MODELLING THE SPREAD OF INVASIVE SPECIES ACROSS HETEROGENEOUS LANDSCAPES

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

by

Joel Peter William Pitt

Lincoln University
2008

Abstract

Invasive species are well known to cause millions of dollars of economic as well as ecological damage around the world. New Zealand, as an island nation, is fortunate because it has the opportunity to regulate and monitor travel and trade to prevent the establishment of new species. Nevertheless foreign species continue to arrive at the borders and continue to cross them, thus requiring some form of management. The control and management of a new incursion of an invasive species would clearly benefit from predictive tools that might indicate where and how quickly the species is likely to spread after it has established. During the process of spread an invasing species must interact with a complex and heterogeneous environment and the suitability of the habitat in a region determines whether it survives. Many dispersal models ignore such interactions and while they may be interesting theoretical models, they are less useful for practical management of invasive species.

The purpose of this study was to create and investigate the behaviour of a spatially explicit model that simulates insect dispersal over realistic landscapes. The spatially explicit model (Modular Dispersal in GIS, **MDiG**) was designed as am open-source modular framework for dispersal simulation integrated within a GIS. The model modules were designed to model an an approximation of local diffusion, long distance dispersal, growth, and chance population mortality based on the underlying suitability of a region for establishment of a viable population. The spatially explicit model has at its core a dispersal module to simulate long distance dispersal based an underlying probability distribution of dispersal events. This study illustrates how to extract the frequency of long distance dispersal events, as well as their distance, from time stamped occurrence data, to fit a Cauchy probability distribution that comprises the dispersal module.

An investigation of the long distance dispersal modules behaviour showed that, in general, it generated predictions of the rate of spread consistent with those of analytical partial differential and integro-difference equations. However, there were some differences. Spread rate was found to be mainly dependent on the measurement technique used to determine the invasion front or boundary, therefore an alternative method to determine the boundary of a population for fat-tailed dispersal kernels is presented. The method is based on the point of greatest change in population density. While previously it was thought that number of foci rather than foci size was more important in stratified dispersal and that finer resolution simulations would spread more quickly, simulations in this study showed that there is an optimal resolution for higher spread rates and rate of area increase. Additionally, much research has suggested that the observed lag at the beginning of an invasion may be due to lack of suitable habitats or low probability of individuals striking the right combination of conditions in a highly heterogeneous environment. This study shows an alternative explanation may simply be fewer dispersal event sources.

A case study is described that involved the application of the spatially explicit dispersal model to Argentine ant spread to recreate the invasion history of that species in New Zealand. Argentine ant is a global invasive pest which arrived in New Zealand in 1990 and has since spread to both main islands of New Zealand, primarily through human mediated dispersal. The spatially explicit simulation model and its prediction ability were compared to that of a uniform spread model based on equivalent total area covered. While the uniform spread model gave more accurate predictions of observed occurrences early in the invasion process it was less effective as the invasion progressed. The spatially explicit model predicted areas of high probability of establishment (hot spots) consistent with where populations have been found but accuracy varied between 40-70% depending on the year of the simulation and parameter selection. While the uniform spread model sometimes slightly outperformed or was equivalent to the simulation with respect to accuracy early in the invasion process, it did not show the relative risk of establishment and was less effective later in the invasion when stochastic random events generated by the simulation model were averaged to represent trends in the pattern of spread. Additionally, probabilistic predictions as generated by the spatially explicit model allow the uncertainty of prediction to be characterised and communicated.

This thesis demonstrates that heterogeneous spread models can give more insight and detail than one dimensional or homogeneous spread models but that both can be useful at different stages of the invasion process. The importance of compiling appropriate data on dispersal and habitat suitability to aid invasion management has been highlighted. Additionally, a number of important hypotheses that need to be addressed to increase understanding of how species interact with the complex environment, have been identified and discussed.

Keywords: dispersal, spread, simulation, model, invasions, invasive insects, Argentine ant, *L. humile*, GIS, framework.

Contents

List of Tables **ix**

List of Figures **x**

1 Introduction **1**

1.1 Biological Invasions 1

1.2 The process of invasion 4

1.3 Factors that influence invasion 5

1.3.1 Landscape effects on the invasion process 7

1.4 Invasion Management 8

1.5 Modelling invasive spread 9

1.6 Heterogeneity 11

1.7 Objectives 11

1.8 Thesis structure 12

2 Dispersal Models **13**

2.1 Simulation 13

2.2 Explicit spatial-temporal modelling 14

2.2.1 Heterogeneity and Homogeneity 16

2.2.2 Scales of time and space 16

2.3 Models of Dispersal and Biological Invasion 17

2.3.1 Spatial dynamics of biological invasions 19

2.3.2 Stratified dispersal 20

2.4 Effect of population growth on species spread 21

2.5 Applicable Fields 22

2.6 Partial Differential Equations 23

2.7 Integrodifference equations 24
2.7.1 Dispersal kernels 25
2.8 Cellular Automata and Coupled Map Lattice models 26
2.9 Metapopulation models 28
2.9.1 Stochastic patch occupancy models 29
2.10 Percolation models 30
2.11 Individual Based Models 30
2.12 Summary 32
2.12.1 Models used for management of invasive species 32

3 A modular dispersal simulation framework **34**
3.1 Using GIS to model dispersal 34
3.1.1 Model Linkage to GIS 36
3.2 Architecture and Design 37
3.2.1 **MDiG** - A framework to manage simulations 38
3.3 Dispersal model modules 41
3.3.1 Local Dispersal 41
3.3.2 Kernel based dispersal 45
3.3.3 Patch Survival 49
3.3.4 Growth 52
3.4 Analysis modules 57
3.4.1 Interpreting results across simulation replicates 57
3.4.2 Area occupied 59
3.4.3 Verification and Validation 59
3.4.4 Rate of spread 60
3.5 Summary 62
3.5.1 Relation to existing modelling approaches 62
3.5.2 Consideration of future direction 63
3.5.3 Vector based models 64

4 Analysing Module Dynamics **65**
4.1 **Local** dispersal dynamics 66
4.2 **Kernel** dispersal 70
4.2.1 Area occupied 70
4.2.2 Spread rate 81

4.3 Comparison of **MDiG** dynamics with other models 88

4.4 Summary 91

5 Modelling the Spread of Argentine Ant in NZ 93

5.1 Introduction 93

5.2 Biology 94

5.2.1 Diet 95

5.2.2 Climatic requirements 95

5.2.3 Habitat 96

5.3 Impacts 96

5.3.1 NZ Impact 98

5.4 Distribution 100

5.4.1 World distribution 100

5.4.2 *L. humile* distribution in NZ 101

5.4.3 Potential Distribution in NZ 102

5.5 A model of Argentine ant spread 104

5.5.1 *L. humile*'s dispersal behaviour 104

5.5.2 Model description 106

5.5.3 Parameter estimation 106

5.5.4 Simulation Region 113

5.5.5 Potential Distribution 113

5.5.6 Uniform spread model 118

5.6 Results 119

5.6.1 Occupancy Envelopes 119

5.6.2 Total Area 121

5.6.3 Comparison of simulation with occurrence data 122

5.6.4 The uniform spread model 123

5.6.5 Hotspots 129

5.7 Discussion and Conclusions 130

6 General Discussion 134

6.1 Modelling framework 134

6.2 Model dynamics 135

6.3 Argentine ant simulation 137

6.4 Data and Parameterisation 138

6.5 Future Research 139
    6.5.1 Argentine ant model 139
    6.5.2 Population models 140
    6.5.3 Lifestage dependent dispersal 141
    6.5.4 Surveillance and Treatment optimisation 141
    6.5.5 Cost estimation 142
    6.5.6 Multi-scale modelling 142
    6.5.7 Numerical Integration 143
    6.5.8 Use of remote imaging to gather data 143
    6.5.9 Conservation 144
    6.5.10 Integrated web based interface 144
6.6 Conclusion 145

References **147**

Appendices **182**

A Gypsy Moth phenology modelling **182**

B The future of MDiG **194**

B.1 Extensions to existing modules 194
    B.1.1 Possible extensions of the local dispersal module 194
    B.1.2 Possible extensions for the kernel dispersal module 196
    B.1.3 Possible extensions to the patch survival module 198
    B.1.4 Possible extensions for the growth module 200
    B.1.5 Possible extensions to the validation system of MDiG 201
    B.1.6 Possible extensions for the rate of spread module 202
B.2 Future module ideas 202
    B.2.1 Vector maps 202
    B.2.2 Gravity models 203
    B.2.3 Human transport networks 203

C New Zealand *L. humile* model **204**

C.1 DispersalML definition 204
    C.1.1 DispersalML for *L. humile* model 204
C.2 New Zealand *L. humile* occurence data 209

C.3 Simulation results 219

List of Tables

2.1 Cellular Automata have been used to model many processes in biology and ecology. Several study areas and related references are shown. 27

3.1 Model modules. 42
3.2 Population growth functions supported by **MDiG**. Where N_t is the population at time t, r is an index of the growth rate, and K is the local carrying capacity (see relevent sections for the definition of any other variables). 54

4.1 Area of invasion as a function of time for different neighbourhood shapes when initiated from a point source. All neighbourhoods have a radius of 1, thus the maximum possible area is when all cells within the radius of 1 are part of the neighbourhood (equivalent to a Moore neighbourhood). The function for this maximal area is $A(t) = (2t - 1)^2$ 67

5.1 Summary of the main documented impacts globally of *L. humile* on native ecosystems. From Harris (2002) 99
5.2 Parameter estimates for the *L. humile* dispersal model. Estimates are for the Cauchy (x_0, γ) and Poisson (λ) distributions that describe the USA dispersal events and nearest neighbour distances respectively. For the Cauchy distribution estimates were made by MLE with the final estimate having a negative log-likelihood of 4479.88.111
5.3 Scheme used for combining land-cover and degree-day suitability maps. H* and T* represent *any* category of habitat and degree-day suitability respectively. 118

C.1 Argentine ant occurence data for New Zealand 209

List of Figures

1.1 Stages of areal increase for an invasion 5

1.2 The spatial layout and structure of a landscape can affect each stage in the invasion process. From With (2002). 7

2.1 The result of stratified dispersal, with the population advancing left to right. The infested zone shows where colonies have coalesced, the transition zone is where isolated colonies establish and grow due to propagules from the infested zone, and the uninfested zone is where the species is generally absent. 21

3.1 The model architecture consists of a controller that reads in a model definition and then manages a series of modules that sequentially modify a population distribution. t indicates the current time-step, and n is the number of modules that comprise the simulation. 39

3.2 Population vs. presence/absence models. Population spread models record individual abundance in each patch (abundance on z-axis of sub-figure a.), whereas presence/absence models purely indicate whether the population exists in a patch. 42

3.3 An example of a **local** module neighbourhood created using a radius of 1 and shape value of 186. This value is equivalent to the binary string `010111010`, with ones indicating what patches in the radius are in the neighbourhod. The binary string maps left to right, top to bottom, for patches in the radius. The black patch represents the current patch being processed. 43

3.4 Subsequent distributions after applying the local dispersal module with shape 186 to a presence-absence distribution for 10 time steps. The gray cells indicate newly occupied patches. 44

3.5 The generic dispersal kernel for several values of s. $s = 0.5$, $s = 1$ (Laplace or negative exponential), $s = 2$ (Gaussian), $s = 3$ (Ribbens). 46

3.6 The Cauchy dispersal kernel for several values of γ. 47

3.7 The Exponential dispersal kernel for several values of β. 48

3.8 The logistic growth curve showing population density N_t (scaled so that $K = 1$) increasing with time t for varying values of the intrinsic growth rate r. 55

3.9 The Beverton-Holt growth curve showing population density N_t increasing with time t for varying values of the population change coefficient r. 55

3.10 The Ricker growth curve showing population density N_t increasing with time t for varying values of the intrinsic growth rate r. 56

3.11 The growth curve from Neubert (1997) showing population density N_t increasing with time t for varying values of the population change coefficient r and a constant threshold $a = 0.9$. 57

3.12 The growth curve from Wang et al. (2002) showing population density N_t increasing with time t for varying values of the growth index r. 58

3.13 The growth curve presented in Keitt et al. (2001) showing population density N_t increasing with time t for varying values of the growth rate r. 58

4.1 Neighbourhood configurations with radius= 1 and various shape values. The shape values have been chosen so that none are rotationally equivalent and that they have a differing number of patches in their neighbourhood. 68

4.2 Neighbourhoods with a radius of r can represent all the neighbourhoods with radii $< r$. Here, the neighbourhood for shape 186 (left) with $r = 1$ is identical to the neighbourhood for shape 1455536 with $r = 2$. 68

4.3 Some neighbourhood configurations are equivalent after rotation. Here 300 (left) is the same as 466 (right) if it is rotated clockwise 90 degrees ($\frac{\pi}{2}$ radians). 69

4.4 The saturation, or proportion of patches present out of maximum area possible, for different neighbourhood shapes and an initial condition of one occupied patch at origin. As model time progresses the saturation of each shape category approaches a limit. 69

4.5 The probability of a dispersal event leaving its source cell varies as both shape and distance parameters of the generic kernel are altered. The surface shows the probability of leaving the source cell, $P_{>r}$, in a region with a resolution of 1. The probability was calculated by numerically evaluating the integral in Equation 3.6. . 71

4.6 Log of mean area occupied by a distribution versus time, for (a) 5 values of the distance parameter, d with $s = 1.0$, and (b) 6 values of the shape parameter, s with $d = 1.0$. The total expected possible area (equivalent to 2^t) is also shown, which assumes every generated dispersal event lands in an unoccupied patch. 72

4.7 Normal plot of area data over 50 replications with parameters $s = 0.5$ and $d = 1.0$. The area values shown are those at the end of the simulation ($t = 25$). 73

4.8 Grid showing the timesteps when a) the distribution of area values between replications conforms to a normal distribution and b) when the area values are log transformed first. The simulations were run for 25 time steps. The shading indicates the P value for accepting the null hypothesis, i.e. the area values for a given parameter combination come from a normal distribution. Black blocks indicate the null hypothesis is rejected ($P < 0.05$). 74

4.9 Mean area resulting from 4 sets of constant parameters in the generic dispersal kernel. Each set was simulated at resolutions that were factors of 20. 76

4.10 Multiple comparisons between area distributions for different resolutions at $t = 15$. Each graph is for different values of the s and d parameters of the generic dispersal kernel. The average group ranks (circles) and ranges (line) are used for comparisons with $p < 0.05$. The intervals are significance ranges, not simultaneous confidence intervals. 77

4.11 The total probability of the dispersal kernel that falls within the source cell for a dispersal event. The total probability volume encompassed by a cell depends on both resolution and kernel parameters. 79

4.12 Mean area with varying s and d dispersal kernel parameters. Solid lines indicate the mean area in a limited space of 200 by 200 cells, the dashed lines are the control simulations in practically unlimited areas. 81

4.13 The mean distance metrics across 50 replications for kernel dispersal with $d = 1.0$ and varying shape parameter. The difference between timesteps (equivalent to the velocity) for the metrics is also shown. 83

4.14 The mean distance metrics across 50 replications for kernel dispersal with $s = 1.0$ and varying distance parameter. The difference between timesteps (equivalent to the velocity) for the metrics is also shown. 84

4.15 The 95% confidence interval for the mean acceleration of distribution boundary as detected by 4 different methods: mean distance, furthest forward, best-cell, and density. If zero is included in the interval then the distribution is not significantly different to zero and thus the distribution boundary is unlikely to be accelerating. 86

4.16 The mean distance metrics across 50 replications for kernel dispersal in a limited area of 200 by 200 patches. The control lines are the results for "unconstrained" simulations (limited to 2400 by 2400). 89

4.17 The mean distance metrics across 50 replications for kernel dispersal (d=5.0,s=0.5) in a large area of 2400 by 2400 units and simulated at different resolutions 90

5.1 *L. humile* workers are small (2-3mm long) and are coloured a uniform light to dark honey-brown. Unlike some ant species there is no strong formic acid smell when they are squashed. Photo courtesy of Landcare Research, NZ. 94

5.2 Known worldwide distribution of Argentine ant. Numbered black circles indicate introduced populations while shaded circles represent presumedly native populations. Used with permission from Suarez et al. (2001). 102

5.3 The distribution of Argentine ant (*L. humile*) in NZ as of 2005. a) Focussed on the Auckland region. b) Showing the entire distribution in New Zealand. Occurrences are shaded based on the year of discovery: 1990-1993 black, 1994-1996 dark gray, 1997-1999 medium gray, 2000-2002 light gray, 2002-2006 white. 103

5.4 The distances travelled by *L. humile* long distance dispersal events in NZ and the USA, using both a) distance-from-origin, and b) nearest-neighbour, methods of distance measurement. Note that there were several events that fall outside of the x-axis range in the distance-from-origin graphs, but which can be seen in the nearest neighbour graphs. This is simply the result of keeping the same scale so that distance measurement techniques can be compared. 110

5.5 The frequency of *L. humile* long distance dispersal events in NZ and the USA. Measuring the a) number events per year, and b) the number of events per year divided by pre-existing sites. 112

5.6 Initial occurrences of *L. humile* in NZ at the end of 1990. White circles indicate occurrence sites. The red dashed line represents State Highway 1, and the green solid lines, railways. The close proximity of the initial sites to major transportation routes may have assisted in *L. humile*'s spread. 114

5.7 Suitability maps for *L. humile* in the North Island of NZ, based on a) land cover classification, b) degree-days above a threshold of 15.9C. 116

5.8 Combined suitability map for *L. humile* in the North Island of NZ. 117

5.9 a) The occupancy envelope generated by a simulation ($\lambda = 0.298$, $\gamma = 83700$, Year 2005) and b) the same occupancy envelope split into thresholded occupancy rates. 120

5.10 Total area encompassed by the *L. humile* simulation envelope at each threshold for the 9 combinations of γ and λ simulated. 121

5.11 The percentage of *L. humile* occurrence sites that lie within the probability envelope of the simulated distributions for each year of the simulation. The percentage that lie anywhere in the probability envelope (i.e. at least one simulation out of the 100 replicates predicts the occurrence of a site) was calculated, as was the percentage for envelopes that have been truncated at the thresholds of 0.05, 0.10, and 0.50. In other words, values below these thresholds are masked out of the envelope. 124

5.12 Map of uniform spread vs **MDiG** simulation. The lines indicating the fronts for the uniform spread begin from the centre circle at 1995 and procede annually to the outer front for 2005. Red dots indicate occurrence sites for *L. humile* as of 2005. The uniform spread model is based on the area increase of the envelope thresholded at 5%. 125

5.13 Comparing the percentage of correctly predicted sites between the **MDiG** *L. humile* simulations and the approximately equivalent linear spread models. A positive $\Delta\%$ indicates the **MDiG** model out-performs the linear spread model. 126

5.14 Prediction for chance of site occupancy for individual sites. Each bar represents an occurrence and are ordered left to right based on ascending distance from invasion origin. Vertical coloured lines correspond to the distance that each uniform spread model reaches for a particular threshold (red 0.50, blue 0.10, green 0.05, yellow none). Graphs are for $\lambda = 0.298$, $\gamma = 83700$. 127

5.15 Snapshots of the dispersal simulation ($\lambda = 0.298$, $\gamma = 83700$) for years 1999, 2001, 2003, and 2005. The map area is coloured according the threshold the probability envelope exceeds (red > 0.5, green > 0.1, blue > 0.05, yellow < 0.05). The coloured circles indicate the boundary for a uniform spread model, the area of which is equivalent to the area encompassed by the thresholded probability envelope. Red dots with black outlines indicate real *L. humile* occurrence sites. 128

5.16 Snapshots of the entire simulation region ($\lambda = 0.298$, $\gamma = 83700$). 129

C.1 Probability envelopes for *L. humile* simulation with parameters, $f = 0.19926$ and scale= 72700 220
C.2 Probability envelopes for *L. humile* simulation with parameters, $f = 0.19926$ and scale= 83673 221
C.3 Probability envelopes for *L. humile* simulation with parameters, $f = 0.19926$ and scale= 94647 222
C.4 Probability envelopes for *L. humile* simulation with parameters, $f = 0.29779$ and scale= 72700 223
C.5 Probability envelopes for *L. humile* simulation with parameters, $f = 0.29779$ and scale= 83673 224
C.6 Probability envelopes for *L. humile* simulation with parameters, $f = 0.29779$ and scale= 94647 225
C.7 Probability envelopes for *L. humile* simulation with parameters, $f = 0.42798$ and scale= 72700 226
C.8 Probability envelopes for *L. humile* simulation with parameters, $f = 0.42798$ and scale= 83673 227
C.9 Probability envelopes for *L. humile* simulation with parameters, $f = 0.42798$ and scale= 94647 228

Chapter 1

Introduction

1.1 Biological Invasions

New Zealand has a unique array of flora and fauna which has arisen because of its isolation from other land masses. Since the advent of colonisation, New Zealand's ecosystems have had to cope with a flood of new species. The burning of large swaths of bush during both pre- and post-settlement resulted in a modified and disturbed environment that would have been very susceptible to the establishment of invasive species (Hobbs 2000). Initially, many plant and animal species were intentionally introduced for agricultural development but have since developed into pests after being found to be well adapted to the environment that exists in New Zealand. Insect species, in addition to being carried by wind and ocean currents (Ward 1992, Gatehouse 1997), were brought in with imported agricultural crops and benefited from the agricultural development of their hosts.

Invasive species are estimated to cause damage in the order billions of dollars per year in the United States alone (Pimentel et al. 2000) mainly from agricultural damage, but increasingly from the cost of protecting ecological reserves. Worldwide, invasive species are a major threat to biodiversity following, in particular, anthropogenic disturbances that result in habitat destruction and fragmentation (Wilcove et al. 1998, Pimentel et al. 2001).

On top of increased carbon dioxide in the atmosphere, widespread changes in land use and land cover, and the hunting and harvesting of natural species populations, invasions are considered to be a significant factor in global environmental change caused by the explosive growth of industry and agriculture in the past two centuries (Vitousek et al. 1996). One of the important consequences of the environmental change caused by biological invasions is that of global species homogenisation

(Olden & Poff 2003) and the replacement of native species by invasive species that can cause significant impacts (Vitousek et al. 1997, Wilcove et al. 1998, Pimentel et al. 2005).

The biotic cost of invasions is most dramatic when non-native species take over a community and completely alter landscape structure and ecosystem function (Vitousek et al. 1996) however not all introduced species establish succesfully in new communities (Williamson & Fitter 1996) nor do all introduced species spread much further than the original incursion site. Some may become harmlessly naturalised into the local resident community (Richardson et al. 2000).

Clearly, it is desirable to prevent species arriving that may become pests and consequently cause serious economic loss through damage to agricultural crops which are part of New Zealand's primary industry, or cause harm to New Zealand's indigenous flora and fauna by disrupting ecosystem processes or out-competing native species. However, the increase in human travel and cargo, both locally (Group 2005), and internationally (Levine & D'Antonio 2003), has increased the chances that many species establish outside of their native range and has made the task of prevention increasingly difficult. Increasing threats are shown in the United States of America by the increase in interceptions on air passenger baggage (Liebhold et al. 2006) and through other vectors of transport (McCullough et al. 2006). One example, among the many, of a species that has had its worldwide distribution greatly enhanced by human-mediated dispersal is the Argentine ant (see Chapter 5 and Suarez et al. 2001).

The time taken to cross large distances, which formerly isolated New Zealand from other landmasses, has been shortened by commercial flight. These flights, along with other transport methods, have been increasing in frequency world-wide (IUCN 1999) and in New Zealand (Whyte 2005*a*) as a result of globalisation. Specifically, over half a million sea containers were shipped to New Zealand in 2004-2005, 57% more containers than 5 years previously. New airlines have been founded and more flights connect to areas containing potentially damaging species such as Asia and South America. These new flights and destinations have increased the number of air passengers arriving in New Zealand to 4 million in 2005, a 30% increase over the previous 5 years (Whyte 2005*b*). The increased connectedness between urban centres has made urban population size more highly correlated with the occurrence of outbreaks than distance between cities, particularly for disease outbreaks (Cliff et al. 1981).

Thankfully there is some respite, as research suggests surprisingly small numbers of potential invaders actually establish or have significant impact. Williamson & Fitter (1996) coined the "tens-rule", based on species introduced to the British flora. In their study Williamson & Fitter (1996) found an estimated 10% of all introduced species actually become established, and of those, only ten percent were invasive enough to spread and be considered pests. This rule, as an approximation, has been supported by analysis of the invasive species of other countries including Germany (Kowarik 1995), the Netherlands (Weeda 1987), and France (Kornas 1990).

Regardless of the small percentage of introductions that have significant impact, the increased risk from large amounts of trade and tourism has increased the sense of urgency to understand factors behind the anthropogenic distribution of propagules (Levine & D'Antonio 2003) and the biological processes leading to establishment (Mack et al. 2000, Kolar & Lodge 2001). Additionally, there is a real need to develop effective tools for invasion risk assessment (Hayes 1997, Simberloff & Alexander 1998, Mack et al. 2000, Kolar & Lodge 2001)

Such growth in travel and trade increases the threat of species invasion, but to combat that threat, detection technologies need to continue to be improved (e.g. the proposed use of sensor technologies as described by Goldson et al. 2002), as well as control and monitoring techniques (Sharov et al. 2002).

The establishment of introduced species is often associated with severe ecological and economic consequences (Mack et al. 2000, Pimentel et al. 2000). Preventing invasive species from arriving in New Zealand is important to preserve the country's endemic flora and fauna because invasive species can reduce biodiversity (Human & Gordon 1997) and displace native speces (Holway 1998*a*). If organisms get past the procedures put in place to intercept them there is a chance that the species may establish a reproducing population that may potentially spread further. If such a population is detected in time and the species is judged to have the potential to cause ecological and economic damage, then it is highly likely a decision will be made to eliminate it. However, the eradication and control of invasive species is a costly process, 800 million dollars are spent in New Zealand annually on invasive phtyphagous insect species alone, covering biosecurity activities including surveying, control and research (Barlow & Goldson 2002).

It is therefore imperative that, upon the arrival of a foreign species, an appropriate government agency carry out a pest risk analysis (if they have not already done so). This rick analysis must be done as quickly as possible, estimating the potential damage the species may have on biological production systems, including agriculture, forestry, and indigenous ecosystems. The resources are not available to inspect all incoming cargo against potential invaders, or to mount an eradication campaign for every species as soon as it is detected, therefore eradication or control actions must be prioritized so as to minimise the damage caused by a species determined to have the potential to spread over a large area and/or have a significant negative impact.

1.2 The process of invasion

Biological invasions are often divided into several stages from arrival to naturalisation. Hastings (1996*a*) identifies three phases: 1) the establishment of the invading species at a single location, 2) the interaction with existing species in the community as a result of population growth, competition and/or predation, and 3) the spatial spread of the invading species from the invasion origin. The first phase can be preceded by an initial introduction phase where propagules of an invasive species arrive but have yet to lead to establishment (Andow et al. 1990).

During the introduction phase, the propagules need to reach a stage where they can sustain a local population. These introduced populations often go undetected because of their small numbers, and many fail to become established because of inconsistently suitable climate (Mack 1995) and Allee effects. Propagule pressure, or the temporal frequency and size of introductions, is clearly an important factor in the establishment of introduced species (Kolar & Lodge 2001, Lockwood et al. 2005, Von Holle & Simberloff 2005). Establishment and the later process of naturalisation have both been shown to correlate with propagule pressure, both in terms of the number of individuals introduced and frequency of introduction events (Williamson & Fitter 1996, Grevstad 1999, Kolar & Lodge 2001). These empirical findings are also consistent with the predictions made in theoretical population biology (Richter-Dyn & Goel 1972).

Invasions, once established are often observed to have a *lag time* between initial introduction and subsequent and obvious population growth. This lag time could be caused by a number of factors, including weak Allee effects, or a small number of foci from which initially the species can spread (Mack 1985). Ewel (1986) suggested lags could be due to a delay in suitable habitat becoming available or a low probability of new propagules striking the right combination of environmental conditions. These lag times, particularly for alien woody species, can last hundreds of years (Kowarik 1995), although for some insect species, like the African bee, the lag time can be almost entirely absent (Kerr 1967).

After a certain time, assuming the alien population continues to maintain itself, an invasive species is usually observed to spread and undergo exponential population growth. It is at this stage that an invasion becomes visible to environmental managers as the population numbers and density increases. This growth and spread continues until further resources and suitable environment become unavailable and then growth and spread slows as the population reaches an equilibrium with its environment.

After a species establishes, the changing spatial pattern of its distribution and the increase in area occupied by its population can usually be partitioned into the following periods (Fig. 1.1): an

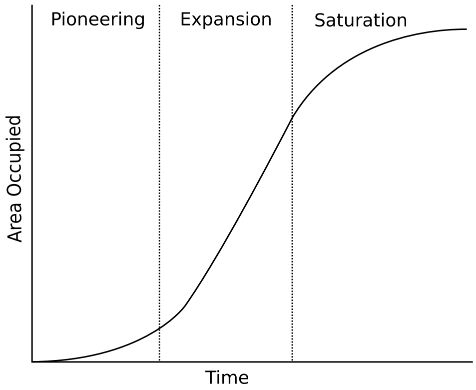


Figure 1.1: The increase in the area occupied by an invasive species' population can be partitioned into 3 stages of pioneering, expansion, and saturation (Shigesada et al. 1995)

early pioneering period as new sub-populations are established, a middle period with rapid change as the population expands to new habitat, and a later period of condensing saturation or fill-in when new habitat and resources become scarce (Mollison 1987, Shigesada et al. 1995). *Bromus tectorum* (cheatgrass) is one example that exhibited this pattern of increase in area as it spread throughout the Great Basin of North America (Radosevich et al. 1997).

The main focus of this work with respect to the spatial spread of a invasive species is its range expansion. However, for many invasions, all three stages can often occur concurrently due to beach-head populations establishing ahead of the main invasion front (see section 2.3.2). These advance populations may arrive via jump dispersal and essentially undergo the same process as the original invasion, at least until the main invasion front catches up and envelopes the isolated population. Since the three stages are all involved throughout the invasion process, they all need to be considered in a model of invasive spread.

1.3 Factors that influence invasion

Many factors influence whether an invasion is successful. Clearly, during the establishment phase, landscape and habitat disturbance is a major factor that favors invasions, particularly invasions of alien plants. Pickett & White (1985) describes disturbance as 'any relatively dicrete event in time that disrupts ecosystem, community or population structure and changes resources or substrate availability or the physical environment'. As such, these disturbances may be caused by large-scale events such as flooding, fire, storms, or at a smaller scale by soil turnover or vegetation removal (Hobbs 1991, Hobbs 2000). Human land-use patterns and the consequent disturbances also enhance invasibility of landscapes (Hobbs 2000) and many environmental weeds of Australia have been associated with at least one type of disturbance that favour their establishment and spread (Humphries et al. 1991).

There are many theories of invasion but perhaps the one that has attracted the most attention has been the biotic resistance theory (Elton 1958). Biotic resistance is the theory that species richness is an important variable controlling the invasibility of native ecosystems. Species poor communities are argued to be more susceptable to invasion and thus they are said to have less biotic resistance (Elton 1958, Pimm 1991). Early reasoning suggested that mainland species had a superior competitive ability to those on smaller islands (Carlquist 1965). However, conversely Huston (1994) and others suggest that diverse native communities, instead of being able to 'resist' invasion, are able to more easily accommodate invaders as high diversity implies greater resource availability and weak interspecific interactions. Empirical evidence has been used to support both conclusions, and Byers & Noonburg (2003) reason that the difference observed in community invasibility correlated with species richness arises from the differences of scale in empirical studies. At large spatial scales more exotics appear to invade species rich communities (e.g. Lonsdale 1999), whereas at small scales, less exotics are observed to invade (e.g. Naeem et al. 1999).

There have also been several attempts to determine statistical associations between invasiveness and life history traits (e.g. Scott & Panetta 1993, Reichard & Hamilton 1997, Kolar & Lodge 2001) or taxonomic relationships (e.g. Daehler 1998, Pysek 1998). Although the invasiveness of a species is still somewhat particular to each case, some biological traits that are considered to be correlated with species invasiveness have been outlined by Rejmánek (2000):

- *fitness homeostasis*, or in other words, an ability to maintain relatively constant fitness over a range of environments.
- small genome size, which is usually associated with short generation time and high relative growth rate.
- easily dispersed by humans and animals, which increases the chances a species is introduced to a new region, greatly increasing the spread rate of the species.

Others suggest that invasions occur because of the "enemy release hypothesis of invasiveness" (Keane & Crawley 2002) which specifies that when introduced species leave their native predators and parasites behind and they are introduced into a new community there is either a reallocation of resources from enemy defence to growth (compensatory release), and/or a direct increase in growth or survivorship (regulatory release) (Colautti et al. 2004). Blossey & Nötzold (1995) have suggested that the absence of native predators might even drive evolution towards greater growth

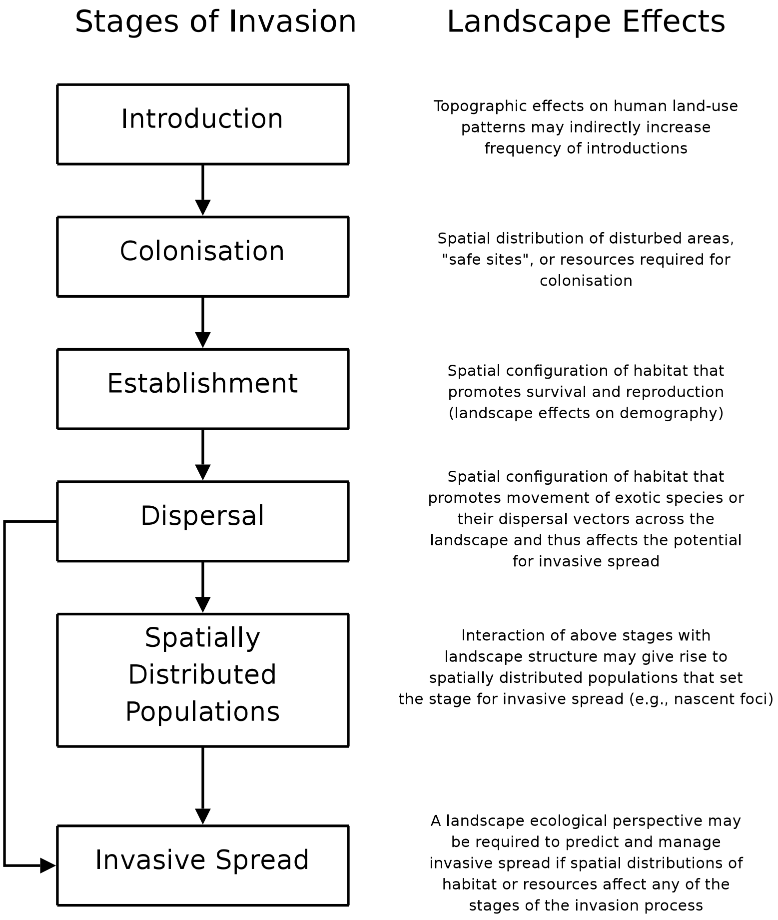


Figure 1.2: The spatial layout and structure of a landscape can affect each stage in the invasion process. From With (2002).

rates . Thus a species that has strong mechanisms for compensatory or regulatory release, or a relatively quick evolutionary rate may be more suited to being invasive compared to other species.

There is no doubt however that both human land-use patterns and global climate change are major factors affecting the spread of invasive species (Mooney & Hobbs 2000). Habitat destruction and fragmentation from anthropogenic disturbances are viewed as the leading threat to biodiversity after the threat posed by invasive species (Wilcove et al. 1998) and consequently increase the risk of invasion.

1.3.1 Landscape effects on the invasion process

The spatial structure of the landscape can impact on the invasion process in many ways (With 2002). Figure 1.2 lists some of the impacts at each stage from introduction to the spread of an invasive species.

The transformation of landscapes by humans has been rapid, and widespread (Whitney 1994), making such rapid change increasingly important in the colonisation of invasives. Not only is there a direct effect of human land-use patterns but such patterns can alter the flow of propagules dispersed by humans indirectly. The distribution of disturbed sites that are prone to invasion along with human land-use patterns that may enhance the invasibility of landscapes (Hobbs 2000) influence whether an alien species will colonise an area.

Also, successful establishment is dependent on the location of resources and habitat for population survival and subsequent growth. The location of habitat and resources also affects the spread of a population as individual dispersal and demography can be affected by landscape structure (e.g. With 1994, With & King 1999, With & King 2001). It is clear that the influence of landscape pattern on the invasion process is inescapable, and spatial models can help us better understand the interaction and potentially predict the direction an invasion will take. Landscape ecology is the study of "the effect of pattern on process" (Turner 1989) or the study of the ecological consequences of spatial pattern. The challenge is to apply the knowledge from landscape ecology to the invasion process.

For example, Eppstein & Molofsky (2007) developed a theoretical model that incorporates propagule pressure, frequency independent growth rates, feedback relationships, resource competition and spatial scale of these interactions. The model predicted one of four outcomes: inability to establish, naturalization, unconditional takeover, or conditional invasion dependent on quantity and spatial distribution of propagules. The latter scenario, where invasion is dependent on the spatial distribution and quantity of propagules, highlights the important role that the spatial pattern of population distributions plays in the success of an invasive species establishment.

1.4 Invasion Management

Decisions with respect to the management of an invasion depends on what stage in the invasion process an alien species has been detected. A key tool for management is prevention through border security practices and quarantine. Because of the uncertainty of establishment and because its success is correlated with propagule pressure, prevention may be one of the most effective management approaches. This is particularly the case for species that are considered a high risk for a country. For example, the Asian gypsy moth is a potential threat to New Zealand, and imported shipping containers and imported cars from high risk regions like Japan and Far East Russia are checked for signs of the moth (MAF 2001*b*, Ross 2004).

Sometimes a species does establish, and may also be fortuitously detected early. Maybe because of a long lag period allowing detection before it undergoes rapid population growth and subsequent spread, or it might be detected by an existing monitoring program for another species, or simply through luck. Eradication is often feasible if an invasive species is detected early and especially if the public are informed of high risk species. For example, an early gypsy moth detection in Hamilton, New Zealand, was eradicated using aerial spraying of Btk (Ross 2004). Informing the Hamilton public to report any sightings of the moth to the Ministry of Agriculture and Fisheries facilitated rapid eradication.

During the rapid population growth and spread of the expansion phase, authorities may not have the option of eradication and are limited to trying to control and contain the spread of the species. This approach was taken by Slow The Spread project, which used strategic eradication of sub-populations to slow the progression of the European gypsy moth across North-Eastern United States of America (Sharov et al. 2002). Decisions about the management of gypsy moth in North America have been well supported by predictive models (Gage et al. 1990, Régnière & Nealis 2002).

1.5 Modelling invasive spread

Once an exotic species establishes a viable population in a new region, the next stage of the invasion process is its spread across the landscape (Hastings 1996*a*). A model that can predict the rate of spread and its direction, and that can be used to optimise the spatial arrangement and frequency of sampling strategies and eradication treatments, would greatly assist government agencies and environmental authorities to design effective monitoring and control measures. However, predicting spread is not an easy task for two reasons. First, data for model parameterisation are usually not available to estimate the rate of spread of a newly detected species, so modeling efforts will often rely on data either from related species, or the same species previously established in another location. Second, many species spread via multiple methods, for example through natural means (a diffusion like process) and large jumps because of human-mediated dispersal (e.g stratified diffusion) (Hengeveld 1989) or rare events such as the potentially extreme distances of wind-borne propagules. Management strategies need to account for such variation in the temporal and spatial aspects of dispersal while the best strategy for slowing or stopping an invaders spread is evaluated. For example, Moody & Mack (1988) have shown that preventing the establishment of new foci for spread can be more effective than attempting to slow spread from existing established populations.

A successful research model for invasive species spread that utilises a realistic landscape could be integrated into frameworks and decision support systems and guide various strategic management approaches. As an example, the 'Slow The Spread' programme has helped the control and supression of Gypsy moth by determining the appropriate size of barrier zones for treatment and monitoring (Sharov et al. 2002).

Interestingly, the total area that is searched for a species' presence can be a much greater factor in cost of control than controlling the density of the species (Susan Timmins, pers. comm.). Clearly models that predict potential distributions and the ecological niches of species can indicate which areas are suitable for that species, hence decreasing the total search area necessary. Spread models incorporate temporal change in a population distribution in addition to information about potential distributions, such that potentially only those regions within communicable distance require monitoring. Furthermore, the cost of eradication is generally recognized as being much greater than the cost of surveillance and containment of an invasive species (Mack et al. 2000). Thus proactive monitoring with appropriate buffer zones based on spread predictions could be more cost-effective than a reactionary approach to high-risk species.

Despite the usefulness of heterogeneous spread models, more general models or indicators of population viability on a landscape are often put forward as having more reliable results (Doak & Mills 1994). Examples are the ecologically scaled landscape indices of Vos et al. (2001) which may be useful for making management decisions. However, these don't address the rate of spread or the speed at which an invasion will occur. This rate is important because any delay in responding to an invasive species may make it impossible to control or eradicate it.

Mooij & DeAngelis (2003) outline how recent progress in model development and theory has shown promise resulting in better parameter estimation for spatially explicit population models. For example, mark-recapture methods have been extended to determine dispersal parameters (e.g. Hanski et al. 2000), and approaches for extracting information from population patterns in space have been developed (e.g. Grimm et al. 1996). Furthermore, Mooij & DeAngelis (1999) and South (1999) showed that error propagation in spatially explicit population models need not be as great as has been previously reported (Ruckelshaus et al. 1997) so there should be some optimism about developing such models. Data however is always a limiting factor in modelling but attempts to develop spatially-explicit spread models, as in all models, should clarify exactly what type of data is necessary.

Kean et al. (2007) indicated the benefits of being able to model the spread of invasive species. Kean et al. (2007) state that the first benefit is that the invasive species and their control efforts can be prioritised. Also decision making becomes more transparent as maps make expected outcomes explicit so that the estimation and communication of the difference made by applying control treatments can be visualised and to some extent quantified. Both control efficacy and surveillance of targeted at-risk areas can be optimised and spread models can help identify research needs, for example what sort of data needs to be collected.

1.6 Heterogeneity

Many models of organism spread have been demonstrated (e.g. Taylor 1980, Banks et al. 1988, Lewis & Kareiva 1993) over several decades, although these usually assume homogeneous environments that limit their accuracy and practicality for invasion research and management. Most methods for predicting spread have focussed on abstract environments, sometimes constrained to one dimension (e.g. Kot et al. 1996). Such lack of realism makes it difficult for researchers to make realistic predications or environmental managers to make decisions regarding species spread. Even the wrong spatial dimension of a model may even impact the model's results. For example, a patchy invasion of a predator or infectious disease can succeed in a model with two-dimensions, whereas the corresponding one-dimensional system (like many analytical models) would result in the species' extinction (Petrovskii et al. 2005). Also of importance, is that abstract models usually exclude the spatial patterning that arises from a population's interaction with the landscape. Turner et al. (1993) argue that spatial heterogeneity is an essential component for predictive models on natural landscapes.

1.7 Objectives

The overall aim of this thesis was to develop and evaluate a flexible research framework that could address the need to incorporate realistic landscapes into spread models for modelling dispersal processes within a GIS (this framework is hereafter called Modular Dispersal in GIS[1] or **MDiG**). The GIS provides the model with a heterogeneous landscape including data on geography, climate, and land cover such that the interaction and behaviour of features taken from traditional population and spread models can be investigated.

The specific aims of the thesis were:

- To develop a spread model integrated within a GIS to give a more detailed and accurate prediction of the process of invasive species spread than models that are less spatially explicit.
- Evaluate the performance of that model over the heterogeneous landscape to increase knowledge on the factors that influence spread rates.

---

[1] http://fruitionnz.com/mdig

The specific objectives of this thesis were:

1. Determine an appropriate modelling system to simulate dispersal behaviour over a realistic landscape and integrate this within a GIS.
2. Evaluate how model behaviour differs from that of traditional analytical and continuous spread models.
3. Develop methods to determine appropriate model parameters and evaluate model performance in relation to observed occurence data for a selected species.
4. Determine the extent to which stochastic models can be used to evaluate risk and predict spread rate and direction for an invading species.

1.8 Thesis structure

Chapter 2 provides an overview of dispersal modelling. Common types of dispersal pattern and the effect of spatial factors on spread are discussed. Existing modelling techniques that can be used to predict the movements of populations or individuals are compared.

Model development and its architecture are covered in Chapter 3. The model was designed in a modular manner, and each module is individually described here.

The dynamics of two spread modules within the model framework are studied in Chapter 4. Dynamics are compared to analytical dispersal model dynamics.

Chapter 5 presents a case study where the spread model is used to recreate the invasion history of Argentine ant in New Zealand. Temporal distribution data and reports from the literature were used to calibrate the model.

The overall discussion in Chapter 6 summarises the findings of this thesis and discusses how these findings address the original objectives. The future of heterogeneous spread modelling is also considered.

Chapter 2

Dispersal Models

> The whole structure of science gradually grows, but only as it is built upon a firm foundation of past research.
>
> – Owen Chamberlain

A model is a simplification of reality that represents a phenomenon through concepts and processes, often with a specific goal in mind. This goal may be to forecast or predict the behaviour of a system, or, if the underlying concepts and processes are not known or well understood, then the purpose of the model may be to gain insight into the behaviour of a process.

Models may be predominantly conceptual in nature where the model ontology is explained in ordinary language, an example of which is the Myers-Briggs personality type (Myers 1980). Alternatively, models can use mathematical and numerical methods for their interpretation and validation. We are interested in the use of these latter models and any further discussion on models refers to mathematical models.

Traditionally purely mathematical models have been the formal method of modelling, using analytical solutions that allow prediction of a system's behaviour from a set of parameters and initial conditions. Computer simulations are more commonly being used to supplement mathematical models for which closed-form analytic solutions are not possible or are overly difficult to obtain.

2.1 Simulation

Simulations aim to imitate the operations of various real world processes (Steyaert 1993) rather than solving sets of equations describing a system. A simulation attempts to represent certain features of physical or abstract systems through the dynamics of another simpler system. They can be modified and record their results to provide potential insights into how the original system being simulated behaves. Simulations are run *in silico* and are commonly orders of magnitude temporally faster than the real system, allowing them to be used for prediction and for observing system dynamics at time scales beyond the ranges normally available to human observers.

Simulation models can be deterministic or stochastic. Deterministic simulation models give a fixed output for a given set of input data and model parameters, whereas stochastic simulation models have at least one stochastic process and consequently the model output is also stochastic (Law & Kelton 1982).

Maynard-Smith (1974) surmised that "for the discovery of general ideas in ecology, different kinds of mathematical descriptions, which may be called models, are called for. Whereas a good simulation should include as much detail as possible, a good model should include as little as possible." However, simulations are now also used for the discovery of general ideas in ecology in conjunction with mathematical models.

2.2 Explicit spatial-temporal modelling

Space and time have always been recognised as crucial components when describing ecological change (Colasanti & Grime 1993) and spatially explicit models are "expected to increase our ability to accurately model populations subject to complex processes" (Balzter et al. 1998). Converting non-spatial models that can be solved analytically to models that include spatial dimensions sometimes results in equations that are intractable to solve and thus have to be simulated instead.

Invading populations are often far from their potential distribution and as such, the way the population changes will be dependent on the structure of suitable habitat which can be explicitly provided in a spatial model. Dispersal is inherently spatial and temporal as it describes the movement of individuals through space and time, as such, explicit spatial and temporal models are useful for modelling the process.

Since spatially explicit models utilise information on landscape details they suffer less from uncertainty than simpler models (for a comparison of event-based, temporally explicit, and spatially and temporally explicit models see Mooij & DeAngelis 2003). Additionally, the spatial dimension of a model can impact results. For example, Petrovskii et al. (2005) found that a patchy invasion of a predator or infectious disease may persist in two-dimensional space, when the corresponding one-dimensional system results in the species' extinction.

Spatio-temporal models have state, rate, and driving variables (Wit 1993). State variables represent the state of the system. Driving variables, or forcing functions, characterise the influence of external factors and are not influenced by the processes within the system. Finally, rate variables are the rate at which state variables change and are influenced by both state and driving variables. Formally, for a general temporally discrete model:

$$X(t+1) = F(X(t), Y(t)) \tag{2.1}$$

where $X(t)$ is the spatial pattern at time t, and $Y(t)$ is a set of variables that may affect the transition function, F. As an example, consider a population model. The population number would be the state variable, the maximum population capacity that the environment will support would be a driving variable, and the intrinsic rate of growth (per capita increase in population when there are no limiting factors) would be a rate variable.

Impact of spatial characteristics on chaotic dynamics

A chaotic system is one whose states diverge quickly if there are any small differences in initial conditions. Even though chaotic systems are deterministic, such behaviour has similarity to the highly stochastic nature of invasions and spread: If a population establishes itself in one area, then it becomes easier for a species to spread to neighbouring regions and these changes compound over time, making it possible for different invasion trajectories to diverge greatly.

Chaos can occur in non-linear systems and is thought to occur in population dynamics. In the past some have argued that chaos is maladaptive because strong population fluctuations increase a population's chance of extinction (Berryman & Millstein 1989). The inclusion of spatial components into a model can lead to more complex behaviour that may exhibit chaotic and non-linear dynamics (Silvertown et al. 1992). Conversely, introducing a spatial component to models, has been demonstrated to also have a stabilising effect (e.g. Travis 2003, Jaggi & Joshi 2001), and at wider spatial scales chaos has been shown to enhance meta-population persistance (Bascompte & Solé 1995).

An important facet of chaos is how minor differences in initial conditions diverge exponentially with time. Thus the probability of local extinction increases, but the fast desychronising rate of chaos leads to short spatial correllation lengths resulting in global stability (Rasmussen & Bohr 1987) termed *chaotic stability* or *homeochaos* (Ikegami & Kaneko 1992). An example of *chaotic stability* are the Turing structures predicted by Alan Turing's seminal paper (Turing 1952) that showed how coupling the processes of reaction and diffusion can induce the formation of stable spatial patterns despite the underlying reagents being in constant flux. Thus the inclusion of space in potentially chaotic systems can both stabilise and promote chaotic behaviour in population dynamics depending on the specifics of the system.

2.2.1 Heterogeneity and Homogeneity

Mathematical models describing dispersal often assume a homogeneous environment because of the difficulty of finding analytical solutions if it were otherwise. This led to the majority of research on dispersal, previous to this decade, ignoring heterogeneous environments. Perhaps not surprisingly, this is the same period when Geographic Information Systems (GIS) became generally accessible to scientists, both in terms of the software for creating models and the hardware required to run models in heterogeneous environments.

There are substantial benefits for using heterogeneous environments in dispersal models. Gardner & Gustafson's (2004) J-walk model, when applied to a heterogeneous environment, showed that model parameters were less sensitive to change than simpler dispersal models that assumed a homogeneous environment. The sensitivity that parameters often show in homogeneous environment models are buffered by irregular landscapes (Mooij & DeAngelis 2003).

As an example of the impact that heterogeneous landscape can have on dispersal, Schröpfer & Engstfeld (1983) found inhospitable mountain ranges formed regional barriers against the spread of the Muskrat (*Ondatra zibethicus*) in Germany. Conversely, an abundance of the Muskrat's preferred wet marsh habitat slows expansion too, since such areas can support large populations and therefore there is less drive for individuals to leave in search for more resources.

2.2.2 Scales of time and space

Grimm et al. (1996) placed emphasis on the need to consider patterns in ecological modelling in general, and others have called for explicit consideration of spatial and temporal scales in ecology (Steele 1989, Weins 1989, Levin 1992). Both these factors are now seen as an important aspect of any ecological system, and in particular more robust relationships in patterns of invasive species spread are likely to emerge at broader spacial scales (Lonsdale 1999).

Depending on the scale at which a system is modelled, the results of some phenomena and species' behaviour will have greater impacts on observable dynamics than others. Such that the model's scale will influence our choice of dynamics to model. The scale of a model also determines the level of abstraction necessary, since complex mechanisms can produce simple structure at a larger scale.

It is important to consider patterns in nature while creating ecological models, as it makes them less arbitrary, and relates them to explicit spatial scales (Grimm et al. 1996). Pattern orientated modelling (Grimm 1994) also assists in 'scaling up' from population ecology to community and ecosystem ecology (Grimm et al. 1996).

Often stated objectives for models are either prediction or understanding. However prediction without understanding represents faith in the power of heuristics and algorithms, and understanding without the ability to predict is illusion (Grimm et al. 1996). An alternative is that a model can be classified according to the point at which it departs from reality and how features are selected and aggregated to represent reality. At one extreme there are complex models that include as much detail about a system as possible. There are many arguments against the usefulness of such models. Studies by Mollison (1986), Starfield & Bleloch (1986), Wissel (1989), and Wissel (1992) all conclude that complex models are unsuitable for the development of understanding and prediction. Often the aggregation and choice of features from reality to include in the model are subjectively selected, based on the background of the modeller which can make it difficult to determine the significance of a model's results unless the modellers bias is known, particularly if the results can't be compared to a well defined performance metric.

Conversely, many classical population ecology models are very simple (such as the Lotka-Volterra model) where the desire is for generality. Levins's (1966) formula 'to sacrifice detail for generality' suggests that detail impedes the aim of generality, but at some stage a modeller must decide what biological aspects and thus how much detail should be included in their model. Too much simplification can lead to scale-free modelling, where extreme idealisation means the model loses its relationship to spatial and temporal scales. Such generic simplified models are also restricted in their ability to produce verifiable predictions (Grimm 1994).

Finally, as suggested above, patterns observed in nature can be used as the basis of a model. Where this pattern is a clearly identifiable structure in nature itself, or in data extracted from nature. Model design through patterns is essentially tautological because science tries to get to the bottom of patterns - "irregular change without pattern lies beyond science" (Grimm et al. 1996).

2.3 Models of Dispersal and Biological Invasion

With respect to dispersal modelling we are at a stage where we understand many of the underlying dynamics through mathematical models applied to real dispersal data, but the inherent complexity of real environments have made practical prediction difficult. Analytical solutions are only available for extremely simplified environments, so to develop practical predictive models simulation is an important tool.

Dispersal is an important aspect of population dynamics allowing individuals to locate new habitat and resources and it is also thought to aid in the stabilisation of populations when in heterogenuous habitats (Ruxton & Rohani 1999) or when predation causes localised extinctions (Goodwin et al. 2005). Dispersal is clearly a major factor in invasion biology - not only does dispersal allow exotic species to invade new regions, it also determines the speed that the initial population distribution will spread to occupy the full extent of suitable environment available, as determined by both abiotic and biotic factors. Furthermore, dispersal from an invader's establishment site is an important step in the process towards naturalisation and has a major influence on the threat an invasive species poses to the environment. The way that a species spreads also determines the practicality of control and eradication attempts.

The simplest method of modelling dispersal is to ignore specific spatial interactions by modeling movement as either a diffusive process (e.g. Andow et al. 1990) or as simple transfer functions (Fahrig & Merriam 1985). Such models provide a holistic view of population dynamics without the need to consider the details of individual dispersal events such that analytical solutions can be developed (e.g. Waser 1985). However these models are not useful when modelling fragmented populations or when investigating the effect of landscape pattern on the rate and direction of spread.

The effect of landscape and habitat pattern on dispersal is dependent on the species being modelled. Gardner & Gustafson (2004) explains that simple mathematical models have often been used for organisms, such as insects, that do not continuously interact with the landscape while dispersing. Whereas a raster or grid-based approach is more appropriate to simulate species that move shorter distances per unit time and interact more strongly with landscape features (e.g. vertebrates and small mammals). However, all organisms eventually interact with the landscape, even if only at the end point where the habitat will determine whether the organism will establish itself.

Both Ruxton & Saravia (1998) and Itami (1994) argue that more biological realism and more deterministic ideas are needed in modelling real world phenomena. However invasion and the process of dispersal are both strongly based on chance occurrences, such as rare long distance dispersal events that have disproportionate impacts on spread rates (Clark et al. 1998). Thus a probabilistic or stochastic method is likely to be more appropriate for studying the spread of invasive species.

2.3.1 Spatial dynamics of biological invasions

There are two components to the spread of a population. The spatial transmission of propagules and the growth of local populations. Both dynamics can have a significant impact on the rate at which a population spreads (Clark et al. 2001).

Pielou (1979) identified three kinds of dispersal: diffusion, jump-dispersal, and secular migration.

Diffusion

Diffusion is the population spread that results from the average movement of indviduals as the population grows and seeks new habitat and resources. Diffusion occurs gradually and across generally hospitable terrain. More specifically it refers to the spread that occurs from individual movement based on random-walks, which gives rise to an expanding Gaussian distribution reflecting population density (Dobzhansky & Wright 1943).

Jump dispersal

Jump dispersal is where propagules travel long distances over what is often inhospitable habitat. Often these events are rare, but still have a large impact on population spread.

Jump dispersal facilitates the spread of invasive species, as these usually radiate from multiple disjunct foci which originate from independent introductions from the species' home-range, or by dispersal with repeated establishment of the alien species from a founder population, or both (Baker 1986).

Jump dispersal is often attributed to volant (winged) organisms (Johnson 1957), but many other species also have some form of jump dispersal, particularly if they associate closely with human activity. In such cases, human-mediated dispersal may occur where humans inadvertantly transfer propagules of a species from one location to another. Often this human transfer is at a distance much further than possible through the species own dispersal mechanisms.

The spread of Oaks through Europe at the end of the Pleistocene epoch (Skellam 1951) was much quicker than possible through an expected seed shadow, as most acorns fall close to the parent, and has been termed Reid's paradox (Clark et al. 1998). Long distance dispersal, possibly by jays moving acorns up to several kilometers (Johnson & III. 1989), may account for some of this.

The frequency and distance of jump-dispersal events are stochastic, difficult to determine, and therefore have rarely been quantified (Higgins & Richardson 1999, Hengeveld 1994) but Veit &

Lewis (1996) and Suarez et al. (2001) make attempts at quantifying them for the House Finch and Argentine ant respectively. Estimates of the rate and distance of long distance dispersal events are however essential for accurate model construction, a limitation that is widely recognised (Shigesada & Kawasaki 1997).

Another example of an attempt at the quantification of jump-dispersal events has been made for zebra mussels (*Dreissena polymorpha*). The potential for over-ground dispersal between watersheds was examined by surveying rates and distances at which recreational boaters travel in Wisconsin (Buchan & Padilla 1999). Approaches such as these offer great promise in the quantification of jump-dispersal events.

Clearly, long distance dispersal is extremely influential for determining spread rate and population spread.

Secular migration

This type of dispersal is similar to diffusion, except that it takes place slow enough that appreciable evolutionary change occurs during the process. Both the environment and natural selection act on the population as the species range shifts over long time intervals. This slow movement of a population may reflect the ongoing dynamic between interacting species, the change of abiotic factors that can affect a spreading species such as climate, and the adaptation of a species to new environments.

2.3.2 Stratified dispersal

Diffusion results in short distance migration that generally has a closed front or continuous spread. Long distance dispersal however progresses more patchily and with broken fronts (Mollison 1977). Both these types of dispersal often occur together and when they do it is called *stratified dispersal* (Hengeveld 1989). Cliff et al. (1981) also distinguished between neighbourhood diffusion and long distance dispersal, but instead used the term “hierarchical diffusion” to describe the joint dispersal pattern. Stratified dispersal includes three things: the establishment of colonies ahead of the advancing population front, population growth in these individual colonies, and colony coalescence that contributes to the advancement of the population front (Figure 2.1).

While the diffusive aspect of stratified dispersal occurs via a species’ own mode of movement, long distance dispersal can occur both through human-mediated dispersal and the species own locomotion, possibly with the assistance of wind or transport down rivers. One potential factor behind differing dispersal distances could be different dispersal morphs, such as different seed size

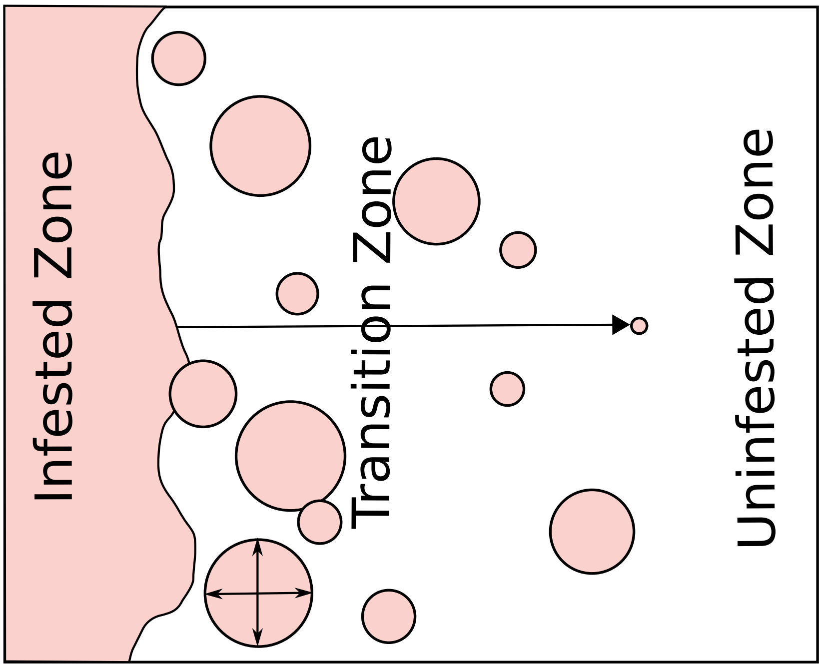


Figure 2.1: The result of stratified dispersal, with the population advancing left to right. The infested zone shows where colonies have coalesced, the transition zone is where isolated colonies establish and grow due to propagules from the infested zone, and the uninfested zone is where the species is generally absent.

in plants (Rose 1978, Southwood 1962) and wing length morphs in insects (Harrison 1980).

2.4 Effect of population growth on species spread

Population increase in a given area can occur either through individuals reproducing or through immigration from external sources, the former being reproductive growth and the latter called linear growth (Zadoks & Schein 1979).

Invasion speed is highly correlated with population growth rate and the rate of spread cannot be predicted from the distribution of dispersal distances alone, as Neubert & Caswell (2000) demonstrate with models for the plant teasel (*Dipsacus sylvestris*) which invades fields in North-Eastern United States. Similarly, Veit & Lewis (1996) found that the spread of House finches (*Carpodacus mexicanus)* throughout eastern United States, following their release from Long island in 1940, was strongly correllated with the rate of population growth near the center of their range.

Allee effects on a population can also slow the spread rate of an invasion because small populations at the invasion front are unable to increase in density without first obtaining support from the bulk of the population (Lewis 1997, Veit & Lewis 1996). Allee effects also obviously impact on the rate of establishment (Drake & Lodge 2006).

2.5 Applicable Fields

In mathematics and physics, diffusion specifically concerns change in the density of something due to uncontrolled movement such as Brownian motion. Brownian motion is random movement of a particle or individual, and results in a Gaussian distribution for the entire population if individuals or particles are all released from a central point. Diffusion is often equated with dispersal (Pielou 1969, Pielou 1977, Pielou 1979) particularly in the humanities, such as in economics (Brown 1981) but also in sciences such as human epidemiology (Cliff et al. 1981).

Diffusion is often described by the power law,

$$\text{msd}(t) \sim 6Dt^{\alpha} \tag{2.2}$$

where D is the diffusion coefficient and t is the elapsed time. Typically, in a diffusive process, the mean squared displacement (msd) of a particle is a linear function of time ($\alpha = 1$). The term *anomalous diffusion* is used to describe a diffusive process with a non-linear dependence on time. More specifically, if $\alpha > 1$, the phenomenon is called super-diffusion. In cellular biology, super-diffusion can be the result of active cellular transport processes (Caspi et al. 2002). If $\alpha < 1$, dispersing particles undergo sub-diffusion. Sub-diffusion has been proposed as a measure of macromolecular crowding in the cytoplasm (Weiss et al. 2004).

One subject in which spread models have become particularly advanced is in the prediction of wild fire behaviour. The models take into account wind direction, underlying vegetation, and even include long distance dispersal events similar to those which are so influential in invasive species spread. In wild fire spread, these long distance dispersal events are referred to as "spotting" (Xu 1994).

Spread models are also important in epidemiology - the study of heath and illness of populations. For example, *SARS* or *Severe Acute Respiratory Syndrome* spreads by common local transmission but also air passenger travel which results in occasional large distance transmission similar to the jump dispersal phenomenon mentioned earlier (Bell 2003).

These varying subjects approach spread of objects or organisms in different ways but also demonstrate dispersal processes occurring at distinctly different scales, such as local growth coupled with diffusion and a stochastic long distance dispersal process.

2.6 Partial Differential Equations

Partial differential equations (PDEs) describe a relation involving an unknown function of several independent variables and its partial derivatives with respect to those variables. They are often used to describe spatial processes and to model physical systems because they can represent the change of variables, such as population density through space and time, to the current value of the variables.

Early mathematical theory on epidemics, random migration, and distribution, was developed by Brownlee (1911). PDE descriptions of reaction-diffusion systems appeared first in population genetics (Fisher 1937, Kolmogorov et al. 1937, Dobzhansky & Wright 1943) before being applied to the spread of invasive populations by Skellam (1951) and Kierstead & Slobodkin (1953).

Skellam (1951) used simple reaction-diffusion equations to describe the spread of muskrats (*Ondatra zibethicus*) in central Europe and these models are still the most commonly used for invasive spread (Andow et al. 1990, Higgins & Richardson 1996, With 2002). Reaction-diffusion equations are based on partial differential equations (PDEs) of the general form:

$$\frac{\delta N}{\delta t} = rN + D\left[\frac{\delta^2 N}{\delta x^2} + \frac{\delta^2 N}{\delta y^2}\right] \tag{2.3}$$

where $N(x, y, t)$ is the population density at time t at point x, y on the landscape, r is the per capita growth rate, and D is the diffusion coefficient which indicates the rate of random movement across the landscape.

The philosophy behind using PDEs is that patterns at a population level do not depend on the individual behaviours of organisms and can be deduced by the statistical properties of a population of individuals. This is similar to the assumptions underlying the theory of molecular diffusion and heatflow allowing PDEs to be used in modelling physical systems. An objection could be that organisms do not follow the assumptions of independent and random movement (Holmes 1993), but Andow et al. (1990) show that simple reaction-diffusion models can still explain population-level patterns even when neglecting individual behaviour and thus such detail may not always be neccessary.

The ability of simple models to explain population-level patterns is important for successful modelling. Despite that telemetry of mobile organisms often successfully describes movement and migration patterns for a variety of species (Gautestad & Mysterud 1993, Walsh et al. 2006), this technology is impractical for smaller propagules, such as insects or plant seeds.

Mathematical analysis of the diffusion equation (Equation 2.3) indicates the velocity of the population front, V, reaches an asymptote (Okubo 1980, Okubo & Levin 2001):

$$V_{t \to \infty} = 2\sqrt{\pi \alpha D} \tag{2.4}$$

A constant velocity for population fronts has been found to apply to experimental data on the spread of many invading organisms (Hengeveld 1989, Shigesada & Kawasaki 1997). Andow et al. (1990) used a simple reaction-diffusion model with parameters taken from seperate studies and then compared them with actual distribution spread rates. The reaction-diffusion equation can give relatively good approximations for some species, such as the muskrat and small cabbage white butterfly, but not for others, such as the cereal leaf beetle (Andow et al. 1990). Species whose spread doesn't fit the reaction-diffusion model probably have their spread driven by several different modes of dispersal (Hengeveld 1989, Shigesada et al. 1995).

An important assumption for the reaction-diffusion equation is that the distances individuals move during a given length of time are drawn from a normal distribution (Kot et al. 1996, Lewis 1997). Empirical data, however, often shows leptokurtic or "fat-tailed" distributions where rare long distance dispersal events occur, in which case, an integrodifference equation with an appropriate kernel that describes the distribution of dispersal distances may be more appropriate (see section 2.7.1 below).

PDEs are generally more difficult to solve analytically than ordinary differential equations, which is why they often only consider simplified model scenarios that make assumptions such as a homogeneous environment and/or confine themselves to one dimensional scenarios.

2.7 Integrodifference equations

Integerodifference equation models (IDEs, also referred to in the literature as integro-differential equations, and integral kernel-based models), unlike reaction-diffusion PDEs which assume reproduction and dispersal occur simultaneously, break dispersal and population dynamics into separate stages. They comprise a difference equation describing population growth at each point in the landscape and an integral operator that accounts for the dispersal of organisms in space (Neubert et al. 1995, With 2002). Thus IDEs are discrete in time, while still treating space as continuous (Wang et al. 2002).

IDEs have the general form:

$$N_{t+1}(x) = \int_{-\infty}^{+\infty} k(x,y) f[N_t(y)] dy \tag{2.5}$$

where $N_{t+1}(x)$ is the population density at point x and time $t+1$. This is calculated as the integral of the new population at each point, $f[N_t(y)]$ (thus f represents a population growth function), and the number of individuals immigrating to point x from point y as determined by the integral operator k, also known as the dispersal kernel.

2.7.1 Dispersal kernels

The dispersal kernel in an IDE determines the probability that a dispersing propagule will travel a given distance and is equivalent to the "dispersion probability field" described by Hengeveld (1989).

A dispersal kernel that has the shape of a Gaussian probability distribution would essentially be the same as a reaction-diffusion PDE, but these could both underestimate the spread rate of the population. For example, the spread of Oak in Europe at the end of the Pleistocene was much faster than predicted by traditional explanations of tree life history combined with restricted Gaussian dispersal (Clark 1998). Instead Clark et al. (2001) used a leptokurtic dispersal kernel, that describe more short and long distance dispersers than a Gaussian kernel that has a comparable mean and variance. Leptokurtic dispersal kernels result in greater spread rate, and can be used to make a greater distinction between the relative number of long versus short distance dispersal events.

Other probability distributions that are often used as dispersal kernels include the exponential distribution and inverse-power functions (Okubo & Levin 2001). Weinberger (1982) showed that if a dispersal kernel is exponentially bounded (i.e. the kernel's tail reaches zero at least as fast as an exponential distribution) then the spread rate reaches a constant and the population front forms a travelling wave (Kot et al. 1996, Neubert et al. 1995), in a similar way to reaction-diffusion systems using Gaussian spread. Such studies show that the shape of the dispersal kernel, especially the tail of it's distribution, is extremely important to spread rate. Those dispersal kernels that are not exponentially bounded are called "fat-tailed", Kot et al. (1996) used an IDE to approximate the rate of spread when such a dispersal kernel is used and found that the rate was strongly sensitive to the shape of the kernel's tail. Particularly when the kernel's moments are finite, although in all cases the population distribution had an accelerating rate of spread. Moreover, Clark, Fastie, Hurtt, Jackson, Johnson, King & Lewis (1998) demonstrated that sensitivity of population spread rate to reproductive rate is increased the fatter the dispersal kernel's tail.

IDE models have also shown that the long-distance component of dispersal ultimately decides invasion speed when supported by adequate population growth, even when long-distance dispersal is rare (Kot et al. 1996, Lewis 1997, Neubert & Caswell 2000).

IDEs are similar to modelling framework used in this study as they both separate the spread of a population into discrete time steps. Additionally, one of the major modules in the framework, the **kernel** module (Chapter 3.3.2) has its conception in the idea IDE dispersal kernels.

2.8 Cellular Automata and Coupled Map Lattice models

Cellular automata (CA) and coupled map lattice models are both examples of spatially explicit population models (SEPMs). SEPMs are population-based or individual-based models of local dynamics and dispersal behaviour that often use a GIS-based description of the habitat (Turner et al. 1995) and simulate the movement of populations or individuals, thus they are distinct from PDEs and IDEs. A lattice of cells is often used to represent the population distribution, with a set of layers describing habitat and life stages with relationships between layers represented by difference equations (Dunning et al. 1995). Note that when partial differential equations are solved numerically then they could be considered similar to SEPMs in the way the use a lattice of cells to describe the population distribution.

SEPMs have been shown to be promising tools for managers to predict population responses to global change as they can easily incorporate landscape complexity (Dunning et al. 1995). Also, the numerical capabilities of modern computers allow cellular automata to be used for the analysis and simulation of spatially heterogenous ecological systems (Plotnick & Gardner 1993). Remote sensing and Geographic Information Systems (Section 3.1) have also driven the development of SEPMs by making large arrays of spatial data available as raster grids.

A CA functions by giving each cell in the lattice a particular state. The state of each cell is then updated to a new value by a function that takes the cell's current state and the states of cells in the neighbourhood of the cell being processed. As CA are based on a lattice of cells, and because all cells are processed in one step before updating them to their new state, CA are spatially and temporally discrete. The states that a cell may adopt are also traditionally discrete, such as the presence, absence, or type of species that occupies a cell, but continuously valued CA also exist, which allow for the cell's state to be a continuous variable.

CA exhibit an array of emergent behaviour unattainable by other models and easily allow for stochasticity. However robust analysis can be more difficult than for diffusion and integrodifference models. Durrett & Levin (1994*b*), Durrett & Levin (1994*a*), Levin & Durrett (1997), and Chave et al. (2002) present general guides for the use of CA in ecology and various areas of ecology have benefitted from their use, as detailed in Table 2.1.

One example of the use of CA in modelling biological spread is in modelling waves of rabies

Table 2.1: Cellular Automata have been used to model many processes in biology and ecology. Several study areas and related references are shown.

Study area	References
Cell colony growth	Eden (1961), Richardson (1973)
Plant populations with multiple modes of reproduction	Harada & Iwasa (1994), Harada et al. (1995)
Forest gap expansion	Kubo et al. (1996), Satake et al. (2004)
Competition	Caswell & Cohen (1991), Caswell & Etter (1992), Etter & Caswell (1994), Tilman et al. (1997), Durrett & Levin (1998), Buttel et al. (2002), Cannas et al. (2003)
Predation	Haswell et al. (1991)
Epidemics	Mollison & Kuulasmaa (1985), Tainaka (1988),Sato et al. (1994),Filipe & Maule (2004)
Game-theory interactions	Nowak et al. (1995), Nakamaru et al. (1996), Nakamaru & Levin (2004)

infection in foxes *Vulpes vulpes* (Grimm et al. 1996). In collection data, the spread of rabies shows 3 to 5 year intervals between waves of infection. After the first wave, where 60 to 80% of the population fall victim, further less pronounced waves occur with a 3 to 5 year period (Anderson et al. 1981, Bacon 1985, Murray et al. 1986). The driving factor behind these waves is the local temporal oscillations in fox density, and CA models were able to effectively model these oscillations and the emergent infection waves.

Although CA may be ideal for modelling the spread of a populations or indviduals over a heterogeneous environment, many CA studies do not explicitly address the speed of invasion or spread rate. Exceptions are studies by Kawasaki et al. (2006), who investigates the effect of colonisation stochasticity on the speed of invasion, and Ellner et al. (1998), who uses the pair-edge approximation for measuring population spread. This approximation technique uses the distance of the furthest forward individuals and approximates the distance travelled by using a random walk.

The inclusion of transition probabilties for change between cell states allows CA to model stochastic processes. Kawasaki et al. (2006) used such a probabilistic CA to demonstrate that the spread rate of an advancing population front is increased by stochasticity. This contrasts with the decrease in spread rate seen by adding stochasticity to one dimensional integro-difference equations (Lewis 2000, Lewis & Pacala 2000, Clark et al. 2001, Snyder 2003).

Itami (1994) argues for use of deterministic models in spatial dynamic simulations using Cellular Automata within GIS. However, because dispersal is stochastic and invasion wavefronts are almost certainly irregular, Kawasaki et al. (2006) suggest that stochastic CA models may be useful for capturing basic aspects of invasion that are difficult to explain using their deterministic or one-dimensional counterparts, such as long distance dispersal and variable habitat suitability, as indeed has been found in this study.

CA and coupled map lattice models are very similar to the modelling framework used in this study. Although this is perhaps not surprising given how general these models are. CA define state transitions based on rules, and these rules can be thought of as analogous to operations that individual modules in the framework perform.

2.9 Metapopulation models

A metapopulation is described by Grimm et al. (1996) as consisting of a number of sub-populations that live in spatially isolated islands of habitat ('patches'). The sub-populations are generally small and potentially threatened with extinction. However, the metapopulation as a whole can survive if local extinctions are cancelled out through recolonization from the remaining sub-populations.

Initially the concept of a metapopulation originates from Levins (1969), although Levins (1969) dealt with metapopulations in a scale-free manner, and ignored dispersal range. For a spatially structured population to be considered and represented as a metapopulation Hanski & Kuusaari (1995) suggest that it should meet several conditions:

- the focal species lives in spatially distinct habitat patches.
- there is no 'mainland' population, and thus all local populations have a significant risk of extinction.
- dispersal is distance-dependent and also limited by distance, which means that the spatial structure of the habitat has significant consequences for dynamics.
- the dynamics of local populations are not completely synchronous.

Metapopulation models have been used to describe mainly populations that are spread across habitats structured as remnant fragments (for a review see Hanski 1999) with relatively little work done on how spatial processes could affect a single population on contiguous suitable habitat (Travis 2003). Although it may be that the question of metapopulations on contiguous suitable habitat can be investigated using other modelling techniques such as Individual Based Models (see Section 2.11, and Law & Dieckmann 2000, Law et al. 2003) and applied geometry (Bolker et al. 2000).

Metapopulation models are used frequently and became known notably for their use for studying butterfly populations since butterflies have well recorded historical data for roost sites that are often geographically patchy. Hill et al. (1996) studied the species *Hesperia comma* and Hanski et al. (2000) modelled the migration of *Melitaea diamina* between habitat patches. The migration between patches not only assists survival of the entire population through re-colonisation of extinct patches, but is also important for geneflow. Migration can impact both the dynamics of local populations (Kuussaari et al. 1998), and entire metapopulations (Gyllenberg & Hanski 1992, Hanski et al. 1995). As with other paradigms for modelling dispersal, the distance travelled by propagules is very important. Hill et al. (1996) showed when predicting occupied patches in a metapopulation of the *Hesperia comma* that an exponential probability distribution representing dispersal distances can underestimate long-distance dispersal.

One benefit of a metapopulation model is that such model's are often based on a discrete network of patches. Limited habitat structured as distinct remnant fragments make the study area easier to define and data collection simpler. Fortunately, populations of some species naturally fit this model, or are forced into it due to patchy habitat and resources (Hanski & Kuusaari 1995).

Metapopulation models, as they don't use contiguous regions of variable habitat but distinct patches of disconnected suitable habitat, are probably the most distinct of the dispersal models from the modelling framework used in this study. Consideration of modelling metapopulations is further discussed in Appendix B.

2.9.1 Stochastic patch occupancy models

A sub-category of metapopulation model are the stochastic patch occupancy models or SPOMs (Moilanen 2004). As the name suggests, these models simply consider only the presence or absence of a population within a habitat patch, with the population size either ignored or inferred by the the size of the habitat patch. If SPOMs use discrete time then they are analogous to, and sometimes called, state transition models.

In SPOMs, the concept of the dispersal parameter represents the chance an individual will travel from one patch to another whereas *connectivity* determines the overall chance of individuals arriving from all other patches in the metapopulation network and is used to determine the chance of colonisation. The chance of extinction is greater for smaller patches but also countered by a rescue effect from connectivity (Moilanen 2004). SPOMs are often used for investigating the effect of habitat loss on species survival (Moilanen 2004) as well as the effect of regional stochasticity in environmental suitability on the same (Moilanen et al. 1998, Moilanen & Cabeza 2002).

2.10 Percolation models

In the mathematical literature Broadbent & Hammersley (1957) introduced the concept of percolation: a model of how fluid particles travel through a random medium. The medium in which the fluid particles are in, and interacts with, dictates the path of the particles.

Percolation theory is often concerned with the connectivity that exists in stochastic structures. Bond percolation employs bonds between sites that have probability p of being open, and probability $(1 - p)$ of being closed (Smythe & Wierman 1978). A network of sites created using bond percolation has the term *cluster* to describe a set of inter-connected sites, and a change of p past a threshold value (called the percolation threshold) will create a sudden increase in cluster size (With 1997). Site percolation is similar to bond percolation, but the sites are classified as open or closed instead of the connections between them.

If a percolation model was to be applied to an ecological situation, such as the connectivity of the landscape, then the potential disruption in the size of the number of connected sites is predicted to occur abruptly, at a threshold level of habitat loss and fragmentation (With 1997). The significance of such a threshold is that invasive spread may occur most rapidly and extensively above a threshold level of disturbance, which may also vary with fragmentation and the pattern of disturbance (Wiens et al. 1997, With 1997).

Percolation has been applied to simulating the competition between two species (Mollison 1987, Durrett & Levin 1994*b*), with distinct invader and defender sites (Plotnick & Gardner 1993), and in the creation of neutral landscape models called spatial correlation landscapes (Schumaker 1996).

Peng (2000) used a site percolation model to model the spread of weeds in New Zealand, and used machine-learning decision trees to generate the probability values of bonds being open or closed. These bonds indicated the relative suitability of the underlying landscape.

As percolation models can be represented as a grid, they can be emulated within the modelling framework developed in this study as discussed within Appendix B.

2.11 Individual Based Models

Complexity is sometimes added to population models through disaggregation of a population into age or size classes, and often this gives more realistic results than their simpler counterparts although at the expense of generality. Continuing to refine the population by size or age still leads to the violation of two aspects of biology. First, all individuals are different, physiologically and behaviourally, and second, interactions between individuals take place locally, which implies that the action of an individual organism is primarily affected by its local spatio-temporal surroundings (c.f. Huston et al. 1988).

The classical approach to modelling population processes aggregates populations and assumes that no significant information is lost averaging individuals. Although this may be true, the classical approach also violates the principal of locality by implicitly assuming that every member of a population has an equal influence on every other member of a population (Huston et al. 1988).

Individual based models (IBMs) consider the individual organism to be a logical basic unit for modelling ecological phenomena (Judson 1994). IBMs represent individuals explicitly and examine the dynamics of a concert of individuals, each acting on simple but biologically relevent rules.

These are two broad categories of IBMs that may be identified: vector-based algorithms to simulate movement as a correlated random walk (Kareiva 1983, Marsh & Jones 1988, Lima et al. 1999), and grid-based representations of landscapes to simulate the interaction of dispersing individuals on a landscape matrix (Allen et al. 1993, Gustafson & Gardner 1996, Wiegand et al. 1999).

It is possible to translate classical population models into IBMs. As an example, Travis (2003) reformulated the Ricker growth equation into an IBM with finite dispersal distances and interaction between individuals limited to their local neighbourhood. This IBM led to both static and dynamic spatial patterns forming, dependent on the growth rate. Limiting the dispersal distance and size of the interaction neighbourhood stabilised global population dynamics. In contrast to metapopulations described earlier this method models a single population in a contiguous landscape, rather than multiple sub-populations across a fragmented landscape. The Ricker growth equation has the potential to be chaotic, but arguably hasn't been found in data collected by experiment or in nature. Travis (2003) suggests the absence of chaotic population dynamics in experimental data is because of the stabilising effect of including spatial interations.

Gardner & Gustafson (2004) created a spatially explicit model of small mammal dispersal within a heterogenuous landscape. The model, called J-walk, considers spatial variability in prey energy and predation risk. J-walk also allows for a spectrum of individual dispersal strategies to be simulated, with each individual's movement affected by their activity level and physiological status.

Kot et al. (2004) analysed two theoretical models previously discussed by Snyder (2003) and showed how branching random-walks could act as a bridge between integrodifference equations and individual-based simulations. However, a branching random-walk formulation only appears applicable for spatially homogenous environments, whereas IBMs and other simulation-based dispersal models have their strength in modelling dispersal across heterogeneous landscapes.

In the past IBMs have been constrained by a limited numbers of individuals due to insuffucient computational resources, but the increase in processor speed and memory capacity, in concert with other developments such as parallel-processing and grid computing makes much larger populations feasible to simulate.

MDiG allows individuals to be represented within a grid, if the simulation resolution is set appropriately, such that a raster cell corresponds to the size of an individual or it's territory. At coarser resolutions, where multiple individuals could coexist within the area represented by a raster cell, SEPMs are a better analogue of a simulation in **MDiG**.

2.12 Summary

This chapter has covered several categories of model that can be used to study dispersal. Specific models however, can have peculiarities that make distinct classification difficult. Despite this, the major features for classifying a model can be reduced to several factors. Whether the model is spatially explicit or not, whether it represents time and space in a discrete or continuous fashion, and how it aggregates populations (either as presence/absence, abundance , or without aggregation in the case of IBMs), and whether the model is stochastic or not.

No class of model is better than another, but certain types of models will be appropriate depending on the questions being investigated, and the form of the data available to support it.

2.12.1 Models used for management of invasive species

Models are frequently used to provide decision support in pest management, often to determine the most efficient or cost effective solution to eradicating or controlling an unwanted species. For example Moody & Mack (1988) used a simple non-spatial stratified dispersal model to demonstrate generally that controlling distant foci is usually more important than controlling the main beachhead population, when attempting to limit the spread of an invasive species.

Similarly, population based models have determined the most cost effective method for controlling invasive populations of a intertidal wetlands grass, *Spartina alterniflora*, in Washington (Taylor & Hastings 2004). Population numbers were important in this study becase the density of the exotic species impacted the effectiveness of its control.

The early stages of an invasion is often characterised by one or a few large foci (Mack 1981, Mooney et al. 1986) which could partly be due to larger foci being easier to detect. Since these larger foci are the first detected, there often become the first and perhaps only target for control (Fenn 1980). However if smaller foci equal the same amount of space as the initally detected large foci and all foci have the same rate of spread, those smaller foci will increase their total area occupied much quicker than a single large focus (Mack 1985). This dynamic may account for the acceleration in range expansion often witnessed in terrestial plant invasions (Ewel 1986, Mack 1985) and thus highlights the importance of attempting to model spread so that attention can be focussed on areas where small foci will potentially establish. Regimes of controlling small foci has been modelled by Moody & Mack (1988) and been an important factor in the success of numerous control programmes (e.g. Eplee 1981, Watson 1985).

In other situations, such as with many terrestrial weeds, the control process can involve much the same effort across a range of plant densities. The main concern then becomes the total area of the range the invasive species occupies and not density of the species within that range (Watson 1985). The importance of area for some invasive species shifts the focus for these decision support models to accurate prediction of the species population distribution, and it's potential for spread. When the distribution of a dispersing invasive species is important for its control, spatially-explicit modelling over heterogeneous landscapes becomes unavoidable.

Chapter 3

A modular dispersal simulation framework

An objective of this thesis was to investigate and develop ways of modelling the spread of invasive species over a heterogeneous environment. This chapter introduces the simulation framework for the spread model that was designed and developed to address that objective.

The architecture of model (**MDiG**) is discussed along with its link to the Geographic Information System GRASS. The individual modules are then introduced with previous related research, and the behaviour of the module is explained. The division of dispersal and population processes into modules allows the framework to be adapted to the specifics of a particular species. One strong influence on the design of the simulation framework was the consideration of possible improvements that may extend **MDiG**'s capacity for dispersal research. These potential improvements are in Appendix B but are not directly required to understand the simulation framework.

3.1 Using GIS to model dispersal

A Geographic Information System (GIS) is a set of tools designed to manage, analyse, and process spatially referenced data that is stored digitally on computer hardware (Goodchild 1993). The benefits of a GIS are many. Spatial data can comprise a range of coordinate systems and datums (shapes representing the surface of the earth) that often make it difficult to directly compare any two sources of data. A GIS manages the process of conversion between projections and datums allowing researchers to focus on the process of data analysis and exporation rather than data conversion. Dynamically generated maps, the most common output for visualising spatial data in a GIS, are considered good communication tools allowing spatial concepts to be conveyed more easily than textual or verbal descriptions (Sequeira 2002).

Another benefit of GIS is that by connecting a GIS with Global Positioning System (GPS) instruments - tools capable of accurately recording positions on the earth's surface - it assists data collection by allowing a measurement of a variable to be easily associated with a spatial location. The combination of GIS and GPS has been used in the area of pest management for the precise tracking of aerial spraying aircraft, and has also been used to associate precise spatial locations with trap catches (Sequeira 2002). Such data, which could potentially be integrated into a real-time model, is obviously important to dispersal models.

Worner (1994) called for models of species establishment and distribution to be integrated with GIS, as even then it had become more accessible to ecologists and environmental managers. Such integration allows researchers to connect models up to heterogeneous spatial data on climate, landcover, and rainfall among others, that can influence whether or not a species will spread and subsequently establish in a given region. To date however, dispersal modellers have been slow to make full use of GIS, likely due to the computational limits at the time and difficulty in obtaining suitable data for model calibration and validation. The lack of data is still a difficulty but not always because the data does not exist, instead the data that has been collected is often kept under restrictive intellectual property clauses that impede research. Pressure on the major GIS systems to provide more simple interfaces to construct models have led to inbuilt scripting tools, and extensions such as ESRI ArcGIS's ModelBuilder (ESRI 2007).

Two example modelling frameworks integrated with GIS are MASIF (Gage et al. 2001) and BioSIM (Régnière & Nealis 2002). These both run models that were originally non-spatial, the results of which are then interpolated over a series of different points in a region. The models make use of spatial data but do not have interaction between simulation points. Clearly such pre-existing models can easily be used, however in these models true spatial dynamics are still missing.

The difficulty of constructing models, whose populations and individuals interact spatially and temporally, is that GIS are primarily designed to process the spatial aspects of reality rather than the temporal aspects. Maps within GIS are generally a snapshot of a spatial area at a particular point in time and don't indicate how the area is changing. An exception are maps that specifically represent an index of change. A sequential series of maps is often used to record temporal change, although this can become prohibitive over large spatial and temporal scales because of data storage restrictions. Data storage capacity however, is constantly increasing (Grochowski & Hoyt 1996, Grochowski & Halem 2003) allowing larger areas and longer timescales to be modelled, along with finer spatial and temporal resolution. Ideally, the extension of current data models and data structures (for an overview of data models and structures see Peuquet 1999) would allow time-variant processes to be more naturally and efficiently supported. One way this can be achieved is to use 3-dimensional raster maps[1], where slices of the 3d map represent fixed points in time.

An investigation on various species in the ecosystem of the Bornhöved Lakes District, using an Individual Based Model (see section 2.11), attempts dynamic spatial modelling in GIS (Breckling et al. 2005). Breckling et al. (2005) simulated the movement of individuals over heterogeneous GIS landscape, although not surprisingly the model is confined to a limited area of 60 ha as simulating many individuals can become computationally expensive. This is especially true for species with high reproductive rates such as insects.

There are several ways of developing a model that makes use of GIS data. Each approach varies with respect to how much it integrates with a given GIS environment. The choice of approach depends on the way in which the model makes use of the spatial data provided by the GIS, but it is also dependent on the modellers preference as well.

3.1.1 Model Linkage to GIS

There are three potential ways in which to connect a model to a GIS, which vary in how closely coupled the model is to the GIS.

Loosely coupled models

In a loosely coupled model the GIS and model remain separate but they use the same format for geographic data files. This makes the geographic information portable so that the model and the GIS can both access and modify it. In a loosely coupled model the user uses the GIS to save information relevent to the model, then the user uses the model to load the same data and carry out the process being modelled. The GIS can then be used to load the model's output for analysis and visualisation.

Closely coupled models

Closely coupled models are intermediary systems for models that need more GIS functionality than simply reading the data in a map. These models use a scripting language provided by the GIS to access GIS functionality.

---

[1] http://grass.itc.it/grid3d/

Sometimes a distinction is made between close and tight coupling (below) however the distinction is small. For example, the GRASS GIS (GRASS Development Team 2006) is driven by console commands - thus normal operations are carried out in what some users might consider a scripting language. Indeed, by combining these operations into a file that lists the relevent commands one can create a closely coupled model.

Tightly coupled models

Another name for these types of model are *integrated* models. Tightly coupled models share a common interface with the GIS, and development of a model has to be done through the use of GIS library calls. This makes it difficult to allow the model to be used in alternative GIS systems as it is specific to the GIS chosen and integrated models are essentially extending the functionality of the GIS.

The reason for tight coupling however, is that the model has more control and it can implement operations that are not a standard part of a GIS. Tight coupling also bypasses any intervening implementation layers between the model and GIS so that the model speed and efficiency tends to be greater than the other linkage methods.

A cautionary note about these models is that, if integration is based on closed monolithic systems, then there is high risk that the model will be costly (Fedra 1996). A tightly coupled model linked to a closed, monolithic GIS is at the mercy of the GIS vendor to keep the GIS maintained and any changes a vendor makes may cause compatibility issues with existing models.

3.2 Architecture and Design

MDiG was developed as a mixture of closely- and tightly-coupled components (see above) for the GRASS (Geographic Resources Analysis Support System) GIS environment (GRASS Development Team 2006). The tightly-coupled components allow direct and fast access to GIS library functions for model behaviours that required rapid execution. At the same time, a closely-coupled model manager was designed to avoid dealing with the intricacies of low-level code. GRASS lends itself to modular development, since every GIS operation is a separate program that runs within the GRASS *shell*. This shell defines information such as the current spatial region that is of interest, working resolution, the database being accessed, and other geographic meta-data. In other words the shell provides an environment in which geographic tools may be used.

Using GRASS as the underlying GIS provides some protection against the closed, monolithic systems mentioned in the previous section. GRASS is open-source software that allows any developer free access to the source code so that they can modify it - so long as they provide any changes made to other developers (Gay 2002). This philosophy means that anyone with the time and inclination can modify and improve GRASS. It also prevents users from being reliant on a piece of software that suddenly becomes antiquated or stops being supported by the vendor. Open-source software is always available to be modified and maintained, even in the unlikely event that all the current developers of GRASS left the project.

Spatial models of populations can treat both space or time as discrete or continuous (Berec 2002). In this thesis a discrete raster format was used to represent population distribution and model dispersal, because this fits more closely with pre-existing metapopulation models. The raster representation of population distribution means that the model is implicitly discrete in both space and time.

Raster maps can easily represent patch- or lattice-based models. In **MDiG**, I use the term *patch*, to refer to information across multiple map layers but within the same spatial region of a raster cell.

3.2.1 MDiG - A framework to manage simulations

MDiG consists of a piece of software referred to as the *controller* (Fig. 3.1) that runs within the GRASS GIS shell. It controls the execution of individual modules that model dispersal and other behaviours of the population. The controller also performs “house-keeping” duties such as naming maps and displaying the progress of the species distribution.

Dispersal models are described as XML (eXtensible Mark-up Language, Cowan et al. 2006) that specify module and simulation parameters (see below). The controller is responsible for updating the XML model definition with simulation results, and allows GIS commands to be carried out on all maps in a simulation replicate, maps in all replications for a given parameter set, or all maps related to the simulations defined in the XML file.

A modular approach

The model architecture is designed to be modular and extensible, to cope with species using different modes of dispersal to spread over a landscape (Neubert et al. 1995, Shigesada & Kawasaki 1997, Buchan & Padilla 1999, Suarez et al. 2001) and also to facilitate future research by allowing influences on dispersal, additional to those investigate here, to be easily incorporated and investigated.

To achieve a modular framework, the main **MDiG** software is designed to be a minimal,

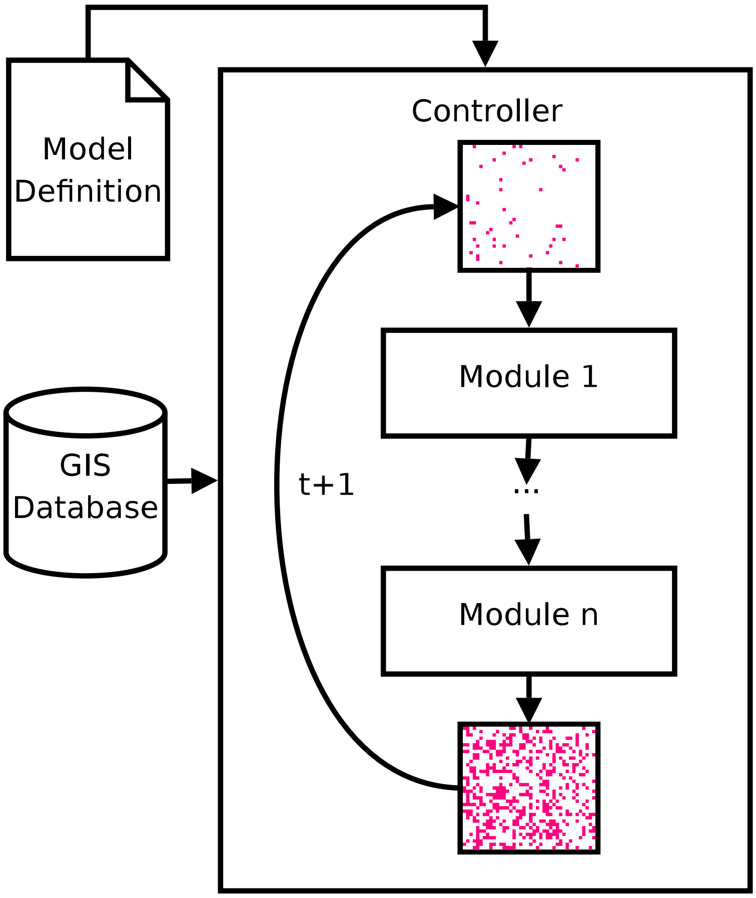


Figure 3.1: The model architecture consists of a controller that reads in a model definition and then manages a series of modules that sequentially modify a population distribution. t indicates the current time-step, and n is the number of modules that comprise the simulation.

administrative system. The dispersal dynamics are carried out by separate modules. Each module takes an input map that represents the current distribution of the species of interest and then modifies this map in a way that reflects the aspect of dispersal the module is designed to simulate. The change between the maps reflects the effect of the module during part of a timestep (since more than one module may run during a timestep, although they are called sequentially), and the length of time this step represents is left to the model designer to decide.

The modular system easily allows new dispersal methods or new aspects that affect the dispersal of a population to be implemented in **MDiG**. Such a system is more easily maintain and editing one module to add new functionality will not inadvertently introduce errors into other modules.

Simulation process

When a simulation is run, the following process is carried out (Figure 3.1):

Initially, a model definition file is read that fully describes the simulation, such as the order of the modules and the number of time-steps the model will run for. The controller takes a map of the initial species distribution and passes it to the first module specified in the definition. Once the module has finished, its output is taken by the controller and passed to the next module. This process is continued until no further defined modules are left to run. Once all modules in a timestep have been completed, the controller increments the timestep and begins the process again using the output from the final module of the previous time step as input to the first module.

At the end of each time-step, analyses may be carried out and statistics can be calculated. Alternatively, analyses can be carried out retrospectively after the simulation has been completed.

A standard for dispersal models?

The model definition file is a portable file format known as XML (eXtensionable Markup Language). The exact layout of this file is specified by an XML schema that defines where and how dispersal events, lifestages, results, and other model aspects are specified. By making this model definition format portable and open, it means researchers who don't use the GRASS can implement the model in another GIS if they choose, thus making **MDiG** interoperable.

An existing pest dispersal simulation tool called PestSpread (Overton et al. 2004) influenced the design of **MDiG** and similarly defines models using XML. Initially the PestSpread format was adhered to, but was found to be too limiting for the needs of spread models used in this study. Features not available in the PestSpread format included lifestages dependent dispersal, multiple replicates, and predefined analyses. Recent renewed interest in PestSpread[2] may see it incorporate

---

[2] J. M. Kean, pers. comm.

some of the extensions in **MDiG** and hopefully a interoperable format can be arrived at.

This file format for specifying dispersal models has been dubbed DispersalML (Dispersal Markup Language) and means that the work around **MDiG** isn't just a specific software implementation but also a proposal for a standard file format that specifies dispersal models in GIS.

3.3 Dispersal model modules

A brief summary of the modules chosen as part of this thesis, and the dispersal and population dynamics that they implement is given in Table 3.1, with in-depth detail given in the following sections. Each module can be parameterised in a number of ways, and can be included more than once (or not at all) in a time-step.

Except for the **growth** module (Fig. 3.1), all modules have the ability to work either with population based models, where the number of individuals in each patch is simulated, or with presence/absence models, where a boolean flag represents whether the species is present in that patch or not (Fig. 3.2). The **growth** module, and some options for other modules only apply to population based models and these are mentioned in the module descriptions.

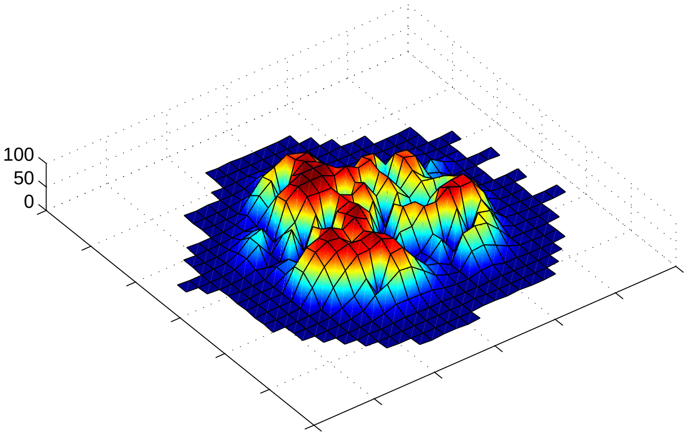
One of the benefits of creating a tightly coupled model within GIS is that the model can run at the current working resolution of the GIS environment, allowing finer scale simulation when desired. Module parameters are rarely resolution independent however (Chapter 4), so care must be taken when using these modules at different resolutions.

Another benefit is that many of the tools within a GIS are, in and of themselves, able to be used as modules. In particular the GRASS command `r.mapcalc` (Shapiro & Westervelt 1991) is able to perform map algebra that can carry out many mathematical operations across multiple maps and implement simple rules such as those that a cellular automata might employ.

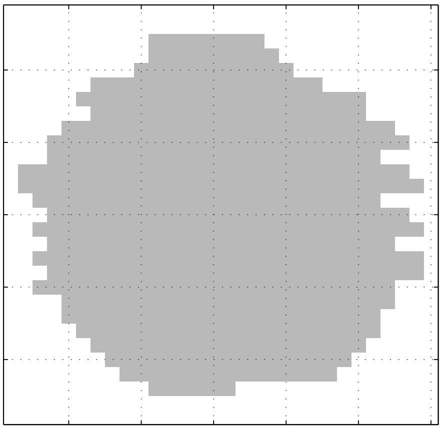
In the module sections, the module and parameters names are shown in bold font for clarity.

3.3.1 Local Dispersal

One of the most ubiquitous dispersal mechanisms seen in modelling is the contiguous diffusion of individuals. Contiguous diffusion is a result of the apparently stochastic behaviour of individuals leading to deterministic diffusion when averaged over many dispersal events (see section 2.6). The **local** module simulates this diffusion without using dispersal kernels. For every patch within the input distribution that contains individuals, the **local** module evenly spreads a proportion of the individuals to the neighbouring patches, defined below.



(a) Population model



(b) Presence/absence model

Figure 3.2: Population vs. presence/absence models. Population spread models record individual abundance in each patch (abundance on z-axis of sub-figure a.), whereas presence/absence models purely indicate whether the population exists in a patch.

Table 3.1: Model modules.

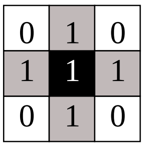
Module	**Description**
growth	Population growth that includes birth, death and density dependent processes.
local	Local contiguous dispersal allowing arbitary neighbourhood shapes. A specified proportion of individuals are dispersed.
kernel	Stochastic dispersal using a Poisson distribution to generate dispersal events and dispersal kernels to determine the distance travelled.
survival	Probability of population within a cell surviving to the following year based on an underlying suitability or survival map.

The **proportion** of individuals that spread from any patch can be specified as a parameter and which patches are considered to be neighbours of any given patch depends on the values of the **radius** and **shape** parameters. The radius parameter indicates the extent of the neighbourhood, where a value of 1 indicates that the surrounding 8 patches are in the extent of the neighbourhood, a value of 2 means the surrounding 24 patches, and so on. The shape parameter is a left to right (west to east), top to bottom (north to south), binary representation of which patches in the neighbourhood's extent are actually included as part of the neighbourhood (Figure 3.3).

Both the shape and radius parameters are not resolution independent as they refer to patches whose size change with resolution. Using the same shape and radius parameters at finer resolutions results in the rate of spread being slower since the distance between patches is smaller, whereas at coarser resolutions the rate of spread is faster. Altering the proportion of the population that spreads however, may compensate for the difference in spread between resolutions.

The **local** module provides an alternative to using a dispersal kernel with a large number of proximate events (see section 3.3.2). Dispersal kernels for contiguous dispersal are not computationally tractable. Sampling a dispersal kernel a sufficient number of times to approach average contiguous diffusion would require more computation than simply occupying neighbouring patches. Also, most literature presents only the average rate of spread giving no indication of the underlying distribution. This means it is simpler to use a model with a constant and deterministically spreading population distribution.

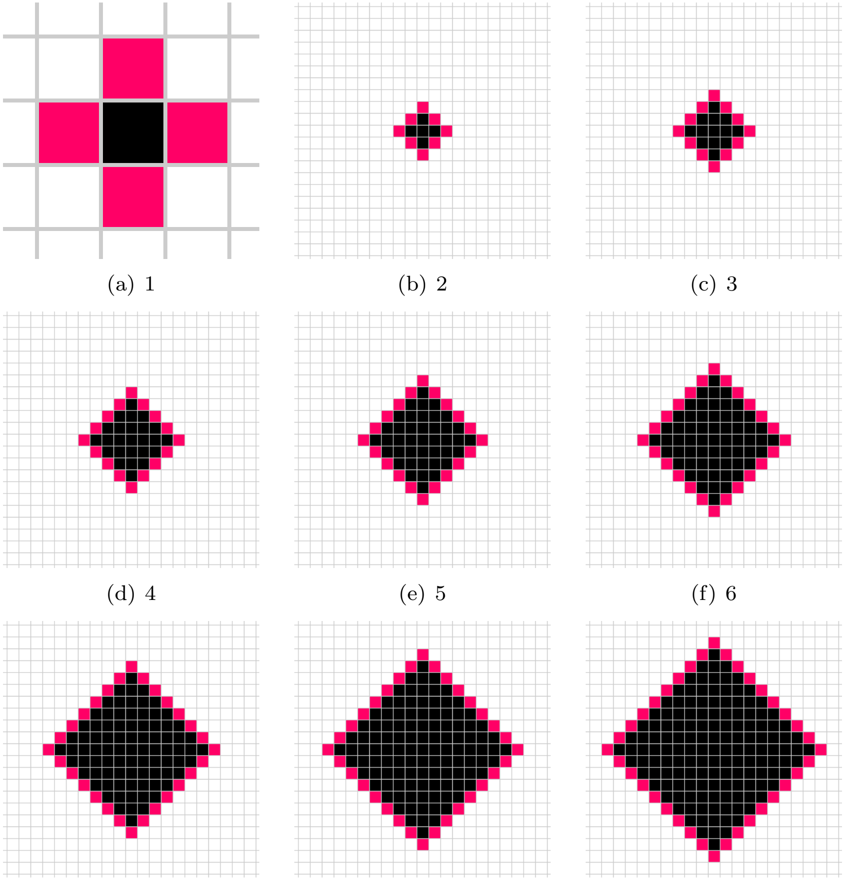
The **local** module can also represent presence or absence values instead of population counts (Figure 3.4). In this case the **proportion** parameter is discarded and the module is analogous to a two state cellular automata, where each patch is occupied in the next time-step if the neighbourhood, when rotated π radians, contains a present patch.



0	1	0
1	1	1
0	1	0

010111010

Figure 3.3: An example of a **local** module neighbourhood created using a radius of 1 and shape value of 186. This value is equivalent to the binary string `010111010`, with ones indicating what patches in the radius are in the neighbourhod. The binary string maps left to right, top to bottom, for patches in the radius. The black patch represents the current patch being processed.


(a) 1 (b) 2 (c) 3

(d) 4 (e) 5 (f) 6

(g) 7 (h) 8 (i) 9

Figure 3.4: Subsequent distributions after applying the local dispersal module with shape 186 to a presence-absence distribution for 10 time steps. The gray cells indicate newly occupied patches.

3.3.2 Kernel based dispersal

The **kernel** module represents jump and long-distance dispersal and is based on the dispersal kernels used in partial differential equations. A dispersal kernel represents the probability a dispersal event travels a certain distance (Hastings et al. 2005). Unlike the **local** dispersal module long distance dispersal usually results in non-contiguous areas of population and is therefore suitable for modelling stratified diffusion (section 2.3.2) as well as rare long distance dispersal events, depending on the kernel distribution chosen and the frequency of dispersal events.

For each occupied raster cell a Poisson process is used to generate how many dispersal events originate from it (similar to the stratified dispersal model of Shigesada et al., (1995)). A Poisson process is ideal for modelling the occurrence of random events in time (Zar 1999). The **frequency** parameter is the mean of the underlying Poisson distribution, λ, and represents the average number of dispersal events that an occupied cell generates. λ can be fitted to distribution data based on the number of new occurence sites detected in a year, divided by the number of pre-existing sites, as this gives an average of how many dispersal events are generated per site.

For each dispersal event generated, the kernel module samples the probability distribution of the kernel to determine the distance travelled by the propagule. The module accepts a **distribution** parameter for specifying the type of distribution used for these the event distances, and accepts parameters for the distribution to fit it to data on the spatial spread of the species being modelled. The model then samples a uniform distribution in the range $[0, 2\pi]$ for the angle or direction that the dispersal event travels in. The destination cell is calculated using the sampled angle and the sampled distance from the source cell and an individual propagule is transferred to that cell.

Kernel distributions

A dispersal kernel decribes the probability of a propagule dispersing a certain distance from a source population. Several kernel distributions are implemented into the kernel module. Others can be added relatively easily.

- A **generic** kernel (Clark, Macklin & Wood 1998), which can represent several distributions:

$$k(x; d, s) = \frac{1}{n} e^{-(\frac{x}{d})^s} \tag{3.1}$$

where $k(x)$ is the likelihood of finding a propagule at distance x from the propagule source,

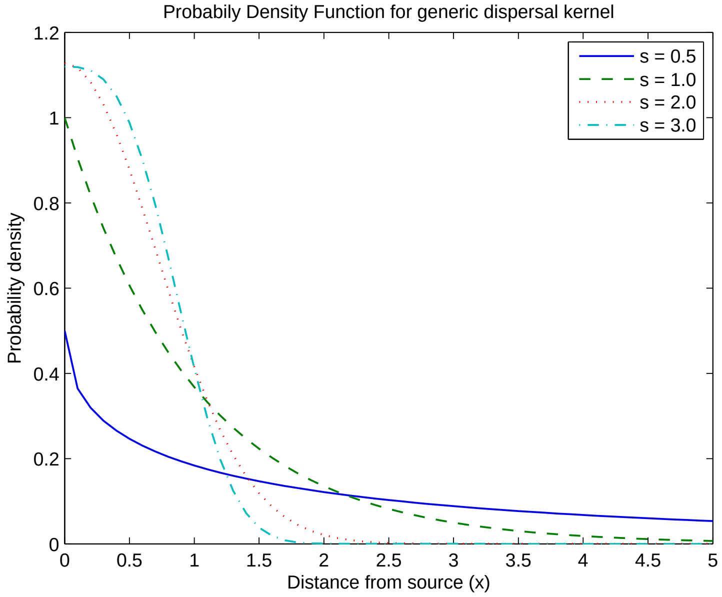


Figure 3.5: The generic dispersal kernel for several values of s. $s = 0.5$, $s = 1$ (Laplace or negative exponential), $s = 2$ (Gaussian), $s = 3$ (Ribbens).

d is the distance parameter, s is a dimensionless shape parameter, and n is a normalisation constant to make the integral of the probability function sum to 1. For one dimension (which is used when sampling the distance of events):

$$n = \frac{\Gamma(\frac{1}{s})d}{s} \tag{3.2}$$

where Γ is the gamma function.

This dispersal kernel is elegant as it can represent several distributions using one expression. Altering the shape parameter s morphs it into replicas of named distributions such as the Exponential ($s = 1.0$) or Gaussian ($s = 2.0$) distributions (Figure 3.5). However, it is difficult to integrate and invert into the form required for sampling (see Section 3.3.2, below). To overcome this, numerical methods are used for calculating integrals when sampling this kernel.

- **Cauchy distribution** - also known, among physicists, as the Lorentz distribution or the Breit-Wigner distribution:

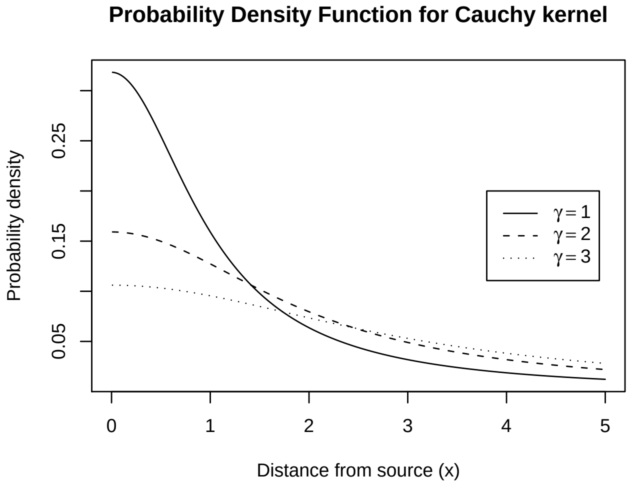


Figure 3.6: The Cauchy dispersal kernel for several values of γ.

$$k(x;\gamma) = \frac{1}{\pi\gamma[1 + (\frac{x}{\gamma})^2]} \tag{3.3}$$

where γ is the scale parameter.

The Cauchy kernel (Figure 3.6) is of particular note as it has infinite moments and can be useful for modelling distributions with rare but extreme values, such as in the case of long-distance dispersal (Shaw 1995).

- **Exponential** or Laplace distribution (Figure 3.7):

$$k(x;\beta) = \begin{cases} \frac{1}{\beta}e^{-x/\beta} & , x \geq 0, \\ 0 & , x < 0. \end{cases} \tag{3.4}$$

where β is the scale parameter.

Estimating parameters

Parameters for the chosen distribution can be estimated by recording the nearest neighbour dispersal distances between occurrences in distribution data from one year to the next (or whatever time-step is appropriate) and then fitting the probability distribution to the frequency distribution of distances.

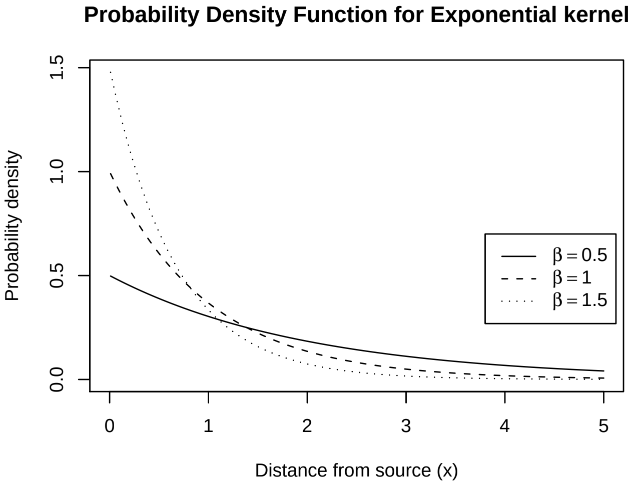


Figure 3.7: The Exponential dispersal kernel for several values of β.

Chapter 5 describes how parameters were derived for Argentine ant from its occurrence data.

Sampling dispersal kernels

Dispersal kernels are probability density functions (PDF) but for random sampling purposes the inverse of the cumulative density function (CDF) is needed. The cumulative probability can be obtained by the integral:

$$\mathrm{CDF}(n) = \int_0^n \mathrm{PDF}(x).dx \tag{3.5}$$

The CDF allows us to utilise the *Golden Rule for Sampling* which originates from a letter from John von Neumann to Stan Ulam in 1947 (Eckhardt 1987):

1. Sample a random number R from $U[0,1]$ - where $U[0,1]$ is a uniform random number in the range $[0,1]$.
2. Equate R with the CDF: $F(x) = R$
3. Invert the CDF and solve for $x : x = F^{-1}(R)$

x is random number drawn from the PDF. In this case x is the distance travelled by a dispersal event.

Impact of forcing a continuous kernel into discrete space

Dispersal events with distances that are insufficient to leave their source patch are discarded. This results in a threshold distance (that varies depending on the angle the dispersal event travels because patches are rectangular instead of circular) before an event becomes potentially effective. As the distance travelled by dispersal events is measured from the centre of the the patch, short distances represented by the distribution kernel are truncated by the boundaries of the source patch. The probability that a dispersal event is ineffective due to this truncation is given by the integral:

$$P_{<r} = 4 \int_0^{\frac{r}{2}} \int_0^{\frac{r}{2}} \frac{k(\sqrt{x^2+y^2})}{2\pi\sqrt{x^2+y^2}}.dx.dy \tag{3.6}$$

where r is the resolution of simulation, $P_{<r}$ is the probability that the event lands within the source patch, and k is the dispersal kernel. Consequently, the probability that a dispersal event is effective is equal to $1 - P_{<r}$, and is influenced by kernel parameters and simulation resolution. The consequences of dispesal kernel truncation and resolution are considered in Section 4.2.

3.3.3 Patch Survival

The **survival** module can utilise either a value, or a map containing values, as a parameter representing the probability of the population in each patch surviving to the next simulation step. For every patch that is occupied, the module samples a uniform distribution in the range $[0, 1]$ and if the value sampled is greater than the probability of survival (i.e. the sampled value indicates the patch didn't survive) then the patch becomes unoccupied. If a map is used, the survival probability is taken from the raster cell corresponding to the location of the patch being processed. The map essentially represents habitat suitability and the module reflects the difficulty that populations have establishing in unsuitable regions. The habitat suitability map can be created using various statistical and machine learning techniques along with manipulation of GIS data (see later in this section).

At the very minimum the **survival** module defines the geographic border of the model by eliminating individuals that disperse into completely unsuitable habitat, for example when terrestial species disperse into bodies of water. The module becomes more useful however, when climate, land-cover, and land-use are incoporated to construct a habitat suitability map.

In a presence-absence simulation the survival probability, whether taken from a constant value or from a map, is used as the probibility that the population in the patch will survive to the next time-step.

Using the survival probabilities directly can also be applied to population-based simulations, or optionally the probabilities can be density dependent. When a population establishes and increases in abundance within a patch, it becomes less likely that the entire patch will become extinct. This behaviour is analagous to an strong Allee effect, since beyond a certain threshold, the chance the population becomes extinct becomes zero. In a population-based simulation with density-dependence, the survival probability is given by:

$$P(Survival) = 1 - \frac{1 - s}{N^x} \tag{3.7}$$

where s is the probability of survival for the patch, N is the population size in the patch, and the parameter x defines the scaling of the extinction risk with population size. The equation is based on the extinction equation often used in metapopulation models (Hanski 1998), although in such models, patch area is used as a surrogate for population size. Hanski (1998) provides justification for this equation, and also demonstrates $s = 2r/v$ as a robust result of his investigation, where r is the intrinsic rate of population growth, and v is the variance in r.

By itself, the **survival** module could potentially result in all populations becoming extinct. However, the addition of dispersal between patches imitates scenarios in metapopulation theory that demonstrate how multiple independent sites, linked by dispersal, can survive (Levins 1969).

An alternative to stochastic extinction used in this module is to supply a map to the **capacity** parameter of the **growth** module (Section 3.3.4). The **capacity** parameter specifies the maximum population or carrying population allowed in any patch. Unsuitable habitat would have zero or very low capacity, and highly suitable habitat would have a much greater capacity.

Creating regional suitability maps

Including different sources of GIS data into a general suitability map before the simulation allows interactions between climate, terrain and other data to be calculated a priori and saves on the amount of computation required during the simulation. Using a suitability or survival map also allows the use of any method that creates such a map.

Estimation of parameters representing dispersal survival and success is difficult. Dispersal success is an important parameter in spatially explicit models that may be prone to error propagation (Ruckelshaus et al. 1997) unless the method of creating the survival map is correct. However, Mooij & DeAngelis (1999) and South (1999) indicate that error propagation in spatially explicit population models need not always be as great as Ruckelshaus et al. (1997) reports.

Suitability maps are also referred to as *potential distributions*, or *niche distributions models.*

The term potential distribution is used because physical barriers, distance, insufficient population, and other factors prevent the species from immediately occupying all suitable regions. Within the context of dispersal and the dynamics of species spread, the potential distribution can be considered the eventual end point, or goal, of the dispersal process, although it may never be reached. Niche models link back to Hutchinson's (1957) fundamental- and realized-niche concepts, by trying to construct a spatial representation of a species' niche.

There are numerous ways to create suitability maps and they can be based on a wide range of data. Some are based on correlations between pre-existing populations and variables such as climate and habitat and then these relationships are used to project a species distribution on other regions (Segurado & Araújo 2004). Others use expert opinion (Harris 2002) or phenological models (Régnière & Nealis 2002, Pitt et al. 2007) to create suitability maps.

Existing distribution data has often been used for prediction of species distributions 1) within the same region and resolution (Guisan et al. 1998, Olden & Jackson 2002, Segurado & Araújo 2004, Elith et al. 2006), 2) within the same region but at a different resolution (Araújo et al. 2005, McPherson et al. 2006), 3) between different regions (Fielding & Haworth 1995, Thuiller, Lavorel, Ara'ujo, Sykes & Prentice 2005, Randin et al. 2006, Segurado & Araújo 2006), or 4) different time periods (Austin 1992, Huntley et al. 1995, Sykes et al. 1996, Berry et al. 2002, Peterson et al. 2002, Thuiller, Richardson, Pysek, Midgley, Hughes & Rouget 2005, Araújo et al. 2006, Harrison et al. 2006).

Niche-based species distribution models have been widely used and are usually based on pattern recognition approaches, where associations between geographic occurrence and a set of predictor variables are explored to support statements about the mechanisms governing species distributions (Araújo & Guisan 2006).

Layers from the different methods of calculating potential distribution may also be combined using layer manipulation within GIS, if necessary weighting them by how much we believe them to be accurate. Some existing tools and methods for generating species distribution models are briefly described below and can be used to create suitability maps for the **survival** module (see Kriticos & Randall, (2001), for a broader and more in-depth comparison between some of the techniques):

- CLIMEX is a computerised system which describes the suitability of the climate for any poikilothermic organism and its potential for growth at any geographical location and/or time (Sutherst & Maywald 1985, Sutherst et al. 1999).

- Neural networks can be trained on existing species distributions and any potentially relevent spatial data, before being applied to new regions. Gevrey & Worner (2006) does this at a coarse scale but it is equally applicable at finer scales. Similarly, other machine-learning methods such as Genetic Algorithms, as used in the GARP (Genetic Algortihm Rule-set Production) software (Stockwell 1999), or decision-trees as used by Peng (2000).

- The lifecycle or phenology of a species might be modelled across a landscape to determine what areas can allow the complete development of the species. For insects, this can range from relatively simple degree-day models which calculate the cumulative number of degrees above a threshold temperature that occur annually (e.g. Hartley & Lester 2003), to mechanistic phenology models based on extensive laboratory development studies. Régnière & Nealis (2002) and Pitt et al. (2007) use temperature driven phenology models to judge whether a region is suitable for Gypsy moth (*Lymantria dispar*), with suitable regions being those that allow a stable seasonality to occur. Further details on the Gypsy moth phenology-based niche model applied to New Zealand are given in Appendix A.
- Where suitable distribution data is not available, expert knowledge is sometimes relied on, such as in the case of Argentine ant (*Linepithema humile*) where Harris (2002) uses literature sources to categorise the land-cover database of New Zealand into different risk levels.

A number of issues should be considered when developing suitability maps or niche-models and Araújo & Guisan (2006) outline some challenges for species distribution modellers, such as sampling methods, model parameterisation, and model evaluation strategies. Araújo & Guisan (2006) also distinguish between niche-models that yield projections of *potential habitats for species* and niche-models that incorporate spatial factors, such as dispersal, that represent the *potential geographic distributions of species.* Since dispersal is modelled separately in the simulation framework, through kernel and nearest neighbour dispersal modules, only niche-models without spatial factors should be used in **MDiG**.

Care must be taken when dealing with an invading species. If the sampled distribution data comes from a species that is still spreading, and is used in model calibration, the niche-models developed from such data may seriously underestimate the potential range of the species' niche (Araújo & Guisan 2006).

3.3.4 Growth

The **growth** module models the change in the number of individuals within a cell from one timestep to the next. Obviously, if a model is only simulating presence-absence dynamics then the growth module isn't necessary as it only works on population numbers. Although the module is named "growth" it may result in a decrease in population, depending on the parameters it is given. Such a decrease may be because of Allee effects or from a patch's population number overshooting it's carrying capacity. The **capacity** parameter indicates the maximum population size possible in a cell, and can be specified as a single global value, or as a map whose values indicate the maximum population size at each patch.

The exact dynamics of the growth module are determined by the underlying difference equation chosen and the values of its associated parameters. A number of growth equations have been described in the literature. A selection that encompasses some of those that have been used in dispersal models have been implemented in the growth module and can be selected by specifying the **function** parameter when using the growth module (Table 3.2 and see below).

Including Allee Effects

Some of the equations can exhibit an Allee effect that is characterised by a positive correlation between population density and the *per capita* growth rate when the population density is low. There are two classes of Allee effect, a weak Allee effect that either maintains population density at low levels or has only very gradual growth before higher population densities are reached (Lewis & Kareiva 1993) and a strong Allee effect that has negative population growth below a critical population density. This critical population density, $n*$, is an unstable equilibirum for deterministic models. For stochastic growth models this critical density usually manifests as an inflection point to the time to extinction: populations that are smaller than $n*$ accelerate towards extinction as they decline (Dennis 2002).

Many species have been shown to exhibit Allee effects, for example plants like the annual herb *Clarkia concinna* (Groom 1998), invertebrates such as the Glanville fritillary butterfly (Kuussaari et al. 1998), and especially sexually reproducing species where mate-finding can limit reproductive success (McCarthy 1997). The impact of Allee effects on invasive species is potentially significant due to beachhead and dispersing populations often arriving at low population densities (Lewis & Kareiva 1993). The existance of Allee effects can be exploited to assist in the eradication of alien species (Liebhold & Bascompte 2003).

Growth Equations

To describe the growth difference equations, the following variables are used. N_t and N_{t+1} are the populations of the patch before and after processing respectively, r is an index of the growth rate, and K is the carrying capacity of a patch. Either a map or a constant may be passed as

Table 3.2: Population growth functions supported by **MDiG**. Where N_t is the population at time t, r is an index of the growth rate, and K is the local carrying capacity (see relevent sections for the definition of any other variables).

Name	Equation	Reference
Logistic growth	$N_{t+1} = N_t + rN_t(1 - \frac{N_t}{K})$	Verhulst (1838)
Beverton-Holt	$N_{t+1} = \frac{rN_t}{(1+\frac{r-1}{K}N_t)}$	Beverton & Holt (1957)
Ricker	$N_{t+1} = N_t e^{(r(1-\frac{N_t}{K}))}$	Ricker (1958)
Neubert	$N_{t+1} = \begin{cases} rN_t & \text{if} N_t < a \\ 1 & \text{if} N_t > a \end{cases}$	Neubert (1997)
Wang	$N_{t+1} = \frac{rN_t^2}{1+((r-1)N_t^2)}$	Wang et al. (2002)
Keitt	$N_{t+1} = N_t + rN_t(\frac{K-N_t}{K})(\frac{N_t-C}{K})$	Keitt et al. (2001)

parameters for **growth** rate (r) and carrying **capacity** (K). If a constant is given it is applied to all patches, but if a map is provided then values are extracted from the map at the same position as the patch being processed in the distribution map. This allows for a heterogenuous environment to affect the growth and population capacity of species if sufficient supporting data exists.

Thus, an alternative way to use the suitability maps described in the section on the **survival** module, is to have them represent a patch specific supportable population. An area not deemed suitable for a species would have a very low capacity for supporting a population, and a suitable area will be able to support a large population of individuals. Since a lower supportable population will inhibit spread, this method would result in a much lower degree of patchiness (resulting from slow even spread) than that which arises from the current method of using chance mortality (resulting in areas becoming extinct, section 3.3.3).

Logistic growth model:

$$N_{t+1} = N_t + rN_t(1 - \frac{N_t}{K}) \tag{3.8}$$

where r is the intrinsic growth rate of a population.

The logistic growth model is the descendent of the unbounded Malthusian growth model. It was initially used by Verhulst (1838) as a demographic model and is an archetypal example of a simple non-linear dynamical equation that can give rise to complex and chaotic dynamics (May 1976).

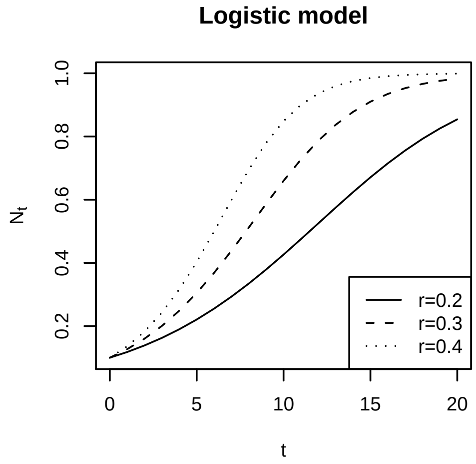


Figure 3.8: The logistic growth curve showing population density N_t (scaled so that $K = 1$) increasing with time t for varying values of the intrinsic growth rate r.

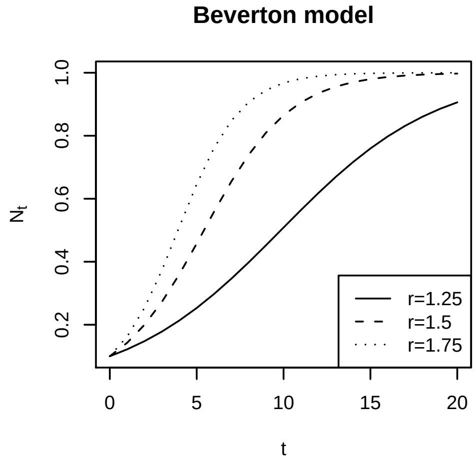


Figure 3.9: The Beverton-Holt growth curve showing population density N_t increasing with time t for varying values of the population change coefficient r.

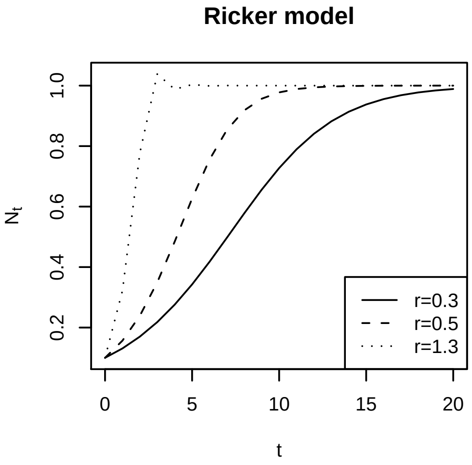


Figure 3.10: The Ricker growth curve showing population density N_t increasing with time t for varying values of the intrinsic growth rate r.

Beverton-Holt model:

$$N_{t+1} = \frac{rN_t}{(1 + \frac{r-1}{K}N_t)} \tag{3.9}$$

The Beverton-Holt model was initially introduced as a population model in the context of fisheries (Beverton & Holt 1957). The model exhibits a weak Allee effect and r indicates the population change coefficient.

Ricker model:

$$N_{t+1} = N_t e^{(r(1-\frac{N_t}{K}))} \tag{3.10}$$

The Ricker growth model (Ricker 1958) in it's general form can exhibit a rich variety of dynamics including limit-cycles and chaos (May 1974) although the consensus from a large body of research indicates that the most commonly found dynamics in nature are more towards stable equilibriums, with a few examples of stable limit cycles (Travis 2003).

Neubert model:

$$N_{t+1} = \begin{cases} Kr\frac{N_t}{K}K & \text{if} N_t < a \\ K & \text{if} N_t > a \end{cases} \tag{3.11}$$

The growth model from Neubert (1997) has a strong Allee effect if $0 \leq r < 1$, a weak Allee effect if $1 < r < \frac{1}{a}$, and no Allee effect if $r = \frac{1}{a}$. It has been used as a particularly simple example of a growth model with Allee effects to explore what these effects do to the speed of travelling waves in dispersing populations (Wang et al. 2002).

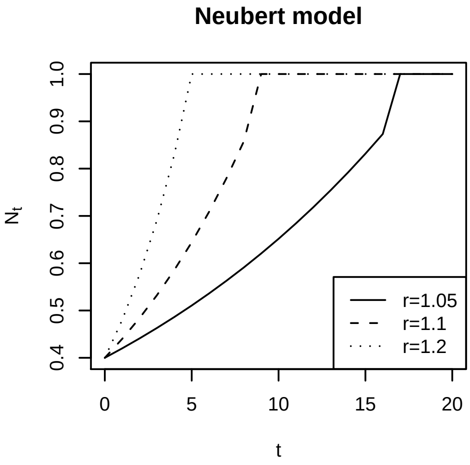


Figure 3.11: The growth curve from Neubert (1997) showing population density N_t increasing with time t for varying values of the population change coefficient r and a constant threshold $a = 0.9$.

Wang:

$$N_{t+1} = \frac{rN_t^2}{1 + ((r-1)N_t^2)} \tag{3.12}$$

where $r > 2$. The growth equation in Wang et al. (2002) has also been used for investigating the impact of Allee effects on travelling wave speed in dispersing populations.

Keitt:

$$N_{t+1} = N_t + rN_t(\frac{K - N_t}{K})(\frac{N_t - C}{K}) \tag{3.13}$$

The growth equation shown in Keitt et al. (2001) is also known as a cubic (Nagumo-type) growth function. It exhibits a strong Allee effect with C giving the threshold population above which population growth is positive.

3.4 Analysis modules

Although the type of analysis that is required by a researcher is very specific to the question being asked, several modules useful in the context of species dispersal were developed.

3.4.1 Interpreting results across simulation replicates

After running a simulation one is left with a group of replicates for each set of parameters that were simulated. Each of the replicates consists of a series of maps representing how the population

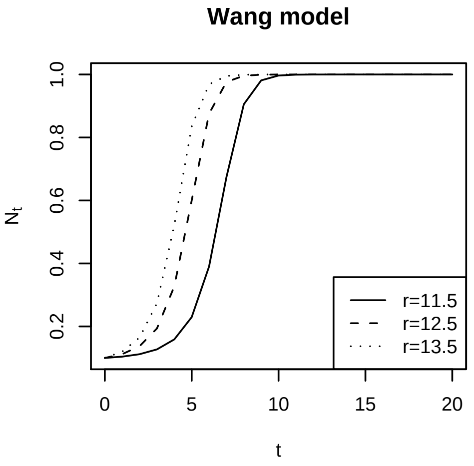


Figure 3.12: The growth curve from Wang et al. (2002) showing population density N_t increasing with time t for varying values of the growth index r.

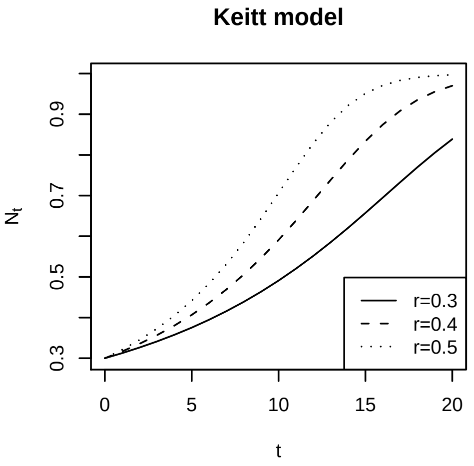


Figure 3.13: The growth curve presented in Keitt et al. (2001) showing population density N_t increasing with time t for varying values of the growth rate r.

distribution changed with time. These maps can all be processed and analysed individually, then the results are averaged across replicates for each time-step.

Alternatively, to assist analysis maps at a particular time-step across all replicates can be combined into an occupancy envelope for presence-absence models, or a mean population envelope for population models. These combined maps can be analysed directly to obtain an averaged result. Such combined maps are also helpful communicate the relative risk between particular areas indicating whether the area might harbour the dispersing population at a particular point in time.

3.4.2 Area occupied

The area occupied by a population is often used to measure the rate of a population's spread (Hengeveld 1989). Through the use of the `r.stats` module, included as part of the GRASS GIS package, the area that a population distribution encompasses can be calculated. `r.stats` can also provide data on the distribution of population counts in each patch if the model is population-based.

3.4.3 Verification and Validation

Verification is the assessment of a model's ability to fit the data used for training or calibrating the model and assessing whether it actually does what it is expected to do. Whereas model validation involves testing a model's ability to predict events using independent data. Both assessments indicate how well a dispersal model's predictions match the observed the data.

Assessment can be carried out either on individual replicate maps or on the combined occupancy envelopes or mean population maps. The analysis is carried out by comparing either distribution maps, occupancy envelope, or mean population map, from each year to all the occurrence sites that exist on or before that year. These occurrence sites are independent from data used to calibrate the model.

For individual replicates of presence-absence models

Assessing the predictive accuracy of presence-absence maps versus site data can be done using confusion matrices. Confusion matrices are constructed by comparing the presence-absence map, for each year's output, to all the occurrence sites that exist on or before that year. All sites that lie within patches that are present in the simulation are true positives, sites that lie in patches that are not present are false negatives, patches that are present but contain no occurences are false positives, and if a patch is not present and does not contain a site it is a true negative.

The measurement of predictive ability can be split into the specificity, or proportion of observed negatives correctly predicted, and the sensitivity, or the proportion of observed positives correctly predicted. The error can then be split into false positives and false negatives (Fielding & Bell 1997, Pearce & Ferrier 2000).

True negatives greatly out weigh all the other measurements since large areas will not be occupied due to the suitability maps used in **Survival**. The result is that the specificity of a model may not be useful as a metric for assessing model performance. Furthermore, occurrence data is often incomplete and contains only a subset of actual sites compounding this bias.

Occupancy envelopes and population-based models

The occupancy maps can be assessed for prediction performance either by using the same method as above, or by setting a threshold on the maps first. The reasoning behind a threshold, is that below a certain population level or chance of occupancy, it is unlikely that the species will be detected. In the case of mean population maps and occupancy envelopes, a threshold also limits the area that the model predicts is occupied, which is important in a management situation where the area requiring monitoring or eradication is an important driver of total cost. Setting a threshold focusses on the regions that have the highest chance of containing a dispersing species, theoretically making the model more precise.

3.4.4 Rate of spread

Rate of spread is an important measure of spatial population performance, as it combines both spatial and demographic processes into a single number. Previous studies that use the area occupied as an indication of the rate of invasion often assume that a distinct population boundary exists. Modelling long distance events leads to a substantially different scenario where the boundary is often stratified with patches of population ahead of the main front (Shigesada et al. 1995). When there are sufficiently few individuals or the model is at a high enough resolution then areas behind what is considered the main front may be patchy with small regions of unsuitable microhabitats that contain no individuals. This effect is exarcebated when a dispersal model is applied within a heterogenous landscape as there are likely to be many areas unsuitable for population survival, fragmenting the distribution front further.

The area occupied and general metrics for spread also tend to assume a uniform spread in all directions, averaging any directional bias. Otherwise distribution spread is assessed qualitatively, or analysed quantatively by hand (e.g Schröpfer & Engstfeld 1983).

Making the assumption the spread is uniform is not appropriate to accurately analyse spread dynamics in heterogeneous landscapes. In addition, qualitative assessment and measurement of spread-rates by hand are not feasible for analysing hundreds of replicated simulations. For this reason the module **ros** (rate of spread) was developed to calculate boundary positions using a variety of heuristics that could be robust given patchy distributions and detect spread that's potentially directed by landscape structure.

To calculate and compare the position of fronts, the **ros** module needs a centre point for all boundary algorithms. The centre point can be provided explicitly based on where an invasion is known to have originated from, or alternatively the module calculates this centre point as the mean of coordinates for all occupied cells (Equation 3.14). If used in an abundance simulation then this mean is weighted by the relative population size (Equation 3.15):

$$\overline{x} = \sum_{i=1}^{N} \frac{x_i}{N} \qquad \overline{y} = \sum_{i=1}^{N} \frac{y_i}{N} \tag{3.14}$$

$$\overline{x}_w = \sum_{i=1}^{N} \frac{w_i x_i}{N} \qquad \overline{y}_w = \sum_{i=1}^{N} \frac{w_i y_i}{N} \tag{3.15}$$

Where $(\overline{x}, \overline{y})$ and $(\overline{x}_w, \overline{y}_w)$ are the mean-centre and mean-weighted-centre coordinates respectively. (x_i, y_i) are the coordinate of the ith patch, N is the total number of occupied patches, and w_i is the weight of the ith patch, given by:

$$w_i = \frac{p_i}{\sum_{i=1}^{N} p_i} \tag{3.16}$$

Here p_i is the population of the ith patch, and the denominator is the total population of the species across patches.

The module then separates all occupied cells into a user-specified number of arcs that originate from the centre, allowing multiple directions and potentially non-uniform spread to be recognised. If only 1 arc is specified with the **arcs** parameter then this is an average boundary calculation suitable only for uniformly circular dispersal patterns. The direction in compass degrees can be specified for the first arc and subsequent arcs are spread evenly around 360°C. For each arc radiating away from the centre point, a number of metrics are calculated:

- **Distance statistics** - mean, standard deviation, average deviation, skew, and kurtosis for the moments of inertia from each occupied patch to the centre.

- **Furthest forward** - finds the furthest forward individual in each arc and uses this distance, and consequently can give very erratic rates of spread as used in Kawasaki et al. (2006).
- **Best cell boundary** - This uses the best cell classification method (Sharov et al. 1995) to place the boundary at a position which minimises the number of incorrectly classified patches. I.e. all patches behind the boundary should be present or contain a population, and all patches beyond the boundary should be absent or contain zero individuals. The method also includes a parameter to specify the relative weight of incorrect positives versus incorrect negatives.
- **Density boundary** - This heuristic takes rings around the centre point and calculates the density for each ring. The border between the rings that have the greatest change between them is used as the invasion front. In a similar way to how ecotones are boundary areas of large changes in ecology or habitat, this method finds boundary areas of large population density change.

The module can also measure the boundary as distances along parallel strips in a particular direction from the center point (as done in Sharov et al. 1995).

3.5 Summary

This chapter has outlined a framework for modelling dispersal phenomena in GIS, specifically focussing on the spread of invasive species. A selection of modules that are likely most useful and generic in application for spread models have been described, along with potential extensions that could further extend the functionality of the framework.

3.5.1 Relation to existing modelling approaches

- *Cellular Automata* - All current modules in **MDiG** can be considered variations of cellular automata (CA). However, for some modules this requires the equivalent CA to have a neighbourhood the size of the entire simulation region and running such a CA is impractical. For example, the **kernel** module, would have extremely complicated rules based on a rule that a given cell changes in relation to all other cells. Such a rule would involve each cell having a separate integral to calculate. In **MDiG**, the **kernel** module works in the opposite direction so that each occupied cell has a probability of creating a dispersal event to another site.

- *Partial differential equations* (PDEs) - **MDiG** mimics the behaviour of PDEs that are numerically solved as a grid. Numerical solutions are often used when an analytical solution cannot be found. Unlike **MDiG**, which uses discrete time and space, PDEs are continuous temporally and spatially.

- *Integrodifference equation models* (IDEs) - IDEs are similar to PDEs except that they use discrete time, and thus are more similar to **MDiG**.

- *Metapopulation models* - Each patch within **MDiG**, whether represented as a cell within a raster map, or as a shape in a vector map, can be considered a separate metapopulation with population processes occurring internally and dispersal processes occurring between patches.

- *Stochastic patch occupancy models* (SPOMs) - Like SPOMs **MDiG** can model the occupancy state of a number of populations, although patches are not necessarily separate in **MDiG**. Also, if vector maps were used to represent the patches of suitable habitat and their occupancy state (see note on vector based dipersal models below), then it would be possible to model the patch dynamics using discrete time. However, specialised SPOM modelling software such as SPOMSIM (Moilanen 2004) could likely be more suitable, since **MDiG**'s use of landscape complexity isn't necessary for such models.

- *Individual based models* - There are two broad categories of spatially explicit IBMs vector-based algorithms (Lima et al. 1999) and grid-based representations of the landscape and individual location (Wiegand et al. 1999). **MDiG** is currently a compromise between these two methods - as it uses patterns observed from diffusive models with aggregated populations to drive a model integrated with the landscape. The **MDiG** framework is extensible enough however to model both these perspectives.

3.5.2 Consideration of future direction

In the development of a broadly applicable framework for dispersal modelling, many factors and pre-existing methods of dispersal modelling were considered. All of the modules outlined had potential for extension or for alternative behaviour that may be more appropriate for some dispersal simulations, these possibilities are described in detail in Appendix B.

Two relatively major directions that **MDiG** could take are described below. First, the use of vector based data to model dispersal, and second, the potential for **MDiG** to be implemented in other Geographic Information Systems.

3.5.3 Vector based models

Creating a vector based dispersal model may be an alternative approach to consider in the future. Vector based models could take the form of an individual based model using individual points to represent individuals (e.g. Gardner & Gustafson 2004, Breckling et al. 2005), or be similar to the raster model by recording the extent of a population with a polygon. Both of these have potential challenges. Modelling a large number of individuals over a large area is computationally expensive and representing a species' distribution with polygons requires complex geometrical processes to split and join polygons as populations split and merge, respectively. Further thoughts on a model with vector maps and potential modules needed for carrying out dispersal behaviours with them are discussed in Section B.2.1. The potential use of vector maps however, has been allowed for in the design of the main **MDiG** framework.

Interoperability

MDiG has been designed with the intention that it can be implemented to support a GIS other than GRASS. Another project called VirGIS tries to abstract away from the specifics of any particular GIS (Bernard & Krüger 2000) and uses an object-orientated layered architecture to provide a common set of GIS services for model objects that can be "plugged-in". VirGIS provides these services by interfacing between a model and the specifics of a particular GIS. VirGIS is a substantially larger project than **MDiG**, and is also reliant on the VirGIS developers implementing the link to each GIS system supported. If VirGIS becomes more established, then **MDiG** could be adapted to use the VirGIS system, which in turn would allow **MDiG** to support all the GIS systems that VirGIS supports.

Each of the modules have been implemented to work with GRASS, and would need to be separately implemented in other GIS systems to allow **MDiG** to be truely interoperable. The modules in this section are those that were thought to be necessary to carry out the research in this thesis, and would likely be of general use to dispersal modellers. They are suggested as a core set of modules to be released with an implentation of **MDiG**.

Alternatively, **MDiG** can be manually reimplemented in other GIS frameworks and, so long as an implementation conforms to the **MDiG** description, could allow model definitions to be shared across frameworks.

Chapter 4

Analysing Module Dynamics

There are many aspects that one can study when using a spatial model and it is important to define what dynamics are of interest. This thesis and **MDiG** are concerned with the dispersal of an invasive organism through a region, which can be measured both in terms of the extent or area the population occupies and the speed at which it spreads.

The increase in a population distribution's area is often measured on experimental invasion data. Frequently a linear increase in the square root of area is observed (Hengeveld 1989). Other rates of spread that are based on the spatial progression of a population front are usually defined as the asymptotic velocity of a point or boundary beyond which a fixed density of individuals lie, in other words the *expectation velocity* of Mollison (1977). In analytical methods of modelling spread, these methods only work for non-linear deterministic models when the tails of any dispersal kernels involved are exponentially bounded. If faced with non-linear stochastic models, a finite velocity of furthest forward individuals known as the *furthest forward velocity* (Mollison 1977) can be used assuming the kernels used have finite variance. Such measurements are not directly applicable to simulation models, but have analagous measurements such as the iterative density method and furthest forward method described in Section 3.4.4.

Simulation models in comparison to mathematical approaches can give different results (Clark et al. 1998), so it is important to research how the simulation modules of **MDiG** behave relative to their analytical counterparts. Knowing the difference in their dynamics can help identify whether patterns in simulation observations might be a consequence of the modelling technique used and what the magnitude of any effect might be.

Analysis of module dynamics also indicates the consequences of parameter error. Ruckelshaus et al. (1997) examined the consequence of parameter error through a generalised sensitivity analysis of patch size, shape and suitable habitat for a dispersing endangered species in an endangered habitat. Ruckelshaus et al. (1997) classified possible errors as:

- Misclassification of habitat suitability.
- Incorrect estimation of how far a disperser can travel.
- Incorrect estimation of the mortality rate of dispersers.

All three of these error classes are of importance in invasive spread models simulated with **MDiG**, however this chapter primarily focusses on the sensitivity of the parameters and modules governing dispersal.

In this chapter, I first investigate the effect of the local module's dispersal neighbourhood on the total area occupied by a species. Then I investigate how the kernel module behaves relative to changes in the dispersal kernel, simulation resolution, and limited environments. The effect of these changes on both the increase in area and rate of spread are investigated.

The simulations all use presence-absence distributions to represent the extent of a population. The simulations also assume a homogeneous environment as most mathematical models, with which we want to compare **MDiG**'s dynamics, make the same assumption.

4.1 Local dispersal dynamics

When utilising the **local** dispersal module, both radius and shape parameters influence how a population of individuals spreads (the **proportion** of dispersing individuals not applicable to presence-absence models). The environment is considered homogeneous so that the pure dynamics of the module are not confounded with environmental factors.

The rate of spread as a conseqeunce of this module alone will always be limited by the radius of the neighbourhood. However, the total area occupied at any given time will involve the interaction of existing occupied patches and the number of unoccupied patches their neighbourhoods overlay.

Hypothesis 4.1 *The number of patches in the neighbourhood* and *their spatial arrangement will determine the area occupied by a population after a given time t with a point of origin at [0,0]*

We assume that the patch on which the neighbourhood is centered is always considered part of the neighbourhood. The originating patch should remain occupied because within an homogenous environment traditional diffusion doesn't create gaps in the distribution. If the model was trying to represent an individual based model - where each patch occupied represented a single individual instead of a population - then this assumption should be revised.

Table 4.1: Area of invasion as a function of time for different neighbourhood shapes when initiated from a point source. All neighbourhoods have a radius of 1, thus the maximum possible area is when all cells within the radius of 1 are part of the neighbourhood (equivalent to a Moore neighbourhood). The function for this maximal area is $A(t) = (2t-1)^2$

Occupied cells	Example shapes	$A(t)$	Saturation ($\lim_{t\to\infty} A(t)$)
2	144	t	0
3	146	$2t-1$	0
3	176, 400	$\frac{t^2+t}{2}$	$\frac{1}{8}$
4	432, 402, 178	t^2	$\frac{1}{4}$
4[a]	284	$2t^2-4t+4$	$\frac{1}{2}$
5	186, 300, 468	$2t^2-2t+1$	$\frac{1}{2}$
5	472	$t^2+\frac{t^2-t}{2}$	$\frac{3}{8}$

---

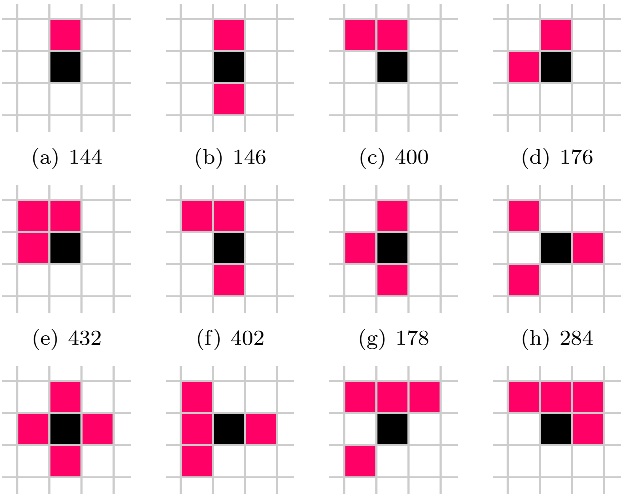
[a] Equation isn't consistant for initial condition.

A variety of shapes for a neighbourhood radius of 1 were considered (Figure 4.1) and the number of patches that the neighbourhood comprised was recorded. Any neighbourhood that can be defined at a radius of 1 can also be defined at radius > 1 by using the correct shape value (Figure 4.2), thus to prove the hypothesis it suffices to show that it is true for a neighbourhood radius of 1.

There are also many shapes that are essentially the same neighbourhood configuration but rotated $\frac{\pi}{2}$, π, or $\frac{3\pi}{2}$ radians (e.g. 466 is shape 300 rotated $\frac{\pi}{2}$ radians, see Figure 4.3). For a simulation model looking at the orientation of dispersal dynamics this distinction between neighbourhood dynamics is important, but when we are considering only the area occupied then rotationally equivalent neighbourhood shapes give the same results.

Each neighbourhood shape was simulated by the deterministic **local** module and the total area recorded at each time step, starting from a single cell occupied in an environment with practically unlimited space. From both the area recorded, and visual inspection of the distribution changes, explicit equations were derived for total area occupied after time t and are presented in Table 4.1. Each equation represents a number of neighbourhood shapes, not only their rotated homologues but also other configurations that result in an equivalent amount of occupied area.

For a **local** module with a given neighbourhood radius of r, the maximum possible occupied area at time t is:



(a) 144 (b) 146 (c) 400 (d) 176

(e) 432 (f) 402 (g) 178 (h) 284

(i) 186 (j) 300 (k) 468 (l) 472

Figure 4.1: Neighbourhood configurations with radius= 1 and various shape values. The shape values have been chosen so that none are rotationally equivalent and that they have a differing number of patches in their neighbourhood.

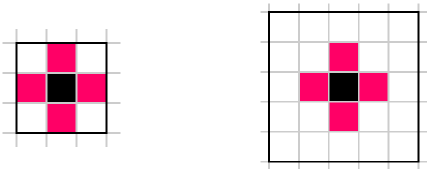


Figure 4.2: Neighbourhoods with a radius of r can represent all the neighbourhoods with radii $< r$. Here, the neighbourhood for shape 186 (left) with $r = 1$ is identical to the neighbourhood for shape 1455536 with $r = 2$.

$$A(t, r) = (2r(t-1)+1)^2 \tag{4.1}$$

This equation is based on all patches within the neighbourhood radius, r, being made part of the neighbourhood. This is equivalent to a shape value of $2^{(2r+1)}$, or a boolean string with $2^{(2r+1)}$ sequential bits of 1.

Comparing the area occupied by a particular neighbourhood shape with the maximal area allows us to calculate the percentage of possible area occupied at any time t. Graphing this percentage over time for each of the area equations in Table 4.1 shows they all tend towards a limit (Figure 4.4). We term this limit value the *saturation*, as it indicates to what degree or proportion of the maximal possible area is saturated by the population's distribution.

In support of the first part of the hypothesis - that the number of patches in the neighbourhood influence total area occupied - the saturation of area by neighbourhoods with fewer patches is

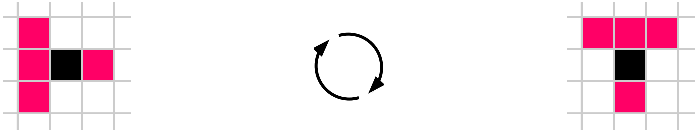


Figure 4.3: Some neighbourhood configurations are equivalent after rotation. Here 300 (left) is the same as 466 (right) if it is rotated clockwise 90 degrees ($\frac{\pi}{2}$ radians).

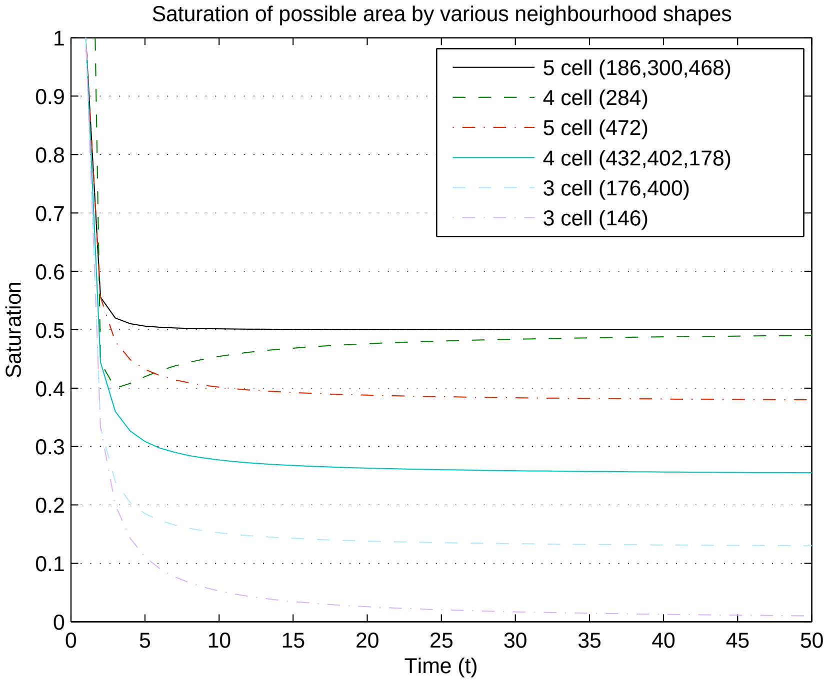


Figure 4.4: The saturation, or proportion of patches present out of maximum area possible, for different neighbourhood shapes and an initial condition of one occupied patch at origin. As model time progresses the saturation of each shape category approaches a limit.

generally less then those neighbourhoods with more patches. One exception to this is for shape 284, whose saturation is the same as the 5 patch neighbourhoods defined by shapes 186, 300, and 468; and whose area occupied is greater than the other 5 patch neighbourhood, 472, when $t > 5$.

The neighbourhood with shape value 284 is also interesting because it approaches its limit from below, unlike the other shape categories that approach their limits from above. This is likely an indication that the 4 patch, shape 284, neighbourhood is disadvantaged at the beginning in comparison with the 5 patch neighbourhood configurations, but has a more optimal neighbourhood for dispersal.

This result supports the second part of the hypothesis - that the spatial arrangement of patches in the neighbourhood also determines the area occupied. The different functions for determining area occupied by a distribution, even though neighbourhoods may have the same number of patches, further supports this part of the hypothesis.

Relevence to dispersal modelling

The relevence of this analysis to practical spread modelling is not immediately obvious as it deals with abstract neighbour shapes that do not have easily conceived real world counterparts. However, skewed shapes that depart from the traditional von Neumann (shape 186) or Moore (shape 511) neighbourhoods may be used in modelling to mimic forces that shape dispersal, such as wind or slope. Knowledge of how a change in neighbourhood shape affects the future distribution of a population helps distinguish what model dynamics arise from which module parameters and avoids false conclusions being made about the cause behind a model observation.

By connecting this module to an underlying map that describes the spread neighbourhood of each location in the region, it's possible to create a form of individual based model, or cellular automata, that has unique dispersal neighbourhoods for each population or individual (depending on the scale of the simulation).

4.2 Kernel dispersal

4.2.1 Area occupied

Here I investigate the area a population distribution spread by the kernel module occupies after each time step. The generic dispersal kernel within the kernel module was selected for this analysis as it is analogous to other dispersal kernels when given an appropriate shape (s) parameter (see Section 3.3.2). First, I determine the effect of varying the parameters, distance (d) and shape (s),

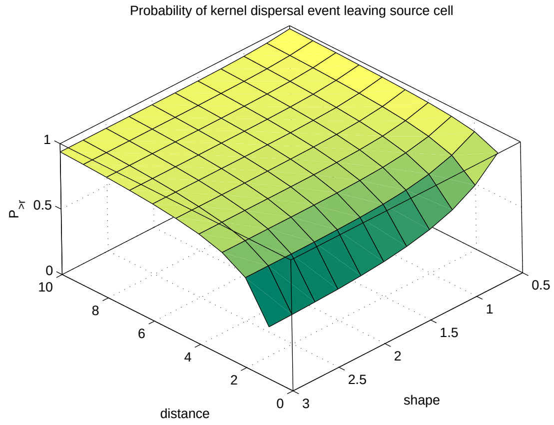


Figure 4.5: The probability of a dispersal event leaving its source cell varies as both shape and distance parameters of the generic kernel are altered. The surface shows the probability of leaving the source cell, $P_{>r}$, in a region with a resolution of 1. The probability was calculated by numerically evaluating the integral in Equation 3.6.

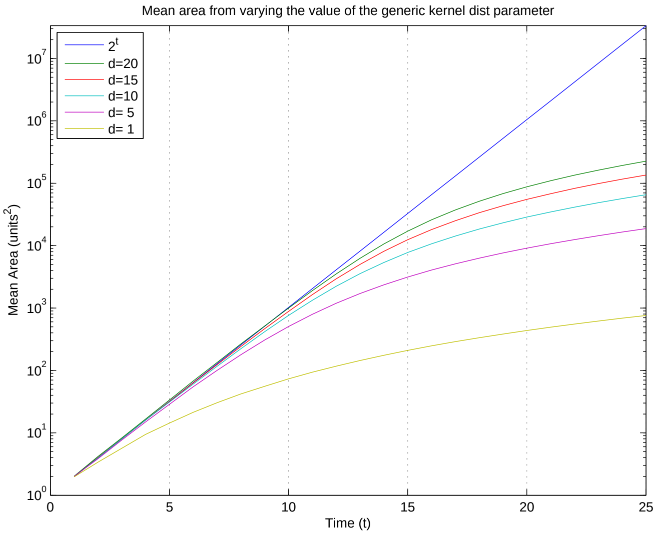
for the generic dispersal kernel, which allows the following two hypotheses to be addressed:

Hypothesis 4.2 *Increasing the distance parameter of the generic dispersal kernel will increase the total area occupied by a spreading population.*

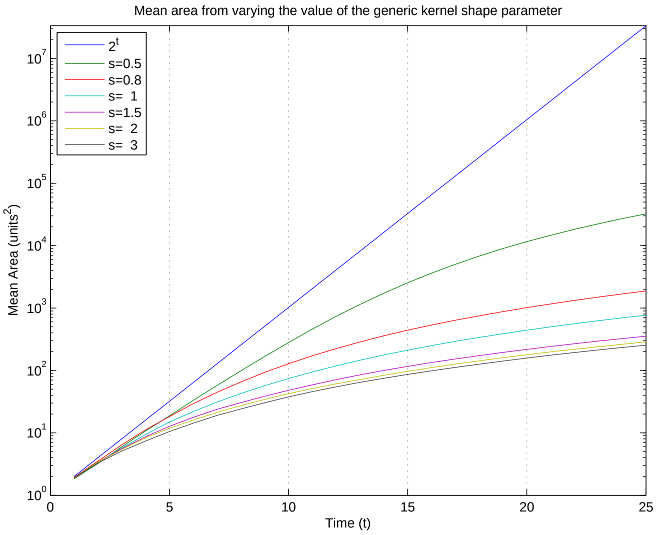
Hypothesis 4.3 *Decreasing the shape parameter of the generic dispersal kernel will increase the total area occupied by a spreading population.*

The reasons for suspecting these hypotheses to be true is that increasing d and decreasing s both stretch the kernel distribution and give rise to a greater region being encompassed by the same cumulative probability (Figure 4.5). The larger the region is, the less likely it is that the destination patch of a dispersal event will be occupied.

Values of 1, 5, 10 and 15 were used for the kernel distance parameter with the shape parameter equal to 1 ($s = 1$). The Poisson distribution for generating dispersal events had a mean of 1. Each value of d was run in the model for 25 time steps and replicated ($n = 50$). The simulation area was constrained to 2400 by 2400 cells, however very few (¡1 per replicate) dispersal events landed outside of the simulation region. For sensitivity analysis of the shape parameter, I set the distance parameter to 1 and tested shape parameter values from 0.5 to 3.0. I also varied the shape parameter over $[0.5, 1.0, 3.0]$, with the distance parameter equal to 20, to observe module dynamics with over an order of magnitude difference in the distance parameter.



(a) Mean area for varying distance values and $s = 1.0$



(b) Mean area for varying shape values and $d = 1.0$

Figure 4.6: Log of mean area occupied by a distribution versus time, for (a) 5 values of the distance parameter, d with $s = 1.0$, and (b) 6 values of the shape parameter, s with $d = 1.0$. The total expected possible area (equivalent to 2^t) is also shown, which assumes every generated dispersal event lands in an unoccupied patch.

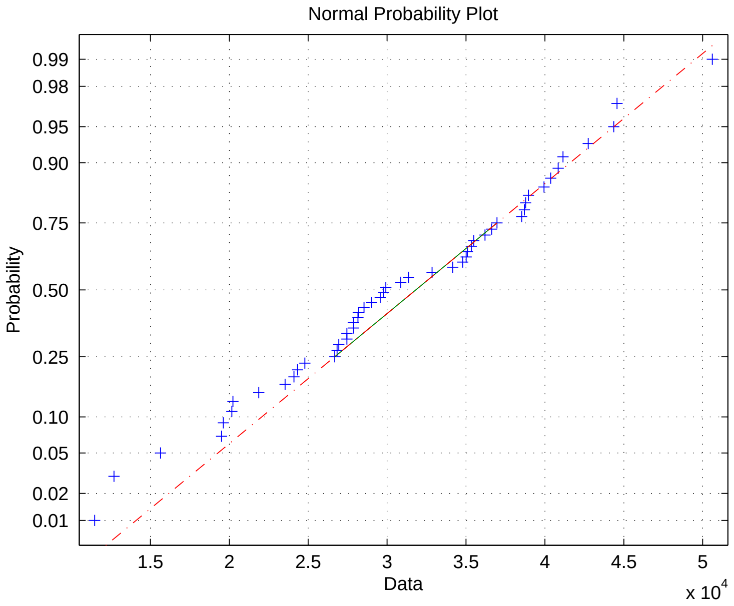


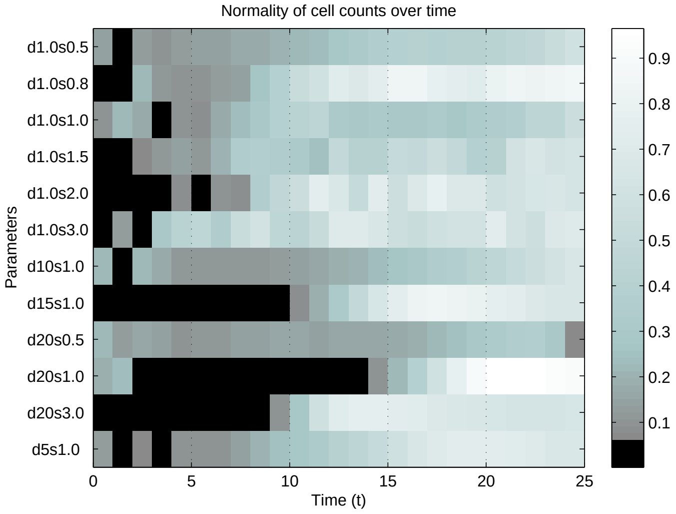
Figure 4.7: Normal plot of area data over 50 replications with parameters $s = 0.5$ and $d = 1.0$. The area values shown are those at the end of the simulation ($t = 25$).

Area distributions

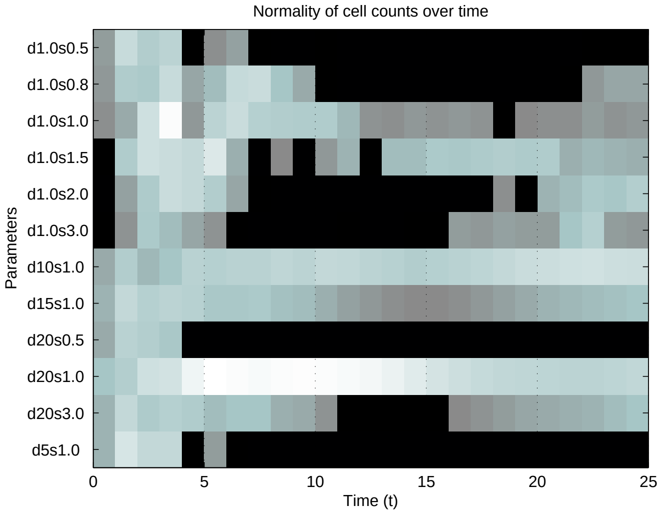
The increase in area is exponential as each existing patch is a potential source for new dispersal events. Those dispersal events which successfully establish are then also sources for further dispersal events in future time-steps.

Non-parametric distribution To accept or reject Hypotheses 4.2 and 4.3 using parametric methods the area distribution across replications for a particular parameter combination need to be normally distributed. Normal probability plots such as Figure 4.7 gave an initial indication that normality holds. The Bera-Jarque Test for normality (Judge et al. 1988) indicated that the distribution of areas is equivalent to a normal distribution ($p = 0.95$) after the simulation has run for several time steps (Figure 4.8). The minimum and maximum outliers were trimmed from the distribution before applying the Bera-Jacque test.

Early on in simulations the areas vary in relatively large amounts between replications. This variation leads to the distributions at some timesteps not being deemed normal ($P < 0.05$) and is likely due to any given dispersal event having a much greater impact on total area in the early stages of a simulation. Further on in the simulation the differences between replications average out due to the large number of dispersal events occuring. Figure 4.8(a) shows the progression of area distribution over time for the various combination of parameter values.



(a) No transformation



(b) Log-transformed data

Figure 4.8: Grid showing the timesteps when a) the distribution of area values between replications conforms to a normal distribution and b) when the area values are log transformed first. The simulations were run for 25 time steps. The shading indicates the P value for accepting the null hypothesis, i.e. the area values for a given parameter combination come from a normal distribution. Black blocks indicate the null hypothesis is rejected ($P < 0.05$).

The standard deviations of the area distributions violate the homogeneiety of variance assumption required by ANOVA (analysis of variance) for testing whether their means are significantly different. A log-transform of the area values could resolve this but as a consequence the distributions would no longer be considered normal by the Bera-Jarque Test (Figure 4.8(b)).

As the assumptions of a normal distribution and homogeneous variance could not be jointly satisfied, a non-parametric test was used for evaluating the significance between area distributions.

Statistical difference The results show that increasing the distance parameter increases the area occupied. This observation can be explained through there being less chance of a propagule arriving at an already occupied patch (because the population is more sparsely distributed). As a consequence, by increasing the distance parameter when it is already relatively large, the resulting increase in mean population area is lessened. This lessening is because the number of new patches created at each step is limited by the rate the Poisson distribution generates new dispersal events rather than being limited by a lack of nearby unoccupied patches. Similarly, decreases in the shape parameter, with the consequence of extending the dispersal kernel's tail, leads to a greater population distribution area as the chance of an occupied destination patch decreases.

Inspection of mean area plots (Figure 4.6) suggest that Hypothesis 4.2 and 4.3 are true. A Kruskal-Wallis test (Gibbons 1985) for values at the final timestep ($t = 25$) supports the visual observations indicating that shape parameter values of 0.5, 0.8, 1.0, 1.5 and 3.0 have significantly different mean areas. Shape parameter values 3.0 and 1.5 are not significantly different from 2.0, but 2.0 is significantly different from shape values lower than 1.5 (all $p < 0.05$). These tests generally support Hypothesis 4.3, that decreasing shape values lead to increases in the area occupied. Although as shape values increase the difference in total area decreases. This results in no significant difference between the means.

Similarly, all simulated values of the generic kernel distance parameter had significantly different areas at $t = 25$, supporting Hypothesis 4.2 - increasing the distance parameter results in a greater area occupied.

Resolution dependence

Another potential influence on the area occupied by a simulated population distribution is the resolution at which it is modelled. While coarse resolutions lead to every patch representing a large occupied area and thus, one assumes, a large source from which insects may disperse, the destination patch has also increased in area by the same amount so that the Poisson mean for dispersal events remains constant and independent of resolution.

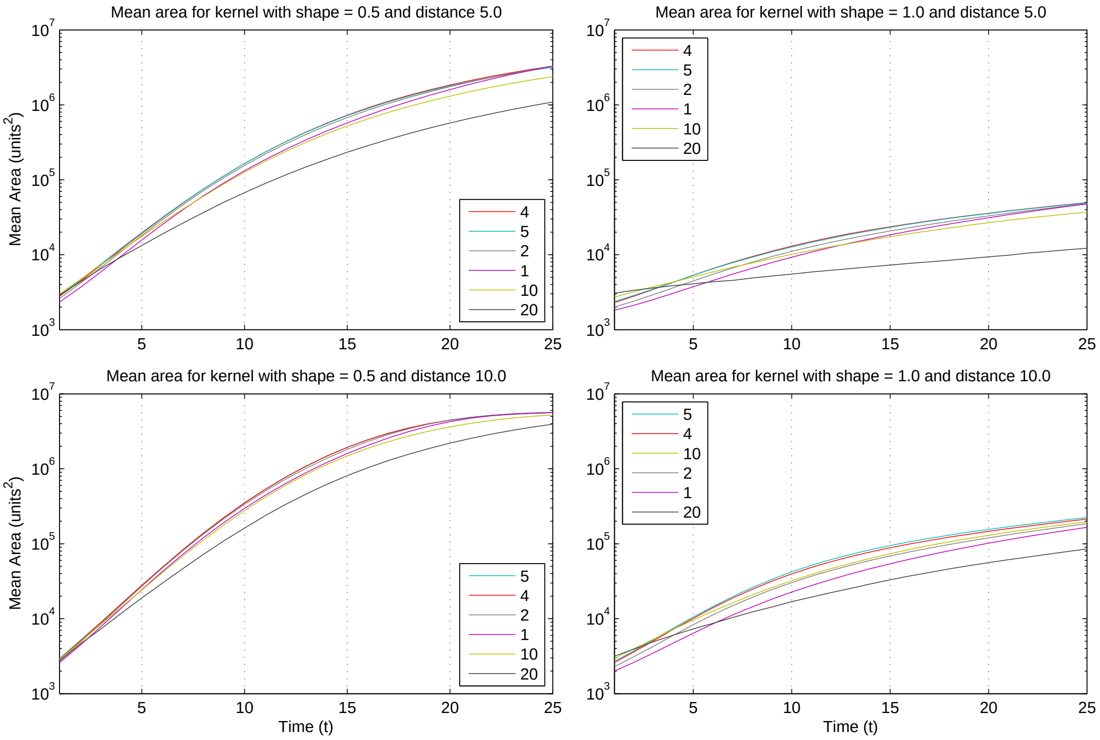


Figure 4.9: Mean area resulting from 4 sets of constant parameters in the generic dispersal kernel. Each set was simulated at resolutions that were factors of 20.

Ideally, a low resolution model would have a coarse increase in area but should be equivalent to a higher resolution model, hence Hypothesis 4.4.

Hypothesis 4.4 *For a generic kernel with set shape and distance parameters, the area occupied, A, after t time-steps will be independent of resolution.*

When working at coarse resolutions (large r) the actual area occupied is different to the number of occupied patches in the simulation and has to be converted to units^2:

$$A = N \times r^2 \tag{4.2}$$

where r is the resolution of the simulation, and N is the number of patches occupied. The greater the value of r the coarser the simulation.

Simulations for all resolutions that were a factor of 20 were replicated 50 times for all kernel parameter combinations across $s = [0.5, 1.0]$ and $d = [5.0, 10.0]$, resulting in 24 simulation configurations and a total of 1200 simulation runs. The initial map consisted of a 40×40 patch centred within a region 2400×2400 and was allowed to run for 25 timesteps. The mean of the Poisson dispersal event distribution, λ, was 1 per patch.

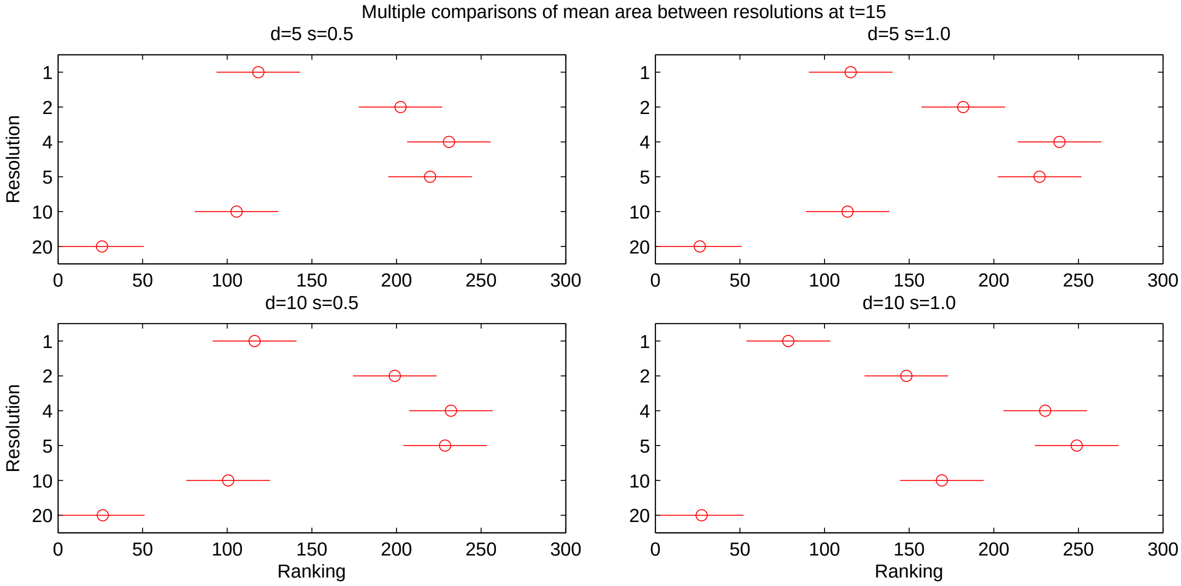


Figure 4.10: Multiple comparisons between area distributions for different resolutions at $t = 15$. Each graph is for different values of the s and d parameters of the generic dispersal kernel. The average group ranks (circles) and ranges (line) are used for comparisons with $p < 0.05$. The intervals are significance ranges, not simultaneous confidence intervals.

The results consistently show the coarsest resolution of $r = 20$ had a mean area lower than all other resolutions (Figure 4.9). The magnitude of this difference depended on the parameters of the generic dispersal kernel, although it was judged to have a significantly different median from the other resolutions using the ranks generated by a Kruskal-Wallis test with the use of multiple comparison (Figure 4.10). This difference is present for all parameter sets at $t = 15$ and this value for t was used for comparisons due to some parameter sets resulting in dispersal that was fast enough to begin being limited by the extents of the simulated area at $t > 15$.

Interestingly, the mean area values did not show a simple ordering. The areas for $r = 1$ were consistently and significantly less than $r = [2, 4, 5]$ and for the parameter set $[s = 1.0, d = 10.0]$ the mean area of $r = 1$ was also less than that of $r = 10$. Resolutions of 4 and 5 had consistently higher mean area followed by $r = 2$.

These results are at odds with Hypothesis 4.4. The reason for $r = 4$ and $r = 5$ resolutions having the greatest increases in area, are not clear although several observations can be made:

As discussed earlier (and shown in Figure 4.5), the total volume of kernel probability that falls within the source patch changes with resolution. Numerically integrating equation 3.6 for the different resolutions and parameter sets that were simulated, shows that a coarse resolution decreases the chance that a dispersal event leaves its source (Figure 4.11), which would lead to less area area occupied for coarse resolutions. This is counteracted by destination patches being much larger, so that when an event does occur the increase in area is much greater. Furthermore, when a destination patch is larger it allows those events that only just reach the boundary of a patch to occupy the destination entirely with further dispersal events occurring from the center of that patch, essentially allowing it to move further then it would otherwise. The large, sudden increases in area at a coarse resolution simulation also leads to a much larger variation in area distributions.

All these factors combine such that certain resolutions give a more rapid increase in area for certain generic kernel parameter sets. Further work needs to be done to rigorously describe the relationship between the increase in area and resolution.

The rapid increase at the start of the simulation (i.e. the initial area at $t = 1$) for coarser resolutions is because the surrounding area has no occupied patches but a single successful dispersal event leads to more area being occupied. As time progresses however, it is easier for new events at fine resolutions to fill in missing patches. Previously it has been thought that the number of foci is more important than their size in stratified diffusion for influencing the rate of expansion in a population (Mack 1985). In that case we'd expect the finest resolution to be optimal. However, these results indicate that there is actually an optimal patch size that isn't the finest resolution for dispersal in a homogeneous (but discrete) environment, when we assume that destination patches are the same size as their source. This is obviously an effect of discretising the environment, but could influence real-world populations that are located in patchy environments.

Limited environmental space The previous experiments used large regions (2400 by 2400 units) to prevent border interactions with the simulation region from overly influencing the results (extreme long distance events inevitably reach outside of the simulation region however). How the dispersal kernel behaves within an environment limited to 200 by 200 unit sized patches is now investigated. A smaller environment makes it possible for a dispersal simulation to completely saturate the region by the end of the simulation. Clearly, for any finite environment over a long enough period, the dispersal kernel will eventually occupy all the patches, although this period may be extremely long depending on the size of the environment. Confining the dispersal kernel to a smaller environment allows the effect of saturation on spread to be investigated.

A restricted environment increases the likelihood that, as spread progresses, more dispersal events will arrive outside of the boundary and hence be lost. This reasoning suggests the following hypothesis:

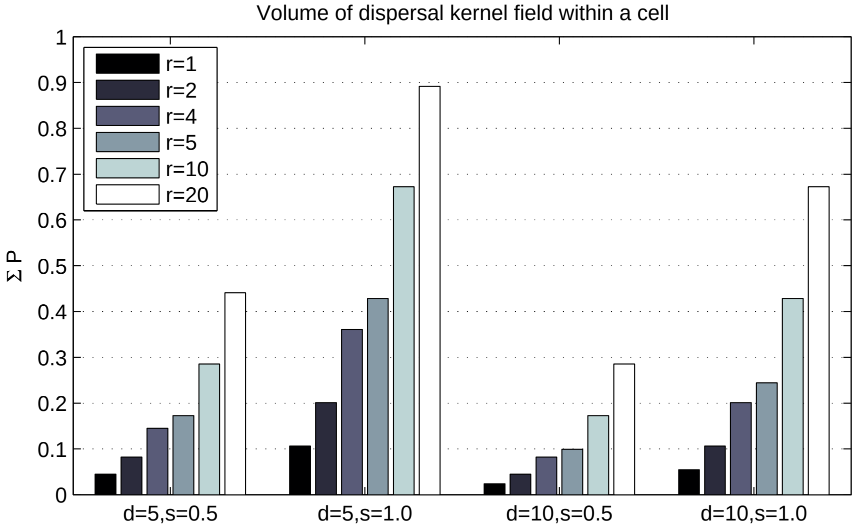


Figure 4.11: The total probability of the dispersal kernel that falls within the source cell for a dispersal event. The total probability volume encompassed by a cell depends on both resolution and kernel parameters.

Hypothesis 4.5 *A limited area will slow the increase in occupied patches relative to the increase in area for a much larger region.*

The following additional hypothesis arises from the observation that spread via the dispersal kernel is an exponential process. When an exponential process such as growth has a limited amount of resource its increase follows a logistic function or S-curve. With respect to dispersal in a limited environment, available patches become scarce and the increase in occupied patches is expected to resemble a logistic function.

Hypothesis 4.6 *When the population is restricted to a limited area, the increase in occupied patches due to the kernel module will follow a logistic rather than exponential growth process.*

To test these hypotheses a region 200 by 200 patches was simulated using a generic dispersal kernel with all four combinations of parameters $d = [1, 10]$ and $s = [0.5, 1.0]$ for 50 time-steps. Additionally, a control experiment was run for each combination of parameters where the region was made to be essentially unlimited. For the control the region was the same as the other kernel experiments (2400 by 2400). These control experiments were run until the area occupied reached 200^2 units^2 or more - in other words, the total area of the limited region.

The results showed that all combinations of parameters, except $d = 1, s = 1$ reached saturation. Those dispersal kernels that were spread over a greater distance (greater probability of larger dispersal distances) have a greater chance of dispersing further and less chance of landing in an already occupied patch (in other words can escape a cluster of already occupied patches), which resulted in their reaching saturation of the region more rapidly (Figure 4.12).

Interestingly, the most diffusive kernel parameters with $d = 10, s = 0.5$ initially increased in occupied patches slower than the kernel with less diffuse parameters $d = 10, s = 1.0$ but then pass each other after the inflection point (Figure 4.12). This is because dispersal events that land outside of the region early in the simulation are lost. The difference can be observed in the equivalent control simulation where these events are not lost and in which the increase in area was substantially quicker.

All parameter combinations that reach saturation also had their control equivalents exceed them in area in accordance with Hypothesis 4.5. The point at which this occurs not only depends on the combination of dispersal parameters but also on the size of the limited environment because the kernel distance parameter, d, scales the probability density function of the dispersal kernel. Scaling the d parameter with the size of environment will give the same temporal increase in the proportion of patches occupied.

For $d = 10, s = 1.0$ the control has less occupied area than the limited environment version until the inflection point at $t = 19$, this difference is not significant however.

The difference between the limited environment simulation and the equivalent control simulation reflects the number of misses or dispersal events that land outside of the limited environment. However this difference is greater than the number of misses, as dispersal events that are successful in establishing in the control equivalent feedback into the total occupied area by generating new dispersal events themselves.

In accordance with Hypothesis 4.6 the increase in area occupied clearly displays the shape of a logistic growth process (Figure 4.12) with slow initial growth before accelerating past the midpoint. The increase in occupied area slows as space becomes limited and then finally reaches a plateau once the entire environment is occupied.

Logistic growth in area of the distribution of a species has been seen in the introduced Red Deer populations around Nelson and the upper South Island of New Zealand. These populations showed a sigmoidal expansion rate in area, with the plateau attributed to the distribution reaching the surrounding sea and an already extensively occupied area filling in (Clarke 1971).

Mollison (1987) describes that at broad scales of biological invasion, the increase in distribution range can be partitioned into periods: an early period of pioneering, a middle period with fastest change, and a later period of condensing saturation or fill in (Shigesada et al. 1995). Ewel (1986) suggested the lag at the beginning to be because of a lack in a suitable habitat becoming available for the invading species or low probability of new propagates striking the right combination of

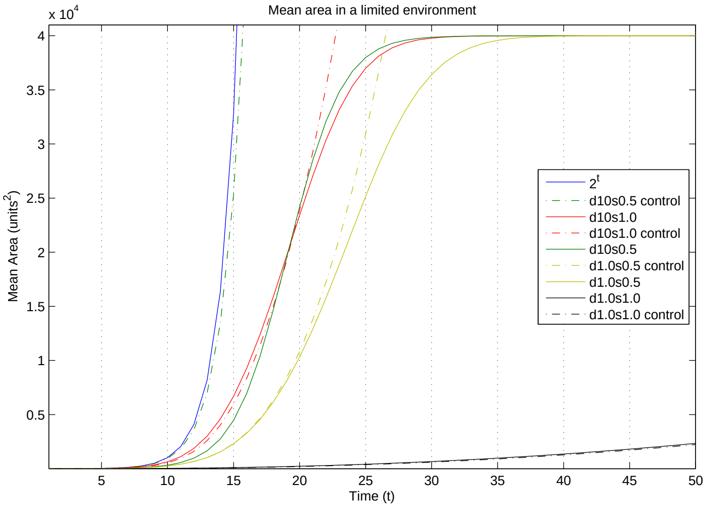


Figure 4.12: Mean area with varying s and d dispersal kernel parameters. Solid lines indicate the mean area in a limited space of 200 by 200 cells, the dashed lines are the control simulations in practically unlimited areas.

environmental conditions. However our results show it could simply be from a lack of sources for dispersal events. An explanation based on environmental heterogeneity is not needed to explain this lag period (although a limited environment, heterogeneous or not, is needed to explain the saturation that occurs).

4.2.2 Spread rate

The total area occupied is just one measure of the spatial distribution of a species. Another measurement of interest is the distance at which occupied patches of the distribution are located relative to the origin of spread. The Rate of Spread module (*ros* module, 3.4.4) calculates several measures of dispersal and spread. Including the statistics on the actual values for distance to each patch, and boundary calculation methods that define the front of a dispersing distribution.

The results from testing Hypotheses 4.2 and 4.3 were used to create similar hypotheses for the affect of the shape and distance kernel parameters on the distance to the distribution front and its rate of spread.

Hypothesis 4.7 *Decreasing the shape parameter will increase the rate at which the distribution front increases.*

Hypothesis 4.8 *Increasing the distance parameter will increase the rate at which the distribution front increases.*

Note that the definition of the distribution front can be determined using several different methods each of which could result in separate conclusions about the hypotheses.

Analytical models often focus on which conditions are required for a front to travel at a constant velocity and provide ways of calculating the velocity from dispersal kernel parameters. Such models also specify how long it takes for the constant velocity to be reached. Some models have accelerating fronts, specifically those that have a distribution kernel with a tail greater than exponential ($s < 1$ for the generic dispersal kernel). Whether these cases apply to our simulation model are explored in the following sections.

Hypothesis 4.9 *The distribution front will exhibit the same velocity and acceleration as predicted by analytical models.*

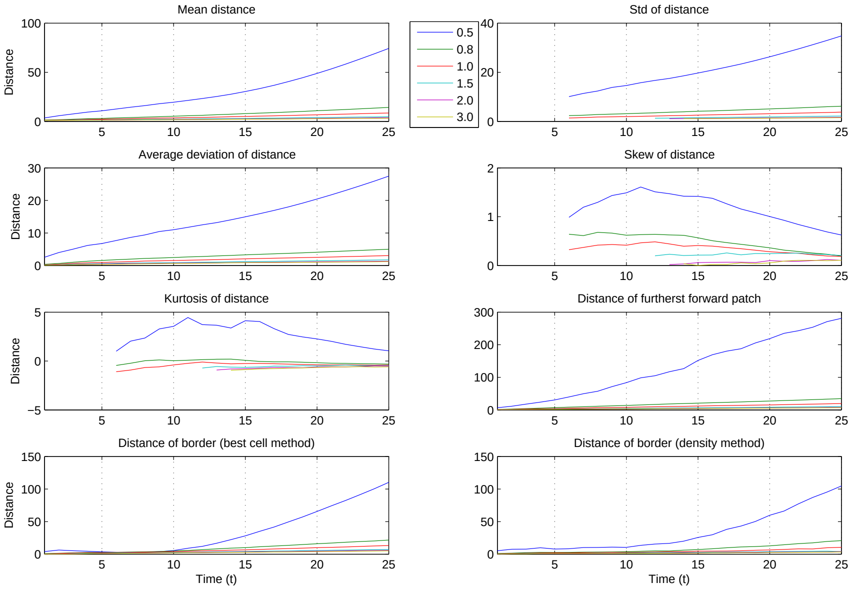
Several measures for the rate of spread are defined. First, the mean distance is the average distance to all occupied patches from the point of origin, and the mean distance velocity is the rate at which mean distance increases from one time-step to the next. The furthest forward distance is the distance to the patch that is furthest from the point of origin. The last two measures for rate of spread, *best-cell* and *density* are based on finding the border for the distribution using iterative methods and comparing the difference in distance between time-steps (For more details, including references, see section 3.4.4).

The simulation output from the area analysis was used. The mean of relevent values over multiple replications were calculated so that average behaviour could be assessed. Thus the mean distance for one replication is the mean distance of patches to the origin, and this mean value is again averaged across replications. The distribution of the mean distance was expected to be normal due to the relatively large number of replications (50) and the Central Limit Theorem.

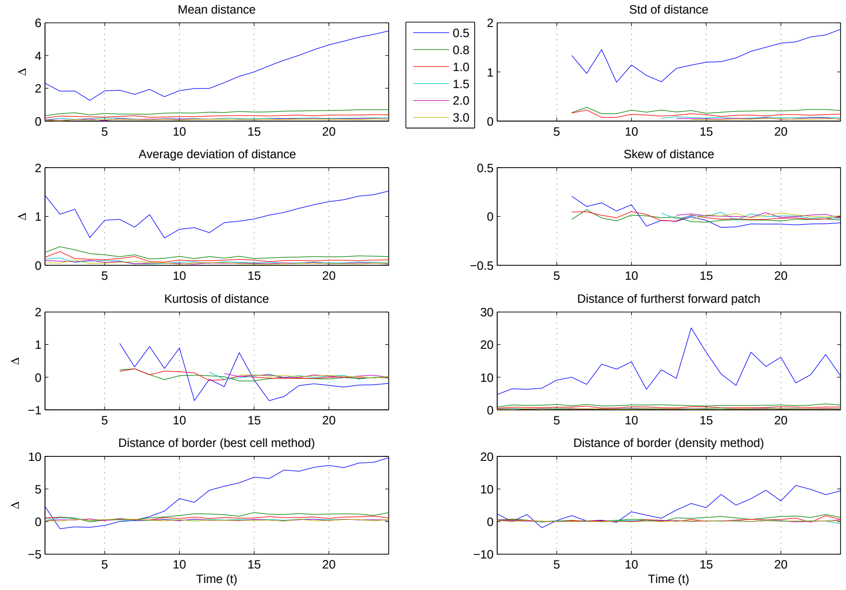
Effect of the shape parameter

Altering the shape parameter results in an almost constant mean distance velocity for all values except 0.5 which accelerates at $t > 12$. This is also an accelerating rather than constant increase in the standard deviation of distance values for $s = 0.5$ (Figure 4.13). As hypothesised (Hyp. 4.7), decreasing the shape parameter leads to greater rates of spread when measuring the distribution's progression by the mean distance method.

Furthest forward measurements show a constant increase with no obvious acceleration for $s = 0.5$, although the rate does fluctuate more as t progresses. Both iterative measures, *density* and *best-cell*, result in acceleration for $s = 0.5$ only after approximately $t = 8$. Contrary to the

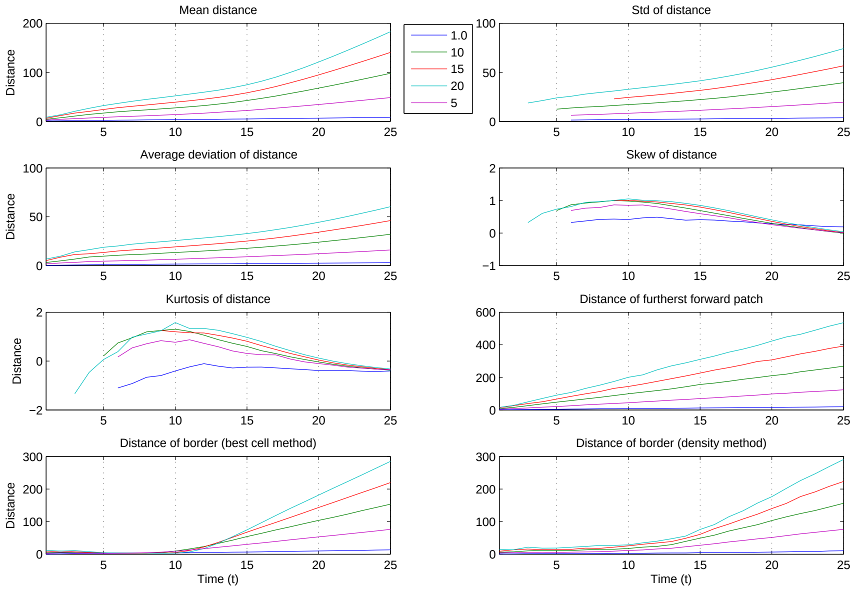


(a) Distance metrics

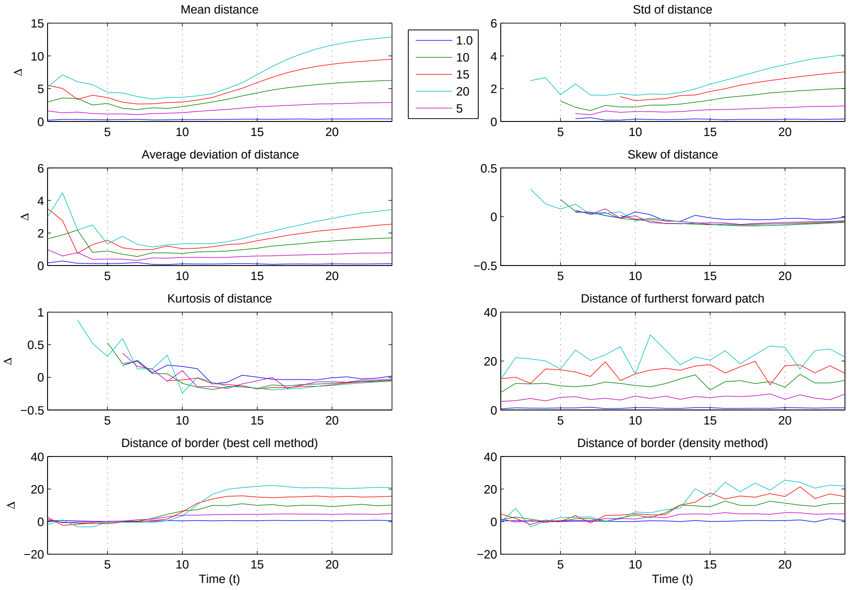


(b) Δ of distance metrics

Figure 4.13: The mean distance metrics across 50 replications for kernel dispersal with $d = 1.0$ and varying shape parameter. The difference between timesteps (equivalent to the velocity) for the metrics is also shown.



(a) Distance metrics



(b) Δ of distance metrics

Figure 4.14: The mean distance metrics across 50 replications for kernel dispersal with $s = 1.0$ and varying distance parameter. The difference between timesteps (equivalent to the velocity) for the metrics is also shown.

other measures, both $s = 1.0$ and $s = 0.8$ seem to accelerate briefly at $t = 9$ for *best-cell* and $t = 12$ for *density*.

In conclusion, the method for measuring rate of spread does not alter the ordering of rates for different shape parameter values and this ordering of rates is consistent with Hypothesis 4.7. Conversely, whether Hypothesis 4.9 holds, in other words whether these rates are constant or not, is dependent on the measurement method.

The skew of distance values and their kurtosis decrease because, as time progresses, the distribution spreads radially away from the point of origin, where there is more space free to be occupied (as area depends on radius, $A = \pi r^2$). There will initially be a skew to lower distance values as the area able to be occupied close to the initial incursion point is limited. However, there are more available cells at larger distances from this initial point and as the population distribution spreads outward this will skew the distribution to higher values (Figure 4.13).

Effect of the distance parameter

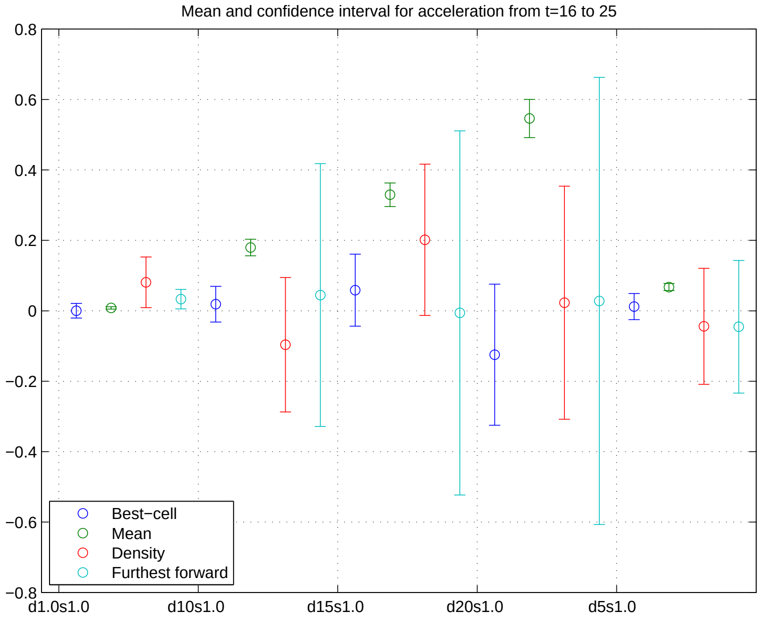
There is an acceleration in the mean distance for values of $d > 1$ at $t > 10$ although this acceleration decreases as time progresses and may eventually lead to a constant velocity. A constant velocity may eventuate because both the iterative measures also accelerate at $t = 10$ but stop accelerating and reach a constant velocity by approximately $t = 15$. Similar to modifying the shape parameter, changing the distance parameter does not result in acceleration for the furthest forward distance measurements but rather an erratic but mostly constant velocity (Figure 4.14).

The resulting order of the distance values, of smaller values having less rapid spread rates than higher values, also indicate that Hypothesis 4.8 is true.

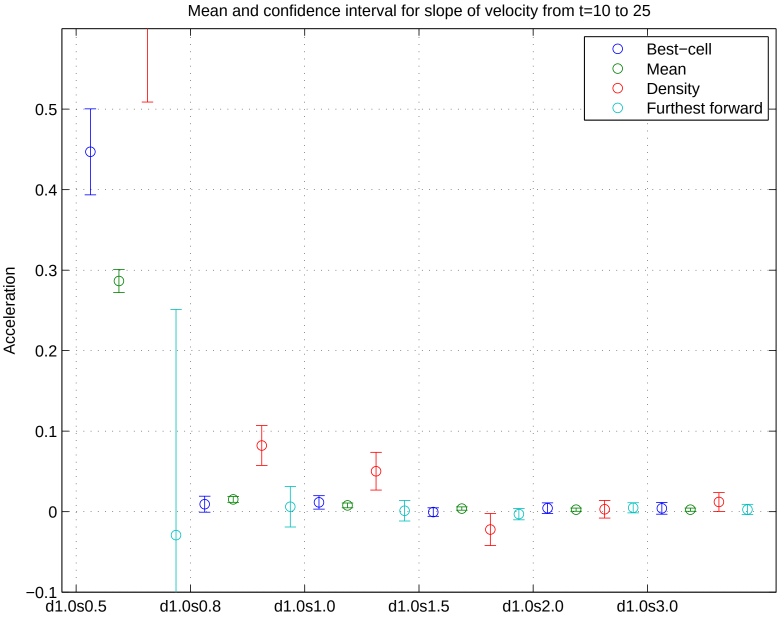
Comparison with analytical models

A frequent observation of analytical models is that they tend to predict a period of acceleration followed by a constant velocity, given appropriate assumptions (Hastings 1996*a*). To compare our simulation results to the behaviour predicted by mathematical models, the difference in distance between time-steps was plotted. This gives a straight line for a constant velocity and a linear increase for constant acceleration. A regression on the difference data for each replication gives a series of slope values. A t-test was then used to test whether the population of slope values, and therefore the population of accelerations, has a mean significantly different from zero since zero slope would represent a constant velocity (Figure 4.15).

For distance, there was no acceleration for the best-cell method, an increasing acceleration with mean distance, and for density and furthest-forward methods, acceleration only occurred



(a) Mean front acceleration for $s = 1$.



(b) Mean front acceleration for $d = 1$.

Figure 4.15: The 95% confidence interval for the mean acceleration of distribution boundary as detected by 4 different methods: mean distance, furthest forward, best-cell, and density. If zero is included in the interval then the distribution is not significantly different to zero and thus the distribution boundary is unlikely to be accelerating.

with $s = 1.0$, $d = 1.0$.

For shape, all measures except for furthest-forward accelerated with $s = 0.5$. The mean distance measure accelerated regardless of s and the density-method accelerated for s values 0.5, 0.8 and 1.0.

Effect of a limited environment on front advancement

Confining a distribution to dispersing via the generic kernel in a limited environment truncates the kernel's probability curve to the extent of the environment. This suggests that a limited environment will decrease the rate of advancement for a distribution front.

Hypothesis 4.10 *Area constraints will limit the rate at which the distribution spreads.*

Using the simulation data from the area analysis in a limited environment showed that a sufficiently diffused dispersal kernel ($d = 10, s = 0.5$) in the 200 by 200 region begins mean distance deceleration immediately whereas in an "unlimited" region it continues accelerating (Figure 4.16). For less diffuse kernels ($d = 10, s = 1.0; d = 1, s = 0.5$) the distance accelerates until the distribution begins to be affected by the region boundary and then deccelerates until no further advancement occurs. This change happens when the entire limited region is occupied by the distribution, and the deceleration occurs because of "filling in", this effect is also seen in the increase of area occupied (section 4.2.1, above). For those kernels that do not result in the distribution being close enough to the boundary in 50 simulated time-steps to be significantly truncated ($s = 1, d = 1$), the behaviour is the same as in an unlimited environment, where there is a constant velocity for mean distance measure. Constraining dispersal events to a fixed area inhibits fat-tailed (or leptokurtic kernels) more than those that are platykurtic and not exponentially bounded. Additional effects of a limited region include a limited standard deviation of the distance values, and skew and kurtosis decreasing rapidly.

Furthest forward measurements demonstrate similar results to mean distance, namely the deceleration observed as the distribution reaches near to the boundary for the dispersal kernel to be significantly truncated. Iterative measurements of front distance (*best-cell* and *density* methods) all have a delay before acceleration, followed by constant velocity and a decceleration once the boundary is reached. Interestingly, the *best-cell* measurement of the population front of the most diffuse dispersal kernel simulated ($d = 10$,$s = 0.5$), lagged behind a less diffuse one ($d = 1.0, s = 0.5$) from $t =$ 10-15. This is because the *best-cell* algorithm needs a sufficiently dense number of patches near the origin before the front can expand and the more diffuse kernel spreads out too quickly at the start for the *best-cell* method to define a population front. As the patches near the origin fill in however, the border quickly accelerates.

Effect of resolution on front advancement.

It would be ideal that the rate of front advancement is be independent of resolution. However, given the results for area, this may be unrealistic but provides a reasonable starting hypothesis:

Hypothesis 4.11 *The rates for the spread of a distribution will be independent of the resolution of the simulation.*

There is less sampling of the dispersal kernel at the beginning of simulation for coarse resolutions. This is because there are fewer patches in the original distribution. For a resolution of 1 there are 40 × 40, or 1600, unit^2 patches while for a resolution of 20 there are 2 × 2, or 4, 400 unit^2 patches. Since the Poisson distribution that generates dispersal events uses a mean that is independent of the patch size area, all resolutions have the same Poisson mean. This independent mean makes it less likely that longer distance events occur in coarse resolutions than for an equal area distribution at a fine resolution with smaller patches. This is reflected by significant measures of furthest forward distances at early t (Figure 4.17).

Resolution also affects the iterative measurements of distance indirectly. Because these measurements are inherently based on the area of patches, a fine resolution gives a gradual increase in the results whereas coarse resolution leads to a stepping effect with step sizes as a function of resolution. The density method in particular cannot successfully detect the population front for the coarsest resolution simulated ($r = 20$).

4.3 Comparison of MDiG dynamics with other models

Certain combinations of modules within **MDiG** are analogous to the behaviour of existing model paradigms. For example the survival module is essentially a Markov chain whose state is the number of occupied patches and whose transition probabilities are made up from the survival probability for each remaining patch. The expected number of surviving patches for a homogeneous survival rate is:

$$E[a_t] = a_{t-1}s \tag{4.3}$$

$$E[a_t] = a_0 s^t \tag{4.4}$$

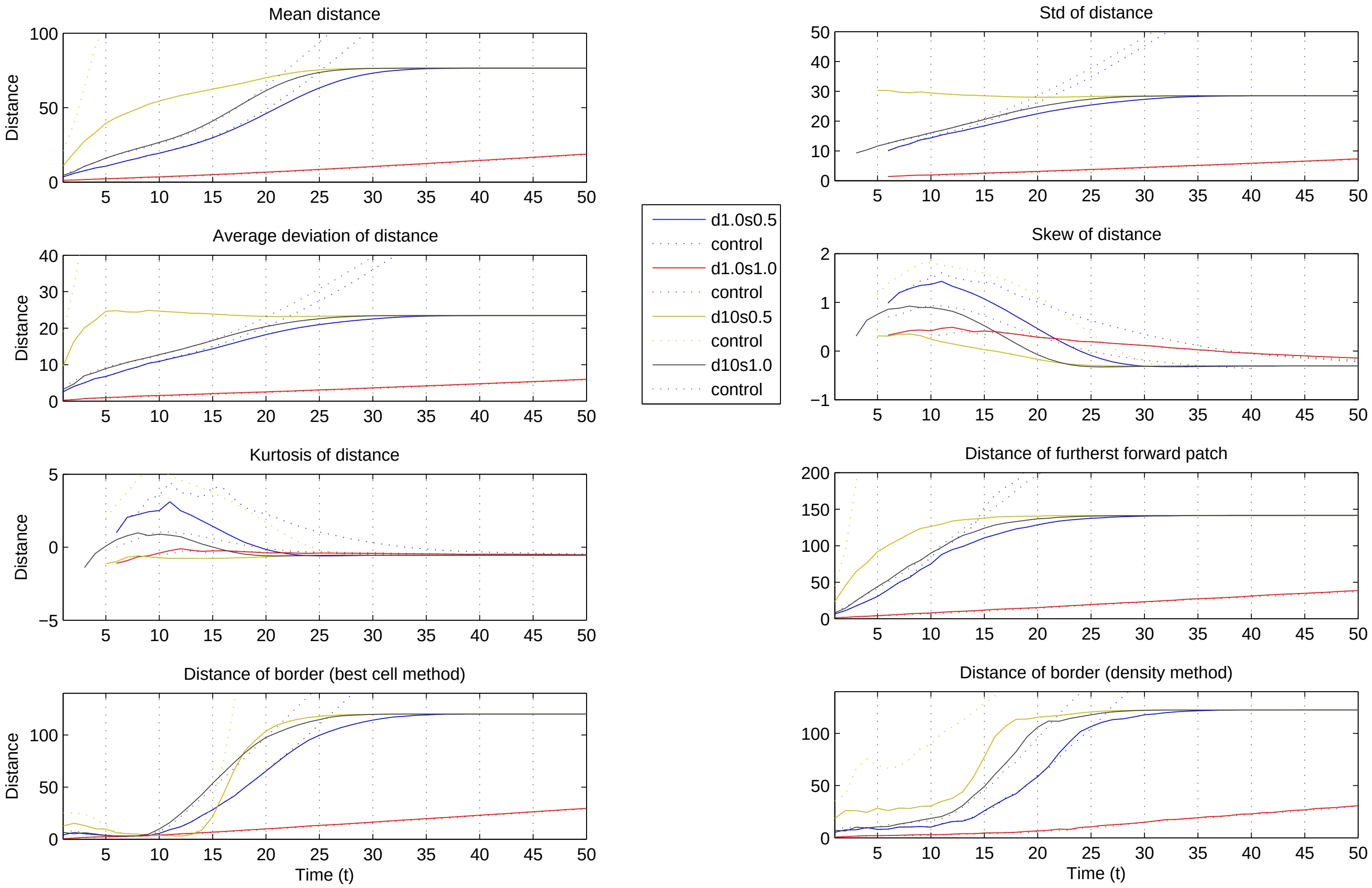


Figure 4.16: The mean distance metrics across 50 replications for kernel dispersal in a limited area of 200 by 200 patches. The control lines are the results for “unconstrained” simulations (limited to 2400 by 2400).

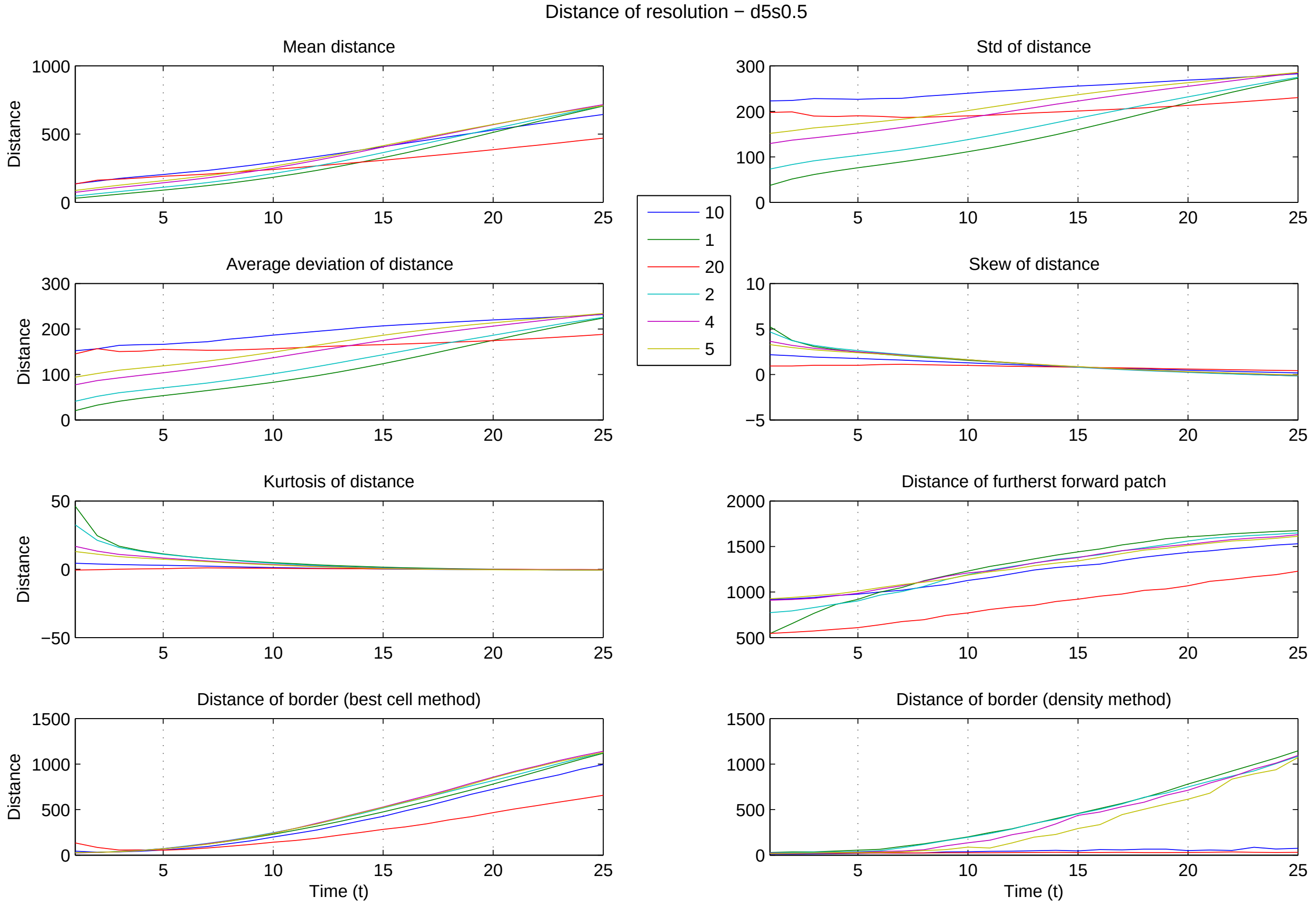


Figure 4.17: The mean distance metrics across 50 replications for kernel dispersal (d=5.0,s=0.5) in a large area of 2400 by 2400 units and simulated at different resolutions

where a_t is the number of patches occupied at time t, and s is the homogeneous surivival rate. In a environment with varied survival rates per patch, s can be replaced with the average survival rate.

Theory based on “open” vs “closed” populations tell us that the eventual result of the survival module used on its own is complete extinction (except in cases where a population exists in patch with 100% survival). The survival module on its own treats each patch as a closed population where introductions are isolated starting events, and these populations theoretically eventually go extinct with a probability of 1. Open populations are where introductions are recurrent, potentially inducing rescue effects (Brown & Kodric-Brown 1977) that decrease the extinction probabilities of well connected patches (Moilanen 2004). Open populations converge to a stationary probability distribution of population sizes, since arriving immigrants are able to restore locally extirpated populations (Drake & Lodge 2006). The addition of the local or kernel module with the survival module, turns the patches into a series of open populations that are all connected to varying degrees depending on the neighbourhood definition (in the case of the local module) and the dispersal kernel (in the case of the kernel module).

The combination of modules can also lead to similarities with other types of model. Combining the local dispersal module and the survival module, gives rise to the equivalent of a site percolation model (as described in section 2.10) when the patches have a survival probabilty of 0 or 1. Existing papers describe how the ratio of suitable to unsuitable patches, and the amount of clustering they exhibit, can influence the time until homeostasis: where the population occupies all sites that can possibly be reached. When this ratio reaches the percolation threshold, a disruption of landscape connectivty occurs, limiting the area the population can occupy at homeostasis (Gardner et al. 1987, With 1997).

4.4 Summary

This chapter examined how the parameters of the local and kernel modules affect the spatial spread of a population. Additionally the dynamics of the kernel module were compared to those of analytical spread models.

The results showed that the local module’s increase in area is dependent not only on the number of patches in the defined neighbourhood but also on the pattern of the neighbourhood.

For the kernel module the results show a clear relationship between the increase in area with changes to the parameters of the generic kernel, which can represent several other probability distributions based on its shape parameter. A more diffuse dispersal kernel and kernels scaled to larger areas (an increase in the distance parameter, d) showed a greater increase in the area of the population distribution, although not all differences were statistically significant. The area measurements were also normally distributed, although variance was not homogeneous between simulations so non-parametric tests were needed to test significance.

Spread rates for the expansion of the population front expands were consistent with what analytical models predict. Dispersal kernels that were not exponentially bounded, so called "fat-tailed" distributions, tended to accelerate. Although whether this held for different cases was dependent on the algorithm used to determine the location of population front.

The impact of resolution on area and spread rate was investigated, and a equivalent ordering of area occupied in relation to resolution was not found. Instead, particular parameter combinations appear more optimal for certain resolutions. This was suspected to be because of the interaction between a continuous dispersal kernel and a discrete patch environment, since the size of these patches are determined by the resolution of the simulation.

When simulations were constrained to a limited region the increase in area and spread rate were restricted compared to that of unbounded spread. As well, the constraining effect impacted the area of the population distribution before saturation of the environment became obvious, particularly for fat-tailed distributions.

There are many further investigations that could be carried out on these modules. Including interactions with heterogeneous environments that have certain characteristics, such as a certain fractal dimension or ratio of suitable to unsuitable habitat. Interactions between modules are also of interest, and provide a wealth of possible research avenues.

Chapter 5

Modelling the Spread of Argentine Ant in NZ

The primary objective of this chapter was to test the modelling framework developed in Chapter 3 using a species that has data available on it's spread such that its invasion history can be reconstructed. Separate data available from outside the study region that can be used to parameterise the model to avoid confounding the model results is also a requirement. The Argentine ant, *Linepithema humile* (Mayr) fulfills these requirements.

The second objective is to compare this model to a simple model of radially uniform diffusion to determine if the extra effort building a more complex model is worthwhile.

5.1 Introduction

The Argentine ant is a worldwide pest that is cited as one of the six worse invasive ants (Holway, Lach, Suarez & Tsutsui 2002) and is listed in "100 of the world's worst invasive alien species" (Lowe et al. 2005). *Linepithema humile* was first recorded in NZ in January, 1990 (Green 1990) and after its discovery at the Mt. Smart, Auckland, site of the 1990 Commonwealth games there was no attempt to control the species because it was considered already well established.

Argentine ant provides a useful case study of an invasive insect introduction in NZ where no control attempts were made until substantially after establishment. Thus the distribution data from monitoring programs show Argentine ant dispersal behaviour mostly without attempts to slow the Argentine ant's spread or eradicate the species from New Zealand.

Figure 5.1: *L. humile* workers are small (2-3mm long) and are coloured a uniform light to dark honey-brown. Unlike some ant species there is no strong formic acid smell when they are squashed. Photo courtesy of Landcare Research, NZ.

5.2 Biology

Argentine ant (*L. humile*, shown in Fig. 5.1) is known as a 'tramp' ant species (Passera 1994), sharing the label with the 'big-headed ant' (*Pheidole megacephala*, Fab.), and the red imported fire ant (*Solenoposis invicta*, Buren). Tramp species tend to have the following features in common that have made them highly successful invaders (Harris 2002):

- **Strong tendencies to move and associate with humans** - in the event of a disturbance caused by the weather, food supply, or human activity, *L. humile* workers are adept at picking up larvae and eggs to relocate their nest. The species is able to survive high levels of disturbance using this behaviour and their close association with human activity means that human transport facilitates their spread by jump-dispersal (Suarez et al. 2001).
- **Unicoloniality** - Adventive populations of *L. humile* effectively act as one large colony. This behaviour, hypothesised to be the result of reduced genetic diversity in introduced populations, means that individuals do not show aggression to individuals from different nests (Chen & Nonacs 2000, Liang & Silverman 2000), even though *L. humile* exhibits inter-colony aggression in their native range (Tsutsui et al. 2000). Recently however, diet-derived nest-mate recognition has been implicated as controlling aggressive interactions (Buczkowski & Silverman 2006). Desite this dietry mechanism, the general lack of inter-colony competition facilitates higher densities of the species and greater foraging activity (Holway 1998*b*).
- **Interspecific aggression** - Workers of *L. humile* display strong aggressive behaviour to other ant species and displace them through sheer weight of numbers, even though individuals

of *L. humile* often lose in aggressive interactions (Holway 1999).

- **Polygyny** - *L. humile* has several queens in a colony (0.1 to 1.6 queens per 100 workers, Keller et al. 1989) resulting in a high reproductive rate, although queens generally live less than 1 year (Keller & Passera 1990).
- **Mating and budding** - Mating occurs in the nest and colonies disperse to new sites by budding, where a group of workers and queens seperate from the main colony and walk to the new site. Laboratory experiments show colonies successfully establishing with a single queen supported by as few as 10 workers (Hee et al. 2000).

5.2.1 Diet

L. humile's diet overlaps greatly with other ant species. Diet is predominantly liquid, with workers feeding primarily on sugar, while queens and larvae feed mainly on protein (Vega & Rust 2001). *Linepithema humile*'s diet includes nectar, insects, seeds, carrion, and honeydew secreted by homopterans (Suarez et al. 1998). *Linepithema humile* and other invasive ants often tend and protect Hemiptera that produce ‘honey dew’ or sweet liquid exudate (Lester et al. 2003). Unusually for ants, the queens of *L. humile* participate in foraging and grooming activities (Vega & Rust 2001).

5.2.2 Climatic requirements

L. humile is most active between 10°C and 30°C. Foraging however ceases when surface temperatures reach 32°C or drop below 15°C (Hedges 1998) - although 10°C has also been reported as the lower range for foraging (Markin 1970*a*).

Oviposition slows during winter and does not occur below a daily mean temperture of 18.3°C. Eggs, larvae, and pupae develop slower and population numbers decline in colder months (Vega & Rust 2001). Through their close association with humans, *L. humile* may persist locally in areas with unfavourably cold climates (e.g. Minnesota and Illinois in the USA, Suarez et al. 2001). This local persistance is due to warmer microclimates near structures and human activities providing regions in which *L. humile* can survive (Suarez et al. 2001).

Environments with high rainfall reduce foraging time and thus possibly reduce establishment success (Vega & Rust 2001), while conversely arid climates can restrict distribution due to a lack of water (Ward 1987, Van Schagen et al. 1993, Kennedy 1998)

5.2.3 Habitat

Globally, most reports of *L. humile* are from urban areas (Suarez et al. 2001). However, this species is also a pest of horticultural land in Australia (Davis & Van Schagen 1993), and low-stature scrub vegetation is susceptable to invasion (Reimer 1993, Way et al. 1997, Human et al. 1998). Forests are not known to be invaded (Reimer 1993, Cole et al. 1992), or only have their bordering margins penetrated by *L. humile* (Suarez et al. 1998, Ward & Harris 2005).

Edge effects, where only the margins of vegetation are colonised, were shown in a study by Suarez et al. (1998). Argentine ant populations in native vegetation fragments near San Diego were found to decrease at distances greater than 100m from the edge. Native ant populations that were possibly displaced were found only at distances greater than 200m into the fragments, thus indicating that *L. humile* likely hadn't established beyond 200m into the vegetation fragments (Suarez et al. 1998). It is unclear if these observations mean that the process of invasion needs more time to progress further, or if competition with native species increases further into the vegetation fragments, or if *L. humile* cannot penetrate further due to lack of water (Suarez et al. 1998). It has been suggested that water run-off from urban developments may provide favourable conditions for *L. humile* and a lack of moisture availability appears to limit *L. humile*'s distribution (Tremper 1976, Ward 1987). Moisture availability will be critical in hot-dry climates where physical distance to water is clearly an important factor (Holway, Suarez & Case 2002).

In NZ, *L. humile* has been found to move only up to 20m into forest habitats. In habitats with a more open canopy, such as mangrove and scrub, the distance penetrated was at least 20m but up to 60m (Ward & Harris 2005).

In contrast, the influence of soil substrate is thought to be the limiting factor affecting the distribution of *L. humile* in western and southern Portugal where it has had a relatively unchanged distribution for 40 years. This species is common in sand and clay loam soils, but largely absent in sandy loam soils despite otherwise favourable conditions (Way et al. 1997).

5.3 Impacts

More so than other ant species, *L. humile* tends to be extremely abundant once established. Interconnected nests and variable methods of trapping, however, make it difficult to accurately assess abundance. *Linepithema humile* in high abundance can result in a decline in local biodiversity for both invertebrates (Cole et al. 1992, Human & Gordon 1997, Bolger et al. 2000) and small vertebrates such as the Californian gnatcather (Sockman 1997) and the horned lizard (Suarez et al. 2000).

Through superior competitive ability, perhaps as a result of greater numbers or the formation of cooperative super colonies, *L. humile* often displaces other ant species (Erickson 1971, Bond & Slingsby 1984, Ward 1987, Cammell et al. 1996, Human & Gordon 1996, Suarez et al. 1998, Holway 1998*a*), but is known to be displaced itself by the red fire ant (*S. invicta*) (Porter et al. 1988). In Bermuda, the big-headed ant, *P. megacephala*, appears to survive in the presence of *L. humile* but does not coexist in the same locations; rather, their territory is in flux and the two species manage to survive in a dynamic equilibrium (Haskins & Haskins 1988). Although *L. humile* reaches high densities through cooperative colony behaviour, areas with an already rich variety of ant fauna present do not show an increase in total ant biomass after *L. humile* establishment. This lack of increase is because of the displacement of existing ant species (e.g. Holway 1998*a*).

The displacement of native ant species can cause existing mutualisms to be disrupted (Bond & Slingsby 1984). For example, a fynbos ecosystem from the southwestern Cape, South Africa, is swept by wildfires every 10-15 years and relies on native ants to bury seeds for regeneration. Argentine ants wiped out two fynbos ant species, *Anoplolepis custodiens* and *Pheidole capensis*, that bury larger seeds. Thus biodiversity is lost after a wildfire because larger seeds are not buried and so the corresponding species cannot regenerate (De Kock 1990).

Ants are often part of soil processes, and these processes could be disrupted by ant species being displaced by *L. humile* (Folgarait 1998).

The arrival of *L. humile* in Japan has led to significantly less species diversity for communities in urban environments where *L. humile* is present compared to where it is absent (Touyama et al. 2003). In parks infested with *L. humile*, species such as *Pheidole noda*, *Pheidole indica*, and *Lasius japonicus*, were absent. Touyama et al. (2003) suggests that *L. humile* is superior to these species in traits such as mobility, recruitment ability, aggressiveness and omnivory. However, two other species, *Paratrechina sakurae* and *Camponotus vitiosus* were less affected by *L. humile*'s presence. Possibly the small body size of *P. sakurae* may reduce its competition for resources such as nesting sites and food with the larger *L. humile* (Touyama et al. 2003) and additionally *L. humile* rarely attacks *P. sakurae* (Miyake et al. 2002). The arboreal nesting habits of *C. vitiosus* may allow it to avoid competition with *L. humile* (Touyama et al. 2003).

Holway (1999) investigated the mechanisms behind *L. humile*'s apparent superior competitive ability. It was found that *L. humile* located baits and recruited other ants to them as quickly or more quickly than native ants, irrespective of whether they occupied the same area. While individual Argentine ant workers were unable to consistently overcome native ant workers, Argentine ant colonies succeeded in displacing native ant colonies from baits, which suggests that numerical advantage is vital to *L. humile*'s superior competitive ability.

Harris (2002) compiled a table on the documented impacts of *L. humile* on community structure and the processes of native ecosystems around the world (Table 5.1).

5.3.1 NZ Impact

In NZ *L. humile* has the potential to become a major household and garden pest reaching much higher colony densities than in its native range (Suarez et al. 1999). The small size of individual ants allow *L. humile* to penetrate food containers (Davis & Van Schagen 1993), furthermore individuals have the potential to spread disease by carrying pathogens on their bodies (e.g. *Staphylococcus*, *Candida*, and *Enterococcus* Ipinza-Regla et al. 1985, Fowler et al. 1993).

L. humile are very aggressive towards other insect species so may threaten endangered native invertebrates. Indirectly, the removal of native ant species and insects could threaten species further up the food chain - especially those that rely on native ants or other invertebrates for food, pollination, or seed dispersal. The extent of such ant interactions are currently unknown in NZ.

Through weight of numbers *L. humile* also directly competes for nectar and honeydew with other insects and birds (Suarez et al. 2000). Large numbers can also kill infant birds in the nest (Newell & Barber 1913), and such attacks have already been observed in New Zealand (V. Van Dyk pers. comm. cited in Harris 2002).

The current ant fauna of NZ consists of 40 species, approximately 11 of which are considered endemic (Valentine & Walker 1991, Harris & Berry 2002) and their local populations could be reduced in competition with *L. humile*. However, researchers consider no native ant species at risk of extinction from *L. humile*, because of their wide geographical distribution and their occurrence in indigenous forest (Harris 2002) which is generally not invaded by *L. humile* (Reimer 1993, Cole et al. 1992).

Lester et al. (2003) have reported that *L. humile* attack honeybee hives in NZ, and competes with honeybees for nectar from the flowers of pohutukawa (*Metrosideros excelsa* A. Richard), New Zealand bottlebrush (*Knightia excelsa* R. Brown), and agapanthus, a common exotic garden plant (*Agapanthus* spp.). Argentine ants have also been observed to kill honeybees, steal honey and destroy hives (Pasfield 1968, Vega & Rust 2001). *Linepithema humile* have been reported by Lach (2005) to harvest nectar from other species such as the Hawaiian ōhiâ tree. It seems *L. humile* has potential to compete strongly against legitimate floral visitors for nectar on other New Zealand plants.

When *L. humile* "farms" and tends Hemiptera such as aphids and scale insects for honeydew,

Table 5.1: Summary of the main documented impacts globally of *L. humile* on native ecosystems. From Harris (2002)

Community Structure	Ant diversity	Interference competition and food monopolisation	Human & Gordon (1996), Holway (1999)
	Abundance and diversity of other invertebrate species	Interference and resource competition; predation on eggs, larvae and adults	Cole et al. (1992), Way et al. (1992), Human & Gordon (1997)
	Abundance of vertebrate species	Interference and resource competition	Suarez et al. (2000)
Community processes	Pollination	Competition for nectar with effective pollinators	Buys (1987), Visser et al. (1996)
	Seed dispersal / regeneration	Displacement of specialist ants that have co-evolved to assist seed dispersal and seedling germination	Bond & Slingsby (1984), Giliomee (1986)
	Decomposition / nutrient cycling	Changing the guild structure of the invertebrate community	Ward (1987), De Kock (1990), Folgarait (1998)

L. humile actively spreads and protects them. This results in *L. humile* being one the worst pests of citrus in Australia (Forest & Bird n.d.) as well as a serious pest for viticulture (Addison & Samways 2000). In NZ this behaviour could interfere with biocontrol of the invasive weed boneseed (*Chrysanthemoides monilifera monilifera*). Argentine ants farm scale insects on boneseed plants, and vigorously protect them. This is expected to result in high mortality in the larvae of the *Tortrix sp.* moth that is planned to be used as a biocontrol agent against boneseed. *Linepithema humile* can also attack herbivores browsing on boneseed leaves (pers. comm. M. Stanley).

Fifteen of NZ's eighteen common horticulturual crops have been found to have *L. humile* tending and feeding on the honeydew exudate of the crop's Hemipteran pests (Lester et al. 2003). Lester et al. (2003) also observed infestations of *L. humile* in New Zealand on walnut (*Juglans* sp.), gum trees (*Eucalyptus spp.*), pine trees (*Pinus* sp.), vegetables such as carrots (*Daucus* carota L.), and tomatoes (*Lycopersicon esculentum* Miller).

Organic produce has been increasing in demand and is growing in economic importance to NZ (Ritchie et al. 2000). The presence of *L. humile* could disrupt organic farming as organic growers rely on natural enemies of Hemiptera to control these pests. With *L. humile* protecting Hemiptera, their natural enemies would no longer be as effective a control (Gonzalez-Hernandez et al. 1999). In particular, Samways et al. (1982) and Reeve & Murdoch (1986) have shown that *L. humile* disrupts the biological control of scale insects and mealybugs on a variety of citrus crops such as lemon, oranges, and grapefruit.

L. humile is not currently found in some Asian countries that NZ actively trades with. If *L. humile* establishes in horticultural areas and ports it could impact our trade with these countries as more intensive monitoring and quarantine procedures are imposed (Forest & Bird n.d.).

5.4 Distribution

5.4.1 World distribution

There are 28 species and subspecies of the genus Linepithema, but only *L. humile* occurs outside of Central and South America. *Linepithema humile* occurs mostly in a latitudinal range of 30-36 degrees in both hemispheres (Majer 1993) and is now common in Mediteranean climates (Williams 1994). Local conditions of sites outside of this range however, such as in urban regions that provide shelter from harsh conditions, can support and allow *L. humile* to thrive.

L. humile is native to Argentina, Uruguay, Paraguay (Suarez et al. 2001) and was also thought to be native to Brazil (Orr & Seike 1998) but genetic sequence data from the mitochondrial cytochrome *b* and nuclear microsatellite allele frequencies led Tsutsui et al. (2001) to believe the Brazilian population to be a different species (Fig. 5.2). Most samples from invasive populations around the world show morphological and genetic similarity to populations from the southern Rio Parana (Tsutsui et al. 2001) and Uruguay (unpublished data cited by Suarez et al. 2001), suggesting that this may have been the original source of the world's invasive population.

Suarez et al. (2001) presents a reconstruction of invasion histories for *L. humile* that shows at least 28 separate introductions to all the continents save Antarctica (Holway, Lach, Suarez & Tsutsui 2002), and to many islands throughout the world (Fig 5.2). The earliest recorded introduction was to Madeira Island in 1882 (Suarez et al. 2001), and many of the currently existing populations worldwide were established by the 1950s. The extent of the invasion by *L. humile* indicates the importance of human mediated dispersal since *L. humile*'s natural dispersal ability is limited (Section 5.5).

Areas that are believed to currently be free of Argentine ant but have been identified as high risk of future establishment include: Madagascar, Taiwan, high-elevation Ethiopia and Yemen, Saint Helena, the Chatham Islands of NZ, Norfolk and Lord Howe Islands of Australia and a number of other oceanic islands. These regions all have highly suitable temperature as predicted by multimodel inference (Hartley et al. 2006). Ecological niche models by Roura-Pascual et al. (2004) using the Genetic Algorithm for Rule-set Prediction (GARP Stockwell 1999) indicate coastal Africa and southeast Asia are vulnerable to invasion. The same models based on future (2050s) climate predictions indicate that the species' range will retract in tropical regions, but expand at higher latitude areas (Roura-Pascual et al. 2004).

5.4.2 *L. humile* distribution in NZ

L. humile's presence in NZ was initially discovered in 1990, when it was found at Mt. Smart, Auckland, the site of the 1990 Commonwealth games (Green 1990). At the time of discovery it was already well established and there was no attempt to eradicate it.

Unlike another tramp ant species, *P. megacephala* (Fab.), which has not spread beyond Auckland and the Kermadec islands since it's discovery in 1942 (Berry et al. 1997), *L. humile* has found its way to many NZ regions in approximately a decade and a half (Fig. 5.3). The NZ Ministry of Agriculture and Fisheries (MAF) undertook a national survey to ascertain the spread of *L. humile* which was completed in September, 2001. The survey identified ten North Island centres where the species is present (Hamilton, Kaitaia, Auckland, Dargaville, Hastings, Morrinsville, Mt Manganui, Warkwarth, Whangaparoa, and Whangarei), whereas in the South Island the ant was

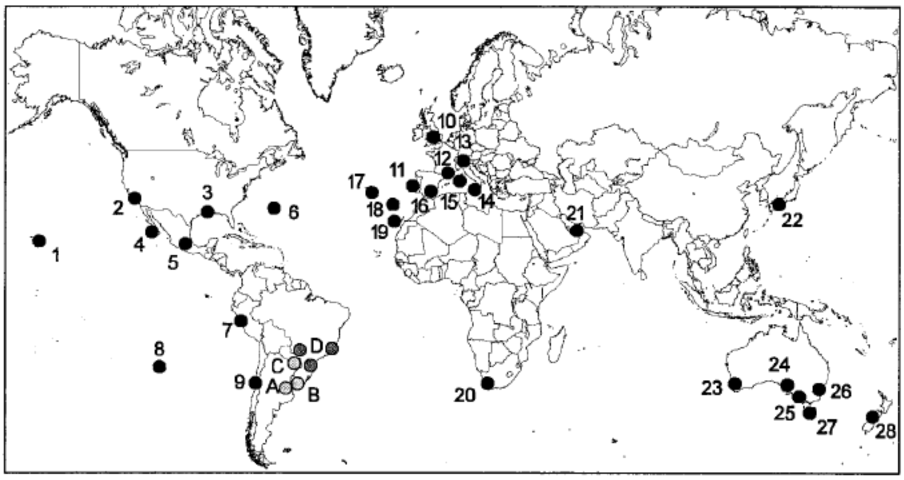


Figure 5.2: Known worldwide distribution of Argentine ant. Numbered black circles indicate introduced populations while shaded circles represent presumedly native populations. Used with permission from Suarez et al. (2001).

found to be mainly absent apart from small populations in Nelson and Christchurch (MAF 2001*a*).

Populations have also been discovered on Tiri Tiri Matangi[1] (Forest & Bird n.d.), a predator-free conservation island in the Hauraki Gulf, and on Great Barrier Island. Both islands have undergone poison baiting treatments to try and eradicate *L. humile* (pers. comm. J. Boow, Biosecurity Team Leader, Auckland Regional Council, and pers. comm. C. Green, Technical Support Officer, Department of Conservation).

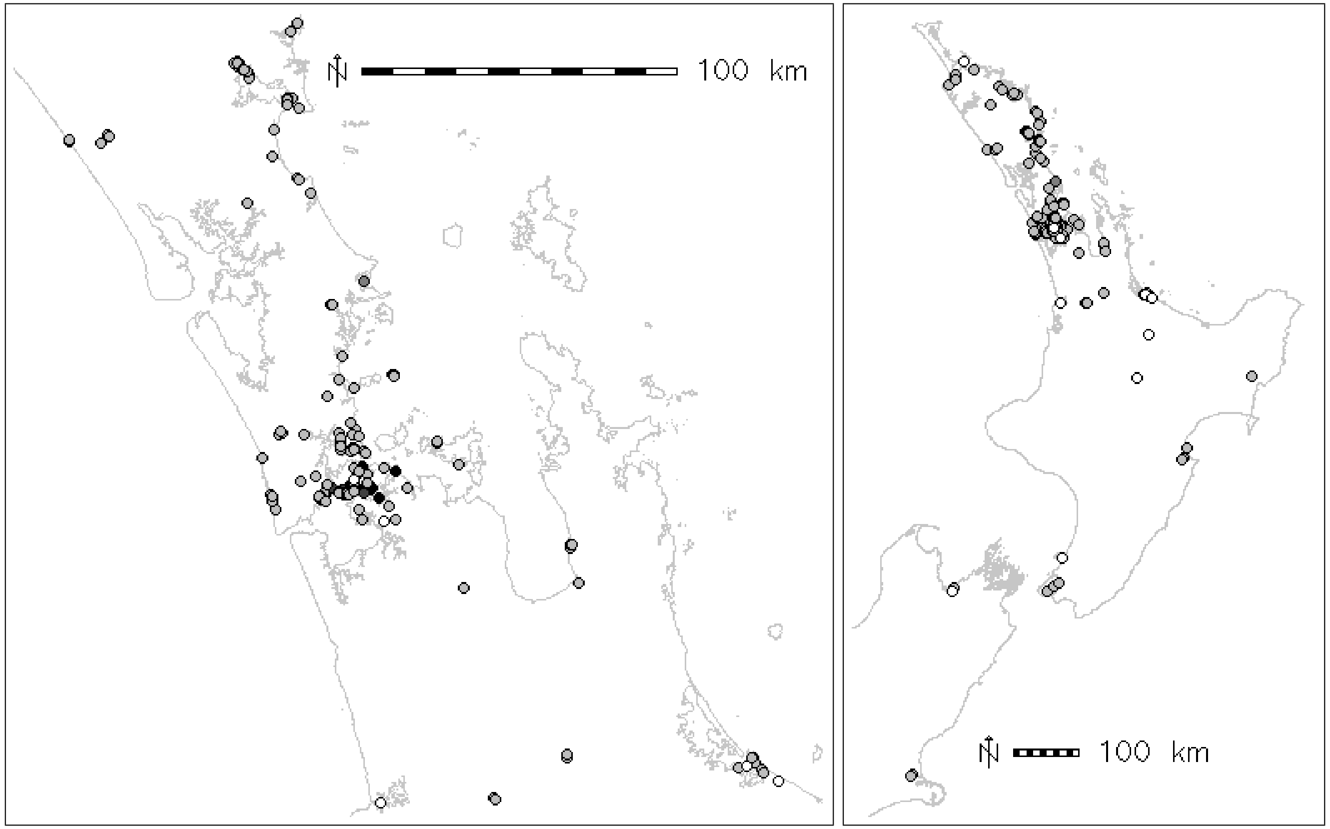
The eradication on Great Barrier Island began during the summer of 2006 to 2007, with two treatments several weeks apart. The programme is expected to be completed at the end of the 2007-2008 summer season. The eradication was a joint Department of Conservation and Auckland Regional Council programme managed by a local contractor but carried out mostly by volunteers (pers. comm. J. Boow).

5.4.3 Potential Distribution in NZ

The potential distribution of a species comprises the areas that the species could survive and establish a reproducing population. Several methods have already been used to judge the potential distribution of *L. humile* in NZ. Harris (2002) used both land cover data and mean annual temperatures to create a risk assessment for the entirety of New Zealand, categorising regions into 5 risk categories: High, medium, low, unsuitable habitat, or too cold.

---

[1] http://www.tiritirimatangi.org.nz/



(a) Auckland region

(b) Entire distribution

Figure 5.3: The distribution of Argentine ant (*L. humile*) in NZ as of 2005. a) Focussed on the Auckland region. b) Showing the entire distribution in New Zealand. Occurrences are shaded based on the year of discovery: 1990-1993 black, 1994-1996 dark gray, 1997-1999 medium gray, 2000-2002 light gray, 2002-2006 white.

Hartley & Lester (2003) used average degree days above a threshold of 15.9°C. as an indicator of survival and establishment with a threshold range of 400-500 cumulative degree days as the most suitable for sustainable *L. humile* populations.

Related work (Ward & Harris 2005), although not specifically focused on determining the potential distribution, found that *L. humile* progression into NZ native forests is limited and could be incorporated as an additional factor for generating *L. humile*'s potential distribution when land cover data is available.

Errors predicting potential distribution, particularly when using a climate matching method, may be partly explained by failing to account for biotic interactions such as competition and availability of hosts or food (Davis et al. 1998, Case et al. 2005). However, we assume distribution limitations due to biotic interactions to be minimal because of *L. humile*'s "tramp" status (See previous section 5.2). *Linepithema humile* is often competitively dominant against other ants and insects and *L. humile* can also form mutualistic associations with wide variety of novel homopteran hosts, despite a lack of co-evolutionary history (Lester et al. 2003). While we make this assumption, it is worth noting that the high diversity of the native *Iridiomyrmex* ants in Australia, combined with hot-dry conditions, has been suggested to be a contributing factor restricting the spread of Argentine ants in Australia (Walters & Mackay 2003, Walters & Mackay 2004, Walters & MacKay 2005).

5.5 A model of Argentine ant spread

5.5.1 *L. humile*'s dispersal behaviour

The dispersal of *L. humile* can be characterised as **stratified diffusion** (section 2.3.2), where an observed patchy distribution results from the interaction of more than one mode of dispersal (Shigesada et al. 1995). The two most influential dispersal modes for *L. humile* appear to be diffusion by budding and jump-dispersal facilitated by transport via humans (Barber 1916, Holway 1995, Suarez et al. 2001).

The prevailing mode of colony reproduction for many ant species involves the queens performing a mating flight before founding a colony independent of their natal nest (Holldobler & Wilson 1990). In contrast, *L. humile* queens are not known to undergo mating flights in their introduced range (Markin 1970*b*, Fluker & Beardsley 1970, Holway 1998*b*). Instead, new nests are formed by budding which involves inseminated queens leaving the nest by foot accompanied by a contingent of workers. As few as 10 workers accompanied by an inseminated queen have been observed to establish a new colony in laboratory studies (Hee et al. 2000).

The rate at which budding occurs results in a spread rate that can be up to 800 m yr^{-1} in highly favourable recently invaded habitat (Holway 1998*b*, Way et al. 1997). However, an average of 150 m yr^{-1} has been calculated for regions where habitat and climate are favourable (Suarez et al. 2001)

L. humile's opportunistic nesting habits and general dietary requirements result in close association with humans (Newell & Barber 1913, Suarez et al. 2001). This association leads to *L. humile* being carried on cargo, cars, pot plants, and other goods transported by humans. The extreme polygyny of *L. humile* increases the likelihood that any particular colony fragment will be viable, and the queens' participation in foraging activities (Vega & Rust 2001) may make the unintentional transport of queens more probable. It is these factors that promote the occurrence of human-mediated jump-dispersal. As a result, the establishment patterns of *L. humile* in many areas follow human transport networks (roads and railways) (Barber 1916, Holway 1995, Human et al. 1998). The distances measured for human-mediated jump-dispersal average three orders of magnitude larger than the range of distances *L. humile* travels by budding (Suarez et al. 2001).

Biotic resistance does not appear to influence rate of spread for established populations (Holway 1998*b*), but may influence success of incipient colonies because numerical advantage appears key to their competitive ability (Tremper 1976, Holway 1999, Walters & MacKay 2005, Rosset et al. 2005).

In studies on the spread of *L. humile* throughout the USA there were clear lag times of several decades before significant dispersal occurred (Suarez et al. 2001). Such lag times are a common features of invasions but their underlying causes often remain unclear (Crooks & Soulé 1996, Hastings 1996*b*). Suarez et al. (2001) provide some explanations such as the inherent features of population growth, environmental changes that benefit invasive species such as increased urbanisation since *L. humile* closely associates with humans, and genetic changes occurring subsequent to introduction that result in fitness increases allowing *L. humile* to reproduce more rapidly. One possibility that isn't mentioned Suarez et al. (2001) however, is that an increase in source sites from which jump-dispersal occurs could be responsible. Clear lag periods are demonstrated in theoretical **MDiG** models that are a result of few source sites at the beginning of the simulations (Section 4.2).

5.5.2 Model description

The model simulates the population distribution of *L. humile* using a raster map where each cell indicates the presence or absence of the species rather than the population number. Population dynamics were not incorporated to simplify the model. As well there were insufficient data to establish robust relationships between *L. humile* dispersal parameters and abundance. It has also been suggested that the presence/absence assumption is necessary in any large-scale ecological study (Steinberg & Kareiva 1997).

The previous section discussed the two major modes of *L. humile* dispersal, budding and human-mediated jump-dispersal. We used the **local** module (Section 3.3.1) to simulate budding, as it is the most appropriate for approximating contiguous diffusion-like processes. Additionally the raster resolution was set to the average budding spread rate (150 m yr^{-1}) and a Von Neumann neighbourhood was used by the **local** module. For human-mediated dispersal, the **kernel** module (Section 3.3.2) was used with a distribution kernel that fitted the observed distribution of jump-dispersal distances in USA occurrence data (section 5.5.3).

The time step for the simulation was a year. At each step the model takes a map representing the current population distribution and generates a distribution map for the following year. The distribution map is generated by passing the previous year's distribution map through the **MDiG** modules in the order: **local**, **kernel**, **survival**.

The scenario for the model was to start the simulation from the initial points of invasion near Mt. Smart in 1990 and running it for 15 years, to validate the model with existing data.

The XML model definition file can be found in Appendix C.1.

5.5.3 Parameter estimation

To calibrate the modules, dispersal and occurrence data from the USA[2] were used. Occurrence data for NZ[3] were used to check that the distances and frequencies of long distance dispersal had occurred at a similar scale to the USA. The explicit spatial locations of New Zealand occurrence data were kept for model validation.

Long distance dispersal distribution

The frequency and distance of long distance dispersal events are difficult to quantify (Higgins & Richardson 1999) but estimates for rate and distance of long distance dispersal events are essential

---

[2] provided by Andrew V. Suarez, Departments of Entomology and Animal Biology, School of Integrative Biology, University of Illinois

[3] provided by Landcare Research, New Zealand, among others, see Acknowledgements

for accurate model construction (Shigesada & Kawasaki 1997).

Previous studies reporting the mean and standard deviation of *L. humile* jump-dispersal distances had been carried out in NZ (Ward et al. 2005) and in the USA (Suarez et al. 2001). However, to simulate the dispersal of *L. humile* more information was needed than merely measures of central tendency and variance. The probability distribution of the distances needed to be estimated so that a realistic distribution of dispersal distances could be sampled (see section 3.3.2).

To establish the distribution of the jump-dispersal distances we extracted distances from both USA and NZ data as follows:

Measurement technique. Two methods were used to estimate the distance of dispersal events. The first method measures the distance of all occurrences from the original invasion sites, the other measures the distances to the nearest pre-existing occurrence of *L. humile.* In other words, any occurrence whose detection year is less than the current one being processed has the distance to it measured, with the smallest these distance being recorded. These two techniques give an upper and lower bound to the probability distributions for distance travelled. When distances are measured from the original invasion points they tend to be over-estimated as not all sites originate from the initial invasion site (it is not however the set of maximum distances possible, as there maybe later sites further away than the origin), likewise, the distances are underestimated when we assume all dispersal events originate only from their nearest pre-existing neighbour.

USA distribution data. The USA occurrence data was recorded at a county scale and for each county the year of Argentine ant discovery was recorded (for further detail on the individual sources see Suarez et al. 2001). The exact geographic location for each occurrence, from which all measurements were made, was the average centroid of all the GIS polygons that belonged to a county, weighted by area size:

$$[x_w, y_w] = \left[\frac{\sum_{i=1}^{N}(x_i a_i)}{\sum_{i=1}^{N} a}, \frac{\sum_{i=1}^{N}(y_i a_i)}{\sum_{i=1}^{N} a} \right] \tag{5.1}$$

Here, $[x_w, y_w]$ are the coordinates of the weighted centroid; N is the number of centroids belonging to the county; x_i, y_i, and a_i are the x coordinate, y coordinate, and area respectively of the ith centroid belonging to the county. Details about how the centroids or measurement points were determined was not specified in Suarez et al. (2001).

The USA data also included occurrences in Hawaii. These were removed from the distance-from-origin measurements as they would distort the data. For the nearest neighbour method, the distance to the first occurrence in Hawaii was discarded, but the subsequent distances to later occurrences in Hawaii were kept since they could be measured relative to the first Hawaii occurrence.

Additionally, in 1907 a population of *L. humile* established on the west coast of North America in California whereas it had previously been confined to the east coast. For dates after 1907, the site in California was considered one of the original invasion points (as per Suarez et al. 2001).

For the USA data there was no obvious correlation between year and the magnitude of dispersal distances. This was unexpected since the occurrence data spans more than a century (the first record was in 1891) and various transportation technologies such as affordable automobiles and commercial airlines have come into use since then. These transport technologies seem to have had little affect on *L. humile* human-mediated dispersal distances, thus a distribution of dispersal distances based on the USA data should also be acceptable for modelling the more recent NZ invasion.

NZ distribution data. The NZ distribution data was more specific than the USA data, but covers a much shorter span of time since *L. humile* only arrived in New Zealand in 1990 (see Figure 5.3). The data consists of 254 records, with 101 of those found in 2001, and 117 found in 2000 and the two years following 2001. These numbers reflect the national survey that was completed at the end of 2001 and perhaps increased awareness of the species. Most occurrences in the data are recorded as GPS coordinates, however some are recorded simply as streets or roads. For these named sites, GPS coordinates were extracted from the place name database available from Land Information New Zealand[4]. Some occurrences were missing the year of discovery, in which case the years were estimated based on the year of detection of neighbouring occurrence sites (see Appendix C for data).

Resulting distributions. A qualitative assessment of dispersal events showed that, for both NZ and the USA, the majority of dispersal events occurred over relatively short distances, the distribution of distances for each country included several large distance dispersal events (Figure 5.4). The USA data had a greater proportion of dispersal distances distributed between 100-700 km.

The distributions of nearest-neighbour distances for each country look similar to an exponential distribution, but the NZ distribution has a much larger λ value (see section 3.3.2). The extreme outliers in both graphs appear to be evidence of a “fat-tailed” distribution. Previous research has suggests that it is often these rare extreme distances that contribute significantly to the speed of

---

[4] http://www.linz.govt.nz/

invasive population spread (Kot et al. 1996, Lewis 1997, Neubert & Caswell 2000), thus a Cauchy distribution, appropriate to describe such data, was used to model the dispersal distances.

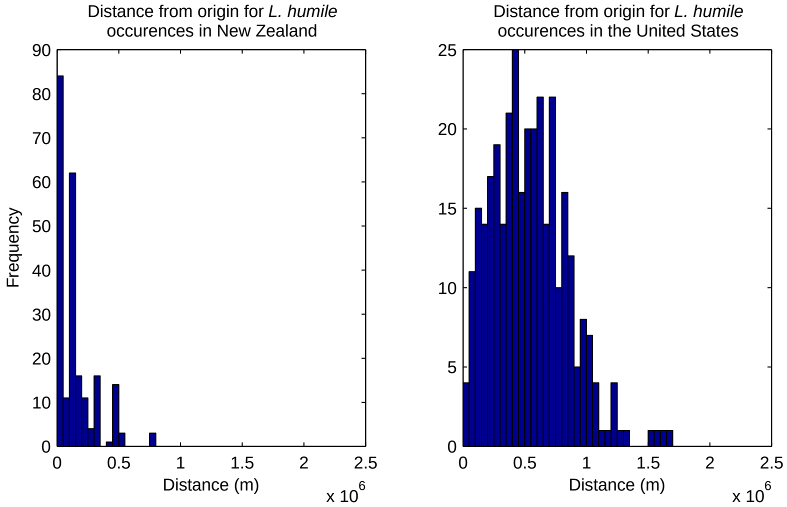
The NZ distance-from-origin graph (Fig. 5.4a) was similar to the nearest-neighbour graphs, however the USA dispersal distances had a peak frequency at around 500km. The greater abundance of larger distances in the USA is because:

- North America is part of a large continent while NZ is an island nation where extremely large distances are not possible. In the case of the nearest-neighbour measurements, the addition of more occurrence sites results in the minimum distance to any previous site shrinking and this effect would be more obvious for smaller regions or countries.
- North America has hosted *L. humile* since 1891, so it has spread for over a hundred years, whereas *L. humile* has only been present in NZ for 16 years. While this shouldn't influence nearest neighbour measurements greatly (except as noted above), but distance-from-origin measurements will increase as time progresses as new populations continue to establish further away from the initial source.
- The data from NZ and USA were mapped in different ways. New Zealand data comprises relatively accurate point occurrences whereas the USA data comprises county centroids. In the USA the distances are limited to a lower bound influenced by the size of the counties involved, even if in reality the dispersal events had occurred over smaller distances. For example, when a dispersal event simply crosses from one side of a county border to the other the distance recorded would be from the centroid of the originating county to the next.

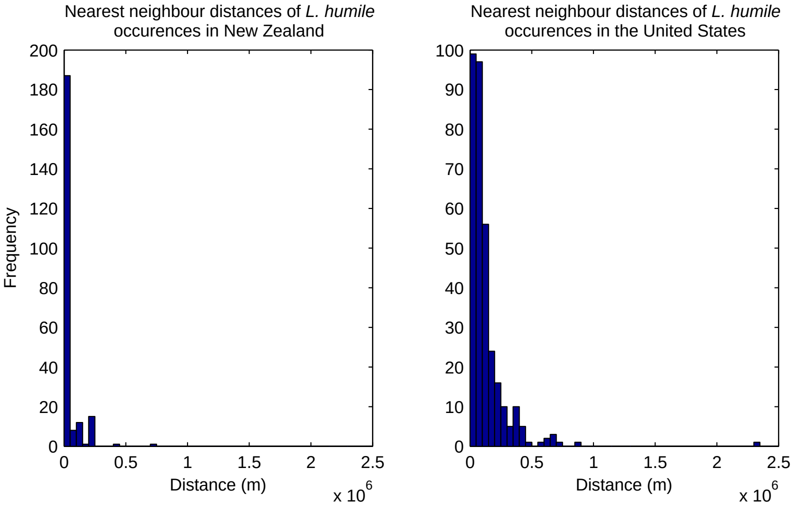
The Cauchy distribution. A Cauchy distribution that describes rare events occuring at extreme distances and that are often important for species dispersal was fitted to the USA nearest neighbour distances (Hengeveld 1989, Higgins & Richardson 1999, Clark et al. 2001, Suarez et al. 2001, Brockmann et al. 2006).

The apparent difference between the distribution of nearest-neighbour distances for each country is less than that difference for distances from the origin. The nearest neighbour distances are also likely to be closer to actual distances travelled by a dispersing species.

To fit the Cauchy distribution to the USA nearest-neighbour dispersal distance distribution, maximum likelihood estimation (MLE) was used with parameters selected in MATLAB®(2007a, The MathWorks, Natick, MA) by an unconstrained non-linear optimization process known as the simplex search method (Lagarias et al. 1998).



(a) Origin



(b) Nearest Neighbour

Figure 5.4: The distances travelled by *L. humile* long distance dispersal events in NZ and the USA, using both a) distance-from-origin, and b) nearest-neighbour, methods of distance measurement. Note that there were several events that fall outside of the x-axis range in the distance-from-origin graphs, but which can be seen in the nearest neighbour graphs. This is simply the result of keeping the same scale so that distance measurement techniques can be compared.

Table 5.2: Parameter estimates for the *L. humile* dispersal model. Estimates are for the Cauchy (x_0, γ) and Poisson (λ) distributions that describe the USA dispersal events and nearest neighbour distances respectively. For the Cauchy distribution estimates were made by MLE with the final estimate having a negative log-likelihood of 4479.88.

Parameter	Estimate (6 s.f.)	95% Confidence Interval (6 s.f.)
x_0, location	1.00000	± 99.6940
γ, scale	83673.4	± 10973.3
λ, Poisson mean	0.297794	$[0.199264, 0.427983]$

The Cauchy distribution's PDF is:

$$f(x; x_0, \gamma) = \frac{1}{\pi\gamma[1 + (\frac{x - x_0}{\gamma})^2]} \tag{5.2}$$

where x_0 is the location parameter and γ is the scale parameter. The parameter values as determined by MLE, along with 95% confidence intervals, are listed in table 5.2.

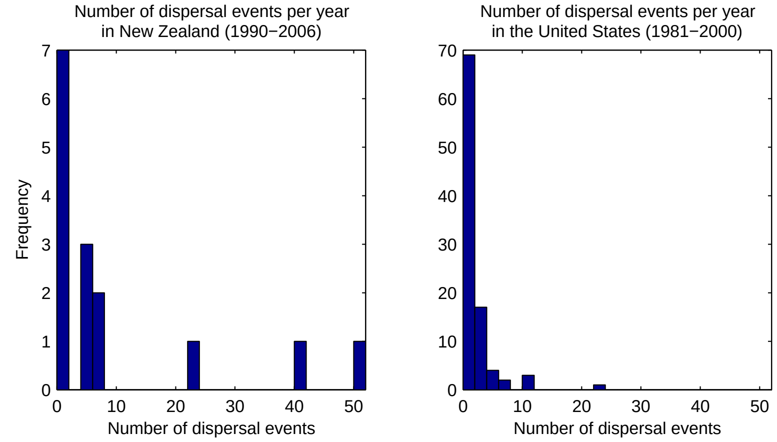
Frequency of dispersal events

Another distribution required to model dispersal is the frequency of long distance dispersal events. A suitable distribution is the Poisson distribution described by its mean λ which is used as a parameter for the **kernel** module (see section 3.3.2).

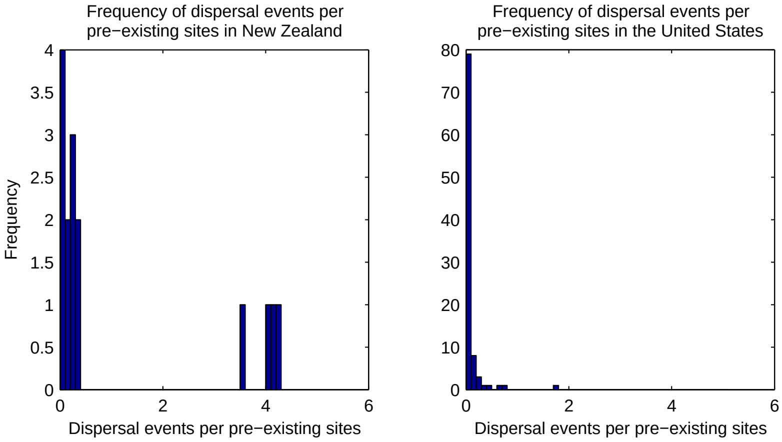
In comparison to the USA data, NZ has far fewer dispersal events but follows a similar trend with a higher frequency of a small number of events per year (Fig. 5.5a). For the New Zealand data, three occurrences of many (> 20) dispersal events in a single year are due to national surveys to determine the extent of *L. humile.* The dispersal events recorded in these surveys were probably present earlier than when they were recorded and so are a little unrealistic.

In the *L. humile* spread model, the number of dispersal events per pre-existing sites each year is modelled by the Poisson distribution. Jump frequency measured this way had mean values less than one for both NZ and USA data (Figure 5.5b).

Unfortunately, the number of sources also increases through the **local** module at a rate of up to $4t - 2$ per year from a single site (Section 4.1). This additional route by which the number of populations can increase means that the rate of increase in the number of occupied sites will be greater than measured from the USA data. However, the **survival** component can also remove potential populations, with a rate dependent on the suitability of the underlying habitat. We make the assumption that the creation of new population sites by the **local** module, and the removal of potential populations by the **survival** module, will cancel one another, preventing the discrepancy



(a) Frequency of dispersal events per year



(b) Frequency of dispersal events per year by pre-existing sites

Figure 5.5: The frequency of *L. humile* long distance dispersal events in NZ and the USA. Measuring the a) number events per year, and b) the number of events per year divided by pre-existing sites.

from becoming too great. Additionally, the measured λ will be less than the true value, due to both undetected populations and the coarse grain of the USA data where potential long distance dispersal events that did not cross county borders won't have been recorded.

In the *L. humile* simulation, the population distribution was represented by a raster map with a resolution of 150m which is the average rate of of population diffusion by budding. This resulution is less coarse than the entire counties used in the USA, but not fine as the NZ point data.

5.5.4 Simulation Region

The region simulated was constrained to the North Island of NZ, an area of approximately 114,000 km^2, at a raster resolution of 150 m. Within the GIS, the simulation was carried out using the 1949 New Zealand Geodetic Datum and the New Zealand Map Grid projection.

To simulate the start of *L. humile*'s real spread, the initial occurrence sites at the Mt. Smart stadium in Auckland and 3 nearby populations also discovered in 1990, the first year of *L. humile*'s arrival in NZ (Figure 5.6), were used.

5.5.5 Potential Distribution

The **survival** module requires a map indicating the suitability of regions for a species' establishment. Several methods for determining the potential distribution and suitability of regions in NZ for *L. humile* were investigated.

Harris (2002) combined manually rated land-cover classes and mean annual temperature ranges to generate a suitability map for NZ. To get a more accurate suitability map, we used the same technique as Harris (2002) but replaced mean annual temperature with suitability measured by degree days required for successful development (Hartley & Lester 2003).

The degree day model specifies the cumulative number of degrees above a developmental threshold required for a species to complete development. For *L. humile*, complete development from egg to adult requires approximately 445 degree-days above a threshold of 15.9°C (Hartley & Lester 2003). Our duplication of Hartley & Lester's (2003) work included data up until 2005, and used air temperature values whereas Hartley & Lester (2003) generated suitability distributions for both air and soil temperatures. Hartley & Lester (2003) note that Maps based on soil temperature data indicated suitable conditions as far south as Central Otago, while air temperature predicted a more limited northerly distribution in Northland and Hawkes Bay. Soil temperatures showed a larger number of suitable areas. Egg to adult worker development time at the 15.9 ± 0.8°C threshold was 445.4 ± 18.9 cumulative degree-days, and more for adult males. On average, across

Figure 5.6: Initial occurrences of *L. humile* in NZ at the end of 1990. White circles indicate occurrence sites. The red dashed line represents State Highway 1, and the green solid lines, railways. The close proximity of the initial sites to major transportation routes may have assisted in *L. humile*'s spread.

sites recorded in New Zealand, soil degree days were 25% greater than air degree days (Hartley & Lester 2003).

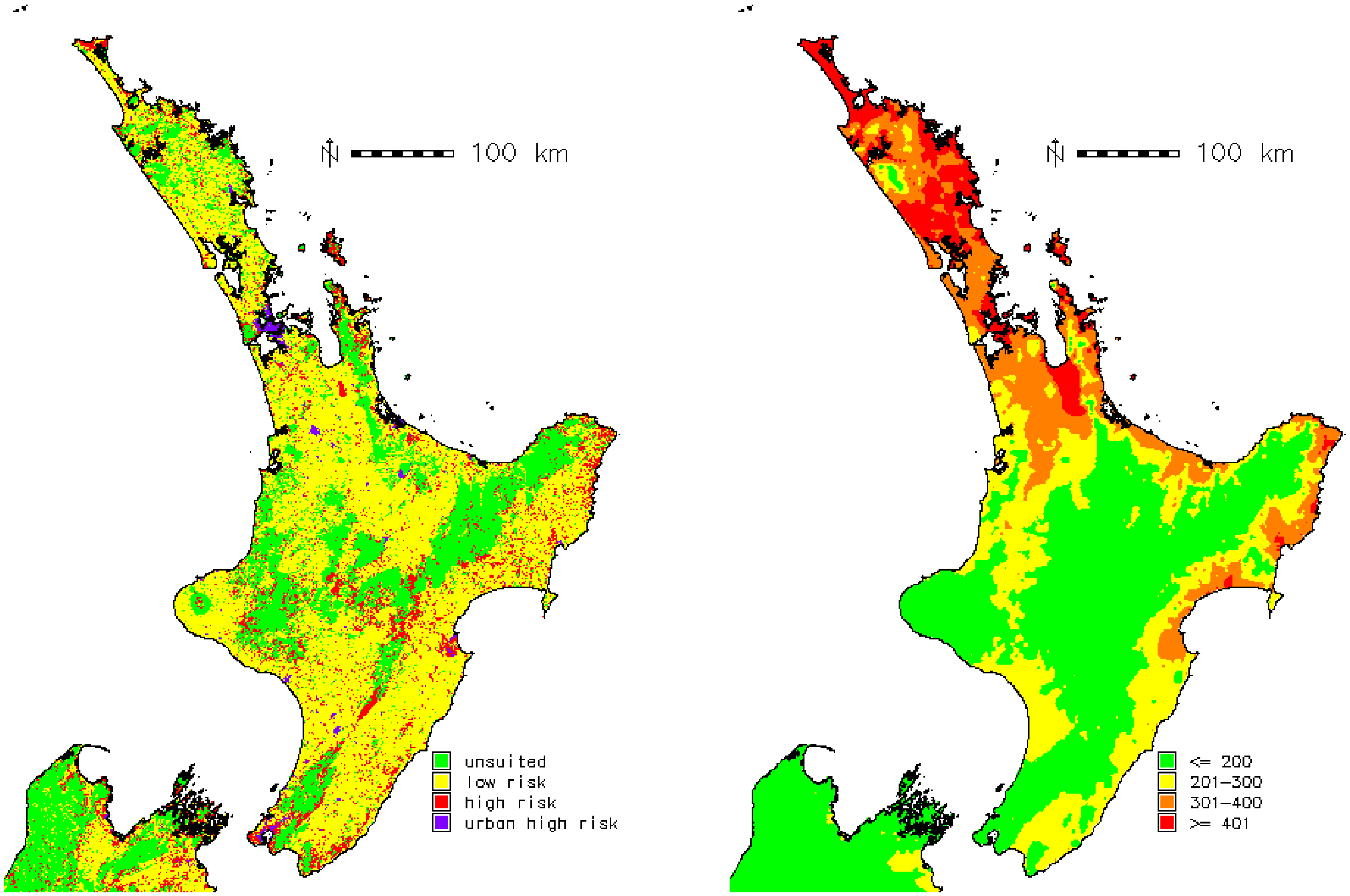
Degree-days based on air temperature clearly limit *L. humile*'s potential distribution map compared with degree-days based on soil temperatures. Air temperatures would restrict foraging and so they are a limiting factor. Air temperatures were also used to determine suitability because country wide data were available for the past 30 years, whereas only a few sites had a sufficient history of soil temperatures.

For the suitability map based on land-cover, the NZ Land Cover Database version 2 was used rather than the first version that was used in Harris (2002). This allowed for changes and corrections in the land-cover database to be included. Land-cover types were divided into 3 categories from unsuitable to high suitability as per Harris (2002) (H0, H1, H2), with a separate category for urban areas (HU) because *L. humile* has been found to survive at lower ambient temperatures in urban environments (Harris 2002), because of their close association with human activity and suitably warmer micro-climates being created by human activity (Suarez et al. 2001). The separation into these categories by (Harris 2002) was based on an extensive review of literature and expert knowledge. For example, both urban environments and scrubland are considered highly suitable (H2), whereas indiginous forest is unsuitable (H1) for *L. humile* survival. Relationships between habitat and *L. humile* were discussed earlier in section 5.2.3.

The suitability map according to land-cover showed a great deal of heterogeneity with the majority of the North Island classified as low suitability (pastoral land as well as pine forest and harvested forest), but interspersed with many small highly suitabile patches (Figure 5.7a). Several large regions of unsuitable habitat are concentrated half-way up the island mainly consisting of indigenous forest, but still interspersed with highly suitable habitat such as scrubland.

Thirty years of historical daily minimum and maximum temperatures interpolated as a grid with a latitudinal and longitudinal resolution of 0.05° were used (data from the National Institute of Water & Atmospheric Research) to create a map of the annual average degree-days above a threshold of 15.9°C available for *L. humile* development (Hartley & Lester 2003). This map was transformed to the NZ Map Grid (NZMG) projection, and interpolated, using inverse distance squared weighting, to a resolution of 150 m. Based on the degree-days required for development, the degree-day map was split into 4 categories of total degree days from low to high suitability ($T0 \leq 200$, $T1 = 201 - 300$, $T2 = 301 - 400$, $T3 \geq 401$).

The degree-day suitability map classified the majority of the central North Island as too cold, bordered by low suitability, with medium and high suitability regions reaching into the upper third of the North Island and bordering the east and north-east coasts (Figure 5.7b).



(a) Land-cover suitability

(b) Degree-days

Figure 5.7: Suitability maps for *L. humile* in the North Island of NZ, based on a) land cover classification, b) degree-days above a threshold of 15.9C.

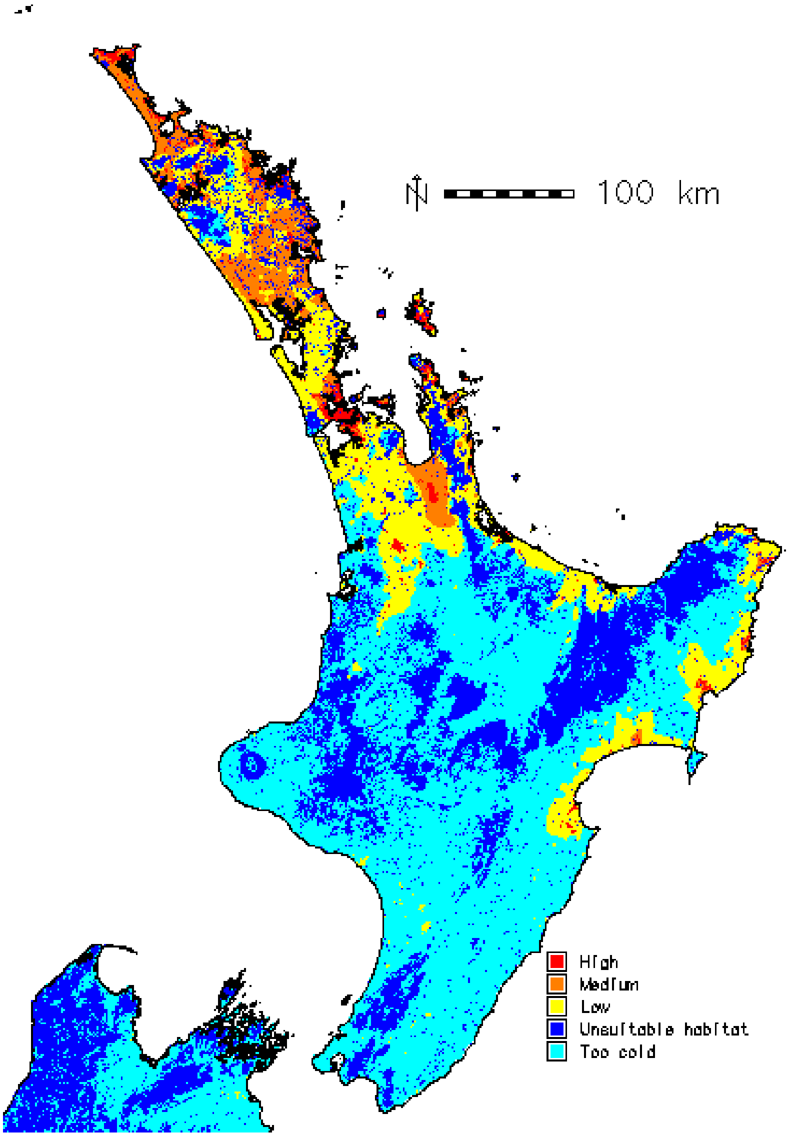


Figure 5.8: Combined suitability map for *L. humile* in the North Island of NZ.

Table 5.3: Scheme used for combining land-cover and degree-day suitability maps. H* and T* represent *any* category of habitat and degree-day suitability respectively.

Combination	Result	p(survival)
T3+H2, T3+HU, T2+HU	High suitability	100
T3+H1	Moderate suitability	80
T1+HU, T1+H1, T2+H1, T1+H2, T2+H2	Low suitability	50
H0+T*	Unsuitable habitat	10
H*+T0	Too cold	10

The land-cover and degree-days suitability maps were combined using the scheme presented by Harris (2002) as shown in Table 5.3. Harris uses mean annual temperature classified into suitability levels, whereas here degree-days have been substituted as they are biologically more relevant and possibly more accurate. The result was a map classified into low, medium, and high suitability, with additional categories for regions with unsuitable landcover and regions that were too cold. Each category was assigned a survival probability indicating the probability of occupants surviving to the next year, with progressively lower probabilities for less suitable categories (Table 5.3). These survival probabilities comprise the biggest assumption in this simulation model as there is no information available on how annual population survival rates relate to habitat suitability.

The overall suitability map was constrained predominantly by the degree-day map, which resulted in most of the central and southern parts of the North Island classified as too cool for *L. humile* development (Fig. 5.8). Much of Northland, along with Auckland and other northern cities and towns, showed medium or high suitability. North-eastern and some Eastern coasts showed low suitability.

5.5.6 Uniform spread model

When monitoring the spread of a species, a large cost factor is the area that needs to be sampled. For this reason it makes sense to try and reduce the area necessary for searching or to distribute the surveying sites more efficiently based on where a species is expected to occur.

To evaluate the performance of the *L. humile* simulation model against a simple spread model, uniform radial spread from the mean centre of the 3 initial invasion sites was modelled. A linear increase in the square root of the range occupied by an introduced species is often observed during the invasion process (Hengeveld 1989). A linear increase in the square root of area is equivalent to a constant spread rate for the boundary or radius, r, of a circular area, A:

$$A = \pi r^2 \sqrt{A} = r\sqrt{\pi} \tag{5.3}$$

The rate of radial increase for the uniform spread model was calculated from the approximate slope of the square root of area covered by the simulation model. Thus, the rate of increase in radius of the uniform spread model is given by:

$$\Delta r = \frac{\sqrt{A_{t1}/\pi} - \sqrt{A_{t0}/\pi}}{t1 - t0} \tag{5.4}$$

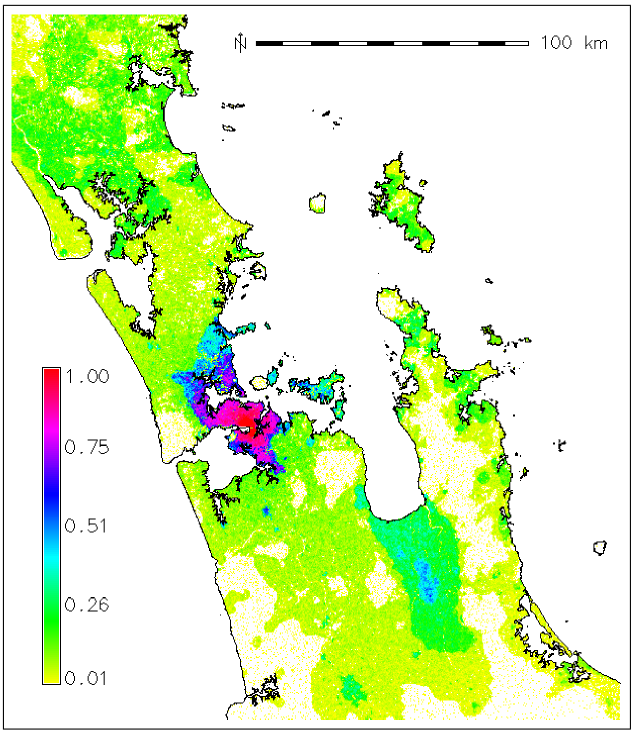
where A_t is the area encompassed by the simulation model at time t, and $t1$ is some time after $t0$. This calculation has the result that both models approximately encompass the same total area at any time-step. Simulation results indicated that at certain phases of increase in occupancy the square root of the area increased at approximately a linear rate (Fig. 5.10), and for comparison $t1$ was the final year of simulation (2005). $t0$ was the initial year of simulation (1990) when the uniform spread model was based on the entire area of the occupancy envelope, and $t0$ was 1995 when the uniform spread model was based on the area of the thresholded occupancy envelope. The different values of $t0$ are necessary as there is insufficient agreement between replicates when using the thresholded occupancy envelopes.

5.6 Results

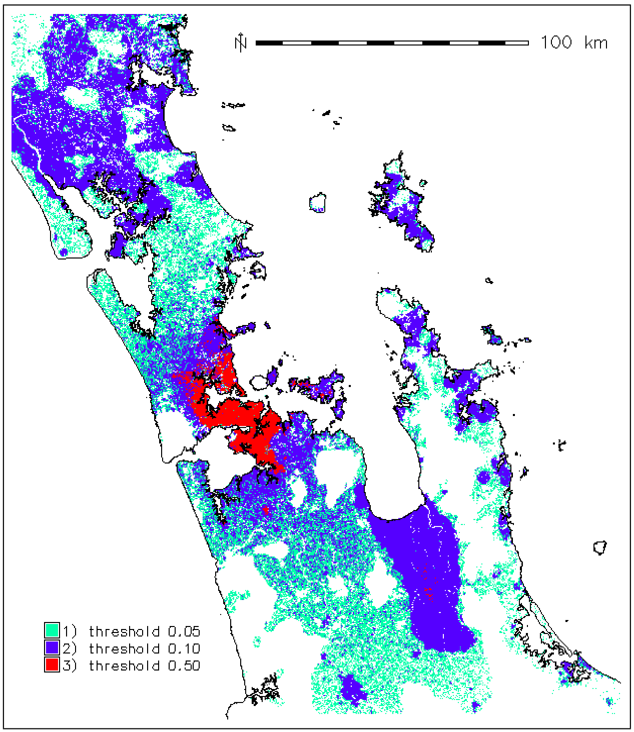
To test the sensitivity of the model to variation in its parameter values, the parameters for the Poisson mean, λ, and the scale parameter, γ, of the Cauchy distribution were simulated for their estimated value and their 95% confidence interval (CI) limits, for a total of 9 parameter combinations. Both the kernel and survival modules involve random sampling from appropriate probability distributions and are therefore stochastic processes. Thus, each parameter combination was simulated 100 times from 1990 to 2005 to give a total of 900 replicates. The experiment took approximately a week to run on a single Pentium 4 3.2 GHz PC.

5.6.1 Occupancy Envelopes

To get an average representation of *L. humile* spread, and a meaningful result for decision support, all the maps across the replicates, for a given year and parameter combination, were averaged to create an occupancy map that indicates the probability that a raster cell will be occupied at a given time. The probability is simply the number of replicates in which the cell was occupied divided by the total number of replicates (100). This occupancy map can then be masked by excluding areas below a given threshold, such that only raster cells with a value greater than or equal to the threshold are present, ignoring regions with a low probability of occupancy (Fig. 5.9).



(a) Occupancy envelope



(b) Thresholds

Figure 5.9: a) The occupancy envelope generated by a simulation ($\lambda = 0.298$, $\gamma = 83700$, Year 2005) and b) the same occupancy envelope split into thresholded occupancy rates.

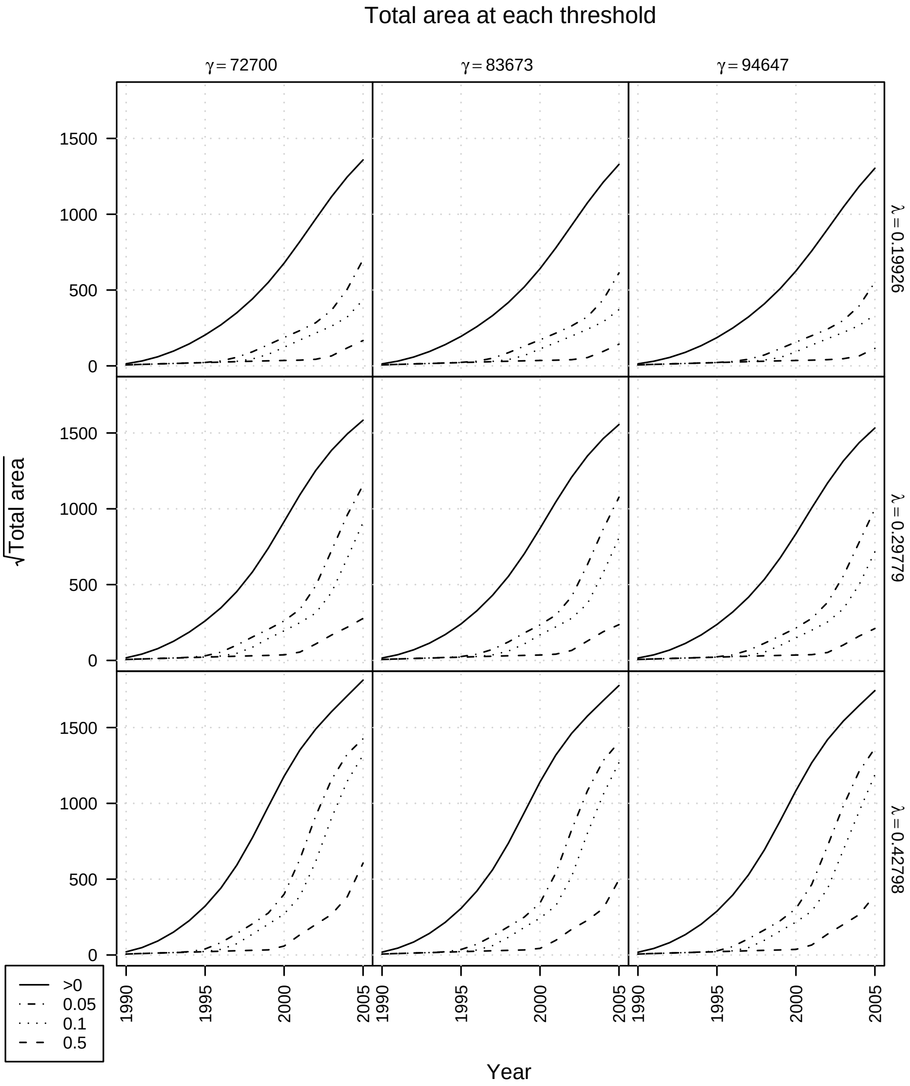


Figure 5.10: Total area encompassed by the *L. humile* simulation envelope at each threshold for the 9 combinations of γ and λ simulated.

5.6.2 Total Area

Figure 5.10 shows how the square root of the occupancy envelope's area ($\sqrt{A}$) increased through time. The rate at which $\sqrt{A}$ increased, for all occupied sites and with no occupancy threshold, accelerated up until 2000-2002 before the rate slows slightly. The graphs for various occupancy thresholds, regardless of the combination of λ and γ, seem to show 3 phases with different rates of increase, the first up until ~1995 for 5% occupancy, ~1996 for 10% occupancy, and ~2001 for 50% occupancy. The initial phase shows a very slow increase in area probably because replicates have little agreement between them. Up until occupancy thresholds reach a $\sqrt{A} = 400$ (Fig. 5.10), occupancy increases almost linearly, before reaching a third phase where the thresholded

$\sqrt{A}$ increases at a faster rate. For $\lambda = 0.42798$, once $\sqrt{A}$ reaches approximately 1200 the rate of increase begins to decelerate.

The maps (Figure 5.15 and Appendix C) show that the different rates in the second and third phases may be the result of the initial invasion area occupying a narrow isthmus between the main land mass of the North Island and Northland. However, once the invasion has spread far enough, it reaches inland regions that contain more area available for occupation. With a uniform environment, the rate that $\sqrt{A}$ increases would expect to accelerate before reaching a linear rate of increase (see section 4.2.1).

Increasing γ from 72700 through 83673 to 94647, decreased the area covered by the occupancy envelope for any given time and delayed when the phases of increase in $\sqrt{A}$ occur. In contrast, increases in λ have the reverse effect, increasing the area occupied at any time an bringing about the onset of the $\sqrt{A}$ phases of increase sooner (Fig. 5.10).

5.6.3 Comparison of simulation with occurrence data

To assess the performance of the simulation model its predictive ability, the percentage of sites that lie within the envelope at each year, was calculated. This percentage was calculated by comparing the occupancy envelope with the observed *L. humile* occurrence sites within New Zealand for that year (Fig. 5.11).

All occupancy thresholds (the threshold value used to create an occupancy envelope) initially included a high percentage of observed sites before the percentage dropped to around 40% at about 1993-1994. The percentage of observed sites within the envelope that includes all sites that have been occupied at least once increases in 1996. The 5% occupancy threshold follows (the envelope that includes sites that are occupied for at least 5% of simulation replicates), peaking in 1999 before dropping in 2001. From 2001 the percentage of observed sites encompassed by all the occupancy envelopes increases.

Across parameter combinations, performance as assessed by the percentage of sites encompassed by the occupancy envelopes showed the same trend of high performance at the start of the invasion, decreasing at around 1992-1993, peaking at 1998, and decreasing slightly again around 2001, before increasing again. Increasing γ generally decreased the overall performance, except in the case of λ=0.199, where performance increased. The upper bound of the λ 95% CI, $\lambda = 0.428$, showed less difference between peaks and lows, and also generally increased performance (Fig. 5.11).

A national survey was carried out by the Department of Conservation in 2001 to determine the extent of *L. humile*'s distribution in New Zealand, which may explain the decrease in predictive ability in that year as well as making it somewhat difficult to assess performance as many of the sites found in 2001 would have established earlier.

5.6.4 The uniform spread model

The simulation and the related uniform spread model are compared in Figure 5.12. The dense conglomerate of observed sites at the initial area of invasion are quickly encompassed by the uniform spread model, but the more dispersed observed sites were only predicted by the simulation model.

When the performance of the simulation model using the mean parameter values ($\lambda = 0.298, \gamma = 83700$) is compared to the simple uniform spread model, the benefits of using a more complex model at different phases of the invasion become apparent (Fig. 5.13). The 5% occupancy envelope predicts the observed sites much earlier than the equivalent uniform radial spread model in all but one year. However, higher occupancy thresholds did not perform as well as their equivalent uniform radial spread model until near the end of the invasion, despite covering the same total area. The trend however, clearly indicated improving performance of the simulation model as the invasion proceeded.

Comparing the performance of the simulation model using other parameter combinations to that of the uniform spread model shows no major differences. The lower limit of the 95% CI of γ resulted in the prediction of observed sites earlier relative to the uniform spread model. The upper 95% CI limit for λ shows less benefit using the complex simulation model at the 5% occupancy level between 1995 and 2000, but thereafter it performs similar to or better than the mean λ. Higher occupancy thresholds (10% and 50%) showed greater performance after 2000 for the upper 95% CI of λ (Fig. 5.13).

An alternative comparison of simulation model to the uniform spread model is to examine the occupancy percentage for each site and order the sites by distance as in Figures 5.14a and 5.14b. These figures also show the point that the uniform spread model encompass the sites for each threshold. As the simulation progresses the sites that lie beyond the uniform spread boundaries become more frequently occupied and demonstrate that these sites are where the simulation model is significantly more informative than just assuming uniform spread.

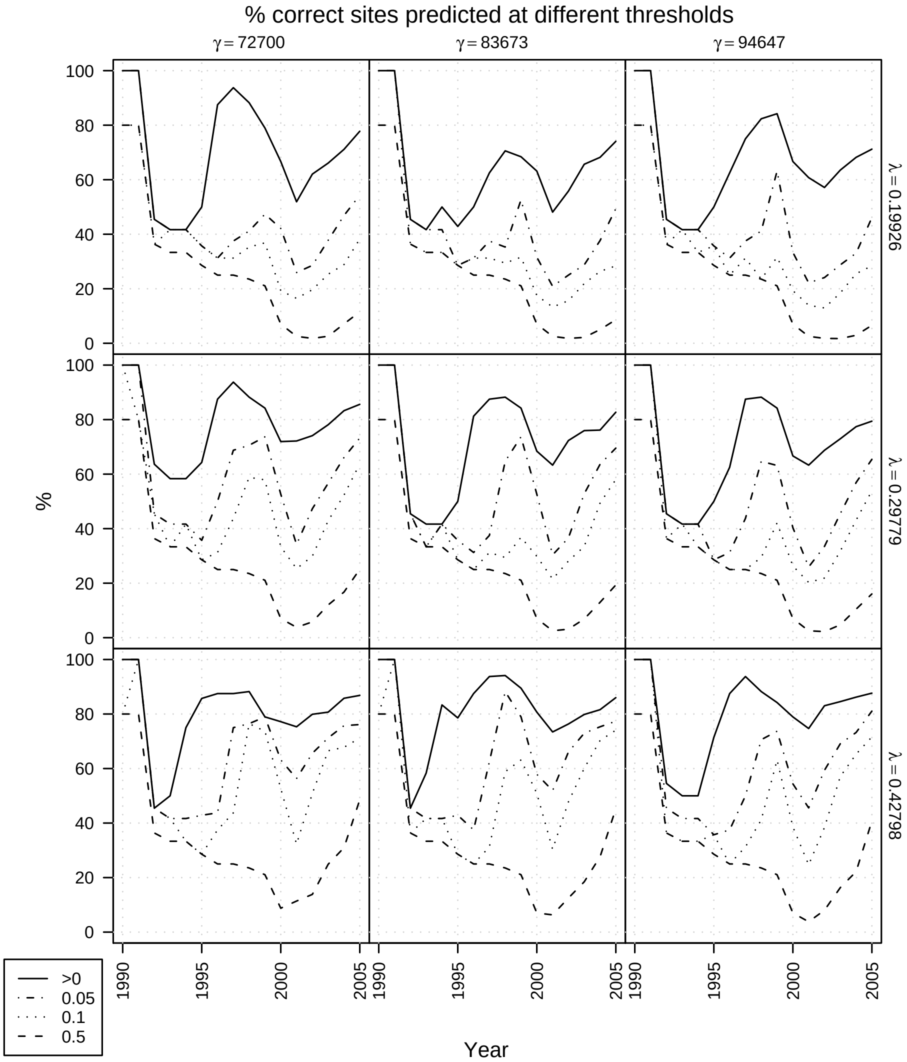


Figure 5.11: The percentage of *L. humile* occurrence sites that lie within the probability envelope of the simulated distributions for each year of the simulation. The percentage that lie anywhere in the probability envelope (i.e. at least one simulation out of the 100 replicates predicts the occurrence of a site) was calculated, as was the percentage for envelopes that have been truncated at the thresholds of 0.05, 0.10, and 0.50. In other words, values below these thresholds are masked out of the envelope.

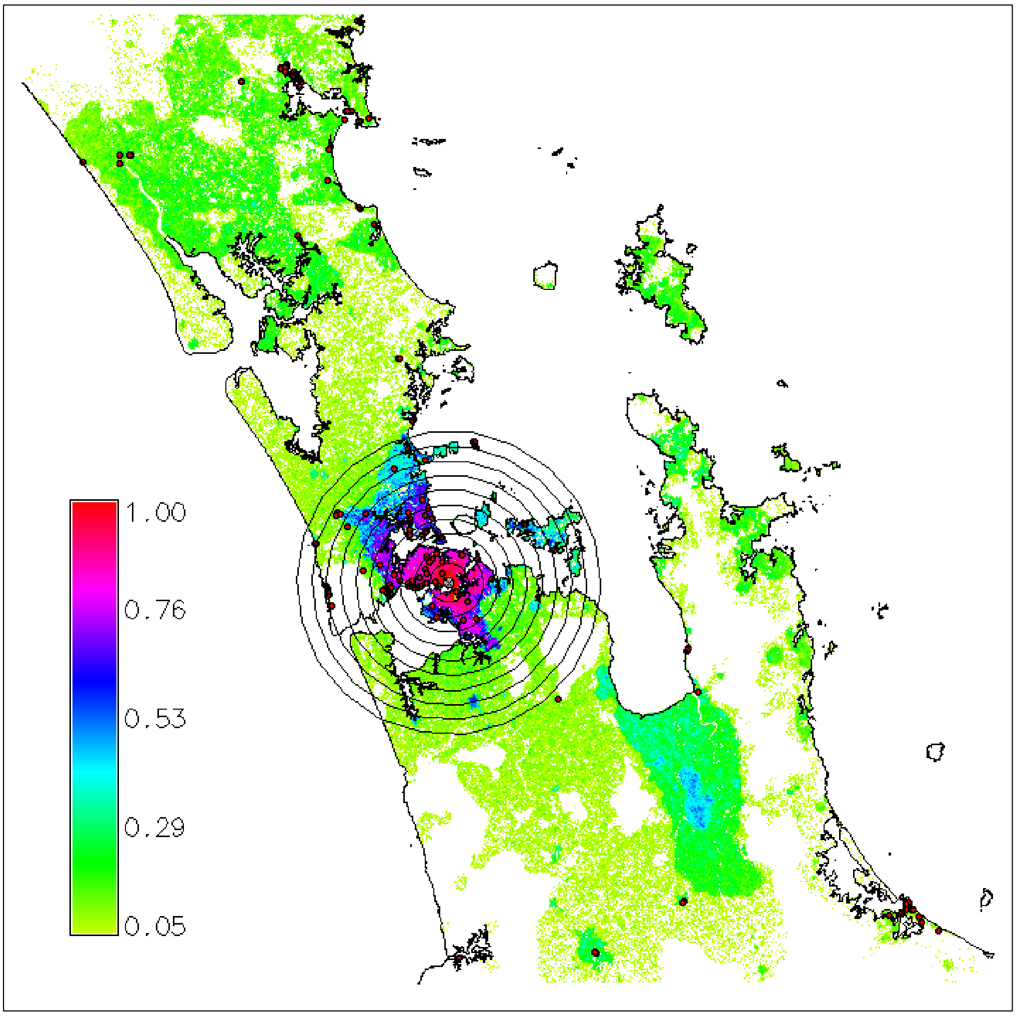


Figure 5.12: Map of uniform spread vs **MDiG** simulation. The lines indicating the fronts for the uniform spread begin from the centre circle at 1995 and procede annually to the outer front for 2005. Red dots indicate occurrence sites for *L. humile* as of 2005. The uniform spread model is based on the area increase of the envelope thresholded at 5%.

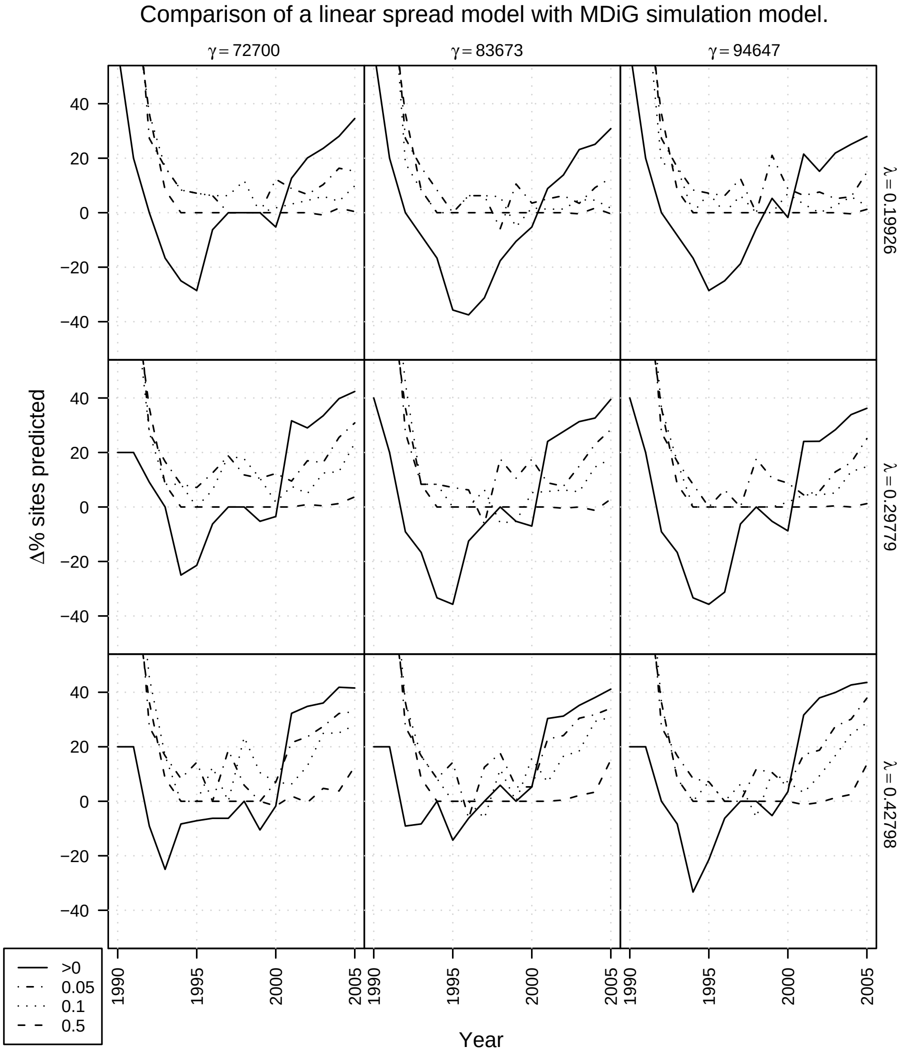
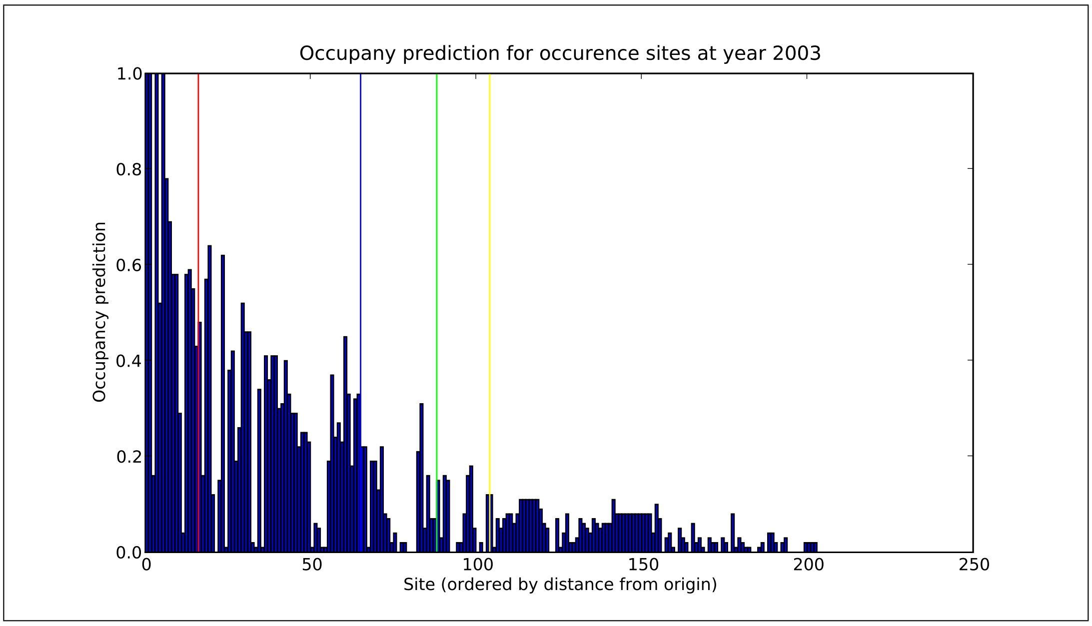
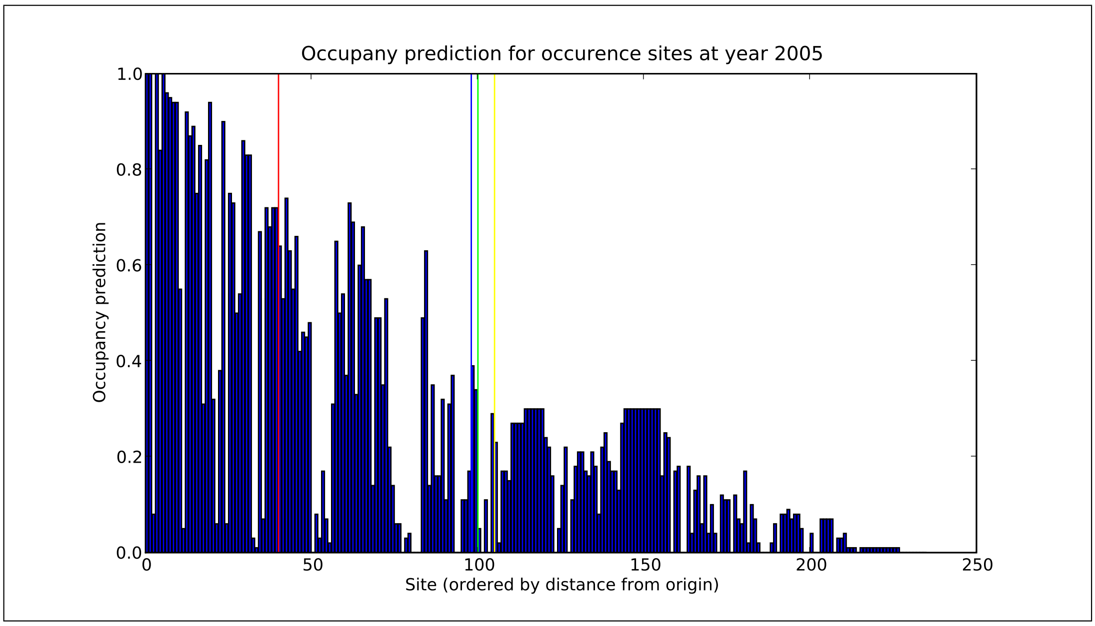


Figure 5.13: Comparing the percentage of correctly predicted sites between the **MDiG** *L. humile* simulations and the approximately equivalent linear spread models. A positive Δ% indicates the **MDiG** model out-performs the linear spread model.



(a) 2003



(b) 2005

Figure 5.14: Prediction for chance of site occupancy for individual sites. Each bar represents an occurrence and are ordered left to right based on ascending distance from invasion origin. Vertical coloured lines correspond to the distance that each uniform spread model reaches for a particular threshold (red 0.50, blue 0.10, green 0.05, yellow none). Graphs are for $\lambda = 0.298$, $\gamma = 83700$.

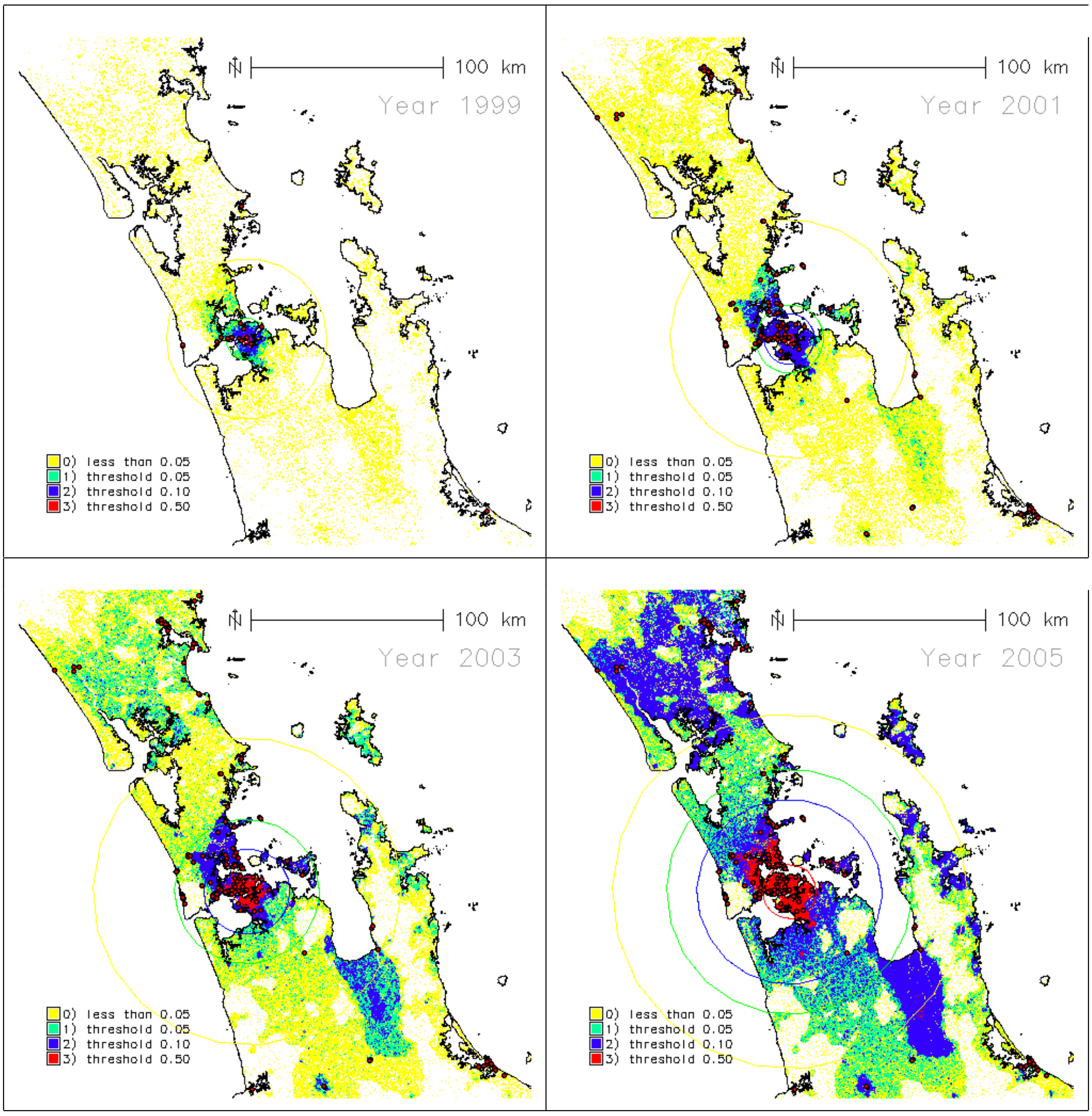


Figure 5.15: Snapshots of the dispersal simulation ($\lambda = 0.298$, $\gamma = 83700$) for years 1999, 2001, 2003, and 2005. The map area is coloured according the threshold the probability envelope exceeds (red > 0.5, green > 0.1, blue > 0.05, yellow < 0.05). The coloured circles indicate the boundary for a uniform spread model, the area of which is equivalent to the area encompassed by the thresholded probability envelope. Red dots with black outlines indicate real *L. humile* occurrence sites.

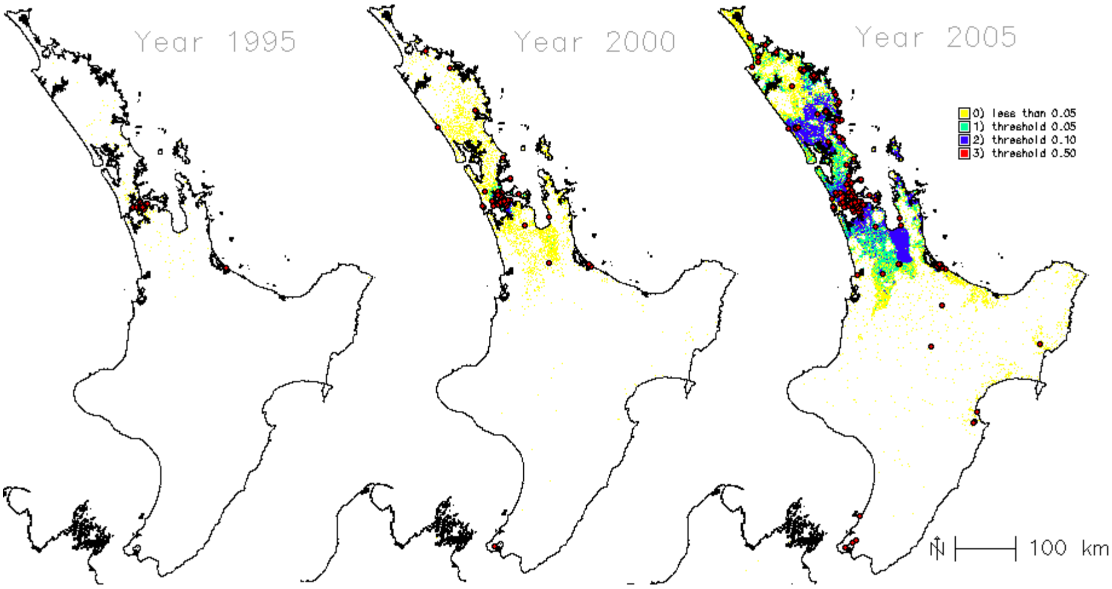


Figure 5.16: Snapshots of the entire simulation region ($\lambda = 0.298$, $\gamma = 83700$).

5.6.5 Hotspots

Hotspots, or areas with high occupancy levels across simulation replicates, unsurprisingly included regions near the invasion epicentre, within and near Auckland city, and nearby cities such as Whangarei and Hamilton, both of which have real-world *L. humile* infestations (Fig. 5.15). Towns near Auckland, such as Pukekohe and Waiuku, were also hot spots, but as yet have no recorded occurrences. Great Barrier Island, Little Barrier Island, Ponui Island, and Tiri Matangi Island are also indicated as hotspots. Great Barrier Island and Tiri Matangi Island have had occurrences recorded, most likely from human-assisted dispersal. Both islands have undergone poison baiting treatments to try to eradicate *L. humile* (pers. comm. J. Boow, Biosecurity Team Leader, Auckland Regional Council and pers. comm. C. Green, Technical Support Officer, Department of Conservation). There was one large hotspot covering most of the Hauraki Plains which has no recorded occurrences of *L. humile*. The land-cover map indicates that in this location there is a large area of scrubland that is a highly suitable habitat for *L. humile* (Suarez et al. 1998) along with several small urban regions. The whole of the Hauraki Plains also had a sufficient number of degree-days for complete *L. humile* development.

No replicates for any combination of variables reached the southern end of the North Island, despite that *L. humile* has been recorded there since 2000. This is probably due to the large area of unsuitable habitat between the invasion epicentre and the south of the North Island. The occurrences in southern end of the North Island likely arrived from Auckland, with populations hitch-hiking on road or rail networks. Explicit modelling of dispersal along transport networks may result in better agreement between occupancy envelopes and these distant southern occurrences. Both Nelson and Christchurch, in the South Island of NZ, had occurrences recorded but are outside of the simulation region.

Over the entire simulation area (Fig. 5.16), populations were mostly excluded from areas classified as unsuitable habitat or too cold (see Table 5.3). As expected high suitability regions were quickly occupied with a greater frequency comparied with low suitability regions. Low suitability regions also took longer to show significant occupancy probabilities as they needed high propagule pressure from high suitability regions. There were a few populations shown at low occupancy levels in areas that were too cold or that had unsuitable habitat. These were probably the result of high propagule pressure.

Lowering γ (72700) resulted in little qualitative difference in the pattern of hotspots and occupancy frequencies, except an increase in occupancy frequency later in the simulation in regions nearer the invasion epicentre. The reverse was the case when a higher γ (94600) was simulated (Appendix C).

At the upper limit of the 95% CI for λ (0.428) the rate of spread increased and occupancy probability was much higher. At $\lambda = 0.428$ the population encroached into areas that were too cold, or that had unsuitable land-cover, because of high propagule pressure into these areas despite the high probability of extinction. Conversely, λ at its lower bound (0.199) resulted in a slower spread with lower occupancy levels (Appendix C).

Despite the differences in model output in response to variation in its parameters, the overall result was unsurprisely a pattern similar to the underlying landscape suitability map spreading from invasion epicentre, but differing mainly in magnitude of occupancy frequency. The simulation was most affected by differences in λ. λ controlled the number of dispersal events, with a cumulative effect such that more events earlier in the simulation resulted in more occupied sites from which successive events could occur.

5.7 Discussion and Conclusions

This study demonstrates that it is possible to estimate how and where an invasive species might spread, by calibrating a stochastic and spatially-explicit presence/absence model to observed data and combining with expert knowledge.

Often the costliest factor involved with monitoring, controlling, or eradicating an invasive species is the extent of the area to be considered. For this reason simulation results were compared with a uniform circular spread model with constant radial spread rate from the invasion epicentre. While the unifrom spread model performed well early in the invasion the simulation model predicted the observed long distance events more rapidly as the invasion proceeded and identified hotspots for searching. The simulation model for *L. humile* matched many observed sites with hotspots and also indicated that the Hauraki Plains in NZ may potentially harbour populations of the ant despite having no recorded occurrences.

No replicates reached the southern end of the North Island, despite records of *L. humile* establishment in urban areas since 2000. This result is consistent with other attempts to predict the spread of other species, where the models often underestimate the rate of spread of invading organisms. This failure is attributed to inaccurate measurement of rare long-distance dispersal events (Andow et al. 1990, Liebhold et al. 1992, Buchan & Padilla 1999, Neubert & Caswell 2000). These results reinforce its importance. Without long-distance dispersal, the spread of *L. humile* would have progressed much more slowly. Indeed, *L. humile* would not have spread beyond two kilometers from each of the three initial occurrences. Numerous studies (e.g. Kot et al. 1996, Lewis 1997, Neubert & Caswell 2000) that use integro-difference equation models for spread show that the long-distance component of dispersal ultimately decides invasion speed, even when long-distance dispersal is extremely rare.

One concern with the spatially-explicit stochastic simulation approach is that early replicate maps will be highly variable, with few areas having occupancy levels sufficient to distinguish hotspots from non-hotspots. However, this variability is realistic, as it is difficult to predict what will happen early in an invasion because of the inherent randomness of the process, and more especially in the presence of long distance dispersal (Higgins & Richardson 1999). On the other hand, the overall long-term result that indicates where species might establish a viable population means the occupancy envelope eventually conforms to the underlying suitability map. When authorities set out to monitor or eradicate an exotic species, the ability to estimate the rate that the invasion occupies suitable regions is preferable to tools that simply indicate the entire potential distribution. The latter, however, is appropriate for pre-border pest risk assessment because at that stage the initial site where the exotic species will arrive and establish is unknown.

Since spatially explicit dispersal models incorporate information on landscape details they suffer less from uncertainty than more simple models (Mooij & DeAngelis 2003) at least later in the invasion process, as was shown here. The occupancy envelopes generated by the simulation had considerable qualitative similarity across parameter combinations. Since the occupancy envelopes themselves would be used to guide the management of invasive species, it is the relative occupancy between regions, indicating those most at risk that is important, rather than the absolute value of the occupancy maps. Similar patterns in the simulation's relative occupancy frequencies matches the observation that landscape heterogeneity can often have a large stabilising effect on ecological models (Kuno 1981, Ruxton & Rohani 1999, Gardner & Gustafson 2004).

The *L. humile* occurrence data for the USA and NZ is in a vector format. New Zealand sites are represented as points, which means every dispersal event and newly established population that occurs can potentially be recorded and there is no coarseness it the occurence data (although in practice, there is coarseness and patchiness in the recording effort). The USA data is different as it's related to counties, so there is an enforced granularity - once *L. humile* has established in a county any dispersal event had to travel far enough to pass the borders of the county. Thus the peak frequency of dispersal events per pre-existing sites is expected to be lower for the USA than for New Zealand, as observed in Figure 5.5b.

There is some controversy about taking data measured at one scale and applying it to another (Levin 1992). In particular, analysis of **MDiG** showed that resolution can have a significant impact on the kernel-based dispersal model (Section 4). The results there showed that coarser resolutions resulted in slower spread due to the granularity of the raster grid preventing events from establishing in already occupied cells. This effect only becomes significant however when the scale of the dispersal kernel was at a similar scale as the simulation resolution, and for the *L. humile* simulation, the Cauchy distribution's scale parameter was more than two orders of magnitude greater than the simulation resolution. Thus the resolution of the simulation is not expected to have a large influence on the results.

The simulation results show similarity to the Skellam (1951) model despite that his model only deals with a contiguously dispersing population from a single, expanding focus. The approximately linear increases in some phases of $\sqrt{A}$ (section 5.6.2) agree with the Skellam's (1951) characterisation of invader spread rate as a linear function of the square root of its range area through time.

Simulating the spread of a species over large spatial scales, such as for the North Island in this study, may also reduce model uncertainty as Lonsdale (1999) suggests that more robust relationships in patterns of invasive species spread are likely to emerge from studies at broader spatial scales. This reduction in uncertainty is due to numerous random events averaging out, although such an assumption can only be made across numerous simulation replicates. Additionally, at broader spatial scales, the range of environmental conditions being considered tends to increase with the extent of the study and thus the relatice difference between “suitable” and “unsuitable” sites become more distinct.

The spatially explicit and stochastic dispersal model used here to predict the invasion history of *L. humile* in NZ shows considerable promise as part of an effective decision support system for pest and invasive species management. The advantage of the use of such a model will become more apparent if the realistic realisations of spread are used to aid sampling, monitoring and eradication programs.

Chapter 6

General Discussion

> Thinking you know something is a sure way to blind yourself.
>
> – Frank Herbert, “Chapterhouse: Dune”

This chapter concludes this thesis and presents an integration of the results from the previous chapters as they relate to the original aims and objectives of the reasarch. Possible directions for future research in dispersal modelling across heterogeneous landscapes are also discussed.

6.1 Modelling framework

The first objective of this study (Chapter 1) was to research, design, and implement a dispersal modelling framework for simulating species spread across the heterogeneous landscape (Chapter 3). The design of the model required detailed investigation of existing dispersal model paradigms where commonalities were sought so that a system capable of fulfilling the research needs of other modellers could be met. The model was also designed with application to invasive species spread in mind.

MDiG was developed as an open source project, using open source libraries and the open source GIS GRASS. Since **MDiG** and the required supporting software are freely available to researchers and because **MDiG** defines a common model definition format, it facilitates collaboration, and allows for community development among researchers.

The **MDiG** system works by representing the population distribution of a species as a raster map. This map is passed to a series of modules, each of which carries out a particular dispersal or population dynamic. The model sits within the GIS GRASS, so has all the functionality the GIS provides, such that existing GIS tools can be incorporated within a spread model. One particularly useful example is Map Algebra which allows map transformations using simple algebra and logic statements. Other GIS tools are helpful in processing results and general analysis.

A modular system, particularly when combined with open source software, enhances research by allowing researchers to implement modules that carry out processes that are necessary for their model. The sequential series of processing steps in **MDiG**, on a raster map (and potentially on vector shapes as well), can mimic many existing model paradigms (Chapter 2) and thus has scope for wide usage.

As well as creating the modelling framework a new method for calculating the position of population front was designed and implemented. The “density” method for calculating the population front calculates the position at which the greatest change in population density occurs. The “density” method gives similar estimates for the population front as the best-cell method (Sharov et al. 1995) although it may not work as well at quite coarse scales (Section 4.2). Despite this, the “density” method is a different way of investigating the spread rate of a patchy, non-continuous population distribution. This new method may be more suitable than others for populations that have gradual changes within the distribution density due to habitat and landscape variability. In such cases, the best cell method may incorrectly place the boundary within the population distribution.

6.2 Model dynamics

Wolfram (2002) provides an encyclopedia of the patterns generated by different varieties of cellular automata. In a similar manner, this thesis investigated the impact of changing the dispersal neighbour of the local spread module. Depending on the shape of the neighbourhood, the total area occupied by the simulated population distribution after a given time can vary, even if the neighbourhoods being compared have the same area (Section 4.1). So even if they potentially have same increase in new populations, the spatial locations of existing populations are important to determine the future area occupied.

When the results of simulation models in **MDiG** are compared to properties predicted by analytical models (second objective, Chapter 1), it was found that they were similar (Section 4.2). Exponentially bounded dispersal kernels reached a constant spread velocity, whereas non-exponentially bounded (“fat-tailed”) kernels sometimes accelerated but this assessment was dependent on the population front measurement technique. Thus the technique used for assessing the location of population front boundaries influences whether or not these predicted accelerating fronts are seen to occur.

The impact of resolution on increase in area and rate of spread, was also investigated and it was found that continuous dispersal movements (distances sampled from continuous dispersal distance probability distributions) forced into a discrete environment meant that an optimal resolution will exist for the greatest spread rate and the greatest rate of increase for the area a population occupies. Despite this, the choice of resolution will often be more influenced by computational limits as some model processes that involve many local interactions do not currently scale to large areas, such as the kernel module with a large λ, where many dispersal events are generated per patch.

Previously it was thought that the number of foci is more important than the foci size in stratified diffusion for influencing the rate of expansion in a population (Mack 1985). In such cases it is expected that the finer resolutions will spread more rapidly. However, the results in Chapter 4 indicate that this is not the case and that there is actually an optimal patch size for dispersal in a homogenous (but discrete) environment when we assume that destination patches are the same size as their source. Thus it is important to characterise whether area occupied relates to long-distance dispersal before assuming all foci generate dispersal events at the same rate as this could lead to erroneous conclusions.

As mentioned earlier in Chapter 4, Mollison (1987) described that, at broad scales of biological invasion, the increase in distribution range can be partitioned into three periods: one of early pioneering, a middle period with fastest change, and a later period of condensing saturation or fill in (Shigesada et al. 1995). Ewel (1986) suggested that the lag at the beginning could be caused by a lack of a suitable habitat available to the invading species or a low probability of new propagules striking the right combination of environmental conditions. The results in Chapter 4 however indicate that this effect can be caused simply by a lack of dispersal event sources. An explanation based on environmental heterogeneity is not necessary.

The analysis of the local module and the kernel module only simulated presence-absence population distributions. One reason for presence/absence simulation is simplicity, as population models would require the inclusion of a growth module, and increase the number of variables. Additionally Steinberg & Kareiva (1997) claim that presence-absence models are the only suitable method for modelling large scale spatio-temporal populations due to the difficulty dealing with local dynamics in a reliable way. However, Clark et al. (1998) and Neubert & Caswell (2000), among others, demonstrate the sensitivity of spread rate to reproductive rate. These studies also indicate that the degree of this sensitivity to growth rate increased the fatter the tail of the dispersal kernel was and fat tailed distributions are frequently used to model long distance dispersal. It would therefore be useful to in future carry out similar analyses as done in Chapter 4 with growth dynamics included.

6.3 Argentine ant simulation

As a case study, the model **MDiG** was applied to recreate the invasion history of Argentine ant in New Zealand (Chapter 5). This case study demonstrated how multiple replications of a presence-absence simulation can be averaged to create occupancy envelopes. Such envelopes can give an indication of how likely it is to find a population of the invasive species within a patch. While only 40-70% of the observed or detected populations were covered by the occupancy envelopes initially (at a 5% threshold using the estimated parameter values) when compared to the equivalent area covered by the uniform spread model, the simulation model was more successful identifying observed populations, particularly later in the simulation.

The occupancy envelopes were not able to predict well early on in an invasion. This was because of a lack of agreement between replicates at the beginning of the simulation, as few patches had a significant probability of matching an observed population. As time progesses, there is greater overlap between replicates, reproducing how invasion spread is more predictable as time progresses. The extreme cases are an incursion that has yet to establish which has low predictability and at the other extreme is an indefinite time in the future where the entire potential distribution of the species becomes completely occupied. From a theoretical point of view, fewer random events occur earlier in an invasion, either an establishment happens or it doesn't from a limited number of propagules. Once a population has established, and has started spreading and growing in numbers, many more propagules and possible establishment events occur. The larger number of random events permits averaging of establishment and spread behaviour in a similar way to how the random movements of particles in a diffusive system eventually give rise to a Gaussian distribution of dispersal distances.

In this study it is clear that **MDiG** is more useful for longer term predictions, and either a continuous uniform spread model or and individual or behaviour based model would be more effective immediately after an invasion is detected. An example is the IBM developed by Guichard et al. (2006). Despite its performance, **MDiG** is a research tool that provides an advanced dispersal simulation framework that is generic and extensible. It could even support a higher resolution individual based model with suitable modules governing individual behaviour and local interaction.

6.4 Data and Parameterisation

Mooij & DeAngelis (2003) describe 3 common questions that arise when estimating parameters:

1. How do we estimate the parameters of interest?
2. What is the degree of confidence in these parameters?
3. If the confidence limits are deemed to be too wide, can we specify what the minimum size of a data set should be, given a desired level of confidence?

The first two questions were addressed in Chapter 5 but the third is difficult, as distribution data used to calculate dispersal parameters is noisy and biased by the suitability of the landscape rather than being a true representation of the dispersal events that occur. In realtiy only the result of successful dispersal events are usually observed. Furthermore, extremely rare events can have a disproportionate effect on the population distribution making it difficult to elucidate the contributions individual dispersal events make (Clark et al. 2001).

While distribution data is incomplete and noisy, it is also more often than not presence-only. This means that sites where the species does not occur are not often recorded. Even if absences are recorded there is a possibility that the species is present but it has been missed due to poor sampling or bad luck. While this kind of error may not unduely influence the statistics or probability distributions representing the distance of dispersal events, it is a major concern for potential distributions and any type of ecological niche model used to create the suitability map required by the *survival* module thus impacting simulated spread.

The problem of presence-only data and it's impact on potential distribution models or ecological niche models has been well studied. Profile techniques focus purely on presence locations, finding commonalities between them. For example environmental envelopes (BIOCLIM uses hyperbox classification, Busby, (1991); HABITAT uses convex hulls, Walker & Cocks, (1991); support vector machines use hyperplane optimisation, Cristianini & Shawe-Taylor, (2000)) and multivariate associations (DOMAIN - utilisation distributions or adaptive kernel methods that describe the intensity of use by animals, Carpenter et al., (1993)). Alternatively some models use pseudo-absences. These methods assume the species is absent in all places the species is not found or takes a random sample of these points. They include regression tree-based techniques (e.g. GARP - Genetic Algorithm for Rule-set Production, Stockwell (1999)), generalised linear models, or generalised additive models.

An alternative to using presence-only and potentially noisy distribution data is to use process-based models that model the development of a species based on time series weather data such as daily minimum and maximum temperatures, and rainfall. For example Pitt et al. (2007) describes the use of a Gypsy moth phenology model to predict what regions of New Zealand had temperatures allowing a stable development cycle from year to year (also see Appendix A). However, such detailed models are only possible after careful study of the biology of the species.

Is important to note, as Oreskes et al. (1994) remind us, that our "access to reality is always incomplete" and thus modellers have often had to, and will continue to, develop models with less than perfect data.

6.5 Future Research

The dispersal model file format is well-defined, and since **MDiG** is modular, open-source, and integrates with the open-source GIS GRASS, it is ideally suited to be modified or improved by other researchers. Modification such as linking human-mediated dispersal to transport networks, wind dispersal of airborne lifestages, local population growth, or cellular-automata based rules can be easily implemented and then shared with other researchers. Of importance is the fact that the modular nature of **MDiG** forces consideration of the various behaviours behind the patterns observed in species spread. Elucidating the underlying mechanisms driving these patterns, makes models less arbitary and links them to explicit spatial scales (Grimm et al. 1996).

6.5.1 Argentine ant model

In the process of constructing the simulation model for *L. humile* spread in NZ, several ways in which the model may be improved became clear. The *L. humile* simulation only recorded presence and absence for the distribution of the species. While it may not be feasible to simulate the population numbers of individual ants, the number of colonies could be used instead. If data on the rate at which budding occurs were available, this rate of increase could be modelled. Despite Hee et al. (2000) reporting colony growth rates, no information could be found on the rate of budding to allow this modification to be used for this study.

MDiG was designed to be modular and extensible, allowing new modes of dispersal to be easily incorporated and compared with other configurations. For further research and for modelling *L. humile* spread there are other dispersal modes that could be included. First, one observed mode of dispersal exhibited by *L. humile* is local spread by rafting downstream along waterways (Barber 1916). Second, patterns relating long distance dispersal with traffic volume along transport networks may exist. Establishment patterns of *L. humile* that follow human transport networks have been reported in the past (Barber 1916, Holway 1995, Human et al. 1998). A more mechanistically realistic model that simulates jump-dispersal along transport networks is clearly possible but was beyond the scope of this thesis.

While there are a variety of methods for generating potential distribution maps such as CLIMEX and GARP (Chapter 3) an interesting study would be to investigate the impact on the performance of the dispersal model of using different methods for generating potential distributions.

One extension of the method used for generating the potential distribution in the *L. humile* model would be to use a specific function to represent how survival might be influenced by temperature, rather than categorising degree-days into high, medium and low suitability. This would remove the sudden changes in suitability between regions that have a degree-day value on either side of a suitability boundary. Such a function would also increase the amount of heterogeneity in the landscape. Determining the appropriate function to represent this relationship would require experimentation.

The conversion from suitability categories to probability of survival is done using an estimate. The probability of surviving in area that is on the edge of climatic suitability is dependent on yearly variations, and the climatic variability within that area. Further research to identify the frequency of micro-climates and their temperature variation compared with the average climate of a region could be used to refine the survival probability.

The land-cover suitability map simply categorises land-cover into different classes. Previous studies have shown that *L. humile* exhibits edge effects at forests (Suarez et al. 1998, Ward & Harris 2005) and by using buffer creation techniques in GIS, the edges of forests could be made relatively more suitable than the inner forest area.

The suitability of region does not remain constant from year to year, and *L. humile* populations exhibit considerable seasonal variations in the density and pattern of their distribution. In favourable conditions, new nests containing queens and workers bud off and the population expands, but in adverse climatic conditions, small nests will merge to form fewer larger colonies (Davis & Van Schagen 1993). The possibility of using stochastic potential distributions, that change in value through time (Moilanen et al. 1998, Moilanen & Cabeza 2002), may mean spread models could allow for this variation.

6.5.2 Population models

Demography is clearly important for determining the rate of invasion spread where population growth rates are highly correlated with invasion speed (Neubert & Caswell 2000, Veit & Lewis

1996). By adding population dynamics to raster cells of a spread simulation, not only can spread of an invasive species be investigated but other types of problems can be addressed. For example, spatially-explicit population models are important tools for conservation studies (Vos et al. 2001) that can help predict the dynamics of endangered populations (Turner et al. 1995) and assist in the design and development of dispersal corridors in wildlife reserves (Collingham & Huntley 2000).

6.5.3 Lifestage dependent dispersal

The different life stages of a species can have unique dispersal dynamics, and not all regions will have a suitable environment for the development of every lifestage. For such species, **MDiG** allows distinct dispersal modes to be assigned each lifestage. Phenology maps may then be used to indicate the timing of recruitment and life stage associated dispersal.

In the past, the shape of the dispersal distribution has been assumed to be more important than demographic parameters for influencing the process of invasion (Van den Bosch et al. 1992), but models that lack stage structured dispersal have been shown to potentially over-estimate invasion speed when not all life stages disperse (Neubert & Caswell 2000).

To create such a stage structured model within **MDiG** process driven models could be used to find the timing of lifestage occurrence. An example is a model to predict the phenology of the Gypsy moth (Régnière & Sharov 1999). Here, maps of life stage timing in New Zealand (Appendix A, Pitt et al. 2007) are used to time the occurence of peak adult moth flight, where individuals are actively dispersing, and the timing of larvae hatch, where individuals use silk threads as a passive dispersal mechanism to be carried by the wind.

6.5.4 Surveillance and Treatment optimisation

Further research utilising a spatially explicit spread model can help design sampling programmes for monitoring and detection of invasive species by investigating different spatial and temporal patterns of sampling in the landscape. The most precise or efficient sampling method could be determined for a species based on the dispersal behaviours that govern its spread. For example, the use of occupancy envelopes that are generated by **MDiG** could be used as an index for the design of a stratified sampling regime.

Similarly, experiments with different eradication treatments, particularly investigating the pattern in which they are applied, can be carried out. Such experiments are rarely possible in reality. Moody & Mack (1988) point out that for species spreading by stratified dispersal, decisions whether to control small distant foci or the main source population can greatly affect the impact and cost of control. Population based simulations could also be carried out when the density of an exotic species impacts the effectiveness of its control, such as in a study by Taylor & Hastings (2004) when controlling the estuary weed *Spartina alterniflora*.

Spread models can also be used to set the size of buffer regions to be monitored or treated with respect to the containment of an invasive species. The Gypsy Moth Slow The Spread (STS) programme used trap catches and stratified dispersal models to delineate a “transition zone”. This zone was further divided into an action and monitoring area. The action area had a finer sampling grid than the monitoring area and was closer to the infested region. Any significant trap catches would initiate an attempt to suppress populations in that local area.

6.5.5 Cost estimation

Leuschner et al. (1996) provided one of the first economic analyses of a invasive species that also incorporated a model of the species’ dispersal. By associating a cost to each patch invaded, the total economic cost of a dispersing population could be estimated. This cost could be based on a combination of factors such as the impact on biodiversity, ecological disturbance, agricultural loss, and social impacts (Pearce & Moran 1994). The cost of treatment could then be weighted against the cost of letting the species spread. Of course, the mapping of economic figures to losses of biodiversity and disruption in ecology can be controversial. However, the management of invasive species is restricted by the economic resource available and some estimate of the value of ecological damage is useful for environmental managers.

6.5.6 Multi-scale modelling

The difficulty of modelling spread on a large scale, and using separate modules to represent different dispersal dynamics, is that multiple scales are forced to run at the same resolution. This results in inefficient algorithms and potentially can impact on model results (Chapter 4).

Cole & Albrecht (1999) used a data structure known as a quad-tree (Samet 1990) for a multiscale CA model in GIS. Long distance dispersal was simulated on large-scale coarse resolution grid, with fine grids within each coarse grid cell to model more complex interactions where necessary. This technique could be used to speed up the simulations, and allow more computationally intensive interactions to be simulated.

6.5.7 Numerical Integration

> "It would appear that we have reached the limits of what it is possible to achieve with computer technology, although one should be careful with such statements, as they tend to sound pretty silly in 5 years."
>
> – Von Neumann, 1903-1957

It is foreseable that one day it may be possible to develop models such that they calculate the contribution of probability distributions to every map cell over time. This would be done by numerically integrating continuous functions such as the dispersal kernel over the area of each patch and recording the chance of establishment. Instead of simulating presence-absence or population numbers and then averaging the results across replications, the model would directly calculate the probability of a insect establishing in an area.

Such a technique is currently intractable at large scales because every patch influences every other one, leading to an algorithm of $O(n^n)$ complexity, where n is proportional to the number of patches in the simulation. Essentially this means that every patch potentially interacts with every other which restricts the region size that can be modelled.

6.5.8 Use of remote imaging to gather data

NASA has been involved in the development of United States wide invasive-species forecasting. Remote sensing technologies are being employed to bolster the accuracy of modelling invasive species movement (Schnase et al. 2002). This technology relies on the fact that certain species have a damaging or widespread enough impact on the landscape that the effects can be observed from imaging satellites. Using this imaging technology along with image processing techniques allows large scale data collection on insect distributions that could potentially be automated. John Schnase, an Earth and Space Data Computing Division senior scientist and principal investigator for the invasive species monitoring programme, confirms the role that predictive dispersal methods could play in the prevention and control of invasive species when he said:

> "When we have effective models and can generate forecasts of where non-native species may move, we can set priorities for prevention and control before the species do damage."[1]:

A similar technique is the use of ground based sensor technologies for detecting insect populations, either through detecting pheromones and other olfactory cues (Settles 2005), or through

---

[1] http://esdcd-news.gsfc.nasa.gov/2003.Winter/02.invasive-species.html

ambient sound profiles (Porter et al. 2005).

6.5.9 Conservation

Habitat loss and fragmentation is generally considered to be the greatest threat worldwide to the survival of species (May & Stumpf 2000, Pimm & Raven 2001). The use of heterogenuous landscape models for modelling threatened species populations and dispersal can assist in managing species, and the design of wildlife reserves (Collingham & Huntley 2000). Metapopulation models can work well for the management of discrete and distinct habitat patches (Hanski 1999), but where habitat suitability is more varied and less distinct across the landscape a more spatially explicit model like **MDiG** would be necessary.

Invasive species establishment and spread are special cases of incorporating a novel species into a new area. Any model developed to predict the spread of invasive species may also be applicable to modelling the spread of genetically modified organisms and potentially the spread of new genetic material. Modelling the spread of an advantageous allele was also the first application of diffusion models (Fisher 1937). Such models may also be applied to predicting the spread of an intentionally introduced biological control agent.

There is clear dichotomy between conservation biology and invasion biology. While conservation biology uses endangered species models to make management decisions to enhance their continued survival, invasive species models are mainly used used to assess methods of eradication. Despite this both disciplines are concerned with monitoring and influencing the state of a population that is potentially far from equilibrium.

6.5.10 Integrated web based interface

Recently much focus has been directed towards on integrated systems that can assist pest risk assessments. One example is NAPPFAST which provides a web based front-end for weather-based mapping of plant pathogens (Magarey et al. 2007).

The Beating Weeds FRST (Foundation for Research, Science, and Technology) project, which has the aim both to predict the spread of weed species in New Zealand and allow different control strategies to be evaluated, is considering an integrated web interface to provide an intuitive interface for managers and researchers to use, utilising a modified version of **MDiG** as the simulation engine (G. Bourdôt pers. com. 2008).

6.6 Conclusion

In conclusion, stochastic spatially-explicit dispersal models integrated with GIS allow predictions to be made about the direction and rate that an exotic population may spread. These predictions vary in accuracy but provide more informative results than uniform spread that does not take into account the underlying heterogeneity of the landscape. Such predictions can assist not only invasion biology, but also conservation biology so that the spread of endangered species can be projected and assist in planning monitoing regimes. Dispersal models can also assist climate change studies by projecting the movement of species into previously climatically hostile areas as they gradually increase in climatic suitability.

The maps generated by models utilising GIS are also a wonderful communication tools. The projections made by the spread model, as occupancy risk envelopes, can be intuitively understood and the relative risk between various areas can easily conveyed to policy and decision makers.

Two decades ago Moody & Mack (1988) called for invasive biology to place more attention on strategy for managing invasive species:

> With few exceptions, we find insufficient attention given to developing [a more effective] strategy ('the art of projecting and directing the larger miltary movements and operations of a campaign') for controlling the spread of alien plants. Instead, current practice emphasizes tactics ('the art of handling forces in battle or in the immediate presence of the enemy', both definitions from the Compact Edition of the *Oxford English Dictionary*) for controlling invasions locally, i.e. the specific procedures used to remove alien plants by burning, excavation, herbicide, etc.

Since then, many researchers have answered that call with analytical and simulation models (e.g. Eiswerth & Johnson 2004, Taylor & Hastings 2004). This thesis has attempted and hopefully provided another substantial step towards the techniques and knowledge required for more effective strategy in the management of invasive species. Additionally it provides a solid foundation for continued research into dispersal simulation across heterogeneous landscapes by providing an open source dispersal framework that researchers can use for collaboration, due to both the free availability of the software and the inclusion of a well-defined model format.

This thesis has also demonstrated that the field is rife with opportunities and varied directions for future research, and it is the hope of the author that future researchers will assist him in exploring them.

Acknowledgements

I would firstly like to thank Dr. Sue P. Worner whose enthusiasm for my research and supportive words have made the 4 years of focussed study and attention bearable. My parents Frank Pitt and Eleonore Pitt for raising me to have a natural curiousity about the world, who guided me to self teach myself programming and computers from a young age, a skill which, as you might imagine, has been invaluable in this age of information.

I want to also thank my ex-partner Greta J. Simpson, whom shared good times and bad with me for the majority of my PhD study, I hope she has a successful and fulfilling life. My friends, particularly Nicola Day, deserve credit for keeping me sane and providing perspective when the task of writing this thesis seemed overwhelming, and for mutual support and hilarity.

The White Elephant (http://www.we.org.nz/), a youth orientated dance group in Christchurch, and the Wellington Dance Collective (http://wdc.org.nz/), an association of like-minded people, have helped me with my hobby as a DJ and musician, allowing my mind to occasionally exercise other talents and come back refreshed for writing and coding.

I would like to thank the numerous people who have provided me with distribution data. For the Argentine ant, Darren Ward (School of Biological Sciences, University of Auckland) and Richard Harris (Landcare Research) provided precompiled lists of Argentine ant occurences. This was supplemented by Kathy Mardon (Nelson city council) who provided maps of where the ant had been reported in Nelson. Lastly, Andrea Booth from the Northland branch of the Department of Conservation for additional sites in Northland.

References

Addison, P. & Samways, M. (2000). A survey of ants (Hymenoptera: Formicidae) that forage in vineyards in the Western Cape Province, South Africa. *African Entomology* **8**, 251–260.

Allen, T. F. H., King, A. W., Milne, B. T., Johnson, A. & Turner, S. (1993). The problem of scaling in ecology. *Evolutionary Trends in Plants* **7**, 3–8.

Anderson, R. M., Jackson, H. C., May, R. M. & Smith, A. M. (1981). Population dynamics of fox rabies in Europe. *Nature* **289**, 765–771.

Andow, D. A., Kareiva, P. M., Levin, S. A. & Okubo, A. (1990). Spread of invading organisms. *Landscape Ecology* **4**, 177–188.

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

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

Araújo, M. B., Thuiller, W. & Pearson, R. G. (2006). Climate warming and the decline of amphibians and reptiles in Europe. *Journal of Biogeography* **33**, 1712–1728.

Austin, M. P. (1992). Modelling the environmental niche of plants - implications for plant community response to elevated CO_2 levels. *Australian Journal of Botany* **40**, 615–630.

Bacon, P. J. (1985). *Population dynamics of rabies in wildlife.* Academic Press, London.

Baker, H. G. (1986). Patterns of plant invasion in North America. *in* H. A. Mooney & J. A. Drake, eds, 'Ecology of Biological Invasions of North America and Hawaii'. Springer, New York. pp. 44–57.

Balzter, H., Braun, P. W. & Kohler, W. (1998). Cellular automata models for vegetation dynamics. *Ecological Modeling* **107**, 113–125.

Banks, H. T., Kareiva, P. M. & Zia, L. (1988). Analyzing field studies of insect dispersal using two-dimensional transport equations.. *Environmental Entomology* **17**, 815–820.

Barber, T. C. (1916). The Argentine ant: Distribution and control in the United States. *United States Department of Agriculture Bulletin* **377**, 1–23.

Barlow, N. D. & Goldson, S. L. (2002). Alien invertebrates in NZ. *in* D. Pimentel, ed., 'Biological Invasions: Economic and environmental costs of alien plant, animal and microbe species'. CRC Press, New York, USA.. pp. 195–216.

Bascompte, J. & Solé, R. V. (1995). Rethinking complexity: modelling spatiotemporal dynamics in ecology. *Trends in Ecology and Evolution* **10**, 361–366.

Bell, D. M. (2003). Public health interventions and SARS spread. *Emerging Infectious Diseases* **10**, 1900–190.

Berec, L. (2002). Techniques of spatially explicit individual-based models: construction, simulation, and mean-field analysis. *Ecological Modelling* **150**, 55–81.

Bernard, L. & Krüger, T. (2000). Integration of GIS and spatio-temporal simulation models: Interoperable components for difference simulation strategies. *Transactions in GIS* **4**(3), 197–215.

Berry, J. A., Green, O. R. & Shattuck, S. O. (1997). Species of *Pheidole* westwood (Hymenoptera, Formicidae) established in New Zealand. *New Zealand Journal of Zoology* **24**, 25–33.

Berry, P. M., Dawson, T. E., Harrison, P. A. & Pearson, R. G. (2002). Modelling potential impacts of climate change on the bioclimatic envelope of species in Britain and Ireland. *Global Ecology and Biogeography* **11**, 453–462.

Berryman, A. A. & Millstein, J. A. (1989). Are ecological systems chaotic - and if not, why not?. *Trends in Ecology and Evolution* **4**, 26–28.

Beverton, R. J. H. & Holt, S. J. (1957). On the dynamics of exploited fish populations. *Fishery Investigations, Series II* **19**, 1533.

Blossey, B. & Nötzold, R. (1995). Evolution of increased competitive ability in invasive non-indigenous plants: a hypothesis. *Journal of Ecology* **83**, 887–889.

Bolger, D. T., Suarez, A. V., Crooks, K. R., Morrison, S. A. & Case, T. J. (2000). Effects of habitat fragmentation and Argentine ants on arthropods in southern California. *Ecological Applications*.

Bolker, B. M., Pacala, S. W. & Levin, S. A. (2000). Moment methods for ecological processes in continuous space. *in* U. Dieckmann, R. Law & J. A. J. Metz, eds, 'The geometry of ecological interactions'. Cambridge University Press, Cambridge. pp. 388–411.

Bond, W. & Slingsby, P. (1984). Collapse of an ant-plant mutualism: the Argentine ant (*Iridomyrmex humilis*) and myrmecochorous Proteaceae. *Ecology* **65**, 1031–1037.

Bossenbroek, J. M., Kraft, C. E. & Nekola, J. C. (2001). Prediction of long-distance dispersal using gravity models: Zebra mussel invasion of inland lakes. *Ecological Applications* **11**, 1778–1788.

Breckling, B., Müller, F., Reuter, H., Hölker, F. & Fränzle, O. (2005). Emergent properties in individual-based ecological models - introducing case studies in an ecosystem research context. *Ecological Modelling* **186**, 376–388.

Broadbent, S. R. & Hammersley, J. M. (1957). Percolation Processes, I. Crystals and Mazes. *in* 'Proceedings of the Cambridge Philosophical Society'. Vol. 53. pp. 629–641.

Brockmann, D., Hufnagel, L. & Geisel, T. (2006). The scaling laws of human travel. *Nature letters* **439**, 462–465.

Brown, J. H. & Kodric-Brown, A. (1977). Turnover rates in insular biogeography: effects of immigration on extinction. *Ecology* **58**, 445449.

Brown, L. A. (1981). *Innovation Diffusion*. Methuan, London.

Brownlee, J. (1911). *The Mathematical Theory of Random Migration and Epidemic Distribution*. Royal Society of Edinburgh.

Buchan, L. A. J. & Padilla, D. K. (1999). Estimating the probability of long-distance dispersal overland dispersal of invading aquatic species. *Ecological Applications* **9**, 254–265.

Buczkowski, G. & Silverman, J. (2006). Geographical variation in Argentine ant aggression behaviour mediated by environmentally derived nestmate recognition cues. *Animal Behaviour* **71**, 327–335.

Bullock, J. M. & Clarke, R. T. (2000). Long distance seed dispersal by wind: measuring and modelling the tail of the curve. *Oecologia* **124**, 506–521.

Busby, J. R. (1991). BIOCLIM - a bioclimatic analysis and prediction system. *in* 'Nature conservation: Cost effective biological surveys and data analysis'. CSIRO, Melbourne.

Buttel, L., Durrett, R. & Levin, S. (2002). Competition and species packing in patchy environments. *Theoretical Population Biology* **61**, 265–276.

Buys, B. (1987). Competition for nectar between Argentine ants (*Iridomyrmex humilis*) and honeybees (*Apis mellifera*) on black ironbark (*Eucalyptus sideroxylon*). *South African Journal of Zoology* **22**, 173–174.

Byers, J. E. & Noonburg, E. G. (2003). Scale dependent effects of biotic resistance to biological invasion. *Ecology* **84**, 1428–1433.

Cammell, M. E., Way, M. J. & Paiva, M. R. (1996). Diversity and structure of ant communities associated with oak, pine, eucalyptus and arable habitats in Portugal. *Insectes Sociaux* **43**, 37–46.

Cannas, A. S., Marco, D. E. & Paez, S. A. (2003). Modelling biological invasions: species traits, species interactions, and habitat heterogeneity. *Math Biosci* **183**, 93–110.

Carlquist, S. (1965). *Island Life*. Natural History Press, Garden City NY.

Carpenter, G., Gillison, A. N. & Winter, J. (1993). DOMAIN: a flexible modelling procedure for mapping potential distributions of plants and animals. *Biodiviersity and Conversation* **2**, 667–680.

Case, T. J., Holt, R. D., McPeek, M. A. & Keitt, T. H. (2005). The community context of species borders: ecological and evolutionary perspectives. *Oikos* **108**, 28–46.

Caspi, A., Granek, R. & Elbaum, M. (2002). Diffusion and directed motion in cellular transport. *Physical Review E* **66**, 011916.

Caswell, H. & Cohen, J. E. (1991). Disturbance, interspecific interaction and diversity in metapopulations. *Biol J Linn Soc* **42**, 193–218.

Caswell, H. & Etter, R. J. (1992). Ecological interactions in patchy environments: from patch-occupancy models to cellular automata. *in* S. A. Levin, T. M. Powell & J. H. Steele, eds, 'Patch dynamics'. Springer, Berlin Heidelberg New York. pp. 93–109.

Chave, J., Muller-Landau, H. C. & Levin, S. A. (2002). Comparing classical community models: theoretical consequences for patterns of diversity. *American Naturalist* **159**, 1–23.

Chen, J. & Nonacs, P. (2000). Nestmate recognition and intraspecific aggression based on environmental cues in Argentine ants (Hymenoptera: Formicidae). *Annals of the Entomological Society of America* **93**, 1333–1337.

Clark, J. S. (1998). Why trees migrate so fast: confronting theory with dispersal biology and the paleo record.. *American Naturalist* **152**, 204–224.

Clark, J. S., Fastie, C., Hurtt, G., Jackson, S. T., Johnson, C., King, G. & Lewis, M. (1998). Reid's paradox of rapid plant migration. *BioScience* **48**, 13–24.

Clark, J. S., Lewis, M. & Horvath, L. (2001). Invasion by extremes: Population spread with variation in dispersal and reproduction. *American Naturalist* **157**(5), 537–554.

Clark, J. S., Macklin, E. & Wood, L. (1998). Stages and spatial scales of recruitment limitation in southern Appalachian forests. *Ecological Monographs* **68**, 213–235.

Clark, J. S., Silman, M., Kern, R., Macklin, E. & HilleRisLambers, J. (1999). Seed dispersal near and far: generalized patterns across temperate and tropical forests. *Ecology* **80**, 1475–1494.

Clarke, C. M. H. (1971). Liberations and dispersal of red deer in northern South Island districts. *N.Z. Journal of Forestry Science* **1**, 194–207.

Cliff, A. D., Haggett, P., Ord, J. D. & Versey, G. R. (1981). *Spatial Diffusion*. Cambridge University Press.

Colasanti, R. & Grime, J. (1993). Resource dynamics and vegetation processes: a deterministic model using two-dimensional cellular automata.. *Functional Ecology* **7**, 169–176.

Colautti, R. I., Ricciardi, A., Grigorovich, I. A. & Macisaac, H. J. (2004). Is invasion success explained by the enemy release hypothesis?. *Ecology Letters* **7**, 721733.

Cole, F. R., Medeiros, A. C., Loope, L. L. & Zuehlke, W. W. (1992). Effects of the Argentine ant on arthropod fauna of Hawaiian high-elevation shrubland. *Ecology* **73**, 1313–1322.

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.

Collingham, Y. C. & Huntley, B. (2000). Impacts of habitat fragmentation and patch size upon migration rates. *Ecological Applications* **10**, 131144.

Cowan, J., Paoli, J., Yergeau, F., Maler, E., Sperberg-McQueen, C. M. & Bray, T., eds (2006). *Extensible Markup Language (XML) 1.1*. second edn. World Wide Web Consortium.

Cristianini, N. & Shawe-Taylor, J. (2000). *An introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge University Press.

Crooks, J. & Soulé, M. E. (1996). Proceedings of the Norway UN conference on alien species. *in* O. Sandlund, P. Schei & A. Viken, eds, 'Lag times in population explosions of invasive species: causes and implications'. Directorate for Nature Management/Norwegian Institute for Nature Research.. Trondheim, Norway. pp. 39–46.

Daehler, C. C. (1998). The taxonomic distribution of invasive angiosperm plants: ecological insights and comparison to agricultural weeds. *Biological Conservation* **84**, 167–180.

Davis, M. A. & Pelsor, M. (2001). Experimental support for a resource-based mechanistic model of invasibility. *Ecology Letters* **4**, 421–428.

Davis, M., Wrage, K. & Reich, P. (1998). Competition between tree seedlings and herbaceous vegetation: support for a theory of resource supply and demand. *Journal of Ecology* **86**, 652–661.

Davis, P. & Van Schagen, J. (1993). Effective control of pest ants. *Journal of Agriculture of Western Australia* **34**, 92–95.

De Kock, A. E. (1990). Interactions between the introduced Argentine ant, *Iridomyrmex humilis* mayr, and two indigenous fynbos ant species. *Journal of the Entomological Society of South Africa* **53**, 107–111.

Dennis, B. (2002). Allee effects in stochastic populations. *Oikos* **96**, 389–401.

Diffendorfer, J. E., Gaines, M. S. & Holt, R. D. (1999). Patterns and impacts of movements and different scales in small mammals. *in* G. W. Barrett & J. D. Peles, eds, 'Ecology of small mammals at the landscape level: experimental approaches'. Springer. pp. 63–88.

Doak, D. F. & Mills, L. S. (1994). A useful role for theory in conservation. *Ecology* **75**, 615–626.

Dobzhansky, T. H. & Wright, S. (1943). Genetics of natural populations. X. Dispersion rates in *Drosophila pseudoobscura*. *Genetics* **28**, 304–340.

Drake, J. M. & Lodge, D. M. (2006). Allee effects, propagule pressure and the probability of establishment: risk analysis for biological invasions. *Biological Invasions* **8**, 365–375.

Duffey, E. (1998). Aerial dispersal in spiders. *in* P. A. Seldon, ed., 'Proceedings of the 17th European Colloquium of Arachnology'. Buckinghamshire: British Arachnological Society. pp. 187–191.

Dunning, J. B. J., Stewart, D. J., Danielson, B. J., Noon, B. R., Root, T. L., Lamberson, R. H. & Stevens, E. E. (1995). Spatially explicit population models: Current forms and future uses. *Ecological Applications* **5**(1), 3–11.

Durrett, R. & Levin, S. (1994*a*). The importance of being discrete (and spatial). *Theoretical Population Biology* **46**, 363–394.

Durrett, R. & Levin, S. (1994*b*). Stochastic spatial models: a user's guide to ecological applications. *Philosophical Transactions of the Royal Society of London B* **343**, 329–350.

Durrett, R. & Levin, S. (1998). Spatial aspects of interspecific competition. *Theoretical Population Biology* **53**, 30–43.

Eckhardt, R. (1987). Stan Ulam, John von Neumann, and the Monte Carlo method. *Los Alamos Science* pp. 131–143.

Eden, M. (1961). A two dimensional growth process. *in* J. Neyman, ed., 'Proceedings of Fourth Berkeley Symposium on Math, Statistics and Probability'. University of California Press, Berkeley, CA. pp. 223–239.

Eiswerth, M. E. & Johnson, W. S. (2004). Managing nonindigenous invasive species: Insights from dynamic analysis. *Environmental and Resource Economics* **23**, 319–342.

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

Ellner, S. P., Sasaki, A., Haraguchi, Y. & Matsuda, H. (1998). Speed of invasion in lattice population models: pair-edge approximation. *Mathematical Biology* **36**, 469–484.

Elton, C. S. (1958). *The Ecology of Invasions by Animals and Plants*. Methuen, London.

Eplee, R. E. (1981). *Striga's* status as a plant parasite in the united states. *Plant Disease* **65**, 951–954.

Eppstein, M. J. & Molofsky, J. (2007). Invasiveness in plant communities with feedbacks. *Ecology Letters* **10**, 253–263.

Erickson, J. M. (1971). The displacement of native ant species by the introduced Argentine ant, *Iridomyrmex humilis* Mayr. *Psyche* **78**, 257–266.

ESRI (2007). 'ArcUser'. January-March Newsletter.

Etter, R. J. & Caswell, H. (1994). The advantages of dispersal in a patchy environment: effects of disturbance in a cellular automaton model. *in* K. J. Eckelbarger & C. M. Young, eds, 'Reproduction, larval biology and recruitment in the deep-sea benthos'. Columbia University Press, New York. pp. 285–305.

Ewel, J. J. (1986). Invisibility: Lessons from south florida. *in* H. A. Mooney & J. K. Drake, eds, 'Ecology of Biological Invasions of North America and Hawaii'. Springer-Verlag, New York. pp. 214–229.

Fahrig, L. & Merriam, G. (1985). Habitat patch connectivity and population survival. *Ecology* **66**, 1762–1768.

Fedra, K. (1996). Distributed models and embedded GIS: Integration strategies and case studies. *in* M. F. Goodchild, L. T. Steyaert & B. O. Parks, eds, 'GIS and Environmental Modeling: Progress and Research Issues'. Fort Collins, CO, GIS World Books. pp. 413–418.

Fenn, J. A. (1980). Control of *hakea* in the western Cape. *in* S. Neser & A. L. P. Cairns, eds, 'Proceedings of the Third National Weeds Conference of South Africa'. Balkema, Cape Town. pp. 167–173.

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**, 38–49.

Fielding, A. H. & Haworth, P. F. (1995). Testing the generality of bird habitat models. *Conservation Biology* **9**, 1466–1481.

Filipe, J. A. N. & Maule, M. M. (2004). Effects of dispersal mechanisms on spatio-temproal development of epidemics. *Journal of Theoretical Biology* **226**, 125–141.

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

Fluker, S. S. & Beardsley, J. W. (1970). Sympatric associations of three ants: *Iridomyrmex humilis*, *Pheidole megacephala*, and *Anoplolepis longipes* in Hawaii. *Annals of the Entomological Society of America* **63**, 1290–1296.

Folgarait, P. J. (1998). Ant biodiversity and its relationship to ecosystem functioning: a review. *Biodiversity and Conservation* **7**, 1221–1244.

Forest & Bird, S. (n.d.). 'Argentine ant fact sheet'.

Fowler, H. G., Alves, L. E. & Bueno, O. C. (1993). Reproductive strategies of the exotic Pharaohs ant *Monomorium pharaonis* (L.) (Hymenoptera: Formicidae) in Brazil. *Invertebrate Reproduction and Development* **23**, 235–238.

Gage, S. H., Colunga-Garcia, M., Helly, J. J., Safir, G. R. & Momin, A. (2001). Structural design for management and visualization of information for simulation models applied to a regional scale. *Computers and Electronics in Agriculture* **33**, 77–84.

Gage, S. H., Wirth, T. M. & Simmons, G. A. (1990). Predicting regional gypsy moth (Lymantriidae) population trends in an expanding population using pheromone trap catch and spatial analysis. *Environmental Entomology* **19**(2), 370–377.

Gardner, R. H. & Gustafson, E. J. (2004). Simulating dispersal of reintroduced species within heterogreneous landscapes. *Ecological Modelling* **171**, 339–358.

Gardner, R. H., Milne, B., O'Neill, R. & Turner, M. (1987). Neutral models for the analysis of broad-scale landscape patterns. *Landscape Ecology* **1**, 19–28.

Gatehouse, A. G. (1997). Behaviour and ecological genetics of wind-borne migration by insects. *Annual Review of Entomology* **42**, 475–502.

Gautestad, A. O. & Mysterud, I. (1993). Physical and biological mechanisms in animal movement processes. *The Journal of Applied Ecology* **30**, 523–535.

Gay, J., ed. (2002). *Free Software, Free Society: Selected Essays of Richard M. Stallman.* Free Software Foundation.

Gevrey, M. & Worner, S. P. (2006). Prediction of global distribution of insect pest species in relation to climate by using an ecological informatics method. *Journal of Economic Entomology* **99**, 979–986.

Gibbons, J. D. (1985). *Nonparametric Statistical Inference.* 2nd edn. M.Dekker.

Giliomee, J. H. (1986). Seed dispersal by ants in the Cape flora threatened by *Iridomyrmex humilis* (hymenoptera: Formicidae). *Entomologia Generalis* **11**, 217–219.

Goel, N. S. & Richter-Dyn, N. (1974). *Stochastic Models in Biology.* Academic Press, New York.

Goldson, S. L., Frampton, E. R., Geddes, N. J. & Braggins, T. J. (2002). The potential of sensor technologies to improve New Zealand's border biosecurity. *in* S. L. Goldson & D. M. Suckling, eds, 'Defending the Green Oasis: New Zealand Biosecurity and Science'. New Zealand Plant Protection Society Inc. Christchurch, New Zealand. pp. 63–72.

Gonzalez-Hernandez, H., Johnson, M. W. & Reimer, N. J. (1999). Impact of *Pheidole megacephala* (F.) (Hymenoptera: Formicidae) on the biological control of *Dysmicoccus brevipes* (Cockerell) (Homoptera: Pseudococcidae). *Biological Control* **15**, 145–152.

Goodchild, M. F. (1993). The state of GIS for environmental problem-solving. *in* M. F. Goodchild, B. O. Parks & L. T. Steyart, eds, 'Environmental Modelling with GIS'. Oxford University Press, New York. pp. 8–15.

Goodwin, B. J., Jones, C. G., Schauber, E. M. & Ostfeld, R. S. (2005). Limited dispersal and heterogeneous predation risk synergistically enhance persistence of rare prey. *Ecology* **12**, 3139–3148.

GRASS Development Team (2006). *Geographic Resources Analysis Support System (GRASS GIS) Software*. ITC-irst. Trento, Italy.

Green, O. R. (1990). Entomologist sets new record at Mt Smart for *Iridomyrmex humilis* established in New Zealand. *The Weta* **13**, 14–16.

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

Grimm, V. (1994). Mathematical models and understanding in ecology. *Ecological Modelling* **75**, 641–651.

Grimm, V., Frank, K., Jeltsch, F., Brandl, R., Uchmanński, J. & Wissel, C. (1996). Pattern-orientated modelling in population ecology. *The Science of the Total Environment* **183**, 151–166.

Grochowski, E. & Halem, R. D. (2003). Technological impact of magnetic hard disk drives on storage systems. *Storage Systems* **42**, 338–346.

Grochowski, E. & Hoyt, R. F. (1996). Future trends in hard disk drives. *IEEE Transections on Magnetics* **32**, 1850–1854.

Groom, M. J. (1998). Allee effects limit population viability of an annual plant. *The American Naturalist* **151**(6), 487–496.

Group, B. M. (2005). Pathway Analysis, Annual Report 2004-2005, BMG 05-06/02. Technical report. Biosecurity, New Zealand.

Guichard, S., Kriticos, D., Worner, S. & Kean, J. (2006). Model of pheromone oriented flight based on moth behaviour and wind fields in a biosecurity context. *in* ‘Australia and New Zealand Entomology Societies Conference, 24-27 September, Adelaide, Australia’.

Guisan, A., Theurillat, J.-P. & Kienast, F. (1998). Predicting the potential distribution of plant species in an alpine environment. *Journal of Vegetation Science* **9**, 65–74.

Gustafson, E. J. & Gardner, R. H. (1996). The effect of landscape heterogeneity on the probability of patch colonization. *Ecology* **77**, 94107.

Gyllenberg, M. & Hanski, I. (1992). Single-species metapopulation dynamics: a structured model. *Theoretical Population Biology* **72**, 35–61.

Haccou, P., Jagers, P. & Vatutin, V. A. (2005). *Branching Processes: Variation, Growth and Extinction of Populations*. Cambridge University Press.

Hambck, P. A. & Englund, G. (2005). Patch area, population density and the scaling of migration rates: the resource concentration hypothesis revisted. *Ecology Letters* **8**, 1057–1065.

Hanley, J. A. & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **143**, 29–36.

Hanski, I. (1998). Connecting the parameters of local extinction and metapopulation dynamics. *Oikos* **83**, 390–396.

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

Hanski, I. & Kuusaari, M. (1995). Butterfly metapopulation dynamics. *in* N. Cappucino & P. Wright, eds, ‘Population Dynamics: New Approaches and Synthesis’. Academic Press, San Diego, CA.

Hanski, I., Alho, J. & Moilanen, A. (2000). Estimating the parameters of survival and migration of individuals in metapopulations. *Ecology* **81**, 239–251.

Hanski, I., Pöyry, J., Kuussaaru, M. & Pakkala, T. (1995). Multiple equilibria in metapopulation dynamics. *Nature* **377**, 618–621.

Hanson, F. B. (1987). Bioeconomic model of the Lake Michigan alewife fishery. *Canadian Journal of Fisheries and Aquatic Sciences* **4**(2), 298–305.

Hanson, F. B. & Tuckwell, H. C. (1997). Population growth with randomly distributed jumps. *Journal of Mathematical Biology* **36**, 169–187.

Harada, Y. & Iwasa, Y. (1994). Lattice population dynamics for plants with dispersing seeds and vegetative propagation. *Researches on Population Ecology* **36**, 237–249.

Harada, Y., Ezoe, H., Iwasa, Y., Matsuda, H. & Sato, K. (1995). Population persistence and spatially limited social interaction. *Theoretical Population Biology* **48**, 65–91.

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.

Harris, R. J. & Berry, J. A. (2002). Confirmation of the establishment of three adventive ants (Hymenoptera: Formicidae) in New Zealand: *Cardiocondyla minutior* Forel, *Ponera leae* Forel, *Mayriella abstinens* Forel. *New Zealand Entomologist* **24**, 53–56.

Harrison, P. A., Berry, P. M., Butt, N. & New, M. (2006). Modelling climate change impacts on species' distributions at the European scale: implications for conservation policy. *Environmental Science and Policy* **9**, 116–128.

Harrison, R. G. (1980). Dispersal polymorphisms in insects. *Annual Revue of Ecological Systems* **11**, 95–118.

Hartley, S. & Lester, P. J. (2003). Temperature-dependent development of the Argentine ant, *Linepithema humile* (Mayr) (Hymenoptera: Formicidae): a degree-day model with implications for range limits in New Zealand.. *New Zealand Entomologist* **26**, 91–100.

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

Haskins, C. & Haskins, E. (1988). Final observations on *Pheidole megacephala* and *Iridomyrmex humilis* in Bermuda. *Psyche Cambridge, Mass* **95**, 177–184.

Hastings, A. (1996*a*). Models of spatial spread: A synthesis. *Biological Conservation* **78**, 143–148.

Hastings, A. (1996*b*). Models of spatial spread: Is the theory complete?. *Ecology* **77**, 1675–1679.

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

Haswell, M. P., Comins, H. N. & May, R. M. (1991). Spatial structure and chaos in insect population dynamics. *Nature* **353**, 255–258.

Hayes, K. (1997). A review of ecological risk assessment methodologies. Technical Report 13. CSIRO Marine Research Technical Report. Collingwood, Victoria, Australia.

Hedges, S. A. (1998). *Field guide for the management of structure-invading ants*. Franzak and Foster, Cleveland, OH.

Hee, J. J., Holway, D. A., Suarez, A. V. & Case, T. J. (2000). Role of propagule size in the success of incipient colonies of the invasive argentine ant. *Conservation Biology* **14**(2), 559–563.

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

Hengeveld, R. (1994). Small step invasion research. *Trends ecology evolution* **9**, 339–342.

Higgins, S. & Richardson, D. M. (1996). A review of models of alien plant spread.. *Ecological Modelling* **87**, 249–265.

Higgins, S. I. & Richardson, D. M. (1999). Predicting plant migration rates in a changing world: The role of long-distance dispersal. *The American Naturalist* **153**, 464–475.

Hill, J. K., Thomas, C. D. & Lewis, O. T. (1996). Effects of habitat patch size and isolation on dispersal by *Hesperia comma* butterflies: implications for metapopulation structure. *Journal of Animal Ecology* **65**, 725–735.

Hobbs, R. J. (1991). Disturbance a precursor to weed invasion in native vegetation. *Plant Protection Quarterly* **6**, 99–104.

Hobbs, R. J. (2000). Land-use changes and invasion. *in* H. A. Mooney & R. J. Hobbs, eds, 'Invasive species in a changing world'. Island Press, Washinton D. C.. pp. 55–64.

Holldobler, B. & Wilson, E. O. (1990). *The Ants*. Harvard University Press, Cambridge, MA.

Holmes, E. E. (1993). Are diffusion models too simple? A comparison with telegraph models of invasion. *American Naturalist* **142**, 403–419.

Holway, D. (1995). Distribution of the Argentine ant (*Linepithema humile*) in northern California. *Conservation Biology* **9**, 1634–1637.

Holway, D. A. (1998*a*). Effect of Argentine ant invasions on ground-dwelling arthropods in northern California riparian woodlands.. *Oecologia* **116**, 252258.

Holway, D. A. (1998*b*). Factors governing rate of invasion: a natural experiment using Argentine ants. *Oecologia* **115**, 206–212.

Holway, D. A. (1999). Competitive mechanisms underlying the displacement of native ants by the invasive Argentine ant. *Ecology* **80**, 238–251.

Holway, D. A., Lach, L., Suarez, A. V. & Tsutsui, N. D. (2002). The causes and consequences of ant invasions. *Annual Review of Ecology and Systematics* **33**, 181–233.

Holway, D. A., Suarez, A. V. & Case, T. J. (2002). Role of abiotic factors in governing susceptability to invasion: a test with argentine ants. *Ecology* **83**, 1610–1619.

Human, K. G. & Gordon, D. M. (1996). Exploitation and interference competition between the invasive Argentine ant, *Linepithema humile*, and native ant species. *Oecologia* **105**, 405–412.

Human, K. G. & Gordon, D. M. (1997). Effects of Argentine ants on invertebrate biodiversity in Northern California. *Conservation Biology* **11**, 1242–1248.

Human, K., Weiss, S., Weiss, A., Sandler, B. & Gordon, D. (1998). Effects of abiotic factors on the distribution and activity of the invasive Argentine ant (*Hymenoptera, Formicidae*). *Environmental Entomology* **27**, 822–833.

Humphries, S. E., Groves, R. H. & Mitchell, D. S. (1991). *Plant invasions: The incidence of Environmental Weeds in*. Australian National Parks and Wildlife Service, Canberra, Australia. chapter Plant invasions and Australian ecosystems: a status review and management decisions, pp. 1–127.

Huntley, B., Berry, P. M., Cramer, W. & McDonald, A. P. (1995). Modelling present and potential future ranges of some european higher plants using climate response surfaces. *Journal of Biogeography* **22**, 967–1001.

Huston, M. A. (1994). *Biological diversity*. Cambridge University Press, Cambridge, UK.

Huston, M., DeAngelis, D. L. & Post, W. M. (1988). New computer models unify ecological theory. *BioScience* **38**, 682–691.

Hutchinson, G. E. (1957). Concluding remarks. *in* 'Cold Spring Harbor Symposia on Quantitative Biology'. Vol. 22. pp. 145–159.

Ikegami, T. & Kaneko, K. (1992). Evolution of host-parasitoid network through homeochaotic dynamics. *Chaos* **2**, 397–407.

Ipinza-Regla, J., Figueroa, G. & Moreno, I. (1985). *Iridomyrmex humilis* (Formicidae) and its role as a possible vector of microbial contamination in the food industry.. *Folia Entomologica Mexicana* **62**, 111–124.

Itami, R. (1994). Simulation spatial dynamics: cellular automata theory. *Landscape & Urban Planning* **30**, 27–47.

IUCN (1999). Guidelines for the prevention of biodiversity loss due to biological invasion. Technical report. International Union for the Conservation of Nature.

Jaggi, S. & Joshi, A. (2001). Incorporating spatial variation in density enhances the stability of simple population dynamics models. *Journal of Theoretical Biology* **209**, 249–255.

Johnson, C. G. (1957). The distribution of insects in the air and the empirical relation of density to height. *Journal of Animal Ecology* **26**, 479–494.

Johnson, W. C. & III., T. W. (1989). The role of blue jays in the postglacial dispersal of fagaceous trees in eastern north america. *Journal of Biogeography* **16**, 561–571.

Judge, G. G., Hill, R. C., Griffiths, W. E., Lutkepohl, H. & Lee., T.-C. (1988). *Introduction to the Theory and Practice of Econometrics.* Wiley.

Judson, O. P. (1994). The rise of the individual-based modelling in ecology. *TREE* **9**(1), 9–14.

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

Kareiva, P. (1985). Finding and loosing plants by *Phyllotreta*: patch size and surrounding habitat. *Ecology* **66**, 1810–1816.

Kareiva, P. M. (1983). Local movement in herbivorous insects: applying a passive diffusion model to mark-recapture field experiments. *Oecologia* **57**, 322–327.

Kawasaki, K., Takasu, F., Caswell, H. & Shigesada, N. (2006). How does stochasticity in colonization accelerate the speed of invasion in a cellular automaton model?. *Ecological Research* **21**, 334–345.

Kean, J. M., Overton, J., Williams, P. & Buxton, R. (2007). Modelling weed spread in heterogeneous landscapes. *in* 'NZIMA Weeds Workshop'.

Keane, R. M. & Crawley, M. J. (2002). Exotic plant invasions and the enemy release hypothesis. *Trends in Ecology and Evolution* **17**, 164170.

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

Keller, L. & Passera, L. (1990). Fecundity of ant queens in relation to their age and the mode of colony founding. *Insectes Sociaux* **37**, 116–130.

Keller, L., Passera, L. & Suzzoni, J. (1989). Queen execution in the Argentine ant, *Iridomyrmex humilis*. *Physiological Entomology* **14**, 157–163.

Kennedy, T. (1998). Patterns of an invasion by Argentine ants (*Linepithema humile*) in a riparian corridor and its effects on ant diversity. *American Midland Naturalist* **140**, 343–350.

Kerr, W. E. (1967). The history of introduction of african bees in brazil. *South African Bee Journal* **39**, 3–5.

Kierstead, H. & Slobodkin, L. B. (1953). The size of water masses containing plankton blooms. *Journal of Marine Research* **12**, 141–147.

Kolar, C. S. & Lodge, D. M. (2001). Progress in invasion biology: predicting invaders. *Trends in Ecology and Evolution* **16**, 199204.

Kolmogorov, A. N., Petrovskii, J. G. & Piskunov, N. S. (1937). Investigation of the diffusion equations connected with an increasing amount of matter and its . Technical report. Bulletin MGU Sect. A.

Kornas, J. (1990). Plant invasions in central europe: historical and ecological perspectives. *in* F. di Castri & A. J. Hansen, eds, 'Biological Invasions in Europe and the Mediterranean Basin'. Kluwer Academic Publishers, Dordrecht, The Netherlands. pp. 19–36.

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

Kot, M., Medlock, J., Reluga, T. & Walton, D. B. (2004). Stochasticity and invasions.

Kowarik, I. (1995). Time lags in biological invasion with regard to the success and failure of alien species. *in* P. Pysek, K. Prach, M. Rejmánek & M. Wade, eds, 'Plant Invasions, General Aspects and Special Problems'. SPB Academic Publishers. pp. 15–38.

Kriticos, D. J. & Randall, R. P. (2001). A comparison of systems to analyse potential weed distributions. *in* R. H. Groves, F. D. Panetta & J. G. Virtue, eds, 'Weed Risk Assessment'. CSIRO Publishing, Melbourne.

Kubo, T., Iwasa, Y. & Furomoto, N. (1996). Forest spatial dynamics with gap expansion: total gap area and gap size distribution. *Journal of Theoretical Biology* **180**, 229–246.

Kuno, E. (1981). Dispersal and the persistence of populations in unstable habitats: A theoretical note. *Oecologia* **49**, 123–126.

Kuussaari, M., Nieminen, M. & Hanski, I. (1996). An experimental study of migration in the butterfly *Melitaea cinxia*. *Journal of Animal Ecology* **65**, 791–801.

Kuussaari, M., Saccheri, I., Camara, M. & Hanski, I. (1998). Allee effect and population dynamics in the glanville fritillary butterfly. *Oikos* **82**, 384–392.

Lach, L. (2005). Interference and exploitation competition of three nectar-theiving invasive ant species. *Insectus Sociaux* **52**, 257–262.

Lagarias, J. C., Reeds, J. A., Wright, M. H. & Wright, P. E. (1998). Convergence properties of the nelder-mead simplex method in low dimensions. *SIAM Journal of Optimization* **9**(1), 112–147.

Lande, R. & Orzack, S. H. (1988). Extinction dynamics of age-structured populations in a fluctuating environment. *Proceedings of the National Academy of Sciences USA* **85**, 7418–7421.

Lande, R., Engen, S. & Sæther, B. E. (2003). *Stochastic Population Dynamics in Ecology and Conservatoin.* Oxford University Press, Oxford, UK.

Law, A. M. & Kelton, W. D. (1982). *Simulation Modelling and Analysis.* McGraw-Hill Book Company, New York.

Law, R. & Dieckmann, U. (2000). A dynamical system for neighbourhoods in plant communities. *Ecology* **81**, 2137–2148.

Law, R., Murrell, D. J. & Dieckmann, U. (2003). On population growth in space and time: the spatial logistic equation. *Ecology* **84**, 252–262.

Lester, P. J., Waring, C. W., Longson, C. G. & Hartley, S. (2003). Argentine and other ants (Hymenoptera: Formicidae) in New Zealand horticultural ecosystems: distribution, hemipteran hosts, and review. *New Zealand Entomologist* **26**, 79–89.

Leuschner, W., Young, J., Walden, S. & Ravlin, F. (1996). Potential benefits of slowing the gypsy moth's spread. *Southern Journal of Applied Forestry* **20**, 6573.

Levin, S. A. (1992). The problem of pattern and scale in ecology. *Ecology* **73**, 1943–1967.

Levin, S. A. & Durrett, R. (1997). From individuals to epidemics. *Philosophical Transactions of the Royal Society of London B* **351**, 1615–1621.

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

Levins, R. (1966). The strategy of model building in population biology. *American Scientist* **54**, 421–431.

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

Lewis, M. A. (1997). Variability, patchiness, and jump dispersal in the spread of an invading population. *in* D. Tilman & P. Kareiva, eds, 'Spatial ecology: the role of space in population dynamics and interspecific interactions'. Princeton University Press, Princeton, New Jersey. pp. 46–69.

Lewis, M. A. (2000). Spread rate for a nonlinear stochastic invasion. *Journal of Mathematical Biology* **41**, 430–454.

Lewis, M. A. & Kareiva, P. M. (1993). Allee dynamics and the spread of invading organisms. *Theoretical Population Biology* **43**, 141–158.

Lewis, M. A. & Pacala, S. (2000). Modeling and analysis of stochastic invasion processes. *Journal of Mathematical Biology* **41**, 387–429.

Liang, D. & Silverman, J. (2000). "You are what you eat": Diet modifies cuticular hydrocarbons and nestmate recognition in the Argentine ant, *Linepithema humile*. *Naturwissenschaften* **87**, 412–416.

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

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

Liebhold, A. M., Work, T. T., McCullough, D. G. & Cavey, J. F. (2006). Airline baggage as a pathway for alien insect species invading the united states. *American Entomologist* **Spring**, 48–54.

Lima, S. L., Zollner, P. A. & Bednekoff, P. A. (1999). Predation, scramble competition, and the vigilance group size effect in dark-eyed juncos (*Junco hyemalis*). *Behavioral Ecology and Sociobiology* **46**, 110–116.

Lockwood, J. L., Cassey, P. & Blackburn, T. (2005). The role of propagule pressure in explaining species invasions. *Trends in Ecology and Evolution* **20**, 223228.

Lonsdale, W. M. (1999). Global patterns of plant invasions and the concept of invasibility. *Ecology* **80**(5), 1522–1536.

Lowe, S., Browne, M. & Boudjelas, S. (2005). *100 of the world's worst invasive alien species: A selection from the Global Invasive Species Database*. Invasive Species Specialist Group.

MacArthur, R. & Wilson, E. O. (1967). *The Theory of Island Biogeography*. Princeton University Press.

Mack, R. N. (1981). Invasion of *Bromus tectorum* l. into western north america: an ecological chronicle.. *Agro-ecosystems* **7**, 145–165.

Mack, R. N. (1985). Invading plants; their potential contribution to population biology. *in* J. White, ed., 'Studies in Plant Demography: A Festschrift for John L. Harper'. Academic Press, New York. pp. 127–142.

Mack, R. N. (1995). Invading plants: their potential contribution to population biology. *in* J. White, ed., 'Studies in Plant Demography'. Academic Press, London. pp. 127–142.

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

MAF (2001*a*). 'Argentine ant survey completed'. Ministry of Agriculture and Forestry. Media Release.

MAF (2001*b*). 'Biosecurity costs regulations'. Ministry of Agriculture and Forestry. MAF Policy.

Magarey, R. D., Fowler, G. A., Borchert, D. M., Sutton, T. B., Colunga-Garcia, M. & Simpson, J. A. (2007). NAPPFAST: An internet system for the weather-based mapping of plant pathogens. *Plant Disease* **91**, 336–345.

Majer, J. (1993). Spread of Argentine ants (*Linepithema humile*), with special reference to Western Australia. *in* D. Williams, ed., 'Exotic ants; biology, impact, and control of introduced species.'. Westview Press. Boulder, CO. pp. 163–173.

Markin, G. . (1970*a*). Food distribution within laboratory colonies of the Argentine ant, *Iridomyrmex humilis* (Mayr). *Insectes Sociaux* **17**, 127–158.

Markin, G. P. (1970*b*). The seasonal life cycle of the Argentine ant, *Iridomyrmex humilis* (Hymenoptera: Formicidae), in southern California.. *Annals of the Entomological Society of America* **63**, 1238–1242.

Marsh, L. M. & Jones, R. E. (1988). The form and consequences of random walk movement models. *Journal of Theoretical Biology* **133**, 113–131.

Matthysen, E. (2005). Density-dependent dispersal in birds and mammals. *Ecography* **28**, 403–416.

May, R. (1976). Simple mathematical models with very complicated dynamics. *Nature* **261**, 459.

May, R. M. (1974). Biological populations with nonoverlapping generations: stable points, stable cycles and chaos. *Science* **186**, 645–647.

May, R. M. & Stumpf, M. P. H. (2000). Ecology: species-area relationships in tropical forests. *Science* **290**, 2084–2086.

Maynard-Smith, J. (1974). *Models in Ecology.* Cambridge University Press.

McCarthy, M. A. (1997). The Allee effect, finding mates and theoretical models. *Ecological Modelling* **103**, 99–102.

McCullough, D. G., Work, T. T., Cavey, J. F., Liebhold, A. M. & Marshall, D. (2006). Interceptions of nonindigenous plant pests at US ports of entry and border crossings over a 17-year period. *Biological Invasions* **8**, 611–630.

McPherson, J. M., Jetz, W. & Rogers, D. J. (2006). Using coarse-grained occurence data to predict species distributions at finer spatial resolutions - possibilities and limitations. *Ecological Modelling* **192**, 499–522.

Miyake, K., Kameyama, T., Sugiyama, T. & Ito, F. (2002). Effect of Argentine ant invasion on Japanese ant fauna in Hiroshima Prefecture, western Japan: a preliminary report (*Hymenoptera: Formicidae*). *Sociobiology* **39**, 465–474.

Moilanen, A. (2004). SPOMSIM: software for stochastic patch occupancy models of metapopulation dynamics. *Ecological Modelling* **179**, 533–550.

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

Moilanen, A., Smith, A. & Hanski, I. (1998). Long-term dynamics in a metapopulation of the American pika. *American Naturalist* **152**, 530–542.

Mollison, D. (1977). Spatial contact models for ecological and epidemic spread. *Journal of the Royal Statistical Society: Series B* **39**, 283–326.

Mollison, D. (1986). Modelling biological invasions: chance, explanation, prediction. *Philosopical Transactions of the Royal Society, B.* **314**, 675–693.

Mollison, D. (1987). *Quantitative apsects of the ecology of biological invasions*. Royal Society, London. chapter Modelling biological invasions: chance, explanation, prediction, pp. 675–693.

Mollison, D. & Kuulasmaa, K. (1985). Spatial epidemic models: theory and simulations. *in* P. J. Bacon, ed., 'Population dynamics of rabies in wildlife'. Academic, London. pp. 291–309.

Moody, M. E. & Mack, R. N. (1988). Controlling the spread of plant invasions: the importance of nascent foci. *Journal of Applied Ecology* **25**, 1009–1021.

Mooij, W. M. & DeAngelis, D. L. (1999). Error propagation in spattially explicit population models: a reassessment. *Conservation Biology* **13**, 903–933.

Mooij, W. N. & DeAngelis, D. L. (2003). Uncertainty in spatially explicit animal dispersal models. *Ecological Applications* **13**, 794–805.

Mooney, H. A. & Hobbs, R. J., eds (2000). *Invasive species in a changing world*. Island Press, Washington D. C.

Mooney, H. A., Hamburg, S. P. & Drake, J. A. (1986). The invasions of plants and animals into California. *in* H. A. Monney & J. A. Drake, eds, 'Ecology of Biological Invasions of North America and Hawaii'. Springer, New York. pp. 250–272.

Murray, J. D., Stanley, E. A. & Brown, D. L. (1986). On the spatial spread of rabies among foxes. *Proceedings of the Royal Society of London* **229**, 111–150.

Myers, I. B. (1980). *Gifts Differing: Understanding Personality Type*. Davies-Black Publishing.

Naeem, S., Byers, D., Tjossem, S. F., Bristow, C. & Li, S. (1999). Plant neighborhood diversity and production. *Ecoscience* **6**, 355–365.

Nakamaru, M. & Levin, S. A. (2004). Spread of two linked social norms on complex interaction networks. *Journal of Theoretical Biology* **230**, 57–62.

Nakamaru, M., Matsuda, H. & Iwasa, Y. (1996). The evolution of cooperation in a lattice-structured population. *Journal of Theoretical Biology* **184**, 65–81.

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

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

Neubert, M. G., Kot, M. & Lewis, M. A. (1995). Dispersal and pattern formation in a discrete-time predator-prey model. *Theoretical Population Biology* **48**, 7–43.

Newell, W. & Barber, T. C. (1913). The Argentine ant. Bulletin 122. Technical report. Bureau of Entomology, U.S. Department of Agriculture, Washington D.C.

Nowak, M. A., May, R. M. & Sigmund, K. (1995). The arithmetics of mutual help. *Scientific American* **272**(6), 76–81.

Okubo, A. (1980). *Diffusion and ecological problems: mathematical models*. Springer-Verlag, Berlin, Germany.

Okubo, A. & Levin, S. A. (2001). *Diffusion and Ecological Problems*. Springer, New York.

Olden, J. D. & Jackson, D. A. (2002). A comparison of statistical approaches for modelling fish species distributions. *Freshwater Biology* **47**, 1976–1995.

Olden, J. D. & Poff, N. L. (2003). Toward a mechanistic understanding and prediction of biotic homogenization. *American Naturalist* **162**, 442–460.

Oreskes, N., Shrader-Frechette, K. & Belitz, K. (1994). Verification, validation, and confirmation of numerical models in the earth sciences. *Science* **236**, 641–646.

Orr, M. R. & Seike, S. H. (1998). Parasitoids deter foraging by Argentine ants (*Linepithema humile*) in their native habitat in brazil. *Oecologia* **117**, 420–425.

Overton, J. M., Kean, J., Price, R., Williams, P. A., Barringer, J. R. F., Barron, M., Cooke, A., Martin, O. & Bellingham, P. J. (2004). Pestspread version 1.0: A prototype model to predict the spatial spread of pests. Contract Report 0405/048. Landcare Research and AgResearch.

P. J. den, B. (1991). Seeing the trees for the wood: random walks or bounded fluctuations of population size?. *Oecologia* **86**, 484–491.

Pahl-Wostl, C. (1996). *The Dynamic Nature of Ecosystems: Chaos and order*. John Wiley & Sons.

Pasfield, G. (1968). Argentine ants. *Australian Natural History* **16**, 12–15.

Passera, L. (1994). Characteristics of tramp species. *in* D. F. Williams, ed., 'Exotic ants: biology, impact and control of introduced species.'. Westview Press, Boulder, Colorado. pp. 22–43.

Pearce, D. W. & Moran, D., eds (1994). *The Economic Value of Biodiversity.* Earthscan Publications Ltd.

Pearce, J. & Ferrier, S. (2000). Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling* **133**, 225–245.

Peng, M. (2000). Spatio-temporal Modelling of Biological Invasions. PhD thesis. University of Auckland.

Peterson, A. T., Ortega-Huerta, M. A., Bartley, J., Sanchez-Cordero, V., Soberón, J., Buddemeier, R. H. & Stockwell, D. R. B. (2002). Future projections for Mexican faunas under global climate change scenarios. *Nature* **416**, 626–629.

Petrovskii, S. V., Malchow, H., Hilker, F. M. & Venturino, E. (2005). Patterns of patchy spread in deterministic and stochastic models of biological invasion and biological control. *Biological Invasions* **7**, 771–793.

Peuquet, D. J. (1999). Time in gis and geographical databases. *in* P. A. Longley, M. F. Goodchild, D. J. Maguire & D. W. Rhind, eds, 'Geographical Information Systems: Principles and Applications'. 2nd edn. Longman, London.

Pickett, S. T. A. & White, P. S. (1985). *The Ecology of Natural Disturbance and Patch Dynamics.* Academic Press, Orlando, FL.

Pielou, E. C. (1969). *An Introduction to Mathematical Ecology.* Wiley, New York.

Pielou, E. C. (1977). *Mathematical Ecology.* John WIley & Sons, New York, New York, USA.

Pielou, E. C. (1979). *Biogeography.* John Wiley & Sons, New York.

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

Pimentel, D., McNair, S., Janecka, J., Wightman, J., Simmonds, C., O'Connell, C., Wong, E., Russel, L., Zern, J., Aquino, T. & Tsomondo, T. (2001). Economic and environmental threats of alien plant, animal, and microbe invasions. *Agriculture, Ecosystems & Environment* **84**, 1–20.

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**, 273288.

Pimm, S. L. (1991). *The balance of nature?.* University of Chicago Press, Chicago Illinois.

Pimm, S. L. & Raven, P. (2001). Extinction by numbers. *Nature* **403**, 843–845.

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.

Plotnick, R. E. & Gardner, R. H. (1993). Lattice and landscape. *in* 'Lecture on Mathematics in the Life Sciences 23: Predicting spatial effects in Ecological Systems'. American Mathematical Society. pp. 129–157.

Porter, J., Arzberger, P., Braun, H.-W., Bryant, P., Gage, S., Hansen, T., Hanson, P., Lin, C.-C., Lin, F.-P., Kratz, T., Michener, W., Shapiro, S. & Williams, T. (2005). Wireless sensor networks for ecology. *BioScience* **55**, 561–572.

Porter, S., Van Eimeren, B. & Gilbert, L. (1988). Invasion of red imported fire ants (Hymenoptera: Formicidae): microgeography of competitive replacement. *Annals of the Entomological Society of America* **81**, 913–918.

Pysek, P. (1998). Is there a taxonomic pattern to plant invasions?. *Oikos* **82**, 282–294.

Rabin, S. (2002). *AI Game Programming Wisdom.* Charles River Media.

Radosevich, S. R., Holt, J. & Ghersa, C. M. (1997). *Weed Ecology, Implications for Management.* John Wiley and Sons. New York, NY.

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

Rasmussen, D. R. & Bohr, T. (1987). Temporal chaos and spatial disorder. *Physics Letters A* **125**, 107–110.

Reeve, R. D. & Murdoch, W. W. (1986). Biological control by the parasitoid *Aphytis melinus*, and population stability of the California red scale. *Journal of Animal Ecology* **55**, 1069–1082.

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

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

Reichard, S. H. & Hamilton, C. W. (1997). Predicting invasions of woody plants introduced into North America. *Conservation Biology* **11**, 193–203.

Reimer, N. (1993). Distribution and impact of alien ants in vulnerable Hawaiian ecosystems. *in* D. Williams, ed., 'Exotic ants: Biology, impact, and control of introduced species'. Westview Press, Boulder, CO. pp. 11–22.

Rejmánek, M. (2000). Invasive plants: approaches and predictions. *Australian Ecology* **25**, 497–506.

Reyes, E., Martin, J. F., White, M. L., Day, J. W. & Kemp, G. P. (2004). Habitat changes in the Mississippi Delta: future scenarios and alternatives.. *in* R. Costanza & A. Voinov, eds, 'Landscape Simulation Modeling: A Spatially Explicit, Dynamic Approach'. Springer. chapter 5.

Rhainds, M., Cries, G. & Saleh, A. (1998). Density and pupation site of apterous female bagworms, *Metisa plana* (Walker) (Ledidoptera: Psychidae), influence the distribution of emergent larvae. *Canadian Entomologist* **130**, 603–613.

Richards, F. J. (1959). A flexible growth function for empirical use. *Journal of Experimental Botany* **10**, 290–300.

Richardson, D. (1973). Random growth in a tessellation. *Proceedings of the Cambridge Philosophical Society* **74**, 515–528.

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

Richter-Dyn, N. & Goel, N. S. (1972). On the extinction of a colonizing species. *Theoretical Population Biology* **3**, 406–433.

Ricker, W. E. (1958). *Handbook of computations for biological statistics of fish populations.* Fisheries Research Board of Canada.

Ritchie, M., Campbell, H. & Sivak, L. (2000). Investigating the market for organic food: Dunedin, new zealand and the world. *in* 'Organics 2020: New Zealand National Conference on Organics'. Unitec Auckland.

Rose, D. J. W. (1978). Epidemiology of maize streak disease. *Annual Revue of Entomology* **23**, 259–282.

Ross, M. G. (2004). Response to a gypsy moth incursion within New Zealand. *in* 'IUFRO conference, Hanmer, 2004'. Biosecurity New Zealand, MAF.

Rosset, H., Keller, L. & Chapuisat, M. (2005). Experimental manipulation of colony genetic diversity had no effect on short-term task efficiency in the Argentine ant *Linepithema humile. Behavioral Ecology and Sociobiology* **58**, 87–98.

Roura-Pascual, N., Suarez, A. V., Gomez, C., Pons, P., Touyama, Y., Wild, A. L. & Peterson, A. T. (2004). Geographical potential of argentine ants (*Linepithema humile* mayr) in the face of global climate change. *Proceedings of the Royal Society of London B* **271**, 2527–2534.

Ruckelshaus, M., Hartway, C. & Kareiva, P. (1997). Assessing the data requirements of spatially explicit dispersal models.. *Conservation Biology* **11**(6), 1298–1306.

Ruxton, G. & Saravia, L. (1998). The need for biological realism in the updating of cellular automata models. *Ecological Modelling* **107**, 105–112.

Ruxton, G. D. & Rohani, P. (1999). Fitness-dependent dispersal in metapopulations and its consequences for persistence and synchrony. *Journal of Animal Ecology* **67**, 530–539.

Rykiel, E. J., J. (1996). Testing ecological models: the meaning of validation. *Ecological Modelling* **90**, 229–244.

Samet, H. (1990). *Applications of Spatial data structures: Computer Graphics, Image Processing and Geographical Information Systems.* Addison-Wesley, Reading, Mass.

Samways, M., Magda, N. & Prins, A. (1982). Ants (Hymenoptera: Formcidae) foraging in citrus trees and attending honeydew-producing homoptera. *Phytophylactica* **14**, 155–157.

Satake, A., Iwasa, Y., Hakoyama, H. & Hubbell, S. P. (2004). Estimating local interaction from spatiotemproal forest data, and Monte Carlo bias correction. *J Theor Biol* **226**, 225–235.

Sato, K., Matsuda, H. & Sasaki, A. (1994). Pathogen invasion and host extinction in lattice structured populations. *J Math Biol* **32**, 251–268.

Schnase, J., Stohlgren, T. & Smith, J. (2002). 'The national invasive species forecasting system: A strategic NASA / USGS partnership to manage biological invasions.'. Earth Observation Magazine.

Schröpfer, R. & Engstfeld, C. (1983). Die ausbreitung des bisams (*Ondatra zibethicus* linné, 1766, rodentia, arvicolidae) in der Bundesrepublik Deutschland. *Z. Angew. Zool* **70**, 1337.

Schumaker, N. H. (1996). Using landcape indices to predict habitat connnectivity. *Ecology* **77**(4), 1210–1225.

Scott, J. K. & Panetta, F. (1993). Predicting the Australian weed status of southern African plants. *Journal of Biogeography* **20**, 87–93.

Segurado, P. & Araújo, M. B. (2004). An evaluation of methods for modelling species distributions. *Journal of Biogeography* **31**, 1555–1568.

Segurado, P. & Araújo, M. B. (2006). Consequence of spatial autocorrelation for niche-based models. *Journal of Applied Ecology* **43**, 433–444.

Sequeira, R. A. (2002). Role of spatial analysis in meeting phytosanitary challenges including risk analysis. *in* 'Proceedings of International Symposium on Pest Risk Analysis'. NAPPO.

Settles, G. S. (2005). Sniffers: Fluid-dynamic sampling for olfactory trace detection in nature and homeland security. *Journal of Fluids Engineering* **127**, 189–219.

Shaffer, M. L. (1981). Minimum population sizes for species conservation. *BioScience* **31**, 131–134.

Shapiro, M. & Westervelt, J. (1991). r.mapcalc: An algebra for gis and image processing. Technical report. U.S. Army Construction Engineering Research Laboratory.

Sharov, A. A., Leonard, D., Liebhold, A. M., Roberts, E. A. & Dickerson, W. (2002). Slow the spread: a national program to contain the gypsy moth. *Journal of Forestry* pp. 30–35.

Sharov, A. A., Roberts, E. A., Liebhold, A. M. & Ravlin, F. W. (1995). Gypsy moth (*Lepidoptera: Lymantriidae*) spread in the central Appalachians: Three methods for species boundaries estimation. *Population Ecology* **24**(6), 1529–1538.

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* **259**, 243–248.

Shigesada, N. & Kawasaki, K. (1997). *Biological Invasions: Theory and Practice*. Oxford University Press.

Shigesada, N., Kawasaki, K. & Takeda, Y. (1995). Modelling stratified diffusion biological invasions.. *The American Naturalist* **146**(2), 229–254.

Sibly, R. M., Barker, D., Hone, J. & Pagel, M. (2007). On the stability of populations of mammals, birds, fish and insects. *Ecology Letters* **10**, 970–976.

Silvertown, J., Holtier, S., Johnson, J. & Dale, P. (1992). Cellular automaton models of interspecific competition for space - the effect of pattern on process. *Journal of Ecology* **80**, 527–534.

Simberloff, D. & Alexander, M. (1998). Assessing risks from biological introductions (excluding gmos) for ecological systems. *in* P. Calow, ed., 'Handbook of Environmental Risk Assessment and Management'. Oxford University Press, Oxford. pp. 147–176.

Sinclair, A. R. E. (1992). Do large mammals disperse like small mammals?. *in* N. C. Stenseth & W. Z. Lidicker, eds, 'Animal dispersal. Small mammals as a model'. London: Chapman & Hall. pp. 229–242.

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

Smythe, R. T. & Wierman, J. C. (1978). *Lecture Notes in Mathematics: First-Passage Percolation on the Square Lattice*. Springer-Verlag, New York.

Snyder, R. E. (2003). How demographic stochasticity can slow biological invasions. *Ecology* **84**, 1333–1339.

Sockman, K. (1997). Variation in life-history traits and nest-site selection affects risk of nest predation in the California gnatcatcher. *The Auk* **114**, 324–332.

Soulé, M. E. (1987). *Viable Populations for Conservation*. Cambridge University Press, Cambridge.

South, A. (1999). Dispersal in spatially explicit models. *Conservation Biology* **13**, 1039–1046.

Southwood, T. R. E. (1962). Migrationof terrestial arthropods in relation to habitat. *Biology Revue* **37**, 171–214.

Stamps, J. A. (2006). The silver spoon effect and habitat selection by natal dispersers. *Ecology Letters* **9**, 1179–1185.

Starfield, A. M. & Bleloch, A. L. (1986). *Building Models for Conservation and Wildlife Management.* Collier Macmillan, London.

Steele, J. H. (1989). Discussion: Scale and coupling in ecological systems. *in* J. Roughgarden, R. M. May & S. A. Levin, eds, 'Perspectives in Ecological Theory'. Princeton University Press, Princeton N.J.. pp. 177–180.

Steinberg, E. K. & Kareiva, P. (1997). Challenges and opportunities for empirical evaluation of "Spatial Theory". *in* D. Tilman & P. Kareiva, eds, 'Spatial Ecology. The Role of Space in Population Dynamics and Interspecific Interactions'. Princeton University Press, Princeton, NJ. pp. 318–332.

Steyaert, L. T. (1993). A perspective for studying of environmental simulation. *in* M. F. Goodchild, B. O. Parks & L. T. Steyaert, eds, 'Environmental Modelling with GIS'. Oxford University Press, New York. pp. 16–30.

Stockwell, D. (1999). The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science* **13**, 143–158.

Suarez, A., Bolger, D. & Case, T. (1998). Effects of fragmentation and invasion on native ant communities in coastal southern California. *Ecology* **79**, 2041–2056.

Suarez, A. V., Holway, D. A. & Case, T. J. (2001). Patterns of spread in biological invasions dominated by long-distance jump dispersal: Insights from Argentine Ants. *Ecology* **98**(3), 1095–1100.

Suarez, A. V., Richmond, J. Q. & Case, T. J. (2000). Prey selection in horned lizards following the invasion of Argentine ants in southern California. *Ecological Applications.*

Suarez, A. V., Tsutsui, N. D., Holway, D. A. & Case, T. J. (1999). Behavioral and genetic differentiation between native and introduced populations of the Argentine ant. *Biological Invasions* **1**, 43–53.

Sutherst, R., Maywald, G., Yonow, T. & Stevens, P. (1999). *CLIMEX: predicting the effects of climate on plants and animals.* Collingwood, Australia CSIRO Publishing.

Sutherst, R. W. & Maywald, G. F. (1985). Climate-matching for quarantine, using CLIMEX. *Plant Protection Quarterly* **6**, 3–7.

Sykes, M. T., Prentice, I. C. & Cramer, W. (1996). A bioclimatic model for the potential distributions of north European tree species under present and future climate. *Journal of Biogeography* **23**, 203–233.

Tainaka, K. (1988). Lattice model for the Lotka-Volterra system. *Journal of the Physical Society of Japan* **57**, 2588–2590.

Taylor, C. M. & Hastings, A. (2004). Finding optimal control strategies for invasive species: a density-structured model for *Spartina alterniflora*. *Journal of Applied Ecology* **41**, 1049–1057.

Taylor, R. A. J. (1980). A family of regression equations describing the density distribution of dispersing organisms. *Nature* **286**, 53–55.

Thuiller, W., Lavorel, S., Ara'ujo, M. B., Sykes, M. T. & Prentice, I. C. (2005). Climate change threatens plant diversity in Europe. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 8245–8250.

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

Tilman, D., Lehman, C. L. & Yin, C. (1997). Habitat destruction, dispersal and deterministic extinction in competitive communities. *American Naturalist* **149**, 407–435.

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

Touyama, Y., Ogata, K. & Sugiyama, T. (2003). The Argentine ant, *linepithema humile*, in Japan: Assessment of impact on species diversity of ant communities in urban environments. *Entomological Science* **6**, 57–62.

Travis, J. M. J. (2003). Neighbourhood size, dispersal distance and the complex dynamics of the spatial Ricker model. *Population Ecology* **45**, 227–237.

Tremper, B. S. (1976). Distribution of Argentine ant, *Iridomyrmex humilis* Mayr, in relation to certain native ants of California: ecological, physiological, and behavioral aspects. PhD thesis. University of California, Berkeley.

Tsutsui, N. D., Suarez, A. V., Holway, D. A. & Case, T. J. (2001). Relationships among native and introduced populations of the Argentine ant (*Linepithema humile*) and the source of introduced populations. *Molecular Ecology* **10**, 2151–2161.

Tsutsui, N., Suarez, A., Holway, D. & Case, T. (2000). Reduced genetic variation and the success of an invasive species. *in* 'Proceedings of the National Academy of Sciences of the United States of America'. Vol. 97. pp. 5948–5953.

Turing, A. (1952). The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London, B.*

Turner, M. G. (1989). Landscape ecology: the effect of pattern on process. *Annual Review of Ecology and Systematics* **20**, 171–197.

Turner, M. G., Arthaud, G. J., Engstrom, R. T., Hejl, S. J., Liu, J., Loeb, S. & McKelvey, K. (1995). Usefulness of spatially explicit population models in land management. *Ecological Applications* **5**, 12–16.

Turner, M. G., Wu, Y. G., Romme, W. H., & Wallace, L. L. (1993). A landscape simulation model of winter foraging by large ungulates. *Ecological Modelling* **69**, 163–184.

Usher, M. B. (1987). Invasibility and wildlife conservation: invasive species on nature reserves. *in* H. Kornberg & M. H. Williamson, eds, 'Quantatative Aspects of the Ecology of Biological Invasions'. The Royal Society, London. pp. 193–207.

Valentine, E. W. & Walker, A. K. (1991). *Annotated Catologue of New Zealand Hymenoptera.* DSIR Plant Protection Report 4. General Printing Services Ltd, Nelson.

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

Van Schagen, J., Davis, P. & Widmer, M. (1993). Ant pests of Western Australia, with particular reference to the Argentine ant (*Linepithema humile*). *in* D. Williams, ed., 'Exotic ants: Biology, impact, and control of introduced species'. Westview Press, Boulder, CO. pp. 174–180.

Vega, S. & Rust, M. (2001). The Argentine ant: A significant invasive species in agricultural, urban and natural environments. *Sociobiology* **37**, 3–25.

Veit, R. R. & Lewis, M. A. (1996). Dispersal, population growth and the Allee effect: dynamics of the House Finch invasion of North America. *The American Naturalist* **148**, 255–274.

Verhulst, P. F. (1838). Notice sur la loi que la population pursuit dans son accroissement. *Correspondance mathématique et physique* **10**, 113–121.

Visser, D., Wright, M. G. & Giliomee, J. H. (1996). The effect of the Argentine ant, *Linepithema humile* (Mayr) (Hymenoptera, Formicidae), on flower-visiting insects of *Protea nitida* Mill (Proteaceae). *African Entomology* **4**, 285–287.

Vitousek, P. M., D'Antonio, C. M., Loope, L. L. & Westbrooks, R. G. (1996). Biological invasions as global environmental change. *American Scientist* **84**, 218–228.

Vitousek, P. M., DAntonio, C. M., Loope, L. L., Rejmánek, M. & Westbrooks, R. (1997). Introduced species: a significant component of human-caused global change. *New Zealand Journal of Ecology* **21**, 1–16.

Von Holle, B. & Simberloff, D. (2005). Ecological resistance to biological invasion overwhelmed by propagule pressure. *2005* **86**, 32123218.

Vos, C. C., Verboom, J., Opdam, P. F. M. & Ter Braak, C. J. F. (2001). Toward ecologically scaled landscape indices. *American Naturalist* **183**, 24–41.

Walker, P. A. & Cocks, K. D. (1991). HABITAT: A procedure for modelling a disjoint environmental envelope for a plant or animal species.. *Global Ecology and Biogeography Letters* **1**(4), 108–118.

Walsh, J. E., Wilson, K.-J. & Elliott, G. P. (2006). Seasonal changes in home range size and habitat selection by Kakapo (*Strigops habroptilus*) on Maud Island. *Notornis* **53**(1), 143–149.

Walters, A. C. & Mackay, D. A. (2003). An experimental study of the relative humidity preference and survival of the Argentine ant, *Linepithema humile* (Hymenoptera, Formicidae): comparisons with a native *Iridomyrmex* species in South Australia. *Insectes Sociaux* **50**, 355–360.

Walters, A. C. & Mackay, D. A. (2004). Comparisons of upper thermal tolerances between the invasive argentine ant (Hymenoptera: Formicidae) and two native Australian ant species. *Annals of the Entomological Society of America* **97**, 971–975.

Walters, A. C. & MacKay, D. A. (2005). Importance of large colony size for successful invasion by Argentine ants (Hymenoptera: Formicidae): Evidence for biotic resistance by native ants. *Australian Ecology* **30**, 395–406.

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

Ward, D. F. & Harris, R. J. (2005). Invasibility of native habitats by Argentine ants, *Linepithema humile*, in New Zealand. *New Zealand Journal of Ecology.*

Ward, D. F., Harris, R. J. & Stanley, M. C. (2005). Human-mediated range expansion of Argentine ants *Linepithema humile* (Hymenoptera: Formicidae) in New Zealand. *Sociobiology* **45**(2), 1–7.

Ward, J. V. (1992). *Aquatic Insect Ecology: Biology and Habitat*. Wiley-Interscience.

Ward, P. (1987). Distribution of the introduced Argentine ant (*Iridomyrmex humilis*) in natural habitats of the lower Sacramento valley and its effects on the indigenous ant fauna. *Hilgardia* **55**, 1–16.

Waser, P. M. (1985). Does competition drive dispersal?. *Ecology* **66**, 1170–1175.

Watson, A. K. (1985). Integrated management of leafy spurge. *in* A. K. Watson, ed., 'Leafy Spurge'. Weed Science Society of America, Champaign. pp. 93–104. Monograph 3.

Way, M., Cammell, M., Paiva, M. & Collingwood, C. (1997). Distribution and dynamics of the Argentine ant *Linepithema (Iridomyrmex) humile* (Mayr) in relation to vegetation, soil conditions, topography and native competitor ants in Portugal. *Insectes Sociaux* **44**, 415–433.

Way, M. J., Cammell, M. E. & Paiva, M. R. (1992). Studies on egg predation by ants (Hymenoptera: Formicidae) especially on the eucalyptus borer *Phoracantha semipunctata* (Coleoptera: Cerambycidae) in Portugal. *Bulletin of Entomological Research* **82**, 425–432.

Weeda, E. J. (1987). Invasions of vascular plants and mosses into the Netherlands. *in* W. Joenje, K. Bakker & L. Vlijm, eds, 'Proceedings of the Koninklijke Nederlandse Akademie van Wetenschapen, Amsterdam. Series C: Biological and Medical Sciences'. pp. 19–29.

Weinberger, H. F. (1982). Long-time behaviour of a class of biological models. *SIAM (Society for Industrial and Applied Mathematics) Journal of Mathematical Analysis* **13**, 353–396.

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

Weiss, M., Elsner, M., Kartberg, F. & Nilsson, T. (2004). Anomalous subdiffusion is a measure for cytoplasmic crowding in living cells. *Biophysical Journal* **87**, 3518–3524.

Whitney, G. (1994). *From coastal wilderness to fruited plain: a history of environmental change in temperate North America from 1500 to the present*. Cambridge University Press, New York.

Whyte, C. (2005*a*). Researching pathways to find the holes and improve the biosecurity system. *in* 'Science and security: informing NZ'. Royal Society of NZ. Wellington.

Whyte, C. (2005*b*). Science and biosecurity - monitoring the effectiveness of biosecurity interventions at new zealands borders. *in* 'Science and security - informing NZ'. Royal Society.

Wiegand, T., Moloney, K. A., Naves, J. & Knauer, F. (1999). Finding the missing link between landscape structure and population dynamics: a spatially explicit perspective. *American Naturalist* **154**, 605–627.

Wiens, J. A., Schooley, R. L. & Weeks, R. D. (1997). Patch landscape and animal movements: do beetles percolate?. *Oikos* **78**, 257–264.

Wilcove, D. S., Rothstein, D., Dubow, J., Phillips, A. & Losos, E. (1998). Assessing the relative importance of habitat destruction, alien species, pollution, over-exploitation, and disease. *BioScience* **48**, 607–616.

Williams, D. F., ed. (1994). *Exotic ants: biology, impact, and control of introduced species.*. Westview Press, Boulder, Colarado.

Williamson, M. & Fitter, A. (1996). The varying success of invaders. *Ecology* **77**(6), 1661–1666.

Wissel, C. (1989). *Theoretische Ökologie - Eine Einf̈hrung.* Springer, Berlin.

Wissel, C. (1992). Aims and limits of ecological modelling exemplified by island theory. *Ecological Modelling* **63**, 1–12.

Wit, C. T. d. (1993). Philosophy and terminology. *in* P. A. Leffeelar, ed., 'On Systems Analysis and simulation of Ecological Processes With Examples in CSMP and Fortran'. Kluwer Academic Publishing. pp. 3–9.

With, K. A. (1994). Using fractal analysis to assess species perceptions of landscape structure. *Landscape Ecology* **9**, 25–36.

With, K. A. (1997). The application of neutral landscape models in conservation biology. *Conservation Biology* **11**(5), 1069–1080.

With, K. A. (2002). The landscape ecology of invasive spread. *Conservation Biology* **16**(5), 1192–1203.

With, K. A. & King, A. W. (1999). Dispersal thresholds in fractal landscapes: a consequence of lacunarity thresholds. *Landscape Ecology* **14**, 73–82.

With, K. A. & King, A. W. (2001). Analysis of landscape sources and sinks: the effect of spatial pattern on avian demography. *Biological Conservation* **100**, 75–88.

Wolff, J. (1997). Population regulation in mammals: an evolutionary perspective. *Journal of Animal Ecology* **66**, 1–13.

Wolfram, S. (2002). *A New Kind of Science*. Wolfram Media.

Worner, S. P. (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. pp. 11–32.

Xu, J. (1994). Simulating the spread of wildfires using a geographic information system and remote sensing. PhD thesis. Rutgers The State University of New Jersey.

Zadoks, J. C. & Schein, R. D. (1979). *Epidemiology and Plant Disease Management*. Oxford University Press.

Zar, J. H. (1999). *Biostatistical Analysis*. fourth edn. Prentice-Hall. chapter Testing for Randomness, pp. 571–578.

Zweig, M. H. & Campbell, G. (1993). Receiver-operator characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry* **39**, 561–577.

Appendix A

Gypsy Moth phenology modelling

This appendix contains a paper written by the thesis author, Joel Peter William Pitt, that was accepted and published in the *International Journal of Biometeorology*, 2007, 51:295 305.

The paper investigates the use of a process driven phenology model to predict which areas of New Zealand have a suitable climate for the species to develop completely and reproduce.

The output of this model would be one method of generating a potential distribution map for input in the survival module of **MDiG** (Section 3.3.3), or if **MDiG** is extended for modelling lifestage dependent dispersal (as is discussed in chapter 6) then the phenology model could be used to time the region-specific occurrences of dispersal associated with each lifestage.

Appendix B

The future of MDiG

B.1 Extensions to existing modules

In the creation of the **MDiG** simulation framework many aspects of dispersal and population processes were considered, such that **MDiG** could be as broadly applicable as possible but still be a useful dispersal framework.

Future directions for existing **MDiG** modules are described below. These ideas were not further researched or implemented, since they were beyond the scope of this thesis. They do however show the potential of the framework.

B.1.1 Possible extensions of the local dispersal module

Local search for suitable habitat.

The **local** module approximates classical diffusion, which is the average of individual movement according to the laws of Brownian motion. Some argue that individual movement is not a truly random process as Brownian motion is, despite this, considering individuals as randomly moving particles is the basis of much analytical work on dispersal that successfully modelled observed patterns of spread.

An alternative would be to define the neighbourhood area in the normal way, but instead of all patches being occupied according to the radius and shape parameters, a suitability map could be referenced (section 3.3.3) and the module would simulate individuals moving from an occupied patch to the patch or patches that are most suitable. The process could be deterministic, i.e. individuals would always move to the patch with the best suitability or the population of dispersers could be split across destination patches in proportion to the patch's suitability. Occupation could alternatively be implemented stochastically, where the destination patch is chosen randomly but each patch's probability of immigration is proportional to its suitability.

If information about habitat or resource selection was available, it might guide estimation of how different habitats should be weighted. For example, the 'silver spoon' effect occurs when dispersers in good condition are more likely to settle in high-quality habitats than dispersers in poor condition. Stamps (2006) suggests that the positive relationship between disperser condition and the quality of post-dispersal habitat is predicted by at least two hypotheses. The competition hypothesis assumes that a disperser's condition affects its chances of competing for space or joining an established group after arriving at a high-quality habitat, while the search hypothesis assumes that a disperser's condition affects its selectivity, and hence its chances of accepting a lower-quality habitat when it is searching for a new habitat. Combining this knowledge with an estimate of the condition of individuals in a population could allow dispersers to be more realistically distributed.

Kernel distribution of individuals

Instead of specifying the **proportion** of individuals to be spread evenly across patches in the neighbourhood, one could specify a dispersal kernel or probability distribution to determine out how to distribute individuals across patches. This would be similar to using the **kernel** module with a very large number of dispersal events, except it would be computationally faster because the proportions dispersing to each cell in the neighbourhood would be determined in advance rather than simulating each dispersal event.

Using maps to specify neighbourhood

Currently the neighbourhood parameter is a single value that is used globally. In other words the parameter applies to every raster cell that is processed. An alternative method would be to use a raster map as the **shape** parameter, where the value in each cell indicates the neighbourhood shape for the equivalent patch in the population distribution.

An example of such a map might be used in the downhill spread of eggs, seeds, or spores. Where the shape value in every cell of the map passed to the **shape** parameter map would indicate the slope and aspect of the terrain.

B.1.2 Possible extensions for the kernel dispersal module

Dispersal events transporting more than one individual.

Currently, in population-based models, the **kernel** module simply transfers a single individual from the source to destination patch. However, for human-assisted long-distance dispersal, that usually is not the case. More often, a group of individuals, for example as part of a nest, egg mass, or otherwise, are transported on cargo or passenger luggage. The transfer of multiple individuals would become especially important if population-based models incorporate an Allee effect (section 3.3.4), since an Allee effect may prevent populations from establishing if only singlular individuals are transferred as a result of long-distance dispersal.

One could also argue that the Poisson distribution generating dispersal events should have a positive density-dependent mean, with a greater number of individuals in an area increasing dispersal. For example, in plants where more source plants result in more seeds and more chance that at least one seed disperses beyond the source patch. Empirical studies suggesting that species dispersal rates depend on population size, with a positive relationship between dispersal rate and population density observed in spiders (Duffey 1998), insects (Rhainds et al. 1998), mammals (Sinclair 1992), and birds (Veit & Lewis 1996). Negative relationships between population density and dispersal rates have also been observed, but mainly in mammals - possibly due to crowding effects impeding free movement (Wolff 1997, Diffendorfer et al. 1999, Matthysen 2005).

Alternatively, there could be a threshold population required before a long distance dispersal event can occur, but with a constant mean for the number of dispersal events generated, regardless of population size. Or, the density of individuals could even affect the number of individuals that are dispersed in an event.

Travis (2003) used the Ricker growth equation to determine the mean of a Poisson process generating the number of new dispersal events. Thus, another possibility is that the **kernel** module could be altered to use growth equations for the generation new dispersal events.

Mixed kernels

Researchers have found it difficult to fit a single distribution to data on dispersal distances (Clark et al. 1999). Models on a fine scale often assume the probability distribution is convex near the dispersal source and platykurtic, but coarse scale models commonly use distributions that are concave near the source and leptokurtic or “fat-tailed” (Clark et al. 1999).

One way to resolve the difficulty in fitting probability distributions to data is by using mixed models, such as a bivariate Student’s t (or “$2Dt$”) distribution (Clark et al. 1999). Implementing mixed model kernels could allow the module to better represent dispersal data. However, the necessity for directly using mixed models may be reduced by the use of multiple **kernel** modules with different distributions in a single timestep, or by combining a **kernel** module with a **local** module. Still, a mixed model kernel could potentially require less computating time to simulate, since only one module would need to be processed instead of at least two, or provide a better fit to dispersal data when individual dispersal mechanisms are unable to be elucidated.

Truncation of long distance dispersal - expansion of region.

When using "fat-tailed" kernels, such as the Cauchy distribution, extremely long distance events can occur. The observed events however are always limited by the extent of the spatial area in which the simulation is running. Rare but influential long distance events greatly effect rates of spread and dynamics (Clark et al. 2001). By constraining the simulation to a limited area, we may be limiting the rate at which the population spreads (see section 4.2.1), since it prevents areas outside of the simulation region from becoming established and subsequently sending individuals back in the simulation region.

The problem could be solved either by the use of a dynamically sized GIS region or another data structure that records populations outside of the region. However, for particular simulation regions, such as islands or waterlocked nations (such as New Zealand), dynamic regions would have limited use. Additionally, a dynamically-sized region is only useful if there is GIS data available for environmental suitability beyond the initial simulation region.

Directed dispersal.

In this study a uniform distribution is used to determine the direction of dispersal events. While many dispersal methods use a kernel to calculate the probability of propagules travelling a given distance, those propagules may not spread evenly in all directions such as where a prevailing wind influences dispersal. Such wind influences affect both seeds (Bullock & Clarke 2000) and insects (Johnson 1957).

One method to include wind direction is to provide relative frequencies for wind travelling in each of 8 compass directions (North, North-east, East, etc.) along with the wind speed in each of these directions. The frequencies can be used to weight the likelihood of a dispersal event travelling a certain direction, and the wind velocity alters the distance travelled (either by altering a kernel distribution's equivalent of a distance parameter or by altering the final value sampled from the distribution) (Overton et al. 2004). However, since there are a variety of ways just to implement effects such as wind, it may be more appropriate to construct a new specialised module rather than extend the **kernel** module.

Wind dispersal was not included in the implentation of **MDiG** used in this thesis because the case study species, *L. humile*, is not know to have it's dispersal significantly influenced by wind (Chapter 5).

Dispersal from boundaries only.

The kernel module used in this model is simple. A Poisson distribution is sampled for every patch with an established population. Sometimes metapopulation models use the perimeter size of contiguous populations to calculate emmigration from a patch, since the greater the perimeter the more points that a species can disperse across (i.e. only cells on the perimeter of a species distribution) (Hambck & Englund 2005).

However, emmigration and immigration rates generally are assumed to be patch area dependent (Kareiva 1985, Hill et al. 1996, Kuussaari et al. 1996), so the current methodology for the module is likely to be the most useful. When perimeter-dependence is observed, it is often due to short dispersal distances. In raster population maps the patches inside of a contiguous region of populations will be less likely to establish new patches.

Large numbers of events

In the current model, each dispersal event from every cell is simulated individually. This works well for rare long-distance events, since the module doesn't have to simulate many events. If many events are simulated however, then the simulations quickly can become almost intractable, requiring several orders of magnitude more processing time. The kernel module could be extended to approximate very large Poisson means, by generating a large number of dispersal events in memory and then applying the same distribution of dispersal events to every cell. This would be useful as an approximation of nearby dispersal events, but would give different results from simulating each event, especially for long tailed kernels, since long-distance dispersal events have more options on where to land, the further away the propagules travel from their source.

This strategy is similar to what PestSpread currently uses (Overton et al. 2004).

B.1.3 Possible extensions to the patch survival module

Percolation Models

Many methods for creating suitability maps lead to maps that have broad areas of either suitable or unsuitable regions. Realistically the environment is made up of many micro-habitats that might provide a suitable haven in a swath of unsuitable environment or vice versa. Generating site percolation maps, where each cell is either suitable or unsuitable, could be done by evaluating the extinction probabilities of a suitability map once at the beginning of the simulation. Cells that are evaluated as extinct are unsuitable microhabitats, while those that are evalulated as surviving are suitable. In a sense this is a deterministic version of the extinction module which implicitly and stochastically creates a different site percolation map for each timestep.

Alternatively, defining the boundary between cells as either traversable or not may be a better approach. A percolation model based on this method is called *bond percolation* and would have to be integrated into the **local** module as it handles the contiguous spread of a population. One difficulty however, would be how to deal with percolation when the **local** module uses a large neighbourhood shape. With larger neighbourhoods it may be possible to find a path around impassible patch boundaries. A potential solution is the use of pathfinding algorithms such as A* (for details see Rabin 2002).

Regions are not static

Ecosystems are open and far from equilibrium and it has been suggested that stable points in nature are rare (Pahl-Wostl 1996). Generally methods for creating suitability maps use current data, or statistical averages, to create a static map indicating areas of suitability. Because the environment is changing, as is the population distribution of a species, a static map is unrealistic. Fluctuations in the abiotic or biotic environment, such as resource availability and competition, have been shown to have a large impact on plant invasion success (Davis & Pelsor 2001). Regional stochasticity can critically influence estimates of metapopulation persistance because correlation between local extinctions greatly increases the probability that the entire metapopulation becomes extinct (Moilanen & Cabeza 2002).

Assuming that the suitability maps are continously valued indices (rather than simply binomial values such suitable or unsuitable) then one method for introducing regional stochasticity into a static raster map is to multiply each cell value by a number sampled from a log-normal distribution as done by Moilanen et al. (1998) and Moilanen & Cabeza (2002). The strength of the regional stochasticity is determined by the standard deviation of the normal distribution used which represents some measure of environmental change. Additionally, the output from landscape models that predict the change of habitat types across a region could be used (Reyes et al. 2004).

Eradication treatment maps

In addition to using a map of habitat and climate suitability to influence a species' distribition and spread we could apply additional maps that represent different methods for controlling or eradicating the population.

Maps of treatment could include allowances for the effectiveness of a treatment method by using appropriate survival probabilities. Different spatial patterns of treatment application could be evalutated based on what we know of a species dispersal characteristics, suggesting appropriate buffer regions and optimising cost versus the chance of failing to control a dispersing species.

B.1.4 Possible extensions for the growth module

Growth as a Stochastic Process

All the growth models used above assume that populations are deterministic in their growth or decline. Frequently however, populations exhibit stochastic dynamics:

In fishery population studies, disasters and bonanzas are important factors in sustainability and their occurrence can be described by a Poisson process with formal solutions that describe the expected time till extinction (Hanson 1987, Hanson & Tuckwell 1997).

Long term studies of carabid populations have shown that their dynamics are indistinguishable from random walks over 20 year periods or more (P. J. den 1991).

Stochastic population growth models in biology are important in the conservation of endangered populations (Shaffer 1981, Soulé 1987, Lande & Orzack 1988) where such stochasticity could have especially severe effects on small populations. Early work also realised the potential importance of stochastic population growth to colonization theory (MacArthur & Wilson 1967, Goel & Richter-Dyn 1974), and thus it is important for invasion biology also.

Stochastic growth models are also important because all parameter measurements are uncertain and have their own measure of variability. This variability can be incorporated in a stochastic model.

Demographic stochasticity is also an important factor for spread processes. Haccou et al. (2005) suggests that the initial phase of a disease epidemic is described by a branching process that, even when above the critical threshold of infection, may go extinct due to stochastic fluctuations in reproduction rate. These fluctuations can also cause the average realised population growth rate to be considerably less than the mean of instantaneous growth rates over time (Lande et al. 2003).

In established populations however, most species tend to remain stable despite year to year fluctuations caused by environmental factors (Sibly et al. 2007). In the absence of density-dependent effects invasion speeds are not affect by stochasticity (Kot et al. 2004), but the presence of density-dependent effects with stochasticity generally decrease invasion speed (Mollison 1977, Lewis 2000, Lewis & Pacala 2000, Clark et al. 2001, Snyder 2003).

More growth equations

Only six growth equations have been implemented as part of the **growth** module. There is always potential for more to included. For example the generalised logistic growth, or Richards' curve, is a widely-used and flexible function in growth modelling (Richards 1959) and may fit the dynamics of a species' population growth better than the currently implemented growth models.

Spatially variable Allee effect

Tobin et al. (2007) demonstrates that Allee effect changes geographically for Gypsy Moth due to wind patterns (more wind disperses individuals, higher allee threshold). although this may depend on the resolution of the model since if spread is already influenced by wind, then variable allee effects due to wind would be included implicitly.

B.1.5 Possible extensions to the validation system of MDiG

Other methods of assessment

To evaluate the probability envelope we can draw from wide discussions on model evaluation philosophy, logic and semantics in the environmental modelling literature (e.g. Rykiel 1996). Often niche-models give a probability of how likely it would be for a species to survive in a particular location. The occupancy envelope is similar to niche-models except that it takes into account spatial spread through time.

Concordance probability

One method of assessment that could be implemented is the concordance probability (CP). The CP compares all possible pairs of observations involving one presence and one absence and is defined as the proportion of comparisons in which the predicted probability is higher for the observed presence than the observed absence (Hanley & McNeil 1982).

The CP however makes an assumption about the relative costs of false positives and false negatives. If these costs can be made explicit then a cost-minimization approach could be included that has the advantage of ensuring a connection to management practice (Zweig & Campbell 1993).

Levels of surveillance

Often occurrence data are biased according to the types of areas that are monitored and this data often only includes presence data. The assessment module could include regions with high and low levels of surveillance and weight the occurrence sites appropriately, since high surveillance areas that don't have occurrence sites would likely have a greater probability of being a true absence.

B.1.6 Possible extensions for the rate of spread module

95% of patches behind line

The mean distance measured by the **ros** (rate of spread) module is not itself a front - but a crude indication of the distribution's progression. Perhaps more useful is the distance below which 95% (or another user-specified percentage) of the distribution can be found.

Alternative for the density boundary heuristic

Currently the density boundary heuristic uses concentric rings (or slices if going in a single direction) of a certain width. A more gradual or smoother method would be to overlap the rings, so that more measurements of density are taken but with some of the measurement area shared between neighbouring rings.

B.2 Future module ideas

In addition to the sections on possible extensions for each modules, there several distinct ideas as well as generic modifications that could be implemented.

B.2.1 Vector maps

Firstly, **MDiG** is a generic framework and although the previously mentioned modules have focussed on raster maps, equivalent modules that work on vector maps could be implemented.

Travelling down this route of using vector maps could lead to metapopulations being simulated by using shape files where each vector polygon represents an individual sub-population. The patch area, perimeter, and distance between patches, could be calculated using the in-built utilities of GIS. This would however be difficult because edges would need to be split and modelled "like a bubble" with converging foci occurring, especially late in invasions (Mack 1981, Usher 1987). There would also be a uniform distribution of the population across the entire polygon. Thus using vectors to represent sub-populations is only really effective for presence absence simulations, or where the distribution can be assumed to be uniformly distributed.

Alternatively vector maps could be used to simulate an individual based approach, with vector points representing each individual and modules could be developed to model local interactions between individuals and their environment.

Individual based models (IBMs) are also be used for modelling the dispersal of organisms. Gardner & Gustafson's (2004) model was a spatically explicit and heterogenuous IBM that investigated the persistance of reintroduced American martens (*Martes americana*). However, for species with large population sizes, such as ants and other invertebrates, and at large spatial scales, such models can become unfeasible with current computing power. One solution is to abstract populations and represent them as a single 'individual' essentially as was done in the Argentine ant study where an 'individual' represents any population that exists within a 150 m by 150 m area.

B.2.2 Gravity models

A module implementing the behaviour of gravity models could allow estimation of long-distance dispersal between discrete points in heterogeneous landscapes, as has been done for zebra mussel (*Dreissena polymorpha*) in inland lakes of the USA (Bossenbroek et al. 2001).

B.2.3 Human transport networks

Long distance dispersal along networks such as roads, railways, or rivers (all of which are commonly available as vectors in GIS) could be implemented as a particular dispersal module in **MDiG**. For many invading species, the key to understanding their dispersal is through measurements of human transport processes, especially for the movement of Argentine ants (Suarez et al. 2001) and disease spors (e.g.*Phytophthora lateralis*, Jules et al., (2002)) by cars and trucks, or, for example, zebra mussels by boats (Buchan & Padilla 1999, Bossenbroek et al. 2001).

Appendix C

New Zealand *L. humile* model

C.1 DispersalML definition

Dispersal simulations in the MDiG framework are defined using the XML model specification format (DispersalML, or Dispersal Markup Language). This indicates to the simulation controller what map or incursion points are to be used for the initial species distribution. For each lifestage, a number of events are described that modify the species distribution map in some way. Each time-step within the model will execute the events in each lifestage, the order they occur is determined by the phenology of lifestage which can be specified as either a integer value or a map of lifestage timing.

A more complete and up-to-date description of the format is available from the MDiG website (`http://fruitionnz.com/mdig/`).

C.1.1 DispersalML for *L. humile* model

```
<model>
    <user><email>pittj2@lincoln.ac.nz</email></user>

    <name> lhumile </name>
    <description> Argentine Ant </description>

    <random>
        <!-- Starting seed for replicates -->
        <initialSeed> 1 </initialSeed>
        <!-- Number of times to repeat simulation for each combination of
            parameters -->
```

```
    <replicates> 10 </replicates>
    <!-- The last state of the random number generator in case the simulation
        didn't
     finish last time -->
    <lastState>25</lastState>
</random>

<regions>
    <!-- Simulation may be run at different resolutions or in different regions
         -->
    <region id=”a” name=”argant_ni”>
    <!-- id is for referral from replicates below, name loads region from file
        -->
        <!-- Resolution to run simulation in -->
        <resolution>150</resolution>
        <!-- Extent of the region in coordinates -->
        <!--<extents north=”6609068” south=”6354012” west=”2568074” east
            =”2829814”/>-->
        <!-- for displaying behind the distribution when display flag is on -->
        <background>
            <!-- accept a variety of map expressions - use same type for other
                places
            a map can be used. -->
            <map>aa_landcover</map><!--<map>aa_background</map>-->
            <!--or <value></value>-->
        </background>
    </region>

</regions>

<period>
    <!-- Time period to run the simulation for -->
    <startTime> 1990 </startTime> <!-- refer to a variable id or use value -->
    <endTime> 2006 </endTime>
    <!--<step> <value> 1 </value> </step>--> <!--Optionally indicate that
        simulation
     steps by step amount-->
</period>

<output>
    <!-- Define how to output results -->
    <!-- Output a map every year -->
```

```
    <raster>
        <interval>1</interval>
        <lifestage>all</lifestage>
    </raster>

    <!-- Output a PNG graphic with overlayed data -->
    <!--<png>
        <interval>1</interval>
        <showTime>true</showTime>
        <showGrid>false</showGrid>
    </png>-->
</output>

<lifestages>
    <!-- Define each lifestage -->
    <lifestage name=”all” populationBased=”false”>
        <!-- Define initial population for region -->
        <initialDistribution region=”a”>
            <!--<map name=”randomsites3” lifestage=”ant” boolean=”true”/>-->
            <sites>
                <s x=”2671900” y=”6474300”/>
                <s x=”2674500” y=”6471400”/>
                <s x=”2672400” y=”6474500”/>
                <s x=”2671900” y=”6474400”/>
                <s x=”2672400” y=”6474500”/>
            </sites>

            <!-- Note: mapcalc can be refreshed if it is a randomly created map
                -->
        </initialDistribution>

        <!-- Define when the lifestage occurs
             can be a value or a map -->
        <phenology region=”a”>
            <value> 1 </value>
            <!--<map> test </map>-->
        </phenology>

        <!-- list dispersal events and there parameters -->
        <event name=”r.dispersal.neighbour”>
            <param name=”shape”>
                <value>186</value>
```

```
    </param>
    <param name=”radius”>
        <value>1</value>
    </param>
    <ifNotPopulationBased>
        <flag name=”b”/>
    </ifNotPopulationBased>
</event>

<event name=”r.dispersal.kernel”>
    <param name=”kernel”>
        <value>cauchy</value>
    </param>
    <param name=”d_b”>
        <variable id=”scale”>
            <value>72700</value>
            <value>83673</value>
            <value>94647</value>
        </variable>
    </param>
    <param name=”frequency”>
        <variable id=”freq”>
            <value>0.19926</value>
            <value>0.29779</value>
            <value>0.42798</value>
        </variable>
    </param>
    <param name=”seed”>
        <seed/>
    </param>
    <ifNotPopulationBased>
        <flag name=”b”/>
    </ifNotPopulationBased>
</event>

<event name=”r.dispersal.survival”>
    <param name=”survival”>
        <value>aa_survival@lhumile</value>
    </param>
    <param name=”seed”>
        <seed/>
    </param>
```

```
                <ifNotPopulationBased>
                    <flag name=”b”/>
                </ifNotPopulationBased>
            </event>

            <!-- Analysis here is for analysis at the end of each year/timestep -->
            <analyses>
                <analysis name=”r.stats”>
                    <param name=”input”>
                        <previousMap/>
                    </param>
                    <flag name=”c”/>
                    <flag name=”n”/>
                    <output>
                        <file generate=”true” ext=”.dat” append=”true” date=”true
                            ”/>
                    </output>
                </analysis>
            </analyses>

        </lifestage>

    </lifestages>
</model>
```

C.2 New Zealand *L. humile* occurence data

Table C.1: Argentine ant occurence data for New Zealand

Crosby district	Location	Easting	Northing	Year
AK	Auckland, 35 Tawa Rd,, Auckland	26462	64885	2001
AK	Auckland regional Botanic Gardens, north entrance, Totara Heights, Auckland	26796	64646	2002
AK	Bishop Park, Kopika and Kohu Rd Corner, Titirangi, Auckland	26573	64726	2002
AK	Opanuku pipeline track, Mountain Rd, Hendersen Valley, Auckland	26502	64770	2002
AK	Estuary Views, Maungamaungaroa Creek, Somerville, Auckland	26836	64749	2002
AK	Karekare Rd, Karekare, Auckland	26421	64679	2002
AK	Titirangi Rudoplh Steiner school, Woodlands Park, Auckland	26564	64711	2002
AK	Bellbird Rise/Penguin Drive Reserve, Murrays Bay, Auckland	26657	64952	2002
AK	Centennial Place, Campbells Bay, Auckland	26672	64932	2002
AK	three streets backing onto Lynfield coastal walkway, Lynfield, Auckland	26631	64734	2002
AK	Sylvania Cresent, Lynfield, Auckland	26637	64727	2002
AK	walkway to Wattle Bay +ACY- Waikowhai Park Cape Horn Rd, Lynfield, Auckland	26646	64727	2002
AK	Hillsborough Reserve, Lynfield, Auckland	26664	64735	2002
AK	Waitea Rd housing, Muriwai, Auckland	26381	64840	2002
AK	Cassino St to Plymouth Reserve walkway, Belmont, North Shore, Auckland	26700	64862	2002
AK	Ngataringa Park walkway, Stanley Bay, North Shore, Auckland	26705	64855	2002
AK	Rotary Res., Lake Rd & Onepoto Stream, Northcote, North Shore, Auckland	26660	64867	2002

AK	Heath reserve to Arahia St. trk, Nothcote Lagoon, North Shore, Auckland	26665	64874	2002
AK	Kauri Park, Chatswood, North Shore, Auckland	26623	64868	2002
AK	Kaipatiki & Beach haven Rd, Beach Haven, North Shore, Auckland	26624	64893	2002
AK	Piha, North Piha Rd, North Side of Bridge over Marawhara Stream, Auckland	26410	64723	2002
AK	Piha, North Piha Beach, Hill side of Rd, Auckland	26406	64728	2002
AK	The Drive reserve, Titirangi, Auckland	26574	64723	2002
AK	Waima Cresent, Titirangi, Auckland	26562	64723	2002
AK	Tirirangi, Lower Woodlands Park Rd, Auckland	26558	64722	2002
AK	Davis Bay track, Paturoa Rd, Titirangi, Auckland	26577	64708	2002
AK	Waiwera Palce, Waiwera, Auckland	26633	65158	2002
AK	Waikowhai Park, around playground, Waikowhai, Auckland	26649	64729	2002
AK	Waikowhai Park, Faulkner Bay boat ramp, Waikowhai, Auckland	26652	64728	2002
AK	The Warehouse Distribution Centre 92 Langley Rd, Wiri, Auckland, Auckland	26763	64642	2003
AK	92 Langley Rd, Wiri, Manurewa, Auckland	26763	64642	2003
AK	Waimauku School, State highway 16, Auckland	26437	64922	2002
AK	Waimauku, 987 State Highway 16, Auckland	26437	64922	2002
AK	Waimauku, 30 Amber Place, Auckland, Auckland	26434	64913	2001
AK	Waimauku, 985 State Highway 16,, Auckland	26445	64917	2000
AK	Piha Beach, start of Marawhara walk, under grass on bridge concrete, Auckland	26410	64723	2000
AK	Piha, 11 Beach Valley Rd, Auckland	26414	64707	2001

AK	Piha, 83 Marine Parade,, Auckland	26411	64715	2001
AK	Piha, Auckland	26412	64705	1999
AK	Piha, sand at top of beach, Auckland	26411	64713	1999
AK	Muriwai Beach, Auckland	26380	64839	2001
AK	Omaha, Auckland	26697	65396	1998
AK	Warkworth, 10 Pound St, , Auckland	26595	65322	2001
AK	Warkworth, Auckland	26598	65320	2001
AK	Forrest Hill, 34 Merryfield Ave, Ak, Auckland	26663	64914	2001
AK	Whangaparaoa Rd, Auckland, Auckland	26621	65086	2001
AK	Tiritiri Matangi Island, Northeast Bay, Auckland	26788	65104	2001
AK	Murrays Beach, 61 Penguin Drive,, Auckland	26658	64952	2001
AK	Glenfield, 88 Lynn Rd, Auckland	26624	64904	2000
AK	Tiritiri Matangi Island, by jetty, Auckland	26794	65098	2000
AK	Tiritiri Matangi Island, North East Bay, Auckland	26793	65103	2000
AK	Cuthill area, Wickham Ave, near Chester Park, Auckland	26620	64920	2000
AK	Milford, 1/17 Saltburn Rd, Auckland	26684	64907	2000
AK	Dairy flat Aerodrome, Auckland	26583	65034	2001
AK	Kumeu, 111 Riverland Rd, Auckland	26512	64916	2002
AK	37 Marsden Ave, Mt Eden, Auckland	26665	64773	2003
AK	Titirangi School, Auckland	26576	64725	1993
AK	Lynfield, Boundary Rd, Auckland	26630	64744	1995
AK	New Lynn, cnr Willerton and Hutcheson Ave, Auckland	26597	64745	1995
AK	New Lynn, Hutcheson Ave, Auckland	26597	64745	1996
AK	Newton, Johnston's Coachlines, Auckland	26667	64808	2001
AK	Newmarket, 371 Khyber Pass, Auckland	26684	64801	2001
AK	Birkenhead, 33 Hebe Place, Auckland	26625	64863	2005
AK	Bucklands Beach, Auckland	26800	64800	1992
AK	Mt Roskill, Auckland	26650	64750	1992
AK	Onehunga, Auckland	26690	64740	1992

AK	Parnell, Auckland	26694	64814	1992
AK	Rumuera, Auckland	26710	64790	1992
AK	Yandarra Primary School, 226 Baids Rd Otara, Auckland	26775	64690	2001
AK	Henderson, 5 Claret Pl, Auckland	26546	64785	2000
AK	Auckland Airport, Auckland	26695	64650	2001
AK	Mangere, 164 Montgomerie (sp?) Rd, Auckland	26684	64677	2001
AK	Northcote, Howard St, near Onepoto demain, Auckland	26663	64870	2000
AK	Glen Eden, Westech Pl, Auckland	26584	64758	2000
AK	15 Rewene Rd, by tennis court, Auckland	26659	64861	2000
AK	Birkdale, 22 Bishopgate St, Auckland	26625	64876	2000
AK	Titirangi, Daffodil St, suburban garden with some large natives, Auckland	26574	64732	2000
AK	Te Whau point, Blockhouse Bay, Auckland	26620	64731	2000
AK	Mt Smart, Auckland	26719	64743	1990
AK	Otahuhu, Nickay St, Auckland	26745	64714	1990
AK	Hillsborough, 14 Goodall St, Auckland	26667	64736	2001
AK	Mt Smart, O'Rourkes Rd, Auckland	26724	64745	1990
AK	Onehunga, Auckland	26700	64730	1996
AK	Onepoto Domain, Auckland, Auckland	26666	64868	2001
AK	One Tree Hill, 36a Waiohua Rd, Auckland, Auckland	26707	64763	2001
AK	Onehunga, Mt Smart Stadium, Auckland	26719	64744	1990
AK	Onehunga, O'Rourkes Rd, Auckland	26724	64745	1990
AK	Titirangi, Bishop Park Scenic Reserve, Atkinson Rd, Auckland	26577	64726	2000
AK	Blockhouse Bay Beach reserve, Auckland	26622	64736	2000
AK	St Heliers, Dingle Dell reserve , Auckland	26759	64809	2000
AK	Mt Eden Domain, Auckland	26677	64790	2000
AK	Titirangi, Woodlands Park school, Auckland	26556	64717	2000
AK	Three Kings, Big King reserve	26668	64763	2000

AK	Waiata reserve (Remuera), Auckland	26710	64790	2000
AK	Waiheke Is, Palm Beech, new house and beech, Auckland	26926	64892	2000
AK	Waiheke Is, 8 Gillies Rd, Auckland	26926	64890	2000
AK	Waiheke Is, Palm Beech, Auckland	26926	64890	2000
AK	Mangatawhiri, Hotel du vin, site 6, on edge of small remnant, Auckland	27010	64437	2000
AK	Mangatawhiri, Hotel du vin, Auckland	27010	64437	2000
BP	Tetana St, Tauranga			2003
BP	Sulphur Point, Tauranga	27897	63878	2003
BP	Papamoa Tavern, Domain Rd, Papamoa Beach, Tauranga	27998	63832	2003
BP	Mt Maunganui, Bayfair infestaton, Tauranga	27947	63868	2001
BP	Mt Maunganui, North Port infestation, Tauranga	27913	63904	2001
BP	Tauranga, Otemoetai Road., Tauranga	27871	63873	2001
BP	Mt Maunganui, Newton St infestation, Tauranga	27927	63886	2001
BP	Mt Maunganui, Tui St infestation, Tauranga	27932	63888	2001
BP	Mt Maunganui, Port infestation, Tauranga	27913	63890	2000
BP	Mt Maunganui, Te Maunga infestation, Tauranga	27953	63855	2000
BP	Mt Maunganui, Totara St, Tauranga	27914	63888	1992
BP	Mt Maunganui, Totara St,, Tauranga	27914	63888	2001
BP	Mt Maunganui, Farm St,, Tauranga	27922	63905	2001
BP	Mt Maunganui, Port offices,, Tauranga	27913	63902	2001
BP	Camp ground Rotorua, Rotorua	27952	63332	2003
CL	Adams Rd, Thornton Bay, Coromandel	27349	64569	2002
CL	Ngar, 17 Patui Ave, Coromandel	27344	64560	2001
CL	Totara, Coromandel	27372	64452	2001
CL	Coromandel, 270 Coast Rd, RD5 Thames, Coromandel	27348	64568	2000
GB	Lake Repongaere, Gisborne	29340	62783	2002

HB	Hastings, juice producer, Hastings	28400	61670	2001
HB	Hastings, Hastings	28400	61670	2001
HB	Hastings, Kitchener St,, Hastings	28400	61670	2001
HB	Hastings, Omahu Rd,, Hastings	28400	61670	2001
HB	Hastings, Omahu Rd,, Hastings	28400	61670	2001
HB	Hastings, Cnr Tomoana Rd and Williams St,, Hastings	28400	61670	2001
HB	Hastings, Warwick Pl,, Hastings	28400	61670	2001
HB	Hastings, Cnr Mayfair Ave and Caroline Rd,, Hastings	28400	61670	2001
HB	Hastings, Harlech St,, Hastings	28400	61670	2001
HB	Hastings, HB A+-P show grounds,, Hastings	28392	61661	2001
HB	Hastings, Coventry Rd,, Hastings	28400	61670	2001
HB	Hastings, Tomoana Rd,, Hastings	28412	61686	2001
HB	Hastings, Williams St,, Hastings	28400	61670	2001
HB	Hastings, Rangiora St,, Hastings	28400	61670	2001
HB	Napier, Napier	28449	61817	2001
MC	Hornby Mall, Honby, Christchurch	24715	57405	2004
MC	Mainfreight International, McAlpine Rd, Christchurch	24745	57405	2002
MC	MC, Riccarton, Christchurch	24775	57415	2000
ND	Whangarei, 52 Sherwood Rd,, Whangarei	26339	66048	2001
ND	George Point Rd, Sherwood Rise, Whangarei, Whangarei	26332	66049	2002
ND	Kauika Rd West, Whangarei, Whangarei	26290	66077	2002
ND	Bayleys Beach, Bayleys Beach	25776	65834	2002
ND	forest remnant at end of Seacrest Rd, Langs Beach, Northland	26495	65711	2002
ND	Whangarei, Riverside Drive,, Whangarei	26322	66068	2001
ND	Whangarei, Handforth St,, Whangarei	26338	66028	2001
ND	Whangarei, Goodwin St,, Whangarei	26343	66042	2001
ND	Whangarei, Station Rd,, Whangarei	25872	65849	2001
ND	Whangarei, Handforth St,, Whangarei	26338	66028	2001

ND	Brookers Bay, Coastal, Brookers Bay			2002
ND	Hatea Drive, Whangarei	26306	66087	2002
ND	Ahipara Beach, Northland	25250	66710	2002
ND	Tokerau Beach, Northland, Northland	25449	67022	2004
ND	SH1 Awanui, Kaitaia	25343	66837	2002
ND	Worth St, Kaitaia	25346	66766	2002
ND	Kaitaia, Puckey Ave,, Kaitaia	25346	66766	2001
ND	Kaitaia, Puckey Ave, Pak 'n' Sav, Kaitaia	25346	66766	2001
ND	Mangonui, beaten from boneseed, Mangonui	25592	66905	2000
ND	Kaitaia, Kaitaia	25346	66766	2001
ND	Kerikeri, beaten from banana passionfruit, Kerikeri	25938	66674	2000
ND	Jacaranda Place, Kerikeri	25973	66634	2002
ND	Swiss Chalet Motel/ Dr Surgery Bay View Rd, Paihia	26101	66570	2002
ND	Hillcrest Rd, Kaikohe	25821	66436	2002
ND	Bayleys Beach, beaten from boneseed, Bayleys Beach	25775	65832	2000
ND	Dargaville, Dargaville	25899	65849	2001
ND	Dargaville, Dargaville	25873	65828	2001
ND	Dargaville, Plunket St,, Dargaville	25897	65850	2001
ND	Dargaville, Plunket/Ranfurly St,, Dargaville	25897	65850	2001
ND	Dargaville, Tirarau St,, Dargaville	25897	65850	2001
ND	Dargaville, Islington St,, Dargaville	25897	65850	2001
ND	Dargaville, Parore St,, Dargaville	25897	65850	2001
ND	Dargaville, Awakino Rd,, Dargaville	25897	65850	2001
ND	Dargaville, Victoria St,, Dargaville	25897	65850	2001
ND	Dargaville, Rimu St,, Dargaville	25897	65850	2001
ND	Dargaville, Ranfurly St,, Dargaville	25897	65850	2001
ND	Dargaville, Gordon St,, Dargaville	25897	65850	2001
ND	Dargaville, Ranfurly St,, Dargaville	25897	65850	2001
ND	Russell Heights Road, Russell	26126	66597	2002
ND	Long Bay (Long Beach?), Russell	26137	66596	2002

ND	Matauwhi Road, Russell	26136	66586	2002
ND	Paroa Bay, Russell	26171	66575	2002
ND	Mill Bay, Mangonui	25590	66901	2002
ND	Grey St East, Mangonui	25592	66905	2002
ND	Russell, Russell	26132	66590	2002
ND	Russell, 46 Oneroa Rd, Russell	26132	66590	2002
ND	Whangarei Heads, 71 Urquharts Bay Rd, Whangarei Heads	26494	65938	2002
ND	Brooker Bay is Whananaki North last little beach on road., Whananaki North	26424	66349	2002
ND	Rockells Bay, Whananaki, Whananaki	26422	66352	2002
ND	Ngunguru, Northland	26468	66177	2002
ND	Tutukaka, Tutukaka	26488	66201	2002
ND	Whananaki, Northland	26436	66314	2001
ND	Whangarei heads, 114 Reotahi Rd +AH4- 500 m from sea, Port, Hewlett St, grass verge, Whangarei Heads	26460	65966	2001
ND	Dutch Point Rd, Norfolk Ave, Reotahi Bay	26475	65966	2001
ND	Whangarei, Whangarei	26290	66075	2001
ND	Whangarei, Bank St outside building, Whangarei	26303	66084	2001
ND	Whangarei, Onerahi, in house, Whangarei	26341	66039	2001
ND	Whangarei, DOC field centre, Whangarei	26301	66067	2001
ND	Whangarei, Port, Hewlett St, grass verge, Whangarei	26324	66056	2000
ND	Whangarei, Fraser St, Port, outside ready mix, Whangarei	26324	66058	2000
ND	Whangarei, Port Rd, outside Astral Marine, Whangarei	26318	66058	2000
ND	Whangarei, Reyburn House, nests under stones along esturay edge, Whangarei	26306	66077	2000
ND	Whangarei, Dent St,, Whangarei	26302	66078	2001
ND	The Heights Rd, Reotahi Bay, Whangarei	26461	65964	2002

ND	DoC carpark and surrounds, Urquharts Bay, Whangarei, Whangarei	26493	65937	2002
ND	Marsden Point oil refinery, Marsden Point	26456	65944	2002
ND	Ruakaka, beach and dunes around township, Ruakaka	26416	65867	2002
ND	Onerahi, Whangarei	26336	66029	2002
ND	Whangarei, Port Rd,, Whangarei	26318	66058	2001
ND	Whangarei, Hewlet St,, Whangarei	26323	66057	2001
ND	Maungatapere, Whangarei	26187	66043	2004
ND	Maungaturoto & Bickerstaff Rd, Northland	26335	65641	2002
ND	Waipu township, Waipu	26413	65785	2002
ND	Langs Beach, Langs Beach	26491	65715	2001
ND	Mangawhai Heads, Mangawhai	26532	65670	2002
ND	Whangaparaoa, Whangaparaoa	26665	65058	2001
ND	Mcleods Bay, McLeods Bay	26998	64821	2002
NN	112 Arapiki Rd, The Ridgeway, Nelson, Nelson	25308	59885	2004
NN	66 Waterhouse Street, Nelson, Nelson	25309	59893	2004
NN	crn Vickerman and CarKeek st, Port Nelson, Nelson	25333	59937	2001
NN	Nelson, Vickerman St, port, edge of infestation near slipway, Nelson	25335	59941	2001
NN	Port crn Collins and Haven Rd, Nelson	25328	59937	2001
NN	Rogers St Port Nelson, Nelson	25333	59939	2001
NN	Graham St, Port Nelson, Nelson	25330	59937	2001
NN	crn Rodgers and Graham St, Port Nelson, Nelson	25331	59940	2001
TO	22 Matai St, Taupo, Taupo	27787	62749	2003
TO	10 Fairview Drive, Taupo, Taupo	27787	62749	2003
WN	Paraparaumu Beach, 36 Arnold Grove, Paraparaumu	26779	60332	2004
WN	Wellington, 128 Jackson St, Petone, Wellington	26674	59963	2001

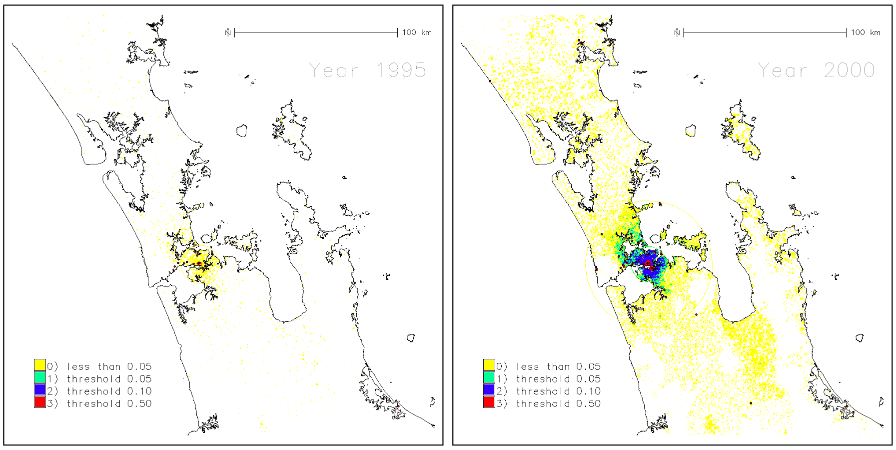
WN	Wellington, 9 Ngaio Rd,, Kelburn	26580	59897	2001
WN	Wellington, 24 Rimu Lane, Kelburn	26578	59896	2001
WN	Jackson St, Petone, Hutt Valley Polytech butchery school, close to Pack and Save, Wellington	26672	59963	2001
WN	Naenae Rd, Lower Hutt, Lower Hutt	26734	59990	2001
WN	Kelburn., Wellington	26575	59890	2000
WO	Waverly Ave, Morrinsville, Morrinsville	27334	63904	2001
WO	Raglan, 1a Daisy St, Raglan	26752	63763	2003
WO	Hamilton, Avalon Drive, Hamilton	27110	63772	2001
WO	Hamilton, Hamilton	27104	63778	2001
WO	Morrinsville, 301 Thames St,Kelly Stott, Morrinsville	27338	63911	2001
WO	Morrinsville (18/9/00) (wasteland beside dairy factory opposite Placemakers), Morrinsville	27338	63908	2000
WO	Morrinsville, Canada St,, Morrinsville	27334	63904	2001
WO	Morrinsville, Park St,, Morrinsville	27334	63904	2001
WO	Morrinsville, Park St,, Morrinsville	27334	63904	2001
WO	Morrinsville, Lindon Street,, Morrinsville	27334	63904	2001
WO	Morrinsville, Thames St,, Morrinsville	27338	63911	2001
WO	Morrinsville, Park Rd / Thames St,, Morrinsville	27334	63904	2001
ND	Aranga Beach	25626	66036	2006
ND	Waipoua Forest Headquarters	25611	66164	2006
ND	Bland Bay	26344	66503	2006
ND	Ocean Beach	26517	65947	2005
ND	Landing Road, Kerikeri (DOC office)	25990	66659	2004
ND	Kaka St, Whangarei (DOC office)	26302	66066	2001
ND	Tokerau Beach	25441	67018	2005
ND	Houhora & rubbish dump	25214	67111	2005
NN	Cambria St, Grove St, Tasman St, Central Nelson	25298	59904	2006

NN	Locking St, Renwick Place	25344	59930	2006
NN	Rotherham? (next to Nelson Golf Club)	25335	59918	2006

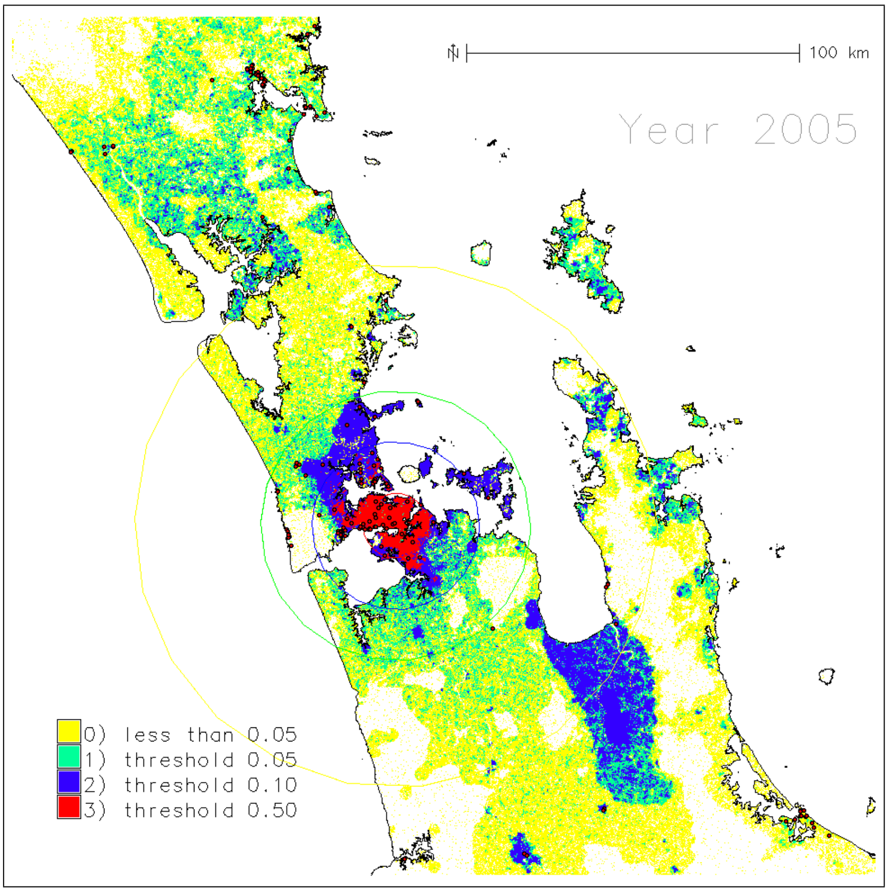
C.3 Simulation results

For the full set of results, including animated GIF files, see the CDROM accompanying this thesis.

Here images from years 1995, 2000, and 2005 are shown for each parameter combination. The region has been cropped to the upper North Island around Auckland.

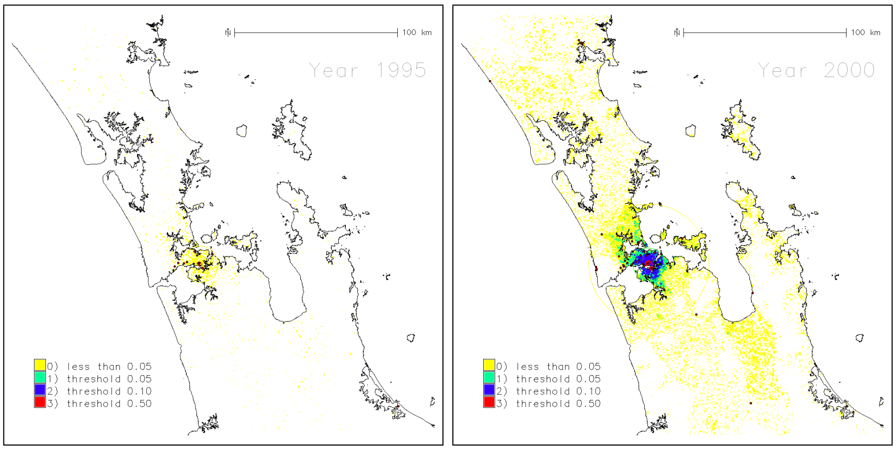


(a) 1995 (b) 2000

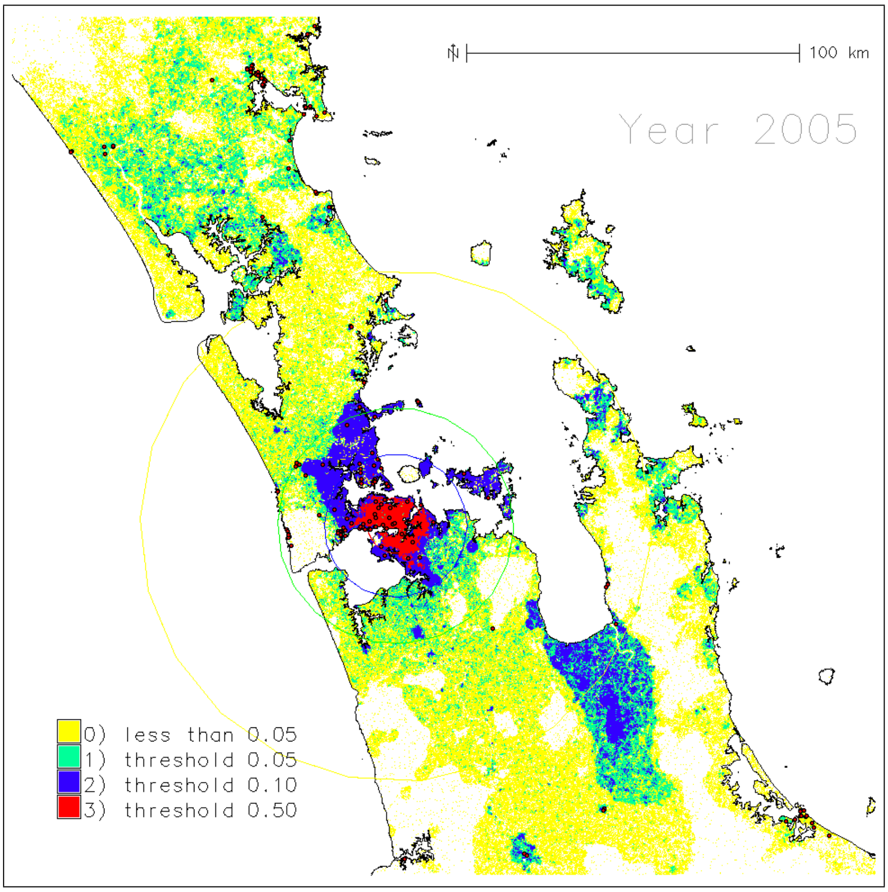


(c) 2005

Figure C.1: Probability envelopes for *L. humile* simulation with parameters, $f = 0.19926$ and scale= 72700

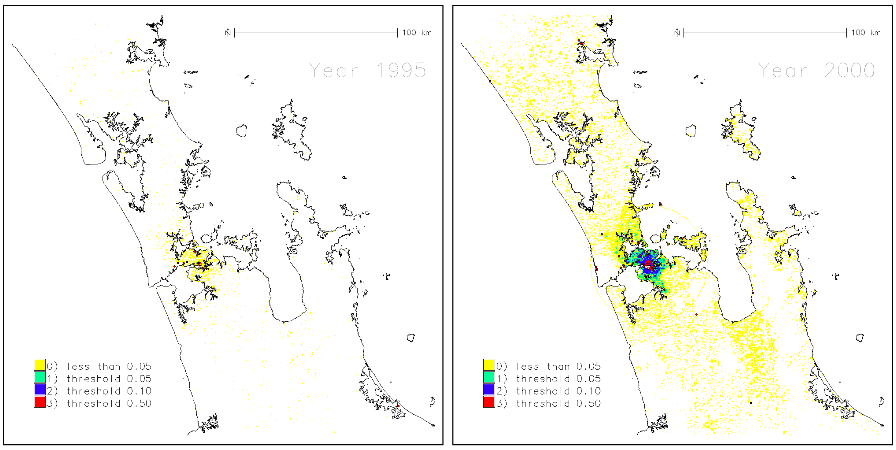


(a) 1995 (b) 2000

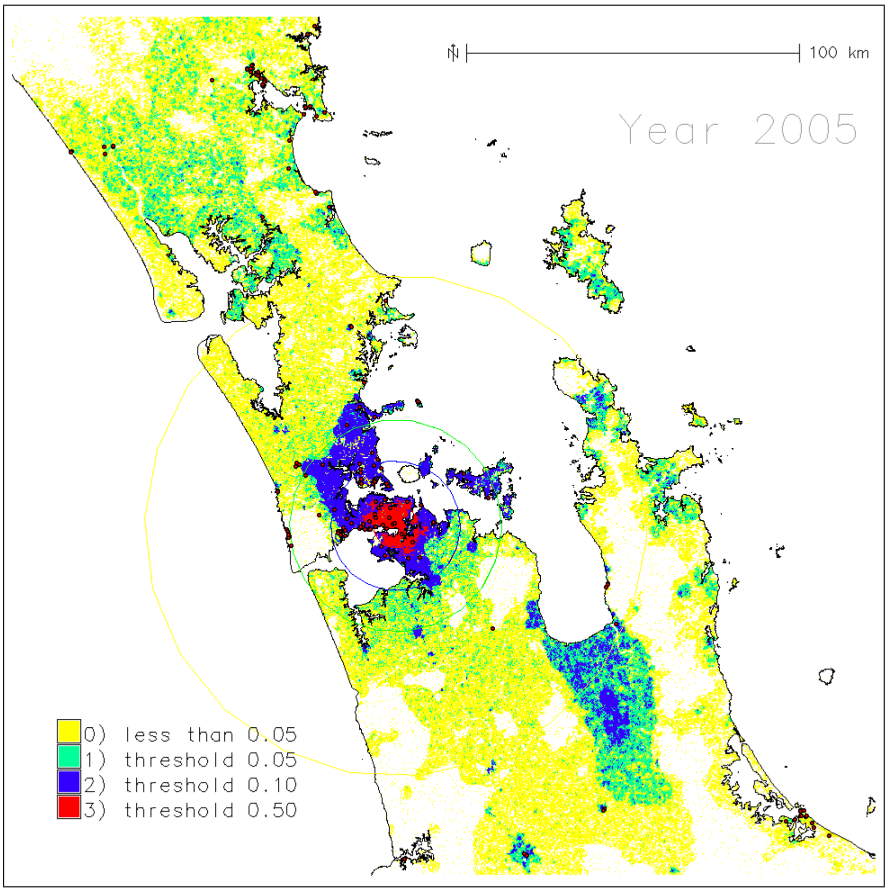


(c) 2005

Figure C.2: Probability envelopes for *L. humile* simulation with parameters, $f = 0.19926$ and scale$= 83673$

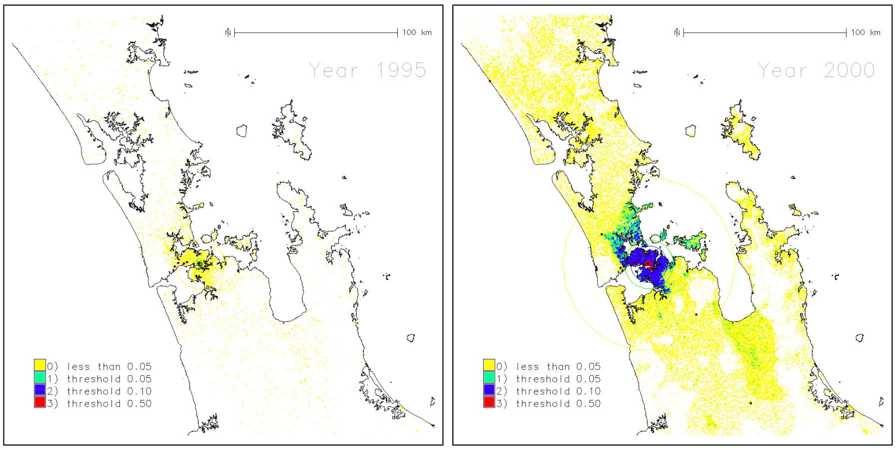


(a) 1995 (b) 2000

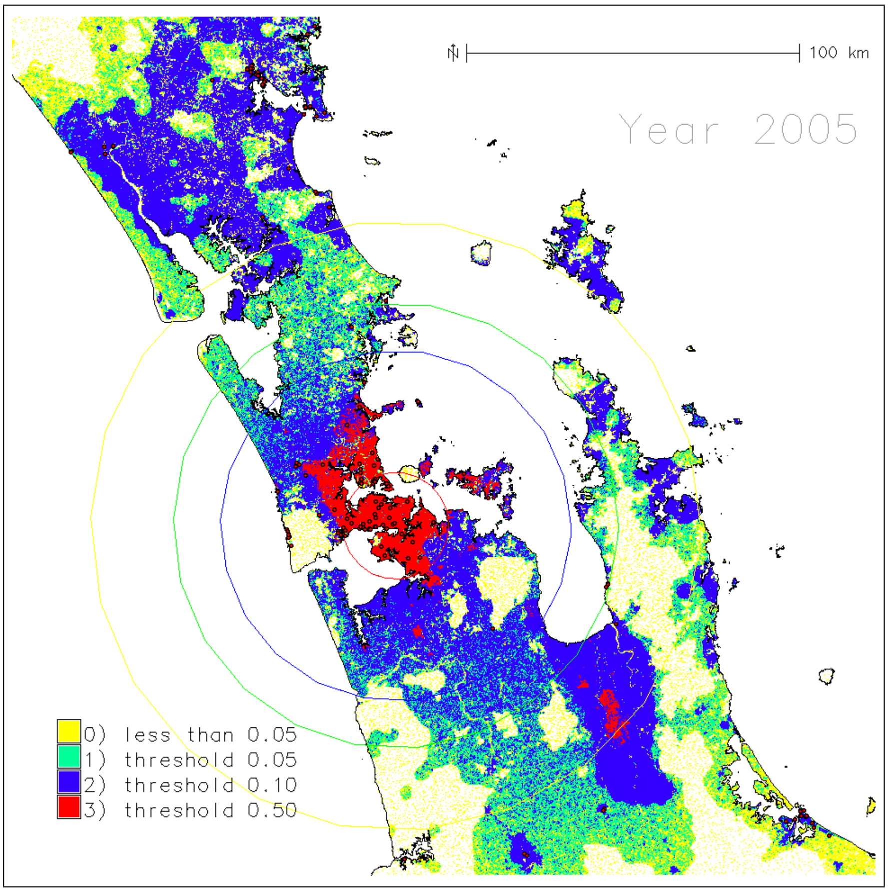


(c) 2005

Figure C.3: Probability envelopes for *L. humile* simulation with parameters, $f = 0.19926$ and scale$= 94647$

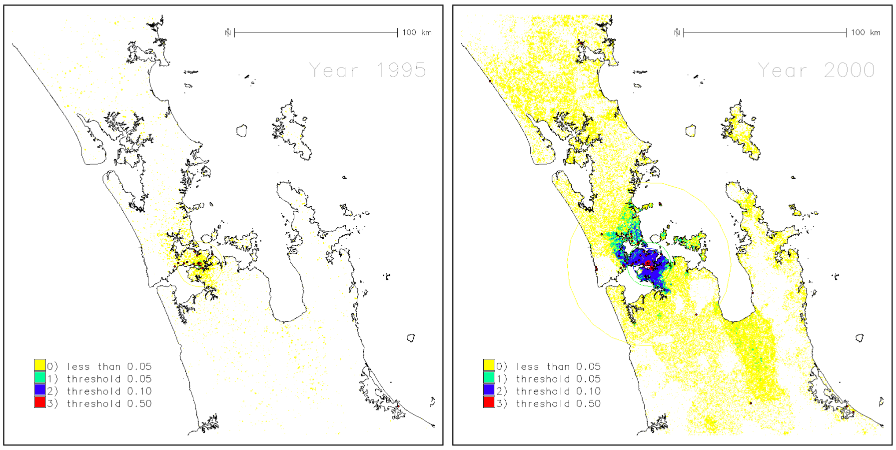


(a) 1995 (b) 2000

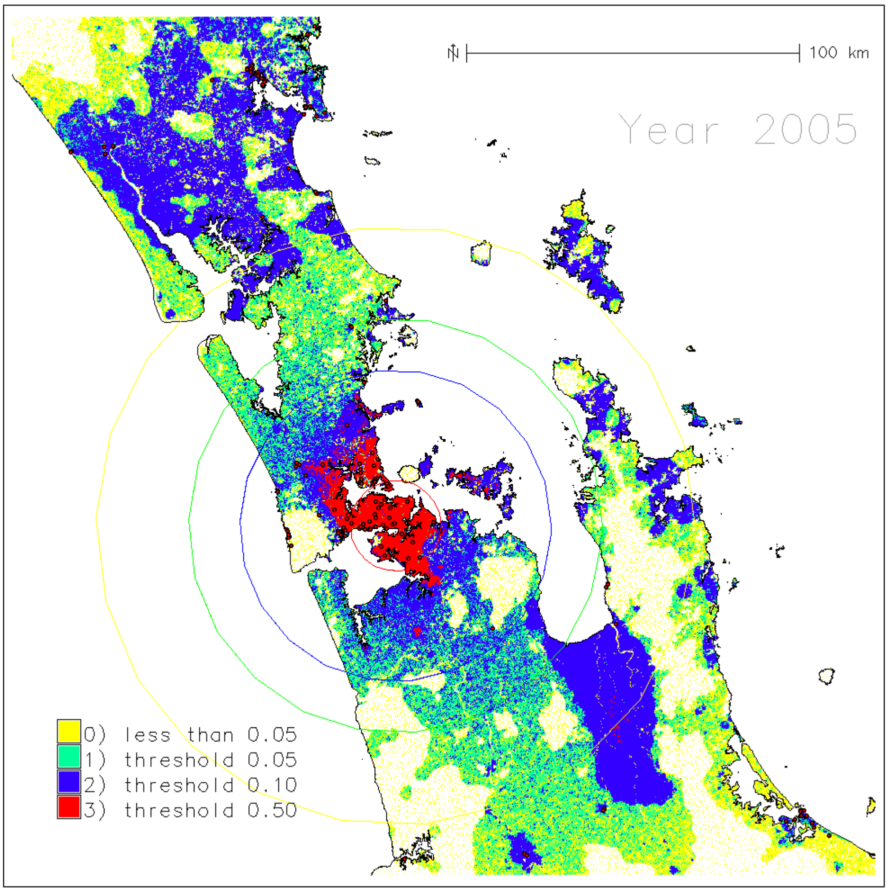


(c) 2005

Figure C.4: Probability envelopes for *L. humile* simulation with parameters, $f = 0.29779$ and scale= 72700

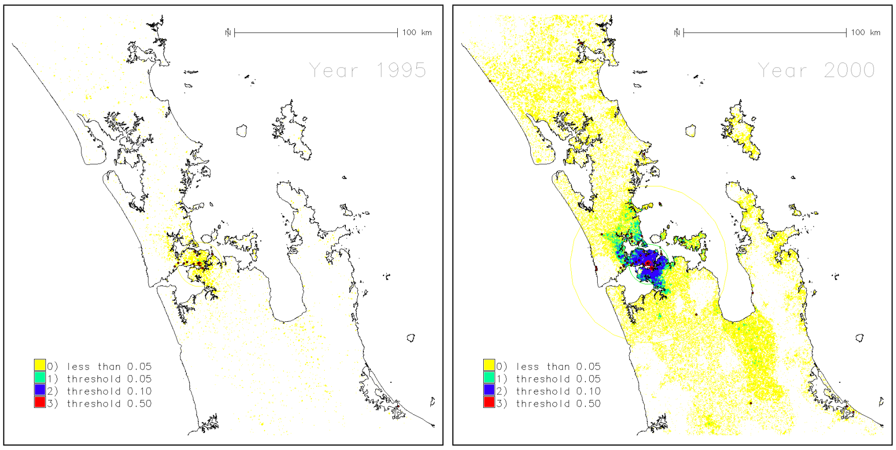


(a) 1995 (b) 2000

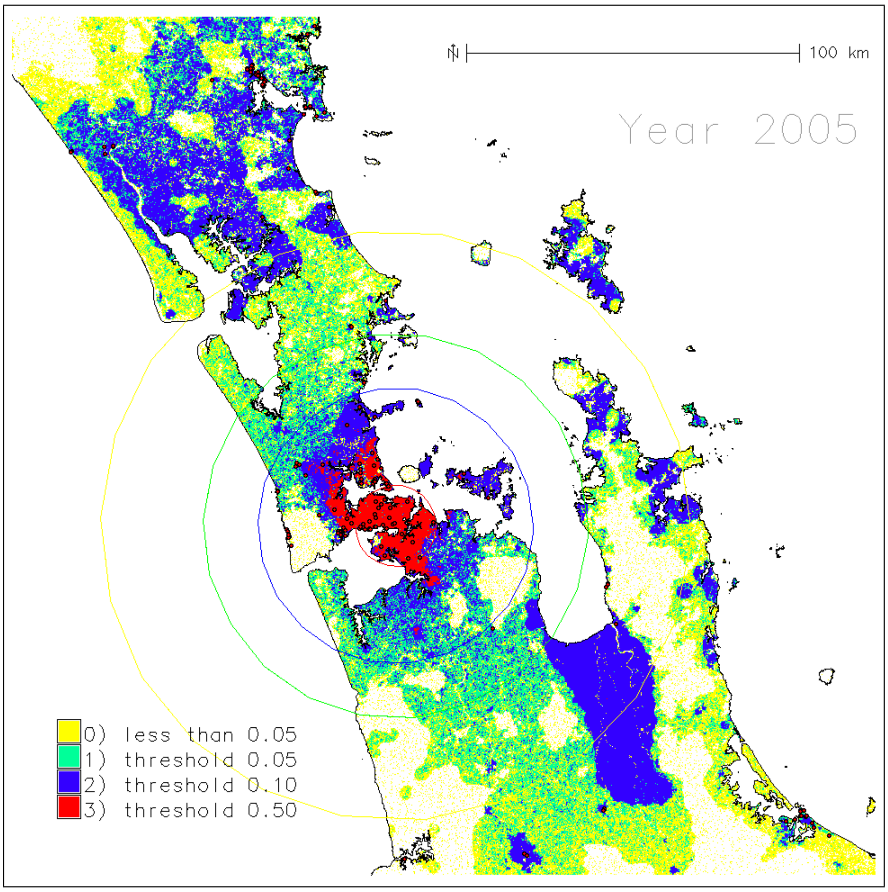


(c) 2005

Figure C.5: Probability envelopes for *L. humile* simulation with parameters, $f = 0.29779$ and scale= 83673

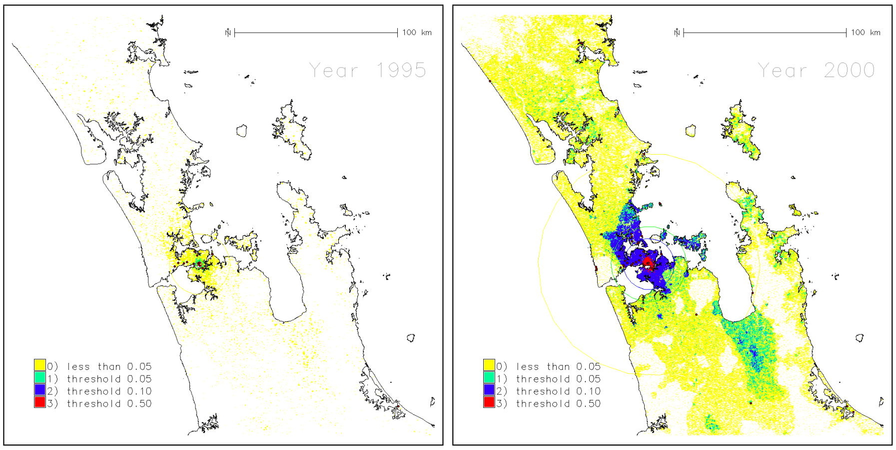


(a) 1995 (b) 2000

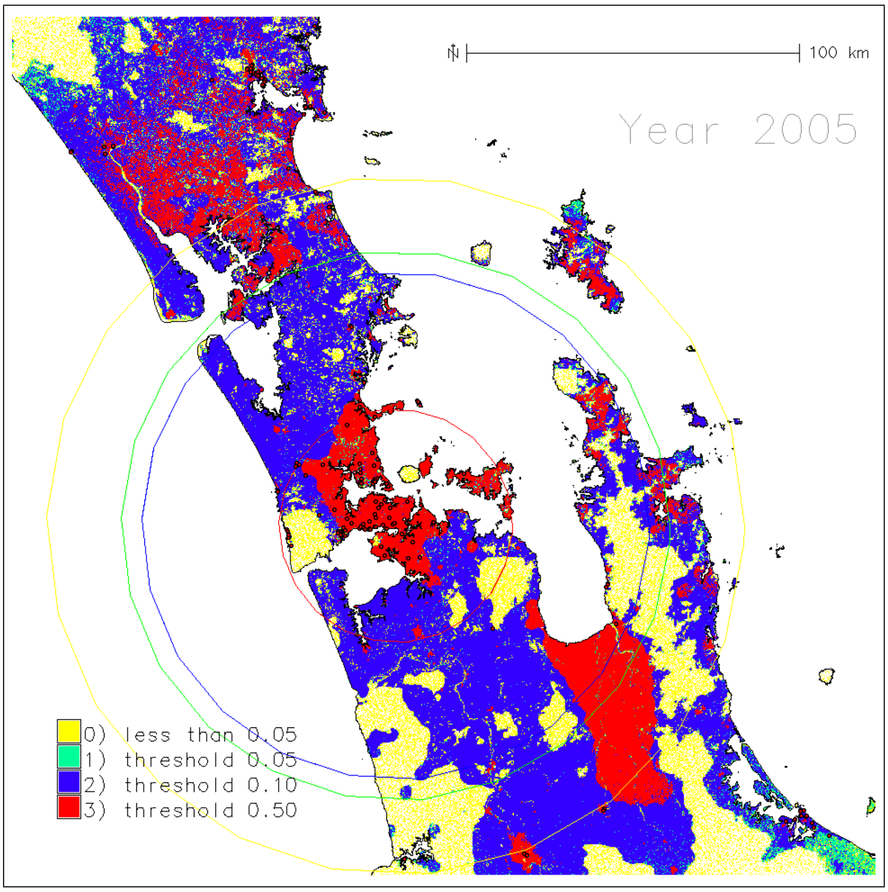


(c) 2005

Figure C.6: Probability envelopes for *L. humile* simulation with parameters, $f = 0.29779$ and scale= 94647

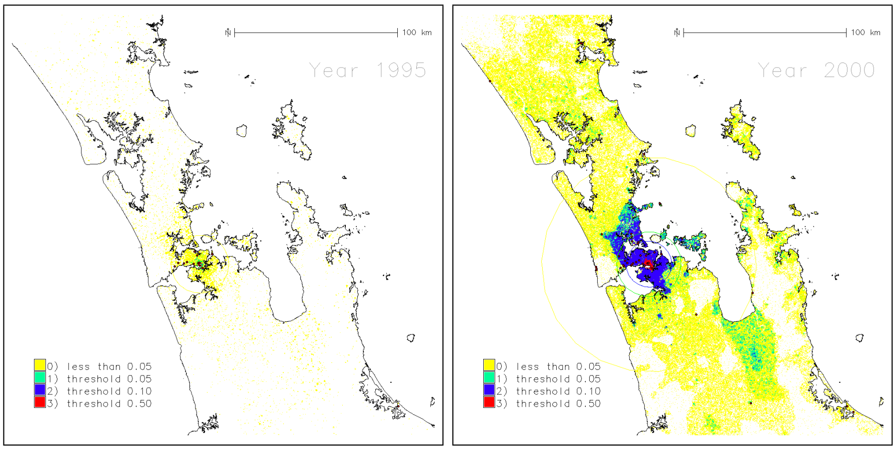


(a) 1995 (b) 2000

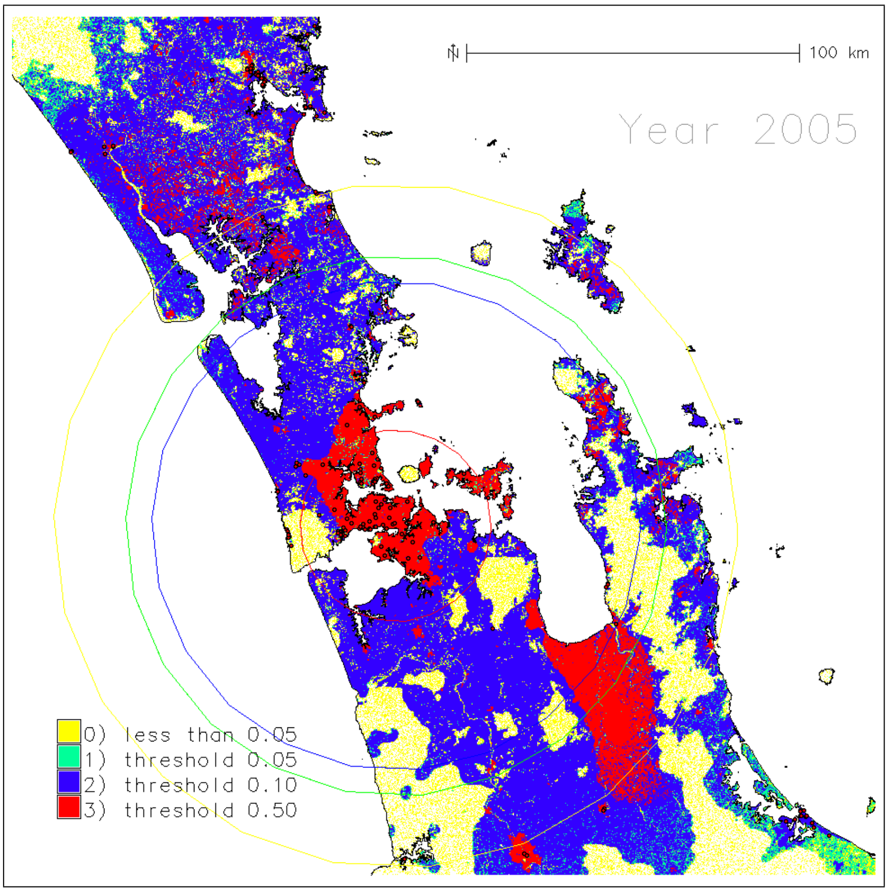


(c) 2005

Figure C.7: Probability envelopes for *L. humile* simulation with parameters, $f = 0.42798$ and scale= 72700

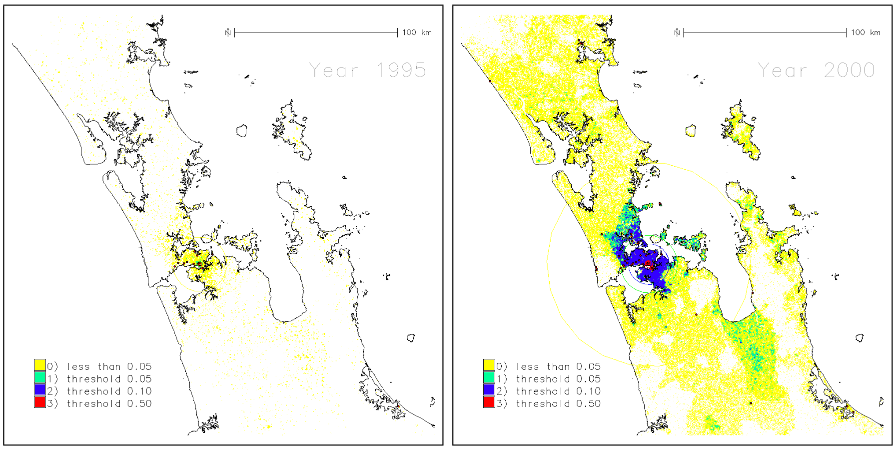


(a) 1995 (b) 2000

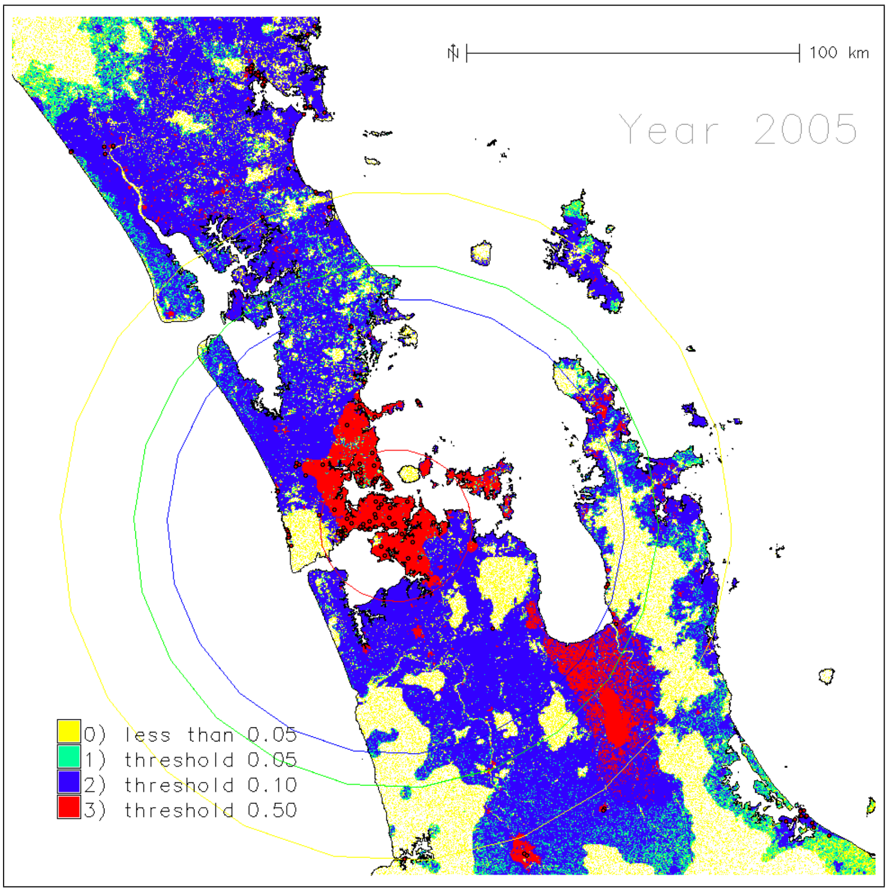


(c) 2005

Figure C.8: Probability envelopes for *L. humile* simulation with parameters, $f = 0.42798$ and scale= 83673



(a) 1995 (b) 2000



(c) 2005

Figure C.9: Probability envelopes for *L. humile* simulation with parameters, $f = 0.42798$ and scale= 94647