular automata do not require much data (just a time series of distributions) so could be used in a system where we lacked demographic data. However, where additional data is available it is desirable to use it, as dispersal may be influenced by demographic factors. Reaction-diffusion models (Skellem 1951) can incorporate such complexity, for example they can utilise additional data such as intrinsic growth rates and carrying capacities (Okubo *et al.* 1989). However, we anticipate that for many invasions only distribution data will be available, so cellular automata will probably be the most useful dispersal model in many cases.

2.9 Appendix 2-4: Modelling the occurrence of the common waxbill: explanatory variable extraction and data analysis

Climate variable extraction and preliminary analysis

Monthly temperature data were interpolated down to 1km^2 resolution. A generalised additive model (GAM) with longitude, latitude, elevation and their interactions as explanatory variables, as well as all valid simplifications, was fitted for each month. We split data into mutually exclusive training (75%) and testing (25%) sets, parameterised models in the training set and assessed their performance at predicting values in the testing set. The best performing model for each month was selected to downscale the data. Precipitation and daily temperature range data were not downscaled as models for those variables had lower explanatory power (r^2<0.9).

10 climatic variables were calculated from the extracted climate data (Appendix table 2.1). Potential explanatory variables were transformed to better approximate normality, and used as explanatory variables in univariate regressions modelling common waxbill occurrence (Appendix table 2.1). GAMs were fitted to check for non-linear relationships, and regression trees were plotted to visualise the interaction structure of the data. The choice of explanatory variables was constrained by strong multicollinearity.

Appendix table 2.1. Univariate logistic regressions between the common waxbill occurrence in 10km^2 UTM grid cells in the Iberian Peninsula and climatic variables.

Variable	Transformation	β	SE	P	r^2
Annual mean temperature		0.546	0.039	<0.001	0.062
Mean temperature of coldest month		0.413	0.026	<0.001	0.070
Mean temperature of warmest month		0.224	0.037	<0.001	0.009
Annual temperature range		-0.498	0.029	<0.001	0.072
Conrad's continentality index	Squared	-0.009	0.000	<0.001	0.086
Mean annual precipitation	Log	0.393	0.128	0.002	0.002
Monthly precipitation coefficient of variation		1.463	0.442	0.001	0.002
Mean daily temp range		-0.564	0.044	<0.001	0.036
Daily temp range coefficient of variation	Squared	-39.619	2.854	<0.001	0.044
Monthly temperature coefficient of variation		-14.736	0.857	<0.001	0.086

Conrad's continentality index, *K*, was calculated by the equation *K*=1.7(annual temperature range/sin(station latitude +10))-14.

Land-cover variable extraction and preliminary analysis

Corine land-cover classes were aggregated into functionally similar habitat classes (Appendix table 2.2). Regression trees were used to identify important land-cover variables, and interactions between variables. Model fit was found to be improved by converting land-cover variables from continuous areas to binary factors denoting presence-absence, so the regression trees were fitted again to the binary land-use data, and land-cover variables were included in the global model as binary factors.

Appendix table 2.2. Aggregation of Corine land-cover classes into groups used in this analysis.

Name	Corine land-cover code
Built	111, 121, 122, 123, 124, 133
Parks and gardens	112, 141, 142
Woody agriculture	221, 222, 223, 244
Irrigated agriculture	212
Rice	213
Bare	All 33s, 131, 132
Arable	211
Heterogeneous agriculture	241, 242, 243
Forest	311, 312, 313
Grassland	321, 231
Scrub and heath	322, 323, 324
Freshwater wetlands	411, 511, 512
Saline wetlands	422, 423, 521, 522

Recorder effort

As spatial variation in detection probability due to varying recorder effort can lead to spurious relationships between species occurrence and environmental variables being identified (McClintock *et al.* 2010; Hill 2011; Kéry 2011)**,** we used the proportion of 10 widespread native species (Eurasian blackbird *Turdus merula*, Sardinian warbler *Sylvia melanocephala*, spotless starling *Sturnus unicolor*, house sparrow *Passer domesticus*, Eurasian chaffinch *Fringilla coelebs*, European serin *Serinus serinus*, European greenfinch *Carduelis chloris*, European goldfinch *Carduelis carduelis*, Eurasian linnet *Carduelis cannabina* and corn bunting *Miliaria calandra*) confirmed to breed in each grid cell as a proxy for recorder effort. The proportion of recorded species confirmed to breed in each grid cell was used as an indication of coverage during the most recent Britain and Ireland bird atlas (British Trust for

Ornithology 2012), and we believe it makes a good proxy for recorder effort as we assume that for non-colonial, generalist, non-migratory species for which breeding cannot be confirmed through specialist survey techniques, variation in recorder effort is more important than habitat suitability in influencing whether the species is confirmed to breed in a grid cell rather than just recorded as present. We note that observers visiting late in the breeding season are more likely to confirm breeding than observers visiting earlier in the breeding season; this will still positively relate the detection probability of common waxbills as our field observations show that common waxbills are more likely to be recorded later in the breeding season (M Sullivan, unpublished data).

Data analysis

Variance inflation factors for climate and land-use global models, calculated using the package car (Fox 2009), were below three, so there was no strong multicollinearity in either global model. We followed a multi-model inference (MMI) approach (Burnham & Anderson 2002) to model selection. As we lacked the *a priori* information to build a small set of plausible candidate models from the global models (Johnson & Omland 2004), we used functions in the package MuMIn (Barton 2011) to fit all valid simplifications (i.e. respecting the principle of marginality) of the climate and land-use global models, and obtain AIC values for each model. ΔAIC values were used to calculate AIC weights for each model across the full set of models. The importance of explanatory variables was calculated by summing the AIC weights of models the variable appeared in. The AIC weighted average of parameter values was calculated for the set of most highly weighted models where AIC weights sum to 0.95 (95% confidence set; Burnham & Anderson 2002). These results are presented in Appendix table 2.3.

.

Appendix table 2.3. Land-use and climate models for the occurrence of the common waxbill in 10km^2 UTM grid cells in the Iberian Peninsula.

Explanatory variable		DWGLM			GLM		
		β	SE	Importance	β	SE	Importance
Land-use							
	Built	-1.36	0.217	1	-2.00	0.26	1
	Irrigated agriculture	0.936	0.205	1	0.59	0.268	1
	Parks and gardens	1.59	0.167	1	1.84	0.184	1
	Rice	0.335	0.601	1	1.67	0.517	1
	Rice: Woody agriculture	2.21	0.675	1	2.17	0.475	1
	Woody agriculture	0.205	0.151	1	0.672	0.123	1
	Built: Rice	-0.734	0.614	0.77	0.03	0.212	0.27
	Built: Irrigated agriculture	-0.116	0.233	0.35	-0.658	0.249	0.98
	Parks and gardens: Irrigated agriculture	-0.067	0.188	0.3	0.403	0.317	0.77
	Parks and gardens: Rice	-0.113	0.314	0.25	-1.45	0.422	1
	Recorder effort	0.026	0.163	0.22	-0.028	0.139	0.27
	Built: Parks and gardens	0.004	0.124	0.14	-0.128	0.236	0.4
Climate							
	MAP	76.5	20.9	1	36.4	15.2	1
	MAP quadratic	-1.88	0.764	1	-1.48	0.47	1
	MAP:MDTR	-5.56	1.49	1	-1.68	1.15	0.81
	MAP:MTCM	-5.44	1.54	1	-1.5	1.2	1
	MAP:MTCM:MDTR	0.614	0.159	1	0.19	0.122	0.81
	MDTR	37.7	9.64	1	19.6	7.41	1
	MTCM	35.5	9.74	1	11.9	7.72	1
	MTCM:MDTR	-3.94	1	1	-1.34	0.777	1
	MDTR quadratic	-0.107	0.084	0.68	-0.358	0.04	1
	Recorder effort	-0.084	0.224	0.24	-0.903	0.268	1

The importance of each variable was calculated by summing the AIC weights of the models in which that variable appeared. Model averaged parameter estimates and standard errors are shown. Note that land-use variables are two level factors, and the parameter values show the change in logit(occupancy) when the land-use class is present in a 10km^2 grid cell.

References

Barton, K. (2011) MuMIn: Multi-model inference. R package version 1.0.0.

British Trust for Ornithology (2012) Bird Atlas 2007-2011. http://blx1.bto.org/atlas-results/percentconfirmed.html. Accessed 20/01/2012.

Burnham, K.P. & Anderson, D.R. (2002) *Model selection and multi-model inference, a practical information-theoretic approach*. Springer, New York.

Costa, H., Elias, G.L. & Farinha, J.C. (1997) Exotic birds in Portugal. *British Birds,* **90,** 562-568.

Ferreira, L.F.F. (1984) Breves notas sobre ocorrência de *Estrilda astirld* (L.) [Aves-*Estrildidae*] no Algarve. *Textos das communicações do 3 Congresso sobre o Algarve*, pp. 569-574.

Fox, J. (2009) car: Companion to Applied Regression. R package version 1.2-14. http://www.r-project.org, http://socserv.socsci.mcmaster.ca/jfox/.

Hengeveld, R. (1989) *Dynamics of biological invasions*. Chapman and Hall, London.

Hill, M.O. (2011) Local frequency as a key to interpreting species occurrence data when recording effort is not known. *Methods in Ecology and Evolution*.

Johnson, J.B. & Omland, K.S. (2004) Model selection in ecology and evolution. *Trends in Ecology & Evolution,* **19,** 101-108.

Kéry, M. (2011) Towards the modelling of true species distributions. *Journal of Biogeography,* **38,** 617-618.

Lever, C. (2005) *Naturalised Birds of the World*. T & AD Poyser, London.

McClintock, B.T., Bailey, L.L., Pollock, K.H. & Simons, T.R. (2010) Unmodeled observation error induces bias when inferring patterns and dynamics of species occurrence via aural detections. *Ecology,* **91,** 2446-2454.

Okubo, A., Maini, P.K., Williamson, M.H. & Murray, J.D. (1989) On the spatial spread of the Grey Squirrel in Britain. *Proceedings of the Royal Society of London. B. Biological Sciences,* **238,** 113-125.

Payne, R.B. (2010) Estrildidae. *Handbook of the Birds of the World* (eds J. del Hoyo, A. Elliott & D.A. Christie). Lynx Edicions, Barcelona.

Reino, L., Moya-Larano, J. & Heitor, A.C. (2009) Using survival regression to study patterns of expansion of invasive species: will the common waxbill expand with global warming? *Ecography,* **32,** 237-246.

Reino, L. & Silva, T. (1998) The distribution and expansion of the Common Waxbill (*Estrilda astrild*) in the Iberian Peninsula. *Biological Consereve e Fauna,* **102,** 163-167.

Silva, T., Reino, L.M. & Borralho, R. (2002) A model for range expansion of an introduced species: the common waxbill Estrilda astrild in Portugal. *Diversity and Distributions,* **8,** 319-326.

Skellem, J.G. (1951) Random dispersal in theoretical populations. *Biometrika,* **38,** 196-218.

3 Dispersal and demographic processes are more important than climate in influencing habitat specificity at expanding range margins of a non-native bird

Submitted to Ecography as Sullivan, M.J.P. and Franco, A.M.A. Dispersal and demographic processes are more important than climate in influencing habitat specificity at expanding range margins of a non-native bird.

Many species are not at equilibrium with their environment. This includes spreading non-native species and species undergoing range shifts in response to climate change. The habitat specificity of these species has been shown to vary between their range margins and range core. This has been attributed to climate, but could also be related to dispersal and demographic processes. We investigate the habitat associations of the non-native common waxbill along multiple expansion axes in the Iberian Peninsula. These have different degrees of climatic suitability, allowing us to disentangle the effects of climate from both dispersal and demographic processes. In contrast to previous studies we find a strong independent effect of dispersal and/or demographic processes, with greater habitat specificity at range margins. The results of this study highlight the importance of considering dispersal and demographic processes when investigating the species-environment relationship of species that are not at equilibrium with their environment, and demonstrates that even a simple additive effect of dispersal and/ or demographic processes can dramatically alter habitat associations in the range core and range margin.

3.1 Introduction

The distributions of many species are not static. Species are shifting their ranges in response to climate change (Hill, Thomas & Huntley 1999; Parmesan & Yohe 2003; Hickling *et al.* 2006), while species transported to new areas by humans are spreading to suitable areas in their non-native range (Sullivan *et al.* 2012; Václavík & Meentemeyer 2012). Species distribution models are commonly used to predict the potential distribution of these species (Peterson 2003; Jimenez-Valverde *et al.* 2011), but typically assume that a species is in equilibrium with their environment (Guisan & Thuiller 2005). Spreading species often violate this assumption (Elith, Kearney & Phillips 2010) as species may fail to occupy all suitable environmental space due to dispersal limitation (Brooker *et al.* 2007; Devictor *et al.* 2012; Sullivan *et al.* 2012). In addition, populations at range margins are often at lower population densities than populations in the range core (Brown 1984), exposed to marginal climates (Oliver *et al.* 2009) and consist of a non-random subset of the genetic diversity of the species (Thomas *et al.* 2001; Waters, Fraser & Hewitt 2013).

Differences in habitat associations between species range core and range margin have been documented (Thomas *et al.* 2001). These could be a consequence of species spreading into marginal climates, as climate has been found to influence habitat associations (Davies *et al.* 2006; Suggitt *et al.* 2012), with unfavourable climate preventing species from occupying poorer-quality habitats (Forister *et al.* 2010; Barnagaud *et al.* 2012; Estrada-Villegas, McGill & Kalko 2012). Butterfly species in the UK have been found to have higher habitat specificity at expanding range margins with marginal climate (Oliver *et al.* 2009). The interaction of climate and habitat in shaping ectotherm distributions has been partially attributed to the microclimates provided by different habitats (Suggitt *et al.* 2012). While this is likely to be less important for endotherms, climate and habitat can still interact as resource rich habitats can enhance survival and breeding success in unfavourable climates (Robb *et al.* 2008).

Differences in habitat specificity between range margins and the range core could also be due to processes due to dispersal and/or demography. Population densities are often lower at range margins (Brown 1984) as populations are yet to reach equilibrium densities. Species may show density dependent habitat use, with the most favourable habitats occupied at low population densities and less favourable habitats only occupied at high population densities (Morris 1987). Populations at range margins are often made up of the descendants of a few long distance dispersers (Dytham 2009; Waters, Fraser & Hewitt 2013), and this reduced genetic diversity may also influence habitat specificity. We collectively call these mechanisms dispersal and/or demographic processes (DDP).

Disentangling the role of climate and DDP in influencing the habitat specificity of range expanding species is challenging. DDP and climate suitability are often confounded, with naturally range expanding species moving into climatically marginal areas. The spread of non-native species provides an opportunity to disentangle the effects of climate and DDP, as species are not necessarily moving into more climatically marginal areas. The expansion of the common waxbill *Estrilda astrild* into areas of varying climatic suitability in the Iberian

Peninsula provides such an opportunity. We assess the importance of climate and DDP in influencing the habitat specificity of the common waxbill in its range core and range margin. To do this, we quantified the habitat associations of common waxbills along three axes of range expansion, which aside from isolated populations in Cataluna and Andalucia covered the species' entire European non-native distribution. While some range expansion is into climatically marginal areas, other axis are expanding into climatically favourable areas, meaning that climatic suitability and residence time are not correlated. We tested whether the habitat associations of common waxbills differed between recently colonised areas and areas that have been colonised for longer, and whether climatic suitability influenced habitat associations.

3.2 Methods

We sampled along three axes of common waxbill range expansion and compared habitat associations in the range core with the range margin. These axes were along the west coast of Portugal from introduction sites near Lisbon and Óbidos, along the south coast of Portugal into south-west Spain from introduction sites in the Algarve, and along the Guadiana valley east into Spain (Silva, Reino & Borralho 2002). This sampling design enabled the influence of residence time to be disentangled from climate, as climate suitability varied between expansion axes. For example, common waxbills introduced to the Lisbon area spread along the west coast of Portugal through areas identified to be climatically suitable by Sullivan *et al.* (2012), and also eastwards into less climatically suitable areas such as Extremadura. We used residence time as a proxy for dispersal, since recently colonised areas are likely to be more dispersal limited than areas that have been colonised for longer as populations have had less time to reach equilibrium densities.

We selected 41 10km squares that contained at least two habitat types that were identified as suitable for common waxbills based on literatures searches (i.e. two or more of irrigated agriculture and rice fields – Corine land-cover (CLC) classes 212 and 213, wetlands and rivers – CLC 411 and 511, and heterogeneous agriculture – CLC level two class 24, Reino and Silva 1998). Point counts were located in or around the selected 10km square (see Figure 3.1 for locations of site centroids). Each set of point counts per 10km square is referred to as a site. Sites could be located in adjacent 10km squares, but point counts in each site were non-overlapping. Sites were assigned a residence time based on the date the 20km by 20km UTM grid-cell they fell in was occupied (Silva, Reino & Borralho 2002). Sites were approximately balanced by residence time date (<10 years, 10 – 20 years, 20 - 30 years, >30 years), and the full ranges of residence times in each expansion axis were sampled. Seasonal effects were controlled for by surveying each expansion axes three times during the fieldwork period (April to June 2011), surveying a third of sites in each residence time strata in each period. At each site, five to 12 point counts were carried out in potentially suitable habitat (wetlands, rivers, rice fields, irrigated agriculture, heterogeneous agriculture), depending on the extent of suitable habitat. In total 349 point counts were performed. Point counts were always >200m apart. At each point count location, the presence or absence of common waxbills during the five minute survey was recorded. All point counts were performed by the same observer (MS). The habitat classes present (see Table 3.1 for habitat classes) at 30m intervals on a grid stretching 90m in each direction from the point count location were recorded (i.e. 49 habitat recording points per point count, see Figure 3.1b for schematic). The presence or absence of a river within 100m of the point count location was noted.

We investigated the microhabitat associations of common waxbills in order to quantify the resources provided by different habitats. This enabled us to score the amount of resources point count locations provided, allowing us to compare the quality of different point counts locations in different habitats. We could then test whether common waxbills avoided lower quality habitats in the range margins or in climatically marginal areas. Microhabitat selection by common waxbills was quantified using focal watches at 68 locations. Habitat availability was recorded in a 180m by 180m square, divided into 30m by 30m sub-squares. The percentage cover was recorded in each sub-square. By recording the amount of habitat in sub-squares at different distances from the observers we were able to adjust the calculation of habitat availability to account for the decline in detectability with distance from observer (see Appendix S1 for details and Figure 3.1c for schematic). Habitat use by common waxbills was recorded in scan samples performed every ten minutes, with the observer allowed to walk up and down a transect crossing the middle of the recording area. During each scan sample the distance from observer, habitat use and activity (feeding or shelter) of each group of common waxbill was recorded. Shelter was defined as any rest activities while not feeding. We quantified the selection of each habitat, given availability, for each activity using Jacobs index (Jacobs 1974).

We quantified the amount of feeding and shelter habitat at each point count location by multiplying the proportion of habitat recording points containing a given habitat class by the Jacobs index of that habitat class (negative values of Jacobs index were set to zero for this calculation), then summing these across habitat classes. This gave us a simple measure of the functional resources provided by different point count locations, allowing us to assess habitat quality.

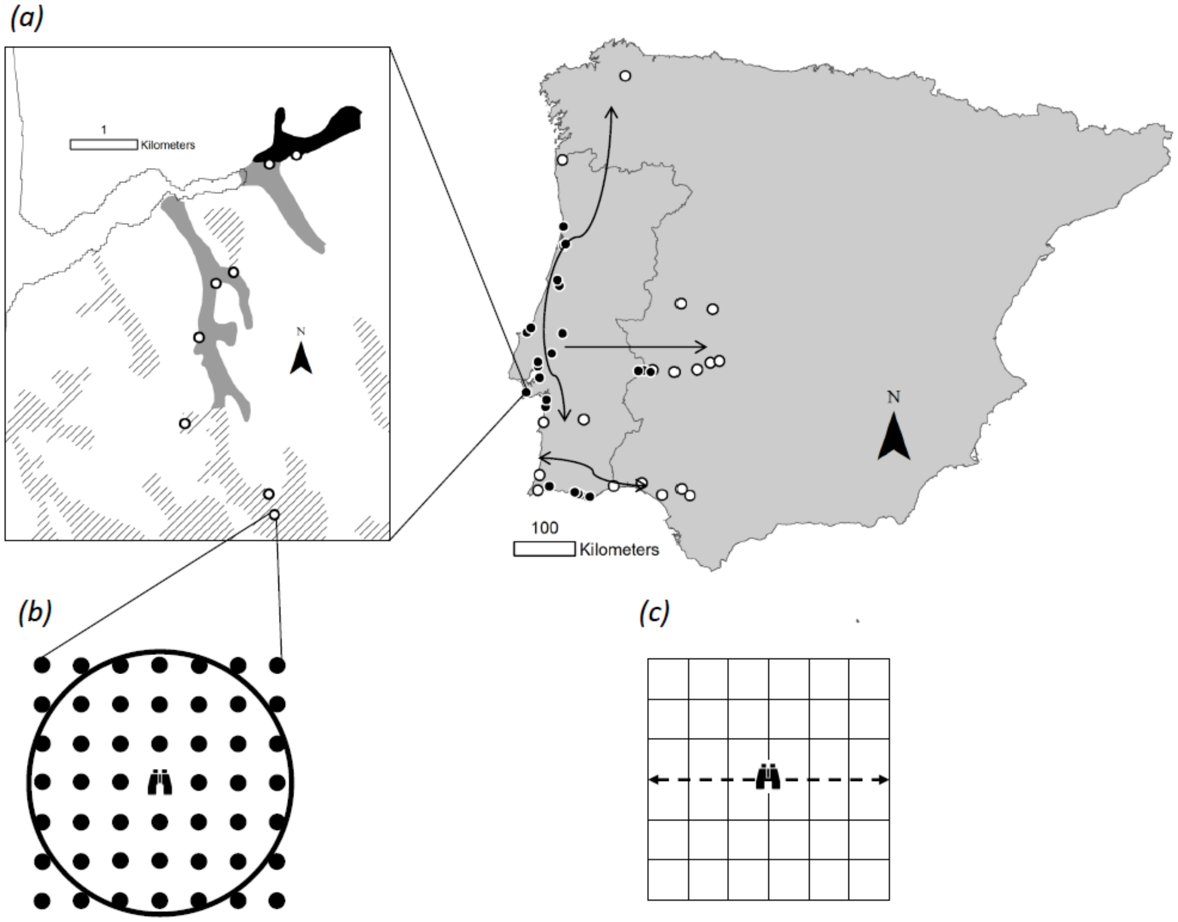


Figure 3.1. (a) Location of survey sites in the Iberian Peninsula. The centroids of each site are plotted. Sites in the range core (i.e colonised by 1990) are shown by filled circles, and sites in the range margin (i.e. colonised after 1990) are shown by open circles. Arrows show axes of range expansion. The insert map shows the location of point counts at one site. Point count locations are shown by open circles. Rice fields are shaded grey, wetlands shaded black, and heterogeneous agriculture (Corine land-cover level two class 24) shown by hashing. The remaining area is largely forestry. (b) Schematic of sampling protocol at each point count. The observer (position shown by binoculars) records birds seen within a 100m radius (shown by circle). Resources are recorded at regularly spaced points (shown by filled circles, habitat also recorded at position of observer). (c) Schematic of sampling protocol at focal watch locations. The observer walks along a central transect (dashed arrow), and records birds and percentage cover of resources in each sub-square.

Table 3.1. Explanatory variables used to model common waxbill occurrence at point counts.

Variable	Abbreviation	Definition
Feeding habitat	F	Proportion of habitat recording points multiplied by Jacob's index value, summed for each habitat with feeding Jacob's index value >0.
Shelter habitat	S	As above but for shelter habitat.
River presence	R	Presence or absence of river within 100m of point count location.
Date	D	Number of days since start of year.
Residence time	Col	Whether 20km grid cell containing point count was colonised before 1990 (i.e. range core) or after 1990 (i.e. range margin).
Climatic suitability	Clim	Probability 10km grid cell containing point count is climatically suitable, obtained from species distribution model described in Sullivan et al. (2012).

We used generalised linear mixed models, with a binomial error distribution and logit link, to model the occurrence of common waxbills at point count locations. These were implemented in lme4 (Bates, Maechler & Dai 2008). We included residence time and its interactions with habitat variables (feeding habitat, shelter habitat and presence of rivers) to test whether habitat associations differed between the range core and the range margin, and included climate suitability (from Sullivan *et al.* 2012) and its interactions with habitat variables to test whether habitat associations varied with climate suitability. With the exception of a null intercept only model, all models contained all habitat variables. We constructed models of progressively increasing complexity from habitat variables only to interactions between habitat and both climate and residence time (see Table 3 for model formulae). Site was included as a random effect. Residence time was simplified to whether or not the grid-cell was colonised before 1990 (i.e. range core and range margin), as models with the more complex variable (decade of first occurrence) did not explain significantly more variation than models with the simple variable. We used sample size corrected Akaike information criterion (AIC_C) to calculate AIC_C weights, which provide measure of the support for a given model being most parsimonious of the candidate set of models (Burnham & Anderson 2002). Variance inflation factors of models were <2, indicating that coliniarity between explanatory variables was sufficiently low as to not confound results. We tested for residual spatial correlation using Moran’s *I*. We did not find significant residual spatial autocorrelation when site was included as a random effect (|*I*|≤0.008, *P*≥0.67).

3.3 Results

Common waxbills selected rough grass, emergent vegetation (including the non-native *Arundo donax*), forbs and houses and gardens for feeding, and emergent vegetation (again including *A. donax*), trees and bushes and to a limited extent forbs for shelter (Table 3.2). Common waxbills were also more likely to be recorded in surveys later in the season (Figure 3.2a) and if rivers were present (Figure 3.2b). The occurrence of common waxbills at point count locations showed positive humped relationships with the amount of feeding and shelter habitat, and was more likely when both feeding and shelter habitat were present (Figure 3.2c and d).

Common waxbills were more likely to occur in their range core than in the range margin (Figure 3.2). Models with interactions between habitat variables and residence time were present in the 95% confidence set of best supported models (Table 3.3. Models of the occurrence of common waxbills. See Table 1 for definition of variable abbreviations., but were less well supported than models with residence time alone, indicating ambiguity about their importance. The biological implications of these interactions were limited, with a small shift in the response curve peak towards areas with more feeding habitat in the range core, towards more shelter habitat in the range margin, and with a greater positive effect of rivers in the range margin (Figure 3.2). In both the range margin and range core common waxbills were most likely to occur at sites with a mixture of feeding and shelter habitat. The independent effect of colonisation time meant that, at any threshold suitability value, common waxbills occurred in a broader range of habitats in the range core than the range margin (Figure 3.2). Although climate suitability appeared in models in the 95% confidence set (Table 3.3), it was less well supported than residence time. Interactions between climate suitability were not well supported, and did not appear in the 95% confidence set of models.

Table 3.2. Microhabitat selection by common waxbills, calculated using Jacobs index (*J*).

Habitat type	Feeding		Shelter	
	N	*J*	*N*	*J*
Rough grass	34	0.44	6	-0.64
Emergent vegetation	19	0.35	44	0.77
Forbs	18	0.25	13	0.08
Houses and gardens	1	0.24	1	-0.23
Arundo donax	3	0.21	12	0.79
Trees and bushes	5	-0.33	17	0.40
Crops	6	-0.59	3	-0.80
Short grass	0	-1.00	0	-1.00

N is the number of observations of each activity in each habitat. In total there were 86 observations of feeding and 96 observations of shelter.

Table 3.3. Models of the occurrence of common waxbills. See Table 1 for definition of variable abbreviations.

Fixed effects	Deviance	AICc	ΔAICc	AIC weight
F + S + F^2 +S^2 + F:S + D + R + Col	387.781	408.432	0.000	0.525
F + S + F^2 +S^2 + F:S + D + R + Col + Clim	387.570	410.353	-1.921	0.201
F + S + F^2 +S^2 + F:S + D + R + Col + Col:F + Col:S + Col:R	383.627	410.714	-2.282	0.168
F + S + F^2 +S^2 + F:S + D + R + Col +Clim + Col:F + Col:S + Col:R	383.505	412.763	-4.330	0.060
F + S + F^2 +S^2 + F:S + D + R + Clim	393.957	414.607	-6.175	0.024
F + S + F^2 +S^2 + F:S + D + R + Col + Clim + Clim:F + Clim:S + Clim:R	386.557	415.814	-7.382	0.013
F + S + F^2 +S^2 + F:S + D + R	399.184	417.715	-9.283	0.005
F + S + F^2 +S^2 + F:S + D + R + Col+ Col:F + Col:S + Col:R + Clim + Clim:F + Clim:S + Clim:R	383.190	419.039	-10.607	0.003
F + S + F^2 +S^2 + F:S + D + R + Clim + Clim:F + Clim:S + Clim:R	392.219	419.305	-10.873	0.002
Intercept only	445.000	449.035	-40.602	0.000

All models are generalised linear mixed effects models with site as a random effect and binomial error structures. Models above the dashed line are in the 95% confidence set of best supported models. : denotes an interaction between variables.

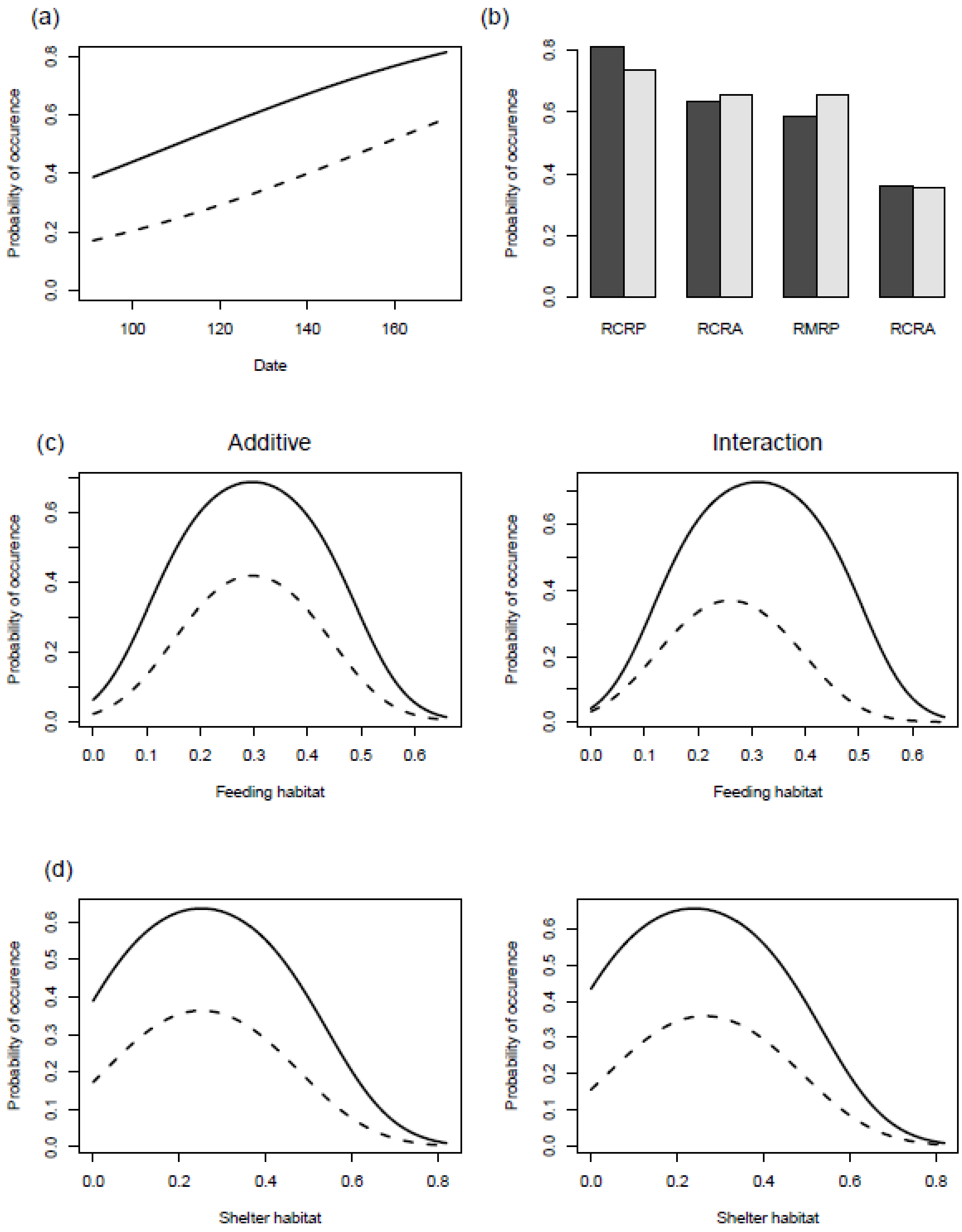


Figure 3.2. Relationship between probability of occurrence and (a) date, (b) river presence, (c) amount of feeding habitat and (d) amount of shelter habitat. Response curves have been constructed by varying the target variables and keeping all other variables at their mean. Solid lines show responses in the range core, dashed lines show responses in the range margin. In all cases except for (c) and (d) the best performing model with residence time as an additive effect was used to construct response curves. In (c) and (d) response curves from

the model with interactions between residence time and habitat variables have also been shown. In (b), RC stands for range core, RM = range margin, RP = river present, RA = river absent.

3.4 Discussion

Common waxbills occupied a broader range of habitats in their range core than in the range margin. Previous studies have found that marginal climates at range boundaries contribute to this pattern (Oliver *et al.* 2009), however, we found a strong effect of residence time independent of climate, and only a weak effect of climate suitability. This study contributes to the growing understanding that species habitat associations vary spatially and through time (Thomas *et al.* 2001).

Although there was some uncertainty over the inclusion of interactions with residence time, the consequences of this were limited, as common waxbills were shown to occupy a broader range of habitats in the range core regardless of inclusion or exclusion of interactions. The broader habitat associations in the range core could be due to greater genetic diversity there (Waters, Fraser & Hewitt 2013) and to selection of less favourable habitats at higher population densities (Morris 1987). Common waxbills were introduced to multiple locations (Silva, Reino & Borralho 2002), and while populations in western Iberian are now part of a contiguous distribution, genetic patterns are likely to be complicated. It is currently not known whether populations at the range margin are less genetically diverse than those in the range core. The range of habitats occupied by common waxbills has increased through time as they have reached higher population densities (Reino & Silva 1998), supporting the role of density dependent habitat selection.

We found limited evidence for the role of climate in shaping habitat associations at the range margin. Our failure to find a strong effect of climate could be because this study looked at an endotherm, while previous studies that have found strong climate-habitat interactions have looked at ectotherms (Oliver *et al.* 2009). These interactions were partially driven by the microclimates provided by different habitats (Suggitt *et al.* 2012); habitat is unlikely to modulate the physiological effects of climate to the same extent in endotherms. Despite this, climate could plausibly interact with the common waxbill's habitat associations in several ways. Firstly, winter survival is related to a bird's energy balance; in order to survive cold weather birds need to increase their food intake (Newton 1998; Siriwardena, Calbrade & Vickery 2008). Winter survival is therefore likely to be higher in higher quality habitats. As common waxbills spread east into central areas of the Iberian Peninsula they encounter colder conditions (Sullivan *et al.* 2012), so this interaction may be increasingly important as the species spreads. Secondly, common waxbills typically breed in mesic habitats (Reino & Silva 1998), and in arid areas in their native range are restricted to wetlands (Barnard 1997). This is unlikely to have occurred in the Iberian Peninsula as precipitation conditions in the Iberian Peninsula are within the species native climatic niche (Sullivan and Franco, unpublished manuscript). Although residence time was more important than climate in influencing habitat associations, the opposite pattern has been found for the common waxbill's behavioural traits (Carvalho *et al.* 2013).

The strong independent effect of residence time means that habitat associations can vary between the range core and the range margin in the absence of interactions with climate. This has implications for the performance of species distribution models of expanding non-native species. These models are likely to underestimate the potential distribution of species, as non-native species initially occupy a subset of potentially suitable habitats. Even for more established species, and range expanding native species, density dependent selection of habitats creates spatial non-stationarity in habitat associations, which can affect model performance if not accounted for (Miller 2012). It is therefore important to explicitly incorporate dispersal related processes in species distribution model fitting procedures (Václavík & Meentemeyer 2009; Sullivan *et al.* 2012).

Large numbers of species are undergoing range shifts in response to anthropogenic change (Hickling *et al.* 2006) and experiencing new conditions in places where they are introduced (Peterson 2003). Understanding their habitat associations and generating accurate predictions of their potential distribution using species distribution models is of crucial importance in conservation planning (Araujo *et al.* 2005). However, the independent effect of residence time on the habitat associations shown here demonstrates that the these species are not at equilibrium with their environment, violating a key assumption of species distribution models (Guisan & Thuiller 2005). Even in the absence of a strong effect of climate or interactions with residence time we found that the range of habitats occupied by a non-native species can vary dramatically between the range core and range margin. These results show that species distribution models that do not account for dispersal related processes are likely to underestimate the future prevalence of the species, with potentially severe impacts for conservation planning. It is thus important to understand the non-equilibrium processes that shape species distributions in order to predict how they will respond to a changing world.

Acknowledgements

We thank Supatchaya Tetachoochert and Chris Sullivan for assisting with fieldwork. MJPS was funded by a NERC PhD studentship.

References

Araujo, M.B., Thuiller, W., Williams, P.H. & Reginster, I. (2005) Downscaling European species atlas distributions to a finer resolution: implications for conservation planning. *Global Ecology and Biogeography,* **14,** 17-30.

Barnagaud, J.-Y., Devictor, V., Jiguet, F., Barbet-Massin, M., Le Viol, I. & Archaux, F. (2012) Relating habitat and climatic niches in birds. *Plos One,* **7**, e32819.

Barnard, P. (1997) Common Waxbill. *The Atlas of Southern African Birds* (eds J.A. Harrison, D.G. Allan, L.G. Underhill, M. Herremans, A.J. Tree, V. Parker & C.J. Brown), pp. 612-613. BirdLife South Africa, Blairgowrie, South Africa.

Barton, K. (2011) MuMIn: Multi-model inference. R package version 1.0.0.

Bates, D., Maechler, M. & Dai, B. (2008) lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-28.

Brooker, R.W., Travis, J.M.J., Clark, E.J. & Dytham, C. (2007) Modelling species' range shifts in a changing climate: The impacts of biotic interactions, dispersal distance and the rate of climate change. *Journal of Theoretical Biology,* **245,** 59-65.

Brown, J.H. (1984) On the Relationship between Abundance and Distribution of Species. *The American Naturalist,* **124,** 255-279.

Burnham, K.P. & Anderson, D.R. (2002) *Model selection and multi-model inference, a practical information-theoretic approach*. Springer, New York.

Carvalho, C.F., Leitão, A.V., Funghi, C., Batalha, H.R., Reis, S., Mota, P.G., Lopes, R.J. & Cardoso, G.C. (2013) Personality traits are related to ecology across a biological invasion. *Behavioral Ecology*.

Davies, Z.G., Wilson, R.J., Coles, S. & Thomas, C.D. (2006) Changing habitat associations of a thermally constrained species, the silver-spotted skipper butterfly, in response to climate warming. *Journal of Animal Ecology,* **75,** 247-256.

Devictor, V., van Swaay, C., Brereton, T., Brotons, L., Chamberlain, D., Heliölä, J., Herrando, S., Julliard, R., Kuussaari, M., Lindström, Å., Reif, J., Roy, D., Schweiger, O., Settele, J., Stefanescu, C., Van Strien, A., Van Turnhout, C., Vermouzek, Z., WallisDeVries, M., Wynhoff, I. & Jiguet, F. (2012) Differences in the climatic debts of birds and butterflies at a continental scale. *Nature Climate Change,* **2,** 121-124.

Dytham, C. (2009) Evolved dispersal strategies at range margins. *Proceedings of the Royal Society B-Biological Sciences,* **276,** 1407-1413.

Elith, J., Kearney, M. & Phillips, S. (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution,* **1,** 330-342.

Estrada-Villegas, S., McGill, B.J. & Kalko, E.K.V. (2012) Climate, habitat, and species interactions at different scales determine the structure of a Neotropical bat community. *Ecology,* **93,** 1183-1193.

Forister, M.L., McCall, A.C., Sanders, N.J., Fordyce, J.A., Thorne, J.H., O'Brien, J., Waetjen, D.P. & Shapiro, A.M. (2010) Compounded effects of climate change and habitat alteration shift patterns of butterfly diversity. *Proceedings of the National Academy of Sciences of the United States of America,* **107,** 2088-2092.

Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters,* **8,** 993-1009.

Hickling, R., Roy, D.B., Hill, J.K., Fox, R. & Thomas, C.D. (2006) The distributions of a wide range of taxonomic groups are expanding polewards. *Global Change Biology,* **12,** 450-455.

Hill, J.K., Thomas, C.D. & Huntley, B. (1999) Climate and habitat availability determine 20th century changes in a butterfly's range margin. *Proceedings of the Royal Society of London. Series B: Biological Sciences,* **266,** 1197-1206.

Jacobs, J. (1974) Quantitative measurement of food selection. *Oecologia,* **14,** 413-417.

Jimenez-Valverde, A., Peterson, A.T., Soberon, J., Overton, J.M., Aragon, P. & Lobo, J.M. (2011) Use of niche models in invasive species risk assessments. *Biological Invasions,* **13,** 2785-2797.

Miller, J.A. (2012) Species distribution models: Spatial autocorrelation and non-stationarity. *Progress in Physical Geography,* **36,** 681-692.

Morris, D.W. (1987) Tests of density-dependent habitat selection in a patchy environment. *Ecological Monographs,* **57,** 270-281.

Newton, I. (1998) *Population Limitation in Birds*. Academic Press, San Diego.

Oliver, T., Hill, J.K., Thomas, C.D., Brereton, T. & Roy, D.B. (2009) Changes in habitat specificity of species at their climatic range boundaries. *Ecology Letters,* **12,** 1091-1102.

Parmesan, C. & Yohe, G. (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature,* **421,** 37-42.

Peterson, A.T. (2003) Predicting the geography of species' invasions via ecological niche modelling. *Quarterly Review of Biology,* **78,** 419-433.

Reino, L. & Silva, T. (1998) The distribution and expansion of the Common Waxbill (Estrilda astrild) in the Iberian Peninsula. *Biologia e Conservazione Della Fauna,* **102,** 163-167.

Robb, G.N., McDonald, R.A., Chamberlain, D.E., Reynolds, S.J., Harrison, T.J.E. & Bearhop, S. (2008) Winter feeding of birds increases productivity in the subsequent breeding season. *Biology Letters,* **4,** 220-223.

Silva, T., Reino, L.M. & Borralho, R. (2002) A model for range expansion of an introduced species: the common waxbill Estrilda astrild in Portugal. *Diversity and Distributions,* **8,** 319-326.

Siriwardena, G.M., Calbrade, N.A. & Vickery, J.A. (2008) Farmland birds and late winter food: does seed supply fail to meet demand? *Ibis,* **150,** 585-595.

Suggitt, A.J., Stefanescu, C., Páramo, F., Oliver, T., Anderson, B.J., Hill, J.K., Roy, D.B., Brereton, T. & Thomas, C.D. (2012) Habitat associations of species show consistent but weak responses to climate. *Biology Letters,* **8,** 590-593.

Sullivan, M.J.P., Davies, R.G., Reino, L. & Franco, A.M.A. (2012) Using dispersal information to model the species–environment relationship of spreading non-native species. *Methods in Ecology and Evolution,* **3,** 870-879.

Thomas, C.D., Bodsworth, E.J., Wilson, R.J., Simmons, A.D., Davies, Z.G., Musche, M. & Conradt, L. (2001) Ecological and evolutionary processes at expanding range margins. *Nature,* **411,** 577-581.

Václavík, T. & Meentemeyer, R.K. (2009) Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? *Ecological Modelling,* **220,** 3248-3258.

Václavík, T. & Meentemeyer, R.K. (2012) Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. *Diversity and Distributions,* **18,** 73-83.

Waters, J.M., Fraser, C.I. & Hewitt, G.M. (2013) Founder takes all: density-dependent processes structure biodiversity. *Trends in Ecology & Evolution,* **28,** 78-85.

4 Why and when do native-trained species distribution models over-predict the potential distribution of non-native species?

Submitted to Diversity and Distributions as Sullivan, M.J.P. and Franco, A.M.A. Why and when do native-trained species distribution models over-predict the potential distribution of non-native species?

Species distribution models (SDMs) are important tools for predicting the spread of non-native species, but their use in this application is challenged by the potential for non-native species to violate the assumption that niches are conserved, and by concerns over their ability to extrapolate to novel environments. We predict the scenarios when SDMs should transfer to new environments, and use simulations and real distribution data to test these predictions. We assessed the contribution of dispersal limitation, niche shifts, and unmodelled environmental variables in influencing species distribution model performance. We hypothesised that SDMs would show highest transferability when both suitable and unsuitable native climate space were found in the non-native range. Our results confirmed this hypothesis, and showed that SDM performance in the native range was a poor predictor of performance in the non-native range. Despite low niche overlap between native and non-native ranges, none of the species included in this study showed evidence for niche shifts. SDMs tended to over-predict species' non-native distributions, largely due to the failure of SDMs to extrapolate temperature responses even in areas with analogous precipitation conditions. SDMs show the best performance at predicting a species non-native distribution when both suitable and unsuitable analogous environmental conditions are found in the non-native range.

4.1 Introduction

Non-native species are major drivers of biodiversity loss (Clavero & García-Berthou 2005), motivating management actions to control some species (Mack *et al.* 2000; Kettenring & Adams 2011). Spatial predictions of their potential distribution, achieved using species distribution models (SDMs), are an important tool for informing management actions (Thuiller *et al.* 2005). For example, assessing the size of non-native species potential distribution is an important component of non-native species risk assessments, as widespread species have more potential to cause conflicts with human activities and native communities (Kumschick & Richardson 2013). However, producing accurate spatial predictions is difficult, as non-native species violate several assumptions of SDMs (Elith & Leathwick 2009). One major concern is the ability of SDMs trained on the native range to accurately predict the distribution of a species in its non-native range. These concerns over the transferability of SDMs arise because the environmental conditions into which a species is introduced may be different from conditions in their native range, so SDMs may be extrapolating into novel environmental space (Wenger & Olden 2012; Zurell, Elith & Schroder 2012). Additionally, species may show different environmental associations in their non-native range due to evolutionary adaptation to new conditions (Muller-Scharer, Schaffner & Steinger 2004; Alexander & Edwards 2010), selection for certain traits prior to establishment (Carrete *et al.* 2012), phenotypic plasticity (Phillimore *et al.* 2012) and genetic drift associated with small founder populations (Hawley *et al.* 2006). Because of these issues, there has been considerable interest in both assessing the transferability of SDMs trained in a species native range (Randin *et al.* 2006; Jimenez-Valverde *et al.* 2011), and assessing whether species occupy different environmental conditions in their non-native range (Broennimann *et al.* 2007; Gallagher *et al.* 2010; Petitpierre *et al.* 2012).

Evaluating whether SDMs perform well in the non-native range is challenging, as non-native species have typically not reached their equilibrium distributions (Jimenez-Valverde *et al.*, 2011; Václavík & Meentemeyer). Methods to evaluate SDM performance typically assess whether patterns of occupancy match predicted suitability. However, spreading non-native species will be absent from some areas that are potentially suitable due to dispersal limitation, and SDMs will be wrongly penalised for these comission errors (Sullivan et al. 2012). Because of this, it has been argued that the ability of SDMs to identify areas where a species is present (its sensitivity) is more important than its ability to identify where a species will be absent (its specificity) (Lobo, Jiménez-Valverde & Hortal 2010; Jimenez-Valverde *et al.* 2011). However, low specificity may arise from genuine model failings, which may mean that the predicted potential distribution of a species is a poor guide for its likely future spread. Identifying whether the apparent over-prediction of non-native species distributions represent genuine model errors is important if SDMs are to be used reliably to guide management actions.

We identify four scenarios, of increasing seriousness for the validity of model predictions, where SDMs predictions poorly relate to the areas occupied by the non-native species. The least serious is for SDMs to over-predict distributions because they predict the potential distribution, while non-native species only occupy the subset of their potential distribution they have been able to disperse to. In this case SDM predictions show the potential spread of a species, but may be a poor guide for their short-term spread. The second scenario is where the SDM correctly characterises a species response in relation to a suite of environmental variables, but these variables do not limit the species distribution in its non-native range, with the non-native distribution instead being limited by unmodelled environmental variables. In this case the SDM predictions are informative, as they correctly show that the non-native range to be suitable for a species for the set of environmental variables used in the SDM, but will perform poorly at classifying which areas will be occupied by a species. The two most serious scenarios are where either the species has undergone a niche shift during the invasion processes (Gallagher *et al.* 2010), or where the SDM has failed to characterise a species' species-environment relationship. In these scenarios the SDM predictions will be incorrect and potentially give misleading indications of the future spread of the species. Identifying which of these scenarios are operating is important if the validity of model predictions is to be evaluated.

Situations where SDMs may perform poorly in the non-native range because they fail to capture the species-environment relationship (i.e. scenario four) are well understood. If the gradients that limit a species distribution are incompletely sampled by the range of environmental variation in the native range the species responses to these variables may not be correctly characterised (Elith and Leathwick 2009). These are likely to become apparent if SDMs are projected into areas with non-analogous environmental conditions (either beyond the absolute range of environmental variables experienced in the training data, or with novel correlations between variables) (Zurell, Elith & Schroder 2012). In contrast, little attention has been paid to identifying the circumstances when SDMs that correctly characterise the species-environment relationship fail to usefully classify between areas that are suitable and unsuitable (scenario two). We hypothesise that SDMs will be able to make useful distinctions between areas of the non-native range that are suitable and unsuitable when both environmental conditions that are suitable and are unsuitable in the native range are present in the non-native range. This is because variables that limit the native distribution are also present in the non-native range (Figure 4.1). When only suitable environmental conditions are present, variables that limit the native distribution are unlikely to limit the non-native distribution. This means that the non-native distribution is therefore likely to be limited by other, unmodelled variables.

We use a simulation to investigate the factors that influence the ability of SDMs trained in the native range to predict equilibrium distributions of species in their non-native range. In particular, we use this simulation to test the hypothesis that SDMs perform better at predicting non-native distributions when both envioronmental conditions that have been shown to be suitable and unsuitable in the native range are present in the non-native range. We then use the case study of the non-native common waxbill *Estrilda astrild* in the Iberian Peninsula to demonstrate how the different factors that can cause poor SDM performance can be identified.

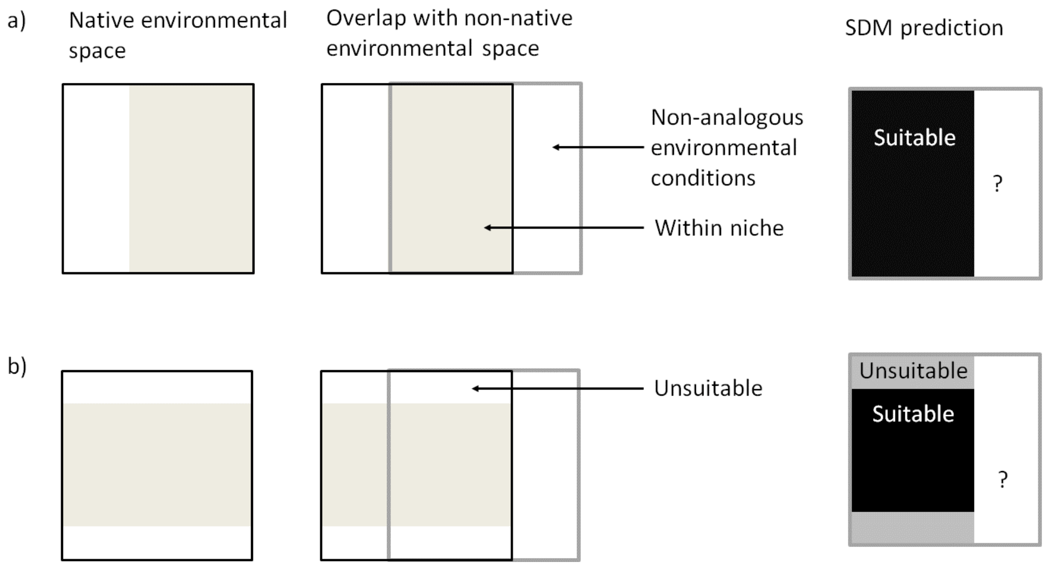


Figure 4.1. Effect of the degree of overlap between native and non-native climate space on the inferences that can be made on the suitability of non-native climate space. Native climate space is shown by the black outlined square, with pale grey shading denoting the native niche. Non-native climate space is shown by the grey outlined square. In (a) non-native climate space is either suitable for the species or non-analogous. Although the same amount of non-analogous climate space is found in (b), both suitable and unsuitable native climates are found in non-native climate space, so native trained SDMs can discriminate between some suitable and unsuitable areas. SDMs will therefore be able to classify areas within analogous climate space as suitable or unsuitable in (b) bur not in (a).

4.2 Methods

4.2.1 Simulation

We explored the factors that influence the transferability of SDMs, and in particular tested the hypothesis that SDMs perform better at predicting the non-native distribution of a species when both environmental conditions that are suitable and unsuitable in the native range are present in the non-native range using a simple simulation. The purpose of this simulation was to test the transferability of SDMs trained on the native range of simulated species with different niches, and test whether the degree of overlap between suitable and available native environmental conditions in the non-native range, and the performance of the SDM in the native range, explained variation in performance in the non-native range. For each simulated species we generated a climate niche, converted that in a stochastic fashion to an ‘observed’ distribution, used this distribution to train a SDM, projected this SDM onto the non-native range, and assessed the correlation between occupied and available native climate space in the species’ non-native range (Figure 4.2).

The climate suitability (S) for a simulated species was given by

$$\mathrm{logit}(S) = \log(S/ (1-S)) = -4 + 1.2 \times (\mathrm{MTCM} + A) - 0.05 \times (\mathrm{MTCM} + A)^2 + 2.5 \times (\mathrm{MTWM} + B) - 0.1 \times (\mathrm{MTWM} + B)^2 + 0.2 \times (\mathrm{dd15} + C) - 0.0008 \times (\mathrm{dd15} + C)^2 + 0.003 \times (\mathrm{MAP} + D) - 0.000001 \times (\mathrm{MAP} + D)^2 + 0.003 \times (\mathrm{ANWB} + E) - 0.000001 \times (\mathrm{AWNB} + E)^2$$

where A to E were randomly selected numbers from a uniform distribution (A = -10 – 5, B = -10 – 10, C = -180 – 40, D = -1000 – 1000, E = -1000 – 2000). The purpose of these was to shift the optima of response curves, so that they varied between being monotonically increasing within the range of environmental variation in the native range to monotonically decreasing. The probability of occurrence was a function of climate suitability and random variation, representing non-climate variability in suitability. The probability of occurrence in a given grid-cell was thus

$$P = S + F$$

where F was randomly selected from a uniform distribution (-0.2 – 0.2). Grid-cells were classed as occupied if these probabilities of occurrence were greater or equal to a number randomly selected from a uniform distribution between zero and one.

The native distribution of simulated species was modelled as a function of the climate variables used in the climate suitability equation, including their quadratic terms. We assessed the ability of SDMs to classify the occupancy status of grid-cells in the native range using the area under the receiver operating characteristic curve (AUC). Models were then projected to the non-native range, and their ability to predict the equilibrium distribution of the non-native species was assessed using AUC. The density of occupied and available areas in the native range were mapped in environmental space, from which we calculated $\mathrm{ND}_{\mathrm{Niche}}$

and $ND_{Available}$ for each grid-cell in the Iberian Peninsula. Simulations were run for 1000 simulated species.

We modelled SDM performance in the non-native range ($Test_{AUC}$) as a function of the correlation between ND_{Niche} and $ND_{Available}$ for temperature (COR_T) and precipitation (COR_P) varaibles (see quantifying climate similarity for details on calculation of these variables), SDM performance in the native range ($Train_{AUC}$) and relative occurrence area (ROA) in the non-native range. The latter was included as a linear and quadratic term in order to control for the known relationship between AUC and ROA. Although these variables were positively correlated, there was sufficient variation to disentangle their contributions, with variance inflation factors ≤ 3.43.

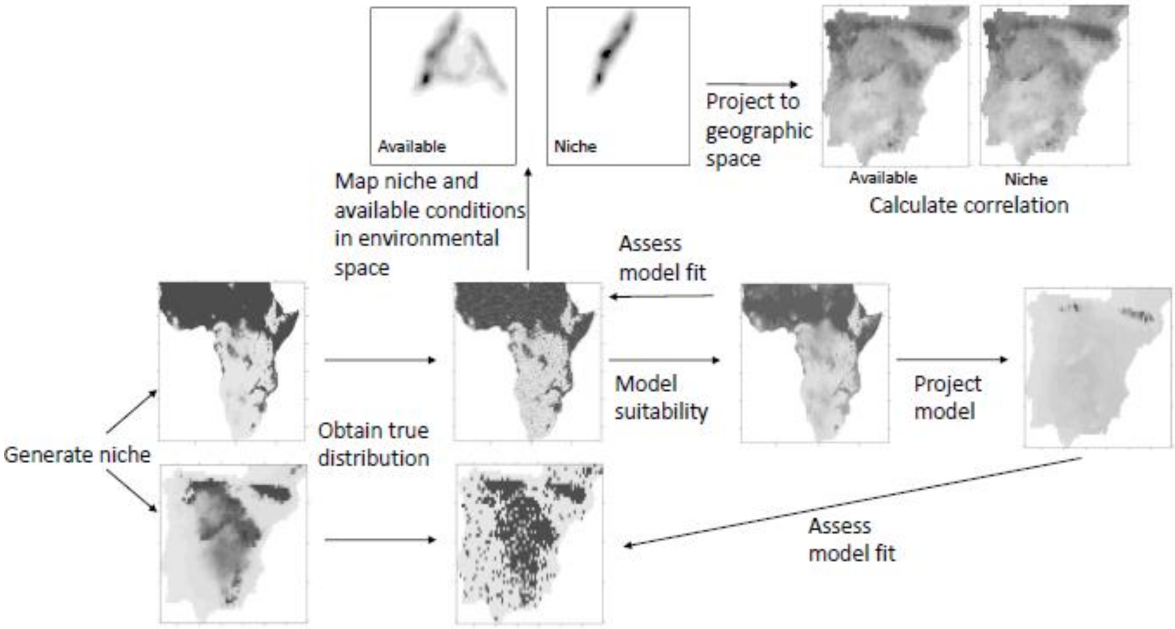


Figure 4.2. Schematic of simulation procedure. Example data from one simulation run are shown.

4.2.2 Case study species

We investigated the climate associations of non-native bird species in the Iberian Peninsula, because recent synchronous atlas surveys of the area have resulted in the availability of fine scale distribution data (Figure 4.3). This allowed us to see if the results of the simulation could be applied to real scenarios. We selected three species, the common waxbill *Estrilda* *astrild* and red avadavat *Amandava amandava* (Passeriformes: Estrildidae), and yellow-crowned bishop *Euplectes afer* (Passeriformes: Plocidae); these were the only established non-native bird species in the Iberian Peninsula not strongly associated with urban areas (and thus urban microclimates) or subject to frequent reintroductions (Matias 2002). The common waxbill and yellow-crowned bishop are native to sub-Saharan Africa, while the red avadavat is native to tropical and subtropical parts of Asia. Yellow-crowned bishops and red avadavats have been established in the Iberian peninsula for over twenty years (Matias 2002), while the common waxbill has been established since the early 1960s (Silva, Reino & Borralho 2002).

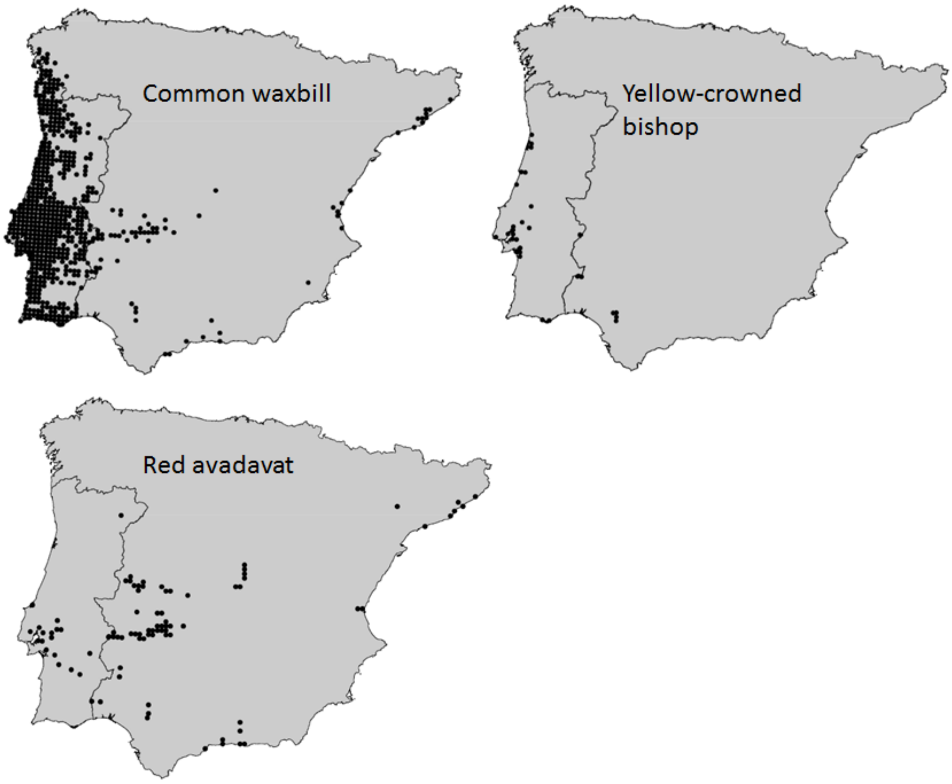


Figure 4.3. Non-native distribution of case study species.

Non-native distributions of the three species were obtained from the most recent Spanish (Marti & de Moral 2003) and Portuguese (Atlas 2008) breeding bird atlases. Distribution data for the yellow-crowned bishop was supplemented with records from a field survey carried out by the authors (Appendix 4-1). The distribution of the common waxbill at two previous time steps, 1984 and 1994, was obtained from Sullivan *et al.* (2012); past distributions of the other species were not available. Non-native distribution data were mapped from 10 x 10 km UTM grid-cells to a 10’grid to match the resolution of climate variables; in cases where the

UTM grid-cell intersected with multiple 10’grid-cells, the occurrence was assigned to the 10’grid-cell with the largest overlap. Native distributions of the three species were obtained from range maps produced by Birdlife International (Birdlife International & NatureServe 2011). 10’grid-cells were classified as occupied if they intersected with a range map polygon.

4.2.3 Climate data

We obtained climate data (averaged over the period from 1960 to 1990) for the native ranges (see Appendix 4-2 for geographical scope) from the CRU CL2.0 10’resolution gridded climatology (New *et al.* 2002). Climate data at the same resolution for the Iberian Peninsula was obtained from the CRU TS1.2 (Mitchell *et al.* 2004), averaged over the ten-year period prior to 2000 (approximately the midpoint of atlas survey work), 1994 and 1984. Seven bioclimatic variables that could plausibly limit the distribution of the species were calculated (Appendix 4-3). These were mean temperature in the coldest (MTCM) and warmest months, number of days above 15°C, annual precipitation, coefficient of variation in monthly precipitation, number of months were potential evapotranspiration exceeds precipitation, and annual balance between precipitation and potential evapotranspiration.

4.2.4 Data analysis

4.2.4.1 Quantifying dispersal limitation

Accounting for dispersal limitation is important in order to investigate its role in niche unfilling, and needs to be accounted for in order to test SDM performance in the non-native range (Sullivan *et al.* 2012). We used a dispersal model, described in full in Sullivan *et al.* (2012), to calculate the probability that each grid cell has been dispersed to. The dispersal model is a cellular automaton where the probability that each grid cell is dispersed to is a function of distance to occupied grid-cells in the previous time step. The decay in dispersal probability with distance was given by a negative exponential dispersal kernel (see Sullivan *et al.* 2012 for details of parameterisation). The model was parameterised could by parameterised for the spread of the common waxbill (the parameterised model was $P_{\text{dispersal}} = e^{-5004x}$, where x is distance from nearest grid-cell colonised in the previous time step, see Sullivan *et al.* 2012 for details). Past distribution data was available only for that species, so we could not parameterise a dispersal model for the other species. Instead, the common waxbill parameterised model was run for one time step from each species’ current distribution. This assumes that the dispersal capacity of the three species is similar, which while plausible is untested. For some analyses we had to select a threshold for selecting whether a grid-cell could be dispersed to. Unless otherwise stated, we set dispersal probabilities ≥0.25 as an arbitrary threshold. Using a higher threshold accentuated the differences between analyses accounting and not accounting for dispersal limitation, but did not change the overall interpretation of the results.

4.2.4.2 Quantifying climate similarity and climatic niches

The methods developed by Broennimann *et al.* (2012) aimed at quantifying niche overlap in environmental space can by extended to map how points in geographic space are to background climate in the native range and to a species native climatic niche. Following Broennimann *et al.* (2012), variation in climate variables across the native and non-native range of a species is collapsed to two dimensions using principal components analysis (PCA). This two dimensional environmental space is divided into a grid of $r \times r$ grid-cells (where r is a arbitrary variable denoting the size of grid cells and thus resolution of environmental space). Geographic locations can be mapped onto this environmental grid based on their PCA scores, and then kernel smoothing functions can be used to obtain the smoothed density of these locations in environmental space. Thus, the smoothed density of available environmental conditions in the native range can be obtained by doing this procedure using all geographic locations in the native range as input, and the smoothed density of a species native niche can be obtained by using geographic locations where the species occurs. The procedure up to this point is described in more detail by Broennimann *et al.* (2012), and R code to implement it is provided in that paper. We note that because PCA space was constructed using climate variables from both the native and non-native range, it is possible to map geographic locations in the non-native range onto this environmental space using their PCA scores. It is then possible to extract both the density of available environmental conditions and the density of a species native niche at that point in environmental space. R code to do this is provided in Appendix S5. We call these metrics $ND_{Available}$ and ND_{Niche}, with the former being the density of available native climate at a point in geographic space, and the latter being the density of a species native climate niche. This method allows geographic maps of available native climate and the density of a species native climate niche to be made. The correlation between them gives an indication of the extent to which areas of analogous climate are within a species niche, with high correlations indicating that almost all analogous climate space is within a species niche.

We did this procedure for the three case study species, using climate space in the Iberian Peninsula and Africa for the common waxbill and yellow-crowned bishop, and climate space in the Iberian Peninsula and south-east Asia for the red avadavat. In both cases we found that the first two axes of PCA space were dominated by precipitation variables (Appendix S4), so we chose to construct separate PCA spaces for temperature and precipitation variables (see Appendix S3 for loadings).

We tested for shifts in the climatic niche of each species, as such niche shifts could lead to poor transferability of SDMs. Using the methods described above, we mapped the density of the native and non-native occurrences in PCA space. We calculated the overlap of the native and non-native climatic niches using D (Schoener 1970), a metric that ranges from 0 (no overlap) to 1 (complete overlap). We tested whether, given available climate space, a species preferentially selected climate space similar to their native climate niche using the niche similarity test (modified from Broennimann *et al.*, 2012, see Appendix S6). This tests whether the overlap between the non-native and native niche is different from the overlap of random niches (selected from available non-native climate space) and the native niche.

4.2.4.3 Species distribution modelling

Generalised linear models (GLMs) and generalised additive models (GAMs) with binomial error structures were trained on the native distribution of each species. Model selection was constrained by strong correlations between climate variables so we used the most complex maximal model permitted by the correlations between climate variables. GLMs were fitted with linear and quadratic terms for each variable, while the complexity of smooth terms in GAMs was selected by generalised cross validation, but constrained to have a maximum complexity of three degrees of freedom (Wood 2006). Models were constructed using logit and complementary log-log links, and the link function that gave the best fit was selected. Non-significant terms were removed from models by stepwise deletion (Crawley 2007). SDMs were projected onto the Iberian Peninsula, and their ability to classify the occupancy status of grid-cells was tested using AUC. This calculation was repeated using only grid-cells with high dispersal probabilities to avoid penalising a SDM for commission errors in areas where the species has not been able to disperse to (Jimenez-Valverde *et al.* 2011).

4.3 Results

4.3.1 Simulation

Simulated species occupied 28.1 ± 0.7 % of their available native range, although this varied between 8.1% and 91.6% between simulations, reflecting differences in the rarity of different combinations of climate variables. Simulated species occupied 63.9 ± 1.0 % of their non-native range. Native trained SDMs showed moderately good performance at classifying the suitability of areas for simulated species ($\text{Train}_{\text{AUC}}$ = 0.760 ± 0.004), but showed poorer performance when transferred to the non-native range (Test_{AUC} = 0.675 ± 0.004). In a univariate regression SDM performance in the native range was negatively related to SDM performance in the non-native range ($\beta = -0.219 \pm 0.035$, $t = -6.3$, $P < 0.0001$), however, this relationship was positive in the multiple regression analysis (Table 4.1). Test_{AUC} was negatively related to COR_{T}, but COR_{P} was not significant (Table 4.1). This indicates that SDM performance in the non-native range was better when both suitable and unsuitable native environmental conditions were present in the non-native range, and that although SDMs that performed better in the native range were also better in the non-native range, performance in the native range was potentially a misleading indicator of SDM performance in the non-native range.

Table 4.1. Model of performance of SDMs of simulated species distributions in the non-native range (Test_{AUC}).

Variable	β	SE	*t*	*P*
ROA	1.849	0.034	53.9	<0.0001
ROA^2	-1.867	0.032	-57.6	<0.0001
COR_{P}	0.015	0.011	1.4	0.173
COR_{T}	-0.257	0.028	-9.2	<0.0001
$\text{Train}_{\text{AUC}}$	0.160	0.027	5.9	<0.0001

$R^2 = 0.831$

4.3.2 Transferability of SDMs

ND_{Niche} and $ND_{Available}$ were strongly correlated for both temperature and precipitation for the common waxbill (r=0.988 and r=0.855 respectively) and yellow-crowned bishop (r=0.911 and r=0.511 respectively), indicating that most analogous conditions were within each species native climatic niche. Spatial congruence of ND_{Niche} and $ND_{Available}$ was lower for the red avadavat (r=0.149 and r=0.298 respectively), indicating that not all areas of analogous climate were within the native climatic niche. We would therefore predict from our framework that SDMs of the red avadavat would perform better at predicting the non-native distribution than SDMs of the other species.

This hypothesis was confirmed, as despite good performance when tested in the native range (AUC = 0.757 – 0.912, Table 4.2), only SDMs of the red avadavat performed well at predicting the non-native distribution (AUC_{GLM} = 0.886, AUC_{GAM} = 0.884, Figure 4.4a, AUC values calculated using dispersal pressure threshold of 0.25, see Table 4.2 for results with other thresholds). SDMs performed poorly at predicting the non-native distribution of common waxbills and yellow-crowned bishops (AUC ≤ 0.571), even when the calculation of AUC was restricted to areas with high dispersal pressure (Table 4.2), as they over-predicted the potential distribution of these species (Figure 4.4b and c).

Table 4.2. Performance of native trained SDMs in the native range (assessed using cross-validation) and in the non-native range.

Species	Model	Native cross-validation AUC	SE	Non-native AUC		
				Dispersal threshold = 0	Dispersal threshold = 0.25	Dispersal threshold = 0.75
Common Waxbill	GLM	0.855	<0.001	0.416	0.509	0.571
	GAM	0.871	<0.001	0.202	0.45	0.542
Red Avadavat	GLM	0.909	<0.001	0.855	0.886	0.992
	GAM	0.912	<0.001	0.862	0.884	0.992
Yellow-crowned Bishop	GLM	0.757	<0.001	0.087	0.296	0.438
	GAM	0.835	<0.001	0.211	0.364	0.45

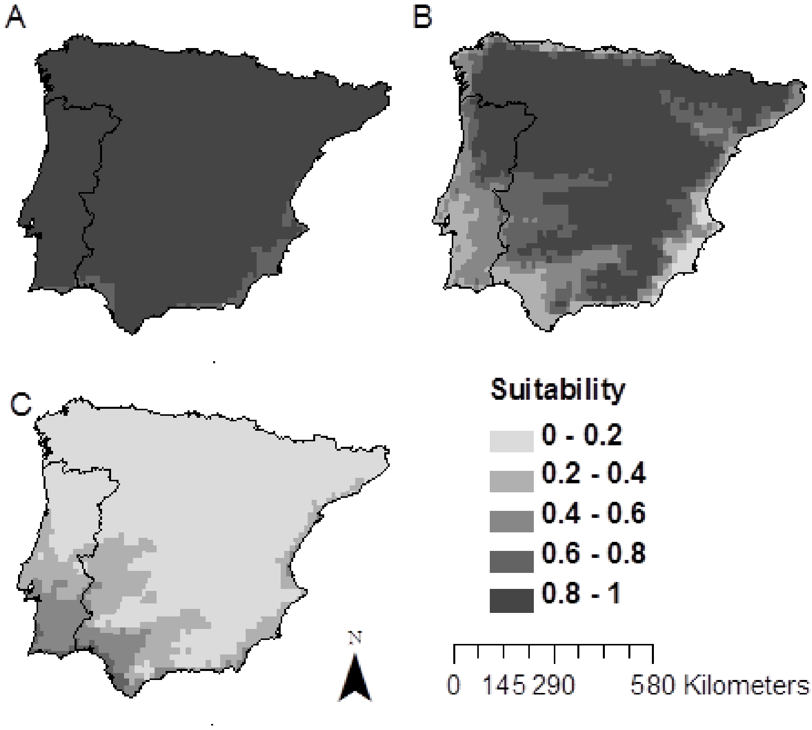


Figure 4.4. Predicted distribution of (A) common waxbill, (B) yellow-crowned bishop and (C) red avadavat by SDMs trained on the native distribution.

4.3.3 Why do SDMs over-predict distributions?

There was considerable unfilling of the climatic niches of all three species (Table 4.3). While some of this was due to dispersal limitation, some niche unfilling remained after accounting for dispersal limitation (Table 4.3). All species selected climatic conditions that were as or more similar to their native niche than expected given the climate in areas where they could disperse to (Table 4.3), indicating an absence of niche shifts.

We looked in detail at the most established non-native species, the common waxbill, to investigate why SDMs over-predicted its distribution. Most of the Iberian Peninsula had temperature conditions that were not analogous to those found in Africa (Figure 4.5), so SDMs were extrapolating into non-analogous climates. SDMs were poor at extrapolating into these areas because they failed to correctly characterise the response to MTCM; native trained SDMs showed that suitability was negatively related to MTCM, while non-native trained SDMs showed a humped relationship (Figure 4.6). In contrast, precipitation conditions within the native climatic niche were widespread in the Iberian Peninsula, indicating that most of the Iberian Peninsula had suitable values of precipitation variables (Fig. 4.5), but there was considerable unfilling of the precipitation niche after accounting for dispersal limitation (Table 4.3).

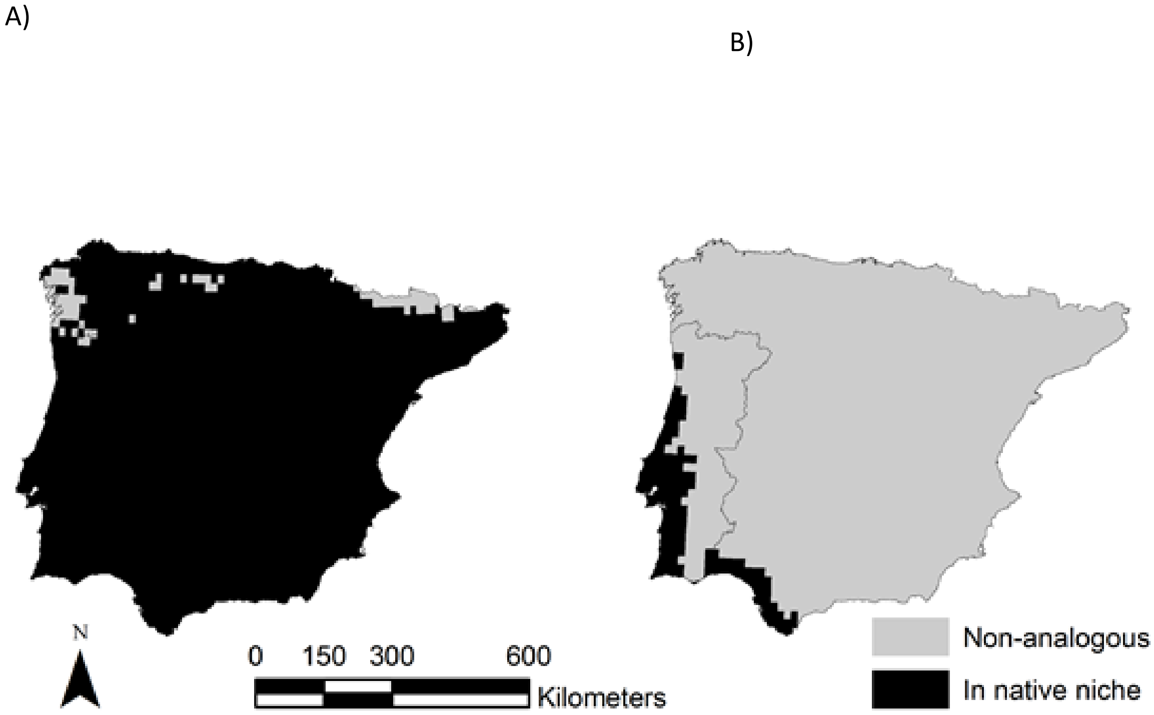


Figure 4.5. Spatial distribution of non-analogous climate conditions and climate conditions in the common waxbill's native niche, mapped for (A) precipitation and (B) temperature. There were no analogous conditions outside the native niche. Producing these maps required selecting an arbitrary threshold for extent of the native niche and analogous conditions. We used the 95th percentile of ND_{Niche} and $ND_{Available}$ values for (A) and the 99th percentile for (B); using the 95th percentile for (B) resulted in all areas being non-analogous.

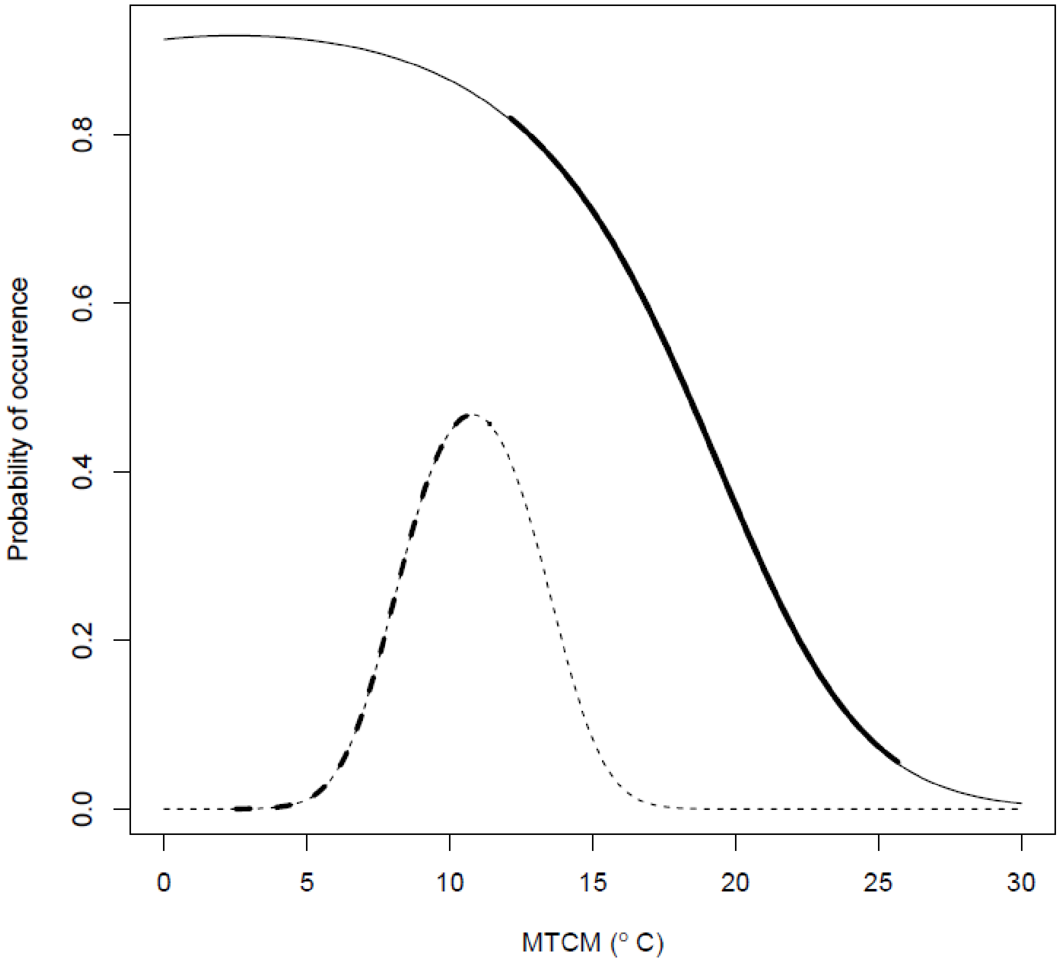


Figure 4.6. Relationship between the probability of common waxbill occurrence and mean temperature in the coldest month (MTCM) derived from SDMs trained on the native (solid line) and non-native (dashed line) distribution. The middle 90% of MTCM values in the training data is shown in bold.

Table 4.3. Unfilling and overlap of the native climatic niche. Expected values of D were calculated using the niche similarity test (see methods for details).

Species	Variable	$D_{Observed}$	$D_{Expected}$	*P*	Unfilling (dispersal)	Unfilling (no dispersal)
Common waxbill	Precipitation	0.423	0.36	0.5	0.152	0.496
	Temperature	0.145	0.147	0.94		
Yellow-crowned bishop	Precipitation	0.359	0.286	0.38	0.609	0.873
	Temperature	0.15	0.05	0.02		
Red avadavat	Precipitation	0.009	0.006	0.36	0.496	0.805
	Temperature	0.057	0.016	0.02	0.725	0.737

4.4 Discussion

We hypothesised that SDMs should show the highest transferability when both suitable and unsuitable native climate space was present in the non-native range. This hypothesis was supported by the results of the simulation, The results of the simulation coincided with a case study focusing on non-native species, where the better performance of native-trained SDMs applied to the red avadavat (where native climate space and the species' niche were weakly correlated in the Iberian Peninsula) than the other species supported our hypothesis. This suggests that the framework presented in Figure 4.1 can be applied to other systems.

We found that SDMs showed poor transferability for two out of the three study species. This reduced performance of SDMs when applied to new regions has been documented previously (Strauss & Biedermann 2007; Barbosa, Real & Vargas 2009; Ervin & Holly 2011). However, in these studies SDMs typically performed better than random when transferred to a new region, while we found worse than random performance of SDMs that performed well in the native range. SDMs over-predicted the potential distribution of our study species. This supports the results of previous studies which have found over-prediction to be common when SDMs are transferred to new regions (Barbosa, Real & Vargas 2009; Stiels *et al.* 2011)**.** Poor SDM transferability was unlikely to be due to niche shifts by non-native species as no statistically significant niche shifts were observed. This supports previous studies, which have found that climatic niche shifts are rare in non-native birds and plants (Petitpierre *et al.* 2012; Strubbe *et al.* 2013).

4.4.1 Why did SDMs over-predict distributions?

None of the study species were at equilibrium with their environment, as niche unfilling was partially attributable to dispersal limitation in all three species. This supports previous studies (Sullivan *et al.* 2012; Václavík & Meentemeyer 2012) that have found that the current distribution of non-native species is largely shaped by dispersal. Although dispersal limitation was an important component in the over-prediction of non-native species distributions, SDMs of common waxbills and yellow-crowned bishops had poor discriminatory power even in areas with high dispersal pressure. This low specificity could result from models poorly characterising the species-environment relationship. Temperature conditions in the Iberian Peninsula were largely non-analogous to those in sub-Saharan Africa, so SDMs were extrapolating beyond the range of the training data. The dangers of such extrapolation has been widely recognised (Elith, Kearney & Phillips 2010), and can lead to both the over-prediction (Barbosa, Real & Vargas 2009) and under-prediction (Barbet-Massin, Thuiller & Jiguet 2010) of species potential distributions. In this study, native trained SDMs modelled the relationship between the occurrence of common waxbills and MTCM as being negative, as common waxbills were absent from parts of the humid tropics with warm winters but widespread in Mediterranean regions of South Africa with cooler winters, so were not limited by cold winter temperatures. Non-native trained SDMs modelled a humped relationship, as common waxbills were absent from areas with cold winters. This difference meant that much of the central Iberian Peninsula, which experiences cold winters, was predicted to be suitable by native trained SDMs and unsuitable by non-native trained SDMs (Sullivan *et al.* 2012).

In contrast to temperature, precipitation conditions in the Iberian Peninsula were largely analogous to those in sub-Saharan Africa. This means that native trained SDMs would be expected to transfer well (Zurell, Elith & Schroder 2012). All species failed to occupy some available precipitation conditions within their native niche, perhaps because these areas had unsuitable temperature conditions. This suggests that either their distribution was limited by other variables, or that their non-native precipitation niche was a subset of their native niche.

The latter mechanism is plausible as local adaptation to climatic conditions means that some native populations have climatic niches that are a subset of the species' niche (Etterson 2004; Phillimore *et al.* 2012). However, we did not find any significant differences between the native and non-native niches of our study species, so this is unlikely. Niche unfilling was more likely to have been driven by an unmodelled, or incorrectly modelled variables. In the case of the common waxbill, this could be habitat availability or temperature. Suitable habitat for the common waxbill was widespread in the Iberian Peninsula (Sullivan *et al.* 2012), so was unlikely to explain niche unfilling. Temperature is a more likely a limiting factor, as non-native trained models have shown that some areas of the Iberian Peninsula with suitable precipitation conditions are too cold in winter for common waxbills (Sullivan *et al.,* 2012).

4.4.2 Interpreting over-predicted potential distributions

We identified that predictions of the potential distribution of the common waxbill were likely to be unreliable if based on SDMs that failed to correctly model the response curve to MTCM. Despite this, we can still obtain useful information from the analyses presented here. Much of the Iberian Peninsula was within the species precipitation niche, so precipitation variables are unlikely to limit its distribution. Therefore, temperature is more likely to be important limiting the species distribution. These variables can be included in future SDMs trained on the non-native distribution.

4.4.3 When are SDMs likely to transfer?

SDMs assume that gradients that limit a species' distribution are adequately sampled (Elith & Leathwick 2009). We demonstrated the poor transferability of SDMs that violate this assumption. SDMs produced very similar estimates of response curves within the range of training data, but produced very different estimates when extrapolated outside this, supporting results of previous studies (Heikkinen, Marmion & Luoto 2012; Wenger & Olden 2012). This explains why the performance of SDMs in the native range (largely interpolating within the range of training data) was a poor predictor of performance when applied to the non-native range (when largely extrapolating outside the range of training data).

Because metrics of SDM performance in the native range are not a guide for non-native performance, caution is needed when evaluating the transferability of SDMs. We suggest the following approach. Response curves of SDMs should be inspected to assess whether gradients have been adequately sampled: humped response curves are likely to indicate that they have been, while linear responses may be the true response they may also indicate that the gradients have not been adequately sampled. If linear response curves are observed, attempt to obtain more occurrence data to ascertain whether the response is likely to be linear or humped. If non-native occurrence data is available, comparing native and non-native response curves can reveal poorly modelled species-environment relationships. Mapping the similarity of the non-native range to occupied and unoccupied native climate space is also a useful tool as it allows areas where SDMs have information about what conditions will be suitable and unsuitable to be visualised. We acknowledge that careful inspection of these is unlikely to be feasible in high throughput modelling exercises involving many species, but suggest that functions that calculate the congruence of suitable and unsuitable environmental conditions are included in modelling workflows.

In this paper we have been concerned with cases where it is not possible to train SDMs on non-native distribution data. Where such data is available, it has been shown to improve SDM performance (Broennimann & Guisan 2008).

4.5 Conclusions

We confirmed our hypothesis that SDM transferability should be highest when non-native climate space overlaps considerably with both occupied and unoccupied climate space. Native-trained SDMs showed poor transferability when only analogous environmental conditions where within a species niche, with a tendency to over-predict the potential distribution of non-native species. This was due to species distributions being limited by either unmodelled or incorrectly modelled environmental variables with poor response curves. Identifying which mechanisms lead to SDMs over-predicting species distributions is essential in order to interpret model predictions.

Acknowledgements

MJPS was funded by a NERC PhD studentship. We thank the ICNF for providing digital atlas data for Portugal, and Birdlife International for supplying species range maps.

References

Alexander, J.M. & Edwards, P.J. (2010) Limits to the niche and range margins of alien species. *Oikos,* **119,** 1377-1386.

Atlas, E. (2008) *Atlas das Aves Nidificantes em Portugal (1999-2005)*. Instituto da Conservacao da Natureza e da Biodiversividade, Sociedade Portuguesa para o Estudo das Aves, Parque Natural da Madeira e Secretaria Regional do Ambiente e do Mar, Lisbon.

Barbet-Massin, M., Thuiller, W. & Jiguet, F. (2010) How much do we overestimate future local extinction rates when restricting the range of occurrence data in climate suitability models? *Ecography,* **33,** 878-886.

Barbosa, A.M., Real, R. & Vargas, J.M. (2009) Transferability of environmental favourability models in geographic space: The case of the Iberian desman (Galemys pyrenaicus) in Portugal and Spain. *Ecological Modelling,* **220,** 747-754.

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.J., Randin, C., Zimmermann, N.E., Graham, C.H. & Guisan, A. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography,* **21,** 481-497.

Broennimann, O. & Guisan, A. (2008) Predicting current and future biological invasions: both native and invaded ranges matter. *Biology Letters,* **4,** 585-589.

Broennimann, O., Treier, U.A., Muller-Scharer, H., Thuiller, W., Peterson, A.T. & Guisan, A. (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters,* **10,** 701-709.

Carrete, M., Edelaar, P., Blas, J., Serrano, D., Potti, J., Dingemanse, N.J. & Tella, J.L. (2012) Don't neglect pre-establishment individual selection in deliberate introductions. *Trends in Ecology & Evolution,* **27,** 67-68.

Clavero, M. & García-Berthou, E. (2005) Invasive species are a leading cause of animal extinctions. *Trends in Ecology & Evolution,* **20,** 110.

Crawley, M.J. (2007) *The R Book*. John Wiley & Sons, Chichester.

Elith, J., Kearney, M. & Phillips, S. (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution,* **1,** 330-342.

Elith, J. & Leathwick, J.R. (2009) Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics,* **40,** 677-697.

Ervin, G.N. & Holly, D.C. (2011) Examining Local Transferability of Predictive Species Distribution Models for Invasive Plants: An Example with Cogongrass (Imperata cylindrica). *Invasive Plant Science and Management,* **4,** 390-401.

Etterson, J.R. (2004) Evolutionary potential of Chamaecrista fasciculata in relation to climate change. 1. Clinal patterns of selection along an environmental gradient in the great plains. *Evolution,* **58,** 1446-1458.

Gallagher, R.V., Beaumont, L.J., Hughes, L. & Leishman, M.R. (2010) Evidence for climatic niche and biome shifts between native and novel ranges in plant species introduced to Australia. *Journal of Ecology,* **98,** 790-799.

Hawley, D.M., Hanley, D., Dhondt, A.A. & Lovette, I.J. (2006) Molecular evidence for a founder effect in invasive house finch (Carpodacus mexicanus) populations experiencing an emergent disease epidemic. *Molecular Ecology,* **15,** 263-275.

Heikkinen, R.K., Marmion, M. & Luoto, M. (2012) Does the interpolation accuracy of species distribution models come at the expense of transferability? *Ecography,* **35,** 276-288.

Hill, M.P., Hoffmann, A.A., Macfadyen, S., Umina, P.A. & Elith, J. (2012) Understanding niche shifts: using current and historical data to model the invasive redlegged earth mite, Halotydeus destructor. *Diversity and Distributions,* **18,** 191-203.

International, B. & NatureServe (2011) Bird species distribution maps of the world. (ed. I. BirdLife). Cambridge, UK and NatureServe, Arlington, USA.

Jimenez-Valverde, A., Peterson, A.T., Soberon, J., Overton, J.M., Aragon, P. & Lobo, J.M. (2011) Use of niche models in invasive species risk assessments. *Biological Invasions,* **13,** 2785-2797.

Jiménez-Valverde, A., Lobo, J.M. & Hortal, J. (2008) Not as good as they seem: the importance of concepts in species distribution modelling. *Diversity and Distributions,* **14,** 885-890.

Kettenring, K.M. & Adams, C.R. (2011) Lessons learned from invasive plant control experiments: a systematic review and meta-analysis. *Journal of Applied Ecology,* **48,** 970-979.

Kumschick, S. & Richardson, D.M. (2013) Species-based risk assessments for biological invasions: advances and challenges. *Diversity and Distributions,* **19,** 1095-1105.

Lobo, J.M., Jiménez-Valverde, A. & Hortal, J. (2010) The uncertain nature of absences and their importance in species distribution modelling. *Ecography,* **33,** 103-114.

Mack, R.N., Simberloff, D., Lonsdale, W.M., Evans, H., Clout, M. & Bazzaz, F.A. (2000) Biotic invasions: Causes, epidemiology, global consequences, and control. *Ecological Applications,* **10,** 689-710.

Marti, R. & de Moral, J.C. (2003) Atlas de las Aves Reproductoras de Espana. Dirección General de Conservación de la Naturaleza-Sociedad Española de Ornithología, Madrid.

Matias, R. (2002) *Aves exóticas que nidificam em Portugal continental*. ICNB, Lisbon.

Medley, K.A. (2010) Niche shifts during the global invasion of the Asian tiger mosquito, Aedes albopictus Skuse (Culicidae), revealed by reciprocal distribution models. *Global Ecology and Biogeography,* **19,** 122-133.

Mitchell, T.D., Carter, T.R., Jones, P.D., Hulme, M. & New, M. (2004) A comprehensive set of high-resoultion grids of monthly climate for Europe and the globe: the observed record (1901-2000) and 16 scenarios (2001-2100). *Working Paper*. Tyndall Centre for Climate Change Research, Norwich, UK.

Muller-Scharer, H., Schaffner, U. & Steinger, T. (2004) Evolution in invasive plants: implications for biological control. *Trends in Ecology & Evolution,* **19,** 417-422.

New, M., Lister, D., Hulme, M. & Makin, I. (2002) A high-resolution data set of surface climate over global land areas. *Climate Research,* **21,** 1-25.

Petitpierre, B., Kueffer, C., Broennimann, O., Randin, C., Daehler, C. & Guisan, A. (2012) Climatic Niche Shifts Are Rare Among Terrestrial Plant Invaders. *Science,* **335,** 1344-1348.

Phillimore, A.B., Stålhandske, S., Smithers, R.J. & Bernard, R. (2012) Dissecting the Contributions of Plasticity and Local Adaptation to the Phenology of a Butterfly and Its Host Plants. *The American Naturalist,* **180,** 655-670.

Randin, C.F., Dirnbock, T., Dullinger, S., Zimmermann, N.E., Zappa, M. & Guisan, A. (2006) Are niche-based species distribution models transferable in space? *Journal of Biogeography,* **33,** 1689-1703.

Schoener, T.W. (1970) Nonsynchronous spatial overlap of lizards in patchy habitats. *Ecology,* **51,** 408-418.

Silva, T., Reino, L.M. & Borralho, R. (2002) A model for range expansion of an introduced species: the common waxbill Estrilda astrild in Portugal. *Diversity and Distributions,* **8,** 319-326.

Stiels, D., Schidelko, K., Engler, J., van den Elzen, R. & Rödder, D. (2011) Predicting the potential distribution of the invasive Common Waxbill Estrilda astrild (Passeriformes: Estrildidae). *Journal of Ornithology***,** 1-12.

Strauss, B. & Biedermann, R. (2007) Evaluating temporal and spatial generality: How valid are species-habitat relationship models? *Ecological Modelling,* **204,** 104-114.

Strubbe, D., Broennimann, O., Chiron, F. & Matthysen, E. (2013) Niche conservatism in non-native birds in Europe: niche unfilling rather than niche expansion. *Global Ecology and Biogeography***, 22**, 962–970.

Sullivan, M.J.P., Davies, R.G., Reino, L. & Franco, A.M.A. (2012) Using dispersal information to model the species–environment relationship of spreading non-native species. *Methods in Ecology and Evolution,* **3,** 870-879.

Thuiller, W., Richardson, D.M., Pysek, P., Midgley, G.F., Hughes, G.O. & Rouget, M. (2005) Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology,* **11,** 2234-2250.

Václavík, T. & Meentemeyer, R.K. (2012) Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. *Diversity and Distributions,* **18,** 73-83.

Wenger, S.J. & Olden, J.D. (2012) Assessing transferability of ecological models: an underappreciated aspect of statistical validation. *Methods in Ecology and Evolution,* **3,** 260-267.

Wood, S.N. (2006) *Generalized Additive Models: An Introduction with R*. Chapman and Hall, London.

Zurell, D., Elith, J. & Schroder, B. (2012) Predicting to new environments: tools for visualizing model behaviour and impacts on mapped distributions. *Diversity and Distributions,* **18,** 628-634.

4.6 Appendix 4-1. Details of field survey of the distribution of yellow-crowned bishops.

The organisers of the Portuguese breeding bird atlas considered that is was likely that there were gaps in the distribution of the yellow-crowned bishop as part of the species' breeding season did not coincide with the main period of atlas fieldwork (Lobo 2008), so a field survey was performed to obtain additional occurrence records. The species is known to inhabit rice fields and irrigated agriculture (Matias 2002), so these habitats in areas near to areas where the species was recorded during the breeding bird atlas (i.e. the lower Teijo and Sado valleys, the Teijo and Sado estuary, the lower Mondego valley and Ria Aveiro) were surveyed during May and June 2011 and 2012, when males were displaying.

4.7 Appendix 4-2. Extent of the native range of study species.

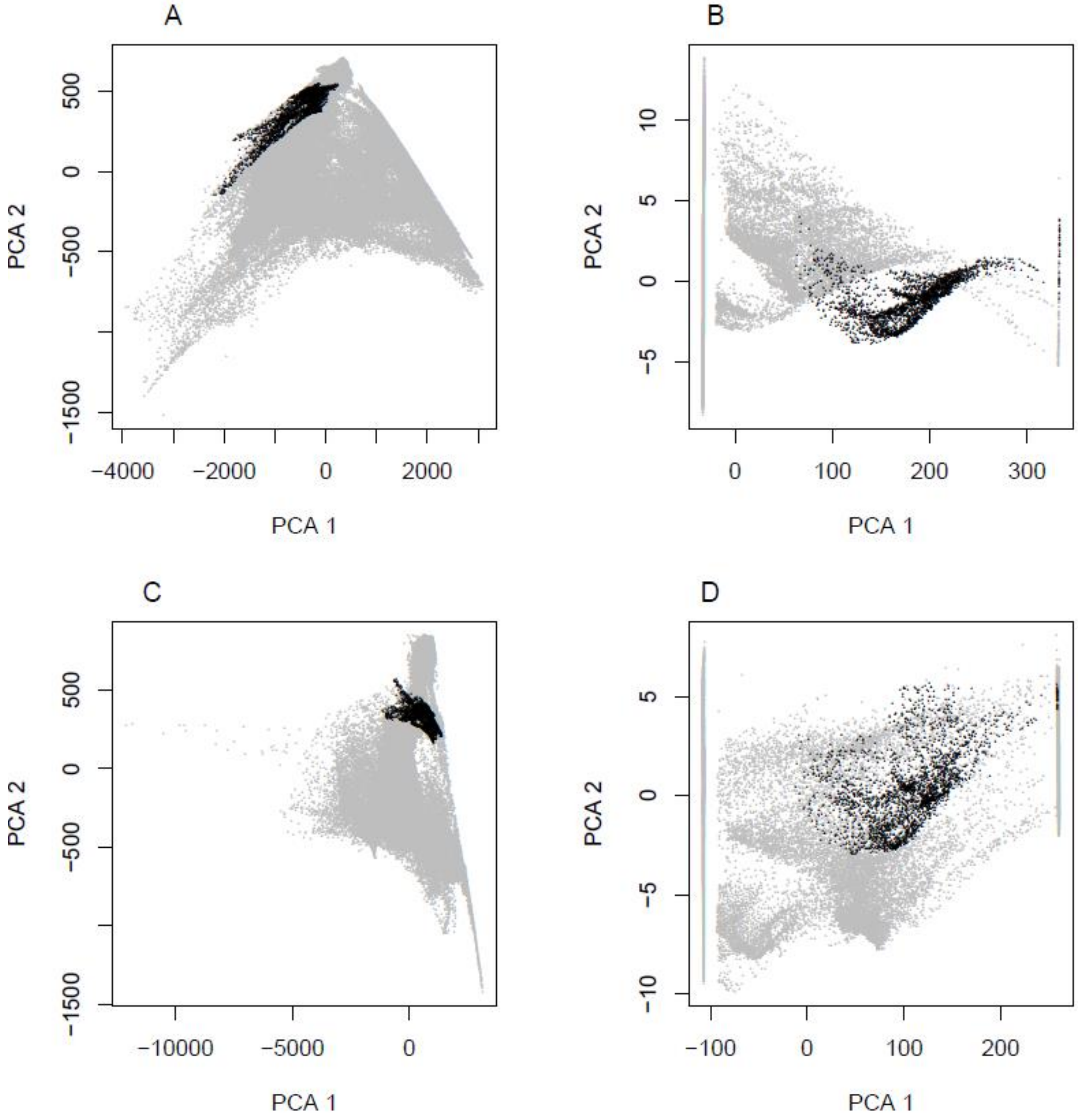
Ideally the native range should be all areas where the species has been able to naturally disperse to. This is not known, so we had to decide on a native range which was likely to encompass all areas where the species could plausibly disperse to but not include areas where the species couldn't reach. For the African species, we set the native range as continental Africa south of 22°N (i.e. sub-Saharan Africa). For the red avadavat, we set the native range as Asia south of 35°N and east of 66°E (i.e. the Indian subcontinent, south-east Asia, Indonesia and the Philippines). Repeating analyses for the common waxbill using continental Africa south of 15°N as the native range had a negligible impact on our results.

4.8 Appendix 4-3. Definition, justification and calculation of bioclimatic variables, and their PCA loadings.

Variable set	Variable	Definition	Reasons for population limitation	Iberia-Africa loading		Iberia-Asia loading	
				PCA 1	PCA 2	PCA 1	PCA 2
Temp	MTCM	Mean temperature in the coldest month	Winter mortality due to cold conditions, low food availability in cold conditions		-0.296		0.61
	MTWM	Mean temperature in the warmest month	Heat stress due to warm conditions, low food availability in cold conditions		-0.954		-0.792
	dd15	Degree days above 15°C	Temperatures below 15°C linked to stress in cage birds, may constrain breeding?	-0.997		-0.995	
Precip	MAP	Annual precipitation	Linked to vegetation types - known to be associated with damp vegetation	-0.435	-0.901	-0.707	-0.707
	PrecipCV	Precipitation coefficient of variation	Seasonality of precipitation linked to seasonality of vegetation and food resources. Seasonal food shortages could cause problems for a sedentary bird				
	lNeg	Number of months with negative water balance	Seasonality of precipitation linked to seasonality of vegetation and food resources. Seasonal food shortages could cause problems for a sedentary bird				
	ANWB	Annual net water balance	Linked to vegetation types - known to be associated with damp vegetation	-0.901	0.435	-0.707	0.707

dd15 was calculated by using a GAM to model monthly mean temperature as a function of date, assuming taking the middle day of each month as the date. The model was used to predict the temperature in every day of the year, and the number of days above 15°C was calculated. PrecipCV is the coefficient of variation of monthly precipitation. lNeg and ANWB were calculated by subtracting total potential evapotranspiration (calculated using the Thornthwaite equation) from total precipitation. Temperature is abbreviated as Temp, and precipitation as precip.

4.9 Appendix 4-4. Distribution of native and non-native climates in PCA space.



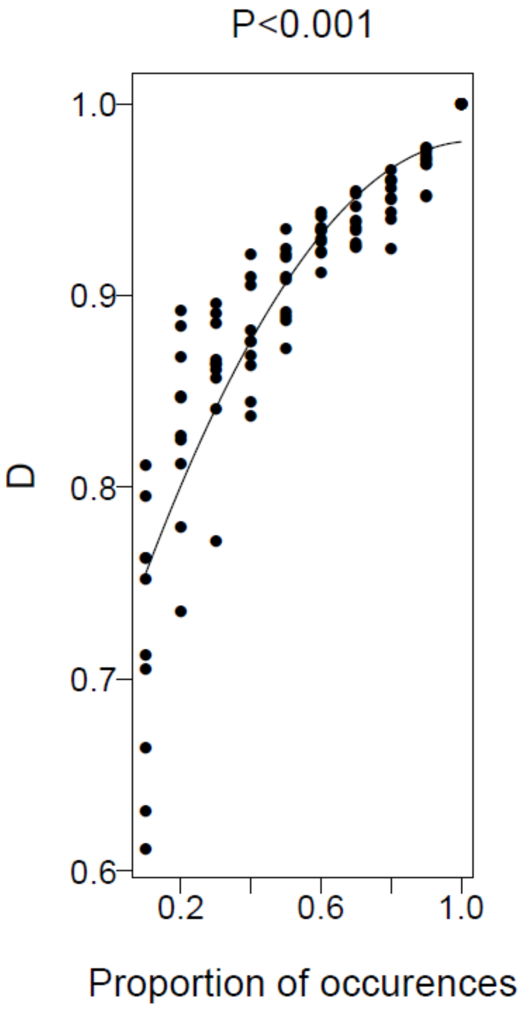
Appendix figure 4.1. Overlap between native (A and B = Sub-Saharan Africa, C and D = Asia) and non-native precipitation (A and C) and temperature (B and D) conditions. Grey dots show native range grid-cells, black dots show non-native range grid cells. In (A) and (C) PCA 1 is negatively correlated with the amount of precipitation an area receives, while PCA 2 is positively correlated with the net water balance of an area relative to the precipitation it receives. Therefore areas in the top right of PCA space are characterised by high precipitation and high surpluses of water, while areas in the bottom right of PCA space have high precipitation but this is offset by high evaporation. In (B) PCA 1 is negatively correlated with temperature, while PCA 2 is negatively correlated with the number of days above 15°C, reflecting seasonality. In (D) PCA 2 is positively correlated with seasonal differences in temperature, while PCA 1 is correlated with number of days above 15°C.

4.10 Appendix 4-5. R code for calculating ND_{Niche} and $ND_{Available}$

```
#Requires function grid.clim in Broennimann et al. 2012 Global Ecology and Biogeography
#Designed to be used in conjunction with functions in that paper
#get.density
#Arguments: scores - subscripts showing rows in PCA scores dataframe in non-native range
#R - parameter to set grid-cell size in kernel smoothing function (see documentation in
Broennimann et al. 2012 Global Ecology and Biogeography)
#zNative - object returned by grid.clim for native range of species
get.density<-function(scores,R,zNative){
x<-zNative$x
y<-zNative$y
spx<-scores[,1]
spy<-scores[,2]
zNative$Z<-zNative$Z/max(zNative$Z)
z.cor<-c()
z.uncor<-c()
Z<-c()
for(i in 1:length(spx)){
xCoord<-length(x[x<spx[i]])+1
yCoord<-length(y[y<spy[i]])+1
if(xCoord>R){
xCoord=R
}
if(yCoord>R){
yCoord=R
}
z.cor[i]<-zNative$z.cor[xCoord,yCoord]
z.uncor[i]<-zNative$z.uncor[xCoord,yCoord]
Z[i]<-zNative$Z[xCoord,yCoord]
}
result<-data.frame("z.cor"=z.cor,"z.uncor"=z.uncor,"Z"=Z,stringsAsFactors=F)
return(result)
}
#Example use
#Dummy dataset - Each row represents a grid cell in geographic space. Columns for
envrionental variables (Temp and Precip), as well as column denoting whether grid cell is in
native or non-native range (Native) and whether the grid cell is occupied (PA)l
dat<-
data.frame("Temp"=rnorm(100,5,3),"Precip"=rnorm(100,100,20),"Native"=c(rep(1,70),rep(0,
30)))
#Giving species highly simplified niche
dat$PA<-ifelse(dat$Temp>4 & dat$Precip<90,1,0)
#Run principal compenets on environmental variables
pcaENV<-princomp(dat[,1:2])
#Store scores of PCA
envscores<-pcaENV$scores
#env12 is PCA scores for native and non-native range
```

```
env12<-envscores[,1:2]
#env1 is PCA scores of native range only
env1<-envscores[dat$Native==1,1:2]
#sp1 is PCA scores for occupied grid cells in native range
sp1<-envscores[dat$Native==1 &dat$PA==1,1:2]
#Set grid cell size for kernel density smoothing
R<-100
#Run grid.clim function from Broennimann et al. 2012 Global Ecology and Biogeography to get smoothed niche densities
zNat<-grid.clim(env12,env1,sp1,R)
#NN.scores if PCA scores in grid-cells in non-native range
NN.scores<-envscores[dat$Native==0,1:2]
#Run get.density function to extract niche densities in each grid cell in non-native range
Niche.density<-get.density(NN.scores,R,zNat)
#Returns niche denisty (z.cor and z.uncor) and density of available environmental conditions(Z) in each grid cell in native range
#If you have lat and long data it is straightforward to map the outputs
#Lat-long data for non-native range
x<-seq(1:5)
y<-seq(1:6)
grid<-expand.grid(x,y)
require(lattice)
levelplot(Niche.density$z.uncor~grid$Var1*grid$Var2)
levelplot(Niche.density$Z~grid$Var1*grid$Var2)
#It is also straightforward to calculate the correlation between occupied and available native environmental conditions
cor(Niche.density$Z,Niche.density$z.uncor)
#In a high throughput analysis this correlation can be linked to errors or warnings
#e.g.
if(cor(Niche.density$Z,Niche.density$z.uncor)>0.5)warning(paste("Correlation is",cor(Niche.density$Z,Niche.density$z.uncor)))
```

4.11 *Appendix 4-6. Sensitivity of D to niche size, and method for performing niche similarity tests.*



Appendix figure 4.2. Relationship between D and niche size. Overlap between the native niche of the common waxbill and different size subsets of occurrences was calculated. *P*-values show the significance of the relationship between D and niche size.

The positive relationship between niche size and D has important implications for the null model used in the niche similarity test. The implementation provided with Broennimann *et al.* (2012) randomly moves the position of the niche, and removes parts of the niche that fall outside non-native climate space. While this preserves niche shape, it means the random niches in the null model may be smaller than the observed niche, so will have lower D values. This results in bias against finding a niche shift. To address this, we modified the procedure for generating random niches so that when part of the niche fell outside non-native climate space, it was reallocated to be within non-native climate space. This preserved niche size, and maintained niche shape as much as possible.

References

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.J., Randin, C., Zimmermann, N.E., Graham, C.H. & Guisan, A. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography,* **21,** 481-497.

Lobo, F. (2008) *Euplectes afer*. *Atlas das aves nidificantes em Portugal* (ed. E. Atlas), pp. 542-543. INCB, Lisbon.

Matias, R. (2002) *Aves exóticas que nidificam em Portugal continental*. ICNB, Lisbon.

5 Do non-native species' niches shift during range expansion? A case study of the common waxbill

Models of the potential distribution of non-native species constructed by projecting their modelled climatic niche onto geographical space are a key tool in the management of these non-native species. These models assume that the niche of a non-native species does not shift during the invasion process. While considerable attention has been paid to documenting niche changes between the native and non-native ranges, the niche dynamics of non-native species during subsequent range expansion are poorly understood. We describe changes in the climatic niche of the common waxbill during its spread through the Iberian Peninsula. We find limited change to the fundamental niche, but considerable changes to the realised niche. These results have implications for the likely performance of species distribution models, as such changes can cause species distribution models to both underestimate and overestimate the potential distribution of non-native species.

5.1 Introduction

Increasing volumes of international trade have contributed to rising numbers of species being introduced to locations outside their native range (Levine & D'Antonio 2003). Predictive modelling of the potential distribution of these non-native species using species distribution models (SDMs) is a key tool for informing their management (Peterson 2003; Thuiller *et al.* 2005). Using SDMs to predict a species' distribution in a different place or time assumes that a species' niche remains constant (i.e. the species exhibits niche conservatism). Because this assumption is critical to the success of predictive distribution modelling, considerable attention has been paid to investigating the niche dynamics of non-native species (Pearman *et al.* 2008).

The realised niche of a species is constrained by the combinations of environmental variables found in their native range. If non-native populations occupy new conditions that are within their fundamental niche, then their realised niche will shift. Fundamental niche shifts can also occur during biological invasions. Non-native populations of species are typically established from small founder populations (Cassey *et al.* 2004), and genetic drift associated with these small populations can lead to niche shifts (Pearman *et al.* 2008). The process of transport and introduction can impose different selection pressures to those experienced by wild populations (Carrete *et al.* 2012), further differentiating non-native and native populations of a species. Multiple introductions from different native populations can combine result in novel genotypes in the non-native range (Lavergne & Molofsky 2007). Finally, selection pressures on the introduced population may be different from those on native populations (Dietz & Edwards 2006). Shifts in the position of the realised niche in environmental space have been documented in non-native species (Broennimann *et al.* 2007; Gallagher *et al.* 2010), although these studies have been criticised for failing to account for niche shape and the availability of conditions in the non-native range (Petitpierre *et al.* 2012). New methods have been proposed to overcome these criticisms (Broennimann *et al.* 2012), and have been used to show that while overlap between the realised niche in the native and non-native range is often low, species typically select environmental conditions in their non-native range that are similar to their native niche and rarely show expansion of their fundamental niche (Petitpierre *et al.* 2012; Strubbe *et al.* 2013).

The niche of a non-native species could also change as it spreads following establishment. The simplest mechanism involves a species expanding its realised niche during range expansion as it encounters a wider range of environmental conditions within its fundamental niche. Expansion of the realised niche many also occur as a species reaches higher population densities due to density dependent selection of less favourable environmental conditions (Brown 1984; Morris 1987). Several processes can cause populations at range margins to have a different fundamental niche to those in the range core. Leading edge populations often consist of the relatives of a few dispersing individuals that increased rapidly due to low intra-specific competition, so contain a subset of the species' genetic diversity (Hewitt 2000; Waters, Fraser & Hewitt 2013). Following Waters *et al.* (2013) we call this the founder takes all process. Populations at range margins can also be different from the range core as adaptation can occur at range margins if populations are exposed to new selection pressures (Thomas *et al.* 2001; Buckley, Butlin & Bridle 2012).

Studies that compare the native and non-native niche at one time period will be unable to detect changes in the non-native niche that could occur during the invasion process. Distribution data at multiple time-points in the invasion process are needed to test whether the non-native niche changes during expansion. Such data are available for the non-native common waxbill *Estrilda astrild* (Passeriformes: Estrildidae) in the Iberian Peninsula (Sullivan *et al.* 2012). We use data on the expansion of the common waxbill to investigate how its realised and fundamental climatic niche has changed through time. We evaluate the evidence for different mechanisms that could have caused these changes, and discuss the consequences of niche change for species distribution modelling.

5.2 Methods

5.2.1 Study species and distribution data

The common waxbill is a granivorous passerine native to sub-Saharan Africa. It was first introduced to Portugal in the 1960s (Silva et al. 2002), and has subsequently spread through much of Portugal and parts of Spain (Reino et al. 2005). The distribution of the common waxbill in the Iberian Peninsula at ten-year time intervals was obtained from Sullivan *et al.* (2012), with the current distribution obtained from the most recent Spanish (Marti & de Moral 2003) and Portuguese (Equipa Atlas 2008) breeding bird atlases. Distribution data were mapped from 10 x 10 km UTM grid-cells to a 10’ grid to match the resolution of climate variables; in cases where the UTM grid-cell intersected with multiple 10’ grid cells, the occurrence was assigned to the 10’ grid cell with the largest overlap. The native distribution of the common waxbill was obtained from range maps produced by Birdlife International (BirdLife International & NatureServe 2011). 10’grid-cells were classified as occupied if they intersected with a range map polygon.

The limits of the native and non-native range should be classed as areas where the species could disperse to (Acevedo *et al.* 2012). For the non-native range, we ran a cellular automata model of common waxbill dispersal from Sullivan *et al.* (2012) on a 10’ grid to calculate the probability that each grid-cell had been dispersed to at each time point. We classed grid-cells with a dispersal probability of ≥0.25 as potentially available. We considered the native range to be continental Africa south of 22°N (using 15°N instead had a negligible effect on results).

5.2.2 Climate data

Climate data at 10’ resolution for the Iberian Peninsula was obtained from the CRU TS1.2 (Mitchell *et al.* 2004), averaged over the ten-year period prior to each time point, with the current time point defined as 2000 (approximately the midpoint of atlas survey work). African climate data (averaged over the period from 1960 to 1990) was extracted from CRU CL2.0 10’resolution gridded climatology (New *et al.* 2002). Seven bioclimatic variables that could plausibly limit the distribution of the species were calculated.). These were mean temperature in the coldest and warmest months, degree days above 15°C, annual precipitation, coefficient in variation in monthly precipitation, number of months were potential evapotranspiration exceeds precipitation, and annual balance between precipitation and potential evapotranspiration.

5.2.3 Quantifying niche change

We followed the framework presented by Broennimann *et al.* (2012) to quantify niche overlap and test for niche shifts. We collapsed variation in climate variables into two dimensional space using principal components analysis (PCA). This was conducted using climate in both the native and non-native range. Kernel density functions were used to map the smoothed density of occupied and available climate space, with the former representing the species realised niche. We used Schoener's D (Schoener 1970; Broennimann *et al.* 2012) to measure the overlap of realised niches in climate space. We used the niche similarity test from Broennimann *et al.* (2012) to test whether the non-native realised niche was more or less similar to the native niche than expected if common waxbills were randomly selecting available climate. To do this, the position of the non-native niche was randomly moved, and the overlap with the native niche was calculated using D. This was repeated 100 times, giving a null distribution of D. Climate availability in the non-native range was accounted for by weighting the selection of niche position by the density of non-native climate. Parts of the niche falling outside non-native climate space were reallocated within it, so that niche size was the same as the observed niche and niche shape was as close as possible given available climate space.

The realised niche in the non-native range can be divided into three components, representing expansion of the native fundamental niche, expansion of the native realised niche and overlap with the native niche (Figure 5.1). Although only expansion of the fundamental niche is typically defined as a niche shift (Petitpierre *et al.* 2012), expansion of the realised niche is important to quantify as it can lead to changes in the species-environment relationship over time. Fundamental niche expansion was calculated as the proportion of the non-native niche that was outside the native niche but within analogous climate, realised niche expansion was the proportion of the non-native niche that was outside the native niche and outside analogous climate, with the remainder of the non-native niche overlapping with the native niche. These calculations require the selection of an arbitrary threshold for defining the extent of niches and analogous climate. We used the 90% percentile as a threshold; repeating analyses with a different threshold (75% or 95%) did not change the temporal changes seen in our results. Niche unfilling was calculated as the proportion of the available native niche in the non-native range that was outside the native niche (Figure 5.1). Niche unfilling was calculated both using the whole Iberian Peninsula and restricting calculations to areas that could be dispersed to by common waxbills.

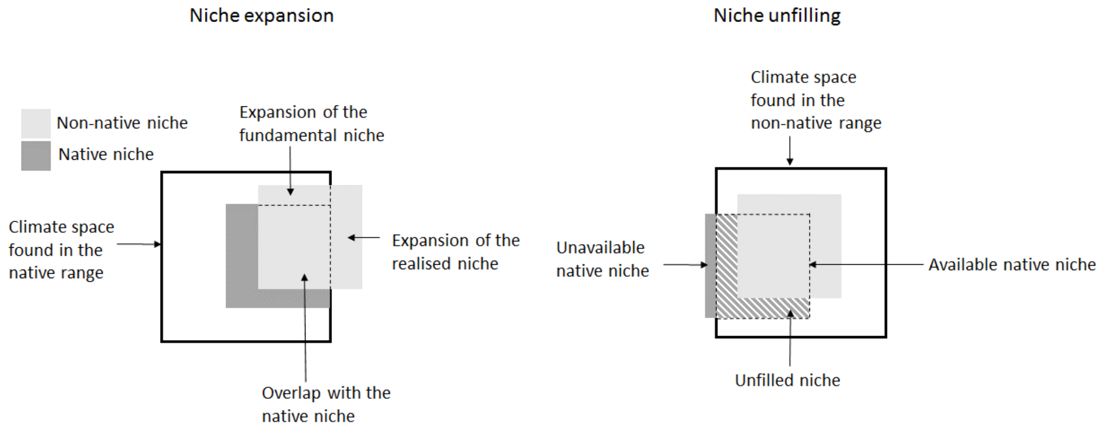


Figure 5.1. Niche expansion and niche unfilling in environmental space. The native realised niche is shown in dark grey, and the non-native realised niche is shown in light grey. The non-native realised niche can be divided into sections that overlap with the native realised niche, areas that represent expansion of the realised niche and areas that represent expansion of the fundamental niche. In a similar fashion, the native realised niche can be divided into areas that are unavailable in the native range and areas that are available. This available native niche can be further divided into areas that overlap with the non-native niche, and areas that are unfilled (shown by hatching).

The similarity of the climate of a grid-cell to a species climatic niche can be calculated by mapping the grid-cell in PCA space based on its climate, then extracting the smoothed niche density at that point. Full methods for calculating this are given in Chapter four. We did this to calculate the similarity of climatic conditions in newly colonised grid-cells to the climatic niche in the previous time step. We used t-tests to test whether these differences were different for 1984 to 1994 than 1994 to the current distribution.

5.2.4 Impact of climate change

We investigated whether climate change caused the non-native niche to become more similar to the native niche. To do this, we calculated niche overlap (using D) between the native range and non-native range using the current distribution of the common waxbill but with the climate in each time period.

5.3 Results

Overlap between native and non-native precipitation and temperature niches increased with time (Figure 5.2). This does not necessarily indicate that common waxbills increasingly selected areas with similar climate to their native niche, as overlap between the native niche and random niches generated by the niche similarity test also increased (Figure 5.2). Overlap between the native and non-native precipitation niches was always greater than the median overlap of random niches (Figure 5.2a), indicating niche conservatism, with significantly greater overlap than expected in 1994 ($P=0.02$). Overlap between native and non-native temperature niches tracked the expected values from the niche similarity test (Figure 5.2b), indicating random selection of available climate space.

While most of the common waxbill's non-native precipitation niche overlapped with their native niche, they did expand their realised niche in all time periods (Figure 5.3a). Expansion of the realised niche was most pronounced between 1984 and 1994, when the proportion of the non-native niche representing realised niche expansion increased from 3.8% to 12.9%. There was a small amount of expansion of the fundamental niche in 1994, with 0.5% of the non-native niche representing fundamental niche expansion (Figure 5.3a). In contrast, the non-native temperature niche mostly represented expansion of the realised niche, although this decreased from 92.9% in 1984 to 84.4% for the current distribution (Figure 5.3b). There was insufficient analogous climate space to calculate temperature niche unfilling. Precipitation niche unfilling increased if dispersal limitation was accounted for (Figure 5.4a) but decreased if it was not accounted for (Figure 5.4b), indicating that while the native precipitation niche has been gradually filled over time, common waxbills are increasingly failing to colonise available native climate space.

The climatic similarity of grid-cells colonised in the current time step and grid-cells colonised in the 1994 time-step was greater than the similarity of grid-cells colonised in the 1994 time-step and the 1984 time-step, given areas available to colonise (Precipitation: Δ1994-1984 = -0.427±0.003, Δ current-1994 = -0.026±0.001, $t_{1149}=155.9$, $P<0.001$; Temperature: Δ1994-1984 = -0.063±0.001, Δ current-1994 = 0.016±0.001, $t_{1991}=55.6$, $P<0.001$). This indicates that the climatic niche in newly colonised areas changed less between the 1994 and current time-steps than the 1984 and 1994 time-steps.

Climate change led to temperature conditions in the Iberian Peninsula becoming more similar to the native niche, although niche overlap remained low (Figure 5.2b). No such directional change occurred for precipitation conditions (Figure 5.2a).

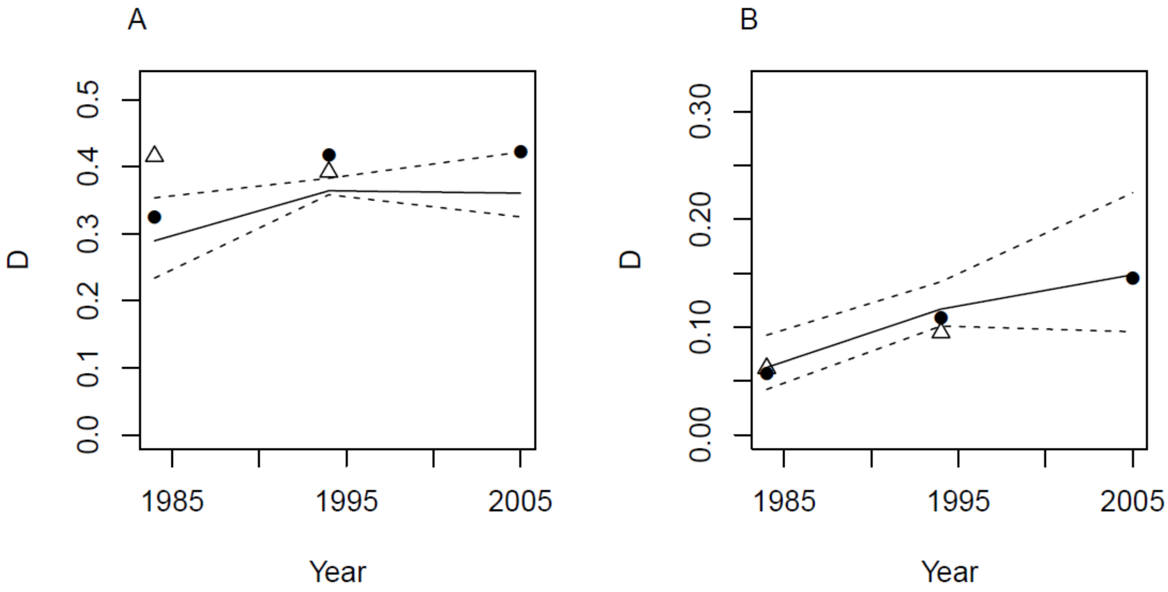


Figure 5.2. Change in overlap between the native and non-native (A) precipitation and (B) temperature niches of the common waxbill. Observed overlap values are marked with filled circles, while the median and interquartile range of overlap values from the niche similarity test null model are shown by the solid and dashed lines respectively. Unfilled triangles show overlap values that would be observed with each time period's climate and the current distribution. Note that scales on the y-axis differ between figures.

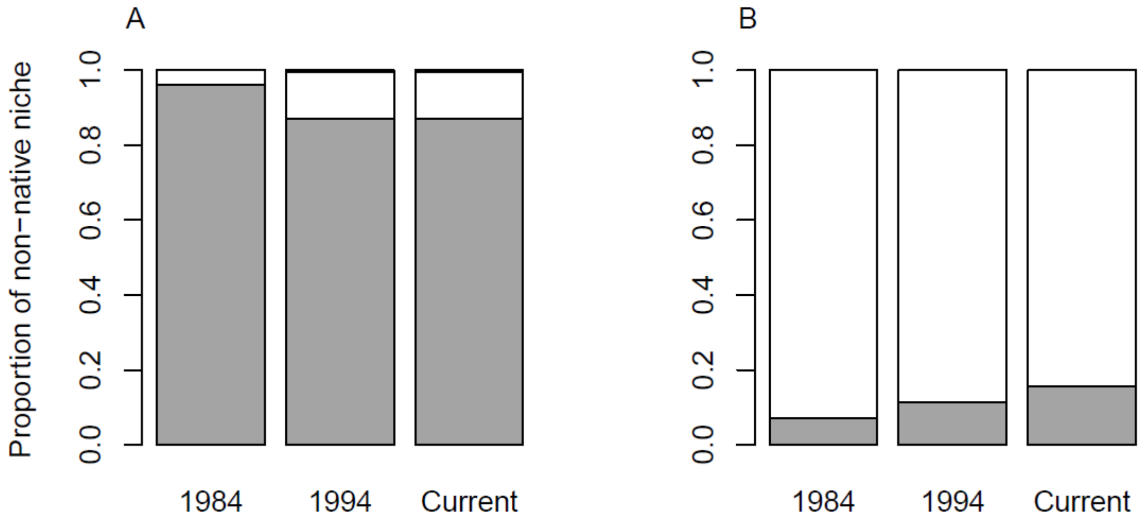


Figure 5.3. Proportion of the non-native niche that overlaps with the native niche (grey bars), represents expansion of the realised niche (white bars) and represents expansion of the fundamental niche (black bars), calculated for (A) precipitation and (B) temperature niches.

Figure 5.4. (A) Proportion of the native precipitation niche that is occupied (grey bars) and unoccupied (white bars) in analogous climate space in the non-native range. The latter represents niche unfilling. (B) As A, but calculated without accounting for dispersal limitation. Niche unfilling could not be calculated for temperature due to insufficient analogous climate space.

5.4 Discussion

The realised niche of the common waxbill broadened during its range expansion. Common waxbills showed very limited expansion of their fundamental niche, and either randomly selected climates or selected areas with similar conditions to their native niche, supporting previous studies that found niche conservatism to be common in non-native birds (Strubbe *et al.* 2013). Observed niche overlap tracked expected niche overlap, indicating that to some extent niche changes were due to exposure to new environmental conditions during range expansion. Changing climate also resulted in increased overlap between the native and non-native niche, with temperature conditions becoming more similar to the common waxbill's native niche.

The climatic niche in newly colonised areas was more similar between the second and third time-steps than the first and second time-steps. This supports the founder takes all process (Waters, Fraser & Hewitt 2013), where populations in the range margins are descendants of a small number of dispersing individuals. Further support for the founder takes all process comes from increasing niche unfilling (after accounting for dispersal limitation) during range expansion, which would be expected if populations in the range margins had a smaller fundamental niche than populations in the range core. While changes in the climatic niche are consistent with the founder takes all process, population genetic studies are needed to give stronger evidence for the operation of specific range expansion mechanisms (e.g. Garroway *et al.* 2011). There was not much support for adaptation at range margins, as only very limited expansion of the fundamental niche occurred. However, there was limited scope for fundamental niche expansion to be demonstrated in this study system, as there was high spatial congruence between climate space analogous to available climates in the native range and climate space analogous to the species' niche in the Iberian Peninsula (Sullivan and Franco, submitted).

These results have implications for the ability of SDMs to predict the potential distribution of non-native species. The general pattern for the realised niche to expand through time indicates that SDMs constructed during the early stages of the invasion process are likely to underestimate the potential distribution of the species. This has been observed in both real and simulated invasions (Vaclavick et al. 2012), and is likely to result from SDMs wrongly identifying suitable areas the species has yet to spread to as unsuitable (Sullivan et al. 2012), as well as genuine expansion of the realised niche. In contrast, the founder-takes-all mechanism could lead to native and early non-native distribution trained SDMs overpredicting the potential distribution of the non-native species, as the fundamental niche of populations at the range margins is a subset of the fundamental niche of native and range core populations.

The inferences on the environmental limits to the common waxbills distribution from this study compliment previous work on the species. Common waxbills tended to select areas with similar climatic conditions to their native niche, although they did expand their realised niche when exposed to new climatic conditions. While the Iberian Peninsula, with the exception of the north-west, was within the native precipitation niche and predicted to be suitable by native trained SDMs (Stiels *et al.* 2011; Sullivan and Franco, submitted), SDMs trained on the non-native distribution predicted that only western and coastal areas had suitable climatic conditions (Sullivan *et al.* 2012). These non-native trained SDMs, which accounted for dispersal limitation, identified that common waxbills were positively related to warmer annual, winter and summer temperatures, annual precipitation and negatively related to continentally (Sullivan *et al.* 2012). We found some unfilling of the common waxbill's native climatic niche in the Iberian Peninsula. This was not due to dispersal limitation, as we had accounted for this in our calculation, and is unlikely to be due to competition with native species as common waxbills occupy a peripheral niche (Batalha, Ramos & Cardoso 2013).

While the fundamental niche of the common waxbill has been largely conserved during range expansion, the realised niche has changed. These changes are likely to have resulted from a combination of exposure to new environmental conditions, density dependent environmental associations and the founder-takes-all process. Together, these mechanisms have the potential to cause native and non-native trained SDMs to over and under predict the potential distribution of common waxbills. It is therefore important to recognise that the realised niche is dynamic, and update SDMs with new data as non-native species expand.

Acknowledgements

We thank Birdlife International for providing species range maps. MJPS was funded by a NERC PhD studentship.

References

Acevedo, P., Jiménez-Valverde, A., Lobo, J.M. & Real, R. (2012) Delimiting the geographical background in species distribution modelling. *Journal of Biogeography,* **39,** 1383-1390.

Batalha, H.R., Ramos, J.A. & Cardoso, G.C. (2013) A successful avian invasion occupies a marginal ecological niche. *Acta Oecologica,* **49,** 92-98.

BirdLife International & NatureServe (2011) Bird species distribution maps of the world. (ed. BirdLife International). Cambridge, UK and NatureServe, Arlington, USA.

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.J., Randin, C., Zimmermann, N.E., Graham, C.H. & Guisan, A. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography,* **21,** 481-497.

Broennimann, O., Treier, U.A., Muller-Scharer, H., Thuiller, W., Peterson, A.T. & Guisan, A. (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters,* **10,** 701-709.

Brown, J.H. (1984) On the Relationship between Abundance and Distribution of Species. *The American Naturalist,* **124,** 255-279.

Buckley, J., Butlin, R.K. & Bridle, J.R. (2012) Evidence for evolutionary change associated with the recent range expansion of the British butterfly, *Aricia agestis*, in response to climate change. *Molecular Ecology,* **21,** 267-280.

Carrete, M., Edelaar, P., Blas, J., Serrano, D., Potti, J., Dingemanse, N.J. & Tella, J.L. (2012) Don’t neglect pre-establishment individual selection in deliberate introductions. *Trends in Ecology & Evolution,* **27,** 67-68.

Cassey, P., Blackburn, T.M., Sol, D., Duncan, R.P. & Lockwood, J.L. (2004) Global patterns of introduction effort and establishment success in birds. *Proceedings of the Royal Society of London. Series B: Biological Sciences,* **271,** S405-S408.

Dietz, H. & Edwards, P.J. (2006) Recognition that causal processes change during plant invasion helpds explain conflicts in evidence. *Ecology,* **87,** 1359-1367.

Equipa Atlas (2008) *Atlas das Aves Nidificantes em Portugal (1999-2005)*. Instituto da Conservacao da Natureza e da Biodiversividade, Sociedade Portuguesa para o Estudo das Aves, Parque Natural da Madeira e Secretaria Regional do Ambiente e do Mar, Lisbon.

Gallagher, R.V., Beaumont, L.J., Hughes, L. & Leishman, M.R. (2010) Evidence for climatic niche and biome shifts between native and novel ranges in plant species introduced to Australia. *Journal of Ecology,* **98,** 790-799.

Garroway, C.J., Bowman, J., Holloway, G.L., Malcolm, J.R. & Wilson, P.J. (2011) The genetic signature of rapid range expansion by flying squirrels in response to contemporary climate warming. *Global Change Biology,* **17,** 1760-1769.

Hewitt, G. (2000) The genetic legacy of the Quaternary ice ages. *Nature,* **405,** 907-913.

Lavergne, S. & Molofsky, J. (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences,* **104,** 3883-3888.

Levine, J.M. & D'Antonio, C.M. (2003) Forecasting biological invasions with increasing international trade. *Conservation Biology,* **17,** 322-326.

Marti, R. & de Moral, J.C. (2003) Atlas de las Aves Reproductoras de Espana. Dirección General de Conservación de la Naturaleza-Sociedad Española de Ornithología, Madrid.

Mitchell, T.D., Carter, T.R., Jones, P.D., Hulme, M. & New, M. (2004) A comprehensive set of high-resoultion grids of monthly climate for Europe and the globe: the observed record (1901-2000) and 16 scenarios (2001-2100). *Working Paper*. Tyndall Centre for Climate Change Research, Norwich, UK.

Morris, D.W. (1987) Tests of Density-Dependent Habitat Selection in a Patchy Environment. *Ecological Monographs,* **57,** 270-281.

New, M., Lister, D., Hulme, M. & Makin, I. (2002) A high-resolution data set of surface climate over global land areas. *Climate Research,* **21,** 1-25.

Pearman, P.B., Guisan, A., Broennimann, O. & Randin, C.F. (2008) Niche dynamics in space and time. *Trends in Ecology & Evolution,* **23,** 149-158.

Peterson, A.T. (2003) Predicting the geography of species' invasions via ecological niche modelling. *Quarterly Review of Biology,* **78,** 419-433.

Petitpierre, B., Kueffer, C., Broennimann, O., Randin, C., Daehler, C. & Guisan, A. (2012) Climatic Niche Shifts Are Rare Among Terrestrial Plant Invaders. *Science,* **335,** 1344-1348.

Schoener, T.W. (1970) Nonsynchronous Spatial Overlap of Lizards in Patchy Habitats. *Ecology,* **51,** 408-418.

Stiels, D., Schidelko, K., Engler, J., van den Elzen, R. & Rödder, D. (2011) Predicting the potential distribution of the invasive Common Waxbill *Estrilda astrild* (Passeriformes: Estrildidae). *Journal of Ornithology***, 152**, 769-780.

Strubbe, D., Broennimann, O., Chiron, F. & Matthysen, E. (2013) Niche conservatism in non-native birds in Europe: niche unfilling rather than niche expansion. *Global Ecology and Biogeography***, 22**, 962–970.

Sullivan, M.J.P., Davies, R.G., Reino, L. & Franco, A.M.A. (2012) Using dispersal information to model the species–environment relationship of spreading non-native species. *Methods in Ecology and Evolution,* **3,** 870-879.

Thomas, C.D., Bodsworth, E.J., Wilson, R.J., Simmons, A.D., Davies, Z.G., Musche, M. & Conradt, L. (2001) Ecological and evolutionary processes at expanding range margins. *Nature,* **411,** 577-581.

Thuiller, W., Richardson, D.M., Pysek, P., Midgley, G.F., Hughes, G.O. & Rouget, M. (2005) Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology,* **11,** 2234-2250.

Waters, J.M., Fraser, C.I. & Hewitt, G.M. (2013) Founder takes all: density-dependent processes structure biodiversity. *Trends in Ecology & Evolution,* **28,** 78-85.

6 Non-native species avoid competition by using underexploited resources in an anthropogenic habitat

Submitted to Proceedings of the Royal Society B as Sullivan, M.J.P., Davies, R.G., Mossman, H.L. and Franco, A.M.A. Non-native species avoid competition by using underexploited resources in anthropogenic habitats.

Anthropogenic modification of habitats may reduce the resources available for native species, leading to population declines and extinction. These same habitats often have the highest richness of non-native species. This may be because human modified habitats provide novel resources that are not accessible to native species but can be exploited by non-native species. We conduct a large-scale test of this hypothesis by comparing the functional diversity and resource use of native and non-native bird communities in a modified habitat (rice fields) and in less modified habitats in the Iberian Peninsula. The functional diversity of native bird communities was lower in rice fields, but non-native birds plugged this gap. Differences in resource use between native and non-native species allowed non-native species to exploit resources that were plentiful in rice fields. Human modified habitats can therefore facilitate the colonisation and expansion of non-native species by providing resources that are not fully exploited by native species.

6.1 Introduction

One of the most fundamental questions in ecology is the extent to which communities are saturated with species (Cornell & Lawton 1992; Loreau 2000). The transport and introduction of non-native species to new areas has provided a natural experiment to investigate this (Crawley 2005). While the ability of many non-native species to establish in communities without a corresponding extinction of native species (Stohlgren *et al.* 2008) demonstrates that communities can often accept more species, competition for resources with native species can influence their ability to become established (Levine, Adler & Yelenik 2004). For example, non-native plants that are distantly related to native species, and therefore likely to have different resource requirements, are more likely to establish successfully (Strauss, Webb & Salamin 2006). Diverse communities may be harder to invade as they use more resources

(Fridley *et al.* 2007), as well as by having more species being more likely to contain superior competitors (Fargione & Tilman 2005). However, observational studies have often found that diverse communities have high non-native species richness, perhaps because they occur where there are more resources (Fridley *et al.* 2007).

Experimental work has linked community invasability to the amount of resources that are unexploited by competitors (Tilman 2004). These can occur when resource pulses temporarily provide surplus resources (Tilman 2004) or when high spatial heterogeneity limits the ability of species to exploit all available resources (Fridley *et al.* 2007). Changes to the state of a system, such as alteration to fire regimes, can also lead to unexploited resources (Moles *et al.* 2012). Such changes act as an environmental filter (Helmus *et al.* 2010), with only a portion of species found in the original community able to persist in the modified one. These local extinctions are often phylogenetically clustered (Helmus *et al.* 2010), so phylogentically conserved functional traits are likely to be lost from the new community (Díaz *et al.* 2013), while loss of any species in communities with low functional redundancy will lead to functional diversity loss (Petchey *et al.* 2007). This loss of functional diversity, and hence diversity of resource use, may reduce the likelihood of species being able to utilise novel resources provided by the altered system. Small-scale experiments have shown that such losses of functional diversity can lower the resistance of communities to invasion (Vaz-Pinto *et al.* 2013), however, it is unknown whether this happens at large spatial scales and in natural ecosystems. The presence of unexploited resources may be brief, as species in the regional species pool able to exploit these novel resources could colonise the community (Kirmer *et al.* 2008). However, if such species are absent from the regional species pool, then unexploited resources may be present for longer. We expect this to be likely when new resources provided by the new system have few local analogues. Human modified landscapes potentially provide an example of this. Conversion of natural habitats to agriculture can dramatically change the resources available (Siriwardena *et al.* 2000) leading to the non-random loss of species (Fritz, Bininda-Emonds & Purvis 2009). While some agricultural habitats, for example wood-pasture, contain functionally similar elements to those in natural habitats (Joffre, Rambal & Ratte 1999), many have no local analogues. However, the ability of human-modified landscapes to provide resources that are underexploited by competitors, and for this to promote invasion, is uncertain.

Human modified landscapes often have high non-native species richness (McKinney 2006). However, it is uncertain whether the presence of unexploited resources is driving this pattern. The number of non-native species established in a given location depends on the number of species that can pass through sufficient stages in the invasion pathway (Blackburn *et al.* 2011) to be recorded as having self-sustaining populations there, so a number of factors aside from community invasability affects non-native species richness. Species first have to be transported to new locations, so areas close to ports and human settlements typically have higher non-native species richness due to high propagule pressure (Cassey *et al.* 2004). The colinearity of propagule pressure and degree of human modification makes disentangling the relative importance of each challenging. Once transported to a new location, the environmental conditions there have to be sufficiently favourable for the non-native species to allow it to survive and reproduce (Blackburn, Lockwood & Cassey 2009), so human-modified landscapes could have more non-native species if they are more climatically suitable for non-native species.

Non-native birds in the Iberian Peninsula are positively associated with rice fields (Matias 2002) at both coarse (Sullivan *et al.* 2012) and fine (Costa, Elias & Farinha 1997) scales. While the coarse scale relationship could be confounded by factors such as climate and propagule pressure, these do not account for fine scale associations. Rice fields are a fairly recent land-use in the Iberian Peninsula; despite some localised cultivation prior to the 20th century, widespread cultivation has only occurred since the 1930s (Lima 1997). Unlike traditional agricultural systems, which comprise of a mosaic of open habitats and wooded features, rice fields comprise open fields that are seasonally flooded, crossed by ditches with wetland vegetation. This combination of open areas and linear wetlands potentially provides a novel resource environment. We compare bird communities in rice fields and in more traditional land-uses containing habitat elements found in rice fields. We investigate whether rice fields provide resources that are not fully exploited by native species, and whether this has facilitated avian invasions in the Iberian Peninsula, providing a large-scale test of the hypothesis that unexploited resources drive invasions. Specifically, we test the hypotheses that (1) resources in rice fields are not fully exploited by native species and (2) differences in resource use for feeding and shelter allow non-native species to exploit these resources.

6.2 Methods

6.2.1 Field survey

Surveys were conducted in Portugal and western Spain in the breeding seasons (April to June) of 2011 and 2012. We surveyed bird communities in rice fields and adjacent agriculture (arable and heterogenous mixtures of fruit crops and arable), rivers and natural wetlands. These adjacent open habitats represent more traditional land-uses and collectively contain the various habitat elements found in rice fields. We selected 51 areas of rice fields, as well as 10 areas containing open habitats but lacking rice fields (Figure 6.1a). At each site we carried out up to 16 point counts (mean = 7.48, sd = 3.26, range = 1 – 16 , but with only one site with 1 point count). These were located >200m apart, sampling rice fields and adjacent open habitats (see Figure 6.1a for an example). In 2011, sites were located across Portugal and western Spain, while in 2012 sites were located in the Tagus and Sado valleys of Portugal. Sites did not overlap, with the exception of two sites that were re-surveyed in 2012 to take advantage of improved access to the rice fields (point counts were never in the same location). Point counts lasted five minutes, and the presence or absence of seed-eating bird species within a 100m radius of the observer was recorded. We recorded the presence of resources (trees and shrubs, emergent vegetation, rough (i.e. ungrazed) grass and forbs) at 30m intervals on a grid stretching 90m in each direction from the point count location (i.e. 49 resource recording points per point count, see Figure 6.1b for schematic). The same observer performed all point counts, including recording resource availability.

Resource selection by target species was quantified in a separate protocol in focal watches at 68 locations. These focal watches were never carried out concurrently with point counts. Resource availability (using the same resource codes as above) was recorded in a 180m by 180m square, crossed midway by a transect that the observer was allowed to move up and down. This square was divided into 30m by 30m sub-squares, and the percentage cover of each resource in these was estimated (see Figure 6.1c for schematic). By recording the spatial configuration of resources in this way, we could adjust availability to account for the decline in detectability of birds with distance from observer (see Appendix S1 for details). Resource use by target species was recorded in scan samples performed every ten minutes. During each scan sample the resource use and activity (feeding or shelter) of each group of target species was recorded. We used Jacobs index (Jacobs 1974) to quantify the selection of each resource for feeding and shelter given the availability of each habitat. Jacobs index was calculated as $J = (u - a)/(u + a - 2ua)$, where u is utilisation and a is availability. By accounting for resource availability we ensure that our calculations of microhabitat selection are not influenced by the availability of resources in different areas. Shelter resources primarily are resources used for resting between intervals of feeding, but also encompasses resources used for nesting and roosting, and has the potential to influence the ability of birds to access food resources (Whittingham & Evans 2004).

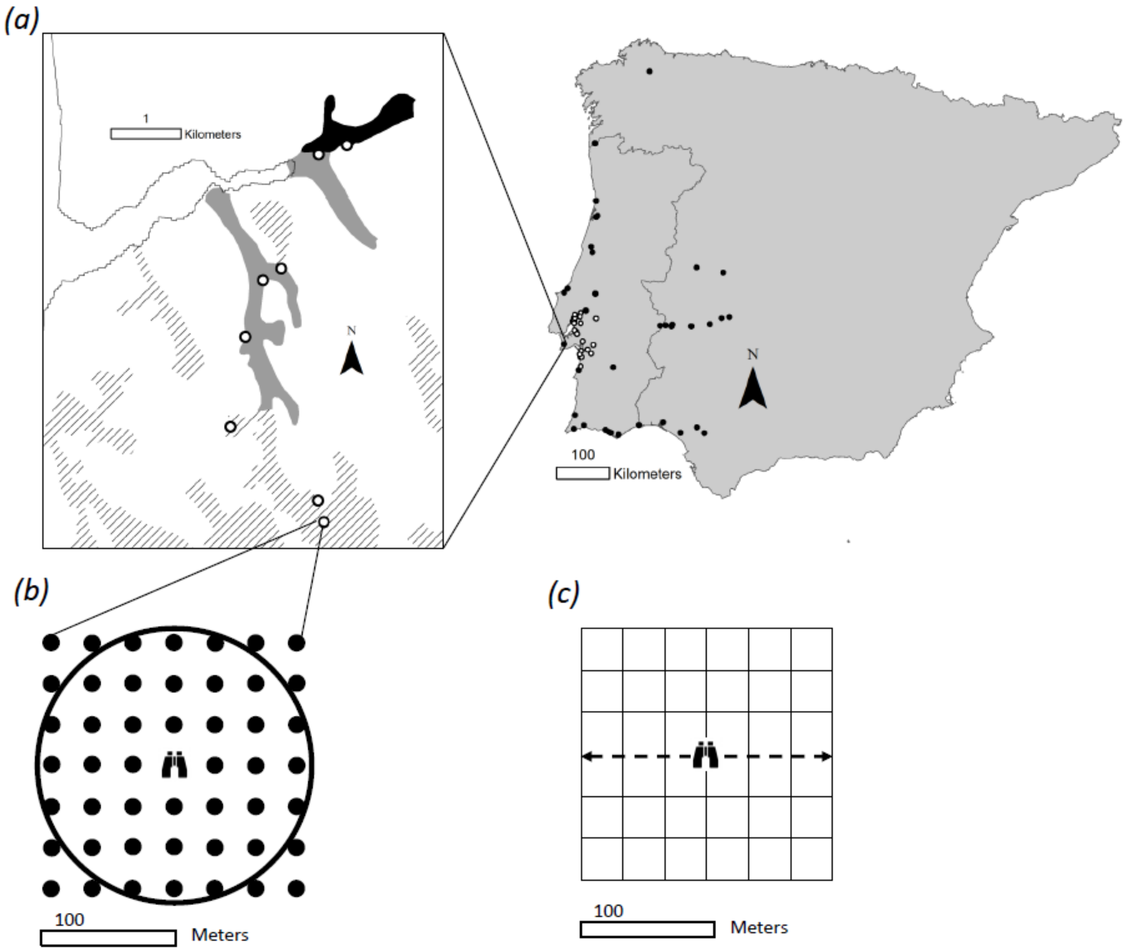


Figure 6.1. (a) Location of survey sites in the Iberian Peninsula. The centroids of each site are plotted. Sites surveyed in 2011 are shown by filled circles, and sites surveyed in 2012 are shown by open circles. The insert map shows the location of point counts at one site. Point count locations are shown by open circles. Rice fields are shaded grey, wetlands shaded black, and heterogeneous agriculture (Corine land-cover level two class 24) shown by hashing. The remaining area is largely forestry. (b) Schematic of sampling protocol at each point count. The observer (position shown by binoculars) records birds seen within a 100m radius (shown by circle). Resources are recorded at regularly spaced points (shown by filled circles, habitat also recorded at position of observer). (c) Schematic of sampling protocol at focal watch locations. The observer walks along a central transect (dashed arrow), and records birds and percentage cover of resources in each sub-square.

We compared the vegetation of rice field margins and adjacent grasslands by surveying vegetation in 0.25m^2 quadrats randomly located in rice field margins and adjacent grasslands. In each quadrat we counted the number of grass and forb species, and measured vegetation height at five points. We measured the weight holding capacity of seed bearing plants by attaching weights to the base of the seed head, and recording how many weights were needed to cause the plant to droop to the ground. We obtained data on mass of bird species from (Cramp & Perrins 1994b; Cramp & Perrins 1994a; Barnard 1997; del Hoyo, Elliot & Christie 2009; del Hoyo, Elliot & Christie 2010) to assess the ability of native and non-native species to access these resources.

It is important to note that we compared sampling units (point counts or quadrats) in rice fields to sampling units in other habitats. Heterogeneity across all sampling units is likely to be higher in other habitats than in rice fields simply because the former category embraces a larger number of habitat types, however, differences in within sampling unit heterogeneity result from genuine differences in the resource composition of habitats.

6.2.2 Functional diversity metrics

We obtained species trait data from published literature, supplemented by field observations (details and data in Appendix S2), to construct a trait matrix containing all native and established non-native species belonging to the seed-eating guild present in Portugal and western Spain. Traits that related to a species' use of resources for feeding and nesting were selected (Table 1). Gower distance was used to convert the trait matrix to a distance matrix, as it can handle a mix of continuous, ordinal and categorical data (Gower 1971). We calculated the functional diversity (FD) of communities following Petchey and Gaston (2006); we used the distance matrix to generate a dendrogram containing all seed-eating bird species, and calculated the FD of the bird community at each point count by dividing the total branch lengths of a dendrogram containing all species in the community by the total branch lengths of a dendrogram containing all target species. We used average linkage to generate a dendrogram as, compared to single and complete linkage, it gave the dendrogram with the highest cophenetic correlation with the distance matrix (c=0.735). We compared observed values of FD with values obtained by running null models where the species composition of communities was changed while maintaining species richness (Petchey & Gaston 2006). Following Mendez *et al.* (2012) we used two null models, one randomly selecting species from the entire seed-eating guild, and one maintaining observed frequencies of species. Null models were run 1000 times. The standardised effect size of functional dispersion was calculated by subtracting the mean FD of null models from observed FD, then dividing this by the standard deviation of FD from null models. Values of greater than one indicate functional over-dispersion, which has been suggested to result from competitive exclusion of similar species and values less than one indicate functional under-dispersion, which can indicate habitat filtering and greater species packing of functionally similar species (Petchey *et al.* 2007). Communities were more under-dispersed when a random null model was used, possibly as a result of broad-scale habitat filtering, but otherwise results were consistent between null models. Therefore only results from the null model maintaining species frequency (FD_{freq}) have been presented.

Table 6.1. Traits used to calculate functional diversity metrics.

Trait	Categories	Data type
Adult diet	Seeds, Green plants, Invertebrates	Ordinal
Nestling diet	Seeds, Green plants, Invertebrates	Ordinal
Feeding agility	Upside down, Vertical stem, Bent stem	Ordinal
Feeding height	Ground, Herb layer, Tree layer	Ordinal
Feeding habitat	Weeds, Cultivated, Trees	Ordinal
Food plants	Grasses, Composite	Ordinal
Morphology	Culmen length, Bill length-depth ratio, Tarsus length, Wing length, Body mass	Continuous
Nest location	Ground, dry low vegetation, wet low vegetation, Tree, Hole nester, Cliff nester	Ordinal
Nest height		Ordinal
Nesting season	January, February, March, April, May, June, July, August, September, October, November, December	Ordinal

6.2.3 *Quantifying habitat associations*

We compared the habitat associations of native species with the habitat associations of non-native species in their native range. To do this, we quantified the strength of a species' association along two environmental gradients (wet to dry habitats and open to closed habitats). We followed the approach of Dolman *et al*. (2012) to quantitatively code habitat associations from qualitative descriptions in literature. The gradient from wet to dry habitats was divided into four categories (extensive wetland, linear or fragmented wetland, damp habitats and dry habitats) and the gradient from open to closed habitats was divided into six categories (low growing vegetation, matrix of low growing vegetation and taller non-woody vegetation, extensive taller non-woody vegetation, matrix of woody and non-woody vegetation dominated by the latter, matrix of woody and non-woody vegetation dominated by the former, and forests and woodland). These categories were chosen to adequately describe the variation in habitat use along these gradients. We mapped habitat types onto these gradients based on the above definitions. The positions of these habitat types on the gradients is given in Appendix S3. We consulted literature references on species' habitat associations (Cramp & Perrins 1994b; Cramp & Perrins 1994a; Barnard 1997; del Hoyo, Elliot & Christie 2009; del Hoyo, Elliot & Christie 2010), and converted these qualitative descriptions into quantitative scores for association strength, where zero corresponded to no association with that habitat type, +1 to qualitative descriptions indicating the species was weakly associated with the habitat, and +2 to strongly associated. These habitat association scores were converted into an overall score for each position on the gradients by taking the maximum score across all habitats at that point. All coding was performed by one person to ensure consistent interpretation of sources. Scores for each species are given in Appendix S3. We compared mean habitat associations of native and non-native species to identify differences in habitat associations across these gradients.

6.2.4 Data analysis

We tested all models involving spatial data for residual spatial autocorrelation using Moran's I tests. In all cases significant residual spatial autocorrelation was found. We addressed this by fitting mixed effects models with site as a random effect. This assumes that observations within the same site are correlated while observations in different sites are independent, and was sufficient to account for residual spatial autocorrelation (Moran's I test *P*-values >0.05) for all models.

All statistical analyses were performed in R (R Development Core Team 2012), using the library lme4 (Bates, Maechler & Dai 2008) to fit mixed effects models. The significance of model terms was assessed using likelihood-ratio tests, where a model containing a given term was compared to a nested model lacking that term.

6.3 Results

17 species (four non-native and 13 native) of seed-eating birds were recorded. All four non-native species were found in a higher proportion of rice field point counts than point counts in adjacent open habitat; the same was true for only three native species (Table 6.2). The non-native common waxbill was the second most widely recorded species, being present in 176 of 456 point counts. House sparrow, goldfinch and serin were the most widely recorded native species, all being present in > 160 point counts (Table 6.2).

Table 6.2**.** Study species recorded, and number of sites and point counts present.

Species	Scientific name	Species code	Number of sites present	Number of point counts in rice fields present	Number of points counts in other open habitats present
Black-headed weaver	*Ploceous melanocephalus*	BH	8	11	8
Bullfinch	*Pyrrhula pyrrhula*	BF	1	0	1
Chaffinch	*Fringilla coelebs*	CH	13	6	21
Common waxbill	*Estrilda astrild*	CW	56	82	94
Corn bunting	*Emberiza calandra*	CB	35	67	75
Goldfinch	*Carduelis carduelis*	GO	54	69	96
Greenfinch	*Carduelis chloris*	GR	42	43	83
Hawfinch	*Coccothraustes coccothraustes*	HA	1	1	0
House sparrow	*Passer domesticus*	HS	60	129	151
Linnet	*Carduelis cannabina*	LI	44	38	50
Red avadavat	*Amandava amandava*	RA	11	13	8
Reed bunting	*Emberiza schoeniclus*	RB	2	1	3
Serin	*Serinus serinus*	SE	50	56	106
Siskin	*Carduelis spinus*	SK	1	0	2
Spanish sparrow	*Passer hispaniolensis*	SS	6	1	9
Tree sparrow	*Passer montnus*	TS	14	11	16
Yellow-crowned bishop	*Euplectes afer*	YB	21	54	13
Total			61	200	256

Rock sparrow *Petronia petronia* (RS), common crossbill *Loxia curvirostra* (CR), ortolan bunting *Emberiza hortulana* (OB) and yellowhammer *Emberiza citronella* (YH) were not recorded in any point counts.

Native seed-eating bird communities had lower functional diversity (χ^2 = 4.329, P = 0.038, Figure 6.2) and near significantly lower species richness (χ^2 = 2.724, P = 0.099) in rice fields than in adjacent open habitats. Non-native species plugged this gap, as the functional diversity of the combined native and non-native seed-eating bird community was not significantly different between rice fields and adjacent open habitats (χ^2 = 0.011, P = 0.918, Figure 6.2), with higher non-native species richness in rice fields (χ^2 = 9.587, P = 0.002). Native bird communities were functionally random (FD_{freq} = 0.02 ± 0.05, t =0.33, P = 0.745), but with lower species packing in rice fields (Δ FD_{freq} = 0.25 ± 0.11, P = 0.024, Figure 6.3a). Functional dispersion was negatively related to species richness (r = -0.13, P = 0.017), indicating that species packing increased with species richness.

Differences in the selection of shelter resources could explain this. Non-native species showed a greater preference for using emergent vegetation for shelter than native species (W=24, P=0.009, Figure 6.4a), while native species primarily used trees and bushes for shelter (Figure 6.4a). Trees and bushes were found less frequently in rice fields than adjacent open habitats (χ^2 = 37.34, P < 0.001, Figure 6.4c), while there was no significant difference in the amount of emergent vegetation (χ^2 = 0.276, P=0.599, Figure 6.4c). This indicates that rice fields contained suitable shelter resources for non-native species but not for native species.

There were also differences in the amount of feeding resources provided by rice fields and adjacent open habitats. Both native and non-native species selected forbs for feeding, while non-native species showed a greater preference for rough grass (W=18, P=0.024, Figure 6.4b). There was no significant difference between the amount of rough grass in rice fields and other sites (χ^2 = 1.525, P = 0.217) but more forbs were found in non-rice field sites (χ^2 = 9.119, P = 0.003, Figure 6.4b).

Plant species richness was lower in rice fields than adjacent grasslands, with 1.452±0.557 fewer species in rice field margins (z=-2.611, P=0.009, Figure 6.4d). Grass species richness did not significantly differ between rice field margins and adjacent grassland (z = -1.01, P=0.313, Figure 6.4d), but forb species richness was significantly lower in rice field margins (z=-2.635, P=0.008, Figure 6.4d). This indicates that grasses made up more of the available food resources in rice fields than in other open habitats. Plant weight holding capacity was 1.9g lower in rice fields than adjacent grasslands (mean weight holding capacity in rice fields = 2.6 ± 0.4g, mean weight holding capacity in other grasslands = 4.5 ± 0.3g, W =215, P=0.002). This indicates that proportionally more food resources in rice fields were only available to light bodied and/or agile species. Non-native species were 11.5 g lighter than native species (non-native body mass = 15.4 ± 3.6g, native body mass = 26.9 ± 2.6g, *W* = 64, *P* = 0.035), with two non-native species (common waxbill *Estrilda astrild* and red avadavat *Amandava amandava*) lighter than any native seed eating bird species.

Rice fields provided conditions that fell within the habitat associations of non-native species in their native range (Figure 6.3b). This indicates that non-native species are pre-adapted to exploit the conditions provided by rice fields. Rice fields occurred at the position in both gradients where there was the greatest difference between the habitat associations of native and non-native species (Figure 6.3b).

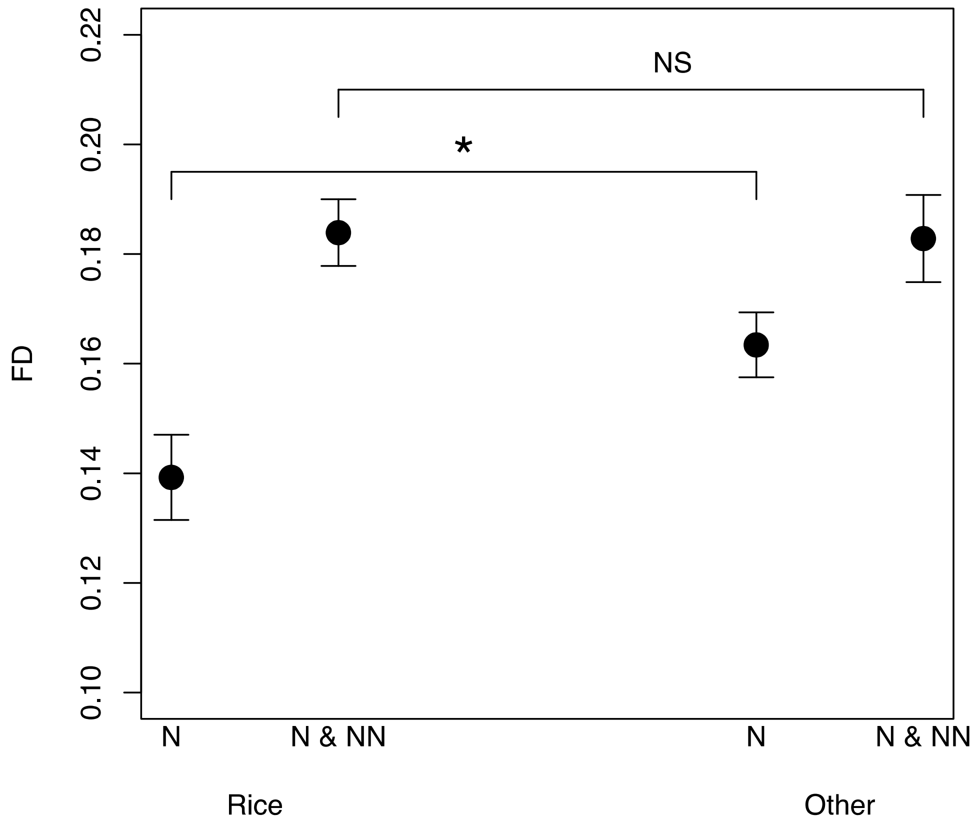


Figure 6.2. Functional diversity (FD) of native (N) and native and non-native (N & NN) seed-eating bird communities in rice fields and other open habitats. * denotes *P*<0.05, NS denotes *P*>0.05.

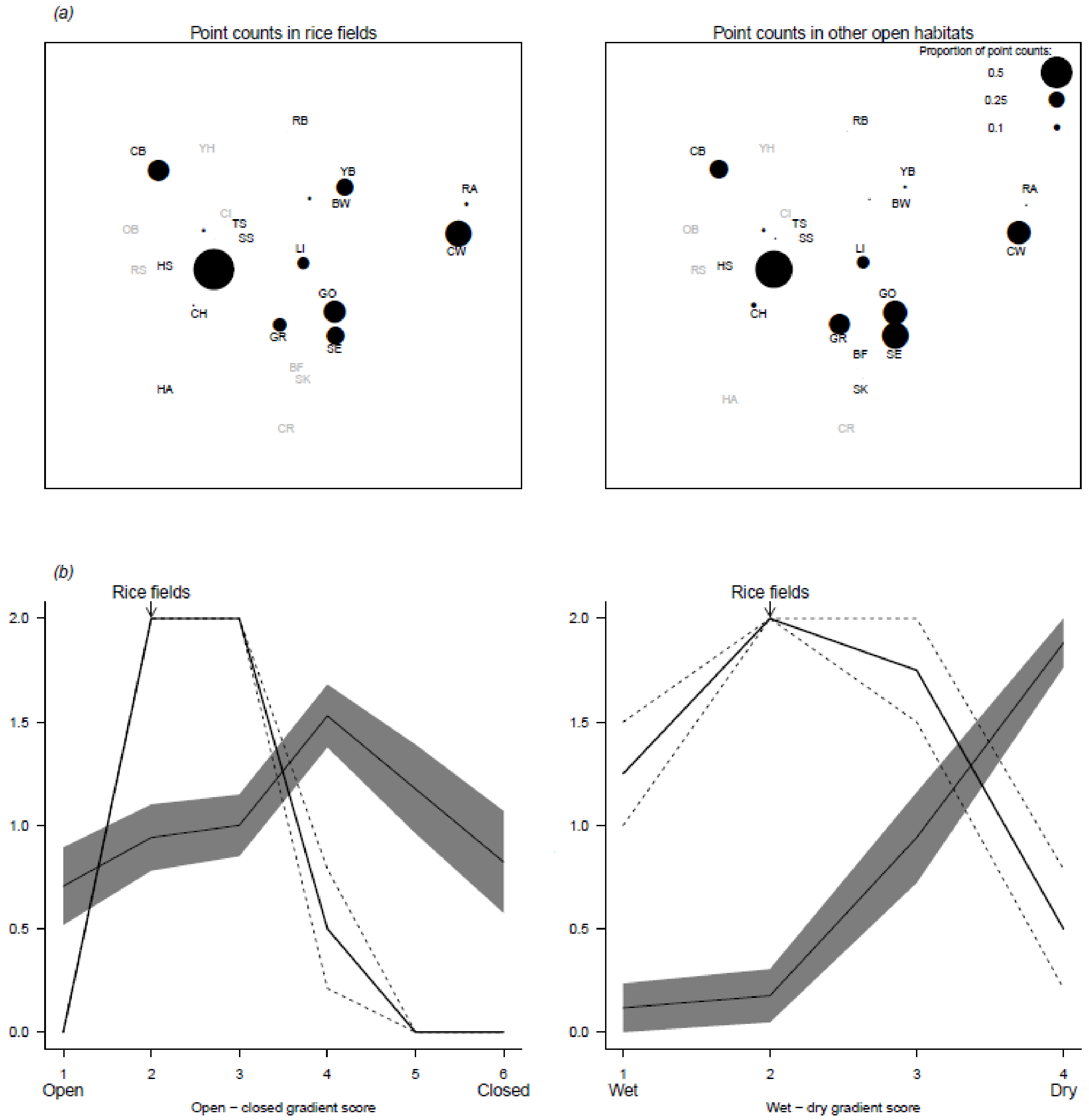


Figure 6.3. (a) Species packing in rice fields and other open habitats. The size of each circle is proportional to the number of point count locations where they were recorded. Species codes are shown next to circles. Species codes in grey denote species not recorded at any point count. (b) Habitat associations of native and non-native species along environmental gradients from open to closed habitats (left) and wet to dry habitats (right). Habitat associations of non-native species are shown by a bold line, with dashed lines showing standard errors. Habitat associations of native species are shown by a solid line, with grey shading showing standard errors Habitat associations for non-native species were based on descriptions of their habitat use in their native range.

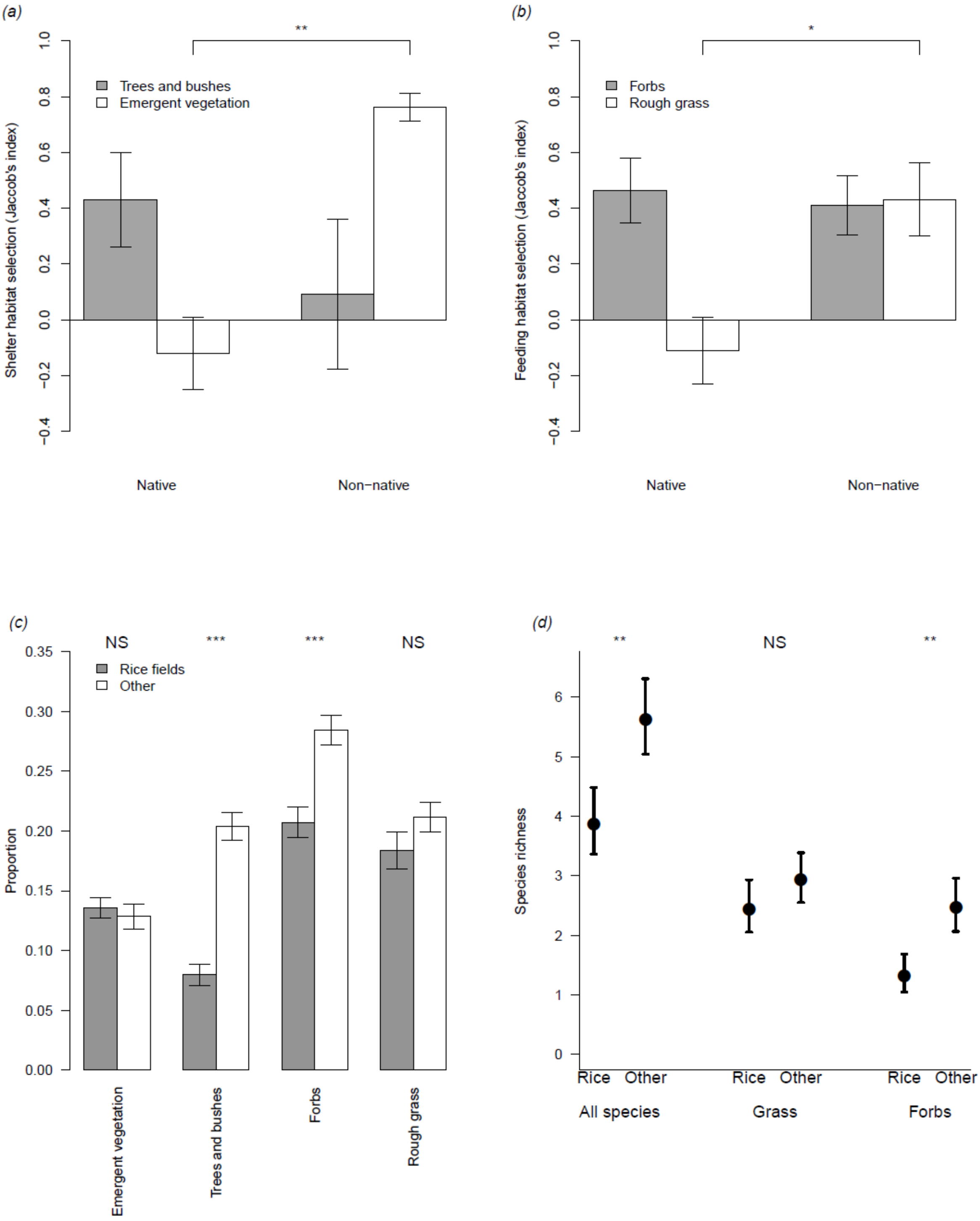


Figure 6.4**.** Selection of resources for (a) shelter and (b) feeding by native and non-native species. Positive values of Jaccob's index indicate that a habitat is selected more than expected given availability, and negative values indicate that it is selected less than expected. (c) Proportion of resource sampling points containing selected habitat types at point count locations in rice fields and other open habitats and (d) species richness of plants in rice field margins and adjacent grassland. *** denotes *P*<0.001, ** denotes *P*<0.01, * denotes P<0.05, NS denotes *P*>0.05.

6.4 Discussion

Rice fields are a recent land-use that provide fewer resources for native seed-eating birds than adjacent open habitats. This was reflected by the lower functional diversity and species packing of native bird communities in rice fields. This scarcity of resources might be expected to hinder the ability of non-native seed eating birds to colonise rice fields. However, non-native species were positively associated with rice fields, and plugged the missing functional diversity in rice fields. The ability of species that are functionally similar to native species to colonise rice fields supports the hypothesis that native species underexploit the resources provided by the novel habitat of rice fields.

Niche differences meant that non-native birds were better able to access these resources. For example, most native seed-eating bird species primarily feed on forbs (Cramp & Perrins 1994b), while non-native species are lighter and more agile so can feed extensively on grasses (Matias 2002), which dominated rice field margins. The high proportion of plant species in rice fields that were grasses supports previous studies (Weerakoon & Gunewardena 1983). Differences in plant species richness and composition between rice field margins and other grasslands could be due to the high herbicide input rice fields receive and the high salinity of soils in rice fields (Ritzema *et al.* 2008). Differences in the selection of shelter habitat was also important for allowing non-native species to exploit rice fields; native species tended to select trees and bushes for shelter, which were rare in rice fields, while non-native species were able to use the emergent vegetation lining drainage ditches in rice fields. This difference in shelter habitat selection is also reflected in breeding habitat selection, as the non-native species often nested in emergent vegetation (Matias 2002) while many of the native species nested in trees and bushes (Cramp & Perrins 1994b).

We found that species-poor native bird communities consisted of functionally dissimilar species; such communities were found in rice fields. Functional overdispersion can result from the competitive exclusion of functionally similar species (Petchey & Gaston 2006), so rice fields may lack sufficient resources to support functionally similar native species. Differences in resource use by non-native species compared to native species was therefore important in allowing them to avoid this competitive exclusion in rice fields.

While the positive association of non-native species with rice fields could be because they are better able than native species to exploit the feeding and shelter habitat provided by rice fields, it is uncertain whether their selection of rice fields was primarily driven by underexploited resources. In other words, would non-native species still be primarily associated with rice fields in the absence of competition from native species in other habitats? On both hydrological and open to closed gradients rice fields lie at the point where non-native species have their most positive habitat association, indicating that rice fields provide a good match to the habitats occupied in the species' native range. Such niche-matching has been widely documented in non-native species, mostly with regard to the climatic niche (Blackburn, Lockwood & Cassey 2009; Strubbe *et al.* 2013) but also in respect to habitat associations (Blackburn, Lockwood & Cassey 2009). However, based on native habitat associations non-native species would also be expected to occur in natural wetlands, river valleys and heterogeneous and abandoned agriculture (Barnard 1997; del Hoyo, Elliot & Christie 2010). While all these habitats are colonised to some extent in the Iberian Peninsula (Matias 2002; Sullivan *et al.* 2012; Sullivan, Grundy & Franco 2014), the results of this study show that non-native species were more likely to be found in rice fields than in these other habitats. In addition, resources non-native species used for feeding and shelter were available in non-rice field habitats, so the positive selection of rice fields was unlikely to be driven by the distribution of these resources. This supports the role of factors other than niche-matching in determining habitat selection by non-native species. Differences between the habitat associations of native and non-native species were greatest in rice fields, indicating the rice fields were potentially the habitat with the greatest amount of underexploited resources (i.e. resources that could not be exploited by native species) that could be accessed by non-native species. Rice fields could therefore release non-native species from competition from native species.

Analyses in this study were restricted to a single guild, seed-eating birds, rather than the entire bird community. Other groups that occasionally eat seeds, such as larks (Alaudidae) and tits (Paridae), could compete with non-native seed eating birds and reduce the degree to which resources in rice fields were underexploited. The extent to which this happened is likely to be limited, as differences in foraging mechanisms and bill morphology mean that these species are functionally distinct to the seed-eating guild considered here. For example, in a previous analysis tits have been found to occupy distinct functional space to the seed-eating guild (Batalha, Ramos & Cardoso 2013, larks were not included in that study). Crested larks *Galerida cristata* were found present in most point counts, but were never observed feeding in field margins, while tits were rarely recorded (M Sullivan, unpublished data). It is therefore unlikely that inclusion of these species would have affected our conclusions.

Conversion of natural and semi-natural habitats to human-modified ones has previously been found to reduce species richness (Seabloom, Dobson & Stoms 2002), with a reduction in functional diversity due to local extinction of species with certain traits (Edwards *et al.* 2013). In addition, intensification of existing agriculture can alter the traits of plant communities to favour a few functional groups (Smart *et al.* 2006). Our work extends this by showing that the functional diversity of native species is lower in a novel land-use than in more traditional human modified land-uses. The ability of species to persist in human-modified habitats is a function of the number of species in the regional species pool that can pass through habitat filters provided by the modified habitats (Díaz *et al.* 2013). We expect traditional land-uses that contain functionally similar habitat elements to natural habitats are likely to pose less of a habitat filter than more novel land-uses. Further work is needed to assess whether the novelty of human-modified habitats (compared to natural habitats) is an important driver of functional diversity loss in other systems. For this loss of functional diversity to result in increased invasability we would expect human-modified habitats to have more underexploited resources. Experimental reduction of functional diversity has been found to enhance community invasability (Vaz-Pinto *et al.* 2013), as the removal of functional groups leaves vacant niches for invaders to exploit. In particular, the loss of species with similar traits to potential invaders is likely to facilitate invasion (Strauss, Webb & Salamin 2006). However, loss of functional diversity in response to land-use change is coupled with the loss of some resources (Siriwardena *et al.* 2000), so it is unclear whether human-modified habitats are easier to invade. Some support for the increased invasability of human modified habitats is provided by observations that non-native plant species richness in Great Britain is higher in human modified habitats, including agricultural systems (Maskell *et al.* 2006). However, non-native species richness is often correlated with native species richness, with the most resource rich habitats having the highest richness of both native and non-native species (Maskell *et al.* 2006; Fridley *et al.* 2007). Therefore, the high non-native species richness of many human modified habitats may be due to the greater provision of certain resources (e.g. through fertiliser addition).

This study demonstrates that novel land-uses can provide resources that were underexploited by native species, but which could be exploited by non-native species. While the selection of rice fields by non-native species was partly a result of niche matching, it is likely that reduced competition from native species facilitated invasion. This indicates that experimental evidence for unexploited resources promoting invasion also applies in wild populations at large spatial scales. Further work is required to investigate whether the results presented here can be generalised to other systems. If novel human-modified habitats typically have underexploited resources, these may provide ‘welcome mats’ to establishing non-native species, and so should be targeted by surveillance programmes.

Acknowledgements

We thank Supatchaya Tetachoochert, James Grundy, Chris Panter and Chris Sullivan for assistance with fieldwork, Tony Davy for assistance with plant identification, and Helder Cardoso, Vitor Encarnacão, the ICNF and Associação PATO for assisting with logistics in the field and providing access to field sites. MJPS was funded by a NERC PhD studentship.

References

Barnard, P. (1997) Common Waxbill. *The Atlas of Southern African Birds* (eds J.A. Harrison, D.G. Allan, L.G. Underhill, M. Herremans, A.J. Tree, V. Parker & C.J. Brown), pp. 612-613. BirdLife South Africa, Blairgowrie, South Africa.

Batalha, H.R., Ramos, J.A. & Cardoso, G.C. (2013) A successful avian invasion occupies a marginal ecological niche. *Acta Oecologica,* **49,** 92-98.

Bates, D., Maechler, M. & Dai, B. (2008) lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-28.

Blackburn, T.M., Lockwood, J.L. & Cassey, P. (2009) *Avian Invasions: The Ecology and Evolution on Exotic Birds*. Oxford University Press, Oxford.

Blackburn, T.M., Pysek, P., Bacher, S., Carlton, J.T., Duncan, R.P., Jarosik, V., Wilson, J.R. & Richardson, D.M. (2011) A proposed unified framework for biological invasions. *Trends Ecology and Evolution,* **26,** 333-339.

Cassey, P., Blackburn, T.M., Sol, D., Duncan, R.P. & Lockwood, J.L. (2004) Global patterns of introduction effort and establishment success in birds. *Proceedings of the Royal Society of London. Series B: Biological Sciences,* **271,** S405-S408.

Cornell, H.V. & Lawton, J.H. (1992) Species interactions, local and regional processes, and limits to the richness of ecological communities - a theoretical perspective. *Journal of Animal Ecology,* **61,** 1-12.

Costa, H., Elias, G.L. & Farinha, J.C. (1997) Exotic birds in Portugal. *British Birds,* **90,** 562-568.

Cramp, S. & Perrins, C. (1994a) *Handbook of the Birds of Europe, the Middle East and North Africa: Buntings to New World Warblers*. Oxford University Press, Oxford.

Cramp, S. & Perrins, C. (1994b) *Handbook of the Birds of Europe, the Middle East and North Africa: Crows to Finches*. Oxford University Press, Oxford.

Crawley, M.J. (2005) ECOLOGY: Learning from the Aliens. *Science,* **310,** 623-624.

del Hoyo, J., Elliot, A. & Christie, D. (2009) *Handbook of the Birds of the World: Bush Shrikes to Old World Sparrows*. Lynx Edicons, Barcelona.

del Hoyo, J., Elliot, A. & Christie, D. (2010) *Handbook of the Birds of the World: Weavers to New World Warblers*. Lynx Edicions, Barcelona.

Dolman, P.M., Panter, C.J. & Mossman, H.L. (2012) The biodiversity audit approach challenges regional priorities and identifies a mismatch in conservation. *Journal of Applied Ecology*, **49**, 986–997.

Díaz, S., Purvis, A., Cornelissen, J.H.C., Mace, G.M., Donoghue, M.J., Ewers, R.M., Jordano, P. & Pearse, W.D. (2013) Functional traits, the phylogeny of function, and ecosystem service vulnerability. *Ecology and Evolution,* **3,** 2958-2975.

Edwards, F.A., Edwards, D.P., Hamer, K.C. & Davies, R.G. (2013) Impacts of logging and conversion of rainforest to oil palm on the functional diversity of birds in Sundaland. *Ibis,* **155,** 313-326.

Fargione, J.E. & Tilman, D. (2005) Diversity decreases invasion via both sampling and complementarity effects. *Ecology Letters,* **8,** 604-611.

Fridley, J.D., Stachowicz, J.J., Naeem, S., Sax, D.F., Seabloom, E.W., Smith, M.D., Stohlgren, T.J., Tilman, D. & Von Holle, B. (2007) The invasion paradox: Reconciling pattern and process in species invasions. *Ecology,* **88,** 3-17.

Fritz, S.A., Bininda-Emonds, O.R.P. & Purvis, A. (2009) Geographical variation in predictors of mammalian extinction risk: big is bad, but only in the tropics. *Ecology Letters,* **12,** 538-549.

Gower, J.C. (1971) General coefficient of similarity and some of its properties. *Biometrics,* **27,** 857-871.

Helmus, M.R., Keller, W., Paterson, M.J., Yan, N.D., Cannon, C.H. & Rusak, J.A. (2010) Communities contain closely related species during ecosystem disturbance. *Ecology Letters,* **13,** 162-174.

Jacobs, J. (1974) Quantitative measurement of food selection. *Oecologia,* **14,** 413-417.

Joffre, R., Rambal, S. & Ratte, J.P. (1999) The dehesa system of southern Spain and Portugal as a natural ecosystem mimic. *Agroforestry Systems,* **45,** 57-79.

Kirmer, A., Tischew, S., Ozinga, W.A., von Lampe, M., Baasch, A. & van Groenendael, J.M. (2008) Importance of regional species pools and functional traits in colonization processes: predicting re-colonization after large-scale destruction of ecosystems. *Journal of Applied Ecology,* **45,** 1523-1530.

Levine, J.M., Adler, P.B. & Yelenik, S.G. (2004) A meta-analysis of biotic resistance to exotic plant invasions. *Ecology Letters,* **7,** 975-989.

Lima, A. (1997) Current situation of rice production in Portugal and the main diseases that occur. *Cahiers Options Méditerranéennes,* **15,** 29-40.

Loreau, M. (2000) Are communities saturated? On the relationship between alpha, beta and gamma diversity. *Ecology Letters,* **3,** 73-76.

Maskell, L., Firbank, L., Thompson, K., Bullock, J. & Smart, S. (2006) Interactions between non-native plant species and the floristic composition of common habitats. *Journal of Ecology,* **94,** 1052-1060.

Matias, R. (2002) *Aves exóticas que nidificam em Portugal continental*. ICNB, Lisbon.

McKinney, M. (2006) Correlated Non-native Species Richness of Birds, Mammals, Herptiles and Plants: Scale Effects of Area, Human Population and Native Plants. *Biological Invasions,* **8,** 415-425.

Mendez, V., Gill, J.A., Burton, N.H.K., Austin, G.E., Petchey, O.L. & Davies, R.G. (2012) Functional diversity across space and time: trends in wader communities on British estuaries. *Diversity and Distributions,* **18,** 356-365.

Moles, A.T., Flores-Moreno, H., Bonser, S.P., Warton, D.I., Helm, A., Warman, L., Eldridge, D.J., Jurado, E., Hemmings, F.A., Reich, P.B., Cavender-Bares, J., Seabloom, E.W., Mayfield, M.M., Sheil, D., Djietror, J.C., Peri, P.L., Enrico, L., Cabido, M.R., Setterfield, S.A., Lehmann, C.E.R. & Thomson, F.J. (2012) Invasions: the trail behind, the path ahead, and a test of a disturbing idea. *Journal of Ecology,* **100,** 116-127.

Newton, I. (1967) The adaptive radiation and feeding ecology of some british finches. *Ibis,* **109,** 33-96.

Petchey, O.L., Evans, K.L., Fishburn, I.S. & Gaston, K.J. (2007) Low functional diversity and no redundancy in British avian assemblages. *Journal of Animal Ecology,* **76,** 977-985.

Petchey, O.L. & Gaston, K.J. (2006) Functional diversity: back to basics and looking forward. *Ecology Letters,* **9,** 741-758.

R Development Core Team (2012) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna.

Ritzema, H.P., Satyanarayana, T.V., Raman, S. & Boonstra, J. (2008) Subsurface drainage to combat waterlogging and salinity in irrigated lands in India: Lessons learned in farmers' fields. *Agricultural Water Management,* **95,** 179-189.

Seabloom, E.W., Dobson, A.P. & Stoms, D.M. (2002) Extinction rates under nonrandom patterns of habitat loss. *Proceedings of the National Academy of Sciences of the United States of America,* **99,** 11229-11234.

Siriwardena, G.M., Crick, H.Q.P., Baillie, S.R. & Wilson, J.D. (2000) Agricultural land-use and the spatial distribution of granivorous lowland farmland birds. *Ecography,* **23,** 702-719.

Smart, S.M., Marrs, R.H., Le Duc, M.G., Thompson, K., Bunce, R.G.H., Firbank, L.G. & Rossall, M.J. (2006) Spatial relationships between intensive land cover and residual plant species diversity in temperate farmed landscapes. *Journal of Applied Ecology,* **43,** 1128-1137.

Stohlgren, T.J., Barnett, D.T., Jarnevich, C.S., Flather, C. & Kartesz, J. (2008) The myth of plant species saturation. *Ecology Letters,* **11,** 313-322.

Strauss, S.Y., Webb, C.O. & Salamin, N. (2006) Exotic taxa less related to native species are more invasive. *Proceedings of the National Academy of Sciences,* **103,** 5841-5845.

Strubbe, D., Broennimann, O., Chiron, F. & Matthysen, E. (2013) Niche conservatism in non-native birds in Europe: niche unfilling rather than niche expansion. *Global Ecology and Biogeography*, **22**, 962–970.

Sullivan, M.J.P., Davies, R.G., Reino, L. & Franco, A.M.A. (2012) Using dispersal information to model the species–environment relationship of spreading non-native species. *Methods in Ecology and Evolution,* **3,** 870-879.

Sullivan, M.J.P., Grundy, J. & Franco, A.M.A. (In Press) Assessing the impacts of the non-native Black-headed Weaver on native *Acrocephalus* warblers. *Ibis*.

Tilman, D. (2004) Niche tradeoffs, neutrality, and community structure: A stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences of the United States of America,* **101,** 10854-10861.

Vaz-Pinto, F., Olabarria, C., Gestoso, I., Cacabelos, E., Incera, M. & Arenas, F. (2013) Functional diversity and climate change: effects on the invasibility of macroalgal assemblages. *Biological Invasions,* **15,** 1833-1846.

Weerakoon, W.L. & Gunewardena, S.D.I.E. (1983) Rice field weed flora of Sri Lanka. *Tropical Agriculturist,* **139,** 1-14.

Whittingham, M.J. & Evans, K.L. (2004) The effects of habitat structure on predation risk of birds in agricultural landscapes. *Ibis,* **146,** 210-220.

6.5 *Appendix 6-1. Calculating habitat selection*

Detectability of birds declines with distance from observer. We therefore modelled the detectability of each species as a function of distance to calculate the effective area. This was done using DISTANCE 6.0 (Thomas *et al.* 2010). We fitted half normal and hazard rate functions to the data (Buckland *et al.* 1993), and selected the function (in all cases a half normal function) with the lowest AIC and best goodness of fit across a range of distances. Detectability can vary between observers, but we had insufficient data to model the effect of observer identity on the detectability of all species. We instead calculated the effect of observer identity on house sparrow *Passer domesticus* detectability, as this was the most frequently recorded species in our dataset, and added this correction to the detection functions for other species. For most species we had insufficient data to robustly fit habitat covariates, so do not model the effect habitat could have on detectability. We believe it unlikely that certain species were more detectable in emergent vegetation while others were more detectable in trees and bushes, but acknowledge that such an interaction would bias the estimation of habitat availability.

We calculated the effective area surveyed by imagining each focal watch sub-square as a three dimensional shape, where the height corresponds to detectability. If every bird was detected (i.e. detectability=1), then the volume (*V*) of each shape = 1x30x30 = 900. This can alternatively be written as equation 1, where *f(x)* is the detection function, *a* is the distance of the nearest side of the sub-square to the observer, and *b* is the distance of the farthest side of the sub-square to the observer.

$$V = \int_a^b f(x)\, dx \times 30 \qquad \text{Equation 1.}$$

The effective area of a resource in each sub-square can thus be calculated by multiplying the percentage cover of the resource in the sub-square by *V*. The effective proportion of each resource can then be calculated by dividing the summed effective area of each resource by the total effective surveyed area (i.e. the sum of *V* across sub-squares).

Resource use at each site was calculated by as the proportion of scan samples where each species was recorded doing an activity where they were using a given resource. The overall proportional use of each resource was calculated by taking the mean of these proportions across all sites, weighted by the number of observations of each activity at each site.

The proportional use of each resource for each activity (r_{ha}) was calculated by taking the mean of the proportion of scan samples where each species was recorded doing an activity ($Ntot_{ai}$) where they were using resource h ($Nres_{hai}$) across all sites, weighted by the number of observations of each activity at each site (equation 2).

$$r_{ha} = \frac{\sum_{i=1}^{n} Ntot_{ai} x_{hai}}{\sum_{i=1}^{n} Ntot_{ai}} \qquad \text{Equation 2a.}$$

Where

$$x_{hai} = \frac{Nres_{hai}}{Ntot_{ai}} \qquad \text{Equation 2b.}$$

Jacobs index was used to quantify resource selection. This was calculated by equation 3, where p_h is the mean effective area of each habitat across focal watch sites.

$$J_{ha} = \frac{r_{ha} - p_h}{((r_{ha} + p_h) - (2 \times r_{ha} \times p_h))} \qquad \text{Equation 3.}$$

6.6 Appendix 6-2. Obtaining trait data

Data on the diet, foraging behaviour, nesting habits and morphology of native species primarily came from (Cramp & Perrins 1994b, a), supplemented by quantitative diet data from Newton (1967) for finches found in the UK (i.e. chaffinch, greenfinch, goldfinch, siskin, linnet, crossbill, bullfinch and hawfinch). Equivalent data for non-native species came from Matias (2002), supplemented with data from Batalha *et al.* (2013) for the common waxbill bill morphology and nesting period. We did find any published data on the diet of black-headed weavers or yellow-crowned bishops in their non-native range, so scored them for these traits based on field observations (Appendix table 6.1), supplemented by descriptions of diet from their native range (del Hoyo, Elliot & Christie 2010).

Quantitative data on diet was not available for all species, so qualitative descriptions also had to be used. Descriptions in literature sources were compared to the qualitative descriptors given in Appendix table 6.2. Rules used to convert quantitative and qualitative data into ordinal scores is given in Appendix table 6.2.

Appendix table 6.1. Field data used to support trait scoring.

Species	Adult diet	Nestling diet	Agility	Grass/ compositae
Black-headed weaver	25 feeding observations where food item noted, 28% invertebrate, 72% seeds. Also one occasion feeding on reed stem, but not known if gleaning invertebrates or eating green material.	7 feeding observations where food seen in bill, 86% insect, 14% seed, however seed thought to be considerably underestimated as young frequently fed from crop after adults feeding on grasses.	76% of feeding observations on ground, but also frequently fed in low vegetation, usually clinging to vertical stems but sometimes also bent stems, clinging to multiple stems to feed.	28% feeding observations on grass, 8% on forbs.
Yellow-crowned bishop	11 feeding observations where food item known – all seed.	Not observed.	Regularly observed feeding on bent stems or upside down. Flushed from the ground, but all quantified feeding observations in low vegetation.	60% of feeding observations on grass, 20% on forbs.

Appendix table 6.2. Rules used to assign scores to traits.

Trait	Code	Quantitative score	Qualitative descriptor
Diet	0	0%	Never/ exceptionally recorded in diet
	1	<5%	Occasional part of diet
	2	5 to 25%	Regular minor component of diet
	3	25-50%	Forms a substantial part of diet, but not primary component
	4	50-75%	Forms the majority of diet, but considerable use of other food indicated
	5	>75% of diet	The dominant part of the diet. Other foods only form minority of diet.
Feeding agility	0		Never/ exceptionally recorded
	1		Rarely recorded
	2		Regularly recorded doing activity, but not a major feeding mode
	3		Common feeding mode
Foraging height	0		Never or rarely recorded
	1		Occasionally forages at this height
	2		Regularly forages at this height, but not primary foraging height
	3		Primary foraging height
Feeding habitat	0		Never or rarely used
	1		Occasionally used
	2		Important habitat, but not clearly dominant
	3		Dominant habitat
Food plants	0		Never/ rarely used
	1		Occasionally used
	2		Important food plant, but not clearly dominant

	3		Dominant food plant
Nest location	0		Never/ rarely used
	1		Sometimes used, but not primary nest location
	2		Main nest location
Nest height	0	<1m	
	1	>1m	
Nesting season	0		Never/ rarely recorded
	1		Some individuals likely to be nesting, but not majority of the population
	2		Main season

6.7 *Appendix 6-3. Quantifying habitat associations*

Appendix table 6.3. Scores given to major habitat types on environmental gradients.

Habitat type	Hydrological gradient	Open - closed gradient	Notes
Arable	4	1	
Gardens	4	4-5	This is given 4 on the open-closed gradient unless gardens are described as wooded or well vegetation, when they are given 5
Hedged field margins	4	4	
Heterogeneous cultivation	4	4	
Natural grassland	4	1	
Non-woody riverine vegetation	2	2	This is given a lower hydrological gradient score than riverine scrub as this category includes riverine emergent vegetation
Open woodland	4	5	
Orchards	4	4	
Pasture	4	1	
Rank grassland	4	2	
Rice fields	2	2	
Riverine scrub	3	4-5	This is given 4 on the open-closed gradient if it is described as light or open scrub, and 5 if it is described as dense, thick or heavy scrub
Savannah	4	4-5	This is given 4 or 5 on the open-closed gradient depending on tree cover
Scrub	4	4-5	This is given 4 on the open-closed gradient if it is described as light or open scrub, and 5 if it is described as dense, thick or heavy scrub
Wet grassland	3	1	
Wet woodland	3	6	
Wetlands	1	3	
Woodland	4	6	
Woodland edge	4	5	

Wood-pasture/ Dehesa	4	4-5	This is given a lower hydrological gradient score than riverine scrub as this category includes riverine emergent vegetation

References

Batalha, H.R., Ramos, J.A. & Cardoso, G.C. (2013) A successful avian invasion occupies a marginal ecological niche. *Acta Oecologica,* **49,** 92-98.

Buckland, S.T., Anderson, D.R., Burnham, K.P. & Laake, J.L. (1993) *Distance Sampling: Estimating Abundance of Biological Populations*. Chapman and Hall, London.

Cramp, S. & Perrins, C. (1994a) Handbook of the Birds of Europe, the Middle East and North Africa: Buntings to New World Warblers. Oxford University Press, Oxford.

Cramp, S. & Perrins, C. (1994b) *Handbook of the Birds of Europe, the Middle East and North Africa: Crows to Finches*. Oxford University Press, Oxford.

del Hoyo, J., Elliot, A. & Christie, D. (2010) *Handbook of the Birds of the World: Weavers to New Word Warblers.* Lynx Edicons, Barcelona.

Matias, R. (2002) *Aves exóticas que nidificam em Portugal continental*. ICNB, Lisbon.

Newton, I. (1967) THE adaptive radiation and feeding ecology of some british finches. *Ibis,* **109,** 33-96.

Thomas, L., Buckland, S.T., Rexstad, E.A., Laake, J.L., Strindberg, S., Hedley, S.L., Bishop, J.R.B., Marques, T.A. & Burnham, K.P. (2010) Distance software: design and analysis of distance sampling surveys for estimating population size. *Journal of Applied Ecology,* **47,** 5-14.

7 Testing multiple pathways for impacts of the non-native Black-headed Weaver *Ploceus melanocephalus* on native birds in the early phase of invasion

Published as Grundy, J.P.B., Franco, A.M.A, Sullivan, M.J.P. (2014) Testing multiple pathways for impacts of the non-native Black-headed Weaver Ploceus melanocephalus *on native birds in the early phase of invasion. Ibis* ***156****: 355-365.*

Material from this chapter also appears in Sullivan, M.J.P., Grundy, J., Franco, A.M.A. (2014) Assessing the impacts of the non-native Black-headed Weaver on native Acrocephalus *warblers. Ibis* ***156****: 231–232.*

Some, but not all non-native species have strong negative impacts on native species. It is desirable to identify whether a non-native species will have a negative impact at an early stage in the invasion process, while management options such as eradication are still available. Although it may be difficult to detect early impacts of non-native species, this is necessary to ensure that management decisions can be based on case-specific scientific evidence. We investigate the impacts of a non-native bird, the Black-headed Weaver *Ploceus melanocephalus*, at an early stage in its invasion of the Iberian Peninsula. To do this we, a priori, identify potential pathways by which competition for shared resources by Black-headed Weavers could lead to population declines in ecologically similar native species, and generate hypotheses to test for evidence of competition along these pathways. Black-headed Weavers could potentially impact native species by displacing them from nesting habitat, or by locally reducing habitat quality. We did not find evidence for either potential competition pathway, suggesting that Black-headed Weavers do not currently compete with native species. However, it is possible that mechanisms that currently allow coexistence may not operate once Black-headed Weavers reach higher population densities or different habitats.

7.1 Introduction

Non-native species are major drivers of avian biodiversity loss (Clavero *et al.* 2009). While some of the most severe impacts have been caused by introduced mammalian predators (Blackburn *et al.* 2004, Hilton & Cuthbert 2010), exotic birds can impact native species through a number of mechanisms, such as predation, hybridisation and transmission of disease (Kumschick & Nentwig 2010). Although rarely demonstrated, non-native birds have also been suspected of competing with native species (Blackburn, Lockwood & Cassey 2009). For example, the establishment of the Common Myna *Acridotheres tristis* in Australia was followed by a decline in the abundance of a number of native bird species (Grarock *et al.* 2012). Whether species compete depends on the degree to which niche differences result in one species limiting their own population more than the populations of other species (Chesson 2000, Adler, HilleRisLamber & Levine 2007) the degree of asymmetry in the competitive weights of species (Adler *et al.* 2007), and the presence of other mechanisms such as predation that limit the population of one species more than others (Griswold & Lounibos 2005). Competitive exclusion is likely when species share similar resource requirements (Ieronymidou, Collar & Dolman 2012), and one species is either dominant at accessing those resources, or has a faster reproduction rate (Chesson 2000, Adler, Ellner & Levine 2010).

It is desirable to identify whether a non-native species will compete with native species early in the invasion process, while the non-native species has a restricted distribution and eradication remains feasible (Lodge *et al.* 2006). However, it is easier to evaluate impacts when an invasion is advanced, as more data are available, allowing competition to be identified with more confidence (Wiens 1989). This leads to a trade-off between early risk assessment and the strength of evidence for demonstrating the existence of an impact. While this has motivated researchers and policy makers to suggest that lack of scientific certainty should not preclude control of non-native species (UNEP 1992, Sixth Conference of the Parties Convention on Biological Diversity 2002, Edelaar & Tella 2012), eradication is costly and poses animal welfare issues (Defra 2003), leading to recognition that it is desirable to have an evidence base to prioritise and justify management actions (Defra 2003, EEA 2010). Thus there is need to use scientific evidence collected early in the invasion process to aid management decisions. This is especially pressing in the Iberian Peninsula, where the number of non-native birds species recorded breeding has increased rapidly since the late 1980s (Matias 2002).

We assess the evidence for competition between the recently established Black-headed Weaver *Ploceus melanocephalus* and two ecologically similar native birds. Black-headed Weavers are native to sub-Saharan Africa, and were first recorded in the Iberian Peninsula in the mid-1990s (Matias 2002). Breeding was confirmed in fewer than ten 10km^2 grid cells in the most recent Portuguese and Spanish breeding bird atlases (Marti & de Moral 2003, Equipa Atlas 2008). Black-headed Weavers nest in emergent vegetation (Colias & Colias 1964) and feed their nestlings on large invertebrates collected primarily by gleaning vegetation (Moreau 1960, Fry & Keith 2004), so share resource requirements with native Great Reed Warblers *Acrocephalus arundinaceus* and Eurasian Reed Warblers *A. scirpaceus* (Graveland 1996, Matias 2002, Cardoso 2008, Leisler & Schulze-Hagen 2011). Black-headed Weavers have been reported behaving aggressively towards both species (Matias 2002). While this could indicate that they are dominant at accessing shared resources, this has not been tested.

At the current stage of invasion we cannot test the influence of Black-headed Weavers on the productivity of native species, so instead focus on detecting behavioural responses to competition. Our approach involves identifying possible pathways by which shared resource requirements could lead to population declines of native species, and generating testable hypotheses for processes along these pathways (Figure 7.1). We test each of the following numbered hypotheses. We first test whether resource requirements of Black-headed Weavers overlap with native species (Figure 7.1, Hypothesis 1). We speculate that this could have a negative impact on native species if Black-headed Weavers exhibit interspecific territoriality and thereby exclude native species (Figure 7.1, Hypotheses 2-4), or locally reduce habitat quality (Figure 7.1, Hypotheses 5-6). Both of these could lead to population declines either by forcing native species to nest in sub-optimal habitat (Figure 7.1, Hypothesis 7), or by directly reducing the space available for native species. By testing for competition at a range of stages along these pathways we can maximise our ability to detect competition, and have a useful framework for assessing the potential for competitive exclusion.

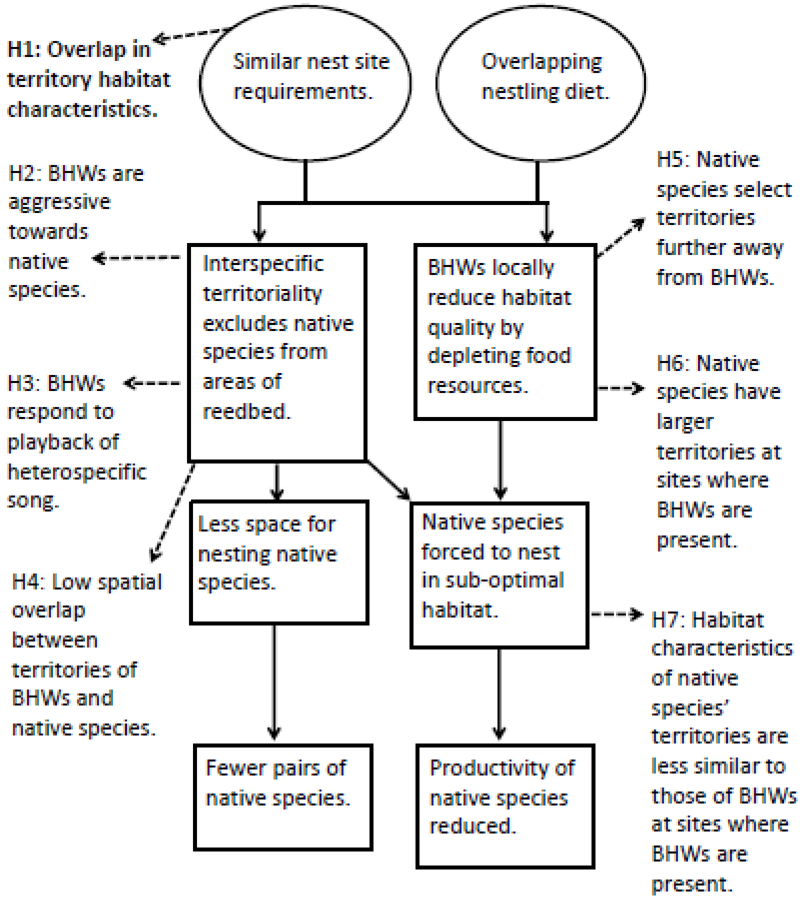


Figure 7.1. Potential pathways by which Black-headed Weavers (BHWs) may impact the population of native *Acrocephalus* warblers. Overlapping resource requirements are shown by ovals, processes are shown by rectangles connected by solid arrows. Hypotheses are linked to the relevant process by dashed arrows. Supported hypotheses (see results) are shown in bold.

7.2 Methods

7.2.1 Study sites

Fieldwork was conducted at four sites in western Portugal. Black-headed Weavers have been established at Paul de Tornada (PT, 39.448° N, 9.135° W) and Barroca d'Alva (BA, 38.729° N, 8.899° W) since the mid-1990s (Matias 2002). Uncolonised sites, with similar habitat and within 20km of colonised sites, were selected as controls. These were Lagoa de Óbidos (LO, 39.385° N, 9.210° W) and Lezíria Grande (LG, 38.931° N, 8.964° W). PT and LO are both extensive wetlands, consisting of patchworks of reedbed (dominated by reed *Phragmities australis)* and open water. BA and LG both consist of reed lined ditches crossing a mix of rice and wheat cultivation and pasture. The colonised study sites selected have high population densities of Black-headed Weavers and native *Acrocephalus* warblers, so potentially provide the best data available on the interaction of Black-headed Weavers and native species. Fieldwork was not conducted at other sites colonised by Black-headed Weavers as they either were unsuitable for Reed Warblers and Great Reed Warblers, or were ecologically sensitive sites.

7.2.2 Playback experiment and aggressive interactions

If Black-headed Weavers exhibited interspecific territoriality towards native species we would expect them to initiate aggressive interactions with native species, and possibly also respond to heterospecific song. To test whether Black-headed Weavers initiated aggressive interactions (defined as fights and chases) with native species (hypothesis two) all incidents of aggression between Black-headed Weavers and native species observed during fieldwork were recorded. Where possible, the species initiating aggression was noted. A binomial test was used to test whether the proportion of aggressive interactions differed from random expectation.

In order to test hypothesis three we conducted a playback experiment to test the reaction of Black-headed Weavers to conspecific and heterospecific song in May 2012, during the weaver breeding season. Songs of Black-headed Weaver, Great Reed Warbler (from Constantine *et al.* 2006) and Eurasian Reed Warbler (from Roche 1997), as well as a recording of background noise made at night at PT, were played from a portable speaker placed five metres away from Black-headed Weaver nests. The quality of warbler recordings was checked by playing these recordings within conspecific territories, and both elicited a reaction. Each recording was played for five minutes, as Catchpole (1978) found this was sufficient time to elicit a response from Eurasian Reed Warblers. Playback experiments were videoed, and the distance of closest approach by Black-headed Weavers during the playback was estimated to the nearest metre.

The responses of Black-headed Weavers from 16 territories (eight at PT and eight at BA) were tested over a three day period to reduce seasonal variation in individual motivation to respond (Dunn *et al.* 2004, Golabek *et al.* 2012). To minimise the effect of habituation, no more than two recordings were played in each territory in one day, with one recording played in the morning and one in the evening. To further control for habituation, the order in which recordings were played was balanced across the 16 territories.

We modelled the distance of approach (m) by Black-headed Weavers as a function of playback treatment using a generalised linear mixed model, with territory identity as a random effect. Data from both sites were pooled as site identity was not significant when included in the previous model ($t_{44} = 0.521$, $P = 0.605$). Due to convergence issues, the model was fitted using quasi-likelihood, with the mean-variance relationship set so that the variance increased with the mean.

7.2.3 Territory and habitat mapping

We made 12 territory mapping visits to each site between early April and late June 2012 (i.e. from territory establishment to nesting for all species (Cramp 1992, Matias 2002)) to record the locations of Black-headed Weaver, Great Reed Warbler and Eurasian Reed Warbler territories. Sites were visited during the morning active period (Robbins 1981), and observations of target species were mapped onto a base map with the aid of a handheld GPS unit. We assigned these observations to territories following Marchant (1983).

We only used observations of singing, fighting or territorial calling birds for determining territory size. Observations were digitised using ArcMap 9.3 (ESRI 2008), and projected onto a Universal Transverse Mercator grid (zone 29N). We calculated the territory centroid by taking the mean of the coordinates of these observations, and delimited territory boundaries by constructing the minimum convex polygon (MCP) that encompassed observations from each territory. Aerial photographs (1 m resolution, Instituto Geográfico Português 2004) were digitised to produce vector maps of reedbed at each site, which were updated based on field observations where there had been large changes in reedbed extent. These maps were used to clip territory MCPs so that they only contained reedbed. We did this so that territories reflected utilisation distributions more closely; areas of open water and agriculture were rarely used by *Acrocephalus* warblers (J.P.B. Grundy, pers. obs.), so contributed very little to the resources available to breeding birds.

To test hypothesis four, territory overlap between pairs of species was calculated by dividing the area occupied by both species by the total area occupied by either species. This calculation was performed on a raster grid (~5m resolution), rather than directly on the vector layers, to aid comparison with a null model. The purpose of the null model was to randomly shift the position of each territory, while maintaining the number of territories at each site, observed territory size and restricting territories to be in reedbed. Further details of the null model mechanism are given in Supporting Information Appendix 7-1. The null model did not restrict intraspecific territory overlap, but overlap of randomly generated conspecific territories was still similar to observed overlap. The null model was run for 1000 iterations, and the overlap between heterospecific territories was calculated in each case, to give a null distribution of overlap values. Competitive exclusion will lead to lower than expected observed values, while selection of similar reedbed habitat will lead to greater observed values than expected. Two-tailed *P*-values were calculated by comparing the observed overlap to quantiles of this null distribution.

We recorded the date of first occupancy of each territory by Great Reed Warblers as this relates to the male's assessment of territory quality (Bensch & Hasselquist 1991). This allowed us to test hypothesis five, as the earliest occupied territories should also be the highest quality ones. We restricted this analysis to Great Reed Warblers as previous studies have shown that the order of territory occupancy relates to territory quality (Bensch & Hasselquist 1991), while it is unknown whether the same holds for Eurasian Reed Warblers. The distance (m) between the centroid of Great Reed Warbler and Black-headed Weaver territories was calculated, and its natural logarithm used to model the date of first occupancy of each territory. As the availability of territories at different distances to Black-headed Weavers varied between sites, site was also included in the model. Territories were not visited every day (median interval between visits = 5.5 days), so a bird may have arrived several days before the recorded occupation date. We tested the sensitivity of our analysis to this measurement error by randomly selecting the date of occupation from the pool of possible dates, and re-running the analysis with 1000 repetitions.

We calculated the size (m^2) of reedbed-clipped MCPs. Some passerines have larger territories when food availability is low (Marshall & Cooper 2004), so food depletion by Black-headed Weavers may cause native species to have larger territories (hypothesis six). Territory size of Eurasian and Great Reed Warblers was modelled as a function of site using a generalised linear model with a gamma distribution to account for the positive mean-variance relationship, with post-hoc Tukey tests performed using the R package multcomp (Hothorn, Bretz & Westfall 2008).

7.2.4 Habitat sampling

Seven territories of each species, corresponding to the minimum number of Great Reed Warbler territories at any one of our study sites, and seven areas of unoccupied reedbed were randomly chosen at each site. At each location habitat variables were measured in one randomly placed 50 x 50 cm quadrat, with the exception of two quadrats being placed in Great Reed Warbler territories because of their larger territory size (Cramp 1992). In each quadrat, we measured the height (cm) of ten new (current season's growth) and ten old (previous season's growth) reeds, the diameter (mm) of ten new and ten old reeds, the density of new and old reeds (measured by counting all reeds within the quadrat), and the percentage cover of reeds, other emergent vegetation, herbaceous plants, woody plants and grasses (estimated visually). These were selected as habitat variables that had been identified as being important for the target species (Dyrcz 1986, Graveland 1996, Martinez-Vilalta *et al.* 2002, Poulin, Lefebvre & Mauchamp 2002), and considered to capture variation in reedbed habitat. Water depth is also an important influence on Great Reed Warbler nest site selection (Graveland 1998), but management of agricultural ditches caused water levels to fluctuate between days at our study sites, so this variable was not included in analyses.

Differences in habitat between species (hypotheses one) were identified using non-metric multidimensional scaling (NMDS), performed in PRIMER v6 (Clarke & Gorley 2006) based on a Euclidean distance matrix generated from the habitat variables. NMDS allows dissimilarities to be mapped in two dimensions. Stress values assess the fit between distances in the distance matrix and those in two dimensional space. Stress values of less than 0.1 indicate a good fit (Clarke & Warwick 1994); the stress value of 0.08 in this study therefore indicates good fit. We investigated how areas of NMDS space related to different habitat characteristics by modelling the matrix of raw habitat variables as a function of NMDS coordinates using the manylm function in the R package mvabund (Wang *et al.* 2012), and plotting the direction of these relationships. We used D (Schoener 1970) to calculate the overlap in habitat associations of the three species. To do this, a kernel density function was used to calculate the density of territories of each species in habitat space. D is then calculated as

$$D = 1 - \tfrac{1}{2}\left(\sum_{ij} |z_{1ij} - z_{2ij}|\right),$$

where z_{1ij} is the standardised territory density of species one and z_{2ij} is the standardised territory density of species two at point ij in environmental space. Full details on the calculation of D are given in Broennimann *et al.* (2012). D ranges from zero to one, with values closer to one indicating higher overlap. We tested whether the overlap between habitat associations of native species shifted to be less similar to those of Black-headed Weavers at sites where Black-headed Weavers are present (hypotheses seven). To do this, we compared observed values of *D* for the overlap between the densities of territories of native species and Black-headed Weavers at sites where Black-headed Weavers were present to values of *D* generated in 1000 iterations of a null model that randomly allocates observations to groups while maintaining the original number of observations in each group (the identity test, Warren, Glor & Turelli 2008).

Having multiple sampling points in Great Reed Warbler territories (due to their larger territory size than other study species) allowed us to test whether variation between territories of the same species was greater than variation within territories. Sampling points within the same Great Reed Warbler territory had more similar habitat characteristics than sampling points in different territories (median Euclidean distance within territories = 77.9, median Euclidean distance between territories = 97.0, Wilcoxon test, $P = 0.076$), justifying the decision to concentrate sampling effort on maximising the number of territories sampled, rather than sampling more points within a territory. Unless otherwise stated, all statistical analyses were performed in R v2.15 (R Development Core Team 2012), with power analyses performed using the package pwr (Champely 2007).

7.3 Results

7.3.1 Do native and non-native species use similar nesting habitat?

Numbers of study species recorded at each site are given in Table 7.1. Great Reed Warblers occupied less habitat space than the other species (Figure 7.2a). Great Reed Warbler territories were characterised by having taller and thicker reeds, although both Eurasian Reed Warblers and Black-headed Weavers also used this habitat (Figure 7.2a & d). Black-headed Weaver and Eurasian Reed Warbler territories overlapped in habitat space more than either species overlapped with Great Reed Warblers (Table 7.2. Overlap (Schoener’s D) between territories of Reed Warbler, Great Reed Warbler and Black-headed Weaver in habitat space.). These results support hypothesis one (Figure 7.1).

Table 7.1. Number of territories of target species recorded at each study site.

Species	Lagoa de Óbidos	Paul de Tornada	Barroca d'Alva	Lezíria Grande
Reed Warbler	27	22	26	29
Great Reed Warbler	8	7	7	7
Black-headed Weaver	0	10	16	0

Table 7.2. Overlap (Schoener’s D) between territories of Reed Warbler, Great Reed Warbler and Black-headed Weaver in habitat space.

	Black-headed Weaver	Great Reed Warbler
Reed Warbler	0.725	0.527
Great Reed Warbler	0.544	

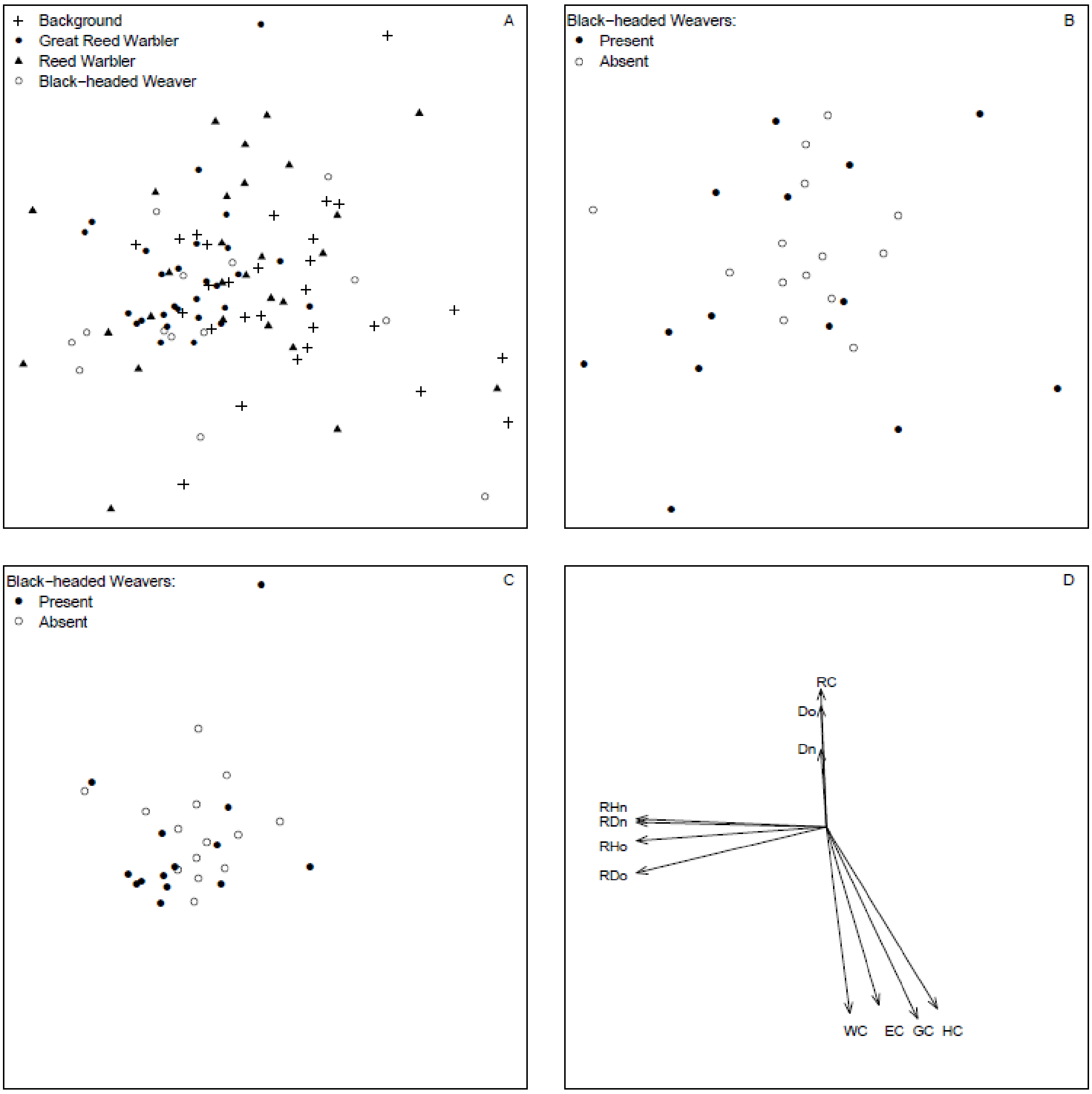


Figure 7.2. NMDS ordination of habitat characteristics in target species territories. Stress=0.08. (A) Position of target species territories and unoccupied background reedbed in NMDS space. (B) Position of Reed Warbler territories in NMDS space at sites where Black-headed Weavers were present and absent. (C) Position of Great Reed Warbler territories in NMDS space at sites where Black-headed Weavers were present and absent. (D) Relationship between habitat variables and the NMDS space. Arrows show the direction of relationships between habitat variables and environmental space. Arrow lengths were only selected for presentation purposes. RHn, height of new reeds (cm); Rho, height of old reeds (cm); RDn, diameter of new reeds (mm); RDo, diameter of old reeds (mm); Dn, density of new reeds; Do, density of old reeds; RC, percentage cover of reeds; HC, percentage cover of herbaceous plants; GC, percentage cover of grasses; EC, percentage cover of emergent vegetation excluding reeds; WC, percentage cover of woody vegetation.

7.3.2 Is there interspecific territoriality?

Limited support was found for hypotheses two and three (Figure 7.1). Aggressive interactions were rarely noted between Black-headed Weavers and native species; in over 120 hours of fieldwork, seven aggressive interactions were observed. In five out of the six occasions where the aggressor was observed, Black-headed Weavers initiated aggression (Binomial test, P = 0.219). Black-headed Weavers approached conspecific song (t_{44}=2.642, P = 0.011, Figure 7.3), but not heterospecific song ($t_{44} \leq 1.723$, $P \geq 0.092$, Figure 7.3) significantly more than background noise.

Observed territory overlap was never lower than expected if territories were randomly distributed, so no support was found for hypothesis four (Figure 7.1). Overlap between Great Reed Warbler and Black-headed Weaver territories was higher than expected if territories were randomly distributed at BA ($\text{Overlap}_{\text{OBS}}$ = 0.256, $\text{Overlap}_{\text{NULL-Median}}$ = 0, P = 0.01) but not significantly different than expected at PT ($\text{Overlap}_{\text{OBS}}$ = 0.011, $\text{Overlap}_{\text{NULL-Median}}$ = 0, P = 0.43). Overlap between Eurasian Reed Warbler and Black-headed Weaver territories was higher than expected if territories were randomly distributed at both PT ($\text{Overlap}_{\text{OBS}}$ = 0.046, $\text{Overlap}_{\text{NULL-Median}}$ = 0, $P < 0.001$) and BA ($\text{Overlap}_{\text{OBS}}$ = 0.327, $\text{Overlap}_{\text{NULL-Median}}$ = 0, $P < 0.001$).

7.3.3 Do Black-headed Weavers reduce habitat quality?

No support was found for hypotheses five, six and seven (Figure 7.1). Great Reed Warbler territory occupation date did not vary significantly between sites ($F_{1,\,10}$ = 1.45, P = 0.256). The distance to the nearest Black-headed Weaver territory did not influence territory occupation date of Great Reed Warblers ($F_{1,\,10} < 0.01$, P = 0.951, Figure 7.4). This result was robust to measurement error caused by gaps between territory mapping visits, as no significant relationships were observed in any permutation of possible occupation dates.

Both Eurasian Reed Warbler and Great Reed Warbler territories were larger in extensive wetland sites than ditch-crossed sites (Figure 7.4). Territory size was not affected by the presence of Black-headed Weavers (Figure 7.4).

Neither Eurasian Reed Warbler (D_{OBS} = 0.791, $D_{\text{NULL-Median}}$ = 0.715, P = 0.164, Figure 7.2b) nor Great Reed Warbler (D_{OBS} = 0.629, $D_{\text{NULL-Median}}$ = 0.546, P = 0.170, Figure 7.2c) territories shifted to be more or less similar to Black-headed Weaver territories at sites where Black-headed Weavers were present. The overlap in territory characteristics of Eurasian Reed Warbler and Great Reed Warbler territories in sites where Black-headed Weavers were present (D_{OBS} = 0.517) and where Black-headed Weavers were absent (D_{OBS} = 0.558) did not differ significantly from the expected overlap from a null model randomly shuffling, for each species, territories in sites with and without Black-headed Weavers ($D_{NULL - Median} = 0.517$, $P \geq 0.66$).

7.3.4 Power analysis

Non-significant results in the direction expected by our hypotheses were found for the response of Black-headed Weavers to native species' songs, and the proportion of aggressive interactions initiated by Black-headed Weavers. We were only able to detect large effect sizes in these analyses; the former analysis had sufficient power to identify mean approaches of ≥ 1.07m as being significantly different from responses to background noise, while the latter analysis would only be significant if all aggressive interactions were initiated by Black-headed Weavers.

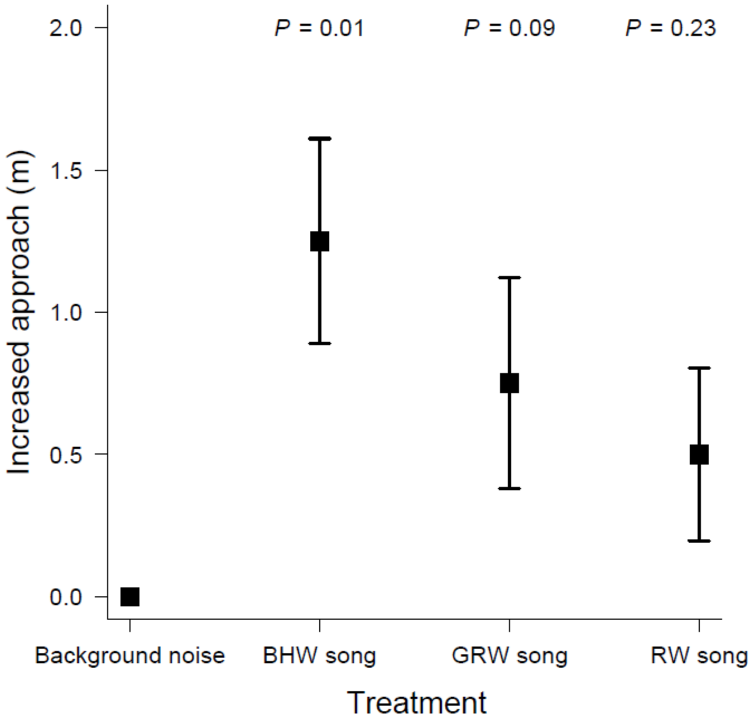


Figure 7.3. Response of male Black-headed Weavers to playback treatments. Mean responses are plotted, with error bars showing the standard error. *P* values show how significant the difference between the response to each treatment was from the response to background noise, and were calculated using a generalised linear mixed model modelling the increased approach as a function of treatment, with territory identity as a random effect. BHW, Black-headed Weaver; GRW, Great Reed Warbler; RW, Reed Warbler.

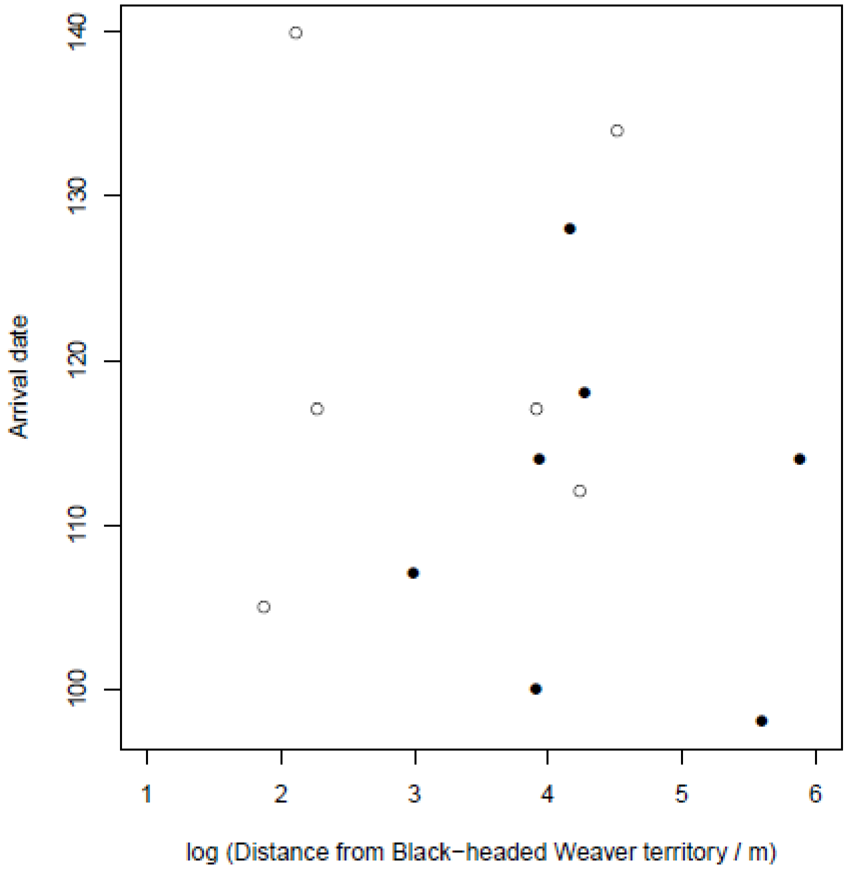


Figure 7.4. Relationship between arrival date of Great Reed Warblers (given as days since the start of the year) and distance of Great Reed Warbler territory from the centroid of the nearest Black-headed Weaver territory. Filled circles show territories from Paul de Tornada, open circles show territories from Barroca d'Alva.

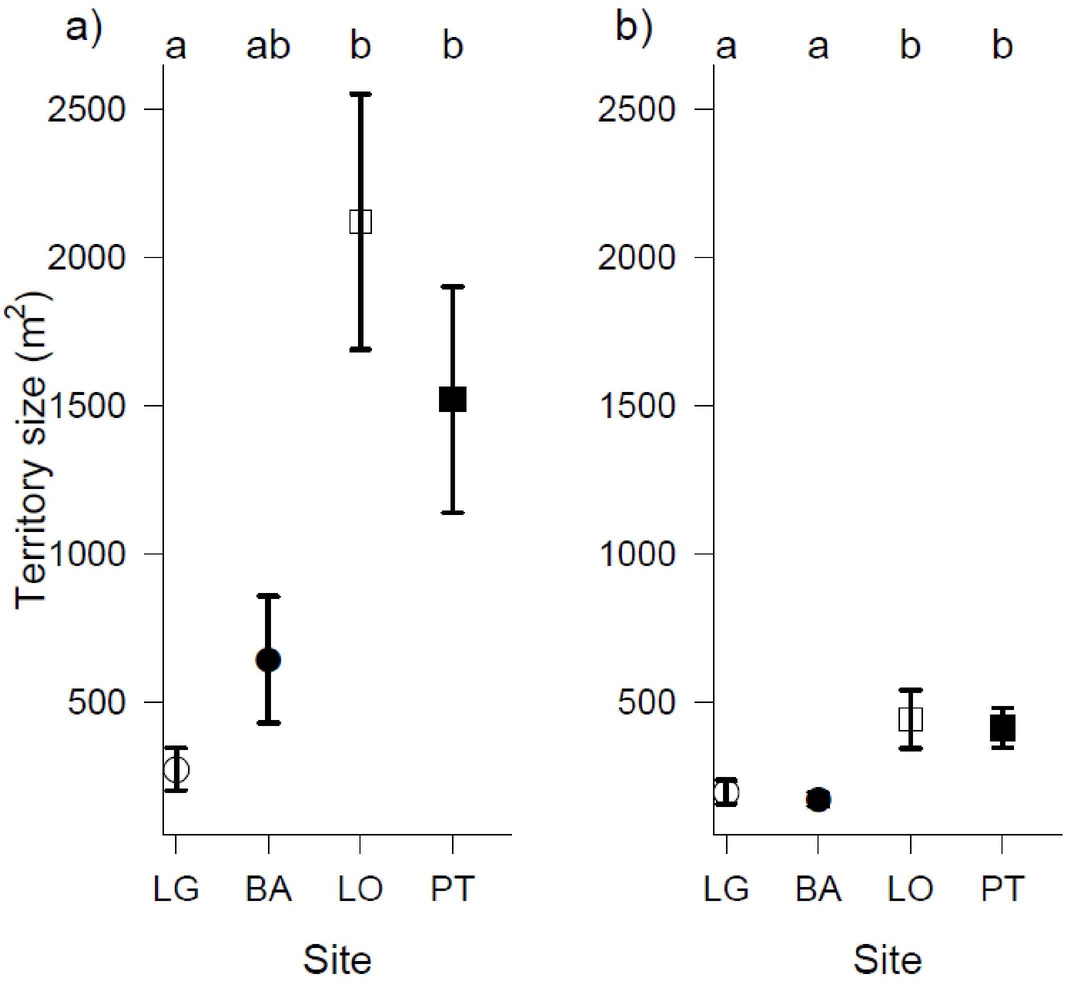


Figure 7.5. Mean ± SE territory sizes of (A) Great Reed Warblers and (B) Reed Warblers at the study sites. Extensive wetland sites are plotted with squares; ditch-crossed sites are plotted with circles. Filled shapes denote sites where Black-headed Weavers are present, and unfilled shapes denote sites where they are absent. Letters indicate sites that did not significantly differ (i.e. $P > 0.05$) in post-hoc tests performed on each species.

7.4 Discussion

7.4.1 Evidence for pathways to competition

Whilst there was overlap in the habitat characteristics of territories of Black-headed Weavers and native *Acrocephalus* warblers, we did not find any statistically significant evidence to support the hypothesis that competition by Black-headed Weavers is currently having population impacts on native species. We therefore conclude that at current population densities (0.43 to 0.70 pairs ha^{-1} in our study sites, Sullivan, Grundy & Franco 2014) Black-headed Weavers are unlikely to have a negative impact on ecologically similar native species.

The habitat characteristics of Eurasian Reed Warbler and Great Reed Warbler territories were similar to those reported in previous studies (Graveland 1996, Leisler & Schulze-Hagen 2011). Great Reed Warblers occupied areas with tall, thick reeds, often associated with the water-facing margin of reedbeds (Graveland 1998). Eurasian Reed Warblers and Black-headed Weavers occupied these areas, but were also found in areas of reedbed that were encroached by terrestrial vegetation (Figure 7.2). Eurasian Reed Warblers were the main species that occupied dense reed, which is often associated with the land-facing margin of reedbeds (Leisler & Schulze-Hagen 2011). Because all three species overlapped in habitat requirements, they are likely to select similar areas of reedbed, which may explain the higher than expected spatial overlap between heterospecific territories at some sites.

The larger size of Eurasian Reed Warbler and Great Reed Warbler territories in extensive reedbeds compared to reed-lined ditches supports previous studies (Dyrcz 1986). Food depletion by Black-headed Weavers could cause native species to increase the size of their territories (Marshall & Cooper 2004), but we did not find any evidence for this.

Although there is anecdotal evidence of Black-headed Weavers displaying aggression towards native species, we found little evidence for this. The results of the playback experiment did not support the hypothesis that Black-headed Weavers respond to native species song. The recordings of Eurasian Reed Warbler and Great Reed Warbler song used in the playback experiment elicit a response from conspecifics, but did not lead to a statistically significant response from Black-headed Weavers. This could be a type II error, as there was a weak tendency for Black-headed Weavers to approach Great Reed Warbler song, but the response was less strong than to conspecific song. It is unlikely that visual stimuli were required to evoke territorial behaviour towards heterospecifics, as aggressive interactions were rarely noted. Additionally, Black-headed Weavers were frequently observed close to native species without being aggressive (J.P.B. Grundy, pers. obs.). Therefore, at present there is little support for territorial defence against reed warblers by Black-headed Weavers.

The proximity to Black-headed Weavers did not influence the attractiveness of territories to returning male Great Reed Warblers. Great Reed Warblers are philopatric to their natal site (Bensch & Hasselquist 1991), so have information about the quality of reedbed patches from previous years. The locations of Black-headed Weaver territories are fairly consistent between years (M.J.P. Sullivan, unpubl. data), so if they reduced Great Reed Warbler productivity this information would be available to returning Great Reed Warblers. Neither Eurasian Reed Warblers nor Great Reed Warblers shifted into habitat less similar to Black-headed Weavers at sites where Black-headed Weavers were present. This does not support the hypothesis that Black-headed Weavers affect native *Acrocephalus* warblers by forcing them into sub-optimal habitat.

We did not directly assess whether Black-headed Weavers reduce the productivity of native species. Due to the restricted distribution of Black-headed Weavers, it would be difficult to disentangle the effects of Black-headed Weavers from other variables on the productivity of native species. Black-headed Weavers could reduce the productivity of native species by competing for nestling food, without causing displacement. In fact, any feeding competition from weavers is likely to be diffused to some extent as although female weavers foraged mainly in their territories, males often foraged outside their territories (J.P.B. Grundy, pers. obs.). Directly testing whether Black-headed Weavers affect the productivity of native species would provide compelling evidence for or against competition acting at territory level, but is not feasible at the present stage in the invasion.

We have only explored a limited range of potential impacts by Black-headed Weavers. Although *Acrocephalus* warblers were the most ecologically similar native species, Black-headed Weavers could also compete for reedbed nesting sites with species such as Savi’s Warblers *Locustella luscinioides*, and for winter food with a range of native granivorous birds. Aside from competition, Black-headed Weavers could have negative impacts by influencing disease transmission, as they are reservoirs for local haemoparasites (Ventim *et al.* 2012).

The apparent coexistence of Black-headed Weavers and *Acrocephalus* warblers may be due to mechanisms that only operate at low population densities. For example, inter-specific territoriality between *Acrocephalus* warblers motivated by factors other than resource defence (Leisler & Schulze-Hagen 2011) reduces their population densities below the resource carrying capacity (Mikami, Kohda & Kawata 2004). This could allow Black-headed Weavers to colonise without impacting native species, however, it is possible that shared resources become limiting when Black-headed Weavers reach higher population densities.

7.4.2 Application to other avian invasions

Pathways from resource overlap to population reduction of native species can be constructed for other non-native species, and could be used to assess the risk posed by newly established species. This can be illustrated using work on two established non-native species as examples.

Both Ring-necked Parakeets *Psittacula krameri* and European Starlings *Sturnus vulgaris* nest in tree cavities and so could compete for this resource with native hole-nesters in Europe and North America respectively. If they are dominant at accessing tree cavities then they can potentially limit the availability of nest sites for native species, which if sufficiently scarce could limit the population of these species (Newton 1994). Small scale studies have demonstrated that both European Starlings (Weitzel 1988) and Ring-necked Parakeets (Strubbe & Matthysen 2009) can displace native species from nest sites. European Starlings may cause native species to alter the timing of their breeding or to nest in sub-optimal cavities, although Koch *et al.* (2012) found limited evidence for this. While these studies have been performed when the species are widespread, similar studies could have been carried out in the early stages of both invasions and used to inform management decisions. Our knowledge of the impacts of a non-native species will be refined as a species spreads, as large-scale studies that could provide stronger evidence for competition are possible. For example, large scale studies have shown that the population level impacts of both European Starlings and Ring-necked Parakeets are limited at current densities (Koenig 2003, Newson *et al.* 2011).

7.4.3 Challenges with informing management decisions

Information on the potential impacts of non-native species is often limited to anecdotal reports, making risk assessment challenging (Strubbe, Shwartz & Chiron 2011). While some researchers argue for a precautionary, zero tolerance approach to non-native species (Edelaar & Tella 2012), others consider that management actions should relate to the amount of evidence that a non-native species has a negative impact (Bauer & Woog 2011). There is a trade-off between statistical power and timely intervention when investigating the impacts of non-native species. For instance, the individual statistical tests used in this study had low statistical power, so would only have been able to detect impacts with large effect sizes. For example, tendencies for Black-headed Weavers to initiate aggression and approach Great Reed Warbler song may have been non-significant due to low statistical power rather than due to the absence of an effect. The failure to find evidence for negative impacts early in an invasion should not be interpreted as conclusive evidence of absence of negative impacts, due to the risk of type II errors, and the fact that coexistence at low population densities may not persist at high population densities. Repeating this study when Black-headed Weavers are more widespread, and hence with a larger sample size, may allow the detection of small impacts that could not be detected in this study. However, eradication becomes increasingly difficult as a species spreads (Lodge *et al.* 2006), so studies that investigate the early impact of non-native species are important.

We recommend taking a pragmatic approach to interpreting the results of studies such as this. As well as testing the statistical significance of hypotheses, we suggest looking at the direction of relationships and magnitude of effect that can be detected given statistical power. This allows identification of species that are showing clear early impacts (i.e. statistically significant results to hypothesis testing), horizon scanning for impacts that may later prove to be significant (i.e. non-significant results in the hypothesised direction), and assessment of uncertainty based on the power of statistical tests. Studies such as this can be performed on multiple species, and the results can be compared in order to prioritise management actions. By testing multiple hypotheses along potential pathways to competitive exclusion we have a clear framework for evaluating the potential for competition, allowing the provision of information to aid management decisions early in the invasion process when eradication is feasible.

Acknowledgements

We thank Hannah Mossman for assistance with fieldwork, Clive Barlow for supplying a recording of the song of Black-headed Weavers, and Helder Cardoso, Vitor Encarnacão, the ICNF and Associação PATO for assisting with logistics in the field and providing access to field sites. Pim Edelaar, Hannah Mossman, and an anonymous reviewer provided comments that improved the quality of this manuscript. This study was supported by a Natural Environment Research Council PhD studentship and a BOU small research grant awarded to MJPS.

References

Adler, P.B., Ellner, S.P. & Levine, J.M. (2010) Coexistence of perennial plants: an embarrassment of niches. *Ecology Letters,* **13**, 1019-1029.

Adler, P. B., HilleRisLambers, J. & Levine, J. M. (2007) A niche for neutrality. *Ecology Letters*, **10**, 95-104.

Bauer, H.-G. & Woog, F. (2011) On the ‘invasiveness’ of non-native bird species. *Ibis*, **153**, 204-206.

Bensch, S. & Hasselquist, D. (1991) Territory infidelity in the polygynous great reed warbler *Aacrocephalus arundinaceus*: the effect of variation in territory attractiveness. *Journal of Animal Ecology*, **60**, 857-871.

Blackburn, T.M., Cassey, P., Duncan, R.P., Evans, K.L. & Gaston, K.J. (2004). Avian extinction and mammalian introductions on oceanic islands. *Science*, **305**, 1955-1958.

Blackburn, T.M., Lockwood, J.L. & Cassey, P. (2009) *Avian Invasions: The Ecology and Evolution on Exotic Birds*, Oxford University Press, Oxford.

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.J., Randin, C., Zimmermann, N.E., Graham, C.H. & Guisan, A. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, **21**, 481-497.

Cardoso, H. (2008) Ploceus melanocephalus. In *Atlas das Aves Nidificantes em Portugal*. (ed. Equipa Atlas), pp. 540-541. ICNB, Lisbon.

Catchpole, C.K. (1972) A comparative study of territory in the Reed warbler (*Acrocephalus scirpaceus*) and Sedge warbler (*A. schoenobaenus*). *Journal of Zoology*, **166**, 213-231.

Catchpole, C.K. (1978) Interspecific territorialism and competition in Acrocephalus warblers as revealed by playback experiments in areas of sympatry and allopatry. Animal Behaviour, **26**, 1072-1080.

Champely, S. (2007) *pwr: Basic functions for power analysis*. R package version 1.1.

Chesson, P. (2000) Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics*, **31**, 343-366.

Clavero, M., Brotons, L., Pons, P. & Sol, D. (2009) Prominent role of invasive species in avian biodiversity loss. *Biological Conservation*, **142**, 2043-2049.

Clarke, K.R. & Gorley, R.N. (2006) *PRIMER v6: User Manual/Tutorial*. PRIMER-E, Plymouth.

Clarke, K.R. & Warwick, R.M. (1994) *Change in marine communities: an approach to statistical analysis and interpretation*. Plymouth Marine Laboratory, Plymouth.

Constantine, M., van den Berg, A.B. & Robb, M. (2006)*The Sound Approach to Birding: A guide to understanding bird sound*, The Sound Approach, Poole.

Cramp, S. (ed.) (1992) *The Birds of the Western Palearctic*, Vol. 6. Oxford University Press, Oxford.

Defra (2003) *Review of non-native species policy: Report of the working group*. Department for Environment, Food and Rural Affairs, London.

Dunn, M., Copelston, M. & Workman, L. (2004) Trade-offs and seasonal variation in territorial defence and predator evasion in the European Robin *Erithacus rubecula*. *Ibis*, **146**, 77-84.

Dyrcz, A. (1986) Factors affecting facultative polygyny and breeding results in the Great Reed Warbler (*Acrocephalus arundinaceus*). *Journal of Ornithololology*, **127**, 447-461.

Edelaar, P.I.M. & Tella, J.L. (2012) Managing non-native species: don't wait until their impacts are proven. *Ibis*, **154**, 635-637.

EEA (2010) Towards an early warning and information system for invasive alien species (IAS) threatening biodiversity in Europe. In *EEA Technical report*. Vol. 5/2010, European Environment Agency, Luxembourg.

Equipa Atlas (2008) *Atlas das Aves Nidificantes em Portugal (1999-2005)*, Instituto da Conservacao da Natureza e da Biodiversividade, Sociedade Portuguesa para o Estudo das Aves, Parque Natural da Madeira e Secretaria Regional do Ambiente e do Mar, Lisbon.

ESRI (2008) *ArcGIS Desktop*. Environmental Systems Research Institute.

Fry, C.H. & Keith, S. (eds.) 2004. The Birds of Africa, London: Christopher Helm.

Golabek, K. A., Ridley, A.R. & Radford, A.N. (2012) Food availability affects strength of seasonal territorial behaviour in a cooperatively breeding bird. *Animal Behaviour*, **83**, 613-619.

Grarock, K., Tidemann, C.R., Wood, J. & Lindenmayer, D.B. (2012) Is It Benign or Is It a Pariah? Empirical Evidence for the Impact of the Common Myna (Acridotheres tristis) on Australian Birds. *PLoS ONE*, **7**, e40622.

Graveland, J. (1996) The decline of an aquatic songbird: The Great Reed Warbler Acrocephalus arundinaceus in the Netherlands. *Limosa*, **69**, 85-96.

Graveland, J. (1998) Reed die-back, water level management and the decline of the great reed warbler Acrocephalus arundinaceus in the Netherlands. *Ardea*, **86**, 187-202.

Griswold, M.W. & Lounibos, L.P. (2005) Does differential predation permit invasive and native mosquito larvae to coexist in Florida? *Ecological Entomology*, **30**, 122-127.

Hilton, G.M. & Cuthbert, R.J. (2010) The catastrophic impact of invasive mammalian predators on birds of the UK Overseas Territories: a review and synthesis. *Ibis*, **152,** 443-458.

Hothorn, T., Bretz, F. & Westfall, P. (2008) Simultaneous Inference in General Parametric Models. *Biometrical Journal*, **50**, 346-363.

Ieronymidou, C., Collar, N.J. & Dolman, P.M. (2012) Endemic Cyprus Warbler *Sylvia melanothorax* and colonizing Sardinian Warbler *Sylvia melanocephala* show different habitat associations. *Ibis*, **154**, 248-259.

Instituto Geográfico Português (2004). *AirPhoto 2004*. Lisbon.

Koch, A.J., Martin, K. & Aitken, K.E.H. (2012) The relationship between introduced European Starlings and the reproductive activities of Mountain Bluebirds and Tree Swallows in British Columbia, Canada. *Ibis*, **154**, 590-600.

Koenig, W.D. (2003) European Starlings and Their Effect on Native Cavity-Nesting Birds. *Conservation Biology*, **17**, 1134-1140.

Kumschick, S. & Nentwig, W. (2010) Some alien birds have as severe an impact as the most effectual alien mammals in Europe. *Biological Conservation*, **143**, 2757-2762.

Leisler, B. & Schulze-Hagen, K. (2011) *The Reed Warblers: Diversity in a uniform bird family*, KNNV Publishing, Utrecht.

Lodge, D.M., Williams, S., MacIsaac, H.J., Hayes, K.R., Leung, B., Reichard, S., Mack, R.N., Moyle, P.B., Smith, M., Andow, D.A., Carlton, J.T. & McMichael, A. (2006) Biological invasions: Recommendations for US policy and management. *Ecological Applications*, **16**, 2035-2054.

Marchant, J.H. (1983) *BTO Common Birds Census Instructions*, BTO, Thetford.

Marshall, M.R. & Cooper, R.J. (2004) Territory size of a migratory songbird in response to caterpillar density and foliage structure. *Ecology*, **85**, 432-445.

Marti, R. & de Moral, J.C. (eds.) (2003) *Atlas de las Aves Reproductoras de Espana*, Dirección General de Conservación de la Naturaleza-Sociedad Española de Ornithologíam, Madrid.

Martinez-Vilalta, J., Bertolero, A., Bigas, D., Paquet, J.-Y. & Martinez-Vilalta, A. (2002) Habitat selection of the passerine community nesting at the Ebro delta reed-beds (NE Spain): some management implications. *Wetlands*, **22**, 318-325.

Matias, R. (2002) *Aves exóticas que nidificam em Portugal continental*, ICNB, Lisbon.

Mikami, O.K., Kohda, M. & Kawata, M. (2004). A new hypothesis for species coexistence: male-male repulsion promotes coexistence of competing species. *Population Ecology*, **46**, 213-217.

Moreau, R.E. (1960) Conspectus and classification of the Ploceine weaver-birds. *Ibis*, **102,** 298-321.

Newson, S.E., Johnston, A., Parrott, D. & Leech, D.I. (2011) Evaluating the population-level impact of an invasive species, Ring-necked Parakeet *Psittacula krameri*, on native avifauna. *Ibis*, **153**, 509-516.

Newton, I. (1994) The role of nest sites in limiting the numbers of hole-nesting birds - a review. *Biological Conservation*, **70**, 265-276.

Poulin, B., Lefebvre, G. & Mauchamp, A. (2002) Habitat requirements of passerines and reedbed management in southern France. *Biological Conservation*,**107**, 315-325.

R Development Core Team (2012) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna.

Robbins, C.S. (1981). Effect of time of day on bird activity. In Ralph, C.J. & Scott, J.M. *Estimating Numbers of Terrestrial Birds*. Cooper Square Press, Lanham, USA.

Roche, J.-C. (1997)*The Bird Songs and Calls of Britain and Europe*, Sittelle, Mens.

Schoener, T.W. (1970) Nonsynchronous spatial overlap of lizards in patchy habitats. *Ecology*, **51**, 408-418.

Sixth Conference of the Parties Convention on Biological Diversity (2002) *Sixth Ordinary Meeting of the Conference of the Parties to the Convention on Biological Diversity*. The Hague, Netherlands.

Strubbe, D. & Matthysen, E. (2009) Experimental evidence for nest-site competition between invasive ring-necked parakeets (*Psittacula krameri*) and native nuthatches (*Sitta europaea*). *Biological Conservation*,**142**, 1588-1594.

Strubbe, D., Shwartz, A. & Chiron, F. (2011) Concerns regarding the scientific evidence informing impact risk assessment and management recommendations for invasive birds. *Biological Conservation*,**144**, 2112-2118.

Sullivan, M.J.P., Grundy, J. & Franco, A.M.A. In press. Assessing the impacts of the non-native Black-headed Weaver on native *Acrocephalus* warblers. *Ibis*.

Unep (1992) *Rio Declaration on Enviornment and Development*. United Nations Environment Programme.

Ventim, R., Mendes, L., Ramos, J., Cardoso, H. & Pérez-Tris, J. (2012) Local haemoparasites in introduced wetland passerines. *Journal of Ornithology*, **153**, 1253-1259.

Wang, Y., Naumann, U., Wright, S. T. & Warton, D. I. (2012) mvabund– an R package for model-based analysis of multivariate abundance data. *Methods in Ecology and Evolution*, **3**, 471-474.

Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, **62**, 2868-2883.

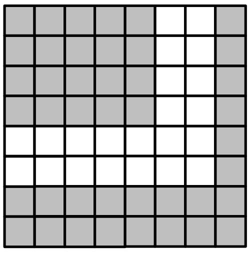
Weitzel, N.H. (1988) Nest-Site competition between the European Starling and native breeding birds in northwestern Nevada. *The Condor*, **90**, 515-517.

Wiens, J.A. (1989) *The Ecology of Bird Communities*, Cambridge University Press, Cambridge.

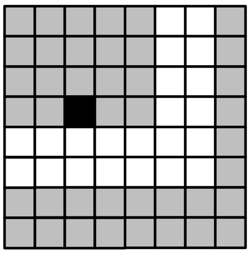
7.5 Appendix 7-1. Details of the method used to generate simulated territories.

We calculated territory overlap under a null model in which we randomly shifted the position of each territory, while maintaining territory size and restricting territories to be in reedbed. Simulated territories were constructed using the following procedure:

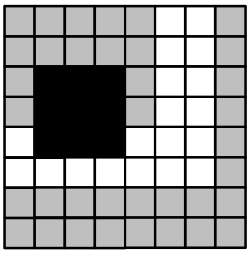
1. We created a raster map of the extent of reedbed at each site, with each grid cell being classified as containing or not containing reedbed. In this demonstration map reedbed-containing grid cells are shown in grey and grid cells that do not contain reedbed are shown in white.



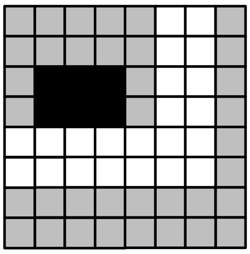
2. For each observed territory, we created a simulated territory of the same size. In this demonstration we will create a simulated territory containing 10 grid cells. The location of each simulated territory was determined by randomly selecting a raster grid cell (black square), with the condition that it contained reedbed, which formed the centre of the simulated territory. If the grid cell did not contain reedbed, a new grid cell was selected.



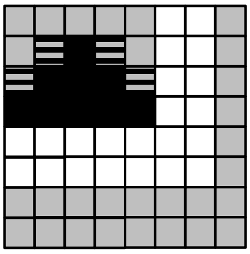
3. We then assigned the remaining nine grid cells to this simulated territory. In extensive wetlands, many territories were approximately square shaped, so we started off by forming the largest possible square given the size of the territory. Note that there is one grid cell remaining that needs to be assigned to this territory.



4. We removed any grid cells that did not contain reedbed from the territory.

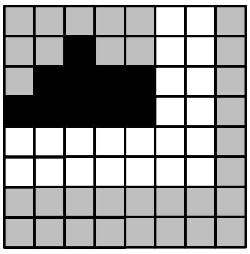


5. This territory now needs four more grid cells to be assigned to it. The closest reedbed-containing grid cells to the territory centroid that were not already part of the territory were assigned to the territory. In this case, there were three reedbed-containing grid cells the same distance away from the territory centroid, and four more grid cells one unit of distance further away. We assigned the

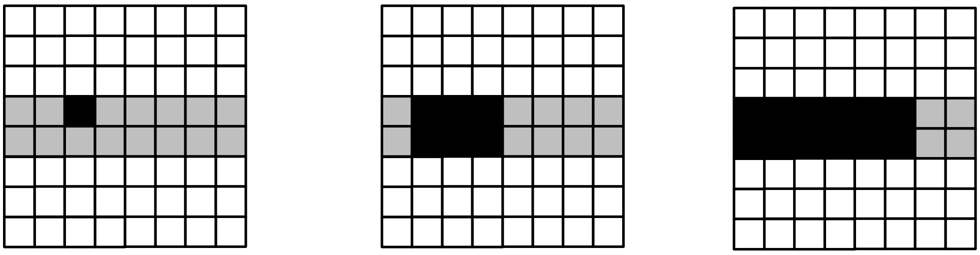


three closest grid cells to the territory, but could assign any one of the four next closest grid cells to the territory (hatched squares).

6. In such cases we randomly assigned one of these grid cells to the simulated territory.



Observed territories in ditch-crossed sites were long and narrow. Note how our procedure for making simulated territories produces long narrow territories when reedbeds are narrow linear features (e.g. along ditches).



8 Concluding remarks

Non-native species have caused enormous economic and environmental damage. For example, mammalian predators introduced to oceanic islands have driven some native bird species to extinction (Blackburn *et al.* 2004), while introduced grey squirrels *Sciurus carolinensis* have caused £10 million damage to Great Britain's timber industry (GB non-native species secretariat 2013)There are concerns that climate and land-use change could exacerbate the potential impact of non-native species by increasing their potential distribution (Sorte, Williams & Zerebecki 2010; Polce *et al.* 2011; Bellard *et al.* 2013), changing species interactions (Sorte *et al.* 2013) and increasing the rate of new introductions by changing patterns of demand in the horticultural trade (Bradley *et al.* 2012).

Species distribution models (SDMs) of the potential distribution of non-native species are a key tool for assessing the current and future threat from non-native species (Peterson 2003; Thuiller *et al.* 2005). Despite their widespread use (Elith & Leathwick 2009), the development of new techniques (Elith *et al.* 2006; Hijmans *et al.* 2011) and frequent papers criticising aspects of SDMs (Guisan & Thuiller 2005; Araújo & Guisan 2006; Jiménez-Valverde, Lobo & Hortal 2008; Vallecillo, Brotons & Thuiller 2009; Barbet-Massin, Thuiller & Jiguet 2010; Gallien *et al.* 2010), a number of issues need to be addressed in the development of SDMs. I will focus here on five issues that I consider to be important to the future development of the field. These are the challenges of modelling species distributions that are not at equilibrium with their environment, the ability of SDMs to transfer to new environments, difficulties in constructing biologically meaningful SDMs, challenges presented by non-ideal input data and issues with uncertainty in SDMs when providing management recommendations.

8.1 Modelling non-equilibrium distributions

Some of the most important applications of SDMs are to predict the potential distribution of range expanding and spreading non-native species. These species violate the assumption that a species is at equilibrium with their environment. Proposed solutions to this problem include delimiting the background area for modelling to just include areas the species can disperse to (Elith, Kearney & Phillips 2010) and using a SDM trained in areas where a species is in equilibrium with their environment (e.g. the native range or the range core) to identify absences that are likely to be due to dispersal limitation (Gallien *et al.* 2012). In chapter two I developed a flexible method to account for dispersal limitation (Sullivan *et al.* 2012). This involved constructing a dispersal model and using it to downweight grid-cells that are unlikely to be dispersed to. This method improved SDM performance, is preferable to Gallien *et al.* (2012) as it includes the biological process of dispersal limitation and has the advantage over Elith *et al.* (2010) in that it quantifies dispersal limitation and recognises that dispersal is a probabilistic process. One issue with the method presented in Sullivan *et al.* (2012) is that the dispersal model is parameterised in a separate stage to the SDM. It would be preferable to simultaneously parameterise the models. A method for doing this is presented in Appendix 8-1.

I showed that the environmental associations of species change through space (chapter three) and time (chapter five). The dispersal ability of species can also vary between the range core and range margin (Phillips, Brown & Shine 2010). These studies indicate that a move from static SDMs to dynamic models that incorporate spatial and temporal non-stationarity in parameters is desirable. Dynamic models of species dispersal have been developed (Palmer, Coulon & Travis 2011; Bocedi *et al.* 2012), which can be combined with integrated dispersal-species distribution models (Sullivan *et al.* 2012) to create dynamic models that are well equipped to deal with the changing nature of species distributions.

8.2 Transferring models to new environments

Predicting the potential distribution of non-native species or native species under environmental change often involves projecting a SDM onto data outside the range encountered in the training dataset. Issues with such extrapolation are well documented (Zurell, Elith & Schroder 2012). In chapter four I extended this work to predict when SDMs should show good discriminative performance in the non-native range. Much attention has been paid to assessing whether statistical relationships should transfer to new environments (Wenger & Olden 2012). However, this ignores another important question, which is whether the mechanisms that limit species distributions in one area still limit the distribution in another. In chapter four I identify areas where the distribution of a species is likely to limited by unmodelled or incorrectly modelled variables, but further work is needed in this area.

8.3 Capturing variables that limit species distributions

Correlative SDMs can accurately describe species distributions (Elith *et al.* 2006). However, so can random spatially structured variables (Bahn & McGill 2007), indicating that variables can explain species distributions without there being mechanistic links. This had led to considerable debate about the extent to which large-scale environmental variables limit species distributions (Beale, Lennon & Gimona 2008; Jimenez-Valverde *et al.* 2011). It should be possible to identify whether SDMs contain meaningful ecological information by testing their performance on independent test data. One way of doing this is retrodiction, where a SDM is constructed using the past distribution of a species and projected to predict the current distribution. Good SDM performance in retrodiction exercises supports SDMs containing meaningful ecological information (Rapacciuolo *et al.* 2012; Rodríguez-Rey, Jiménez-Valverde & Acevedo 2013). However, this is largely a result of good SDM performance in areas occupied in the first time period (Rapacciuolo *et al.* 2012), with SDMs performing no better at identifying newly occupied grid-cells that either chance (Rapacciuolo *et al.* 2012) or simple dispersal models (Rodríguez-Rey, Jiménez-Valverde & Acevedo 2013).

SDMs are vulnerable to identifying variables that happen to have higher or lower values inside a species distribution than outside it as important for limiting species distributions. The risk of this happening is higher when species distributions are spatially structured by unmodelled factors such as dispersal (Sullivan *et al.* 2012), or when environmental variables are included in SDMs with little thought for their likely mechanistic role (Rodder & Lotters 2009); for example including densities of pensioners as a predictor of monk parakeet occurrence (Rodriguez-Pastor *et al.* 2012). Good practice to increase the probability of SDMs only containing variables that mechanistically limit a species distribution is to carefully select predictor variables that are likely to limit a species distribution (Elith & Leathwick 2009), account for processes that spatially structure species distributions (Sullivan *et al.* 2012) or account for spatial autocorrelation (Václavík, Kupfer & Meentemeyer 2012), and test models on truly independent datasets (i.e. breaking local patterns of spatial autocorrelation) (Wenger & Olden 2012).

These correlative SDMs can still be criticised for their failure to explicitly incorporate mechanisms such as species interactions that limit species distributions (Schmitz *et al.* 2003; Higgins, O'Hara & Romermann 2012). It is possible to incorporate mechanistic processes in hybrid SDMs (Higgins, O'Hara & Romermann 2012), however, identifying them remains challenging. Analyses that relate time series of population trends to variation in resources and climate can disentangle different factors affecting populations (Pearce-Higgins *et al.* 2010). However, lagged responses to climate extremes will not be apparent in these time series, reducing the extrapolative performance of such models under sustained directional forcing (Sullivan *et al.* in prep.). The analysis of datasets that vary both in time and space possibly offer the best opportunity to disentangle the roles of different factors that act on species populations and distributions.

8.4 Overcoming imperfect input data

SDMs are typically constructed at coarse spatial resolutions (grid-cells of 1km^2 or more) due to the availability of occurrence and environmental data (Vaughan & Ormerod 2003). However, species may only occupy a small portion of a grid-cell (Araujo *et al.* 2005), and are perhaps limited by fine-scale variation in environmental variables (Suggitt *et al.* 2011). While this problem is best solved by obtaining fine-scale distribution data, the development of statistical techniques that reduce regression dilution are promising (McInerny & Purves 2011). Mobile species will more frequently occupy some areas more than others. When species can be tagged their utilisation distributions can be obtained, and used as input for SDMs (Carroll *et al.* 2013)

Distribution data from biological atlases are subject to uneven detection probabilities (Lobo, Jiménez-Valverde & Hortal 2010), which can result in underestimates of species prevalence and, more seriously, failure to correctly characterise species environmental associations (Kéry 2011). The latter problem can be reduced by including recorder effort as a covariate (chapter two) or using hierarchical models (Kéry 2011). These methods need to be more widely adopted by the species distribution modelling community.

8.5 Uncertainty and policy relevance

The dominant output of SDMs is a probabilistic map showing where a species is likely to occur. Thresholds are sometimes used to convert these probabilities to produce a binary map of suitable and unsuitable areas (Jiménez-Valverde & Lobo 2007). Such maps are useful for policy makers as they simply show which areas are suitable for a species, but they hide the uncertainty associated with these predictions. Uncertainty due to model extrapolation is sometimes quantified and mapped (Zurell, Elith & Schroder 2012), and I present a method for doing this in chapter four. There is considerable potential to map uncertainty arising from other aspects of model development, such as predictor variable selection, and map areas that are suitable in all modelling scenarios, suitable only in some modelling scenarios and unsuitable in all modelling scenarios. Developing methods to quantify uncertainty is important to ensure uncertainty in SDM predictions is propagated in derived analyses (e.g. reserve selection) (Moilanen *et al.* 2005). However, increasing the uncertainty in mapped outputs decreases the ease in which they are interpreted by policy makers. The domination of maps in the discourse of species distribution modelling has led to other outputs, such as the quantification of the species-environment relationship and qualitative descriptions of species response to environmental change, to be somewhat neglected. Some of these simpler outputs will have less uncertainty than mapped outputs, but may provide sufficient information to inform policy makers.

8.6 Impacts and management

While many non-native species have dramatic negative impacts, others do not. This has prompted debate amongst ecologists as to whether it is necessary for impacts to be demonstrated before non-native species are controlled (Bauer & Woog 2011; Davis *et al.* 2011; Simberloff 2011; Edelaar & Tella 2012; Thomas 2013). The argument on one side of the debate is that control of non-native species is easiest when they are newly established, and associated ethical issues minimised, so it is justified to take a precautionary approach and routinely control newly established non-native species (Edelaar & Tella 2012). The argument on the other side of the debate is that species should not be judged by their origins as many non-native species do not have negative impacts (Davis *et al.* 2011; Thomas 2013). There is an implicit assumption in this debate that managers have sufficient resources to control all non-native species if necessary. This is unlikely, so the goal of ecologists should be shifted from deciding whether to control a species or not to prioritising which species should be controlled given available resources.

One barrier to providing useful management information is the tendency for ecologists to conduct impact assessments on long-established species (Tanner *et al.* 2013), when identifying the early impacts of non-native species is more useful as it is more feasible to control these species (Lodge *et al.* 2006). I developed a framework for investigating the early impacts of non-native species (chapter seven). The main issue with attempting to investigate the early impacts of non-native species is a lack of statistical power. However, species with large impacts can be detected early on, while by pragmatically interpreting the results of early impact studies they can be used to help prioritise management actions.

Another way ecologists can provide management relevant information is by predicting the impact of non-native species before they are introduced. A number of trait based risk assessments have been constructed (Leung *et al.* 2012; Kumschick & Richardson 2013), a further development of these to incorporate taxa-specific, impact-specific and context dependent variation in impact will make these a powerful tool in the future.

Ecologists may avoid investigating the early impacts of non-native species due to the risks of not finding significant results and subsequent difficulties getting published. The establishment of a journal dedicated to documenting impact assessments of non-native species, in a similar vein to Conservation Evidence, may help persuade ecologists to undertake more policy relevant science.

8.7 Future directions

Considerable progress has been made in developing methodologies to predict the potential distribution and impacts of non-native species (Václavík & Meentemeyer 2009; Bradley *et al.* 2010; Magee *et al.* 2010; Sullivan *et al.* 2012; Václavík & Meentemeyer 2012; Kumschick, Bacher & Blackburn 2013; Kumschick & Richardson 2013; Sullivan, Grundy & Franco In Press). Large scale multi-taxa studies have been able to explore how global change will affect non-native species (e.g. Bellard *et al.* 2013). However, providing a generalised assessment of the effects of global change on non-native species using current tools is a daunting prospect. It is possible to take a step back from SDMs and look at patterns of climatic similarity; this approach has been used to identify areas of climatic endemism and future rarity (Ohlemüller *et al.* 2006). A similar approach could be applied to predict changes in invasion risk under climate change. I hypothesise that if climate change increases the global similarity of a location’s climate it will increase its invasion risk, as the pool of potential invaders is larger. If this hypothesis is supported it would allow a global assessment of the effect of climate change on invasion risk.

In chapter six I found that a novel land-use facilitated an avian invasion by providing underexploited food resources. Future studies should test the generality of this result. Future land-use change could lead to non-random extinctions (Fritz, Bininda-Emonds & Purvis 2009), potentially providing vacant functional space for new non-native species. The effect of global change on the functional composition of communities (Díaz *et al.* 2013) and their subsequent resistance to invasion (Levine, Adler & Yelenik 2004) needs further study.

Bradley *et al.* (2012) highlighted the how emerging trade partnerships can increase the risk of new non-native species becoming established. More work is needed to link patterns of trade connectivity with environmental suitability to identify the extent to which trade-connectivity has limited the number of non-native species a location receives, and to identify areas that would receive more non-native species under shifting trade patterns. This could be combined with work described above investigating the roles of climate similarity and native species trait composition at influencing non-native species richness to create a powerful model of invasion risk that could be applied at a global scale.

Advances in our ability to predict the distribution and impacts of non-native species can be made at the local scale as well. Presently, maps of potential impact are rare. Integrating context dependent assessments of impact with dynamic species distribution models would allow the production spatially explicit maps of potential impact. Producing these models will require further development of species distribution modelling and impact assessment methods. In addition to the potential developments in species distribution modelling mentioned earlier, further work has to be carried out to develop an understanding of how the environmental context of a non-native population affects its impacts.

Horizon scanning exercises using models of the potential distribution and impacts of non-native species can be used to identify potentially damaging species before they are introduced. Doing this will mean that scientific evidence for the importance of controlling non-native species can be provided at the start of the invasion process.

References

Araujo, M.B., Thuiller, W., Williams, P.H. & Reginster, I. (2005) Downscaling European species atlas distributions to a finer resolution: implications for conservation planning. *Global Ecology and Biogeography,* **14,** 17-30.

Araújo, M.B. & Guisan, A. (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography,* **33,** 1677-1688.

Bahn, V. & McGill, B.J. (2007) Can niche-based distribution models outperform spatial interpolation? *Global Ecology and Biogeography,* **16,** 733-742.

Barbet-Massin, M., Thuiller, W. & Jiguet, F. (2010) How much do we overestimate future local extinction rates when restricting the range of occurrence data in climate suitability models? *Ecography,* **33,** 878-886.

Bauer, H.-G. & Woog, F. (2011) On the ‘invasiveness’ of non-native bird species. *Ibis,* **153,** 204-206.

Beale, C.M., Lennon, J.J. & Gimona, A. (2008) Opening the climate envelope reveals no macroscale associations with climate in European birds. *Proceedings of the National Academy of Sciences,* **105,** 14908-14912.

Bellard, C., Thuiller, W., Leroy, B., Genovesi, P., Bakkenes, M. & Courchamp, F. (2013) Will climate change promote future invasions? *Global Change Biology*.

Blackburn, T.M., Cassey, P., Duncan, R.P., Evans, K.L. & Gaston, K.J. (2004) Avian extinction and mammalian introductions on oceanic islands. *Science,* **305,** 1955-1958.

Bocedi, G., Pe’er, G., Heikkinen, R.K., Matsinos, Y. & Travis, J.M.J. (2012) Projecting species’ range expansion dynamics: sources of systematic biases when scaling up patterns and processes. *Methods in Ecology and Evolution*, **3,** 1008-1018.

Bradley, B.A., Blumenthal, D.M., Early, R., Grosholz, E.D., Lawler, J.J., Miller, L.P., Sorte, C.J.B., D'Antonio, C.M., Diez, J.M., Dukes, J.S., Ibanez, I. & Olden, J.D. (2012) Global change, global trade, and the next wave of plant invasions. *Frontiers in Ecology and the Environment,* **10,** 20-28.

Bradley, B.A., Blumenthal, D.M., Wilcove, D.S. & Ziska, L.H. (2010) Predicting plant invasions in an era of global change. *Trends in Ecology & Evolution,* **25,** 310-318.

Carroll, M., Ewing, S., Owen, E. & Bolton, M. (2013) Modelling the impacts of climate change on seabirds in the UK. *European Ornithologists' Union Conference 2013*. Norwich.

Davis, M.A., Chew, M.K., Hobbs, R.J., Lugo, A.E., Ewel, J.J., Vermeij, G.J., Brown, J.H., Rosenzweig, M.L., Gardener, M.R., Carroll, S.P., Thompson, K., Pickett, S.T.A., Stromberg, J.C., Tredici, P.D., Suding, K.N., Ehrenfeld, J.G., Philip Grime, J., Mascaro, J. & Briggs, J.C. (2011) Don't judge species on their origins. *Nature,* **474,** 153-154.

Díaz, S., Purvis, A., Cornelissen, J.H.C., Mace, G.M., Donoghue, M.J., Ewers, R.M., Jordano, P. & Pearse, W.D. (2013) Functional traits, the phylogeny of function, and ecosystem service vulnerability. *Ecology and Evolution,* **3,** 2958-2975.

Edelaar, P.I.M. & Tella, J.L. (2012) Managing non-native species: don’t wait until their impacts are proven. *Ibis,* **154,** 635-637.

Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberon, J., Williams, S., Wisz, M.S. & Zimmermann, N.E. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography,* **29,** 129-151.

Elith, J., Kearney, M. & Phillips, S. (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution,* **1,** 330-342.

Elith, J. & Leathwick, J.R. (2009) Species Distribution Models: Ecological explanation and prediction across space and Time. *Annual Review of Ecology Evolution and Systematics*, pp. 677-697. Annual Reviews, Palo Alto.

Fritz, S.A., Bininda-Emonds, O.R.P. & Purvis, A. (2009) Geographical variation in predictors of mammalian extinction risk: big is bad, but only in the tropics. *Ecology Letters,* **12,** 538-549.

Gallien, L., Douzet, R., Pratte, S., Zimmermann, N.E. & Thuiller, W. (2012) Invasive species distribution models - how violating the equilibrium assumption can create new insights. *Global Ecology and Biogeography,* **21,** 1126-1136.

Gallien, L., Münkemüller, T., Albert, C.H., Boulangeat, I. & Thuiller, W. (2010) Predicting potential distributions of invasive species: where to go from here? *Diversity and Distributions,* **16,** 331-342.

GB non-native species secretariat (2013) http://www.nonnativespecies.org/home/index.cfm.

Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters,* **8,** 993-1009.

Higgins, S.I., O'Hara, R.B. & Romermann, C. (2012) A niche for biology in species distribution models. *Journal of Biogeography,* **39,** 2091-2095.

Hijmans, R.J., Phillips, S., Leathwick, J.R. & Elith, J. (2011) Package 'dismo'. Available online at: http://cran.r-project.org/web/packages/dismo/index.html.

Jimenez-Valverde, A., Barve, N., Lira-Noriega, A., Maher, S.P., Nakazawa, Y., Papes, M., Soberon, J., Sukumaran, J. & Peterson, A.T. (2011) Dominant climate influences on North American bird distributions. *Global Ecology and Biogeography,* **20,** 114-118.

Jiménez-Valverde, A. & Lobo, J.M. (2007) Threshold criteria for conversion of probability of species presence to either–or presence–absence. *Acta Oecologica,* **31,** 361-369.

Jiménez-Valverde, A., Lobo, J.M. & Hortal, J. (2008) Not as good as they seem: the importance of concepts in species distribution modelling. *Diversity and Distributions,* **14,** 885-890.

Kumschick, S., Bacher, S. & Blackburn, T.M. (2013) What determines the impact of alien birds and mammals in Europe? *Biological Invasions,* **15,** 785-797.

Kumschick, S. & Richardson, D.M. (2013) Species-based risk assessments for biological invasions: advances and challenges. *Diversity and Distributions,* **19,** 1095-1105.

Kéry, M. (2011) Towards the modelling of true species distributions. *Journal of Biogeography,* **38,** 617-618.

Leung, B., Roura-Pascual, N., Bacher, S., Heikkila, J., Brotons, L., Burgman, M.A., Dehnen-Schmutz, K., Essl, F., Hulme, P.E., Richardson, D.M., Sol, D. & Vila, M. (2012) TEASIng apart alien species risk assessments: a framework for best practices. *Ecology Letters,* **15,** 1475-1493.

Levine, J.M., Adler, P.B. & Yelenik, S.G. (2004) A meta-analysis of biotic resistance to exotic plant invasions. *Ecology Letters,* **7,** 975-989.

Lobo, J.M., Jiménez-Valverde, A. & Hortal, J. (2010) The uncertain nature of absences and their importance in species distribution modelling. *Ecography,* **33,** 103-114.

Lodge, D.M., Williams, S., MacIsaac, H.J., Hayes, K.R., Leung, B., Reichard, S., Mack, R.N., Moyle, P.B., Smith, M., Andow, D.A., Carlton, J.T. & McMichael, A. (2006) Biological invasions: Recommendations for US policy and management. *Ecological Applications,* **16,** 2035-2054.

Magee, T.K., Ringold, P.L., Bollman, M.A. & Ernst, T.L. (2010) Index of alien impact: a method for evaluating potential ecological impact of alien plant species. *Environmental Management,* **45,** 759-778.

McInerny, G.J. & Purves, D.W. (2011) Fine-scale environmental variation in species distribution modelling: regression dilution, latent variables and neighbourly advice. *Methods in Ecology and Evolution,* **2,** 248-257.

Moilanen, A., Franco, A.M.A., Early, R.I., Fox, R., Wintle, B. & Thomas, C.D. (2005) Prioritizing multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proceedings of the Royal Society B: Biological Sciences,* **272,** 1885-1891.

Ohlemüller, R., Gritti, E.S., Sykes, M.T. & Thomas, C.D. (2006) Towards European climate risk surfaces: the extent and distribution of analogous and non-analogous climates 1931–2100. *Global Ecology and Biogeography,* **15,** 395-405.

Palmer, S.C.F., Coulon, A. & Travis, J.M.J. (2011) Introducing a ‘stochastic movement simulator’ for estimating habitat connectivity. *Methods in Ecology and Evolution,* **2,** 258-268.

Pearce-Higgins, J.W., Dennis, P., Whittingham, M.J. & Yalden, D.W. (2010) Impacts of climate on prey abundance account for fluctuations in a population of a northern wader at the southern edge of its range. *Global Change Biology,* **16,** 12-23.

Peterson, A.T. (2003) Predicting the geography of species' invasions via ecological niche modelling. *Quarterly Review of Biology,* **78,** 419-433.

Phillips, B.L., Brown, G.P. & Shine, R. (2010) Evolutionarily accelerated invasions: the rate of dispersal evolves upwards during the range advance of cane toads. *Journal of Evolutionary Biology,* **23,** 2595-2601.

Polce, C., Kunin, W.E., Biesmeijer, J.C., Dauber, J., Phillips, O.L. & Network, A.F.S. (2011) Alien and native plants show contrasting responses to climate and land use in Europe. *Global Ecology and Biogeography,* **20,** 367-379.

Rapacciuolo, G., Roy, D.B., Gillings, S., Fox, R., Walker, K. & Purvis, A. (2012) Climatic Associations of British Species Distributions Show Good Transferability in Time but Low Predictive Accuracy for Range Change. *PLoS ONE,* **7,** e40212.

Rodder, D. & Lotters, S. (2009) Niche shift versus niche conservatism? Climatic characteristics of the native and invasive ranges of the Mediterranean house gecko (Hemidactylus turcicus). *Global Ecology and Biogeography,* **18,** 674-687.

Rodriguez-Pastor, R., Senar, J.C., Ortega, A., Faus, J., Uribe, F. & Montalvo, T. (2012) Distribution patterns of invasive Monk parakeets (Myiopsitta monachus) in an urban habitat. *Animal Biodiversity and Conservation,* **35,** 107-117.

Rodríguez-Rey, M., Jiménez-Valverde, A. & Acevedo, P. (2013) Species distribution models predict range expansion better than chance but not better than a simple dispersal model. *Ecological Modelling,* **256,** 1-5.

Schmitz, O.J., Post, E., Burns, C.E. & Johnston, K.M. (2003) Ecosystem responses to global climate change: Moving beyond color mapping. *Bioscience,* **53,** 1199-1205.

Simberloff, D. (2011) Non-natives: 141 scientists object. *Nature,* **475,** 36.

Sorte, C.J.B., Ibanez, I., Blumenthal, D.M., Molinari, N.A., Miller, L.P., Grosholz, E.D., Diez, J.M., D'Antonio, C.M., Olden, J.D., Jones, S.J. & Dukes, J.S. (2013) Poised to prosper? A cross-system comparison of climate change effects on native and non-native species performance. *Ecology Letters,* **16,** 261-270.

Sorte, C.J.B., Williams, S.L. & Zerebecki, R.A. (2010) Ocean warming increases threat of invasive species in a marine fouling community. *Ecology,* **91,** 2198-2204.

Suggitt, A.J., Gillingham, P.K., Hill, J.K., Huntley, B., Kunin, W.E., Roy, D.B. & Thomas, C.D. (2011) Habitat microclimates drive fine-scale variation in extreme temperatures. *Oikos,* **120,** 1-8.

Sullivan, M.J.P., Davies, R.G., Reino, L. & Franco, A.M.A. (2012) Using dispersal information to model the species–environment relationship of spreading non-native species. *Methods in Ecology and Evolution,* **3,** 870-879.

Sullivan, M.J.P., Grundy, J. & Franco, A.M.A. (2014) Assessing the impacts of the non-native Black-headed Weaver on native *Acrocephalus* warblers. *Ibis*, **156**, 231-232

Tanner, R.A., Varia, S., Eschen, R., Wood, S., Murphy, S.T. & Gange, A.C. (2013) Impacts of an invasive non-native annual weed, *Impatiens glandulifera*, on above- and below-ground invertebrate communities in the United Kingdom. *PLoS ONE,* **8,** e67271.

Thomas, C.D. (2013) The Anthropocene could raise biological diversity. *Nature,* **502,** 7-7.

Thuiller, W., Richardson, D.M., Pysek, P., Midgley, G.F., Hughes, G.O. & Rouget, M. (2005) Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology,* **11,** 2234-2250.

Vallecillo, S., Brotons, L. & Thuiller, W. (2009) Dangers of predicting bird species distributions in response to land-cover changes. *Ecological Applications,* **19,** 538-549.

Vaughan, I.P. & Ormerod, S.J. (2003) Improving the quality of distribution models for conservation by addressing shortcomings in the field collection of training data. *Conservation Biology,* **17,** 1601-1611.

Václavík, T., Kupfer, J.A. & Meentemeyer, R.K. (2012) Accounting for multi-scale spatial autocorrelation improves performance of invasive species distribution modelling (iSDM). *Journal of Biogeography,* **39,** 42-55.

Václavík, T. & Meentemeyer, R.K. (2009) Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? *Ecological Modelling,* **220,** 3248-3258.

Václavík, T. & Meentemeyer, R.K. (2012) Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. *Diversity and Distributions,* **18,** 73-83.

Wenger, S.J. & Olden, J.D. (2012) Assessing transferability of ecological models: an underappreciated aspect of statistical validation. *Methods in Ecology and Evolution,* **3,** 260-267.

Zurell, D., Elith, J. & Schroder, B. (2012) Predicting to new environments: tools for visualizing model behaviour and impacts on mapped distributions. *Diversity and Distributions,* **18,** 628-634.

8.8 Appendix 8-1

Hierarchical models that estimate the probability a grid-cell has been dispersed to and the probability the grid-cell is suitable as a latent variables can be used to simultaneously parameterise dispersal and species distribution models.

The probability a grid-cell has been dispersed to can be estimated as $z_i = e^{-d/a}$, where d is the distance of the grid-cell to the nearest grid-cell occupied in the previous time-step and a is a scale parameter of the negative exponential distribution. Other distributions (e.g. Gaussian) can be used for the dispersal kernel. The probability a grid-cell being suitable is found by a standard logistic regression model of the form $\mathrm{logit}(\Psi_i) = \alpha + \beta.x_i$, where α and β are parameters for the intercept and slope respectively. The occupancy status of a grid-cell is the product of the probability of dispersal and the probability of occupancy, $y_i \sim \mathrm{Bernoulli}\ (z_i \times \Psi_i)$. This model can be parameterised in a Bayesian framework.