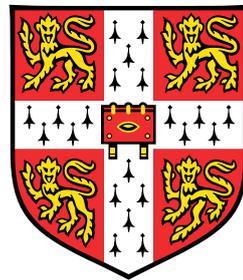


Modelling spatial strategies for the durable deployment of crop disease resistance



Benjamin Morgan Watkinson-Powell

Department of Plant Sciences

University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Declaration

This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements.

The contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification at the University of Cambridge, or at any other university or similar institution.

This dissertation does not exceed the prescribed word limit of 60,000 words (excluding bibliography, figures and appendices) as specified by the Degree Committee of the Faculty of Biology.

Benjamin Morgan Watkinson-Powell

December 2019

Abstract

Modelling spatial strategies for the durable deployment of crop disease resistance

Benjamin Morgan Watkinson-Powell

Maximising the durability of crop disease resistance genes in the face of pathogen evolution is a major challenge in modern agricultural epidemiology. Spatial diversification in the deployment of resistance genes, where susceptible and resistant fields are more closely intermixed, is predicted to drive lower epidemic intensities over evolutionary timescales. This is due to an increase in the strength of dilution effects, caused by pathogen inoculum challenging host tissue to which it is not well-specialised. The factors that interact with and determine the magnitude of this spatial effect are not currently well understood however, leading to uncertainty over the pathosystems where such a strategy is most likely to be cost-effective.

We initially use a spatially explicit model, incorporating seasonality and localised reservoirs of inoculum, to explore disease dynamics within landscapes containing a mixture of fields planted with either susceptible or resistant cultivars. When the spatial diversification of these fields is maximised, with lower aggregation of similar fields, the overall intensity of the landscape scale epidemic is reduced. The strength of this spatial effect however depends strongly on the pathogen dispersal characteristics, any fitness cost(s) of the resistance breaking trait, the efficacy of host resistance, and the length of the timeframe of interest.

The conclusions drawn from this initial work, about how multi-strain disease dynamics respond to the scale of spatial diversification in a multi-host landscape, allow us to construct a general spatially implicit model that captures these fundamental dynamics. This new model features a novel method for incorporating spatial structure using an intuitive spatial aggregation metric that can be easily estimated from spatially explicit landscape data. The model is simple enough to be amenable to mathematical invasion analysis, while being flexible enough that questions of resistance durability can be thoroughly explored. In particular, results demonstrating interaction between spatial heterogeneity and cultivar cropping ratio are

presented, an investigation that was not easily possible in our earlier more complex model. These results indicate that optimal spatial deployment strategies depend on a variety of factors, and may not necessarily be constant over time.

Overall, these models allow us to make general predictions of the types of system for which spatial diversification is most likely to be cost-effective, paving the way for potential economic modelling and pathosystem specific evaluation. In addition, this approach for capturing detailed spatial structure and multi-species interactions within simple mathematical models could be applied to a wide variety of ecological and evolutionary systems. This study highlights the importance of studying the effect of genetics on landscape scale spatial dynamics within host-pathogen disease systems, as well as providing new mathematical tools to do so.

I dedicate this thesis to my mum, who sadly left us during my PhD, but from whom I even now hear words of encouragement.

Acknowledgements

I would like to sincerely thank my supervisor Nik Cunniffe, for all the guidance, support and advice he has provided me with over the last four years. Your kindness, patience, brilliance, rigour and friendship have all been hugely appreciated. I am also extremely grateful to my co-supervisor Chris Gilligan for his council, wisdom and perspective.

I thank the BBSRC, who have funded my PhD through a DTP studentship, and the University of Cambridge Department of Plant Sciences, for hosting me throughout my studies. My thanks also go out to the many staff and students at Darwin College that I have come to know over the years, and who have always provided a welcome distraction, along with genuine support and friendship.

Thank you also to the many members, past and present, of the Theoretical and Computational Epidemiology group, and the Epidemiology and Modelling group in the Department of Plant Sciences. They have together created a fantastic environment in which to work, and which I will be sorry to leave. In particular, I would like to thank Elliott Bussell, Sarah Mitchell, Richard Stutt, Andrew Craig, Eleftherios Avramidis, Clare Allen, Sally Hames, Ruairi Donnolly, and Yevhen Suprunenko, for their help, encouragement and friendship.

Finally, I express my deepest thanks to my family and closest friends. This thesis is as much a product of your love and support as it is anything else.

Table of contents

List of figures	x
List of tables	xx
1 General Introduction	1
1.1 Background and motivation	1
1.2 Major gene resistance	2
1.2.1 Mathematical modelling of qualitative resistance durability	4
1.2.2 Effect of crop rotation	6
1.3 Spatial heterogeneity	6
1.3.1 Landscape scale spatial effects	7
1.4 Multi-allelic and quantitative resistance	9
1.5 Combining resistance traits	11
1.6 Thesis aims	12
2 When does spatial diversification usefully maximise the durability of crop disease resistance?	14
2.1 Introduction	14
2.2 Materials and Methods	17
2.3 Results	23
2.3.1 Effect of the kernel parameter η	24
2.3.2 Effect of the cost of the <i>rb</i> trait δ	25
2.3.3 Partial resistance	28
2.3.4 Effect of timescale	30
2.4 Discussion	32
2.5 Appendix 1 - Fully susceptible model	39

2.6	Appendix 2 - Interior landscape Edge/Area ratio	42
2.7	Appendix 3 - Additional figures	43
3	From explicit to implicit - capturing spatial structure	44
3.1	Introduction	44
3.2	Methods	49
3.2.1	Developing a suite of test landscapes of varying degrees of aggregation	49
3.2.2	Representing pathogen spread in a two-patch model	52
3.3	Results	58
3.3.1	Landscape generation	58
3.3.2	Estimation of p and its independence from the cropping ratio ϕ	60
3.4	Discussion	66
3.5	Appendix 1 - Proof that the model from Eqns (3.3) and (3.4) can reduce to a simple SIS model	72
3.5.1	Proof for unlinked patches ($p = 0$)	72
3.5.2	General proof (for any value of p)	74
4	How does spatial heterogeneity interact with the cropping ratio for achieving maximal resistance durability?	79
4.1	Introduction	79
4.2	Methods	82
4.2.1	The two-patch two-strain model	82
4.2.2	Invasion analysis	84
4.2.3	Model simulation	89
4.3	Results	90
4.3.1	Invasion analysis	90
4.3.2	Model simulation	97
4.4	Discussion	107
4.5	Appendix 1 - Finding the wt only equilibrium	115
4.6	Appendix 2 - Stability analysis for the model in which only resistant crop is planted	116
4.6.1	Case in which coexistence is not possible (i.e. $\gamma \neq (1 - \delta)$)	117
4.6.2	Case in which coexistence is possible (i.e. $\gamma = (1 - \delta)$)	119

5 General Discussion	121
5.1 Take home messages on resistance durability	122
5.2 Methodological developments	123
5.2.1 Assumptions and limitations	124
5.3 Applications and future work	125
5.4 Concluding remarks	128
References	129

List of figures

2.1	Example random field arrangements for two patch templates within the agricultural landscape.	23
2.2	Seasonal disease progress curves for each infection class (genotype combination) in an example landscape simulation replicate. The intensity of disease changes between seasons due to the changing contribution of primary infection from the local reservoirs, which in turn is based on the local epidemic intensities in the previous season. The two field types in the landscape are highly mixed (edge/area ratio = 1.4), the cost of the <i>rb</i> trait $\delta = 0.3$ and the <i>R</i> host is completely resistant to the <i>wt</i> strain $\gamma = 0$. Only 25 seasons are shown here so that the individual disease progress curves can see clearly seen.	24
2.3	The effect of the kernel parameter η on the relationship between landscape edge/area ratio and epidemic intensities. Epidemic intensities over evolutionary timescales are reduced at smaller scales of spatial heterogeneity. The spatial scale at which a spatial strategy is effective depends on the scale of pathogen dispersal. For all simulations presented here, the number of seasons $n_y = 40$, the cost of the <i>rb</i> trait $\delta = 0.3$ and the <i>R</i> host is completely resistant to the <i>wt</i> strain $\gamma = 0$.	25
2.4	The effect of the fitness cost of the resistance breaking trait (δ) on the relationship between landscape edge/area ratio and epidemic intensities. The strength of the spatial suppressive effect increases as the fitness cost of the resistance breaking trait is increased from low values. The pathogen genotype frequencies depend on the scale of spatial heterogeneity. For all simulations presented here, the number of seasons $n_y = 40$, the <i>R</i> host is completely resistant to the <i>wt</i> strain $\gamma = 0$ and the kernel parameter $\eta = 2$.	26

- 2.5 The effect of the fitness cost of the resistance breaking trait (δ) on the reduction in epidemic intensities from using a high compared to a low landscape edge/area ratio. **The strength of the spatial suppressive effect depends on the fitness cost of the resistance breaking trait, with a peak effect strength at an intermediate cost.** The average epidemic intensities resulting from the low and high ends of the E/A ratio scale are shown in (a), while the absolute differences between the results for these two E/A ratio values are shown in (b). Note that using the proportional differences in epidemic intensity produces a qualitatively similar pattern. Error bars show the 5th and 95th percentiles of the simulation replicates with stochastic landscape generation. For all simulations presented here, the number of seasons $n_y = 40$, the R host is completely resistant to the wt strain $\gamma = 0$ and the kernel parameter $\eta = 2$ 27
- 2.6 The effect of the susceptibility of the R host to the wt pathogen strain (γ) on the reduction in epidemic intensities from using a high compared to a low landscape edge/area ratio. **The strength of the spatial suppressive effect is decreased at lower efficacies of the host R gene, as long as the R gene is not so weak that the resistance breaking strain is not fit enough to invade the landscape.** The average epidemic intensities resulting from the low and high ends of the E/A ratio scale are shown in (c), while the absolute differences between the results for these two E/A ratio values are shown in (d). Note that using the proportional differences in epidemic intensity produces a qualitatively similar pattern. The relationship between the full range of landscape edge/area ratios and epidemic intensity, for each host/pathogen genotype combination, is shown for two values of γ in (a) and (b), and their corresponding positions along the x axis in (c) are marked. Error bars show the 5th and 95th percentiles of the simulation replicates with stochastic landscape generation. For all simulations presented here, the number of seasons $n_y = 40$, the cost of the rb trait $\delta = 0.3$ and the kernel parameter $\eta = 2$ 29

- 2.7 The combined effects of the fitness cost of the resistance breaking trait (δ), and the susceptibility of the R host to the wt pathogen strain (γ), on the absolute reduction in epidemic intensities from using a high compared to a low landscape E/A ratio. The baseline epidemic intensities from using a low E/A ratio are shown in Appendix 3 Fig. 2.11 in section 2.7. Note that using the proportional differences in epidemic intensity produces a qualitatively similar pattern. **The strongest spatial suppressive effect occurs with a 100% effective resistance gene that imposes intermediate fitness costs on the resistance breaking pathogen strain. The spatial effect operates in the reverse direction if the resistance breaking strain is not fit enough to invade the landscape, so that only the wild-type strain is present, and the R gene is of intermediate to lower efficacy.** In the region where the rb strain does not invade (defined arbitrarily as when the rb epidemic intensity < 0.01), there is no consistent response to changes in δ on the horizontal axis (due to the absence of the rb strain). For all simulations presented here, the number of seasons $n_y = 40$ and the kernel parameter $\eta = 2$ 30
- 2.8 The effect of the number of seasons (n_y) on the reduction in epidemic intensities from using a high compared to a low landscape edge/area ratio, at different fitness costs of the resistance breaking trait (δ). **The strength of the spatial suppressive effect depends on the length of the time frame of interest. There is the potential for a peak effect strength over an intermediate number of seasons, which occurs over a greater number of seasons with a higher fitness cost of the resistance breaking trait.** The average epidemic intensities resulting from the low and high ends of the E/A ratio scale are shown in (a), (b) and (c), while the absolute differences between the results for these two E/A ratio values are shown in (d), (e) and (f). Note that using the proportional differences in epidemic intensity produces a qualitatively similar pattern. Error bars show the 5th and 95th percentiles of the simulation replicates with stochastic landscape generation. For all simulations presented here, the R host is completely resistant to the wt strain $\gamma = 0$ and the kernel parameter $\eta = 2$ 31

- 3.2 Diagram representing the structure of the two-patch epidemiological model. The arrows represent the probabilities with which inoculum disperses within and between host variety patches. The state variable Y is the number of infected hosts in the left-hand side patch with population size $(1 - \phi)N$, while Z is the number of infected host in the right-hand side patch with population size ϕN 52
- 3.3 Probability tree showing how the inoculum produced in a given host patch may potentially be dispersed. If dispersal is local the inoculum will remain on the same patch type, whereas if it is global the host type it lands on will be governed by the cropping ratio ϕ 54
- 3.4 Flow diagram with pseudoequations showing the estimation process for the aggregation parameter p . The faded out arrows represent the alternative ways in which the value of p can be derived from the expressions for the dispersal probability coefficients in the model. In all cases presented in this chapter the value of p is however estimated from the proportion of inoculum dispersing from susceptible fields that infects resistant fields. The calculation for the total kernel weighted distances between the different types of field is also shown in Eqns 3.11-3.14. 57
- 3.5 Examples of host landscapes generated using the cellular automaton algorithm. Landscape run 1 (a-f) and run 2 (g-l) are each a time series of snapshots taken from single iterations of the algorithm run. The cropping ratio ϕ and edge/area ratio values are shown for each landscape. The landscape snapshots in each group were chosen to show a roughly evenly spaced time series between the minimum and maximum (to the extent that practicality allows) levels of aggregation. The estimated p values for each landscape are given for a single set of field sampling in each landscape with the number of fields $n_f = 100$ and a negative exponential dispersal kernel with mean dispersal distance = 1. The landscape dimensions are 100x100. 59

- 3.6 Violin plots showing for different cropping ratios (ϕ) the distribution of differences between pairs of aggregation parameter p estimates (the maximum possible difference being 1). Each pair of estimates is taken from a given landscape (with $\phi = x$) and its inverted or 'flipped' counterpart (with $\phi = 1 - x$). All p estimates in these plots were generated using random sampling of 1000 field locations for each landscape. The central points give the mean difference and the error bars show \pm 1 standard deviation. The mean dispersal distance (mmd) of the kernel is 1 in (a) and 3 in (b) and (c) (within landscape dimensions 100x100). Plots (a) and (b) were generated from the landscapes modelled as a flat areas with edge effects, whereas in (c) each landscape is modelled as a torus in order to remove any edge effects. 63
- 3.7 Violin plots showing for different cropping ratios (ϕ) the distribution of differences between pairs of aggregation parameter p estimates (the maximum possible difference being 1). Each pair of estimates is taken from a given landscape (with $\phi = x$) and its inverted or 'flipped' counterpart (with $\phi = 1 - x$). All p estimates in these plots were generated using random sampling of 1000 field locations for each landscape (of dimensions 100x100). The central points give the mean difference and the error bars show \pm 1 standard deviation. The mean dispersal distance of the kernel used to generate these plots is 3, and the landscapes are modelled as flat areas with edge effects. The edge/area ratio range for the data in each plot is shown. This figure is a partitioning of Fig. 3.6b. 64
- 3.8 Individual differences, across all cropping ratios, between pairs of aggregation parameter p estimates (the maximum possible difference being 1) as a function of edge/area (E/A) ratio. Each pair of estimates is taken from a given landscape (with $\phi = x$) and its inverted or 'flipped' counterpart (with $\phi = 1 - x$). All p estimates in these plots were generated using random sampling of 1000 field locations for each landscape (of dimensions 100x100). The landscapes are modelled as flat areas with edge effects, and the mean dispersal distance (mdd) of the kernel used to generate each plot is given. Plots (a) and (b) were generated from the landscapes modelled as a flat areas with edge effects, whereas in (c) each landscape is modelled as a torus in order to remove any edge effects. The lines of best fit, the coefficients for which are given in parentheses for each plot, were fitted using generalised least squares. . 65

- 3.9 The landscape template that produced the largest difference in paired aggregation parameter p estimates (out of those in figure (3.8a)) is shown in (a). Further paired p value estimates for different numbers of fields (n_f) are shown in (b), with 30 field location sampling replicates for each value of n_f . Each pair of estimates in (b) is taken from a the landscape shown in (a) (with cropping ratio $\phi = 0.1$) and its inverted or 'flipped' counterpart (with $\phi = 0.9$). The black line in (b) gives the mean difference between paired p value estimates for each value of n_f . The mean dispersal distance of the kernel used is 1, and the edge/area ratio = 3.505. The landscape is modelled as a flat area with edge effects. 66
- 4.1 Example plots showing how the pairwise invasion thresholds, for each pathogen strain invading the other strain at its endemic equilibrium (a and b), can be used to identify combinations of parameter values under which coexistence will/will not occur (c). The fitness cost of the resistance breaking trait $\delta = 0.4$, and the susceptibility of the resistant host to the wt pathogen strain $\gamma = 0.2$. The infection rate $\beta = 2$ per host tissue unit per unit time, the harvesting/replanting rate $\eta = 1$ per unit time, and the overall host population size $N = 10000$ (although results are independent of N), are used for all Figs unless otherwise stated. . . . 92
- 4.2 Plots showing the conditions required for invasion and coexistence of the wt and rb pathogen strains, as a function of the continuous parameters ϕ (cropping ratio of the resistant crop) and p (the aggregation parameter). The plots are arranged in a grid as a function of the discretised parameters δ (fitness cost of the resistant breaking trait) in the horizontal axis and γ (susceptibility of the resistant host to the wt strain) in the vertical axis. 94
- 4.3 Plots showing the conditions required for invasion and coexistence of the wt and rb pathogen strains, as a function of the continuous parameters δ (fitness cost of the resistant breaking trait) and γ (susceptibility of the resistant host to the wt strain). The plots are arranged in a grid as a function of the discretised parameters ϕ (cropping ratio of the resistant crop) in the horizontal axis and p (the aggregation parameter) in the vertical axis. 95

- 4.4 Plot showing the critical values of the aggregation parameter p (the proportion of inoculum that disperses according to the global cropping ratio) where invasion of the wt equilibrium by the rb strain will be prevented at the widest range of potential cropping ratios (ϕ). This is presented as a function of the parameters δ (fitness cost of the resistant breaking trait) and γ (susceptibility of the resistant host to the wt strain). For each value of γ , the median of the range of potential values for critical p is plotted. These critical p values are used to plot a smoothing curve (using the loess method) for each value of δ . Critical p values are only given for δ and γ value combinations where rb invasion can possibly occur. 96
- 4.5 Plot showing the values of the parameters δ (fitness cost of the resistant breaking trait) and γ (susceptibility of the resistant host to the wt strain) where the two pathogen strains will invade and coexist in a simplified version of the model where the cropping ratio $\phi = 1$, meaning that only resistant crop is planted (and so when p is irrelevant). 96
- 4.6 The effect of the fitness cost of the rb trait (δ) on the reduction in average seasonal epidemic intensities (calculated from AUDPCs) from using a high compared to a low degree of mixing (p) in the host landscape. Epidemic intensities as a function of δ at three different values of the landscape aggregation parameter p are shown in (a). The absolute differences in epidemic intensities from using a high compared to a low p value, in pairs of these p values, is shown in (b). The the landscape cropping ratio $\phi = 0.5$, the number of seasons $n_y = 40$, the R host is completely resistant to the wt strain ($\gamma = 0$), and the initial frequency of the rb strain = 0.01. 98
- 4.7 The effect of the susceptibility of the R host to the wt pathogen strain (γ) on the reduction in average seasonal epidemic intensities (calculated from AUDPCs) from using a high compared to a low degree of mixing (p) in the host landscape. Epidemic intensities as a function of γ at three different values of the landscape aggregation parameter p are shown in (a). The absolute differences in epidemic intensities from using a high compared to a low p value, in pairs of these p values, is shown in (b). The the landscape cropping ratio $\phi = 0.5$, the number of seasons $n_y = 40$, the fitness cost of the rb trait $\delta = 0.3$, and the initial frequency of the rb strain = 0.01. 99

- 4.8 The effect of the fitness cost of the *rb* trait (δ) and the susceptibility of the *R* host to the *wt* pathogen strain (γ) on the reduction in average seasonal epidemic intensities (calculated from AUDPCs) from using a high compared to a low degree of mixing (p) in the host landscape. The absolute differences in epidemic intensities from using high compared to low p values, in pairs of p values, are shown in (a), (b), (c) and (d). The parameter conditions required for invasion and coexistence of the two strains are shown in (e), (f) and (g). The the landscape cropping ratio $\phi = 0.5$, the number of seasons $n_y = 40$, and the initial frequency of the *rb* strain = 0.01. 101
- 4.9 The effect of the number of seasons (n_y) on the reduction in average seasonal epidemic intensities (calculated from AUDPCs) from using a high compared to a low degree of mixing (p) in the host landscape. The absolute differences in epidemic intensities from using $p = 0.5$ compared to $p = 0.1$ are shown for a range of values of the fitness cost of the *rb* trait (δ) and the susceptibility of the *R* host to the *wt* pathogen strain (γ). The the landscape cropping ratio $\phi = 0.5$, and the initial frequency of the *rb* strain = 0.01. 102
- 4.10 The effect of the cropping ratio ϕ and the aggregation parameter p on equilibrium disease frequency (simulation results at $y = 1000$). This is shown for different fitness costs of the *rb* trait (δ), and different susceptibilities of the *R* host to the *wt* pathogen strain. The ϕ and p values where disease frequency is lowest, for given values of δ and γ , is shown as red point. The values of ϕ where disease frequencies are lowest for given values of p are shown as a dotted red line. The values of p where disease frequencies are lowest for given values of ϕ are shown as a green line. The yellow contour lines, for aiding the visualisation of the change in disease frequency, have a bin width of 0.02. The number of seasons $n_y = 1000$, and the initial frequency of the *rb* strain = 0.01. 104
- 4.11 The effect of the cropping ratio ϕ and the aggregation parameter p on average seasonal epidemic intensities calculated from AUDPCs between $y = 0$ and $y = n_y = 40$. This is shown for different fitness costs of the *rb* trait (δ), and different susceptibilities of the *R* host to the *wt* pathogen strain. The ϕ and p values where epidemic intensity is lowest, for given values of δ and γ , is shown as red point. The values of ϕ where epidemic intensities are lowest for given values of p are shown as a red line. The values of p where epidemic intensities are lowest for given values of ϕ are shown as a green line. The yellow contour lines, for aiding the visualisation of the change in epidemic intensity, have a bin width of 0.02. The number of seasons $n_y = 40$, and the initial frequency of the *rb* strain = 0.01. 105

- 4.12 The effect of the cropping ratio ϕ and the aggregation parameter p on average seasonal epidemic intensities calculated from AUDPCs between $y = 10$ and $y = n_y = 40$. This is shown for different fitness costs of the rb trait (δ), and different susceptibilities of the R host to the wt pathogen strain. The ϕ and p values where epidemic intensity is lowest, for given values of δ and γ , is shown as red point. The values of ϕ where epidemic intensities are lowest for given values of p are shown as a red line. The values of p where epidemic intensities are lowest for given values of ϕ are shown as a green line. The yellow contour lines, for aiding the visualisation of the change in epidemic intensity, have a bin width of 0.02. The number of seasons $n_y = 40$, and the initial frequency of the rb strain = 0.01. 107

List of tables

2.1	Parameters and variables used in the mathematical model. Parameters set to specific values are generally examples taken from Fabre et al. (2012) and/or are set for arbitrary reasons which are explained in the main text.	22
2.2	The interior edge/area ratios of the landscapes at each of the six scales of spatial heterogeneity (see also Fig. 2.1 and Appendix 2 in section 2.6). The interior edge of a landscape is the summed length of the straight lines (excluding the outside edges) that are used to divide the landscape into a given number of evenly sized patches. This summed edge length is then divided by the total area of the landscape to give the edge/area ratio.	22
4.1	Parameters and variables used in the mathematical model	84
4.2	The figures in chapters 2 and 4 that demonstrate the same qualitative effect of a given parameter(s) on the durability of disease resistance.	100
4.3	Summary of the long-term behaviour of the model in which coexistence is not possible ($\gamma \neq (1 - \delta)$).	119

Chapter 1

General Introduction

1.1 Background and motivation

Plant disease accounts for an estimated 13% loss in annual crop yields, with estimates of losses to major crops ranging as high as 30% for rice (Oerke, 2006; Oerke and Dehne, 2004; Savary et al., 2019). The breeding and deployment of host resistance in crops, in an attempt to combat such losses, is therefore a major area of focus for both research and industry (Nelson et al., 2018). Disease resistant crops are a valuable and effective tool, especially when used as part of an integrated pest management approach, however their use can present various challenges (Mundt et al., 2002). Pathogen populations have the ability to evolve in response to the deployment of host resistance, in a way that often allows them to overcome or circumvent the particular resistance mechanism so that infection can still occur. In addition to the challenge of improving food security, there is also environmental concern attached to this issue, as a lower efficacy of resistant crops generally leads to a greater reliance on chemical pesticides, with potential negative impacts on non-agricultural ecosystems (Brown, 2015).

Understanding the ecological and evolutionary processes by which pathogen populations respond to the use of resistant crop cultivars is therefore of prime importance, and will aid both the breeding of resistance traits and their optimal deployment (Mundt, 2014). Generally, it is hoped that such an improved understanding will increase what is known as the durability of crop disease resistance, defined as the period of time that resistance remains effective with its widespread cultivation, which is an area of interest to both plant breeders and epidemiologists (Johnson, 1979). The durability of disease resistance varies widely depending on factors such as the molecular mechanism of resistance, the specific resistance gene(s) involved, the mode

of pathogen reproduction, and the pathogen population size and dispersal scale (McDonald and Linde, 2002). Resistances to fungal and bacterial diseases are generally thought to break down faster than those for viruses, in part due to the often lower mutational fitness costs in cellular pathogens (Garcia-Arenal and McDonald, 2003). The resultant effect of resistance breakdown can be a significant reduction in crop yields, underlining the need to understand how host resistances can be effectively deployed.

1.2 Major gene resistance

The interaction between plant disease resistance and corresponding pathogen virulence, defined here as the qualitative ability of a pathogen to infect a host, often takes the form of a gene-for-gene relationship (Flor, 1971). This genetic system is defined by a single major resistance (R) gene in the host that recognises, and provides qualitative resistance to, a single corresponding avirulence gene in the pathogen. Often this avirulence gene has an important role as an effector in the mechanism of infection (Cui et al., 2015). Mutations in the avirulence gene can cause it to evade the corresponding R gene and become 'resistance breaking' or 'virulent' (Flor, 1971). Mutation rates have been estimated at 10^{-5} - 10^{-7} per gene, meaning that mutation in an avirulence gene leading it to become resistance-breaking should occur in one in every 100,000 to 10 million cells per generation (Stam and McDonald, 2018).

Major gene resistance in some crops has historically broken down relatively rapidly, and has been described as exhibiting boom-and-bust dynamics as new cultivars need to be repeatedly developed (McDonald and Linde, 2002). The *Mla13* resistance gene to powdery mildew on barley was overcome within 6-7 years after its introduction in the early 1980s (Wolfe and McDermott, 1994), while the *Lr11* and *Lr24* winter wheat leaf rust resistance genes were overcome within 1-2 years (Kolmer, 1996). In contrast, other major resistance genes have remained durable over very long periods of time. One notable example is the *Sr26* resistance gene for wheat stem rust which has maintained its effectiveness ever since its development in 1961 (Knott, 1961; Qureshi et al., 2018). This is one of the few genes that resists the stem rust pathogen race TTKSK (Ug99) along with all other races of this pathogen (Qureshi et al., 2018).

In theoretical population genetic models resistance durability is measured by calculating the changing frequencies of host genotypes (susceptible and resistant), and pathogen genotypes

(wild-type and resistance breaking). Some early models emphasised sustained oscillations of host and pathogen genotypes, or the eventual dominance of so-called pathogen 'super races' that are virulent against all host varieties (Groth, 1976; Jayakar, 1970; Leonard, 1969, 1977; Leonard and Czochor, 1980). More recent theoretical work by Tellier and Brown (2007) however suggests that gene-for-gene interactions can often be stabilised by negative direct frequency-dependent selection (ndFDS), resulting in a stable polymorphism in which resistance breaking and wild-type pathogen genotypes coexist at fixed frequencies. This form of selection is generated here by ecological factors that reduce the degree of synchronous coupling between the life cycles and oscillating gene frequency dynamics of the host and pathogen populations (Brown and Tellier, 2011). Agricultural practice tends to reduce many of the factors that promote this genetic diversity however, and therefore the resistance breaking genotype can often take over the pathogen population (Brown and Tellier, 2011; Sun and Yang, 1999). The deployment of large host monocultures serves to synchronise the host and pathogen life cycles in space and time, meaning that the effects of ndFDS are lost (Brown and Tellier, 2011). Host and pathogen dispersal scales are also important factors in the maintenance of genetic polymorphisms (Gandon, 2002). This is because migration between locally adapted patches of hosts and pathogens may reintroduce alleles that had previously been lost due to selection and/or stochastic effects. Negative frequency-dependent selection may then cause these reintroduced alleles to increase in frequency, resulting in increased polymorphism (Gandon, 2002). It is also worth noting however that developed agricultural practices tend to limit the effects of coevolutionary dynamics in crop disease systems (Burdon and Laine, 2019). This is because the genetic composition of the host crop population is usually under the control of the grower, and therefore does not change over time in a direct response to the population genetics of the pathogen. Coevolutionary dynamics are arguably therefore of greater relevance to wild plant pathosystems, and in agricultural systems where seed saving practices are common. In this latter example, commonly used by subsistence farmers, a portion of the seeds produced by a crop in a given season are used to plant the crop for the next season, thereby allowing the genetic composition of the host population to change between seasons (Montúfar and Ayala, 2019).

There is substantial evidence however that resistance breaking traits in plant pathogens can have high reproductive fitness costs, with examples in plant viruses (Fraile et al., 2011), and basidiomycetes (Bruns et al., 2014; Thrall and Burdon, 2003). These fitness costs generally

occur if the mutations in the pathogen avirulence gene, that allow it to evade host recognition, also compromise that gene's function to a certain degree. Given that this gene's function is typically the secretion of an effector protein that plays some role in the infection process, changes due to mutation can result in a reduced infection efficiency (Leach et al., 2001). It is important to note that these fitness costs do not always occur. Mutations in avirulence genes that have no clear measurable role in pathogen infectivity (Bai et al., 2000), are compensated for due to functional redundancy (Collmer, 1998), or have clearly separable avirulence (host recognition) and infectivity functions (Shan et al., 2000), will often result in zero fitness cost mutations (Leach et al., 2001).

If present however, fitness costs have the potential to prevent the resistance breaking strain from completely dominating when a mixture of host cultivars are deployed. On a fully susceptible host, which can be infected by both wild-type and resistance breaking pathogen strains, the costs of the resistance breaking trait cause it to potentially experience negative selective pressure. In the absence of immigration or new mutation, this would ultimately lead to loss of the resistance breaking strain from the pathogen population. Such loss of unnecessary virulence is seen in field experimental evolution studies, where the lack of positive selection can result in a reduced frequency of individual resistance breaking pathogen strains (Bousset et al., 2018). On the other hand, mutation to virulence can occur quite frequently, with wild-type viruses for example becoming resistance breaking with as few as one or two nucleotide substitutions in the genes affecting avirulence (Harrison, 2002). This can lead to resistance breaking strains persisting at very low background frequencies in pathogen populations, even in the absence of selection from the corresponding resistant hosts (Stam and McDonald, 2018).

1.2.1 Mathematical modelling of qualitative resistance durability

As with any process which occurs over relatively long periods of time and over large spatial scales, mathematical modelling is a useful tool for examining the factors affecting resistance durability. In recent years, modelling in this area has integrated population genetics and epidemiology to consider durability largely in terms of the overall amount of host plant material infected in a given time period. van den Bosch and Gilligan (2003) introduced the number of additional uninfected host growth days as a measure of resistance durability, moving the focus away from simply looking at pathogen genotype frequencies as had often been done previously (Brown, 1995; Lannou and Mundt, 1996). The traditional metric, of time until the

resistance breaking pathogen reached a certain frequency in the population, encouraged the use of lower cropping ratios (of the resistant cultivar), as this minimises the selective pressure favouring the resistance breaking trait. However, at low cropping ratios, yield loss due to the widespread infection of the susceptible cultivar by the wild-type pathogen makes such a strategy uneconomical. The trade-off between resistance durability and the additional yield from the use of the resistant cultivar means that, in this model, yield is largely independent of cropping ratio (van den Bosch and Gilligan, 2003). This prediction is overturned however when demographic stochasticity is added to the model. Lo Iacono et al. (2013) found that high cropping ratios could increase the risk of wild-type/avirulent pathogen extinction due to large stochastic effects at certain frequencies of periodic perturbation, such as intermittent fungicide application. Intermediate optimal cropping ratios can also emerge if the susceptible variety would otherwise outyield the resistant variety in the absence of disease (Vyska et al., 2016).

The work by van den Bosch and Gilligan (2003) was further developed by Fabre et al. (2012), who represented the seasonally disturbed nature of the agricultural environment with a semi-discrete model (Mailleret and Lemesle, 2009). In the Fabre et al. (2012) model, disease spread occurs in continuous time during cropping seasons but is reset upon harvesting, with dynamics in subsequent years being affected by a reservoir component which allows the pathogen to overwinter between discrete seasons (Burdon and Thrall, 2008). The model also incorporated a mutation-selection balance which determines the frequency of the resistance breaking pathogen strain in fields containing the susceptible cultivar (Ribeiro et al., 1998). The genotype frequency of the resistance breaking strain is assumed in this approach to be at a very low but stable equilibrium due to a combination of negative selection pressure as a result of fitness costs, and repeated mutations towards virulence. This approach contrasts with that taken by studies such as van den Bosch and Gilligan (2003), which involve explicitly modelling the densities of the different pathogen genotypes. The approach of Fabre et al. (2012) makes the characterisation of the genetic system simpler, however it risks overlooking the role of dynamic allele frequency changes, that would be accounted for by a more explicit modelling approach.

Fabre et al. (2012) also implicitly characterised landscape spatial structure by modifying the relative contributions of within field, between field, and reservoir driven infection in a single equation governing the total level of infection in an average field of a given variety. These relative contributions were found to influence the optimal cropping ratio for minimising long

term epidemic intensity, which varied from intermediate to high proportions of the resistant crop. While the landscape characterisation captured the overall degree of connectedness between different fields, along with the relative importance of ongoing primary infection from the reservoir component, it took no account of explicit spatial effects such as dispersal distances or the scale or pattern of spatial heterogeneity. Results from the study by Fabre et al. (2012) specifically highlighted the impact of resistance gene choice on durability, in terms of how easily it can be overcome by mutations in the pathogen avirulence gene, and with what associated fitness costs. In particular they found that using resistance genes that were more difficult for a pathogen to overcome led to lower epidemic intensities and higher optimal cropping ratios.

1.2.2 Effect of crop rotation

Later work using the model developed by Fabre et al. (2012) demonstrated that cultivar mixtures and 'rotation', which involved periodically changing the resistant cropping ratio, could be used to limit epidemic intensities (Fabre et al., 2015). The efficacy of these different strategies depended on the relative strength of pathogen transmission between fields, within fields, and from the reservoir. These results tie in with the idea that multiple and fluctuating selection pressures can be used to disrupt the evolution of pathogen populations and prevent them from becoming fully adapted to their hosts (Burdon et al., 2016). This principle has also been applied to the control of fungicide resistance evolution in pathogen populations (Oliver, 2016). Crop rotation is an example of temporal disease control, which changes the direction of selection over time so that the fittest pathogen strain changes between cropping seasons (Zhan et al., 2015). Temporal changes in selection can affect not only qualitative virulence but also aggressiveness (the quantitative ability to reproduce on a host), with data for wheat leaf rust in France showing adaptation of the pathogen towards newly released cultivars at the expense of reduced fitness on a pre-existing cultivar (Papaix et al., 2011).

1.3 Spatial heterogeneity

It is known from both theoretical and experimental studies that the mixing of cultivars with differing resistance properties can reduce the rate of disease spread in single and multi-pathogen strain systems (Mundt, 2002; Zhan and McDonald, 2013). This can occur through dilution effects, where some of the force of infection from a given pathogen strain is wasted due to

its reduced ability, or even inability, to infect the portions of the host population upon which it is not specialised (Mundt, 2002). The case that the use of such mixtures will lead to the evolution and spread of complex pathogen races virulent on a wide range of cultivars is not a strong one. Chin and Wolfe (1984) found that although a wide diversity of host cultivars led to selection for virulence flexibility in the pathogen population, the overall reduction in the pathogen population size meant that the absolute frequency of these complex races was actually lower than in pure stands. Meanwhile Kolmer (1995) found selection for intermediate levels of virulence flexibility in populations of *Puccinia recondita f. sp. tritici* on Thatcher wheat multilines, a trade-off that suggests the presence of resistance breaking gene costs. The idea that increasing the number of resistance varieties in a mixture reduces overall disease severity, when pathogens exhibit at least partial host specialisation, is supported by Mikaberidze et al. (2015).

It has been found in some theoretical model pathosystems that within-field cultivar mixing is more effective at controlling disease spread than planting some fields wholly with susceptible plants and others with resistant plants (Skelsey et al., 2010). There are however frequent logistical complications associated with the use of within-field mixtures, including phenotypic differences between host varieties, such as in harvest dates, along with the perceived or mandated by consumers need for product purity (Burdon et al., 2016). These factors, which in part explain the limited use of mixtures in agriculture in the developed world (Smithson and Lenne, 1996), highlight the need to understand the effects on disease dynamics of employing host diversification at larger scales than within a single field, i.e. at the landscape or between field scale.

1.3.1 Landscape scale spatial effects

The number of studies looking at the effects of landscape scale host diversification combined with long range pathogen dispersal is limited, no doubt held back by the largely field scale nature of empirical work in this area (Johnson, 1981; Plantegenest et al., 2007). However landscape scale spatial effects are undoubtedly an important consideration when investigating resistance durability due to their important effects on crop disease epidemics in general (Gilligan and van den Bosch, 2008). This viewpoint is supported by evidence from modelling studies of the evolution of fungicide resistance which points to different combinations of factors driving regional versus within field dynamics of resistant and sensitive strains (Parnell et al., 2006; van den

Bosch and Gilligan, 2008). This topic has also received a degree of attention in theoretical studies however, with Skelsey et al. (2010) finding that the clustering of potato cultivation reduced the spread of late blight between clusters, but increased overall epidemic intensity due to higher spread within clusters. It has been posited that this scenario can potentially create a trade-off between within and between patch dispersal, resulting in maximised dispersal at intermediate scales of spatial clustering (Skelsey et al., 2013). The effect of field size on landscape scale disease spread has previously been considered by Vanderplank (1948), while the role that host population fragmentation plays in the evolution of host disease resistance has been highlighted by Carlsson-Granér and Thrall (2002). Spatial structure is also relevant over longer timescales, with Papaïx et al. (2013) showing, in a general theoretical metapopulation study, that spatial clustering of habitat patches facilitates specialisation within a population, in addition to driving increased evolutionary speeds.

One component of landscape scale spatial structure that can potentially be optimised for disease control and resistance durability is the scale of spatial heterogeneity in the deployment of different host cultivars. The mechanism behind the efficacy of mixtures suggests that mixing host genotypes at smaller scales of spatial heterogeneity, and thereby creating smaller genotype unit areas (GUAs), would benefit disease control by decreasing connectivity between patches of the same host cultivar (Mundt, 2002). A modelling study by Papaïx et al. (2014b), concerning a single pathogen strain, found similar benefits to mixing crop genotypes at smaller spatial scales when using major genes conferring complete resistance. However these authors also found that for some levels of incomplete resistance, small patches of partially resistant crop could act as sinks for nearby pathogen populations in susceptible patches, and thereby act as stepping stones to increase overall disease incidence in the landscape with greater field mixing.

Some studies have looked more explicitly at resistance durability against multiple pathogen strains over evolutionary timescales. Sapoukhina et al. (2009) showed, using a reaction-diffusion model, that random mixtures in the host landscape provide greater long term disease suppression compared with patchy mixtures at larger scales of spatial heterogeneity. This is supported by a recent modelling study by Papaïx et al. (2018), who showed that low levels of spatial aggregation, in a mixed landscape of fields planted with either susceptible or resistant cultivars, reduced epidemic intensities over both short term epidemiological timescales and at the long term evolutionary equilibrium. A field experimental evolution study by Bousset et al. (2018) provides a degree of empirical support for this general idea, with results suggesting that higher

'genetic connectivity' within a host population facilitates higher levels of infection. Arguably, a potential weakness of this study was the implicit representation of 'genetic connectivity' by greater experimental inoculation of host variety patches by their specialist pathogen strains, which means that it did not truly demonstrate a landscape scale spatial effect.

1.4 Multi-allelic and quantitative resistance

Much of the past work concerning resistance durability has concerned qualitative, gene-for-gene, interactions, at least partially due to the relative ease with which this genetic system can be modelled. Furthermore, attention is often focussed on simple two allele systems, despite the existence of several multi-allelic R genes, with each allele recognising different pathogen avirulence proteins (Haltermann et al., 2003). In diploid organisms, possession of multiple alleles can confer heterozygote advantage, due to protection against multiple pathogen races, and may help to promote R gene diversity (Brown and Tellier, 2011; Ye et al., 2003). A number of traits however confer quantitative resistance, also referred to as polygenic, partial, field, or horizontal resistance, which reduces disease severity but does not entirely prevent infection (Nelson, 1978). This resistance is usually non-specific, and so is at least partially effective against a wide range of potential pathogens. Rather than being controlled by a single gene, quantitative resistance is usually associated with a number of quantitative trait loci (QTL), the exact number of which and their relative contributions to the phenotype are often unknown (Young, 1996). Quantitative resistance has been selected for in breeding programs against a wide variety of crop diseases, such as powdery mildew in barley (Parlevliet and Van Ommeren, 1988), rice blast (Liu et al., 2004), and various foliar pathogens of maize (Asea et al., 2009). It is particularly useful in systems where qualitative resistance has historically not been durable, and against necrotrophic pathogens where major R-genes are rarely effective (Poland et al., 2009).

The complicated polygenic nature of these traits means that they can vary in their degree of effectiveness, based on the infective ability or aggressiveness of the pathogen, which must accumulate multiple mutations in order to overcome the resistance to a noticeable degree. As a result of this, quantitative resistance is usually considered to be more durable than *R*-gene mediated qualitative resistance, though exceptions to this can be found (Stuthman et al., 2007). Andrivon et al. (2007) found that *Phytophthora infestans* could rapidly adapt to overcome

the polygenic partial resistance of locally dominant potato cultivars. This occurred via an approximately 50% increase in spore production in isolates collected only three weeks apart during the epidemic. This contrasts with some of the examples of durable qualitative resistance given in earlier sections, however it should be noted that effective qualitative resistance to *Phytophthora infestans* does not generally exist (Andrivon et al., 2007).

Previous models of the durability of quantitative resistance traits have found that conclusions drawn from the more commonly studied qualitative traits are not always transferable. Lo Iacono et al. (2012), in a deterministic model, found that there was greater 'Healthy Area Duration' (HAD) gain when qualitative resistance was deployed at a high cropping ratio. This meant that the benefits to crop yield of increased control outweighed the effect of greater selection pressure for the evolution of resistance breaking traits in the pathogen population. This is in contrast to the findings of van den Bosch and Gilligan (2003) that the yield benefit is largely independent of cropping ratio with a qualitative resistance trait. Studies such as that of Lo Iacono et al. (2012) model quantitative resistance as a single continuous trait, the level of protection it provides being subject to evolutionary change depending on the selection pressures acting on both the host and pathogen populations. Change in a pathogen's ability to infect a host can be considered as a change in the host resistance genotype and/or a change in the quantitative virulence of the pathogen population. If costs of virulence to the pathogen are present, then there may be a trade-off in the the ability of the pathogen to infect alternative hosts. Gudelj et al. (2004) modelled a system in which the ability of a pathogen to infect one host type negatively correlated with its ability to infect another host type. Adaptive dynamics was used to show that the evolutionary behaviour of the pathogen population was strongly dependent on the shape of the fitness trade-off curve that affected its transmission on the two alternative hosts. Under certain conditions, selective pressure could lead to the evolution of a specialist or generalist population, with the further possibility of an evolutionary branching event forming two specialist populations. Virulence on resistant host varieties can also result in trade-offs against various aspects of the pathogen's life cycle, such as the length of the infective period (Laine and Barrès, 2013).

1.5 Combining resistance traits

In an effort to maximise the durability of resistance, different resistance traits are often combined by plant breeders (Mundt et al., 2002). Such strategies aim to increase durability by both reducing the pathogen population size and limiting selection for resistance breaking pathogen races. This can be achieved by pyramiding multiple qualitative R genes, which both protects against multiple pathogen races, and ensures that resistance will remain effective if one R gene is overcome (Kiyosawa, 1982; Mundt et al., 2002). The exact number of R genes required for the long term success of this strategy, and the factors this depends on, is largely unknown, however there is some evidence that certain specific R gene combinations are more durable than others (Mundt et al., 1990). Quantitative resistance traits can be used to increase the durability of qualitative resistance genes, as has been demonstrated experimentally by Brun et al. (2010). This is because the partial resistance provided by a quantitative trait reduces the pathogen population size, thereby limiting the speed at which virulent genotypes can emerge and spread to overcome the qualitative resistance. Such a result however requires that the quantitative trait is durable and maintains its effectiveness against the pathogen. A further strategy, that is often considered durable, is the pyramiding of multiple quantitative resistance traits or minor R genes that each provide partial disease resistance (Nelson et al., 2018). It has been posited by some that pyramiding a small number of such traits, which tend to combine their efficacy in an additive fashion, will provide sufficient resistance to the Ug99 strain of wheat stem rust (Singh et al., 2006).

Most modelling studies that have investigated the combination of qualitative and quantitative resistance types have treated the quantitative trait as a simple reduction in pathogen infection rate, rather than as a continuous trait with the potential to evolve (Kiyosawa, 1982; Pietravalle et al., 2006). However the Vertifolia effect, described by Vanderplank (1963), shows how the effectiveness of quantitative resistance could decay in a host population when combined with single or multiple effective R genes. This is due to the presence of resistance costs which, if the qualitative trait is successful in preventing infection of the host, selects for a less costly and less effective quantitative trait. This decay means that if a new pathogen genotype emerges that can break the plant's R-gene mediated resistance, disease progression is much faster than it would be had the quantitative trait remained as effective (Vanderplank, 1963). The selective pressure causing such a quantitative decay will usually be imposed by plant breeding programmes, however seed saving practices, more commonly used in developing

regions, may enable host population genetic changes between generations driven by within field pathogen dynamics (Montúfar and Ayala, 2019). Resistance costs are partly made up of the cost associated with the R gene itself, the mean reduction of fitness of which has been estimated at 3.5% in a meta-analysis by Bergelson and Purrington (1996). There is a much greater cost associated with the actual expression of inducible of host defences by the resistance trait. However as this cost is usually only incurred in the presence of a potentially infective pathogen, it is usually outweighed by the benefits of protection (Brown and Rant, 2013).

1.6 Thesis aims

A number of recent theoretical studies have begun comparing and contrasting the various available strategies for the optimal deployment of resistance genes, such as using mosaics (between field spatial diversification), mixtures, rotations and pyramids (Djidjou-Demasse et al., 2017; Fabre et al., 2015; Rimbaud et al., 2018). We believe however that the specific role of spatial diversification and the dynamics which affect this strategy are not currently well understood. While existing studies consistently point to smaller scales of spatial heterogeneity being optimal for durable and effective disease control in multi-strain systems, there is generally little investigation of the factors that influence the strength of this spatial effect. In order for such spatial strategies to be employed in commercial agriculture, they will need to be cost effective, in that the benefit to resistance durability of planting fields of different cultivars at smaller scales of spatial heterogeneity must offset the likely increased financial cost and operational difficulty of farming in this manner. A principal aim of this thesis is therefore to examine the factors interacting with, and influencing the strength of, any such spatial effect. These factors include the dispersal characteristics of the pathogen, the fitness costs associated with the resistance breaking trait, the efficacy of the host resistance gene, the cropping ratio, and the length of the timescale of interest. Aside from the examination of this spatial effect strength, we also seek to identify overall optimum strategies for maximising resistance durability. These strategies will be defined by the optimum scale of spatial heterogeneity and resistance cropping ratio for a given fitness cost of the resistance breaking trait and efficacy of the resistance gene. The precise manner in which the scale of spatial heterogeneity interacts with the cropping ratio, a key focus of chapters 3 and 4, is another area that has not been thoroughly addressed in existing studies.

Given the complex nature of the dynamic interactions involved in this topic, we restrict our models to looking at pairs of pathogen strains interacting with pairs of resistance varieties. While this system may primarily be applicable to the study of gene-for-gene qualitative resistance, conclusions may also be of relevance to other forms of resistance. This is because all that is required in a given system is that one pathogen strain is fitter than the other on one host, and the other strain is fitter on the other host. This increased fitness could be due to a major R gene having been broken, or alternatively due to a series of minor genes, QTLs, or elements of a complex regulatory resistance pathway having been overcome. In order to untangle the role of spatial heterogeneity, we first develop a spatially explicit model (chapter 2), which includes various forms of ecological complexity, before reducing and simplifying the essential dynamics from this approach into a more flexible and general theoretical framework (chapter 3). This framework is then applied to investigating the effects of spatial heterogeneity on resistance durability (chapter 4), before overall conclusions and avenues for future work are discussed (chapter 5).

Chapter 2

When does spatial diversification usefully maximise the durability of crop disease resistance?

2.1 Introduction

As was stated in the overall introduction to this thesis, our general goal is to understand how spatial heterogeneity in an agricultural host landscape affects the durability of crop disease resistance. Previous work by Fabre et al. (2012) looked at some effects of landscape structure, in terms of the relative contributions of within field, between field, and reservoir driven infection. However by using a non-spatial model these authors did not incorporate explicit spatial effects or attempt to capture the degree of spatial mixing between host varieties. Some other studies have begun to look at the overall effects of explicit spatial heterogeneity on durability, but have not investigated the detailed interactions between such heterogeneity and the other various parameters that govern epidemiological systems (Papaix et al., 2018; Sapoukhina et al., 2009). This may potentially be because the model systems used in these studies were not designed to allow the scale of spatial heterogeneity to be varied as a continuous parameter. By instead simply defining arbitrarily high and low levels of heterogeneity, or by using random versus 'patchy' host mixing, these studies generally conclude that smaller scales of spatial heterogeneity or lower degrees of aggregation will always lead to lower disease incidence over evolutionary timescales. How the magnitude of this general suppressive effect on disease responds to changes in the parameters of the system is a key question however that remains unaddressed.

Of particular interest is the role of the parameters that define the genetic nature of the host-pathogen interaction, and thereby determine the relative fitnesses of the two pathogen strains on the two host varieties. Given that these parameters will have different values for different specific pathosystems, an understanding of how they interact with spatial heterogeneity will allow us to identify the systems where such spatial effects are likely to have the greatest impact. This rationale also applies to factors such as the dispersal characteristics of the pathogen, and the timeframe of interest. This timeframe defines the number of seasons over which policy makers or growers want to maximise the improved crop yield from the use of a specific resistant variety, and is related to the frequency with which new resistance varieties become available. If the breeding of new varieties is a rapid process for example, then a policy maker or grower will only need to maximise the yield gain from the use of a resistant variety over a relatively short timeframe. If however, spatial diversification only provides a significant yield improvement over a much longer timeframe, then it might be rationally concluded that such a strategy is not worth pursuing.

In order to investigate the potentially complex effects of spatial heterogeneity we adapt and extend the model published by Fabre et al. (2012). This model is well known within the field of resistance durability modelling, and is clearly defined as a system of ordinary differential equations (ODEs). The use of ODEs particularly lends itself to use within a spatially explicit framework due to the history of coupled systems of such equations being used to represent dynamic metapopulations (Keeling and Rohani, 2002). While alternative spatially explicit approaches, such as individual based or stochastic models, could be used to investigate this topic, these methods tend to prioritise realistic complexity over simplicity and ease of analysis or investigation. While such more realistic approaches can be highly useful, particularly for predictive modelling, given that we are attempting to untangle complex dynamics in a general, theoretical, and somewhat abstract system it makes sense to embrace simplicity wherever possible. The seasonal nature of the model, which includes primary infection of variable intensity originating from the reservoir host, is also retained for ease of comparison with the previous work of Fabre et al. (2012).

Following Fabre et al. (2012), the within season model presented here has a Susceptible-Infected (*SI*) structure with a fixed overall population size for the within season component. The crop harvesting and subsequent planting of new uninfected hosts by the beginning of the next season means that the model is effectively using a Susceptible-Infected-Susceptible

(*SIS*) structure. (Note that, with the exclusion of this paragraph, in this thesis the term 'susceptible' generally refers to the host variety genotype, and not to the uninfected host as it does here.) This is one of the simplest forms of compartmental epidemiological model, as it excludes any consideration of disease free host growth, infected host removal or other loss of infectivity during seasons, or a latently infected period. The approach of distinguishing within versus between field infection used by Fabre et al. (2012) is also useful for inclusion in a spatially explicit model. This is due to the fact that the distances between fields are used to calculate the magnitude of between field transmission, while the within field local population dynamics are assumed to occur under mean field conditions. In order to represent the dispersal of pathogen inoculum between fields we use a negative exponential dispersal kernel, where the contribution that infections in a given field make to the occurrence of new infections in another field declines rapidly with the distance between the two fields.

In addition to the inclusion of spatial field locations, the other major departure from the Fabre et al. (2012) model is the explicit separation of the dynamics of the wild-type and resistance breaking pathogen strains. This is in comparison to the previous assumption of a mutation-selection balance which allowed Fabre et al. (2012) to only include a single equation in the system of ODEs for the density of the pathogen on each of the two host varieties. This is reliant however on the further assumption that the frequency of the resistance breaking strain in susceptible fields is at a very low equilibrium frequency, necessarily constant at all times, driven by the greater reproductive fitness of the wild-type strain on these hosts. While this frequency is meant to reflect the fitness cost of the resistance breaking strain, it would in reality also be affected by the scale of between field dispersal of inoculum. This is particularly relevant if we are investigating the scale of spatial heterogeneity in the landscape, as high levels of resistance breaking inoculum dispersal from nearby resistant fields for example would inevitably increase the resistance breaking frequency on susceptible fields. Furthermore, there is no clear reason why the assumption of an equilibrium frequency is appropriate, as we might well expect the frequency to change over the course of the epidemic as the resistance breaking strain spreads. While explicit separation of the strain dynamics means we must effectively double the number of equations in our ODE system, this added complexity is necessary for questions we investigate in this study.

The aim of this chapter is to investigate the factors that influence the strength of any effect that results from mixing the different cultivars at smaller scales of spatial heterogeneity.

This will involve measuring epidemic intensities over evolutionary timescales, which acts as our method of evaluating resistance durability. Initially we will show how the intensity of epidemics responds to changes in the scale of spatial heterogeneity, and also how this epidemic intensity is distributed between the different host-pathogen genotype combinations. We will also show how the scale of pathogen dispersal effects these patterns. Following this, we will calculate the change in epidemic intensity from using small versus large scales of spatial heterogeneity in the host landscape, and will show how this change responds to various factors. These factors are the fitness cost of the resistance breaking trait, the efficacy of the resistance gene, and the length of the timescale of interest. For some combinations of these factors we will also show how they interact to determine the spatial effect strength.

2.2 Materials and Methods

Our *SI* (Susceptible, Infected) model tracks two pathogen strains, a 'wild-type' (*wt*) and a 'resistance breaking' (*rb*) genotype (the principal variables and parameters used in this model are summarised in table 2.1). The underlying host landscape consists of a number ($n_f = 100$) of cultivated fields, each with a constant number of plants/plant tissue units ($n_p = 1000$) of either a susceptible (*S*) or resistant (*R*) cultivar type (note that the use of continuous state variables combined with the method of epidemic parameterisation makes the results ultimately independent of n_p). The proportion of resistant fields (the cropping ratio) is set evenly with that of susceptible fields at $\phi = 0.5$.

Resistance acts in a gene-for-gene system where the *wt* strain can freely infect *S* fields, but has reduced fitness (that may be zero) when infecting *R* fields. The *rb* strain has equal infective ability on both host genotypes, but may have a reproductive fitness cost (δ) (expressed on both host genotypes) associated with its resistance breaking trait. The focus of our model on disease spread at the landscape scale, with kernel driven pathogen dispersal, makes it appropriate for application to any gene-for-gene crop disease system with a foliar wind dispersed pathogen (such as many rusts or powdery mildews).

The overall number of infected plants in each field is simulated for $n_d = 120$ days, in each of n_y seasons in a semi-discrete modelling approach. The choice of 120 growing days per season is arbitrary, since the method of epidemic parameterisation (see Appendix 1 in section 2.5) makes the results ultimately independent of n_d . In this, continuous time dynamics in ordinary

differential equations (ODEs) are used for the within season component, and discrete dynamics are used for the pathogen infecting and overwintering in the reservoirs (Mailleret and Lemesle, 2009). The reservoir components are localised to each field, and represent non-crop plants that act as alternative hosts for the pathogen. These non-crop plants can either grow within the fields or in the surrounding field margins and hedgerows, and may include crop volunteers, wild weed relatives or unrelated wild plants that are targeted by the same pathogen (Burdon and Thrall, 2008). Example alternative hosts that act as disease reservoirs include *Berberis* spp. (barberry) (Jin et al., 2010) and *Hordeum* spp. (barley grass) (Wellings et al., 2000) for the wheat stripe rust pathogen *Puccinia striiformis* f.sp. *tritici*, and *Abutilon theophrasti* (velvetleaf) for a number of soybean fungal parasites (Hepperly et al., 1980). For some pathosystems such as aggregate sheath spot in rice caused by *Rhizoctonia oryzae-sativae*, the reservoir component could alternatively represent the pathogen overwintering in field stubble (Lanoiselet et al., 2005). The importance and role of reservoir hosts varies between disease systems and depends upon a number of factors such as the the pathogen mating system and genetic diversity, along with local agricultural management practices (Burdon and Thrall, 2008; Fabre et al., 2012). In all simulations, $I_{wt,x,y}(0) = I_{rb,x,y}(0) = 0$ (i.e. the number of infected plants within fields is set to zero at the beginning of each season), with all epidemics in each season started by primary infection from the reservoir of inoculum.

The total area under the disease progress curve (AUDPC) over the n_y seasons is used as the measure of epidemic intensity (Madden et al., 2007). As a measure of resistance durability this is equivalent to the number of uninfected host growth days used by van den Bosch and Gilligan (2003). The AUDPC is normalised to a value between 0 and 1 to obtain the average proportion of plants infected across the sequence of epidemic seasons (epidemic intensity = $AUDPC / (n_f n_p n_y n_d)$). Healthy plants can become infected through three alternative routes: from infected plants in the same field at rate β_F , from infected plants in other fields at rate β_C , and from the reservoir of inoculum at rates $\alpha_{wt,y}$ or $\alpha_{rb,y}$ in season y for the *wt* and *rb* pathogen strains respectively. The pathogen population size in the reservoir is not explicitly modelled but is represented by these rate parameters ($\alpha_{wt,y}$ and $\alpha_{rb,y}$) by scaling the baseline rate of infection from the reservoir component α_E .

The values of the transmission parameters β_F , β_C and α_E were optimised by following Fabre et al. (2012) to calculate the relative contributions of the three infection routes to, and the overall intensity of, a baseline epidemic scenario (see Appendix 1 in section 2.5).

The version of the model used for this optimisation uses a landscape with susceptible fields only ($\phi = 0$). The values of the transmission parameters are such that in this case the three transmission routes have an equal contribution to maintaining the epidemic, and the mean proportion of plants infected during a season is 0.5.

The fields are located within a square landscape of 10x10 arbitrary distance units, where the distances (d) between each pairwise combination of fields are calculated and used in a normalised negative exponential dispersal kernel (of the form $K = \frac{\eta^2}{2\pi} e^{-\eta d}$ with mean dispersal distance $1/\eta$) to calculate each pairwise strength of spatial coupling. This form of kernel is commonly used to represent dispersal in epidemiological systems where relatively short ranged dispersal is of primary importance (Parnell et al., 2015, 2010, 2009). Although our model could represent a landscape scale epidemic, the vast majority of dispersal still occurs over relatively short distances. The importance of rare long distance dispersal events is limited due to the endemic disease scenario with both strains of pathogen already present in all spatial locations, albeit with the resistance-breaking strain at a low initial density. Each simulation was repeated 100 times, with different random field locations in each replicate.

To explore the effect of the scale of spatial heterogeneity in the arrangement of S and R fields on epidemic intensities, the landscape area is arranged (with cropping ratio $\phi = 0.5$) into a number of patches. Each patch only contains fields of a single host genotype (S or R). At the largest scale the landscape is split in half, with all the S fields in one half and all the R fields in the other. At the smallest scale the landscape is split into 64 patches with alternating host genotypes. Four intermediate scales are also used, for a total of six landscape templates with different patch sizes and scales of spatial heterogeneity (table 2.2). We characterise the scale of spatial heterogeneity via the interior edge/area (E/A) ratio of each landscape (see Appendix 2 in section 2.6). This metric captures the degree of contact between the two different host patch types, relative to the overall size of the landscape. It serves as an effective proxy for the overall proximity of the two host field types to each other, and can also be applied to landscapes with irregular patch shapes, variable patch sizes, and different total areas. The pattern of host genotype patches is used as a template to place the two types of field, using random coordinates, in alternating patches within the landscape (Fig. 2.1).

The spread of the two pathogen strains in a given field within a season follows the general pseudo-equation form:

$$\text{Rate of change of infected plants} = \left(\text{Uninfected plants} \right) \left(\begin{array}{l} \text{Rate of} \\ \text{primary (reservoir)} \\ \text{infection} \end{array} + \begin{array}{l} \text{Rate of} \\ \text{between field} \\ \text{infection} \end{array} + \begin{array}{l} \text{Rate of} \\ \text{within field} \\ \text{infection} \end{array} \right) \quad (2.1)$$

and is described by the deterministic ODE system:

$$\frac{dI_{wt,x,y}}{dt} = \gamma^* (n_p - I_{wt,x,y} - I_{rb,x,y}) \left(\alpha_{wt,x,y} + \beta_C \left(\sum_{z=1}^{n_f} I_{wt,z,y} K[z,x] \right) + \beta_F I_{wt,x,y} \right) \quad (2.2)$$

$$\frac{dI_{rb,x,y}}{dt} = (1 - \delta) (n_p - I_{wt,x,y} - I_{rb,x,y}) \left(\alpha_{rb,x,y} + \beta_C \left(\sum_{z=1}^{n_f} I_{rb,z,y} K[z,x] \right) + \beta_F I_{rb,x,y} \right), \quad (2.3)$$

in which

$$\gamma^* = \begin{cases} 1 & \text{if } x \text{ is of type } S \\ \gamma & \text{if } x \text{ is of type } R \end{cases}, \quad (2.4)$$

in Eqn (2.2).

In Eqns (2.2) and (2.3), x indicates variables pertaining to a particular field, $K[z,x]$ is the dispersal kernel coupling field z to field x , and $z \neq x$. Here $I_{wt,x,y}$ is the number of plants infected by the wt strain while $I_{rb,x,y}$ is the number infected by the rb strain (both being in field x and in season y). The cost of the resistance breaking trait is δ , which is assumed act on the rate of sporulation on both the primary and reservoir host, while γ is the susceptibility of the R host to the wt strain. If resistance to the wt strain is complete $\gamma = 0$, meaning that the wt epidemic in an R field (Eqn (2.2) where $\gamma^* = \gamma$) does not occur.

The terms $\alpha_{wt,x,y}$ and $\alpha_{rb,x,y}$ represent the specific infection rates of each pathogen genotype from the local reservoir, where the reservoir host is assumed to be selectively neutral to both pathogen genotypes. These rates are subject to change between seasons from initial values of $\alpha_{wt,x,0} = \alpha_E$ and $\alpha_{rb,x,0} = \theta \alpha_E$, where θ is the initial frequency of the rb genotype. The contributions of the reservoir components to the epidemic in season y are given by:

$$\alpha_{wt,x,y} = \lambda \frac{\alpha_E (A_{wt,x,y-1})}{A_0/n_f} + (1 - \lambda) \alpha_{wt,x,y-1} \quad (2.5)$$

$$\alpha_{rb,x,y} = \lambda \frac{\alpha_E (A_{rb,x,y-1})}{A_0/n_f} + (1 - \lambda) \alpha_{rb,x,y-1} \quad (2.6)$$

Here, $A_{wt,x,y-1} = \int_0^{n_d} I_{wt,x,y-1}(t) dt$ and $A_{rb,x,y-1} = \int_0^{n_d} I_{rb,x,y-1}(t) dt$, the AUDPCs for the epidemics in the previous season caused by the *wt* and *rb* pathogen genotypes respectively in a given individual field x . The baseline landscape AUDPC $A_0 = \sum_{x=1}^{n_f} (\int_0^{n_d} I_{S,x,y}(t) dt)$, calculated from one season in the fully susceptible baseline model, is used to scale the previous season's epidemic to measure the proportional reduction in epidemic intensity due to the presence of R fields.

The parameter $\lambda \in [0, 1]$ characterises the reservoir. High values of λ indicate a rapidly changing reservoir with primary infection largely depending on the intensity of the previous season's field epidemics. This scenario could potentially be caused by a short reservoir host lifespan (annual plants), a low rate of secondary spread within the reservoir, or a small reservoir size (Fabre et al., 2012). Low values of λ indicate a more 'stable' reservoir, with a larger effect of older epidemics damping changes to $\alpha_{wt,x,y}$ and $\alpha_{rb,x,y}$. A value of $\lambda = 0.5$, equally weighting both terms in Eqns (2.5) and (2.6), is used here. At present, there is a lack of clear evidence from epidemics in crops and their reservoir hosts that could be used to parameterise this measure of reservoir stability.

Table 2.1 Parameters and variables used in the mathematical model. Parameters set to specific values are generally examples taken from Fabre et al. (2012) and/or are set for arbitrary reasons which are explained in the main text.

Symbol	Parameter/Variable Description	Constraints/Values
$I_{wt,x,y}$	Number of plants in field x and season y infected by the wt strain	
$I_{rb,x,y}$	Number of plants in field x and season y infected by the rb strain	
n_f	Number of fields	100
n_p	Number of plants/plant tissue units per field	1000
ϕ	Cropping ratio (proportion of resistant fields)	0.5
n_d	Number of days in a season	120
n_y	Number of seasons	$1 \leq n_y \leq 80$
θ	Background equilibrium frequency of rb strain on S host	0.01
δ	Fitness cost of the rb trait	$0 \leq \delta \leq 1$
γ	Susceptibility of the R host to the wt strain	$0 \leq \gamma \leq 1$
$K[z,x]$	2D normalised dispersal kernel coupling field z to field x	$z \neq x, K = \frac{\eta^2}{2\pi} e^{-\eta d}$
η	Dispersal kernel parameter	3 levels: 1, 2, 3
A_0	Baseline AUDPC for one season in a fully susceptible landscape	$A_0 = \sum_{x=1}^{n_f} (\int_0^{n_d} I_{S,x,y}(t) dt)$
β_F	Within field infection rate	(see text and SI)
β_C	Between field infection rate	(see text and SI)
α_E	Baseline rate of infection from the reservoir	(see text and SI)
$\alpha_{wt,y}$	Rate of wt infection from the reservoir	
$\alpha_{rb,y}$	Rate of rb infection from the reservoir	
$A_{wt,x,y-1}$	AUDPC for the wt epidemic in field x in the previous season	$\int_0^{n_d} I_{wt,x,y-1}(t) dt$
$A_{rb,x,y-1}$	AUDPC for the rb epidemic in field x in the previous season	$\int_0^{n_d} I_{rb,x,y-1}(t) dt$

Table 2.2 The interior edge/area ratios of the landscapes at each of the six scales of spatial heterogeneity (see also Fig. 2.1 and Appendix 2 in section 2.6). The interior edge of a landscape is the summed length of the straight lines (excluding the outside edges) that are used to divide the landscape into a given number of evenly sized patches. This summed edge length is then divided by the total area of the landscape to give the edge/area ratio.

Number of patches	Edge/area (E/A) ratio
2	$10/10^2 = 0.1$
4	$20/10^2 = 0.2$
8	$40/10^2 = 0.4$
16	$60/10^2 = 0.6$
32	$100/10^2 = 1$
64	$140/10^2 = 1.4$

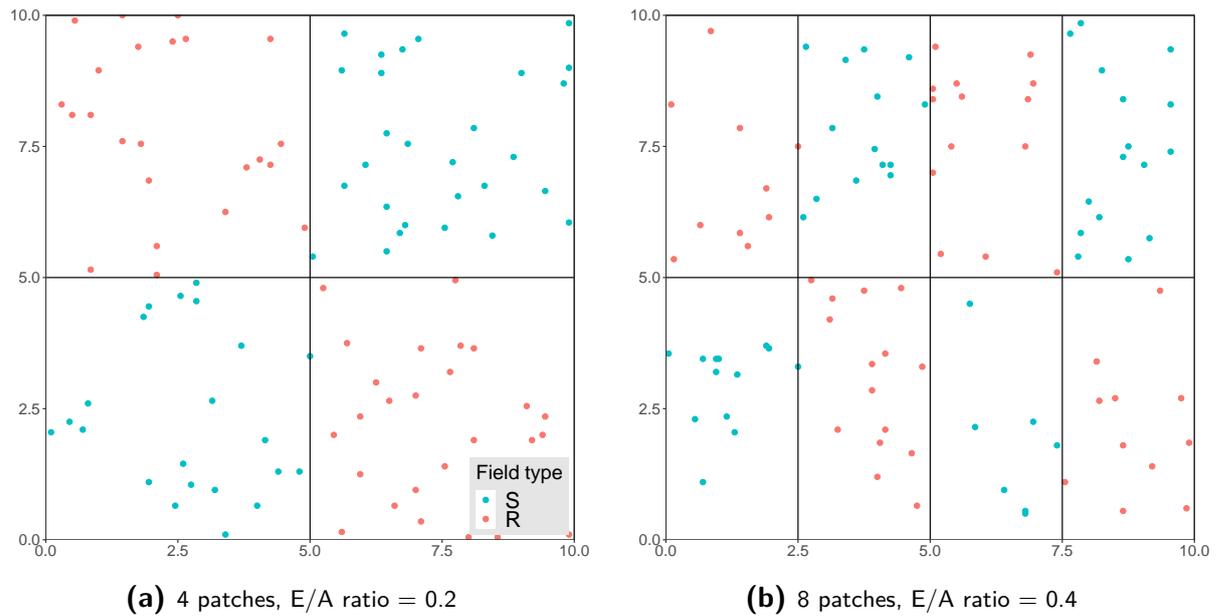


Fig. 2.1 Example random field arrangements for two patch templates within the agricultural landscape.

2.3 Results

An illustrative time series of seasonal disease progress curves for an example landscape with a high edge/area ratio is shown in Fig. 2.2. The average seasonal epidemic intensity (average proportion of plants infected throughout the epidemic) over a 40 season time period invariably decreases as the edge/area (E/A) ratio of the landscape is increased (Fig. 2.3). This is due to the larger dilution caused by greater mixing of the two host genotype field types at smaller scales of spatial heterogeneity. Epidemic intensity is initially measured here over a 40 season time period, to balance short term and long term evolutionary dynamics, although the effect of varying the time frame of interest is described in section 2.3.4. The reduction in epidemic intensities due to the planting of S and R fields at smaller scales of spatial heterogeneity within the landscape (higher E/A ratios) is referred to as the ‘spatial suppressive effect’ throughout the remainder of this chapter.

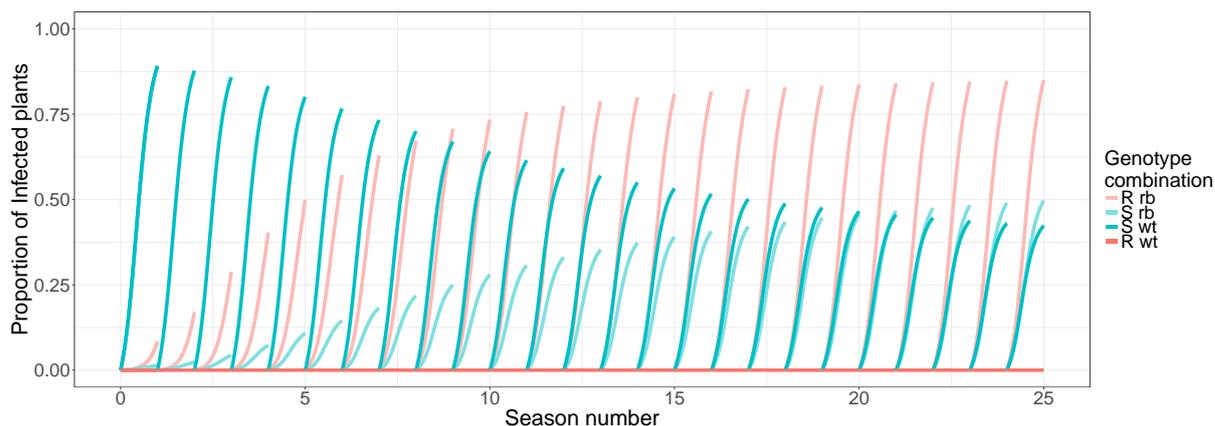


Fig. 2.2 Seasonal disease progress curves for each infection class (genotype combination) in an example landscape simulation replicate. The intensity of disease changes between seasons due to the changing contribution of primary infection from the local reservoirs, which in turn is based on the local epidemic intensities in the previous season. The two field types in the landscape are highly mixed (edge/area ratio = 1.4), the cost of the *rb* trait $\delta = 0.3$ and the *R* host is completely resistant to the *wt* strain $\gamma = 0$. Only 25 seasons are shown here so that the individual disease progress curves can see clearly seen.

The *wt* epidemic is suppressed at higher E/A ratios (Fig. 2.3) due to the smaller spatial grain causing an increase in the proportion of dispersed *wt* inoculum from *S* fields that is wasted as it lands on, but is unable to infect, *R* fields. The *rb* epidemic in *R* fields meanwhile is suppressed by a corresponding process in which there is an increase in the proportion of *rb* inoculum that lands on *S* fields. The *rb* inoculum is in direct competition with *wt* inoculum for uninfected host tissue within *S* fields, resulting in a lower intensity *rb* epidemic than would take place in an *R* field. The consequent *rb* genotype dispersal from these *S* fields back onto the nearby *R* fields therefore has a lower force of infection than in *R* field to *R* field transmission over the same distance. The reduction in the intensity of the *rb* epidemic on *R* fields is compensated to a certain extent by the increase in the frequency of the *rb* genotype on *S* fields as the two field genotypes are more closely mixed together in space. Whether this compensation ultimately increases or decreases the intensity of the overall landscape *rb* epidemic, as the E/A ratio is increased, depends upon the genetic parameters (Fig. 2.4).

2.3.1 Effect of the kernel parameter η

The gradient of the spatial suppressive effect (the rate of change in the response of average seasonal epidemic intensity to E/A ratio) decreases as the E/A ratio is increased (Fig. 2.3). For a higher mean dispersal distance, the rate of change in epidemic intensity is quite steep at low E/A ratios, but then almost completely flattens out at high E/A ratios. In contrast, for a lower mean dispersal distance, there is comparatively little change in the gradient of

the spatial suppressive effect as the E/A ratio is increased. This implies that higher mean dispersal distances cause the strength of the spatial suppressive effect to be washed out at small scales of spatial heterogeneity. This occurs because the high mean dispersal distance of the pathogen limits the impact of any further decrease in the scale of spatial heterogeneity. In general, this means that in the reverse direction, as the mean dispersal distance decreases, the spatial suppressive effect is relevant over a larger range of E/A ratios. Aside from this effect, we also see that there is a generally higher disease epidemic intensity with lower mean dispersal distances (Fig. 2.3c). This is due to less dispersed inoculum being lost over the landscape edges than occurs with a higher mean dispersal distance.

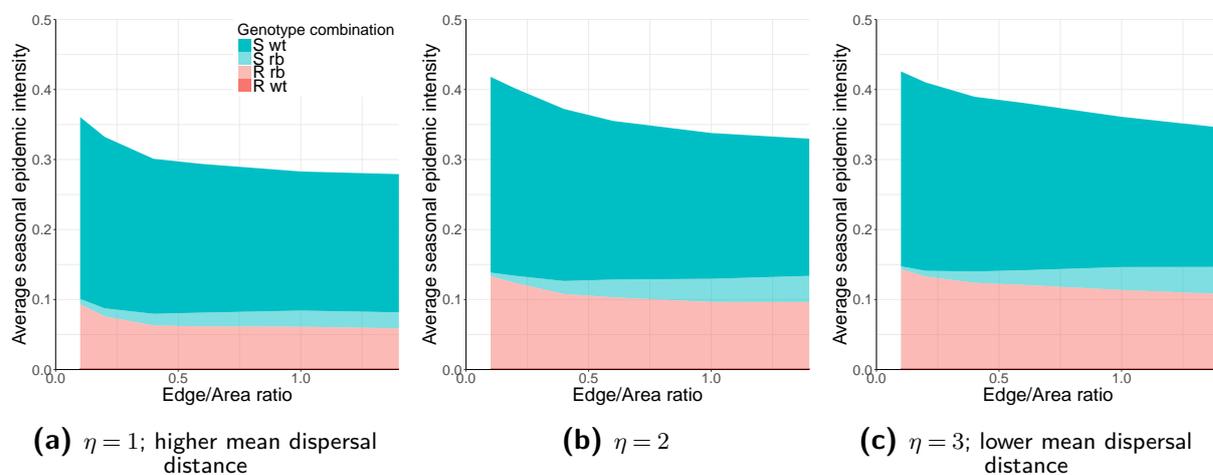


Fig. 2.3 The effect of the kernel parameter η on the relationship between landscape edge/area ratio and epidemic intensities. **Epidemic intensities over evolutionary timescales are reduced at smaller scales of spatial heterogeneity. The spatial scale at which a spatial strategy is effective depends on the scale of pathogen dispersal.** For all simulations presented here, the number of seasons $n_y = 40$, the cost of the *rb* trait $\delta = 0.3$ and the *R* host is completely resistant to the *wt* strain $\gamma = 0$.

2.3.2 Effect of the cost of the *rb* trait δ

The overall size of the spatial suppressive effect increases as the cost of the *rb* trait δ is increased from 0 to 0.4 (Fig. 2.4). Furthermore, the extent to which the pathogen genotype frequencies on the *S* host change, as we move from larger to smaller scales of spatial heterogeneity, also depends on the cost of the *rb* trait. When $\delta = 0$ (Fig. 2.4a) the *rb* genotype is able to take advantage of the greater proportion of host fields that it can infect, and the close proximity of the two field types at smaller spatial scales of heterogeneity, allowing it to outcompete the *wt* genotype. This replacement of pathogen genotypes at different E/A ratios is seen to a lesser extent when $\delta = 0.2$ (Fig. 2.4b), and is almost absent when $\delta = 0.4$ (Fig. 2.4c). As the fitness cost δ increases the *rb* genotype is unable to compete as effectively with the *wt* genotype on

the S host at small scales of spatial heterogeneity. This is despite the close proximity of large numbers of R fields, which act as a major source of rb inoculum, to the S fields.

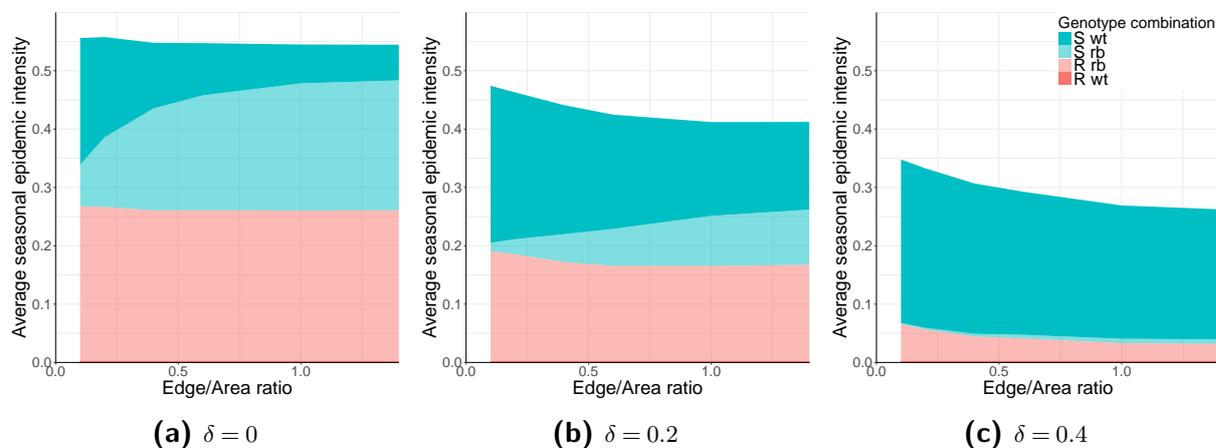


Fig. 2.4 The effect of the fitness cost of the resistance breaking trait (δ) on the relationship between landscape edge/area ratio and epidemic intensities. **The strength of the spatial suppressive effect increases as the fitness cost of the resistance breaking trait is increased from low values. The pathogen genotype frequencies depend on the scale of spatial heterogeneity.** For all simulations presented here, the number of seasons $n_y = 40$, the R host is completely resistant to the wt strain $\gamma = 0$ and the kernel parameter $\eta = 2$.

Plotting overall epidemic intensity against the full range of values for the fitness cost of the rb trait (δ), at both the low and high ends of the E/A ratio scale (i.e. for E/A = 0.1 and 1.4), allows these contrasting spatial scenarios to be compared (Fig. 2.5). For both E/A ratios, epidemic intensities decrease as δ is increased, up to $\delta = 0.6$ where the rb trait is too expensive for that pathogen genotype to invade and there is no further effect of increasing δ (Fig. 2.5a). The variability and strength of the spatial suppressive effect can be ascertained by plotting the difference between the epidemic intensities for the two E/A ratio values (Fig. 2.5b). The strength of the spatial suppressive effect increases from $\delta = 0$ to 0.3, but decreases from $\delta = 0.4$ to 0.6. A small spatial suppressive effect is still seen at $\delta = 0$ due to the small scale of spatial heterogeneity disrupting the transient wt epidemic, before the wt strain is outcompeted by the rb and reaches its near zero evolutionary equilibrium frequency.

The initial increase in the strength of the spatial suppressive effect is due to a steeper gradient of change, in the response of overall epidemic intensity to changes in δ , at small scales of spatial heterogeneity (E/A ratio = 1.4) (Fig. 2.5a). This steeper change with δ is in turn caused by rapidly changing rb dynamics on the S host, which are a larger driver of system sensitivity with greater field mixing (Appendix 3 Fig. 2.10 in section 2.7). Here, any increase in δ reduces the competitive ability of the rb genotype against the wt on susceptible hosts, which consequentially increases the amount of rb inoculum that is ‘wasted’ as it disperses onto these S hosts, thereby increasing the spatial suppressive effect strength.

The subsequent fall in the strength of the spatial suppressive effect is correspondingly due to a steeper gradient of change with δ at large scales of spatial heterogeneity (E/A ratio = 0.1) (Fig. 2.5a). In this range of δ values there is an increased sensitivity to changes in δ of the rb epidemic on the R host, relative to that on the S host (Appendix 3 Fig. 2.10 in section 2.7). This is primarily because the rb genotype is already unable to compete effectively with the wt on the susceptible host in this range, and therefore does not respond to further changes in the fitness cost. These rapidly changing R rb dynamics are a larger driver of system sensitivity with less field mixing.

The value of δ for which the gradient of overall epidemic intensity change with δ is equal at both small and large scales of spatial heterogeneity ($\delta = 0.3$ to 0.4), is the point of maximum spatial suppressive effect on epidemic intensities (i.e. the maximum distance between the curves). At this point the combined sensitivity effects, of rb dynamics on both S and R hosts to changes in δ , have the same net result at high and low landscape E/A ratios.

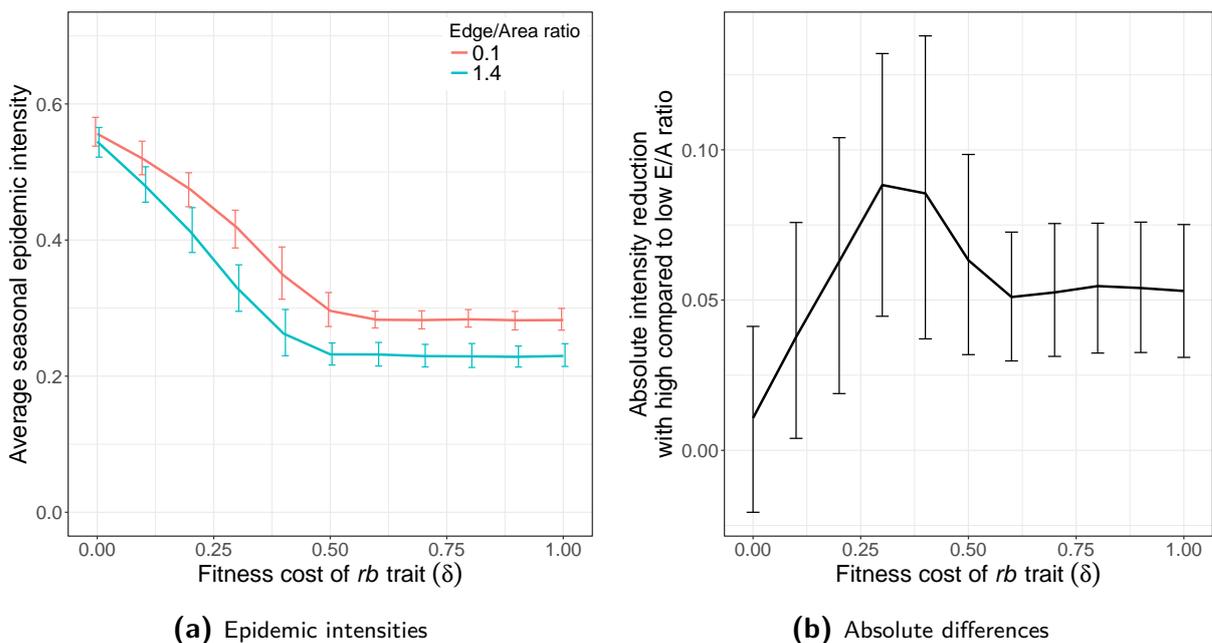


Fig. 2.5 The effect of the fitness cost of the resistance breaking trait (δ) on the reduction in epidemic intensities from using a high compared to a low landscape edge/area ratio. **The strength of the spatial suppressive effect depends on the fitness cost of the resistance breaking trait, with a peak effect strength at an intermediate cost.** The average epidemic intensities resulting from the low and high ends of the E/A ratio scale are shown in (a), while the absolute differences between the results for these two E/A ratio values are shown in (b). Note that using the proportional differences in epidemic intensity produces a qualitatively similar pattern. Error bars show the 5th and 95th percentiles of the simulation replicates with stochastic landscape generation. For all simulations presented here, the number of seasons $n_y = 40$, the R host is completely resistant to the wt strain $\gamma = 0$ and the kernel parameter $\eta = 2$.

2.3.3 Partial resistance

Here, we relax the assumption that the *wt* strain cannot infect resistant hosts (Fig. 2.6). As γ is increased above 0, the *wt* genotype becomes able to infect the *R* host, and at higher frequencies with smaller scales of spatial heterogeneity (high E/A ratios) (Fig. 2.6a,b). The reduced efficacy of the resistance gene allows the *wt* strain to compete more effectively with the *rb* strain on the *R* host, particularly at high E/A ratios where the field types are more greatly mixed in space. The strength of the spatial suppressive effect, from using a high rather than a low E/A ratio, is shown in Fig. 2.6d. For an intermediate cost of the *rb* trait ($\delta = 0.3$), the strength of the spatial suppressive effect decreases to zero as γ is increased from 0 to 0.6. This effect is due to a reduction in the proportion of *wt* inoculum that is 'wasted' in its increased dispersal onto *R* fields at smaller scales of spatial heterogeneity. When $\gamma = 1$ there is no effect of the scale of spatial heterogeneity, as the landscape is then made up of entirely susceptible hosts, and therefore the *wt* genotype is able to fully outcompete the *rb* genotype.

There is a range of γ values, from 0.6 to 0.9 for $\delta = 0.3$ (Fig. 2.6d), for which the strength of the spatial suppressive effect dips below zero and becomes negative, indicating that smaller scales of spatial heterogeneity increase epidemic intensities. This increase only occurs at δ and γ combinations where the *rb* strain is not fit enough to invade, or is only present at extremely low frequencies (Fig. 2.7). In addition to this, the increased epidemic intensities are only observed at medium to high values of γ , and are most apparent at medium γ values. This effect is due to the inability of the *wt* pathogen genotype to sustain epidemics in the *R* host fields without the presence of nearby *S* fields to act as a source of *wt* inoculum. At higher E/A ratios however, where the fields types are mixed at smaller scales of spatial heterogeneity, the closer proximity of *S* field *wt* sources enables greater infection of the *R* field sinks, thereby increasing the overall proportion of infected plants in the landscape, despite the reduction in epidemic intensities on the *S* fields themselves. At very high values of γ , the *wt* genotype is fit enough to better sustain epidemics at large scales of spatial heterogeneity (low E/A ratios), so the negative spatial suppressive effect is reduced in strength (and cannot occur at all with $\gamma = 1$). At low values of γ , spatial suppression of epidemic intensities are still observed, despite the non-invasion of the *rb* strain, due to the very low fitness of the *wt* genotype on the *R* host, which severely limits *wt* epidemics in *R* fields at any scale of spatial heterogeneity.

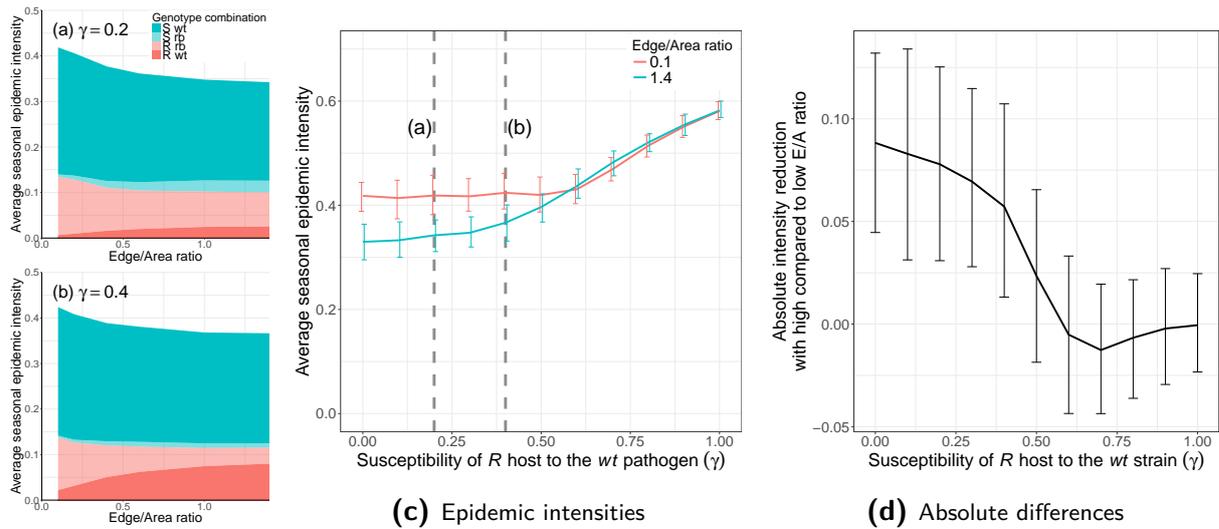


Fig. 2.6 The effect of the susceptibility of the *R* host to the *wt* pathogen strain (γ) on the reduction in epidemic intensities from using a high compared to a low landscape edge/area ratio. **The strength of the spatial suppressive effect is decreased at lower efficacies of the host *R* gene, as long as the *R* gene is not so weak that the resistance breaking strain is not fit enough to invade the landscape.** The average epidemic intensities resulting from the low and high ends of the E/A ratio scale are shown in (c), while the absolute differences between the results for these two E/A ratio values are shown in (d). Note that using the proportional differences in epidemic intensity produces a qualitatively similar pattern. The relationship between the full range of landscape edge/area ratios and epidemic intensity, for each host/pathogen genotype combination, is shown for two values of γ in (a) and (b), and their corresponding positions along the x axis in (c) are marked. Error bars show the 5th and 95th percentiles of the simulation replicates with stochastic landscape generation. For all simulations presented here, the number of seasons $n_y = 40$, the cost of the *rb* trait $\delta = 0.3$ and the kernel parameter $\eta = 2$.

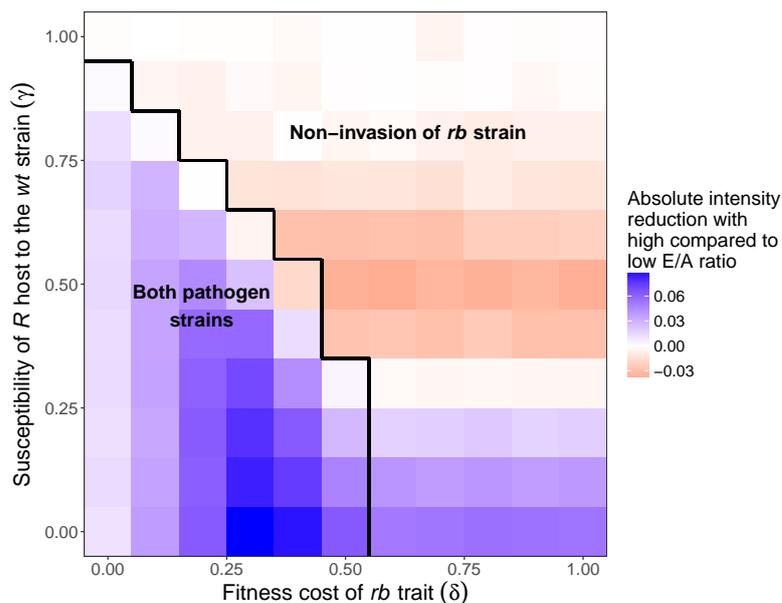


Fig. 2.7 The combined effects of the fitness cost of the resistance breaking trait (δ), and the susceptibility of the R host to the wt pathogen strain (γ), on the absolute reduction in epidemic intensities from using a high compared to a low landscape E/A ratio. The baseline epidemic intensities from using a low E/A ratio are shown in Appendix 3 Fig. 2.11 in section 2.7. Note that using the proportional differences in epidemic intensity produces a qualitatively similar pattern. **The strongest spatial suppressive effect occurs with a 100% effective resistance gene that imposes intermediate fitness costs on the resistance breaking pathogen strain. The spatial effect operates in the reverse direction if the resistance breaking strain is not fit enough to invade the landscape, so that only the wild-type strain is present, and the R gene is of intermediate to lower efficacy.** In the region where the rb strain does not invade (defined arbitrarily as when the rb epidemic intensity < 0.01), there is no consistent response to changes in δ on the horizontal axis (due to the absence of the rb strain). For all simulations presented here, the number of seasons $n_y = 40$ and the kernel parameter $\eta = 2$.

2.3.4 Effect of timescale

By relaxing the assumption of a 40 season time period, we can measure the strength of the spatial suppressive effect, or the difference in epidemic intensities between low and high landscape E/A ratios, across a wide range of eco-evolutionary timescales. The average seasonal epidemic intensity generally increases as the number of seasons is increased (Fig. 2.8a,b,c). This is due to the increasing frequency of the rb genotype, which facilitates greater infection of R fields as the system approaches its long term evolutionary equilibrium. In the case where the fitness cost of the rb trait $\delta = 0$, there is no significant spatial suppressive effect on epidemic intensities over 80 seasons (i.e. the red and blue curves converge in Fig. 2.8a). Here, the rb strain is able to completely outcompete the wt on both host genotypes at the long term evolutionary equilibrium, meaning that the scale of spatial heterogeneity has no effect. A transient spatial suppressive effect does however occur over a low to medium number of seasons, as the system has not yet reached its equilibrium state and the wt strain is still present at significant frequencies.

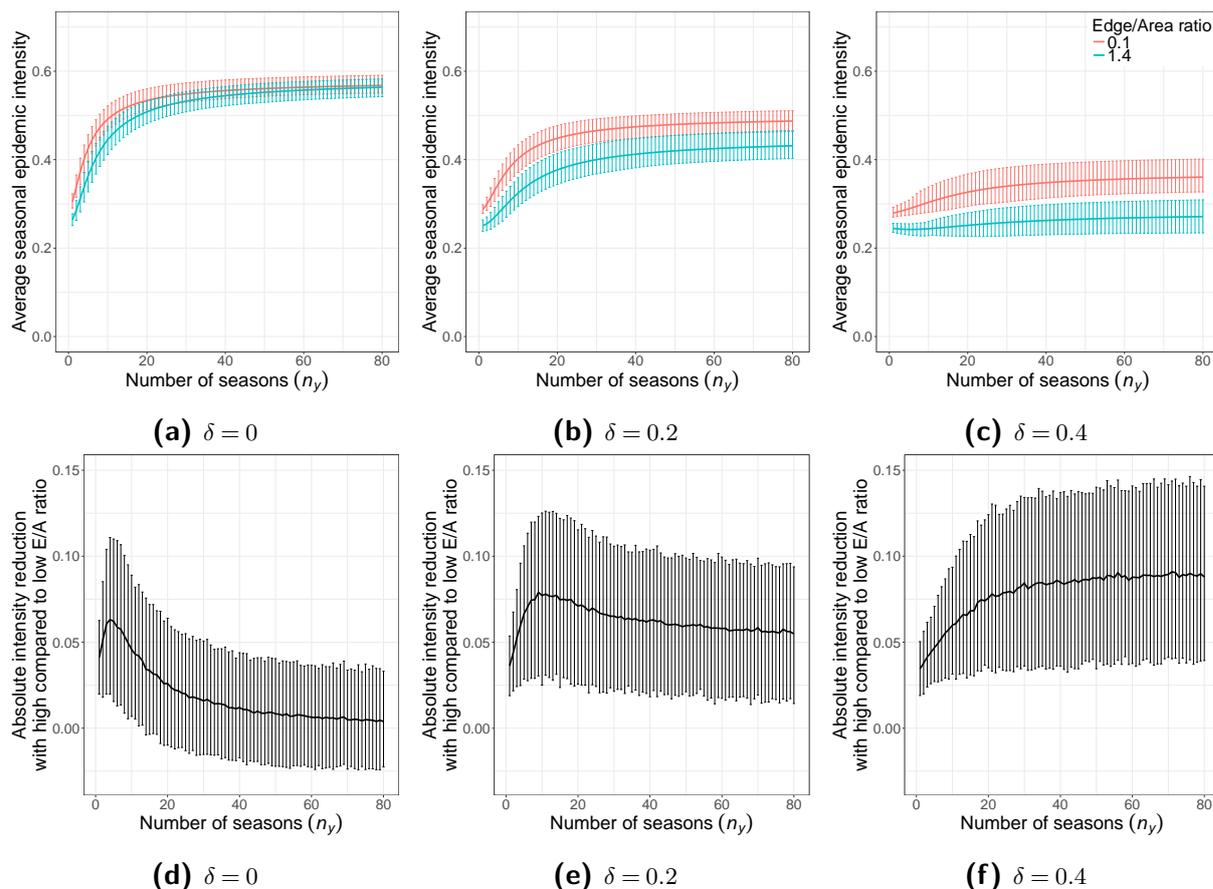


Fig. 2.8 The effect of the number of seasons (n_y) on the reduction in epidemic intensities from using a high compared to a low landscape edge/area ratio, at different fitness costs of the resistance breaking trait (δ). **The strength of the spatial suppressive effect depends on the length of the time frame of interest. There is the potential for a peak effect strength over an intermediate number of seasons, which occurs over a greater number of seasons with a higher fitness cost of the resistance breaking trait.** The average epidemic intensities resulting from the low and high ends of the E/A ratio scale are shown in (a), (b) and (c), while the absolute differences between the results for these two E/A ratio values are shown in (d), (e) and (f). Note that using the proportional differences in epidemic intensity produces a qualitatively similar pattern. Error bars show the 5th and 95th percentiles of the simulation replicates with stochastic landscape generation. For all simulations presented here, the R host is completely resistant to the wt strain $\gamma = 0$ and the kernel parameter $\eta = 2$.

There is a specific number of seasons, for a given fitness cost of the rb trait (δ), over which the greatest spatial suppressive effect on epidemic intensities can be achieved (Fig. 2.8d,e,f). For $\delta = 0$ and $\delta = 0.2$ (Fig. 2.8d,e) this peak effect strength occurs over a relatively short time period, approximately 4 and 11 seasons respectively, whereas with $\delta = 0.4$ the greatest spatial suppressive effect is observed as the system approaches its long term evolutionary equilibrium.

The initial increase in the strength of the spatial suppressive effect, from the first season to the intermediate peak, is due to a faster increase in epidemic intensities at a low E/A ratio compared with a high E/A ratio. This is because the newly emergent rb genotype is able to propagate rapidly in the highly aggregated R fields. The rb genotype initially spreads less quickly in S fields, and therefore at high E/A ratios, due to the high level of competition with

the coexisting *wt* pathogen genotype. Beyond the peak in spatial suppression, if it is present, the strength of the spatial suppressive effect declines as epidemic intensities begin to increase faster at high compared to low E/A ratios. This is due to the increased importance of the *rb* strain spread on S hosts to the sensitivity of the system to changes in season number. The overall *rb* frequency in the landscape is higher, than it is over a lower number of seasons, resulting in a faster growth rate on the S host as it competes more effectively with the *wt* strain. The increase in the intensity of the S *rb* epidemic increases the overall connectivity for the *rb* strain in landscapes with a high E/A ratio, thereby increasing the epidemic intensity for the R *rb* epidemic as well. The *rb* spread on the R host with a low E/A ratio however, is already well advanced at this stage, causing its rate of spread to decline.

The pattern here is that of a trade-off in the relative sensitivity to change in season number of the *rb* epidemics in the S and R hosts, with the peak in the strength of spatial suppression being where the combined sensitivity effect of the two processes is the same at both small and large scales of spatial heterogeneity. This peak occurs over a higher number of seasons with larger values of δ because the greater trait cost means that the *rb* strain spreads more slowly on the R host, and takes longer to become competitive with the *wt* strain on the S host. There is no intermediate peak with $\delta = 0.4$ (Fig. 2.8f) because the *rb* trait is too costly for the *rb* strain to spread as effectively on the S host (Fig. 2.4c).

The variation in epidemic intensities due to the stochastic placement of fields in the landscape (size of the error bars), is highest at the points where the strength of the spatial suppressive effect is greatest (Fig. 2.8). This implies that the effect of the specific stochastic arrangement of fields, and therefore the precise degree of spatial heterogeneity, is greatest when the general sensitivity to spatial dynamics is maximised.

2.4 Discussion

Our model shows that planting susceptible and resistant host crop fields at smaller scales of spatial heterogeneity reduces epidemic intensities over a wide range of eco-evolutionary timescales. Such spatial strategies would therefore increase the durability of disease resistance, using a definition of durability analogous to the additional yield measured by the number of uninfected host growth days (van den Bosch and Gilligan, 2003). The underlying mechanism is similar to the dilution effect that has been reported to reduce short term epidemic intensities

in within field cultivar mixtures (Mundt, 2002). Smaller scales of spatial heterogeneity reduce the local density of a given crop cultivar, meaning that some of a pathogen strain's potential force of infection is wasted as its inoculum disperses onto other nearby cultivars on which it has a lower reproductive fitness. This 'wasted inoculum' effect generally suppresses the epidemics caused by each pathogen strain on the host cultivar where they are specialised and have the greatest fitness, while to a smaller extent boosting epidemics of that strain on their less preferred host. Whilst this study has focussed on qualitative plant resistance genes in a gene-for-gene system, these dynamics should be generally applicable to any system where multiple pathogen strains have different relative fitnesses when infecting multiple different host genotypes. Indeed, a similar basic spatial suppressive effect on resistance durability was observed by Papaix et al. (2018), in a model in which the resistance trait gradually eroded due to progressive small mutations in the pathogen population.

Our results also shed light on a number of the factors determining the strength of this spatial suppressive effect. This has significant implications for the potential effectiveness of any such spatial strategy if it were implemented by growers, as the benefit in terms of crop yield must outweigh any potential economic costs of farming in such a manner. Growing monoculture crops in very large fields, with consequentially large scales of spatial heterogeneity, is the norm in many systems of developed agriculture, largely for reasons of economic efficiency. It is therefore important that we characterise the fundamental eco-evolutionary processes that interact with host spatial structure, as these will ultimately play a crucial role in identifying the specific pathosystems where such spatial strategies are most likely to succeed.

The scale of pathogen dispersal within the agricultural landscape must correspond in some sense to the scale of heterogeneity implemented in that landscape in order to maximise the effectiveness of a spatial diversification strategy. Given that lower mean dispersal distances lead to the spatial suppressive effect being relevant over a wider range of scales (Fig. 2.3), pathogens with more restricted ranges of landscape scale dispersal are more likely to be effectively controlled in this manner. The underlying idea of requiring different spatial diversification strategies to control pathogens with different dispersal characteristics is supported by Sapoukhina et al. (2010). That work, however, focussed on the qualitative differences between local short ranged dispersal through diffusion, and stratified dispersal that also included a separate long distance component, rather than looking at spatial heterogeneity as a continuous scale.

The spatial suppressive effect of cropping pattern on epidemic spread is maximised at an intermediate value for the fitness cost associated with the resistance breaking trait (Fig. 2.5). This peak in effect strength is ultimately driven by landscapes with different scales of spatial heterogeneity generating different frequencies of interaction between the various host and pathogen genotypes in the system. The epidemic intensities for these different infection classes respond at different rates to changes in the cost of the resistance breaking trait, depending on the current value of that trait (Appendix 3 Fig. 2.10 in section 2.7). It is these different rates of change that drive the variable strength of the spatial suppressive effect, and create the peak effect strength at intermediate fitness cost values. The exact fitness cost value for this peak in the spatial suppressive effect is lower with a less effective resistance gene (Fig. 2.7).

A less effective resistance gene lowers the strength of the spatial suppressive effect, as long as the resistance breaking strain has a high enough fitness to be able to invade the agricultural landscape (Figs 2.6, 2.7). If this is not the case, and only the wild-type strain is present, intermediate to lower efficacy *R* genes can drive a reverse spatial suppressive effect, where epidemic intensities are higher with smaller scales of spatial heterogeneity. This occurs because wild-type epidemics are only able to sustain themselves in partially resistant fields when there are nearby susceptible fields to act as inoculum sources. The effect is similar to that observed by Papaïx et al. (2014b), who showed in a single strain system that the directional effect of spatial aggregation depended on the R_0 value for the disease epidemic on the resistant variety. It is worth noting here that if the fitness cost of the resistance-breaking trait is greater than the efficacy of the resistance gene, the resistance-breaking strain will never be able to invade or persist in the landscape as the wild-type will be fitter even on the resistant host.

The combined genetic context in this system is created by the combination of the fitness cost of the resistance breaking trait and the efficacy of the resistance gene. From this we can conclude that a spatial diversification strategy is most likely to be cost effective when using a 100% effective major resistance gene that imposes intermediate fitness costs on a resistance breaking pathogen strain. Despite this, a spatial strategy is still likely to be at least partially effective in any genetic context, as long as the resistance breaking pathogen strain is fit enough to invade and persist within the landscape (Fig. 2.7). The fact that spatial diversification can actually worsen epidemics when only a wild-type strain is present, and a partially effective resistance gene is used, highlights the necessity of understanding the state of the pathogen community and the genetic nature of the system before implementing such control strategies. A

spatial strategy will be less effective when there are no fitness costs for the resistance breaking strain, however an effect is still observed due to the time required for this strain to fully take over the pathogen population. This naturally becomes particularly apparent when looking over a lower number of seasons.

In this study the fitness cost of the resistance-breaking trait and the efficacy of the resistance gene are assumed to have fixed values, implying that no further relevant mutations are occurring in the pathogen population within the timescale of the epidemic. It is known in some disease systems however, such as in *Xanthomonas campestris* pv. *vesicatoria* (bacterial leaf spot) infection in peppers, that virulent (resistance-breaking) strains can accumulate point mutations over time that maintain their infective fitness while decreasing their detection by major resistance genes (Gassmann et al., 2000). This phenomenon could be included in our model by representing the fitness cost of the resistance-breaking trait and the efficacy of the resistance gene as continuous traits which can vary over time in response to selection. This would likely result in the model behaving somewhat like a model for quantitative polygenic resistance, despite the focus still being on major-gene qualitative resistance.

A critical factor that is generally neglected within the study of resistance durability is the length of the time frame of interest. If we consider any improvement in durability to be the yield gain achieved, this will naturally depend on the time period over which we measure such gains, which in practice should itself depend on factors such as the frequency with which new resistant cultivars are developed (varietal replacement rate). This timescale will obviously vary for different crop disease systems and will also depend on the country in question. For example, the average age of hybrid maize varieties (a measure of varietal replacement rate) in the US is only 3-4 years, whereas for wet-season rice in India it is 28 years (Atlin et al., 2017). How long or short this timescale is will play a significant role in determining whether a spatial strategy has a large enough effect to be economical for practical use. The strength of the spatial suppressive effect depends on the number of seasons over which it is measured, with the potential for a peak in spatial suppressive effect strength over an intermediate number of seasons (Fig. 2.8), i.e. in between a short term epidemiological timescale and the long term evolutionary equilibrium. In a similar manner to the effect of the fitness cost of the resistance breaking trait, this occurs because the different infection classes in the system respond at different and varying rates to changes in season number. The frequency of these host-pathogen genotype combinations, and therefore the effect they have on overall epidemic intensity, depends on the

scale of spatial heterogeneity in the distribution of host genotypes within the landscape. The resultant peak spatial suppressive effect occurs over a higher number of seasons with a higher cost of the resistance breaking trait, due to the slower spread rate of this resistance breaking strain. Generally this means that a spatial strategy is most likely to be effective over short timescales for resistance breaking strains that carry little or no fitness costs, and over longer timescales for more costly traits.

In the current study we have restricted the cropping ratio of the susceptible and resistant cultivars to 50 : 50, in order to avoid having to consider potential interactions between the effects of the scale of spatial heterogeneity and the amount of resistant crop deployed. The potential ways in which the patterns we have described might be influenced by different cropping ratios is a valid area for further study however, as is the way that optimal cropping ratios might in turn be influenced by spatial dynamics. The non-spatial model of Fabre et al. (2012) demonstrated that the optimal cropping ratio (i.e. the proportion of resistant fields) varied from intermediate to high values, and depended among other factors on the relative contributions of within field, between field and reservoir driven infection. Demographic stochasticity, which has been shown to bias optimal cropping ratios towards higher values, is another potential route for further investigation (Lo Iacono et al., 2013). The associated chance of pathogen strain extinction at low frequencies or under periodic perturbation could potentially interact with the effects of patch size and spatial heterogeneity on disease dynamics.

As was mentioned in earlier sections, the compartmental epidemiological model which we have used is built upon a fundamentally simple *SI* model, featuring a fixed within season host population size. With the replanting of hosts between seasons this is similar to an *SIS* model. The justification for this is that the spatially explicit aspect of the model, combined with multi-species interactions, already means there is considerable complexity in the system and a large number of parameters. While the inclusion of host demography and additional epidemiological compartments may be relevant for a number of specific pathosystems, in this initial general theoretical model it made sense to focus on the simplest case possible so as to enable a clearer and more thorough analysis (as was also done in Fabre et al. (2012)). It is also worth noting that the initial frequency of the resistance breaking strain we used, relative to that of the wild-type, was reasonably high at $\theta = 0.01$. This means that we were less concerned by the initial period of exponential growth of the resistance breaking strain that would occur when the number of infected individuals is extremely low. The exclusion of a latent period in

our model is therefore of limited relevance, as the major effect of this period is often that of modifying this initial rate of exponential growth (Cunniffe et al., 2012; Madden et al., 2007).

The initial frequency of the resistance breaking strain in the pathogen population could represent the background equilibrium frequency resulting from a mutation-selection balance in the absence of positive selection from the deployment of the resistant host (Fabre et al., 2012). The frequency $\theta = 0.01$ used in our study is relatively high compared to some estimates for rare pathogen mutants under a mutation-selection balance. These frequencies can be several orders of magnitude lower depending on the number of required number of mutational events (Ribeiro et al., 1998). The practical reason for using a higher initial frequency was to ensure that the spread of the resistance breaking could occur in a reasonable timeframe, without the need for additional parameters controlling the pathogen generation rate. The use of a separate pathogen death rate term combined with a higher infection rate, to allow rapid spread from a lower initial frequency, would have produced qualitatively similar results. The only likely difference that this change would make is to potentially increase the period of time that the resistance braking pathogen remained at a very low frequency at the beginning of the epidemic, thereby slightly reducing epidemic intensities. The higher initial frequency could also be seen as due to it being a background frequency in the reservoir host that is not the result of a mutation-selection balance. The absence of negative selection on the resistance breaking strain in this case could be explained by the pathogen using different effectors for the infection of these alternative hosts (Deslandes and Rivas, 2012; Raffaele et al., 2010). This is because any fitness cost associated with mutations in a pathogen effector, that make it resistance breaking, might not be expressed if that specific effector is not used for the infection of a certain host. An additional alternative scenario that our model could correspond with is one where the model has been allowed to burn in to the point where the resistance breaking strain becomes high enough to be detectable. Only once this detection occurs will the maximisation of resistance durability, as measured by long term epidemic intensity, become a priority for a community of growers.

We have also assumed that the initial frequency of the resistance breaking strain is uncorrelated with the fitness cost associated with this trait. This is in contrast with Fabre et al. (2012) who used a mutation-selection balance to represent the fitness cost of the resistance breaking trait and also to set the initial background frequency of this trait. The assumption of Fabre et al. (2012) is an important component of their study, as their model used the

mutation-selection to represent the discounted fitness of the resistance breaking strain on the susceptible host, without explicitly modelling the separate pathogen strains. Our model on the other hand uses an explicit fitness cost with a separate equation for each strain in each field, and as such does not heavily rely on a mutation-selection balance assumption. If we were to change the initial frequency to reflect the fitness cost of the resistance breaking trait in a mutation-selection balance, we would find that the higher fitness cost trait would begin the epidemic at lower initial frequencies. This would mean that these trait would spread more slowly, and as such would simply exaggerate the effect of already having a higher fitness cost. The fact that these two effects of having a higher fitness cost operate in the same direction means that epidemic intensities would likely just respond more rapidly to changes in this cost, and as such would be unlikely to produce qualitatively different results patterns.

Furthermore, the probability distribution of fitness effects for non-lethal pathogen mutants, as shown by Carrasco et al. (2007) for single nucleotide substitutions in *Tobacco etch potyvirus*, reveals that higher fitness cost mutants can be much more common than those with a lower fitness cost. This trend in the effective mutation probability of different fitness traits might be expected to interfere with, and potentially counteract, the trend in equilibrium frequencies produced by a mutation-selection balance that assumes a constant mutation rate. Given this uncertain complexity, and its likely limited impact on the behaviour of our model, we chose to use the simplified assumption of a fixed initial strain frequency. Our assumption could also fit with the scenario where the initial frequency of the resistance breaking strain is due to it originating from the reservoir host, rather than from the the within field epidemic. This is because infection of the reservoir host may not require the use of the specific pathogen effector that incurs a fitness cost in the crop host when it acquires resistance breaking mutations.

In this study we used a negative exponential dispersal kernel function, which is commonly used to represent dispersal in epidemiological systems where relatively short ranged dispersal is of primary importance (Parnell et al., 2015, 2010, 2009). While data on the dispersal of inoculum in some systems can be captured better with fatter tailed dispersal kernels such as power law functions, the primary benefit of these functions is that they can better represent the occurrence of very long distance dispersal events (Grosdidier et al., 2018). We are concerned in this study with an endemic disease scenario where both strains of pathogen are already present in all spatial locations, albeit with the resistance breaking strain at a low initial density. This

means that the frequency of long dispersal events, and the resultant isolated foci of inoculum spread captured by fat tailed kernels, is of limited importance.

In conclusion, this study has demonstrated the key effect that spatial structure can have on disease resistance durability. The diversification of resistance genes at small scales of spatial heterogeneity is a potentially valuable strategy for improving long term crop yields, depending on whether the strength of the spatial suppressive effect leads to such a strategy being economical. Factors such as the pathogen dispersal scale, the genetic properties of the host-pathogen interaction, and the time frame of interest play a crucial role, and highlight the need for a thorough understanding of any disease system to which this strategy is applied.

2.5 Appendix 1 - Fully susceptible model

A landscape made up of only susceptible fields forms the baseline epidemiological context to which all other simulations (with resistance included) are compared. In this model the only state variable of interest is $I_{S,x,y}$, the number of infected plants in a representative S field in season y . Healthy plants can become infected through three alternative routes: from the $I_{S,x,y}$ infected plants in the same field at rate β_F , from infected plants in other fields at rate β_C , and from the reservoir at rate α_E . When combined these rates give the ODE for the change in the number of infected plants as:

$$\frac{dI_{S,x,y}}{dt} = (n_p - I_{S,x,y})(\alpha_E + \beta_C \left(\sum_{z=1}^{n_f} I_{S,z,y} K[z,x] \right) + \beta_F I_{S,x,y}) \quad (2.7)$$

where x indicates variables pertaining to a particular field, $K[z,x]$ is the dispersal kernel coupling field z to field x , and $z \neq x$. The dispersal kernel follows a normalised negative exponential distribution of the form $K = \frac{\eta^2}{2\pi} e^{-\eta d}$. In all simulations, $I_{S,x,y}(0) = 0$ (i.e. the number of infected plants within fields is set to zero at the beginning of each season), with all epidemics started by infection from primary reservoir inoculum. Solving Eqn 2.7 numerically, then integrating with respect to time, gives the area under the disease progress curve (AUDPC). The AUDPC for this baseline model (where the proportion of R fields $\phi = 0$) is $A_0 = \sum_{x=1}^{n_f} (\int_0^{n_d} I_{S,x,y}(t) dt)$, where n_f is the number of fields and n_d is the number of days in a season. This is used for comparison with later simulations to measure the reduction in epidemic intensity with the inclusion of resistant fields and the resistance breaking pathogen. The epidemic behaviour and intensity in this fully susceptible model does not change between

seasons as there is no transformation in the composition of the reservoir component over time. In the full model from the main paper on the other hand, the rates of primary infection from the reservoirs, derived from the constant α_E value and given by $\alpha_{wt,x,y}$ and $\alpha_{rb,x,y}$, do change over time.

The values of the parameters β_F , β_C and α_E used in the simulations were determined by calculating the relative contributions of the three infection routes to, and the intensity of, the overall landscape epidemic. The overall intensity of the epidemic, which describes the proportion of infected plants, averaged over a season, is given by $\Omega_{int} = A_0/(n_f n_p n_d)$. The relative contribution of each route of infection is described by the ‘epidemic profile’, which is given by $\Omega_{pfl} = (\Omega_{pfl}^1, 1 - \Omega_{pfl}^1 - \Omega_{pfl}^3, \Omega_{pfl}^3)$, where the three components are the relative contributions of the reservoir, the between field contacts, and the within field contacts respectively.

The following equations are used to calculate the contributions of the reservoir (Eqn. 2.8) and within field infection (Eqn. 2.9) at time t .

$$\frac{dI_{\bar{S},y}^1}{dt} = (n_p - I_{\bar{S},y}(t))\alpha_E \quad (2.8)$$

$$\frac{dI_{\bar{S},y}^3}{dt} = (n_p - I_{\bar{S},y}(t))\beta_F I_{\bar{S},y}(t) \quad (2.9)$$

Here, $I_{\bar{S},y}(t)$ is the average number of infected plants per field in the spatially explicit baseline model (Eqn. 2.7) at time t . Taking the AUDPCs of Eqns 2.8 and 2.9, we calculate $\Omega_{pfl}^1 = \frac{n_f}{A_0} \int_0^{n_d} I_{\bar{S},y}^1(t) dt$ and $\Omega_{pfl}^3 = \frac{n_f}{A_0} \int_0^{n_d} I_{\bar{S},y}^3(t) dt$, which gives the proportion of the total number of infected plants that are infected by the reservoir and by within field transmission respectively. The proportion of infected plants that are infected by between field transmission is calculated by taking the remaining proportion of infected plants that did not become infected by one of the other two routes of transmission ($\Omega_{pfl}^2 = 1 - \Omega_{pfl}^1 - \Omega_{pfl}^3$). This is done because the value of Ω_{pfl}^2 is more difficult to compute directly as it involves spatially explicit interactions that cannot be derived from average field outcomes as in Eqns 2.8 and 2.9.

The R optimiser function *nlimb* was used to find the β_F , β_C and α_E values that gave $\Omega_{int} = 0.5$ and $\Omega_{pfl} = (1/3, 1/3, 1/3)$. The objective function for this optimisation is given by $E = (\Omega_{int} - 0.5)^2 + (\Omega_{pfl}^1 - 0.333)^2 + (\Omega_{pfl}^2 - 0.333)^2$, where $\Omega_{pfl}^2 = 1 - \Omega_{pfl}^1 - \Omega_{pfl}^3$. In this way, for a fully susceptible model, the three transmission routes are of equal importance in

maintaining the epidemic, and half of the plants in the landscape become infected over a season.

Optimising the model to set infection rate parameters for every random replicate set of field coordinates is computationally expensive, and undesirable for the purposes of reproducibility. Therefore, the fields in the fully susceptible model used to optimise the parameters and provide a baseline epidemic scenario were arranged in a regular square grid pattern (with 1 arbitrary distance unit gaps between fields) within the 10x10 landscape. This of course means that the random landscape used for each simulation replicate produces a slightly different epidemic intensity and profile. This is of limited concern however, as the purpose of the epidemic parameter optimisation is merely to create epidemics of a general size and balanced structure that allow convenient manipulation and comparison.

2.6 Appendix 2 - Interior landscape Edge/Area ratio

The interior edge/area ratio of a landscape is calculated by summing the total length of the edges separating different patch types, and dividing by the total area of the landscape. For example in a 10×10 landscape divided into quarters by two lines perpendicular to the landscape perimeter (Fig. 2.9), the total interior edge is of length 20, and so the edge/area ratio is $20/10^2 = 0.2$.

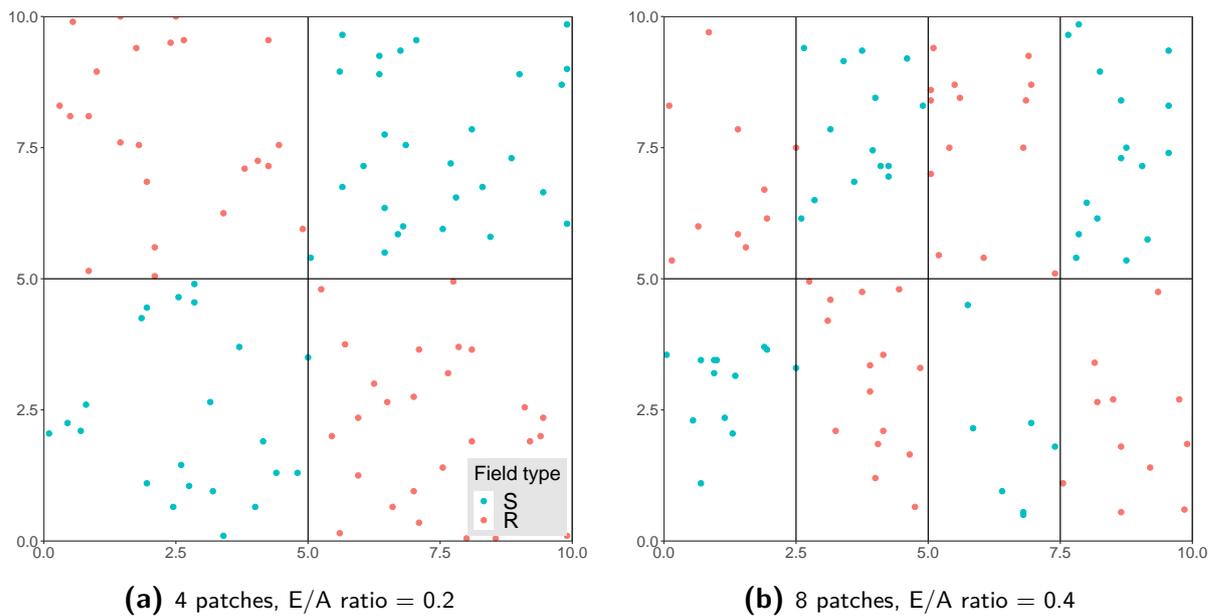


Fig. 2.9 Example random field arrangements for two templates that divide the agricultural landscape into four and eight patches respectively. The labelled dimensions of the landscape templates are used to calculate the interior edge/area ratios (4 patches: $20/10^2 = 0.2$, 8 patches: $40/10^2 = 0.4$), where the numerator of the fraction is the length of the internal edge (4 patches: = 20, 8 patches: = 40) and the denominator is the overall landscape area (= 10^2). This is a repeat of Fig. 2.1.

The patches created by the landscape template are categorised as being either of type S or type R in an alternating pattern. In each replicate simulation, individual fields are then designated (according to the cropping ratio $\phi = 0.5$) as being either S or R , and are then assigned a location with random coordinates drawn from a uniform distribution within the patches pertaining to that genotype.

2.7 Appendix 3 - Additional figures

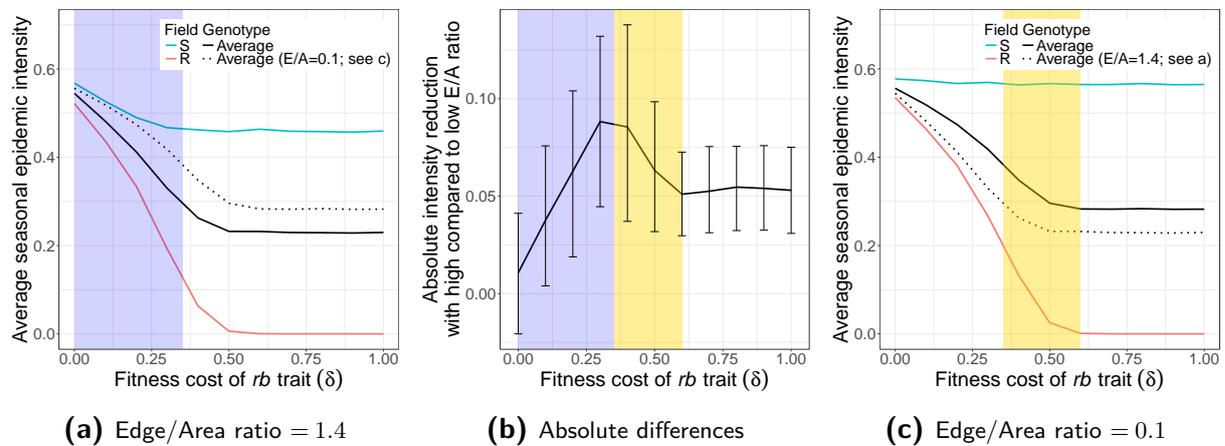


Fig. 2.10 The effect of the fitness cost of the resistance breaking trait (δ) on the reduction in epidemic intensities from using a high compared to a low landscape edge/area ratio. The average epidemic intensities, separated into field types (coloured lines), resulting from a high E/A ratio are shown in (a), and from a low E/A ratio in (c). The total epidemic intensities across both field types (i.e. the mean of the red and blue responses; also shown in main text Fig. 2.5a) are given as a solid black line for the E/A ratios in (a) and (c). The total epidemic intensities for the opposite ends of the E/A ratio scale are also included as dotted lines in (a) and (c). The absolute differences between the total epidemic intensities for these two E/A ratio values are shown in (b). Note that using the proportional differences in epidemic intensity produces a qualitatively similar pattern. Error bars show the 5th and 95th percentiles of the simulation replicates with stochastic landscape generation. **In the blue shaded region, the system changes more rapidly with δ at a high E/A ratio (a), due to the activity of the rb pathogen on the S host (blue line in (a)). Beyond the intermediate peak in (b), in the gold shaded region, the system changes more rapidly with δ at a low E/A ratio (c), due to the activity of the rb pathogen on the R host (red line in (c)).** For all simulations presented here, the number of seasons $n_y = 40$, the R host is completely resistant to the wt strain $\gamma = 0$ and the kernel parameter $\eta = 2$.

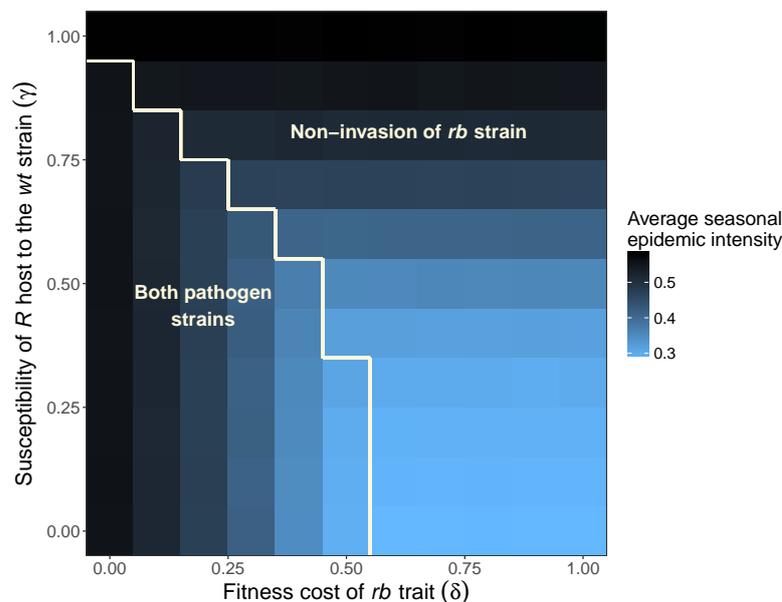


Fig. 2.11 The combined effects of the fitness cost of the resistance breaking trait (δ), and the susceptibility of the R host to the wt pathogen strain (γ), on the average epidemic intensities resulting from a low E/A ratio = 0.1. This plot serves as a baseline for the absolute reduction in epidemic intensities from using a high compared to a low E/A ratio (main text Fig. 2.7). In the region where the rb strain does not invade (defined arbitrarily as when the rb epidemic intensity < 0.01), there is no consistent response to changes in δ on the horizontal axis (due to the absence of the rb strain). For all simulations presented here, the number of seasons $n_y = 40$ and the kernel parameter $\eta = 2$.

Chapter 3

From explicit to implicit - capturing spatial structure

3.1 Introduction

In the previous chapter, we used a model that featured a spatially explicit between field component in order to demonstrate how the scale of spatial heterogeneity affects the durability of disease resistance. This involved generating spatially explicit point data for the locations of the centres of fields within a landscape, with each field containing one of two different varieties of crop. Using an underlying landscape 'template' that conformed to a specific value of a landscape metric, the edge/area ratio, we demonstrated how the outcome of the epidemic, which we measured via the area under the disease progress curve (AUDPC), is related to the value of this landscape metric. In our specific case we have shown how distributing fields of different crop varieties at smaller scales of spatial heterogeneity through a higher landscape template edge/area ratio improves the control of a disease system with a wild-type and a resistance breaking strain over evolutionary timescales. The edge/area ratio as used in the previous chapter refers to the division of a landscape into simple patterns such as halves and quarters etc. Although simple, this has the clear advantage of being easy to understand, while making abundantly clear the nature of the spatial heterogeneity being generated.

Landscapes and field locations generated in this manner are somewhat artificial and inflexible however, since real agricultural landscapes are clearly not divided up in this way (Margosian et al., 2009). This leads to the question of how to measure or approximate the scale of spatial heterogeneity or degree of aggregation in irregular landscapes that do not necessarily

contain a predetermined number of regularly sized discrete patches for each crop variety. A real agricultural landscape might contain a number of irregularly shaped fields of unequal sizes, only some of which might contain crops that are hosts of a specific plant disease. Realism for its own sake is not necessarily the most vital element of this study, since we are focussing on developing generally applicable theoretical principles. However, it is nonetheless very important to show that the spatial relationships we have described are not specific to the particular and highly artificial form of spatial heterogeneity we have considered thus far.

Another issue relates to the use of templates for the placement of field coordinates. This means that the edge/area landscape metric does not relate directly to the exact spatial locations of the fields, but instead relates only to the background template. Repeated or high density sampling for field coordinates, by using a larger number of fields, is likely to improve the strength of the relationship between the landscape metric and the model outcome. However this requires a great deal more computational time and is still only an approximation. Imagine also the reverse process by which we might take real data for field locations from an aerial photo for example and use it to calculate the value of a spatial metric. Using the edge/area ratio based on an underpinning template, there is currently no obvious way by which a landscape might be divided up into different patches or regions for the different crop varieties. It is therefore impossible for us to calculate our key metric for a real landscape. Given all of this, a much preferable solution would be one that can be used to calculate a landscape metric directly from the exact field locations. This would mean that the results of a model would be directly related to the metric that measures the degree of spatial mixing in the landscape.

Arguably the key lever for a community of growers when attempting to maximise the durability of crop disease resistance is the cropping ratio at which the resistant variety is deployed (i.e. the proportion of fields in which the resistant crop is grown). This was not addressed in the previous chapter. This was partly driven by a desire to investigate the effects of spatially heterogeneity by itself without the complication of an additional variable parameter. The complex and often counterintuitive nature of some interactions in this previous study perhaps justify this approach. However it is ultimately necessary to include the effects of cropping ratio in the study of resistance durability due to the ease with which groups of growers can manipulate this parameter. Other studies have previously shown that the cropping ratio of an agricultural landscape can have pronounced effects on resistance durability (Fabre et al.,

2012; Lo lacono et al., 2013; Papaix et al., 2018). The quantitative interaction between this variable and the scale of spatial heterogeneity in the landscape is still uncharacterised however.

The major reason why cropping ratio was not included in the previous chapter is that the edge/area ratio metric of a landscape is not independent of the cropping ratio. Alternative pre-existing metrics that measure the scale of spatial heterogeneity in the strictly defined manner that we require, and are also independent of the cropping ratio, were not found. This means that we cannot vary the cropping ratio while maintaining the same scale of spatial heterogeneity in the landscape, since moving the cropping ratio away from 50:50 in either direction will, all other things being equal, tend to reduce the length of the edges between the two patch types. The E/A ratio works as proxy for the scale of spatial heterogeneity at any given cropping ratio, but it does not have a consistent definition across different cropping ratios. Dividing the E/A ratio by the cropping ratio of the least used host variety was tested as a potential correction for this effect, however the interdependence of the two parameters remains. To avoid this problem, and to allow the effects of cropping ratio and spatial heterogeneity to be studied independently, any new metric for the scale of spatial heterogeneity or degree of aggregation must therefore be explicitly constructed to work at any cropping ratio. This would allow the cropping ratio and degree of spatial heterogeneity to each be changed independently without any in-built covariance.

There are a wide variety of methods in the mathematical biology literature for considering different kinds of spatial dynamics at different spatial scales. These methods vary in their complexity and ease of implementation, however many are focussed essentially on reducing complex dynamics using simplified mathematical formulations that allow at least some degree of analytical investigation. These contrast to the generally more realistic but less tractable approaches of using coupled systems of ordinary differential equations, as found in chapter 2 (Levin, 1974; Tilman, 1994), or individual based models (Cunniffe et al., 2015). Reaction-diffusion models using partial differential equations (PDEs) are a well known method that incorporate parameters governing rates or birth, death and movement in a population of organisms that vary as a function of both time and space (Holmes et al., 1994). PDEs can be a useful tool for studying questions such as how patch size affects the critical persistence of a population, and also how coexistence of multiple species can be facilitated by differences in dispersal ability (Kierstead and Slobodkin, 1953; Shigesada et al., 1979; Skellam, 1951). These models are however focussed on the movement of organisms through continuous space,

with necessarily local dispersal, and as such are not the best method for modelling agricultural landscapes with discrete host patches and potentially long distance aerial dispersal of inoculum.

In systems where the local spatial interactions of individuals within a population are of vital importance, models based on spatial moment equations can be used to reduce spatial information to more implicit levels of consideration (Bolker and Pacala, 1997). These models seek to describe the overall behaviour of a system in terms of the average population density (1st moment) and the correlation densities of pairs and triplets of individuals within a population (2nd and 3rd order moments respectively). Higher orders of spatial interaction are usually ignored through the use of what is known as a moment closure approximation, approximating the true density of higher order moments by a simple function of lower order moments (Morozov and Poggiale, 2012; Webb et al., 2007). Pairwise approximation equations are a variant use of spatial moments that focus on local dispersal and nearest neighbour interactions within discrete space (Filipe and Gibson, 1998; Matsuda et al., 1992). While such methods are powerful and fit well to spatially explicit models in some cases, they can be difficult to implement and understand, and might not be the most applicable method when larger scale interactions between populations of organisms are being considered. Spatial moment approximations are also generally reliant on the assumption that there is no environmental heterogeneity, and focus instead on dynamic spatial interactions between individuals over time (Morozov and Poggiale, 2012). They are therefore of limited use when considering the impact of static spatial structure in a host population.

At the extreme end of model simplicity and implicit characterisation of space, is what is often known as the 'modified mean field approach' using non-linear incidence functions (Liu et al., 1986). These are very similar to standard non-spatial models with random mixing, but with the contact rates within the system modified to somehow take into account the density-dependent effects of spatial clustering (Gubbins and Gilligan, 1997; Pascual et al., 2011; Roy and Pascual, 2006). Commonly used possibilities involve replacing the 'standard' mass action transmission term βSI by a nonlinear proxy involving additional parameters, e.g. $\beta SI/(1 + \gamma I)$ or $\beta S^\gamma I^\delta$. The approach is highly phenomenological, and requires the fitting or estimation of a number of 'mixing parameters' (i.e. δ and/or γ above) with somewhat ambiguous definitions. Certainly there is no clear prescription for how the mixing parameters should be chosen to match a given scale of spatial heterogeneity while accounting for interactions between species. Other strategies for the implicit consideration of space include aggregation

methods that describe the distribution of multiple species between different patches as a function of certain global parameters such as the population sizes of these species (Iwasa et al., 1987). This approach is linked to early ideas by Hassell and May (1973) about the distribution of parasitoids within a host population, and is related to the ecological concept of the ideal free distribution (Lampert, 2005). However again there is no clear guide from the literature on functional forms or parameter groupings to adopt for the question under consideration here. We therefore reject these methods as unworkable for the problem in hand.

Returning to the model used in the previous chapter, the results demonstrate that the key dynamics are driven by the frequency with which the two different pathogen strains infect the host varieties to which they are well-specialised, relative to the frequency with which they infect their less preferred host variety. These relative frequencies are determined by the overall spatial proximity of the two host varieties in the landscape to each other. This is because there will be a great deal more cross infection of the resistance breaking pathogen strain infecting the susceptible host, for example, when fields with resistant hosts are located close by on average to fields with susceptible hosts, so that they can act as a major source of resistance breaking inoculum. The importance of within and between variety interactions suggests a useful simplification of the system, that is more specifically suited to the representation of host heterogeneity than any of the modelling approaches described above. The relative frequency with which infections on one host variety lead to infections on another, as opposed to further infections within the original host variety population, can be included in a simple two patch model. These two patches would not represent real discrete spatial patches, but instead all fields of a given type. The rates of spread between patches would represent the degree to which two host populations are intermixed. A significant challenge in constructing this form of model, however, lies in developing a theoretical framework that intuitively links together the various rates at which infections occur between and within the patches in the model. Developing and testing this theoretical framework for the capture of spatial information within a simple model is the focus of this chapter. In order to simplify this task, and focus primarily at this stage on the spatial structure of the model rather than questions of resistance durability, we use a simplified model with two host varieties but only one generalist pathogen strain. We will use the model in the context of resistance durability in the next chapter.

3.2 Methods

3.2.1 Developing a suite of test landscapes of varying degrees of aggregation

In order to test any new metric that attempts to capture the scale of spatial heterogeneity in a landscape, it is important that we have a range of at least semi-realistic landscapes available. Sourcing real agricultural landscape data is impractical at the current stage for this purpose as there are factors, considered in the chapter discussion, that would complicate the estimation of spatial heterogeneity without further model development. The primary issue with real data is the likely uneven distribution of host crop fields in space, with additional space taken up for non-host crops or other uses. Since there is a wide variety of possible spatial characteristics, only some of which we are primarily concerned with analysing here, we needed to find a way to generate a range of hypothetical landscapes that systematically vary only in their degrees of spatial aggregation and cropping ratios. In order to achieve this, we constructed a cellular automaton that begins with a randomly generated, highly mixed, landscape at a specified cropping ratio, before gradually aggregating over time according to a simple set of rules that govern the behaviour of the cells in the landscape. As aggregation progresses, snapshots of the landscape are taken at specific values of the edge/area ratio. The edge/area ratio was calculated in a similar manner as in chapter 2, but with the interior edge of the more complex generated landscapes derived from the total.edge value calculated by the ClassStat function of the R package SDMTools. Since the edge/area ratio is being constantly updated, a time series of related landscapes is produced that varies from highly mixed to highly aggregated. This approach differs from a global optimisation method such as simulated annealing (Kirkpatrick et al., 1983; Papaix et al., 2014a) in that decisions for changes in the landscape are not based on the value of the output landscape metric at any given time, but instead progress due to an unchanging set of rules that have been designed to ultimately move the system towards a certain state or outcome. A shared benefit of this approach with simulated annealing is that both algorithms allow cell changes that move the output metric in the undesired direction, which in this case means moving the landscape towards lower levels of aggregation. In optimisation terms, this prevents the landscape from becoming stuck in patterns that produce local minima for the edge/area ratio aggregation metric, and increases the likelihood that something close to the global minima of the metric, giving almost complete

aggregation, will be reached. It is important to note that although the algorithm we have developed is both powerful and flexible, the precise nature of how the landscapes are generated is relatively unimportant to the epidemiological model presented in this chapter. All that is needed is a workable method for generating a series of random landscapes that could potentially represent the spatial heterogeneity of a host plant population in a relatively realistic manner.

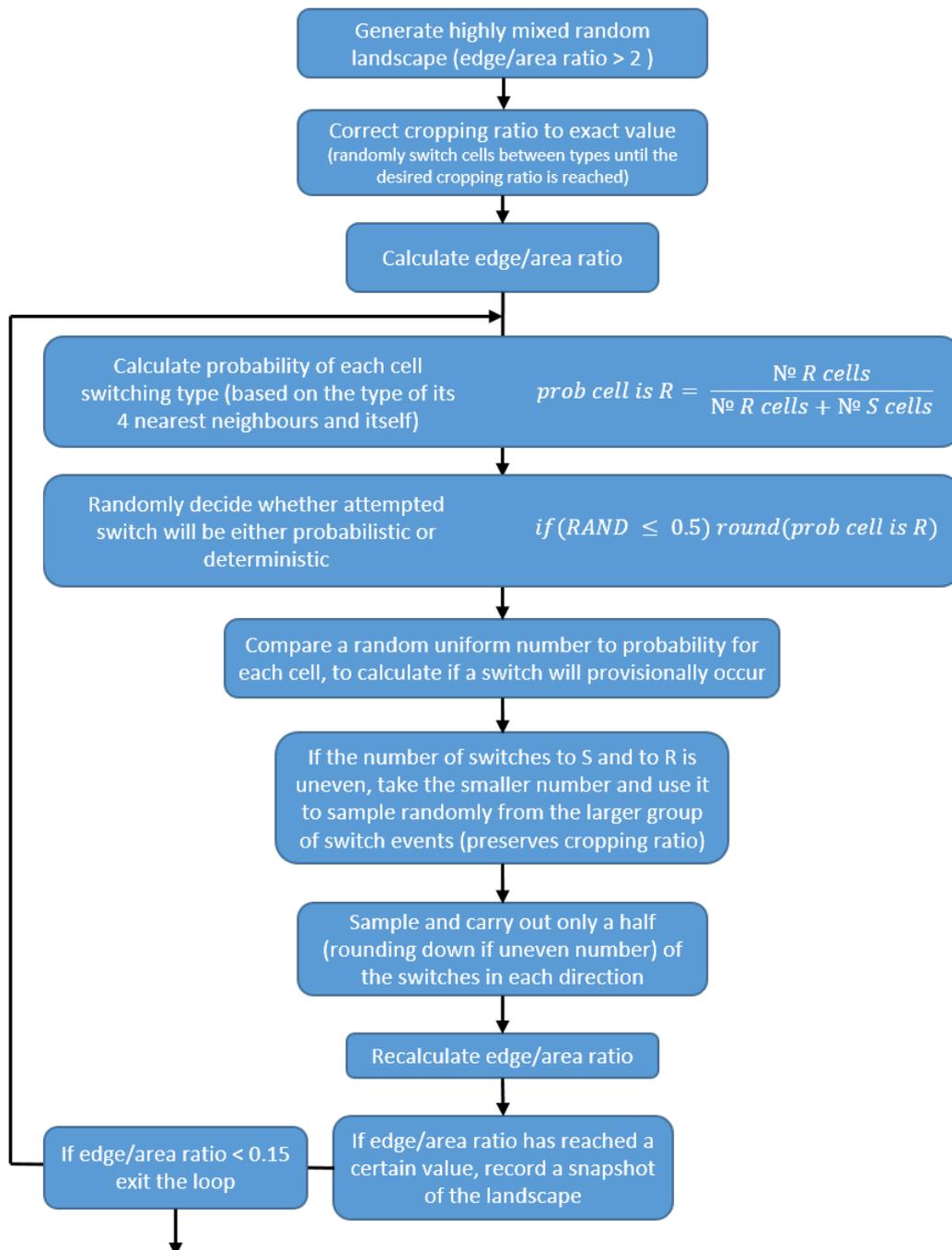


Fig. 3.1 Flow diagram showing the cellular automaton irregular landscape generator algorithm.

The algorithm begins by randomly allocating landscape cells as either susceptible (S) or resistant (R) according to the cropping ratio. A number of such landscapes are generated until one is found that has an edge/area ratio of > 2 , so as to start with an adequately high level of mixing. Furthermore, as each initial random generation likely gives a landscape with a slightly different cropping ratio, individual random cells are switched between types until the desired cropping ratio is reached. From this point, the probability of each cell either remaining as or switching to R in the first step is calculated, based on a cell's own type and that of up to four nearest neighbours (Fig. 3.1). Then with a 50% chance this probability may be rounded up to 1 or down to 0. This step is done because it allows the algorithm to strike a balance between deterministic and probabilistic behaviour, which in turn maximises the overall speed of aggregation while avoiding local minima which stall the algorithm's progress. This set of probabilities is then used to provisionally calculate whether or not a switch in type occurs for each cell in the landscape. Inevitably the number of switches to S and to R will not be equal, so in order to maintain the cropping ratio the number of switches from the smaller of these two groups is used to sample from the larger group. From the potential switches from each of these two now evenly sized groups, 50% of these are then sampled as the final switches for this step that will actually take place. This is done because having a very high number of nearby synchronised switches in the landscape tends to lead to transient structures such as repeatedly inverting chessboard patterns in certain local regions. Such structures, which are completely disrupted by increasing the level of asynchrony, would slow down the overall progress of the algorithm and produce artificial looking patterns. Once a set of cell switches has taken place, the edge/area ratio of the landscape is recalculated and the aggregation step is then repeated, continuing until further aggregation in the landscape no longer occurs within a practical timeframe. To generate a diverse library of example landscapes, we ran the cellular automaton a total of 20 times, recording snapshots of the aggregating landscapes every 0.05 E/A ratio interval until an E/A ratio of approximately 0.15 was reached. This produced a total of 3699 landscapes (the number of landscapes recorded in a given algorithm run depends on the exact E/A ratio of the initial randomly generated landscape).

3.2.2 Representing pathogen spread in a two-patch model

The two-patch model

In a simple initial case with only one strain of pathogen present, we can represent the progress of disease on two, equally susceptible, host varieties with a two patch metapopulation model (Fig. 3.2). Here, N is the overall number of host tissue units, and the relative sizes of the two host patches is determined by the cropping ratio ϕ . As stated earlier, the degree of aggregation or scale of spatial heterogeneity in a landscape determines the frequencies with which infections on one host variety cause infections on the other host variety, relative to the frequency with which they cause further infections on the same host variety. These relative frequencies have been identified in chapter 2 as key components driving the effects of spatial heterogeneity. The frequency with which infections Y in the first host patch cause infections Z in the second host patch is given by p_1 , while the frequency with which infections Y in the first host patch lead to further infections in that same host patch is given by $1 - p_1$. The same pattern also applies to the infections in each host patch caused by existing infections Z in the second host patch using the rate parameter p_2 .

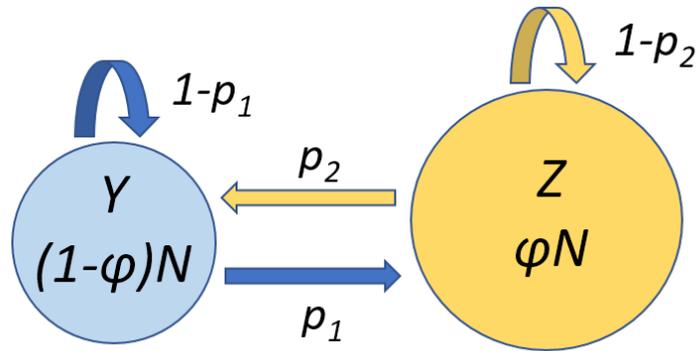


Fig. 3.2 Diagram representing the structure of the two-patch epidemiological model. The arrows represent the probabilities with which produces inoculum disperses within and between host variety patches. The state variable Y is the number of infected hosts in the left-hand side patch with population size $(1 - \phi)N$, while Z is the number of infected host in the right-hand side patch with population size ϕN .

The epidemiological system described above is expressed as an *SIS* model in a system of differential equations:

$$\frac{dY}{dt} = \beta \left(\frac{(1-\phi)N - Y}{(1-\phi)N} \right) ((1-p_1)Y + p_2Z) - \eta Y, \quad (3.1)$$

$$\frac{dZ}{dt} = \beta \left(\frac{\phi N - Z}{\phi N} \right) (p_1Y + (1-p_2)Z) - \eta Z, \quad (3.2)$$

Here β is an infection rate term that encapsulates the rate of inoculum production by an infected plant ω , multiplied by the probability that infection occurs σ , once the inoculum has dispersed and landed on an uninfected host tissue unit. The coefficients related to p_1 and p_2 described above govern the relative rates of inoculum dispersal within and between the two host types, while the probability that the host tissue the inoculum disperses to is uninfected is given in the large set of parentheses in each equation. This proportion of uninfected hosts must be given in frequency dependent terms as a proportion of the relative host population sizes in the two patches. This is because these proportions are conditional on the inoculum having already landed on a particular host type, the probability of which is separately governed by p_1 and p_2 . The semi-discrete seasonal component of the spatially explicit model from the first chapter is replaced by continuous harvesting and replanting, the rate of which is governed by the parameter η .

Linking the dispersal between types of field to ϕ and p

While the above pair of equations describes how the two-patch model works at a basic level, it does not incorporate any individual parameter that governs the scale of spatial heterogeneity or degree of aggregation in the host landscape that the model is approximating. To address this, we replace the rate coefficients p_1 and p_2 with expressions that use the cropping ratio ϕ and a new parameter p to link together the four coefficients in Eqns (3.1) and (3.2).

$$\frac{dY}{dt} = \beta \left(\frac{(1-\phi)N - Y}{(1-\phi)N} \right) ((1-p\phi)Y + p(1-\phi)Z) - \eta Y, \quad (3.3)$$

$$\frac{dZ}{dt} = \beta \left(\frac{\phi N - Z}{\phi N} \right) (p\phi Y + (1-p(1-\phi))Z) - \eta Z, \quad (3.4)$$

$$\mathbb{P}(\text{Inoculum lands on } Y | \text{Emitted by } Y) = (1-p) + p(1-\phi) \quad (3.5)$$

$$= 1 - p\phi \quad (3.6)$$

$$\mathbb{P}(\text{Inoculum lands on } Z | \text{Emitted by } Y) = p\phi \quad (3.7)$$

$$\mathbb{P}(\text{Inoculum lands on } Y | \text{Emitted by } Z) = p(1-\phi) \quad (3.8)$$

$$\mathbb{P}(\text{Inoculum lands on } Z | \text{Emitted by } Z) = (1-p) + p\phi \quad (3.9)$$

$$= 1 - p(1 - \phi) \quad (3.10)$$

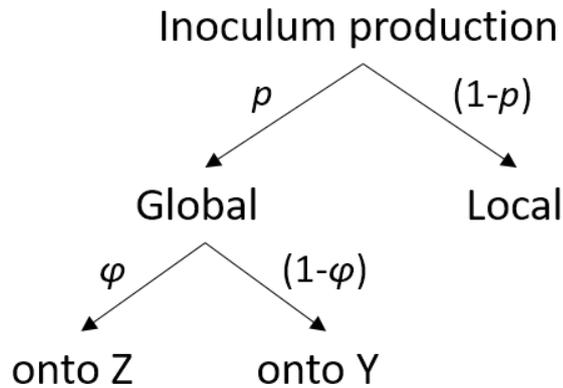


Fig. 3.3 Probability tree showing how the inoculum produced in a given host patch may potentially be dispersed. If dispersal is local the inoculum will remain on the same patch type, whereas if it is global the host type it lands on will be governed by the cropping ratio ϕ .

The parameter p represents the proportion of inoculum produced by a given host variety that disperses onto the two host varieties in the landscape at a ratio equal to the cropping ratio. This parameter is used to form the infection four dispersal probability coefficients in Eqns 3.3 and 3.4 using the probability tree in Fig. 3.3 and Eqns 3.5-3.10. If for example $p = 1$, then all inoculum produced in the landscape disperses according to the cropping ratio, and therefore there is no preferential dispersal to hosts of the same variety as you would expect if the varieties were at all aggregated. In this way $p = 1$ represents a non-spatial model with dispersal governed by mass action only. In the case where $p = 0$ however, no inoculum disperses according to the cropping ratio, and all preferentially disperses onto hosts of the same variety in a locally aggregated patch. This case would represent two completely aggregated and separated host variety populations where no cross infection occurs. In reality, a real host landscape with any degree of spatial structure would give p values between 0 and 1, as there is likely to always be some degree of cross-infection occurring between varieties. The expressions for these coefficients capture the fact that as p is reduced from 1, an increasing proportion of inoculum disperses onto the same variety that it originated from, as the size of local aggregated patches increases. It should be noted that the model in Eqns (3.3) and (3.4) reduces to a simple SIS model when the values of Y and Z are proportional to the cropping ratio ϕ (Appendix 1 in section 3.5). This is because in this case, where the pathogen is equally able to infect both host varieties, the spatial structure of the host population is essentially invisible to pathogen population, and therefore disease progress is independent of both p and ϕ . The proportional

distribution of infected individuals between host varieties is equivalent to the even distribution of infection across the landscape.

For the relationship between the four coefficients for Y and Z in Eqns (3.3) and (3.4) to hold, it is assumed that the size of the landscape is very large compared to the size of any local aggregated patches. This is required because the effective cropping ratio that determines dispersal outside of the locally aggregated patch must equal the overall cropping ratio, regardless of the size of the locally aggregated patch. In a small landscape with large scales of aggregation this assumption would not hold, however adjusting for this would require explicitly accounting for the finite scale of the landscape, which is neither practical nor necessary for an initial theoretical model such as this. It is also assumed that dispersal off the edge of the landscape produces no edge effects, and that there is no spatial variation in non-specific host field density. Breaking either of these assumptions would mean that the amount of inoculum lost in the gaps between fields or over the edge of the landscape would likely not be proportional to the cropping ratio of the two host varieties.

Estimating the value of p

For the sake of simplicity, when explaining the estimation for p we will describe the two patches as susceptible (S) and resistant (R), with Y and Z infected individuals respectively. The aggregation parameter p , in combination with the cropping ratio, ties together the four infection coefficients in this two-patch model, while giving an intuitive and specific explanation for the spatial effect this parameter captures. It would however be very difficult to directly measure the value of p in a landscape, as this would require defining discrete patches or regions in the landscape where inoculum disperses preferentially onto each host variety. Thankfully we can estimate p very easily in a theoretical model by using a dispersal kernel to calculate for each host variety the probability at which inoculum disperses onto the same variety and the probability at which it disperses onto the other variety. These four directional pairwise dispersal probabilities are the same as the infection coefficients p_1 , $1 - p_1$, p_2 and $1 - p_2$ used in Eqns. (3.1) and (3.2). Once we have these probabilities we can calculate the aggregation parameter using $p = p_1/\phi$. Note that p can be derived from any of the four probability coefficients, and will give the same value as long as the assumptions of the model are upheld. In reality these assumptions are likely to be broken when estimating p from a finite landscape, and so deriving p from the four coefficients will give slightly different results. Taking the mean value of these

for the estimation of p could work as a method to minimise this error. We do not do this here because the measurement of this error is a key feature of our test for the independence of p and the cropping ratio ϕ .

In order to calculate the rate at which susceptible hosts disperse onto resistant hosts for example in our generated landscapes, we would first need to sample locations for individual susceptible and resistant ‘fields’ within the relevant patch types. These ‘fields’ could represent whole or partial crop fields, depending on the overall scale of the landscape being represented. Assigning a field to every possible spatial coordinate within the landscape would give us the most accurate estimation for p , however random coordinate sampling (a poisson point process) using a smaller number of fields over a number of p estimation replicates could potentially give equivalent results using fewer computational resources. We would then calculate each pairwise distance between all susceptible and resistant fields before feeding each of these values into a dispersal kernel (in our case a negative exponential kernel). The mean dispersal distance of this kernel would be dictated by the particular disease being studied. The sum of these kernel weighted distances between susceptible and resistant fields (Eqn 3.12) would then be divided by the sum of the kernel weighted distances between susceptible fields and all types of field (Eqns 3.11 and 3.12). This gives the proportion of overall inoculum dispersing from susceptible fields that specifically lands on resistant fields. The total kernel weighted distances between the different types of field are calculated as follows:

$$\text{Total } S \text{ to } S = \sum_{x=1}^{n_f} \sum_{\substack{z=1 \\ z \neq x}}^{n_f} K[z, x], \text{ if } z \text{ is } S \text{ and } x \text{ is } S \quad (3.11)$$

$$\text{Total } S \text{ to } R = \sum_{x=1}^{n_f} \sum_{\substack{z=1 \\ z \neq x}}^{n_f} K[z, x], \text{ if } z \text{ is } S \text{ and } x \text{ is } R \quad (3.12)$$

$$\text{Total } R \text{ to } S = \sum_{x=1}^{n_f} \sum_{\substack{z=1 \\ z \neq x}}^{n_f} K[z, x], \text{ if } z \text{ is } R \text{ and } x \text{ is } S \quad (3.13)$$

$$\text{Total } R \text{ to } R = \sum_{x=1}^{n_f} \sum_{\substack{z=1 \\ z \neq x}}^{n_f} K[z, x], \text{ if } z \text{ is } R \text{ and } x \text{ is } R \quad (3.14)$$

$$(3.15)$$

where $K[z, x]$ is the dispersal kernel weighted distance from field z to field x .

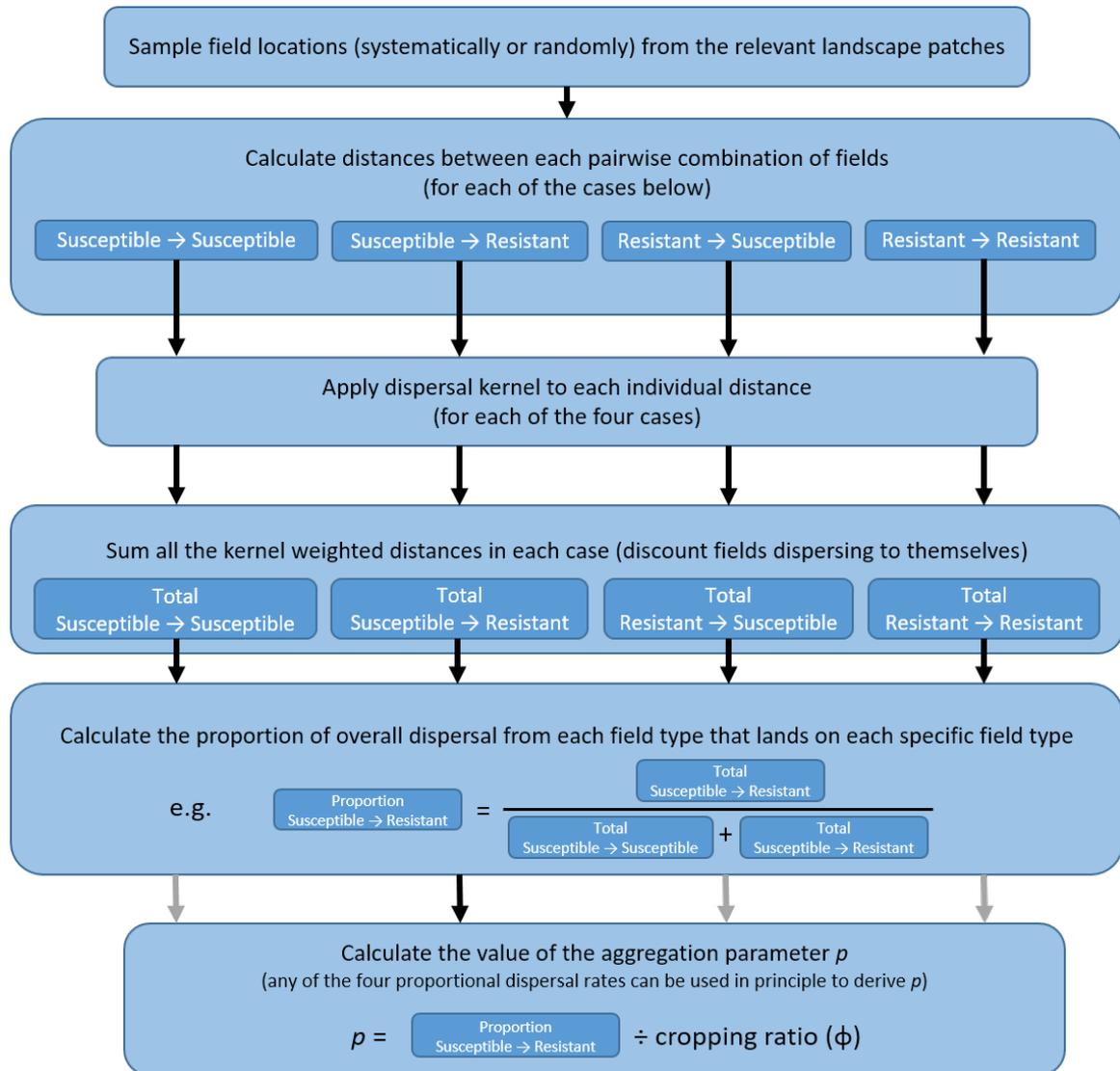


Fig. 3.4 Flow diagram with pseudoequations showing the estimation process for the aggregation parameter p . The faded out arrows represent the alternative ways in which the value of p can be derived from the expressions for the dispersal probability coefficients in the model. In all cases presented in this chapter the value of p is however estimated from the proportion of inoculum dispersing from susceptible fields that infects resistant fields. The calculation for the total kernel weighted distances between the different types of field is also shown in Eqns 3.11-3.14.

Testing the independence of p and ϕ

As was stated earlier, it is important that a metric attempting to capture the scale of spatial heterogeneity or degree of aggregation in a landscape functions independently of the cropping ratio. The aggregation metric p has been specifically designed so that the same value has a consistent meaning across all possible cropping ratios, however for this to be tested we need to use generated landscapes of different cropping ratios that are already known to have exactly the same degree of aggregation. This can be achieved simply by taking any given generated landscape and ‘flipping’ the field varieties that belong to the different patches, meaning that

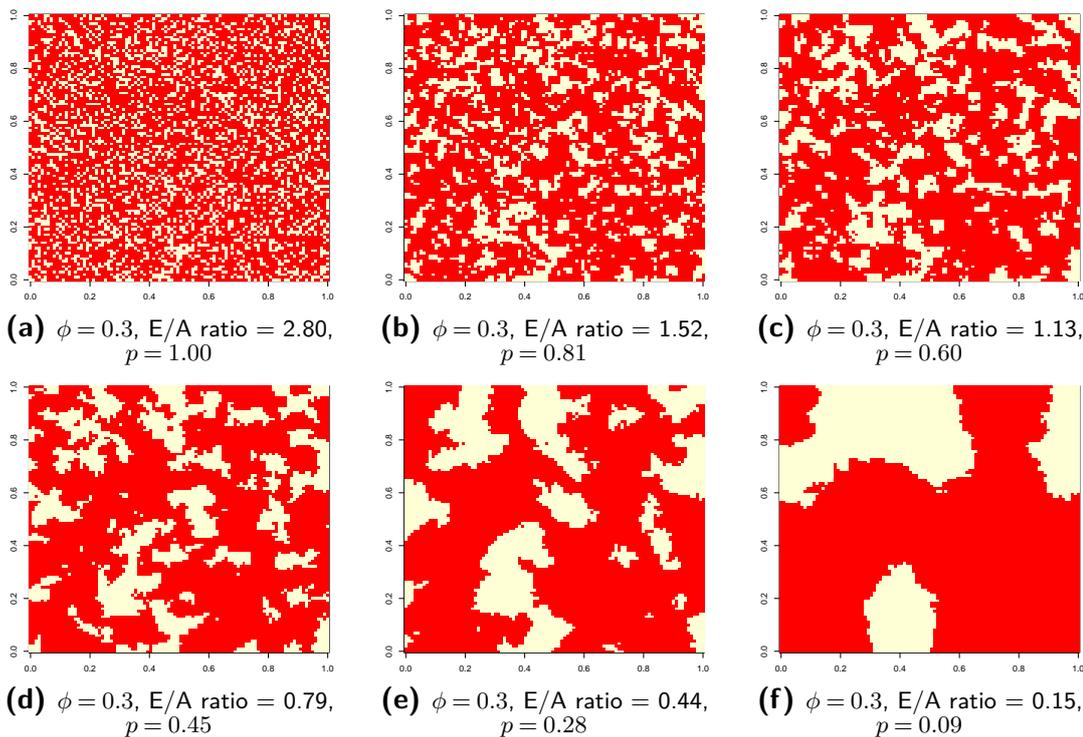
susceptible fields become resistant and resistant fields become susceptible. Using the method described in Fig. 3.4, the estimated values for p in pairs of these 'flipped' landscapes can then be compared to see if they match. As was stated earlier, this p comparison test is the reason why here we do not take the mean value from the four different probability coefficients that can be derived in Fig. 3.4. The proportion dispersing from resistant to susceptible in a 'flipped' landscape is the same as the proportion dispersing from susceptible to resistant in its 'non-flipped' counterpart. This means that if we took the mean of the p estimates taken from the four possible coefficients for each of the 'flipped' and 'non-flipped' landscapes, the resultant p values would always be identical.

3.3 Results

3.3.1 Landscape generation

As can be seen from Fig. 3.5, the cellular automaton algorithm is capable of generating a wide range of unique landscapes, with different cropping ratios and scales of spatial heterogeneity/degrees of aggregation. The patches produced as the cells in the landscape aggregate during the algorithm run are irregular in shape with somewhat rough edges. Generating patches with either a smoother or a more fragmented shape is possible by respectively increasing or decreasing the degree of determinism in the transformation of cells over time. This is done by increasing or decreasing the probability (currently 50%) that an attempted patch type switch is carried out deterministically (see Fig. 3.1). We consider the current balance within the algorithm as shown here to be reasonable for our purposes however, as excess determinism or stochasticity can both limit the overall degree of aggregation which is possible within a reasonable computational time frame. In the case of too much determinism, patches aggregate very rapidly in the initial stages of each algorithm run, but can then become highly resistant to further aggregation as the smoothed patch edges make further shape changes very slow and difficult. This scenario is somewhat analogous to a classical optimisation algorithm becoming trapped at a local minimum. In the the case of excess stochasticity, the continual disruption of patch shape by the introgression and spread of the other cell type simply hinders aggregation by increasing the number of required time steps for a certain level of aggregation to be achieved.

Run 1



Run 2

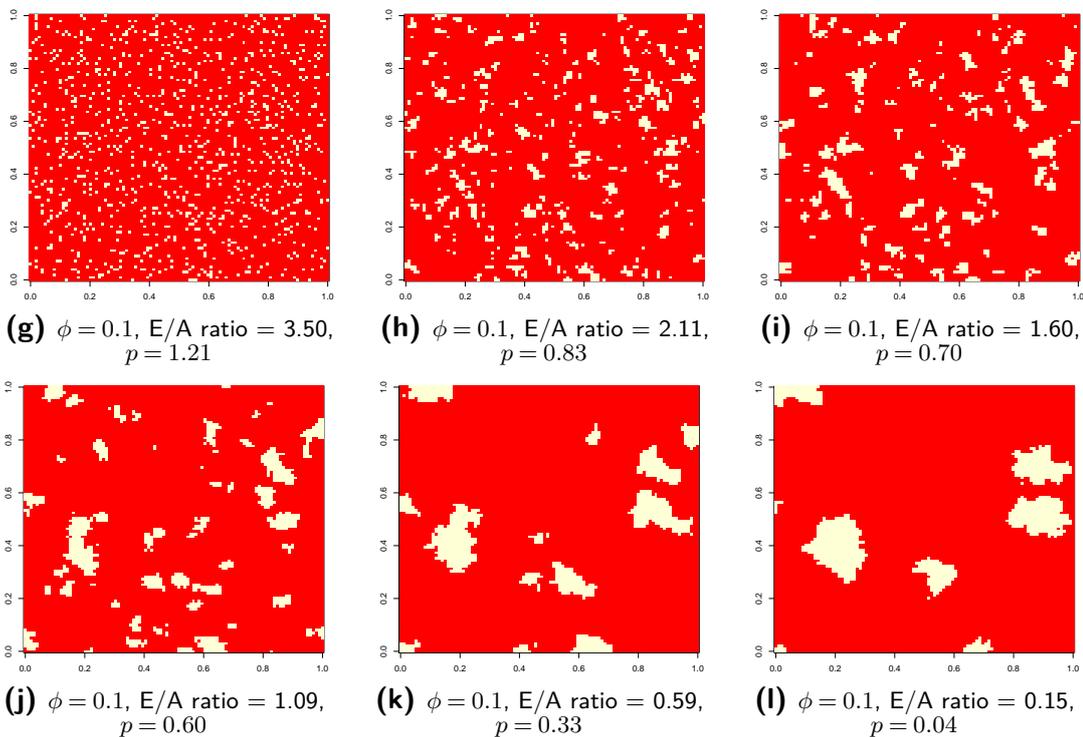


Fig. 3.5 Examples of host landscapes generated using the cellular automaton algorithm. Landscape run 1 (a-f) and run 2 (g-l) are each a time series of snapshots taken from single iterations of the algorithm run. The cropping ratio ϕ and edge/area ratio values are shown for each landscape. The landscape snapshots in each group were chosen to show a roughly evenly spaced time series between the minimum and maximum (to the extent that practicality allows) levels of aggregation. The estimated p values for each landscape are given for a single set of field sampling in each landscape with the number of fields $n_f = 100$ and a negative exponential dispersal kernel with mean dispersal distance = 1. The landscape dimensions are 100x100.

3.3.2 Estimation of p and its independence from the cropping ratio

ϕ

It should be noted that when estimating the value of the aggregation parameter p , it is possible for values > 1 to be returned (Fig. 3.5g). This is despite the theoretical definition where $0 \leq p \leq 1$, as it is assumed that the maximum level of mixing in the landscape is where dispersal occurs in a manner proportional to the cropping ratio ϕ . A case where $p > 1$ refers to a situation where a greater proportion of inoculum dispersed from susceptible fields lands on resistant fields than would be expected if dispersal were proportional to the cropping ratio. As seen in Fig. 3.5g, this can occur at low cropping ratios and very high degrees of mixing due to the highly isolated nature of the susceptible patches in the landscape. The fields within these patches will then disperse proportionally very little inoculum to other susceptible fields, an effect which is exacerbated by using a tight dispersal kernel. There is also a higher probability in these cases that individual susceptible and resistant fields will randomly be located very close to one another, and so with a tight kernel there could be proportionally very high dispersal between these fields. Deriving the estimations for p from all four infection probability coefficients and taking the mean value reduces the impact of this effect and brings p closer to 1 in these cases.

Effect of cropping ratio ϕ

As stated earlier, pairs of aggregation parameter p estimates were taken from each generated landscape and its inverted or 'flipped' counterpart. The distribution of differences between these pairs of p estimates (flipped p - 'non-flipped' p) for all generated landscapes, covering a the full range of E/A ratios, is shown for a range of cropping ratios (ϕ) in Fig. 3.6. In these plots, if the estimation of p were completely accurate, the differences between the pairs of p values should be 0, as the inverted nature of the landscape pairs means that they in fact have identical levels of aggregation, regardless of their cropping ratios. These results generally show that the random sampling of field locations (in this case 1000 fields within a 100x100 landscape) can create significant variation in the accuracy of p estimation. This is because even though both landscapes in a pair use the same general field locations (ignoring the field types), the spatial distribution and density of fields is not consistent across the landscape, and therefore is not consistent between the two patch types. This means for example that if the susceptible fields are by chance closer on average to other fields than the resistant fields

are, then more of their dispersed inoculum (measured using the kernel weighted between field distances) will land on other fields rather than falling into the gaps between fields.

The accuracy range for p estimation is greater for more uneven cropping ratios (the lower cropping ratios shown in Fig. 3.6), and for a tighter dispersal kernel (Fig. 3.6a). The first of these observations is likely because uneven cropping ratios produce more isolated fields of the less frequent crop variety that may, due to the low density random field location sampling, be very far away from another field of the same variety. This would mean that there is very little dispersal between fields of the less frequent variety, and therefore that a very high proportion of dispersal lands on the other variety. This would in turn lead to an overestimate of p in flipped landscapes, leading to a more positive difference between this estimate and that from the 'non-flipped' landscape (where p is calculated from the dispersal from the more frequent variety). Using a very tight dispersal kernel magnifies the importance of small scale spatial effects that result from the random field sampling process, so that fine differences in field density between the different variety patches can have large effects on p estimation.

Across all 3699 pairs of generated landscapes however the mean difference between paired p estimates is close to 0, indicating that in general the p value of a landscape works independently of the cropping ratio as a true measure of aggregation in a landscape. When we increase the mean dispersal distance of the kernel from 1 (Fig. 3.6a) to 3 (Fig. 3.6b) we find that the p estimation for the 'flipped' landscape (where the resistant crop is the most frequent) is on average higher than for the 'non-flipped' version, particularly for more uneven cropping ratios. This systematic difference is mostly caused by differences in p estimation at greater degrees of aggregation (lower E/A ratios in Fig. 3.7). It is the nature of the landscape aggregation algorithm that highly aggregated patches of the lower frequency variety will tend to form more frequently near the edges of the landscape as the algorithm progresses (Figs 3.5f, 3.5l). This is because the edges of the landscape partially protect an aggregated patch from disruption and break up due the random introgression of the other patch type. These landscapes will also have a particularly low E/A ratio and p value due to the small contact surface between the two patch types. The proportional over representation on average of the less frequent patch type near the landscape edges consequently means that more inoculum from this patch type is lost to the empty space beyond these edges. This lost inoculum means that, for the 'flipped landscapes' the overall magnitude of dispersal from the minority susceptible fields is reduced, below the level that would be expected proportional to the cropping ratio. Given that

the magnitude of dispersal between the two field types will always be equal in the 'flipped' and 'non-flipped' landscapes (as it is the same set of measurements), this means that the proportion of overall dispersal from the susceptible fields that lands on the majority resistant fields is increased. This increased proportional dispersal from susceptible to resistant fields in the flipped landscapes is what then drives the increased p value estimation. The idea that the over estimation of p at cropping ratios over 0.5 (the 'flipped' landscapes) is driven by edge effects is supported by Fig. 3.6c, which shows that the systematic differences in p estimation are no longer clearly apparent when we model each landscape as a torus (by wrapping around the landscape edges so that they connect to the opposite side) to remove edge effects. The range in p estimation accuracy is also reduced in this case compared to when inoculum is being lost over the landscape edges.

The effect of aggregation on p estimation accuracy

We can also look at the accuracy of p estimation as a function of the degree of aggregation as measured by the edge/area ratio (Fig. 3.8). This shows that for a very tight dispersal kernel (Fig. 3.8a) the estimation accuracy is greatest in highly aggregated landscapes, but decreases as the field varieties become more mixed. This is because small patches in a highly mixed landscape are effectively highly isolated when the dispersal kernel is very tight (Fig. 3.9a). In a similar manner as to how uneven cropping ratios decrease estimation accuracy, this would mean that there is very little dispersal between fields of the less frequent variety, which would in turn lead to an overestimation of p in 'flipped' landscapes. For a wider dispersal kernel (Fig. 3.8b) this same pattern follows up to an intermediate degree of between field mixing, with accuracy subsequently increasing as mixing is further increased. This directional change in trend at moderate E/A ratios is likely due to the closer proximity of the small patches of the less frequent variety to each other at greater degrees of between field mixing. With a greater mean dispersal distance, this increased proximity with greater mixing is enough to increase connectivity between fields of this type, despite the fact that the individual patches are getting smaller.

Effect of sampling density

The landscape that produced the largest difference in estimated p values between the inverted landscape pairs from Fig. 3.8a, where we use a tight dispersal kernel and a flat landscape with

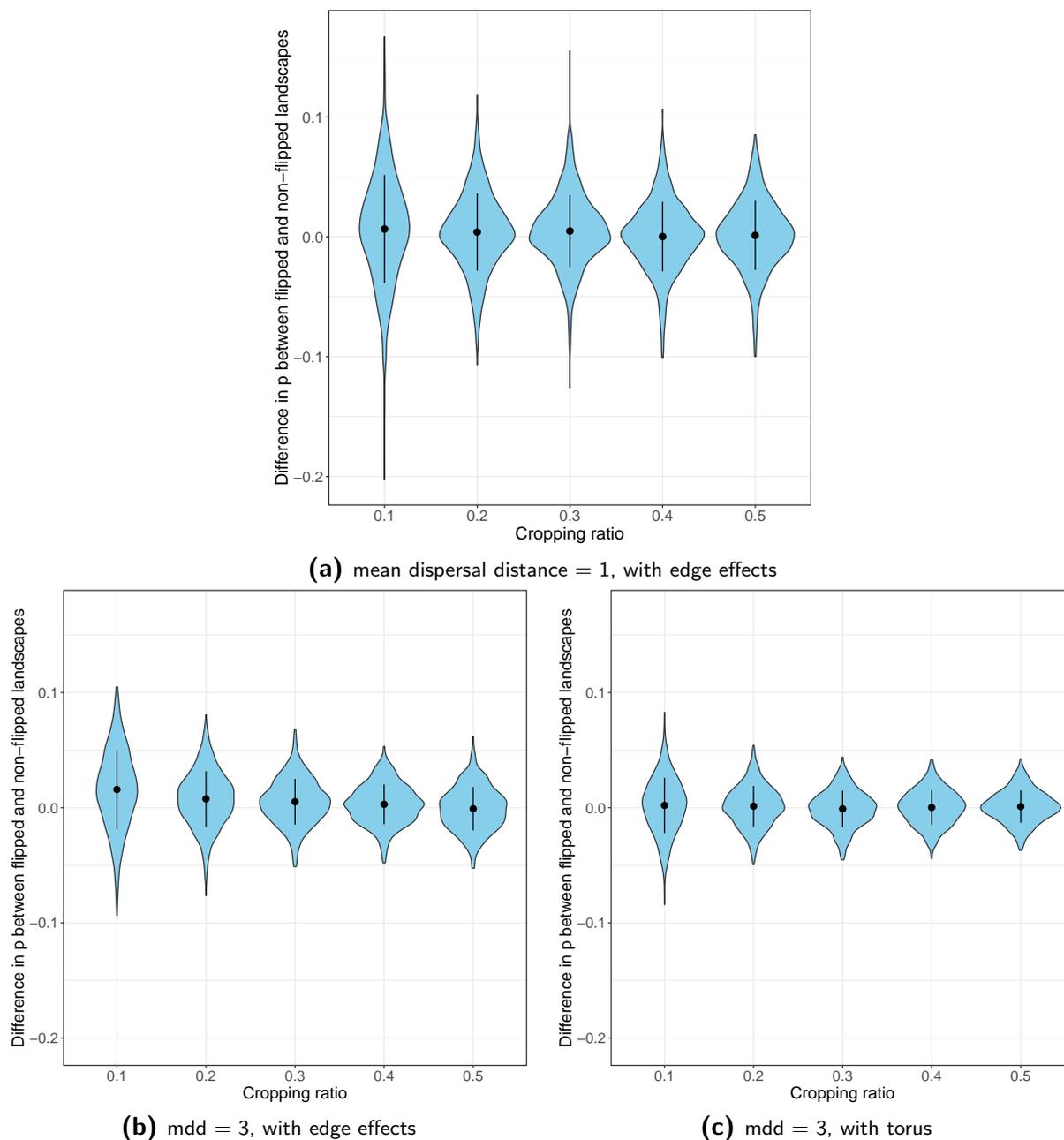


Fig. 3.6 Violin plots showing for different cropping ratios (ϕ) the distribution of differences between pairs of aggregation parameter p estimates (the maximum possible difference being 1). Each pair of estimates is taken from a given landscape (with $\phi = x$) and its inverted or 'flipped' counterpart (with $\phi = 1 - x$). All p estimates in these plots were generated using random sampling of 1000 field locations for each landscape. The central points give the mean difference and the error bars show ± 1 standard deviation. The mean dispersal distance (mmd) of the kernel is 1 in (a) and 3 in (b) and (c) (within landscape dimensions 100x100). Plots (a) and (b) were generated from the landscapes modelled as a flat areas with edge effects, whereas in (c) each landscape is modelled as a torus in order to remove any edge effects.

edge effects, is shown in Fig. 3.9a. The large difference in p values here of over 0.2 is only the result of a single set of randomly sampled field locations, however the very uneven cropping ratio and low level of aggregation suggests that this landscape is also generally likely to give inaccurate estimations for the true p value. We use this flat landscape to show that even very

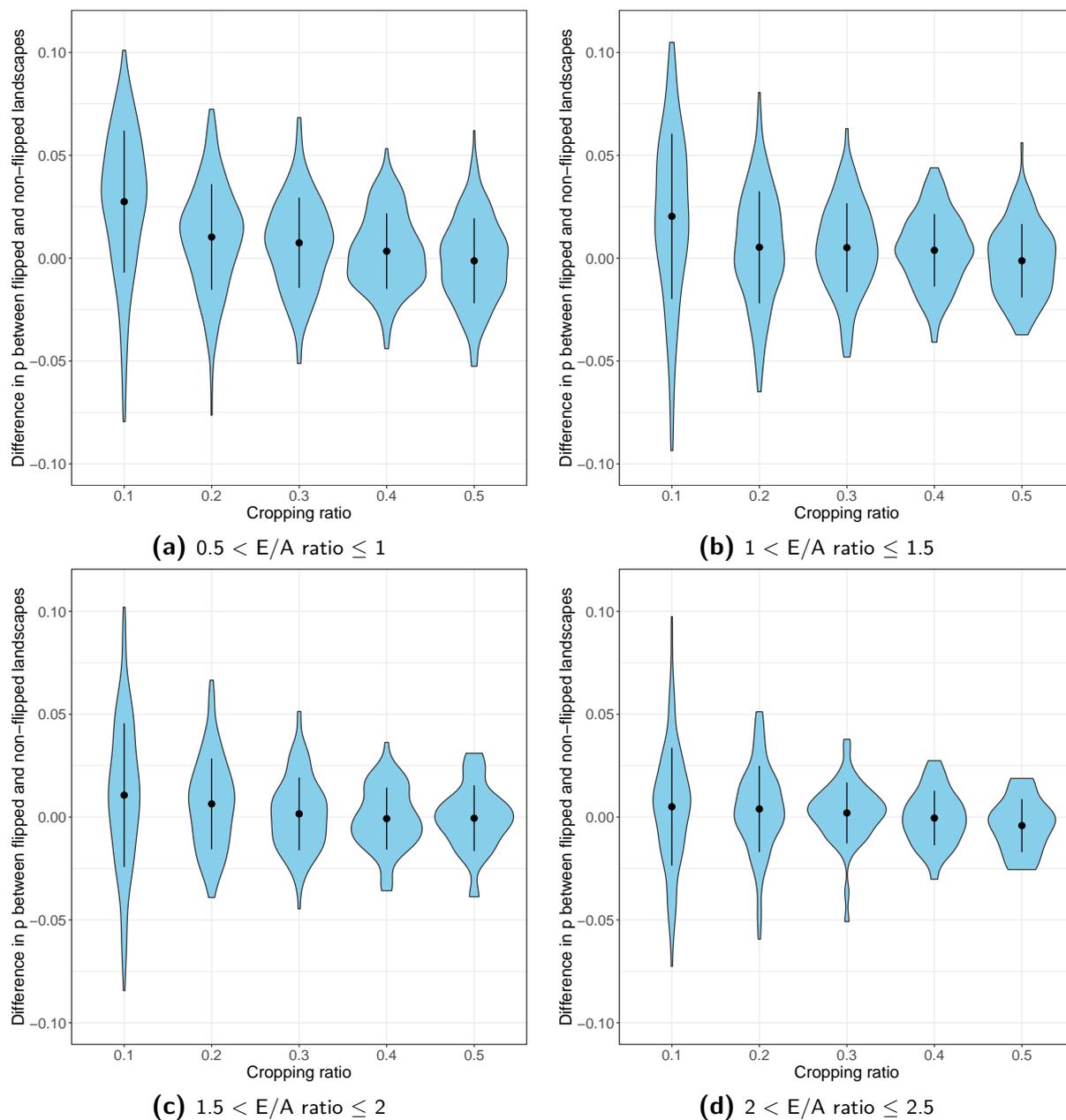


Fig. 3.7 Violin plots showing for different cropping ratios (ϕ) the distribution of differences between pairs of aggregation parameter p estimates (the maximum possible difference being 1). Each pair of estimates is taken from a given landscape (with $\phi = x$) and its inverted or ‘flipped’ counterpart (with $\phi = 1 - x$). All p estimates in these plots were generated using random sampling of 1000 field locations for each landscape (of dimensions 100x100). The central points give the mean difference and the error bars show ± 1 standard deviation. The mean dispersal distance of the kernel used to generate these plots is 3, and the landscapes are modelled as flat areas with edge effects. The edge/area ratio range for the data in each plot is shown. This figure is a partitioning of Fig. 3.6b.

large estimation errors, partly driven by edge effects, can be largely corrected with a higher sampling density.

By increasing the number of sampled fields (n_f) we can increase the amount of spatial information being captured, and decrease errors in p estimation due to inconsistent field density across the landscape (Fig. 3.9b). From this plot we can see that the range of differences in

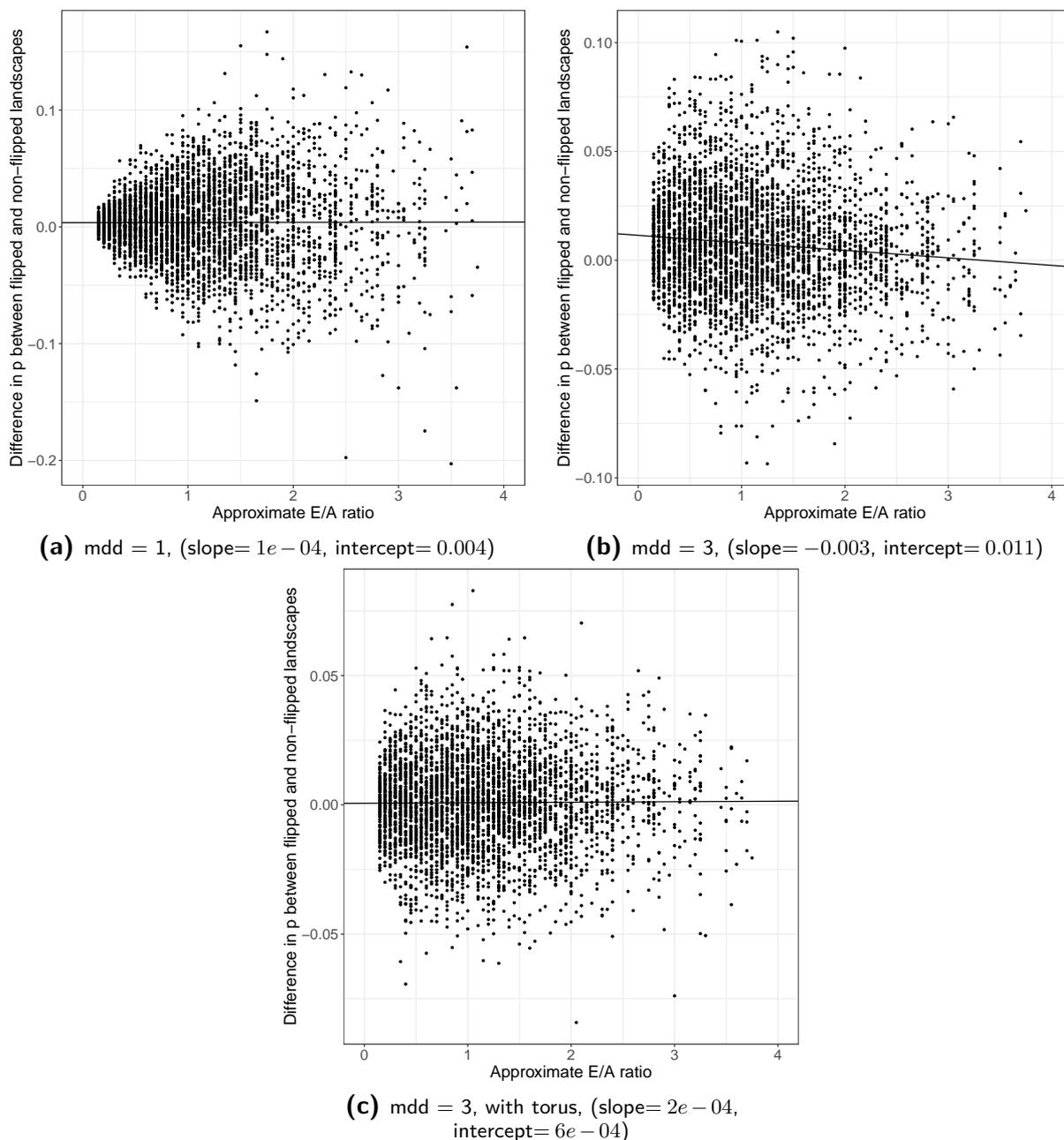


Fig. 3.8 Individual differences, across all cropping ratios, between pairs of aggregation parameter p estimates (the maximum possible difference being 1) as a function of edge/area (E/A) ratio. Each pair of estimates is taken from a given landscape (with $\phi = x$) and its inverted or 'flipped' counterpart (with $\phi = 1 - x$). All p estimates in these plots were generated using random sampling of 1000 field locations for each landscape (of dimensions 100x100). The landscapes are modelled as flat areas with edge effects, and the mean dispersal distance (mdd) of the kernel used to generate each plot is given. Plots (a) and (b) were generated from the landscapes modelled as a flat areas with edge effects, whereas in (c) each landscape is modelled as a torus in order to remove any edge effects. The lines of best fit, the coefficients for which are given in parentheses for each plot, were fitted using generalised least squares.

paired estimates for p decreases as n_f is increased, to the point that they converge at a single value when the landscape is comprehensively sampled by 10,000 evenly spaced fields. This converged p value difference is slightly below 0 due to the over-representation of susceptible (red) fields near the landscape edges, which would have been introduced in the random

landscape patch generation process. By taking the mean of the paired p value differences using 1000 fields across 30 random field sampling replicates however we can come very close to the 'true' difference in this landscape captured by comprehensively sampling with 10,000 fields (see the mean difference black line in Fig. 3.9b). Due to the lower number of pairwise between field calculations required, this repeated low density sampling appears to take significantly less computational time than sampling with a very high density of fields over fewer replicates.

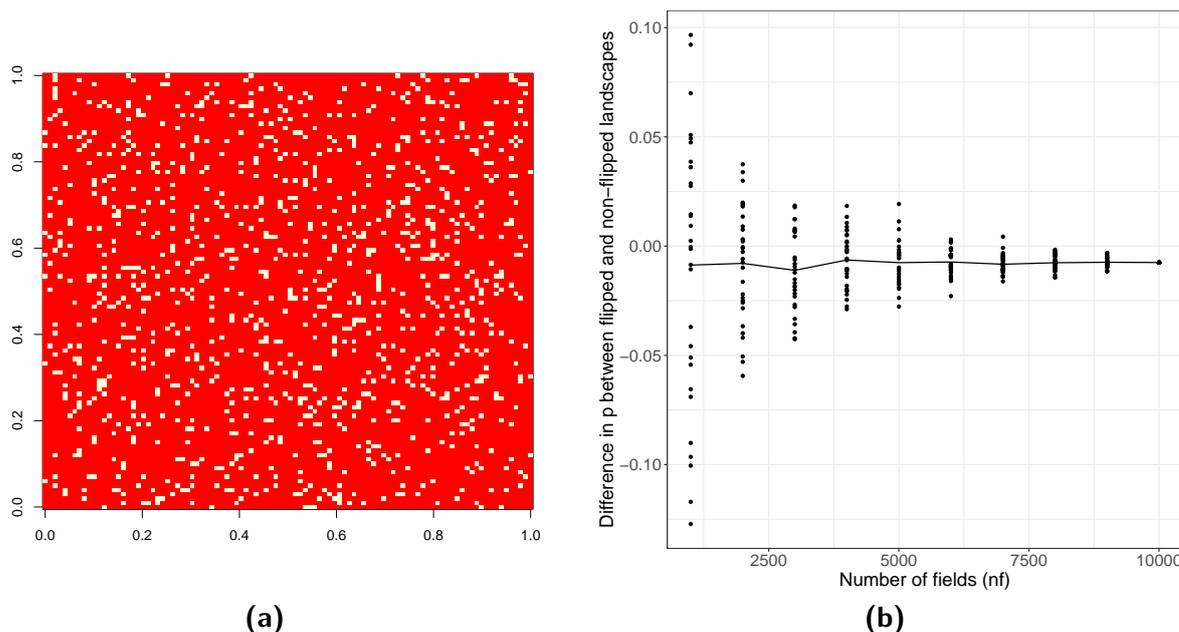


Fig. 3.9 The landscape template that produced the largest difference in paired aggregation parameter p estimates (out of those in figure (3.8a)) is shown in (a). Further paired p value estimates for different numbers of fields (n_f) are shown in (b), with 30 field location sampling replicates for each value of n_f . Each pair of estimates in (b) is taken from a the landscape shown in (a) (with cropping ratio $\phi = 0.1$) and its inverted or 'flipped' counterpart (with $\phi = 0.9$). The black line in (b) gives the mean difference between paired p value estimates for each value of n_f . The mean dispersal distance of the kernel used is 1, and the edge/area ratio = 3.505. The landscape is modelled as a flat area with edge effects.

3.4 Discussion

The two-patch model that we have presented in this chapter is intended to represent epidemiological dynamics at different scales of spatial heterogeneity in a host organism population. This is achieved by using the probability (p), with which inoculum disperses in a manner proportional to the cropping ratio of the two host varieties, as the parameter representing spatial aggregation within the host population. A value close to 1 for this probability means that the host population is well mixed, and is closer to mean field conditions, while a low value means that the relative frequency of dispersal between hosts of the same variety within locally aggregated patches is increased. This new metric for the measurement of within population

spatial heterogeneity both has an intuitive biological meaning, and can be easily estimated from host location point data.

The specific system used for the application of this model is that of a generic aerially dispersed plant disease epidemic within the crop fields of an agricultural landscape. Such an epidemic, featuring gene-for-gene resistance dynamics, could be caused by many rusts or powdery mildews. The ultimate goal is for the model to be used in the study of the durability of crop disease resistance (Fabre et al., 2012; van den Bosch and Gilligan, 2003). Adaptation of the model for this purpose is a simple matter of adding additional equations to the system of ODEs that represent the spread of a second pathogen strain. The differential ability of the strains to infect the two host varieties would then be modelled by adding parameters to scale the infection rate on a given host.

Despite this target system, the general nature of the epidemiological model means that this approach to modelling spatial heterogeneity could be used for any system where dynamic changes to the pathogen population occur on a much shorter timescale to those of the host population. Furthermore, the method for calculating pathogen dispersal between and within host patches, and therefore also for the calculation of the aggregation parameter p , is also extremely general. This could potentially be applied to a much wider variety of ecological and evolutionary systems, which might simply be seeking to represent population dynamics within a spatially heterogeneous habitat landscape, and need not necessarily involve the study of disease spread (Aars and Ims, 2000; Freckleton and Watkinson, 2002; Nachman, 2006; Wu, 2009).

A major goal for the development of this model was that the aggregation parameter and the cropping ratio, which determines the relative sizes of the two host variety patches, were to be entirely independent of one another. The spatial heterogeneity metric described in chapter 2, the landscape edge/area (E/A) ratio, which was used to study the durability of crop disease resistance, faced the problem that the cropping ratio could not be altered without changing the meaning of the landscape metric. This covariance with the E/A ratio, prevented rigorous investigation of the interaction between the scale of spatial heterogeneity and the cropping ratio of the landscape. In the results for the two-patch model described in this chapter, we have shown that the aggregation parameter p and the cropping ratio ϕ are entirely independent of one another, and as such can be freely varied in further analysis of the model. This independence is a direct consequence of the design of the aggregation parameter, and is

dependent on the dispersal within and between host variety patches being considered in terms of how proportional this dispersal is to the cropping ratio of the system.

The estimation process for p is simple and is not restricted to any particular form of dispersal kernel, meaning that this approach can be applied to a very wide variety of crop disease pathosystems. For example, a negative exponential kernel has been used to represent dispersal in citrus greening (Huanglongbing) (Parnell et al., 2015) and Asiatic citrus canker (Parnell et al., 2009), while an exponential power kernel has been used in the case of the banana plant fungus *Mycosphaerella fijiensis* (Rieux et al., 2014). As this model is still highly theoretical in its formulation, it does depend on a series of assumptions that affect the accuracy with which the aggregation parameter can be calculated from sample host landscapes. In particular, the strict definition of the aggregation parameter p depends on the assumption that the spatial area of the landscape either infinite or modelled as a torus, so that there are no edge effects caused by inoculum dispersing over the edges of the landscape. It also assumes that the area of the landscape is very large compared to the size of the locally aggregated patches so that the effective cropping ratio at which inoculum disperses when it spreads beyond these local patches is effectively equal to the global cropping ratio of the landscape. While for landscapes of a finite area with varying scales of spatial heterogeneity these assumptions will inevitably be broken to some degree, our results show that p estimation accuracy for our simulated landscapes is still very good, and the independence of p and the cropping ratio ϕ is still largely maintained, particularly for low mean dispersal distances. For the purposes of theoretical model analysis, the integrity of the biological meaning for aggregation parameter p is best maintained by using a landscape where the scale of the landscape dimensions are much larger than the scale of pathogen dispersal, or alternatively where the edges are wrapped around so that the landscape is modelled as a torus (Ovaskainen and Cornell, 2006).

The model also assumes that the crop fields, ignoring specific host variety, are evenly distributed throughout space. This, along with the assumption that there are no edge effects, is driven by the need for the overall dispersal of inoculum from the two crop varieties to occur proportionally to the cropping ratio, once the relative infection densities in the two host patches is taken into account. Another way of saying this is that one host field type must not disproportionately lose dispersed inoculum into gaps in the landscape where there are no fields. This assumption would be broken when one field variety is over represented near the landscape edges, or has more individual fields which are particularly isolated in the landscape. If this

model were to be applied to real agricultural landscape data, where the spatial structure and pathogen dispersal properties could not be controlled or manipulated, it is probable that the model presented in this chapter could be modified so that the finite area of the landscape and uneven host distribution could be explicitly accounted for. This could potentially be achieved partly by adding terms representing dispersal to empty or non-host occupied space, and by explicitly scaling the value of p to account for the changing effective cropping ratio as the scale of local aggregation increases. Detailed real agricultural landscape data could be obtained via access to the EU's Integrated Administration and Control System (IACS) database, or the Rural Land Register (RLR) in the UK.

A further assumption, that is inherent to the definition of the aggregation parameter p , is that the maximum degree of between host variety mixing in the landscape is dispersal is entirely proportional to the landscape cropping ratio and follows the law of mass action. The estimation of p from the sample generated landscapes however shows that a returned value of $p > 1$ is possible, particularly at a very low cropping ratio, high levels of mixing, and low mean dispersal distances. This is due to the highly isolated fields of the less frequent variety dispersing a disproportionately high amount of inoculum to the surrounding fields of the more frequent variety. A similar pattern would also be found in a regular landscape lattice, with varieties arranged in a chessboard layout, where dispersal occurred on a largely nearest neighbour basis. While this scenario does highlight a limitation of the definition of the aggregation parameter p , we argue that landscapes of this spatial structure would be extremely unlikely to occur in reality. This is because in many crop disease systems, for example in barley powdery mildew and wheat leaf rust, a large proportion of inoculum dispersal occurs over very short distances such as within the same crop field or even the same host plant (Mundt, 2009). Such local dispersal effects mean that it is very unlikely that a disproportionate amount of dispersed inoculum would land on fields of a different variety. Even if varieties were mixed within fields, dispersal frequencies would likely only approach proportionality to the cropping ratio. It should be noted that in the p estimation process presented in this chapter it was assumed that there was no within field component of inoculum dispersal, so as to simplify the application of the dispersal kernel. If it were instead assumed that a certain proportion of inoculum remained within the originator field, then the proportion of dispersal between hosts of the same variety would be markedly increased and the estimated value of p would always be significantly below 1.

The cellular automaton algorithm that we have developed and used for creating a suite of irregular landscapes has proved a flexible tool for generating varying degrees of within host population spatial heterogeneity. The large library of landscapes has subsequently been a vital component for demonstrating the independence of the aggregation parameter and the cropping ratio in our two-patch epidemiological model. While the algorithm is entirely artificial in its construction, the pattern of cells aggregating over time based on the local density of each cell type is designed to mimic similar clustering processes that occur in nature as a population of organisms grows and spreads. We therefore believe that this method of landscape generation has the potential to be used for a wider variety of ecological systems where spatial heterogeneity is of interest. Further developments to this algorithm are undoubtedly possible, with the potential for including additional patch types and unoccupied space. While approaches such as simulated annealing (Kirkpatrick et al., 1983; Papaix et al., 2014a) could be used to produce similar landscape patterns, the use of a cellular automaton potentially has an advantage in that it does not rely on making decisions to change the landscape structure in response to the value of a global output metric, as is the case in classical optimisation algorithms. The subsequently 'blind' nature of the cellular automaton which nevertheless increases the aggregation of the landscape over time, could potentially give computational speed benefits in some cases, although this would need to be tested. In order to accurately represent the degree of aggregation in a given generated landscape, the p estimation process must include a sufficiently high sampling density of field point locations. This can be achieved by using a very large number of fields, potentially up to the point where every possible spatial coordinate has been sampled. Alternatively, a much lower number of sampled fields can be applied over a number of replicates, which gives a similarly high level of accurate landscape representation while reducing computational intensity compared to sampling every coordinate.

The particular form of spatial structure that this model is intended to capture, which is based on the dispersal kernel weighted distances between fields of the same and different host varieties, has been chosen because we consider this to be the main driver of the dynamics influencing resistance durability shown in chapter 2. This method contrasts with many of the spatial modelling approaches described in the introduction of this chapter which are most frequently used to describe the spatial interactions between infected and uninfected individuals within a host population. In our model however we ignore the spatial distribution of the pathogen population within hosts of the same variety, and assume that infected and uninfected

host tissue units are well mixed within each host variety population. This assumption also necessarily applies to the initial conditions of an epidemic. We focus on implicitly capturing spatial information for the distribution of the host population, and then infer the spatial distribution of the pathogen population from this host structure.

In reality, we might expect that infected and uninfected host tissue units, and the two pathogen strains, would be unevenly distributed within an individual host variety population. More specifically it is reasonable to assume that infections of hosts by the pathogen strain to which it is most susceptible would be highly concentrated in aggregated local patches of that host variety. Infections of that host by the other pathogen strain meanwhile would be concentrated in the less aggregated regions where there would be a increased local density of the other host variety, which could act as a source of inoculum for that strain. This means that the effective degrees of aggregation experienced by each specific host variety/pathogen strain would be different. Overall, due to the predominance of host infections caused by the pathogen strains to which they are most susceptible, we expect that the overall strength of aggregation effects in a given landscape would increase over time as the global infection density increased. Our current model most likely therefore underestimates the effects of spatial aggregation. Future work on this model could seek to account for this effect by scaling the value of the aggregation parameter p by varying degrees in the different host/pathogen dispersal probability coefficients, potentially so that they depend on the overall level of infection. The values of any added scaling parameters would need to be optimised by comparing the two-patch model with an equivalent spatially explicit model. In chapter 4 we compare the results of the two-patch model with the equivalent results from the spatially explicit model used in chapter 2 to demonstrate qualitative behavioural similarity. We therefore consider that the above model alterations and parameterisation would only be necessary if very close quantitative similarity is required.

The advantage of our approach is that it uses a set of simplifying assumptions to reduce the spatial complexity of the system so that it captures only the hypothesised underlying ecological drivers of the resistance durability dynamics described from the spatially explicit model used in the previous chapter. This contrasts with the spatial information that would be captured if we had used a spatial moment approximation approach, which would usually seek to capture the dynamically changing spatial interactions between infected and uninfected individuals over time (Bolker and Pacala, 1997; Morozov and Poggiale, 2012). It is likely that the use of spatial moments for a multi-host multi-strain epidemiological system, would have resulted in

significantly more difficult implementation and the potential capture of superfluous spatial information. This extra spatial information may have overcome some of the model limitations described above, however it would also have introduced significant complexity in excess of our requirements. If we were to express our system in terms of spatial moments, the distribution of the different host varieties within the landscape is equivalent to second order spatial moments that remain static over time. This is due to the fact that the aggregation parameter is calculated from the kernel weighted distances between pairs of field in the landscapes. The proportion of inoculum dispersal that is proportional to the cropping ratio, and the distribution of the pathogen population within host variety patches meanwhile is equivalent to the first order moments.

In conclusion, the two-patch epidemiological model we have presented in this chapter effectively captures spatially explicit information on the distribution of different host varieties, and represents this information in an intuitive spatially implicit framework. This framework models the scale of spatial heterogeneity in a landscape of different host varieties in a way that allows the parameter governing this spatial aggregation to be varied completely independently of the cropping ratio. This independence, combined with the general and flexible nature of the model, will allow us to further our investigation of the dynamical interactions between spatial heterogeneity and the durability of crop disease resistance. Furthermore, this model also represents a novel development in the mathematical representation of spatial dynamics in ecological systems, which has the potential to be applied in a wide variety of contexts.

3.5 Appendix 1 - Proof that the model from Eqns (3.3) and (3.4) can reduce to a simple SIS model

3.5.1 Proof for unlinked patches ($p = 0$)

First we need to prove that if the relative number of infections in the two patches is proportional to the patches' cropping ratio, then this ratio will remain unchanged over time, even if the patches are completely unlinked.

$$\frac{dY}{dt} = \beta \left(\frac{(1-\phi)N - Y}{(1-\phi)N} \right) ((1-p\phi)Y + p(1-\phi)Z) - \eta Y, \quad (3.16)$$

$$\frac{dZ}{dt} = \beta \left(\frac{\phi N - Z}{\phi N} \right) (p\phi Y + (1 - p(1 - \phi))Z) - \eta Z. \quad (3.17)$$

If we assume that $p = 0$, meaning that there is no mixing between the two host variety patches, then:

$$\frac{dY}{dt} = \dot{Y} = \beta \left(\frac{(1 - \phi)N - Y}{(1 - \phi)N} \right) Y - \eta Y, \quad (3.18)$$

$$\frac{dZ}{dt} = \dot{Z} = \beta \left(\frac{\phi N - Z}{\phi N} \right) Z - \eta Z. \quad (3.19)$$

We introduce a new variable:

$$\epsilon = \frac{Y}{Y + Z}. \quad (3.20)$$

which is the number of infections in the first patch Y as a proportion of the total number of infections $Y + Z$. We can then show using the quotient rule how this proportion ϵ changes over time:

$$\frac{d\epsilon}{dt} = \frac{(Y + Z)\dot{Y} - (\dot{Y} + \dot{Z})Y}{(Y + Z)^2}, \quad (3.21)$$

$$(Y + Z)^2 \frac{d\epsilon}{dt} = Y\dot{Y} + Z\dot{Y} - Y\dot{Y} - \dot{Z}Y, \quad (3.22)$$

$$= Z\dot{Y} - \dot{Z}Y, \quad (3.23)$$

$$= Z \left[\beta \left(1 - \frac{Y}{(1 - \phi)N} \right) Y - \eta Y \right] - Y \left[\beta \left(1 - \frac{Z}{\phi N} \right) Z - \eta Z \right], \quad (3.24)$$

$$= Z\beta Y - \frac{Z\beta Y^2}{(1 - \phi)N} - \eta Y Z - Z\beta Y + \frac{Y\beta Z^2}{\phi N} + \eta Y Z, \quad (3.25)$$

$$(Y + Z)^2 \frac{d\epsilon}{dt} = \beta Y Z \left(\frac{Z}{\phi N} - \frac{Y}{(1 - \phi)N} \right), \quad (3.26)$$

$$\frac{d\epsilon}{dt} = \beta \frac{Y}{Y + Z} \frac{Z}{Y + Z} \left(\frac{Z}{\phi N} - \frac{Y}{(1 - \phi)N} \right), \quad (3.27)$$

$$= \beta \epsilon (1 - \epsilon) \left(\frac{Z}{\phi N} - \frac{Y}{(1 - \phi)N} \right). \quad (3.28)$$

Therefore:

$$\text{if } \frac{Z}{\phi N} = \frac{Y}{(1 - \phi)N}, \text{ then } \frac{d\epsilon}{dt} = 0, \text{ and so } \epsilon = \frac{Y}{Y + Z} \text{ is constant} \quad (3.29)$$

We then need to prove that if the relative number of infections in the two patches is always proportional to the patches' cropping ratio, then the epidemic conforms to the dynamics of a non-spatial model.

If $I = Y + Z$, then we can sum Eqns (3.18) and (3.19) to give:

$$\dot{I} = \dot{Y} + \dot{Z} = \beta(Y + Z) - \beta \left(\frac{Y^2}{(1-\phi)N} + \frac{Z^2}{\phi N} \right) - \eta(Y + Z), \quad (3.30)$$

$$= \beta I - \beta \left(\frac{\phi Y^2 + (1-\phi)Z^2}{(1-\phi)\phi N} \right) - \eta I. \quad (3.31)$$

If the values of Y and Z at time t are proportional to cropping ratio of their respective host patches (Eqn 3.29), then:

$$Y(t) = (1-\phi)\theta, \quad Z(t) = \phi\theta, \quad I(t) = \theta. \quad (3.32)$$

where θ is the overall number of infected plants in the landscape at time t . From this we can show that:

$$\dot{I}(t) = \beta\theta - \beta \left(\frac{\phi(1-\phi)^2\theta^2 + (1-\phi)\phi^2\theta^2}{(1-\phi)\phi N} \right) - \eta\theta, \quad (3.33)$$

$$= \beta\theta - \beta \left(\frac{(1-\phi)\theta^2 + \phi\theta^2}{N} \right) - \eta\theta, \quad (3.34)$$

$$= \beta\theta - \beta \frac{\theta^2}{N} - \eta\theta, \quad (3.35)$$

$$= \beta \left(1 - \frac{\theta}{N} \right) \theta - \eta\theta. \quad (3.36)$$

This means that when the relative number of infections in the two patches is proportional to the patches' cropping ratio, the epidemic conforms to the dynamics of a simple non-spatial *SIS* model.

3.5.2 General proof (for any value of p)

First we need to prove that if the relative number of infections in the two patches is proportional to the patches' cropping ratio, then this ratio will remain unchanged over time.

$$\frac{dY}{dt} = \dot{Y} = \beta \left(\frac{(1-\phi)N - Y}{(1-\phi)N} \right) ((1-p\phi)Y + p(1-\phi)Z) - \eta Y, \quad (3.37)$$

$$\frac{dZ}{dt} = \dot{Z} = \beta \left(\frac{\phi N - Z}{\phi N} \right) (p\phi Y + (1-p(1-\phi))Z) - \eta Z. \quad (3.38)$$

We introduce a new variable:

$$\epsilon = \frac{Y}{Y + Z}. \quad (3.39)$$

which is the number of infections in the first patch Y as a proportion of the total number of infections $Y + Z$. We can then show using the quotient rule how this proportion ϵ changes over time:

$$\frac{d\epsilon}{dt} = \frac{(Y+Z)\dot{Y} - (\dot{Y} + \dot{Z})Y}{(Y+Z)^2}, \quad (3.40)$$

$$(Y+Z)^2 \frac{d\epsilon}{dt} = Y\dot{Y} + Z\dot{Y} - Y\dot{Y} - \dot{Z}Y, \quad (3.41)$$

$$= Z\dot{Y} - \dot{Z}Y, \quad (3.42)$$

$$= Z \left[\beta \left(1 - \frac{Y}{(1-\phi)N} \right) ((1-p\phi)Y + p(1-\phi)Z) - \eta Y \right] - Y \left[\beta \left(1 - \frac{Z}{\phi N} \right) (p\phi Y + (1-p(1-\phi))Z) - \eta Z \right], \quad (3.43)$$

$$= \beta Z(1-p\phi)Y + \beta Z^2 p(1-\phi) - \frac{\beta Z(1-p\phi)Y^2}{(1-\phi)N} - \frac{\beta Y p(1-\phi)Z^2}{(1-\phi)N} - \eta Y Z - \beta p\phi Y^2 - \beta Y(1-p(1-\phi))Z + \frac{\beta Z p\phi Y^2}{\phi N} + \frac{\beta Y(1-p(1-\phi))Z^2}{\phi N} + \eta Y Z, \quad (3.44)$$

$$= \beta(1-p)YZ + \beta p(1-\phi)YZ + \beta p(1-\phi)Z^2 - \beta p\phi Y^2 - \beta(1-p)YZ - \beta p\phi YZ + \beta Y Z \left(\frac{p\phi Y}{\phi N} + \frac{(1-p)Z}{\phi N} + \frac{p\phi Z}{\phi N} + \frac{(1-p)Y}{(1-\phi)N} - \frac{p(1-\phi)Y}{(1-\phi)N} - \frac{p(1-\phi)Z}{(1-\phi)N} \right), \quad (3.45)$$

$$= \beta p \left((1-\phi)YZ + (1-\phi)Z^2 - \phi Y^2 - \phi YZ \right) + \beta Y Z (1-p) \left(\frac{Z}{\phi N} - \frac{Y}{(1-\phi)N} \right), \quad (3.46)$$

$$(Y+Z)^2 \frac{d\epsilon}{dt} = \beta p \left((1-\phi)YZ + (1-\phi)Z^2 - \phi Y^2 - \phi YZ \right) + \beta Y Z (1-p) \left(\frac{Z}{\phi N} - \frac{Y}{(1-\phi)N} \right), \quad (3.47)$$

$$\frac{d\epsilon}{dt} = \beta p \left((1-\phi) \frac{Y}{Y+Z} \frac{Z}{Y+Z} + (1-\phi) \left(\frac{Z}{Y+Z} \right)^2 - \phi \left(\frac{Y}{Y+Z} \right)^2 - \phi \frac{Y}{Y+Z} \frac{Z}{Y+Z} \right) + \beta \frac{Y}{Y+Z} \frac{Z}{Y+Z} (1-p) \left(\frac{Z}{\phi N} - \frac{Y}{(1-\phi)N} \right), \quad (3.48)$$

$$= 2\beta p \left((1-\phi) \frac{Z}{Y+Z} - \phi \frac{Y}{Y+Z} \right) + \beta \epsilon (1-\epsilon)(1-p) \left(\frac{Z}{\phi N} - \frac{Y}{(1-\phi)N} \right), \quad (3.49)$$

$$= \frac{2\beta p\phi(1-\phi)N}{Y+Z} \left(\frac{Z}{\phi N} - \frac{Y}{(1-\phi)N} \right) + \beta \epsilon (1-\epsilon)(1-p) \left(\frac{Z}{\phi N} - \frac{Y}{(1-\phi)N} \right), \quad (3.50)$$

$$= \left(\frac{2\beta p\phi(1-\phi)N}{Y+Z} + \beta \epsilon (1-\epsilon)(1-p) \right) \left(\frac{Z}{\phi N} - \frac{Y}{(1-\phi)N} \right). \quad (3.51)$$

$$(3.52)$$

$$(3.53)$$

Therefore:

$$\text{if } \frac{Z}{\phi N} = \frac{Y}{(1-\phi)N}, \text{ then } \frac{d\epsilon}{dt} = 0, \text{ and so } \epsilon = \frac{Y}{Y+Z} \text{ is constant} \quad (3.54)$$

We then need to prove that if the relative number of infections in the two patches is always proportional to the patches' cropping ratio, then the epidemic conforms to the dynamics of a non-spatial model.

If $I = Y + Z$, then we can sum Eqns (3.37) and (3.38) to give:

$$\dot{I} = \dot{Y} + \dot{Z} = \beta((1-p\phi)Y + p(1-\phi)Z + p\phi Y + (1-p(1-\phi))Z) \quad (3.55)$$

$$- \beta \left(\frac{(1-p\phi)Y^2}{(1-\phi)N} + \frac{p(1-\phi)YZ}{(1-\phi)N} + \frac{p\phi YZ}{\phi N} + \frac{(1-p(1-\phi))Z^2}{\phi N} \right) \quad (3.56)$$

$$- \eta(Y + Z), \quad (3.57)$$

$$= \beta(Y + Z) \quad (3.58)$$

$$- \beta \left(\frac{(1-p\phi)Y^2}{(1-\phi)N} + \frac{p(1-\phi)YZ}{(1-\phi)N} + \frac{p\phi YZ}{\phi N} + \frac{(1-p(1-\phi))Z^2}{\phi N} \right) \quad (3.59)$$

$$- \eta I. \quad (3.60)$$

If the values of Y and Z at time t are proportional to cropping ratio of their respective host patches (Eqn 3.54), then:

$$Y(t) = (1-\phi)\theta, \quad Z(t) = \phi\theta, \quad I(t) = \theta. \quad (3.61)$$

where θ is the overall number of infected plants in the landscape at time t . From this we can show that:

$$\dot{I}(t) = \beta\theta - \beta \left(\frac{(1-p\phi)(1-\phi)^2\theta^2}{(1-\phi)N} + \frac{p(1-\phi)^2\theta^2\phi}{(1-\phi)N} \right) \quad (3.62)$$

$$+ \frac{p\phi^2(1-\phi)\theta^2}{\phi N} + \frac{(1-p(1-\phi))\phi^2\theta^2}{\phi N} - \eta\theta, \quad (3.63)$$

$$= \beta\theta - \beta \left(\frac{(1-\phi)\theta^2}{N} + \frac{\theta^2\phi}{N} \right) - \eta\theta, \quad (3.64)$$

$$= \beta\theta - \beta \frac{\theta^2}{N} - \eta\theta, \quad (3.65)$$

$$= \beta \left(1 - \frac{\theta}{N}\right) \theta - \eta \theta. \quad (3.66)$$

This means that when the relative number of infections in the two patches is proportional to the patches' cropping ratio, the epidemic conforms to the dynamics of a simple non-spatial *SIS* model.

Chapter 4

How does spatial heterogeneity interact with the cropping ratio for achieving maximal resistance durability?

4.1 Introduction

In the previous chapter, we introduced a novel method for incorporating the scale of spatial heterogeneity between different host varieties in a simple epidemiological model. The scale of spatial heterogeneity can also be thought of as the degree of 'mixing' between the different varieties, the degree of spatial aggregation in the individual varieties, or the level of spatial diversification within the host population. The intention for this new model is to enable us to study the effects of spatial heterogeneity on epidemics using a spatially implicit approach that is flexible and at least partially analytically tractable. The model is based around the spatial aggregation parameter p , which is the proportion of inoculum that disperses proportionally to the cropping ratio, with the proportion $1 - p$ dispersing preferentially onto the same host variety as where the inoculum was produced. The extent to which inoculum disperses preferentially onto the same host variety represents the extent to which different plants of the same variety are aggregated together in space, in combination with the relative scale of pathogen dispersal. This parameter p is used to determine the frequency with which inoculum produced on the different varieties leads to new infections on the same and on other varieties.

In chapter 3 we demonstrated the flexibility of the new model, in terms of allowing the aggregation parameter p and the cropping ratio to vary independently of one another. This

overcomes an issue with the spatially explicit model, as used in chapter 2, where using the edge/area ratio of the agricultural landscape as our measure of spatial heterogeneity constrained us to looking at a single value for the cropping ratio of the system. Having a constrained cropping ratio in this initial study did have the benefit of reducing the number of parameters allowed to freely vary in an otherwise highly complex system, and allowed us to investigate the specific effects of spatial heterogeneity in significant depth. Furthermore, the conclusions gathered from this spatially explicit work allowed us to identify some of the key drivers behind the effects of spatial heterogeneity, which provided the inspiration for the formulation of the new spatially implicit model. However, the new model considered in this chapter allows us to ask questions that were not possible in the previous explicit space model. Principally we ask, how does the degree of aggregation in the system interact with the resistance cropping ratio, and how does this interaction change with different fitness costs and resistance gene efficacies? In addition, although the previous edge/area metric was also a continuous parameter, the clearer and more precise definition of the aggregation parameter p will lend itself to more definitive biological interpretation.

The cropping ratio of varieties has previously been shown in various theoretical studies to have a significant effect on the durability of crop disease resistance. Using epidemic intensity over evolutionary timescales as their measure of durability, in a similar manner our current work, Fabre et al. (2012) showed that plotting epidemic intensity against cropping ratio typically produced a u-shaped graph. These u-shaped responses produced intermediate to high optimal cropping ratios for resistance durability in their model, depending on the epidemiological parameters and the mutation-selection balance of the resistance breaking strain. Papaïx et al. (2018) also indicated that intermediate cropping ratios were optimal, this time for minimising epidemic intensity at the evolutionary equilibrium, which is the closest of their resistance durability metrics to our current work. These authors also suggested that a low levels of spatial aggregation between varieties was optimal for minimising the disease equilibrium at all cropping ratios. An alternative metric for durability used by Papaïx et al. (2018), indicated that using either very high or low cropping ratios was best for minimising the time until the initial emergence of a resistance breaking strain through mutation. This is because a very high cropping ratio drastically reduces the size of the wild-type population, and therefore decreases the likelihood that a more virulent mutant will emerge, while a low cropping ratio reduces the

selection pressure for the subsequent establishment of the new mutant population (van den Bosch and Gilligan, 2003).

Both the cropping ratio and the scale of spatial heterogeneity are levers that are at least partially under the control of communities of growers and are not entirely dependent on the specific pathosystem. This means that it may simply be enough in some cases to identify the optimum values for both of these parameters that gives the best disease control outcome. It may also be possible on the other hand that one or both of these parameters will be constrained to a particular value or range of values. The smallest possible scale of spatial heterogeneity may for example be practically limited by the financial cost and operational difficulty of such practices. Conversely, the cropping ratio may be limited by the financial cost or availability of resistant plants or seeds. In these circumstances it may be most relevant to identify the optimum value of one parameter that can be practically freely varied, for a given value of the other parameter.

The next logical step in this programme of research is to apply the spatially implicit 'two-patch' model to the question of resistance durability that was previously studied using the spatially explicit approach in chapter 2. This will involve adding an extra pair of equations, to the system of ODEs described in chapter 3, in order to describe the dynamics of a second pathogen strain. The behaviour of the two pathogen strains on the two host varieties will then be determined by parameters that govern their infectivity and associated fitness costs. The simpler nature of the implicit model means that we can partially approach the question by finding the invasion thresholds and required conditions for coexistence in the model, using a combination of mathematical analysis and analytically guided numerical work. We have previously carried out analysis on a simplified version of the model where both host varieties are treated equally by a pathogen strain. This analysis showed that the two equation system of ODEs reduces to a simple *SIS* model in this special case, due to the fact that a completely generalist pathogen will not be affected by the degree of aggregation or cropping ratio of the two host varieties. This simple model closely corresponds to the behaviour of the resistance breaking strain in our current system when the wild-type is not present. The wild-type dynamics on the other hand will be more complex as it has a reduced ability to infect the resistant variety, with potential knock on effects for the dynamics of the resistance breaking strain. Knowledge about the invasion thresholds for the two strains, individually and when in competition, will answer practical questions about how to prevent the invasion of an undesirable strain in an

agricultural landscape. It will also hopefully aid in the interpretation of the inevitably more complex results from an exhaustive set of numerical simulations.

Carrying out numerical simulations and finding disease progress curves, in a similar manner to before, will remain the main measure by which we assess the durability of crop disease resistance. The area under the disease progress curve, which is normalised to give the epidemic intensity averaged over a given period of time, is a measure of resistance durability that is equivalent to similar measures used by Fabre et al. (2012, 2015) and van den Bosch and Gilligan (2003). The rationale for the use of this metric is that the optimal disease management strategy will be one that minimises crop yield losses over the evolutionary time period of interest. During this time period the frequencies of the two pathogen strains will change depending on their relative ability to infect plants throughout the agricultural landscape. The hope is that the dual methods of analysis and simulation, in combination with the clearer definition for the new spatial metric p , will enable a more comprehensive study of the role that spatial heterogeneity plays in the durability of crop disease resistance. A number of the results in chapter 2 from the spatially explicit were highly complex, unintuitive and a challenge to explain, so a deeper understanding of how the dynamics involved operate will be invaluable. This is particularly true for the non-unimodal responses of spatial diversification efficacy to changes in the fitness cost of the resistance breaking trait, the efficacy of the resistance gene and the length of the time period of interest.

4.2 Methods

4.2.1 The two-patch two-strain model

We combine the host-parasite gene-for-gene system introduced in chapter 2 and the spatially implicit model presented in chapter 3. The result is an *SIS* model in which a wild-type (*wt*) and a resistance breaking (*rb*) pathogen strain infect two 'patches' of host crop, where one patch contains a susceptible (*S*) host variety and the other contains a resistant host (*R*) variety. As is detailed in chapter 3, the probabilities with which infections in one patch lead to infections in the other patch, and also further infections in the same patch, are driven by the scale of spatial heterogeneity in the landscape being modelled. The manner in which these probabilities are characterised allows us to independently vary the parameter controlling the

scale of spatial heterogeneity (or degree of variety specific aggregation) and the cropping ratio in the landscape.

The four state variables in the system are given below, where I_S^{wt} is the number of infections caused by the wt strain on the S variety, I_S^{rb} is the number caused by the rb strain on the S variety, I_R^{wt} is the number caused by the wt strain on the R variety, and I_R^{rb} is the number caused by the rb strain on the R variety. For the sake of simpler mathematical notation, these state variables are represented by the letters X , Y , W and Z respectively:

$$I_S^{wt} \rightarrow X, \quad I_R^{wt} \rightarrow W, \quad (4.1)$$

$$I_S^{rb} \rightarrow Y, \quad I_R^{rb} \rightarrow Z. \quad (4.2)$$

This full model then is represented by the following system of ordinary differential equations (ODEs):

$$\frac{dX}{dt} = \beta \left(\frac{(1-\phi)N - X - Y}{(1-\phi)N} \right) ((1-p\phi)X + p(1-\phi)W) - \eta X, \quad (4.3)$$

$$\frac{dY}{dt} = \beta(1-\delta) \left(\frac{(1-\phi)N - X - Y}{(1-\phi)N} \right) ((1-p\phi)Y + p(1-\phi)Z) - \eta Y, \quad (4.4)$$

$$\frac{dW}{dt} = \beta\gamma \left(\frac{\phi N - W - Z}{\phi N} \right) (p\phi X + (1-p(1-\phi))W) - \eta W, \quad (4.5)$$

$$\frac{dZ}{dt} = \beta(1-\delta) \left(\frac{\phi N - W - Z}{\phi N} \right) (p\phi Y + (1-p(1-\phi))Z) - \eta Z. \quad (4.6)$$

where β is the net infection rate, δ is the fitness cost of the rb trait, γ is the susceptibility of the R host to the wt pathogen strain, N is the overall number of host tissue units in the landscape, ϕ is the cropping ratio of the R host, p is the aggregation parameter (the probability with which produced inoculum disperses according to the cropping ratio), and η is the harvesting/replanting rate. The variables and parameters used in the model are summarised in table 4.1.

The fitness cost of the resistance breaking trait δ is assumed to be expressed when the rb pathogen strain infects either host variety patch, meaning that the rb strain essentially acts as a generalist pathogen. This contrasts to the specialist wt strain which infects the S host with no fitness costs, but is only able to infect the R host as far as is permitted by the parameter

γ . When $\gamma = 0$ the R host is entirely resistant to the wt strain, and so no infection occurs, whereas if $\gamma = 1$ then the host resistance gene has no function.

Table 4.1 Parameters and variables used in the mathematical model

Symbol	Parameter/Variable Description	Constraints/Values
$I_S^{wt} \rightarrow X$	Number of susceptible host units infected by the wt pathogen strain	
$I_S^{rb} \rightarrow Y$	Number of susceptible host units infected by the rb pathogen strain	
$I_R^{wt} \rightarrow W$	Number of resistant host units infected by the wt pathogen strain	
$I_R^{rb} \rightarrow Z$	Number of resistant host units infected by the rb pathogen strain	
β	Infection rate (inoculum production \times infection probability)	2
η	Harvesting/replanting rate	1
N	Overall host population size (host tissue units)	10000
ϕ	Cropping ratio (proportion of resistant fields)	$0 \leq \phi \leq 1$
p	Probability that produced inoculum disperses proportionally to the cropping ratio ϕ	$0 \leq p \leq 1$
δ	Fitness cost of the rb trait	$0 \leq \delta \leq 1$
γ	Susceptibility of the R host to the wt strain	$0 \leq \gamma \leq 1$
n_y	Number of seasons	

As is detailed in chapter 3, the number of uninfected host tissue units in each patch, in the transmission terms of the model, must be expressed as a proportion of the total number of host tissue units in that patch. This means that the model is entirely frequency dependent (Keeling and Rohani, 2011), and so the results will be independent of the overall population size N (assuming that the state variables are proportional to N). The coefficients within the smaller brackets in Eqns (4.3-4.6), expressed in terms of p and ϕ , that govern the probability of inoculum dispersal within and between the two patches are also taken directly from the single strain model given in chapter 3, and capture the degree to which the two host varieties are aggregated in the agricultural landscape, while allowing the cropping ratio to be varied independently.

4.2.2 Invasion analysis

In order to determine the conditions under which the two pathogen strains will coexist or not in the landscape, we need to conduct a pairwise invasion analysis. This asks whether a very small number of infections by one strain will spread and invade the endemic equilibrium of the other strain (which was reached in the absence of the invading strain). If both strains can invade the other's equilibrium it follows that they will both coexist in the landscape. If one strain will invade the other's equilibrium but the reverse is not true, then only the first strain will ever remain the system at equilibrium (when both strains are initially present). There is

also the potential for one or both strains to be entirely unable to invade the landscape even when they are the only strain present.

***wt* invading the *rb* equilibrium**

rb invading $(Y, Z) = (0, 0)$ Before conducting the pairwise invasion analysis we need to find the equilibria for each pathogen strain by itself in the landscape. For the *rb* strain, this involves solving the following system of ODEs which results when we assume that $X = W = 0$ in Eqns (4.4) and (4.6):

$$\frac{dY}{dt} = \beta(1-\delta) \left(\frac{(1-\phi)N - Y}{(1-\phi)N} \right) ((1-p\phi)Y + p(1-\phi)Z) - \eta Y, \quad (4.7)$$

$$\frac{dZ}{dt} = \beta(1-\delta) \left(\frac{\phi N - Z}{\phi N} \right) (p\phi Y + (1-p(1-\phi))Z) - \eta Z. \quad (4.8)$$

As shown in the Appendix to chapter 3, the model in this form, where the pathogen has the same fitness when infecting either host variety reduces to a simple *SIS* model when Y and Z are proportional to the cropping ratio. However this proportionality follows from the assumption that the pathogen is initially evenly distributed in the landscape. If we denote the total number of infections by $I = Y + Z$ then, as shown in chapter 3, we can sum Eqns (4.7) and (4.8) to give:

$$\frac{dI}{dt} = \beta(1-\delta) \left(1 - \frac{I}{N} \right) I - \eta I \quad (4.9)$$

This *SIS* has a basic reproductive number of:

$$R_0^{rb} = \frac{\beta(1-\delta)}{\eta} \quad (4.10)$$

If $R_0^{rb} > 1$ then the *rb* can invade the $(0, 0)$ equilibrium, and it eventually reaches an endemic equilibrium of:

$$I^* = \frac{(\beta(1-\delta) - \eta)N}{\beta(1-\delta)} \quad (4.11)$$

wt **invading** $(X, Y, W, Z) = (0, Y^*, 0, Z^*)$ The Jacobian for the system given in Eqns (4.3-4.6), when linearised around the equilibrium in which only the *rb* strain is present (i.e. invasion of the *wt* pathogen strain) is given by:

$$J = \begin{pmatrix} J_{1,1} & J_{1,2} & J_{1,3} & J_{1,4} \\ J_{2,1} & J_{2,2} & J_{2,3} & J_{2,4} \\ J_{3,1} & J_{3,2} & J_{3,3} & J_{3,4} \\ J_{4,1} & J_{4,2} & J_{4,3} & J_{4,4} \end{pmatrix} \quad (4.12)$$

where:

$$J_{1,1} = \beta \left(1 - p\phi - \frac{(1-p\phi)Y}{(1-\phi)N} \right) - \eta, \quad (4.13)$$

$$J_{1,2} = 0, \quad (4.14)$$

$$J_{1,3} = \beta p \left(1 - \phi - \frac{Y}{N} \right), \quad (4.15)$$

$$J_{1,4} = 0, \quad (4.16)$$

$$J_{2,1} = -\beta(1-\delta) \left(\frac{(1-p\phi)Y}{(1-\phi)N} + \frac{pZ}{N} \right), \quad (4.17)$$

$$J_{2,2} = \beta(1-\delta) \left(1 - p\phi - \frac{2(1-p\phi)Y}{(1-\phi)N} - \frac{pZ}{N} \right) - \eta, \quad (4.18)$$

$$J_{2,3} = 0, \quad (4.19)$$

$$J_{2,4} = \beta p(1-\delta) \left(1 - \phi - \frac{Y}{N} \right), \quad (4.20)$$

$$J_{3,1} = \beta \gamma p \left(\phi - \frac{Z}{N} \right), \quad (4.21)$$

$$J_{3,2} = 0, \quad (4.22)$$

$$J_{3,3} = \beta \gamma \left(1 - p(1-\phi) - \frac{(1-p(1-\phi))Z}{\phi N} \right) - \eta, \quad (4.23)$$

$$J_{3,4} = 0, \quad (4.24)$$

$$J_{4,1} = 0, \quad (4.25)$$

$$J_{4,2} = \beta(1-\delta)p \left(\phi - \frac{Z}{N} \right), \quad (4.26)$$

$$J_{4,3} = -\beta(1-\delta) \left(\frac{pY}{N} + \frac{(1-p(1-\phi))Z}{\phi N} \right), \quad (4.27)$$

$$J_{4,4} = \beta(1-\delta) \left(1 - p(1-\phi) - \frac{2(1-p(1-\phi))Z}{\phi N} - \frac{pY}{N} \right) - \eta. \quad (4.28)$$

The equilibrium values for the *rb* strain in the two patches are then substituted into this matrix as $Y = (1 - \phi)I^*$ and $Z = \phi I^*$, where I^* comes from Eqn (4.11). For a given set of parameter values being tested, the leading eigenvalue of this matrix can then be calculated numerically to determine stability. If the leading eigenvalue is positive then the system is unstable and the *wt* strain will invade the *rb* equilibrium. This gives an unequivocal numerical method to determine whether or not invasion will occur for any set of parameters.

***rb* invading the *wt* equilibrium**

***wt* invading** $(X, W) = (0, 0)$ In order to obtain the *wt* endemic equilibrium when it invades the landscape by itself we must solve the following system of ODEs which results when we assume that $Y = Z = 0$ in Eqns (4.3) and (4.5):

$$\frac{dX}{dt} = \beta \left(\frac{(1 - \phi)N - X}{(1 - \phi)N} \right) ((1 - p\phi)X + p(1 - \phi)W) - \eta X, \quad (4.29)$$

$$\frac{dW}{dt} = \beta\gamma \left(\frac{\phi N - W}{\phi N} \right) (p\phi X + (1 - p(1 - \phi))W) - \eta W. \quad (4.30)$$

Due to the more asymmetrical nature of this system when compared to Eqns (4.7) and (4.8), sufficient analytical progress could not be made when attempting to find equilibria and determine stability. Therefore we instead use a computer algebra system to convert Eqns (4.29) and (4.30) into a 4th order polynomial which can be solved numerically for a given set of parameter values to find the values of X and W at the endemic *wt* equilibrium. This process is detailed in Appendix 1. Comprehensive numerical testing showed that there is only ever a single endemic *wt* equilibrium per set of parameters.

***rb* invading** $(X, Y, W, Z) = (X^*, 0, W^*, 0)$ These values are then substituted into the following Jacobian which results when the system given in Eqns (4.3-4.6) is linearised around the equilibrium in which only the *wt* strain is present (i.e. invasion of the *rb* pathogen strain):

$$J = \begin{pmatrix} J_{1,1} & J_{1,2} & J_{1,3} & J_{1,4} \\ J_{2,1} & J_{2,2} & J_{2,3} & J_{2,4} \\ J_{3,1} & J_{3,2} & J_{3,3} & J_{3,4} \\ J_{4,1} & J_{4,2} & J_{4,3} & J_{4,4} \end{pmatrix} \quad (4.31)$$

where:

$$J_{1,1} = \beta \left(1 - p\phi - \frac{2(1-p\phi)X}{(1-\phi)N} - \frac{pW}{N} \right) - \eta, \quad (4.32)$$

$$J_{1,2} = -\beta \left(\frac{(1-p\phi)X}{(1-\phi)N} + \frac{pW}{N} \right), \quad (4.33)$$

$$J_{1,3} = \beta p \left(1 - \phi - \frac{X}{N} \right), \quad (4.34)$$

$$J_{1,4} = 0, \quad (4.35)$$

$$J_{2,1} = 0, \quad (4.36)$$

$$J_{2,2} = \beta(1-\delta) \left(1 - p\phi - \frac{(1-p\phi)X}{(1-\phi)N} \right) - \eta, \quad (4.37)$$

$$J_{2,3} = 0, \quad (4.38)$$

$$J_{2,4} = \beta p(1-\delta) \left(1 - \phi - \frac{X}{N} \right), \quad (4.39)$$

$$J_{3,1} = \beta \gamma p \left(\phi - \frac{W}{N} \right), \quad (4.40)$$

$$J_{3,2} = 0, \quad (4.41)$$

$$J_{3,3} = \beta \gamma \left(1 - p(1-\phi) - \frac{pX}{N} - \frac{2(1-p(1-\phi))W}{\phi N} \right) - \eta, \quad (4.42)$$

$$J_{3,4} = -\beta \gamma \left(\frac{pX}{N} + \frac{(1-p(1-\phi))W}{\phi N} \right), \quad (4.43)$$

$$J_{4,1} = 0, \quad (4.44)$$

$$J_{4,2} = \beta(1-\delta)p \left(\phi - \frac{W}{N} \right), \quad (4.45)$$

$$J_{4,3} = 0, \quad (4.46)$$

$$J_{4,4} = \beta(1-\delta) \left(1 - p(1-\phi) - \frac{(1-p(1-\phi))W}{\phi N} \right) - \eta. \quad (4.47)$$

For the given set of parameter values we can then numerically calculate the leading eigenvalue for this matrix. If the leading eigenvalue is positive then the system is unstable and the *rb* strain will invade the *wt* equilibrium.

Numerical calculations

For both sets of pairwise invasion analyses the invasion criteria were tested with a wide range of parameter values by numerically simulating the system of ODEs given by Eqns (4.3-4.6). We used 10890 systematically chosen sets of parameter values for these and all subsequent tests, spanning the full range of potential values for ϕ , p , δ , and γ as detailed in table 4.1. These

simulation tests confirmed that the conditions required for a very small amount one strain to spread, when the other strain was at its endemic equilibrium, agreed with the conditions indicated by the invasion analysis. The expressions for the endemic equilibria, of each strain by itself, were also tested by comparing them against the long term results from simulations of Eqns (4.7) and (4.8) for the *rb* equilibrium, and Eqns (4.29) and (4.30) for the *wt*. The simulations were also repeated for the full system of ODEs with 100 sets of random initial conditions per tested set of parameter values. This confirmed that in all cases there is a single unique endemic equilibrium, results do not depend on initial conditions, and there is no evidence of backwards bifurcation behaviour (where the conditions for coexistence might not match the pairwise invasion thresholds) (Gilligan and van den Bosch, 2008; van den Driessche and Watmough, 2002).

4.2.3 Model simulation

When simulating the full system of ODEs for the purpose of calculating AUDPCs (Area Under Disease Progress Curves), the initial condition for the overall incidence of disease was set at the endemic equilibrium of the *wt* strain in the case where the cropping ratio $\phi = 0$. This is the endemic equilibrium of the *wt* in a system where all hosts are fully susceptible. In order to represent the initial invasion of the landscape by the *rb* strain, the majority of initial infections are caused by the *wt* with only a small percentage caused by the *rb* strain. The initial condition for the infection frequency of the *rb* strain was set at one infected host tissue unit per 100 infected host tissue units (frequency of 0.01) in each patch. This was done so as to match the initial conditions used in chapter 2, where the justification for the use of this value can also be found. For all parameter combinations tested in both the invasion analyses and simulations the infection rate $\beta = 2$ per host tissue unit per unit time, the harvesting/replanting rate $\eta = 1$ per unit time, and the overall host population size $N = 10000$ (although results are independent of N). This is so that harvesting/replanting occurs once per season, and also so that half of the hosts in the landscape are infected at the *wt* equilibrium when all hosts are susceptible. This latter condition matches the simulations carried out in chapter 2, and is intended to give moderate baseline level of infection before the modification of the parameters governing the landscape structure and host-parasite genetic interaction.

In order to test whether the two-patch model produced similar outcomes to the spatially explicit model from chapter 2, we estimated values for the aggregation parameter p that

correspond to the regularly divided landscapes used in that earlier chapter. Using the method described in chapter 3 section 3.2.2, we estimated p from the landscape templates that correspond to the low ($E/A=0.1$) and high ($E/A=1.4$) ends of the edge/area ratio scale used in chapter 2. This was done using the number of sampled fields as $n_f = 1000$, and taking the mean of the estimated p values from 30 replicates for each landscape template. To minimise the estimation error we also took a mean of the two possible p value derivations for each set of sampled fields (calculated from the proportion of dispersal from susceptible to resistant fields, and from the proportion of dispersal from resistant to susceptible fields). The mean dispersal distance used in chapter 2 was $mdd = 1/\eta = 0.5$ arbitrary distance units within a 10×10 landscape. The estimation for p on the other hand was calculated in a 100×100 landscape, as is the case throughout chapter 3, meaning that the mean dispersal distance used for this estimation was scaled up to $mdd = 5$. This ensures that the scale of dispersal relative to the size of the landscape is the same as was used in chapter 2. This p estimation process only accounts however for the between field component of dispersal in the epidemics simulated in chapter 2. In the spatially explicit model used in this earlier chapter, the epidemiological parameters were optimised such that approximately $1/3$ of total infection over a $n_y = 40$ season time period was due to between field transmission, with the remainder being caused by within field and reservoir driven transmission. This means that the value of p , which is a proportion of the between field transmission component only, must be divided by 3 to obtain the actual proportion of inoculum dispersing according to the cropping ratio in the spatially explicit model.

4.3 Results

4.3.1 Invasion analysis

Pairwise invasion threshold plots for the two pathogen strains can be used to determine the conditions under which the strains will coexist in the landscape, or under which one will outcompete the other (Fig. 4.1). The values for the cropping ratio of the resistant crop (ϕ) and the aggregation parameter p (the proportion of inoculum that disperses according to the global cropping ratio) where the rb strain can invade the resident wt strain at its endemic equilibrium, at given values for the cost of the resistance breaking trait (δ) and the susceptibility of the resistant host to the wt strain (γ), are shown in Fig. (4.1a). The parameter region where the wt strain will not invade the landscape by itself are also shown here. Correspondingly,

the parameter values where the *wt* strain will invade the resident *rb* strain at its endemic equilibrium are shown in Fig. (4.1b). Combining these invasion thresholds gives the region of parameter space where both strains of pathogen will invade the endemic equilibrium of the other, meaning that the two strains will coexist in the landscape (Fig. 4.1c). These invasion analysis results have been confirmed by numerical simulations using 100 sets of random initial conditions per combination of parameter values shown, the results from which give identical parameter values under which coexistence does or does not occur. These simulations also indicate that there exists only one unique endemic equilibrium for each set of parameter value combinations, i.e. that the long-term behaviour of the system can be fully and conveniently characterised via this invasion analysis.

Conditions for invasion and coexistence

Plots showing the pairwise invasion thresholds and conditions for coexistence in the full model as a function of the parameters ϕ and p , at different combinations of the discretised parameters δ and γ , are shown in Fig. 4.2. The *wt* pathogen strain can invade the *rb* strain at a wider range of parameter values when the cropping ratio (ϕ) is lower (since there is more susceptible crop in the system), and also when the aggregation parameter p is low (since greater aggregation promotes the *wt* because of the more clustered S host). There is a frequently large region of parameter space where the *wt* strain will invade a landscape in the absence of the *rb* strain, but not when the two strains are in competition (Fig. 4.2 green *rb* only region). This indicates that an endemic *rb* strain in the landscape will inhibit *wt* invasion in many cases.

In some cases, the threshold for the *rb* strain invading the *wt* strain equilibrium follows a correspondingly similar pattern to the *wt* invasion of the *rb* equilibrium. In these, the *rb* strain invades at a wider range of parameter values when the cropping ratio ϕ is higher (since there is more resistant crop in the system), and when the aggregation parameter p is lower (since the R host is more aggregated). At some values of δ and γ however, including those shown in Fig. 4.1, this invasion threshold is non-monotonic as a function of the aggregation parameter p . This means that in these cases it is at an intermediate value of p where *rb* invasion will not occur at the widest range of potential cropping ratio (ϕ) values. This contrasts to the alternative situation (such as for $\delta = 0.1$, $\gamma = 0.6$ in Fig. 4.2) where invasion of the *rb* strain by the *wt* is always most easily prevented at $p = 1$ where dispersal occurs according to the global cropping ratio (panmixis).

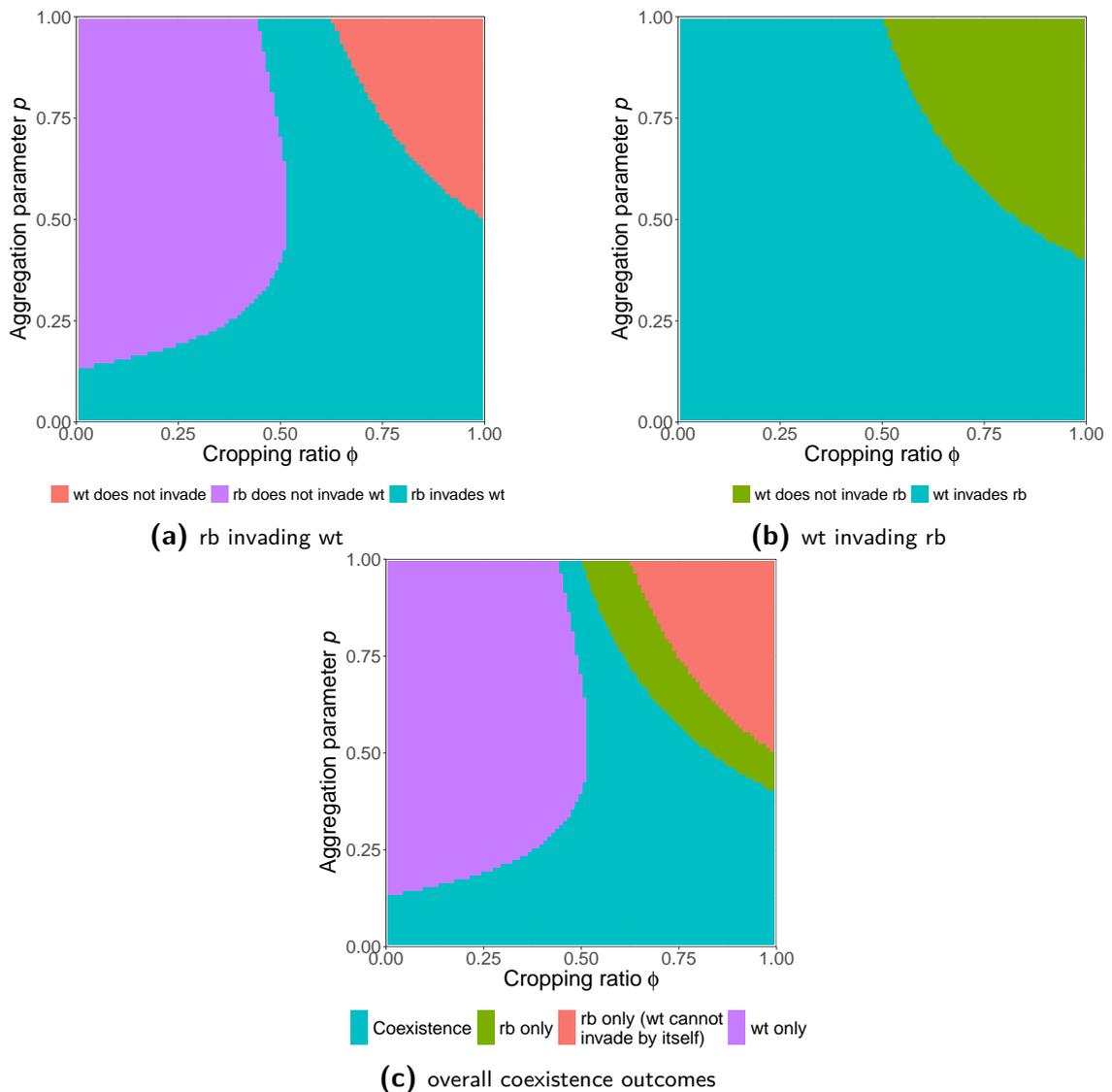


Fig. 4.1 Example plots showing how the pairwise invasion thresholds, for each pathogen strain invading the other strain at its endemic equilibrium (a and b), can be used to identify combinations of parameter values under which coexistence will/will not occur (c). The fitness cost of the resistance breaking trait $\delta = 0.4$, and the susceptibility of the resistant host to the *wt* pathogen strain $\gamma = 0.2$. The infection rate $\beta = 2$ per host tissue unit per unit time, the harvesting/replanting rate $\eta = 1$ per unit time, and the overall host population size $N = 10000$ (although results are independent of N), are used for all Figs unless otherwise stated.

Coexistence of the two pathogen strains in the landscape is generally achieved at the widest range of parameter values when the cropping ratio ϕ is at an intermediate value. The exact position of this intermediate ϕ value depends on the parameters δ and γ which determine the relative competitive ability of the two pathogen strains. When the values of δ and γ are such that neither strain has a large competitive advantage over the other in the landscape, then coexistence is more likely. In general, coexistence is promoted by greater degrees of aggregation in the landscape (low values of p), however the non-monotonic *rb* invasion threshold as a function of p means that there are a number of parameter value combinations where decreasing the value of p (increasing aggregation) will in fact prevent *rb* invasion. Somewhat unsurprisingly

the general response of the system to increasing the cost of the resistant breaking trait (δ) is that the *rb* strain will invade, and *rb* will resist *wt* invasion, at a reduced range of parameter values. Correspondingly, as the susceptibility of the resistant host to the *wt* strain (γ) is increased, the *wt* strain is able to invade the *rb* equilibrium, and the *wt* resists *rb* invasion, at a wider range of parameter values. The patterns described here can also be observed in Fig. 4.3, where δ and γ are shown across their full range as continuous parameters, while the parameters p and ϕ are discretised. There is a wider range of potential qualitative outcomes in Fig. 4.3 due to the wider range of δ and γ values shown. One of these outcomes, that only occurs when $\delta = 0$ and $\gamma = 1$, is that either strain can persist in the landscape depending on their initial conditions. This is because both strains have identical fitness in this case as they can both fully infect both host varieties without any fitness costs. Due to their equal fitness, the frequency of the two strains will remain constant over time and neither will spread at the expense of the other.

The critical p value that most ‘resists’ invasion of the *rb* strain

The non-monotonic threshold for the *rb* pathogen strain invading the *wt* equilibrium produces a peak value for the aggregation parameter p at which *rb* invasion will not occur for the widest range of cropping ratios (ϕ) (Fig. 4.2). The estimated values for this critical p value peak are shown as a function of δ and γ in Fig. 4.4. The general trend is that this critical p is lower when the cost of the resistance breaking trait δ is high, and is higher when the susceptibility of the resistant host to the *wt* strain is high. This results picture is somewhat complicated by the fact that a given value of γ , the peak in the *rb* invasion threshold is relatively flat over a varying range of values for p (Fig. 4.2 e.g. panel $\delta = 0.3$, $\gamma = 0.4$). This means that, even with high levels of numerical precision, the estimated critical p covers a range of potential values, which in many cases extends to $p = 1$. We believe that changes to the range of potential critical p values (flatness of the invasion threshold peak) at different values of δ and γ , are responsible for the dipping trend in critical p when γ reaches higher values, indicating that this dip is likely a numerical artefact. Despite this complication however, we can clearly see a general trend of critical p against δ and γ as stated previously.

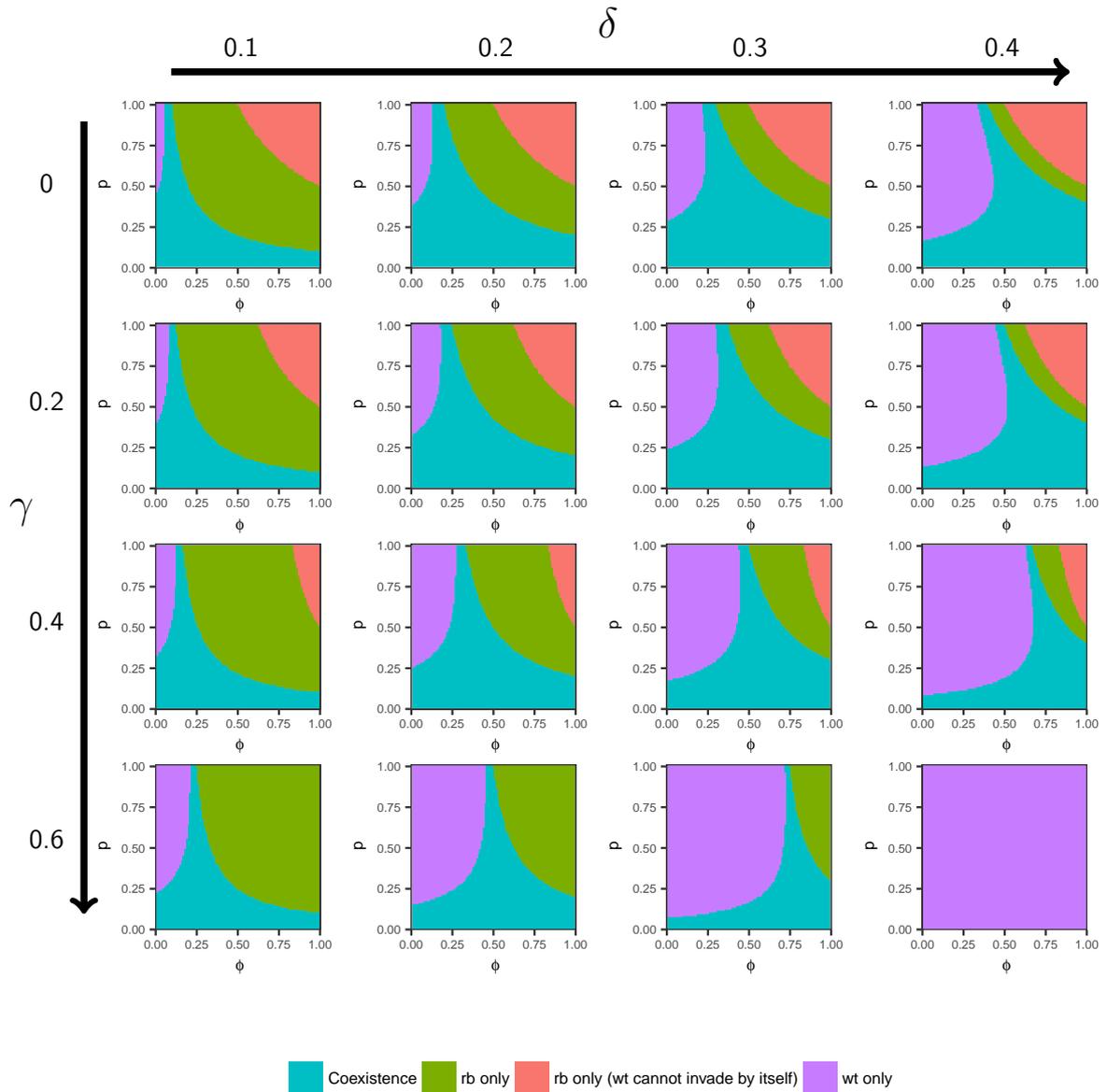


Fig. 4.2 Plots showing the conditions required for invasion and coexistence of the *wt* and *rb* pathogen strains, as a function of the continuous parameters ϕ (cropping ratio of the resistant crop) and p (the aggregation parameter). The plots are arranged in a grid as a function of the discretised parameters δ (fitness cost of the resistant breaking trait) in the horizontal axis and γ (susceptibility of the resistant host to the *wt* strain) in the vertical axis.

Simple model where the cropping ratio $\phi = 1$

The invasion thresholds and coexistence conditions in the simplified two equation model that results when the cropping ratio $\phi = 1$ (meaning that only the resistant crop is planted) are shown in Fig. 4.5. The pattern shown here does not depend on the aggregation parameter p because there is only one variety of host present in the system. In this case, given that the infection rate $\beta = 2$ per host tissue unit per unit time, the harvesting/replanting rate $\eta = 1$ per unit time and the basic $R_0 = \beta/\eta = 2$, the invasion thresholds δ can be described

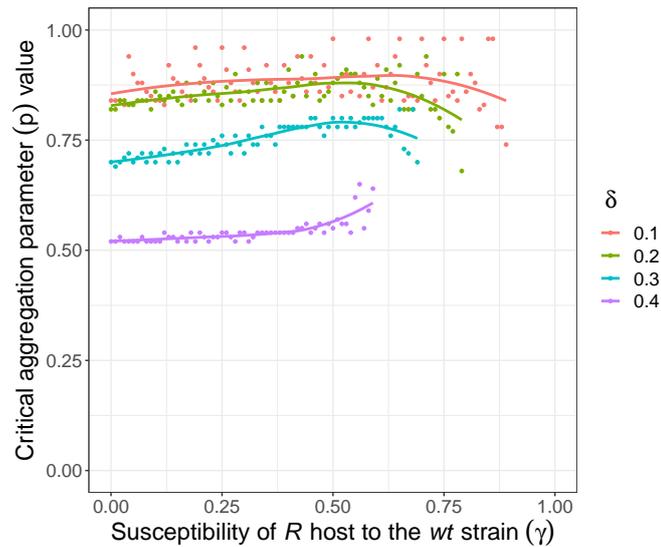


Fig. 4.4 Plot showing the critical values of the aggregation parameter p (the proportion of inoculum that disperses according to the global cropping ratio) where invasion of the wt equilibrium by the rb strain will be prevented at the widest range of potential cropping ratios (ϕ). This is presented as a function of the parameters δ (fitness cost of the resistant breaking trait) and γ (susceptibility of the resistant host to the wt strain). For each value of γ , the median of the range of potential values for critical p is plotted. These critical p values are used to plot a smoothing curve (using the loess method) for each value of δ . Critical p values are only given for δ and γ value combinations where rb invasion can possibly occur.

identical niches (Holt and Dobson, 2006), which here is due to the lack of a susceptible host variety.

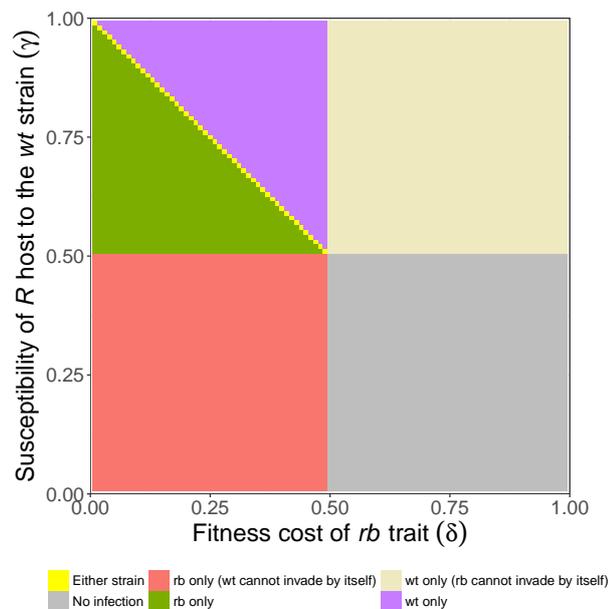


Fig. 4.5 Plot showing the values of the parameters δ (fitness cost of the resistant breaking trait) and γ (susceptibility of the resistant host to the wt strain) where the two pathogen strains will invade and coexist in a simplified version of the model where the cropping ratio $\phi = 1$, meaning that only resistant crop is planted (and so when p is irrelevant).

4.3.2 Model simulation

Effect of the fitness cost of the *rb* trait δ

Results from numerical simulations of the two-patch model use the average seasonal epidemic intensity, which is the average proportion of infected plants over a given number of seasons, as the output metric with which epidemic outcomes are measured. This is the same as the area under the disease progress curve (AUDPC), normalised so that it is expressed as a proportion between 0 and 1. The response of epidemic intensity, at a 50:50 cropping ratio ($\phi = 0.5$), and with an *R* gene that is completely effective against the *wt* strain ($\gamma = 0$), to the fitness cost of the *rb* strain δ is shown in Fig. 4.6a. This is shown at three different values for the aggregation metric p , in order to demonstrate how epidemics differ at high, low and intermediate levels of host variety mixing in the landscape. The timeframe here, of $n_y = 40$, is chosen to match that from the spatially explicit model in chapter 2. This timeframe is intended to capture a mixture of short term epidemiological and longer term evolutionary responses to disease management. In general, epidemic intensities decrease as δ is increased, up until the point where the response completely flattens at intermediate to high values of δ . The point at which this flattening occurs is the point at which the *rb* is no longer fit enough to invade the landscape. Epidemic intensities are also shown to be lower with higher levels of host mixing (higher p) for this range of parameter values.

If we look at the absolute differences between the epidemic intensities at different values of p we can evaluate the reduction in disease due to mixing the host varieties at smaller scales of spatial heterogeneity (Fig. 4.6b). From this we see that the general trend is that of increasing mixing efficacy at higher values of δ , up until the point where the *rb* strain will not invade at either value of p being compared. There is however a marked non-unimodal response when comparing $p = 0.5$ with $p = 0.1$, and a much smaller non-unimodal response when comparing $p = 0.9$ with $p = 0.1$. These responses produce intermediate values of δ at which increased spatial mixing at the given values of p provide the greatest reduction in disease incidence. Whether or not this intermediate peak in the response occurs seems to be partly driven by the lowest value of δ at which the *rb* strain is not able to invade the *wt* in the landscape. It is seen in Fig. 4.6a that this critical value of δ is lower at $p = 0.5$ compared to the other values of p shown. This difference is particularly clear when comparing the green $p = 0.5$ line in 4.6a and the red $p = 0.1$ line, where the *rb* strain is still able to invade at higher values of δ . Note that the general patterns shown in Fig. 4.6 are very similar to those shown in Fig. 5 of chapter

2, particularly when comparing $p = 0.5$ with $p = 0.1$. This suggests that, when looking at the relevant scales of spatial heterogeneity, the qualitative behaviour of the two-patch model from this chapter is very similar to the that of the spatially explicit model used previously.

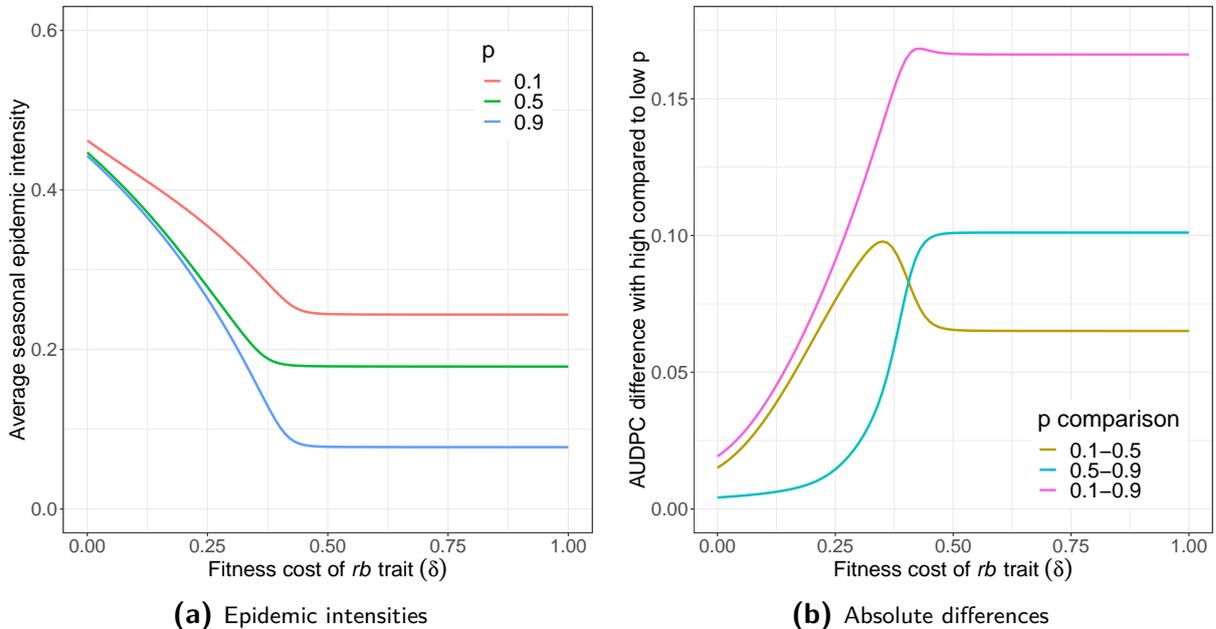


Fig. 4.6 The effect of the fitness cost of the rb trait (δ) on the reduction in average seasonal epidemic intensities (calculated from AUDPCs) from using a high compared to a low degree of mixing (p) in the host landscape. Epidemic intensities as a function of δ at three different values of the landscape aggregation parameter p are shown in (a). The absolute differences in epidemic intensities from using a high compared to a low p value, in pairs of these p values, is shown in (b). The the landscape cropping ratio $\phi = 0.5$, the number of seasons $n_y = 40$, the R host is completely resistant to the wt strain ($\gamma = 0$), and the initial frequency of the rb strain = 0.01.

Partial resistance

The response of epidemic intensities to the susceptibility of the R host to the wt strain (γ) at different values of the aggregation metric p is shown in Fig. 4.7a. This is shown for an intermediate fitness cost of the rb trait ($\delta = 0.3$), where the effects of increased spatial mixing (as shown in Fig. 4.6b) are relatively strong, at a 50:50 cropping ratio ($\phi = 0.5$). In all cases shown here, epidemic intensities increase as the value of γ increases. Increasing the degree of spatial mixing by using a higher value of p decreases epidemic intensities at low values of γ , but can potentially increase intensities at higher values of γ . When looking at the reduction in epidemic intensities from using a higher value of p , the dominant trend is that of the spatial effect decreasing in magnitude towards zero as γ is increased. Due to the crossing over of some lines in Fig. 4.7a however, there is a negative effect of increased spatial mixing when looking at $p = 0.5$ compared to $p = 0.1$, and $p = 0.9$ compared to $p = 0.1$. This means that

in these cases epidemic intensities increase with the greater degree of between host mixing. While the comparison of $p = 0.9$ vs $p = 0.5$ does not produce this negative effect, it does give a non-unimodal response at lower values of γ that produces an intermediate peak in the efficacy of increased spatial mixing that is not present with the other p value comparisons.

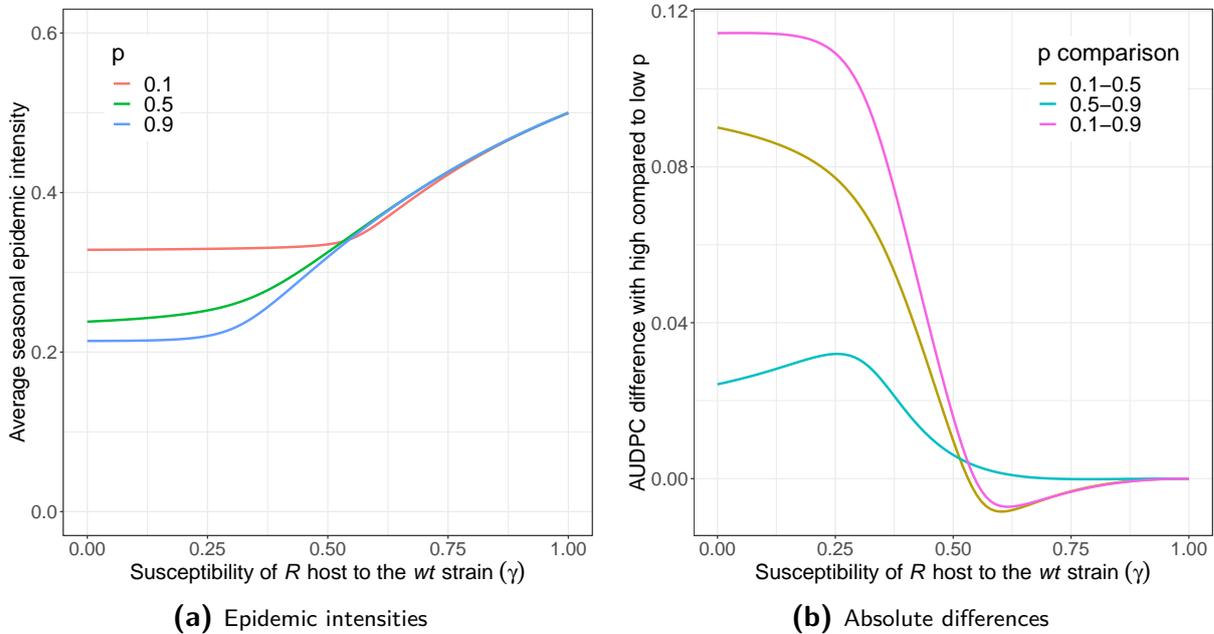


Fig. 4.7 The effect of the susceptibility of the R host to the wt pathogen strain (γ) on the reduction in average seasonal epidemic intensities (calculated from AUDPCs) from using a high compared to a low degree of mixing (p) in the host landscape. Epidemic intensities as a function of γ at three different values of the landscape aggregation parameter p are shown in (a). The absolute differences in epidemic intensities from using a high compared to a low p value, in pairs of these p values, is shown in (b). The landscape cropping ratio $\phi = 0.5$, the number of seasons $n_y = 40$, the fitness cost of the rb trait $\delta = 0.3$, and the initial frequency of the rb strain = 0.01.

The regions of δ and γ parameter space where the maximum efficacies of disease control, through increased host spatial mixing, occur with different p value comparisons are seen in Fig. 4.8. These results, which Figs 4.6b and 4.7b are slices of, are also for a 50:50 cropping ratio ($\phi = 0.5$) and a timeframe of $n_y = 40$. From these heatmaps we can clearly see that there is frequently a band of parameter space (darker blue) that produces an intermediate peak in the efficacy of disease suppression through increased spatial mixing in both the δ and γ dimensions. This non-monotonic response is the same as those seen in Figs 4.6b and 4.7b. Note that these non-monotonic regions are somewhat related to the parameter values required for the coexistence of the two pathogen strains in the landscape at one or both of the p values being compared (Figs 4.8e f and g). The regions of parameter space, at intermediate to high values of γ , where a negative spatial suppressive effect occurs (orange region) can be also clearly seen (Figs 4.8a c and d). Whether or not the intermediate peaks or negative regions of spatial disease suppression efficacy occur depends on the particular degrees of host mixing (p

values) that are being compared. For example, regions of negative efficacy are not observed when comparing intermediate to high levels of between host mixing ($p = 0.9$ vs $p = 0.5$ - Fig. 4.8b). Fig. 4.8d shows a comparison between the p values that were calculated for the regularly divided landscapes (using either extreme value of the edge/area ratio metric) compared in chapter 2 Fig. 2.7. This demonstrates that the qualitative behaviour of the two-patch model is very similar to that of the spatially explicit model used previously, when looking at the relevant scales of spatial heterogeneity in the host population. The figures in chapter 2 and the current chapter that show the same qualitative response of resistance durability to a given parameter, or combination of parameters, are shown in table 4.2.

Effect of timescale

If we relax the assumption of a fixed timeframe over which we measure the efficacy of disease control through increased spatial mixing, we can show the strength of this spatial suppressive effect as a function of the length of the timeframe of interest (Fig. 4.9). These results compare the use of intermediate levels of spatial mixing compared to low levels ($p = 0.5$ vs $p = 0.1$). As is also seen in chapter 2 Fig. 2.8 using the spatially explicit model, there is frequently an intermediate number of seasons over which the maximum spatial suppressive effect occurs. This intermediate peak occurs at low fitness costs of the resistance breaking trait ($\delta = 0$ and $\delta = 0.2$) but not at the higher values of δ shown where the spatial effect strength simply increases as it approaches its equilibrium value. Where the intermediate peak spatial effect does occur, it is observed over a smaller number of seasons with a low fitness cost of the rb trait (δ). Note that for $\delta = 0$, the rb strain (which can equally infect both host varieties) will always completely replace the wt strain at equilibrium (unless $\gamma = 0$), so any spatial suppressive effect seen (red line) is only ever transient while the wt is still present in the landscape. The complex effects of timescale length demonstrated here mean that the values of δ and γ which are most amenable to spatial disease control vary as a function of the timescale.

Table 4.2 The figures in chapters 2 and 4 that demonstrate the same qualitative effect of a given parameter(s) on the durability of disease resistance.

Chapter 2	Chapter 4	Parameter
Fig. 2.5	Fig. 4.6	Effect of the fitness cost of the rb trait δ
Fig. 2.6	Fig. 4.7	Effect of the susceptibility of the R host to the wt strain γ
Fig. 2.7	Fig. 4.8	Combined effect of δ and γ
Fig. 2.8	Fig. 4.9	Effect of the number of seasons n_y

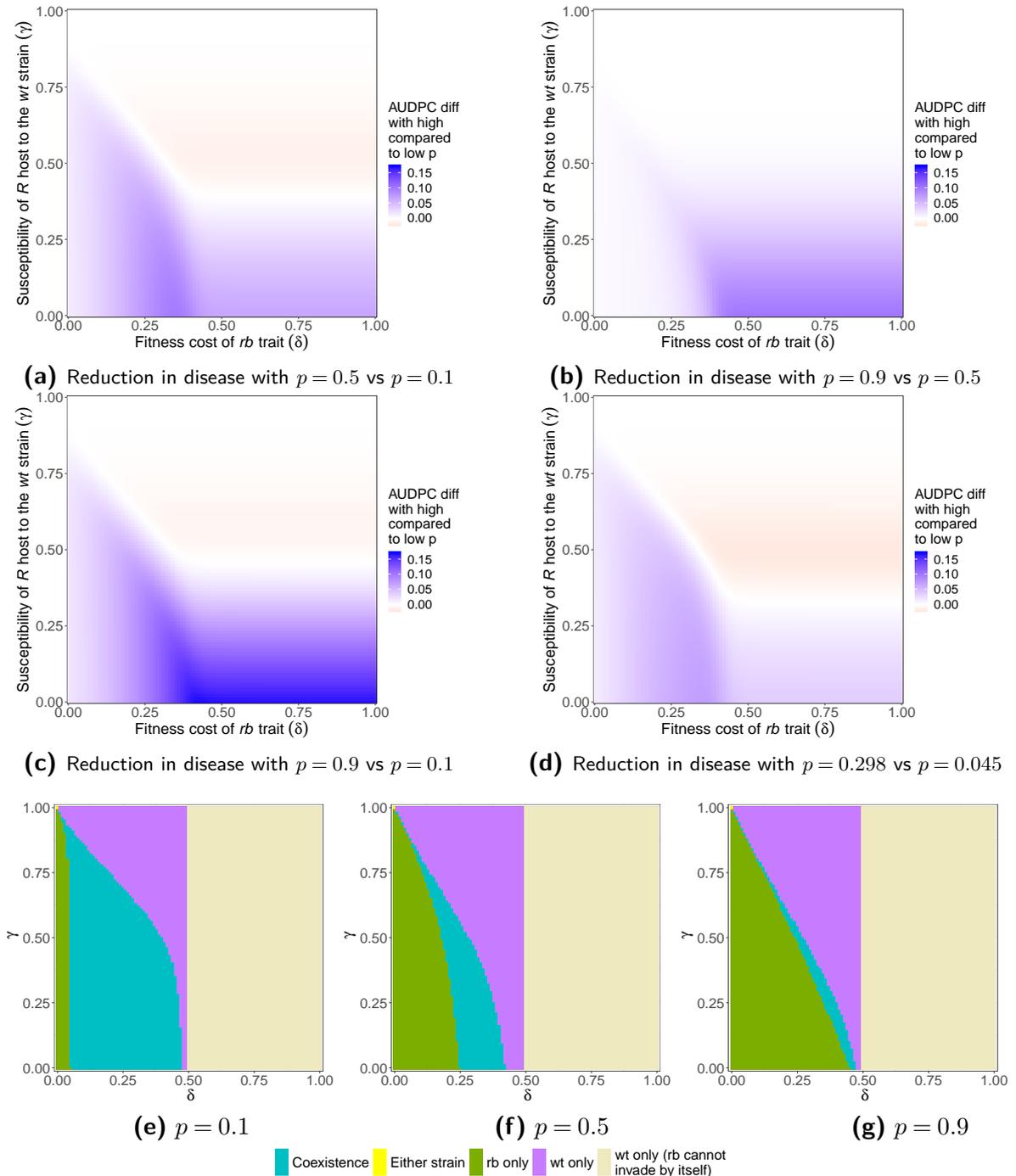


Fig. 4.8 The effect of the fitness cost of the *rb* trait (δ) and the susceptibility of the *R* host to the *wt* pathogen strain (γ) on the reduction in average seasonal epidemic intensities (calculated from AUDPCs) from using a high compared to a low degree of mixing (p) in the host landscape. The absolute differences in epidemic intensities from using high compared to low p values, in pairs of p values, are shown in (a), (b), (c) and (d). The parameter conditions required for invasion and coexistence of the two strains are shown in (e), (f) and (g). The the landscape cropping ratio $\phi = 0.5$, the number of seasons $n_y = 40$, and the initial frequency of the *rb* strain = 0.01.

Effect of the cropping ratio ϕ on the optimal strategy for resistance durability

By allowing the cropping ratio ϕ to freely vary, we can observe how this parameter interacts with the degree of aggregation in the landscape, at different values for the fitness cost of the

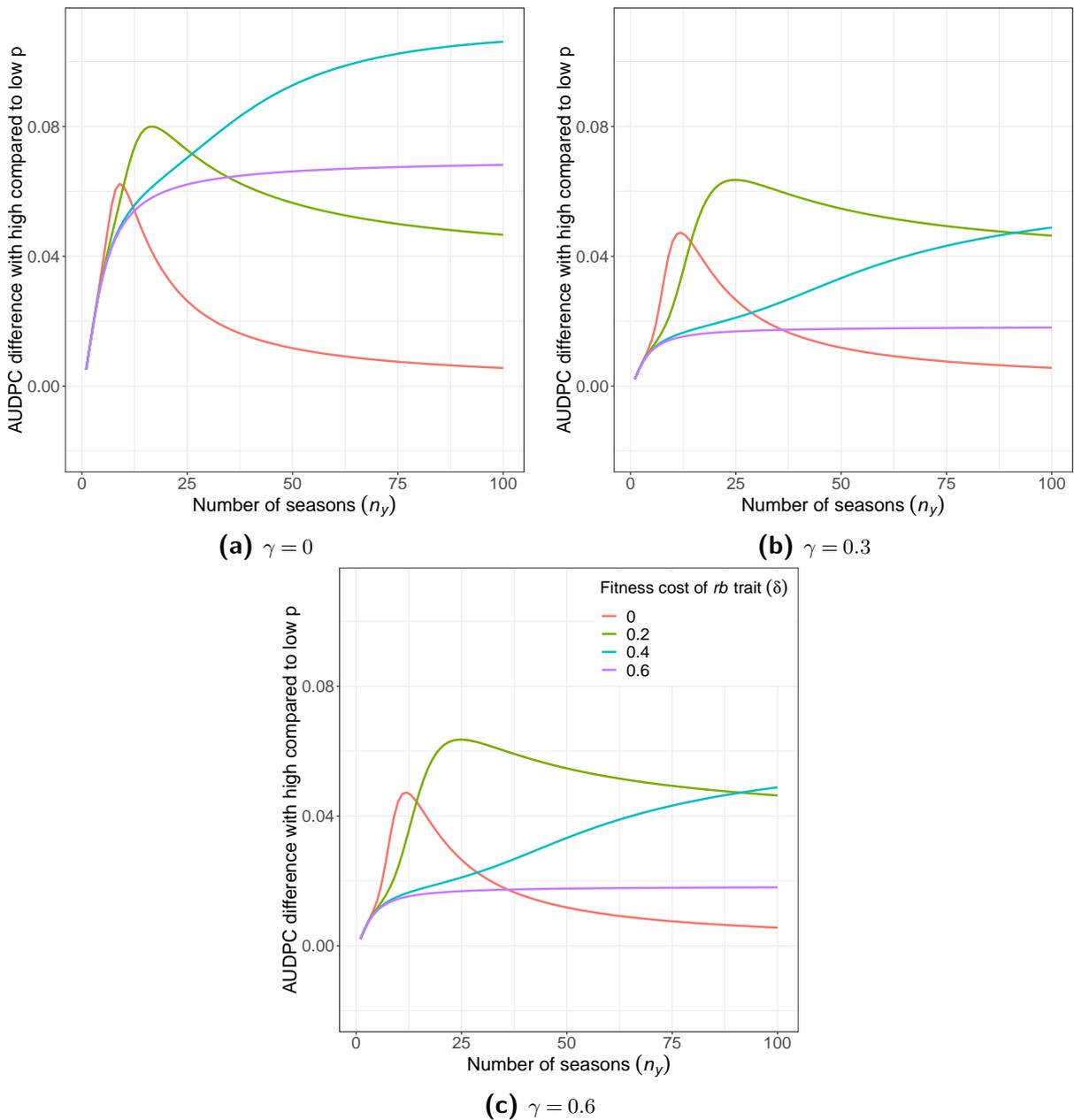


Fig. 4.9 The effect of the number of seasons (n_y) on the reduction in average seasonal epidemic intensities (calculated from AUDPCs) from using a high degree of mixing (p) in the host landscape. The absolute differences in epidemic intensities from using $p = 0.5$ compared to $p = 0.1$ are shown for a range of values of the fitness cost of the rb trait (δ) and the susceptibility of the R host to the wt pathogen strain (γ). The landscape cropping ratio $\phi = 0.5$, and the initial frequency of the rb strain = 0.01.

rb strain (δ) and the susceptibility of the R host to the wt strain (γ). The equilibrium disease frequency resulting from these parameter combinations is shown in Fig. 4.10. These plots show that the optimal cropping ratio ϕ , giving the lowest equilibrium disease frequency for a given value of the aggregation parameter p (red line), is either $\phi = 1$ or follows almost perfectly the wt strain invasion thresholds shown in Fig. 4.2. Since increasing the value of ϕ above this threshold will have no effect on the rb strain by itself when the wt is not present, these

values of optimal ϕ must produce the same epidemic outcome as when $\phi = 1$. The overall optimum cropping ratio (red dot) is either at or close to the threshold where the wt is driven extinct. Any visible difference between the overall optimum ϕ and this wt invasion threshold is simply due to numerical error (caused by the simulation taking a very long time to reach equilibrium). This was revealed by closely comparing additional simulations results for sets of parameter values close to the overall optimum. It is the case that the effect of increasing the degree of mixing in the landscape (using a higher p) is stronger at intermediate cropping ratios (as moving from low to high p here crosses the most contour lines). Despite this effect at intermediate cropping ratios however, it is still optimal to simply use as much resistant crop (high ϕ) as possible.

If the the cropping ratio is the restricting factor on the other hand, we are interested in the optimal degree of aggregation p for a given value of ϕ (green line). This is at or very close to $p = 1$ (total mixing) at the overall ϕ and p optimum (red dot). At greater given ϕ values than this overall optimum the optimum p follows the wt invasion threshold as before and so must be higher enough that the wt strain is forced extinct. At constrained ϕ values that are lower than the overall optimum, the optimal $p = 1$ for low values of γ . For the higher values of γ shown here however, the optimal p (for lower ϕ) is frequently a low to intermediate value, especially at higher values of δ . This means that increasing the degree of mixing in the landscape in these parameter regions can increase the severity of disease, which is the same phenomenon shown by the negative regions in Fig. 4.8. The line showing the optimal p value mostly follows the invasion threshold for the rb pathogen strain shown in Fig. 4.2. This implies that the optimum degree of host mixing is high enough that the the rb strain is not able to invade but no higher.

Fig. 4.11 shows the average disease epidemic intensity, calculated from AUDPCs between $y = 0$ and $y = n_y = 40$. The pattern here is very similar to that given by the equilibrium disease frequencies (Fig. 4.10), with the exception that the lines representing the optimal p and ϕ values no longer follow exactly the the rb and wt strain invasion thresholds. This is due to the transient effect of the initial large number of individuals infected by the resident wt strain, meaning that a higher cropping ratio ϕ or degree of mixing p is required in order to force the wt extinct within the $n_y = 40$ timeframe. There may be an intermediate optimum cropping ratio at very high values of p , however the differences in epidemic intensity between these optima and when $\phi = 1$ are vanishingly small. Alternatively, if preventing the invasion of the rb strain is the priority, due to the use of a low cropping ratio ϕ , the optimal degree of mixing

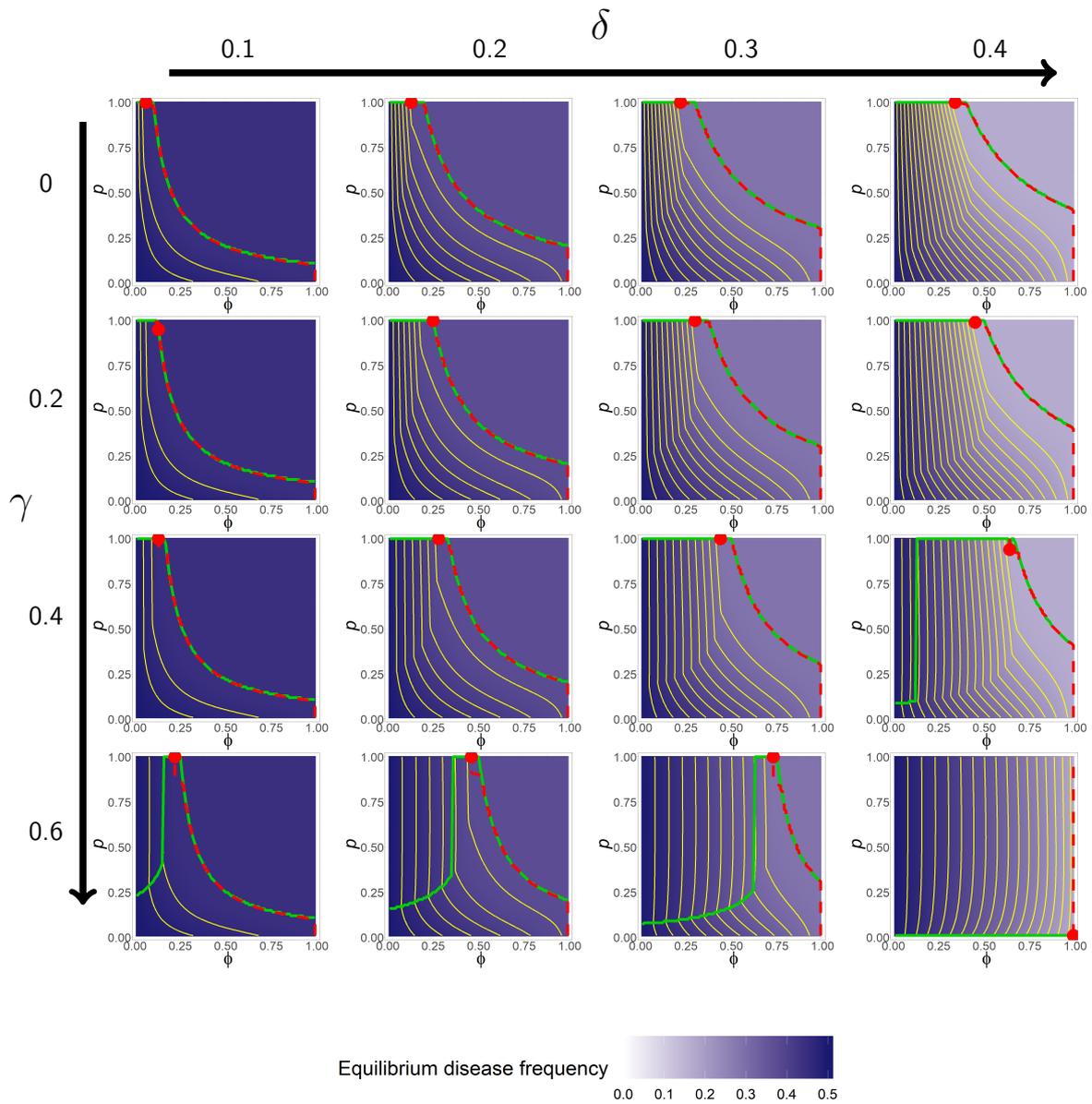


Fig. 4.10 The effect of the cropping ratio ϕ and the aggregation parameter p on equilibrium disease frequency (simulation results at $y = 1000$). This is shown for different fitness costs of the rb trait (δ), and different susceptibilities of the R host to the wt pathogen strain. The ϕ and p values where disease frequency is lowest, for given values of δ and γ , is shown as red point. The values of ϕ where disease frequencies are lowest for given values of p are shown as a dotted red line. The values of p where disease frequencies are lowest for given values of ϕ are shown as a green line. The yellow contour lines, for aiding the visualisation of the change in disease frequency, have a bin width of 0.02. The number of seasons $n_y = 1000$, and the initial frequency of the rb strain = 0.01.

p is lower than at equilibrium (Fig. 4.10). This is due to the slow spread of the rb strain in the face of the resident wt population. Qualitatively similar results to these were obtained when using different numbers of seasons (ranging from $n_y = 10$ to $n_y = 100$).

Taking the average seasonal epidemic intensity from $y = 10$ to $y = n_y = 40$, thereby excluding the importance of controlling the epidemic in the first 10 seasons, we see a more complex picture for optimal management strategies (Fig. 4.12). Here the rings formed by the

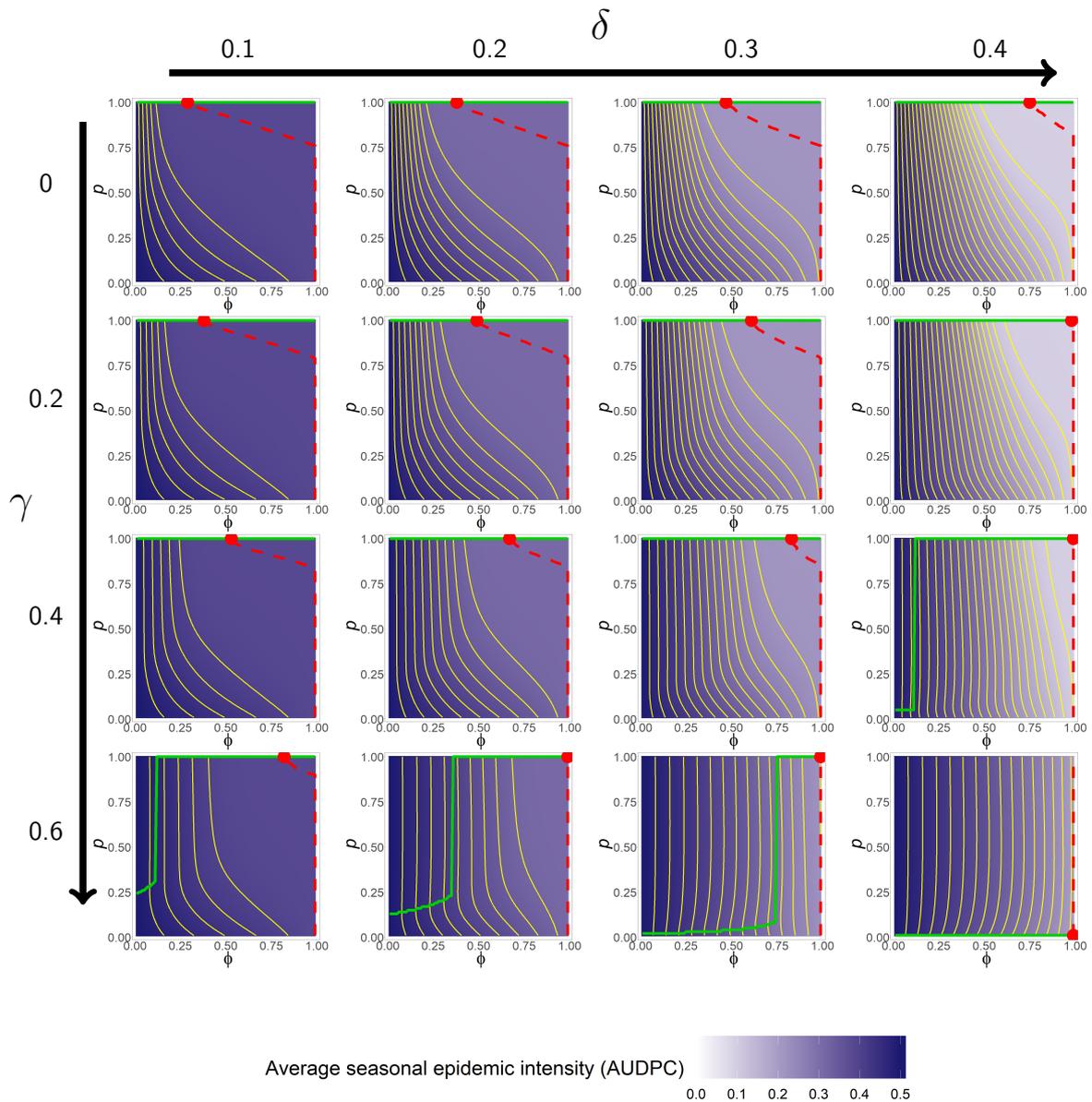


Fig. 4.11 The effect of the cropping ratio ϕ and the aggregation parameter p on average seasonal epidemic intensities calculated from AUDPCs between $y = 0$ and $y = n_y = 40$. This is shown for different fitness costs of the rb trait (δ), and different susceptibilities of the R host to the wt pathogen strain. The ϕ and p values where epidemic intensity is lowest, for given values of δ and γ , is shown as red point. The values of ϕ where epidemic intensities are lowest for given values of p are shown as a red line. The values of p where epidemic intensities are lowest for given values of ϕ are shown as a green line. The yellow contour lines, for aiding the visualisation of the change in epidemic intensity, have a bin width of 0.02. The number of seasons $n_y = 40$, and the initial frequency of the rb strain = 0.01.

contour lines show that there are true intermediate optimum cropping ratios (ϕ), with higher optimum ϕ values at lower p values (less mixing). There are also true intermediate optimum p values (degrees of mixing) at cropping ratios above the overall ϕ optimum. These optimum p values, which were at or no different to $p = 1$ in Figs 4.10 and 4.11, are lower for higher cropping ratios. The intermediate optimum ϕ values are higher for higher values of both δ and γ . Optimal p values at cropping ratios higher than the overall ϕ optimum are also higher with

higher values of δ and γ . Optimal p values at cropping ratios lower than the overall ϕ optimum on the other hand are lower with higher values of δ and γ . If we measure the AUDPCs for the epidemic from a slightly later season number, such as from $y = 20$ to $y = n_y = 40$ for example (not shown), then the non-unimodal responses in the ϕ and p dimensions (contour line rings) are more prominent at higher δ and γ values compared to Fig. 4.12, and less prominent at lower values. These results indicate that the optimal strategy for the maintenance of resistance durability depends on the timescales over which minimisation of epidemic intensity is measured, and may therefore not be constant over time.

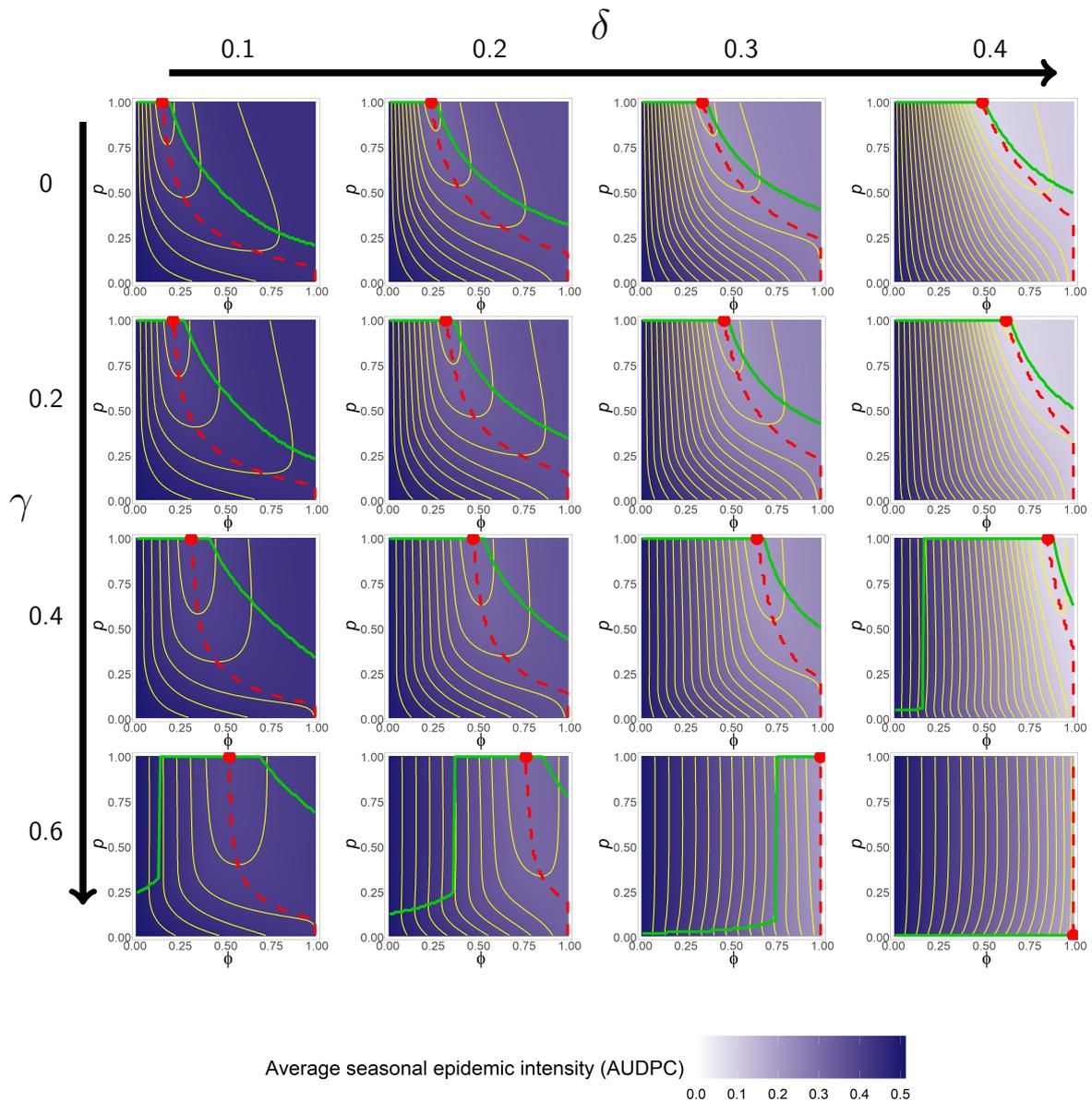


Fig. 4.12 The effect of the cropping ratio ϕ and the aggregation parameter p on average seasonal epidemic intensities calculated from AUDPCs between $y = 10$ and $y = n_y = 40$. This is shown for different fitness costs of the *rb* trait (δ), and different susceptibilities of the *R* host to the *wt* pathogen strain. The ϕ and p values where epidemic intensity is lowest, for given values of δ and γ , is shown as red point. The values of ϕ where epidemic intensities are lowest for given values of p are shown as a red line. The values of p where epidemic intensities are lowest for given values of ϕ are shown as a green line. The yellow contour lines, for aiding the visualisation of the change in epidemic intensity, have a bin width of 0.02. The number of seasons $n_y = 40$, and the initial frequency of the *rb* strain = 0.01.

4.4 Discussion

In this work we have applied our spatially implicit framework, for representing the effect of the scale of spatial heterogeneity between two host varieties on disease progress, to the question of how to maximise the durability of crop disease resistance. Our results show that resistance durability, defined by the extent that epidemic intensity can be controlled over evolutionary timescales, is significantly affected by a variety of factors that are included in the model. Most

importantly we have demonstrated the complex nature of the interaction between the scale of spatial heterogeneity and the resistance cropping ratio. We are able to explore this interplay in a rigorous and clearly defined manner due to the specific construction of the spatially implicit framework, which ensures that the two relevant parameters are independent of one another. These two parameters represent two of the main levers that can be adjusted by growers to control disease, given the specific pathosystem and resistances available to them. We have also shown how this interaction between the scale of spatial heterogeneity and the cropping ratio depends in turn on the two major parameters that define the nature of the interaction between the host and pathogen genotypes. These are the fitness cost of the resistance breaking trait and the susceptibility of the resistant host to the wild-type pathogen strain, which can also be understood as measuring the efficacy of the resistance gene.

When compared to the spatially explicit model studied in chapter 2, the mathematically simpler nature of our spatially implicit framework has enabled us to carry out a partially analytical invasion analysis as part of our study of resistance durability. The information gleaned from this analysis both aids in the interpretation of our later simulation results, and also demonstrates how growers might seek to prevent the introduction of certain disease strains in the first place. The extent to which an agricultural landscape is able to 'resist' the invasion or introduction of a resistance breaking strain could be seen as a measure of resistance durability in of itself. This is similar in some ways to measures focussed on the time until emergence of a newly resistance breaking or virulent strain (Papaix et al., 2018; van den Bosch and Gilligan, 2003). The time until emergence is usually thought of as depending on two factors: the probability that a new resistance breaking genotype will emerge through mutation, and the probability that this new genotype will become established. The invasibility of the landscape by the resistance breaking strain is similar to this probability of establishment component, given that both of these processes are made easier with high proportions of the resistant crop that the new strain is adapted to infecting.

The dominant pattern in our results is that coexistence of the wild-type and resistance breaking strains within the landscape is facilitated by large scales of spatial heterogeneity, where there is a great deal of spatial aggregation within the individual varieties, and by intermediate cropping ratios. This relatively intuitive result is due to the fact that the dilution effect, where inoculum disperses frequently onto host tissue where it is less specialised and reduces infection efficiency, is weaker when hosts of the same variety are highly aggregated (Mundt, 2002). The

same effect of the scale of spatial heterogeneity was also seen in chapter 2, due to the same set of processes. Intermediate cropping ratios in turn mean that neither pathogen strain has an overwhelming fitness advantage in the landscape due to the greater availability of its preferred host material and so neither is able to drive its competitor strain to extinction. The exact range of cropping ratios and scales of spatial heterogeneity that enable coexistence are affected by strongly by the fitness cost of the resistance breaking strain and the efficacy of resistance. This is due to the fact that change to these variables directly changes the relative fitnesses of the two strains.

The most interesting finding to come from these invasion analyses is however that the threshold for the resistance breaking strain invading the wild-type does not always respond in a unimodal fashion to the scale of spatial heterogeneity. This means that there are some parameter values where increasing the degree of mixing between the host varieties will actually make invasion by the resistance breaking invasion strain easier, in contrast to the usual pattern where increased mixing hinders invasion due to increased infection dilution. This occurs because although increased mixing leads to greater proportional dispersal of the resistance breaking inoculum onto the susceptible cultivar, it also means that the size of the resident wild-type population is greatly reduced by its own reduced infection efficiency as it disperses more frequently onto to the resistant crop. The advantage to the resistance breaking strain of having less competition from the wild-type outweighs the negative effects of increased mixing in these cases. This can also be seen as the increased between host mixing having a more severe impact on the wild-type, due to the fact that the resistance breaking strain is able to equally infect both host varieties when the wild-type is absent.

The non-unimodal response of the invasion threshold means that in these cases there is an intermediate degree of between host mixing where the landscape is most able to resist the invasion of the resistance breaking strain. This intermediate degree of mixing is lower with a greater costs of the resistance breaking trait. This is likely because as the resistance breaking trait becomes less competitive it has more to gain from the resident wild-type population being suppressed, and so there is a wider range of degrees of spatial mixing where the resistance breaking strain benefits from increased mixing. Conversely as the efficacy of the resistance gene is reduced the intermediate degree of mixing where resistance breaking invasion is most difficult increases. This could be because suppression of the resident wild-type population with

greater mixing is reduced when the wild-type is better able to infect both host varieties, and so there is less benefit to the invading resistance breaking strain.

We have carried out a range of numerical simulations for the purpose of comparing with and confirming the results taken from the spatially explicit model studied in chapter 2. From these simulations we used the epidemic intensity, from the area under the disease progress curve, as our measure of resistance durability. The results from these simulations (Figs 4.6, 4.7, 4.8, 4.9) are generally very similar to those previously found (chapter 2 Figs 2.5, 2.6, 2.7, 2.8), but have the added benefit of being able to compare epidemic intensities across a wider range of scales of spatial heterogeneity, due to the use of a more clearly defined spatial aggregation parameter. The key results from the spatially explicit model used in chapter 2 showed that the efficacy of mixing different varieties at smaller scales of spatial heterogeneity varied as a function of three parameters. These were the cost of the resistance breaking trait, the efficacy of the resistance gene and the length of the timeframe of interest. In particular there was shown to be a maximum efficacy of mixing at intermediate values for the cost of the resistance breaking trait and the timeframe length. These non-unimodal responses have been replicated in the two-patch spatially implicit model, but only when comparing low versus intermediate levels of mixing. The specific degrees of mixing in the new model that correspond to the values of the edge/area ratio metric used in the spatially explicit model have also been calculated and tested, again producing very similar results.

When comparing intermediate to high degrees of mixing on the other hand however we frequently see that the non-unimodality of results is weaker. An example of this is that, with a very effective resistance gene, increasing the fitness cost of the resistance breaking trait simply increases the efficacy of greater spatial mixing, rather than producing an intermediate peak. Whether or not the intermediate peak occurs seems to be related to the shape of the resistance breaking strain invasion threshold, as the maximum value of the fitness cost of the resistance breaking strain that permits invasion of that strain is significantly lower with intermediate levels of mixing compared to low levels. When moving from intermediate to high levels of mixing on the other hand the maximum value of this fitness cost that permits invasion instead increases slightly. The efficacy of spatial disease control depends on whether increasing the degree of mixing from one level to another significantly changes the position of the resistance breaking strain's invasion threshold, such that invasion is more difficult. If this is the case there will be an intermediate peak in spatial effect close to the parameter value where this difference in

invasion threshold with different degrees of mixing occurs. If however increasing the degree of mixing does not make it harder for the resistance breaking strain to invade, or if it makes invasion easier, then the control strategy will not be especially effective at any intermediate fitness cost of the resistance breaking strain.

This theme of non-unimodality being related to whether changing the level of mixing leads to significant movement of an invasion threshold also applies to the efficacy of the resistance gene. Here we see a previously unobserved intermediate peak in the efficacy of moving from intermediate to high levels of mixing. In this case the greater level of mixing forces the wild-type strain to be driven towards extinction at the high efficacies of the resistance gene where we see the intermediate peak. The non-unimodal responses to both the fitness cost of the resistance breaking trait and the efficacy of the resistance gene combine to form a band of maximum efficacy of increased mixing at intermediate values for these traits (Fig. 4.8). The prominence of this intermediate band, and therefore the presence of a maximum effect at intermediate parameter values, depends again on the specific levels of mixing being compared. In a similar manner to the patterns described above, the existence of this intermediate band seems predicated on whether increasing the level of mixing from one level to another significantly changes the invasion thresholds, and therefore conditions required for coexistence, in the model.

As was the case in the earlier spatially explicit model, there is a region of parameter space, at intermediate to high levels of the fitness cost of the resistance breaking trait and intermediate to low efficacies of the resistance gene, where increased mixing can worsen an epidemic and lead to greater yield loss. This phenomenon occurs only at parameter values that prevent the invasion of the resistance breaking strain, leaving only the wild-type, in at least one of the two levels of mixing being compared. This pattern, which closely replicates that seen using the spatially explicit model in chapter 2, is only observed again when comparing low to intermediate levels of mixing, and is absent when comparing intermediate to high levels. The benefit to the disease of increased mixing, is that it allows the susceptible hosts in the landscape to act as stepping stones and increase disease incidence on the nearby partially resistant hosts. The lack of this effect when moving to even higher levels of mixing indicates that in this parameter range there is no further benefit to even more closely mixed stepping stones.

In general, the replication of the complex and frequently non-intuitive results from chapter 2, using our spatially implicit framework, demonstrates the validity of our new approach as a method of representing the scale of spatial heterogeneity. Furthermore, the comparison of a

wider range of mixing scales using a more clearly defined spatial parameter, and the analysis of invasion thresholds, had provided us with further insight into the nature and occurrence of these patterns. The key take home point from this exercise is that when maximising resistance durability it is highly important to know the values for the fitness parameters of the specific pathosystem, as well as the precise scales of spatial heterogeneity that are being compared. Depending on the values of these parameters, highly complex behaviours may emerge, with major implications for the efficacy of spatial variety mixing as a disease control strategy for resistance durability.

In our study of how the scale of spatial heterogeneity interacts with the resistance cropping ratio we generally move away from measuring the strength of the spatial effect, and instead focus on identifying the optimum conditions for disease control. This is because, with two parameters that could be manipulated by a grower in a specific disease system, it becomes impractical to compare spatial effect strengths when the baseline level of infection at a given degree of mixing varies as a function of the cropping ratio. If we are concerned with minimising the disease frequency at equilibrium, the overall optimum strategy is to use a sufficiently high cropping ratio and degree of mixing that forces the wild-type strain to become extinct. This means that there is no intermediate optimum cropping ratio that minimises disease incidence at equilibrium. The likely reason for this result is that at the disease equilibrium there is no benefit to either controlling the initial transient wild-type epidemic or slowing down the spread of the resistance breaking strain. Instead the priority is to minimise the loss of yield to the stable pathogen population at equilibrium. Due to the fitness cost of the resistance breaking trait, this strain leads to fewer infections overall than the wild-type strain does, despite the resistance breaking strain's ability to infect both varieties. This conclusion contrasts with the findings of Papaïx et al. (2018), where maximum disease control at equilibrium was achieved using an intermediate cropping ratio. A potential reason for this is that Papaïx et al. (2018) modelled the efficacy of resistance as a continuous trait that eroded over time in a manner similar to a quantitative resistance, as opposed to using a gene-for-gene qualitative resistance as we have done here. Due to the highly complex nature of the model used by Papaïx et al. (2018) there may also be other factors behind this difference in conclusions that are not easily possible to elucidate. If the cropping ratio is constrained to a low or intermediate value, perhaps by the limited availability of resistant plant material, then there may be an optimum intermediate scale of spatial heterogeneity when using a partially effective resistance gene. This intermediate

optimum is a result of the negative spatial effect of increased mixing when only the wild-type is present as described earlier for a 50:50 cropping ratio. The optimum largely follows the invasion threshold for the resistance breaking strain, meaning that the optimum degree of mixing is enough that it prevents the invasion of that strain, but no more.

The response of the disease incidence at equilibrium to changing parameter values helps us to understand the behaviour of the model, and in particular how this relates to the two strains' invasion thresholds. It is however a less relevant measure of resistance durability since it may take a large number of seasons for the disease equilibrium to be reached, during which time the importance of controlling transient disease dynamics is discounted (Galvani, 2003; White and Gilligan, 2006). Measuring durability instead by the epidemic intensity over a given evolutionary timeframe, we in fact see a very similar qualitative pattern of results as at equilibrium. It is again always optimal to use a high enough resistance cropping ratio or mix the two crop varieties at small enough scales of spatial heterogeneity so that the wild-type strain is forced to extinction. If the cropping ratio is limited to low values, there is again an intermediate scale of spatial heterogeneity that provides optimal disease control. The difference with the results at equilibrium is that, due to the initial transient resident wild-type population, a higher cropping ratio or greater degrees of mixing are required to sufficiently suppress this initial wild-type population. These results contrast with those from Fabre et al. (2012) who found that there is frequently an intermediate optimum cropping ratio that minimises epidemic intensity. Although the study of Fabre et al. (2012) was not looking at different scales of spatial heterogeneity, they did also model a simple gene-for-gene interaction, and used the same measure of resistance durability as in our study. One difference however was the use of a semi-discrete seasonal model with primary infection from a reservoir component, as was included in our spatially explicit model from chapter 2. The more complex representation of spatial structure and the explicit inclusion of resistance-breaking fitness costs in our model may also be factors behind the differing sets of conclusions.

We did however show that there may be a true intermediate optimal cropping ratio if the importance of controlling the initial transient wild-type population is reduced. Without doing this, the need to control the large initial wild-type epidemic, combined with the advantage of using a high cropping ratio at equilibrium, is so great that it washes out any benefit of planting the susceptible crop. The intermediate optimum may come from a trade-off between the conflicting advantages of controlling the wild-type epidemic and of slowing down the spread of

the resistance breaking strain. If totally mean field mixing of the host varieties is not possible, due to the cost or practicality of such an approach, then the optimal cropping ratio increases as the varieties become more aggregated. This is because with less variety mixing to control the wild-type epidemic, a higher proportion of the resistant crop is required to compensate. The optimal cropping ratio is higher with a greater fitness cost of the resistance breaking trait as the importance of slowing down the spread of this strain is reduced when it has a lower fitness. Higher cropping ratios are also needed with a reduced efficacy of the resistance gene in order to provide adequate control of the wild-type epidemic. The fact that the optimal cropping ratio changes depending on the number of seasons after which the efficacy of disease control is measured potentially suggests that the optimal cropping ratio is generally not constant over time. An alternative explanation for this pattern however is that the intermediate optimum cropping ratio is an artefact of the fact that the initial conditions are different for different cropping ratios at the time when we begin measuring the area under the disease progress curve. A truer test for this effect might involve beginning a simulation with all resistant crop planted, to control the initial wild-type epidemic, before switching to an alternative strategy for which the presence of an intermediate optimal cropping ratio could be tested. Another test might involve the cropping ratio dynamically changing in a continuous manner over time to achieve the optimum strategy. Optimal control theory could be used for this purpose in future research (Bussell et al., 2019; Forster and Gilligan, 2007; Hocking, 1991; Sethi and Staats, 1978).

In conclusion, this chapter has successfully demonstrated that a simplified spatially implicit model can be used to study the complex effects of spatial heterogeneity on the durability of disease resistance. These effects were previously explored in chapter 2 using a spatially explicit model, and our new approach has replicated these earlier results while providing greater flexibility and more analytical possibilities. We have shed greater light on the nature of the complex dynamics that effect the efficacy of spatial variety mixing as a control strategy to improve resistance durability. Furthermore, we have also explored the interaction between the scale of spatial heterogeneity and the resistance cropping ratio, two variables that have not previously been clearly disentangled.

4.5 Appendix 1 - Finding the *wt* only equilibrium

If we eliminate W from the *wt* only pair of ODEs given by Eqns (4.29) and (4.30) using the computer algebra program *wxMaxima* we get the following quartic polynomial:

$$0 = a_4 X^4 + a_3 X^3 + a_2 X^2 + a_1 X \quad (4.48)$$

where:

$$a_4 = -(p-1)(p\Phi-1)\beta^3\gamma \quad (4.49)$$

$$a_3 = -N(\Phi-1)\beta^2(p^2\Phi^2\gamma\eta - p^2\Phi\gamma\eta + 2p\gamma\eta - 2\gamma\eta - p^2\Phi^2\eta + p\Phi\eta) \quad (4.50)$$

$$+ 3p^2\Phi\beta\gamma - 3p\Phi\beta\gamma - 2p\beta\gamma + 2\beta\gamma) \quad (4.51)$$

$$a_2 = -N^2(\Phi-1)^2\beta(p\Phi\gamma\eta^2 - p\gamma\eta^2 + \gamma\eta^2 - p\Phi\eta^2 + 2p^2\Phi^2\beta\gamma\eta - 2p^2\Phi\beta\gamma\eta) \quad (4.52)$$

$$+ p\Phi\beta\gamma\eta + 2p\beta\gamma\eta - 2\beta\gamma\eta - 2p^2\Phi^2\beta\eta + 2p\Phi\beta\eta + 3p^2\Phi\beta^2\gamma - 3p\Phi\beta^2\gamma) \quad (4.53)$$

$$- p\beta^2\gamma + \beta^2\gamma) \quad (4.54)$$

$$a_1 = N^3p(\Phi-1)^3\Phi\beta(\eta^2 - p\Phi\beta\gamma\eta + p\beta\gamma\eta - \beta\gamma\eta + p\Phi\beta\eta - \beta\eta - p\beta^2\gamma) \quad (4.55)$$

$$+ \beta^2\gamma) \quad (4.56)$$

X can be factored out of this polynomial, indicating $X = 0$ is always an equilibrium (as expected). The remaining roots are given by solutions to $a_4 X^3 + a_3 X^2 + a_2 X + a_1 = 0$. This polynomial was solved numerically using the *polyroot* function in *R*. Each was substituted into the following expression, which comes from the *wt* only pair of ODEs (Eqns 4.29 and 4.30), to find the corresponding roots for W .

$$W = \frac{\eta X(1-\phi)N - \beta((1-\phi)N - X)(1-p\phi)X}{\beta((1-\phi)N - X)p(1-\phi)} \quad (4.57)$$

For all parameter combinations shown in the results section, and those from the testing process described in section 4.2.2, there is only one real and positive root (or there are repeated roots) for X and W .

4.6 Appendix 2 - Stability analysis for the model in which only resistant crop is planted

The two-patch model when the cropping ratio $\phi = 1$ (so there is only resistant crop planted) is given by:

$$\frac{dW}{dt} = \beta\gamma \left(1 - \frac{W+Z}{N}\right) W - \eta W, \quad (4.58)$$

$$\frac{dZ}{dt} = \beta(1-\delta) \left(1 - \frac{W+Z}{N}\right) Z - \eta Z. \quad (4.59)$$

where $W = I_R^{wt}$ and $Z = I_R^{rb}$.

At equilibrium, Eqn (4.58) indicates either

$$W = 0, \quad (4.60)$$

or

$$\beta\gamma \left(1 - \frac{W+Z}{N}\right) - \eta = 0. \quad (4.61)$$

Similarly, Eqn (4.59) implies either

$$Z = 0, \quad (4.62)$$

or

$$\beta(1-\delta) \left(1 - \frac{W+Z}{N}\right) - \eta = 0. \quad (4.63)$$

At any equilibrium either Eqn (4.60) or Eqn (4.61), as well as either Eqn (4.62) or Eqn (4.63), must simultaneously be satisfied.

For coexistence of the two strains at non-zero density:

$$\left(1 - \frac{W+Z}{N}\right) = \frac{\eta}{\beta\gamma} = \frac{\eta}{\beta(1-\delta)}. \quad (4.64)$$

where the first part of the equality comes from Eqn (4.61) and the second part from Eqn (4.63). Coexistence is therefore only possible if the parameters are such that $\gamma = (1-\delta)$. We handle this degenerate case separately in the below, first concentrating on the more general case in which $\gamma \neq (1-\delta)$ and so coexistence is not possible.

4.6.1 Case in which coexistence is not possible (i.e. $\gamma \neq (1 - \delta)$)

The individual R_0 values for each pathogen strain are given by:

$$R_0^W = \frac{\beta\gamma}{\eta}, \quad (4.65)$$

$$R_0^Z = \frac{\beta(1-\delta)}{\eta}. \quad (4.66)$$

Since $\gamma \neq (1 - \delta)$ these two R_0 values are different, because each of the two strains has a different fitness on the R host. As demonstrated above, there is then no coexistence equilibrium.

The three equilibria of the system are:

$$(W, Z) = (0, 0), \quad (4.67)$$

$$= (W^*, 0), \quad (4.68)$$

$$= (0, Z^*), \quad (4.69)$$

where

$$W^* = N \left(1 - \frac{1}{R_0^W} \right), \quad (4.70)$$

$$Z^* = N \left(1 - \frac{1}{R_0^Z} \right). \quad (4.71)$$

The Jacobian for Eqns (4.58) and (4.59) is given by:

$$J = \begin{pmatrix} \beta\gamma \left(1 - \frac{2W}{N} - \frac{Z}{N} \right) - \eta & -\frac{\beta\gamma W}{N} \\ -\frac{\beta(1-\delta)Z}{N} & \beta(1-\delta) \left(1 - \frac{W}{N} - \frac{2Z}{N} \right) - \eta \end{pmatrix}. \quad (4.72)$$

We use this to analyse the stability of each of the three equilibria in Eqns (4.67)-(4.69) in turn.

Trivial equilibrium $(W, Z) = (0, 0)$

Here the Jacobian reduces to

$$J = \begin{pmatrix} \beta\gamma - \eta & 0 \\ 0 & \beta(1-\delta) - \eta \end{pmatrix}, \quad (4.73)$$

so the eigenvalues are:

$$\lambda_1 = \beta\gamma - \eta = \eta \left(\frac{\beta\gamma}{\eta} - 1 \right) = \eta (R_0^W - 1), \quad (4.74)$$

$$\lambda_2 = \beta(1 - \delta) - \eta = \eta \left(\frac{\beta(1 - \delta)}{\eta} - 1 \right) = \eta (R_0^Z - 1). \quad (4.75)$$

The system will therefore have at least one positive eigenvalue at the trivial equilibrium, and so be unstable, if either $R_0^W > 1$ or $R_0^Z > 1$.

Equilibrium with only W present, $(W, Z) = (W^*, 0)$

First we note from Eqn (4.70) that a pre-condition for this equilibrium to be biologically-meaningful, with $W^* > 0$, is that

$$\frac{1}{R_0^W} < 1 \quad \text{i.e.} \quad R_0^W > 1. \quad (4.76)$$

When $(W, Z) = (W^*, 0)$, considering that

$$\beta\gamma \left(1 - \frac{2W^*}{N} \right) - \eta = \beta\gamma \left(1 - \frac{W^*}{N} \right) - \eta - \frac{\beta\gamma W^*}{N} = -\frac{\beta\gamma W^*}{N}, \quad (4.77)$$

(using Eqn (4.61)) and also that:

$$\left(1 - \frac{W^*}{N} \right) = \left(1 - \frac{N}{N} \left(1 - \frac{1}{R_0^W} \right) \right) = \frac{1}{R_0^W}, \quad (4.78)$$

(using Eqn (4.70)), the Jacobian becomes

$$J = \begin{pmatrix} -\frac{\beta\gamma W^*}{N} & -\frac{\beta\gamma W^*}{N} \\ 0 & \beta(1 - \delta) \frac{1}{R_0^W} - \eta \end{pmatrix}. \quad (4.79)$$

So the eigenvalues are:

$$\lambda_1 = -\frac{\beta\gamma W^*}{N} < 0 \quad (\text{since } W^* > 0), \quad (4.80)$$

$$\lambda_2 = \eta \left(\frac{\beta(1 - \delta)}{\eta} \frac{1}{R_0^W} - 1 \right), \quad (4.81)$$

$$= \eta \left(\frac{R_0^Z}{R_0^W} - 1 \right). \quad (4.82)$$

This means that the system will be stable at this equilibrium whenever

$$\frac{R_0^Z}{R_0^W} - 1 < 0 \text{ i.e. } R_0^Z < R_0^W. \quad (4.83)$$

The condition for the *wt* strain equilibrium being stable to invasion by the *rb* strain is therefore simply that the reproductive number for the *wt* strain is larger than that for the *rb* strain.

Equilibrium with only *Z* present, $(W, Z) = (0, Z^*)$

A virtually identical argument indicates that the equilibrium in which only *Z* is present is biologically-meaningful only if $R_0^Z > 1$ and is stable to invasion by the *wt* strain whenever

$$R_0^W < R_0^Z. \quad (4.84)$$

Summary of the analysis of the model when coexistence is not possible

The results of the analysis are summarised in Table 4.3.

Table 4.3 Summary of the long-term behaviour of the model in which coexistence is not possible ($\gamma \neq (1 - \delta)$).

$R_0^W > 1$	$R_0^Z > 1$	$R_0^W > R_0^Z$	Pathogen species present at long-term at equilibrium
x	x	irrelevant	neither <i>wt</i> nor <i>rb</i>
✓	x	must be true	<i>wt</i> only
x	✓	can never be true	<i>rb</i> only
✓	✓	✓	<i>wt</i> only
✓	✓	x	<i>rb</i> only

4.6.2 Case in which coexistence is possible (i.e. $\gamma = (1 - \delta)$)

If we assume that the two strains have the same fitness on the *R* host, with

$$\gamma = (1 - \delta) = \zeta, \quad (4.85)$$

where ζ is introduced for notational convenience in the below, then the system is given by:

$$\frac{dW}{dt} = \beta\zeta \left(1 - \frac{W+Z}{N}\right) W - \eta W, \quad (4.86)$$

$$\frac{dZ}{dt} = \beta\zeta \left(1 - \frac{W+Z}{N}\right) Z - \eta Z. \quad (4.87)$$

If $I = W + Z$ then:

$$\frac{dI}{dt} = \zeta \left(1 - \frac{I}{N}\right) I - \eta I. \quad (4.88)$$

We introduce a new variable:

$$\epsilon = \frac{W}{W + Z}, \quad (4.89)$$

which is the proportion of infections caused by the *wt* pathogen strain. We can then use the quotient rule to show how this proportion ϵ changes over time Eqns (4.86) and (4.87):

$$\frac{d\epsilon}{dt} = \frac{(W + Z)\frac{dW}{dt} - \left(\frac{dW}{dt} + \frac{dZ}{dt}\right)W}{(W + Z)^2}, \quad (4.90)$$

$$= \frac{Z\frac{dW}{dt} - W\frac{dZ}{dt}}{(W + Z)^2}. \quad (4.91)$$

Given that:

$$Z\frac{dW}{dt} = W\frac{dZ}{dt} = \beta\zeta \left(1 - \frac{W + Z}{N}\right) WZ - \eta WZ, \quad (4.92)$$

this means that:

$$\frac{d\epsilon}{dt} = 0. \quad (4.93)$$

Therefore the proportion of infections caused by the *wt* and *rb* pathogen strains remains constant over time whenever $\gamma = (1 - \delta) = \zeta$.

This means that – because the number of infections of each strain simply increases logistically at the same rate – although the two strains will coexist under these conditions, neither strain is able to increase in density relative to the other, with the relative population sizes of the two strains entirely dependent on their initial conditions. Whether or not each pathogen population tends to a non-zero value depends on whether $R_0^W = R_0^R \leq 1$.

Chapter 5

General Discussion

The combination of studies in this thesis have together led to an advancement in our understanding of how spatial heterogeneity affects the durability of crop disease resistance. In chapter 2 we built a spatially explicit model that highlighted the key effects of spatial heterogeneity on durability at a given cropping ratio, which for simplicity was set at 50:50, while also helping to identify the essential drivers of these effects. Our results showed that mixing susceptible and resistant crops at small scales of spatial heterogeneity could improve resistance durability, but that the strength of this effect depended strongly on the parameters governing the specific pathosystem. The frequently complex effects of these parameters could best be explained as being driven by the frequencies of interaction between the various host and pathogen genotypes.

This knowledge contributed towards the construction of a novel, much simpler, spatially implicit model that can be used to represent spatial heterogeneity in a generic epidemiological setting. By this we mean that it is not necessarily specific to plant disease systems. This second model can be used to estimate the scale of spatial heterogeneity from spatially explicit data, is amenable to some mathematical analysis, and crucially allows the cropping ratio to be varied independently of the parameter that controls the scale of spatial heterogeneity in the system. The testing of this model was aided by the use of a novel method, based on a cellular automaton, for the stochastic generation of spatial landscape patterns with varying scales of heterogeneity. We then applied the new model to the study of resistance durability, allowing us to confirm, clarify and extend the results of the spatially explicit model, while adding insight into the previously unexplored role of the cropping ratio. Our work has focussed on a relatively simple gene-for-gene interaction, with two host varieties and two pathogen strains, but has the potential to be applied to more complex systems. All of the models used in thesis feature

both susceptible and resistant crop hosts, together with wild-type and resistance breaking pathogen strains. The resistance breaking strain can infect both host varieties equally, but potentially carries a fitness cost, while the wild-type strain is less able to infect the resistant host, depending on the efficacy of the resistance gene.

5.1 Take home messages on resistance durability

The somewhat linear nature of the research programme and thesis - with various threads from earlier chapters unified in chapter 4 - means that the overall take home messages are essentially the same as those which are described in depth in the discussion to chapter 4. The general pattern for these messages is that the optimal spatial deployment strategy for the maximisation of resistance durability can depend strongly on the genetic fitness related parameters that govern the host-pathogen interaction, while also depending on the timeframe across which the benefit of control is measured. The scale of pathogen dispersal also plays a key role as it changes the effective scale of spatial heterogeneity experienced by the pathogen population at any given landscape topology. More specifically, if there is no restriction of the availability and/or cost to growers of the resistant crop, and if the cropping ratio is to remain constant over time, then our results indicate that it is the optimal strategy to prioritise the control of the resident wild-type population. This is achieved by using a combination of a high resistance cropping ratio and high levels of spatial variety mixing, with the goal of driving the wild-type rapidly to extinction. If greater variety mixing at smaller scales of spatial heterogeneity carries any kind of financial cost or operational difficulty then the optimal strategy is simply to plant all resistant crop as this renders the degree of mixing irrelevant.

If alternatively, the availability of resistant plant material is limited, perhaps due to financial cost, cultural practices or the slow propagation of a new variety, there may be an intermediate optimal degree of variety mixing that is sufficient to prevent the spread of a resistance breaking strain, but is not any higher. With a restricted cropping ratio, the strategy of mixing varieties at smaller scales of spatial heterogeneity to maximise resistance durability has variable degrees of effectiveness, depending on the pathogen dispersal scale, the genetic properties of the host-pathogen interaction, and the length of the time scale of interest. This effectiveness is often maximised at intermediate cropping ratios and fitness costs, and over intermediate time scale lengths, where there is a degree of balance between the competitiveness of the two strains

in the landscape. In these cases, the change to using greater spatial mixing of varieties can have a critical effect on the ability of one pathogen strain to invade or persist in the landscape. Effectiveness also depends on the scale of pathogen dispersal being somewhat similar to the scale of spatial heterogeneity in the landscape. If the availability of resistance is unlimited, then optimal intermediate cropping ratios, which have been suggested by previous studies such as Fabre et al. (2012) and Papaïx et al. (2018), are only present if the importance of controlling the initial resident wild-type epidemic is reduced. The difference between our results and these previously published studies could be due to variations in the consideration of spatial structure, fitness costs or seasonality. Our results potentially suggest that it may be optimal to change the cropping ratio over time, by planting all resistant crop at first, but increasing the proportion of the susceptible crop as the resistance breaking strain spreads. Further testing would be required to confirm this effect however. The intermediate optimal cropping ratio, that results from reducing the importance of controlling the initial wild-type epidemic, increases if the degree of spatial variety mixing is reduced.

5.2 Methodological developments

Aside from the greater insight we now have into the durability of crop disease resistance, our work has also led to methodological developments in terms of how we represent the scale of spatial heterogeneity. This new method allows us to model the complex effects of spatial heterogeneity, using a simple and intuitive parameter that can be reliably estimated from spatially explicit data. This parameter alters the frequency with which the different host and pathogen genotypes interact with one another, so as to represent the relative spatial proximity of the different host varieties to one other as conditioned by pathogen dispersal. It was these frequencies of interaction that were reasoned to be the main drivers behind the effects of spatial heterogeneity in our initial spatially explicit model. The simplicity of the new spatially implicit model allows mathematical invasion analysis, using a combination of symbolic and numeric approaches. Furthermore, the design of the model ensures that the relative proportions of the two varieties of host can be altered without affecting the given scale of spatial heterogeneity. The new model differs from other existing methods of capturing spatial structure using simple expressions, such as non-linear incidence functions (Liu et al., 1986; Roy and Pascual, 2006), in that it is tailored to the representation of spatial heterogeneity between different varieties

of host organism. Compared with more complex and difficult to implement methods such as spatial moment equations (Bolker and Pacala, 1997; Morozov and Poggiale, 2012), our approach has an advantage in that it captures the fundamental drivers influencing pathogen strain dynamics without including excessively detailed and less relevant spatial information. This information would include details such as the dynamic spatial interactions between infected and uninfected host individuals over time.

The model is currently focussed on representing the spread of disease within different varieties of host crop, however the generic nature of the aggregation parameter means that application in a variety of other contexts is possible. These could include wild plant pathogen epidemics, or even the dynamics of other organism populations within spatially heterogeneous environments (Aars and Ims, 2000; Wu, 2009). Further work could seek to identify other ecological and evolutionary systems where the scale of spatial heterogeneity in patchy habitats can be used to infer complex population dynamics.

5.2.1 Assumptions and limitations

Any model that seeks to reduce complex spatial information into a simpler model is inevitably based upon a series of assumptions and carries with it certain limitations. A simplification inherent in our spatially implicit model from chapters 3 and 4 is that infections caused by the different pathogen strains are assumed to be homogeneously distributed within the host individuals of a given crop variety. For the study of resistance durability this assumption applies to both the within and between field scales. The spatial information that we do capture is about the heterogeneity between the different host varieties, and the pathogen infection dynamics are then inferred from this. It might however be expected in reality that the two pathogen strains would experience different effective degrees of aggregation in the host populations due to the spatial location of the host individuals they more frequently infect. These effective degrees of aggregation, which would be due to the different strains being more common in either the more or less aggregated regions in the landscape, may also change over time as the epidemic progresses. Future work may develop the existing model to account for this by scaling the aggregation parameter differently for each of the two strains, and also so that it depends on the overall level of infection. This would be done alongside the direct testing of the spatially implicit model against an equivalent spatially explicit model, so as to determine how closely the results can be scaled to match. We have already demonstrated in chapter 4 the qualitative

similarity of results from spatially implicit and explicit models, however quantitative similarity would require further testing and potentially model development.

There are also further assumptions that, while not affecting the fundamental dynamics being studied, may limit the ability of the model in its current form to accurately replicate the effects of heterogeneity in real spatial data. One of these is that inoculum is lost, due to it dispersing and landing in the spaces between plants or over the edge of the measured landscape, in a manner that is modelled as being relative to the proportions of the two host varieties. For a specific landscape this might not be the case however, simply due to the uneven distribution of different plants or fields. If this model was used in the future to estimate the scale of spatial heterogeneity from real landscape data, modifications could likely be made to account for any disproportionate dispersal into non-host space. Strictly speaking our model also depends on the scale of the landscape being very large relative to the size of locally aggregated patches of a given host variety. This is so that the effective global cropping ratio, that inoculum disperses proportionally to outside of a locally aggregated patch, is unaffected by the specific size of that patch. We have shown in chapter 3 that our model is likely to be relatively robust to the breaking of these assumptions, however achieving maximum accuracy may require the model to be altered so that it accounts for the specific finite size of the landscape being represented.

5.3 Applications and future work

The principal value of our model system at present is its exploration of the fundamental dynamics which determine the effects of spatial heterogeneity on crop disease resistance durability in a wide range of potential epidemiological contexts. In future however, with some of the modifications mentioned above, the model could be used to identify optimal strategies for control in specific pathosystems that feature gene-for-gene interactions. This would require obtaining estimates for the parameter values that determine model behaviour, such as the infection rate, the fitness cost of the resistance breaking trait, and the efficacy of the resistance gene. The infection rate in our model is the product of the rate of inoculum production by an infected host tissue unit and the probability that infection occurs once inoculum has landed on uninfected host tissue. It may be possible to obtain experimental estimates for these parameters from published studies, however there is often an inherent difficulty in applying data from a laboratory setting to a complex real world epidemic (Cunniffe and Gilligan, in press). A more

reliable approach would be to estimate the infection rate from disease progress data taken from a field study. Fitness costs of resistance breaking traits can be estimated by comparing infection in fully susceptible hosts caused by strains with and without the trait in question. Various pathogen life history traits, such as spore density and lesion growth rate, can be recorded and used to calculate the basic reproductive number for each strain (Montarry et al., 2010). These can then be compared and used to evaluate fitness costs. Mixed strain epidemics can also be used to measure the change in the frequencies of the two strains over time (Huang et al., 2006). Resistance efficacies are commonly measured by methods such as comparing the area under the disease progress curve for different epidemics, although again to infer the value of a parameter in our model, some model fitting would be required (Griffey et al., 1993; Jeger and Viljanen-Rollinson, 2001).

Currently there is a lack of available field or experimental data that could be used to parameterise the models presented in this study, and validate the dynamics which we have described. We would require disease progress data, separated by host cultivar/pathogen strain combination, for epidemics in host landscapes with a range of cropping ratios and scales of spatial heterogeneity. For larger pathogen dispersal scales this data could be obtained by conducting field trials, however this may be impractical due to the number and scale of trials that would be required. For this reason it might be more practical to conduct greenhouse experiments on a plant disease system with a pathogen that disperses at smaller spatial scales. The usefulness of such data for validation purposes would be increased if a number of different resistance traits, with different associated pathogen fitness costs, could be tested.

Additional questions that could be addressed in future work include how well our simple model fits to more complex patterns of spatial disease spread. In all of our current work we have assumed that both the wild-type and the resistance breaking strain are initially present, the latter at the same low frequency in all host locations, and we are largely concerned with changes in infection frequency from this initial state. This is intended to reflect an endemic disease scenario where the resistance breaking strain is already present at a background frequency before the introduction of positive selection. An alternative scenario however might involve a novel strain arriving through dispersal at a particular location in the landscape, where it then potentially spreads out from. How the scale of spatial heterogeneity in the host landscape affects the ability of a resident wild-type pathogen population to 'resist' the invasion of this new strain in this context would be an interesting addition to the results presented here. Adapting

our models, particularly the two-patch models presented in chapters 3 and 4, to represent this significantly more complex scenario of spatial spread is not trivial however, and warrants its own research project. The inclusion of demographic stochasticity in the pathogen population and the testing of fatter tailed dispersal kernels may also be valuable additions to a study involving more complex spatial spread dynamics. This is due to the fact that rare long distance dispersal events, that can be better represented by fatter tailed kernels, are likely to play a more important role in a scenario where a pathogen must spread out from a particular location.

Stochasticity may also play an important role if we were to consider the initial emergence of the resistance breaking strain through random mutation events in the wild-type population. The time until this initial emergence occurs is often thought of as another measure of resistance durability (van den Bosch and Gilligan, 2003), and has been previously shown to respond differently to spatial heterogeneity compared to the longer term measures that we have focussed on here (Papaix et al., 2018). If the emergence of the resistance breaking strain occurred through a single dispersal or mutational event, then the initial frequency of that strain would be extremely low. In this context it might be important to study the effect of a latent period on the initial rate of exponential growth (Cunniffe et al., 2012; Madden et al., 2007). We could also test the current deterministic and endemic model with a lower initial background frequency of the resistance breaking strain. In order to ensure that this strain still spreads within a reasonable timeframe, despite the lower initial frequency, the model may need to include a separate pathogen death rate term combined with a higher infection rate. This would enable rapid spread of a pathogen strain from a low initial frequency. Including features such as a lower initial resistance-breaking frequency or variable latent period would provide additional insight on the role of spatial heterogeneity in maintaining resistance durability. We however consider it unlikely that these additions would significantly alter the core fundamental dynamics we have described in this thesis, hence why they have not been included in the current study.

The foundation provided by this thesis could be used as a platform to study the effects of spatial heterogeneity on more complex disease resistance traits. In particular our two-patch model could potentially be modified to incorporate a polygenic quantitative resistance that decreases in efficacy as one pathogen strain increases its virulence through mutational events (Lo Iacono et al., 2012; Nelson, 1978; Young, 1996). The inclusion of multiple pyramided resistance traits (Djidjou-Demasse et al., 2017; Kiyosawa, 1982), or sexual recombination in the pathogen population (Brown, 1995), are also promising avenues for future development. It may

also be interesting to consider how spatial heterogeneity interacts with temporal heterogeneity through crop rotation (Fabre et al., 2015), or whether simultaneous fungicide application and resistance management leads to any change in optimal strategy (Carolan et al., 2017). Preliminary work we have undertaken on the dual roles of spatial and temporal heterogeneity has indicated that while both strategies can be useful in suppressing disease over evolutionary timescales, their combined effects are less than additive. This is likely due to the increase in cultivar and pathogen strain heterogeneity caused by one strategy resulting in the other strategy having a less dramatic effect than it would have done had it been acting on a non-heterogeneous baseline scenario. Finally, dynamic disease management strategies that change over time, as suggested by our results on optimal cropping ratios and scales of spatial heterogeneity, could also be investigated through the use of optimal control theory (Forster and Gilligan, 2007; Sethi and Staats, 1978).

5.4 Concluding remarks

In this thesis, we have demonstrated how complex spatial phenomena can be studied by building simple models based on an in depth knowledge of the fundamental dynamics involved. The initial development of more complex models can aid in the early study of these dynamics, provided that effort is invested to understand the pattern of results produced from a mechanistic perspective. Our approach contrasts with the frequently used one of building an initial simple model, before adding complexity and realism at a later stage. Simple mathematical models may ignore many aspects of biological realism, that might be required for the application of predictive models in a practical capacity. They do however allow for the isolation and study of fundamental underlying ecological and evolutionary processes that might otherwise be obscured. Simple model also have the advantage of having fewer freely varying parameters, meaning that dynamics can be investigated in a wide variety of potential epidemiological contexts. Our work has advanced the study of spatial heterogeneity and its effects on resistance durability in plant disease systems. In particular, we highlight the importance of obtaining accurate parameter estimates for the pathosystem in question, since a one size fits all approach is unlikely to be successful.

References

- Aars, J. and Ims, R. A. (2000). Population dynamic and genetic consequences of spatial density-dependent dispersal in patchy populations. *The American Naturalist*, 155(2):252–265.
- Andrivo, D., Pilet, F., Montarry, J., Hafidi, M., Corbiere, R., Achbani, E. H., Pelle, R., and Ellisseche, D. (2007). Adaptation of *Phytophthora infestans* to partial resistance in potato: Evidence from French and Moroccan populations. *Phytopathology*, 97(3):338–343.
- Asea, G., Vivek, B. S., Bigirwa, G., Lipps, P. E., and Pratt, R. C. (2009). Validation of consensus quantitative trait loci associated with resistance to multiple foliar pathogens of maize. *Phytopathology*, 99(5):540–547.
- Atlin, G. N., Cairns, J. E., and Das, B. (2017). Rapid breeding and varietal replacement are critical to adaptation of cropping systems in the developing world to climate change. *Global Food Security*, 12:31–37.
- Bai, J., Choi, S.-H., Ponciano, G., Leung, H., and Leach, J. E. (2000). *Xanthomonas oryzae* pv. *oryzae* avirulence genes contribute differently and specifically to pathogen aggressiveness. *Molecular Plant-Microbe Interactions*, 13(12):1322–1329.
- Bergelson, J. and Purrington, C. B. (1996). Surveying patterns in the cost of resistance in plants. *American Naturalist*, 148(3):536–558.
- Bolker, B. and Pacala, S. W. (1997). Using moment equations to understand stochastically driven spatial pattern formation in ecological systems. *Theoretical Population Biology*, 52(3):179–197.
- Bousset, L., Sprague, S. J., Thrall, P. H., and Barrett, L. G. (2018). Spatio-temporal connectivity and host resistance influence evolutionary and epidemiological dynamics of the canola pathogen *Leptosphaeria maculans*. *Evolutionary Applications*, 11(8):1354–1370.

- Brown, J. K. (2015). Durable resistance of crops to disease: a darwinian perspective. *Annual Review of Phytopathology*, 53:513–539.
- Brown, J. K. and Tellier, A. (2011). Plant-parasite coevolution: bridging the gap between genetics and ecology. *Annual Review of Phytopathology*, 49:345–367.
- Brown, J. K. M. (1995). Recombination and selection in populations of plant-pathogens. *Plant Pathology*, 44(2):279–293.
- Brown, J. K. M. and Rant, J. C. (2013). Fitness costs and trade-offs of disease resistance and their consequences for breeding arable crops. *Plant Pathology*, 62:83–95.
- Brun, H., Chevre, A.-M., Fitt, B. D. L., Powers, S., Besnard, A.-L., Ermel, M., Huteau, V., Marquer, B., Eber, F., Renard, M., and Andrivon, D. (2010). Quantitative resistance increases the durability of qualitative resistance to *Leptosphaeria maculans* in *Brassica napus*. *New Phytologist*, 185(1):285–299.
- Bruns, E., Carson, M. L., and May, G. (2014). The jack of all trades is master of none: a pathogen's ability to infect a greater number of host genotypes comes at a cost of delayed reproduction. *Evolution*, 68(9):2453–2466.
- Burdon, J. J. and Laine, A.-L. (2019). *Evolutionary dynamics of plant pathogen interactions*. Cambridge University Press, Cambridge, UK.
- Burdon, J. J. and Thrall, P. H. (2008). Pathogen evolution across the agro-ecological interface: implications for disease management. *Evolutionary Applications*, 1(1):57–65.
- Burdon, J. J., Zhan, J., Barrett, L. G., Papaïx, J., and Thrall, P. H. (2016). Addressing the challenges of pathogen evolution on the world's arable crops. *Phytopathology*, 106(10):1117–1127.
- Bussell, E. H., Dangerfield, C. E., Gilligan, C. A., and Cunniffe, N. J. (2019). Applying optimal control theory to complex epidemiological models to inform real-world disease management. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 374(1776):20180284.
- Carlsson-Granér, U. and Thrall, P. H. (2002). The spatial distribution of plant populations, disease dynamics and evolution of resistance. *Oikos*, 97(1):97–110.

- Carolan, K., Helps, J., van den Berg, F., Bain, R., Paveley, N., and van den Bosch, F. (2017). Extending the durability of cultivar resistance by limiting epidemic growth rates. *Proceedings of the Royal Society B: Biological Sciences*, 284(1863):20170828.
- Carrasco, P., de la Iglesia, F., and Elena, S. F. (2007). Distribution of fitness and virulence effects caused by single-nucleotide substitutions in tobacco etch virus. *Journal of Virology*, 81(23):12979–12984.
- Chin, K. and Wolfe, M. (1984). Selection on *Erysiphe graminis* in pure and mixed stands of barley. *Plant Pathology*, 33(4):535–546.
- Collmer, A. (1998). Determinants of pathogenicity and avirulence in plant pathogenic bacteria. *Current Opinion in Plant Biology*, 1(4):329–335.
- Cui, H., Tsuda, K., and Parker, J. E. (2015). Effector-triggered immunity: from pathogen perception to robust defense. *Annual Review of Plant Biology*, 66:487–511.
- Cunniffe, N., Stutt, R., van den Bosch, F., and Gilligan, C. (2012). Time-dependent infectivity and flexible latent and infectious periods in compartmental models of plant disease. *Phytopathology*, 102(4):365–380.
- Cunniffe, N. J. and Gilligan, C. A. (in press). Use of mathematical models to predict epidemics and to optimise disease detection and management. In Ristaino, J., editor, *Emerging plant diseases and global food security*. American Phytopathological Society (APS Press), St. Paul, USA.
- Cunniffe, N. J., Stutt, R. O., DeSimone, R. E., Gottwald, T. R., and Gilligan, C. A. (2015). Optimising and communicating options for the control of invasive plant disease when there is epidemiological uncertainty. *PLoS Computational Biology*, 11(4):e1004211.
- Deslandes, L. and Rivas, S. (2012). Catch me if you can: bacterial effectors and plant targets. *Trends in Plant Science*, 17(11):644–655.
- Djidjou-Demasse, R., Moury, B., and Fabre, F. (2017). Mosaics often outperform pyramids: insights from a model comparing strategies for the deployment of plant resistance genes against viruses in agricultural landscapes. *New Phytologist*, 216(1):239–253.

- Fabre, F., Rousseau, E., Mailleret, L., and Moury, B. (2012). Durable strategies to deploy plant resistance in agricultural landscapes. *New Phytologist*, 193(4):1064–1075.
- Fabre, F., Rousseau, E., Mailleret, L., and Moury, B. (2015). Epidemiological and evolutionary management of plant resistance: optimizing the deployment of cultivar mixtures in time and space in agricultural landscapes. *Evolutionary Applications*, 8(10):919–932.
- Filipe, J. and Gibson, G. (1998). Studying and approximating spatio-temporal models for epidemic spread and control. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 353(1378):2153–2162.
- Flor, H. H. (1971). Current status of the gene-for-gene concept. *Annual Review of Phytopathology*, 9:275–296.
- Forster, G. A. and Gilligan, C. A. (2007). Optimizing the control of disease infestations at the landscape scale. *Proceedings of the National Academy of Sciences*, 104(12):4984–4989.
- Fraile, A., Pagan, I., Anastasio, G., Saez, E., and Garcia-Arenal, F. (2011). Rapid genetic diversification and high fitness penalties associated with pathogenicity evolution in a plant virus. *Molecular Biology and Evolution*, 28(4):1425–1437.
- Freckleton, R. and Watkinson, A. (2002). Large-scale spatial dynamics of plants: metapopulations, regional ensembles and patchy populations. *Journal of Ecology*, 90(3):419–434.
- Galvani, A. P. (2003). Epidemiology meets evolutionary ecology. *Trends in Ecology & Evolution*, 18(3):132–139.
- Gandon, S. (2002). Local adaptation and the geometry of host–parasite coevolution. *Ecology Letters*, 5(2):246–256.
- Garcia-Arenal, F. and McDonald, B. A. (2003). An analysis of the durability of resistance to plant viruses. *Phytopathology*, 93(8):941–952.
- Gassmann, W., Dahlbeck, D., Chesnokova, O., Minsavage, G. V., Jones, J. B., and Staskawicz, B. J. (2000). Molecular evolution of virulence in natural field strains of *Xanthomonas campestris* pv. *vesicatoria*. *Journal of Bacteriology*, 182(24):7053–7059.
- Gilligan, C. A. and van den Bosch, F. (2008). Epidemiological models for invasion and persistence of pathogens. *Annual Review of Phytopathology*, 46:385–418.

- Griffey, C., Das, M., and Stromberg, E. (1993). Effectiveness of adult-plant resistance in reducing grain yield loss to powdery mildew in winter wheat. *Plant Disease*, 77(6):618–622.
- Grosdidier, M., Iosif, R., Husson, C., Cael, O., Scordia, T., and Marçais, B. (2018). Tracking the invasion: dispersal of *Hymenoscyphus fraxineus* airborne inoculum at different scales. *FEMS Microbiology Ecology*, 94(5):fiy049.
- Groth, J. (1976). Multilines and “super races”: a simple model. *Phytopathology*, 66(937):9.
- Gubbins, S. and Gilligan, C. A. (1997). A test of heterogeneous mixing as a mechanism for ecological persistence in a disturbed environment. *Proceedings of the Royal Society B: Biological Sciences*, 264(1379):227–232.
- Gubbins, S. and Gilligan, C. A. (1999). Invasion thresholds for fungicide resistance: deterministic and stochastic analyses. *Proceedings of the Royal Society B: Biological Sciences*, 266(1437):2539–2549.
- Gudelj, I., van den Bosch, F., and Gilligan, C. A. (2004). Transmission rates and adaptive evolution of pathogens in sympatric heterogeneous plant populations. *Proceedings of the Royal Society B: Biological Sciences*, 271(1553):2187–2194.
- Halterman, D. A., Wei, F. S., and Wise, R. P. (2003). Powdery mildew-induced *Mla* mRNAs are alternatively spliced and contain multiple upstream open reading frames. *Plant Physiology*, 131(2):558–567.
- Harrison, B. D. (2002). Virus variation in relation to resistance-breaking in plants. *Euphytica*, 124(2):181–192.
- Hassell, M. P. and May, R. M. (1973). Stability in insect host-parasite models. *The Journal of Animal Ecology*, pages 693–726.
- Hepperly, P., Kirkpatrick, B., Sinclair, J., et al. (1980). *Abutilon theophrasti*: wild host for three fungal parasites of soybean. *Phytopathology*, 70(4):307–310.
- Hocking, L. M. (1991). *Optimal control: an introduction to the theory with applications*. Oxford University Press, Oxford, UK.
- Holmes, E. E., Lewis, M. A., Banks, J., and Veit, R. (1994). Partial differential equations in ecology: spatial interactions and population dynamics. *Ecology*, 75(1):17–29.

- Holt, R. D. and Dobson, A. P. (2006). Extending the principles of community ecology to address the epidemiology of host-pathogen systems. In Collinge, S. and Ray, C., editors, *Disease ecology: community structure and pathogen dynamics*, page 6. Oxford University Press Oxford, UK.
- Huang, Y., Evans, N., Fitt, B., Li, Z., Rouxel, T., and Balesdent, M. (2006). Fitness cost associated with loss of the avrM4 avirulence function in *Leptosphaeria maculans* (phoma stem canker of oilseed rape). In B.D.L., F., N., E., B.J., H., and B.M., C., editors, *Sustainable strategies for managing Brassica napus (oilseed rape) resistance to Leptosphaeria maculans (phoma stem canker)*, pages 77–89. Springer, Dordrecht, The Netherlands.
- Iwasa, Y., Andreasen, V., and Levin, S. (1987). Aggregation in model ecosystems. i. perfect aggregation. *Ecological Modelling*, 37(3-4):287–302.
- Jayakar, S. (1970). A mathematical model for interaction of gene frequencies in a parasite and its host. *Theoretical Population Biology*, 1(2):140–164.
- Jeger, M. and Viljanen-Rollinson, S. (2001). The use of the area under the disease-progress curve (AUDPC) to assess quantitative disease resistance in crop cultivars. *Theoretical and Applied Genetics*, 102(1):32–40.
- Jin, Y., Szabo, L. J., and Carson, M. (2010). Century-old mystery of puccinia striiformis life history solved with the identification of berberis as an alternate host. *Phytopathology*, 100(5):432–435.
- Johnson, R. (1979). Concept of durable resistance. *Phytopathology*, 69(3):198–199.
- Johnson, R. (1981). Durable resistance: definition of, genetic control, and attainment in plant breeding. *Phytopathology*, 71(6):567–568.
- Keeling, M. J. and Rohani, P. (2002). Estimating spatial coupling in epidemiological systems: a mechanistic approach. *Ecology Letters*, 5(1):20–29.
- Keeling, M. J. and Rohani, P. (2011). *Modeling infectious diseases in humans and animals*. Princeton University Press, Princeton.
- Kierstead, H. and Slobodkin, L. B. (1953). The size of water masses containing plankton blooms. *Journal of Marine Research*, 12(1):141–147.

- Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220(4598):671–680.
- Kiyosawa, S. (1982). Genetics and epidemiological modeling of breakdown of plant disease resistance. *Annual Review of Phytopathology*, 20(1):93–117.
- Knott, D. (1961). The inheritance of rust resistance. vi. the transfer of stem rust resistance from *Agropyron elongatum* to common wheat. *Canadian Journal of Plant Science*, 41(1):109–123.
- Kolmer, J. (1995). Selection of *Puccinia recondita f. sp. tritici* virulence phenotypes in three multilines of thatcher wheat lines near isogenic for leaf rust resistance genes. *Canadian Journal of Botany*, 73(7):1081–1088.
- Kolmer, J. (1996). Genetics of resistance to wheat leaf rust. *Annual Review of Phytopathology*, 34(1):435–455.
- Laine, A.-L. and Barrès, B. (2013). Epidemiological and evolutionary consequences of life-history trade-offs in pathogens. *Plant Pathology*, 62(S1):96–105.
- Lampert, W. (2005). Vertical distribution of zooplankton: density dependence and evidence for an ideal free distribution with costs. *BMC Biology*, 3(1):10.
- Lannou, C. and Mundt, C. C. (1996). Evolution of a pathogen population in host mixtures: Simple race-complex race competition. *Plant Pathology*, 45(3):440–453.
- Lanoiselet, V., Cottier, E., Ash, G., Hind-Lanoiselet, T., Murray, G. M., and Harper, J. (2005). Prevalence and survival, with emphasis on stubble burning, of *Rhizoctonia* spp., causal agents of sheath diseases of rice in australia. *Australasian Plant Pathology*, 34(2):135–142.
- Leach, J. E., Vera Cruz, C. M., Bai, J., and Leung, H. (2001). Pathogen fitness penalty as a predictor of durability of disease resistance genes. *Annual Review of Phytopathology*, 39(1):187–224.
- Leonard, K. (1969). Selection in heterogeneous populations of *Puccinia graminis f. sp. avenae*. *Phytopathology*.
- Leonard, K. (1977). Selection pressures and plant pathogens. *Annals of the New York Academy of Sciences*, 287(1):207–222.

- Leonard, K. and Czocho, R. (1980). Theory of genetic interactions among populations of plants and their pathogens. *Annual Review of Phytopathology*, 18(1):237–258.
- Levin, S. A. (1974). Dispersion and population interactions. *The American Naturalist*, 108(960):207–228.
- Liu, B., Zhang, S., Zhu, X., Yang, Q., Wu, S., Mei, M., Mauleon, R., Leach, J., Mew, T., and Leung, H. (2004). Candidate defense genes as predictors of quantitative blast resistance in rice. *Molecular Plant-Microbe Interactions*, 17(10):1146–1152.
- Liu, W., Levin, S. A., and Iwasa, Y. (1986). Influence of nonlinear incidence rates upon the behavior of SIRS epidemiological models. *Journal of Mathematical Biology*, 23(2):187–204.
- Lo Iacono, G., van den Bosch, F., and Gilligan, C. A. (2013). Durable resistance to crop pathogens: An epidemiological framework to predict risk under uncertainty. *PLoS Computational Biology*, 9(1):e1002870.
- Lo Iacono, G., van den Bosch, F., and Paveley, N. (2012). The evolution of plant pathogens in response to host resistance: Factors affecting the gain from deployment of qualitative and quantitative resistance. *Journal of Theoretical Biology*, 304:152–163.
- Madden, L. V., Hughes, G., and van den Bosch, F. (2007). *The Study of Plant Disease Epidemics*. American Phytopathological Society (APS Press), St. Paul, USA.
- Mailleret, L. and Lemesle, V. (2009). A note on semi-discrete modelling in the life sciences. *Philosophical Transactions of the Royal Society A: Mathematical Physical and Engineering Sciences*, 367(1908):4779–4799.
- Margosian, M. L., Garrett, K. A., Hutchinson, J. S., and With, K. A. (2009). Connectivity of the american agricultural landscape: assessing the national risk of crop pest and disease spread. *BioScience*, 59(2):141–151.
- Matsuda, H., Ogita, N., Sasaki, A., and Satō, K. (1992). Statistical mechanics of population: the lattice lotka-volterra model. *Progress of Theoretical Physics*, 88(6):1035–1049.
- McDonald, B. A. and Linde, C. (2002). Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology*, 40(1):349–379.

- Mikaberidze, A., McDonald, B. A., and Bonhoeffer, S. (2015). Developing smarter host mixtures to control plant disease. *Plant Pathology*, 64(4):996–1004.
- Montarry, J., Hamelin, F. M., Glais, I., Corbière, R., and Andrivon, D. (2010). Fitness costs associated with unnecessary virulence factors and life history traits: evolutionary insights from the potato late blight pathogen *Phytophthora infestans*. *BMC Evolutionary Biology*, 10(1):283.
- Montúfar, R. and Ayala, M. (2019). Perceptions of agrobiodiversity and seed-saving practices in the northern andes of ecuador. *Journal of Ethnobiology and Ethnomedicine*, 15(1):35.
- Morozov, A. and Poggiale, J.-C. (2012). From spatially explicit ecological models to mean-field dynamics: the state of the art and perspectives. *Ecological Complexity*, 10:1–11.
- Mundt, C. et al. (1990). Probability of mutation to multiple virulence and durability of resistance gene pyramids. *Phytopathology*, 80(3):221–223.
- Mundt, C. C. (2002). Use of multiline cultivars and cultivar mixtures for disease management. *Annual Review of Phytopathology*, 40:381–410.
- Mundt, C. C. (2009). Importance of autoinfection to the epidemiology of polycyclic foliar disease. *Phytopathology*, 99(10):1116–1120.
- Mundt, C. C. (2014). Durable resistance: a key to sustainable management of pathogens and pests. *Infection, Genetics and Evolution*, 27:446–455.
- Mundt, C. C., Cowger, C., and Garrett, K. A. (2002). Relevance of integrated disease management to resistance durability. *Euphytica*, 124(2):245–252.
- Nachman, G. (2006). A functional response model of a predator population foraging in a patchy habitat. *Journal of Animal Ecology*, 75(4):948–958.
- Nelson, R., Wiesner-Hanks, T., Wisser, R., and Balint-Kurti, P. (2018). Navigating complexity to breed disease-resistant crops. *Nature Reviews Genetics*, 19(1):21.
- Nelson, R. R. (1978). Genetics of horizontal resistance to plant diseases. *Annual Review of Phytopathology*, 16:359–378.
- Oerke, E. (2006). Crop losses to pests. *The Journal of Agricultural Science*, 144(1):31–43.

- Oerke, E. and Dehne, H. (2004). Safeguarding production—losses in major crops and the role of crop protection. *Crop Protection*, 23(4):275–285.
- Oliver, R. (2016). Fungicide resistance management in practice; mixtures, alternations and cross resistance patterns. *Journal of Plant Pathology*, 98(4sup):11.
- Ovaskainen, O. and Cornell, S. J. (2006). Space and stochasticity in population dynamics. *Proceedings of the National Academy of Sciences*, 103(34):12781–12786.
- Papaïx, J., Adamczyk-Chauvat, K., Bouvier, A., Kiêu, K., Touzeau, S., Lannou, C., and Monod, H. (2014a). Pathogen population dynamics in agricultural landscapes: The Ddal modelling framework. *Infection, Genetics and Evolution*, 27:509–520.
- Papaïx, J., David, O., Lannou, C., and Monod, H. (2013). Dynamics of adaptation in spatially heterogeneous metapopulations. *PloS One*, 8(2):e54697.
- Papaïx, J., Goyeau, H., Du Cheyron, P., Monod, H., and Lannou, C. (2011). Influence of cultivated landscape composition on variety resistance: an assessment based on wheat leaf rust epidemics. *New Phytologist*, 191(4):1095–1107.
- Papaïx, J., Rimbaud, L., Burdon, J. J., Zhan, J., and Thrall, P. H. (2018). Differential impact of landscape-scale strategies for crop cultivar deployment on disease dynamics, resistance durability and long-term evolutionary control. *Evolutionary Applications*, 11(5):705–717.
- Papaïx, J., Touzeau, S., Monod, H., and Lannou, C. (2014b). Can epidemic control be achieved by altering landscape connectivity in agricultural systems? *Ecological Modelling*, 284:35–47.
- Parlevliet, J. and Van Ommeren, A. (1988). Accumulation of partial resistance in barley to barley leaf rust and powdery mildew through recurrent selection against susceptibility. *Euphytica*, 37(3):261–274.
- Parnell, S., Gottwald, T., Cunniffe, N., Alonso Chavez, V., and van den Bosch, F. (2015). Early detection surveillance for an emerging plant pathogen: a rule of thumb to predict prevalence at first discovery. *Proceedings of the Royal Society B: Biological Sciences*, 282(1814):20151478.
- Parnell, S., Gottwald, T., Gilligan, C., Cunniffe, N., and van den Bosch, F. (2010). The effect of landscape pattern on the optimal eradication zone of an invading epidemic. *Phytopathology*, 100(7):638–644.

- Parnell, S., Gottwald, T., van den Bosch, F., and Gilligan, C. (2009). Optimal strategies for the eradication of asiatic citrus canker in heterogeneous host landscapes. *Phytopathology*, 99(12):1370–1376.
- Parnell, S., van den Bosch, F., and Gilligan, C. A. (2006). Large-scale fungicide spray heterogeneity and the regional spread of resistant pathogen strains. *Phytopathology*, 96(5):549–555.
- Pascual, M., Roy, M., and Laneri, K. (2011). Simple models for complex systems: exploiting the relationship between local and global densities. *Theoretical Ecology*, 4(2):211–222.
- Pietravalle, S., Lemarie, S., and van den Bosch, F. (2006). Durability of resistance and cost of virulence. *European Journal of Plant Pathology*, 114(1):107–116.
- Plantegenest, M., Le May, C., and Fabre, F. (2007). Landscape epidemiology of plant diseases. *Journal of the Royal Society Interface*, 4(16):963–972.
- Poland, J. A., Balint-Kurti, P. J., Wisser, R. J., Pratt, R. C., and Nelson, R. J. (2009). Shades of gray: the world of quantitative disease resistance. *Trends in Plant Science*, 14(1):21–29.
- Qureshi, N., Kandiah, P., Gessese, M. K., Nsabiyera, V., Wells, V., Babu, P., Wong, D., Hayden, M., Bariana, H., and Bansal, U. (2018). Development of co-dominant kasp markers co-segregating with ug99 effective stem rust resistance gene sr26 in wheat. *Molecular Breeding*, 38(8):97.
- Raffaele, S., Farrer, R. A., Cano, L. M., Studholme, D. J., MacLean, D., Thines, M., Jiang, R. H. Y., Zody, M. C., Kunjeti, S. G., Donofrio, N. M., Meyers, B. C., Nusbaum, C., and Kamoun, S. (2010). Genome evolution following host jumps in the irish potato famine pathogen lineage. *Science*, 330(6010):1540–1543.
- Ribeiro, R. M., Bonhoeffer, S., and Nowak, M. A. (1998). The frequency of resistant mutant virus before antiviral therapy. *AIDS*, 12(5):461–465.
- Rieux, A., Soubeyrand, S., Bonnot, F., Klein, E. K., Ngando, J. E., Mehl, A., Ravigne, V., Carlier, J., and de Bellaire, L. d. L. (2014). Long-distance wind-dispersal of spores in a fungal plant pathogen: estimation of anisotropic dispersal kernels from an extensive field experiment. *PLoS One*, 9(8).

- Rimbaud, L., Papaïx, J., Barrett, L. G., Burdon, J. J., and Thrall, P. H. (2018). Mosaics, mixtures, rotations or pyramiding: What is the optimal strategy to deploy major gene resistance? *Evolutionary Applications*, 11(10):1791–1810.
- Roy, M. and Pascual, M. (2006). On representing network heterogeneities in the incidence rate of simple epidemic models. *Ecological Complexity*, 3(1):80–90.
- Sapoukhina, N., Durel, C.-E., and Le Cam, B. (2009). Spatial deployment of gene-for-gene resistance governs evolution and spread of pathogen populations. *Theoretical Ecology*, 2(4):229.
- Sapoukhina, N., Tyutyunov, Y., Sache, I., and Arditi, R. (2010). Spatially mixed crops to control the stratified dispersal of airborne fungal diseases. *Ecological Modelling*, 221(23):2793–2800.
- Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., and Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nature Ecology & Evolution*, 3(3):430.
- Sethi, S. P. and Staats, P. W. (1978). Optimal control of some simple deterministic epidemic models. *Journal of the Operational Research Society*, 29(2):129–136.
- Shan, L., He, P., Zhou, J.-M., and Tang, X. (2000). A cluster of mutations disrupt the avirulence but not the virulence function of *avrpt0*. *Molecular Plant-Microbe Interactions*, 13(6):592–598.
- Shigesada, N., Kawasaki, K., and Teramoto, E. (1979). Spatial segregation of interacting species. *Journal of Theoretical Biology*, 79(1):83–99.
- Singh, R. P., Hodson, D. P., Jin, Y., Huerta-Espino, J., Kinyua, M. G., Wanyera, R., Njau, P., Ward, R. W., et al. (2006). Current status, likely migration and strategies to mitigate the threat to wheat production from race Ug99 (TTKS) of stem rust pathogen. *CAB reviews: perspectives in agriculture, veterinary science, nutrition and natural resources*, 1(54):1–13.
- Skellam, J. G. (1951). Random dispersal in theoretical populations. *Biometrika*, 38(1/2):196–218.
- Skelsey, P., Rossing, W. A. H., Kessel, G. J. T., and van der Werf, W. (2010). Invasion of *Phytophthora infestans* at the landscape level: How do spatial scale and weather modulate

- the consequences of spatial heterogeneity in host resistance? *Phytopathology*, 100(11):1146–1161.
- Skelsey, P., With, K. A., and Garrett, K. A. (2013). Why dispersal should be maximized at intermediate scales of heterogeneity. *Theoretical Ecology*, 6(2):203–211.
- Smithson, J. B. and Lenne, J. M. (1996). Varietal mixtures: a viable strategy for sustainable productivity in subsistence agriculture. *Annals of Applied Biology*, 128(1):127–158.
- Stam, R. and McDonald, B. A. (2018). When resistance gene pyramids are not durable - the role of pathogen diversity. *Molecular Plant Pathology*, 19(3):521–524.
- Stuthman, D. D., Leonard, K. J., and Miller-Garvin, J. (2007). Breeding crops for durable resistance to disease. *Advances in Agronomy*, 95:319–367.
- Sun, P. and Yang, X. B. (1999). Two properties of a gene-for-gene coevolution system under human perturbations. *Phytopathology*, 89(9):811–816.
- Tellier, A. and Brown, J. K. M. (2007). Stability of genetic polymorphism in host-parasite interactions. *Proceedings of the Royal Society B: Biological Sciences*, 274(1611):809–817.
- Thrall, P. H. and Burdon, J. J. (2003). Evolution of virulence in a plant host-pathogen metapopulation. *Science*, 299(5613):1735–1737.
- Tilman, D. (1994). Competition and biodiversity in spatially structured habitats. *Ecology*, 75(1):2–16.
- van den Bosch, F. and Gilligan, C. A. (2003). Measures of durability of resistance. *Phytopathology*, 93(5):616–625.
- van den Bosch, F. and Gilligan, C. A. (2008). Models of fungicide resistance dynamics. *Annual Review of Phytopathology*, 46:123–147.
- van den Driessche, P. and Watmough, J. (2002). Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences*, 180(1-2):29–48.
- Vanderplank, J. (1948). The relation between the size of fields and the spread of plant-disease into them. part i. crowd diseases. *Empire Journal of Experimental Agriculture*, 16:134–142.

- Vanderplank, J. E. (1963). *Plant diseases: epidemics and control*. Academic Press, New York, USA.
- Vyska, M., Cunniffe, N., and Gilligan, C. (2016). Trade-off between disease resistance and crop yield: a landscape-scale mathematical modelling perspective. *Journal of The Royal Society Interface*, 13(123):20160451.
- Webb, S. D., Keeling, M. J., and Boots, M. (2007). Host–parasite interactions between the local and the mean-field: How and when does spatial population structure matter? *Journal of Theoretical Biology*, 249(1):140–152.
- Wellings, C., Burdon, J., McIntosh, R., Wallwork, H., Raman, H., Murray, G., et al. (2000). A new variant of *Puccinia striiformis* causing stripe rust on barley and wild hordeum species in australia. *Plant Pathology*, 49(6).
- White, K. J. and Gilligan, C. (2006). The role of initial inoculum on epidemic dynamics. *Journal of Theoretical Biology*, 242(3):670–682.
- Wolfe, M. and McDermott, J. (1994). Population genetics of plant pathogen interactions: the example of the *Erysiphe graminis-Hordeum vulgare* pathosystem. *Annual Review of Phytopathology*, 32(1):89–113.
- Wu, J. (2009). Ecological dynamics in fragmented landscapes. In Levin, S., editor, *Princeton guide to ecology*, pages 438–444. Princeton University Press, Princeton, USA.
- Ye, T. Z., Yang, R. C., and Yeh, F. C. (2003). Coevolution in natural pathosystems: Effects of dominance on host–pathogen interactions. *Phytopathology*, 93(5):633–639.
- Young, N. D. (1996). QTL mapping and quantitative disease resistance in plants. *Annual Review of Phytopathology*, 34:479–501.
- Zhan, J. and McDonald, B. A. (2013). Field-based experimental evolution of three cereal pathogens using a mark–release–recapture strategy. *Plant Pathology*, 62(S1):106–114.
- Zhan, J., Thrall, P. H., Papaix, J., Xie, L., and Burdon, J. J. (2015). Playing on a pathogen’s weakness: using evolution to guide sustainable plant disease control strategies. *Annual Review of Phytopathology*, 53:19–43.