Complex systems analysis of invasive species in heterogeneous environments

---

A thesis
submitted in partial fulfilment
of the requirements for the Degree of
Doctor of Philosophy
at
Lincoln University

by

Audrey Lustig

Lincoln University, New Zealand
2016

To my parents,

For being so loving and supportive.

– A.L.

Abstract

Given that so many ecosystems currently face major habitat change conducive to the rising global spread of invasive species, there is growing awareness of the need to adopt proactive management strategies. Among models most used to predict future changes in distribution of invasive species, few explicitly incorporate characteristic of the population dynamics at the invasion front and the spatial heterogeneity of the environment. In particular, the influence of landscape composition and configuration on population dynamics and ecosystem susceptibility to invasion, remain dependant on broad generalization.

The purpose of this study was to investigate how biological characteristics of invasive species interact with the structure of the landscape, to determine establishment and spread success. Critical to this research was the development of a unique spatially-explicit model that allowed for a systemic investigation of the impact of landscape structure on population dynamics of a species. The modelling framework has three components,1) a spatially-explicit, individual-based dispersal simulation framework, 2) a landscape generator allowing independent change in the composition and configuration of landscape components, and, 3) appropriate landscape measures that establish a quantitative relationship between demography, dispersal and the environment. The framework allowed a shift in focus from an individual species, to a more general approach where the pattern of invasion over multiple species and landscape scenarios were used to infer key drivers of invasion.

To identify appropriate landscape measures for this research, a multi-scale analysis of widely used landscape metrics was carried out. That analysis highlighted that landscape metrics are sensitive to complex interactions between the intrinsic characteristics of a landscape, and scale-dependent factors, making it difficult to isolate landscape pattern driven effects from the effects of changing spatial scale. As a solution, the self-organising map (SOM) clustering approach is proposed as an efficient way to disentangle the relationships among landscape metrics and spatial scale when accurate characterization of landscape pattern is a key input in spatially explicit ecological models.

The investigation of the effect of landscape structure on the establishment and spread of invasive species showed that both population density and rate of spread are affected in significant ways, and sometimes interactively, by landscape based components such as, suitable habitat amount, habitat patch aggregation, core area, edge density and habitat shape complexity. A key result of this research suggests that areas that are vulnerable to invasion can be better predicted by quantifying the elements of the landscape that significantly influence the density and spread of a species. However, the identification of an optimal set of landscape metrics for a species will require case specific study as clearly different species will respond to landscape structure in different ways.

This research also demonstrated that dominant processes shaping population density and spread of invasive species can be identified and prioritized, as well as those of secondary importance. Variables representing an Allee effect, the intrinsic rate of increase and propagule pressure were found to be the three most influential on both density and spread over multiple invasive species scenarios. The frequency of long-distance dispersal events, in combination with measures of suitable habitat amount and habitat aggregation, was found to be a better indicators of population dynamics than the intrinsic abilities of a species to disperse in fragmented landscapes. This research has shown that the presence of an Allee effect leads to a balance between the effectiveness of spread and invasion success. Spread is maximized at an intermediate dispersal rate and inhibited at both low and high rates. The configuration and composition of the landscape, by either increasing or mitigating or the dispersal abilities of a species, can lead to a rate of spread under a dispersal threshold for which density and spread is at the highest.

This research highlights how complex interactions between propagule pressure, species traits and habitat characteristics can determine patterns of invasion across fragmented landscapes. Successful management of invasive species, particularly for prioritisation and design efficient surveillance and control strategies, will depend on understanding this context dependent effect across habitats. More important, however, the research highlights the need for implementing multi-scenario modelling frameworks to reduce model uncertainty and to identify optimal trade-offs between model precision and complexity. Such development is in its infancy, and further research to correctly and consistently assess, and communicate uncertainty, surrounding spread modelling is needed. Informative assessments and clear communication of uncertainty will allow end-users and practitioners to make more informed decisions about the potential for invasive species establishment

Acknowledgment

There are a number of people behind this work who deserve to be both acknowledged and thanked here: committed supervisors, generous research group members, inspiring family, a fantastically supportive partner, and patient friends.

I extend abundant and sincere thanks to my main academic supervisor Assoc. Prof. Susan P. Worner for the pleasure of working together and exchanging ideas. You have been an excellent supervisor and a caring mentor, who have routinely gone beyond your duties to instil confidence in both myself and my work. Your passion for science and feminism, your wide knowledge and logical way of thinking have been extremely helpful. It is due to your valuable guidance, cheerful enthusiasm and ever-friendly nature that I was able to complete my PhD research. I am humbly grateful.

To my co-supervision team, Dr. Crile Doscher, Assoc. Prof. Daniel B. Stouffer, and Dr. Julien Cucherousset, thank you very much for your contribution to this project, your useful suggestions, and unrelenting support throughout this process. I am ever grateful that you gave me the opportunity to work in your respective labs and for continuously challenging me with new ideas. To Dr. Joel Pitt, thank you for actively explaining and allowing me to apply your general modular dispersal framework (MdiG). To Peter Maxwell, thank you for your support to enhancing the performance of my software codes and accessing the University of Auckland's high performance computing (HPC) facilities and related eScience infrastructure. Dr. Hossein Ali Narouei Khandan, Jack Linton, Marona Capdevilla-Rovira, Mariona Roige, Ursual Torres, and Dr. Senait D. Senay provided invaluable help into the project development, statistical support, results interpretation and assistance in realizing the final manuscript. To Dr. Karl Suiter, Yu Takeuchi, Yulu Xia and Steve Hong thank you for impating me your valuable advice and comments, especially on our joint work on uncertainty and spread modelling, which helped to further refine my work towards a fruitful completion.

To Dr. Hannah Buckley, Prof. Jason Tyliankis, Camille Coux, Johanna Voinopol-Sassu, Mauricio Ganzalez Chang, Dr. Timothy Curran and Jenny Dent, I am ever grateful that you gave me the opportunity to work on a great variety, and always challenging projects. To the Te Pūnaha Matatini staff and Whānau, to all colleges and staff in the Bio-Protection Research Centre, thank you for your help and encouragement. In particular, I would like to thanks Dr. Andrew Holyoake, Susan Paterson, Dianne Fyfe, Eunice Wheal, Diane Dinnis and Dr. Caroline Pratt for having provided all types of support whenever needed and in record time: computer, quiet office, remote connection, administrative and financial support to cite some. To my friends, your friendship means a lot to me. Thank you for mutual support and hilarity, for regularly bringing coffee, wine and smiles! You let me own a happy family in Christchurch, New Zealand.

I am very much thankful to my family for their continued support. In particular, to my parents, Claudine and Jean-Jacques, my grand-parents, Jacqueline, Marcel and Marthe, and my brother and sisters Elsa, Chloé and Quentin for being so loving and supportive. To my partner Sam for his unremitting encouragement and for being willing to spend many weeks thousands of kilometres apart from me (maybe he thinks that was not such a bad deal after all). ‘Chinna Chinna Aasai Siragadikum Aasai. Muthu Muthu Aasai Mudinthuvaitha Aasai. Vennilavu Thottu Muthamida Aasai. Ennai Intha Bhoomi Sutrivara Aasai’

This project was granted by a postgraduate Research scholarship provided by the Bio-Protection Research Centre, Lincoln University, New Zealand and funded by the Tertiary Education Commission (TEC), New Zealand. In addition, Better Border Biosecurity (B3), the Ministry for Primary Industry (MPI) and the Te Pūnaha Matatini Research Centre provided equipment and supplemental funding. Computer time for the simulations performed in this thesis were provided in part by the High Performance Computing Centre of the New Zealand eScience Infrastructure, NeSI.

Contents

List of Figures

List of Tables

Chapter 1

General introduction

1.1 Biological invasions in changing ecosystems

Invasive alien species are species that are introduced in a location outside their natural geographical range, where they become established and then proliferate and spread to the detriment of natural systems and human interests (Richardson *et al.*, 2000; Blackburn *et al.*, 2011). The on-going, global mass invasion by invasive alien species comprises a rapidly growing array of unique ecological, economic and social challenges of unprecedented magnitude (Carpenter *et al.*, 2006). The negative ecological impacts of invasive species have been reported at multiple levels of biological organization, ranging from genetic and evolutionary changes in individual organisms, disruption of native population dynamics, local and regional extinctions of native species, to the alteration of biochemical pools and fluxes of nutrients, fire regimes and water quality (Simberloff *et al.*, 2013). Invasive species have now affected nearly every type of ecosystem on the planet and pose one of the biggest threats to biodiversity worldwide and associated ecosystems services (Carpenter *et al.*, 2006; Bellard *et al.*, 2016).

Invasive species have significant socio-economic impacts, inevitably affecting human well-being. They can reduce yields from agriculture, fisheries, and forestry, sometimes exacerbating the already grave situation of food security in developing countries (Steiner, 2010). The conversion of native communities to invasive-dominated communities has aesthetic and cultural impacts (Nuñez & Simberloff, 2005). Some of the large epidemics that caused humanitarian and economic crises, such as malaria (*Plasmodium spp.*), Bubonic plague (*Yersinia pestis*) and West Nile virus (*Flavivirus spp.*), are known to be spread by invasive species (Lounibos, 2002). Various publications report large sums of money spent with respect to the cost of detection, control and eradication of invasive species. According to De Poorter *et al.* (2007), USD 1.5 trillion/year is incurred globally (5% of the global economy), the majority of which is due to crop losses and the application of herbicides and pesticides to reduce invasive weeds and pests. Pimentel *et al.* (2000, 2005) estimated damage costs associated with invasive species effects and their control to be approximately USD 120 billion/year in the United States alone. Even for a small island nation in the Southern Hemisphere, Giera & Bell (2009) concluded that close to $2 billion dollars are spent in New Zealand annually on invasive alien species alone. Invasive species therefore contribute to social instability and economic hardship, placing constraints on sustainable development, economic growth, food security and public health.

As humans have broken down the major biogeographic barriers by the rapid expansion of enhanced transport technology and trade routes among previously disconnected countries, the frequency and magnitude of biological invasions and their costs to society have been increasing at an alarming rate (Meyerson & Mooney, 2007; Perrings *et al.*, 2005; Banks *et al.*, 2015). It is also becoming clear that the human driven changes in regional land-use, environment and climate, through eutrophication, urbanization, pollution, and non-sustainable harvesting, for example, can increase the numerical abundance and/or geographic range size of an invasive species (Didham *et al.*, 2007; Thuiller *et al.*, 2008). Land-use changes, through habitat loss and fragmentation, continue to open new areas to invasive species and alter disturbance regimes that are conducive to the establishment and spread of invasive species (Didham *et al.*, 2005). Increase in nutrients such as nitrogen from land-use, also promote the growth of opportunistic species. Taranu *et al.* (2015), for example, reported that an increase in nitrogen from agricultural activity promotes the expansion of harmful cyanobacteria throughout lakes in the northern temperate-subarctic regions, posing a serious threat to drinking water sources. Climate change is projected to alter temperature and precipitation regimes at a regional scale, extending the climatic suitability for some species, in particular, tropical and sub-tropical invasive species (Hellmann *et al.*, 2008).

The prolific nature of the invasive species problem has made biological invasions an enormous ecological and societal challenge. Prevention and early detection, coupled with a rapid assessment and a quick and effective response, is viewed as one of the most cost-effective approaches to combating invasive species (Pyšek & Richardson, 2010). While real progress has been made towards identifying key drivers and processes that facilitate or prevent biological invasions, the on-going and future change in the distribution of invasive species remains extremely difficult to predict. For example, we are still no closer to determining why some species establish self-sustaining populations in an area and others do not. Nor are we any closer to determining why a species may become overwhelmingly invasive in a particular environment, while others will spread slowly or not at all. As a result, management strategies of invasive species are mostly reactive. However, given the projected rate of environmental change and associated risk of biological invasion, there is a growing interest in pre-emptively managing ecosystems to mitigate negative impacts of invasive species before significant damage occurs (Larson *et al.*, 2011; Thuiller *et al.*, 2008).

Proactive responses are limited without the existence of effective forecasting methods and the availability of advanced models to do this. Species distribution models (SDMs) and prediction of spread have become increasingly important tools to investigate hypotheses about the processes and drivers of invasions (Venette, 2015). Mechanistic models, in particular, provide excellent tools for evaluating the relative importance of different drivers of biological invasion because they involve, 1) inventories of the biological, ecological, environmental and human factors that impact the establishment and spread of invasive species, and, 2) quantifying the probabilities of invasion associated with these factors (Robledo-Arnuncio *et al.*, 2014; Venette, 2015). In practice, however, risk predictions of establishment and spread typically lack explicit consideration of the interaction between multiple drivers of invasion (Catford *et al.*, 2011; Gallien *et al.*, 2014). Current theoretical and empirical studies lead to conclusions most applicable to the specific conditions under which they are developed, which limit inference across species and ecosystems.For example, the relative importance of population growth, dispersal abilities and propagule pressure versus habitat structure remains poorly understood (González-Moreno *et al.*, 2013a). Robledo-Arnuncio *et al.* (2014) emphasized the urgent need for more integrative approaches and broader modelling frameworks that focus on cross-study comparisons at different spatio-temporal scales to make better predictions of the vulnerability of organisms and ecosystems to invasion.

In this research I broadly review the recent achievements in developing forecasts of spread of invasive species and address some of their limitations. I devote particular attention to the challenge of accounting for the relative contribution of multiple drivers of invasion which comprises an essential basis for explaining and predicting spatial patterns of spread in changing ecosystems. The broad aim of this thesis is to undertake an integrated assessment based on a spatially-explicit model of spread of an invasive species under different scenarios of habitat composition and configuration. The integrated assessment and its methodological approach are intended to be useful for decision-making within New Zealand‘s existing biosecurity system and can be easily applied elsewhere. The remainder of this Chapter provides a broad overview of the research topic and thesis organization. It provides the rationale for the present study and its significance, as well as a statement of the research objectives and tasks undertaken. In particular, Section 1.2 gives an overview of the research topic relating to drivers of establishment and spread of invasive species. Section 1.3 broadly reviews the achievements in developing models of spread. Section 1.4 emphasises the knowledge gaps and future direction of research, and details the research development of this thesis.

1.2 Drivers of establishment and spread of invasive species

1.2.1 The process of invasion

Biological invasions involve four sequential stages, 1) the arrival phase, generally driven by a small number of individuals (propagules), 2) the establishment phase with low spread rate, 3) an expansion phase characterized by high spread rates and exponential population growth, and, 4) a saturation phase when spread rates reach a limit (Lockwood *et al.*, 2013). I focus on the second and third stages where spatial-modelling has the most to offer. These invasion stages are perceived as being divided by ecological filters that species must overcome before proceeding to the next stage. The invading population has to disperse into a suitable environment (1st filter), where it can survive and breed successfully (2nd filter), establish in recipient communities through competitive interactions (3rd filter), and finally promote further survival and reproduction by expanding its range to suitable habitats (4th filter). Among others, Catford *et al.* (2009) and Gurevitch *et al.* (2011) identified more than a dozen leading hypotheses that have been proposed and tested to characterize the vulnerability of organisms and ecosystems to invasion. These hypotheses typically consider the size and frequency of introduction (propagule pressure), the life-history traits of the invaders, and the abiotic and biotic characteristics of the recipient ecosystems, all of which vary in time and space. Surprisingly, despite the general consensus on the importance of investigating these three processes and their interactions, they are usually studied independently (e.g. Gallien *et al.*, 2014).

1.2.2 Propagule pressure

A predominant observation in a variety of taxonomic groups is that, as propagule pressure increases, so does establishment and spread success (Lockwood *et al.*, 2005, 2009). Intuitively, increasing the size and number of propagules enables the incipient population to overcome limitations associated with small populations, such as Allee effects, environmental stochasticity and demographic stochasticity. Reflecting on the generality of the relationship between propagule pressure and invasion success, Colautti *et al.* (2006) have advocated that propagule pressure is the key driver of invasion and should form a foundational element of all establishment and spread risk analysis.

The relative effect of a single, versus multiple incursions remains, however, relatively unexplored. Hopper & Roush (1993), Grevstad (1999) and Liebhold & Bascompte (2003) argued that Allee effects, such as reduced fitness when conspecific density is low, favour fewer introductions of more individuals, while high environmental stochasticity favours a larger number of introductions. Nevertheless, Brook (2004) suggested that repeated introductions of small populations may systematically promote establishment by supplementing genetic variation adaptive for new habitats. Using a simulation model to explore the probability of establishment of exotic birds, Cassey *et al.* (2014), found that one introduction event always outperformed multiple introduction events that summed to the same size. In addition, while the size of propagule was undoubtedly important for ensuring population growth in the novel environment, the authors found that demographic traits had the greatest effect on establishment. This finding echoes several other studies that have also questioned the generality of the positive relationship between propagule pressure and invasion success (Nuñez *et al.*, 2011; Yeates *et al.*, 2012). The initial introduction stage of most biological invasions is poorly understood, in part, because there is often no reliable records of the species that were introduced but failed to establish (Haydar & Wolff, 2011; Jerde & Lewis, 2007; Ferrari & Lookingbill, 2009). McGregor *et al.* (2012) emphasized that non-random patterns of introduction may result in the pool of introduced species being a biased subset with regard to key species attributes and life-history traits, potentially biasing our perception of which factors are important for determining establishment and spread success.

1.2.3 The importance of life-history traits

Successful invaders often possess life-history traits that promote invasiveness. Different combinations of life-history traits might be required depending on the characteristics of the invaded habitat (Delatte *et al.*, 2009) and may vary across invading species (Sakai *et al.*, 2001). For example, current theory on optimal insect development involves a direct trade-off between the disadvantage of a longer juvenile development time and the fitness advantages of large body size. However, it is widely recognized that rapid demographic growth and rapid dispersal rates promote the establishment and spread of an invading population in a novel environment (Kolar & Lodge, 2001; Kneitel & Chase, 2004; Pyšek & Hulme, 2005).

Upon arriving in a novel environment, invading populations are likely to be small and therefore subject to demographic and environmental stochasticity, as well as Allee effects (Taylor & Hastings, 2005). Rapid demographic growth may promote establishment by counteracting such negative effects, while populations with low growth rates may face extinction. On the other hand, spread of invading organisms is often the result of short-distance dispersal (natural dispersal abilities of the organism into surrounding adjacent areas) in combination with long-distance dispersal (e.g. wind-borne transfer of small organisms, human transportation). Both isolated colonies formed by long-range dispersal and populations standing on the boundaries of the population front are generally of low density and therefore subjected to stochasticity and Allee effects, as in the establishment phase. Keitt *et al.* (2001) and Taylor & Hastings (2005) suggested that the intrinsic rate of population increase represents a critical factor regulating the rate of range expansion, by counteracting demographic stochasticity and Allee effects once the population is established.

On the other hand, high dispersal rates have been proposed to facilitate spread, suggesting a positive feedback between range expansion with selection for increased dispersability (Hill *et al.*, 2011). Smith *et al.* (2014), however, showed that when Allee effects apply, high dispersal rates can also act as a drain on the introduced population, which can become too small to be maintained. The authors concluded that spread was maximal for intermediate dispersal rates but was inhibited at low and high dispersal rates. Dispersal rates are determined to a large extent by long-distance dispersal events, particularly in patchy or fragmented environments because long-distance dispersal events can connect habitat patches that are no longer physically linked (Clobert *et al.*, 2012). Robledo-Arnuncio *et al.* (2014) suggest that it is clear that the frequency and length of long-distance dispersal events are important for explaining differences in propagule deposition, but they only explain a small fraction of the variation observed in population density and rate of spread of invasive species across habitats. In many studies, such variation has been attributed to temporal and spatial variation in habitat conditions (abiotic and biotic interactions and habitat structure) that can act to limit or exacerbate the growth and dispersal abilities of invading organisms (e.g. Robledo-Arnuncio *et al.*, 2014).

1.2.4 Vulnerability of ecosystems to invasion

Facon *et al.* (2006) and Barney & Whitlow (2008) emphasize the central role of a match between species requirements and the characteristics of the new environment invaded. This match can either be present at initial contact between a species and its novel environment, or may develop subsequently through changes to the invasive species (evolution), the environment (land-use or climate change for example) or both (Facon *et al.*, 2006). At global and regional scales, it is generally well accepted that abiotic conditions, such as temperature and precipitation, and resource availability (e.g. host availability), are the main drivers of species distributions. This acceptance is related to the fact that species will not be able to invade an area that has abiotic conditions outside of its physiological tolerance levels (Petitpierre *et al.*, 2012). At a local scale, invasive species typically have higher density in their novel habitat than in their native ranges (Vilà *et al.*, 2007; Ibáñez *et al.*, 2009). These differences have been partly explained by biotic interactions and human influence, such as biotic resistance and human generated disturbance.

The biotic resistance hypothesis proposes that habitats with higher biodiversity are more resistant against invaders than habitats with low biodiversity (Elton, 2000; Shea & Chesson, 2002). Because of their geographic and evolutionary isolation, islands often present lower biodiversity than continents and are expected to be more prone to invasion (island susceptibility hypothesis) (Elton, 2000). Alternatively, the 'enemy release hypothesis' (Colautti *et al.*, 2004) specifies that invasive species is often not be regulated in the new environment due to the absence of their native enemies. The 'novel weapons hypothesis' proposes that some invasive species gain advantages over native species by possessing a trait, such as better defence mechanism or competitive ability that is new to the resident community of native species, that can affect them negatively (Keane & Crawley, 2002; Colautti *et al.*, 2004). Finally, the 'meltdown hypothesis' claims that the establishment of one invasive species in a new environment makes it easier for other non-native species to invade (Simberloff & Von Holle, 1999).

All the hypotheses and theories have connections and points upon which various authors disagree, along with gaps and limitations in what they encompass. Ultimately, propagule pressure, abiotic factors and biotic interactions, both with ecological and evolutionary components, can act as filters that control the match of a population of invading species to its novel habitat (Worner, 2002; Catford *et al.*, 2009; Gurevitch *et al.*, 2011). As such, disturbance, defined in this study as any event that disrupts ecosystem structure and function (Pickett & White, 2013), is widely recognized to offer new opportunities for a species to enter an ecosystem by encouraging a shift in abiotic conditions, biotic composition and disrupting source pools of natural enemies. Increasing levels of human driven change in regional land-use and climate, coupled with the dramatic increase in the transport of biota across the world (propagule pressure), have created types of disturbance that are unlike anything in the evolutionary history of many species. Early in the development of the field of invasion science, several approaches were considered to address how alteration of landscapes might promote invasive establishment and spread (Fisher, 1937). Among these, the utility of landscape ecology, that aims to better understand the effect of landscape structure on population dynamics and spread of invasive species, has long been recognized (With, 2002). However, it is only recently that it has seen a resurgence of interest (Vilà & Ibáñez, 2011).

1.2.5 Landscape effects on the invasion process

Human activities have become key drivers of invasive species establishment and dispersal, through their direct contribution to the transport of propagules and intense land-use changes (Mack & Lonsdale, 2001; Zimmermann *et al.*, 2014). Land-use change can be seen as a landscape-level disturbance underlying habitat loss and habitat fragmentation (Hobbs, 2000). Urban and semi-urban areas are particularly vulnerable to biological invasion. Cities are centres of intentional and accidental species immigration, for example through ornamental plantings, green waste dumping and domestic animals and serve as major sources of propagule pressure due to the amount of commodities arriving or passing through for trade and commercial activities (Vitousek *et al.*, 1996; Garden *et al.*, 2006; Niinemets & Peñuelas, 2008). Human infrastructures, such as roads, highways, hiking trails and railways, create new potential habitat for invasion, as well as potential for propagule transportation along transport routes (Mortensen *et al.*, 2009).

The modern human modification of open range lands, converts complex natural ecosystems to simplified managed ecosystems with increased nutrient availability that encourages the establishment of invasive species (Chytrỳ *et al.*, 2008; Pyšek *et al.*, 2010). In agricultural ecosystems for example, modern methods have generated monocultures that are amalgamated and enlarged to enhance farming efficiency. Landscape simplification, coupled with the intense use of fertilizers, increases the concentration of suitable resources that are available to a particular invasive species and can reduce habitat types that support invasive species' natural enemies (Jonsson *et al.*, 2015; Tscharntke *et al.*, 2012; Rigot *et al.*, 2014). Fragmentation of wild-land habitat, resulting from expanded agricultural and urban development, diminishes habitat size and connectivity among native populations and increases edge habitat between natural and transformed areas (Harper *et al.*, 2005; González-Moreno *et al.*, 2014; La Morgia *et al.*, 2011). Habitat edges are disturbed habitats that have different environmental and species composition than interior habitats that can modify nutrient transport (Peterjohn & Correll, 1984), affecting species persistence and biodiversity and nurturing invasive species (Kennedy *et al.*, 2002). Because fragmentation tends to give invasive species a foothold in the landscape, it can have devastating effects even on interior habitats.

Empirical studies are strongly supported by simulation models showing that the local density of invasive species is determined not only by local habitat quality, but also by the spatial structure of the habitat in the surrounding landscape (With, 2002, 2004). In this context, landscapes comprised of more or less isolated patches of suitable habitat within a matrix of less suitable habitat. From the perspective of a particular species, this patchwork of various habitat types may offer different resources for foraging and reproduction, and different suitability for species dispersal (Hanski, 1999). Within landscape theory there are two important aspects of landscape spatial structure capable of impacting the growth and movement of species (Hansson *et al.*, 2012). First is landscape connectivity, defined as the combined effect of landscape elements that facilitate or disrupt the movement of individuals between habitat patches, and, second, environmental heterogeneity (i.e. variance and range of environmental conditions and their spatial autocorrelation). However, we only have a limited understanding of how these factors interact to affect the establishment and spread of an invasive species.

Surprisingly, theoretical and empirical approaches lack explicit consideration of the effect of landscape features on the success of establishment and spread of species. Little is known about how invasive species move through the landscape. This lack of understanding may be partly due to the difficulty measuring long-distance dispersal across species and habitats (Robledo-Arnuncio *et al.*, 2014). Even less is known about the relative importance of different features of landscape structure, for example, patch size, shape, edge density, patch aggregation, compared with propagule pressure and life-history traits in shaping the distribution and abundance of invasive species. Identifying which factors have most influence on the rate of spread and population density is nevertheless fundamental. The ability to identify the importance of each factor will not only contribute to an improved understanding of the underlying process, but will also facilitate the development of efficient invasion predictors and the identification of factors that may be most effective reducing population density and dispersal rates.

1.2.6 Summary and future perspectives

The distribution and spread of invasive species is the result of complex combinations of factors. The size and frequency of introduction (propagule pressure), the life-history traits of the invaders, population interactions and suitability of the abiotic recipient environment create the conditions that lead to demographic changes resulting in population establishment and growth. Processes such as disturbance that provide physical space or resource availability encourage shifts in community composition and increased the susceptibility to invasion. Interest in the effects of broad scale environmental changes has spurred the effort by landscape ecologists to decipher the effect of habitat loss and fragmentation on various ecological processes, including the establishment and spread of invasive species. Surprisingly, despite the general consensus about the importance of investigating these different processes and their interactions on the success of invasion, risk predictions of establishment and spread typically lack explicit consideration of the interaction between multiple drivers (e.g. Gallien *et al.*, 2014). In particular, spatially-explicit analyses of invasive spread at a landscape level are very few and current empirical and theoretical studies traditionally focus on a single-species approach and habitat requirements, which limits inference across species and habitats (e.g. Catford *et al.*, 2011). As a result, the relative role of habitat structure (composition and configuration) versus propagule pressure and life-history traits of the invasive species remain unclear.

Correctly establishing causality through a mechanistic understanding of habitat modification on species invasion is crucial to achieve better prediction and management of invasive-species in changing ecosystems. For example, if the spread of a local invasion is habitat limited (e.g. if space is saturated), the potential for further impacts, such as the displacement of native species, is low. On the contrary, if marginal habitat limits the invasion by providing natural barriers to dispersal, the risk of invasion of adjacent, potentially suitable habitat remain important and the potential for further impact is much higher. The outcome of such spatial analysis, however, remains limited by constraints in the ability to carry out a comprehensive quantitative synthesis of landscape effects on the spread of invasive species. In particular, empirical studies fail to account for large gradients of habitat complexity, such as differing patch size, shape, juxtaposition, interpatch connectivity or habitat aggregation, and therefore offer insufficient or even no replication across habitat, in part because of an unavoidable trade-off between spatial scope, sampling intensity and accuracy (Robledo-Arnuncio *et al.*, 2014).

To incorporate landscape ecological expertise in risk assessment of biological invasion, tools and models are required that allow for a systematic investigation of the impact of changes in landscape structure on population dynamics. Understanding the landscape ecology of invasive spread may afford new insights and opportunities for managing and restoring landscapes to control the spread of invasive species and minimize the invasibility of communities.

1.3 Models of establishment and spread of invasive species

1.3.1 Species distribution models

A major division has existed between the approaches employed for modelling large-scale species distribution and those for modelling local-scale population spread. Broad scale projections of species' distribution are largely based on static approaches linking current species distribution (SDMs) to environmental variables such as temperature, precipitation and elevation, and resource distribution (Guisan & Thuiller, 2005). These models can be used to project future distribution and impacts from future land-use or climate scenarios (Thomas *et al.*, 2004). Techniques for fitting SDMs to observed data have developed rapidly and offer practical advantages over more mechanistic modelling approaches due to the relative simplicity of their data requirements, their relative ease of use within freeware packages, and the range of the interactions (biotic and abiotic) they can detect and characterize (Guisan & Thuiller, 2005; Elith & Leathwick, 2009). Consequently, correlative SDMs have been used widely for invasive species risk screening and conservation applications (e.g. Moilanen & Wintle, 2007). However, despite the status of dispersal and population dynamics as key ecological constraints of species distribution and spread, as well as the role of spatial structure in limiting or exacerbating the dispersal abilities of an invading population, these factors have rarely been incorporated into SDMs (Guisan & Thuiller, 2005; Franklin, 2010; Huntley *et al.*, 2010; Schurr *et al.*, 2012; Thuiller *et al.*, 2013). In contrast, mechanistic models of spread incorporate, to differing degrees, both demography and dispersal at their core.

1.3.2 Analytical models of spread

Historically, spread models relate the spatial location of reproducing individuals to the spatial location of their parents through the definition of a dispersal kernel, a probability function describing the distance from the original location after dispersal. Partial differential equations (PDEs) and reaction-diffusion (RD) models assume spatial homogeneity, random reproduction and random movement during an individual life time (Fisher, 1937; Skellam, 1951). These basic approaches were then extended to include the role of demographic heterogeneity using, for example, age-dependency of movement and reproduction (Mollison, 1991; Van den Bosch *et al.*, 1992). Later, discrete-time integro-difference equations (IDEs) were developed to account for different forms of dispersal kernel and, in particular, to accommodate the fat-tailed shape of observed long-distance dispersal data (Kot *et al.*, 1996). More detailed models link an age- or stage-structured matrix IDE and thus account for age dependent dispersal (Caswell, 2001). Adding demographic complexity to spatial population models has greatly improved the understanding of how local dynamics can affect spatial spread (Jongejans *et al.*, 2008). However, both RD models and the IDE framework again assume spatial homogeneity of the environment.

Theoretical examination of the influence of spatial heterogeneity in modelling efforts begins with a binary classification of suitable and unsuitable sites (Shigesada *et al.*, 1986). Persistence conditions as a function of the fraction of favourable area have been found and Shigesada *et al.* (1995) showed that rate of spatial spread is dependent on the mean spread rate in different environments. Their approach has been extended to include sinusoidally varying environments (Kinezaki *et al.*, 2006), two dimensional domains (Kinezaki *et al.*, 2003; Roques *et al.*, 2008), more general periodic environments in multiple dimensions (Weinberger, 2002; Berestycki *et al.*, 2005; Roques & Hamel, 2007; Hamel *et al.*, 2010), directional movements (Petrovskii & Li, 2003; Almeida *et al.*, 2006) and many other phenomena (Berestycki *et al.*, 2009; Kinezaki *et al.*, 2010; Roques & Chekroun, 2010; Garnier *et al.*, 2012; Kawasaki *et al.*, 2012; Vergni *et al.*, 2012). The effect of periodic fragmentation has also been examined in the framework of an IDE model (Shigesada & Kawasaki, 1997; Van Kirk & Lewis, 1997) incorporating spatially varying growth rates and varying dispersal kernels (Weinberger *et al.*, 2008). Recent developments in IDE incorporate fluctuating environments (Neubert *et al.*, 2000), Allee effects (Veit & Lewis, 1996; Wang *et al.*, 2002; Dewhirst & Lutscher, 2009), interspecies competition (Owen & Lewis, 2001; Samia & Lutscher, 2010) and other phenomena associated with spatio-temporal heterogeneity (Caswell *et al.*, 2011; Schreiber & Ryan, 2011; Brown *et al.*, 2012). There are clear benefits to using theoretical oriented approaches in that they provide exact solutions in many cases and allow for important comparisons of model outputs between study systems. However, at present, it is not clear how an analytically tractable model can formally represent a real landscape in the form of a quality function and thus consideration of biological and geographic realism is still limited.

1.3.3 Simulation-based models

Simulation-based models have been used as a possible alternative as they can easily incorporate important characteristics of large numbers of individual organisms and their interactions with resources and ecological processes (Hastings *et al.*, 2005; Jongejans *et al.*, 2008; Hui *et al.*, 2011). The way in which landscape structure and habitat factors have been incorporated into such models has changed markedly over time (Fahrig & Nuttle, 2005). Early models were developed under the scope of metapopulation theory which considers the habitat spatial structure as populations subdivided into spatially isolated suitable habitat patches, functionally connected by distance-dependant dispersal fluxes (Levins, 1969; Hanski, 1991). The dynamics of the system results in a balance between local extinctions and re-colonisation but ignores local demographic dynamics and assumes independence of habitat patches and constant colonization probabilities (Hanski, 1999). These models are built based on the perspective of identifying specific conditions in habitat patches were their pattern in a landscape has implications for population dynamics or persistence (Fahrig & Nuttle, 2005). Several extensions have considered partially explicit representations of the spatial relationships among habitat patches by including, for example, stochastic patch occupancy (Moilanen, 1999; Hanski & Ovaskainen, 2000; Facon & David, 2006), variation in patch size (Howe *et al.*, 1991), variation in patch isolation (Adler & Nuernberger, 1994; Day & Possingham, 1995) or by considering situations where colonization and extinction rates are not at equilibrium (Marvier *et al.*, 2004; Seno & Koshiba, 2005). Network models also consider spatially located subpopulations with their own dynamics, but with variability in the connection structure of subpopulations (Keeling, 1999; Minor & Gardner, 2011; Ferrari *et al.*, 2014).

More recently, models that represent landscapes based on geographical maps where groups of individuals or single individuals are located explicitly in a cell (habitat patch) within a regular grid representing the environment, have been created. One approach is represented by cellular automaton models where each cell is attributed a state ('occupied' or 'empty') which is updated in every time step by examining the state of the neighbouring cells and the previous state using well-defined transition rules (Cannas *et al.*, 2003; Soons *et al.*, 2005; Ferrari & Lookingbill, 2009; Crespo-Pérez *et al.*, 2011). Percolation theory is another approach that examines how spatial heterogeneity affects the invasion process by determining whether the organism reaches all suitable sites or is limited to a subset of suitable sites by distance or barriers (Grassberger, 1983; Ming & Albrecht, 2004; With, 2004; Oborny *et al.*, 2007). An important recurrent theme has been the emergence of critical threshold responses to landscape structure. The connectivity between habitat patches in a landscape may decline suddenly once a certain amount of habitat loss and fragmentation is reached. The effect on invasive spread is that it may dramatically declines past some critical level of habitat loss, which depends upon the species dispersal abilities and degree of habitat fragmentation (With, 2002). However, these methods do not incorporate species-specific demographic processes that are known to contribute to invasive spread.

The general evolution of combining demographic and spatial complexity in spatial population models has taken the form of spatially-explicit, raster-based or individual-based models (IBMs) (Nehrbass & Winkler, 2007; Pitt *et al.*, 2009; Renton *et al.*, 2011; Guichard *et al.*, 2012). Each individual is explicitly modelled and acts in response to limited information that may come from intra-specific interactions (breeding) or interactions with the environment (habitat suitability) that drives movement across the landscape. Correlative SDMs often act as templates on which mechanistic model can operate. The parallel development of digital representation of landscape structures using Geographical Information Systems (GIS) provides insight into how real-world landscape heterogeneity can affect the spatial patterns of spread of invasive species (Ming & Albrecht, 2004; Pitt *et al.*, 2009; Bocedi *et al.*, 2014; Meier *et al.*, 2014). We now have the ability to develop models of range expansion that incorporate different levels of complexity such as, propagule pressure, important individual characteristics, population dynamics and dispersal, and their interaction with resources and species community over a realistic spatial and temporal landscape (see for example Keesing *et al.*, 2006; Keith *et al.*, 2008; Meentemeyer *et al.*, 2011; Carrasco *et al.*, 2012; Parry *et al.*, 2013). Yet, only a few studies have employed them to simulate species establishment and spread, and even fewer have derived projections for species under different land-use scenarios.

1.3.4 Detecting and predicting the response to landscape changes

The explicit consideration of spatial heterogeneity in a model of spread is currently recognised as a major advance in contemporary ecology (Fahrig *et al.*, 2011). At the foundation of this analytical advance is the basic need to characterize and quantify spatial heterogeneity to detect change in landscape structure and to investigate the relationship between landscape structure and demographic processes. Rapid fragmentation and habitat loss due to anthropogenic activities spurred the development of a variety of landscape metrics to quantify spatial composition (e.g. habitat type and abundance) and configuration (e.g. shape complexity, edge density) of landscapes. Yet, these metrics have been rarely used as explanatory factors of invasion success (With, 2002; Sebert-Cuvillier *et al.*, 2008; Robinson *et al.*, 2014; Smith *et al.*, 2014; Morel-Journel *et al.*, 2015).

The general perception is that three basic problems are found with the utilization of landscape metrics. These are, 1) the metrics are highly correlated, 2) there may be ambiguous responses to different spatial processes, and, 3) there are sensitive to changes in spatial scale (resolution and extent). Quantifying the specific effect of habitat configuration on spread success, for example, is difficult because many configuration metrics are correlated with habitat amount (Kupfer, 2012). Such limitations can often be addressed, or put into perspective through careful data manipulation, analysis and interpretation. More fundamental, however, spatial modelling in invasion ecology is confronted with difficulties arising from a lack of a conceptual framework to investigate the relationships between quantitative measures of spatial heterogeneity and the spread of invasive species.

1.4 Challenges for invasive species dispersal modelling

The ability to predict when and where an already introduced or potentially invasive species will go next in the landscape has great advantages for the development of subsequent monitoring and eradication strategies. In particular, time and money can be prioritized effectively if comparatively easy-to-measure life-history traits, propagule pressure, ecosystem characteristics or landscape factors can be used to refine assessments of the relative risks of invasiveness among a pool of non-native species and habitats. However a number of theoretical and technical factors still challenge the use of complex spatial models of spread. Two of these factors important to this thesis are discussed below.

1.4.1 Limited taxonomic and geographic scope of current research

One of the factors currently limiting our understanding of the effect of habitat changes on population dynamics of invasive species is the relatively limited taxonomic and geographic scope of current research. It is acknowledged that for many, if not all invasive species, multiple drivers contribute to their establishment and dispersal, but as suggested previously, their combined effects are rarely investigated. Most empirical and theoretical studies evaluate questions and hypotheses about the role of landscape structure within a single landscape and thus provide no replication. The majority of these studies focus on a single species and thus do not provide insight into the response of multiple species in the same landscape. Similarly, spatially-explicit simulation models of invasive spread generally integrate specific characteristics of the studied species, for example, specific dispersal strategies or population dynamics, to obtain ‘realism’, but the result of the simulations are closely linked to the choice of the parameters rather than providing any insight into general principles. More holistic models are based on assumptions about dispersal but tend to ignore important demographic processes, such as Allee effects at the leading edge of geographical ranges. While applied studies of single species of interest are important and will continue to be useful, building a better quantitative understanding of how interactions among drivers might mitigate or exacerbate establishment and spread of non-indigenous organisms, demands more than the current focus on a single-species approach and specific environmental requirements (Robinet *et al.*, 2012b; Wang & Jackson, 2014). Critical to this effort is the development of a framework that should facilitate a systematic investigation into the impact of spatial heterogeneity on spatially-explicit ecological processes. Such studies are in their infancy, but a more rigorous framework is needed to advance predictive landscape structure so that it can more reliably inform future actions to address the effect global land-use change (Robledo-Arnuncio *et al.*, 2014).

1.4.2 Uncertainty in spread modelling

Another central issue involved with predicting future species distribution and abundance in a spatially-explicit context, is to understand and deal with a vast array of uncertainties. Poor quality data, estimation errors, unpredictable climate or environmental disturbances, model types, and a lack of understanding of the fundamental mechanisms of invasive spread, are among the many different factors causing the uncertainty in assessing spread risk, and more particularly, spatially-explicit spread forecasts. For example, a major problem with any model of spread dynamics is fitting a dispersal kernels because spread patterns are very sensitive to the behaviour of long-term distance dispersal, for which reliable observation data are less available (Robledo-Arnuncio *et al.*, 2014). Tackling such uncertainty is particularly challenging for pest risk managers to incorporate into their decision-making process. It has often been stated that complex spatial models used to synthesise available knowledge of the dynamics of the invasion process, are too hard to parametrize, test and understand because they contain too many poorly known parameters and cannot be evaluated using traditional statistical approaches. This criticism has especially been targeted towards such complex models for guiding effective management of invasive species at a regional scale (Robinet et al., 2012b; Venette, 2015). However, models are often most useful for challenging what we think we understand. Investigating the use of different modelling approaches (different model structures and parameters) and the impact on projections performance could help identify the optimal trade-offs between precision and complexity. Such development has been aided by the development of consensus on parametrization, evaluation methods (see for example the pattern oriented modelling framework (POM) developed by Grimm *et al.*, 2005) along with the parallel progress in data availability, computation, statistics and formal quantitative measurement of uncertainty. While initial investigation showed the incorporation of uncertainty adds to the credibility to pest risk maps, systematically quantifying uncertainty or at least specifying it, is not yet routine. Venette *et al.* (2010) and Venette (2015) emphasized the urgent need for substantial improvement in the quantification and communication of uncertainty with respect to invasive spread modelling.

1.4.3 Overall aim and Objectives

The overall aim of this thesis is to determine how biological characteristics of invasive species interact with the abiotic variables and resource distribution to determine establishment success and spread in a varying environment. The specific objectives are:

- **Objective 1**: To develop a general dispersal simulation framework that allows for systematic investigation of changes in the landscape structure on population dynamics of invasive species

- **Objective 2**: To identify a set of key generic landscape metrics that enable a concise characterization of independent aspects of landscape structure, regardless the spatial scale at which the patterns are represented.

- **Objective 3**: To disentangle the independent and interactive effects of landscape composition and landscape configuration on the establishment and spread of invasive species.

- **Objective 4**: To determine the relative importance of dispersal, reproductive rate, propagule pressure and habitat structure on population growth and spread of invasive species.

- **Objective 5**: To review the sources of uncertainty associated with the development of a spread model, and, to investigate a selected set of methods for evaluating spatially-explicit spread model performance and uncertainty.

1.4.4 Thesis structure

Chapter 2 provides a general modelling framework that allows for a systemic investigation of the impact of landscape structure on population dynamics of invasive insects. It comprises of a spatially-explicit dispersal simulation framework, a landscape generator allowing independent change in the composition and configuration of landscape components, and appropriate landscape measures that establish a quantitative relationship between landscape structure and population dynamics.

Chapter 3 discusses the influence of spatial resolution and extent on the characterization of spatial patterns. It describes novel methods to identify a set of key generic landscape metrics that enable a concise characterization of independent aspects of landscape structure, regardless the spatial scale at which the patterns are described.

Chapter 4 identifies the essential aspects of landscape that interact with dispersal and demographic processes based on multi-species dispersal scenarios and discusses the role of landscape structure shaping the rate and pattern of spread.

Chapter 5 evaluates the relative role of biological characteristics of invasive species compared with abiotic variables and resource distribution in determining establishment success and spread in a varying environment.

Chapter 6 provides a brief literature review of the sources of uncertainty associated with the development of species distribution models and spread models, and proposes a selected set of methods for quantifying spatially-explicit spread model performance and uncertainty.

Chapter 7 presents a general discussion of all the results reported in this thesis and their contribution to addressing the research topics and issues outlined in the objectives. Concluding remarks as well as recommendation for future research to improve models of spread for a better predictions of invasive species, especially under global changes, are given.

Chapter 2

MDiG: a modelling platform for the establishment and spread of invasive species in heterogeneous environments

Contribution of authors

This chapter describes a collaboration between the current author with Joel P. W. Pitt, who developed the general modular dispersal framework (MDiG), and Senait D. Senay, who collected the data, parametrized and applied the MDiG to model the spread of *Pieris brassicae* in the South Island, New Zealand. This work is integrated and extended by the current author to further illustrate MDiG's capability by investigating the effect of change in landscape composition and configuration on population density and the dispersal of an invasive species using the European gypsy moth, *Lymantria dispar*, as a case study. This chapter has been submitted for publication with J. P. W. Pitt, S. D. Senay, D. B. Stouffer, C. Doscher and S. P. Worner as authors to *Methods in Ecology and Evolution*, 2016.

Abstract

Human activities have become key drivers of invasive species dispersal through their direct contribution to the transport of the species and through intensive land-use changes, such as habitat loss and fragmentation. Yet, only recently has the influence of the surrounding landscape on invasive species spread started to be considered. The scientific community increasingly recognizes the need for broader modelling framework that focuses on cross-study comparisons at different spatio-temporal scales. Such studies are expected to make better prediction regarding the vulnerability of ecosystems to invasion in changing environments. Using two illustrative examples, we introduce MDiG, a general modelling framework that allows for a systematic investigation of the effect of habitat change on invasive species establishment and spread. The essential parts of the framework are, 1) a spatially explicit model (MDIG), that allows population dynamics and dispersal to be modelled in geographical information system (GIS), 2) a landscape generator that allow replicated landscape patterns with partially controllable spatial properties to be generated, and, 3) landscape metrics that depict the essential aspects of landscape with which dispersal and demographic processes interact. The modelling framework provides functionality for a wide variety of applications ranging from predictions of the spatio-temporal spread of real species and comparison of potential management strategies, to theoretical investigation of the effect habitat change on population dynamics.

Keywords

Heterogeneous landscape, invasive species, spatial modelling, stratified dispersal, population dynamics, environmental change, landscape metrics

2.1 Introduction

Biological invasions can be both the result of global change as well as the drivers of that change, posing extraordinary ecological, economic and health challenges world-wide (Didham *et al.*, 2007; Lockwood *et al.*, 2013). Human activities, in particular, have become key drivers of invasive species establishment and dispersal, through their direct contribution to the transport of propagules and through regional changes in habitats and climates (Hobbs, 2000; Hellmann *et al.*, 2008; Robledo-Arnuncio *et al.*, 2014). Understanding establishment and spread mechanisms and the ability to predict when and where an already introduced alien species will go next in the landscape, is central to controlling their spread and mitigating their impacts in rapidly changing environments (Venette, 2015). The outcome of such spatial analysis, however, remains limited by constraints in the ability to carry out a comprehensive quantitative synthesis of habitat change on the spread of invasive species (With *et al.*, 1997; With, 2002, 2004; Vilà & Ibáñez, 2011; Robledo-Arnuncio *et al.*, 2014) and by the relatively limited taxonomic and geographic scope of current research (Catford *et al.*, 2011; Robledo-Arnuncio *et al.*, 2014).

Preventing the spread of invasive species into new habitats requires an awareness of the types of species that might pose a threat to an ecosystem, and which ecosystems are especially vulnerable to invasion in the face of increasing land-use and climate change. Theory indicates that the distribution and spread of invasive species is the result of a complex combination of factors (see for example the invasion framework described in Catford *et al.*, 2011). These factors include the size and frequency of introduction (propagule pressure), species specific traits that are thought to confer high fitness such as high reproductive and efficient dispersal abilities, and the abiotic and biotic characteristics of the recipient ecosystems that may limit or facilitate the establishment of invasive species. Particular characteristics of a species demography and dispersal, have long been recognised as the main factors influencing the survival of invasive organisms in fragmented landscapes, such that the capability of individuals to grow and disperse between small subpopulations allows the system to function as a viable metapopulation (Hanski, 1999). Demographic stochasticity, caused by chance events influencing individual mortality, reproduction and dispersal, results in random fluctuations in population size that also varies according to local population density and environmental conditions. Surprisingly, despite the general consensus that establishment and spread of invasive species in heterogeneous landscapes depends on the interplay between spatio-temporal heterogeneity, stochastic dispersal and demographic processes, and idiosyncratic propagule pressure, dynamic risk predictions of establishment and spread typically lack explicit consideration of the interaction between these multiple drivers of invasion (Catford *et al.*, 2011; Gallien *et al.*, 2014). In particular, spatio-temporal predictions of establishment and spread, across different species demography and dispersal characteristics and environmental conditions, are very few (Worner, 1994; Guisan & Thuiller, 2005; Franklin, 2010; Huntley *et al.*, 2010; Dormann *et al.*, 2012; Schurr *et al.*, 2012; Thuiller *et al.*, 2013). In addition, the relative effect of habitat change versus propagule pressure, demography and dispersal on invasion success remains unclear.

Early progress in the development of models of establishment and spread were divided into approaches used for modelling large-scale species distribution, versus those for modelling local-scale population spread (Hastings *et al.*, 2005; Hui *et al.*, 2011). Broad scale projections of species' distribution, which have dominated the recent literature, are largely based on static approaches linking current species occurrences to environmental variables such as temperature, precipitation and resource distribution (SDMs: Guisan & Zimmermann, 2000). These models can be used to project future distribution and impacts from future land-use or climate scenarios (Thomas *et al.*, 2004), yet typically overlook important demographic and dispersal processes. On the contrary, phenomenological/mechanistic models of spread, such as reaction-diffusion processes (Fisher, 1937; Skellam, 1951), integro-difference equations (Kot *et al.*, 1996), matrix models (Caswell, 2001; Ramula *et al.*, 2008), metapopulation models (Hanski, 1999; Hanski & Ovaskainen, 2000) or cellular automata (Higgins *et al.*, 2000), incorporate, to differing degrees, both demography and dispersal processes as their basis, but they assume a relatively homogeneous environment. The general evolution of combining demographic, dispersal and spatial complexity in spatial population models has resulted in spatially-explicit dispersal models integrated within geographical information systems (GIS) or process-based SDMs (Nehrbass & Winkler, 2007; Pitt *et al.*, 2009; Renton *et al.*, 2011; Guichard *et al.*, 2012; Bocedi *et al.*, 2014). These models are thought to be more robust to extrapolation to novel habitat and climate conditions because they rely on the characterization of processes regulating the probability of a population surviving to reproduce and disperse in response to local environment conditions. Therefore such models account for the effect of landscape characteristics on the mobility and survival of invading species (Ewers & Didham, 2006; Pitt *et al.*, 2009).

Despite much progress having been made regarding our understanding of the effects of spatial heterogeneity on establishment and spread of invasive species (see for example the review of Hastings *et al.*, 2005; Blackwood *et al.*, 2010), generalised insights of landscape effects that are needed to obtain reliable predictions of species response to habitat changes, are lacking. Particular issues are that most empirical and theoretical studies evaluate questions and hypotheses about the role of landscape structure within a single landscape and thus provide no replication (but see With, 2002; Vilà & Ibáñez, 2011; González-Moreno *et al.*, 2013b). Additionally, the majority of these studies focus on a single species and thus do not provide insight into trait variability in the same landscape (but see Catford *et al.*, 2011; Robinet *et al.*, 2012b; Wang & Jackson, 2014). More fundamental, in a thorough review of contemporary plant dispersal ecology, Robledo-Arnuncio *et al.* (2014) emphasized that spatial-explicit spread modelling is confronted with difficulties arising from a lack of a conceptual framework to investigate the relationships between quantitative measures of spatial heterogeneity and the spread of species.

The primary aim of this study was therefore to construct a general modelling framework that allows for a systematic investigation of the impact of habitat change (e.g. patch characteristics, habitat corridors or landscape permeability) on invasive species establishment and spread. The essential parts of the framework are, 1) a spatially explicit model (MDIG), that allows population dynamics and dispersal to be modelled in geographical information system (GIS) (Pitt, 2008), 2) a landscape generator that allow replicated landscape patterns with partially controllable spatial properties to be generated, and, 3) landscape metrics that depict the essential aspects of landscape with which dispersal and demographic processes interact. Such a framework will give deeper insights into species traits and landscape features that lead to establishment and spread success, and may be key to preventing new incursions and the development of efficient monitoring, surveillance, control and eradication programmes. The framework is unique in two key aspects. First, it includes the capability for much greater realism when modelling reproduction and dispersal processes as it accounts for that inter-individual variability and key stochasticities in demographic and dispersal processes. Second, MDiG explicitly relates demographic and dispersal processes to the landscape in which these processes occur, using the open-source GIS program GRASS (`http://grass.osgeo.org`). MDiG therefore offers possibilities for a broad range of simulation-based modelling experiments, from basic theoretical investigations of invasion dynamics, to strategic modelling of spatio-temporal species distribution and management options. The software was implemented using the programming languages C and Python, and packaged as a freely available, standalone application for Linux and Microsoft Windows. A web implemented interface facilitates initial model exploration. In this study, we describe the general structure of the model and demonstrate its capabilities with two examples. Further details about the model and its implementation can be found in both Pitt (2008) and the official website (`http://github.com/ferrouswheel/mdig`).

2.2 Model description

2.2.1 A spatially explicit dispersal model (MDiG)

The modular dispersal model (MDiG) was originally developed by Pitt (2008). MDiG uses GRASS-GIS raster maps to represent either the presence/absence or abundances of the species under study in raster cells. Initial distribution data can be imported or defined by the user. The model architecture was designed to be extensible to many different taxa, characterized by different population dynamics and dispersal strategies, over realistic landscapes. At a population-scale, a demographic sub-model provides different levels of complexity to mimic the fate of individual organisms, by simulating the life-history events of birth and death. The population can have discrete generations or population structure, for which fecundity, survival and development can also be stage dependent. At the landscape-scale, a dispersal sub-model provides explicit rules that determine the pattern of dispersal for each individual. Demographic stochasticity (intraspecific variation in growth and dispersal traits) and spatio-temporal heterogeneity (spatial or temporal variation in population dynamics traits, dispersal abilities, carrying capacity and Allee effect) can be easily modelled by modifying growth and dispersal traits as a function of local density, local habitat quality or life-stages. Finally, the framework contains a management sub-model that allows different *in situ* management strategies that impact the spread of population distribution to be tested, either by adding an additional event to a life-stage, or by modifying life-history attributes or dispersal traits. These treatments can be global or restricted to a certain region. The model definition file is specified in a file formatted in XML, that defines when and how demography, dispersal, simulation results and other model aspects are specified.

2.2.2 The population sub-model

The growth module, 'r.mdig.growth', is designed to represent local growth or the number of individuals within each cell, from one time step to another. The definition of a time step is up to the user, but applies to the entire model. The carrying capacity parameter indicate the maximal number of individuals in a cell. It can be specified as a global value, for each land-cover category or for each cell of the map, to accommodate for spatio-temporal variation in habitat quality. The population growth dynamic is determined by a difference equation chosen and parametrized by the user. The options include logistic growth (Verhulst, 1838), the Beverton-Holt equation (Beverton & Holt, 1957), the Ricker equation (Ricker, 1958), Neubert equation (Neubert, 1997), the Wang equation (Wang *et al.*, 2002) and the Keitt equation (Keitt *et al.*, 2001). Population-based simulations with different life-stages can apply a matrix-based population model using the life-stage module, 'r.mdig.agepop'. The module is designed to account for growth and dispersal age dependency. Population behaviour is programmed based on knowledge about the biology and the life cycle of the species collected from the literature or from experiments.

2.2.3 The dispersal sub-model

2.2.3.1 Local dispersal

The neighbourhood module, 'r.mdig.neighbour', is designed to represent local spread or a diffusion process based on random walks to surrounding adjacent cells. The proportion of individuals that spread from any cell can be specified as a parameter. Both the shape, which defines the direction of the neighbourhood of a cell (East, North, West, South) and the range, which pre-determines the radius of the neighbourhood around the focal cell (2 or 4 cells), are used to represent local random movement of individuals. Example of neighbourhoods include the Von Neumann and the Moore shapes (Figure 2.1).

2.2.3.2 Long-distance dispersal events

The kernel module, 'r.mdig.kernel' is designed to represent long-distance dispersal events, resulting, for example, from wind disturbances, animal dispersal or through human transportation. A Poisson process is used to approximate the number of long-distance dispersal events that are generated from an occupied cell, while the user can parametrise a Cauchy (Shaw, 1995) or exponential (Mollison, 1972) dispersal kernel to determine the distances travelled from the occupied site (Nathan *et al.*, 2002; Levin *et al.*, 2003). Finally, a uniform distribution in the range of $[0, 2\pi]$ is sampled to determine the direction of each generated long-distance event (Pitt, 2008). Currently, the model does not allow one

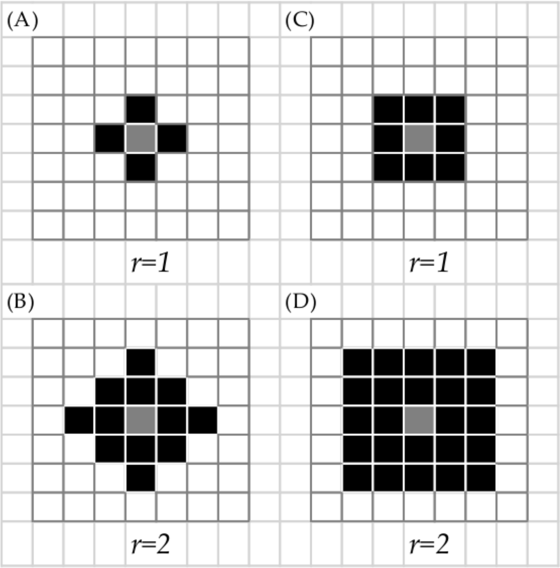

Figure 2.1: From Senay (2014), with permission. Local dispersal neighbourhoods, (A) Von Neumann shape with radius =1, (B) Von Neumann shape with radius = 2, (C) Moore shape with radius = 1, (D) Moore shape with radius = 2.

to explicitly explore the relative contribution of multiple vectors to particular dispersal pathways. While representing such pathways explicitly would generate more realistic long-distance dispersal, it is usually very difficult to find such detailed information empirically. If such information was available, then dispersal kernels characterised by mixed probability distributions could be used to model them (Gilbert *et al.*, 2004). Both local and long-dispersal events can be programmed based on knowledge about the dispersal ability of the species under study collected from the literature or from experimental studies, such as capture-recapture data.

2.2.3.3 Dispersal survival

The survival module, 'r.mdig.survival', allows the species-landscape interactions to be incorporated. The user specifies a habitat suitability map, that can be either realistic maps based on known habitat suitability generated in GIS (Pitt, 2008) or artificial maps produced by a landscape generator, in the form of survival probability map ranging from 0-1 reflecting the difficulty that populations have establishing in each raster cell. The framework is totally asynchronous: the individuals modelled through the local dispersal and dispersal kernel modules are passed through the survival module to determine the population in each cell surviving to the next simulation step based on the underlying suitability value. It is also possible to provide a single survival value if the landscape is homogeneous, such as with a monoculture in an agricultural field or glasshouse (Pitt, 2008).

2.2.4 Modelling habitat suitability

2.2.4.1 Generating habitat suitability maps or the survival layer

There are numerous ways to create habitat suitability maps and they can be based on a wide range of data (Guisan & Zimmermann, 2000; Guisan & Thuiller, 2005). The most common methods are based on static approaches linking current species distribution to environmental variables such as climate, elevation, biotic interactions, vegetation or human disturbance (Guisan & Thuiller, 2005; Thuiller *et al.*, 2008). Others used phenological models (Régnière & Nealis, 2002; Pitt *et al.*, 2007) or expert opinion (Harris, 2002). In a thorough review of the ecological principles and assumptions underpinning habitat suitability modelling, Guisan & Thuiller (2005), Araújo & Guisan (2006) and Elith *et al.* (2009) have all highlighted the key steps in good habitat-suitability-modelling practice including gathering the relevant data, dealing with correlated predictor variables, selecting an appropriate modelling algorithm, fitting and evaluating the model performance and predictive performance. A number of programs, such as the python-based ArcGIS toolbox – SDMtoolbox developed by Brown (2014) or the open-source R package *dismo* developed by Hijmans *et al.* (2016), provide ample opportunity to apply and improve existing approaches and ultimately to develop new ones.

2.2.4.2 Computationally generated landscapes

Another option for creating habitat suitability maps comprises using a landscape simulator that provides a framework for generating replicated landscape patterns with partially controllable spatial properties, particularly with respect to their composition and configuration of components (With & King, 1997; Turner, 1990). When combined with a population dynamic model such as MDiG, these artificial landscapes serve as a template to systematically investigate the effect of landscape structure in fragmented and heterogeneous landscapes (Turner, 2005). The successful application of computer generated landscapes has also led to the development of software designed to create them using a variety of algorithms. Examples include standalone software such as RULE (Gardner, 1999), the subsequent QRULE (Gardner & Urban, 2007), SIMMAP (Saura, 2003), as well as software packages such as a the *ecomodtools* package for R (Chipperfield *et al.*, 2011) or the python-based NLMpy package (Etherington *et al.*, 2014). These frameworks provide easy integration with GIS data and can be integrated within MDiG, providing outstanding opportunities for the design of models ranging from a very simple static spatial establishment and spread model to very complex dynamic ones.

2.2.4.3 Characterization of landscape structure

The ability to quantitatively describe landscape structure is a prerequisite to detect changes and to investigate the relationship between landscape structure and, demographic and dispersal processes. The structure of the landscape – that is the composition and configuration of its components – can be described and quantified by means of landscape metrics for the purpose of understanding the influence of different landscape components such as habitat size, shape, and abundance on demographic and dispersal processes. The plethora of metrics available means that an exhaustive review of all published metrics is beyond the scope of this study. To date, the most comprehensive overview of formulae and domains of traditional metrics has been provided by McGarigal *et al.* (2012). The general perception is that there are three important problems associated with the use of landscape metrics. They are, 1) a high degree of correlation in between the metrics themselves, 2) ambiguous responses to different spatial processes, and, 3) sensitivity to changes in spatial scale (resolution and extent). Quantifying the specific effect of habitat configuration on spread success, for example, is difficult because many configuration metrics are correlated with habitat amount (Kupfer, 2012). Such limitations can often be addressed, or put in perspective, through careful data manipulation, analysis and interpretation (Kupfer, 2012; McGarigal *et al.*, 2012; Uuemaa *et al.*, 2013; Lustig *et al.*, 2015). Landscape metrics can be used in conjunction with both geospatial data and computationally generated landscapes, allowing for comparison between the two sources of data. To calculate landscape metrics, computer programs have been developed such as Fragstats (McGarigal *et al.*, 2012), the python-based vLATE and Patch Analyst 4.1 modules implemented on the very well-known ArcGIS (Rempel *et al.*, 2012), or the plugin LecoS (`http://www.qgis.org/en/site/`) for the Quantum GIS freeware and the two open-sources modules r.le and r.li (Rocchini *et al.*, 2013) implemented in GRASS-GIS.

2.3 Example of applications

2.3.1 The effect of spatial heterogeneity on the establishment and spread of *P. brassicae*

MDiG was developed to model invasive species spread in heterogeneous environments (Pitt, 2008) and has been employed in variety of applications (Kriticos *et al.*, 2008; Pitt *et al.*, 2009, 2011; Worner *et al.*, In Press). In this study, we use the great white butterfly, *Pieris brassicae*, to illustrate some of the capability of MDiG for modelling dynamic range expansion at a regional scale. In this example, we investigate how different representations of spatial heterogeneity in urban landscapes can change the final projected species' distribution.

P. brassicae is an oligophagous butterfly, native to Europe and Asia that feeds on members of the family Brassicacea, commonly found in home gardens and as crops in agriculture (Phillips *et al.*, 2014). Additionally, New Zealand has a number of threatened native Brassicae species. The species was first detected in Nelson, New Zealand in May 2010, and has since been the focus of intense monitoring and eradication efforts (Phillips *et al.*, 2014). We investigated the spread dynamic of the species in five administrative districts in the South Island, New Zealand, that were either in contact with or near to the locations invaded by *P. brassicae.* These districts were Buller, Tasman, Nelson City, Marlborough and Kaikoura (comprising $12,466$ sq. ha).

For simplicity, we focus on a dynamic presence/absence model. The initial dispersal site was set in a cell close to Nelson port which is suspected to be the site of *P. brassicae* unintentional introduction. The cell resolution was set to 100 m to approximate the median distance of local movement of *P. brassicae* as reviewed in Feltwell (1982). Estimates of the median distance and average frequency of long-distance dispersal events were obtained from the dispersal history of *P. brassicae* in the United Kingdom for which, well referenced temporal presence data were found in Feltwell (1982), Heath *et al.* (1984) and the Global Biodiversity Information Facility (GBIF) database. Two survival layers were developed to investigate the effect of urban landscape structure on invasive species spread. The first survival layer ($Surv_1$) included four data sources: climate suitability, degree days, land cover, and high resolution remotely sensed data (Figure 2.2 - 2.3). The second survival layer ($Surv_2$) included all components used in $Surv_1$, except the high resolution remotely sensed data (Figure 2.3). High resolution remotely sensed data are needed to distinguish highly suitable home gardens, public parks and untended green spaces from human-made structures such as houses and roads. Underestimation of the complexity of the urban landscape could lead to an over-estimation of the spread ability of *P. brassicae*.

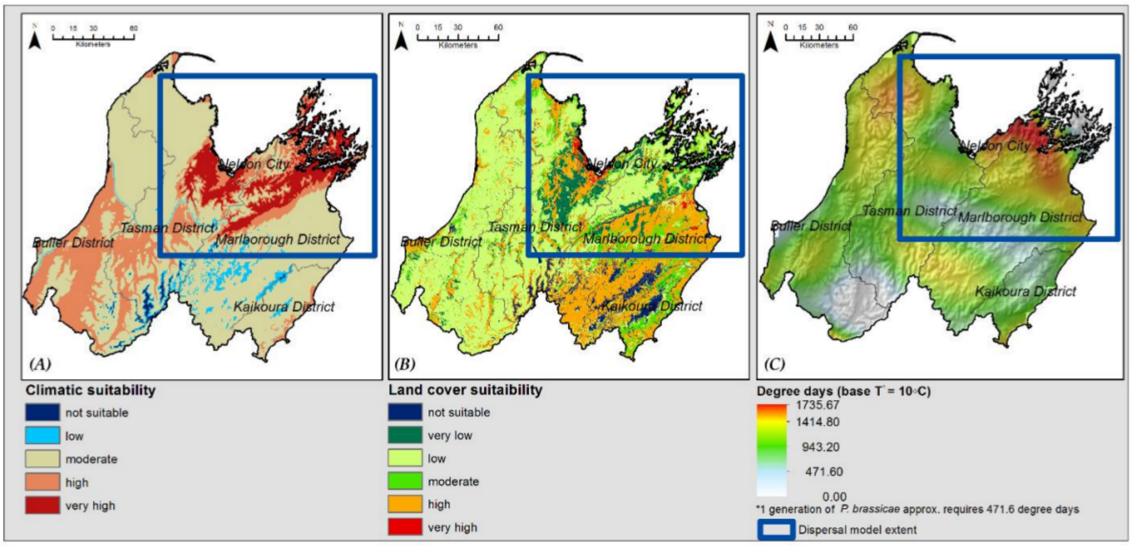

Figure 2.2: From Senay (2014), with permission. Suitability maps used to build the survival layer of *P. brassicae*: (A) hybrid climate model, (B) land-cover suitability layer (C) accumulated growing degree days layer

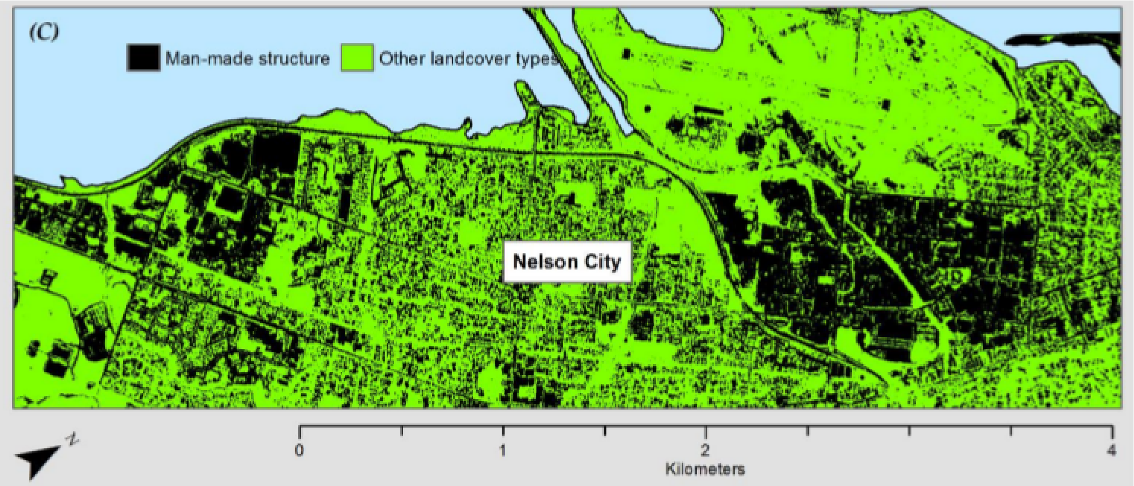

Figure 2.3: From Senay (2014), with permission. Zoom on the high resolution SPOT Map® data used for generating the survival layer $Surv_1$, in which man-made structures can be identified.

Sixteen years of simulations were undertaken representing dispersal from the year 2010 to 2025. The simulation was replicated 1000 times to account for dispersal stochasticity. Three thresholds $[5\%, 10\%, 50\%]$ that corresponded to the number of times a cell was occupied during dispersal for all the replications was used to estimate probability of dispersal into a cell (Pitt *et al.*, 2009). The New Zealand data set of *P. brassicae* detections and absences, obtained from the Department of Conservation (Phillips *et al.*, 2014), was used to compare the first three years of the dispersal occupancy envelopes of both $Surv_1$ and $Surv_2$ dispersal model outputs with field data. Three performance measures - accuracy, sensitivity and specificity - were used to estimate the mean performance of the two dispersal models. Further explanation about the datasets, the parametrization methods and performance measures can be found in Appendix A and Senay (2014).

Both models closely simulated the progression of the Nelson inner city invasion, according to patterns observed from the occurrence data (Figure A.2). For 2011, however, there were more actual occurrences than presence locations predicted by the spread model. This discrepancy could result from high stochasticity in dispersal patterns at early stages of the invasion process (Pitt *et al.*, 2009). Nevertheless, while we assumed that *P. brassicae* was first introduced in New Zealand in 2010, it is highly possible that the species had already completed a generation or two before it was detected, which could explain the more dispersed surveillance data when compared to the few presence locations predicted by the models for the year 2011.

Accounting for the complexity of the urban landscape resulted in a substantial increase in the accuracy (68% versus 27%) and specificity (69% versus 25%) of the model predictions, but a lower sensitivity score (48% versus 83%). The high sensitivity obtained from the model using the survival layer $Surv_2$ was due to the high survival value given to all urban areas. However, unsuitable sites were also incorrectly labelled as suitable, so specificity was low. Higher precision in mapping unsuitable patches among highly suitable urban areas slowed dispersal. When the actual dispersal maps are compared there is an apparent delay in occupation of suitable areas when the first survival layer ($Surv_1$) is used. By 2020, for example, the model using the survival layer $Surv_2$ predicted high-risk areas of invasion ($> 50\%$) that reached Renwick and Blenheim, and covered extensive areas beyond the Wairau valley in the Marlborough district; in contrast, the model using the survival layer $Surv_1$ predicted only limited dispersal within the $10 - 50\%$ envelope that reached beyond Wairau valley (Figure 2.4). It is also notable, that by the end of 2025, high-risk areas predicted by the model using the survival layer $Surv_2$ covered extensive areas in the bays, islands and peninsulas of Marlborough Sounds while these areas were still not covered by the high percentage envelope generated by the model using the survival layer $Surv_1$ (Figure 2.4).

2.3.2 Impact of landscape structure on mean population size and mean dispersal distances

MDiG enables a broad range of hypotheses related to single and interaction effects of habitat change, spatio-temporal scale (spatial extent and resolution), propagule pressure, demography and dispersal, on species abundance, spread and persistence, to be investigated. In this section, we illustrate MDiG's capability further by investigating how change in habitat composition and configuration can affect population density and the dispersal abilities of the Asian gypsy moth, *L.dispsar*.

L. dispar is a notorious insect defoliator, native to Europe and Asia, feeding on over 500 tree and shrub species. The European strain was accidentally introduced to North America near Boston, USA, in 1869 and has subsequently invaded much of the susceptible forests of north-eastern America (Tobin & Blackburn, 2007). We used the computer

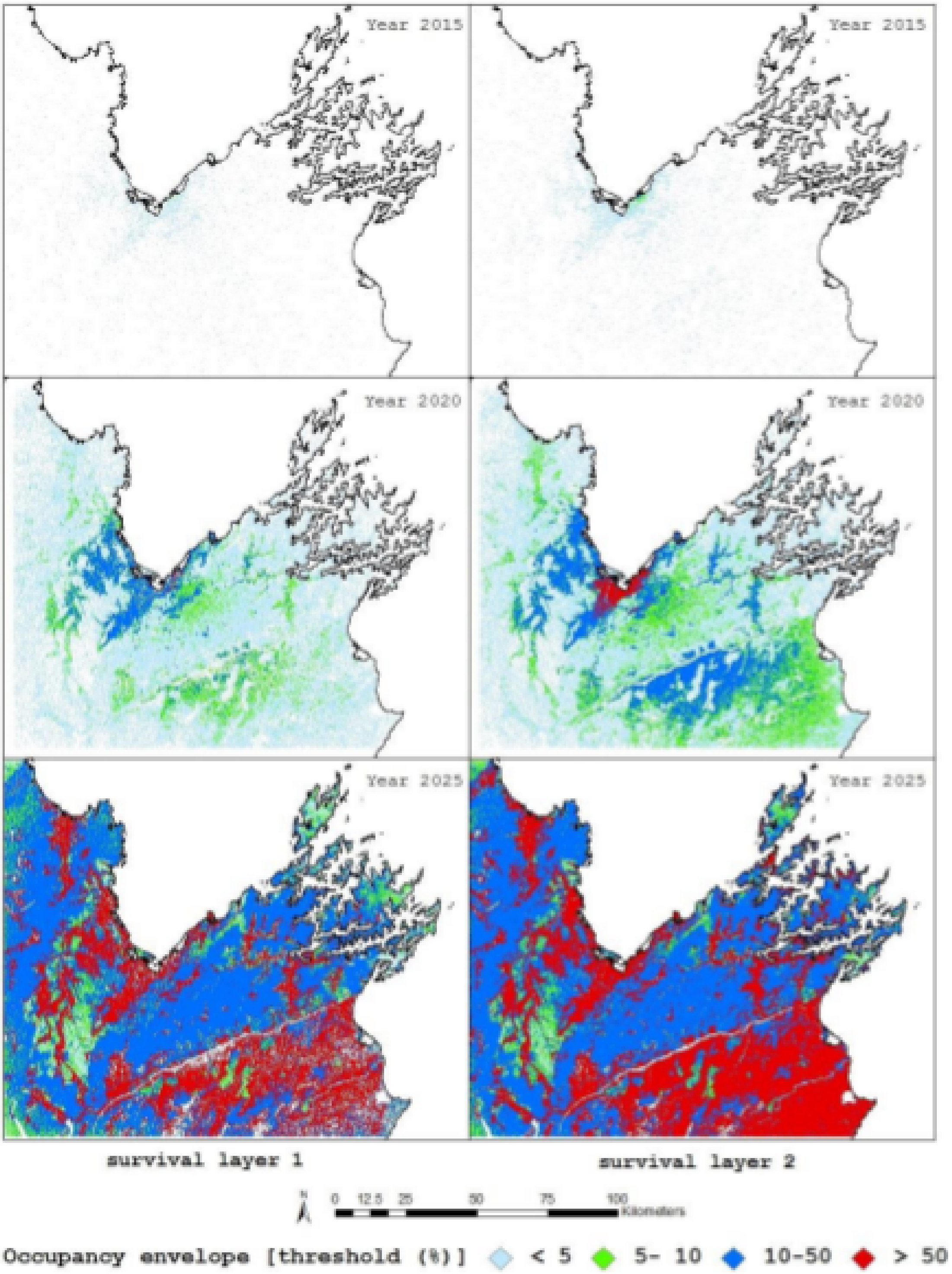

Figure 2.4: From Senay (2014), with permission. Dispersal coverage for the year 2015, 2020 and 2025 based on survival layer $Surv_1$ (left panel) and survival layer $Surv_2$ (Right panel)

program Qrule 4.2 to generate binary (suitable, unsuitable) landscapes, in which habitat configuration (described here as the degree of spatial autocorrelation, $H = 0.3, 0.5, 0.7$) and habitat amount (described here as the percentage of suitable habitat cover, $P = 35\%, 55\%, 75\%$) can be systematically and independently controlled (Gardner & Urban, 2007) (Figure 2.5). The extent of the study covered 128 × 128 raster grid cells (13,384 km^2). The cell resolution was set to 10 km to approximate the median distance of local movements of *L. dispar* as shown in Johnson *et al.* (2006) and Liebhold *et al.* (1992). The average frequency of long-dispersal events emerging from each occupied cell was drawn from a Poisson process, while the median distance travelled was approximated by a Cauchy distribution. The functions were fit based on a thorough review of dispersal abilities of Jankovic & Petrovskii (2013). The initial dispersal site was arbitrarily set to the biggest habitat patch in the landscape. Following Johnson *et al.* (2006) and Liebhold & Bascompte (2003), the local density of *L. dispar* (population density per raster cell) was approximated by a deterministic Allee logistic growth model. Further explanations about the datasets and parametrization methods can be found in Appendix A.

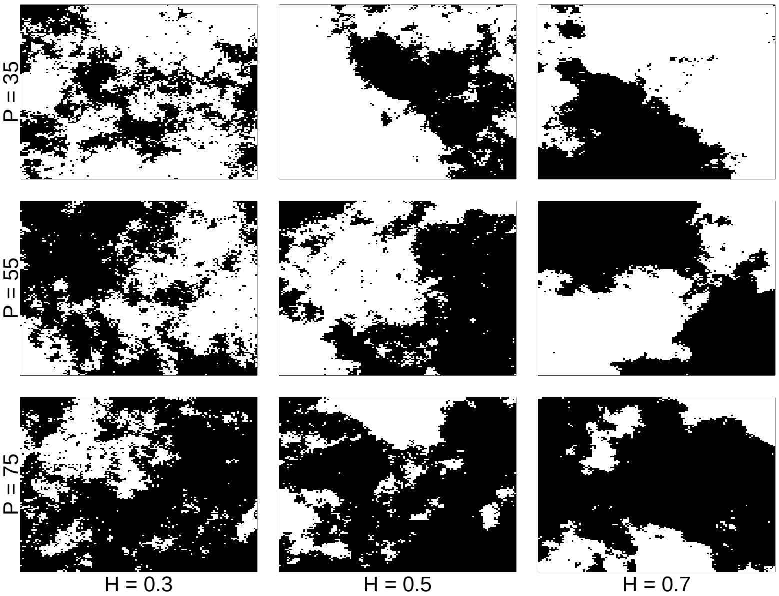

Figure 2.5: Example of survival layers used in *L. dispar* dispersal model. The landscapes were simulated across a three-step gradient of habitat fragmentation (H) and a three-step gradient of habitat amount (P).

For each landscape, the species was allowed to expand its range for 30 years and simulations were replicated 10 times to account for dispersal stochasticity. Each landscape was characterized by two commonly used landscape metrics: a measure of the percentage of suitable habitat cover in the landscape (PLAND), that describe habitat composition, and a measure of the degree of connectedness among suitable patches on a landscape (CONNECT), that describes habitat configuration. These two metrics were calculated by the software FRAGSTAT 4.2 (McGarigal *et al.*, 2012). Finally, MDiG was used to quantify the dependency of population density (d) and rate of spread (ROS) on the landscape characteristics (PLAND and CONNECT), the population traits dispersal ability ($dist$) and intrinsic rate of increase (r) (See details in Appendix A).

For a single introduction of five individuals into the landscape, the simulated average population density (average no. individuals per raster cell) remains relatively low during the first 10 years before the population density increases exponentially towards the habitat carrying capacity ($K = 100$) (Figure 2.6). The rate of spread (average no. of new sites occupied per time step) is characterised by an initial phase with a relatively low spread. In particular, when population density is low, if a large proportion of available sites are colonized (for example there is a large number of newly colonized sites at $t = 5$ and 6), few empty sites can be colonized in the next time step (i.e., there are few newly colonized sites at $t = 6$ and 7), and so spread is reduced. New sites are further colonized only after the population locally grows in the newly invaded sites ($t = 10$) and starts to produce new propagules that can sustain the wave of advancement (Figure 2.6). This sequence of lower-higher rate of spread occurs throughout the invasion process. The existence of such dynamics in rate of spread have long been reported and may occur for several reasons. For example, individuals must overcome Allee effects that may constrain growth in newly invaded sites before generating propagules for further invasion, potentially imposing limits on totally unregulated spread. The dynamic was observed for both slow-fast reproducing and dispersal species. However, the time lag before the rate of spread exponentially increases, is longer for species with lower intrinsic rate of increase as expected (Figure 2.6).

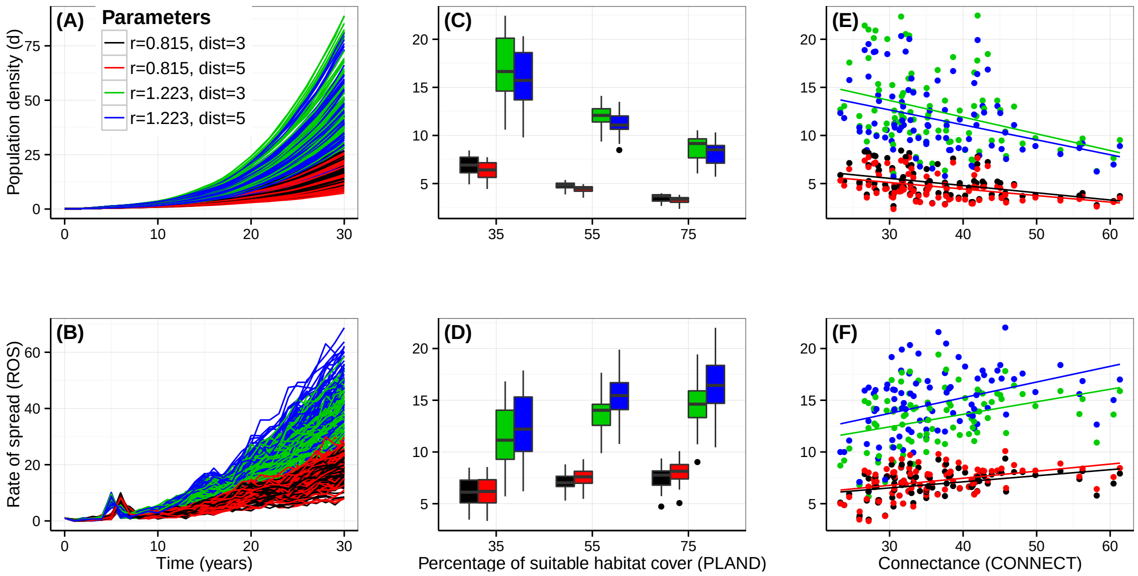

Figure 2.6: European gypsy moth, *L. dispar*, range expansion in changing environments. The first column shows (A) the population density and (B) the rate of spread (no. of new occupied cells/year) over a period of 30 years. Each time series depicts one of the 10 replicate runs for a different combination of intrinsic rate of increase (r) of the species and the median distance ($dist$) of long-distance dispersal event. In the middle, (C) the average population density (d) and (D) rate of spread (ROS) for each combination of r and $dist$ is represented as a function of the percentage of suitable habitat in the landscapes (PLAND), while the column on the right (E-F) represents the same output as a function of the connectivity index (CONNECT).

Increasing the percentage of suitable habitat cover in the landscape (PLAND) and the degree of connectedness among suitable patches on a landscape (CONNECT) resulted, on average, in a more prolific spread but reduced the local density of the population (Figure 2.6). This asymmetry in the response to changes in the structure of the landscape suggests that species that have limited dispersal opportunities tend to maximise their populations locally but will be limited for establishing a population over a large area. On the other hand, species that have high dispersal opportunities may spread but face the added risk of not establishing or going extinct due to lower population density and consequent Allee effects. Interestingly, the response of the species to change in habitat structure was independent of the intrinsic dispersal abilities of the species. However, species with a higher intrinsic rate of increase (r) systematically outperformed species with a lower rate of intrinsic increase.

2.4 Discussion

The MDiG dispersal framework introduced in this study, is a direct response to a call for integrated dynamic models of establishment and spread of invasive species (Worner, 1994; Huntley *et al.*, 2010; Morales *et al.*, 2010; Schurr *et al.*, 2012; Thuiller *et al.*, 2013). It represents population dynamics and dispersal behaviours within a modelling framework integrated in GRASS-GIS (Pitt, 2008). This simulation approach accommodates for intraspecific variability, spatio-temporal heterogeneity and is partially stochastic, addressing three out of the eight important challenges in dispersal modelling suggested in Robledo-Arnuncio *et al.* (2014). The spatially explicit spread simulator can easily be modified for species with structurally different demographic and dispersal behaviours to generalize the use of the framework to many different taxa. The framework also allows the manipulation of fragmented and/or anthropogenic landscapes to test predictions regarding landscape modifications or regional climate changes for invasive species management and conservation purposes.

In the first example, we illustrated the potential of MDiG for modelling the spatio-temporal spread of a real system using *P. brassicae* as a model species and accounting for habitat heterogeneity. A major challenge for establishment and spread modelling is to decide the appropriate extent, spatio-temporal scale and the complexity of the study. In this study, the effect of spatial resolution on the occupied area predicted was investigated. Recoding the heterogeneity of the urban area to delimit unsuitable man-made structures was necessary to gain accuracy and specificity of projections. Otherwise, the whole urban area would have unrealistically facilitated dispersal by overestimating the dispersal rate. The most important implication of such a result is that overestimating future dispersal might incorrectly discourage an eradication attempt (Senay, 2014). There are many situations where selective recoding of the landscape could be applied. For example, if the target species was a forest pest with a specific host in a diverse forest landscape, then, keeping all other land cover data constant, and using a medium resolution but hyper-spectral image of forest areas could be used to map the particular host species by giving host trees higher survival probability than other tree species. Utilizing a modelling platform such as MDiG, which can account for such spatial complexity, is crucial to assess the reliability of projections at different spatio-temporal scales and as a basis for improved invasive species management planning.

The second example was illustrative of the many potential theoretical applications of MDiG. We investigated how changes in landscape composition and configuration can affect population density and the dispersal abilities of an invasive specie. Inferring the relative role of habitat structure on the course of invasion relies on a judicious combination of methods, particularly the use of a landscape generator and landscape metrics that can be combined in an iterative process. The analysis was approached from a case-specific viewpoint, parametrising a population sub-model based on detailed demographic and dispersal attributes of the well studied European gypsy moth, *L. dispar*. In this way, our model had sufficient biological details to reproduce realistic parameter ranges. MDiG was then used to explore different ranges of demographic and dispersal behaviours in changing environments. This particular study exemplified the likely complexity of demographic and dispersal responses to environmental changes and highlighted how a model such as MDiG, with the capability of providing raster cell-based outputs at high spatio-temporal resolution, can help in gain better understanding of these dynamics. The role of the landscape structure in biological invasions has not yet been formally incorporated into risk analysis and management of biological invasions. Most efforts related to preventing and controlling invasions are conducted at the local scale, where the influence of land-use on invasion is seldom explored (Vilà & Ibáñez, 2011). In particular, understanding the influence of landscape composition, configuration, and their spatio-temporal dynamics on ecosystem susceptibility to invasion, rests on broad generalization (Vilà & Ibáñez, 2011). For example, it is well recognised that ecological corridors that facilitate dispersal increase invasion risks or that anthropogenic ecosystems are generally more invaded than natural ecosystems. However, there are differences in the abundance and rate of spread of invasive species even within a particular land-use type, as growth and dispersal can be influenced by small-grain landscape differences such as as man-made structures in urban areas (Vilà & Ibáñez, 2011; González-Moreno *et al.*, 2013a; Senay, 2014), habitat edges (Bartuszevige & Gorchov, 2006), and habitat connectivity (Thiele *et al.*, 2008) variables. Accurate accounts of the invasion process and effective conservation programs will depend on such considerations.

The design of MDiG exploits recent advances in population dynamics and dispersal theory of invasive species. Demography and dispersal processes are clearly key determinants of species' spatial dynamics and responses to rapid environmental changes. However, insufficient representation of dispersal at the landscape scale is still a major limitation in many approaches used for species distribution modelling (Baguette & Van Dyck, 2007; Clobert *et al.*, 2012; Travis *et al.*, 2013). In contrast, MDiG allows demography and dispersal to be modelled explicitly to explore different context dependencies such as density dependence, and responses to landscape structure. What is encouraging, the availability of movement data is increasing in number rapidly, in particular long-distance dispersal in heterogeneous landscapes as well as meaningful characterisation of average growth and dispersal patterns across temporal scales (Cagnacci *et al.*, 2010; Morales *et al.*, 2010; Robledo-Arnuncio *et al.*, 2014). High-quality data on how multiple species grow and move across complex landscapes will provide data to better parametrize the framework. In turn, the MDiG framework can help to generate hypotheses to be tested empirically and determine how these hypothesis scale over time and space.

Chapter 3

Towards more predictable and consistent landscape metrics across spatial scales

Contribution of authors

This chapter describes a collaboration between the current author with Mariona Roigé, who helped conceive the self-organising map (SOM) analysis and, Associate Professors Daniel B. Stouffer and Susan P. Worner, who provided statistical advice and comments on the manuscript.
The results of this chapter are published as: Lustig, A., Stouffer, D. B., Roigé, M. & Worner, S. P. (2015). Towards more predictable and consistent landscape metrics across spatial scales. *Ecological Indicators*, 57, 11-21.

Abstract

Habitat change and fragmentation are considered key drivers of environmental change and biodiversity loss. To understand and mitigate the effects of such spatial disturbances on biological systems, it is critical to quantify changes in landscape pattern. However, the characterization of spatial patterns remains complicated in part because most widely used landscape metrics vary with the amount of usable habitat available in the landscape, and vary with the scale of the spatial data used to calculate them. In this study, we investigate the nature of the relationship between intrinsic characteristics of spatial pattern and extrinsic scale-dependent factors that affect the characterization of landscape patterns. To do so, we used techniques from modern multivariate statistics to disentangle widely used landscape metrics with respect to four landscape components: extent (E), resolution (R), percentage of suitable habitat cover (P), and spatial autocorrelation level (H). Our results highlight those metrics that are less sensitive to change in spatial scale and those that are less correlated. We found, however, significant and complex interactions between intrinsic and extrinsic characteristics of landscape patterns that will always complicate researcher's ability to isolate purely landscape pattern driven effects from the effects of changing spatial scale. As such, our study illustrates the need for a more systematic investigation of the relationship between intrinsic characteristics and extrinsic properties to accurately characterize observed landscape patterns.

Keywords

Landscape patterns, landscape metrics, multi-scale analysis, PERMANOVA, SOM, interactions

3.1 Introduction

Landscape pattern refers to discrete landscape features of an ecosystem (composition) and their spatial arrangement (configuration) within the landscape. Biotic and abiotic determinants, as well as human activities, have been shown as driving forces that shape landscape patterns (Turner, 1990). Furthermore, the rate, extent and magnitude of human alteration of the earth's terrestrial surface is greater now than ever in history, driving unprecedented change in ecosystem processes (Lambin *et al.*, 2001). Such changes range from biodiversity loss and climate change to important modification of ecosystem services (Foley *et al.*, 2005). Accurately quantifying and characterizing landscape pattern has therefore become a major priority for addressing a wide range of spatial analysis applications (Turner, 2005).

In this regard, a plethora of quantitative metrics have been developed to ostensibly provide simple quantitative measurements of the composition and configuration of a landscape (Baker & Cai, 1992; McGarigal & Marks, 1995; O'Neill *et al.*, 1988; Turner, 1990). In general, the calculation of these landscape metrics requires the use of a categorical map, often indicating land-cover or land-use. Typically, these metrics are then used to investigate the relationship between landscape pattern and ecological processes, or as an indicator of ecological condition and risk (O'Neill *et al.*, 1997; Uuemaa *et al.*, 2013).

They are also of key importance for identifying or detecting critical spatial and temporal changes in landscape patterns to anticipate abrupt ecological transition (Johnson & Patil, 2007). The outcome of such spatial analyses, however, remains limited by constraints in our ability to quantify the changes in landscape pattern (Turner, 2005; Uuemaa *et al.*, 2013). In particular, the characterization of landscape patterns depends not only on the patterns themselves but also on the way they are represented (Wu, 2013).

Multiple scale-dependent factors can affect the characterization of a landscape pattern. For example, most landscape metrics are sensitive to changes in the resolution (grain size) of the spatial data (Wiens, 1989; He *et al.*, 2000; Wickham & Rhtters, 1995; Wu, 2004; Saura, 2004; Li *et al.*, 2005; Frohn & Hao, 2006), the extent of the area under investigation (Frohn & Hao, 2006; Li *et al.*, 2005; Saura & Martinez-Millan, 2001; Szabó *et al.*, 2014; Wu, 2004; Wu *et al.*, 2002), or the classification scheme of categorical maps (Bailey *et al.*, 2007; Buyantuyev & Wu, 2007; Castilla *et al.*, 2009; Li *et al.*, 2005; Peng *et al.*, 2010). There are many examples of studies that have investigated the sensitivity of landscape metrics to change in spatial scales (Saura & Martinez-Millan, 2001; Wu, 2004; Wu *et al.*, 2002). Such studies typically target a small set of landscape metrics and base conclusions about the effect of spatial scale on landscape metrics on unique case studies, investigating a single or two scale-dependent factors in isolation (Lechner *et al.*, 2013). Thus far, limited consideration has been given to the vexing question of interaction between scale dependent-factors and change in the landscape patterns (Lechner *et al.*, 2013; Peng *et al.*, 2010).

Additionally, the use and application of landscape metrics is hampered by several characteristics of the metrics themselves (Uuemaa *et al.*, 2013). Many landscape metrics are strongly correlated with the proportion of habitat cover on the landscape (Neel *et al.*, 2004). As a consequence, metrics used to characterize particular aspects of the configuration of the landscape pattern cannot be easily interpreted if the proportion of habitat cover on the landscape is different (Neel *et al.*, 2004; Remmel & Csillag, 2003; Wang & Cumming, 2011). Furthermore, no single metric can fully capture and describe intricate landscape pattern. On the other hand, reducing the number of metrics by correlation and ordination techniques has failed to render the ecological meaning of the latent metric to the practitioner (Turner, 2005). Several suggestions have been made for a minimum set of metrics that capture independent elements of the variation in observed landscape patterns while minimizing redundancy and capturing the desired qualities (Riitters *et al.*, 1995; Cushman *et al.*, 2008). Nonetheless, no general framework exists that permits a particular component of landscape patterns to be unambiguously linked to specific landscape metrics.

To address these persistent challenges, most previous research has been directed toward developing a more rigorous statistical interpretation of landscape metrics. The development of the neutral landscape model (Gardner *et al.*, 1987; With *et al.*, 1997) has provided a framework for generating replicated landscape patterns with partially controllable spatial properties, particularly with respect to their composition and configuration of components (Turner, 2005). Inspection of the relationships among landscape metrics revealed that many were nonlinear and often not monotonic across composition and configuration scenarios (Neel *et al.*, 2004; Remmel & Csillag, 2003). However, most of these studies were limited to maps of the same spatial extent and resolution to avoid the confounding effects of these extrinsic scale-dependent factors. There is a dearth of studies that explicitly assess the relative importance of scale-dependent factors versus changes in intrinsic characteristics of landscape patterns on the characterization of spatial patterns (Estreguil *et al.*, 2014; Lechner *et al.*, 2013). Yet, it is critical to determine whether a change in spatial scale has the same effect in all spatial patterns or whether particular types of spatial patterns (e.g. those with high fragmentation level) are more sensitive to a change in spatial scale than others.

The primary aim of this study was to investigate the nature of the relationship between intrinsic characteristics of spatial patterns and extrinsic scale-dependent factors that affect the characterization of landscape patterns. This research is motivated by the need to identify a set of key generic landscape metrics that enable concise characterization of independent aspects of spatial patterns regardless of the scale at which the patterns are represented (Lindenmayer *et al.*, 2008). In terms of scale-dependent factors affecting the representation of landscape patterns, we investigated the role of spatial resolution (R) and spatial extent (E). These scale-dependent factors were tested in relation to the intrinsic characteristics of the landscape patterns themselves as described by the spatial autocorrelation (H) and the percentage of suitable habitat cover (P). We first tested the statistical significance of the interaction between landscape pattern and scale-dependent factors to assess the magnitude of these interactions and their statistical effect on landscape metrics. Second, we showed how a self-organizing map (SOM) can be used to identify less correlated subsets of landscape metrics thereby providing a robust alternative to traditional ordination techniques.

3.2 Material and methods

3.2.1 Landscape patterns and landscape metrics generation

We used the computer program Qrule 4.2 to generate a wide range of landscape patterns, in which fragmentation (measured as the degree of spatial autocorrelation) and proportion of habitat cover can be systematically and independently controlled (Gardner, 1999; Gardner & Urban, 2007). We considered a binary distinction between suitable and unsuitable habitat type. Qrule uses a midpoint displacement algorithm (Saupe, 1988) to generate multi-fractal maps in which the degree of spatial autocorrelation among adjacent cells (H) can be controlled. We generated landscape patterns in a full factorial design across an 11-step gradient in spatial autocorrelation ($H = 0-1$ in increments of 0.1, 0 being close to random and 1 being completely clustered) and a 10-step gradient in proportion of suitable habitat habitat cover ($P = 5 - 95\%$ in 10% increments) with 100 replicate landscapes for each of the 110 factor combinations (Table 3.1). In order to analyse the influence of spatial extent and resolution on landscape metrics, we generated binary landscape patterns for 40, 20 and 10 m cell size raster and three different extents of 640×640, 1280×1280, and 2560×2560 m^2 (Table 3.1). We used independent realizations for each spatial scale to assure the statistical independence of the estimates corresponding to different resolution and extent.

For each sample landscape, we calculated 101 landscape metrics using the computer program FRAGSTATS 4.2 (McGarigal *et al.*, 2012). The metrics were defined for the suitable habitat cover only and are commonly referred as class-level metrics. McGarigal *et al.* (2012) categorized these metrics into five groups corresponding to the aspect of landscape structure emphasized. These include area/edge/density, shape, core area, con-

Predictor variables	**Measures**
Intrinsic characteristic	
Landscape spatial autocorrelation (Fragmentation)	$H = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1$
Percentage of suitable habitat cover in a binary scheme	$P = 5, 15, 25, 35, 45, 55, 65, 75, 85, 95\%$
Scale-dependent factor	
Spatial extent	$E = 640, 1280, 2560m^2$
Spatial resolution or pixel size	$R = 10, 20, 40m$

Table 3.1: Landscape patterns - list of predictor variables tested and their values.

trast and aggregation (Table 3.2). Metric calculation was based on a 80 m edge depth affecting metrics related to core area, a 400 m search radius affecting metrics based on the distribution of suitable habitat cells within a specified distance of a focal point and an eight-neighbour rule.

3.2.2 Permutational multivariate analysis of variance

To test the null hypothesis of no statistical difference between landscape metrics for four predictor variables, spatial extent (E), spatial resolution (R), percentage of suitable habitat cover (P), and spatial aggregation (H), we used the permutational multivariate analysis of variance (PERMANOVA) (Anderson, 2001). This method provides an alternative to traditional analysis of variance (ANOVA/MANOVA) that is distribution free and insensitive to the correlation among response variables (in our case, landscape metrics). PERMANOVA involves the construction of a distance matrix which reflects the similarity/dissimilarity of each pair of simulated landscape patterns with respect to their landscape metric values (Anderson, 2001). The PERMANOVA analysis is then based on partitioning the multivariate variation of the distance measures. We standardized landscape metric values and used euclidean distances to model the relationships among simulated landscape patterns. The final test statistic is a multivariate analogue to the traditional parametric univariate Fisher's F-ratio (Anderson, 2001). However, the hypothesis (H_0) of no difference among a priori defined groups of landscapes is evaluated using a Fisher's F-test based on sequential sums of squares from 999 unconstrained permutations that randomized the observations (landscape samples) among classification levels.

Response variables	Measures	Range	Response variables	Measures	Range
Area/density/edge metrics			***Core metrics***		
Total suitable area	CA/TA	1.78 - 624.03	Total core area	TCA	0.00 - 531.55
Percentage of Landscape	PLAND	4.34 - 96.01	Core area percentage of landscape	CPLAND	0.00 - 81.11
Number of patches	NP	1.0 - 1355.4	Number of disjunct core areas	NDCA	0.00 - 35.97
Patch density	PD	0.24 - 256.45	Disjunct core area density	DCAD	0.00 - 531.12
Total edge	TE	750 - 31.55×10^4	Core area distribution	CORE_MN	0.00 - 420.04
Edge density	ED	3.76 - 540.04		CORE_AM	0.00 - 531.41
Landscape shape index	LSI	1.22 - 49.19		CORE_MD	0.00 - 399.31
Normalized landscape shape index	nLSI	0.01 - 0.75		CORE_RA	0.00 - 511.58
Largest patch index	LPI	0.62 - 95.98		CORE_SD	0.00 - 198.35
Patch area distribution	AREA_MN	0.03 - 492.58		CORE_CV	0.00 - 2869.97
	AREA_AM	0.18 - 623.65	Disjunct core area distribution	DCORE_MN	0.00 - 488.24
	AREA_MD	0.01 - 467.98		DCORE_AM	0.00 - 531.12
	AREA_RA	0.00 - 622.48		DCORE_MD	0.00 - 482.21
	AREA_SD	0.00 - 243.01		DCORE_RA	0.00 - 421.25
	AREA_CV	0.00 - 2243.21		DCORE_SD	0.00 - 180.97
Radius of gyration distribution	GYRATE_MN	6.93 - 769.62		DCORE_CV	0.00 - 374.41
	GYRATE_AM	17.55 - 976.70	Core area index distribution	CAI_MN	0.00 - 67.33
	GYRATE_MD	5.00 - 732.29		CAI_AM	0.00 - 85.18
	GYRATE_RA	0.00 - 972.27		CAI_MD	0.00 - 66.02
	GYRATE_SD	0.00 - 372.62		CAI_RA	0.00 - 82.18
	GYRATE_CV	0.00 - 587.09		CAI_SD	0.00 - 31.90
				CAI_CV	0.00 - 2858.56
Shape metrics					
Perimeter-area ratio distribution	PARA_MN	79.36 - 3613.27	***Aggregation metrics***		
	PARA_AM	21.55 - 2741.45	Percentage of like adjacencies	PLADJ	17.88 - 99.41
	PARA_MD	79.36 - 4000.00	Clumpiness index	CLUMPY	0.18 - 0.98
	PARA_RA	0.00 - 3972.13	Aggregation index	AI	25.46 - 99.81
	PARA_SD	0.00 - 1521.89	Landscape division index	DIVISION	0.08 - 10.99
	PARA_CV	0.00 - 65.28	Splitting index	SPLIT	1.09 - 21343
Shape index distribution	SHAPE_MN	1.09 - 2.90	Effective mesh size	MESH	0.01 - 593.82
	SHAPE_AM	1.18 - 28.58	Patch cohesion index	COHESION	34.21 - 99.98
	SHAPE_MD	1.00 - 2.74	Connectance index	CONNECT	0.00 - 88.76
	SHAPE_RA	0.00 - 30.24	Proximity index distribution	PROX_MN	0.00 - 12476
	SHAPE_SD	0.00 - 1.94		PROX_AM	0.00 - 1255.84
	SHAPE_CV	0.00 - 131.17		PROX_MD	0.00 - 1255.85
Fractal index distribution	FRAC_MN	1.02 - 1.15		PROX_RA	0.00 - 14864
	FRAC_AM	1.03 - 1.43		PROX_SD	0.00 - 15553
	FRAC_MD	1.00 - 1.14		PROX_CV	0.00 - 5743.96
	FRAC_RA	0.00 - 0.46	Similarity index distribution	SIMI_MN	0.00 - 12476
	FRAC_SD	0.00 - 0.10		SIMI_AM	0.00 - 1255.84
	FRAC_CV	0.00 - 9.19		SIMI_MD	0.00 - 1255.85
Related circumscribing circle distribution	CIRCLE_MN	0.19 - 0.47		SIMI_RA	0.00 - 14864
	CIRCLE_AM	0.38 - 0.71		SIMI_SD	0.00 - 15553
	CIRCLE_MD	0.00 - 0.49		SIMI_CV	0.00 - 5743.96
	CIRCLE_RA	0.00 - 0.87			
	CIRCLE_SD	0.00 - 0.32			
	CIRCLE_CV	0.00 - 151.29			
Contiguity index distribution	CONTIG_MN	0.08 - 0.90			
	CONTIG_AM	0.16 - 0.99			
	CONTIG_MD	0.00 - 0.90			
	CONTIG_RA	0.00 - 0.99			
	CONTIG_SD	0.00 - 0.38			
	CONTIG_CV	0.00 - 195.36			
Contrast metrics					
Contrast-weighted edge density	CWED	6.76 - 540.04			
Total edge contrast index	TECI	23.31 - 99.01			
Edge contrast index distribution	ECON_MN	23.31 - 99.25			
	ECON_AM	23.31 - 98.80			
	ECON_MD	32.31 - 100.00			
	ECON_RA	0.00 - 57.71			
	ECON_SD	0.00 - 23.01			
	ECON_CV	0.00 - 28.14			

Table 3.2: List of response variables tested and their ranges. A detailed description of each metric can be found in McGarigal *et al.* (2012). Twelve metrics characterize the aggregation properties of the patches (cells) belonging to the suitable area only. The first- and second-order statistics are as follows: mean (MN), area-weighted mean (AM), median (MD), range (RA), standard deviation (SD), and coefficient of variation (CV).

First, we investigated the main effects of each predictor variable (E, R, P, and H) independently on the joint response of landscape metrics. Second, we analysed the significance of the two-, three-, and four-way interactions to quantify how E, R, P, and H interact to affect the joint response of landscape metric values. Finally, we conducted separate PERMANOVA analyses for each landscape metric separately. We carried out the PERMANOVA using the function `adonis` in the package `vegan` (`R` Development Core Team 2012).

We further assessed interactions between intrinsic and extrinsic characteristics of landscape patterns using interaction plots. The interaction plots describe the mean of the response variable for each level of one predictor variable (e.g. the percentage of suitable habitat cover, P) plotted over the levels of a second predictor variable (e.g. the spatial autocorrelation, H). The Y axis is the dependent variable. Two-way interaction is indicated by non-parallel lines in the resulting plots.

3.2.3 The self-organizing map clustering technique

To identify less correlated subsets of landscape metrics, we used a self-organizing map (SOM) clustering technique. SOM is a heuristic statistical tool based on methods from machine learning that explores large, complex data sets to detect linear and nonlinear patterns. A detailed description of the SOM methodology can be found in Kohonen (1988). Typically, a SOM is employed as a data reduction and visualisation technique that performs a nonlinear projection of multidimensional data onto a map of nodes. The system learns to represent the input data (landscape metrics) in a way that reflects the statistical structure of the overall collection of the input patterns. Highly correlated landscape metrics are clustered together in specific nodes, whereby nodes that are close together in the map are more alike than nodes that are farther away.

We standardized landscape metrics values. We chose the number of nodes in the output map using the formula $c = 5\sqrt{N}$, where c is the number of nodes and N is the number of landscape metrics (Vesanto *et al.*, 1999). We then displayed the U-matrix (unified distance matrix) to visualize the local Euclidean pair-wise distances between neighbouring nodes. To assess the reliability of the results of the SOM methods, we also estimated two commonly used quality measures. We used the quantization error to provide a measure of how well the map reflects the statistical structure of the overall collection of landscape metrics. The quantization error is equal to the average distance between each of the landscape metrics and its best matching nodes (Kohonen, 2001). Second, we used the topographic error to measure how well the topology is preserved by the map. It measures the proportion of all landscape metrics for which first and second best matching node are not adjacent units (Kohonen, 2001). We carried out the SOM analysis using the SOM Toolbox for Matlab 5 computing environment (Vesanto *et al.*, 1999).

3.3 Results

3.3.1 Multivariate response of landscape metrics

The PERMANOVA provided a novel statistical framework for testing the significance of differences between landscapes. In particular it allowed us to consider the non-linear and complex relations between predictor and response variables. According to the PERMANOVA results, both the extrinsic scale-dependent factors, spatial extent (E) and spatial resolution (R), and the intrinsic characteristics of landscape patterns, percentage of suitable habitat cover (P) and spatial autocorrelation (H), significantly contributed to the variation in the similarity/dissimilarity of landscape patterns (Table 3.3). We observed the same results when testing the independent predictors individually. Overall, the multi-way interactions were significant but showed a smaller contribution to the variation in landscape metric values.

3.3.2 Univariate response of landscape metrics

The individual responses of landscape metrics to different combinations of the four predictor variables E, R, P, and H differed among landscape metrics (Figure 3.1). The main effects of the percentage of suitable habitat cover, P, or the spatial autocorrelation, H, were significant for all landscape metrics indicating that all the selected landscape metrics in this study quantify spatial patterns effectively. In total, only four metrics, the landscape

	DF	SS	MS	Pseudo-F	R^2	P(perm)
P	1	18792	18791.6	744.80	0.18812	0.001
H	1	10921	10921.1	432.86	0.10933	0.001
R	1	11229	11228.8	445.05	0.11241	0.001
E	1	15701	15700.5	622.29	0.15718	0.001
$P \times H$	1	2177	2176.6	86.27	0.02179	0.001
$P \times R$	1	2763	2763.0	109.51	0.02766	0.001
$H \times R$	1	1183	1183.2	46.90	0.01185	0.001
$P \times E$	1	5940	5940.2	235.44	0.05947	0.001
$H \times E$	1	1953	1953.3	77.42	0.01956	0.001
$R \times E$	1	2210	2210.5	87.61	0.02213	0.001
$P \times H \times R$	1	266	266.2	10.55	0.00266	0.001
$P \times H \times E$	1	578	578.1	22.91	0.00579	0.001
$P \times R \times E$	1	1138	1137.9	45.10	0.01139	0.001
$H \times R \times E$	1	353	353.3	14.00	0.00354	0.001
$P \times H \times R \times E$	1	110	110.2	4.37	0.00110	0.003
Residuals	974	24574	25.2		0.25	
Total	989	99889			1.0	

Table 3.3: Results of the multivariate PERMANOVA based on an Euclidean similarity/dissimilarity of spatial patterns in relation to four predictor variables P (percentage of suitable habitat cover), H (spatial autocorrelation), R (resolution), and E (extent) and their interactions. Df is the degrees of freedom; SS is the sum of squares; MS is the mean sum of squares; Pseudo - F value by permutation based on 999 permutations and P(perm) is the P-value (lowest P-value possible is 0.001).

division index (DIVISION), the largest patch index (LPI), the coefficient of variation of the perimeter-area ratio index (PARA_CV) and the splitting index (SPLIT), were not significantly affected by the scale-dependent factors, H and P.

The metrics fell into three groups. First, eighteen of the 101 metrics were significantly affected by spatial autocorrelation, H, but not significantly affected by the proportion of suitable habitat cover, P. Examples include the clumpiness index (CLUMPY) or the edge density (ED) that mainly quantify the spatial configuration of landscape patterns as measured by spatial autocorrelation, H. However, with the exception of the standard deviation of the perimeter-area ratio index (PARA_SD), all metrics in this group were also significantly affected by a change in both spatial resolution, R, and extent, E. Second, fifteen of the 101 metrics were significantly affected by the proportion of suitable habitat cover, P, but not significantly affected by spatial autocorrelation, H. Examples include the effective mesh size (MESH) or the total core area (CA). These metrics mainly quantify the spatial composition of landscape patterns as measured by the percentage of suitable habitat cover, P. With the exception of the landscape division index (DIVISION), all metrics in this group were also significantly affected by either a change in extent, E, only or by a change in both spatial resolution, R, and extent, E. Third, sixty-eight of the 101 metrics responded significantly to changing both the proportion of suitable habitat cover, P, and spatial autocorrelation, H, indicating that the information content of each metric is not a single spatial component, but a complex of several spatial components together. In general, the highest F-ratios were reported for a change in extent, E, and percentage of suitable habitat cover, P. In comparison the change in spatial autocorrelation, H, and in spatial resolution, R, had a smaller effect on the response variables.

3.3.3 Interactions of E, R, P, and H on the univariate response of landscape metrics

The results of the PERMANOVA presented a range of two-, three- and four-way interactions. Two metrics, the normalized shape index (nLSI) and the landscape division index (DIVISION), did not respond significantly to factor interactions. Only a few metrics, such as the edge density (ED) or the largest patch index (LPI), responded to a limited number of two-way interactions. Most of the landscape metrics had larger interactions at two or more levels. Generally, the combination of predictor variables $P \times E$, $P \times R$, and, $P \times H$ had greater effect than other combinations of predictor variables. The combination of predictor variables $H \times R$ had on average the smallest two-way interaction effect. The magnitudes of the three-way and four-way interactions were consistently smaller than the two way interactions.

We further analysed the interaction plots for each of the landscape metrics in order to examine the interactions in more detail. Here we found that the metrics fall into three qualitative groups. (1) Metrics, such as the aggregation index (AI), that presented simple monotonic interaction patterns (Figure 3.2). As an example, an increase in the percentage of suitable habitat cover, P, is always associated with a decrease in AI value. However, this change is more important at smaller extent and higher resolution. The interaction trend remained consistent regardless of the level of the interaction (no cross-over or shape bell were observed). (2) Metrics, such as the coefficient of variation of the related circumscribing circle (CIRCLE_CV), that presented complex interaction patterns (cross-over/bell-shaped) associated with a change in intrinsic characteristics of the landscape, H and P, only (Figure 3.3). While the spatial extent, E, and the spatial resolution, R, had a significant effect on the value of CIRCLE_CV (Table 3.3), an increase in spatial extent and spatial resolution was always associated with an increase in CIRCLE_CV without crossed-over in between spatial scale levels. (3) Metrics, such as the connectance index (CONNECT), that showed significant complex two-way interactions associated with both a change in intrinsic characteristics of the landscape, H and P, and scale-dependent factors, E and R (Figure 3.4).

Figure 3.1: Summary table of the main effect, two-, three- and four-way interactions of the four predictor variables P (percentage of suitable habitat cover), H (spatial autocorrelation), R (resolution), and E (extent). Blank cells indicate no significant interactions (P-value ≥ 0.05), light and dark grey and black cells indicate significant interactions. The highest F-ratio reported for the combination of explanatory variables per response variable is denoted by the black cells. The second and third highest F-ratio are respectively denoted by dark and light grey cells.

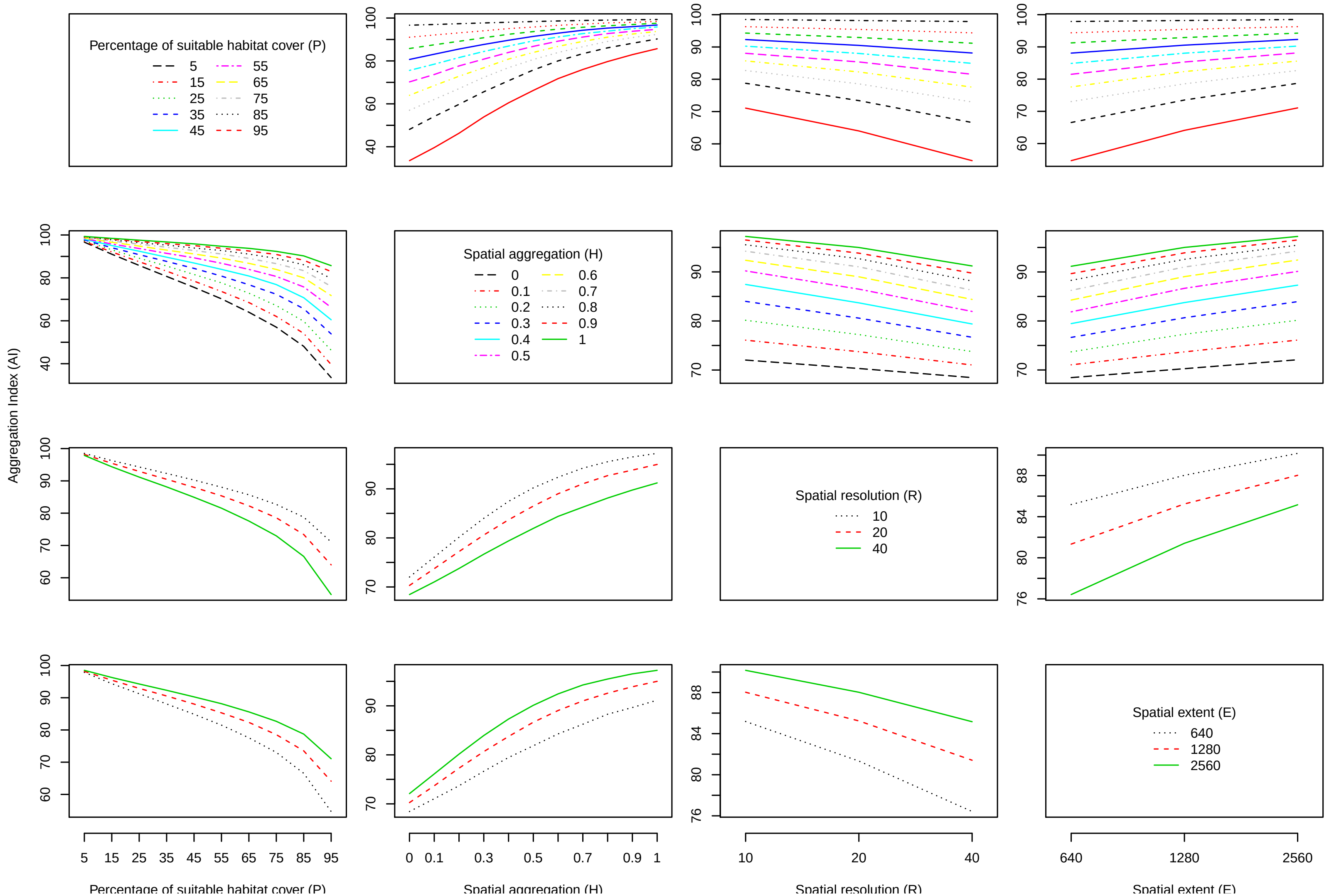

Figure 3.2: Two-way interaction plot of the the aggregation index (AI) as a representative example of group 1. The profile plots of group 1 are characterized by non-linear and non-parallel lines indicating possible interactions. However, no complex two-way interactions (bell-shaped curve or cross-over) are observed. The value of the metrics is calculated for all combinations of the four predictor variables P (percentage of suitable habitat cover), H (spatial autocorrelation), R (resolution) and E (extent).

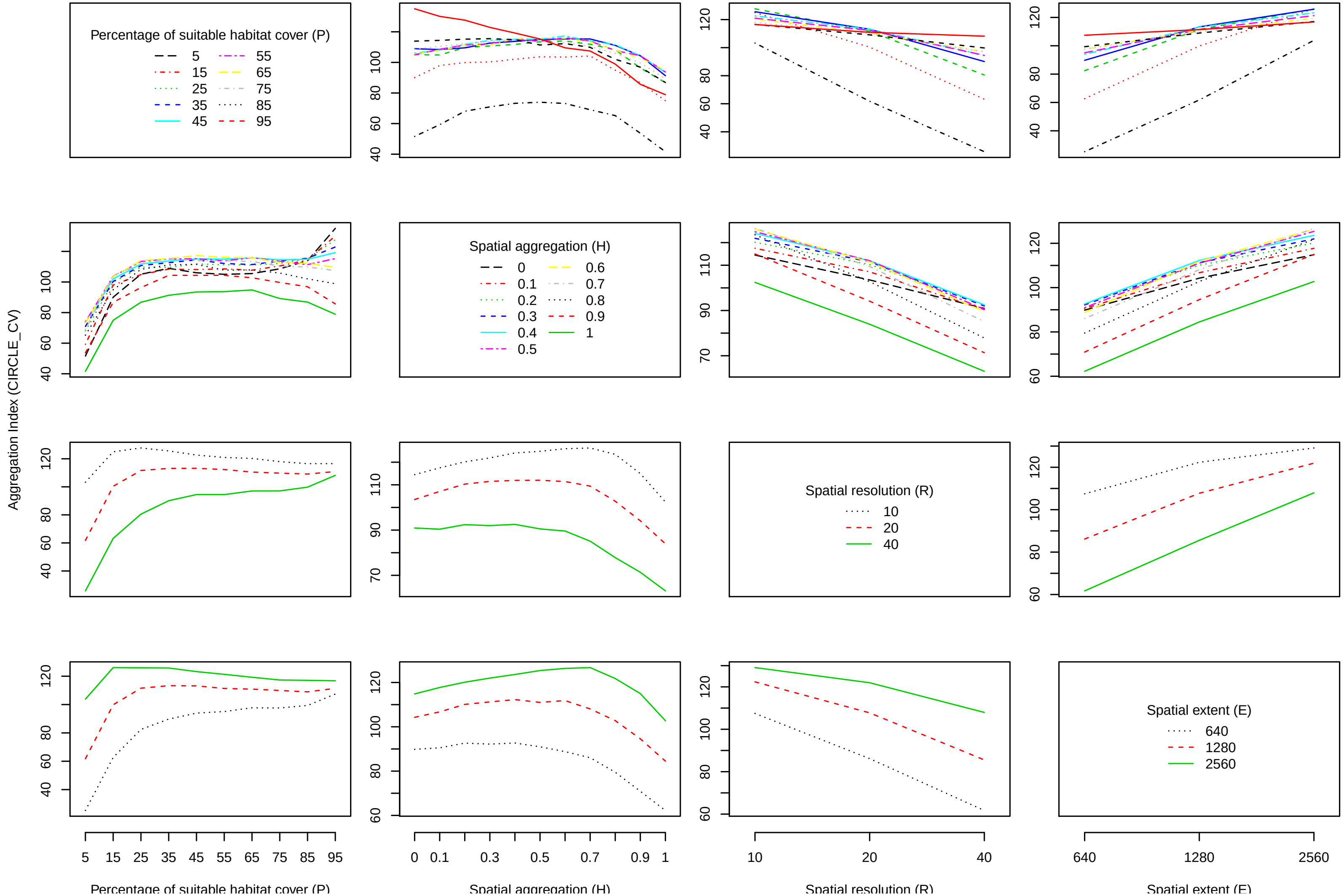

Figure 3.3: Two-way interaction plot for the related circumscribing circle metric as a representative example of group 2. The profile plots of group 2 are characterized by non-linear and non-parallel lines indicating possible significant interactions. No complex two-way interactions are observed when considering changes in R and E (two bottom lines). However, cross-over and a bell-shape curve are observed when considering changes in P and H (two top lines). The value of the metrics is calculated for all combinations of the four predictor variables P (percentage of suitable habitat cover), H (spatial autocorrelation), R (resolution) and E (extent).

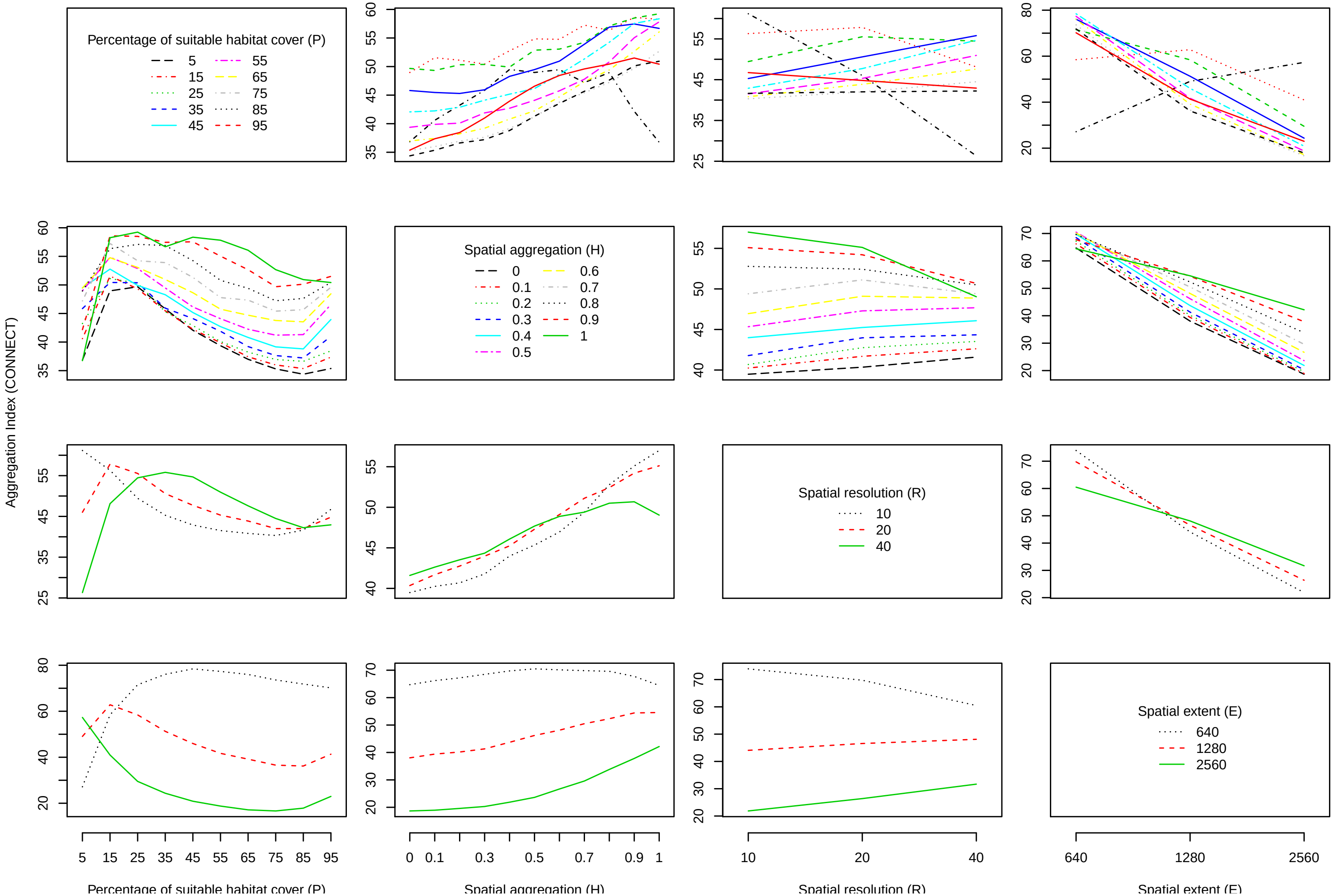

Figure 3.4: Two-way interaction plot for the connectance index as a representative example of group 3. The profile plots of the group 3 are characterized by non-parallel and non-monotonic lines indicating complex interactions. Cross-overs are observed for most combinations of predictor variables. The value of the metrics is calculated for all combination of the four predictor variables P (percentage of suitable habitat cover), H (spatial autocorrelation), R (resolution) and E (extent).

3.3.4 Extraction of landscape metric patterns

We used the SOM map of 4×13 nodes to organise the metrics into fifty-two output units identifying metrics with a similar response pattern associated with a change in landscape intrinsic conditions and spatial scales (Figure 3.5). The clustering of landscape metrics in the SOM analysis was supported by a negligible topographic error indicating a very good preservation of the initial landscape metric topology.

To visualize the effect of the four predictor variables on the classification of the landscape metrics in the SOM, we calculated, for each output node, the average F-ratio of each predictor variables and their two-way interactions (Figure 3.6). Dark nodes represent high values while light nodes represent low values. Globally, the average F-ratio associated with a change in spatial autocorrelation, H, increases from left to right, and the average F-ratio associated with a change in both scale-dependent factors, R and P, increases from the top to the bottom. Looking more closely, the bottom left corner of the map represented metrics strongly affected by spatial resolution, R, or the interaction factor $R \times H$ (Figure 3.6). The bottom right corner of the map identified metrics strongly affected by a change in spatial extent, E, or by the interaction factor $R \times E$. In the upper left, nodes represented metrics most affected by a change in spatial aggregation, H, or the interaction factor $H \times P$. Finally, the upper right identified metrics most affected by the interaction factor $P \times E$.

The U-matrix reveals that metrics strongly affected by a change in spatial scale (bottom of the map) are closer together, in terms of response to change in predictor variables, than metrics located in the upper part of the map (Figure 3.5). The main effect of the spatial extent (E), spatial resolution (R), percentage of suitable habitat cover (P), and spatial aggregation (H) and the two-way interaction factors $P \times H$, $H \times R$ and $P \times E$ showed the clearest gradient among predictor variables. This suggests that the two-way interactions are important factors for the classification of landscape metrics, while the effect of the three-way and four-way interaction have less importance.

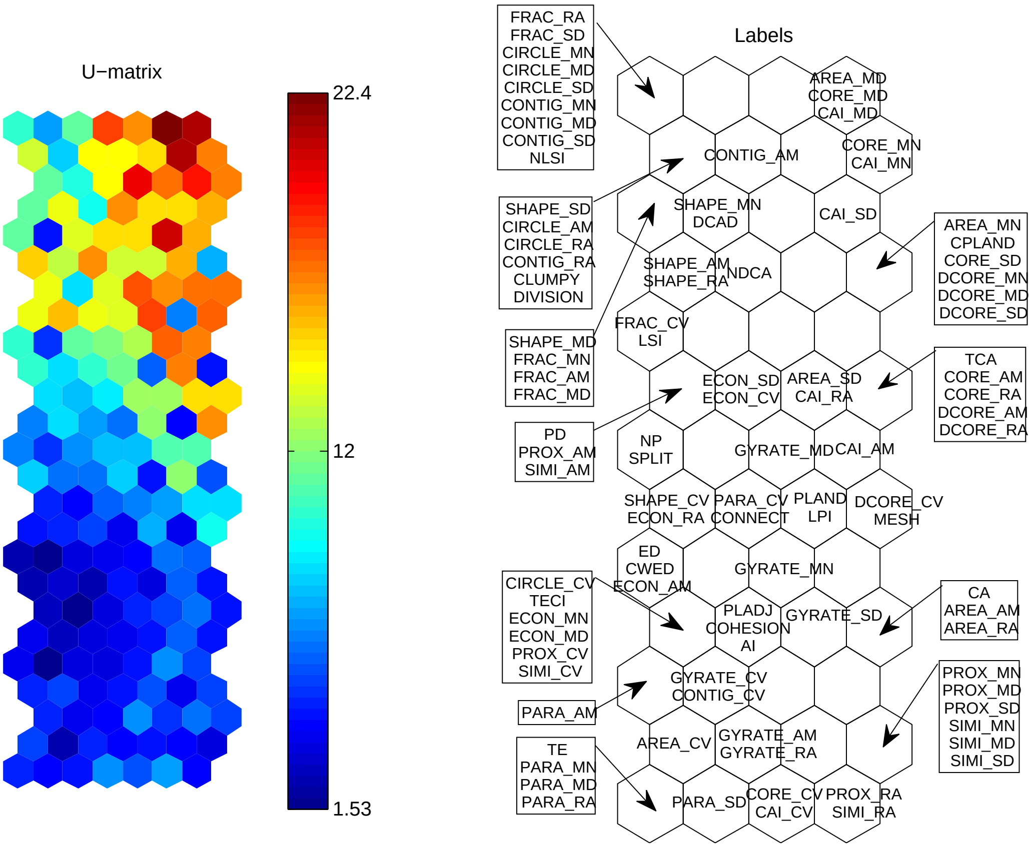

Figure 3.5: Clustering of landscape metrics identified with SOM analysis. The panel on the left is a representation of the local Euclidean pair-wise distances between neighbouring SOM nodes depicted in a color scale image. On the right, the name of the landscape metrics are associated with the SOM nodes.

3.4 Discussion

3.4.1 Determining the significance of differences between spatial patterns

This study provides clear evidence that the percentage of suitable habitat cover, P, and spatial extent, E, accounted for the highest amount of variation of the landscape metrics when compared to spatial autocorrelation, H, and spatial resolution, R. Variation in these two predictors also tended to result in large main effects and large interactions. On the other hand, each landscape metric appeared to have a unique behaviour in terms of the main effect and interactions of the four predictor variables. As a consequence, results are consistent with previous studies that observed high variability in landscape metric

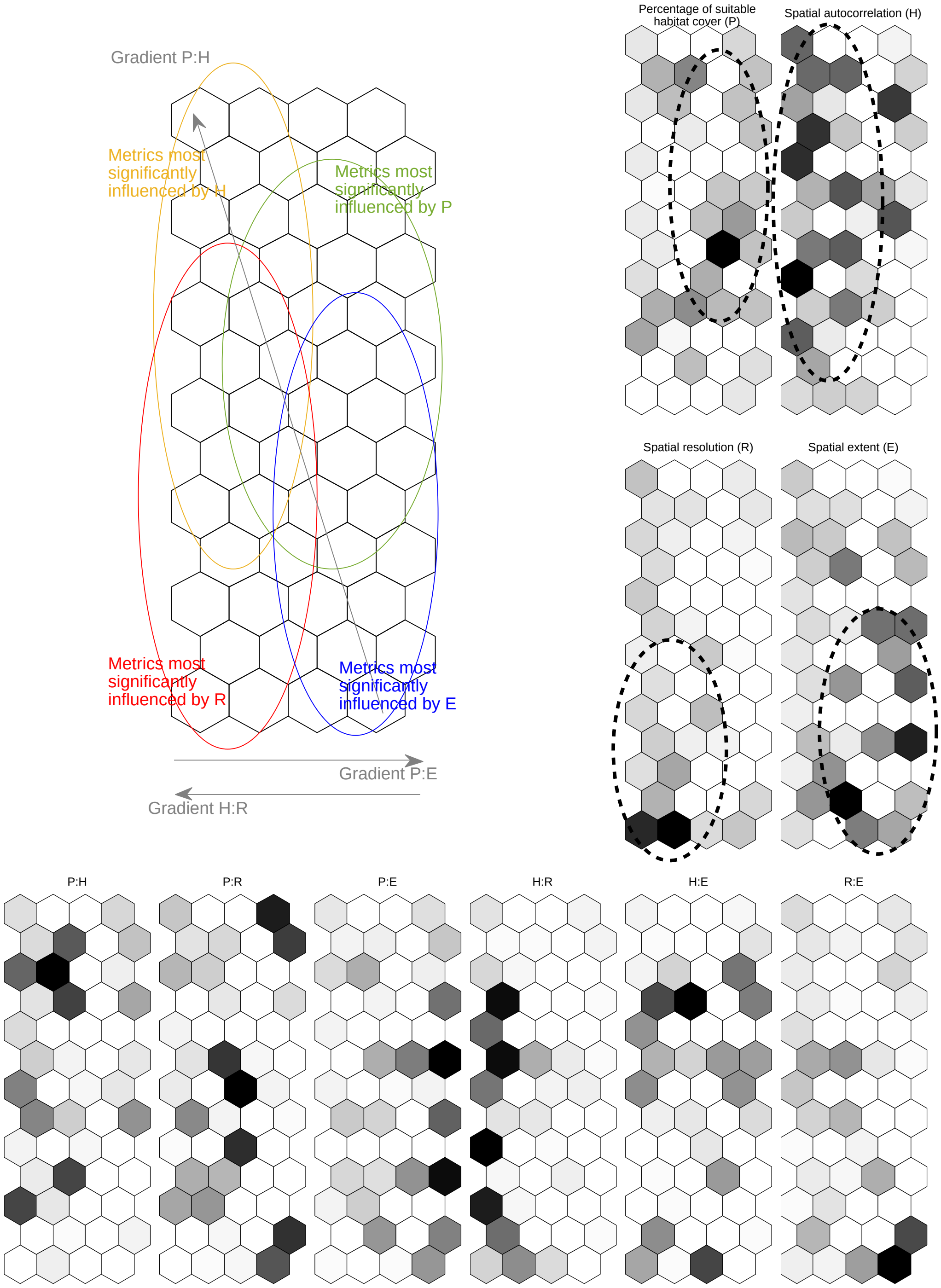

Figure 3.6: Visualization of the average F-ratio obtained in the PERMANOVA analysis on the SOM map. We calculated the mean F-ratio in each node of the SOM map for the main effect of the four predictor variables and their two-way interactions. Dark nodes represent high values while light nodes represent low values. The areas with the highest F-ratio are marked with a dashed circle. The upper-left graph summarizes the effect of the four predictor variables and their two way interactions in clustering landscape metrics.

responses to change in spatial conditions (Lechner *et al.*, 2013; Neel *et al.*, 2004) or spatial scale (Wu *et al.*, 2002; Wu, 2004; Šímová & Gdulová, 2012). More broadly, however, the results of our analysis demonstrated that the way in which landscape metrics capture spatial patterns is as much a property of that pattern intrinsic characteristics as it is of scale-dependent factors.

3.4.2 The importance of interactions in clustering landscape metrics

Our SOM analysis helped identify relationships among metrics that more traditional statistical methods might overlook. It is worth pointing out that the resulting final set of clusters was different than groups of landscape metrics based on conceptual similarity (McGarigal *et al.*, 2012) according to the aspect of landscape they supposedly measure. It is common for the practitioner to select metrics from each of these conceptual groups (patch dominance, patch dispersion, nearest neighbour distance, aggregation, patch complexity, edge, contrast, and neighbourhood similarity) to quantify different aspects of the landscape (Cushman *et al.*, 2008; Estreguil *et al.*, 2014). Our results confirmed those of a previous study that emphasize the importance of also considering the behavioural grouping (responses to change in spatial patterns and scale) of landscape metrics (Neel *et al.*, 2004).

More importantly, landscape metrics constitute a highly redundant data-set. To reduce this redundancy, much research focused on a priori selection of independent landscape metrics by evaluating landscape metric response to change of particular components of the landscape (Estreguil *et al.*, 2014; Neel *et al.*, 2004; Turner, 2005). Our results here suggest that landscape metrics present high-order correlations, indicating a need to consider correlations that involve more than just two features. In particular, the two-way interactions between intrinsic characteristics and scale-dependent factors are important factors in patterning landscape metrics.

3.4.3 Multiple drivers and their interactions

Examination of the interaction plots showed only a small subset of landscape metrics with a monotonic response to interaction factors. For these reduced set of metrics, as long as the final spatial scale is the same, only the magnitude of the difference between the landscapes should change. For example, when comparing two spatial patterns, more heterogeneous spatial patterns will always be characterized as more heterogeneous regardless of the scale of the data. However, most of the metrics presented complex bell-shaped or crossed-over interactions across spatial gradients and spatial scales. This makes extrapolation of responses across spatial scales challenging at best and impossible at worst. The presence of complex interactions indicates that it will be exceedingly difficult to isolate purely landscape pattern driven effects from the effects of changing spatial scale when conducting ecological analyses.

Most previous studies have focused on the main effect of scale-dependent factors and intrinsic characteristics but the interactions among the varied explanatory factors remain poorly understood. The presence of significant complex interactions highlights the importance of considering multiple perspectives for characterizing landscape structure using landscape metrics. For example, one could argue that some landscape metrics exhibit power scaling relationships with change in spatial extent or spatial resolution (Wu, 2004), but this would ignore that numerically speaking, the two-way interactions of scale-dependent factors and intrinsic characteristics are as important as the main effects. Thus, it would be dangerous to conclude from such a biased interpretation about the predictability of landscape metrics behaviour across spatial scale. Such statements are not a criticism of the ability of ecologists to make general predictions about the behaviour of landscape metrics across spatial scales. Rather, our results reaffirm the urgent need to consider the complementary effects of intrinsic and extrinsic characteristics of landscape patterns to conduct integrated landscape pattern assessment (Lechner *et al.*, 2013; Turner, 2005).

3.5 Conclusion

In this study, we tested the adequacy of 101 landscape metrics to quantify concisely independent aspects of spatial pattern regardless of the scale at which they are represented. The findings highlight the sensitivity of landscape metrics to changes in the intrinsic characteristics (spatial autocorrelation, H, and percentage of suitable habitat cover, P), and scale-dependent factors (spatial extent, E, and spatial resolution, R), individually but also to their interactions. Landscape metric classification essentially resulted from the main effect of the predictor variables and their two-way interactions. The presence of significant complex interactions between intrinsic and extrinsic characteristics of landscape patterns makes it difficult to isolate purely landscape pattern driven effects from the effects of changing spatial scale. As such, our study illustrates the need for more systematic investigation of the relationship between intrinsic characteristics and extrinsic properties when accurate characterization of landscape pattern is a key input in spatially-explicit ecological models.

Chapter 4

A comprehensive synthesis of the effect of landscape structure on the spread of invasive species

Contribution of authors

This chapter describes a collaboration between the current author with Doctor Crile Doscher and Associate Professors Daniel B. Stouffer and Susan P. Worner, who provided statistical advice and comments on the manuscript. The results of this chapter are intended to be submitted as follow:
Lustig, A., Stouffer D. B., Doscher C. & Worner, S. P. (2016) A comprehensive synthesis of the effect of landscape structure on the spread of invasive species, *Landscape Ecology.*

Abstract

With accelerated land-use change and habitat fragmentation throughout the world, understanding the relative effects of landscape composition and configuration on biological systems and bioinvasion in particular, is needed to design effective management strategies. However, this topic is poorly understood in part because empirical studies often fail to account for large gradients of habitat complexity and offer insufficient or even no replications across habitat. We construct and explore a spatially-explicit modelling framework that allows for systematic investigation of the impact of changes in landscape composition (i.e. amount of suitable habitat) and landscape configuration (e.g. patch size, shape, juxtaposition, interpatch connectivity or habitat aggregation) on establishment and spread of invasive species. Our results suggest that the presence of an Allee effect, the intrinsic rate of increase, frequency of long-distance dispersal events and proportion of suitable habitat are the primary factors influencing population growth and spread of invasive species.

This research has shown that the presence of an Allee effect leads to a balance between the effectiveness of spread and invasion success. Spread is maximized at an intermediate dispersal level and inhibited at both low and high levels of dispersal. The configuration or composition of the landscape, by either increasing or mitigating or the dispersal abilities of a species, can lead to the rate of spread under a dispersal threshold for which density and spread is at the highest. This study proposes the differences in landscape structure as an additional explanation to the highly variable spread dynamics observed in natural and anthropogenic landscapes. A landscape-scale perspective, which systematically analyses the effect of a core set of landscape structures (proportion of suitable habitat, patch aggregation, patch shape, core area and edge density) on population growth and spread could significantly improve spread risk assessment, impact assessment and control or containment.

Keywords

Invasive insects, heterogeneous landscape, landscape metrics, population dynamics, invasive spread, spatially-explicit individual centred models

4.1 Introduction

Throughout the world, habitat fragmentation and land-use change are among the most critical threats to biodiversity and ecosystems services (Millennium Ecosystem Assessment, 2005; Cardinale *et al.*, 2012). Both processes result in highly heterogeneous landscapes, showing different composition (habitat types and proportion) and configuration (spatial arrangement of habitat types) (Fahrig *et al.*, 2011). Advances in theoretical and empirical approaches to analysing range expansion of invasive species have highlighted the importance of such spatial heterogeneity as a critical factor that can influence the invasion process (e.g. Betts *et al.*, 2014; Bradley, 2010; Hastings *et al.*, 2005; Vilà & Ibáñez, 2011). By interacting with the distribution of invasive species, levels of spatial heterogeneity can either promote the establishment and spread of invasive species (La Morgia *et al.*, 2011; Thies *et al.*, 2011; With, 2002) or, alternatively, can act as a barrier to spread (Jules *et al.*, 2002; Mundt *et al.*, 2011) or host colonization (Zhang & Schlyter, 2004). However, it is not fully understood why some new populations spread rapidly across the landscape while others spread slowly or not at all. Whether the variability observed in nature or in experimental ensembles might be accounted for by systematic differences between landscape structures remains an open research question (Meier *et al.*, 2014; Richter *et al.*, 2013; Sutherland *et al.*, 2013).

Research on landscape-level effects on the establishment and spread of invasive species has focused predominantly on the contribution of natural and semi-natural habitat types surrounding urban, forest and agricultural ecosystems. Human based land-use promotes habitat disturbance and human assisted dispersal that can increase propagule pressure and promote invasive species establishment and spread (González-Moreno *et al.*, 2013b; Pyšek & Richardson, 2010). The urban landscape supports a diverse fauna ranging from native species to opportunistic species (both native and invasive) which exploit modified habitats (Garden *et al.*, 2006). For example, exotic phytophagous insects may find suitable hosts in urban gardens near an airport or port such as the exotic Queensland fruit fly (*Bactrocera tryoni*) found in Whangerei and Auckland, New Zealand in 2011 and 2012. In agricultural and plantation forestry ecosystems, modern methods have generated monocultures that are continuous in their extent and dominated by few crop species, leading to homogenized landscapes (Margosian *et al.*, 2009). Landscape simplification increases the concentration of suitable resources that are available to particular invasive species (Jonsson *et al.*, 2015; Tscharntke *et al.*, 2012; Rigot *et al.*, 2014) and can reduce the number of habitat types that support invasive species' natural enemies (Cardinale *et al.*, 2012; Chaplin-Kramer *et al.*, 2011). Furthermore, the fragmentation of wild-land habitat resulting from agricultural and urban development also affects the spread of invasive species (Harper *et al.*, 2005; González-Moreno *et al.*, 2014; La Morgia *et al.*, 2011). Urban forests and parklands represent an increasing percentage of our remaining near-natural habitats. Because of their proximity to sites of introduction and their (often) large ratio of edge to interior habitat, they are prime habitat for introduced plant and animal species which can then spread into less urbanised areas (Martin *et al.*, 2008). Landscape disturbance features such as hiking trails or roads can also promote invasion spread, in part by creating edges where invasive species can easily establish (Christen & Matlack, 2006);

Syntheses of research across multiple taxa and habitat types offer strong evidence for landscape composition shaping the establishment and spread of invasive species (e.g Robledo-Arnuncio *et al.*, 2014; Tscharntke *et al.*, 2012). Empirical studies are strongly supported by simulation models showing that the local density of invasive species is determined not only by local habitat quality, but also by the spatial structure of the habitat in the surrounding landscape (e.g. With, 2002, 2004). Management strategies for invasive species that target different aspects of host/habitat patterns (e.g., abundance, aggregation, isolation, quality) modified the composition and configuration of the landscape. Understanding the spatially-explicit risks from land manipulation is therefore critical for planning effective land management (Bradley, 2010) and especially strategies to eradicate or contain an invasive species. The outcome of such spatial analysis, however, remains limited by constraints on our ability to carry out a comprehensive quantitative synthesis of landscape effects on the spread of invasive species. In particular, empirical studies often fail to account for large gradients of habitat complexity (e.g. patch size, shape, juxtaposition, inter-patch connectivity or habitat aggregation) and offer insufficient or even no replications across habitat, in part because of an unavoidable trade-off between spatial scope, sampling intensity and accuracy (Robledo-Arnuncio *et al.*, 2014).

Spatially-explicit models, on the other hand, have provided a good basis for generating replicated landscape patterns with partially controllable spatial properties (With, 2004). These models traditionally focus on the effect of a limited number of landscape attributes, such as habitat density or patch connectivity (e.g. Morel-Journel *et al.*, 2015; Sebert-Cuvillier *et al.*, 2008; Smith *et al.*, 2014; With, 2002), where connectivity is classically defined as the degree to which landscape features facilitate or impede the rate of movement of species between habitat patches. Depending on the species considered, these different landscape features have been shown to modify dispersal rates between habitat patches (Calabrese & Fagan, 2004) or the pattern of dispersion (Jonsen & Taylor, 2000). Theoretical studies, such as that by Hanski & Gaggiotti (2004) suggest that as connectivity increases it facilitates the formation of a metapopulations which can increase the persistence of local populations through the dynamics of source-sink population. On the other hand, Smith *et al.* (2014) and Morel-Journel *et al.* (2015) suggested that increasing connectivity during dispersal allows more effective spread, but simultaneously can decrease population density at the source, which can accentuate demographic stochasticity and Allee effects. Morel-Journel *et al.* (2015) further suggested that increasing connectivity during dispersal allows more effective spread, but simultaneously can decrease population density at the source, which can accentuate demographic stochasticity and Allee effects. Such results offer theoretical insights into how landscape structure influences spread and survival and population persistence. However, we are far from having a complete picture about which features of the landscape can affect population dynamics and invasive spread. Furthermore, although there is a substantial literature on demographic analysis of invasive population that helps shed light on specific life-history traits contributing to invasions, only a minority of spatially-explicit models pose demographic processes as essential to a general understanding of invasion. Nevertheless, many studies report that the magnitude and direction of landscape effects on each species depends on that species life-history attributes, sensitivities to environmental disturbances and dispersal abilities (e.g. Guisan & Thuiller, 2005; Huntley *et al.*, 2010; Schurr *et al.*, 2012; Robledo-Arnuncio *et al.*, 2014). Clearly, there is urgent need for more integrative approaches that link demographic processes and dispersal strategies to the landscape to better understand the consequences of increasing anthropogenically driven land-use change, and global environmental and climatic change on species spread (Bocedi *et al.*, 2014; Robledo-Arnuncio *et al.*, 2014; Wang & Jackson, 2014).

The primary aim of this study was to disentangle the independent and interactive effects of landscape composition and landscape configuration on the establishment and spread of invasive species. Critical to this effort is the development of a spatially-explicit model to forecast the spread of species in relation to landscape structure. We first tested the relative importance of the quantity of suitable habitat, its configuration and their potential interactions as predictors of density and rate of spread of gypsy moth as a model invasive species. Second, using a multidimensional scaling ordination technique, we explore the consistency of the effect of landscape composition and configuration on invasive species establishment and spread across species with different life-strategies. This study allowed a generic core of landscape features that significantly affect biological invasion success to be identified. Such features are essential to inform recommendations in pest risk management.

4.2 Material and methods

4.2.1 A spatially-explicit spread model of gypsy moth

The central part of the modelling framework is provided by a spatially-explicit model, MDiG, that represent current understanding of the ecology of invasive insect spread and establishment (Pitt *et al.*, 2009). We used the gypsy moth (*Lymantria dispar dispar*) as a model invasive species. The gypsy moth is native to the temperate forests of Europe and Asia and is a notorious insect defoliator, occasionally causing extensive tree mortality. A European strain was accidentally introduced to North America near Boston, USA in 1869 and has subsequently invaded much of the susceptible forest of north-eastern America (Tobin & Blackburn, 2007). As part of the campaign to slow the gypsy moth spread across the United States, intensive monitoring efforts have been focused around the invasion front. As a result, the spread of the gypsy moth across North America is, perhaps, the most thoroughly studied biological invasion, providing a unique opportunity to explore spatio-temporal variability in patterns and rates of spread (Tobin *et al.*, 2015).

To initialize a simulation, individuals were located on a square lattice (simulation arena) comprised of suitable and unsuitable habitat types. Following Johnson *et al.* (2006) and Liebhold & Bascompte (2003), the local density of gypsy moth (density per raster cell) was approximated by a deterministic Allee logistic growth model:

$$N_{t+1} = N_t \exp\left[r\left(1 - \frac{N_t}{K}\right)\left(\frac{N_t - C}{K}\right)\right] \tag{4.1}$$

where C is the Allee threshold, r the intrinsic growth rate and K the carrying capacity. The values of these parameters were based on previous estimates reported in Johnson *et al.* (2006) and Liebhold & Bascompte (2003) from capture-release data collected from 1988 to 2004 at the invasion front (Table 4.1).

Spatially-explicit rules determine the pattern of local dispersal and generic long distance dispersal spread. For every occupied raster cell (or habitat patch), a proportion of the individuals spread evenly to the surrounding patches. In addition, a Poisson process is used to generate how many long distance dispersal events originate from each habitat patch (Pitt *et al.*, 2009). The frequency of these long-distance dispersal events correspond to the mean of the underlying Poisson distribution and was set to 0.05. The distance travelled by the individuals is approximated by a Cauchy probability distribution as it allows for long, rare dispersal distances which may occur naturally (wind dispersal for example) or which may represent human assisted dispersal (Hastings *et al.*, 2005). Following Johnson *et al.* (2006), we fixed the median dispersal distances of long-distance dispersal events (median of the Cauchy distribution) at 5 raster cells. The direction of dispersal was sampled from a uniform distribution (Pitt *et al.*, 2009).

4.2.2 Generating and quantifying landscape heterogeneity

We generated binary (suitable, unsuitable habitat) landscape patterns for 10 km cell size raster grids of extents 1280×1280 km^2 to approximate local movements of larvae and male adults gypsy moths (Jankovic & Petrovskii, 2013). The computer program Qrule 4.2 (Gardner, 1999) was used to generate a gradient of landscape complexity, in which fragmentation (measured as the degree of spatial autocorrelation) and proportion of habitat cover can be systematically and independently controlled (Gardner & Urban, 2007). Qrule uses a midpoint displacement algorithm (Saupe, 1988) to generate multi-fractal maps in which the degree of spatial autocorrelation among adjacent cells (H) can be controlled. We generated landscape patterns in a full factorial design across a five-steps gradient in spatial autocorrelation ($H = 0.1 - 1$ in increments of 0.2, 0 being close to random and 1 being completely clustered) and a six-steps gradient in proportion of suitable habitat habitat cover ($P = 5 - 95\%$ in 10% increments) with 10 replicate landscapes for each of the 35 factor combinations (Table 4.1).

For each sample landscape, we calculated landscape metrics using the computer program FRAGSTATS (McGarigal *et al.*, 2012). These metrics ostensibly provide simple quantitative measurements of the composition and configuration of a landscape. The landscape metrics were defined for suitable habitat cover only and are commonly referred to as class-level metrics. McGarigal *et al.* (2012) categorized these metrics into seven groups corresponding to the aspect of landscape structure emphasized. These include patch area/density/edge, shape, aggregation, contrast, core, isolation/proximity and connectivity metrics (Table 4.2). In addition, the percentage of the suitable habitat (PLAND), considered the most universal measure of landscape composition, was included in the

Parameters	**Abbreviated code**	**Parameter value**
Population demography		
Growth rate	r	0.815, 1.223*, 1.63 per capacity growth
Allee threshold	C	0, 2*, 5 individuals per raster cell
Carrying capacity	K	30, 50*, 100 individuals per raster cell
Propagule pressure	nbp	5, 10*, 15 individuals per raster cell
long-distance dispersal behaviour		
Median distance	λ	3, 5*, 7 raster cells
Frequency	f	0.01, 0.05*, 0.09
Landscape structure		
Landscape spatial autocorrelation (fragmentation)	H	H= 0,0.3,0.5,0.7,0.9
Percentage of suitable habitat patches in a binary scheme	P	P = 5,15,35,55,75,95

Table 4.1: Parameter values for the stochastic spatially-explicit spread model, MDiG. Parameters are subdivided into the three main model 'components'. Those parameters marked with * indicate the baseline parameter values for a spread model of gypsy moth.

analysis to enable a comparison between the relative effect of landscape configuration, landscape composition and their potential interactions.

4.2.3 Species scenario testing

Insect species can be placed along a 'slow-fast life-history continuum', where changes in population growth rate arise mainly from variability in reproductive rates in fast species and in survival rates for slow species (Herrando-Pérez *et al.*, 2012). The position of a species along this continuum could also reflect the propensity of a species to invade a new environment. Similarly, changes in dispersal distances, frequencies of dispersal events, carrying capacity and Allee threshold are also expected among species. We used MdiG to broadly capture this variation, by creating twenty-seven different combinations of species' life-history attributes and dispersal abilities in a mono-factorial design (Table 4.1). The main reason for doing this was to evaluate how change in the intrinsic growth rate, carrying capacity, Allee effect threshold, and the frequency and median distance of long-distance dispersal events affect the dynamics of spread, and to quantify a general relationship across life-history attributes to changes to landscape variables. For each life-history scenario and each sample landscape, we modelled the fate of a single release of 5 to 20 individuals randomly localized in the largest suitable patch. We varied each parameter over three values (Table 4.1) and studied the effects of these changes on the dynamics of the simulated invasion. For each landscape, the species was allowed to expand its range for 30 years, one model time step represents one year, and simulations were replicated 25 times to account for dispersal stochasticity. A number of response variables were recorded over time: population density – d (number of individuals in the population), the occupied area – OA (number of raster cells occupied/total number of raster cells), the rate of spread – ROS (number of new cells occupied per simulation run), the average dispersal distance – avDist, and the maximum dispersal distance – maxDist.

4.2.4 Statistical analysis

4.2.4.1 Quantifying the effect of landscape metrics as indicators of invasive spread

To investigate the role of landscape metrics as indicators of invasive insect spread, we used a multivariate regression model (Burnham & Anderson, 2002). A multiple model analysis helps to improve inferential properties compared with traditional single model approaches, by allowing an ensemble of plausible candidates to be ranked and integrated using information theory. This approach accounts for the uncertainty associated with a single model by providing confidence interval coverage of variable estimates.

The model outputs comprising average population density (d), occupied area (OA), rate of spread (ROS), average dispersal distance (avDist), and, maximum distance dispersal (maxDist) for each species scenario and landscape structure scenario were analysed using generalized linear models (GLMs) with a Gaussian error distribution for the species scenario as a random factor. The predictor variables included 84 metrics of landscape configuration and composition. Prior to conducting the multivariate regression, the collinearity among landscape metrics was reduced by selecting independent groups of metrics identified in Lustig *et al.* (2015). The degree of mulicollineraty among predictor variables in each group was further assessed by calculating the generalized variance inflation factor (GVIF) for all predictors (Dormann *et al.*, 2013). Predictors giving high GVIF were identified by calculating the pair-wise correlation between all predictors using a non-parametric Spearman's rank correlation. Highly correlated variables were removed from the analysis until all GVIF values were smaller than 10 (Table 4.2).

Metrics	**Abbreviated code**	**Range**	**Description**
Area/density/edge metrics			
Percentage of suitable habitat	PLAND	4.34 − 96.01	
Edge density	ED	3.76 − 540.04	Measure of edge length on a per unit area
Normalized shape index	nLSI	0.01 − 0.75	Measure of total length of edge normalized by the number of cell surface. It provides a simple measure of class aggregation or clumpiness.
Radius of gyration distribution	GYRATE_CV	0.00 − 587.09	Measure of the variance in patch extent (mean distance each cell in the patch). It is effected by both patch size and patch compaction.
Shape metrics			
Perimeter-area ration distribution	PARA_MD	79.36 − 4000.00	Median of perimeter-area ratio (simple measure of shape complexity).
Shape index distribution	SHAPE_MD	1.00 − 2.74	Measure of the overall shape complexity.
Related circumscribing circle distribution	CIRCLE_AM	0.38 − 0.71	Average (AM), median (MA) and range (RA) of patch elongation.
	CIRCLE_RA	0.00 − 0.87	
	CIRCLE_CV	0.00 − 151.29	
Fractal index distribution	FRAC_MD	1.00 − 1.14	Measure that reflect shape complexity (median - MD and coefficient of variation -CV of the fractal dimension) across a range of spatial scale
	FRAC_CV	0.00 − 9.19	
Aggregation metrics			
Splitting index	SPLIT	1.09 − 21343	Measure of the cumulative area distribution, SPLIT = 1 when the landscape consist of one patch and increases as the focal patch is subdivided into smaller patch.
Percentage of like adjacencies	PLADJ	17.88 − 99.41	Frequency with which different pairs of patch types appear side-by-side.
Contrast metrics			
Edge contrast index	ECON_RA	1.09 − 21343	Mediam (MD), standard deviation (SD) and range (RA) of the amount of contrast along the patch perimeter.
	ECON_MD	32.31 − 100.00	
	ECON_SD	0.00 − 23.01	
Core metrics			
Disjunct core area density	DCAD	0.00 − 531.12	Number of disjunct core area on a per unit area basis.
Core are distribution	CORE_MD	0.00 − 399.31	Median area in focal patch greater than a specified depth-of-edge distance
Disjunct core area distribution	DCORE_MD	0.00 − 482.21	Median (MD) and standard deviation (SD) area in disjunct patch greater than a specified depth-of-edge distance C
	DCORE_SD	0.00 − 180.97	
Core are index distribution	CAI_CV	0.00 − 2858.56	Coefficient of variation of patch core area as a percentage of patch area.
Isolation/proximity metrics			
Proximity index distribution	PROX_AM	0.00 − 1255.84	Average (AM) and coefficient of variation (CV) o f the proximity of all patches whose edges area within a specified search radius
	PROX_CV	0.00 − 5743.96	
Connectivity metrics			
Connect index	CONNECT	0.00 − 88.76	Number of functional joining between patches of the corresponding patch type.
	PROX_CV	0.00 − 5743.96	

Table 4.2: List of landscape metrics considered as useful predictors of invasive spread. A detailed description of each metric can be found in McGarigal *et al.* (2012).

We then performed a forward-backward step-wise variable selection procedure based on the Akaike information criterion (AICc) with a correction for finite sample sizes (Burnham & Anderson, 2002). Landscape metrics were mean-centred to facilitate the interpretation of the main effect and the outputs were log transformed to improve the normal distribution of the models' residuals. First, we identified the landscape variables that significantly improved model performance (lowest AICc values). The models that accounted for 95% of the cumulative sum of Akaike weights (Akaike weights represent the likelihood of a model) were retained. Second, we calculated the averaged parameter estimates of the model, and their 95% confidence intervals, using a model averaging algorithm. The relative importance of each landscape metric as a predictor of invasive spread was determined, based on the sum of Akaike weights, with 1 being the most important (the metric variable is present in all model candidates for each species scenario) and 0 the least important. Possible predictor variables were considered to contribute to improve model performance if their AIC weight summation was relatively high ($w > 0.7$).

4.2.4.2 Rank similarities among landscape and species scenarios

Finally, we used the non-metric multidimensional scaling (nMDS) method, to identify groupings in the response data. In essence, values of the response variables (average density, rate of spread, occupied area and dispersal distances) that are nearby in nMDS space are more similar to each other than those further apart. This technique is flexible enough to accommodate for non-linearity in the data. In addition, a 'biplot' was overlaid in the NMDS space to show how the position of different simulation runs in the ordination space are related to the selected landscape metrics and life-history traits. Statistical analyses were performed using the R statistical system v 3.2.2 (R Development Core Team 2013). Model selection for mixed models was conducted using 'lme4' package (Bates *et al.*, 2015) and 'MuMIn' package for model-averaging of coefficients (Bartoń, 2016). The ordination analysis was conducted using the 'vegan' package (Oksanen *et al.*, 2016).

4.3 Results

4.3.1 Aspects of landscape structure influencing the spread of invasive insects

Reduction of collinearity among the landscape metrics that are considered useful predictors of invasive spread, resulted in the selection of 23 metrics of landscape configuration and composition (Table 4.2). These metrics had a pair-wise Spearman correlation lower that 0.8 and a variation inflation factor lower than 10. The selected set of metrics cover all seven pattern aspects identified in McGarigal *et al.* (2012), including area/density/edge, shape, aggregation, contrast, core, isolation/proximity and connectivity metrics.

Across all scenarios, average population density (d), rate of spread (ROS), mean dispersal distance (avDist) and maximum dispersal distance (maxDist) were best predicted by three landscape metrics: proportion of suitable habitat in the landscape (PLAND) - the most common measure of habitat composition; edge density (ED) and normalized shape index (nLSI) – a simple measure of habitat patch shape and clumpiness. These three landscape metrics collectively represent a gradient from areas with few, small and clustered habitat patches towards an area with high inter-dispersion and covered by a large amount of habitat with elongated patches. The effect of the proportion of suitable habitat in the landscape was the most important variable affecting all aspects of spread (Table 4.3). Increase in density, rate of spread, mean dispersal distance and maximum dispersal distance were all positively related to PLAND, while these four components of spread were negatively related to the normalized shape index. An increase in population density, rate of spread and maximum dispersal distance were positively related to edge density, while average dispersal distance was negatively related to edge density.

The percentage of like adjacencies (PLADJ) and the splitting index (SPLIT) are both measures of patch aggregation that were found to significantly affect population density and rate of spread. In addition, the mean circumscribing circle (CIRCLE_AM, a measure of patch elongation), the connectivity index (CONNECT), and the amount of disjunct core areas (DCDA) also had a significant effect on population density and rate of spread. Although the effects were significant, the actual effect size of these three landscape metrics on population density were small, with lower summed Akaike weights (0.01, 0.01 and 0.02 respectively) and model partial slope coefficient near 0 (Table 4.3).

The core metrics, in particular, the median of the disjunct patch area DCORE_MD, only show a significant effect on the mean dispersal distance (avDist) and the maximum dispersal distance (maxDist). A number of shape metrics (FRAC_CV, SHAPE_MD, PARA_MD), core metrics (CAI_CV and DCAD) and aggregation/connectivity metrics (CONNECT and SPLIT) were found to have a significant effect on dispersal distance. However, the effect size of these configuration metrics were relatively small as shown by lower summed Akaike weights and a model partial slope coefficient near 0 for all metrics

Metrics	**Insect density (d)**				**Occupied area (OA)**				**Rate of spread (ROS)**			
	w	β	Lower CI	Upper CI	w	β	Lower CI	Upper CI	w	β	Lower CI	Upper CI
PLAND	1	0.1745	0.161	0.188	1	0.018	0.0170	0.0191	1	0.2232	0.2053	0.2412
ED	0.62	0.0277	0.0146	0.0407					1	0.0279	0.0140	0.0419
nLSI	0.96	-0.0453	-0.057	-0.033	0.6	-0.001	-0.002	-0.0002	1	-0.0432	-0.0587	-0.0278
PLADJ	0.67	0.0647	0.0425	0.0869	0.4	0.001	0.0002	0.002	1	0.0488	0.0258	0.0720
SPLIT	0.59	0.0190	0.0083	0.0297					1	0.0200	0.0098	0.0303
CIRCLE_AM	0.02	0.0149	0.0033	0.0266					0.82	0.0168	0.0057	0.0278
CONNECT	0.01	-0.0099	-0.020	-0.0003					1	-0.0121	-0.0215	-0.0028
DCDA	0.08	0.0129	0.0040	0.0219					0.17	0.0118	0.0034	0.0202
PLAND:SPLIT	1	-1.0682	-1.5140	-0.6223					1	-1.3602	-1.8091	-0.9112
PLAND:CIRCLE_AM	0.08	0.0015	0.0066	0.0333					1	0.0297	0.0210	0.0385
PLAND:ED	0.04	-0.0014	-0.0456	-0.0160					1	-0.0389	-0.0508	-0.0269
PLAND:DCAD	0.03	-0.0003	-0.0212	-0.0008								
PLAND:PLADJ									0.53	0.3554	0.02603	0.6848

Metrics	**Mean dispersal distance (AvDist)**				**Max dispersal distance (MaxDist)**			
	w	β	Lower CI	Upper CI	w	β	Lower CI	Upper CI
PLAND	1	0.0419	0.0384	0.0455	1	0.1948	0.1790	0.2106
ED	1	-0.0028	-0.0048	-0.001	0.99	0.0222	0.0106	0.0336
nLSI	1	-0.0084	-0.0105	-0.0064	1	-0.0486	-0.0640	-0.0333
PLADJ					0.45	0.0288	0.0109	0.0466
SPLIT					0.45	0.0126	0.0059	0.0193
CIRCLE_AM					0.74	0.0138	0.0061	0.0214
CONNECT					0.45	-0.0100	-0.0159	-0.0040
DCDA					0.1	0.0087	0.0020	0.0153
DCORE_MD	1	-0.0043	-0.0066	-0.002	0.85	-0.0148	-0.0226	-0.0070
DCORE_SD	1	-0.0019	-0.0036	-0.0003				
CAI_CV					0.03	-0.0065	-0.0118	-0.0011
SHAPE_MD					0.02	0.0046	0.0005	0.0088
PARA_MD					0.02	-0.0045	-0.008	-0.0004
PLAND:SPLIT	1	-0.1669	-0.2224	-0.1114	1	-1.4083	-1.8890	-0.9275
PLAND:ED					1	-0.0395	-0.0538	-0.0252
PLAND:CIRCLE_AM					0.73	0.0293	0.02013	0.0386
PLAND:DCORE_SD					0.28	-0.0353	-0.0514	-0.0192
PLAND:DCORE_MD					0.18	-0.0171	-0.0318	-0.0024
PLAND: DCDA					0.01	-0.0145	-0.0251	-0.0039
PLAND:CAI_CV					0.01	0.0168	0.0039	0.0298

Table 4.3: Model-average partial regression coefficient (β) and unconditional 95% coefficient intervals (CI) from models of total population density, occupied area, rate of spread, mean dispersal distance and maximum dispersal distance in relation to landscape variation. Coefficients are based on mean-centred data. AICc weights (w) indicate relative importance of variable j based on summing weights across models were variable j occurs.

(Table 4.3).

Furthermore, there was some evidence of significant interactions between the proportion of suitable habitat in the landscape and landscape configuration metrics, which were stronger and better supported for rate of spread and dispersal distances than for population density (Table 4.3). The average influence of the proportion of suitable habitat (PLAND) on rate of spread, dispersal distance and population density decreased when the degree of fragmentation in the landscape (SPLIT) and edge density (ED) - length of borders between patches of different habitat type in a given area, increased. On the contrary, the average influence of the proportion of high-quality habitat (PLAND) on rate of spread and dispersal distances was enhanced when the related circumscribing circle (CIRCLE_AM) - a measure of patch elongation, increased. This last index may be particularly useful for distinguishing habitat patches that are both narrow and elongated. These interactions predicted that the marginal increase of the proportion of suitable habitat within a landscape is on average less when the habitat is highly fragmented.

We also found significant effects of the interactions between the proportion of high-quality habitat (PLAND) and core metrics (DCORE_MD, DCORE_SD, CAI_CV and DCAD) on the maximum dispersal distance, and between the proportion of high-quality habitat (PLAND) and measure of patch aggregation (PLADJ) on the rate of spread. However, the effect of these latest interaction were relatively weak, with lower summed Akaike weights (Table 4.3).

4.3.2 Impact of landscape structure on model behaviour

We used non-metric multidimensional scaling (nMDS) to identify how values of the response variables (average density, rate of spread, occupied area, dispersal distances) in ordination space relates to landscape variation as represented by selected landscape metrics and variable life-history traits. We distinguish two distinct set of simulations: simulations with an Allee effect present (Figure 4.1), and simulations without an Allee effect (Figure 4.2).

Figure 4.1 suggests two broad categories of model dynamic when an Allee effect is present. Axis 1 of the ordination space broadly represents the rate of spread (ROS), the average dispersal distance (avDist), and the maximal dispersal distance (maxDist). These three components of spread increased towards the right of the space with increasing Allee effect and were strongly correlated with the frequency of dispersal events f. Axis 2 of the ordination space broadly relates to population density (d). The diffuse cluster toward the bottom of the ordination space represents 'failed' or less populated establishment, while the cluster toward the top of the axis 2 represents the most successful invasion. The biplot suggests that the most successful invasions (high density, high rate of spread and long dispersal distances) were positively correlated with the rate of increase r and negatively correlated with dispersal abilities *dist* and carrying capacity K. However, we observed a high variability within the successful invasions in terms of density, the occupied area and rate of increase, which are further correlated to the metrics of landscape structure.

In particular, the realized average dispersal distance (avDist) and rate of spread (ROS) are positively correlated with landscape metrics that decrease with habitat fragmentation (proportion of suitable habitat - PLAND, percentage of like adjacencies - PLADJ, connectivity index - CONNECT, number of disjunct core areas - DCDA, and variation in disjunct patch area - DCORE_MD and DCORE_SD). The realized average dispersal distance (avDist) and rate of spread (ROS) are negatively correlated with metrics which increase with fragmentation (edge density - ED, dplitting index - SPLIT, circumscribing circle - CIRCLE_AM).

In the absence of an Allee effect (Figure 4.2), the biplot suggests that landscape structure has a higher impact on the course of invasion. In particular, the most successful invasions are positively correlated to the proportion of suitable habitat (PLAND), the percentage of like adjacencies (PLADJ), the connectivity index (CONNECT), the number of disjunct core areas (DCDA) and the disjunct patch area DCORE_MD and DCORE_SD, while negatively correlated to the splitting index (SPLIT- a measure of fragmentation) and a measure of patch elongation (CIRCLE_AM). Both the rate of increase and the carrying capacity K are positively correlated with the most successful invasion, in particular with the highest density d. The median dispersal abilities *dist* and average frequency of dispersal events f are negatively correlated to the most successful invasion. Overall, in both studies (with and without Allee effects), the rate of spread and the realised average dispersal distance are the two factors most correlated to measures of landscape composition and configuration.

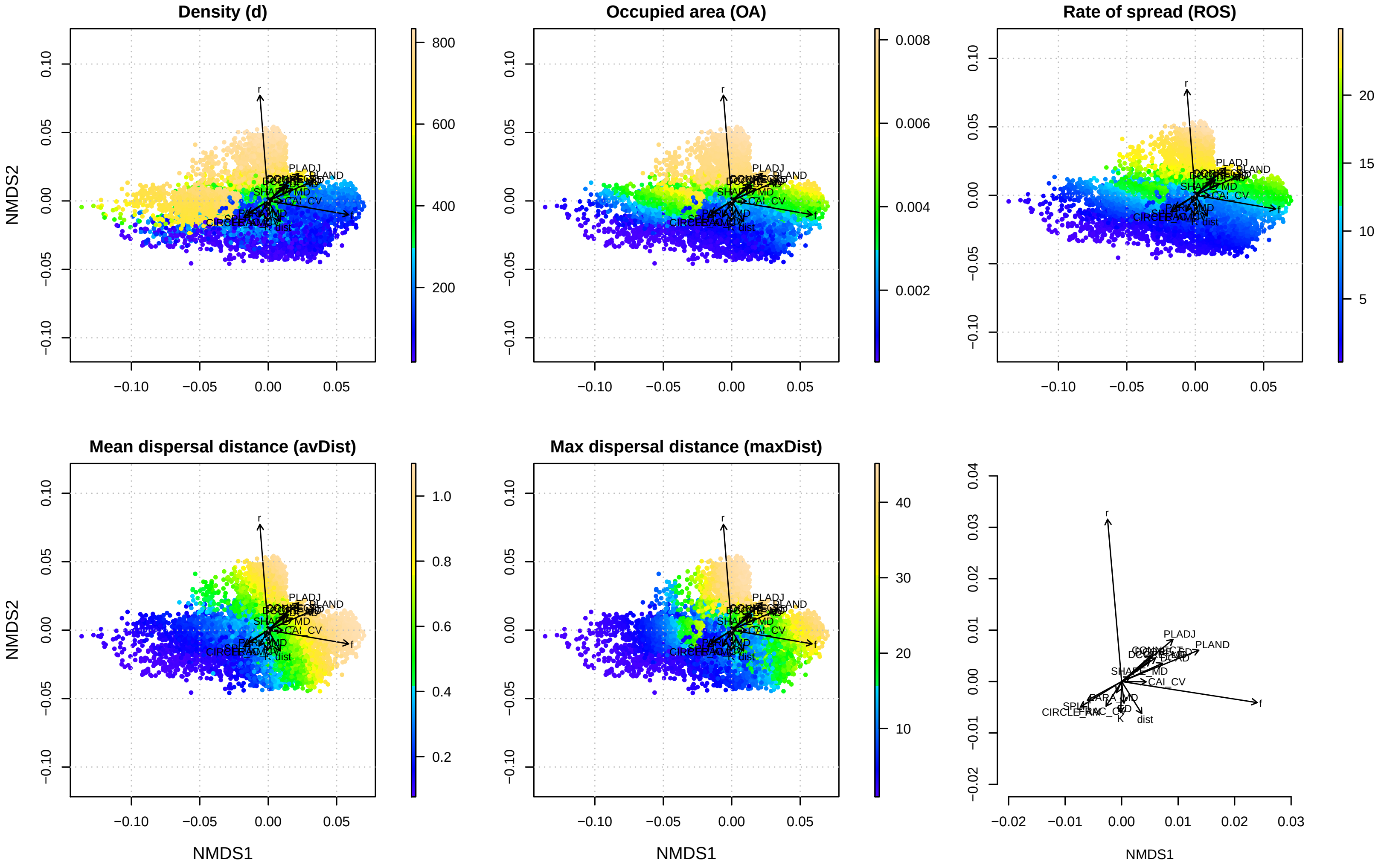

Figure 4.1: Non-metric multidimensional scaling (nMDS) profiles of the dispersal experiments with Allee effects over a gradient of spatial heterogeneity. Each point represents a dispersal experiment. The gradient of color indicates how the range of average population density (d), rate of spread (ROS), mean dispersal distance (avDist) and maximum dispersal distance (maxDist) associated with each dispersal experiment are distributed in the ordination space. The 'biplot' shows how the position of different dispersal experiments in the ordination space are related to the selected landscape metrics and life-history traits.

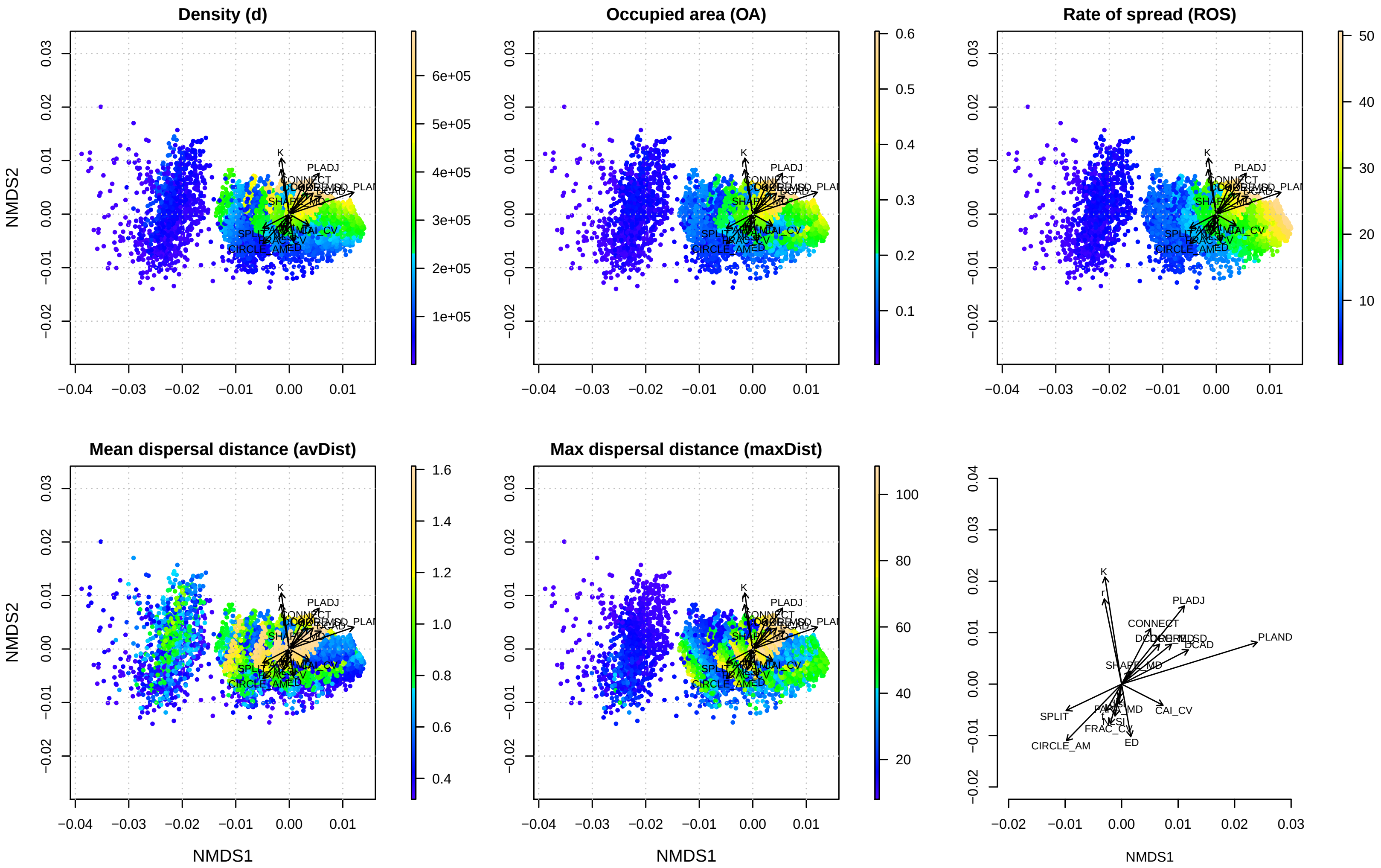

Figure 4.2: Non-metric multidimensional scaling (nMDS) profiles of the dispersal experiments without Allee effects over a gradient of spatial heterogeneity. The gradient of color indicates how the range of average population density (d), rate of spread (ROS), mean dispersal distance (avDist) and maximum dispersal distance (maxDist) associated with each dispersal experiment are distributed in the ordination space. The 'biplot' shows how the position of different dispersal experiments in the ordination space are related to the selected landscape metrics and life-history traits.

4.3.3 Balancing effectiveness of spread and invasion success

When an Allee effect was absent, an increase in the proportion of suitable habitat (PLAND) results an increase in the mean dispersal distances, and subsequently to an increase in the rate of spread and population density (Figure 4.3). The population may undergo spread and growth regardless of the proportion of suitable habitat.

In the presence of an Allee effect, the range of dispersal distance that allow for optimal spread was considerably reduced compared to the case without an Allee effect. At lower realised mean dispersal distances (avDist), spread was limited as evidenced by a lower realized mean rate of spread, but growth happened as indicated by a local maxima in the population density curve (Figure 4.3). A two-fold increase in the mean dispersal distance leads to the highest rate of spread and population density. The range of permissible rate of spread and population density was dramatically reduced when the mean dispersal distance went over a certain dispersal threshold (Figure 4.3). These results suggested that invasive species subjected to an Allee effect may present two different growth patterns when arriving to a new environment. On one hand, species that have limited dispersal opportunities will maximise their populations locally but will fail to establish a population over a large area. On the other hand, species that have high dispersal opportunities may spread but face the added risk of not establishing or going extinct. Overall, population density and rate of spread were maximised for intermediate dispersal abilities.

The results also showed that an increase in the proportion of suitable habitat (PLAND) in the landscape resulted in an increase in the probability of observing a mean dispersal distance higher than the dispersal threshold that can limit growth and spread (Figure 4.3). In a landscape with more than 70% of suitable habitat, 5% of the simulations resulted in a mean dispersal distance higher than the dispersal threshold. In a landscape with more than 80% of suitable habitat, 10% of the simulations resulted in a mean dispersal distance higher than this dispersal threshold. Although a high proportion of suitable habitat would lead to more successful spread (highest mean dispersal distances) when the source population density is high enough, it may also reduce spread success and population growth when the source population is too small. This result provides new insights into the proportion of suitable habitat as an additional explanation to the highly variable spread rate observed in both natural and anthropogenic landscapes.

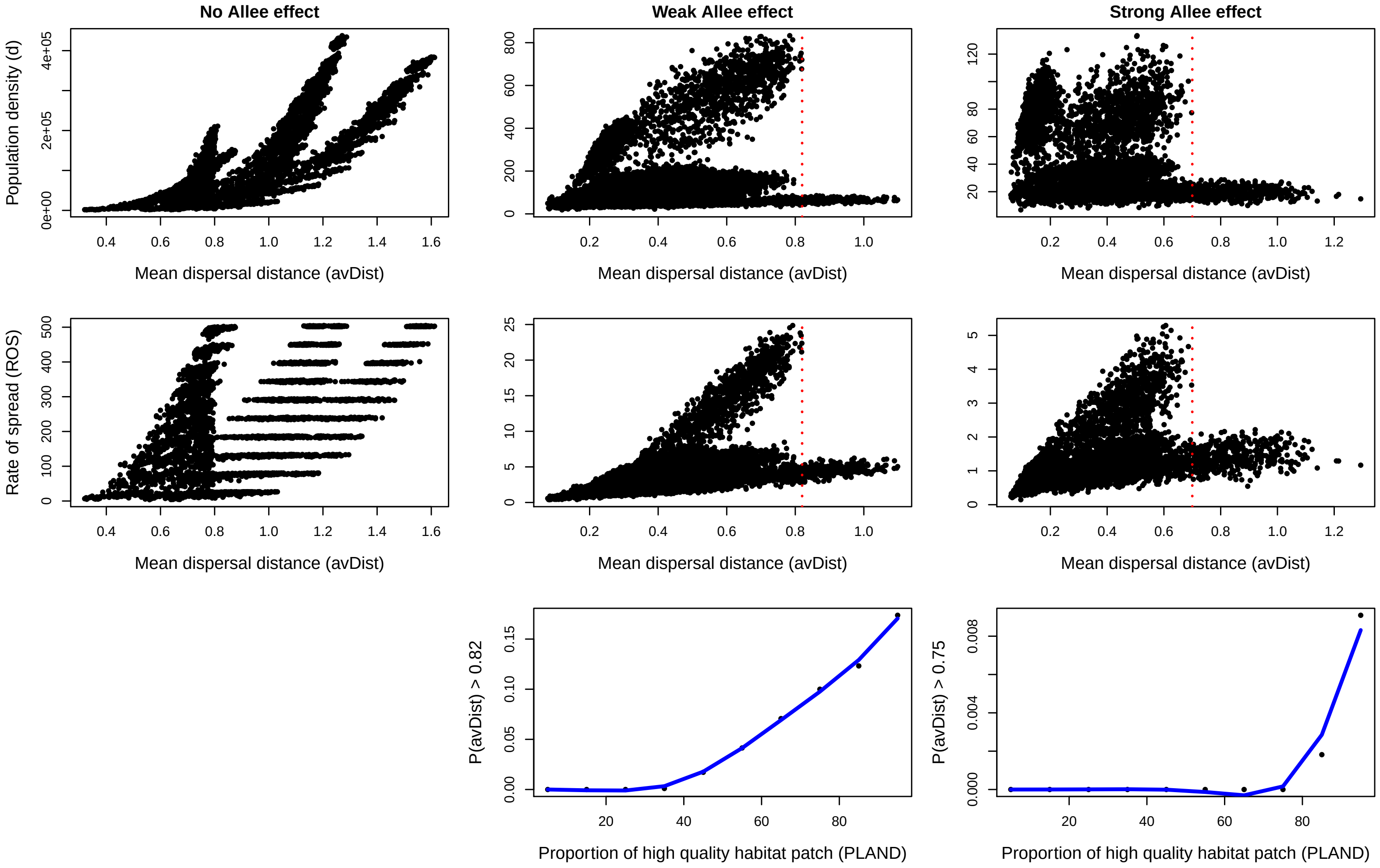

Figure 4.3: Balancing effectiveness of spread and invasion success in heterogeneous landscapes. The top panels show how population density (d) and rate of spread (ROS) respond to different ranges of mean dispersal distances for three different scenarios: without Allee effect, with weak Allee effect and with strong Allee effect. Above a dispersal threshold (red dotted line), growth and rate of spread were considerably reduced or not observed. The bottom panel represents the probability of observing a mean dispersal distance higher than the dispersal threshold that can limit growth and spread as a function of the proportion of suitable habitat (PLAND).

4.4 Discussion

The influence of landscape structure on the spread of invasive species is generally acknowledged but is often difficult to quantify. In this study, we investigated the effect of landscape structure as a predictor of patterns of spread of invasive insects. We found that both landscape composition and landscape configuration influenced population density and rate of spread in significant and sometimes interactive ways. However, the particular effect of landscape structure on invasive spread depends on several life-history characteristics such as the intrinsic rate of increase, dispersal ability, and frequency of long dispersal events as well as the presence or absence of an Allee effect.

4.4.1 Landscape metrics as indicators of spread

This research has shown that the density and rate of spread of invasive insects are more impacted by the proportion of suitable habitat than by habitat configuration. This has significant implications for the prediction and management of invasive species in general. In all species scenarios studied here, the increase in density and rate of spread of the species was positively related to an increasing proportion of suitable habitat in the landscape (PLAND). The direct effect of the amount of suitable habitat on invasive species density and spread has been reported in several studies, particularly in agroecosystems (Chaplin-Kramer *et al.*, 2011). Our results suggest that such an effect may be caused by a greater aggregation of suitable habitat patches (as measured by nLSI, PLADJ, SPLIT) and a slight decrease in edge density (ED). Different spread rates in different landscapes can also be caused by variation in connectivity of habitat patches (CONNECT), in habitat patch elongation (CIRCLE_AM) and by variation in the ratio of edge-to-interior habitat (DCORE_MD and DCORE_SD). In other words, landscapes containing larger patches of suitable habitat with simple, circular or square shape are expected to be more conducive to the rapid spread of invasive insects. Previous research also confirms that the metrics identified in this study that characterize habitat density (Sebert-Cuvillier *et al.*, 2008; Tscharntke *et al.*, 2012; With, 2002), connectivity (e.g. Morel-Journel *et al.*, 2015; Sebert-Cuvillier *et al.*, 2008; With, 2002; Wang *et al.*, 2011), path aggregation (Wang *et al.*, 2011; With & King, 1999b) and edge density (Radeloff *et al.*, 2000; Tscharntke *et al.*, 2012; With,

2002) are important aspects of the landscape that can either influence or exacerbate the spread of insects.

Overall, we found that the main effect of the configuration of habitats at a landscape scale had less impact on total density and rate of spread of the invasive species when compared to the main effect of the amount of habitat. Such a finding is consistent with and confirms the finding of many studies that habitat homogenisation is one of the key drivers of global increase in pest and disease establishment and spread (Jonsson *et al.*, 2015). Nevertheless, it is recognized that population density and rate of spread of invasive species can be significantly affected by complex interactions between the composition and configuration of the landscape in the introduction area.

4.4.2 Fragmentation *versus* proportion of suitable habitat in the landscape

The results in this study show that density and spread of invasive insects were best predicted by an interaction between the proportion of suitable habitat in the landscape (PLAND) and the splitting index (SPLIT). Moreover, evidence was found of a significant interaction between the proportion of suitable habitat in the landscape (PLAND) and respectively, the effects of edge density (ED) and average patch elongation (CIRCLE_AM) on rate of spread and dispersal distances.

The splitting index is a measure of habitat fragmentation designed to quantify the extent to which habitat patches are aggregated in space (Jaeger, 2000). Our results contrast with theoretical studies that suggest that suitable habitat connectivity at the introduction site is a determining factor of patterns of spread (Morel-Journel *et al.*, 2015). On the contrary, we found that invasive spread might be limited more by the size and distribution of non-habitat gaps than by a measure of habitat connectivity (the degree to which landscape features facilitate or impede the rate of movement of species between habitat patches). Ultimately, the notion of patch aggregation is embedded within the notion of connectivity such that, when habitat patch aggregation increases, connectivity also increases. Our results are in agreement with the results of With & King (1999a) and Wang *et al.* (2011)

who also highlighted the critical role of non-habitat gaps in reducing the spread of species over the landscape.

On the other hand, the effect of edge density (ED) and habitat patch elongation (CIRCLE_AM) was particularly challenging to assess. In this study, a decrease in edge density led to a slight increase in population density and rate of spread of the gypsy moth, but a slight decrease in density. Other studies have found that depending on whether species are moving primarily within or between habitat types, the micro-environments between habitat fragment may critically increase (Radeloff *et al.*, 2000; Tscharntke *et al.*, 2012; With, 2002) or decrease (Tscharntke *et al.*, 2012) the potential of spread. Ewers *et al.* (2007) further demonstrated that an increase in habitat shape complexity is associated with an increase in the amount of edge habitat that results in a higher edge density (ED). Large patches are more likely to be of a complex shape. We found that as shape complexity increases, patches become less circular (CIRCLE_AM), and the probability of spread and growth of invasive species decreases.

Overall, the results demonstrate that growth and spread of invasive species can be significantly affected by the combined landscape based components: suitable habitat amount, habitat patch aggregation and habitat shape complexity. We found that population density and rate of spread were positively correlated with landscape metrics that increase with decreasing habitat fragmentation, and negatively correlated with metrics which increase with increasing habitat fragmentation. To our knowledge, most theoretical studies that have focused on the spread of invasive species at the colonization front have also found a non negligible effect of the spatial arrangement of habitat on local density and rate of spread (Morel-Journel *et al.*, 2015; With, 2002). Fahrig *et al.* (2011) and Tscharntke *et al.* (2012), however, proposed that species respond idiosyncratically to habitat fragmentation versus habitat loss, which may suggest that highly mobile species can tolerate habitat fragmentation as long as the amount of total habitat is sufficient. Their finding suggest that the configuration of the landscape might not always be important for predicting the density and rate of spread of invasive species. Results from this study might highlight the need for data on the foraging and movement patterns of invasive insects in different habitat types and landscape contexts to validate mechanistic predictions.

4.4.3 Impact of landscape structure on population dynamics

Clearly, the overall outcome of spread critically depends on several life-history characteristics, dispersal ability as well as landscape factors. The presence of an Allee effect, the intrinsic rate of increase, the frequency of long-distance dispersal events and the proportion of suitable habitat were the primary discriminants of successful establishment and spread. As reflected by Figure 4.3, our results support two important hypothesis about population spread and growth in a heterogeneous environment. First, in the absence of an Allee effect, and in the absence of other influences (predation, for example), a population can spread and grow regardless of the proportion of suitable habitat. Increasing the proportion of suitable habitat in the landscape increases the overall population density and rate of spread exponentially (Figure 4.3; Shigesada & Kawasaki (1997)). Second, in the presence of an Allee effect, a population experiences a balancing effect between the effectiveness of spread and invasion success (Figure 4.3; Smith *et al.* (2014)): growth and spread are maximized at intermediate dispersal abilities but inhibited at high dispersal abilities. Species that have a high intrinsic rate of increase quickly exceed the size of population that is vulnerable to Allee effects, ensuring a higher probability of persistence of the introduced propagule(s) (Figure 4.1, Morel-Journel *et al.* (2015)).

Furthermore, Smith *et al.* (2014) used a programmed Allee effect in bacteria to demonstrate that an increase in the number of suitable habitat patches can result in more prolific spread but simultaneously increases the chances that each population could fail to establish due to Allee effect. Our simulations based on insect life-history parameters are in broad agreement with these results. We found evidence of a dispersal threshold above which population spread and growth can be inhibited (Section 4.3.3). Additionally, the results of this study echo the 'all or nothing' strategy proposed by Heimpel & Asplen (2011) and Morel-Journel *et al.* (2015), where they propose that a population would either go extinct quickly or form a stable population expanding across highly connected landscape. Therefore, species with an Allee effect follow unique dynamics of spread that can be significantly affected by the structure of the landscape. This study further revealed that by either allowing or preventing the dispersal of a species from the introduction site (or effective containment) (Section 4.3.1) complex interactions between the composition or configuration of the landscape may push these species into a range of dispersal that allows optimal spread and growth. Specifically, the results in this study showed that in a landscape with more than 70% of suitable habitat, 5% of the simulations resulted in a mean dispersal distance higher than a dispersal threshold that can limit growth and spread. Such an observation may offer an additional explanation to the highly variable spread dynamics observed in both natural and anthropogenic landscapes. It can also explain why, during an outbreak, spread can be initially slow manifesting itself as a time lag followed by an increase over time. Species must establish a population density that can exceed the size vulnerable to Allee effects before effective spread will occur.

4.5 Conclusion

This study offers new insights for a management programs that aim to limit the spread of invasive species. Strong evidence was found for landscape composition and configuration shaping the successful establishment and rate of spread of a species. Management strategies targeting different aspects of the landscape, including barrier zones (natural, anthropogenic, environmental barriers), modification of the host/habitat pattern (e.g., abundance, aggregation, isolation, quality), or regulation of the dispersal vector, modifies the shape of the relationship. Adopting a strategy that systematically reduces the dispersal ability of a new introduced species by reducing dispersal between suitable habitat patches, for example (Brown *et al.*, 2006), might inadvertently help them thrive by pushing the species into a range of dispersal that allows optimal spread and growth. We therefore recommend that management strategies for containment or eradication of invasive species set spatial priorities of control at the landscape scale. These spatial priorities should account for specific landscape characteristics, the stage of the invasion and the life-history characteristics of the species. The analysis in this study provides a basis for the development of a core set of structural landscape metrics as indicators of invasive insect spread. It is proposed that the seven metrics identified in this study (PLAND, nLSI, PLADJ, SPLIT, ED, CIRCLE_AM, and DCORE_MD) together include the majority of the information provided by the other metrics. However, optimal choice of metrics will require case specific exploration of their indicator values as there is no one-case-fits-all strategy that can ensure optimal risk assessment and management control for all species and landscapes. The overall outcome of this study is that strategic planning for management of invasive species must account for the effect of a core set of landscape features, the stage of invasion and species traits, as there is no single control option that ensures optimal risk assessment and management for all species and landscapes.

Chapter 5

The relative effects of propagule pressure, habitat structure and life-history traits on the establishment and spread of invasive species

Contribution of authors

This chapter describes a collaboration between the current author with Doctor Crile Doscher and Associate Professors Daniel B. Stouffer and Susan P. Worner, who provided statistical advice and comments on the manuscript. The results of this chapter are intended to be submitted as follow:
Lustig, A., Stouffer D. B., Doscher C. & Worner, S. P. (2016). The relative effects of propagule pressure, habitat structure and life-history traits on the establishment and spread of invasive species , *Journal of Ecology.*

Abstract

Propagule pressure, habitat structure of the introduction sites, and the demographic and dispersal attributes of invasive species have all been shown to influence their establishment, density and spread. Yet, their relative importance explaining invasive spread success and failure and their generality determining invasion outcomes across different regions and for different species have not been firmly established. Using a spatially-explicit model, we performed an boosted regression tree analysis to explore the relative importance of these factors on the density, rate of spread and mean dispersal distances of invasive insects. Variables representing an Allee effect, the intrinsic rate of increase and propagule pressure were found to be the three most influential variables in each model. Surprisingly, we did not find a significant effect of intrinsic dispersal abilities on response to habitat changes. Our results rather suggested that the frequency of long-distance dispersal events in combination with measures of suitable habitat amount and habitat aggregation was a better indicator of population dynamics than the intrinsic abilities of a species to disperse in fragmented landscapes. This implies that establishment and spread is most strongly related to reproduction, which determines the total number of potential colonists, but can be in turn exacerbated or mitigated by complex interactions between habitat structure and long-distance dispersal events. This study highlighted how complex interactions between propagule pressure, species traits and habitat characteristics can determine patterns of invasion across fragmented landscapes. Successful management of invasive species will depend on understanding this context dependent effect across habitats.

Keywords

Invasive insect spread, spatially-explicit models, landscape structure, Allee effects, propagule pressure, population dynamics

5.1 Introduction

A fundamental question in invasion biology is how the establishment, distribution and density of invasive organisms are influenced by various interacting factors, including propagule pressure, habitat characteristics at the introduction sites and species life-history traits (Thomas & Moloney, 2013). Although the influences of these different factors are well recognized, their relative importance in the success or failure of establishment and spread of invasive species remains unclear (Catford *et al.*, 2011; Cassey *et al.*, 2014; Robledo-Arnuncio *et al.*, 2014).

A predominant observation in a variety of taxonomic groups is that, as the number of invasive organisms introduced into a new area and the frequency of these releases increase, so does establishment and spread success (see, for example, the review of Lockwood *et al.*, 2005; Hayes & Barry, 2008; Lockwood *et al.*, 2009; Blackburn *et al.*, 2011). Intuitively, increasing the size and availability of propagules enables the incipient population to overcome behavioural limitations, such as mate access, or other limitations associated with small populations, including Allee effects and demographic stochasticity. Yet, several studies have questioned the generality of the positive propagule pressure – invasion success relationship. In particular, empirical experiments on insects (Sagata & Lester, 2009), plants (Yeates *et al.*, 2012), birds (Sol *et al.*, 2012) and marine species (Chadwell & Engelhardt, 2008; Haydar & Wolff, 2011) have found low or no statistical evidence for the positive influence of propagule pressure on establishment success. Jerde & Lewis (2007) discussed the fact that the initial introduction stage of most biological invasions is poorly understood, in part because there is often no reliable records of the species that were introduced but failed to establish. McGregor *et al.* (2012) further suggested that non-random patterns of introductions may result in the pool of introduced species being a biased subset with regard to key species attributes, potentially biasing perceptions of which factors are important for determining establishment and spread success. Clearly, one must therefore be cautious about assuming that reduced propagule pressure will lead to lower probability of establishment and spread.

Invasion success is also determined by the ability of the arriving propagules to overcome various pressures of the new host environment. At global scales, it is generally well accepted that abiotic conditions and resource availability are the main drivers influencing species distributions. Clearly, a species will not be able to invade an area that has abiotic conditions outside of its physiological tolerance levels (Petitpierre *et al.*, 2012). At a local scale, however, patterns of invasive species density and rate of spread differ between invaded habitats (Vilà *et al.*, 2007; Ibáñez *et al.*, 2009). These differences have been partly explained by the complex interplay of patterns and processes related to the variation and influence of spatial, temporal, biotic and anthropogenic drivers (Catford *et al.*, 2009). Biotic interactions, and their variations in time and space, are known to affect species' spatial patterns through several mechanisms such as competition, predation, parasitism or symbiosis, that can ultimately impede or facilitate invasion by new organisms (Mitchell *et al.*, 2006; Tylianakis *et al.*, 2008). In addition, human activities, such as urbanization or agricultural and forestry activities, promote invasion by increasing human-mediated propagule pressure and landscape disturbances (e.g. fluctuating resource availability, habitat loss and fragmentation) that encourage a shift in biotic communities and can benefit non-indigenous establishment (Hobbs, 2000). Across the globe, many native ecosystems have now been replaced by human-dominated mosaic landscapes, wherein a patchwork of sharply contrasted human land-use patterns has been superimposed on preexisting patterns of heterogeneity in environmental conditions. In this context, landscapes are a composition of more or less isolated patches of suitable habitat within a matrix of less suitable habitats (Hanski & Gaggiotti, 2004). From the perspective of a particular species, this patchwork of various habitat types may offer different resources for foraging and reproduction, and different suitability for species dispersal, emphasizing movement among distinct habitats as an essential facet of the ecology of many species (Fahrig & Nuttle, 2005). Understanding how landscape structure (i.e. composition and configuration) moderates population growth and natural dispersal of invasive species is therefore crucial to comprehending their dynamics in increasingly human-dominated landscapes (With, 2002; Vilà & Ibáñez, 2011; González-Moreno *et al.*, 2013a).

We understand from theory that there are two important aspects of the landscape spatial structure with potential to impact invasive species density and spread: landscape connectivity, defined as the combined effect of landscape elements that facilitate or disrupt the movement of individuals between habitat patches, and environmental heterogeneity (i.e. variance and range of environmental conditions and their spatial autocorrelation) (Fahrig & Nuttle, 2005; Fahrig *et al.*, 2011; Fagan *et al.*, 2013). More specifically, we proposed in Chapter 4 the differences in landscape structure, which can ostensibly characterized by a core set of eight landscape metrics, as an additional explanation to the highly variable spread dynamics observed in natural and anthropogenic landscapes. However, an understanding of these effects is still very limited, and it will be major challenge to derive generalities across systems. In particular, perception of the environment and surrounding landscape differs between species and the great range of dispersal and reproductive abilities complicates the interpretation of interaction patterns. For example, Tscharntke *et al.* (2012) suggested that species that are mobile enough should be less susceptible to habitat structure than less mobile species. However, some empirical studies, such as León-Cortés *et al.* (2003) and Van Houtan *et al.* (2007), have found the opposite, that more mobile species are more sensitive to habitat structure.

Nevertheless, empirical and theoretical studies suggest that species susceptibility to the interactive effects of propagule pressure and habitat structure depends on their demography. Sakai *et al.* (2001) observed that the possession of growth traits, which allow for fast intrinsic rates of increase, such as short generation time or high fecundity, is common to many successful invaders across taxa. In particular, Grevstad (1999) found that the consequences of an Allee effect and spatial heterogeneity on insect biocontrol agents was exacerbated for populations with lower reproductive rates. Similarly, Holland *et al.* (2005) found a negative association between reproductive rate and minimum habitat amount required for presence across a group of woodboring beetles, while Vance *et al.* (2003) described a similar relationship to explain the occurrence of forest bird species. In a synthesis across taxa in wetland ecosystems, Quesnelle *et al.* (2014) suggested that dispersal between habitat patches is less important than species' reproductive rates for population persistence in fragmented landscapes. On the other hand, Cassey *et al.* (2014) have investigated the relative importance of life-history traits on influencing the probability of invasion of exotic birds along with the combined effect of propagule pressure and environmental stochasticity. The authors found that while propagule size and environmental stochasticity undoubtedly influence establishment and spread success, demographic traits have an equal or greater effect on invasion success of birds. All these studies predict that species with lower reproductive rates are more sensitive to suitable habitat amount in the landscape than species with higher reproductive rates. Yet, these studies did not investigate, the relative importance of habitat structure and its influence on invasion success at the introduction site in detail.

In consideration of the above, the main objective of this study was to determine the relative importance of dispersal ability, reproductive rate, propagule pressure and habitat structure for determining population establishment, growth and spread of invasive insects. A spatially-explicit framework was used to explore a range of relationships between invasive success and propagule pressure, habitat structure and life history traits. We expand on previous studies by focusing on the response of a population model to change in landscape composition and configuration at the introduction site. Developing a better understanding of the relative effects of these different factors on establishment and spread of invasive species is an essential step to inform more targeted management actions. Not only is it intended that this study yield an improved understanding of the underlying invasion processes, it should also facilitate the development of efficient predictors that are not species- or context-specific. For example, if the interplay between dispersal abilities and habitat structure drives density and spread of invasive species, eradication programmes should focus on limiting dispersal. In contrast, if reproductive rate drives the response, the focus should be on limiting reproductive stages. To investigate these questions, we performed an uncertainty analysis to explore the relative importance of different factors on the density, rate of spread and mean dispersal distances of invasive insects.

5.2 Material and methods

5.2.1 MdiG simulations

Introductions of non invasive species were simulated using the spatially explicit model, MdiG (Pitt *et al.*, 2009). The model attempts to represent in great detail the complexity of the ecology of invasive insects' establishment and spread. At the population scale, the model integrates a parameter set to simulate life history events of birth and death, following a deterministic Allee logistic growth model. Environmental stochasticity (i.e. spatial variation in population growth) is included by modifying mortality as a function of local insect density and habitat quality. At the landscape level, spatially explicit rules determine the pattern of local (random walk) and generic long distance dispersal spread (Cauchy probability distribution). Additional details about the modelling framework and the parametrization of the model can be found in Chapter 2 and Chapter 4, respectively.

We set the parameters of the models for a generic gypsy moth (*Lymantria dispar dispar*) (Chapter 4). Invasion by gypsy moth across North America is one of the most thoroughly studied biological invasions, providing a unique opportunity to explore spatio-temporal variability in patterns and rates of spread (Jankovic & Petrovskii, 2013). The landscape, or simulation arena, consists of a binary distinction between suitable and unsuitable habitat types. As described in Chapter 2, we generated landscape patterns in a full factorial design across a gradient in spatial autocorrelation and gradient in the proportion of suitable habitat cover. We used a core set of eight structural landscape metrics (PLAND, nLSI, PLADJ, SPLIT, ED, CIRCLE_AM, DCORE_MD, and CONNECT), that enable a quantitative comparison of landscape structure, and have been shown to be useful indicators of invasive insect density and spread (Chapter 4).

Parameters	**Abbreviated code**	**Parameter value**
Population demography		
Growth rate	r	0.815, 1.223*, 1.63 per capacity growth
Allee threshold	C	0, 2*, 5 individuals per raster cell
Carrying capacity	K	30, 50*, 100 individuals per raster cell
Propagule pressure	nbp	5, 10*, 15 individuals per raster cell
long-distance dispersal behaviour		
Median distance	λ	3, 5*, 7 raster cells
Frequency	f	0.01, 0.05*, 0.09
Landscape structure		
Landscape spatial autocorrelation (fragmentation)	H	H= 0,0.3,0.5,0.7,0.9
Percentage of suitable habitat patches in a binary scheme	P	P = 5,15,35,55,75,95

Table 5.1: Parameter values for the stochastic spatially-explicit spread model, MDiG. Parameters are subdivided into the three main model ‘components’. Those parameters marked with * indicate the baseline parameter values for a spread model of gypsy moth.

We used MdiG to broadly capture twenty-seven different combinations of species' life-history attributes and dispersal abilities in a mono-factorial design (Table 5.1). The main reason for doing this was to evaluate how change in intrinsic growth rate, carrying capacity, Allee effect threshold, and the frequency and median distance of long-distance dispersal events affect the dynamics of spread, and to quantify a general relationship across life-history attributes to changes to landscape variables. For each life-history scenario and each sample landscape, we modelled the fate of a single release of 5 to 20 individuals randomly localized in the largest suitable patch. We varied each parameter over three values (Table 5.1) and studied the effects of these changes on the dynamics of the simulated invasion. For each landscape, the species was allowed to expand its range for 30 years, one model time step represents one year and simulations were replicated 25 times to account for dispersal stochasticity. Population density – d (number of individuals per raster cell), the rate of spread - ROS (number of new cell occupied per simulation run), and the average dispersal distance – avDist were recorded over time.

5.2.2 Boosted regression trees

To evaluate the relative importance of dispersal ability, reproductive rate, propagule pressure and habitat structure for determining population establishment, growth and spread of invasive insects, we used generalized boosted regression trees models (BRTs). BRTs are a type of machine-learning statistical tool that have been increasingly used to analyse multivariate simulation data that have complex non-linear interaction effects on response data (Elith *et al.*, 2009). The BRT analysis allows quantification of the relative importance of each predictor in explaining variation in the response variables under study.

Boosting works by building a sequence of regression models (the boosting algorithm), each fitted to the residuals of the previous model with larger residuals given more weight (Elith *et al.*, 2009). Predictions are then combined into a final ensemble prediction. The idea behind this method is to partition the space of response variables into homogeneous groups based on a series of binary rules (splits) constructed from the predictor variables. The contribution (or importance) of a variable is based on the number of times the variable is selected for splitting. Model performance was assessed using the predictive deviance expressed as a percentage of the null deviance, and, the cross-validated correlation which evaluates the accuracy with which the model fits the response variables (Elith *et al.*, 2009). Visualization of fitted functions is achieved using partial dependence plots, which show the relationship between a focal predictor and a response variable controlling for the average effect of all other variables in the model. Advantages offered by BRT models include their ability to accommodate different types of predictor values and its immunity to the effects of extreme outliers. BRT also facilitate the study of interactions among predictor variables, by varying the number of splits (size) of the individual regression trees. The interaction strength among variables is assessed using residual-based statistics defined by Friedman & Popescu (2008), where larger values indicate stronger dependencies.

In this study we considered two different models for each of the response variables: a model without interactions among predictor variables (interaction depth = 1) and a model that accounted for up to 5-way interactions between predictor variables (interaction depth = 5). Prior to conducting regression analysis, predictor variables were centred and normalized, while the response variables were log transformed. To calibrate the models a 10-fold cross-validation (CV) was initially performed, for which 50% of the simulated data were withheld from the model-building process (bag fraction of 0.5) and subsequently used for testing (Elith *et al.*, 2009). We tried different learning rates ranging from 0.0005 to 0.1, interaction depths ranging from 1 to 10 and the number of trees ranging from 50 to 10000. A learning rate of 0.001, 1000 trees and a Gaussian error structure were found to achieve the smallest predictive deviance while still being computationally feasible. All BRTs were

Model	**Complexity**	**No. regression tree**	**No. splits**	**Cross-validated mean residual deviance (±SE)**	**Cross-validated correlation (±SE)**
Density (d)	1	2490	4	1.340 ($\pm 6.63 \times 10^{-3}$)	0.957 ($\pm 3.24 \times 10^{-4}$)
	5	1390	16	0.247 ($\pm 2.57 \times 10^{-3}$)	0.992 ($\pm 7.60 \times 10^{-5}$)
Rate of spread (ROS)	1	2380	4	0.446 ($\pm 4.69 \times 10^{-4}$)	0.951 ($\pm 6.64 \times 10^{-5}$)
	5	1390	16	0.078 ($\pm 2.34 \times 10^{-4}$)	0.992 ($\pm 3.09 \times 10^{-5}$)
Mean dispersal distance (avDist)	1	3430	4	0.008 ($\pm 1.28 \times 10^{-5}$)	0.930 ($\pm 1.46 \times 10^{-4}$)
	5	1840	16	0.002 ($\pm 5.42 \times 10^{-6}$)	0.980 ($\pm 5.10 \times 10^{-5}$)

Table 5.2: Boosted regression tree model performance relating population density and spread of invasive species to propagule pressure, demography, dispersal and landscape structure metrics. Table values indicate, 1) the complexity or maximum number of interactions that can be captured by the statistical model, 2) the number of trees fitted, 3) the number of partition or splits used to partition the space of the response variable, 4) the cross-validated deviance and correlation, to evaluate the accuracy with which, the model calibrated on a training data set, fits the sampled response variables, based on a 10-fold cross validation (Elith *et al.*, 2009).

fitted in R (R Development Core Team) using the *gbm* package (Ridgeway, 2015).

5.3 Results

5.3.1 Non-interaction regression models for population density and spread

The three non-interaction models (Table 5.2, model complexity = 1) yielded good agreement between the training dataset and the remaining data used for testing (Table 5.2, deviance < 1.5), indicating that the predicted relationships between the response and predictor variables were reliable. The BRT model related to mean dispersal distance (avDist) produced the lowest deviance, while the BRT model related to density (d) the highest. All models showed a cross-validated correlation higher than 0.9 (Table 5.2).

The Allee effect threshold had the highest explanatory power for all models accounting for more than 84.14% of the variation in population density (d), 82.73% of the variation in rate of spread (ROS), and 49.94% of variation in mean dispersal distance (avDist) (Figure 5.1). As the Allee threshold decreases, population density and the rate of spread increase (Figure B.1). Moreover, all response variables were also linked to population growth and dispersal ability, propagule pressure and landscape structure. In particular, relative influence metrics revealed that the propagule size(nbp), the intrinsic rate of increase (r), the carrying capacity of the environment (K), together with Allee threshold accounted for more than 95% of model performance when forecasting population density (d) (Figure 5.1). Rate of spread (ROS), however, was better predicted by considering the percentage of suitable habitat in the landscape (PLAND), the intrinsic rate of increase (r), and the propagule size (nbp), together with the Allee threshold. Finally, the Allee threshold and the frequency of long-distance dispersal events were of primary importance explaining variation in mean dispersal distance (avDist), while the intrinsic rate (r) of increase and the splitting index (SPLIT) exerted lesser influence (Figure 5.1). Similar patterns were observed in the simulations with an Allee effect alone (Allee threshold > 0) (Figure B.2). As shown previously the model predicted higher population density and rate of spread when propagule pressure increases, the intrinsic rate of increase increases and the Allee threshold decreases. The importance of the Allee threshold, however, was lower confirming the importance of propagule pressure and rate of increase as key factors in determining invasion success.

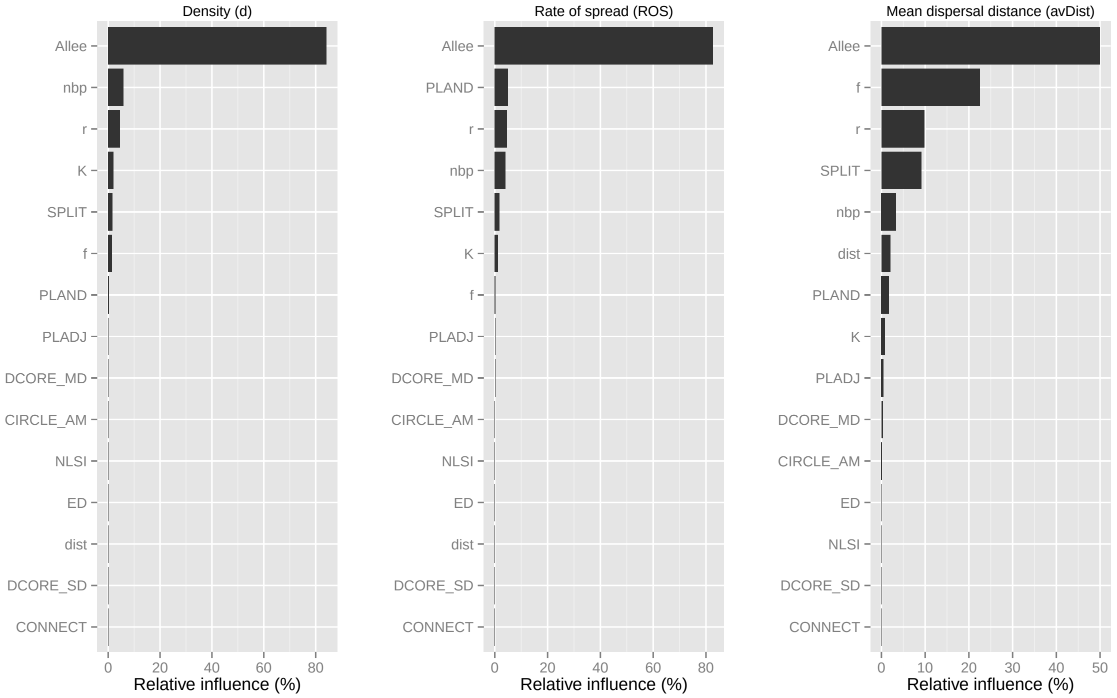

Figure 5.1: Relative influence (%) of propagule pressure, demography, dispersal and landscape structure on log transformed population density (d), rate of of spread (ROS) and average dispersal distance (avDist) of invasive insects (Non-interaction regression models).

It is notable that the average distance of long-distance dispersal events ($dist$), in other words, the intrinsic dispersal abilities of the species under study, was only retained in the model considering the realized mean dispersal distance (avDist) (Figure 5.1). Moreover, among the eight metrics of landscape structure considered in this study, only the percentage of suitable habitat in the landscape (PLAND) and a measure of patch aggregation (SPLIT) contributed significantly to the performance of all three models. Core area metrics (DCORE_MD, CIRCLE_AM), patch adjacency (PLADJ) accounted for less than 2% of model performance, while the edge density index (ED), the connectivity index (CONNECT) and the standard deviation of core area (DCORE_SD) were omitted by the sequential boosted learning models related to population density (d) and the rate of spread (ROS), assigning them a zero coefficient and thus discarding them from the resulting formula. This result suggests that a relatively simple measure of habitat configuration

(SPLIT) and composition (PLAND) can account for most of the explanatory power of landscape structure as a driver of spread rate (ROS) and population density (d).

5.3.2 Interaction regression models for population density and spread

The addition of simple interaction terms improved the mean deviance explained and the mean root square correlation of all three BRT models, indicating the importance of interactions between predictor variables in explaining population density (d), rate of spread (ROS) and mean dispersal distance (avDist) (Table 5.2). Table 5.3 shows that there are some significant interactions between parameters that are particularly strong in models for population density (d) and rate of spread (ROS).

Rank	**Interaction terms**	**Int. strength**
Population density (d)		
1	Intrinsic rate of increase (r) × Allee threshold (Allee)	151.34
2	Propagule size (nbp) × Allee threshold (Allee)	151.34
3	Frequency of long-distance dispersal events (f) × Allee threshold (Allee)	55.42
4	Intrinsic rate of increase (r) × Propagule size (nbp)	44.02
5	Splitting Index (SPLIT) × Allee threshold (Allee)	22.69
Rate of spread (ROS)		
1	Intrinsic rate of increase (r) × Allee threshold (Allee)	71.69
2	Percentage of suitable habitat (PLAND) × Allee threshold (Allee)	51.82
3	Intrinsic rate of increase (r) × Propagule size (nbp)	18.76
4	Propagule size (nbp) × Allee threshold (Allee)	14.55
5	Carrying capacity (K) × Allee threshold (Allee)	7.11
Average dispersal distance (avDist)		
1	Intrinsic rate of increase (r) × Allee threshold (Allee)	0.76
2	Splitting Index (SPLIT) × Frequency of long-distance dispersal events (f)	0.46
3	Propagule size (nbp) × Allee threshold (Allee)	0.20
4	Frequency of long-distance dispersal events (f) × Intrinsic rate of increase (r)	0.17
5	Intrinsic rate of increase (r) × Propagule size (nbp)	0.14

Table 5.3: Interaction strength between the most important parameters for log-transformed population density (d), rate of of spread (ROS) and average dispersal distance (avDist). Interaction strength (Int. strength) is quantified using the residual-based methods outlined in Elith *et al.* (2009) and quantifies the excess of residual variance explained under the null hypothesis of no interaction effects. The interaction rank was used to compare interaction strength between one model and another.

The interaction between the intrinsic rate of increase (r) and the Allee threshold (Allee) was identified as the strongest interaction for all thee BRT models. In each case, the intrinsic rate of increase (r) and propagule size (nbp) appear to have little influence on model outputs when the parameters are considered in isolation (relative contribution $< 5\%$), however, together with the Allee threshold (Allee) or with each other, they feature among the strongest interactions. Intuitively, population density and rate of spread increase with increasing intrinsic rate of increase, increasing propagule size and lowering the Allee effect.

Overall, the interactions between predictor variables showed a simple monotonic interaction pattern (Figure B.3). However, it is worth noting that the effect of propagule size (nbp) and intrinsic rate of increase (r), on both population density (d) and rate of spread (ROS), appear to be more important for simulations with a weak or strong Allee effect (Allee equal to 0 or 1 respectively), than for simulations where the Allee effect was absent (Allee equal to -1). In contrast, the effect of the percentage of suitable habitat cover (PLAND) on rate of spread (ROS) appears to be more important for simulations without Allee effect (Allee equal to -1), than for simulations with a weak or strong Allee effect. Surprisingly, population density (d) was positively related to the frequency of long-distance dispersal events (f) in simulations without Allee effect, but negatively related to this same parameter for simulation with weak or strong Allee effect (Figure B.3). This suggests that an increase in the frequency of long-distance dispersal events (f) can result in more prolific spread by increasing the chance to find suitable habitat patches but can simultaneously increase the chance that each population could fail to establish due to Allee effect or dispersal mortality. Finally, when the value of the splitting index (SPLIT, a measure of patch aggregation) decreases, population density (d) increases particularly at low value of SPLIT. In this scenario, it is at a higher degree of patch aggregation, that population growth is clearly enhanced. The average dispersal distance (avDist) was also enhanced at higher levels of patch aggregation, probably due to better success in the local spread of individuals.

5.4 Discussion

One of the factors currently limiting our understanding of biological invasion is the relatively limited taxonomic and geographic scope of current research (Catford *et al.*, 2011). It is acknowledged that for many, if not all invasive species, multiple drivers contribute to their successful establishment and dispersal, but their combined effects are rarely investigated (Catford *et al.*, 2009; Thomas & Moloney, 2013; Catford *et al.*, 2009; Cassey *et al.*, 2014). This study has demonstrated that it is possible to take a general approach which is not species or habitat specific to help identify general predictors of invasion. Particular attention was given to the challenge of accounting for the relative contribution of multiple drivers of invasion that form the key basis for explaining and predicting spatial patterns of spread in changing ecosystems (Thuiller *et al.*, 2008; Catford *et al.*, 2011; Robledo-Arnuncio *et al.*, 2014).

5.4.1 Primary importance of the Allee threshold, propagule size and rate of increase

Virtually every animal or plant species that requires sexual recombination for reproduction may be expected to exhibit some form of Allee effect on population growth, particularly in fragmented landscapes. In a thorough review of the consequences of an Allee effect in biological invasion, Taylor & Hastings (2005) reported how an Allee effect can influence many aspects of population dynamics including the probability of successful invasion, the rate of spread and even the final range distribution, and genetic diversity of the invasive species. They proposed the Allee effect as a central dynamic of biological invasions, that can influence optimal control decisions, costs of control and the estimation of the risk posed by potentially invasive species. The results in this study reaffirm the critical role of an Allee effect for determining population density and spread of invasive species, as we found the Allee threshold to be the primary determinant of simulation outcomes.

More importantly, strong interactions between the Allee threshold, intrinsic population rate of increase, and propagule size were the three most influential variables in each model. While initial founder size was undoubtedly an important factor for ensuring population growth and spread, we found the population intrinsic rate of increase to have an equal or greater influence on population density and invasive spread. These results run contrary to the well-established empirical prediction that propagule pressure is the most influential factor determining invasion success (see for example the review by Lockwood *et al.*, 2005). Possible reasons could be that empirical studies focus on a limited subset of life-history traits as generally little or no information on invasion failures are readily available, biasing our perception of which factors are important for determining invasion success. This suggestion also supports the recent study by Cassey *et al.* (2014) which also identified demographic traits to be the most important factor influencing the probability of invasion of exotic birds. However, unlike our study, Cassey *et al.* (2014) found no evidence of the influence of an Allee effect in their model. They suggested that environmental stochasticity, that can also lead to extinction of low-density populations, appears to matter more than Allee dynamics. Environmental stochasticity was not explicitly considered in our study. Theoretical and empirical studies that have investigated the relative role of Allee dynamics, environmental stochasticity, and even demographic stochasticity in invasive species within different habitats are still rare and suggests a direction for future research.

Interestingly, our study found a greater effect of growth traits than movement-related traits on both population growth and rate of spread in fragmented landscapes. Such a result is consistent with and confirms the finding of several studies that have identified species' reproductive rate as one of the major factors affecting habitat-amount requirements for species persistence. For example, With & King (1999a) found that reproductive rate has a much larger effect on the amount of habitat required for invasive establishment and spread than the species' dispersal abilities. In line with these results, a recent meta-analysis of habitat loss effects across wetland vertebrates species by Quesnelle *et al.* (2014) suggested that dispersal traits are less important than species' reproductive rates for population persistence in fragmented landscapes. A possible explanation is that higher reproductive rates compensate for increased dispersal mortality in fragmented landscapes and reduce local extinction risk Fahrig & Nuttle (2005). Overall, while dispersal abilities, environmental stochasticity and propagule pressure can all be important for determining population growth and spread of an invasive species, this study confirms that demographic traits, such as the intrinsic rate of increase, are of critical importance.

5.4.2 The contribution of long-distance dispersal events

Our results suggested that the intrinsic dispersal abilities of a species is not a useful predictor of invasive spread. Indeed, we found no evidence to support the prediction that species with greater dispersal abilities were better invaders than species with lower dispersal abilities. A possible explanation for the apparent lack of influence of natural dispersal ability of a species on invasion success could be that the ability of each species to move in the landscape varies widely among landscape structures, such that their ability changes as the landscape structure changes. In particular, in Chapter 4 the realized average dispersal distance was shown to be a product of both species traits and landscape structure. Similar to that result, Ferraz *et al.* (2007) found no effect of dispersal ability on patch occupancy for 55 non-indigenous birds. Additionally, Andrew & Ustin (2010) found no effect of dispersal ability on the spread of a perennial peperweed, *Lepidium latifolium*, in anthropogenic landscapes throughout California, while Quesnelle *et al.* (2014) found no influence of dispersal ability on wetland vertebrates response to habitat loss. Fahrig (2007) further suggested that species dispersal ability cannot be estimated independently of landscape structure and that to test the relative importance of dispersal ability, the landscape context should match the location where the data were collected.

Nevertheless, it is worth noting the critical role of long-distance dispersal events for determining the realized average dispersal distance in the landscape and population density. Long-distance dispersal has been shown to be central to many key population processes, such as metapopulation dynamics, diversity and dynamics in species communities, especially in fragmented landscapes (Cain *et al.*, 2000; Trakhtenbrot *et al.*, 2005). Most importantly, long-distance dispersal events can determine to a large extent the rate of spread and the final range distribution of invasive species, and therefore is fundamental to species responses to rapid climate and habitat changes (Hill *et al.*, 2011). Neubert & Caswell (2000) and Woolcock & Cousens (2000), using integro-difference models of spread in heterogeneous landscapes, found that the average distance of long-distance dispersal events define invasion rates, which were remarkably insensitive to the frequency of long-distance dispersal. However, our simulation results, which accounted for a more complex representation of landscape structure, suggested that the frequency of long-distance dispersal events could be a better indicator of population dynamics than the intrinsic dispersal ability of a species. We found that, in the absence of an Allee effect, population growth and average dispersal distance increased as the frequency of long-distance dispersal increased. In the presence of an Allee effect, however, average dispersal distance increased with the frequency of long-distance dispersal events, while population density decreased with the frequency of long-distance dispersal events. This counterintuitive finding can be explained by the fact that repeated dispersal events of small populations may be insufficient to establish growth in the recipient area, but collectively can reduce the size of the source population at the introduction site. In Chapter 4, the magnitude of this effect was shown to be dependent on the landscape structure and the species reproductive ability. Despite this explanation being plausible, there is no experimental support. Long-distance dispersal events are inherently hard to measure, and there are only a few data sets that can be used to describe the tails of a long-distance dispersal kernel (Robledo-Arnuncio *et al.*, 2014). Furthermore, because estimating the relative contribution of diverse variables to dispersal and their variation in time and space, is very difficult, few studies even try to do so (Robledo-Arnuncio *et al.*, 2014). Results from this study reaffirm the need for more data on movement patterns, particularly on long-distance movement, of invasive species in different habitat types and landscape contexts, to provide a better understanding the ecological consequences of long-distance dispersal events in changing environments.

5.4.3 The role of landscape structure

Measures of habitat amount (PLAND) and habitat aggregation (SPLIT) were shown to be the most important landscape factors for determining, population density, rate of spread, and the realized dispersal distance. Habitat amount has been shown to be very important to spread dynamics and to mediate the effects of landscape configuration on population density and spread. In particular, With (2002) demonstrated that spread rate increased with increasing habitat availability above a certain threshold after which habitat configuration becomes unimportant. Our simulations based on insect life history strategies are in broad agreement with this result, as we found the rate of spread to be higher when the percentage of suitable habitat in the landscape was higher, and the measures of habitat configuration had a relatively weak effect on simulation outcomes. However, the variability in dispersal success in the landscape was best accounted for by differences in the aggregation of habitat patches as measure by the splitting index (SPLIT), with closer and larger patches having significantly greater exchange of dispersing organisms, with subsequent impact on population density.

Of the parameters tested in this analysis, the measure of edge density (ED) and connectivity (CONNECT) had the weakest effect on spread rate and population density. We suggested in Chapter 4 that population growth and spread might be limited more by the size and distribution of non-habitat gaps than by a measure of habitat connectivity. That observation would explain the relative high influence of the splitting index and relatively weak effect of the connectivity index. The weak effect of edge density is more surprising as habitat edges are often thought to give invasive species a foothold in the landscape (Vilà & Ibáñez, 2011). It could be that edge density, when measured at the landscape level, is a relatively ineffective predictor of rate of spread. Alternatively, edges may be most effective when they actively and directly enhance dispersal, for example by higher wind disturbance at habitat edges enhancing long-distance dispersal events, rather than simply providing access to habitat. Such mechanisms were not considered in this study.

Finally, conservation studies have often stressed the importance of a habitat core zone for the persistence of native species. In this context, management of invasive species is continually undermined by re-invasion at the periphery of the treated area, so that benefits for eradication or biodiversity maintenance in the case of conservation, are realized only within a smaller core near the center of the treated area (see the study by Glen *et al.*, 2013, for more details). Our study showed that core area metrics (DOCRE_MD, CIRCLE_AM and DCORE_SD) accounted for less than 2% of model performance, suggesting that any conservation strategy should consider more than a critical core habitat for invasive species management. In particular, management of invasive species should also consider the landscape adjacent to these core areas (aggregation to other suitable areas) because propagule pressure or successful traits are likely to be different depending on the surrounding landscape.

5.5 Conclusion

Elucidating the factors that determine establishment and spread success of invasive species involves disentangling the influence of characteristics of the species and the environment, and the idiosyncrasies of specific introduction events. A wide range of organisms with differing demographic and dispersal abilities, and different introduction scenarios, can be simulated using the approach presented in this study to increase our understanding of how landscape structure affects organism dispersal. From the array of predictor variables considered in this study, strong interactions between an Allee effect, rate of increase and propagule size are emerging as consistent correlates of establishment and spread success. All three processes controlled the local population density at the introduction site. The primary importance of an Allee effect suggests that any eradication strategy should consider the extent to which life histories of the species contribute to the Allee effect and examine how eradication programmes can exploit such dynamics by maintaining population densities under Allee threshold.

Surprisingly, we found no evidence that the intrinsic dispersal ability influences species establishment or spread. This implies that spread rate is more strongly related to reproduction, which determines the total number of organisms participating in dispersal, than it is to a species' intrinsic dispersal ability. From an invasive species management perspective, our results suggest that priority should be placed on species with high intrinsic rate of increase and that the eradication programmes focus should be on limiting reproductive stages as a priority. Population density and rate of spread were linked, to a lesser extent, to three main landscape processes, including the frequency of long-distance dispersal events, a measure of habitat amount and a measure of habitat aggregation. All three parameters influenced the ability of the species to create new founder populations that resulted in more prolific spread by increasing the chance to find suitable habitat patches. However, this result simultaneously increases the chances that each population could fail to establish due to an Allee effect or high mortality during dispersal. Just at it is necessary to move from site/species-specific to landscape level and multi-taxon analyses to understand the ecological relationship between the population growth and spread of invasive species within their environments, it is also necessary to base risk assessment and management strategies on the characteristics of the landscape matrix of the recipient region. It is important to recognize that establishment and spread will always be a stochastic process and like most things cannot be predicted with absolute certainty, however, our understanding of invasive risk will increase by understanding context-dependent effects across habitats at the different stages of the invasion process.

Chapter 6

Sources of uncertainty when modelling invasive species establishment and spread

Abstract

Estimates of establishment, spread, and impacts of invasive species, as well as future prediction about the effect of anthropogenic land-use and climate change, rely on projections from a broad range of models. Uncertainties associated with such estimates have often limited the prevention, report, and control of biological invasions, particularly at local scales. In this study, I review the sources of uncertainty associated with the development of an establishment and spread model, and discuss a set of methods for quantifying model predictive performance as well as uncertainty. Part of this uncertainty is irreducible because of the internal variability of natural systems. However, uncertainty associated with data, model parametrization and model structures can potentially be reduced. Six procedures emerged as important for reducing or eliminated uncertainty in establishment and spread modelling. I discuss how understanding the errors and uncertainties that occur during the process of modelling the establishment and spread of invasive species, along with accurate estimation and communication of the information content of any models, is key to improving the value of predictive models for practitioners.

Keywords

Uncertainty, biosecurity, forecasting, expert judgement, scientific error, model averaging

6.1 Introduction

The increasing economic, social and environmental losses caused by invasive species intensify the need to identify and implement optimal management strategies to prevent, detect, and control them (Funk, 2015). Invasive species distribution maps and prediction of spread are important for assessing the risk of invasive species. Such maps and predictions help illustrate the probability of invasion by an alien species and how the potential consequences of that invasion vary temporally and spatially. Such maps are powerful tools that can be used to assist policymakers make decisions with respect to biosecurity surveillance (Cacho *et al.*, 2010; Demon *et al.*, 2011; Epanchin-Niell *et al.*, 2012) and invasive species incursion responses (Coutts *et al.*, 2011; Venette, 2015). However, uncertainty is an inevitable component of assessing risk, particularly for spatially explicit spread forecasts. Such uncertainty is a challenge for invasive species risk managers to incorporate in their decision-making process (Carrasco *et al.*, 2012; Liu *et al.*, 2010; Yemshanov *et al.*, 2010). One of the biggest impediments to the incorporation of uncertainty into risk assessment and the prioritization of risks, has been the lack of techniques to effectively represent and communicate uncertainty, whatever its source (Benke *et al.*, 2011).

Projecting the potential spread and impact of newly introduced species requires the use of a population spread model. Since the seminal work of Fisher (1937) and Skellam (1951), the focus in general ecology has been mainly the development of species spread models, to understand the mechanisms of spread or to forecast the rate, direction and success of future dispersal events (reviewed in Hastings *et al.*, 2005; Higgins *et al.*, 1996). In the area of invasive species risk assessment, spread models help to identify the role of different pathways of spread (Robinet *et al.*, 2009, 2011, 2012a) and have been used to assess the value of the adoption of a strategic control zones to slow the spread of invasive species (Buckley *et al.*, 2005; Sharov & Liebhold, 1998; Sharov, 2004). The most recent spread modelling approaches incorporate different levels of complexity such as the dynamics of population growth, stratified dispersal, propagule pressure, Allee effects, as well as realistic habitat landscape heterogeneity using a geographic information systems (GIS) framework, climatic variability or host diversity (see for example Carrasco *et al.*, 2010; Keesing *et al.*, 2006; Keith *et al.*, 2008; Meentemeyer *et al.*, 2011; Parry *et al.*, 2013; Pitt *et al.*, 2009). Methods that integrate invasive species spread and climate suitability with economic factors have been developed to assess the degree to which economic policies or control strategies can be used to optimize management decisions of invasive species invasions (Barbier & Shogren, 2004; Cook *et al.*, 2007b; McDermott *et al.*, 2013; Richards *et al.*, 2010). At the very least, initial investigations suggest that invasive species risk maps can narrow the set of geographical locations that would need to be targeted, thus reducing costly inspections and public outreach activities, while the incorporation of uncertainty analysis would add to their credibility (Venette, 2015).

Along with the increasing application of such approaches, criticism has been raised concerning the usefulness of such models for guiding effective management of invasive species at a regional scale (Robinet *et al.*, 2012a). Spatially explicit spread models are considered to be biologically and technically complex, and require too many poorly known parameters for their projections to be reliable in practice (Hartig *et al.*, 2011). In particular, complex spatially explicit spread models might not provide credible quantitative estimates of where a species might establish and spread, particularly at local scales (Caley *et al.*, 2008; Robinet *et al.*, 2012a). Despite this, new analytical methods are being developed to provide formal quantitative measurements of uncertainty and to address the perceived aversion to the concept of model uncertainty of some biosecurity decision-makers, but they are not yet routine (Cook *et al.*, 2012; Venette, 2015). However, Venette *et al.* (2010) called for substantial improvement in visual decision-support model documentation, communication of uncertainty, and improved training with respect to invasive species spread modelling. There is a large body of literature that has increased our understanding of the propagation of error arising from model inputs, to output and predictions, as well as methods that aim to make the best decision in the face of extreme uncertainty (Gould *et al.*, 2014; Kriticos *et al.*, 2013; Matott *et al.*, 2009; McGeoch *et al.*, 2012; Regan *et al.*, 2002; Yemshanov *et al.*, 2010). Nonetheless, no general guidelines exist that permit a specific invasive species spread map to be unambiguously linked to uncertainty analysis and therefore provide guidance for decision makers with respect to their interpretation.

The primary aim of this study was to, 1) review the sources of uncertainty associated with the development of a spread model, and, 2) discuss a selected set of methods for gauging spatially-explicit, individual-based spread model performance and uncertainty. I argue that a clearly defined role of science in the decision-making process, along with accurate estimation and communication of the information content of any spread model, will determine the success of a spread forecast initiative. I review current good practice and the research that is required to develop an improved future capability.

6.2 Sources of uncertainty in invasive species spread models

6.2.1 Typologies and sources of uncertainty

Uncertainties are *intrinsic* to all risk map analyses of invasive species and is *unavoidable*, even if some forms of uncertainty can be reduced (Cook *et al.*, 2012). The question is therefore not how to remove uncertainty, but how to most appropriately represent it in a model to evaluate it in a way that is reproducible (Haefner, 2005; Matott *et al.*, 2009; O'Sullivan & Perry, 2013; Refsgaard *et al.*, 2007). In other words, the challenge is to demonstrate that a particular model forecasts the course of invasion spread better than could be obtained by chance alone (Venette, 2015).

To better understand the sources of uncertainty associated with models in ecology Regan *et al.* (2002), Matott *et al.* (2009) and Drescher *et al.* (2013) categorized uncertainty as *inherent uncertainty* that is associated with the natural variability of the studied system, or *epistemic uncertainty* that is derived from incomplete knowledge about the system of interest. While inherent uncertainty, which results particularly from strong nonlinearities and stochasticity in the modelled system, is irreducible; epistemic uncertainty can, in theory, be reduced through additional data or empirical research. Among the main sources of uncertainty in biological modelling, McMahon *et al.* (2009) include: process error (the way constructs are structured, formulated and used to represent observed processes), measurement error, intraspecific variability, spatial and temporal heterogeneity, uncertainty about initial conditions, and even user uncertainty and software implementation uncertainty (Table C.1). Walker *et al.* (2003) described epistemic uncertainty as anything which '*provides a deviation from the unachievable ideal of completely deterministic knowledge of the relevant system*'. They attempt to classify uncertainty further, to develop a clear understanding of the various types of uncertainty, how each arises and how different types might be best dealt with. Such classification relies on the *location* of the uncertainties (for example with respect to the original question, parametrization of the model, model structure, etc.), the *level* of uncertainty (from epistemic knowledge to complete ignorance) and the *nature* of uncertainty (natural variability as opposed to measurement error). Ascough *et al.* (2008) considered three additional types of uncertainty: *variability uncertainty*, that is generated by the inherent variability in natural and human system (e.g. institutional and social context dependency, human factor uncertainty), *linguistic uncertainty*, that arises because our natural language is vague and ambiguous, and *decision-making uncertainty*, that emerges whenever there is conflict or ambiguity about how to quantify and compare economic, societal and environmental objectives.

6.2.2 Uncertainty and models of species spread

With respect to spread models in particular, McGeoch *et al.* (2012) provided a comprehensive review of the sources of uncertainty when identifying and prioritizing invasive species. This study concluded that most error associated with species occurrence and environmental data are epistemic in nature, and in most cases, resulted from a fundamental lack of knowledge of risk elements or measurement error. For example, parameters such as rate of increase, Allee threshold, and dispersal rate, are notoriously difficult to infer. The public and decision-makers often need swift action when a new invasive species is detected in a given area. Such urgency often does not allow time to acquire the necessary knowledge to unambiguously characterize the potential spread and impact of the species (Venette, 2015). In addition, the relative role of spatial and temporal heterogeneity, as well as processes such as intraspecific variation, biotic interaction, and/or rapid evolution in shaping the establishment and spread of invasive species may be relatively unknown (Gould *et al.*, 2014).

Systematic characterization of uncertainty in a spatially-explicit model of spread, is further complicated by uncertainties associated with spatial data (O'Sullivan & Perry, 2013). Such uncertainties include, locational error (uncertainty about where an invasive species is in space), classification error (error associated with the nature of the entity, such as mis-classification of habitat or climate suitability), and errors related to scales (grain/resolution and extent of the spatial data). Spatial or locational uncertainties in spatial data have been thoroughly reviewed in Zhang & Goodchild (2002) and Lu & Weng (2007). Furthermore, Gould *et al.* (2014) provided a review of different sources of uncertainty in species distribution modelling and provided a conceptual framework for understanding and communicating the impact of uncertainty on such models. Importantly, uncertainty in models of spread represents a cumulative measure that may arise from a variety of sources that can be spatially structured (i.e. spatially autocorrelated), and that propagates in complex ways through spatial simulation (Ascough *et al.*, 2008). The different sources of uncertainty with respect to spread models are summarized in Table C.1.

While the relevance of characterizing uncertainty when characterizing or communicating spread risk is widely acknowledged, in practice, it is often overlooked (Cook *et al.*, 2012). In addition, the prediction of spread traditionally focuses on a single-species approach leading to conclusions that are most applicable to the specific conditions under which the model has been developed, and are therefore not readily be transferred to another situation (Robinet *et al.*, 2012a). This lack of generality is a major impediment to the development of a coherent invasive species risk assessment that would represent an understanding of when and why a particular set of factors is required to explain the dynamics of invasive spread (Venette, 2015). Overcoming this impediment requires a new culture of analysis and presentation of spread models (Matott *et al.*, 2009). Instead of focusing on making sure that a model appears to correctly represent the set of processes to explain certain observations, modellers should more often discuss or demonstrate, when and why a particular model is not appropriate (Hartig *et al.*, 2011; O'Sullivan & Perry, 2013; Parry *et al.*, 2013).

Furthermore, there exists a wide array of analytical approaches to evaluate model performance, ranging from various ways of comparing model output against observed patterns of spread, to more qualitative approaches. Basically, the purpose of all such methods is to identify uncertainty in the model, and assess which uncertainty matters the most (O'Sullivan & Perry, 2013). Despite that there is substantial analytical capacity to evaluate model performance, most software packages for species distribution modelling or risk map analyses lack good methods to represent uncertainty (Hartig *et al.*, 2011; Matott *et al.*, 2009). A special challenge for species distribution model outputs (usually a map) is that their reliability may need to be judged at multiple scales, such as the entire map, particular regions, or perhaps, most critically, for individual and sometimes very specific geographic locations (O'Sullivan & Perry, 2013). The ways to deal with these uncertainty will be further discussed below. The ways to deal with uncertainty in spatially explicit spread models of invasive species will be further discussed below.

6.3 Identifying and quantifying uncertainty in invasive spread forecasts

A plethora of methods have been proposed to evaluate model performance or error, and therefore indirectly quantify uncertainties in ecological models. Haefner (2005), Matott *et al.* (2009) and Hartig *et al.* (2011) have provided a thorough overview of concepts and methods for model evaluation as well as a catalogue of software tools available to facilitate the process. In addition, Jager & King (2004) and Lu & Weng (2007) proposed a taxonomy of approaches for specifically evaluating uncertainties in the particular context of spatial modelling. Similarly, Evans (2011) and O'Sullivan & Perry (2013) reviewed the issues involved with understanding error in spatially stochastic (agent-based) modelling, covering a broad range of methodologies and viewpoints from across spatial modelling science (see Table 6.1). In the following sections, I briefly summarize a few of the methodologies that have been applied to spatially-explicit models of spread and/or to forecast the spread of invasive species in heterogeneous landscapes.

6.3.1 Model based-uncertainty

6.3.1.1 Error analysis (EA)

EA aims to quantify the influence of different sources of input estimation error on a given model's dynamics and output (Haefner, 2005; Jager & King, 2004; Matott *et al.*, 2009). EA also helps to understand how errors combine within the system (error amplification or compensation). Parysow *et al.* (2000) and Mcgwire & Fisher (2001) describe error analysis and, in particular, spatial error budget analysis, as a method to systematically partition the contribution of different sources of error introduced by each parameter in the

Methods	**Description**	**Selected references**
Model based-uncertainty		
Error analysis (EA)	Identification of the sources of error that cause the largest variation in model forecast	Haefner (2005); Hartig *et al.* (2011); Evans (2011); Jager & King (2004); Matott *et al.* (2009)
Sensitivity analysis (SA)	Identification of model output components most sensitive to local (spatially distributed) input variables	Crosetto *et al.* (2000); Haefner (2005); Jager & King (2004); Matott *et al.* (2009)
Uncertainty analysis (UA)	Identification of how uncertainty in multiple (interacting) parameters and their representation influence uncertainty in model forecast	Hartig *et al.* (2011); Jager & King (2004); Matott *et al.* (2009)
Bayesian networks (BNs)	Combine prior distributions of uncertainty to yield an updated (posterior) set of distributions	O'Sullivan & Perry (2013); Matott *et al.* (2009); Railsback & Grimm (2011)
Spatial data analysis (SDA)	Detecting and quantifying characteristics of geographical data and, specifically, on spatial autocorrelation, spatial heterogeneity and scale-dependence structure	Dormann *et al.* (2013); Evans (2011); Jager & King (2004)
Robustness analysis (RA)	Analysis of the extent to which different representational decision influence model dynamics	Evans (2011); Matott *et al.* (2009)
Confrontational methods		
Visualization and difference measure	Visual comparison of empirical observations and model predictions (non-spatial and spatial measures)	Fox & Hendler (2011); Grimm (2002); Spiegelhalter *et al.* (2011)
Statistical methods	Quantitative comparison and analysis of predictions and observations (via linear regression models, correlation, etc.)	Mayer & Butler (1993)
Exploratory/heuristics		
Pattern oriented modelling (POM)	Use of multiple observed patterns to evaluate and refine models and select between alternate representation	Grimm & Railsback (2012b); Railsback & Grimm (2011)
Multi-Model Analysis (MMA)	Generate ensemble predictions via consideration of multiple plausible models	Burnham *et al.* (2011); Burnham & Anderson (2002)
Participatory modelling (PM)	Methods involving expert opinion into model design and evaluation	Krueger *et al.* (2012); Martin *et al.* (2012); Millington *et al.* (2011)

Table 6.1: Selected approaches and tools for the evaluation of spatially-explicit models (modified from O'Sullivan & Perry, 2013)

spread model. A clear detailing of the technique from the point of view of tracking input errors can be found in the Joint Committee for Guides in Metrology report, JCGM (2008).

6.3.1.2 Sensitivity analysis (SA)

SA is among the most widely used methodologies for assessing uncertainty. SA seeks to rank input parameters by their relative influence on variation and uncertainty in the target output variable. SA involves systematic alteration of model parameter values and evaluating their effect on model outputs. In the case of a spread model, this might involve the population rate of increase, Allee threshold, mean long-distance dispersal and/or the classification of habitat suitability. In its traditional form, SA is often conducted using a 'local' approach where the parameters of interest are systematically varied one-at-a-time (no interaction considered) by some small amount. Among others, Fassò & Perri (2006), Saltelli *et al.* (2008) and Zajac (2010) reviewed a large variety of methods and tools available for sensitivity testing, but only a few are well suited for spatial models. Some headway has been made developing tools for evaluating uncertainties in the spatial context (see for example Congalton & Green, 1993; Crosetto *et al.*, 2000; Kocabas & Dragicevic, 2006; Pontius, 2002), but rigorous evaluation for spatially explicit models remains a real challenge due to the large number of factors and interactions between components of the models at different spatial scales (Jager & King, 2004; O'Sullivan & Perry, 2013).

6.3.1.3 Uncertainty analysis (UA)

UA is a more general approach that seeks to quantify the variation in model prediction caused by uncertainty in multiple, potentially interacting input parameters (Jager & King, 2004; Matott *et al.*, 2009; Railsback & Grimm, 2011). UA involves generating a probability density function for each parameter of interest, and quantifying the impact of input uncertainties on the empirical distribution of the model output. Many different approaches for conducting UA have been developed and reviewed in Matott *et al.* (2009) and Hartig *et al.* (2011). Among them, the Monte Carlo approach, which does not require assumptions about model structure, has been the most widely applied to spatially-explicit data (Crosetto *et al.*, 2000). A good introduction to Monte Carlo techniques in a spatial context can be found in Walker *et al.* (2003), along with a wide range of references to studies of sensitivity testing. A more generic study of uncertainty testing, concentrating on statistical summaries, can be found in Bobashev & Morris (2010).However, as emphasized in Railsback & Grimm (2011) and O'Sullivan & Perry (2013), the computational cost of covering the parameter space of complex models such as most spread models, rapidly becomes impractical.

To reduce the computational cost of global UA, two different approaches have been developed: the approximation and sampling methods (Evans, 2011; Matott *et al.*, 2009). Approximation methods characterize model output uncertainty by propagating one or more statistical moments (e.g., mean, variance, skewness, and kurtosis) of the various input distributions through the modelling system. Examples include error propagation equations (Gertner, 1987), point estimate methods (Tsai & Franceschini, 2005), and various reliability methods (Hamed *et al.*, 1996; Portielje *et al.*, 2000; Skaggs & Barry, 1996). On the other hand, sampling methods guide the selection of a structured parameter space that allows the extraction of a large amount of uncertainty with a relatively small input sample size. Helton *et al.* (2006) and Helton (2008) provided a thorough survey of sampling-based methods for uncertainty and sensitivity analysis. Among them, Latin hypercube sampling or quasi-random sampling have been the most widely used. With respect to estimating risks for emerging invasive species threats, probability models may be inadequately formulated because of the very high importance of rare events (i.e., events associated with the extreme tails of the distribution), which most probability models do not describe well (Kriticos *et al.*, 2013).

6.3.1.4 Bayesian Networks (BN)

BNs are probabilistic graphical models that combine prior distributions of input errors with general knowledge and site specific data to yield an updated (posterior) set of distributions. BNs can simultaneously represent uncertainty in input data and response data, as well as in model parameter distributions, model code, structure and resolution (Clark, 2005; Clark & Gelfand, 2006). Developing a BN involves, 1) defining a directed acyclic graph that specifies the conditional probability dependencies in the data, 2) defining prior probability distributions for all graph nodes (i.e. sources of uncertainty), and, 3) defining a likelihood function and sampling strategy (e.g. Markov chain Monte Carlo - MCMC) for inducing a posterior distribution based on prior distributions. Credal networks are regarded as an extension of BNs, where credal sets replace probability mass functions in the specification of the network variables (Cozman, 2000). These credals are groups of probability distribution that represent uncertainty about the probability model that should be used. Therefore credal networks allow the representation and manipulation of uncertainty in graphic models, where probability values may be imprecise or indeterminate.

6.3.1.5 Spatial Data analysis (SDA)

SDA refers to analytical, statistical and graphical procedures for evaluating and summarizing spatial input data. It typically comprises characterization of the spatial and temporal structure of input data. A key difficulty with spatial data is the presence of scale-dependent spatio-temporal correlation structures. Spatial and temporal autocorrelation can have a significant effect on the apparent sample size by introducing redundancy (Getis, 2007). These issues become particularly important when datasets for validating models are drawn from the same area by sample splitting (Araújo *et al.*, 2005), resulting in, for example, positive autocorrelation between sample units that can falsely reduce error levels. An overview of challenges arising from cross-scale analysis is provided by Fekete *et al.* (2010) and implications for the prioritization of intervention areas in the context of climate change can be found in Hagenlocher *et al.* (2014). The development of neutral landscape models by Gardner *et al.* (1987) and With *et al.* (1997) has also provided a new framework for generating replicated landscape patterns with partially controlled spatial properties. Neutral models allow hypothesis testing about how variation in spatial structure can affect model forecasts. Neutral models are used for generating alternative categorical landscapes such as in error analysis, with the exception that the generated spatial patterns do not represent deviations from a reference map.

6.3.1.6 Robustness analysis (RA)

As well as errors and uncertainty associated with input data, there are also epistemic uncertainties associated with model structure, in particular, with respect to the choice of the functional structure of the model and the choice of variables. The RA replaces the 'entire' model or submodel components with a different representation (or construct) to identify how the model behaves under different functional forms (Beven & Binley, 1992; Weisberg, 2006). Assessing model uncertainty has become the subject of considerable attention within the context of statistics (Burnham & Anderson, 2002; Johnson & Omland, 2004; Link & Barker, 2009; Lukacs *et al.*, 2010) and is currently an area of rapid development for assessing stochastic spatial simulation models (Grimm *et al.*, 2005; Hartig *et al.*, 2011; Schurr *et al.*, 2012; Wood, 2010; Thiele & Grimm, 2015)

Robustness analysis of stochastic models is particularly challenging, in part because their likelihood functions cannot usually be calculated explicitly. It is therefore difficult to couple such models to well-established statistical theory such as maximum likelihood or Bayesian statistics. A number of new methods, among them, genetic programming, approximate Bayesian computing (ABC), Metropolis-Hasting Markov chain Monte Carlo (MCMC), pattern-oriented modelling (POM), and synthetic likelihood, bypass that limitation (Clark, 2005; Clark & Gelfand, 2006; Matott *et al.*, 2009; Poli *et al.*, 2008). These methods share three main principles, 1) aggregation of simulated and observed data via summary statistics, 2) likelihood approximation based on the summary statistics, and, 3) efficient sampling.. Bolker *et al.* (2009) and Hartig *et al.* (2011) provide thorough overviews of RA techniques.

6.3.2 Visualization and confrontational approaches

6.3.2.1 Visualization techniques

Explanation and communication of uncertainty to an audience with a wide range of scientific knowledge presents a particularly significant challenge. As noted in Grimm (2002), MacEachren *et al.* (2005), Fox & Hendler (2011), Rocchini *et al.* (2011) and Spiegelhalter *et al.* (2011), visualization of a model's output and uncertainties should play a vital role in evaluating and communicating model-based science. Visualizing data and their uncertainties appeals to intuition and emotion. The ways different types of visualizations are processed and understood clearly depends on factors such as the personality and numeracy of members of the audience. Fortunately, increasing availability of data and public interest has led to a golden age of interactive visualizations and multiple types of representation (infographics) that can be adjusted to the need of the audience (see for example, the New York Times and the Guardian, as well as IBM's Many Eyes – `http://www.boostlabs.com/`)-, and plotly – `http://plot.ly/` that allows visitors to create their own visualization). New approaches, however, are needed to determine how best to visualize particular kinds of scientific data. A significant contribution in this direction can be seen in the “Periodic table of Visualization Methods” (Lengler and Epler, `www.visual-literacy.org`) which associates a number of visualization techniques, coupled with emerging Web-based technologies, organized by type of data they apply to and the complexity of their application.

6.3.2.2 Difference measures

Visualization techniques can be supported by difference measures that summarise and quantify the predictive accuracy of a model. Such metrics quantify the ‘distance’ between a set of observations and the corresponding model predictions. In a spatial context, a simple comparison of categorical maps comprises cross tabulated predicated and observed values to form an error or confusion matrix such that the diagonal elements of the matrix correspond to the correct predications, while the off-diagonal elements represent incorrect predictions (Allouche *et al.*, 2006; Fielding & Bell, 1997). From such a matrix a number of model performance measures can be derived, including:

- the sensitivity index – proportion of correctly predicted positive observations (Allouche *et al.*, 2006; Fielding & Bell, 1997)
- the specificity index - proportion of correctly predicted negative observations (Allouche *et al.*, 2006; Fielding & Bell, 1997)
- the AUC - area under the receiver operating characteristic curve (ROC) plot that measures the discriminatory capacity of classification models (Fielding & Bell, 1997; Jiménez-Valverde, 2012; Lobo *et al.*, 2008)
- the kappa statistic – the proportion of correctly classified locations after accounting for the probability of chance agreement (Cohen, 1960)
- the precision index - calculated by both bootstrapping (re-sampling of the data) and cross validation (partitioning data) to test the model’s ability to generalize or predict new data (Worner *et al.*, 2010)

A thorough review of these methods can be found in Freeman & Moisen (2008) and Senay (2014). The main issue with such approaches is that they tend to confuse errors in quantification with those in location. For spatial modellers, it is key to understand the distribution of uncertainty in space and time. Uncertainty can, therefore, be usefully displayed on maps. Therefore, efforts have been made to develop metrics such as fuzzy set map comparison (Hagen, 2003) or hierarchical fuzzy pattern matching (Power *et al.*, 2001), and have been reviewed in Pontius (2000, 2002), Visser & De Nijs (2006) and Hartley *et al.* (2006). In the specific case of biological invasion, long-term observational data sets documenting range expansion are rare, therefore there have been few direct comparisons of predictive accuracy from spatially-explicit models to long-term records of spread (Fitzpatrick *et al.*, 2012; Pitt *et al.*, 2009).

6.3.2.3 Formal statistical tests

Alongside visualization techniques and difference measures, frequentist statistical tests have been used to evaluate models against observations, including (non-linear) principal component analysis, linear and non-linear regression, (hierarchical) variance partitioning and (permutational) univariate and multivariate analysis of variance (Bolker *et al.*, 2009; Haefner, 2005; Matott *et al.*, 2009). As McCarthy *et al.* (2001) commented, the use of statistical tests of model performance can lead to a view of a model as being 'certain' once it has passed sufficient conditions. In other word, a modeller can simply increase the sample size of his data until he/she gets the desired p-value. To prevent that, O'Sullivan & Perry (2013) concluded that in the evaluation of any model, a pluralist approach that combines careful visualization and appropriate quantitative tools, that can be adjusted to the need of the audience, is likely to be the most successful.

6.3.3 Exploratory/heuristic approaches

6.3.3.1 Multi-Model Analysis (MMA)

Commonly, multiple runs of the same model using different initial conditions, parameters, as well as structural forms allow for an assessment of the variation in the model outputs by quantification of output error (Evans, 2011). It allows statistical summaries to be generated, along with confidence statistics (e.g. Pitt *et al.*, 2009; Senay *et al.*, 2013). Quantitative MMA assigns performance scores to each candidate model (see section 6.3.1 and 6.3.2) (Burnham & Anderson, 2002; Ye *et al.*, 2008; Senay, 2014). The difference between model outputs can be used to make statements about the observed data and to make forecasts based on model averaging or a consensus approach (ensemble forecasting).

Ensemble forecast systems address, at least in part, uncertainties due to model structure (Burnham *et al.*, 2011). The potential of combining different models using a frequentist approach has been reviewed in Anderson (2008) and Burnham *et al.* (2011), while Link & Barker (2009) provided an overview from a Bayesian perspective. Hartig *et al.* (2011) extended previous work by considering methods for comparing dynamic simulation models. Furthermore, Collins (2007) gave an insight into the state of the art of ensemble modelling in much of the field of complex systems modelling. A good starting point is the generalized likelihood uncertainty (GLUE) framework, developed in Beven (2006), which provides a means of weighting models in ensembles based on their likelihood. Alternatively, Piou *et al.* (2009) outlined an approach derived from the deviance information criterion (DIC) for comparing simulation models in a multi-model framework. Finally, Araújo & New (2007) provided an overview of ensemble forecasting to the analysis of bioclimatic ensembles and discuss their uses and limitations for supporting policy decisions in biodiversity conservation. In general, combining multiple outputs improves forecast reliability (Johnson & Omland, 2004). In addition, the combination of predictions allows forecast sharpness (the closeness of forecasts) to be assessed as an additional uncertainty measure (Gneiting & Raftery, 2005).

These techniques, nevertheless, require observational data or at least knowledge of the structure of uncertainty, which for new invasive pests is typically lacking. Therefore, quantitative MMA in the context of forecast of spread for invasive species is more likely to rely on expert elicitation and assurance or quality control procedures (rather than on quantitative measures) to assess the relative merits of alternative models. An alternative approach, that aims to overcome the lack of data, consists in using models to conduct experiments in a 'virtual (*in silico*) laboratory' (O'Sullivan & Perry, 2013). Pattern-oriented modelling proposed by Grimm & Railsback (2012a) fits well within this approach.

6.3.3.2 Pattern-oriented modelling (POM)

Grimm & Railsback (2012a) defined POM as '*the multi-criteria design, selection and calibration of models of complex systems*'. POM is intended to make use of a set of patterns observed at multiple scales and levels of organization to inform the design, testing, and evaluation of detailed simulations (Wiegand *et al.*, 2003; Grimm & Railsback, 2012a). These multiple patterns are used to guide three key phases of the modelling process, 1) determining the scales, variables and processes to be included in the model, 2) choosing between different representations of key low-level processes, and, 3) calibrating the model. A formal likelihood based, or Bayesian framework can be used for step two (see section 6.3.3.1), but it could be equally a more informal procedure in which different hypotheses are tested. For example, different population models can be developed with or without considering an Allee effect. As Wood (2010) noticed, an important decision in this process is the choice of appropriate summary statistics that will be used to describe the observed patterns. He further recommended the use of more than one statistic to limit information loss and provide a brief overview of the types of measures that can be used. POM formalizes the use of multiple patterns for developing and testing stochastic models that helps to communicate models more effectively (Grimm *et al.*, 2005).

6.3.3.3 Participatory modelling

There exists a wide range of methodologies for involving expert knowledge in model design and assessment. At the simplest level, it involves (formal or informal) expert validation of parameters and processes. For example, expert knowledge can be used to constrain parameter values and ranges (Janssen *et al.*, 2010) or to develop priors for Bayesian treatments of uncertainty (Clark, 2005; Kavetski *et al.*, 2006; Kuhnert *et al.*, 2010). More sophisticated approaches formally incorporate expert opinion into automated procedures. For example, Martin *et al.* (2012) and Krueger *et al.* (2012) provided a critical review of the role of expert opinion in environmental modelling. They conclude that 'participation' of experts is a key element of contemporary approaches in complex environmental modelling. However, the use of domain experts in modelling will strongly benefit from '*formal, systematic and transparent procedures*'. Krueger *et al.* (2012) called for '*a paradigmatic shift from traditional unbiased and impartial experts towards unbiased processes of expert elicitation and a plurality of expertise and eventually models*'. They further highlight the role for science to maintain and enhance general procedures that will improve accuracy and information content and ensure uncertainty is captured accurately. Just as the reliability of empirical data depends on the rigour with which it was acquired, so too does that of expert knowledge.

6.4 Model uncertainty in relation to risk assessment

There has been considerable progress in recognizing and communicating the limitations and uncertainties associated with invasive species distribution and spread modelling forecasts. Buisson *et al.* (2010), Hartley *et al.* (2006), Gould *et al.* (2014), Pitt *et al.* (2009), Marmion *et al.* (2009), Senay (2014) and Thuiller *et al.* (2006) have identified sources of uncertainty in invasive pest or insect species distribution models. Their approaches use a multi-model inference to generate confidence intervals that incorporate both the uncertainty involved in model selection as well as the error associated with model fitting. Hartley *et al.* (2006) further quantified the costs of making false negative versus false positive assessments in order to directly connect modelling to decision-making. Alternatively, Heikkinen *et al.* (2006), Latimer *et al.* (2006), Marion *et al.* (2012) formulated and fitted species distribution models into a Bayesian framework to incorporate uncertainty about dynamic processes such as reproduction, mortality and dispersal into spatial and temporal projections of species distribution and abundances. Cook *et al.* (2007a) and Bierman *et al.* (2010) developed a Bayesian approach to parameter estimation in a stochastic spatio-temporal model of the spread of invasive species. The methods were used to account for critical spatial heterogeneity. Along the same line, Eiswerth & Van Kooten (2002) developed a stochastic model for invasive weed fuzzy sets that recognizes several sources of uncertainty, including lack of data, measurement error, variability in rate of spread, and impact of management measures. Finally, Ben-Haim (2006), Carrasco *et al.* (2012), Moilanen & Wintle (2006) and Yemshanov *et al.* (2010) applied information-gap theory to invasive pest distribution models. Information-gap theory aims to make the 'best' decision in the face of extreme uncertainty. It accommodates non-statistical uncertainties, such as the subjective choice of candidate variables and the structural assumptions embedded in spatial analysis, to account for unknown levels of potentially extreme uncertainty.

Assuming that pest risk modellers integrate their public role into decision-making process by acknowledging their responsibility to communicate limitations and uncertainties, a second, more difficult challenge then is to consider how to display or show uncertainty to policy makers and the general public (Evans, 2011; Mouquet *et al.*, 2015). This is a contentious area that challenges the relationship between scientists, decision-makers, and the public (see for review Brown, 2010). Unfortunately, scientific uncertainty can be converted into policy reticence, even when the science points strongly to action. If policymakers fail to act because they think the possibility of harm too uncertain – and they are wrong – then the failure to act could be catastrophic. Equally, if policymakers assume the possibility of harm near certain, and they act to mitigate it, they may turn out to be wrong and have committed costly resources and investment in a wasteful fashion. In contrast with scientific goals, the central goal of policy makers is not necessarily to understand the sources of uncertainty *per se*, but rather to understand the acceptability of random uncertainties with respect to the acceptability of the risks associated with them. However, determining the acceptability of risks is a complex process where a number of factors, for example, economic, societal, and environmental needs and constraints, may influence whether or not a risk is judged acceptable. Given this goal, the question then becomes, how should uncertainty be treated in risk assessment modelling?

Shackley & Wynne (1996) and McCarthy *et al.* (2001) discussed briefly some of the mechanisms by which scientists mitigate the effects of uncertainty. Walker *et al.* (2003) proposed a taxonomy to communicate uncertainties to decision-makers. For example, they identified *scenario uncertainty*, in which it is not clear what scenario is going to occur. Such taxonomy helps to overcome linguistic uncertainty (as described in sectionn 6.2.1 and Table C.1) and facilitate the communication of uncertainty with results to decision-makers. Along the same lines, Cornélis & Brunet (2002) examined spatial uncertainties from the point of view of decision-makers. Morss *et al.* (2008) discussed how the public understand uncertainty and want it displayed, while Dewulf *et al.* (2005), Mouquet *et al.*

(2015) and Refsgaard *et al.* (2007) provided a detailed discussion on the weight of scientific methodology and decision-makers engagement in the decision-making process. On the other hand, Brown (2010) and Couclelis (2003) provided useful discussions on the place of uncertainty in science as a social process, and uncertainty's place in scientific self-reflection and knowledge production. Most of these studies concluded by stating that the diversity of tools and approaches, as well as the complexity of the interactions between scientific, decision-makers, and the public at large, has lead to a clear and on-going lack of consistent standards and tools that limit our ability to deal with uncertainty and effectively communicate uncertainties to decision makers.

6.5 Discussion

Currently, there is considerable interest in uncertainty within the context of environmental and ecological decision-making (Mouquet *et al.*, 2015; Schindler & Hilborn, 2015). This is motivated by a recognition that there is a pressing need for reliable forecasts of ecosystem state, ecosystem services, and natural capital (Clark *et al.*, 2001). Within the broad areas that are now 'environmental and ecological modelling', there has been considerable progress in data availability, computation and statistics, which means that the purpose for which models are designed are much more diverse than they were a few decades ago. This diversity of uses and the recognition of fundamental problems in seeing models in terms of robust forecasts, has lead to the diversification of approaches to evaluate models and uncertainties (Hartig *et al.*, 2011; Matott *et al.*, 2009). In this study I briefly reviewed the sources of uncertainty about species distribution and spread risk modelling, methods of representing them, and how these different approaches have been applied to one important contemporary policy question — biological invasion. Although we are constantly making incremental progress to understand, characterize, and forecast the invasion process, it is clear that uncertainty, whatever its source, is here to stay. Uncertainty must therefore be addressed in any comprehensive risk assessment and decision-making situation. Initial investigations suggest that the incorporation of uncertainty analysis and the appropriate communication of the limits associated with an invasive species spread risk forecast adds to the credibility of an invasive species or risk map (Venette, 2015). Despite these sentiments, considerably more could be done to effectively characterize and communicate uncertainty about spread risk forecasts and consequences for policymakers.

There has already been an increasing focus on errors associated with input data and processes, and how these can have an overwhelming effect on model uncertainty and can alter the treatment decision. Measures appropriate for many sources of uncertainty are already available, both for quantifying uncertainty, through, for example, sensitivity analysis, uncertainty analysis and performance measures and for incorporating uncertainty into the modelling effort, for example, Bayesian approaches. However, prediction of spread traditionally focuses on a single-species and environment approach, using a single class of spread models, leading to conclusions most applicable to the specific conditions under which the model has been developed. Structural uncertainty, or uncertainty associated with the choice of the functional structure of the model and variables, tends to be overlooked. Consideration of a combination of different models can be seen as an opportunity to make forecasts at different spatio-temporal scales and organisational levels. Combined models can facilitate correct interpretation of spread, where differences and similarities between different classes of distribution model can be identified pointing to when and why a particular set of factors should be considered for accurate forecast. However, projections by alternative models can give highly variable forecasts that can lead to policy reticence. A solution to communicating/representing inter-model variation is the development of methods that combine several separate forecasts together; such as by consensus, averaging or bounded forecasts (Araújo & New, 2007; Hartig *et al.*, 2011). If used appropriately these approaches can enable more robust decision making in the face of uncertainty. Combining forecasts, however, remains dependent on individual predictions thus improved individual forecasts will yield a better combined forecast.

Essential to the concept of robust forecasting is also the need for data to test and validate models. Ecology is undergoing a major transformation with the accumulation of large data sets. There is therefore an urgent need to create standards for data acquisition, storage and sharing, but also standards to communicate uncertainty associated with these data. A good starting point is the eXtensible Markup Language schemata that promises to take uncertainty recording and manipulation from the current level of the dataset down to the specific datum, storing detailed uncertainty information with each data point. A notable example for spatial modellers is UncertML (`http://www.uncertml.org/`) which has the potential to be used with the Geographical Markup Language, along with the web-based framework supplied to aid in its more general use, UncertWeb (`http://www.uncertweb.org/`).

But more importantly, the development of formal approaches to uncertainty that engages scientists and decision-makers will increase our ability to communicate forecasts. As in Evans (2011) uncertainty representation could be embedded within the demands of quality assurance (QA) guidelines. See, for example, Refsgaard *et al.* (2007) and Joint Committee for Guides in Metrology report, JCGM (2008) who described schemes designed under the ISO 9000 standards family. Such guidelines should also include detailed frameworks for decision-making under uncertainty. Interactions with other researchers in other disciplines, including statistics and social scientists will help decide the most appropriate analytical tools. Any decision must be a pragmatic component of the infrastructure and dynamics of the decision-making process. Such goals and efforts will encourage enhanced tool interoperability and facilitate the communication of uncertainty with results. As models (whatever their form) are used more often to support and develop policy, I believe that methods that encourage the participation of stakeholders - so called participatory modelling - will facilitate the communication of uncertainty around models. Additionally, such methods will also facilitate the development processes such as adaptive management and improved evaluation methods.

6.6 Conclusion

The growing societal need for predictions of current and future biological invasions has led to the development of new modelling approaches to more closely link science, decision makers and the public. Although we are constantly making incremental progress to understand, characterize, and forecast the invasion process, it is clear that uncertainty, whatever its source, is here to stay. Uncertainty must therefore be addressed in any comprehensive invasive species risk assessment and decision-making situation. Any species distribution model or spread model used for decision support should include

- Development of standardized measures for evaluating model performance that are understood and accepted by the risk assessors and decision makers (e.g. type I/type II/type III errors or false positive/true positive measures)

- Development of standardized methods for assessing uncertainty in species distribution models and spread risk forecasts (e.g. data analysis to characterize input and response data, sensitivity analysis to determine the most important set of parameters and uncertainty analysis to establish the range of predicted outcome)
- Development of standardized datasets to allow inter-model comparison of uncertainty and development of appropriate open-source software for increasing computational efficiency and optimization
- Development of standardized methods for showing and communicating uncertainty (e.g. ensemble modelling, taxonomy of uncertainty, participatory modelling)
- Improvement of methods that incorporate assessment of uncertainty, whatever its source, into risk models through the use of fuzzy-rule based model or Bayesian probability approaches
- Incorporation of adaptive management practices which will help risk assessor and decision makers maintain flexibility in their decisions, knowing that uncertainties exist and so provides the latitude to adjust direction to improve progress towards desired outcomes (e.g. participatory modelling).

Chapter 7

General discussion

In this chapter, the major findings from the previous chapters are discussed in relation to the aims and objectives of the thesis. In addition, a number of recommendations are given for future research regarding the effect of landscape structure on the establishment and spread of invasive species. Finally, concluding remarks are given.

7.1 Research context and objectives

Chapter 1 provided a broad review of the factors governing the establishment and spread of invasive species in changing environments and of the existing modelling approaches used to describe and predict rate of spread and local population density of invasive species. Given that so many ecosystems face rapid and major environmental changes that can be conductive to the establishment and spread of invasive species, there is a growing awareness of the need to adopt proactive management strategies that attempt to forecast species responses to future environments (Guisan & Thuiller, 2005; Thuiller *et al.*, 2008; Mouquet *et al.*, 2015). Few of the models most used to predict future distribution of invasive species explicitly integrate dispersal processes or characteristics of species population dynamics. In particular, little attention has been given to simulating demographic and dispersal processes at the leading edge of the distribution where range expansion occurs (see, for example, the review by Thuiller *et al.*, 2008; Sexton *et al.*, 2009). In addition, most efforts related to the prevention and control of invasions are conducted at the local scale, where the influence of the landscape structure on the process of invasion is seldom explored (Vilà & Ibáñez, 2011). Understanding the influence of landscape composition, configuration, and their spatio-temporal interactions on ecosystem susceptibility to invasion, is currently confined to broad generalization. For example, it is well recognised that while ecological corridors facilitate dispersal, they can also increase invasion risks and that anthropogenic ecosystems are generally more invaded than natural ecosystems (Vilà & Ibáñez, 2011; González-Moreno *et al.*, 2014). Such generalisation ignores the differences in the population density and rate of spread of invasive species within a particular land-use type, as growth and dispersal can be fine-tuned by small-grain landscape differences such as habitat size, habitat edge, and habitat connectivity variables (With, 2002). Accurate accounts of the invasion process and effective eradication or conservation programs will depend on such considerations (With, 2002; Vilà & Ibáñez, 2011). Assessing the effect of landscape structure on species spread poses a great challenge to distribution modelling because characterising the effects requires quantitative information on the relationship between demography, dispersal and the environment in which these processes occur.

In this thesis, I proposed four main avenues to progress the understanding and prediction of the establishment and spread of invasive species in changing environments. First, based on an existing model described by Pitt (2008), I developed a simulation framework that allows for systematic investigation of changes in landscape structure on population density and spread of invasive species (Chapter 2). Second, I identified a set of key generic landscape metrics that allow a concise characterization of independent aspects of the landscape structure, regardless of the scale and resolution at which the patterns are represented (Chapter 3). Third, I investigated the independent and interactive effects of the different aspects of the landscape structure on the establishment and spread of invasive species (Chapter 4). Finally, I determined the relative importance of the effect of landscape structure in relation to dispersal ability, reproductive rate and propagule pressure on population growth and spread of invasive species (Chapter 5).

7.2 The importance of a general spatially-explicit model of spread

The dispersal modular framework, MDiG, introduced in Chapter 2, was developed in response to a call for explicitly accounting for population dynamics and dispersal behaviours within a modelling framework integrated in geographic information systems

(GIS) (Worner, 1994; Guisan & Thuiller, 2005). MDiG accommodates for spatio-temporal heterogeneity and is partially stochastic (Pitt, 2008).

The framework is very flexible in that it could be applied at multiple spatial scales and can easily be modified for species with structurally different demographic and dispersal behaviours to generalize the use of the framework to many plant, insect or other invertebrate species (Pitt, 2008). It provides a means of synthesising all our available knowledge to predict establishment, spread and impacts of invasive species. In Chapter 2, the great white butterfly, *Pieris brassicae*, was used to illustrate some of the capability of MDiG for modelling dynamic range expansion at a regional scale. Accurately estimating the distribution of an invading organism at any time in the future, including the time it takes to reach an equilibrium within its new environment, is of paramount importance for planning eradication strategies or even to decide whether any eradication effort is necessary or possible.

In addition to incorporating realistic landscapes, the MDiG framework allows the manipulation of fragmented and/or anthropogenic landscapes to test predictions regarding landscape modification or regional climate change for invasive species management purposes. The challenge to distribution modelling lies in predicting how demographic and dispersal responses will be modified by environmental conditions (Thuiller *et al.*, 2008). In this study, the MDiG framework was integrated and extended to further illustrate MDiG's capability by investigating the effect of change in landscape composition and configuration on population density and the dispersal of an invasive species, using the European gypsy moth, *Lymantria dispar*, as a case study. Inferring the relative role of habitat structure on the course of invasion relied on a judicious combination of methods, particularly the use of a landscape generator and landscape metrics that can be combined in an iterative process, to provide quantitative information on the relationship between demography, dispersal and the environment in which these processes occur. A model such as MDiG, with the capability of providing individual-based or raster cell-based outputs at high spatio-temporal resolution, can result in better understanding of these dynamics.

MDiG as a modelling framework integrates several components recognised as fundamental for tackling species establishment and spread under environmental change (Pitt, 2008). The framework was developed as a modular, open source project, facilitating collaboration and providing a solid foundation on which additional components can be easily added. Suggestions for future development of the framework are given in Section 7.7.

7.3 Detecting and quantifying landscape structure changes

The ability to quantitatively describe landscape structure is a prerequisite to detect changes and to investigate the relationship between landscape structure and, demographic and dispersal processes. In this regard, a plethora of quantitative metrics have been developed to provide simple quantitative measurements of the composition and configuration of a landscape (McGarigal *et al.*, 2012). However, the general perception is that there are three basic problems with using landscape metrics. They are, 1) a high degree of correlation between the metrics themselves, 2) ambiguous responses to different spatial processes, and, 3) sensitivity to change in spatial scale (resolution and extent) (Leitão & Ahern, 2002). Quantifying the specific effect of habitat configuration on establishment and spread success, for example, is difficult because many configuration metrics are correlated with habitat amount (Kupfer, 2012). Such limitations can often be addressed, or put in perspective, through careful data manipulation, analysis and interpretation.

In Chapter 3, the adequacy of widely used landscape metrics was tested to quantify concisely, independent aspects of spatial pattern, regardless of the scale at which they are represented. The findings highlight the sensitivity of landscape metrics to changes in the intrinsic characteristics of a landscape (spatial autocorrelation and percentage of suitable habitat cover), and scale-dependent factors (spatial extent and spatial resolution), individually but also to their interactions. The presence of significant complex interactions between intrinsic and extrinsic characteristics of landscape pattern makes it difficult to isolate purely landscape pattern driven effects from the effects of changing spatial scale. As such, this research illustrated the need for more systematic investigation of the relationship between intrinsic characteristics and extrinsic properties when accurate characterization of landscape pattern is a key input within spatially explicit ecological models. The self-organising map (SOM) clustering approach is proposed as an efficient way to identify relationships among metrics that more traditional statistical methods might overlook, by accounting for linear and non-linear interactions among predictors. Efficient analytical tools, such as SOM, may thus help to identify key generic landscape metrics to be used for quantifying the relationship between landscape structure and ecological processes, such as demographic and dispersal processes.

7.4 The importance of habitat amount and patch aggregation

In both Chapter 4 and Chapter 5, the effect of landscape composition and configuration on population density and spread of a invasive species was investigated. First, I used a multi-model averaging approach to explicitly test the degree to which different landscape elements such as suitable habitat amount, habitat patch size, isolation and quality, can explain population density and spread of invasive insects. Second, I used a boosted regression tree analysis to identify the relative importance of landscape processes in relation to dispersal ability, reproductive rate and propagule pressure. Both analyses were case-specific, where the MDiG modelling framework (Chapter 2) was parametrized on detailed demographic and dispersal attributes of the well studied European gypsy moth, *L. dispar*. In this way, the model had sufficient biological detail to reproduce realistic parameter ranges for a large flying insect. MDiG was then used to explore different ranges of demographic rates, dispersal behaviours and propagule pressure in changing environments. A patch based, binary view of the landscape, comprising suitable and unsuitable areas, was adopted. In these landscapes, habitat configuration, described as the degree of spatial autocorrelation, and habitat amount, described as the percentage of suitable habitat in the landscape, were systematically controlled.

The multi-model averaging approach used in Chapter 4 provided a basis for the development of a core set of structural metrics as indicators of invasive insect establishment and spread. Both population density and spread of invasive species can be affected in significant, and sometimes interactive ways, by the landscape based components: suitable habitat amount, habitat patch aggregation, core area, edge density and habitat shape complexity. In particular, the interaction between the percentage of suitable habitat in the landscape (PLAND) and a measure of habitat patch aggregation (SPLIT) had the greatest effect on both population density and rate of spread. To my knowledge, only a handful of studies have tested for the significance of interactions between the landscape variables (see for example Ewers *et al.*, 2007). As such, the research has illustrated the need for a more systematic investigation of the relationship between multiple landscape processes on the process of invasion.

The importance of suitable habitat amount and habitat patch aggregation in determining population density and spread was also highlighted by the boosted regression tree analysis used in Chapter 5. Contrary to expectations, both statistical approaches indicated a relatively weak effect of habitat patch connectivity (CONNECT)and edge density (ED) on population density and spread rate. This result reaffirms that population growth and spread might be limited more by the size and distribution of non-habitat gaps than by a measure of habitat connectivity as suggested in With & King (1999b) and Wang & Cumming (2011). The weak effect of edge density is more surprising as habitat edges are often thought to give invasive species a foothold in the landscape. For example, in a review of the landscape effects on the local population density and richness of invasive plants, Vilà & Ibáñez (2011) found that invasive plants are more abundant at habitat edges than in the interior of habitat. Edge density may be more important when habitat edge has a direct effect on growth and dispersal. For example, micro-climate variations at the habitat edge may enhance growth or higher wind disturbance may enhance long-distance dispersal events, rather than simply providing access to habitat. Such mechanisms were not considered in this research but could be easily integrating using buffering technique in GRASS-GIS to attribute a higher degree of suitability to the edge of the habitat patch than the inner core area when building a survival layer.

This work confirms the possibility that by quantifying the elements of the landscape that influence the population density and spread of invasive species, areas that are vulnerable to invasion can be better predicted. Subsequently, such analyses can be used to prioritize surveillance and control strategies. Furthermore, landscape metrics can be useful for proactive land-use planning to help establish the capacity of the landscape to maintain ecological processes, as discussed in Leitão & Ahern (2002) and Leitão *et al.* (2012). However, the capacity to generalise and compare findings from a range of ecosystems and range of taxa will help progress the identification of consistent landscape predictors.

Interestingly, it was found in both Chapters 4 and 5 that the particular effect of landscape structure on invasive spread depends on several life history characteristics of the species, such as the intrinsic rate of increase, and frequency of long-distance dispersal events as well as the presence or absence of an Allee effect. Therefore choice of an appropriate set of landscape metrics will require case specific exploration of their indicator values as it is unlikely that all species will respond to landscape structure in a similar way. Besides, variation across taxa would be useful for generating hypotheses and testing the degree to which various life-history and ecological traits influence sensitivity to landscape structure.

7.5 Ecology of invasion and implications for management strategies

The boosted regression tree analysis used in Chapter 5 indicated dominant processes shaping population density and spread of invasive species that can be prioritized as well as those that are of secondary importance. This simplification is needed in the context of pressing invasive species management, where simple yet effective advice is needed. Variables indicating Allee threshold, intrinsic rate of increase and propagule size were rated in as three most influential variables in each model. The collective influence of these three factors, as well as the effect of suitable habitat amount and habitat patch aggregation, is consistent with previous studies (see, for example, the review by Catford *et al.*, 2009). Despite these findings, this research offers new insights for management programs that aim to limit the establishment and spread of invasive species.

First, while propagule size was undoubtedly an important factor for ensuring population growth and spread in a new environment, demographic traits had an equal or greater influence on population density and invasive spread. These results are well in line with a recent discussion around the role of propagule pressure in determining invasion success. In particular, a predominant observation in a variety of taxonomic groups is that, as propagule pressure increases, so does establishment and spread success (Lockwood *et al.*, 2005, 2009). Intuitively, increasing the size and number of propagules enables the incipient population to overcome limitations associated with small populations, such as Allee effects, environmental stochasticity and demographic stochasticity. However modelling has shown that propagule pressure supplements demography, dispersal and life-history more generally (Cassey *et al.*, 2014). Possible reasons could be that because little or no information on invasion failure is readily available because empirical studies focus on a limited subset of life-history traits, biasing our perception of which factors are important for determining invasion success (McGregor *et al.*, 2012). Clearly, one must be cautious about assuming that reduced propagule pressure will lead to lower probability of establishment and spread, thereby lowering the assessment of risk.

Second, the primary importance of an Allee effect suggests that any eradication strategy should examine how eradication programmes can exploit such dynamics by maintaining population density under an Allee threshold (Chapter 5; Johnson *et al.*, 2006; Liebhold & Tobin, 2010). It is likely that management of invasive populations with Allee effects should be different from those without Allee effects. In particular, it was shown in Chapter 4 that in the absence of an Allee effect, population density and rate of spread is positively correlated with the proportion of suitable habitat, reproductive rate, dispersal ability and propagule pressure. In the presence of an Allee effect, however, spread is maximized at an intermediate dispersal level and inhibited at both low and high levels of dispersal. This result also highlights the existence of a dispersal threshold above which, both population spread and growth can be inhibited. This is a non-trivial result for management purposes. Adopting a strategy that systematically reduces the dispersal ability of an invading species that is occurring close to its dispersal threshold, may result in overcompensation leading to a very high rate of spread. However, the analysis undertaken in this research accounted for only three Allee threshold values (no Allee effect, weak versus strong Allee effect), and would benefit from further investigation of different threshold values to investigate how important Allee effects are in invasive species.

Third, a greater effect of growth-related traits than dispersal-related traits in both population density and rate of spread in heterogeneous landscapes, was found (Chapter 5). This result implies that spread rate is more strongly related to intrinsic rate of increase, which determines the total number of individuals participating in dispersal, than it is to a species' intrinsic dispersal ability. From an invasive species management perspective, this result suggests that priority should be placed on species with high intrinsic rate of increase and that eradication programmes should focus on limiting reproductive stages as a priority.

Fourth, no evidence was found to support the prediction that species with greater dispersal abilities have a higher rate of spread than species with lower dispersal abilities (Chapter 5). This result suggests that the intrinsic dispersal ability of a species is not a useful predictor of invasive spread. Indeed, analyses conducted in Chapters 4 and 5 showed that the realised dispersal distance in a landscape is a product of both species dispersal traits and landscape structure. The research reaffirms, as suggested in Fahrig (2007), that species dispersal ability cannot be estimated independently of landscape structure and that to test the relative importance of dispersal ability, the landscape context should match the location where the data were collected. It is further suggested that simple measures of landscape structure, characterising the percentage of suitable habitat and habitat patch aggregation in the landscape, may encapsulate the effect of landscape structure on population density and rate of spread of invasive species.

Finally, the frequency of long-distance dispersal events were a better indicator of population dynamics than the intrinsic dispersal abilities. However, in the absence of an Allee effect, population growth and average dispersal distance increased as the frequency of long-distance dispersal increased. However, in the presence of an Allee effect, while average dispersal distance increased with the frequency of long-distance dispersal events, population density decreased with the frequency of such events. This counter-intuitive finding can be explained by the fact that repeated dispersal events of small populations may be insufficient to establish growth in the recipient area, and collectively, can reduce the size of the source population at the introduction site. Despite that this explanation is plausible, there is no experimental support available in the literature. Long-distance dispersal events are inherently hard to measure, and there are only a few data sets that can be used to describe the tails of a long-distance dispersal kernel. Results from this study reaffirms there is a need for more data on movement patterns, particularly on long-distance movement, of invasive species in different habitat types and landscape contexts (Robledo-Arnuncio *et al.*, 2014).

7.6 Uncertainty in establishment and spread modelling

The modular dispersal framework MDiG was initially developed to support strategic forecasts of spatio-temporal invasive species distribution and management options (Chapter 2). When used for spread forecast purposes, the modelling framework integrates specific characteristics of the studied species, for example, specific dispersal strategies or habitat requirements, to obtain 'realism'. But the result of the simulations remain closely linked to the choice of the parameters rather than providing any insight into general principles. In this research, however, a more holistic approach was adopted where the pattern of invasion generated by multiple species scenarios within different landscape structure, was used to infer key drivers of population density and spread that are not taxon- or region-specific (Chapters 2, 3, 4 and 5). One key aspect of this assessment was to identify which parameter uncertainties were likely to have a large impact on population density and spread estimates. The analyses were built around a series of relatively simple assumptions regarding the characteristics of the species and the environment, such as a random walk for approximating local diffusion of the insects, a single Cauchy distribution for approximating long-distance dispersal events or a binary distinction of suitable and unsuitable habitat. Despite this simplicity, the model as a whole is structurally complex and produced a rich variety of plausible range expansion dynamics that remain to be tested empirically.

As discussed in Chapter 6, uncertainty is an inevitable component of any spread modelling exercise. Models typically carry substantial error margins due to structural (model specification) uncertainty and parameter (data) uncertainty as well as inherent (natural) stochasticity of ecological dynamics. These uncertainties need to be quantified to draw inferences about the robustness of model results and subsequent ecological hypotheses. In this research, I have largely side-stepped model structural analysis. Such uncertainty is likely to be high. For example, major differences regarding the effect of landscape structure on population density and spread were found when considering simulations with an Allee effect and those without an Allee effect (Chapters 2, 4 and 5). Some discrepancy between the results of this present study and those from Cassey *et al.* (2014) were also highlighted in Chapter 5. Using a competitive modelling framework (VORTEX, Lacy 2000) for simulating establishment success of exotic birds, Cassey *et al.* (2014) found evidence for a greater effect of growth-related traits than propagule pressure and dispersal-related traits, indicating consistency with this study. However, unlike the results in this research, Cassey *et al.* (2014) found no evidence of the influence of an Allee effect in their model. The authors suggested that environmental stochasticity, that can also lead to extinction of low-density populations, appears to matter more than Allee dynamics. Environmental stochasticity was not explicitly considered in this thesis. Whether explicitly integrating environmental stochasticity in the MDiG modelling framework will result in a decrease influence of an Allee effect or whether such inconsistency emerges because of others structural differences in the two models (MDiG and VORTEX) remains an open question.

Choosing the appropriate degree of abstraction of species demography, dispersal, propagule pressure and ecosystem characteristics to keep a balance between maintaining reality and reducing model complexity, is a fundamental challenge to establishment and spread modelling. It was suggested in Chapter 6 that methods such as pattern-oriented modelling (POM) can help to decide how structurally complex models such as MDiG need to be, to increase its predictive power. Thuiller *et al.* (2008) suggested that such a decision is scale-dependent. Complex models are likely to be more accurate at finer resolutions, whereas simple models are likely to offer useful and parsimonious solutions at broader scales. Yet, the development of complex models remain necessary not only to help understand the relative importance of different drivers and their interactions on the population density and spread of invasive species, but also as an aid to optimising the trade-off between precision and model complexity. Suggestions for future development of the MDiG modelling framework are given in Section 7.7.

7.7 Perspectives

Based on the recent review of research frontiers in species distribution modelling by Thuiller *et al.* (2008) and in plant dispersal by Robledo-Arnuncio *et al.* (2014), I discuss a set of research questions for which a modelling framework such as MDiG could be used to help progress our understanding of the establishment and spread of invasive species in heterogeneous landscapes.

7.7.1 What is the appropriate model resolution, extent or habitat complexity to use?

A challenge for establishment and dispersal modelling is to decide the appropriate extent, resolution and complexity of the study. The resolution and the extent over which dispersal studies are carried out can affect the accuracy of results. For example, the level of landscape detail (complexity) depends on the spatial resolution of the study, and subsequently affect estimation of dispersal rate and population density. A good example from this research was that recoding high-resolution data of an urban area to delimit unsuitable man-made structures was necessary to gain accuracy and specificity in the projections of *P. brassicae* in South Island, New Zealand (Chapter 2). Otherwise the model would have enhanced spread estimates by overestimating the dispersal rate in a uniform urban area generated from larger resolution data. Increasing the spatial resolution also increases the relative habitat amount and habitat patch isolation in the landscape (Chapter 3), which was found to have a significant effect on population density and spread rate (Chapters 4).

A modelling framework such as MDiG can be used to investigate how the relation between attributes of the landscape and population dynamics scale across spatial resolution and extent. While selective re-coding of certain areas of the landscape based on species attributes as demonstrated in Chapter 2 could greatly benefit studies case by case, a generalized and exhaustive study that could elucidate a possible relationship between species attributes and mode of dispersal with optimum landscape resolution, configuration and composition is greatly needed. An example of such procedure can be found in Skelsey & Garrett (2013). As there is no one optimum spatial extent and resolution that can be used for all cases, each of these parameters have to be determined based on the characteristics of the species studied (Senay, 2014).

7.7.2 What is the contribution of different dispersal vectors to rate of spread?

It is acknowledged that for many plants, insects and other invertebrate species, multiple vectors contribute to long-distance dispersal (Bowler & Benton, 2005). For example, natural dispersal of the European gypsy moth *L. dispar*, used as case study, is primarily by the wind-borne dispersal of the first instar larvae and adult flight by males (CABI database - `http://www.cabi.org/`). A number of anthropogenic-driven dispersal vectors have also been reported including clothing/footwear and possessions, land vehicles, wood containers and packaging and plant trade (CABI database). These different vectors may fluctuate over different spatial and temporal scales and may respond differently to environmental change (Bowler & Benton, 2005). As discussed in Chapter 2, MDiG currently does not explicitly allow the contribution of multiple vectors to particular dispersal pathways to be simulated. While a mechanistic approach that simulates long-distance dispersal along a transport network is possible, it is usually very difficult to find such detailed information. Alternatively, dispersal kernels characterised by mixed probability distributions could be used to investigate the contribution of different dispersal vectors on population density and spread of invasive species (Gilbert *et al.*, 2004).

The main difficulty accounting for different dispersal dynamics in MDiG is that multiple temporal and spatial scales are forced to use the same spatial resolution. While this results in efficient alogrithms, it can potentially impact simulation outputs (Pitt, 2008). Methods for encoding raster data, such as the *quadtree* data structure used in Cole & Albrecht (1999), can be used to improve data storage requirements and access speeds. In the Cole & Albrecht (1999) study, the authors encoded long distance dispersal on a large-scale coarse resolution grid, with fine grids within each coarse grid cell to model more complex interactions and finer-scale processes when necessary. As such, MDiG could be used to identify the critical vectors of long-distance dispersal with disproportionate impact on population density and rate of spread of invasive species. The model can also be used to better understand how the interplay between long-distance dispersal and dynamic heterogeneous environments determines the spread of invasive species under climate and land-use change scenarios.

7.7.3 What is the implication of landscape-dependent variation on demography and dispersal ability?

Dispersal, as integrated in MDiG, is regarded as an unconditional process, with a unique Poisson distribution indicating the numbers of the population dispersing at each generation from each occupied raster cells. However, dispersal processes are clearly dependent on a range of factors such as habitat quality, density-dependence or habitat patch size (Clobert *et al.*, 2009, 2012). Gathering empirical information on such factors is an essential step that will contribute to the construction of more realistic assumptions regarding different dispersal processes. The MDiG framework already integrates a suitability map in the form of survival probability map ranging from 0-1 reflecting the difficulty that populations have establishing in each raster cell. An additional layer could be used in the form of dispersal probability map ranging from 0-1 reflecting the difficulty that populations have to disperse in each raster cell. As such, MDiG could be used to investigate the key environmental factors that have a disproportionate impact on the population density and/or spread of an invasive species.

7.7.4 What are the implications of demographic and environmental stochasticity on rate of spread?

A further step toward the realism representing dispersal and demographic processes in MDiG would be to account for demographic and environmental stochasticity. Four types of stochastic effects are generally recognised. First, demographic stochasticity arises from intra-specific variation in individual probability of reproduction and death. The combined influence of an Allee effect and demographic stochasticity contribute to population extinction when populations are at very low densities, therefore strongly influencing the successful establishment of invasive species (Liebhold & Bascompte, 2003). Second, stochastic genetic differences between source populations may lead to lag-time in invasion that are independent of population density and landscape structure. Third, empirical evidence suggests that intra-specific variation in dispersal ability (dispersal stochasticity), that is associated with other key life history traits such as morphological or behavioural traits (Clobert *et al.*, 2009, 2012), can have consequences on both spatial dynamics and dispersal evolution (Baguette *et al.*, 2012, 2014). Fourth, environmental stochasticity arises from random spatial variations that similarly affect birth, death and dispersal rates of all individuals in the populations.

Technically, it is not complicated to see how MDiG could simulate the reproduction, dispersal and death event for every single organism in the landscape to account for intraspecific variability in demographic, genetic and dispersal ability, yet some costs would be incurred with respect to uncertainty and interpretability. On the other hand, accounting for environmental stochasticity simply requires including the effects of additive environmental noise that affects the population density and dispersal events in each raster cell (see, for example, Liebhold & Bascompte, 2003). As such, MDiG could be used to investigate how the relative role of environmental stochasticity and Allee dynamics may influence population and spread of invasive species in heterogeneous environments.

7.7.5 What is the implication of temporal variation in demographic and dispersal ability on species spread?

Dispersal and demographic processes not only vary over space and among individuals of the same population (see Section 7.7.3 and 7.7.4), but also over time. Temporal variation in endogenous factors, such as life-stages, and exogenous factors such as seasonal variation in climate, influence reproduction, death and dispersal rates (see for example Robledo-Arnuncio *et al.*, 2014). Temporal characterization of propagule dispersal, which examines variation in fecundity, death, dispersal rates and their association with local environment characteristics, can be an enormous task. In its simple form, MdiG already integrates different dispersal modes that can be assigned to each life-stage. Phenological maps could also be used to indicate the timing of propagule recruitment and life stage associated dispersal. A good starting point is the model to predict the phenology of Gypsy moth, *L. dispar* (Régnière & Sharov, 1999), in which maps of life stage timing in New Zealand were used to time the occurrence of peak adult moth flight, when individuals are actively dispersing (Pitt *et al.*, 2007).

In addition, the suitability of a region does not remain constant from year to year due to gradual change in land-use or rapid changes in temperature, for example. The possibility of using stochastic potential suitability that changes in value through time (Moilanen & Cabeza, 2002) may mean spread models could allow for such variation. However, in my opinion, understanding variation in demography and dispersal and their consequences on population dynamics in heterogeneous landscapes will probably require conceptual and methodological advances to establish a clear partition into temporal, spatial and environmental components over different spatial and temporal scales.

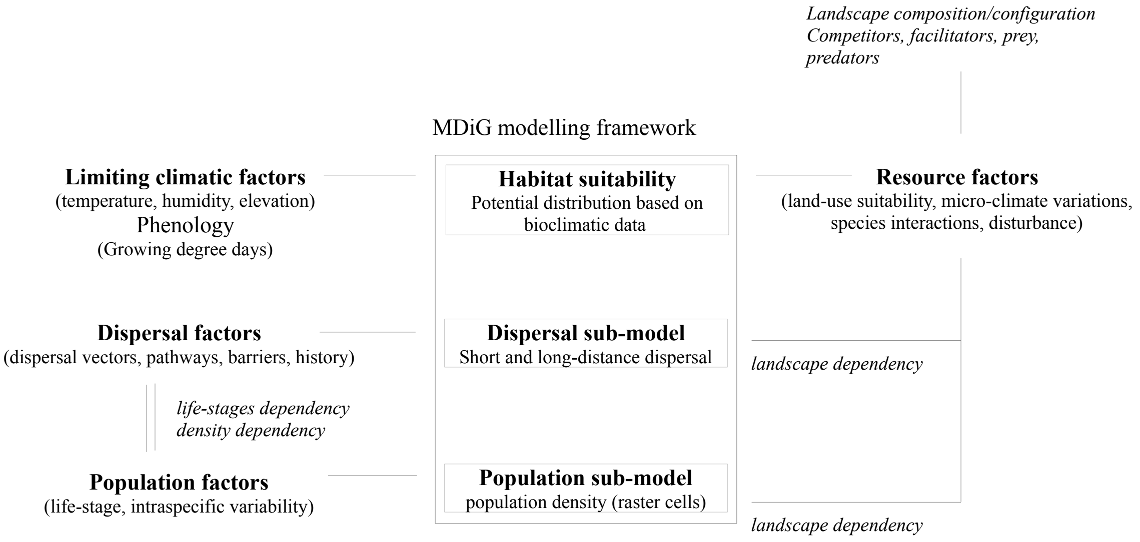

Figure 7.1: General modular dispersal framework, MDiG.

7.8 Concluding remarks

Advances modelling the establishment and spread of invasive species will be determined by the ability to surmount the challenges of spatial scale and heterogeneity, temporal variation, and system complexity (Robledo-Arnuncio *et al.*, 2014). New inferential and predictive methods, such as the MDiG modelling framework, will be applied to better represent population growth, density, rate of spread and trajectories of invasive species over different taxonomic groups and spatial scales, in environmentally and demographic explicit contexts. Such approaches rely on a judicious combination of techniques that provide information useful for decision-makers to explore often uncontrollable, irreducible uncertainty about the future. Such information offers resource managers, a method for developing more resilient conservation or management strategies, especially at the landscape level, which can integrate key climate change, land use change, invasive species ecology and landscape-ecological linkages.

Although we are constantly making incremental progress to understand, characterize, and forecast the invasion process in heterogeneous environment, it is clear that uncertainty, is here to stay no matter what information is used to make decisions (Chapter 6). Criticism has been raised concerning the usefulness of such complex models for guiding effective management of invasive species at a regional scale (Robinet *et al.*, 2012b; Venette, 2015). However, models are often most useful for challenging what we think we understand. Investigating the use of different modelling approaches (different model structures and parameters) and the impact on projections performance could help identify the optimal trade-offs between precision and complexity. Such development is in its infancy, and further research to correctly and consistently assess as well as communicate uncertainty around spread modelling is needed, so that end-users can make more informed decisions about the potential for invasive species establishment and spread into new areas.

References

Adler, F. R., & Nuernberger, B. (1994). Persistence in patchy irregular landscapes. *Theoretical Population Biology*, *45*(1), 41–75.

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (tss). *Journal of Applied Ecology*, *43*(6), 1223–1232.

Almeida, R. C., Delphim, S. A., & Costa, M. I. d. S. (2006). A numerical model to solve single-species invasion problems with Allee effects. *Ecological Modelling*, *192*(3), 601–617.

Anderson, D. R. (2008). *Model based inference in the life sciences: a primer on evidence*. Springer Science and Business Media. 184 pp.

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, *26*(1), 32–46.

Andrew, M. E., & Ustin, S. L. (2010). The effects of temporally variable dispersal and landscape structure on invasive species spread. *Ecological Applications*, *20*(3), 593–608.

Araújo, M. B., & Guisan, A. (2006). Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, *33*(10), 1677–1688.

Araújo, M. B., & New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, *22*(1), 42–47.

Araújo, M. B., Pearson, R. G., Thuiller, W., & Erhard, M. (2005). Validation of species–climate impact models under climate change. *Global Change Biology*, *11*(9), 1504–1513.

Ascough, J. C., Maier, H. R., Ravalico, J. K., & Strudley, M. W. (2008). Future research challenges for incorporation of uncertainty in environmental and ecological decision-making. *Ecological Modelling*, *219*(3), 383–399.

Baguette, M., Legrand, D., Fréville, H., Van Dyck, H., & Ducatez, S. (2012). Evolutionary ecology of dispersal in fragmented landscape. (pp. 381–391).

Baguette, M., Stevens, V. M., & Clobert, J. (2014). The pros and cons of applying the movement ecology paradigm for studying animal dispersal. *Movement Ecology*, *2*(1), 13.

Baguette, M., & Van Dyck, H. (2007). Landscape connectivity and animal behavior: functional grain as a key determinant for dispersal. *Landscape Ecology*, *22*(8), 1117–1129.

Bailey, D., Billeter, R., Aviron, S., Schweiger, O., & Herzog, F. (2007). The influence of thematic resolution on metric selection for biodiversity monitoring in agricultural landscapes. *Landscape Ecology*, *22*(3), 461–473.

Baker, W. L., & Cai, Y. (1992). The r.le programs for multiscale analysis of landscape structure using the GRASS geographical information system. *Landscape Ecology*, *7*(4), 291–302.

Banks, N. C., Paini, D. R., Bayliss, K. L., & Hodda, M. (2015). The role of global trade and transport network topology in the human-mediated dispersal of alien species. *Ecology Letters*, *18*(2), 188–199.

Barbier, E. B., & Shogren, J. F. (2004). Growth with endogenous risk of biological invasion. *Economic Inquiry*, *42*(4), 587–601.

Barlow, N., & Dixon, A. F. G. (1980). *Simulation of lime aphid population dynamics.*. Centre for Agricultural Publishing and Documentation. 165 pp.

Barney, J. N., & Whitlow, T. H. (2008). A unifying framework for biological invasions: the state factor model. *Biological Invasions*, *10*(3), 259–272.

Barry, S., & Elith, J. (2006). Error and uncertainty in habitat models. *Journal of Applied Ecology*, *43*(3), 413–423.

Bartoń, K. (2016). MuMIn: multi-model inference, version 1.15.6. *R-Cran package*. `https://cran.r-project.org/web/packages/MuMIn/index.html`.

Bartuszevige, A. M., & Gorchov, D. L. (2006). Avian seed dispersal of an invasive shrub. *Biological Invasions*, *8*(5), 1013–1022.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., & Eigen adn Grothendieck, G. (2015). lme4: Linear mixed-effects models using 'eigen' and s4, version 1.1-10. *R-Cran package*. `https://cran.r-project.org/web/packages/lme4/index.html`.

Beale, C. M., & Lennon, J. J. (2012). Incorporating uncertainty in predictive species distribution modelling. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1586), 247–258.

Bellard, C., Cassey, P., & Blackburn, T. M. (2016). Alien species as a driver of recent extinctions. *Biology letters*, *12*(2), 20150623.

Ben-Haim, Y. (2006). *Info-gap decision theory: decisions under severe uncertainty*. Academic Press. 368 pp.

Benke, K. K., Steel, J. L., & Weiss, J. E. (2011). Risk assessment models for invasive species: uncertainty in rankings from multi-criteria analysis. *Biological Invasions*, *13*(1), 239–253.

Berestycki, H., Diekmann, O., Nagelkerke, C., & Zegeling, P. (2009). Can a species keep pace with a shifting climate? *Bulletin of Mathematical Biology*, *71*(2), 399–429.

Berestycki, H., Hamel, F., & Roques, L. (2005). Analysis of the periodically fragmented environment model: Ii—biological invasions and pulsating travelling fronts. *Journal de Mathématiques purés et appliquées*, *84*(8), 1101–1146.

Betts, M. G., Fahrig, L., Hadley, A. S., Halstead, K. E., Bowman, J., Robinson, W. D., Wiens, J. A., & Lindenmayer, D. B. (2014). A species-centered approach for uncovering generalities in organism responses to habitat loss and fragmentation. *Ecography*, *37*(6), 517–527.

Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology*, *320*(1), 18–36.

Beven, K., & Binley, A. (1992). The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, *6*(3), 279–298.

Beverton, R. J. H., & Holt, S. J. (1957). On the dynamics of exploited fish populations. *1*(2). UK Ministry of Agriculture. Fisheries, and Food, London, UK.

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*(3), 451–460.

Blackburn, T. M., Pyšek, P., Bacher, S., Carlton, J. T., Duncan, R. P., Jarošík, V., Wilson, J. R., & Richardson, D. M. (2011). A proposed unified framework for biological invasions. *Trends in Ecology and Evolution*, *26*(7), 333–339.

Blackwood, J., Hastings, A., & Costello, C. (2010). Cost-effective management of invasive species using linear-quadratic control. *Ecological Economics*, *69*(3), 519–527.

Blunck, H. (1954). Beobachtungen über Wanderflüge von *Pieris brassicae L. Beiträge zur Entomologie*, *4*, 485–528.

Bobashev, G. V., & Morris, R. J. (2010). Uncertainty and inference in agent-based models. In *Second International Conference on Advances in System Simulation (SIMUL)*, (pp. 67–71). IEEE.

Bocedi, G., Palmer, S. C., Pe'er, G., Heikkinen, R. K., Matsinos, Y. G., Watts, K., & Travis, J. M. (2014). RangeShifter: a platform for modelling spatial eco-evolutionary dynamics and species' responses to environmental changes. *Methods in Ecology and Evolution*, *5*(4), 388–396.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J.-S. S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology and Evolution*, *24*(3), 127–135.

Bowler, D. E., & Benton, T. G. (2005). Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. *Biological Reviews*, *80*(02), 205–225.

Bradley, B. A. (2010). Assessing ecosystem threats from global and regional change: hierarchical modeling of risk to sagebrush ecosystems from climate change, land use and invasive species in Nevada, USA. *Ecography*, *33*(1), 198–208.

Brook, B. W. (2004). Australasian bird invasions: accidents of history? *Ornithological Science*, *3*(1), 33–42.

Brown, G. P., Phillips, B. L., Webb, J. K., & Shine, R. (2006). Toad on the road: use of roads as dispersal corridors by cane toads (*Bufo marinus*) at an invasion front in tropical Australia. *Biological Conservation*, *133*(1), 88–94.

Brown, J. D. (2010). Prospects for the open treatment of uncertainty in environmental research. *Progress in Physical Geography*, *34*(1), 75–100.

Brown, J. L. (2014). SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *Methods in Ecology and Evolution*, *5*(7), 694–700.

Brown, R., James, A., Pitchford, J., & Plank, M. (2012). Habitat fragmentation: Simple models for local persistence and the spread of invasive species. *Journal of Theoretical Biology*, *310*, 231–238.

Buckley, Y. M., Brockerhoff, E., Langer, L., Ledgard, N., North, H., & Rees, M. (2005). Slowing down a pine invasion despite uncertainty in demography and dispersal. *Journal of Applied Ecology*, *42*(6), 1020–1030.

Buisson, L., Thuiller, W., Casajus, N., Lek, S., & Grenouillet, G. (2010). Uncertainty in ensemble forecasting of species distribution. *Global Change Biology*, *16*(4), 1145–1157.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Science and Business Media. 488 pp.

Burnham, K. P., Anderson, D. R., & Huyvaert, K. P. (2011). AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral Ecology and Sociobiology*, *65*(1), 23–35.

Buyantuyev, A., & Wu, J. (2007). Effects of thematic resolution on landscape pattern analysis. *Landscape Ecology*, *22*(1), 7–13.

Cacho, O. J., Spring, D., Hester, S., & Mac Nally, R. (2010). Allocating surveillance effort in the management of invasive species: a spatially-explicit model. *Environmental Modelling and Software*, *25*(4), 444–454.

Cagnacci, F., Boitani, L., Powell, R. A., & Boyce, M. S. (2010). Animal ecology meets gps-based radiotelemetry: a perfect storm of opportunities and challenges. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1550), 2157–2162.

Cain, M. L., Milligan, B. G., & Strand, A. E. (2000). Long-distance seed dispersal in plant populations. *American Journal of Botany*, *87*(9), 1217–1227.

Calabrese, J. M., & Fagan, W. F. (2004). Lost in time, lonely, and single: reproductive asynchrony and the Allee effect. *The American Naturalist*, *164*(1), 25–37.

Caley, P., Groves, R. H., & Barker, R. (2008). Estimating the invasion success of introduced plants. *Diversity and Distributions*, *14*(2), 196–203.

Cannas, S. A., Marco, D. E., & Páez, S. A. (2003). Modelling biological invasions: species traits, species interactions, and habitat heterogeneity. *Mathematical Biosciences*, *183*(1), 93–110.

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., Narwani, A., Mace, G. M., Tilman, D., & Wardle, D. A. (2012). Biodiversity loss and its impact on humanity. *Nature*, *486*(7401), 59–67.

Carpenter, S. R., De Fries, R., Dietz, T., Mooney, H. A., Polasky, S., Reid, W. V., & Scholes, R. J. (2006). Millennium ecosystem assessment: research needs. Tech. rep., American Association for the Advancement of Science, USA.

Carrasco, L. R., Baker, R., MacLeod, A., Knight, J. D., & Mumford, J. D. (2010). Optimal and robust control of invasive alien species spreading in homogeneous landscapes. *Journal of the Royal Society Interface*, *7*(44), 529–540.

Carrasco, L. R., Cook, D., Baker, R., MacLeod, A., Knight, J. D., & Mumford, J. D. (2012). Towards the integration of spread and economic impacts of biological invasions in a landscape of learning and imitating agents. *Ecological Economics*, *76*, 95–103.

Cassey, P., Prowse, T. A., & Blackburn, T. M. (2014). A population model for predicting the successful establishment of introduced bird species. *Oecologia*, *175*(1), 417–428.

Castilla, G., Larkin, K., Linke, J., & Hay, G. J. (2009). The impact of thematic resolution on the patch-mosaic model of natural landscapes. *Landscape Ecology*, *24*(1), 15–23.

Caswell, H. (2001). *Matrix population models: construction, analysis, and interpretation*. Wiley Online Library. 722 pp.

Caswell, H., Neubert, M. G., & Hunter, C. M. (2011). Demography and dispersal: invasion speeds and sensitivity analysis in periodic and stochastic environments. *Theoretical Ecology*, *4*(4), 407–421.

Catford, J. A., Jansson, R., & Nilsson, C. (2009). Reducing redundancy in invasion ecology by integrating hypotheses into a single theoretical framework. *Diversity and Distributions*, *15*(1), 22–40.

Catford, J. A., Vesk, P. A., White, M. D., & Wintle, B. A. (2011). Hotspots of plant invasion predicted by propagule pressure and ecosystem characteristics. *Diversity and Distributions*, *17*(6), 1099–1110.

Chadwell, T. B., & Engelhardt, K. A. (2008). Effects of pre-existing submersed vegetation and propagule pressure on the invasion success of *Hydrilla verticillata*. *Journal of Applied Ecology*, *45*(2), 515–523.

Chaplin-Kramer, R., O'Rourke, M. E., Blitzer, E. J., & Kremen, C. (2011). A meta-analysis of crop pest and natural enemy response to landscape complexity. *Ecology Letters*, *14*(9), 922–932.

Chipperfield, J. D., Dytham, C., & Hovestadt, T. (2011). An updated algorithm for the generation of neutral landscapes by spectral synthesis. *PloS One*, *6*(2), e17040.

Christen, D., & Matlack, G. (2006). The role of roadsides in plant invasions: a demographic approach. *Conservation Biology*, *20*(2), 385–391.

Chytrỳ, M., Maskell, L. C., Pino, J., Pyšek, P., Vilà, M., Font, X., & Smart, S. M. (2008). Habitat invasions by alien plants: a quantitative comparison among mediterranean, subcontinental and oceanic regions of europe. *Journal of Applied Ecology*, *45*(2), 448–458.

Clark, J. S. (2005). Why environmental scientists are becoming Bayesians. *Ecology Letters*, *8*(1), 2–14.

Clark, J. S., Carpenter, S. R., Barber, M., Collins, S., Dobson, A., Foley, J. A., Lodge, D. M., Pascual, M., Pielke Jr, R., & Pizer, W. (2001). Ecological forecasts: an emerging imperative. *Science*, *293*(5530), 657–660.

Clark, J. S., & Gelfand, A. E. (2006). A future for models and data in environmental science. *Trends in Ecology and Evolution*, *21*(7), 375–380.

Clobert, J., Baguette, M., Benton, T. G., Bullock, J. M., & Ducatez, S. (2012). *Dispersal Ecology and Evolution*. Oxford University Press. 462 pp.

Clobert, J., Galliard, L., Cote, J., Meylan, S., & Massot, M. (2009). Informed dispersal, heterogeneity in animal dispersal syndromes and the dynamics of spatially structured populations. *Ecology letters*, *12*(3), 197–209.

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

Colautti, R. I., Grigorovich, I. A., & MacIsaac, H. J. (2006). Propagule pressure: a null model for biological invasions. *Biological Invasions*, *8*(5), 1023–1037.

Colautti, R. I., Ricciardi, A., Grigorovich, I. A., & MacIsaac, H. J. (2004). Is invasion success explained by the enemy release hypothesis? *Ecology Letters*, *7*(8), 721–733.

Cole, V., & Albrecht, J. (1999). Exploring geographic parameter space with a gis implementation of cellular automata. In *The 11th Annual Colloquium of the Spatial Information Research Centre*. University of Otago, Dunedin, New Zealand, Citeseer.

Collins, M. (2007). Ensembles and probabilities: a new era in the prediction of climate change. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *365*(1857), 1957–1970.

Congalton, R. G., & Green, K. (1993). A practical look at the sources of confusion in error matrix generation. *Photogrammetric Engineering and Remote Sensing*, *59*(5), 641–644.

Conn, A., Gould, N., & Toint, P. L. (2000). *Trust-Region Methods, Society for Industrial and Applied Mathematics*, vol. 1. Society for Industrial and Applied Mathematics, SIAM, Philadelphia. 959 pp.

Cook, A., Marion, G., Butler, A., & Gibson, G. (2007a). Bayesian inference for the spatio-temporal invasion of alien species. *Bulletin of mathematical biology*, *69*(6), 2005–2025.

Cook, D., Liu, S., Aurambout, J.-P., Villalta, O., Kriticos, D. J., Edwards, J., Diggle, A., Siddique, A.-B., Hurley, M., & Lowell, K. (2012). Communicating xxuncertainty in biosecurity adaption. Final report CRC10162, Cooperative Research Centre for National Plant Biosecurity, Australia.

Cook, D. C., Thomas, M. B., Cunningham, S. A., Anderson, D. L., & De Barro, P. J. (2007b). Predicting the economic impact of an invasive species on an ecosystem service. *Ecological Applications*, *17*(6), 1832–1840.

Cornélis, B., & Brunet, S. (2002). A policy-maker point of view on uncertainties in spatial decisions. In W. Shi, P. Fisher, & M. F. Goodchild (Eds.) *Spatial data quality*, chap. 12, (pp. 168–185). Taylor and Francis, London, UK.

Couclelis, H. (2003). The certainty of uncertainty: GIS and the limits of geographic knowledge. *Transactions in GIS*, *7*(2), 165–175.

Coutts, S. R., van Klinken, R. D., Yokomizo, H., & Buckley, Y. M. (2011). What are the key drivers of spread in invasive plants: dispersal, demography or landscape: and how can we use this knowledge to aid management? *Biological Invasions*, *13*(7), 1649–1661.

Cozman, F. G. (2000). Credal networks. *Artificial Intelligence*, *120*(2), 199–233.

Crespo-Pérez, V., Rebaudo, F., Silvain, J.-F., & Dangles, O. (2011). Modeling invasive species spread in complex landscapes: the case of potato moth in ecuador. *Landscape Ecology*, *26*(10), 1447–1461.

Crosetto, M., Tarantola, S., & Saltelli, A. (2000). Sensitivity and uncertainty analysis in spatial modelling based on gis. *Agriculture, Ecosystems and Environment*, *81*(1), 71–79.

Cushman, S. A., McGarigal, K., & Neel, M. C. (2008). Parsimony in landscape metrics: strength, universality, and consistency. *Ecological Indicators*, *8*(5), 691–703.

Davies, C., & Gilbert, N. (1985). A comparative study of the egg-laying behaviour and larval development of *Pieris rapae L.* and *P. brassicae L.* on the same host plants. *Oecologia*, *67*(2), 278–281.

Day, J. R., & Possingham, H. P. (1995). A stochastic metapopulation model with variability in patch size and position. *Theoretical Population Biology*, *48*(3), 333–360.

De Poorter, M., Pagad, S., & Ullah, I. M. (2007). Invasive alien species and protected areas - Part I: scoping the scale and nature of invasive alien species threats to protected areas, impediments to IAS management and means to address those impediments. Tech. rep., World Bank Report, The global Invasive Species Programme (GISP).

Delatte, H., Duyck, P.-F., Triboire, A., David, P., Becker, N., Bonato, O., & Reynaud, B. (2009). Differential invasion success among biotypes: case of *Bemisia tabaci*. *Biological Invasions*, *11*(4), 1059–1070.

Demon, I., Cunniffe, N., Marchant, B., Gilligan, C., & van den Bosch, F. (2011). Spatial sampling to detect an invasive pathogen outside of an eradication zone. *Phytopathology*, *101*(6), 725–731.

Dewhirst, S., & Lutscher, F. (2009). Dispersal in heterogeneous habitats: thresholds, spatial scales, and approximate rates of spread. *Ecology*, *90*(5), 1338–1345.

Dewulf, A., Craps, M., Bouwen, R., Taillieu, T., & Pahl-Wostl, C. (2005). Integrated management of natural resources: dealing with ambiguous issues, multiple actors and diverging frames. *Water Science and Technology*, *52*(6), 115–124.

Didham, R. K., Tylianakis, J. M., Gemmell, N. J., Rand, T. A., & Ewers, R. M. (2007). Interactive effects of habitat modification and species invasion on native species decline. *Trends in Ecology and Evolution*, *22*(9), 489–496.

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 and Evolution*, *20*(9), 470–474.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., & Leitão, P. J. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 27–46.

Dormann, C. F., Schymanski, S. J., Cabral, J., Chuine, I., Graham, C., Hartig, F., Kearney, M., Morin, X., Römermann, C., & Schröder, B. (2012). Correlation and process in species distribution models: bridging a dichotomy. *Journal of Biogeography*, *39*(12), 2119–2131.

Drescher, M., Perera, A. H., Johnson, C. J., Buse, L. J., Drew, C. A., & Burgman, M. A. (2013). Toward rigorous use of expert knowledge in ecological research. *Ecosphere*, *4*(7), art83.

Dwyer, G., & Elkinton, J. S. (1993). Using simple models to predict virus epizootics in gypsy moth populations. *Journal of Animal Ecology*, *62*(1), 1–11.

Eiswerth, M. E., & Van Kooten, G. C. (2002). Uncertainty, economics, and the spread of an invasive plant species. *American Journal of Agricultural Economics*, (pp. 1317–1322).

Elith, J., Burgman, M. A., & Regan, H. M. (2002). Mapping epistemic uncertainties and vague concepts in predictions of species distribution. *Ecological Modelling*, *157*(2), 313–329.

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*(1), 677.

Elith, J., Leathwick, J. R., & Hastie, T. (2009). A working guide to boosted regression trees. *Journal of Animal Ecology*, *77*(4), 802–813.

Elton, C. S. (2000). *The ecology of invasions by animals and plants*. University of Chicago Press. 196 pp.

Epanchin-Niell, R. S., Haight, R. G., Berec, L., Kean, J. M., & Liebhold, A. M. (2012). Optimal surveillance and eradication of invasive species in heterogeneous landscapes. *Ecology Letters*, *15*(8), 803–812.

Estreguil, C., De Rigo, D., & Caudullo, G. (2014). A proposal for an integrated modelling framework to characterise habitat pattern. *Environmental Modelling & Software*, *52*, 176–191.

Etherington, T. R., Holland, E. P., & O'Sullivan, D. (2014). NLMpy: a python software package for the creation of neutral landscape models within a general numerical framework. *Methods in Ecology and Evolution*, *6*(2), 164–168.

Evans, A. (2011). Uncertainty and Error. In A. J. Heppenstall, A. T. Crooks, L. M. See, & M. Batty (Eds.) *Agent-based models of geographical systems*, (p. 760 pp.). Springer Science and Business Media.

Ewers, R. M., & Didham, R. K. (2006). Confounding factors in the detection of species responses to habitat fragmentation. *Biological Reviews*, *81*(01), 117–142.

Ewers, R. M., Thorpe, S., & Didham, R. K. (2007). Synergistic interactions between edge and area effects in a heavily fragmented landscape. *Ecology*, *88*(1), 96–106.

Facon, B., & David, P. (2006). Metapopulation dynamics and biological invasions: a spatially explicit model applied to a freshwater snail. *The American Naturalist*, *168*(6), 769–783.

Facon, B., Genton, B. J., Shykoff, J., Jarne, P., Estoup, A., & David, P. (2006). A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology and Evolution*, *21*(3), 130–135.

Fagan, W. F., Lewis, M. A., Auger-Méthé, M., Avgar, T., Benhamou, S., Breed, G., LaDage, L., Schlägel, U. E., Tang, W.-w., Papastamatiou, Y. P., *et al.* (2013). Spatial memory and animal movement. *Ecology letters*, *16*(10), 1316–1329.

Fahrig, L. (2007). Non-optimal animal movement in human-altered landscapes. *Functional Ecology*, *21*(6), 1003–1015.

Fahrig, L., Baudry, J., Brotons, L., Burel, F. G., Crist, T. O., Fuller, R. J., Sirami, C., Siriwardena, G. M., & Martin, J.-L. (2011). Functional landscape heterogeneity and animal biodiversity in agricultural landscapes. *Ecology Letters*, *14*(2), 101–112.

Fahrig, L., & Nuttle, W. K. (2005). Population ecology in spatially heterogeneous environments. In G. Lovett, C. Jones, M. Turner, & K. Weathers (Eds.) *Ecosystem function in heterogeneous landscapes*, chap. 6, (pp. 95–118). Springer.

Fassò, A., & Perri, P. F. (2006). Sensitivity analysis.

Fekete, A., Damm, M., & Birkmann, J. (2010). Scales as a challenge for vulnerability assessment. *Natural Hazards*, *55*(3), 729–747.

Feltwell, J. (1982). *Large White butterfly: the biology, biochemistry, and physiology of Pieris brassicae (Linnaeus)*, vol. 18. Springer Science and Business Media. 564 pp.

Ferrari, J. R., & Lookingbill, T. R. (2009). Initial conditions and their effect on invasion velocity across heterogeneous landscapes. *Biological Invasions*, *11*(6), 1247–1258.

Ferrari, J. R., Preisser, E. L., & Fitzpatrick, M. C. (2014). Modeling the spread of invasive species using dynamic network models. *Biological invasions*, *16*(4), 949–960.

Ferraz, G., Nichols, J. D., Hines, J. E., Stouffer, P. C., Bierregaard, R. O., & Lovejoy, T. E. (2007). A large-scale deforestation experiment: effects of patch area and isolation on amazon birds. *Science*, *315*(5809), 238–241.

Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, *24*(01), 38–49.

Fisher, P. F. (1999). Models of uncertainty in spatial data. *Geographical Information Systems*, *1*, 191–205.

Fisher, R. A. (1937). The wave of advance of advantageous genes. *Annals of Eugenics*, *7*(4), 355–369.

Fitzpatrick, M. C., Preisser, E. L., Porter, A., Elkinton, J., & Ellison, A. M. (2012). Modeling range dynamics in heterogeneous landscapes: invasion of the hemlock woolly adelgid in eastern North America. *Ecological Applications*, *22*(2), 472–486.

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, F. S., Coe, M. T., Daily, G. C., & Gibbs, H. K. (2005). Global consequences of land use. *Science*, *309*(5734), 570–574.

Fox, P., & Hendler, J. (2011). Changing the equation on scientific data visualization. *Science*, *331*(6018), 705–708.

Franklin, J. (2010). Moving beyond static species distribution models in support of conservation biogeography. *Diversity and Distributions*, *16*(3), 321–330.

Freeman, E. A., & Moisen, G. G. (2008). A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecological Modelling*, *217*(1), 48–58.

Friedman, J. H., & Popescu, B. E. (2008). Predictive learning via rule ensembles. *The Annals of Applied Statistics*, *2*(3), 916–954.

Frohn, R. C., & Hao, Y. (2006). Landscape metric performance in analyzing two decades of deforestation in the Amazon basin of Rondonia, Brazil. *Remote Sensing of Environment*, *100*(2), 237–251.

Funk, J. L. (2015). Invasive species: A global problem in need of a global solution. *BioScience*, *65*(6), biv053.

Gallien, L., Mazel, F., Lavergne, S., Renaud, J., Douzet, R., & Thuiller, W. (2014). Contrasting the effects of environment, dispersal and biotic interactions to explain the distribution of invasive plants in alpine communities. *Biological Invasions*, *17*(5), 1407–1423.

Garden, J., Mcalpine, C., Peterson, A., Jones, D., & Possingham, H. (2006). Review of the ecology of australian urban fauna: a focus on spatially explicit processes. *Austral Ecology*, *31*(2), 126–148.

Gardner, R. H. (1999). RULE: map generation and a spatial analysis program. In *Landscape ecological analysis*, (pp. 280–303). Springer.

Gardner, R. H., Milne, B. T., Turnei, M. G., & O'Neill, R. V. (1987). Neutral models for the analysis of broad-scale landscape pattern. *Landscape Ecology*, *1*(1), 19–28.

Gardner, R. H., & Urban, D. L. (2007). Neutral models for testing landscape hypotheses. *Landscape Ecology*, *22*(1), 15–29.

Garnier, J., Roques, L., & Hamel, F. (2012). Success rate of a biological invasion in terms of the spatial distribution of the founding population. *Bulletin of Mathematical Biology*, *74*(2), 453–473.

Gertner, G. (1987). Approximating precision in simulation projections: An efficient alternative to Monte Carlo methods. *Forest Science*, *33*(1), 230–239.

Getis, A. (2007). Reflections on spatial autocorrelation. *Regional Science and Urban Economics*, *37*(4), 491–496.

Giera, N., & Bell, B. (2009). Economic costs of pests to new zealand. Tech. Rep. 2009/31, MAF Biosecurity New Zealand Technical Paper. 84 pp.

Gilbert, M., Grégoire, J.-C., Freise, J. F., & Heitland, W. (2004). Long-distance dispersal and human population density allow the prediction of invasive patterns in the horse chestnut *leafminer Cameraria ohridella*. *Journal of Animal Ecology*, *73*(3), 459–468.

Glen, A. S., Pech, R. P., & Byrom, A. E. (2013). Connectivity and invasive species management: towards an integrated landscape approach. *Biological Invasions*, *15*(10), 2127–2138.

Gneiting, T., & Raftery, A. E. (2005). Weather forecasting with ensemble methods. *Science*, *310*(5746), 248–249.

González-Moreno, P., Diez, J. M., Ibáñez, I., Font, X., & Vilà, M. (2014). Plant invasions are context-dependent: multiscale effects of climate, human activity and habitat. *Diversity and Distributions*, *20*(6), 720–731.

González-Moreno, P., Pino, J., Carreras, D., Basnou, C., Fernández-Rebollar, I., & Vilá, M. (2013a). Quantifying the landscape influence on plant invasions in mediterranean coastal habitats. *Landscape Ecology*, *28*(5), 891–903.

González-Moreno, P., Pino, J., Gassó, N., & Vilá, M. (2013b). Landscape context modulates alien plant invasion in mediterranean forest edges. *Biological Invasions*, *15*(3), 547–557.

Gould, S. F., Beeton, N. J., Harris, R., Hutchinson, M. F., Lechner, A. M., Porfirio, L. L., & Mackey, B. G. (2014). A tool for simulating and communicating uncertainty when modelling species distributions under future climates. *Ecology and Evolution*, *4*(24), 4798–4811.

Grassberger, P. (1983). On the critical behavior of the general epidemic process and dynamical percolation. *Mathematical Biosciences*, *63*(2), 157–172.

Grevstad, F. S. (1999). Experimental invasions using biological control introductions: the influence of release size on the chance of population establishment. *Biological Invasions*, *1*(4), 313–323.

Grimm, V. (2002). Visual debugging: a way of analyzing, understanding and communicating bottom-up simulation models in ecology. *Natural Resource Modeling*, *15*(1), 23–38.

Grimm, V., & Railsback, S. F. (2012a). Designing, formulating, and communicating agent-based models. In *Agent-based models of geographical systems*, chap. 17, (pp. 361–377). Springer Science and Business Media.

Grimm, V., & Railsback, S. F. (2012b). Pattern-oriented modelling: a 'multi-scope'for predictive systems ecology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1586), 298–310.

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W. M., Railsback, S. F., Thulke, H.-H., Weiner, J., Wiegand, T., & DeAngelis, D. L. (2005). Pattern-oriented modeling of agent-based complex systems: lessons from Ecology. *Science*, *310*(5750), 987–991.

Guichard, S., Kriticos, D. J., Leriche, A., Kean, J. M., & Worner, S. P. (2012). Individual-based modelling of moth dispersal to improve biosecurity incursion response. *Journal of Applied Ecology*, *49*(1), 287–296.

Guisan, A., & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, *8*(9), 993–1009.

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, *135*(2), 147–186.

Gurevitch, J., Fox, G., Wardle, G., & Taub, D. (2011). Emergent insights from the synthesis of conceptual frameworks for biological invasions. *Ecology Letters*, *14*(4), 407–418.

Haefner, J. W. (2005). *Modeling Biological Systems: Principles and Applications*. Springer Science and Business Media. 475 pp.

Hagen, A. (2003). Fuzzy set approach to assessing similarity of categorical maps. *International Journal of Geographical Information Science*, *17*(3), 235–249.

Hagenlocher, M., Lang, S., Holbling, D., Tiede, D., & Kienberger, S. (2014). Modeling hotspots of climate change in the Sahel using object-based regionalization of multidimensional gridded datasets. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *7*(1), 229–234.

Hamed, M. M., Bedient, P. B., & Dawson, C. N. (1996). Probabilistic modeling of aquifer heterogeneity using reliability methods. *Advances in Water Resources*, *19*(5), 277–295.

Hamel, F., Fayard, J., & Roques, L. (2010). Spreading speeds in slowly oscillating environments. *Bulletin of Mathematical Biology*, *72*(5), 1166–1191.

Hanski, I. (1991). Single-species metapopulation dynamics: concepts, models and observations. *Biological Journal of the Linnean Society*, *42*(1), 17–38.

Hanski, I. (1999). *Metapopulation ecology*. Oxford University Press. 328 pp.

Hanski, I., & Ovaskainen, O. (2000). The metapopulation capacity of a fragmented landscape. *Nature*, *404*(6779), 755–758.

Hanski, I. A., & Gaggiotti, O. E. (2004). *Ecology, genetics and evolution of metapopulations*. Academic Press. 696 pp.

Hansson, L., Fahrig, L., & Merriam, G. (2012). *Mosaic landscapes and ecological processes*. Springer Science and Business Media. 356 pp.

Harper, K. A., Macdonald, S. E., Burton, P. J., Chen, J., Brosofske, K. D., Saunders, S. C., Euskirchen, E. S., Roberts, D., Jaiteh, M. S., & Esseen, P.-A. (2005). Edge influence on forest structure and composition in fragmented landscapes. *Conservation Biology*, *19*(3), 768–782.

Harris, R. J. (2002). Potential impact of the Argentine ant (*Linepithema humile*) in New Zealand and options for its control. *Science for Conservation*, *196*(36), 1–36.

Hartig, F., Calabrese, J. M., Reineking, B., Wiegand, T., & Huth, A. (2011). Statistical inference for stochastic simulation models–theory and application. *Ecology Letters*, *14*(8), 816–827.

Hartley, S., Harris, R., & Lester, P. J. (2006). Quantifying uncertainty in the potential distribution of an invasive species: climate and the argentine ant. *Ecology Letters*, *9*(9), 1068–1079.

Hastings, A., Cuddington, K., Davies, K. F., Dugaw, C. J., Elmendorf, S., Freestone, A., Harrison, S., Holland, M., Lambrinos, J., & Malvadkar, U. (2005). The spatial spread of invasions: new developments in theory and evidence. *Ecology Letters*, *8*(1), 91–101.

Haydar, D., & Wolff, W. J. (2011). Predicting invasion patterns in coastal ecosystems: relationship between vector strength and vector tempo. *Marine Ecology Progress Series*, *431*, 1–10.

Hayes, K. R., & Barry, S. C. (2008). Are there any consistent predictors of invasion success? *Biological Invasions*, *10*(4), 483–506.

He, H. S., DeZonia, B. E., & Mladenoff, D. J. (2000). An aggregation index (AI) to quantify spatial patterns of landscapes. *Landscape Ecology*, *15*(7), 591–601.

Heath, J., Pollard, E., & Thomas, J. A. (1984). *Atlas of butterflies in Britain and Ireland*. Viking. 160 pp.

Heikkinen, R. K., Luoto, M., Araújo, M. B., Virkkala, R., Thuiller, W., & Sykes, M. T. (2006). Methods and uncertainties in bioclimatic envelope modelling under climate change. *Progress in Physical Geography*, *30*(6), 751–777.

Heimpel, G. E., & Asplen, M. K. (2011). A 'Goldilocks' hypothesis for dispersal of biological control agents. *BioControl*, *56*(4), 441–450.

Hellmann, J. J., Byers, J. E., Bierwagen, B. G., & Dukes, J. S. (2008). Five potential consequences of climate change for invasive species. *Conservation Biology*, *22*(3), 534–543.

Helton, J. C. (2008). Uncertainty and sensitivity analysis for models of complex systems. In F. Graziani (Ed.) *Computational Methods in Transport: Verification and Validation*, vol. 62, (pp. 207–228). Springer. Lecture Notes in Computational Science and Engineering.

Helton, J. C., Johnson, J. D., Sallaberry, C. J., & Storlie, C. B. (2006). Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliability Engineering and System Safety*, *91*(10), 1175–1209.

Herrando-Pérez, S., Delean, S., Brook, B. W., & Bradshaw, C. J. (2012). Density dependence: an ecological Tower of Babel. *Oecologia*, *170*(3), 585–603.

Higgins, S. I., Richardson, D. M., & Cowling, R. M. (1996). Modeling invasive plant spread: the role of plant-environment interactions and model structure. *Ecology*, *77*(7), 2043–2054.

Higgins, S. I., Richardson, D. M., & Cowling, R. M. (2000). Using a dynamic landscape model for planning the management of alien plant invasions. *Ecological Applications*, *10*(6), 1833–1848.

Hijmans, R., Phillips, S., Leathwick, J., & Elith, J. (2016). Dismo: Species distribution modeling, version 1.0-15. *R-Cran package*. `https://cran.r-project.org/web/packages/dismo/index.html`.

Hill, J. K., Griffiths, H. M., & Thomas, C. D. (2011). Climate change and evolutionary adaptations at species' range margins. *Annual Review of Entomology*, *56*, 143–159.

Hobbs, R. J. (2000). Land-use changes and invasions. In H. A. Mooney, & R. J. Hobbs (Eds.) *Invasive species in a changing world*, chap. 3, (pp. 55–64). Island Press: Washington, DC, USA.

Holland, J. D., Fahrig, L., & Cappuccino, N. (2005). Body size affects the spatial scale of habitat-beetle interactions. *Oikos*, *110*(1), 101–108.

Holland, R. A., Wikelski, M., & Wilcove, D. S. (2006). How and why do insects migrate? *Science*, *313*(5788), 794–796.

Hopper, K. R., & Roush, R. T. (1993). Mate finding, dispersal, number released, and the success of biological control introductions. *Ecological Entomology*, *18*(4), 321–331.

Howe, R. W., Davis, G. J., & Mosca, V. (1991). The demographic significance of 'sink' populations. *Biological Conservation*, *57*(3), 239–255.

Hui, C., Krug, R. M., & Richardson, D. M. (2011). Modelling spread in invasion ecology: a synthesis. In D. M. Richardson (Ed.) *Fifty Years of Invasion Ecology: The Legacy of Charles Elton. Wiley-Blackwell, Chichester, West Sussex*, chap. 25, (pp. 329–343). Blackwell Publishing Ltd, Chichester, West Sussex.

Huntley, B., Barnard, P., Altwegg, R., Chambers, L., Coetzee, B. W., Gibson, L., Hockey, P. A., Hole, D. G., Midgley, G. F., & Underhill, L. G. (2010). Beyond bioclimatic envelopes: dynamic species' range and abundance modelling in the context of climatic change. *Ecography*, *33*(3), 621–626.

Ibáñez, I., Silander Jr, J. A., Wilson, A. M., LaFleur, N., Tanaka, N., & Tsuyama, I. (2009). Multivariate forecasts of potential distributions of invasive plant species. *Ecological Applications*, *19*(2), 359–375.

Jaeger, J. A. (2000). Landscape division, splitting index, and effective mesh size: new measures of landscape fragmentation. *Landscape Ecology*, *15*(2), 115–130.

Jager, H. I., & King, A. W. (2004). Spatial uncertainty and ecological models. *Ecosystems*, *7*(8), 841–847.

Jankovic, M., & Petrovskii, S. (2013). Gypsy moth invasion in north america: A simulation study of the spatial pattern and the rate of spread. *Ecological Complexity*, *14*, 132–144.

Janssen, J., Krol, M., Schielen, R., Hoekstra, A., & de Kok, J.-L. (2010). Assessment of uncertainties in expert knowledge, illustrated in fuzzy rule-based models. *Ecological Modelling*, *221*(9), 1245–1251.

Jerde, C. L., & Lewis, M. A. (2007). Waiting for invasions: a framework for the arrival of nonindigenous species. *The American Naturalist*, *170*(1), 1–9.

Jiménez-Valverde, A. (2012). Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography*, *21*(4), 498–507.

Johnson, D. M., Liebhold, A. M., Tobin, P. C., & Bjørnstad, O. N. (2006). Allee effects and pulsed invasion by the gypsy moth. *Nature*, *444*(7117), 361–363.

Johnson, G. D., & Patil, G. P. (2007). *Landscape pattern analysis for assessing ecosystem condition*, vol. 1. Springer. 130 pp.

Johnson, J. B., & Omland, K. S. (2004). Model selection in Ecology and Evolution. *Trends in Ecology and Evolution*, *19*(2), 101–108.

Joint Committee for Guides in Metrology report, JCGM (2008). Evaluation of measurement data-guide to the expression of uncertainty in measurement. Tech. Rep. JCGM 100:2008, (GUM 1995 with minor corrections) for the Joint Committee for Guides in Metrology (JCGM).

Jongejans, E., Skarpaas, O., & Shea, K. (2008). Dispersal, demography and spatial population models for conservation and control management. *Perspectives in Plant Ecology, Evolution and Systematics*, *9*(3), 153–170.

Jonsen, I. D., & Taylor, P. D. (2000). Fine-scale movement behaviors of calopterygid damselflies are influenced by landscape structure: an experimental manipulation. *Oikos*, *88*(3), 553–562.

Jonsson, M., Straub, C. S., Didham, R. K., Buckley, H. L., Case, B. S., Hale, R. J., Gratton, C., & Wratten, S. D. (2015). Experimental evidence that the effectiveness of conservation biological control depends on landscape complexity. *Journal of Applied Ecology*, *52*(5), 1274–1282.

Jules, E. S., Kauffman, M. J., Ritts, W. D., & Carroll, A. L. (2002). Spread of an invasive pathogen over a variable landscape: a nonnative root rot on Port Orford cedar. *Ecology*, *83*(11), 3167–3181.

Kavetski, D., Kuczera, G., & Franks, S. W. (2006). Bayesian analysis of input uncertainty in hydrological modeling: 2. application. *Water Resources Research*, *42*(3).

Kawasaki, K., Asano, K., & Shigesada, N. (2012). Impact of directed movement on invasive spread in periodic patchy environments. *Bulletin of Mathematical Biology*, *74*(6), 1448–1467.

Kean, J. (2013). How accurate are methods for predicting phenology in New Zealand? *New Zealand Plant Protection*, *66*, 124–131.

Kean, J., & Phillips, C. (2013). Phenology and diapause research for great white butterfly *(Pieris brassicae)*. Tech. Rep. RE400/2012/447, Agresearch, Lincoln, New Zealand: (Report for Ministry for Primary Industries). 41 pp.

Keane, R. M., & Crawley, M. J. (2002). Exotic plant invasions and the enemy release hypothesis. *Trends in Ecology and Evolution*, *17*(4), 164–170.

Keeling, M. J. (1999). The effects of local spatial structure on epidemiological invasions. *Proceedings of the Royal Society of London B: Biological Sciences*, *266*(1421), 859–867.

Keesing, F., Holt, R. D., & Ostfeld, R. S. (2006). Effects of species diversity on disease risk. *Ecology Letters*, *9*(4), 485–498.

Keith, D. A., Akçakaya, H. R., Thuiller, W., Midgley, G. F., Pearson, R. G., Phillips, S. J., Regan, H. M., Araújo, M. B., & Rebelo, T. G. (2008). Predicting extinction risks under climate change: coupling stochastic population models with dynamic bioclimatic habitat models. *Biology Letters*, *4*(5), 560–563.

Keitt, T. H., Lewis, M. A., & Holt, R. D. (2001). Allee effects, invasion pinning, and species' borders. *The American Naturalist*, *157*(2), 203–216.

Kennedy, T. A., Naeem, S., Howe, K. M., Knops, J. M., Tilman, D., & Reich, P. (2002). Biodiversity as a barrier to ecological invasion. *Nature*, *417*(6889), 636–638.

Kinezaki, N., Kawasaki, K., & Shigesada, N. (2006). Spatial dynamics of invasion in sinusoidally varying environments. *Population Ecology*, *48*(4), 263–270.

Kinezaki, N., Kawasaki, K., & Shigesada, N. (2010). The effect of the spatial configuration of habitat fragmentation on invasive spread. *Theoretical Population Biology*, *78*(4), 298–308.

Kinezaki, N., Kawasaki, K., Takasu, F., & Shigesada, N. (2003). Modeling biological invasions into periodically fragmented environments. *Theoretical Population Biology*, *64*(3), 291–302.

Kneitel, J. M., & Chase, J. M. (2004). Trade-offs in community ecology: linking spatial scales and species coexistence. *Ecology Letters*, *7*(1), 69–80.

Kocabas, V., & Dragicevic, S. (2006). Assessing cellular automata model behaviour using a sensitivity analysis approach. *Computers, Environment and Urban Systems*, *30*(6), 921–953.

Kohonen, T. (1988). Self-organization and associative memory. *1*.

Kohonen, T. (2001). *Self-organizing maps*. Springer-Verlag Berlin Heidelberg. 502 pp.

Kolar, C. S., & Lodge, D. M. (2001). Progress in invasion biology: predicting invaders. *Trends in Ecology and Evolution*, *16*(4), 199–204.

Kot, M., Lewis, M. A., & van den Driessche, P. (1996). Dispersal data and the spread of invading organisms. *Ecology*, *77*(7), 2027–2042.

Kriticos, D. J., Leriche, A., Bulman, L. S., Kimberley, M. O., Alcaraz, S., & Richardson, B. (2008). Modelling the efficacy of different sampling strategies for estimating disease levels and detecting the spread of new pests. Tech. Rep. Unpublished, Ensis: Australia's CSIRO Forest Biosciences and New Zealand Crown Research Institute, Scion, Rotorua, New Zealand. 37 pp.

Kriticos, D. J., Leriche, A., Palmer, D., Cook, D. C., Brockerhoff, E. G., Stephens, A. E., & Watt, M. S. (2013). Linking climate suitability, spread rates and host-impact when estimating the potential costs of invasive pests. *PloS One*, *8*(2), e54861.

Kriticos, D. J., Webber, B. L., Leriche, A., Ota, N., Macadam, I., Bathols, J., & Scott, J. K. (2012). CliMond: global high-resolution historical and future scenario climate surfaces for bioclimatic modelling. *Methods in Ecology and Evolution*, *3*(1), 53–64.

Krueger, T., Page, T., Hubacek, K., Smith, L., & Hiscock, K. (2012). The role of expert opinion in environmental modelling. *Environmental Modelling & Software*, *36*, 4–18.

Kuhnert, P. M., Martin, T. G., & Griffiths, S. P. (2010). A guide to eliciting and using expert knowledge in bayesian ecological models. *Ecology Letters*, *13*(7), 900–914.

Kupfer, J. A. (2012). Landscape ecology and biogeography: rethinking landscape metrics in a post-FRAGSTATS landscape. *Progress in Physical Geography*, *39*.

La Morgia, V., Malenotti, E., Badino, G., & Bona, F. (2011). Where do we go from here? dispersal simulations shed light on the role of landscape structure in determining animal redistribution after reintroduction. *Landscape Ecology*, *26*(7), 969–981.

Lambin, E. F., Turner, B. L., Geist, H. J., Agbola, S. B., Angelsen, A., Bruce, J. W., Coomes, O. T., Dirzo, R., Fischer, G., & Folke, C. (2001). The causes of land-use and land-cover change: moving beyond the myths. *Global Environmental Change*, *11*(4), 261–269.

Larson, D. L., Phillips-Mao, L., Quiram, G., Sharpe, L., Stark, R., Sugita, S., & Weiler, A. (2011). A framework for sustainable invasive species management: Environmental, social, and economic objectives. *Journal of Environmental Management*, *92*(1), 14–22.

Latimer, A. M., Wu, S., Gelfand, A. E., & Silander Jr, J. A. (2006). Building statistical models to analyze species distributions. *Ecological applications*, *16*(1), 33–50.

Lechner, A. M., Reinke, K. J., Wang, Y., & Bastin, L. (2013). Interactions between landcover pattern and geospatial processing methods: Effects on landscape metrics and classification accuracy. *Ecological Complexity*, *15*, 71–82.

Leitão, A. B., & Ahern, J. (2002). Applying landscape ecological concepts and metrics in sustainable landscape planning. *Landscape and urban planning*, *59*(2), 65–93.

Leitão, A. B., Miller, J., Ahern, J., & McGarigal, K. (2012). *Measuring landscapes: A planner's handbook*. Island press. 272 pp.

León-Cortés, J. L., Lennon, J. J., & Thomas, C. D. (2003). Ecological dynamics of extinct species in empty habitat networks. 1. the role of habitat pattern and quantity, stochasticity and dispersal. *Oikos*, *102*(3), 449–464.

Levin, S. A., Muller-Landau, H. C., Nathan, R., & Chave, J. (2003). The ecology and evolution of seed dispersal: a theoretical perspective. *Annual Review of Ecology, Evolution, and Systematics*, *64*, 575–604.

Levins, R. (1969). Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society of America*, *15*(3), 237–240.

Li, X., He, H. S., Bu, R., Wen, Q., Chang, Y., Hu, Y., & Li, Y. (2005). The adequacy of different landscape metrics for various landscape patterns. *Pattern Recognition*, *38*(12), 2626–2638.

Liebhold, A., & Bascompte, J. (2003). The Allee effect, stochastic dynamics and the eradication of alien species. *Ecology Letters*, *6*(2), 133–140.

Liebhold, A., & Mastro, V. (1989). Learning from the legacy of Leopold Trouvelot. *Bulletin of the Entomological Society of America*, *35*(2), 20–22.

Liebhold, A. M., Halverson, J. A., & Elmes, G. A. (1992). Gypsy moth invasion in North America: a quantitative analysis. *Journal of Biogeography*, *19*(5), 513–520.

Liebhold, A. M., & Tobin, P. C. (2006). Growth of newly established alien populations: comparison of North American gypsy moth colonies with invasion theory. *Population Ecology*, *48*(4), 253–262.

Liebhold, A. M., & Tobin, P. C. (2010). Exploiting the Achilles heels of pest invasions: Allee effects, stratified dispersal and management of forest insect establishment and spread. In *New Zealand Journal of Forestry Science*, vol. 40, (pp. 25–33). New Zealand Forest Research Institute.

Lindenmayer, D., Hobbs, R. J., Montague-Drake, R., Alexandra, J., Bennett, A., Burgman, M., Cale, P., Calhoun, A., Cramer, V., & Cullen, P. (2008). A checklist for ecological management of landscapes for conservation. *Ecology Letters*, *11*(1), 78–91.

Link, W. A., & Barker, R. J. (2009). *Bayesian inference: with ecological applications*. Academic Press. 400 pp.

Liu, S., Proctor, W., & Cook, D. (2010). Using an integrated fuzzy set and deliberative multi-criteria evaluation approach to facilitate decision-making in invasive species management. *Ecological Economics*, *69*(12), 2374–2382.

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*(2), 145–151.

Lockwood, J. L., Cassey, P., & Blackburn, T. (2005). The role of propagule pressure in explaining species invasions. *Trends in Ecology and Evolution*, *20*(5), 223–228.

Lockwood, J. L., Cassey, P., & Blackburn, T. M. (2009). The more you introduce the more you get: the role of colonization pressure and propagule pressure in invasion ecology. *Diversity and Distributions*, *15*(5), 904–910.

Lockwood, J. L., Hoopes, M. F., & Marchetti, M. P. (2013). *Invasion ecology*. John Wiley & Sons. 312 pp.

Lounibos, L. P. (2002). Invasions by insect vectors of human disease. *Annual Review of Entomology*, *47*(1), 233–266.

Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, *28*(5), 823–870.

Lukacs, P. M., Burnham, K. P., & Anderson, D. R. (2010). Model selection bias and Freedman's paradox. *Annals of the Institute of Statistical Mathematics*, *62*(1), 117–125.

Lustig, A., Stouffer, D. B., Roigé, M., & Worner, S. P. (2015). Towards more predictable and consistent landscape metrics across spatial scales. *Ecological Indicators*, *57*, 11–21.

MacEachren, A. M., Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M., & Hetzler, E. (2005). Visualizing geospatial information uncertainty: What we know and what we need to know. *Cartography and Geographic Information Science*, *32*(3), 139–160.

Mack, R. N., & Lonsdale, W. M. (2001). Humans as global plant dispersers: Getting more than we bargained for current introductions of species for aesthetic purposes present the largest single challenge for predicting which plant immigrants will become future pests. *BioScience*, *51*(2), 95–102.

Margosian, M. L., Garrett, K. A., & Hutchinson, J. S. (2009). Connectivity of the american agricultural landscape: assessing the national risk of crop pest and disease spread. *BioScience*, *59*(2), 141–151.

Marion, G., McInerny, G. J., Pagel, J., Catterall, S., Cook, A. R., Hartig, F., & O'Hara, R. B. (2012). Parameter and uncertainty estimation for process-oriented population and distribution models: data, statistics and the niche. *Journal of Biogeography*, *39*(12), 2225–2239.

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*(1), 59–69.

Martin, P. H., Canham, C. D., & Marks, P. L. (2008). Why forests appear resistant to exotic plant invasions: intentional introductions, stand dynamics, and the role of shade tolerance. *Frontiers in Ecology and the Environment*, *7*(3), 142–149.

Martin, T. G., Burgman, M. A., Fidler, F., Kuhnert, P. M., Low-Choy, S., McBride, M., & Mengersen, K. (2012). Eliciting expert knowledge in conservation science. *Conservation Biology*, *26*(1), 29–38.

Marvier, M., Kareiva, P., & Neubert, M. G. (2004). Habitat destruction, fragmentation, and disturbance promote invasion by habitat generalists in a multispecies metapopulation. *Risk Analysis*, *24*(4), 869–878.

Matott, L. S., Babendreier, J. E., & Purucker, S. T. (2009). Evaluating uncertainty in integrated environmental models: a review of concepts and tools. *Water Resources Research*, *45*(6).

Mayer, D. G., & Butler, D. G. (1993). Statistical validation. *Ecological Modelling*, *68*(1), 21–32.

Mayo, J. H., Straka, T. J., & Leonard, D. S. (2003). The cost of slowing the spread of the gypsy moth (*Lepidoptera: Lymantriidae*). *Journal of Economic Entomology*, *96*(5), 1448–1454.

McCarthy, M. A., Possingham, H. P., Day, J. R., & Tyre, A. (2001). Testing the accuracy of population viability analysis. *Conservation Biology*, *15*(4), 1030–1038.

McDermott, S. M., Irwin, R. E., & Taylor, B. W. (2013). Using economic instruments to develop effective management of invasive species: insights from a bioeconomic model. *Ecological Applications*, *23*(5), 1086–1100.

McGarigal, K., Cushman, S., & Ene, E. (2012). FRAGSTATS v4: spatial pattern analysis program for categorical and continuous maps.
URL `http://www.umass.edu/landeco/research/fragstats/fragstats.html`

McGarigal, K., & Marks, B. J. (1995). Spatial pattern analysis program for quantifying landscape structure. Gen. Tech. Rep. PNW-GTR-351. US Department of Agriculture, Forest Service, Pacific Northwest Research Station.

McGeoch, M. A., Spear, D., Kleynhans, E. J., & Marais, E. (2012). Uncertainty in invasive alien species listing. *Ecological Applications*, *22*(3), 959–971.

McGregor, K. F., Watt, M. S., Hulme, P. E., & Duncan, R. P. (2012). What determines pine naturalization: species traits, climate suitability or forestry use? *Diversity and Distributions*, *18*(10), 1013–1023.

Mcgwire, K. C., & Fisher, P. (2001). Spatially variable thematic accuracy: Beyond the confusion matrix. In C. T. Hunsaker (Ed.) *Spatial uncertainty in ecology: implications for remote sensing and GIS applications*, (pp. 308–329). Springer.

McMahon, S. M., Dietze, M. C., Hersh, M. H., Moran, E. V., & Clark, J. S. (2009). A predictive framework to understand forest responses to global change. *Annals of the New York Academy of Sciences*, *1162*(1), 221–236.

Meentemeyer, R. K., Cunniffe, N. J., Cook, A. R., Filipe, J. A., Hunter, R. D., Rizzo, D. M., & Gilligan, C. A. (2011). Epidemiological modeling of invasion in heterogeneous landscapes: spread of sudden oak death in California (1990-2030). *Ecosphere*, *2*(2, Article 27), 1–24.

Meier, E. S., Dullinger, S., Zimmermann, N. E., Baumgartner, D., Gattringer, A., & Hülber, K. (2014). Space matters when defining effective management for invasive plants. *Diversity and Distributions*, *20*(9), 1029–1043.

Meyerson, L. A., & Mooney, H. A. (2007). Invasive alien species in an era of globalization. *Frontiers in Ecology and the Environment*, *5*(4), 199–208.

Millennium Ecosystem Assessment (2005). Ecosystems and human well-being. millennium ecosystem assessment 2005. ecosystems and human well-being synthesis.

Millington, A., Blumler, M., & Schickhoff, U. (2011). *The SAGE handbook of biogeography*. Sage. 624 pp.

Ming, P., & Albrecht, J. (2004). Integrated framework for the simulation of biological invasions in a heterogeneous landscape. *Transactions in GIS*, *8*(3), 309–334.

Minor, E. S., & Gardner, R. H. (2011). Landscape connectivity and seed dispersal characteristics inform the best management strategy for exotic plants. *Ecological Applications*, *21*(3), 739–749.

Mitchell, C. E., Agrawal, A. A., Bever, J. D., Gilbert, G. S., Hufbauer, R. A., Klironomos, J. N., Maron, J. L., Morris, W. F., Parker, I. M., & Power, A. G. (2006). Biotic interactions and plant invasions. *Ecology Letters*, *9*(6), 726–740.

Moilanen, A. (1999). Patch occupancy models of metapopulation dynamics: efficient parameter estimation using implicit statistical inference. *Ecology*, *80*(3), 1031–1043.

Moilanen, A., & Cabeza, M. (2002). Single-species dynamic site selection. *Ecological Applications*, *12*(3), 913–926.

Moilanen, A., & Wintle, B. A. (2006). Uncertainty analysis favours selection of spatially aggregated reserve networks. *Biological Conservation*, *129*(3), 427–434.

Moilanen, A., & Wintle, B. A. (2007). The boundary-quality penalty: a quantitative method for approximating species responses to fragmentation in reserve selection. *Conservation Biology*, *21*(2), 355–364.

Mollison, D. (1972). The rate of spatial propagation of simple epidemics. In *Proceedings of the sixth Berkeley symposium on mathematical statistics and probability*, vol. 3, (pp. 579–614). Berkeley and Los Angeles University, USA and University of California Edinburgh, UK.

Mollison, D. (1991). Dependence of epidemic and population velocities on basic parameters. *Mathematical Biosciences*, *107*(2), 255–287.

Morales, J. M., Moorcroft, P. R., Matthiopoulos, J., Frair, J. L., Kie, J. G., Powell, R. A., Merrill, E. H., & Haydon, D. T. (2010). Building the bridge between animal movement and population dynamics. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *365*(1550), 2289–2301.

Morel-Journel, T., Girod, P., Mailleret, L., Auguste, A., Blin, A., & Vercken, E. (2015). The highs and lows of dispersal: how connectivity and initial population size jointly shape establishment dynamics in discrete landscapes. *Oikos*, *0*, 001–009.

Morin, R. S., Liebhold, A. M., Luzader, E. R., Lister, A. J., Gottschalk, K. W., & Twardus, D. B. (2005). Mapping host-species abundance of three major exotic forest pests. Tech. Rep. NE-726, Department of Agriculture, Forest Service, Northeastern Research Station, Newtown Square, PA: U.S.

Morss, R. E., Demuth, J. L., & Lazo, J. K. (2008). Communicating uncertainty in weather forecasts: a survey of the US public. *Weather and Forecasting*, *23*(5), 974–991.

Mortensen, D. A., Rauschert, E. S., Nord, A. N., & Jones, B. P. (2009). Forest roads facilitate the spread of invasive plants. *BioOne Research Evolved*, *2*(3), 191–199.

Mouquet, N., Lagadeuc, Y., Devictor, V., Doyen, L., Duputié, A., Eveillard, D., Faure, D., Garnier, E., Gimenez, O., Huneman, P., Jabot, F., Jarne, P., Joly, D., Julliard, R., Kéfi, S., Kergoat, G. J., Lavorel, S., Le Gall, L., Meslin, L., Morand, S., Morin, X., Morlon, H., Pinay, G., Pradel, R., Schurr, F. M., Thuiller, W., & Loreau, M. (2015). Review: Predictive ecology in a changing world. *Journal of Applied Ecology*, *52*(5), 1293–1310.

Mundt, C. C., Sackett, K. E., & Wallace, L. D. (2011). Landscape heterogeneity and disease spread: experimental approaches with a plant pathogen. *Ecological Applications*, *21*(2), 321–328.

Nathan, R., Katul, G. G., Horn, H. S., Thomas, S. M., Oren, R., Avissar, R., Pacala, S. W., & Levin, S. A. (2002). Mechanisms of long-distance dispersal of seeds by wind. *Nature*, *418*(6896), 409–413.

Neel, M. C., McGarigal, K., & Cushman, S. A. (2004). Behavior of class-level landscape metrics across gradients of class aggregation and area. *Landscape Ecology*, *19*(4), 435–455.

Nehrbass, N., & Winkler, E. (2007). Is the Giant Hogweed still a threat? An individual-based modelling approach for local invasion dynamics of *Heracleum mantegazzianum*. *Ecological Modelling*, *201*(3), 377–384.

Neubert, M. (1997). A simple population model with qualitatively uncertain dynamics. *Journal of Theoretical Biology*, *189*(4), 399–411.

Neubert, M. G., & Caswell, H. (2000). Demography and dispersal: calculation and sensitivity analysis of invasion speed for structured populations. *Ecology*, *81*(6), 1613–1628.

Neubert, M. G., Kot, M., & Lewis, M. A. (2000). Invasion speeds in fluctuating environments. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, *267*(1453), 1603–1610.

Niinemets, Ü., & Peñuelas, J. (2008). Gardening and urban landscaping: significant players in global change. *Trends in Plant Science*, *13*(2), 60–65.

Nuñez, M. A., Moretti, A., & Simberloff, D. (2011). Propagule pressure hypothesis not supported by an 80-year experiment on woody species invasion. *Oikos*, *120*(9), 1311–1316.

Nuñez, M. A., & Simberloff, D. (2005). Invasive species and the cultural keystone species concept. *Ecology and Society*, *10*(1), r4.

Oborny, B., Szabó, G., & Meszéna, G. (2007). Survival of species in patchy landscapes: percolation in space and time. In D. Storch, P. Marquet, & J. Brown (Eds.) *Scaling biodiversity*, chap. 20, (pp. 409–440). Cambridge University Press, Cambridge.

Oksanen, J., Blanchet, F., Kindt, R., Legendre, P., O'Hara, B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Wagner, H., & Oksanen, J. (2016). vegan: Community ecology package, version 2.3-3. *R-Cran package*. `https://cran.r-project.org/web/packages/vegan/index.html`.

O'Neill, R., Krummel, J., Gardner, R., Sugihara, G., Jackson, B., DeAngelis, D., Milne, B., Turner, M., Zygmunt, B., Christensen, S., Dale, V., & Graham, R. (1988). Indices of landscape pattern. *Landscape Ecology*, *1*(3), 153–162.

O'Neill, R. V., Hunsaker, C. T., Jones, K. B., Riitters, K. H., Wickham, J. D., Schwartz, P. M., Goodman, I. A., Jackson, B. L., & Baillargeon, W. S. (1997). Monitoring environmental quality at the landscape scale. *BioScience*, (pp. 513–519).

O'Sullivan, D., & Perry, G. L. (2013). *Spatial simulation: exploring pattern and process*. John Wiley and Sons. 332 pp.

Owen, M., & Lewis, M. (2001). How predation can slow, stop or reverse a prey invasion. *Bulletin of Mathematical Biology*, *63*(4), 655–684.

Parry, H. R., Sadler, R. J., & Kriticos, D. J. (2013). Practical guidelines for modelling post-entry spread in invasion ecology. *NeoBiota*, *18*, 41–66.

Parysow, P., Gertner, G., & Westervelt, J. (2000). Efficient approximation for building error budgets for process models. *Ecological Modelling*, *135*(2), 111–125.

Peacock, L., & Worner, S. P. (2008). Biological and ecological traits that assist establishment of alien invasive insects. *New Zealand Plant Protection*, *61*, 1–7.

Peng, J., Wang, Y., Zhang, Y., Wu, J., Li, W., & Li, Y. (2010). Evaluating the effectiveness of landscape metrics in quantifying spatial patterns. *Ecological Indicators*, *10*(2), 217–223.

Perrings, C., Dehnen-Schmutz, K., Touza, J., & Williamson, M. (2005). How to manage biological invasions under globalization. *Trends in Ecology and Evolution*, *20*(5), 212–215.

Peterjohn, W. T., & Correll, D. L. (1984). Nutrient dynamics in an agricultural watershed: observations on the role of a riparian forest. *Ecology*, *65*(5), 1466–1475.

Petitpierre, B., Kueffer, C., Broennimann, O., Randin, C., Daehler, C., & Guisan, A. (2012). Climatic niche shifts are rare among terrestrial plant invaders. *Science*, *335*(6074), 1344–1348.

Petrovskii, S., & Li, B.-L. (2003). An exactly solvable model of population dynamics with density-dependent migrations and the Allee effect. *Mathematical Biosciences*, *186*(1), 79–91.

Petrovskii, S. V., & McKay, K. (2010). Biological invasion and biological control: a case study of the gypsy moth spread. In *What makes an alien invasive? Risk and policy responses*. The George Hotel, Edinburgh, 7-8 December 2010: Association of Applied Biologists.

Phillips, C., Brown, K., Green, C., Walker, G., Broome, K., Toft, R., Vander Lee, B., & King, M. (2014). *Pieris brassicae* (great white butterfly) eradication annual report. Tech. rep., Whakatū/Nelson Office, Department of Conservation, New Zealand (Report for Ministry for Primary Industries). 37 pp.

Pickett, S. T., & White, P. S. (2013). *The ecology of natural disturbance and patch dynamics*. Elsevier. 472 pp.

Pimentel, D., Lach, L., Zuniga, R., & Morrison, D. (2000). Environmental and economic costs of non-indigenous species in the United States. *BioScience*, *50*(1), 53–65.

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*(3), 273–288.

Piou, C., Berger, U., & Grimm, V. (2009). Proposing an information criterion for individual-based models developed in a pattern-oriented modelling framework. *Ecological Modelling*, *220*(17), 1957–1967.

Pitt, J. P. W. (2008). *Modelling the spread of invasive species across heterogeneous landscapes*. PhD, Lincoln University, New Zealand. 232 pp.

Pitt, J. P. W., Kriticos, D. J., & Dodd, M. B. (2011). Temporal limits to simulating the future spread pattern of invasive species: *Buddleja davidii* in Europe and New Zealand. *Ecological Modelling*, *222*(11), 1880–1887.

Pitt, J. P. W., Régnière, J., & Worner, S. (2007). Risk assessment of the gypsy moth, *Lymantria dispar* (L), in New Zealand based on phenology modelling. *International Journal of Biometeorology*, *51*(4), 295–305.

Pitt, J. P. W., Worner, S. P., & Suarez, A. V. (2009). Predicting argentine ant spread over the heterogeneous landscape using a spatially explicit stochastic model. *Ecological Applications*, *19*(5), 1176–1186.

Poli, R., Langdon, W. B., McPhee, N. F., & Koza, J. R. (2008). *A field guide to genetic programming*. Lulu.com. 250 pp.

Pontius, R. G. (2000). Quantification error versus location error in comparison of categorical maps. *Photogrammetric Engineering and Remote Sensing*, *66*(8), 1011–1016.

Pontius, R. G. (2002). Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. *Photogrammetric Engineering and Remote Sensing*, *68*(10), 1041–1050.

Portielje, R., Hvitved-Jacobsen, T., & Schaarup-Jensen, K. (2000). Risk analysis using stochastic reliability methods applied to two cases of deterministic water quality models. *Water Research*, *34*(1), 153–170.

Power, C., Simms, A., & White, R. (2001). Hierarchical fuzzy pattern matching for the regional comparison of land use maps. *International Journal of Geographical Information Science*, *15*(1), 77–100.

Pyšek, P., Bacher, S., Chytrỳ, M., Jarošík, V., Wild, J., Celesti-Grapow, L., Gassó, N., Kenis, M., Lambdon, P. W., & Nentwig, W. (2010). Contrasting patterns in the invasions of european terrestrial and freshwater habitats by alien plants, insects and vertebrates. *Global Ecology and Biogeography*, *19*(3), 317–331.

Pyšek, P., & Hulme, P. E. (2005). Spatio-temporal dynamics of plant invasions: linking pattern to process. *Ecoscience*, *12*(3), 302–315.

Pyšek, P., & Richardson, D. M. (2010). Invasive species, environmental change and management, and health. *Annual Review of Environment and Resources*, *35*, 25–55.

Quesnelle, P. E., Lindsay, K. E., & Fahrig, L. (2014). Low reproductive rate predicts species sensitivity to habitat loss: a meta-analysis of wetland vertebrates. *PloS One*, *9*(3), e90926.

Radeloff, V. C., Mladenoff, D. J., & Boyce, M. S. (2000). The changing relation of landscape patterns and jack pine budworm populations during an outbreak. *Oikos*, *90*(3), 417–430.

Railsback, S. F., & Grimm, V. (2011). *Agent-based and individual-based modeling: a practical introduction*. Princeton University Press. 352 pp.

Ramula, S., Knight, T. M., Burns, J. H., & Buckley, Y. M. (2008). General guidelines for invasive plant management based on comparative demography of invasive and native plant populations. *Journal of Applied Ecology*, *45*(4), 1124–1133.

Reaser, J. K., Meyerson, L. A., & Von Holle, B. (2008). Saving camels from straws: how propagule pressure-based prevention policies can reduce the risk of biological invasion. *Biological Invasions*, *10*(7), 1085–1098.

Refsgaard, J. C., van der Sluijs, J. P., Højberg, A. L., & Vanrolleghem, P. A. (2007). Uncertainty in the environmental modelling process: a framework and guidance. *Environmental Modelling & Software*, *22*(11), 1543–1556.

Regan, H. M., Colyvan, M., & Burgman, M. A. (2002). A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecological Applications*, *12*(2), 618–628.

Régnière, J., & Nealis, V. (2002). Modelling seasonality of gypsy moth, *Lymantria dispar* (*Lepidoptera*: *Lymantriidae*), to evaluate probability of its persistence in novel environments. *The Canadian Entomologist*, *134*(06), 805–824.

Régnière, J., & Sharov, A. (1999). Simulating temperature-dependent ecological processes at the subcontinental scale: male gypsy moth flight phenology as an example. *International Journal of Biometeorology*, *42*(3), 146–152.

Remmel, T. K., & Csillag, F. (2003). When are two landscape pattern indices significantly different? *Journal of Geographical Systems*, *5*(4), 331–351.

Rempel, R. S., Kaukinen, D., & Carr, A. P. (2012). Patch analyst and patch grid. Tech. rep., Ontario Ministry of Natural Resources. Centre for Northern Forest Ecosystem Research, Canada.

Renton, M., Savage, D., & Chopard, J. (2011). A general spatially-explicit model to inform rapid response to new biological invasions: why do we need one and what should it look like. In *MODSIM2011, 19th International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand*.

Richards, T. J., Ellsworth, P., Tronstad, R., & Naranjo, S. (2010). Market-based instruments for the optimal control of invasive insect species: B. tabaci in arizona. *Journal of Agricultural and Resource Economics*, (pp. 349–367).

Richardson, D. M., Pyšek, P., Rejmánek, M., Barbour, M. G., Panetta, F. D., & West, C. J. (2000). Naturalization and invasion of alien plants: concepts and definitions. *Diversity and Distributions*, *6*(2), 93–107.

Richter, R., Dullinger, S., Essl, F., Leitner, M., & Vogl, G. (2013). How to account for habitat suitability in weed management programmes? *Biological Invasions*, *15*(3), 657–669.

Ricker, W. E. (1958). Handbook of computations for biological statistics offish populations. *Bulletin of the Fisheries Research Board of Canada*, *119*, 1–300.

Ridgeway, G. (2015). gbm: Generalized boosted regression models, version 2.1.1. *R-Cran package*. `https://cran.r-project.org/web/packages/gbm/index.html`.

Rigot, T., van Halder, I., & Jactel, H. (2014). Landscape diversity slows the spread of an invasive forest pest species. *Ecography*, *37*(7), 648–658.

Riitters, K. H., O'neill, R. V., Hunsaker, C. T., Wickham, J. D., Yankee, D. H., Timmins, S. P., Jones, K. B., & Jackson, B. L. (1995). A factor analysis of landscape pattern and structure metrics. *Landscape Ecology*, *10*(1), 23–39.

Robinet, C., Imbert, C.-E., Rousselet, J., Sauvard, D., Garcia, J., Goussard, F., & Roques, A. (2012a). Human-mediated long-distance jumps of the pine processionary moth in Europe. *Biological Invasions*, *14*(8), 1557–1569.

Robinet, C., Kehlenbeck, H., Kriticos, D. J., Baker, R. H., Battisti, A., Brunel, S., Dupin, M., Eyre, D., Faccoli, M., & Ilieva, Z. (2012b). A suite of models to support the quantitative assessment of spread in pest risk analysis. *Plos One*, (e43366).

Robinet, C., Lance, D. R., Thorpe, K. W., Onufrieva, K. S., Tobin, P. C., & Liebhold, A. M. (2008). Dispersion in time and space affect mating success and Allee effects in invading gypsy moth populations. *Journal of Animal Ecology*, *77*(5), 966–973.

Robinet, C., Roques, A., Pan, H., Fang, G., Ye, J., Zhang, Y., & Sun, J. (2009). Role of human-mediated dispersal in the spread of the pinewood nematode in China. *PLoS One*, *4*(2), e4646.

Robinet, C., Van Opstal, N., Baker, R., & Roques, A. (2011). Applying a spread model to identify the entry points from which the pine wood nematode, the vector of pine wilt disease, would spread most rapidly across Europe. *Biological Invasions*, *13*(12), 2981–2995.

Robinson, R. A., Morrison, C. A., & Baillie, S. R. (2014). Integrating demographic data: towards a framework for monitoring wildlife populations at large spatial scales. *Methods in Ecology and Evolution*, *5*(12), 1361–1372.

Robledo-Arnuncio, J. J., Klein, E. K., Muller-Landau, H. C., & Santamaría, L. (2014). Space, time and complexity in plant dispersal ecology. *Movement Ecology*, *2*(1), 16.

Rocchini, D., Delucchi, L., Bacaro, G., Cavallini, P., Feilhauer, H., Foody, G. M., He, K. S., Nagendra, H., Porta, C., & Ricotta, C. (2013). Calculating landscape diversity with information-theory based indices: A GRASS GIS solution. *Ecological Informatics*, *17*, 82–93.

Rocchini, D., Hortal, J., Lengyel, S., Lobo, J. M., Jimenez-Valverde, A., Ricotta, C., Bacaro, G., & Chiarucci, A. (2011). Accounting for uncertainty when mapping species distributions: the need for maps of ignorance. *Progress in Physical Geography*, *35*(2), 211–226.

Roer, H. (1959). Über Flug-und Wandergewohnheiten von *Pieris brassicae L. Zeitschrift für angewandte Entomologie*, *44*(3), 272–309.

Roques, L., Auger-Rozenberg, M.-A., & Roques, A. (2008). Modelling the impact of an invasive insect via reaction-diffusion. *Mathematical Biosciences*, *216*(1), 47–55.

Roques, L., & Chekroun, M. D. (2010). Does reaction-diffusion support the duality of fragmentation effect? *Ecological Complexity*, *7*(1), 100–106.

Roques, L., & Hamel, F. (2007). Mathematical analysis of the optimal habitat configurations for species persistence. *Mathematical Biosciences*, *210*(1), 34–59.

Sagata, K., & Lester, P. J. (2009). Behavioural plasticity associated with propagule size, resources, and the invasion success of the Argentine ant *Linepithema humile*. *Journal of Applied Ecology*, *46*(1), 19–27.

Sakai, A. K., Allendorf, F. W., Holt, J. S., Lodge, D. M., Molofsky, J., With, K. A., Baughman, S., Cabin, R. J., Cohen, J. E., & Ellstrand, N. C. (2001). The population biology of invasive specie. *Annual Review of Ecology and Systematics*, (pp. 305–332).

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2008). *Global sensitivity analysis: the primer*. John Wiley and Sons, Ltd. 304 pp.

Samia, Y., & Lutscher, F. (2010). Coexistence and spread of competitors in heterogeneous landscapes. *Bulletin of Mathematical Biology*, *72*(8), 2089–2112.

Saupe, D. (1988). Algorithms for random fractals. In *The science of fractal images*, (pp. 71–136). Springer.

Saura, S. (2003). Simmap 2.0 landscape categorical spatial patterns simulation software: User's manual. *Universidad Politecnica de Madrid*.

Saura, S. (2004). Effects of remote sensor spatial resolution and data aggregation on selected fragmentation indices. *Landscape Ecology*, *19*(2), 197–209.

Saura, S., & Martinez-Millan, J. (2001). Sensitivity of landscape pattern metrics to map spatial extent. *Photogrammetric Engineering and Remote Sensing*, *67*(9), 1027–1036.

Schindler, D. E., & Hilborn, R. (2015). Prediction, precaution, and policy under global change. *Science*, *347*(6225), 953–954.

Schreiber, S. J., & Ryan, M. E. (2011). Invasion speeds for structured populations in fluctuating environments. *Theoretical Ecology*, *4*(4), 423–434.

Schurr, F. M., Pagel, J., Cabral, J. S., Groeneveld, J., Bykova, O., O'Hara, R. B., Hartig, F., Kissling, W. D., Linder, H. P., & Midgley, G. F. (2012). How to understand species' niches and range dynamics: a demographic research agenda for biogeography. *Journal of Biogeography*, *39*(12), 2146–2162.

Sebert-Cuvillier, E., Simon-Goyheneche, V., Paccaut, F., Chabrerie, O., Goubet, O., & Decocq, G. (2008). Spatial spread of an alien tree species in a heterogeneous forest landscape: a spatially realistic simulation model. *Landscape Ecology*, *23*(7), 787–801.

Senay, S. D. (2014). *Modelling invasive species-landscape interactions using high resolution, spatially explicit models*. PhD, Lincoln University, New Zealand. 352 pp.

Senay, S. D., Worner, S. P., & Ikeda, T. (2013). Novel three-step pseudo-absence selection technique for improved species distribution modelling. *PloS One*, *8*(8).

Seno, H., & Koshiba, S. (2005). A mathematical model for invasion range of population dispersion through a patchy environment. *Biological Invasions*, *7*(5), 757–770.

Sexton, J. P., McIntyre, P. J., Angert, A. L., & Rice, K. J. (2009). Evolution and ecology of species range limits. *Annual Reviews Ecology, Evolution, and Systematics*, *40*, 415–436.

Shackley, S., & Wynne, B. (1996). Representing uncertainty in global climate change science and policy: Boundary-ordering devices and authority. *Science, Technology and Human Values*, *21*(3), 275–302.

Sharov, A. A. (2004). Bioeconomics of managing the spread of exotic pest species with barrier zones. *Risk Analysis*, *24*(4), 879–892.

Sharov, A. A., & Liebhold, A. M. (1998). Model of slowing the spread of gypsy moth (*Lepidoptera: Lymantriidae*) with a barrier zone. *Ecological Applications*, *8*(4), 1170–1179.

Shaw, M. W. (1995). Simulation of population expansion and spatial pattern when individual dispersal distributions do not decline exponentially with distance. *Proceedings of the Royal Society of London B: Biological Sciences*, *259*(1356), 243–248.

Shea, K., & Chesson, P. (2002). Community ecology theory as a framework for biological invasions. *Trends in Ecology and Evolution*, *17*(4), 170–176.

Shigesada, N., & Kawasaki, K. (1997). *Biological invasions: theory and practice*. Oxford University Press, UK. 218 pp.

Shigesada, N., Kawasaki, K., & Takeda, Y. (1995). Modeling stratified diffusion in biological invasions. *The American Naturalist*, *146*(2), 229–251.

Shigesada, N., Kawasaki, K., & Teramoto, E. (1986). Traveling periodic waves in heterogeneous environments. *Theoretical Population Biology*, *30*(1), 143–160.

Simberloff, D., Martin, J.-L., Genovesi, P., Maris, V., Wardle, D. A., Aronson, J., Courchamp, F., Galil, B., García-Berthou, E., & Pascal, M. (2013). Impacts of biological invasions: what's what and the way forward. *Trends in Ecology and Evolution*, *28*(1), 58–66.

Simberloff, D., & Von Holle, B. (1999). Positive interactions of nonindigenous species: invasional meltdown? *Biological Invasions*, *1*(1), 21–32.

Skaggs, T., & Barry, D. (1996). Assessing uncertainty in subsurface solute transport: efficient first-order reliability methods. *Environmental Software*, *11*(1), 179–184.

Skellam, J. G. (1951). Random dispersal in theoretical populations. *Biometrika*, (pp. 196–218).

Skelsey, P., & Garrett, K. A. (2013). Pest and disease management: Why we shouldn't go against the grain. *PloS one*, *8*(9), e75892.

Smith, R., Tan, C., Srimani, J. K., Pai, A., Riccione, K. A., Song, H., & You, L. (2014). Programmed Allee effect in bacteria causes a tradeoff between population spread and survival. *Proceedings of the National Academy of Sciences*, *111*(5), 1969–1974.

Sol, D., Bartomeus, I., & Griffin, A. S. (2012). The paradox of invasion in birds: competitive superiority or ecological opportunism? *Oecologia*, *169*(2), 553–564.

Soons, M., Messelink, J., Jongejans, E., & Heil, G. (2005). Habitat fragmentation reduces grassland connectivity for both short-distance and long-distance wind-dispersed forbs. *Journal of Ecology*, *93*(6), 1214–1225.

Spiegelhalter, D., Pearson, M., & Short, I. (2011). Visualizing uncertainty about the future. *Science*, *333*(6048), 1393–1400.

Spieth, H. R., & Cordes, R. (2012). Geographic comparison of seasonal migration events of the large white butterfly, *Pieris brassicae*. *Ecological Entomology*, *37*(6), 439–445.

Steiner, A. (2010). Counting the cost of alien invasions. *BBC news*. `http://news.bbc.co.uk/2/hi/science/nature/8615398.stm`.

Sutherland, W. J., Freckleton, R. P., Godfray, H. C. J., Beissinger, S. R., Benton, T., Cameron, D. D., Carmel, Y., Coomes, D. A., Coulson, T., & Emmerson, M. C. (2013). Identification of 100 fundamental ecological questions. *Journal of Ecology*, *101*(1), 58–67.

Szabó, S., Túri, Z., & Márton, S. (2014). Factors biasing the correlation structure of patch level landscape metrics. *Ecological Indicators*, *36*, 1–10.

Taranu, Z. E., Gregory-Eaves, I., Leavitt, P. R., Bunting, L., Buchaca, T., Catalan, J., Domaizon, I., Guilizzoni, P., Lami, A., & McGowan, S. (2015). Acceleration of cyanobacterial dominance in north temperate-subarctic lakes during the anthropocene. *Ecology Letters*, *18*(4), 375–384.

Taylor, C. M., & Hastings, A. (2005). Allee effects in biological invasions. *Ecology Letters*, *8*(8), 895–908.

Thiele, J., Schuckert, U., & Otte, A. (2008). Cultural landscapes of germany are patch-corridor-matrix mosaics for an invasive megaforb. *Landscape Ecology*, *23*(4), 453–465.

Thiele, J. C., & Grimm, V. (2015). Replicating and breaking models: good for you and good for ecology. *Oikos*, *124*(6), 691–696.

Thies, C., Haenke, S., Scherber, C., Bengtsson, J., Bommarco, R., Clement, L. W., Ceryngier, P., Dennis, C., Emmerson, M., & Gagic, V. (2011). The relationship between agricultural intensification and biological control: experimental tests across europe. *Ecological Applications*, *21*(6), 2187–2196.

Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J., Collingham, Y. C., Erasmus, B. F., De Siqueira, M. F., Grainger, A., & Hannah, L. (2004). Extinction risk from climate change. *Nature*, *427*(6970), 145–148.

Thomas, S. M., & Moloney, K. A. (2013). Hierarchical factors impacting the distribution of an invasive species: landscape context and propagule pressure. *Landscape Ecology*, *28*(1), 81–93.

Thuiller, W., Albert, C., Araújo, M. B., Berry, P. M., Cabeza, M., Guisan, A., Hickler, T., Midgley, G. F., Paterson, J., & Schurr, F. M. (2008). Predicting global change impacts on plant species' distributions: future challenges. *Perspectives in Plant Ecology, Evolution and Systematics*, *9*(3), 137–152.

Thuiller, W., F Midgley, G., Rougeti, M., & M Cowling, R. (2006). Predicting patterns of plant species richness in megadiverse South Africa. *Ecography*, *29*(5), 733–744.

Thuiller, W., Münkemüller, T., Lavergne, S., Mouillot, D., Mouquet, N., Schiffers, K., & Gravel, D. (2013). A road map for integrating eco-evolutionary processes into biodiversity models. *Ecology Letters*, *16*(s1), 94–105.

Tobin, P. C., & Blackburn, L. M. (2007). Slow the spread: a national program to manage the gypsy moth. Tech. Rep. NRS-6, General Technical Report-Northern Research Station, USDA Forest Service.

Tobin, P. C., Liebhold, A. M., & Anderson, E. (2015). Estimating spread rates of non-native species: the gypsy moth as a case study. In R. C. Venette (Ed.) *Pest Risk Modelling and Mapping for Invasive Alien Species*, chap. 9, (pp. 131–144). CAB International and USDA.

Tobin, P. C., Liebhold, A. M., & Anderson Roberts, E. (2007). Comparison of methods for estimating the spread of a non-indigenous species. *Journal of Biogeography*, *34*(2), 305–312.

Trakhtenbrot, A., Nathan, R., Perry, G., & Richardson, D. M. (2005). The importance of long-distance dispersal in biodiversity conservation. *Diversity and Distributions*, *11*(2), 173–181.

Travis, J. M., Delgado, M., Bocedi, G., Baguette, M., Bartoń, K., Bonte, D., Boulangeat, I., Hodgson, J. A., Kubisch, A., & Penteriani, V. (2013). Dispersal and species' responses to climate change. *Oikos*, *122*(11), 1532–1540.

Tsai, C. W., & Franceschini, S. (2005). Evaluation of probabilistic point estimate methods in uncertainty analysis for environmental engineering applications. *Journal of Environmental Engineering*, *131*(3), 387–395.

Tscharntke, T., Tylianakis, J. M., Rand, T. A., Didham, R. K., Fahrig, L., Batary, P., Bengtsson, J., Clough, Y., Crist, T. O., & Dormann, C. F. (2012). Landscape moderation of biodiversity patterns and processes-eight hypotheses. *Biological Reviews*, *87*(3), 661–685.

Turner, M. G. (1990). Spatial and temporal analysis of landscape patterns. *Landscape Ecology*, *4*(1), 21–30.

Turner, M. G. (2005). Landscape ecology: what is the state of the science? *Annual Review of Ecology, Evolution, and Systematics*, (pp. 319–344).

Tylianakis, J. M., Didham, R. K., Bascompte, J., & Wardle, D. A. (2008). Global change and species interactions in terrestrial ecosystems. *Ecology Letters*, *11*(12), 1351–1363.

Uuemaa, E., Mander, Ü., & Marja, R. (2013). Trends in the use of landscape spatial metrics as landscape indicators: A review. *Ecological Indicators*, *28*, 100–106.

Van den Bosch, F., Hengeveld, R., & Metz, J. (1992). Analysing the velocity of animal range expansion. *Journal of Biogeography*, *19*(2), 135–150.

Van Houtan, K. S., Pimm, S. L., Halley, J. M., Bierregaard, R. O., & Lovejoy, T. E. (2007). Dispersal of amazonian birds in continuous and fragmented forest. *Ecology Letters*, *10*(3), 219–229.

Van Kirk, R. W., & Lewis, M. A. (1997). Integrodifference models for persistence in fragmented habitats. *Bulletin of Mathematical Biology*, *59*(1), 107–137.

Vance, M. D., Fahrig, L., & Flather, C. H. (2003). Effect of reproductive rate on minimum habitat requirements of forest-breeding birds. *Ecology*, *84*(10), 2643–2653.

Veit, R. R., & Lewis, M. A. (1996). Dispersal, population growth, and the allee effect: dynamics of the house finch invasion of eastern north america. *The American Naturalist*, *148*(2), 255–274.

Venette, R. C. (2015). *Pest Risk Modelling and Mapping for Invasive Alien Species*. CABI.

Venette, R. C., Kriticos, D. J., Magarey, R. D., Koch, F. H., Baker, R. H., Worner, S. P., Raboteaux, N. N. G., McKenney, D. W., Dobesberger, E. J., & Yemshanov, D. (2010). Pest risk maps for invasive alien species: a roadmap for improvement. *BioScience*, *60*(5), 349–362.

Vercken, E., Kramer, A. M., Tobin, P. C., & Drake, J. M. (2011). Critical patch size generated by Allee effect in gypsy moth, *Lymantria dispar (L.)*. *Ecology Letters*, *14*(2), 179–186.

Vergni, D., Iannaccone, S., Berti, S., & Cencini, M. (2012). Invasions in heterogeneous habitats in the presence of advection. *Journal of Theoretical Biology*, *301*, 141–152.

Verhulst, P.-F. (1838). Notice sur la loi que la population poursuit dans son accroissement. *Correspondance Mathématique et Physique*, *10*, 113–121.

Vesanto, J., Himberg, J., Alhoniemi, E., & Parhankangas, J. (1999). Self-organizing map in Matlab: the SOM toolbox. In *Proceedings of the Matlab DSP conference*, vol. 99, (pp. 16–17). Helsinki: Laboratory of Information and Computer Science. `http://www.cis.hut/fi/projects/somtoolbox/documentation/somlag.shtml`.

Vilà, M., & Ibáñez, I. (2011). Plant invasions in the landscape. *Landscape Ecology*, *26*(4), 461–472.

Vilà, M., Pino, J., Font, X., & Bruelheide, H. (2007). Regional assessment of plant invasions across different habitat types. *Journal of Vegetation Science*, *18*(1), 35–42.

Visser, H., & De Nijs, T. (2006). The map comparison kit. *Environmental Modelling and Software*, *21*(3), 346–358.

Vitousek, P. M., D'Antonio, C. M., Loope, L. L., & Westbrooks, R. (1996). Biological invasions as global environmental change. *American Scientist*, *84*(5), 468–478.

Šímová, P., & Gdulová, K. (2012). Landscape indices behavior: a review of scale effects. *Applied Geography*, *34*, 385–394.

Walker, W. E., Harremoës, P., Rotmans, J., van der Sluijs, J. P., van Asselt, M. B. A., Janssen, P., & Krayer von Krauss, M. P. (2003). Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment*, *4*(1), 5–17.

Wang, L., & Jackson, D. A. (2014). Shaping up model transferability and generality of species distribution modeling for predicting invasions: implications from a study on *Bythotrephes longimanus*. *Biological Invasions*, *16*(10), 1–25.

Wang, M.-H., Kot, M., & Neubert, M. G. (2002). Integrodifference equations, Allee effects, and invasions. *Journal of Mathematical Biology*, *44*(2), 150–168.

Wang, X., & Cumming, S. G. (2011). Measuring landscape configuration with normalized metrics. *Landscape Ecology*, *26*(5), 723–736.

Wang, Z., Wu, J., Shang, H., & Cheng, J. (2011). Landscape connectivity shapes the spread pattern of the rice water weevil: a case study from Zhejiang, China. *Environmental Management*, *47*(2), 254–262.

Weinberger, H. F. (2002). On spreading speeds and traveling waves for growth and migration models in a periodic habitat. *Journal of Mathematical Biology*, *45*(6), 511–548.

Weinberger, H. F., Kawasaki, K., & Shigesada, N. (2008). Spreading speeds of spatially periodic integro-difference models for populations with nonmonotone recruitment functions. *Journal of mathematical biology*, *57*(3), 387–411.

Weisberg, M. (2006). Robustness analysis. *Philosophy of Science*, *73*(5), 730–742.

Wickham, J. D., & Rhtters, K. H. (1995). Sensitivity of landscape metrics to pixel size. *International Journal of Remote Sensing*, *16*(18), 3585–3594.

Wiegand, T., Jeltsch, F., Hanski, I., & Grimm, V. (2003). Using pattern-oriented modeling for revealing hidden information: a key for reconciling ecological theory and application. *Oikos*, *100*(2), 209–222.

Wiens, J. A. (1989). Spatial scaling in ecology. *Functional Ecology*, *3*(4), 385–397.

With, K. A. (2002). The landscape ecology of invasive spread. *Conservation Biology*, *16*(5), 1192–1203.

With, K. A. (2004). Assessing the risk of invasive spread in fragmented landscapes. *Risk Analysis*, *24*(4), 803–815.

With, K. A., Gardner, R. H., & Turner, M. G. (1997). Landscape connectivity and population distributions in heterogeneous environments. *Oikos*, *78*(1), 151–169.

With, K. A., & King, A. W. (1997). The use and misuse of neutral landscape models in ecology. *Oikos*, *79*(2), 219–229.

With, K. A., & King, A. W. (1999a). Dispersal success on fractal landscapes: a consequence of lacunarity thresholds. *Landscape Ecology*, *14*(1), 73–82.

With, K. A., & King, A. W. (1999b). Extinction thresholds for species in fractal landscapes. *Conservation Biology*, *13*(2), 314–326.

Wood, S. N. (2010). Statistical inference for noisy nonlinear ecological dynamic systems. *Nature*, *466*(7310), 1102–1104.

Woolcock, J. L., & Cousens, R. (2000). A mathematical analysis of factors affecting the rate of spread of patches of annual weeds in an arable field. *Weed Science*, *48*(1), 27–34.

Worner, S. (1994). Predicting the establishment of exotic pests in relation to climate. In J. L. Sharp, & G. J. Hallman (Eds.) *Quarantine treatments for pests of food plants*. Westview Press Boulder, CO.

Worner, S. P. (2002). Predicting the invasive potential of exotic insects. In G. J. Hallman, & C. P. Schwalbe (Eds.) *Invasive arthropods in agriculture: problems and solutions*, (pp. 119–137). Science Publishers, Inc. Enfield, New Hampshire.

Worner, S. P., Gevrey, M., Ikeda, T., Leday, G., Pitt, J., Schliebs, S., & Soltic, S. (2014). Ecological informatics for the prediction and management of invasive species. In N. K. Kasabov (Ed.) *Springer Handbook of Bio-Neuroinformatics*, (pp. 565–583). Springer.

Worner, S. P., Ikeda, T., Leday, G., & Joy, M. (2010). Surveillance tools for freshwater invertebrates. (p. 116 p.). Ministry of Agriculture and Forestry, New Zealand. MAF Biosecurity New Zealand Technical Paper No: 2010/21 prepared for MAFBNZ Operational Research.

Worner, S. P., Lankin, G., Lustig, A., Khandan, H. A. N., & Senay, S. D. (In Press). Being better than average: the application of computational intelligence in pest management and biosecurity. In R. M. Beresford, K. J. Froud, J. M. Kean, & S. P. Worner (Eds.) *The plant protection data toolbox*. New Zealand Plant Protection Society Inc., New Zealand: Proceedings of New Zealand Plant Protection Society Symposium.

Wu, J. (2004). Effects of changing scale on landscape pattern analysis: scaling relations. *Landscape Ecology*, *19*(2), 125–138.

Wu, J. (2013). Key concepts and research topics in landscape ecology revisited: 30 years after the Allerton park workshop. *Landscape Ecology*, *28*(1), 1–11.

Wu, J., Shen, W., Sun, W., & Tueller, P. T. (2002). Empirical patterns of the effects of changing scale on landscape metrics. *Landscape Ecology*, *17*(8), 761–782.

Ye, M., Meyer, P. D., & Neuman, S. P. (2008). On model selection criteria in multimodel analysis. *Water Resources Research*, *44*(3).

Yeates, A. G., Schooler, S. S., Garono, R. J., & Buckley, Y. M. (2012). Biological control as an invasion process: disturbance and propagule pressure affect the invasion success of *Lythrum salicaria* biological control agents. *Biological Invasions*, *14*(2), 255–271.

Yemshanov, D., Koch, F. H., Ben-Haim, Y., & Smith, W. D. (2010). Robustness of risk maps and survey networks to knowledge gaps about a new invasive pest. *Risk Analysis*, *30*(2), 261–276.

Zajac, Z. B. (2010). *Global sensitivity and uncertainty analysis of spatially distributed watershed models*. Ph.D. thesis, University of Florida.

Zhang, J., & Goodchild, M. F. (2002). *Uncertainty in geographical information*. CRC press. 266 pp.

Zhang, Q.-H., & Schlyter, F. (2004). Olfactory recognition and behavioural avoidance of angiosperm nonhost volatiles by conifer-inhabiting bark beetles. *Agricultural and Forest Entomology*, *6*(1), 1–20.

Zimmermann, H., Brandt, P., Fischer, J., Welk, E., & von Wehrden, H. (2014). The human release hypothesis for biological invasions: human activity as a determinant of the abundance of invasive plant species. *Version 2. F1000Research*, *3*(109).

Appendix A

Supplement to Chapter 2

A.1 Modelling dynamic range expansion of *P. brassicae*

A.1.1 Economy, biology and ecology of *P. brassicae*

P. brassicae, a butterfly originating from Europe and Asia, was first introduced in Nelson city, New Zealand in May 2010, where it is considered as Unwanted Organism under the Hazardous Substances and New Organisms Act 1996. In 2010, the Ministry of Primary Industry (MPI), New Zealand responded to *P. brassicae* incursion with a monitoring programme that aimed to slow its spread. In 2012, the Department of Conservation (DOC), New Zealand, initiated an eradication attempt due to the risk it posed to New Zealand native Brassicaceae, collecting presence records through both passive and active surveillance (Phillips *et al.*, 2014). By this time, *P. brassicae* was firmly established in Nelson area and spreading in a South-Western direction. The Department of Conservation (DOC), New Zealand, estimated the damage costs associated with *P. brassicae* and their control - if the species spread throughout New Zealand - to be approximately NZD 7.4 million/annum.

P. brassicae has two generations per year with four main life-stages: egg, larva (caterpillar), pupa, and adult. *P. brassicae* is naturally nomadic. It does not live in permanent colonies but breeds wherever suitable conditions are encountered, showing preferences for cultivated areas where species of Brassicaceae are cultivated and urban gardens. Under optimum conditions, enormous numbers of eggs can build up (laid upright in clusters of 40-100) which then explode outwards in strong migrations. Reason for migration in *P. brassicae* is unclear (Holland *et al.*, 2006) but it seems to be an obligate behaviour of the butterfly to leave its hatching site within 2–3 days (Blunck, 1954). *P. brassicae* exhibits both local and long-distance dispersal. Feltwell (1982) concluded that the average local dispersal distances differed from a mere 7 m/year to a worrying 350 m/year, reported as the average distance *P. brassicae* larvae cover while looking for pupation sites. Unlike long distance dispersal, there is no evidence that *P. brassicae* exhibits directional bias during local dispersal (Davies & Gilbert, 1985). Previous reports of long-distance, mass migrations revealed that *P. brassicae* tend to fly in one direction, showing a proportional (1:1) sex ratio during migratory long flights, reducing Allee effects on migrating groups (Spieth & Cordes, 2012). Topographic barrier can force migrating butterflies to drastically redirect their flight. Spieth & Cordes (2012) reported that mountains ranges may be crossed using mountain passes that provide access to the preferred flight direction, however coastal barriers hindering access pose a much greater threat to survival. Wind currents can assist insects in their long flights. Nevertheless *P. brassicae* has been observed flying upwind (Blunck, 1954; Roer, 1959), suggesting that wind is not necessarily the cause for *P. brassicae* long distance dispersal. Long-range movement is most often due to anthropogenic transportation, such as cars, cargo containers or host plants, of diapausing eggs and pupaes providing a second mode of spread (Feltwell, 1982).

A.1.2 Study area

The extent of the study covers the five administrative districts in the South Island, New Zealand, that were either in contact with- or nearby the locations invaded by *P. brassicae*. These districts were Buller, Tasman, Nelson City, Marlborough and Kaikoura (12, 466 sq. ha).

A.1.3 Parametrizing the local dispersal sub-model for *P. brassicae*

A Von Neumann shape with range = 1 was chosen to represent the uniform, local neighbourhood for spread of *P. brassicae* within one time step. We chose a cell resolution of 100 m to approximate the median distance of local movements of larvae and adults reviewed in Feltwell (1982). The initial dispersal site was set in a cell close to Nelson port which is suspected to be the site of *P. brassicae* introduction.

A.1.4 Parametrizing the long-distance dispersal sub-model for *P. brassicae*

Estimate of the median distance and average frequency of long-dispersal events were built from the dispersal history of *P. brassicae* in the United Kingdom for which, well referenced temporal data were found in Feltwell (1982), Heath *et al.* (1984) and the Global Biodiversity Information Facility (GBIF) database. First, a map of *P. brassicae* distribution in the United Kingdom Feltwell (1982) was scanned and rectified using the ArcGIS software according to the UTM Zone 30 projection information. The provisional version of the original atlas of Heath *et al.* (1984) described the sampling scheme used to collect the distribution data and indicated that each occurrence point represented a 10 km sq. grid on the map. Accordingly, a cell size of 10 km was used as the standard data processing resolution for dispersal parameter estimation. Second, the GBIF data were resampled using 10 km sq. intervals, and were projected into the British national grid coordinate system. The resampled GBIF *P. brassicae* occurrence data ($n = 505$) along with the points digitized from Heath *et al.* (1984) ($n = 351$) were used to estimate dispersal parameters for *P. brassicae.* Information on dates of records of *P. brassicae* occurrences from Feltwell (1982) and GBIF database was used to characterise the 856 points with a year of introduction into their respective localities.

The Cauchy distribution, commonly used to account for the fat-tailed characteristics of the distribution of rare long distance events (Kot *et al.*, 1996; Cain *et al.*, 2000; Higgins *et al.*, 2000), was chosen to approximate the long-distance dispersal behaviour of *P. brassicae.* The Cauchy probability density function is given below.

$$f(x|x_0, \gamma) = \frac{1}{\pi}\left[\frac{\gamma}{(x - x_0)^2 + \gamma^2}\right] \tag{A.1}$$

where x_0 is the location parameter, specifying the location of the peak of the distribution and is used as a proxy of the median distance of long-distance dispersal events, and γ is the scale parameter of the Cauchy distribution.

The median distance of long-distance dispersal events for *P. brassicae* was parametrized by producing the nearest neighbour distance vector between occurrence points for each time period as described in Robinet *et al.* (2009). A random uniform noise within a range of [-146 m, 146 m] was applied to the nearest neighbour distance vector extracted from the occurrence points with several replications ($n = 1000$) to account for uncertainty from the digitized points (Pitt *et al.*, 2011). Parameters were estimated by fitting the noised distance data ($n = 1000$) to the Cauchy distribution using the maximum likelihood estimator with the trust-region-reflective optimization algorithm (Conn *et al.*, 2000) using MatLab software. The mean and standard deviation of the parameter estimates over the 1000 replicates was used to assess the stability of the estimated Cauchy location and scale parameters.

The frequency of the dispersal events was estimated from the United Kingdom temporal occurrence data. The historical *P. brassicae* presence data was first classed into time periods. The ratio of the number of new *P. brassicae* sites to the number of existing sites was calculated for each period. The resulting vectors of ratios, R, reflect the minimum number of dispersal events that needed to be generated from each cell to achieve the number of occupied cells in the next time period. These vector were then weighted by the number of sites for each period as described by Pitt *et al.* (2011) as following:

$$R = \frac{\sum_{t=t_0+1}^{T} N_t - N_{t-1}}{\sum_{t=t_0+1}^{T} N_{t-1}} \tag{A.2}$$

R is the vector of average weighted ratios calculated by dividing the number of newly invaded cells to the number of existing cells, t_0 is the first year with occurrence data, T is the last year, and N_t is the number of cells that are occupied within time t (Pitt *et al.*, 2011).

The resulting vector was fitted to the Poisson distribution, where the expected mean frequency of the distribution (λ) was estimated from Equation A.3.

$$f(k|\lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \tag{A.3}$$

k is a vector of discrete integers $[0, 1, 2, \dots]$ and λ is the expected mean of the Poisson distribution, with $\lambda > 0$ condition.

A.1.5 Building a survival layer for *P. brassicae*

Two survival layers were developed to investigate the effect of landscape on invasive species spread. The first survival layer ($Surv_1$) included four data sources: climate suitability, degree days, land cover, and high resolution remotely sensed data. The second survival layer ($Surv_2$) included all components used in $Surv_1$, except the high resolution remotely sensed data. Brief explanations of these datasets are given below.

The dataset included 35 bioclimatic variables downloaded from the CLIMOND website (Kriticos *et al.*, 2012) and four topographic variables derived from the SRTM global digital elevation model dataset (NASA-GSFC, 2000) accessed through the WORLDCLIM data portal (`http://www.worldclim.org/`), were used to generate the climate suitability layer. First, the selection of relevant environmental variables was performed using a random forest classifier. Second, the multi-model framework developed by Worner *et al.* (2014) was used to predict the potential distribution of *P. brassicae* for the study area. The data were projected to New Zealand Geographic Datum 2000 coordinate system and re-sampled at 100 m, which was the raster resolution set for the dispersal model.

A 30 years daily temperature data obtained from NIWA (`https://www.niwa.co.nz`) was used to generate the growing degree days (GDD) layer for the study area. The lower temperature threshold for the development of a *P. brassicae* pupae, i.e. 10 °C (Kean & Phillips, 2013) was selected as a base temperature to generate the GDD value for two main reasons: 1) the pupae had the highest temperature development threshold of all the life-stages (Kean & Phillips, 2013), therefore it is impossible for *P. brassicae* to complete its life cycle unless temperatures exceed this threshold, 2) the pupae stage is a stationary stage therefore do not move to more suitable locations. To calculate the GDD value, we used the Barlow method (Barlow & Dixon, 1980) as it is expected to give the least error when validated with real data Kean (2013). The degree day data calculated for 509 points was interpolated into a raster surface using spline interpolation.

The New Zealand Land Cover Dataset, LCBD2 (`Koordinates.com`) was used to generate the land cover data layer for the study area. The land cover data was produced by Landcare Research based on SPOT imagery (resolution 15 m) and the pan-sharpened Landsat 7 ETM+ imagery (resolution 15 m). The dataset has 43 types of land covers. The probability scheme used in Senay (2014) to assign survival percentages according to

Land cover name	**ID**	**Description**
Built-up area (urban areas or settlements)	1	Very high suitability
Orchard and other perennial crops	1	Very high suitability
Short-rotation cropland	1	Very high suitability
Urban parkland/ open space	1	Very high suitability
Alpine grass/ herbfield	2	High suitability
Depleted tussock grassland	2	High suitability
High producing exotic grassland	2	High suitability
Low producing grassland	2	High suitability
Tall tussock grassland	2	High suitability
Broadleaved indigenous hardwood	3	Moderate suitability
Deciduous hardwoods	3	Moderate suitability
Manuka or kanuka	3	Moderate suitability
Matagouri	3	Moderate suitability
Gorse and broom	4	Low suitability
Herbaceous freshwater vegetation	4	Low suitability
Herbaceous saline vegetation	4	Low suitability
Indigenous forest	4	Low suitability
Major shelterbelts	4	Low suitability
Mixed exotic shrubland	4	Low suitability
Sub alpine shrubland	4	Low suitability
Vineyard	4	Low suitability
Afforestation (imaged post LCDB 1)	5	Very low suitability
Fernland	5	Very low suitability
Flaxland	5	Very low suitability
Grey scrub	5	Very low suitability
Other exotic forest	5	Very low suitability
Pine forest - closed canopy	5	Very low suitability
Pine forest - open canopy	5	Very low suitability
Alpine gravel and rock	6	Not suitable
Coastal sand and gravel	6	Not suitable
Dump	6	Not suitable
Estuarine open water	6	Not suitable
Lake and pond	6	Not suitable
Landslide	6	Not suitable
Permanent snow and ice	6	Not suitable
River	6	Not suitable
River and lakeshore gravel and rock	6	Not suitable
Surface mine	6	Not suitable
Transport infrastructure	6	Not suitable

Table A.1: Land-cover re-classification schemes to characterize habitat suitability for *P. brassicae*.

the land cover classes were applied to grouped and re-classed the land cover types into six classes according to their suitability for *P. brassicae* (Table A.1). The re-classed ESRI polygon dataset was then converted to raster using 100 m resolution.

P. brassicae is know to be spreading in Nelson city by breeding in home gardens that are available in residential blocks. The land-cover layer, that classified urban area as built-up and homogeneous landscape, is likely to over-estimate the initial dispersal of *P. brassicae*. High resolution remotely sensed data based on SPOT33 Maps 2.5 m resolution satellite imagery was used to characterise the geographical detail in urban areas of the survival layer, $Surv_1$. A single layer labelled 'man-made structures' was generated from the satellite image to update the homogeneous 'built-up' class in the land cover data.

The landscape components were rescaled to produce a $0-1$ survival probability layer. The land cover data was used as the base on which extra values from the climatic, GDD layers were added. The values in the land cover data layer were rescaled between the ranges $1 - 0$, where high local variation is given low survival probability. For the growing degree day dataset the layer was rescaled between $0.5 - 1$. The minimum probability of 0.5 was given to the GDD dataset as it requires 471.6 cumulative growing degree days for *P. brassicae* to complete its life cycle at least once. The first survival layer ($Surv_1$) was also recoded with the man-made structures data identified from the SPOT image classification where all areas that were overlaid by the man-made structure data were set to a survival of zero probability (Figure A.1).

A.1.6 Model evaluation and performance measures

The New Zealand data set of *P. brassicae* detections and absences obtained from the Department of Conservation (Phillips *et al.*, 2014) were used to compare the first three years of the dispersal occupancy envelopes of both $Surv_1$ and $Surv_2$ dispersal model outputs with field data. We used cross tabulated predicated and observed values to form a confusion matrix such that the diagonal elements of the matrix correspond to the correct predications of absence/presence locations, while the off-diagonal elements represent incorrect predictions (Allouche *et al.*, 2006; Fielding & Bell, 1997). From such a matrix a three performance measures were used to estimate the mean performance of the two dispersal models:

- the accuracy index, $\frac{TP+TN}{TP+TN+FP+FN}$, is the proportion of correctly predicted presence and absence locations to the total number of locations
- the sensitivity index, $\frac{TP}{TP+FN}$, proportion of correctly predicted presence locations to the total number of presence locations
- the specificity index, $\frac{TN}{TN+FP}$, proportion of correctly predicted absence locations to the total number of absence locations

where, TP stands for true positive, TN for true negative, FP for false positive and FN for False negative. A thorough review of these methods can be found in Freeman & Moisen (2008) and Senay (2014).

Parameters	**Abbreviated code**	**Parameters value**
Population sub-model		
Presence/absence model		
Local dispersal sub-model (uniform)		
Mean distance of local dispersal events	R	100 meters
Long-distance dispersal sub-model (Cauchy distribution)		
Median distance of long dispersal events	x_0	154.22 meters
Cauchy scale	λ	99.43
Frequency of long dispersal events	f	0.48
Propagule pressure		
Propagule size	nbp	1 presence cell
Landscape structure		
Survival layer including high resolution remotely sensed data	$Surv_1$	
Survival layer excluding high resolution remotely sensed data	$Surv_2$	

Table A.2: Baseline values for parameters of a spatially-explicit, individual based model of *P. brassicae*

A.1.7 Parameter summary for *P. brassicae*

The MDiG parameters used in this study are given in Table A.2. Sixteen years of simulations were undertaken representing dispersal from the year 2010 to 2025. The simulation was replicated 1000 times to account for dispersal stochasticity. Replicated maps were merged into an occupancy envelope (an average maps for each time steps) using three threshold values [5%, 10%, 50%] that corresponded to the number of times a cell was occupied during dispersal.

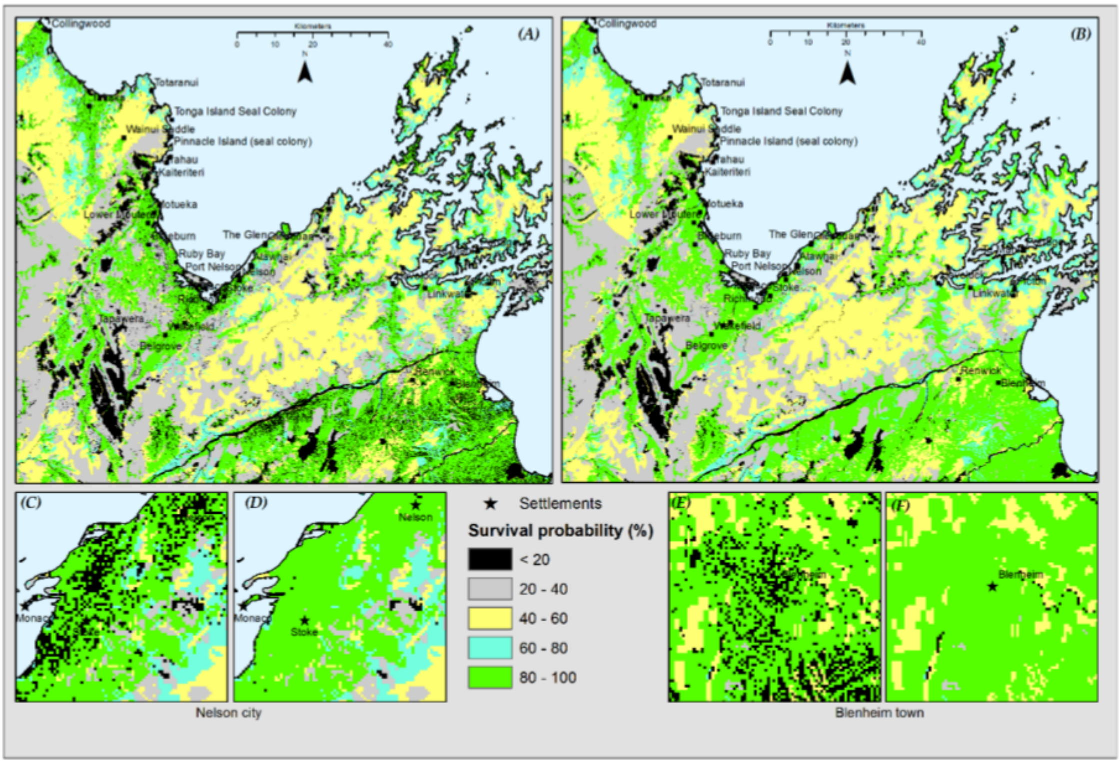

Figure A.1: From Senay (2014), with permission. Survival layers used in *P. brassicae* dispersal model. (A) Survival layer $Surv_1$ including climate, land cover, growing degree day and high resolution man-made structure layer to increased the detail of urban areas. (B) Survival layer $Surv_2$ composed of all suitable layers included in $Surv_1$ except the man-made structure layer. (C) and (D) show a zoom on Nelson city for $Surv_1$ and $Surv_2$ respectively, and (E) and (F) show a zoom on Blenheim town for $Surv_1$ and $Surv_2$.

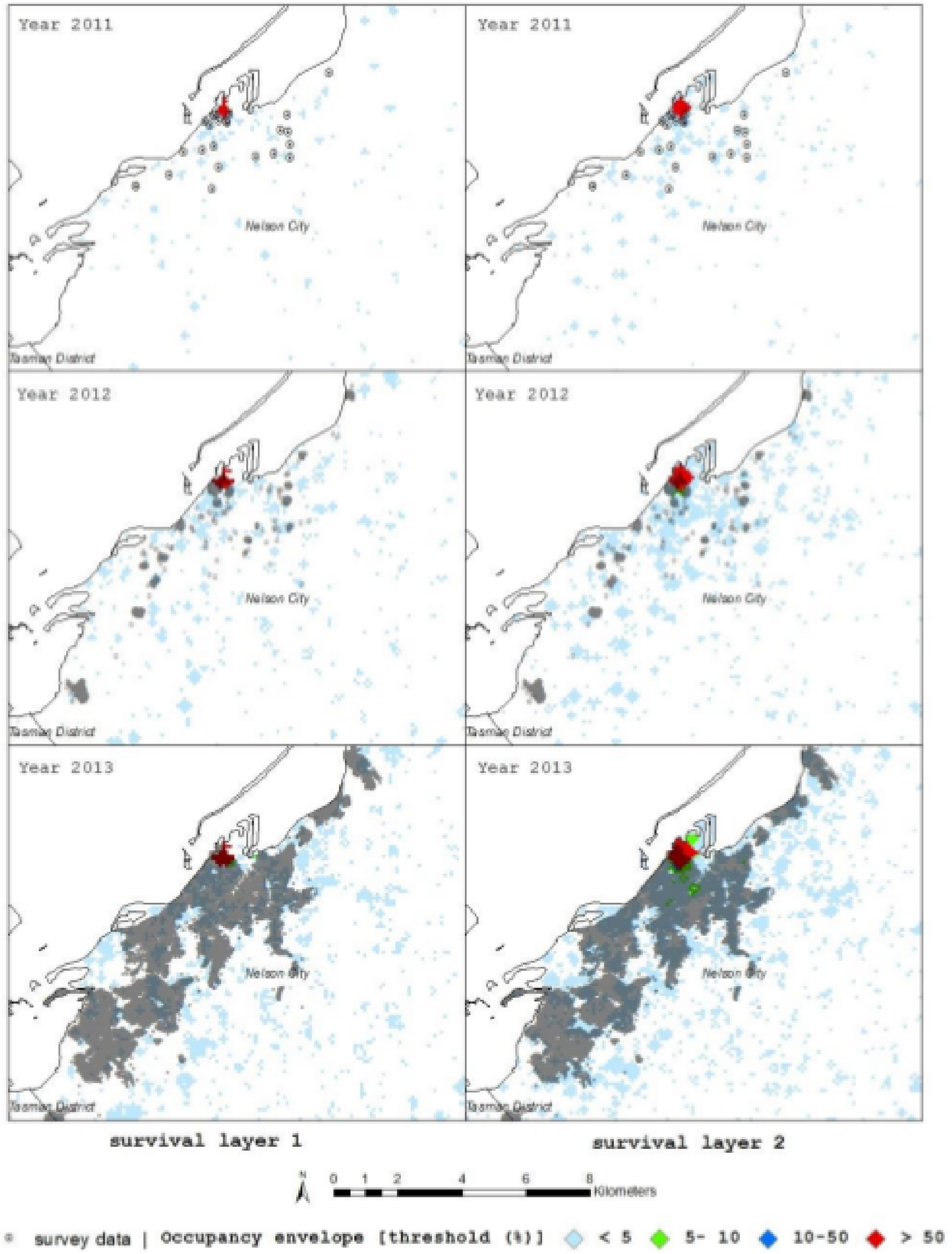

Figure A.2: From Senay (2014), with permission. Dispersal maps overlaid with *P. brassicae* presences from field survey data

A.2 Modelling dynamic range expansion of *L. dispar*

A.2.1 Economy, biology and ecology of *L. dispar*

L. dispar, commonly know as the Asian Gypsy moth, is considered internationally to be among the most serious of all forest insect pests (Liebhold *et al.*, 1992). Their presence can destroy the aesthetic beauty of an area by causing large-scale defoliation and occasionally intense tree mortality, and covering the area with their waste products and silk. The species originally evolved in the temperate forests of Europe and Asia but was accidentally introduced to North America outside of Boston, Massachusetts, USA, in 1869. Since then it has subsequently invaded much of the susceptible forest of north-eastern North America, from Ontario to North Carolina and Nova Scotia to Wisconsin (Liebhold & Mastro, 1989; Morin *et al.*, 2005). An extensive containment programme - named *Slow-the-Spread*, for which more than $194 million was spent on management and monitoring between 1985 and 2004 alone (Johnson *et al.*, 2006; Mayo *et al.*, 2003), has dramatically reduce the spread of the species (Tobin & Blackburn, 2007). However, gypsy moth invasion continues and, still threatens to establish throughout most North America (Tobin *et al.*, 2015). The spread of the gypsy moth across eastern North America is, one of the most thoroughly studied biological invasion, providing a unique opportunity to explore spatio-temporal variability in rates of spread (Johnson *et al.*, 2006). Each year, around 150,000 pheromone-baited traps are deployed in the transition zone that separate well established population and areas in which the gypsy moth is still absent (Tobin & Blackburn, 2007). In this transition zone, populations are low and somewhat discontinuous. Traps are typically placed at 2-, 3- or 8-km inter-trap distances up to 150 km from the leading edge of the gypsy moth distribution, which has been shown to be sufficient to detect low-density isolated colonies (Tobin & Blackburn, 2007).

The gypsy moth has one generation per year with four main life-stages: egg, larva (caterpillar), pupa, and adult. In spring the eggs hatch into larvae, which crawl up to the tree tops, then suspend themselves on silk threads and are passively dispersed by the wind (Liebhold *et al.*, 1992). It is at this stage that they feed on more than 300 hundreds species of trees and shrub. 4–6 weeks later in early summer, they enter a pupal stage. 10 to 14 days later, adult gypsy moths emerge and are present in July and August. Adult female gypsy moths are flightless, and ballooning of 1st instars usually occurs only over short distances. Soon after emergence, adult females mate and lay their eggs, which overwinter. Egg masses can be found on the branches or trunk of trees, fences, buildings, and vehicles and contains around 100-1000 eggs (Liebhold *et al.*, 1992).

Long-range movement is most often due to anthropogenic transportation of life-stages, providing a second mode of spread (Liebhold & Tobin, 2010). This leads to stratified diffusion. Whilst natural dispersal is limited to early instars, artificial dispersal affects all life-stages but most frequently involves egg masses. Long distance dispersal leads to the formation of isolated colonies ahead of the initially infested area, which may grow and coalesce thereby increasing the rate of spread. By analysing available historical country level quarantine data on gypsy moth invasion, (Liebhold *et al.*, 1992) concluded that the spread rates differed, throughout the past century, from a mere 2.82 km/year to a worrying 20.78 km/year. More recently, the value of 0.003 km^2/generation was referred to by several authors (e.g. Liebhold & Tobin, 2006; Robinet *et al.*, 2008) sparking an inconsistency in terms of units. Finally, Tobin *et al.* (2007), using various spread rate estimation techniques, came up with even a broader range of spread rates as 2.6–28.6 km/year.

Another factor shown to affect the gypsy moth spread is the Allee effect (e.g. Liebhold & Bascompte, 2003; Vercken *et al.*, 2011). New isolated colonies of gypsy moth ahead of the initially infested area, are of low abundance and highly prone to Allee effects and extinction (Liebhold & Bascompte, 2003). However, estimates of the Allee threshold are usually approximate: high observation error and demographic stochasticities are inevitable consequences of low abundance (Johnson *et al.*, 2006; Vercken *et al.*, 2011). The current gypsy moth containment programme offers an exception because of its extensive grids of pheromone-baited traps, which are sensitive to extremely low moth densities along the invasion front. In Wisconsin, Allee effect threshold value was estimated at 2.2 moths/trap (Vercken *et al.*, 2011). A much higher value of 20.7 moths/trap was established in West Virginia and North Carolina. Johnson *et al.* (2006) provided overall estimates of 17 moths/trap for the Allee threshold.

The programme also allows to estimate habitat carrying capacity. Tobin *et al.* (2007) pinpoint the carrying capacity at around 283 moths/trap in Wisconsin and 673 moths/trap in West Virginia and North Carolina. In earlier studies, Dwyer & Elkinton (1993) estimated the Allee threshold as 500 larvae/m^2 and Sharov & Liebhold (1998) as 200,000 egg masses/km^2. The latest values were difficult to compare with the estimates mentioned above as it refers to a different life-stage. Johnson *et al.* (2006) provided overall estimates of 687 moths/trap for the carrying capacity. In a similar study, having analysed pheromone trapping data from a large-scale field study in Washington, Liebhold & Bascompte (2003) estimated the Allee threshold as 106.7 moths/colony, which is much higher than the estimate by Johnson *et al.* (2006).

Propagule pressure has long be suggested as one of the most important factor for determining both the scale of invasion extent and impact. Although the US Department of Agriculture's Animal and Plant Health Inspection Service (APHIS) does not have a regulation prohibiting the entry of vessels that are high risk for *L. dispar* infestation, its Plant Pest and Quarantine (PPQ) division has requested that the shipping industry not bring ships into US ports that have been in Far East Russian ports between July 15 and September 30 of the previous year or high-risk Japanese ports during the high-risk hatching period (i.e., when the propagule number of *L. dispar* is likely to be highest) (Reaser *et al.*, 2008). In order to estimate the propagule pressure, a list of all phytophagous insect species intercepted from 1993 to 1999 at the New Zealand border was complied from interception data held by the Ministry of Agriculture and Forestry, New Zealand (Peacock & Worner, 2008). From this list, 238 interceptions corresponded to the discovery of gypsy moth eggs/larvae/pupae in between August 1998 and March 2002. Most of the eggs/larvae/pupae were found at the border security in vehicles (tyres) imported from Japan. Among the 238 interceptions, 37 introductions consisted of the introduction of alive individuals (9 larvae, 3 pupae and 97 alive egg mass). On average 3 to 4 introductions were reported each year (min-max: 1-25). More than half of the propagule consisted of 2 egg mass (100-1000 eggs).

Prof. A. Liebhold suggested that the considerable variation reported for Allee thresholds, population dynamics and dispersal abilities can reflect geographical variation in the habitat that affects growth rates, dispersal rates and carrying capacity (personal communication). Dispersal probably is important because it affects the year-to-year vari-

Parameter name	**Parameter estimates**	**References**
Dispersal abilities, λ	0.003 km^2 0.1 ha/year 2.82 - 20.78 km/year 2.6 - 28.6 km/year	Liebhold & Tobin (2006) Robinet *et al.* (2008) Liebhold *et al.* (1992) Johnson *et al.* (2006)
Carrying capacity, K	283 moths/trap 673 moths/trap 687 moths/trap 500 larvae/m^2 200,000 egg masses/km^2	Tobin *et al.* (2007) Tobin *et al.* (2007) Johnson *et al.* (2006) Dwyer & Elkinton (1993) Sharov & Liebhold (1998)
Allee threshold, C	2.2 moths/trap 20.7 moths/trap 17 moths/trap 106.7 moths/colony	Tobin *et al.* (2007) Tobin *et al.* (2007) Johnson *et al.* (2006) Liebhold & Bascompte (2003)
Intrinsic growth rate, r	100-1000 eggs/female 700 egg/female 1.65-4.6 per capacity growth	Petrovskii & McKay (2010) Tobin *et al.* (2007) Johnson *et al.* (2006)
Propagule pressure (size and number)	1-25 introductions/years (median 3) 1-9 eggs mass intercepted (median 2)	NZ interception data (unpublished) NZ interception data (unpublished)

Table A.3: Review of parameter estimates for a population model of gypsy moth, *L. dispar*

ation in trap capture. Gypsy moth females are flightless, so dispersal is only by males. However, at the expanding population front, population growth is limited by mate-finding and thus the local abundance of males limits mating and thus population growth. For analysing the spread of the Asian gypsy moth, Prof. A. Liebhold further suggested to use the range of value given in Johnson *et al.* (2006). These parameters were estimated using the data from the expanding gypsy moth population front - the so-called transition zone. Most of the other values were estimated from the state of Washington which is far from the gypsy moth invasion front, but where the insect is occasionally accidentally introduced and forms isolated colonies that must be eradicated.

A.2.2 Building a survival layer for *L. dispar*

We used the computer program Qrule 4.2 to generate binary (suitable, unsuitable) landscapes, in which fragmentation (measured as the degree of spatial autocorrelation) and proportion of suitable habitat cover can be systematically and independently controlled (Gardner & Urban, 2007). Qrule uses a midpoint displacement algorithm (Saupe, 1988) to generate multi-fractal maps in which the degree of spatial autocorrelation among adjacent cells (H) can be controlled. We generated landscape across a three-step gradient in spatial autocorrelation ($H = 0.3, 0.5, 0.7$) and a three-step gradient in the proportion of suitable habitat cover ($P = 35, 50, 75$), with 10 replicate landscapes for each factor combinations. The extent of the study covers 128 *times* 128 raster grid cells (13,384 sq.).

Each sample landscape was characterized by two commonly used landscape metrics generated by the computer program FRAGSTATS 4.2 (McGarigal *et al.*, 2012). The percentage of suitable habitat cover (PLAND) was used to quantify the proportional abundance of each patch type in the landscape (measure of habitat composition), while the connectance index (CONNECT) was used to measure the connectivity between suitable patches. In this study, the connectance was calculated using the average Euclidean distance from cell centre to cell centre.

A.2.3 Parametrizing the local dispersal sub-model for *L. dispar*

A Von Neumann shape with range = 1 was chosen to represent the uniform, local neighbourhood for spread of *L. dispar* within one time step. We chose a cell resolution of 10 km/year to approximate the median distance of local movements of larvae and adults as shown in in Johnson *et al.* (2006) and (Liebhold *et al.*, 1992).

A.2.4 Parametrizing the long-distance dispersal sub-model for *L. dispar*

Long-distance dispersal events were approximated by a Cauchy probability distribution. Each colony (occupied cell) was assumed to give rise to a Poisson number of offspring colonies that were initiated with a random fraction of the parent's abundance (stochastic long-distance jumps). Based on the study of Johnson *et al.* (2006), we fixed the median of long-dispersal distance at 50 km and frequency of long dispersal at 0.01.

A.2.5 Parametrizing the population sub-model for *L. dispar*

Following Johnson *et al.* (2006) and Liebhold & Bascompte (2003), the abundance of gypsy moth was approximated by a deterministic Allee logistic growth model:

$$N_{t+1} = N_t \exp\left[r\left(1 - \frac{N_t}{K}\right)\left(\frac{N_t - C}{K}\right)\right] \tag{A.4}$$

where N is the number of individuals at time t, C is the Allee threshold, r the intrinsic growth rate and K the carrying capacity. The values of these parameters were based on previous estimates reported in Johnson *et al.* (2006) and Liebhold & Bascompte (2003) to fit release -capture data collected from 1988 to 2004.

References	**Intrinsic growth rate**, r **(per capacity growth)**	**Allee threshold**, C **(number of individuals)**	**Carrying capacity**, K **(number of individuals)**
Liebhold & Bascompte (2003)	$r * C/K = 1.740 \pm 0.122$	106.7	$r/K = 0.0163 \pm 0.0163$
Johnson *et al.* (2006)	1.5	0.01 (virtually no Allee effect) 2 (high Allee threshold)	100

Table A.4: Estimates of intrinsic growth rate, Allee threshold and carrying capacity for a population model for gypsy moth, *L. dispar*

A.2.6 Parameter summary for *L. dispar*

The MDiG parameters used in this study are given in Table A.5. The analysis was approached from a case-specific view point, parametrising a population sub-model based on detailed demographic and dispersal attributes of the well studied European gypsy moth, *L. dispar*. In this way, our model had sufficient biological details to reproduce realistic parameter ranges. MDiG was then used to explore four different scenarios of demographic and dispersal behaviours in changing environments (TableTable A.5): a specie with 1) slow reproducing and short dispersal abilities, 2) slow reproducing and long dispersal abilities, 3) fast reproducing and short dispersal abilities and, 4) fast reproducing and short dispersal abilities. The main reason for doing this was to evaluate how changes in the intrinsic growth rate, median distance of long-distance dispersal events and landscape structure affect the population density – d (number of individuals per raster cells) and the rate of spread – ROS (number of new cells occupied per simulation run). Thirty years of simulations were undertaken representing dispersal, one model time step represents one year. Simulations were replicated 10 times to account for dispersal stochasticity.

Parameters	**Abbreviated code**	**Parameters value**
Population sub-model (logistic model wit Allee effect - Equation A.4)		
Growth rate	r	0.815 (slow) per capacity growth 1.223 (fast) per capacity growth
Allee threshold	C	2 individuals per raster cell
Carrying capacity	K	50 individuals per raster cell
Local dispersal sub-model (uniform)		
Mean distance of local dispersal events	R	1 raster cell (10 km)
Long-distance dispersal sub-model (Cauchy distribution)		
Median distance of long dispersal events	λ	3 (short) raster cells 5 (long) raster cells
Frequency of long dispersal events	f	0.05
Propagule pressure		
Propagule size	nbp	5 individuals
Landscape structure (binary landscapes)		
Spatial autocorrelation (fragmentation)	H	0, 0.5, 1
Percentage of suitable habitat cover	P	25, 50, 75

Table A.5: Baseline value of MDiG parameters for modelling the spread of gypsy moth, *L.dispar*, in computer generated landscapes.

R

R

Appendix B

Supplement to Chapter 5

B.1 Partial dependence plots for population abundance and rate of spread

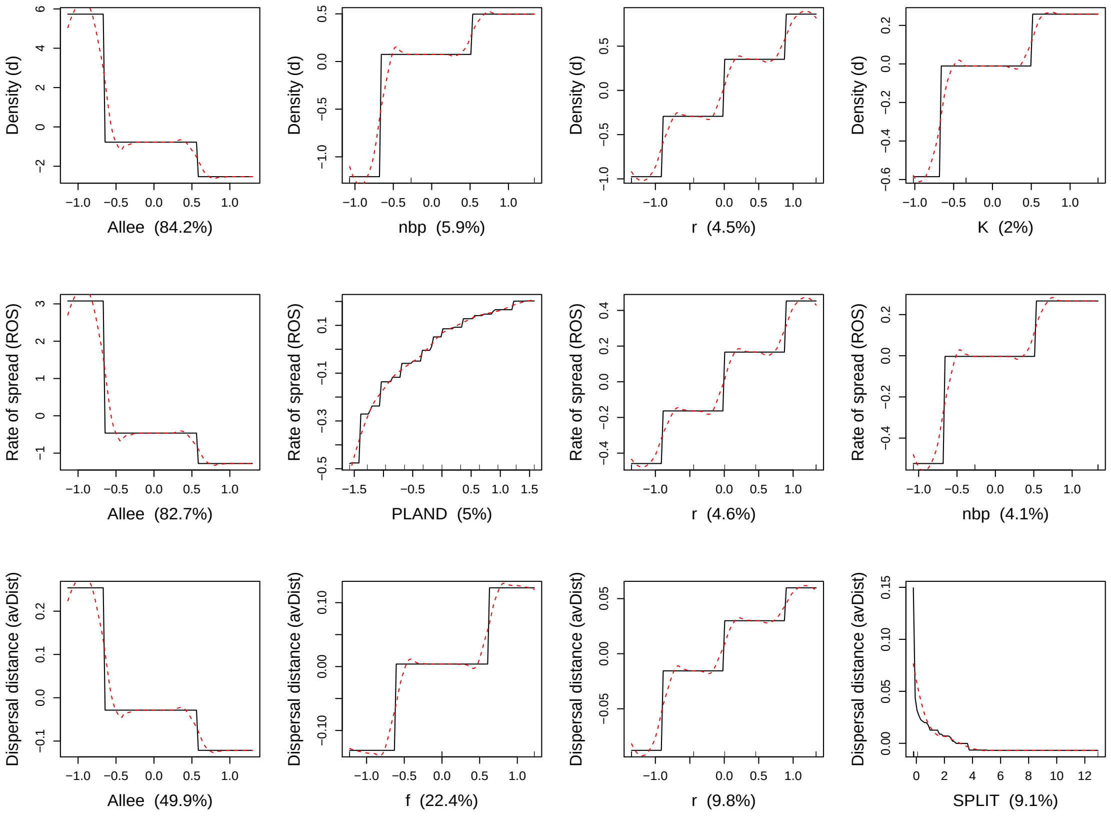

Figure B.1: Partial dependence plot showing the marginal effect of the three most important variables (centred and normalized) that determine the log transformed density (d), rate of of spread (ROS) and average dispersal distance (avDist) of invasive insects.

B.2 Variable importance for simulations including an Allee effect

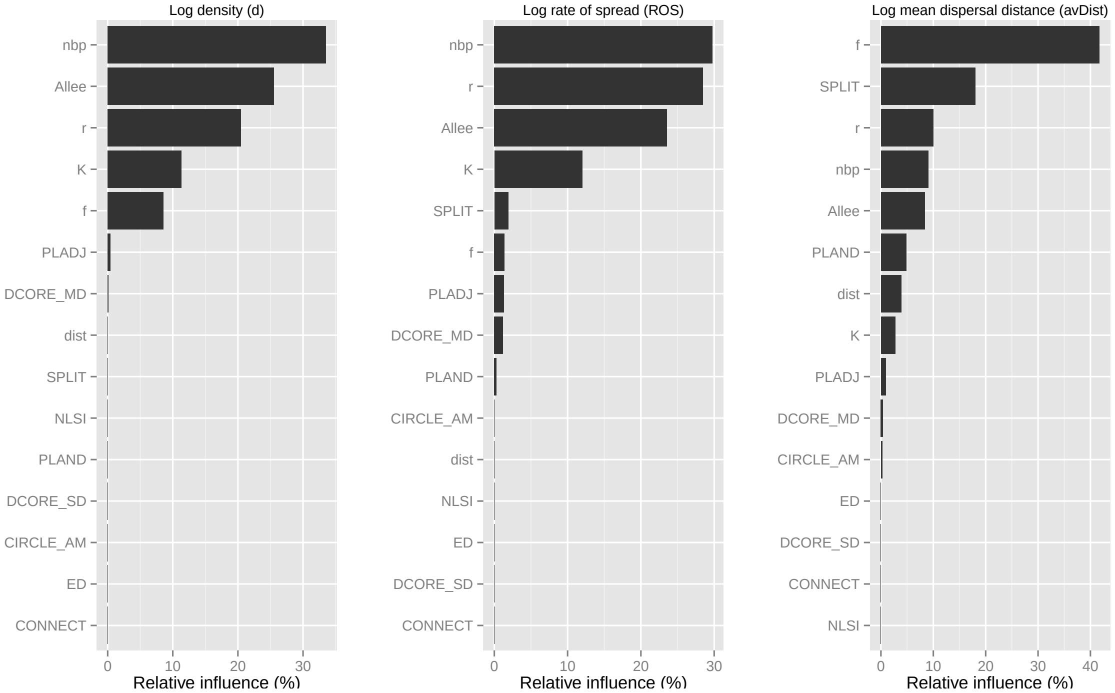

Figure B.2: Relative influence (%) of propagule pressure, demography, dispersal and landscape structure on log transformed density (*d*), rate of of spread (ROS) and average dispersal distance (avDist) for simulations with an Allee effect).

B.3 Variable interactions for simulations including an Allee effect

Rank	**Interaction terms**	**Int. strength**
Population density (d)		
1	Intrinsic rate of increase (r) × Allee threshold (Allee)	94.13
2	Propagule size (nbp) × Carrying capacity (K)	66.68
3	Propagule size (nbp) × Allee threshold (Allee)	46.89
4	Intrinsic rate of increase (r) × Propagule size (nbp)	26.89
5	Carrying capacity (K) × Allee threshold (Allee)	26.30
Rate of spread (ROS)		
1	Intrinsic rate of increase (r) × Allee threshold (Allee)	28.58
2	Intrinsic rate of increase (r) × Propagule size (nbp)	23.64
3	Carrying capacity (K) × Propagule size (nbp)	23.25
4	Carrying capacity (K) × Intrinsic rate of increase (r)	21.65
5	Carrying capacity (K) × Allee threshold (Allee)	1.71
Average dispersal distance (avDist)		
1	Intrinsic rate of increase (r) × Allee threshold (Allee)	0.44
2	Splitting Index (SPLIT) × Frequency of long-dispersal events (f)	0.37
3	Propagule size (nbp) × Intrinsic rate of increase (r)	0.33
4	Frequency of long-dispersal events (f) × Intrinsic rate of increase (r)	0.17
5	Carrying capacity (K) × Allee threshold (Allee)	0.17

Table B.1: Interaction strength between the most important parameters assessed in the non-interaction models log-transformed population density (d), rate of of spread (ROS) and average dispersal distance (avDist) when only simulations accounting for an Allee effect were considered. Interaction strength (Int. strength) is quantified using the residual-base methods outlined in Elith *et al.* (2009).

B.4 Pair-wise interactions effects on population abundance

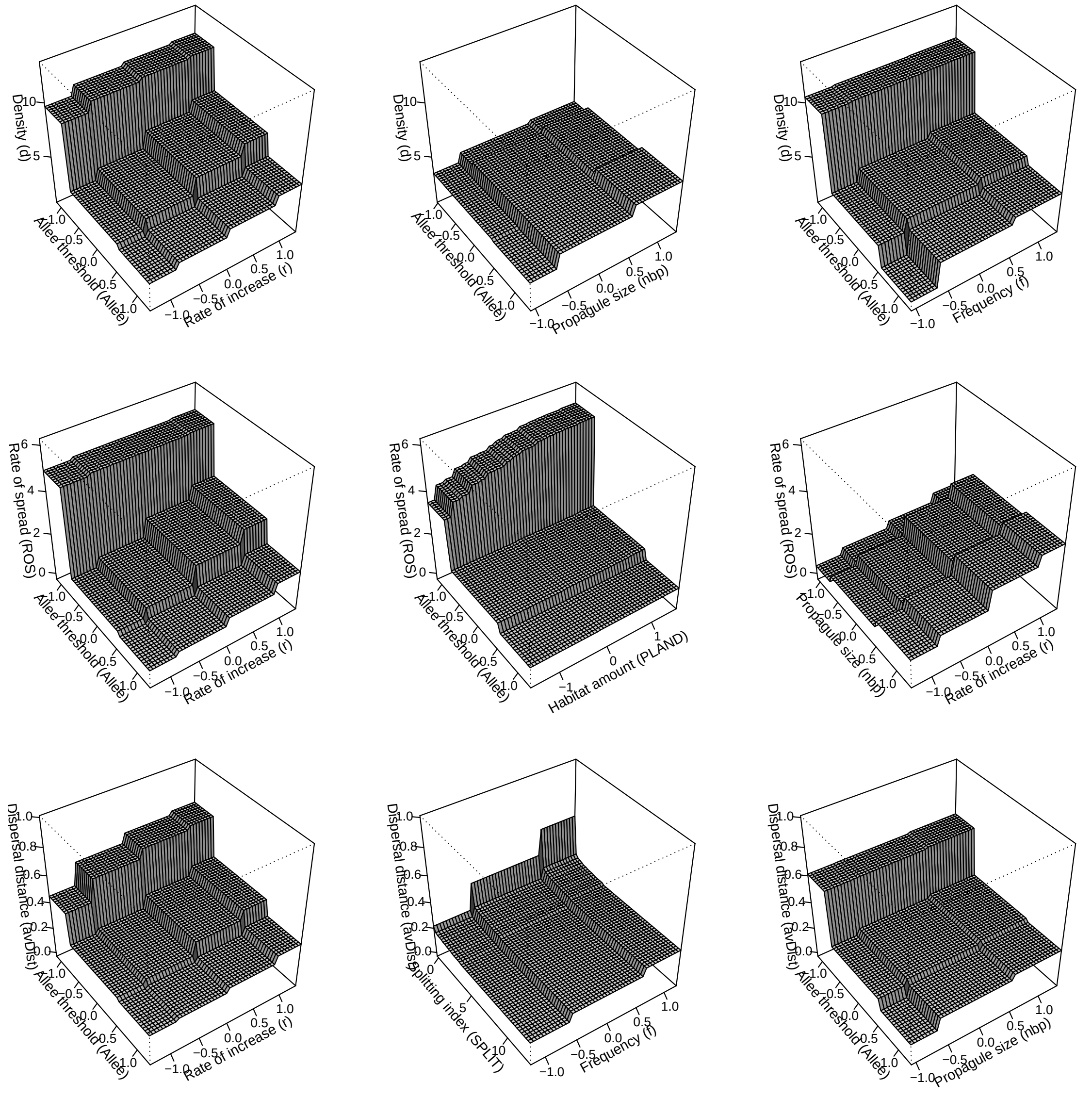

Figure B.3: Interaction plots showing the interactive effect of the three most important variables (centred and normalized) for determining the log transformed density (d), rate of of spread (ROS) and average dispersal distance (avDist) of invasive insects.

Appendix C

Supplement to Chapter 6

C.1 Potential sources of uncertainties in invasive species establishment and spread modelling

Source of uncertainty	**Description**	**References**
1. Knowledge uncertainties (Epistemic or Reducible)		
1.1 Process understanding *(related to structure uncertainty)*		
Disciplinary history and attendant form of available scientific knowledge (incomplete understanding of species' ecology or inability to reflect ecological complexity)	It reflects on a lack of knowledge of which factors constrain the distribution of species. Examples include: missing environmental covariate, physiological requirements at different life stages, knowledge about dispersal behaviour and pathways, knowledge about the relative species and biotic interaction	Ascough *et al.* (2008); Gould *et al.* (2014)
Temporal variation in species' ecology	Identifying factors that constrain the distribution of species is further complicated by temporal heterogeneity: development of non-analogous environmental conditions, altered outcome of species interactions, adaptation and evolutionary change, phenotypic plasticity, niche shifts	Gould *et al.* (2014)
Spatial variation in species' ecology	Identifying factors that constrain the distribution of species is further complicated by spatial heterogeneity: population specific local optima and variation in limiting factors across species range	Gould *et al.* (2014)
Spatial and temporal scales at which data knowledge applies	Spatial and temporal mismatch between input data and species' ecology	Ascough *et al.* (2008); Gould *et al.* (2014)
Precision	Qualitative versus quantitative nature of understanding across different temporal and spatial scales and level of organization	Ascough *et al.* (2008)
Availability of data to construct, calibrate and test predictive models	A major issue is related to a lack of false positive (absence occurrence data) that lead to inaccurate identification of the attributes of unsuitable sites	Ascough *et al.* (2008); Gould *et al.* (2014); McGeoch *et al.* (2012)
1.2 Model uncertainty **1.2.1 Parametric/Data uncertainty** ***Species occurrence data and listening***		
Human error (reliability)	Erroneous information in lists and databases, resulting from error entry	McGeoch *et al.* (2012)
Incomplete information (completeness)	Searches of data sources are not comprehensive, resulting in incomplete information	Gould *et al.* (2014); McGeoch *et al.* (2012)
Species identification (reliability)	Misidentification of alien species resulting from taxonomic uncertainty	Gould *et al.* (2014); McGeoch *et al.* (2012)
Insufficient survey	Insufficient survey data resulting in failure to recognize invasive species or uncertainty about where a invasive species is in space and time	McGeoch *et al.* (2012)

Source of uncertainty	**Description**	**References**
Measurement error	Imperfect measurements or technique produce random error in estimation of species traits and characteristics (e.g. propagule pressure, initial densities, rate of increase, Allee threshold, dispersal rate and distance driven by rare, long dispersal events)	Gould *et al.* (2014)
Scale of the data (geographical and temporal correlation)	Overestimation due to the coarse resolution of species distribution maps (spatial correlation). Extralimital species species are often not recognized as invasive species. Lack of absence data also limit the creation of models	McGeoch *et al.* (2012)
Data and knowledge not documented (completeness)	Data are not available in the form of publications	McGeoch *et al.* (2012)
Data not accessible (transferability)	Language barriers to information and data transfer. There is often no single comprehensive data source that encompasses either a broad range of taxa and/or region	Elith *et al.* (2002)
Lack of baseline information (vagueness)	Inadequate indigenous range information, such as cryptogenic or translocated species, resulting in subjective interpretation of species being either alien of not. There is a need for clarifying meaning	Elith *et al.* (2002); McGeoch *et al.* (2012)
Environmental data		
Measurement error	Imperfect measurements or techniques produce random variation in results. Examples include available equipment that may not record location precisely	Barry & Elith (2006)
Systematic error	Methods produce biased data, e.g. sampling is close to roads(edge effect)	Elith *et al.* (2002)
Human error or entry error	Data can be miscoded during entry to GIS	Lu & Weng (2007)
Spatial classification error (class generalization)	Error in the nature of the spatial entity, such as misclassification of habitat or climate habitat suitability and details to include (road, edges)	Beale & Lennon (2012); Barry & Elith (2006)
Spatial error interpolation	Generalization of the cartographic representation of the object before digitizing, including displacement, simplification	Elith *et al.* (2002)
Temporal differences	The object changes character between the time of data collection and of database usage In the course of data processing (such as rasterizing vector data) and error arises because of rounding or algorithm error	Fisher (1999)
Data processing	In the course of data processing (such as rasterizing vector data) and error arises because of rounding or algorithm error	Fisher (1999)
Future climate data	Climatic variability, GCM model differences or emissions scenarios differences	Gould *et al.* (2014)
1.2.1 Parametric/Data uncertainty		
Modelling method including model parametrization and fitting	Among other, variable selection and model structure (from approximation and functional forms, equations, and mathematical expression to model fitting and parametrization) are often a key source of uncertainty as model projection vary according to the variables including in the model and the way there are related	Gould *et al.* (2014); Hartig *et al.* (2011); Matott *et al.* (2009)
Model selection and evaluation	Model selection and evaluation is particularly problematic in the case of spatial projection of species distribution and spread, as no independent data are available for future conditions and the likelihood of complex model can usual not be expressed	Burnham *et al.* (2011); Gould *et al.* (2014); Hartig *et al.* (2011); Matott *et al.* (2009); Railsback & Grimm (2011)
Technical uncertainty	Uncertainty generated by software or hardware errors	Ascough *et al.* (2008)
2. Variability uncertainty (Aleatory or irreducible)		
Natural variability	Uncertainty related to the inherent randomness/ stochasticity of natural and anthropogenic systems (chaotic and unpredictable)	Ascough *et al.* (2008)

Source of uncertainty	**Description**	**References**
Human variability (subjective judgement)	Uncertainty associated with human input, such as the level of knowledge, education, cultural bias, moral beliefs, concerns and interests of the invasive species risk analysis manager/decision-maker, stakeholder and lobby group can have a significant impact at all stages of the decision-making process. Even the 'technical aspect' of the modelling exercise vary significantly when different modellers are presented with the same data and context (experience and preferences of the modellers)	Ascough *et al.* (2008)
Institutional variability	Uncertainty due to the current social, economic, cultural and political climate	Ascough *et al.* (2008)
3. Linguistic uncertainty (reducible)		
Vagueness	Sharp boundaries and homogeneous classes do not represent reality	Elith *et al.* (2002); Regan *et al.* (2002)
Ambiguity	Words can have several meaning. For example, a wide range of alternative definition exist for alien and invasive species, and the adoption of alternative definitions results in differences in IAS listening	Elith *et al.* (2002); McGeoch *et al.* (2012); Regan *et al.* (2002)
Underspecificity	Quantitative and qualitative data may have unwanted generality, e.g. location not precisely reported	Elith *et al.* (2002); Regan *et al.* (2002)
4. Decision uncertainty		
Uncertainty in the acceptability of risk	Uncertainty due to frequent trade-offs between complex and often competing economic, societal and environmental needs and constraints that may influence whether or not a risk is judged acceptable	Ascough *et al.* (2008)

Table C.1: Potential sources of uncertainty in spatially explicit models of spread (modified from Ascough *et al.*, 2008; Elith *et al.*, 2002; Fisher, 1999; Gould *et al.*, 2014; McGeoch *et al.*, 2012)