Effects of Land Management upon Species Population Dynamics: A Spatially Explicit, Individual-based Model

by
Hazel Ruth Parry

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds
School of Geography
October 2006

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

This PhD was funded by a grant from the Central Science Laboratory Seedcorn fund.

I would like to say thank you to everyone who has given me advice and support over the last four years, in particular my primary supervisors Dr Derek Morgan[1,2] and Dr Andrew Evans[3]. Other members of my research support group at the University of Leeds and at the Central Science Laboratory have made invaluable contributions to the development of this work, notably Dr Jason Noble[4], Dr Keith Walters[1], Prof. Stuart Lane[3] and Dr Steve Carver[3]. Others who have generously given their time to help me include: my brother Daniel Parry, who takes credit for my understanding of Java programming; Dr Karim Djemame[4], who advised on parallel computing; Phil Northing[1] for his feedback; Dr Hans Baveco[5], who advised on 'super-individuals'; Dr Sim Reaney[3], who advised on landscape generation; Dr Andrew Ross[6] for last-minute help with LaTeX; Garry Fry[1] for the photographs of *R. padi* and *S. avenae*; Stéphane Pietravalle[1] for statistical advice; and Prof. Stephen Compton[7] for the use of his aphid literature. I'm also grateful to the Multi-agent Systems and Simulation (MASS) research group[3] for inspiration, debates and cake.

Special thanks to Dr Alison Heppenstall[3], who has been a great friend, conference companion and advisor, as well as the first person to read this thesis. Thanks also to all the other wonderful people I know in Leeds and York, who have made the past four years fly by. Finally, huge thanks to Beryl and David Parry - my Mum and Dad.

---

[1] Central Science Laboratory
[2] Health and Safety Laboratory
[3] School of Geography, University of Leeds
[4] School of Computing, University of Leeds
[5] Alterra
[6] School of Environment, University of Leeds
[7] School of Biology, University of Leeds

Abstract

Individual-based approaches in ecology provide a new approach to spatially explicit modelling. They are paralleled by the emergence of agent-based modelling in the field of artificial intelligence (AI) that is manifest in object-based approaches in a number of geographical disciplines, from hydrology to sociology. An individual-based approach to the simulation of organisms in a spatial context allows for a greater understanding of how individual-level behaviour and interactions result in population-level phenomena at the landscape-scale. Such models are inherently flexible and adaptable to other species or systems, and the model can be parameterised from biological behavioural information widely available in the literature.

This research constructs, analyses and experiments with an individual-based model of aphid (*Rhopalosiphum padi*) population dynamics in agricultural landscapes during the autumn and winter. The model combines deterministic equations governing the development of the aphids with stochastic, behavioural rules. Several stages of model assessment validate the model: assessment at the conceptual, developmental and operational stages. The need for a solution for the model to cope with large population sizes led to experimentation with both mathematical and computational solutions to this problem. It was found that parallel computing to distribute the simulation across a 30-node Beowulf cluster was the most effective at increasing model efficiency whilst preserving model behaviour. Key scenarios are presented, that show the power of this approach in predicting potential impacts of agricultural landscape change, including the effects of crop management, marginal habitat configuration and pesticide regime. This research clearly demonstrates the potential of spatially explicit individual-based modelling to predict scenarios that may advise policy decision-makers as a landscape management tool.

Contents

List of Tables

List of Figures

Publications from this Thesis

Reference	**Relevant thesis chapter**
Journal publications	
Parry, H.R., Evans, A.J. and Morgan, D. (2006) Aphid population response to agricultural landscape change: A spatially explicit, individual-based model. Ecological Modelling, *199, 451-463.*	Chapters 2, 4 and 9.
Parry, H.R. (2008) Agent Based Modeling, Large Scale Simulations. *Invited chapter*, Encyclopedia of Complexity and System Science, Meyers, R.A. (*ed.*), Springer-Verlag, Germany	Chapter 6.
Parry, H.R., Evans, A.J. and Morgan, D. (-) Millions of agents: pushing the limits of agent-based simulation. *In prep.*	Chapter 6.
Peer-reviewed conference proceedings	
Parry, H.R., Evans, A.J. and Heppenstall, A.J. (2006) Millions of agents: parallel simulations with the Repast agent-based toolkit. In R.Trappl (ed.) "Cybernetics and Systems 2006". Vienna: Austrian Society for Cybernetic Studies, April 2006.	Chapter 6.
Parry, H.R., Evans, A.J. and Morgan, D. (2004) Aphid Population Dynamics in Agricultural Landscapes: An Agent-based Simulation Model. In Pahl-Wostl, C., Schmidt, S. and Jakeman, T. (eds) iEMSs 2004 International Congress: "Complexity and Integrated Resources Management". International Environmental Modelling and Software Society, Osnabrueck, Germany, June 2004.	Chapter 4.

CHAPTER 1

Introduction

1.1 Context of the research

Pest control in the UK is a key element of arable agricultural management. A major arable agricultural pest is the aphid, causing damage to crops both by direct feeding damage and indirectly by the transmission of plant viruses. In particular, Barley Yellow Dwarf Virus (BYDV), transmitted by over twenty species of aphid, is a devastating and widespread disease of cereals worldwide (Bailey et al., 1995). The risk of the spread of this disease may be increasing with climate change in the UK, due to milder winters and wetter summers leading to greater survival in winter and larger summer populations (Harrington, 2002). Economic losses due to cereal aphids alone were calculated to be around £120m per annum in the UK in the 1980s (Tatchell, 1989). Control of aphid pests historically relies on insecticide application. In the autumn, this application is primarily prophylactic spraying against aphids for BYDV control. A more rational method of control of aphids would be an economic and efficient application of insecticide, at the right time and place (Johnson, 1962). Better understanding of population dynamics, movement and risk would potentially lead to more rational use of pesticides in this context. In addition, as early as 1962, Johnson acknowledged that alternative methods of control existed, including arranging planting dates to minimise impacts and control by natural enemies. Advances in scientific knowledge to develop greater understanding of pest population dynamics were required to improve insecticide application management and explore alternative forms of control, research which continues today.

In order to further understanding of pest population dynamics at the landscape scale, three key influences on landscape-insect interactions should be considered. Firstly, how pest populations are spatially structured. Spatial structure may be evident at differing landscape scales; thus, it is important to consider both the regional distribution and the local abundance of the pest species. The movement and migration of the species in relation to the landscape structure will determine how the pest population will spread across the landscape, and how outbreaks may occur. Secondly, at the landscape scale, changing landscape structure will alter the dynamics of pest outbreaks. Key landscape factors are the size and configuration of fields, the fragmentation of crop habitat and the presence of intermediary, non-crop habitat. The structure of the landscape may also be affected by management regimes, such as pesticide spray or field boundary management that will impact upon the pest population dynamics both locally and regionally. Thirdly, the importance of spatial scale and the way in which the landscape-insect interactions may alter with changes in scale should be considered. A key method that might be employed to understand such complex spatial landscape-insect interactions is simulation modelling.

In ecology, it has become evident that geography is important in simulation modelling: there is a need for simulation models to take into account spatial variation in landscapes, as well as changes over time. In addition, models are increasingly required to be flexible enough to function at differing landscape scales, and to be adaptable to multiple species. Such models aim to develop greater understanding of the behavioural processes that lead to population dynamics and distributions, and are useable as landscape management tools. Development of more flexible spatial models has required significant advances in the computational skills of ecologists and application of complex computer science methods in recent years. This thesis presents a primary example of this novel integrated approach.

1.2 Aims and objectives of the research

The overall aim of this research is to develop and apply a simulation model of aphid pest population dynamics in agricultural landscapes, using detailed biological information about an aphid pest species and their localised behaviour to predict potential outcomes from landscape change scenarios at a larger scale. This model is intended as a research tool to aid understanding of insect response to landscape change and how individual insect behaviour may lead to population-level dynamics. It is also intended as a broad-scale management tool for testing pest management strategies (e.g. for policy makers).

To achieve this aim, the following research objectives were formulated:

1. Review and discuss current understanding of insect pest population dynamics in the context of changing agricultural landscapes in England, highlighting potential areas for ecological simulations to aid research.
2. Evaluate the modelling techniques that could be employed to construct a generic, multi-scale simulation of ecologically-important organisms in agricultural landscapes.
3. Use spatially explicit individual-based modelling (IBM) techniques to build a model to simulate the spatial and temporal population dynamics of *Rhopalosiphum padi*, an aphid pest of cereals, in agricultural landscapes.

4. Conduct model assessment including conceptual, developmental and operational validation and sensitivity testing.
5. Develop and evaluate techniques for simulating large numbers of individuals.
6. Demonstrate the application of the model to study insect pest population outbreaks, both spatially and temporally, in relation to agricultural management practices. In particular, to examine how populations are spatially structured and how changing landscape structure alters the dynamics of pest outbreaks.
7. Provide recommendations for future work in the form of a research agenda.

1.3 Research methods

This study requires an inter-disciplinary approach and thus it spans three disciplines: computer science, ecology (pest management) and geography.

1.3.1 Computer science and simulation modelling

Computational science has been used in this thesis to: (1) develop the simulation model (including coding the model using object-oriented Java programming and the use of the Repast toolkit); (2) develop computational solutions to modelling problems (through parallel computing and super-individuals); (3) construct mathematical equations for key model parameters.

Over the last ten years computing has made significant progress. However, in fields such as geography and ecology, scholars and practitioners are reluctant to fully embrace new technologies (Openshaw and Turton, 2000), instead preferring to continue to run models constructed years ago on machines that are now much faster, as such models that once took days to run now run in seconds. Yet, there is more to the change in technology than increased processor speeds (Openshaw and Turton, 2000). For example, by using more recent object-oriented programming techniques and parallelisation, more powerful models that fully exploit the power of twenty-first century technology can be realised. With a return to more quantitative approaches in geography in the last decade (for example the importance of Geographical Information Systems in geographical research) geographers are starting to embrace technology and computers have become used intensively as a research tool.

Since the late 1980s, IBM has grown as a self-aware sub-discipline within the field of ecology (DeAngelis and Gross, 1992; Grimm and Railsback, 2005), drawing upon parallel advances in other fields such as computer science and social science. IBM (known as multi-agent simulation (MAS) in other fields) allows for representation of individuals, their behaviour and their interactions. IBM has been used in this thesis as it is a flexible approach that enables the creation of biologically-realistic simulation models to investigate the behaviour of organisms, and enables models to function more easily at multiple landscape scales.

1.3.2 Ecology and pest management science

In this thesis ecological science has been employed to: (1) ensure the biological realism of the model in relation to empirical studies in the literature; (2) parameterise the model from the information in such studies and published field results; (3) compare the model to field datasets to validate the model; (4) develop and analyse model scenarios; (4) construct mathematical equations for key model parameters.

In agricultural ecosystems, individual-based models can help increase understanding of plant-insect interactions at the landscape scale and how such interactions may be best manipulated to minimise the impacts of pests. Individual-based behavioural models enable exploration of the link between individual and higher-level processes, such as the relationship between local interactions (e.g. density dependent morph determination) with regional population fluctuations.

1.3.3 Geography

The thesis uses geographical science to: (1) analyse the outcomes of the model: in particular scenario generation at the landscape scale; (2) manipulate landscapes with which to test the model behaviour; (3) explore relationships between individual behaviour, local interactions and regional population dynamics.

A spatial, geographical approach is taken to the studies presented in this thesis. Exploration of the population structure in terms of regional and local distributions as well as migration is important. Geography is at the core of this thesis, linking individual behaviour and local responses to landscape change with regional population dynamics.

1.4 Organisation of the thesis

The structure of the thesis relates to the research aims and objectives. Chapter 2 expands the key themes highlighted in the research context outlined above in section 1.1, and provides the background to current understanding of insect pest population dynamics in agricultural landscapes of the UK. In addition, the role that modelling can play in improving crop protection is assessed. A range of modelling techniques are critically reviewed in this chapter, and an evaluation is made of current aphid population models. Subsequently, an IBM approach is identified as the most flexible and suitable technique to employ for this research.

Prior to the core model development chapters, the data used as a basis for the simulation development and model analysis is reviewed in chapter 3. Following this chapter, the remainder of the thesis is split equally between model development and demonstration of the use of the model to analyse agri-management options in a series of model 'scenarios'.

Chapters 4, 5 and 6 are the core model development chapters. These are technical chapters that present firstly, the model development (chapter 4); secondly, model analysis and validation (chapter 5); thirdly, an evaluation of techniques to improve the model's capabilities as a consequence of the limitations of computing power to model large numbers of individuals, using both a mathematical and a computational solution (chapter 6).

Following the three modelling chapters are three chapters (chapters 7, 8 and 9) that explore the possible effects of agricultural change upon insect population dynamics, using the model. These chapters focus on key agricultural practices that are known to impact upon aphid population dynamics: chapter 7 compares the effects of altering the timing of crop introduction and crop configuration; chapter 8 explores the role of non-cropped habitat in the movement and survival of aphids in the landscape and chapter 9 examines the implications of varying pesticide regimes upon the population dynamics of the aphid.

The final chapter summarises the findings in the thesis and evaluates the model in more general terms computationally and scientifically, as well as in terms of agri-management purposes.

CHAPTER 2

Background

2.1 Introduction

In relation to pest management, this chapter reviews key literature on: agricultural change (section 2.2); ecological modelling in an agricultural landscape context (section 2.3) and the rise of the novel modelling approaches of multi-agent simulation (MAS) and individual-based modelling (IBM) (section 2.4). IBM is evaluated as a key methodology with the potential to develop more generic and powerful simulation models of pests species in agricultural landscapes. An overview of the advantages of using modelling toolkits to develop IBM simulations is given (section 2.5), qualifying the use of the MAS development toolkit Repast in the development of the simulation model in this thesis. The final part of this chapter focuses on the chosen study organism, aphids, and evaluates historic and current approaches to the simulation of aphid pest population dynamics (section 2.6). The potential advances that an individual-based approach will make in the field of aphid modelling are highlighted.

Overall, this chapter meets the first and second objectives of this thesis: firstly, to review and discuss current understanding of pest population dynamics in the context of changing agricultural landscapes in England, highlighting potential areas for ecological simulations to aid research; secondly, to evaluate the modelling techniques that could be employed to construct a generic, multi-scale simulation of ecologically-important organisms in agricultural landscapes.

2.2 Agricultural change and insect pests

2.2.1 Introduction

The June 2004 agricultural survey showed that the total crop area in the UK was 4.6 million hectares (`http://www.defra.gov.uk/esg/work_htm/publications/cs/`

farmstats_web/default.htm). Of this, 2.0 million hectares were given to wheat. In England, most arable farming is in the South and East (figure 2.1), where the agricultural landscape is characterised by large, intensively farmed fields separated by narrow hedgerows.

Since the end of World War II, changes in arable farming have led to intensification of production, resulting in a 'dramatic reduction' in landscape diversity (Robinson and Sutherland, 2002). Since 1990, there has been a trend towards fewer, larger farms in the more intensively farmed areas. Farmers have diversified their sources of incomes through other activities such as farm-based tourism, agricultural contracting and off-farm working, although individual farmers tend to remain specialised in their range of agricultural enterprises (DEFRA, 2000). Thus, fewer, larger, highly mechanised farms now exist, that produce approximately four times the yields of 1945 (Robinson and Sutherland, 2002). Changes since World War II have impacted on a large number of species significantly, and in general there has been a decline in all farmland taxa: mammals, plants, birds and invertebrates (Robinson and Sutherland, 2002). In recent years, national and international government policies have worked to deintensify farming in such landscapes (for example 'set-aside', the environmental stewardship and conservation headland schemes), and a significant organic movement has altered the way some agricultural land is now managed. Integrated Farm Management (IFM) is also gaining popularity as a means to maintain efficient food production with minimum environmental impact (e.g. `http://www.leafuk.org/leaf/`).

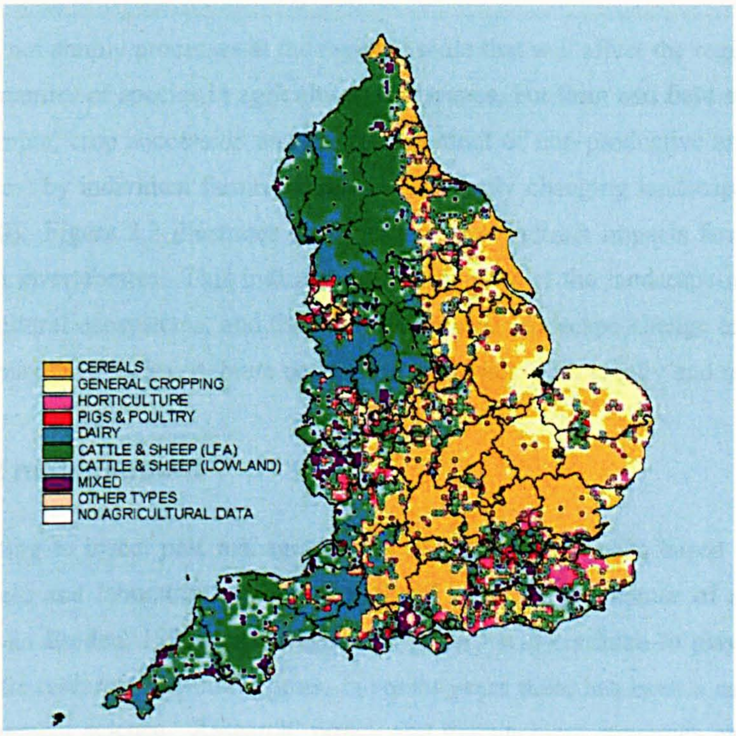


Figure 2.1: Dominant farm types in England by 5 kilometre grid square. ©Crown Copyright 2004 Reproduced from the Ordnance Survey Defra license no. GD272631 Department for Environment, Food and Rural Affairs.

2.2.2 The structure and complexity of agricultural landscapes

Large-scale intensive agriculture has resulted in increased uniformity of landscape structure, which is both spatial, with large areas of mono-crop, and temporal, as old rotation systems have been replaced by the growth of the same crop year after year. Landscape structure is important to the population dynamics of invertebrate species. Both temporal and spatial changes to the structure of the agricultural landscape will influence the distribution of pest and predator populations. Östman et al. (2001) show that altering regional landscape structure can enhance the biological control of aphid pests, where aphid aggregation depends upon crop type, crop system, environmental region and surrounding landscape.

Arable agricultural landscapes are typically characterised as a patchwork of large fields of mono-crops with low species diversity, intersected by corridors of higher diversity such as hedgerows and woodland. This landscape structure inevitably influences the ecology of insects that live within it. Conservation studies tend to treat this landscape as bi-polar: as islands of favourable habitat within a sea of hostile, chemically maintained mono-culture (e.g. Dennis et al., 2000). Much of this philosophy derives from theories of island biogeography (MacArthur and Wilson, 1967) and metapopulations (Levins, 1969), for example many theories about population dynamics and nature reserve configuration are based upon these ideas (Cabeza et al., 2004; Hanski, 1997; Hanski and Simberloff, 1997; Harrison, 1994).

However, it is not simply processes at the regional scale that will affect the regional and local population dynamics of species in agricultural landscapes, but farm and field scale processes also. For example, crop succession and the management of non-productive areas (including field boundaries) by individual farmers creates a constantly changing landscape mosaic (Burel et al., 2000). Figure 2.2 illustrates the key direct and indirect impacts farming practices might have on invertebrates. This indicates the complexity of the landscape-insect relationships in agricultural ecosystems, and the multiple impacts landscape change and agricultural management may have on invertebrate population dynamics, both locally and regionally.

2.2.3 Insect pest management

Research relating to insect pest management to date has largely been based upon isolated, small scale field and laboratory studies or national monitoring schemes of a few selected species (e.g. van Emden, 1972). Although such studies will continue to play an important role in scientific research into insect pests, in recent years there has been a call for changes in pest management science. Research, policy and management strategies need to become more integrated in order to address contemporary questions. Such questions regard insect movement and integrated pest management at larger spatial and temporal scales, and across differing levels of ecological organisation. They should be addressed in a "transdisciplinary" manner, where research is coordinated between research groups and interested parties, and

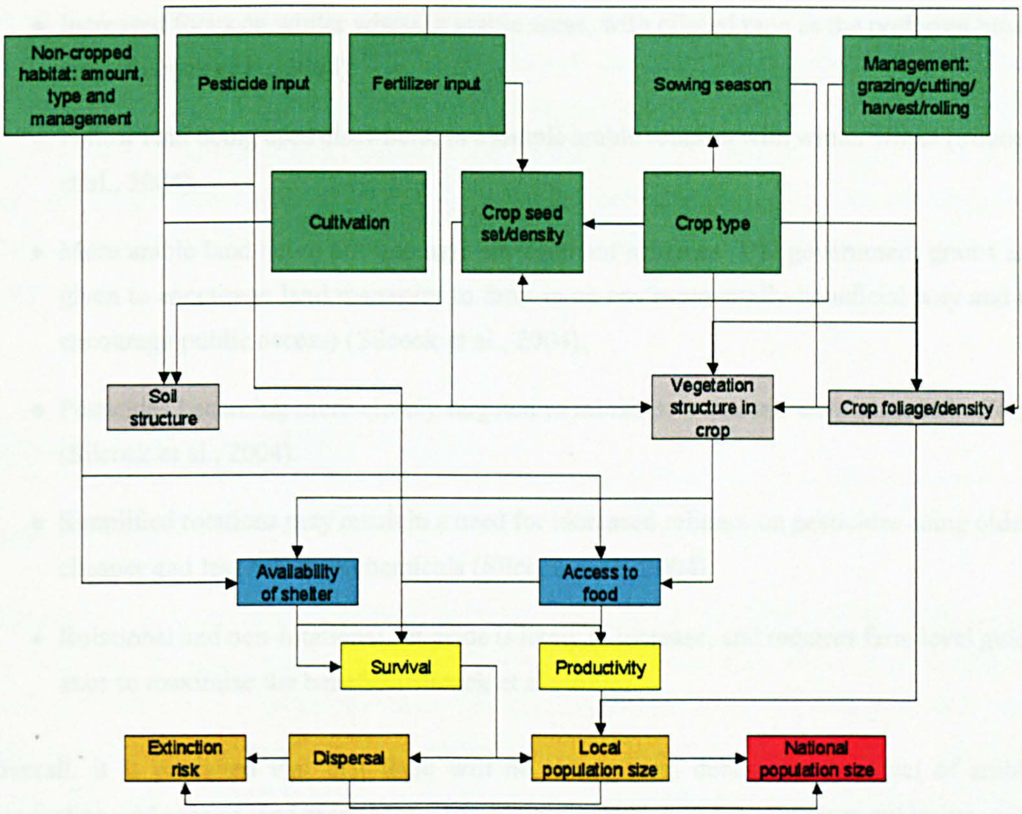


Figure 2.2: The complex relationship between farming practices and their impact on invertebrate populations (after Benton et al., 2003).

also between disciplines, in order that research and insect pest control programs can be effectively implemented at the landscape scale (Barrett, 2000).

In this, simulation models are to play an important role, as they provide a means to test hypotheses and carry out studies that address questions at greater temporal and spatial scales, as well as enable analyses at multiple scales (Barrett, 2000). However, for an integrated approach to be instigated, such models will face challenges of data availability at both the regional and individual behavioural scale. This may cause problems for both model parameterisation and validation, and is further discussed in chapter 3.

2.2.4 Contemporary issues

The agricultural landscape is highly dynamic, changing constantly from season to season. However, general trends in agricultural practice can be identified, and perhaps also predicted in relation to policy reforms. In particular, changes to the Common Agricultural Policy (CAP) detailed within the reform agreement of June 2003 are believed to potentially have wide ranging implications for land management in the UK. The following arable landscape changes have been suggested as potential outcomes of the CAP reform agreement (Silcock et al., 2003, 2004):

- Increased focus on winter wheat in arable areas, with oilseed rape as the preferred break crop (Silcock et al., 2004).
- Fallow land being used elsewhere, in a simple arable rotation with winter wheat (Silcock et al., 2004).
- More arable land being put into agri-environment schemes (UK government grants are given to encourage land managers to farm in an environmentally beneficial way and to encourage public access) (Silcock et al., 2004).
- Pesticides becoming more closely targeted to minimise costs and environmental effects (Silcock et al., 2004).
- Simplified rotations may result in a need for increased reliance on pesticides using older, cheaper and less selective chemicals (Silcock et al., 2004).
- Rotational and non-rotational set-aside is likely to increase, and requires farm level guidance to maximise the benefits (Silcock et al., 2003).

Overall, it is predicted that that there will be a long-term decline in the level of arable production and area of land cropped, and arable agriculture will become more extensive, with increased fallow land and reduced inputs, due to decoupling and Single Farm Payments (GFA-RACE and IEEP, 2003). At the same time this may lead to some regional specialisation and concentration, as already evident (DEFRA, 2000). Enhanced funding for agri-environment schemes and rural development schemes should lead to more environmentally sensitive farming, reductions in pesticide use, protection and management of landscape features such as hedgerows and also diversification (Silcock et al., 2003). However, specialisation and concentration in some areas could have the inverse effect on the landscape, with loss of hedgerows and continued use of chemicals. Specialisation and simplification of cropping systems will result in a move to larger areas of winter wheat (GFA-RACE and IEEP, 2003). Not all areas may diversify, and a decline in mixed farming is also possible, with more homogeneous cropping that will reduce landscape diversity.

2.2.5 Summary

In agricultural systems in the UK, there is an interplay between the effects of regional scale landscape dynamics such as land use and landscape structure, that influence resource availability, landscape connectivity and permeability, and smaller scale dynamics at a farm and field scale. At the farm scale, crop succession and the management of boundaries or non-productive areas will affect insect population dynamics, and at the field scale farming practices will also have an effect, all due to their impacts on habitat quality and availability (Burel et al., 2000).

In terms of the management of pest populations, field studies can be used in conjunction with modelling approaches to address multi-scale research questions. The development of reliable, flexible modelling tools to predict trends in pest population dynamics is critical to aid the development of pest control programs at the landscape scale. The following sections detail the modelling approaches in the literature that have been applied to agricultural landscapes in the past, and recent developments that have the potential to provide the more flexible tools required.

2.3 Ecological modelling

2.3.1 Introduction

Simulation has been used extensively to explain ecological and geographical phenomena, by analysing the properties of theoretical models of the surrounding world. Traditionally in ecology, these models have been mostly based upon mathematical 'top-down' relationships between variables, derived from empirical measurements. In particular, differential equations have been used in many models, based on general empirical laws (e.g. the Lotka-Volterra equation of predator-prey dynamics: Lotka, 1925; Volterra, 1926).

Firstly, in terms of ecological modelling, this has meant many spatial models prior to the 1990s have focused on populations or species groups, rather than individual animals, thus describing the 'average' individual (e.g. Renshaw, 1991). In particular, due to the mathematical top-down relationships between variables, models in ecology have often failed to take into account the complexity of the multiple concurrent interactions in ecosystems (Laval, 1996). By ignoring individual behaviour, important factors are not taken into account, including reproduction and competition at the individual level, which may greatly influence general population trends.

Secondly, much of classical modelling has been highly theoretical, and many predictions have been unreliable due to uncertainty about ecological processes. There is a significant need for ecological models to address real-world management problems, but issues include the lack of transferability of traditional models and their uncertainty (Conroy et al., 1995).

Thirdly, there has been a disappointing lack of genericness. This includes problems with models such as a unique structure, development by an individual or small group, poor documentation, lack of transportability across platforms or to other users, and models maintained only by their developers (if at all) (Reynolds and Acock, 1997). However, some attempts have been made to address this problem using object-oriented and individual-/agent-based models, of which Reynolds and Acock (1997) give examples (further examples can be found in more recent work, e.g. Grimm and Railsback, 2005). The inherent flexibility of an agent-based, object-oriented approach enables modellers to attempt to create more generic models.

Overall, there are several significant problems with the classical approach to ecological modelling. They include a 'sealed-off nature of level of analysis', a lack of 'complexity and realism of parameters', 'difficulty of modelling action (of individuals or groups)', and 'qualitative deficiency' (Ferber, 1999: pp. 35-36). Section 2.4 explores the ways in which relatively new 'individual-based' approaches in ecological simulation provide a means with which to address these issues.

2.3.2 Theoretical constructs of the agricultural landscape

Metapopulation theory

Metapopulation concepts are often applied in the context of agricultural landscapes in the UK. Metapopulation theory is based upon the idea that discrete patches of wildlife habitat holding breeding sub-populations are connected by corridors through an almost untraversable matrix. This generally assumes that there are empty patches waiting to be colonised (Hanski, 1999). The sub-populations that make the metapopulation have asynchronous population dynamics, where the subpopulations have different demographical fates. Much ecological research has been founded on an examination of the 'connectivity' and travel 'costs' between these patches, as well as patch composition, 'quality' and distribution (Green, 1994; Henien et al., 1998; LaGra, 1991). Hunter (2002) describes key landscape features that influence insects as being habitat edge-interior ratio, isolation of fragments, patch area, patch quality, patch diversity and micro-climate.

The emphasis on a divided, fragmented landscape with identifiable patches of favourable habitat in a hostile matrix is clear. However, when viewed in this way a landscape can only have a limited range of possible dynamics occurring within and between the subdivided populations (Wiegand et al., 1999). In reality habitat patches are not entirely discrete with a less distinct separation between habitat and matrix, and it is also hard to define when a patch is 'empty'. Metapopulation models therefore face the issues of inferring process from pattern.

Alternatives to the metapopulation paradigm

Although the metapopulation theory has been shown empirically to reflect the population dynamics of some species of insect, primarily butterflies (e.g. Hanski's studies of *Melitaea cinxia* (Hanski et al., 1994) and Hill's study of *Hesperia comma* (Hill et al., 1996)), it is questionable whether it can be applied to all species or in the context of agricultural landscapes and annual crops (Tscharntke and Brandl, 2004). Aphids, for example, are found not just in prime habitat, but also on intermediary hosts as they migrate across the landscape. High rates of migration could mean that subpopulations of aphids, even if they exist, have dynamics that are highly correlated. It is therefore undesirable to use metapopulation theory for all species, without prior consideration of what is already known about the spatial dynamics of the species (Elmhagen and Angerbjörn, 2001). In fact, so few species really fit into the metapopulation model neatly, that it is perhaps better to begin by modelling the influence of landscape structure by other means.

In contrast to the entrenchment of metapopulation ideas in wildlife management, a more integrated approach is evident in recent research into pest management. The interaction between the crop habitat and the wider matrix is not only acknowledged, but is seen as fundamental to understanding population dynamics, not just of pest species but of the valuable species that depend on them (Bommarco and Fagan, 2002; Holland et al., 2004). It is also acknowledged that agricultural land itself can be important to the survival of conservation species today. Such changes in the way the agricultural landscape is perceived are reflected in the increasing popularity of diffusion models (Hanski, 1999) and individual-based modelling (IBM) (DeAngelis and Gross, 1992; Grimm and Railsback, 2005), not only in pest management but also in modelling of other species dynamics (for example, aquatic ecosystems (Ferreira, 1995; Laval, 1995)).

Modelling populations in agricultural landscapes

Individual-based models (IBMs) take local, within-patch, population dynamics into account, which have been shown to impact the wider distribution of the species (Melbourne et al., 2004). From such models, that are not based upon a preconceived idea about the species' population-level behaviour, it is possible to better replicate complex spatial dynamics produced by emergent phenomena in real systems. Rather than prescribing an inability of particular species to traverse areas of 'hostile' landscape, IBMs treat the landscape as a continuum to which the individual responds in relation to its biological characteristics and needs. In reality the individuals have no global knowledge of 'patches' and 'connectivity', so why should a model prescribe dispersal ability based upon such constructs? Importantly, such models are better able to take into account the complexity of the landscape, and the changes that occur continuously over space and time. A landscape cannot realistically be categorised into permanently 'favourable' habitat patches within an unfavourable matrix. Indeed, some suggest that such a patch-based, binary view of the world should be abandoned (Wiegand et al., 1999). Variable habitat availability, edge effects, a myriad of farming systems and practices, crop rotation and variation as well as catastrophic events such as hedge removal and harvesting are all intrinsic in the arable landscape. Models that can take such factors into account will inevitably be better at representing the interaction between species and their environment, as well as predicting the response of such species to change.

Landscape studies have previously focused upon finding general rules that relate species diversity to habitat and landscape features; however, no such rules really exist, the relationships depend upon the organism in question and the region. Spatial ecology and dispersal varies between organisms, and other factors such as the influences of management and surrounding landscapes can also play an important role (Jeanneret et al., 2003). It is believed that an IBM approach provides a better mechanism with which to explore these processes and their multi-scale relationship with the population dynamics of insect pests.

2.4 Individual-based modelling and multi-agent simulation

2.4.1 Introduction

In order to produce models that are more capable of furthering understanding of the processes that influence population dynamics spatially and temporally, as well as forecasting the effects of management or other human activity on population distributions, new modelling approaches have developed in the past decade (Grimm and Railsback, 2005). Models have become more spatially explicit (Conroy et al., 1995; Dunning et al., 1995), and attempts have been made to link these to real landscapes via Geographical Information Systems. Section 2.3 presented a critique of the techniques that have been traditionally used to simulate species population dynamics, and which still maintain a high degree of popularity in the field today. This critique showed that these techniques have key shortcomings which limit their application in this thesis, such as a top-down approach, a lack of transferability and lack of genericness. Individual-based modelling (IBM) approaches were identified as a potential methodology for the simulation of pest population dynamics in agricultural landscapes. The foundations of this relatively new approach are reviewed in the following section.

2.4.2 The origins of individual-based simulation

Based upon object-oriented (Booch, 1994) and other software engineering concepts, a recent software 'synthesis' of 'agent-based computing' has arisen (Jennings, 2000). Agent-based ideas were first developed in Artificial Intelligence (AI) research in the late 1970s, when Hewitt (1977) proposed the concept of self-contained 'actors' concurrently communicating through messages (Heppenstall, 2004). The original software engineering concepts were built upon to search for a paradigm that would define less rigid relationships between computational entities and provide mechanisms to represent the inherent organisational structure of complex, distributed systems (Jennings, 2000), such as ecological systems.

As early as the 1950s, cybernetics theorists acknowledged that 'complex systems' could not be analysed as separate components in isolation from one another (e.g. Ashby, 1956). The global behaviour of such systems was thought to be 'more than the sum of its parts' (Parrott and Kok, 2000). Around the same time, the founder of modern ecology, Eugene Odum, was also putting forward the same idea that an ecosystem should be viewed as 'greater than the sum of its parts'. Thus, it was agreed that a purely reductionist approach was not adequate to describe a complex system such as an ecosystem.

Therefore, on the one hand, research was centred in the cognitive, natural and social sciences in terms of modelling, explaining and simulating natural phenomena to provide a complex system model. On the other hand the research was focused on development of a practical method and technique: complex computing systems (Ferber, 1999). Thus multi-agent simulation (MAS) and IBM have grown from a significant interdisciplinary movement (Bousquet and Le Page, 2004).

In ecology occasional, fairly obscure and independent modelling research into the effects of individual behaviour upon system-level properties has arisen (e.g. Kaiser, 1979; Myers, 1976; Newnham, 1964). It was not until the late 1980s that ecologists established these ideas as a self-conscious discipline of 'Individual-based Modelling', with the article written by Huston et al. (1988) being most influential. The later paper of Hogeweg and Hesper (1990) more clearly defines the pragmatic and paradigmatic potential of IBM in ecology, and synthesises the ideas with preceding research in Artificial Life. Two years later the fundamental literature was gathered following a conference in 1990 at Knoxville, Tennessee "Individual-based Models and Approaches in Ecology": DeAngelis and Gross (1992). The recent publication of "Individual-based Modelling and Ecology" by (Grimm and Railsback, 2005) presents the state of the science in ecology at the present time.

2.4.3 Definitions: 'agents', 'individuals' and 'objects'

A popular definition of an agent, particularly in AI research, is that of Wooldridge (1999: pp. 29):

> "an agent is a computer system that is situated in some environment, and that is capable of autonomous action in this environment in order to meet its design objectives."

In particular, it is the autonomy, flexibility, inter-agent communication, reactivity and proactiveness of the agents that distinguishes the paradigm and gives power to the software that implements it (Heppenstall, 2004; Jennings, 2000).

Agents and MAS have been much more explicitly defined in the literature than the IBM associated with ecology. Yet, in many ways MAS and IBM can be thought of as essentially the same thing, the difference in terminology simply resulting from parallel developments in different fields of research. However, MAS do give more emphasis to decision-making processes of the agents and their social organisation. Also agents may not simply be individuals, but higher levels of organisation (e.g. a village, a shoal of fish etc.) (Bousquet and Le Page, 2004). That is not to say that IBMs do not incorporate such ideas, for example research into social behaviour of insects (Saffre et al., 1999) and the concept of 'Super-Individuals' (Scheffer et al., 1995).

2.4.4 The key features of MAS: agent behaviour, interactions and the environment

It is the interaction of agents, both with one another and with their environment, that is key to the development of the simulation and landscape distribution pattern (Ferber, 1999; Jennings, 2000). These interactions range from very simple operations that send small amounts of information between agents as required, to much more complex interactions that can include cooperation, competition and negotiation. This is commonly done by a message-passing system in object-oriented programming. These interactions may also undergo dynamic changes, as the organisational context between the agents is modified.

Figure 2.3 shows how internal states and behavioural rules of an organism equates to the internal structure of an agent. These may be fixed, or alter depending on the agent's interactions with other agents or the environment (Epstein and Axtell, 1996).

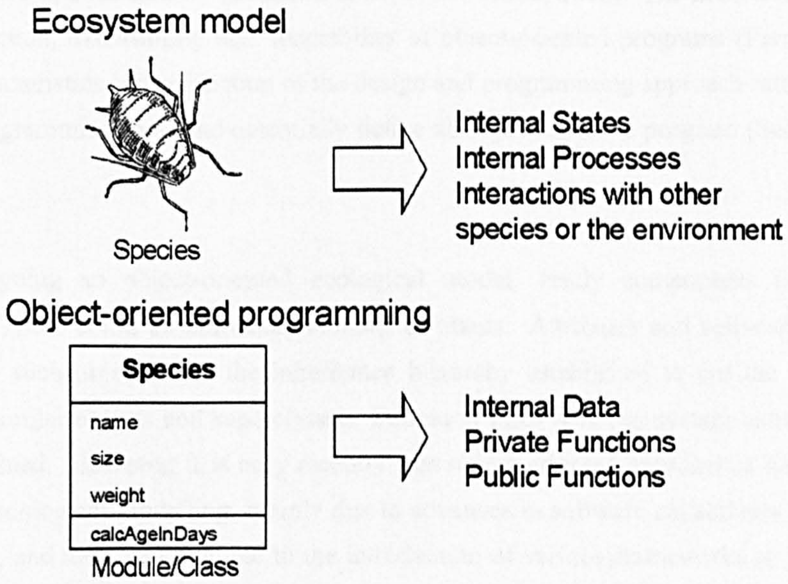


Figure 2.3: The similarities between the concept of an organism of a particular species and of a module within an object-oriented program (after Büssenschütt and Pahl-Wostl, 2000).

In ecological IBMs the environment, which represents the real ecological environment, plays an important part in the model. There are a number of properties the environment may have, which can have a significant impact on the outcome of the simulation. For example, firstly there may not simply be one environment, but several, or perhaps multiple layers of environmental information. Secondly, the environment may be static and unchanging throughout the simulation, or dynamic. Thirdly, the environment may have deterministic or non-deterministic effects upon the agent. Finally, the environment may be either continuous or discrete, and the information about the environment actually available to the agent may be limited (Russell and Norvig, 1995).

These factors combine to influence the agent behaviour in the simulation. Thus it is important to consider the implications of assigning particular properties to the environment, the implications of modelling interactions in a particular fashion, and how certain internal states and behaviour rules may have important outcomes. The simulation developed for this thesis is analysed in relation to these key features by conducting a sensitivity analysis, see section 5.3.

2.4.5 Building a MAS architecture

An 'implementation language' (Ferber, 1999) is used to develop a MAS. This includes representation of the agents and environment, as well as communications, behaviour and perception rule generation. It will also provide the user interface and tools required to finalise a MAS. The languages used are commonly 'classic' programming languages, such as Java and C++, which generally use an object-oriented structure. The key elements of object-oriented programming techniques for MAS include data abstraction, inheritance, encapsulation, polymorphism and modularity (Sequeira et al., 1997; Wood, 2002). The main advantages are the construction, extensibility and adaptability of object-oriented programs (Ferreira, 1995). These characteristics 'are a function of the design and programming approach rather than of a specific programming tool' and essentially define an object-oriented program (Sequeira et al., 1997).

When designing an object-oriented ecological model, firstly components (objects) are identified. These could be individual animals or plants. Attributes and behaviours are then assigned to such objects, and the inheritance hierarchy established to put the objects into classes of similar objects and superclasses. Structural links between system components can then be defined. However, it is only recently that object-oriented approaches have begun to be used in ecological modelling, mainly due to advances in software capabilities (Parrott and Kok, 2000), and most recently due to the introduction of various frameworks to aid software development (Grimm and Railsback, 2005; Lorek and Sonnenschein, 1998, 1999); see section 2.5. Examples of ecological models that implement object-oriented programming include the HOOFS animal foraging model (Beecham and Farnsworth, 1998; Beecham et al., 2002), the winter deer browsing model (Congleton et al., 1997), the ATLSS species response to hydrological change model (DeAngelis et al., 1998), the GEMACE farming-hunting-duck model (Mathevet et al., 2003), the ECOWIN aquatic ecosystem model (Ferreira, 1995), the HOOD pelagic Tunicate bloom model (Laval, 1995, 1996), and the SHALOM species dynamics model (Ziv, 1998a,b).

Multi-Agent Systems are also able to consider both expert opinion and empirical parameters, and have the capacity to integrate quantitative variables, differential equations and rule based behaviour into the same model. Modifications to the model are also quite straightforward (such as adding another species). It therefore helps in the search for a model, rather than simply model implementation and response analysis.

2.4.6 Suitability for modelling complex systems

Ecosystem and multi-agent modelling have both been much influenced by the adoption of ideas arising from studies of complex systems (Parrott and Kok, 2000).

Complexity and complex systems pertain to ideas of randomness and irregularity in a system, where individual-scale interactions may result in either very complex or surprisingly simple patterns of behaviour at the larger scale. Complex systems are therefore usually made up of agents interacting in a non-linear fashion. The agents are capable of generating emergent behavioural patterns, of deciding between rules and of relying upon data across a variety of scales. The concept allows for studies of interaction between hierarchical levels rather than fixed levels of analysis.

Some key concepts from complexity theory are important in ecological modelling discourse today. These include ideas of self-organisation, emergence, adaptivity, critical levels of connectivity and in some living systems, autopoiesis (Parrott and Kok, 2000). The exhibition by natural ecosystems of complex dynamics and structures has prompted a change in the general approach to ecological modelling today: instead of a ‘top-down’ approach, ‘bottom-up’ modelling strategies are now favoured, such as MAS, as a direct result of this school of thought (Parrott and Kok, 2001). These models are also able to better represent environmental heterogeneity and demographic stochasticities (Beecham et al., 2002). Figure 2.4 illustrates how an ecosystem may be modelled as a complex system, as a logical progression from conceptual model to representation as objects and finally as computational constructs.

IBM brings together both reductionist and holistic approaches. A reductionist approach, on the one hand, would attempt to understand things by dissection. On the other hand, a holistic approach aims to understand a whole system at once, and is the approach generally taken by ecologists. Through IBM and MAS it is possible to consider individuals, whilst also gaining greater understanding of the system as a whole. Whereas in most conventional approaches, ecosystems are studied only at one particular hierarchical level or spatial scale, the approach of IBM, MAS and complex systems ideas provides a framework in which entities at different

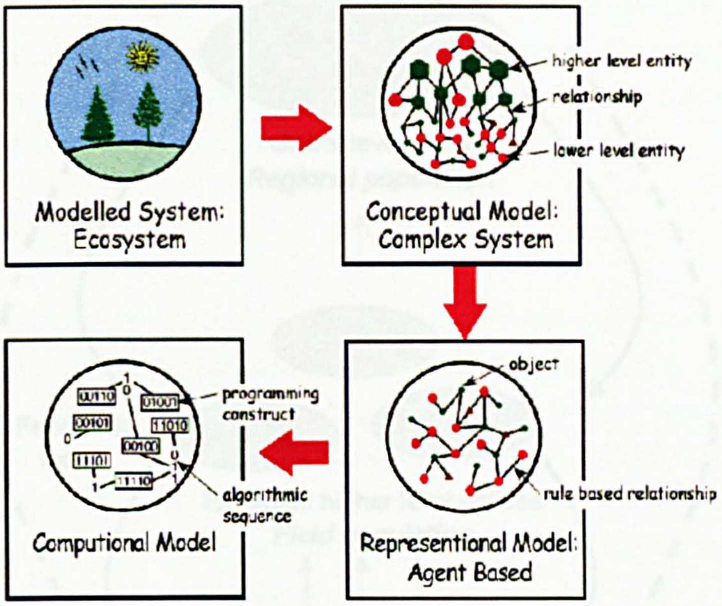


Figure 2.4: From ecosystem to agent-based computational model (after Parrott and Kok, 2000).

levels, and relationships between them, can be accommodated (Flake, 1998; Parrott and Kok, 2000). This relates importantly to the issues of spatial scale, and by using an agent-based modelling framework it is possible to incorporate differing landscape scales and scaling factors into the model itself. This allows analysis of interactions between entities at differing landscape scales and enables production of a model that will be essentially scale-free. Figure 2.5 illustrates the relationships between scales (hierarchical levels) in a complex system such as an insect community.

There is little ecological research that has actually explicitly utilised the potential of multi-agent systems in this manner, a notable exception being the CORMAS model of Mathevet et al. (2003). However, such models are not simple to implement when dealing with a complex system, as many authors have argued that complex systems are generally characterised by the presence of 'dissimilar spatial and temporal features at different scales' (Bar-Yam, 1999; Parrott, 2002).

2.4.7 Potential deficiency of individual-based simulation

Jennings (2000) outlines two major drawbacks relating to the use of an IBM or MAS approach. Firstly, the patterns produced by individual runs of the simulation are usually stochastic, requiring multiple runs to obtain a mean value. Secondly, possible emergent behaviour makes it difficult to associate system behaviour with the actions of the constituent components. Thus, although simulations are realistic as local decisions are made which may not be globally optimal, the resultant system dynamics are difficult to analyse.

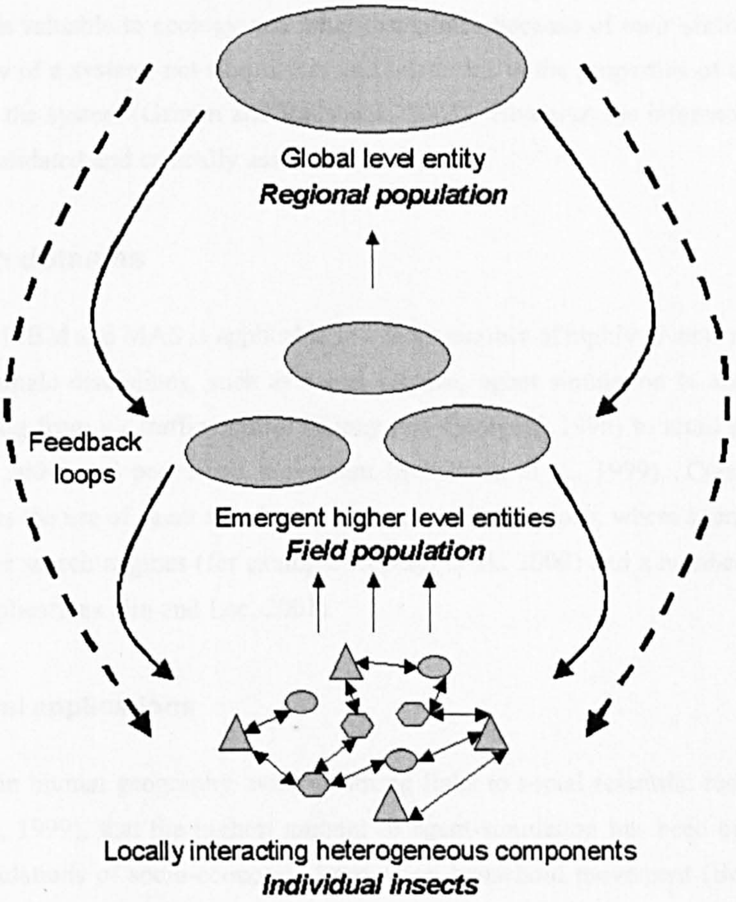


Figure 2.5: Hierarchy of scales and relationships between scales in a complex system, for example different levels at which to perceive an insect community (after Parrott, 2002).

The autonomy of the agents and stochasticity of agent-simulation leads to high levels of uncertainty in the system (Ferber, 1999), issues of validation and an inability to establish ‘a mathematical proof’ of the model results (Bousquet and Le Page, 2004), see chapter 5. Models can also be hard to communicate, as they are difficult to describe in the universal language of mathematics (Grimm and Railsback, 2005). Data requirements can be high, and some authors criticise IBMs for being ‘data hungry’, particularly in a spatial context (Beissinger and Westphal, 1998).

Important deficiencies found with IBMs in ecology, according to Grimm (1999), are that IBMs: produce results for a single species that are not generalisable; are complex, but do not address this complexity with techniques to deal with it; and are too elaborate to be communicated efficiently. It is important that IBM applications are driven as much by paradigmatic motivations as by pragmatic ones (Grimm and Railsback, 2005). That is to say, that development of such simulations should be as much about adding to general ecological theory as it is about prediction of single species distributions and population dynamics (Grimm et al., 1999).

IBMs are most valuable to ecology and other disciplines because of their ability to represent the complexity of a system, not simplify it, and relate this to the properties of the individuals that comprise the system (Grimm and Railsback, 2005). However, the inferences made must be carefully validated and critically assessed.

2.4.8 Application domains

The concept of IBM and MAS is applicable in a large number of highly diverse research areas. Even within single disciplines, such as social science, agent simulation is used to simulate systems ranging from air traffic control (Kinny and Georgeff, 1996) to retail petrol markets *(Heppenstall, 2004) and pedestrian movement (Schelhorn et al., 1999).* Computational *research includes the use of agent simulation for Internet applications, where agents can be used* as part of price search engines (for example Kephart et al., 2000) and a number of electronic commerce applications (Jin and Lee, 2001).

Geographical applications

Overall, it is in human geography, with its strong links to social scientific research (Gilbert and Troitzsch, 1999), that the highest amount of agent-simulation has been conducted. For example, *simulations of socio-economic impacts on household movement* (Benenson et al., 2001) and urban dynamics and growth (Portugali, 1999).

However, there are physical and biogeographical applications of agent simulation, for example research in landcover change such as the SLUCE project ("Spatial Land Use Change and Ecological Effects at the Rural-Urban Interface" `http://cscs.umich.edu/research/projects/sluce`) and integrated land management (Pahl-Wostl, 2002). There are similar concepts used in fluvial and hydrological research (for example Favis-Mortlock, 1998; Schmeeckle and Nelson, 2003).

Research into MAS in both human and physical geography is increasing rapidly, in line with other disciplines. Much of the development is limited by the programming and software development capabilities of the researchers; thus, unlike previous quantitative movements in geography, the development of software (such as toolkits, see chapter 2, section 2.5) is guiding agent formulation.

Ecological applications

Section 2.4.2 explains the origins of IBM in ecology and some of the key early texts. Further examples of simulations that use an object-oriented approach are given in section 2.4.5. Throughout the 1990s, a vast amount of ecological literature has accumulated, detailing many different IBMs, much of which has been published in the journal 'Ecological Modelling' (for a review see Grimm, 1999). Key simulations often referred to include ATLSS (Across Trophic Level System Simulation) (Abbott et al., 1995; DeAngelis et al., 1998), simulations of birds and mammals by Sibly et al. (e.g. 2005) and Topping et al. (e.g. 2003, 2005) and fish simulations by DeAngelis et al. (1980) (see also DeAngelis et al., 1993; Tyler and Rose, 1994; Van Winkle et al., 1997).

In terms of simulations of insects and insect behaviour, research has focused primarily on social insects such as ants and bees. Some key papers include: multi-agent simulation of honey-bee colonies by Sumpter and Broomhead (1998), social insect task partitioning by Ratnieks and Anderson (1999) and termite survival by Miramontes and DeSouza (1996). In terms of pest simulation, there is little research in this area. One example is research at the University of Queensland (Room et al., 1996), which models growth of individual plants, either singly or in small stands, and interactions with individual insects as they crawl on or fly between the virtual plants. A recent paper by Walters et al. (2006) explores the use of an IBM to simulate changes in dispersal patterns under climate change for an insect population, including an assessment of the effects of climate change on long-distance dispersal. Such work indicates the potential of an IBM methodology to be used to increase understanding of the possible impacts of climate change upon pest population dynamics.

Agent-based systems are highly adaptable and flexible, with research in many different fields using the technique at an ever increasing rate. Thus, with a view to understanding the way in which local processes and interactions affect regional population dynamics in a particular agro-ecosystem, agent simulation provides the most useful tool. However, despite a large number of simulations in existence, individual-based simulation in ecology is still a rudimentary science (Grimm and Railsback, 2005). Lessons can be learnt from other fields and software development in computer science can aid the building of a more coherent theoretical and conceptual framework for IBMs. A broad view is taken of the science of 'agent-based simulation' in this thesis, with the aim to draw upon research in a number of disciplines to develop the model.

2.5 Agent-based modelling 'frameworks' and 'toolkits'

Researchers approach ecological modelling from highly diverse backgrounds, with variable experience in software development (if any). Models are becoming increasingly complex, and a shift is occurring towards simulation development that aims to increase our understanding of how ecosystems function from the interactions of the individuals constituting the ecosystem: individual-based modelling. This has created significant problems for model development and credibility, as researchers struggle to master novel simulation techniques, at the same time as trying to make their models easily understandable, transferable and communicable.

The predominance of the development of software from scratch by inexperienced programmers for IBMs in ecology has repercussions for the development of the science as a whole (Grimm and Railsback, 2005). In particular, it has resulted in too much time being dedicated to programming rather than ecological science, errors have not been managed effectively, credibility and usefulness issues have arisen as important model components have been unobservable and untestable and models have not been used to their full potential because of their uniqueness, resulting in difficulties of model communication, sharing and understanding (Grimm and Railsback, 2005).

To aid the development of IBMs, there are many simulation toolkits available. To date, these are primarily used in social science 'agent' simulation research (for an evaluation, see Tobias and Hofmann, 2004). At the present time, the best approach to individual-based modelling and software development in ecology is to make the best possible use of model frameworks, techniques and code libraries already freely available (Grimm and Railsback, 2005).

This section gives an overview of the most popular simulation toolkits currently used by model developers in a range of research fields. It then introduces Repast, a toolkit developed for social scientists, that can be used as an equally successful and useful tool for the development of ecological models.

2.5.1 Definition of frameworks and toolkits

Agent-based modelling frameworks are defined as 'a standardised, general design for implementing a class of models..., a framework provides the overall model structure with the programmer writing new code to fill in the details for a particular model' (Grimm and Railsback, 2005: pp. 268). Frameworks are typically implemented as code libraries or simulation 'toolkits'.

Simulation toolkits are a set of re-useable code (generally in the form of object-oriented classes) that can be customised for a specific model (Grimm and Railsback, 2005). Toolkits for a number of computer languages are quite readily available on the Internet (see table 2.1 for a summary of some popular software). Such toolkits can relieve some of the burden of programming the technical aspects of a simulation for inexperienced programmers, such as simulation control, input and output procedures or GUI development. They can also increase the reliability and efficiency of the model, as the components of the toolkit have been created and optimised by professional software developers.

Name	**Authors**	**Brief description**	**Programming**	**Pros**	**Cons**	**Website**	**Publications**
Swarm	Santa Fe Institute	A set of 'portable libraries' that can be used to help simulate a variety of agent-based models. The modeller can define semantics for agents with the help of the SWARM simulator, so that agents that are created can perform various actions and interaction. Hierarchical scheduling and random number generation makes complex and realistic simulations. There are also ways of monitoring the average states of agents, individual agents or the system entropy.	Objective C or Java interface. Runs on Windows, Linux or Sun	Solid flexible agent framework. Good support.	Requires programming skills. Simplistic agents.	`http://wiki.swarm.org/`	Daniels (1999); Minar et al. (1996); Terna (1998)
Repast	Origins at the University of Chicago. Currently managed by the Repast Organisation for Architecture and Development (ROAD).	Borrows concepts from Swarm, but is differentiated by its multiple pure implementations in several languages and built-in adaptive features such as genetic algorithms and regression. It is also easier to install and use.	Java (Repast J), Microsoft.Net (Repast.Net) and Python Scripting (Repast Py).	Freely available, flexible, well established framework. Good user support network, demos and instructions make it very user-friendly.	Requires programming skills.	`http://repast.sourceforge.net/`	Collier (2003); Collier et al. (2003)
StarLogo	MIT	Based on the 'Logo' language that creates drawings and animations by giving commands to a graphic 'turtle' on the screen. Capacity for the control of thousands of turtles. Patches form an environment, and turtles and patches can interact. Handles density dependent functions and behaviour rules.	Simple scripting language. Runs on any platform.	Simple, easy to use, and can build reasonably complex models.	Limitations in the scripting language limits its range of models, e.g. interaction is minimal	`http://education.mit.edu/starlogo/`	Stevens Colella et al. (2001)
AScape	Brooking Institute	Integrated agent framework that will run models on a variety of topologies. 'Scapes' are built up from collections of agents, and can also themselves act as agents. Behaviours occur across Scapes, determined by 'rules'. AScape manages graphical views and collection of statistics for Scapes, and also mechanisms to control and alter parameters	Java	Nice interface, with good reporting capabilities and realism	For original model development requires a knowledge of Java	`http://www.brook.edu/es/dynamics/models/ascape/default.htm`	`http://jasss.soc.surrey.ac.uk/4/1/5.html`; Parker (2001)
Zeus	Open source (developed by British Telecom	Rule engine, backward chaining planner and visualisation tools are included in a simple graphical environment to build distributed agent systems. Library of pre-written and pre-tested agent components. Various 'editors' enable descriptions of agents, tasks, organisational relationships and strategies. Report and statistics tools in the viewer.	Java	Very flexible, high quality toolkit. Manual and help facilities. Extendable in Java.	Perhaps too flexible: complex to construct models	`http://labs.bt.com/projects/agents/zeus/`	Collis et al. (2000)
AgentSheets	AgentSheets	Commercial software. Flexible with basic user interaction and own script. Simple to build via interface.	None, windows based	Simple, easy to use and run.	Cost $99. Not very extensible.	`http://www.agentsheets.com`	`http://jasss.soc.surrey.ac.uk/3/3/forum/1.html`; Repenning et al. (2000)

Table 2.1: Toolkits and frameworks available for agent simulation development.

2.5.2 Ecological modelling toolkits

There are few software packages that aid model development created explicitly for ecological modelling. Ecological modelling tools in existence tend to comprise programmes created for modelling a specific aspect of species population dynamics, such as population viability, size, growth rate and migration rates by a particular means, and are essentially complete models with the option to vary certain parameters (e.g. GARP `http://www.lifemapper.org/desktopgarp`, LAMARC `http://evolution.gs.washington.edu/lamarc/`).

An example of modelling software that combines population viability analysis with metapopulation dynamics is RAMAS GIS (`http://www.ramas.com`; see also Lindenmayer et al., 1995). It is a more dynamic and flexible form of software tool that covers the same aspects of population dynamics as the models mentioned above, but can also handle GIS-generated maps representing real landscapes to simulate possible future scenarios. This model framework allows input of species-specific parameters, but can only handle one species at a time. It also ignores species-landscape feedback and temporal habitat change. As the framework is fundamentally based on the principle of metapopulation dynamics, it is perhaps constrained by this theory (i.e. it assumes *a priori* certain population-level movement patterns and patch dynamics). It certainly does not allow for simulations based upon individual movements.

2.5.3 Generic toolkits

At this point in time, many agent-based simulation toolkits are in the early stages of development and very few can be termed 'user friendly'. Repast, 'Recursive Porus Agent Simulation Toolkit' (`http://repast.sourceforge.net/`) is a particular toolkit that has received high credentials from social science users and reviewers (e.g. Tobias and Hofmann, 2004). It is based on the well-established Swarm toolkit (`http://wiki.swarm.org/`), developed for over ten years to date by the Santa Fe Institute.

Swarm

Swarm is a freely-available, pioneering, interdisciplinary software package for multi-agent simulation of complex systems. Its popularity is reflected by a large and dynamic Swarm user community and an annual conference. However, the choice of the authors to construct the software in Objective C, a relatively obscure computer language, could be said to have somewhat hampered its uptake in the modelling community. Swarm developer Paul Johnson remarks in the website Frequently Asked Questions (FAQs) "If I were to start today, I'd use Java", but at the start of the development of Swarm in 1994 this was not an option. Today Swarm has a Java programming interface to the Objective C libraries, but this lacks flexibility and requires awkward compilation to native code. However, several recent toolkits have been developed that borrow much from the Swarm simulation toolkit yet are developed in a more accessible manner (e.g. Gecko `http://www.cbc.yale.edu/old/cce/gecko.html`, ASCAPE `http://www.brook.edu/es/dynamics/models/ascape/default.htm`and Repast).

Repast

Repast is an agent-based modelling toolkit developed using the object-oriented programming language Java (`http://java.sun.com`). Object-oriented programming (OOP) languages such as Java and C++ have facilitated the development of agent-based models greatly. This is because they reflect the structure of the system being modelled (for a discussion of OOP in comparison to procedural and other approaches in this context see Grimm and Railsback, 2005: chapter 8).

Repast has many key advantages that have shown it to 'win out over the other Java based programming libraries' (Tobias and Hofmann, 2004). According to Tobias and Hofmann (2004), the key features of Repast that make it very attractive are:

- It is released under the open-source BSD licence and thus can be freely downloaded with full source code.
- It is extensively documented, with 'How to' documents, FAQs online and a popular newsgroup that answer questions promptly. There are also many free publications detailing research that uses the toolkit available on the website.
- The user base is medium and growing rapidly. There is a good website and many projects that use the toolkit.
- The software is constantly being developed and extended, and is assured for the next five years.
- There are a number of tools that make visualisation and data editing simple.
- Simple stepping methods can be extended and made more dynamic by the user.
- Control and recording of the model are highly advanced compared to other toolkits. There are useful classes that can be implemented to create Quicktime movies and real-time graphs can be drawn as the simulation runs. The display can also be made probe-able and zoomable.
- The GUI is quite easily constructed, and makes adding parameters and starting the simulation very straightforward.
- Repast puts few limitations on the possibilities for agents.

This user-friendliness has meant that although the framework is designed with the social scientist in mind, it is highly transferable to other disciplines. Interesting future developments planned include interoperability with Geographical Information Systems (already there are classes that simplify reading in ASCII files that represent the environment) and a distributed computing capability, that will allow parallel computing of simulations.

Applications and existing simulations

Although Repast is primarily a social simulation toolkit, there are several examples of how it has been used to simulate ecological systems and landscape change. These include agent-based modelling of "Spatial Land Use Change and Ecological Effects at the Rural-Urban Interface" (SLUCE) (`http://cscs.umich.edu/research/projects/sluce`) and the simulation of freshwater fish population dynamics (Dumoulin, 2004).

2.5.4 Potential drawbacks

Individual-based simulation development is not simple by its very nature. There is often a conflict between the desire to analyse system complexity in detail and the computational limitations, imposed by either the ability of the researcher or the hardware available. Much experimentation is usually needed and it is important to start with the simplest model and only then to add levels of complexity gradually to the simulation (Inchiosa and Parker, 2002). For an ecologist to use a simulation toolkit effectively, it is necessary to understand a classic programming language such as Java or C++ (in which many of the available toolkits are written), and to learn the principles of object-oriented programming. It is also important that the toolkit is fully understood, and the software is constantly tested and debugged. Repast provides some of the best support and documentation for any modelling toolkit available; however, even Repast's documentation is not particularly comprehensive. This can make it rather opaque at times to understand precisely how some methods function.

Overall, the potential drawbacks of using a simulation toolkit are far outweighed by the benefits discussed above, and on the basis of this Repast was chosen to develop the model for this thesis.

2.6 Aphid specific models

2.6.1 Early aphid simulations

Over the past forty years there have been a large number of simulation models constructed to investigate the problems of controlling aphids. In particular, these studies set out to simulate the interaction of pest species with crops and with natural enemies, with what has been referred to as a 'tri-trophic system view of arable farming' (Gosselke et al., 2001).

Early approaches were based upon the 'variable life-table' approach, pioneered in Australia by Hughes and Gilbert (1968). This approach focused on the cabbage aphid, *Brevicoryne brassicae* (L.), both in the original work by Hughes and Gilbert (1968) and in subsequent simulation models (Gilbert and Hughes, 1971). Several similar models were constructed to simulate *Sitobion avenae* (L.) in a number of Western European countries, including the Netherlands, France, the UK and Germany throughout the 1980s (Carter et al., 1982; Carter and Rabbinge, 1980; Mann and Wratten, 1987; Pierre and Dedryver, 1984; Rabbinge et al., 1979; Vorley and Wratten, 1985). These mathematical models are mainly deterministic, analytical models, based upon equations representing relationships between organisms and the key driving variable, temperature, derived from laboratory and field data. These models are all aspatial, modelling population dynamics over time at the scale of a single tiller. Most of these models function at a daily time-step.

One of the earliest models of *Rhopalosiphum padi* was developed in Sweden by Wiktelius and Pettersson (1985). This was used to examine the likely effects of plant resistance on aphid numbers. However, this model suffers from a lack of quantitative validation (Morgan, 1990). Later models, such as those constructed by Morgan (1990) and Ekbom et al. (1992), are more robust, having undergone field validation as well as sensitivity analysis. Morgan (1990) dealt with population dynamics during the autumn in the UK that had not previously been simulated effectively despite this being the most important time for the spread of Barley Yellow Dwarf Virus (BYDV) by *R. padi* populations in the UK.

2.6.2 The most recent aphid simulations

Table 2.6.3 gives a brief overview of the qualities of the most recent (post-1990) aphid population dynamic models for the species *R. padi* and *S. avenae*. Some of these are based on earlier models mentioned above, and most are deterministic and aspatial. They also tend to estimate population dynamics based upon parameters quantified in the laboratory, without taking into account environmental conditions experienced by the aphids in the field. Only Fabre et al. (2006) attempts to account for the effects of environmental variability and potential feedback loops, through the use of Bayesian inference to aid parameter estimation.

Previous work by Hopkins and Morse (1996) used a model based upon that developed by Morgan (1990) to construct a simple model of the spread of BYDV in a barley field, explicitly considering individual plants and aphids in a spatial context. This model was developed primarily as a demonstration of computing techniques, to improve simulation performance, in particular the use of data structures for efficient storage of individual aphid data using Fortran 90. With large numbers of individuals, storage and processing become key to the viability of models. More flexible storage is now readily available in languages such as Java (for example ArrayLists); however, results from this model indicated at an early stage that distributed parallel computing or other technological solutions would have to be employed to cope with individual-based simulations of such systems. The technological solutions to improve memory availability and simulation performance of the model in this thesis are detailed in chapter 6.

2.6.3 Evaluation and research objectives for future aphid models

Despite the development of aphid models for over forty years, it has been shown that there has been little change from the original phenomenological, deterministic methodology (table 2.6.3). These models have limited predictive capacity, as predictions may only be possible for the range of conditions over which the model was parameterised. A behavioural modelling approach allows for greater flexibility for predictions in novel environments, since current models make no assumptions about demographic responses to changing environments: demographic responses result from individual responses to localised change, such as density dependent feedback. In addition, in order to predict population response to changing landscapes, it is necessary to link demography to resource availability. Models need to be spatially explicit, and able to adapt to changing environments, as well as to novel landscape configurations.

2.7 Background summary

Primarily, this research is driven by a need to understand pest population dynamics in the context of changing agricultural landscapes through ecological simulation. The development of reliable, flexible modelling tools to predict trends in pest population dynamics is critical to aid the development of pest control programs at the landscape scale.

Complexity can be introduced into all models of population dynamics by considering heterogeneous landscapes, competition, knowledge of habitat suitability and site-dependent mortality. Such an addition will enhance a basic understanding of population dynamics (Wiegand et al., 1999). For 'generalist' pest species populations, this is best done without the preconceptions of the metapopulation paradigm, but by using an individual-based approach that takes local population dynamics into account in a continuous landscape.

The background to individual- and agent-based systems has been presented, enabling an assessment of the suitability of the paradigm to model complex ecological systems, in comparison with more established methods. This thesis builds upon previous methodologies in the simulation of aphid pest populations (particularly the approach of Morgan, 1990), introducing a novel individual-based (agent-based), spatially explicit approach to the field.

CHAPTER 3

Data

3.1 Insect data

3.1.1 Aphid pests

Cereal aphids were chosen as the organism with which this model would be developed for a number of reasons. Firstly, aphids have been a significant agricultural pest in the UK since the late 1960s (Dean, 1973a). Yield losses have been caused by both direct feeding damage and also by the transmission of plant pathogenic viral diseases. Since this time there has been a huge amount of research, data collection (both within the laboratory and in the field), and simulation effort to forecast the timing and severity of outbreaks in the UK.

Due to this large amount of research, a second reason to model aphid populations was the availability of data with which to parameterise and also validate the model, as well as the availability of existing models that could form the basis for the model as well as provide a useful comparison.

Thirdly, the behaviour of aphids can be simplified, despite the complex morphologies that develop, lifestage transitions and the influence of the wind on movement. It has been found by a number of studies that temperature is a key driver for mortality as well as other processes such as development (Dean, 1974a,b; Williams, 1980).

Cereal aphid numbers are generally highest in the spring and summer. In autumn, there is a second wave of migrant alates in some species. *Rhopalosiphum padi* tends to migrate to its winter host at this time (the bird-cherry tree) and *Sitobion avenae* overwinters holocyclically on grass. As climate is getting warmer however, more *R. padi* are overwintering in the field as vivparous apterae or alatae, particularly in the South of England.

3.1.2 Insect pest monitoring

National monitoring schemes for invertebrates are largely limited to certain species in the UK, especially Lepidoptera and aphids. For Lepidoptera, species that occur on farmland (generalists) tend to be sustained, but there are possible declines in species with more restricted distributions (specialists), for example species found on unimproved grassland (Robinson and Sutherland, 2002). However, key Lepidoptera surveys, such as the Butterfly Monitoring Scheme (BMS) (Asher et al., 2001), as well as other invertebrate surveys, are often biased to the species' prime habitat, ignoring farmland populations.

Aphid population distribution and monitoring

Aphid monitoring in the UK is part of a wider, international monitoring scheme called 'Examine'. The Rothamsted Insect Survey are responsible for the UK monitoring, with a network of sixteen 12ft suction traps (figure 3.1). The traps have been running for many years (the one at Rothamsted began in 1965). The traps are emptied daily during the 'aphid season' and the aphids identified to species in most cases. Each trap is representative of what is flying over an area of radius approximately 80 km, but there is considerable local variation in aphid density at ground level. The data are used for fundamental studies on factors affecting the dynamics of aphid populations and to provide sponsors with information that aids aphid control decisions. They have also been interpolated to map aerial aphid abundance at the national scale (Woiwod and Tatchell, 1984).

Cereal aphid populations in the UK landscape as a whole did not change much between 1969 and 1988, and a few species increased (Woiwod, 1991). Species that significantly increased in this time period include *S. avenae* ($P<0.001$ Woiwod, 1991) and *R. padi* exhibited a similar upward trend (figures 3.2 and 3.3). However, studies that focus on aphid numbers within cereal fields show that arable management change has greatly reduced the number of aphids within fields since 1970 (Robinson and Sutherland, 2002). There is therefore a complex relationship between local and national trends, where national changes in cropping such as a shift from spring-sown to winter-sown cereal may increase national aphid abundance, but have varying impact in the context of the phenological relationships between the crop and the pests, so that increases may not be observed in all species in all regions (Woiwod, 1991).

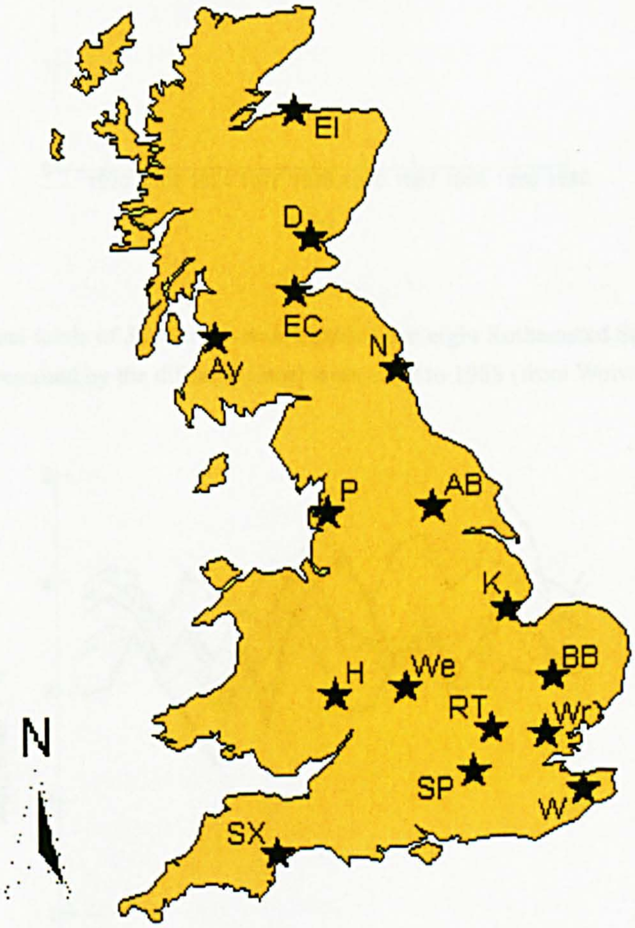


Figure 3.1: Current suction trap sites in the UK: El - Elgin, D - Dundee, EC - East Craigs, Ay - Ayr, N - Newcastle, P - Preston, AB - Askham Bryan, K - Kirton, BB - Broom's Barn, We - Wellesbourne, H - Hereford, RT - Rothamsted Tower, Wr - Writtle, SP - Silwood Park, W - Wye, SX - Starcross.

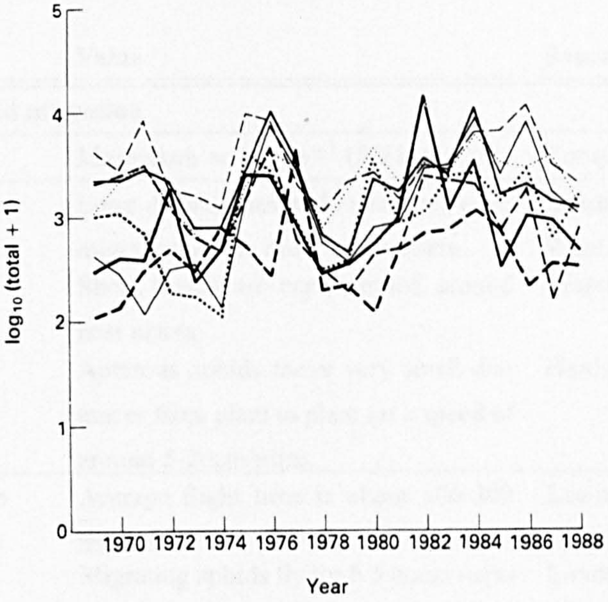


Figure 3.2: Annual totals of *Sitobion avenae* aphids from eight Rothamsted Survey suction traps (represented by the different lines) from 1969 to 1988 (from Woiwod, 1991).

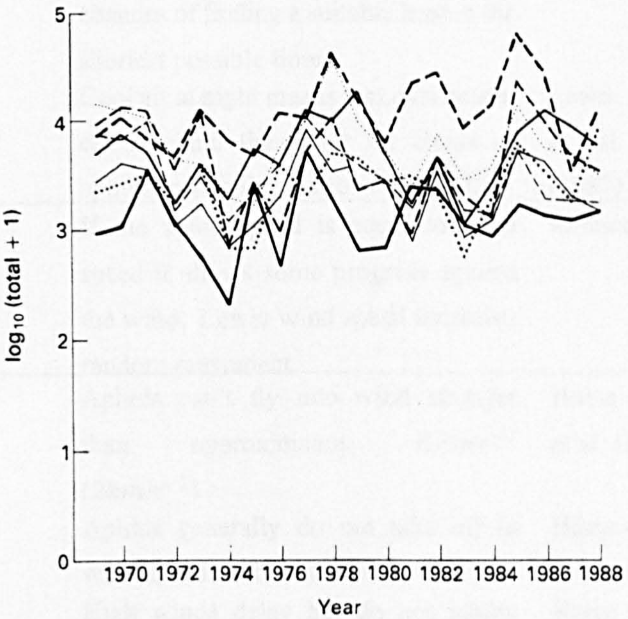


Figure 3.3: Annual totals of *Rhopalosiphum padi* aphids from eight Rothamsted Survey suction traps (represented by the different lines) from 1969 to 1988 (from Woiwod, 1991).

The following summaries of the biology of *R. padi* and *S. avenae* are based upon information given in the Rothamsted Insect Survey Aphid Encyclopaedia entries for *R. padi* and *S. avenae*, `http://www.rothamsted.ac.uk/insect-survey`. Table 3.1 gives key information of aphid behaviour observed both in the laboratory and in the field, that forms a basis for the parameterisation of the model in a biologically realistic manner.

Parameter	**Value**	**Source**
Movement and migration		
Flight speed	Maximum = $0.9ms^{-1}$ ($3.24kmh^{-1}$).	Compton (2002)
Flight distance	Long distance migration is rare, most migration of the order of 20-50km.	Loxdale et al. (1993); Ward et al. (1998)
	Short flights are concentrated around host plants.	Kennedy et al. (1959)
	Apterous aphids move very small distances from plant to plant (at a speed of around 5-20cm/min).	Hardie (1994)
Flight duration	Average flight time is about 100-200 minutes.	Lewis (1965)
	Migrating aphids fly for 6.5 hours maximum.	Loxdale et al. (1993)
Flight phases	Two kinds of flight: Foraging flight at low wind speed and long distance flight (> 20km), which is much rarer. Aphids usually try to maximise their chances of finding a suitable host in the shortest possible time.	Loxdale et al. (1993); Ward et al. (1998).
	Cool air at night means that over southern England the upper air clears of aphids (probably true for whole UK).	Lewis (1965); Loxdale et al. (1993); Robert (1987)
Flight path	If the aphid speed is equal to wind speed it allows some progress against the wind. Lower wind speed increases random movement.	Kennedy et al. (1959)
Wind effects	Aphids can't fly into wind stronger than approximately $0.5ms^{-1}$ ($2kmh^{-1}$).	Haine (1955); Loxdale et al. (1993)
	Aphids generally do not take off in wind speeds above $8kmh^{-1}$.	Haine (1955)
	High winds delay but do not inhibit take-off.	Haine (1955); Walters and Dixon (1984)
Boundary layer	Above 1m from the ground aphid movement is controlled by the wind.	Broadbent (1948); Campbell and Ridout (2001); Compton (2002); Taylor (1974)
Aphids and landscape features		

Table 3.1 – continued from previous page

Parameter	Value	Source
Altitude, latitude and longitude	Altitude is not significant in aphid incidence, but longitude and latitude are (higher incidence has been observed in the west of mainland Britain, and at lower latitudes: greatest incidence is found at about 50 degrees North with a second peak at 54 degrees North).	Foster et al. (2004)
Land use (farmland)	*Rhopalosiphum padi* prefer Barley, Wheat and improved grassland.	Dean (1973b); Leather et al. (1989); Mann et al. (1996)
Land use (non-farmland)	Higher aphid incidence found on non-arable grassland, moorland, woodland and uncultivated/wasteland. Also next to set-aside and in the absence of hedgerows. Disturbed land, indicated by buildings and main roads, is also associated with higher incidence of aphids.	Foster et al. (2004); Leather and Dixon (1982)
Other factors (excluded from the simulation)		
Field size	Aphid densities highest in fields between 2 and 4 ha. (small and large fields have lowest incidence).	Foster et al. (2004)
Seasonal variation	*R. Padi* have three seasons of flight activity: Spring, summer and autumn. In England, Spring = week 17-25, Summer = week 24-33.5, Autumn = week 32.5-50 from January 1st.	Wiktelius (1981)
	The temperature threshold for flight varies with season for R.Padi: SPRING: min temp = 14°C; 50% takeoff = 16-17°C; 100% takeoff = 13-14°C SUMMER: min temp = 11°C; 50% takeoff = 13-14°C; 100% takeoff = 15°C AUTUMN: min temp = 3°C; 50% takeoff = 9-10°C; 100% takeoff = 13°C.	Walters and Dixon (1984); Wiktelius (1981)

Table 3.1 – continued from previous page

Parameter	Value	Source
Long distance flight termination	Renewed visual responsiveness to plant-related wavelengths occurs, especially to yellow. Aphids are attracted to wavelengths $> 500\text{m}\mu$, especially yellow, also green and orange. Actively bypass blue to ultra violet spectrum. Tested in the field, but some species are less sensitive.	Compton (2002); Kennedy et al. (1961); Loxdale et al. (1993); Robert (1987)
	Host density is unimportant to aphid decision to land.	Ward et al. (1998)
	Precipitation may wash some aphids from the air.	Morgan (2000)

Table 3.1: Key information on aphid behaviour, including source.

3.1.3 *Rhopalosiphum padi*

Host plants/Life cycle

R. padi is a host alternating aphid species (figure 3.4). The eggs of this aphid overwinter on bird cherry trees, *Prunus padus*, and hatch from April onwards. After several generations, winged forms are produced which migrate in May/early June to numerous species of Gramineae, including all the major cereals and pasture grasses. After several generations are produced, once summer cereals ripen, the species migrates away to wild grasses, particularly in eastern England. As day length decreases in autumn alates migrate back to the primary host *P. padus* to lay eggs between September and November. In this species the last migration is usually the biggest of the year.

However, as winters become warmer and cereals are winter sown, a significant proportion of *R. padi* populations in the Midlands and the South overwinter as mobile stages on early sown winter cereals and grasses (such as winter wheat).

Pest status/damage

R. padi is a major pest on wheat, barley, oats and maize, and a minor pest on rye. Direct feeding damage is rare due to low numbers, only occurring in some years where heavy infestations

Figure 3.4: Photograph of *Rhopalosiphum padi*.

develop on maize/sweetcorn in late summer. However, as a key vector of Barley Yellow Dwarf Virus (BYDV) it has achieved pest status. *R. padi* may provide the primary source of BYDV infection in early sown winter cereals, with possible over-wintering in the field as further generations develop from autumn migrants under mild conditions.

3.1.4 *Sitobion avenae*

Host plants/Life cycle

S. avenae overwinters in the field, thus spending its entire year on cereals and grasses (figure 3.5). Mostly this is as mobile stages on wild grasses or winter cereals, which then develop rapidly in warm springs. A small proportion overwinters as eggs on Gramineae. Apterous aphids develop first, then alate forms usually develop and fly in late May/June. Colonies rarely become dense before late June. In continuing hot dry conditions, these colonies can increase quickly. Alates continue to be produced throughout the summer in response to increasing population density and declining food quality, moving to re-infest crops or other grasses. It is from these that a comparatively small autumn migration arises, which infest early sown winter cereals as well as wild grasses.

Figure 3.5: Photograph of *Sitobion avenae*.

Pest status/damage

S. avenae is a major pest on wheat, a moderate pest on barley and oats, and a minor pest on maize. The species causes direct feeding damage through May, June and early July. It also has pest status in winter-sown cereals in September/October, and throughout mild winters, as a virus vector of BYDV. It is more cold-hardy than *R. padi*, and thus more significant in the secondary spread of BYDV in winter cereals.

3.1.5 Model parameters

Algorithms are used as a basis for the model to describe the individual aphid's development, temperature dependent survival, reproduction and morph determination in the model, as well as overall immigration rates (for details see chapter 4 and for equations see appendix A.4). In addition, stochastic rules are applied to simulate the individual's behaviour, especially in relation to its environment (further detail on the landscape data is given in section 3.2). These rules are based upon a wide range of literature sources (table 3.1), a number of which relate to several studies across a number of years to verify the parameters used. Rules are derived from the literature to model the aphid's movement and migration behaviour (including their response to wind) and their survival in relation to the environment. Importantly, light and temperature in England after dusk tends to drop below the threshold for flight activity; thus, it is a fair assumption that few aphids are able to fly at night (Johnson, 1962; Loxdale et al., 1993). The use of this data to parameterise the model and formulate model rules is detailed in chapter 4, section 4.3.2.

The wind model used in the scenarios and sensitivity tests for the thesis is not the full wind model developed and included in the code (see appendix E.2). It has been simplified, so that wind direction is constant or wind has been removed altogether, in order to focus on the effects of the landscape alone. However, the full wind processes that can be used for the model are described in detail in chapter 4, and would be valuable in further studies relating explicitly to migration of the aphids by wind.

3.1.6 Evaluation of the insect data

The model presented in this thesis uses existing biological knowledge, much of which is summarised in table 3.1 to construct a spatially explicit individual-based model (IBM) of aphid population dynamics in agricultural landscapes, that can be used to explore the relationships between aphid populations and landscape features.

Data that forms the basis for the behaviour rules in the model (table 3.1) is from a highly diverse range of studies that have taken place both in the field and the laboratory, across a number of decades. However, this also means that many of the key values, such as the height of the hypothetical 'boundary' layer for aphid take-off (Broadbent, 1948; Campbell and Ridout, 2001; Compton, 2002; Taylor, 1974), have been determined (and thus verified) to be a similar value by a number of studies. Therefore, there is reasonable confidence in the value of the majority of the parameters that are derived from the literature due to such replication of studies.

However, it should be noted that many studies into aphid migration have taken place under artificial conditions in a laboratory where flight is likely to be sustained for longer; 'normal' aphid flight behaviour 'maximizes their chances of finding a suitable host(s) in the shortest possible time' (Loxdale et al., 1993: pp. 305). Therefore, parameters used to estimate flight distances in this model may over-estimate aphid movement somewhat.

3.2 Landscape Data

3.2.1 Theoretical data

The creation of artificial 'neutral' landscapes in ecology is a well established research methodology. A number of theoretical ecological studies use such landscapes as input to simulation models (e.g. Keitt and Johnson, 1995; Moloney and Levin, 1996; Palmer, 1992), some with particular reference to the simulation of the spread of invasive species in heterogeneous landscapes (Ming and Albrecht, 2004; With, 2004).

This thesis uses a fragmented landscape of favourable, marginal and unfavourable habitat. The generation of a fragmented landscape of a given habitat cover percentage is based upon a methodology described in Tischendorf and Fahrig (2000), based upon ideas adopted from percolation theory (Gardner et al., 1987). This is a simple, three stage process, as shown in figure 3.6. The Java code for the generation of a fragmented landscape is given in appendix C.

3.2.2 Real landscape data

The regions most likely to be affected by alterations to crop and fallow patterns are the South-East, East Midlands, East of England and Yorkshire and Humberside (Silcock et al., 2004). Therefore a study using realistic landscapes is focused in this region. An area close to Rothamsted, Hertfordshire, was chosen, as this is where field data were gathered and used to help validate the model (section 3.3). The study region was 5×5 km, and the landscape grid derived from an ASCII raster taken from a Land Cover Map 2000 (LCM2000) dataset of Hertfordshire, England (origin $51°51'12''N, 0°19'3''W$), with cells of dimension 25 × 25m. This dataset was categorised into areas of favourable, marginal and unfavourable habitat, based on information available in the literature (table 3.2).

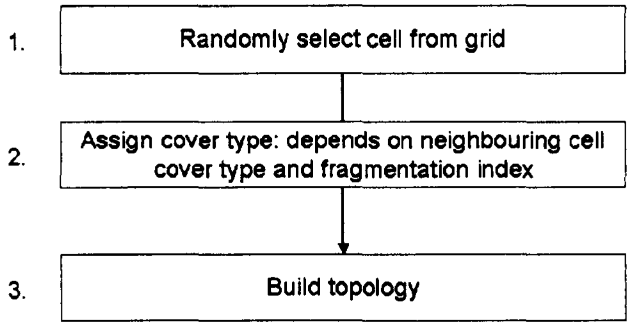


Figure 3.6: Algorithm to generate artificial landscapes on a grid.

3.2.3 Evaluation of the landscape data

LCM2000 was chosen, as it is closest to providing a source of 'habitat' data, although it does not describe each landscape class in detail or indicate habitat quality. LCM2000 `http://www.ceh.ac.uk/sections/seo/lcm2000_home.html` is derived from a computer classification of satellite scenes, obtained mainly from Landsat satellites (Fuller et al., 2002). The datset used was a 25m ASCII raster dataset compiled by the Centre for Ecology and Hydrology (CEH), derived from the original vector dataset. LCM2000 was calibrated to a field survey undertaken by Countryside Survey 2000 (CS2000) `http://www.cs2000.org.uk` (Haines-Young et al., 2000), that provides some ground-truthing and error estimation for the dataset. In general, it was found that the structure of the landscape is well represented by LCM2000, with discrepancies between the data occurring when the lower spatial resolution of LCM2000 affects the detail.

With particular reference to agricultural land, arable and horticultural land covers just over 23.4% of the UK according to LCM2000 and 21.5% by CS2000 field survey estimates. LCM2000s higher estimate relates in part to small features such as woodlands, prevalent in arable landscapes but generalised out by LCM2000. About 70% of LCM2000 arable and horticultural land is coincident with the CEH field survey 'arable' class. Some confusion also occurred when the field survey and the satellite survey took place in different years (as rotation farming may cause land use to change slightly, from arable to horticultural or improved grassland). There is also some confusion between arable and built up land, as erroneous classification of satellite images sometimes occurs where part-grown or ripening crops have spectral signatures readily confused with those of partly vegetated suburban areas (Fuller et al., 2000).

Habitat category	LCM2000 class	Source
Favourable	Arable and Horticultural: barley	Leather et al. (1989)
	Arable and Horticultural: wheat	Dean (1973b)
	Improved grassland: intensive	Mann et al. (1996)
Marginal	Improved grassland: grass setaside	Leather and Dixon (1982)
	Calcareous grassland: managed	Mann et al. (1996)
	Calcareous grassland: rough	Leather and Dixon (1982)
Unfavourable	Broad-leaved woodland: deciduous	
	Broad-leaved woodland: mixed	
	Coniferous woodland: conifers	
	Arable and Horticultural: bare ground	
	Arable and Horticultural: field beans	
	Arable and Horticultural: horticulture	
	Arable and Horticultural: linseed	
	Arable and Horticultural: potatoes	
	Arable and Horticultural: peas	
	Arable and Horticultural: oilseed rape	
	Arable and Horticultural: sugarbeet	
	Arable and Horticultural: unknown	
	Arable and horticultural: setaside (bare)	
	Built-up areas, gardens: suburban/rural development	
	Built-up areas, gardens: urban residential/commercial	
	Built-up areas, gardens: urban industrial	

Table 3.2: Habitat category and values associated with it, as derived from LCM2000 in relation to the literature.

3.3 Field data

Field data with which to compare the model as part of the validation process was gathered during the autumns and winters of 1985/6, 1986/7 and 1987/8 at Rothamsted by Derek Morgan (Morgan, 1990). The data used are aphid population densities surveyed on winter barley, cv Igri, without insecticide application. The total number of aphids were counted by observation on four 0.5m lengths of crop row within two unsprayed plots. Observations began in September for each year at weekly intervals, until no aphids were found in the crop.

Four species of aphid were found in the samples, the most common species was *Rhopalosiphum padi* in the first two years, and *Sitobion avenae* was the most common in the final year. The other species were *Rhopalosiphum maidis* and *Rhopalosiphum insertum*. Immigration occurred during September and October. The largest number of *R. padi* caught in the suction trap was in 1985, and the least in 1987 (Morgan, 1990). Contemporary suction trap data were obtained from the Rothamsted Insect Survey.

3.3.1 Evaluation of the field data

This field dataset, although surveyed over several years, is limited spatially. Although suitable for temporal comparisons (when scaled appropriately), it does not provide a spatial comparison with which to compare output from the model simulation runs. The compilation of such a spatial dataset is considered beyond the scope of this thesis in terms of time and resources (see chapter 5), but is a consideration for the further development of this work and future studies. For example, survey techniques for the development of multi-scale spatial datasets have recently been developed over a period of several years by researchers in Germany (e.g. Clough et al., 2005; Thies et al., 2005).

CHAPTER 4

Model Development

4.1 Development stages

This chapter is the first of the three 'core model development' chapters outlined in the thesis introduction (chapter 1). This chapter describes the stages in the conceptual and technical development of the model, including the use of the Repast toolkit. It also shows how the model has been constructed, details the parameters and algorithms used, and describes how the model has been tested and validated. This relates to the third objective of the thesis, demonstrating the use of IBM techniques to build a model to simulate the spatial and temporal population dynamics of *Rhopalosiphum padi*, an aphid pest of cereals, in agricultural landscapes.

There are five key stages in simulation development:

1. Define the aims of the simulation
2. Develop a framework
3. Build up the framework with appropriate classes and code
4. Test the simulation and validation
5. Develop further with more advanced features

This is a non-linear development strategy: for example, following testing there can be a return to an earlier stage of the development process. The multi-agent simulation toolkit Repast plays an important role in this simulation development. The toolkit aids model development at stage 2, by providing an existing framework designed to simplify the creation of an agent-based simulation (see chapter 2, section 2.5.3). It also aids the model development at stage 3, with the library of Java classes that simplify complex actions in a model (for example extracting data from an ASCII file) (see section 4.4). Repast simplifies the addition of more advanced features at stage 5, including batch simulation or movie capture.

4.2 The aims of the simulation

The primary aim of the simulation is to provide a tool that will meet the aims and objectives of the thesis, as outlined in chapter 1. In order to do this, the specific aims of the simulation were to:

1. Build a model that represents the population dynamics and behaviour of a species of aphid.
2. Enable the model to function at differing landscape scales.
3. Demonstrate the use of the model in both theoretical and real landscapes.
4. Investigate the creation of a generic model applicable to many different species of insect across a wide range of scales.
5. Apply the model to study and make recommendations for the management of insect pest population outbreaks, both spatially and temporally, as stated in the overall objectives of this thesis.

4.3 Model framework

4.3.1 Conceptual framework

Several aims of the simulation have been achieved by the use of individual-based (agent-based) modelling techniques (see chapter 2 for background). Within a Multi-agent system (MAS), individual agents (insects) can be created and assigned individual rule sets (characteristics and behaviour). Interaction amongst agents, and between the agents and the landscape, can be built into the model. The individual-based model (IBM) emulates as closely as possible the biological parameters and rules that govern the behaviour of individual insects. The literature has provided detailed information on these parameters and rules, as detailed in chapter 3. Overall, an agent methodology is particularly suitable for this application for the following reasons:

- Flexibility. Different rule sets can be assigned to different insects, subspecies, and individuals; new rules are easily integrated into the model.
- The model structure reflects ecosystem structure, and the real behaviour of organisms.
- The agent framework allows for interactions at different temporal and spatial scales and the simulation of emergent ecosystem properties.
- It has been possible to parameterise the model with extensive, biological information derived from individual-scale studies available in the literature. Such parameters can often be derived from multiple studies, giving greater confidence.

- Individual life histories can be examined and contribute to the model in a way not possible in statistical models.

When compared to more traditional techniques such as differential equations used to simulate population dynamics in ecology, these factors illustrate that an individual-based approach is a powerful technique to use. For example, agent simulation makes it possible to examine the impact of individual, unique, behaviours at different geographical and temporal scales (see chapter 2 for a full comparison).

Temporal and spatial scale

R. padi is a major pest of cereal crops in many parts of the world, particularly as a vector of the causative agents (Cereal yellow dwarf virus and Barley yellow dwarf virus) of the disease Barley Yellow Dwarf Virus (BYDV). In Southern England, the virus has most impact in cereals sown early in autumn (Morgan, 2000). Thus the simulations are run from when the species is observed in suction traps in September until the population dies out about six months later (see also chapter 3 section 3.1). The model is run on a daily time-step, as this is a sufficient timeframe within which to update all the key parameters.

The model is designed to be flexible enough that it can potentially be used to simulate population dynamics across a region of any spatial extent, with grid cells of any resolution. However, most of the simulations have been run for a region 5 × 5km, with grid resolution 25 × 25m. The extent has been chosen as this is a scale at which it is possible to meaningfully isolate the source of an aphid outbreak (such as a field) and to examine the spatial dynamics across the immediate region. Alate aphids generally move a maximum distance of a few kilometres, with long-distance migration the exception rather than the rule. However, the model could be used in future to look at occasional long-distance migration of aphids from a source.

The grid cell size of 25 × 25m is used as this is the resolution of Land Cover Map 2000 (LCM2000), the source of the landscape data (see chapter 3). The implication of using this cell size is that the population density of aphids is assumed to be uniform across this area. The effect this assumption has upon the simulation is shown in chapter 5, section 5.5.1.

4.3.2 Model rules

The full model code and documentation is included with this thesis on a separate CD, see appendix E.

At each daily time-step of the aphid simulation, a sequence of events takes place in the model (figure 4.1) and for each individual aphid within the model (figure 4.2). These events are organised according to the structure of the Repast framework: beginning the simulation with methods 'setup' and 'buildModel', followed by 'preStep', 'step' and 'postStep' methods within which the aphid model code is positioned. The inputs were habitat data, daily minimum temperature, maximum temperature and mean temperature, wind speed and wind direction. Wind direction can be varied randomly; however, a single value of 270° representing the prevailing westerly wind has been used predominantly in this thesis, and wind speed was varied quasi-randomly according to an equation derived from BADC data recorded at Rothamsted (see appendix A.4, section A.4). More detail on the values and sources of these inputs can be found in chapter 3. Data may also be input on when crops are sown and which fields, or when pesticide is sprayed and which fields, as files. Classes that represent a species of insect structure the model, each implementing an 'Insect' interface. Information about any Insect agent includes a unique ID tag for the agent, the agent's 'age' as a proportion of adult development (where nymphs become adult at 1.00) and the agent's position in three-dimensional space. In addition, for aphid agents, information on whether or not the agent has undergone migration and the agent's morphology (alate or apterous) is also recorded.

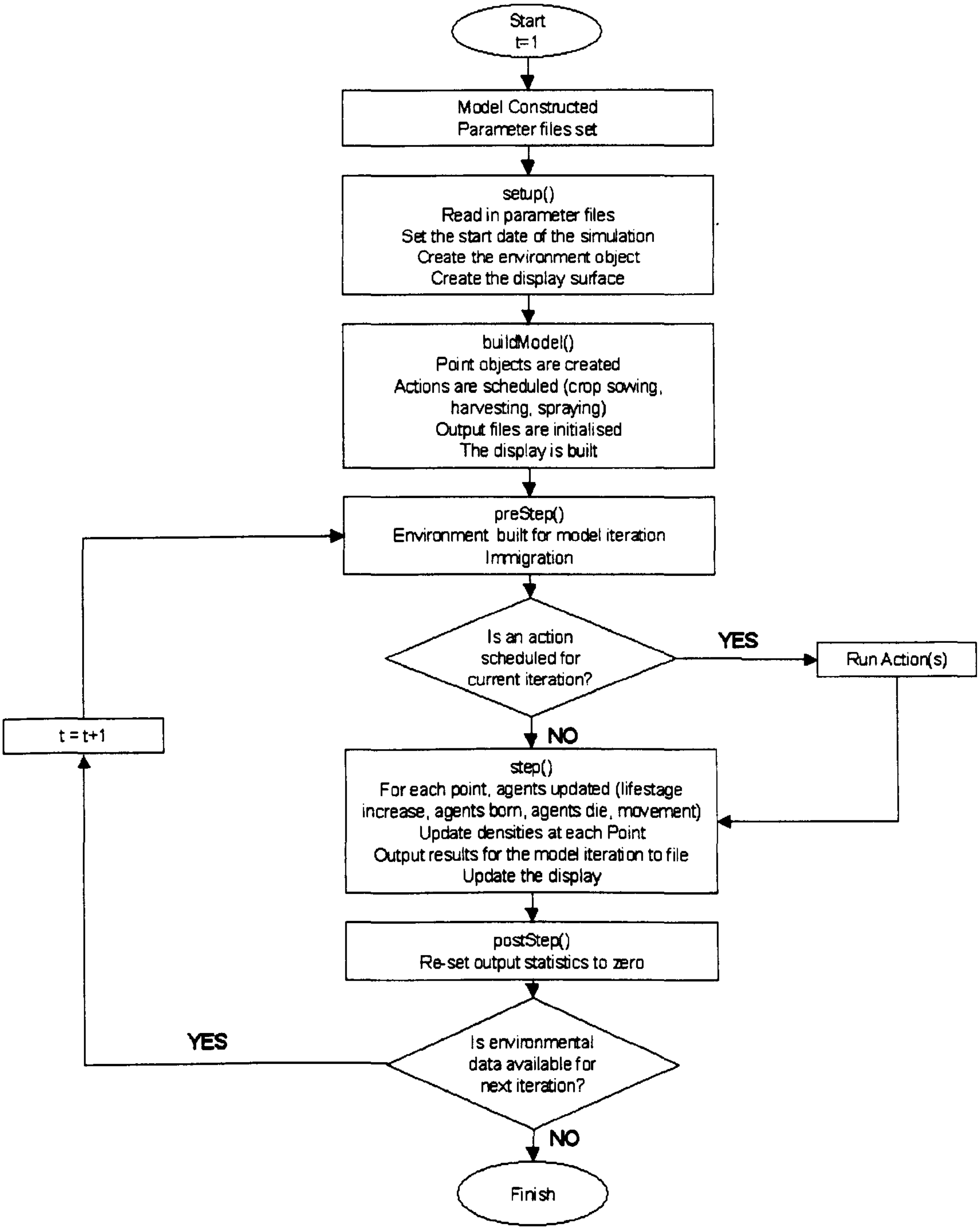


Figure 4.1: Flowchart illustrating the operation of rules at each stage of a model run.

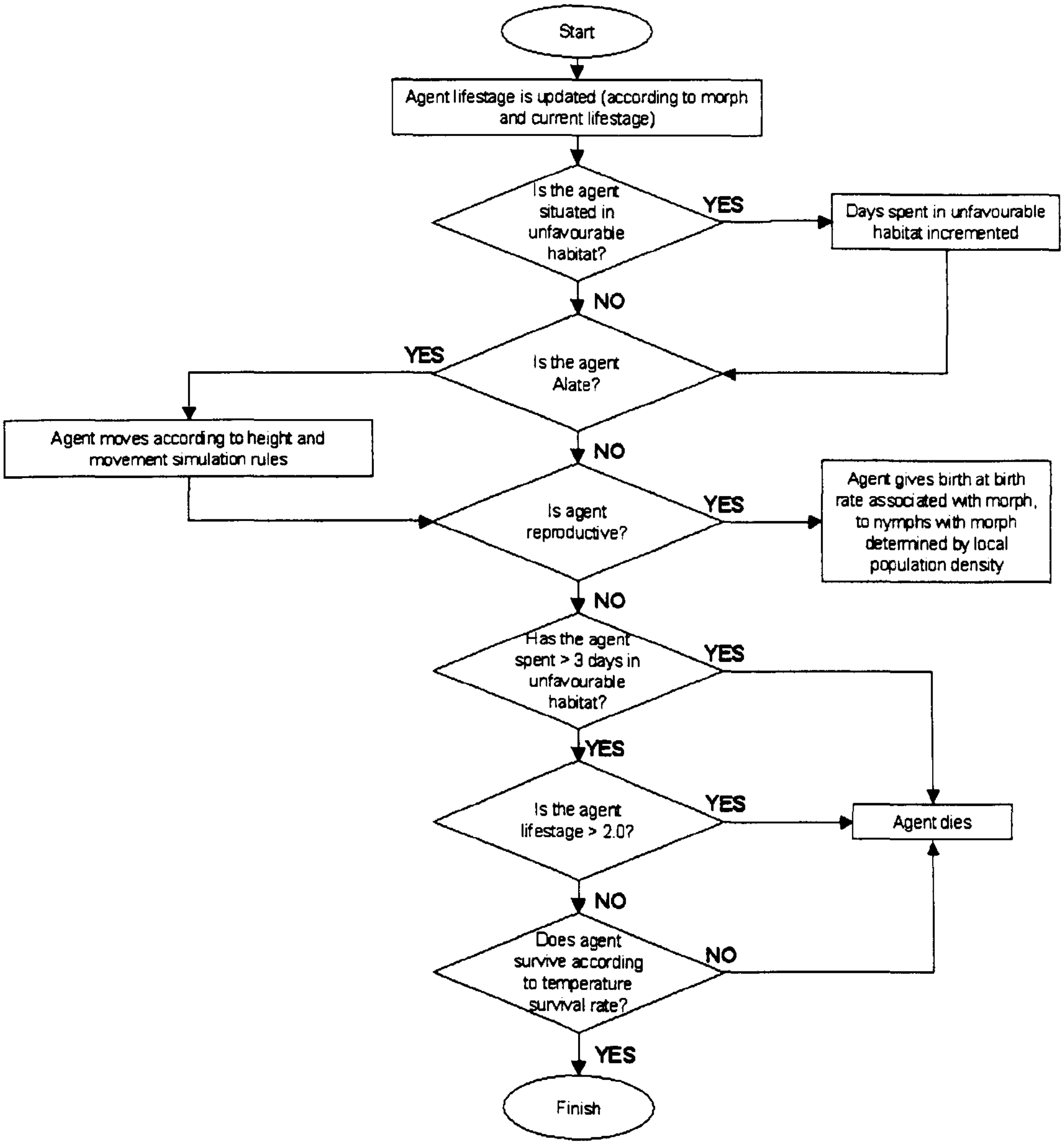


Figure 4.2: Flowchart illustrating the rule set for each agent.

Initial Immigration

For aphids, the immigrants are assumed to be reproductive alate adults, of uniform age. They are also assumed to have undergone 'migration', thus will probably not have a desire to migrate long distances again (Kennedy and Booth, 1963). Before the simulation is started, initial immigration can be input in a number of ways. Immigrants are either input as a single number of immigrants in a discrete event, or as a continuous influx at each time-step. Once immigrants enter the region, they can be randomly distributed across the region, or randomly distributed within an area of the region (such as a single field). The type of immigration is decided before the simulation is run and the appropriate input files chosen; it is not altered during the simulation.

If immigration is input as an estimate from suction trap data, the daily immigration rate of *R. padi* into a cereal field is calculated using a similar procedure to Carter et al. (1982) (see equation A.1, Appendix A.4). Estimates are made from suction trap records following the research of Taylor and Palmer (1972). The estimate relates numbers of alate exules found in the suction trap to deposition rates in the field per hectare of crop: it was assumed that for every aphid caught in the suction trap, assuming an average flight duration and height-density gradient, 237 immigrants colonise 1ha of crop (Taylor and Palmer, 1972: pp. 204). Data were taken from nearby Rothamsted Insect Survey 12.2m suction trap samples of female migrants. This is a very rough estimate that is not species-specific; however, it has been successfully used in previous simulation models of aphids (Carter et al., 1982; Morgan, 2000). It is shown in the sensitivity analysis for this parameter (section 5.3) that varying the number of immigrants alters the temporal trend linearly, and only has a significant influence on spatial and temporal output when this parameter is altered by an order of magnitude.

Reproduction

Aphid agents become reproductive once the agent achieves the appropriate physiological age for reproduction. For alate aphids this is when the proportion of development is 0.9522; for apterous this is 0.9463. The birthrate depends on the morphology of the reproductive aphid, and the daily minimum temperature, maximum temperature and mean temperatures (Wratten, 1977). A logistic curve was used by Morgan (2000) (based upon the data of Dean, 1974b; Dixon, 1976; Leather and Dixon, 1982) to derive an equation for the reproductive rate of apterous adults (see equation A.8, Appendix A.4). Alate adults are assumed to be 1.3 times less fecund than apterous adults (Wratten, 1977) (see equation A.9, Appendix A.4).

Nymphs are located at the same location as their parent when born. The stimulus to produce alates capable of dispersal is related to crowding and/or tactile responses to the nutrient quality of the host (Loxdale and Lushai, 1999). The aphid density per m^2 at the location nymphs are born therefore determines the morphology of the nymphs created (see equation A.10, Appendix A.4).

Ageing and Mortality

Aphid agents at any life-stage may die depending on a survival rate affected by the number of day-degrees below 2.8°C for the day. The survival rate of the aphid agents is calculated from the daily minimum, maximum and mean temperatures (see equation A.3, Appendix A.4).

Mortality also occurs when the aphid agents reach the maximum age of 2.00 (the number of days that this will take depends again on temperatures, see below), and when they remain on unfavourable habitat for more than three days (at present the absence of research in this area makes this an estimate of the aphid's ability to survive poor conditions). The age of the aphid agent increases each day, at a rate determined by the daily temperatures (see equations A.4-A.7, Appendix A.4). The average proportion of development of an aphid was then calculated each day by Simpson's rule (Barlow and Dixon, 1980) (see equation A.2, Appendix A.4).

Other abiotic factors such as rainfall may be relevant (Araya and Fereres, 1991; Dean and Wilding, 1971; Dhalival and Singh, 1975), as well as the effects of predation and parasites or fungi. These are not accounted for in this study, but are acknowledged as an area for future development. It is anticipated that other abiotic factors will not be as significant as temperature, but predation and parasites may have a significant impact on the population dynamics, particularly at the local scale.

Movement

Migration is an important element in a spatial simulation of insect population dynamics, but it can be highly complex. Aphids divide into two classes, those that remain on similar types of plant, and those that are 'host-alternating' (Johnson, 1962). The case-study species used in this thesis, *R. padi*, is host-alternating, and generally migrates back to the primary host *Prunus padus* to lay eggs between September and November. This model focuses on alate exules that remain in the field during this time, as increasingly the species overwinters as mobile stages on early sown winter cereals and grasses. These populations pose a significant threat of the increased spread of BYDV in winter wheat in the context of global warming, particularly in the South of England. This removes the need to consider migration from the primary host (*Prunus padus*); however, the movement and migration of the aphid between secondary hosts is complex, involving various behavioural responses (Loxdale et al., 1993; Robert, 1987).

Much research in the 1950s to 1960s was founded on the belief that the flight of alate aphids can be separated into two phases. The first is a migratory phase followed by a foraging phase (Kennedy and Booth, 1963; Moericke, 1955). However, it is now maintained that alate aphids tend to move only short distances, of the order of 20m in favourable habitat and 100m in poor habitat. Migration is an infrequent occurrence, the exception rather than the rule and there is a gradation from 'non-flyers' to local flyers and then migrants (figure 4.3). In general, if an aphid encounters an unsuitable host it will move on (Loxdale et al., 1993).

The rules of migratory flight used in this model, figure 4.4, follow four principles: firstly, alate aphids will move voluntarily if wind speed is not above 8kmh^{-1} (Haine, 1955; Johnson, 1962; Kennedy and Booth, 1963). Second, aphid migration will take place within a day and during daylight hours (thus a migration event will complete within a single iteration of the model, as this functions on a daily basis) (Loxdale et al., 1993). Thirdly, an individual can only migrate a distance of several kilometres once (if at all) during its lifetime (Ward et al., 1998). Finally, migration will last for a random duration of up to 6.5 hours (Loxdale et al., 1993) during which time the aphid will be carried by the wind a distance determined by the flight duration multiplied by the wind speed, in the direction of the wind's movement (Haine, 1955; Loxdale et al., 1993). It is also assumed that a 'boundary layer' at a height of 1m exists, below which the aphid is unaffected by the wind and free to move at will and above which the aphid's movement is controlled by the wind (Taylor, 1974).

In the model the aphid's vertical position (in metres) is only relevant when the aphid is alate and adult. Thus, for the majority of the population, height (z) is set to zero. If an aphid is alate and adult, but is below the boundary layer threshold, it acts according to the wind speed: if the wind speed is lower than the wind speed halting take-off (8 kmh^{-1}) the aphid moves upwards above the boundary layer; if not, the aphid remains below the boundary layer and moves accordingly (see figure 4.4). Once the aphid's z position is above the boundary layer, if it is a potential migrant (i.e. has not previously migrated) it will continue to move vertically by up to 1m (either up or down). If the wind speed is stronger than its flight speed, it is carried by the wind, otherwise, it moves randomly (see figure 4.4). For aphids that have migrated their vertical position is re-set to zero.

Aphids lose control of their flight at wind speeds of around 2kmh^{-1} (Haine, 1955; Loxdale et al., 1993). Thus it can be inferred that foraging flight may occur at low wind speeds (2kmh^{-1} or less), taking the form of increasingly 'random movement' as wind speeds lower, and short flights tend to be concentrated around host plants (Kennedy et al., 1959). The speed of these movements is set to be the aphid maximum flight speed of 0.9ms^{-1} (3.24kmh^{-1}) (Compton, 2002). To obtain the distance flown this is then multiplied by the average flight time of an aphid, which is about 100-200 minutes (Lewis, 1965).

Low temperatures have been shown to inhibit the take-off of cereal aphids in the field (Walters and Dixon, 1984). However, the impacts of this, particularly for movement on a daily timescale, are limited. *R. padi* adapt to the conditions at the season of their development and

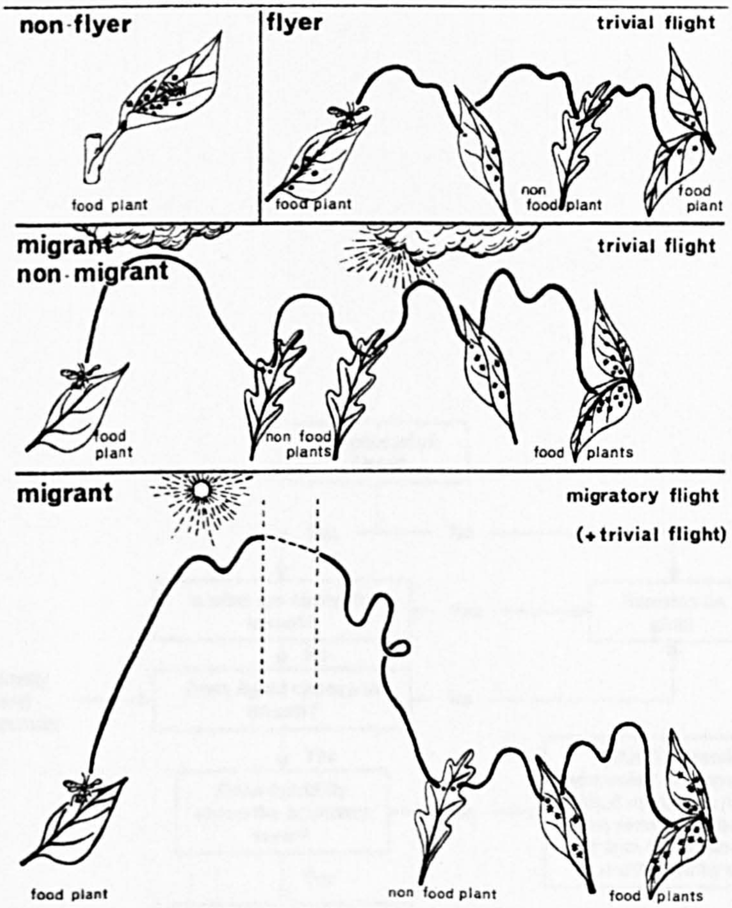


Figure 4.3: Possible patterns of an aphid's flight behaviour according to the aphid's status, the plants encountered, and the weather conditions (Robert, 1987: after Harrewijn et al. (1981); Moericke (1955); Shaw (1970)).

any delay experienced due to adverse temperatures lasts only for a matter of hours, and is not believed to actually prevent movement of the alate morphs (Walters and Dixon, 1984).

Table 4.1 gives an overview of the parameters described above, the model processes and their default values. Default values may be specified as single values or variables upon which the parameter or process depends if it is derived from other variables (primarily temperature dependent). These values are used throughout the thesis, unless otherwise specified.

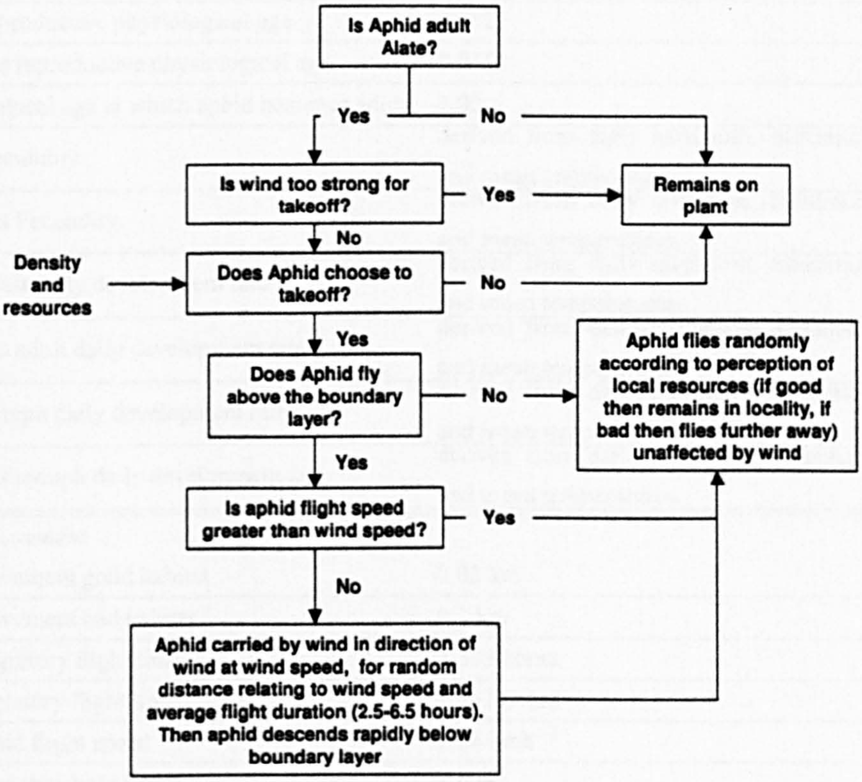


Figure 4.4: Flowchart illustrating the rule set for the wind sub-model.

Parameter	**Value**
Aphid biology	
Survival rate	derived from daily minimum, maximum and mean temperatures.
Adult longevity	2.0
Immigration (number of aphids)	input from file, either a single event or across several days. Alate adults.
Alate reproductive physiological age	0.9522
Apterous reproductive physiological age	0.9463
Physiological age at which aphid becomes adult	1.00
Alate Fecundity	derived from daily minimum, maximum and mean temperatures.
Apterous Fecundity	derived from daily minimum, maximum and mean temperatures.
Alate adult daily development rate	derived from daily minimum, maximum and mean temperatures.
Apterous adult daily development rate	derived from daily minimum, maximum and mean temperatures.
Alate nymph daily development rate	derived from daily minimum, maximum and mean temperatures.
Apterous nymph daily development rate	derived from daily minimum, maximum and mean temperatures.
Aphid movement	
Max movement good habitat	0.02 km
Max movement bad habitat	0.1 km
Max migratory flight time	3.333 hours
Min migratory flight time	1.667 hours
Max aphid flight speed	3.24 kmh^{-1}
Boundary threshold height	1.0 m
Environment	
Wind speed	derived from julian day
Wind speed halting take-off	8 kmh^{-1}
Wind direction	270.0°
Cell size	25×25 m
Region size	50×50 cells

Table 4.1: Overview of parameters and their default values used in the model

4.4 Construction of the IBM

The model has been developed in Java using object-oriented programming, due the conceptual similarity of object-oriented code to both the agent simulation and ecosystem structure (see chapter 2 for more detailed evaluation of the choice of object-oriented individual-based simulation). Java is easily developed, particularly with the use of NetBeans, an Interactive Development Environment (IDE) for Java (`http://www.netbeans.org/`). Java is multi-platform and freely available, and is the primary language in which the Repast toolkit has been developed.

In the early stages of model development, an agent-model was constructed completely from scratch without the aid of a toolkit. This resulted in a model that performed the functions required but with a rather basic Graphical User Interface (GUI) (figure 4.5). It also had an inelegant structure, where many agent parameters such as birth-rate and survival were part of the environment class. Better structuring of parameters into the relevant classes was needed, and also a means to make the model more flexible (for example to make it possible to add different agents representing other species). Repast (see chapter 2, section 2.5.3) provides a framework that aids the development of an agent-based model significantly, using a library of Java classes. This approach was subsequently adopted, to improve the functionality of the GUI and speed model development. It can be seen clearly from a comparison between the original and the final GUI that Repast has made a significant difference to the appearance and simplicity of use of the model (figure 4.6). The creation of a GUI that is easily useable and readable is fundamental to the understanding, analysis and communication of the model (Grimm, 2002). Experiments with run-times have also shown that the equivalent Repast model is a great deal faster than the original model developed, able to handle over 10,000 agents on a 200 x 200 cell grid in just a few minutes using 300MB of RAM.

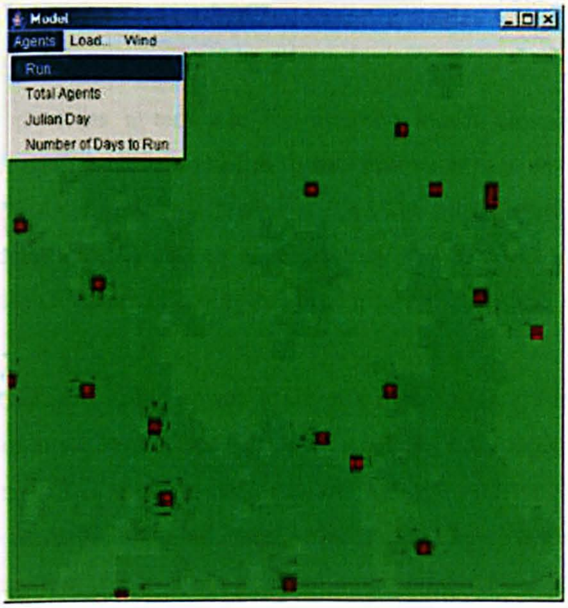
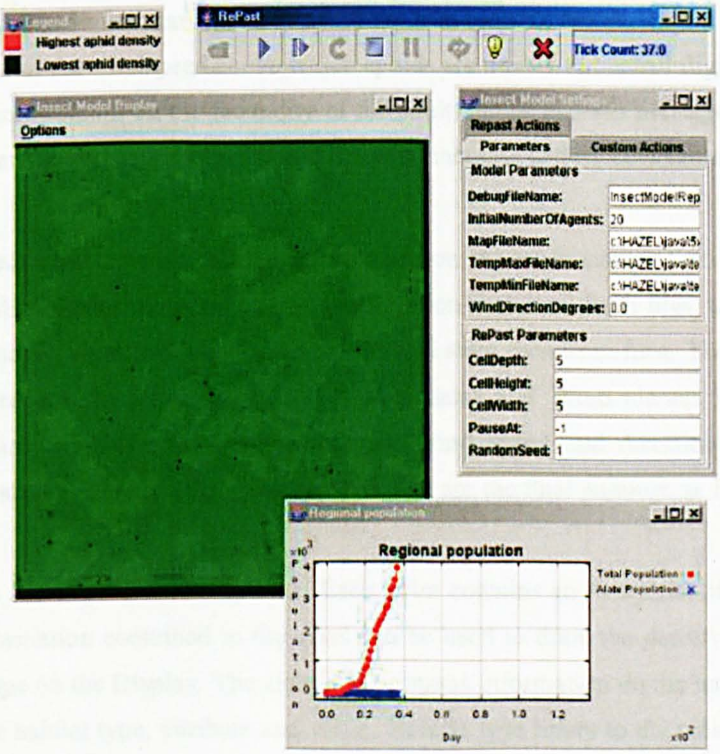


Figure 4.5: Model GUI constructed without the aid of a toolkit.



Model Parameters

DebugFileName: InsectModelRep
InitialNumberOfAgents: 20
MapFileName: c:\HAZEL\java\5
TempMaxFileName: c:\HAZEL\java\te
TempMinFileName: c:\HAZEL\java\te
WindDirectionDegrees: 0.0

RePast Parameters

CellDepth: 5
CellHeight: 5
CellWidth: 5
PauseAt: -1
RandomSeed: 1

Figure 4.6: Model GUI constructed with Repast. Features windows showing model display, Legend, GUI and chart (generated in real-time). Also includes GUI facilities for taking model snapshot and making Quicktime movies.

Appendix E.1 (included CD) shows a Unified Modelling Language (UML) schematic of the model structure, with method and class interaction and variables.

The Insect interface provides a template for creating insect classes. This contains all essential methods that any insect will require to function such as to set an agent's ID, spatial co-ordinates (x,y,z), lifestage and reproductive status. The main benefit of implementing the Insect interface in species specific classes is consistency and control of the program structure, as well as to ensure that all essential biological functions are included.

The Aphid class represents an aphid agent. It inherits from the Insect interface, thus contains those methods and some additional species specific ones: information on the morphology of the aphid, how long it has spent in unfavourable habitat, and whether it has migrated is specific to aphids. To construct an aphid agent, information about its environment is required. This is useful if the agent wants to find out about environmental conditions. The coordinate class stores coordinate objects for the Aphid class. This makes it simple to recall the x or y location of the aphid.

The AphidInfo class simply contains the aphid specific parameters for the model. These are the boundary threshold for migratory flight, the windspeed above which aphids will not take off, the maximum and minimum speed of aphid flight in kmh^{-1}, the maximum and minimum flight duration, and the windspeed above which aphids are unable to control flight (see chapter 3). It also has information on the fecundity of the adult aphids and the lifestage increases for each development stage (both fecundity and lifestage increase differs between morphs).

The Environment class contains all the information on the environment that the agents are exposed to. This includes temperature and immigration data read from files, using the TemperatureData and ImmigrationData classes to read and store data from files. The Environment class also references the RasterMap (ASCII map data) and Wind classes which contain information from the GIS and wind parameters (wind speed and direction) respectively. Survival rate and day degrees below a boundary value are the final parameters for the class.

The Point class stores geographical points. Each point contains an ArrayList of agents at that point. The information contained in the class can be used to draw the density of the agents and the landscape on the Display. The class also contains information on the number of births at the point, the habitat type, attribute and value. Habitat type refers to the values assigned in the ASCII file. This is then translated to a habitat attribute by looking up the key from the value file, if necessary.

InsectModel is the main class for the model. It controls the flow of the program (figure 4.1) and is structured following the RePast framework (see chapter 2, section 2.5.3). This class contains a lot of variables, as it is best to keep them all in one place to edit easily, and by locating the variables in this class they can also be altered by the GUI once the model is running. Thus it contains all the file names for the model, i.e. those that are used by the TemperatureData, RasterMap and ImmigrationData classes to read in data, and the filenames to be written to as output from the simulation. The class also contains the date of sowing and date of harvesting of the crop, as well as pesticide spray date and fields to be sprayed.

For each daily iteration (called a 'tick' in Repast) the model rules are run (figure 4.1). First the environment is built for the iteration. However, if there is insufficient information to build the environment (for example there is no more temperature data) the simulation is halted. Once the environment is built immigration then takes place if requested, and immigrant aphids are added to the agent list associated with the point object they have immigrated to. Actions take place as the agents at each point are iterated. The display surface is then updated.

4.5 Model development summary

A non-linear strategy for the development of the model is presented in this chapter, and details are given on the construction of the IBM of the behaviour and movement of *R. padi* in changing agricultural landscapes. The use of the Repast toolkit is detailed, and the conceptual framework for the model is presented. The primary reasons that an individual-based approach is appropriate are identified as flexibility, a reflection of real ecosystem structure, multi-scale interactions, ease of parameterisation and the possibility to examine individual life histories. The choice of spatial extent and resolution at which to present the model is discussed, and further assessment is made of the implications of this scale choice in chapter 5.

This chapter has described the model construction. The following chapter assesses the model to show the stability and validity of the model; thus fulfilling objective 4 of the thesis. The abundance of the simulated pest species *R. padi* also required further model development to handle large numbers of individuals (the necessity of which made it objective 5 of the research). The techniques developed to handle large numbers of individuals are detailed in chapter 6.

CHAPTER 5

Model Analysis

5.1 Introduction

This chapter presents the model assessment. This includes validation of the model concepts, development and operation, as well as sensitivity testing to meet objective 4 of the thesis. Validation is taken to be a single component of overall model evaluation; hence the term 'model assessment' is used to encompass the range of techniques employed in this chapter. The concept of validation, and how one should go about validating a model, are much debated in the literature; some consider validation essential, whilst others consider it impossible (e.g. see Lane and Richards, 2001; Rykiel, 1996). An empirical, 'positivist', approach is generally adopted, demanding that models be validated by comparison to field data (e.g. Wilder, 1999). However, this approach only forms part of the broader model assessment conducted for this thesis (after Lane and Richards, 2001).

An important first step is to identify and define clearly the *purpose* of the model (as in chapter 4, section 4.2). Most ecological models are built for scientific research purposes, but also increasingly for forecasting and management purposes. Key considerations are:

- Whether the model is acceptable for its intended use: i.e. is it comparable to the real-world well enough for its purpose (Giere, 1991)?
- How confident one can be that the model results can be used to make correct inferences about the real system (Curry et al., 1989).

Model assessment is continuous as the model is refined and developed. Due to the complexity of individual-based ecological models and the environmental systems they simulate, Lane and Richards (2001) state that falsification is:

- dependent upon the criteria used to falsify,
- inevitable due to complexity,

- of no real use unless it enables the modeller to understand why the model fails.

Not only is it important to identify when and why a model fails, but also to then take predictions that do not conform to existing theoretical and empirical understanding and use them to perhaps view knowledge in a different way, and advise further modelling, field and experimental activities (Kirkby, 1996; Lane and Richards, 2001). Many models are primarily of heuristic value (Lane and Richards, 2001: pp. 435), providing guidance for further study. Thus, models should challenge existing knowledge, be it empirical or theoretical, and should not necessarily be compared to it by 'validation and verification' and then rejected if they do not agree. Empirical validation is included as part of a broader model assessment for this thesis, and it was found that there is a good agreement between the model output and field observations at the plant-scale (see section 5.2.4).

5.2 Model assessment

'Validation is a demonstration that a model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model' (Rykiel, 1996: pp. 233).

Rykiel (1996) makes two key arguments that relate to research models. Firstly, that validation is not an essential activity to evaluate a model intended for research, but it is important for building credibility in the user community. Second, models can be validated as acceptable for pragmatic purposes, but theoretical validity is always provisional. Model assessment methods are applied to 'critical stages in the model building process' (Rykiel, 1996):

- Design: the theoretical, observational, intuitive and general knowledge basis for the model (*Conceptual validation*).
- Implementation: Empirical testing of the model's assumptions (where possible). Also verification of the computation and data (*Development validation*).
- Operation: Input-output relationships of the model are compared with a real system (*Model operation validation*).

It should also be noted that a model may arguably not require empirical validation. It may not be possible due to lack of data, and a model's behaviour can quite legitimately be explored without validation. "Model development is a significant scientific contribution in itself without any validation test being undertaken in addition" (Rykiel, 1996: pp. 241). Ideally, validation should be undertaken to improve confidence in a model, either by the modeller or other scientists, but may not be possible at the time of development: a paradox of scientific endeavor. In particular, this is true when the purpose of the model is purely theoretical and established to describe knowledge. This thesis presents empirical validation at the plant-scale, but was constrained by time and data availability to empirically validate the landscape-scale outputs: this could form a study in itself.

5.2.1 Assessment of the aphid simulation

Lane and Richards (2001) advocate an approach that moves beyond positivist validation against empirical data (which implicitly assumes that empirical data better represents reality), to validation and assessment of the model in which a wide range of different quantitative and qualitative criteria are used to assess the model's capabilities to represent particular natural process dynamics. They term this a 'post-positivist' approach (see also Brown, 1996).

The assessment is non-linear, rather than a simple procedure at each stage of the model development. For example, if the model is found to be inadequate at the developmental stage (perhaps the behaviour is highly unrealistic) a return to assessment at the conceptual stage may be in order, and perhaps also model calibration. Similarly, if the model is found not to meet operational requirements in the last stage, it is necessary to reassess the concepts and development once more.

5.2.2 Validation at the conceptual stage

It is important to validate the logic of a model's scientific content, which cannot be done by simply comparing simulated data with empirical data (Usunoff et al., 1992). In particular this is done at the exploratory/theoretical model development stage. It ensures that the theories and assumptions underlying the conceptual model are made explicit, and are justifiable, if not correct.

In an ecological model this could constitute a justification of the spatial and temporal scale of the model, and why certain processes are necessary to include in the model, ensuring 'the model representation of the problem or system, its structure, logic, mathematical, and causal relationships, are reasonable for the model's intended use' (Rykiel, 1996: pp. 234). The model must be grounded in existing knowledge, and it should be clearly identified what the model can, and cannot, achieve and at which spatial and temporal scales.

The conceptual model can also indicate where the model may fail, and thus where other forms of validation must be rigorous (Lane and Richards, 2001). The fundamental logic of the aphid model's scientific content is assessed throughout this thesis. The key elements considered are:

- Justification of an individual-based modelling (IBM) approach (as opposed to population-level, procedural modelling, etc.).

- Justification of spatial scale used in the simulation (the minimum cell size, the spatial extent of study region, ability to look at multiple-scales).
- Justification of temporal scale used in the simulation (winter months due to the importance of the spread of BYDV in winter crops, daily time-steps).
- Explanation of why certain processes are included in the model: wind, landscape change, density dependence, temperature change etc.
- Justification of why equations or stochastic methods are used.
- Demonstration of how the model relates to other models; of the species (aphid), insects in general and ecological models as a whole.
- Clear definition of the aims and objectives of the model, in order to identify what it should achieve.
- Identification of any potential weaknesses of the model.
- Evaluation of the behaviour of the model qualitatively: does it seem to represent reality adequately? Does the output follow realistic trends? Does it compare to other models (model-to-model comparison)?

The majority of these points have been addressed by the Background chapter (chapter 2) and the Model Development chapter (chapter 4). A conceptually valid model will not, however, be guaranteed to make accurate predictions.

5.2.3 Validation at the development stage

At this stage assessment mainly involves checking and verifying the parameters and coding of the model. It is a continual process throughout the model development, as each component of the model is assessed before, during and after construction. This involved:

- Clear assessment of the quality of the parameters used in the model, the way in which they are derived and calibrated, and any assumptions that are made.
- Checking of computer code for errors.
- Testing that the consistency of stochastic output is appropriate.
- Tracing the behaviour of specific variables through the model for realistic behaviour and accuracy (e.g. aphid life cycles).
- Ensuring documentation of code, and clear diagrammatic representation.

The data for the model, assumptions and parameters used are described and evaluated in this thesis (chapter 3). The code was constantly debugged throughout its development, and version control was exercised to maintain records and quality, as well as clear documentation (see javadocs, appendix E.3). This also involved examination of individual behaviour by output of individual life histories.

5.2.4 Validation at the model operation stage

Once development of the model was complete, the model was then evaluated in terms of its sensitivity to certain parameters, the 'validity' of any predictions was assessed and uncertainties were evaluated. Such an assessment is a test of the robustness of the model, as it may lead to the formulation of alternative hypotheses and important parameters and processes being identified.

The purpose of this simulation was to generate scenarios showing the present and possible future effects of current countryside and agri-management practices at various scales (see section 4.2). Such scenarios are highly theoretical as the objective was to test the model for realistic behaviour in a landscape context. However, comparison with real systems was necessary. If the model is to claim that it could be used as a management tool, it must achieve a higher degree of 'validity' against a real system (often empirical), for it to be accepted by the management user-group.

Validation against empirical data: plant scale

The model was validated against independent field data collected at the plant scale (scaled to $1m^2$, assuming 300 plants per m^2), (figure 5.1). The simulation was run for a single $1m^2$ crop cell and compared to observed *Rhopalosiphum padi* populations in 1985/6 at Rothamsted (field data from D. Morgan, *pers. comm.*: see chapter 3, section 3.3). Immigration was predicted from the suction trap data (see chapter 4, section 4.3.2). Aphid densities were under-predicted by the model, but follow a very similar trend to the field; populations increased rapidly from very low numbers and peaked around 40 days later. Thereafter numbers declined gradually, although aphids were present throughout the winter, albeit at low density. The percentage error between the model and the field density varies between 32% and 67% of the field value. However, the 95% confidence limits derived from the standard error of the field data are such that the model predictions fall largely within these. Overall, based upon the comparison with the available field data, the model appears to make reasonable predictions of aphid density and follows the same trend as reality (indicating 'event validity', where the occurrence, timing and magnitude of simulated and actual events is comparable).

The model's relationship with abiotic variables, in particular temperature, was important. This reflects empirical evidence that shows aphid population dynamics are particularly sensitive to temperature change (e.g. Dean, 1974b; Walters and Dixon, 1984). As temperatures alter and decrease through the season, aphid populations fluctuate and then also begin to decline when

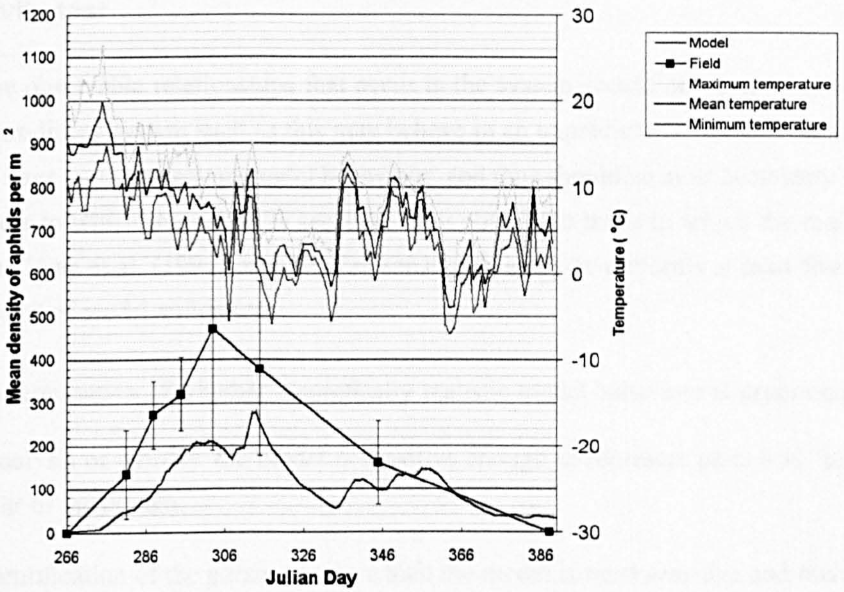


Figure 5.1: Primary axis: Simulation for single $1m^2$ crop cell (solid line) and observed (■) *Rhopalosiphum padi* populations in 1985 at Rothamsted, with 95% confidence limits (1.96× standard error). Field data from D Morgan, *pers. comm.*. Secondary axis: Temperatures measured during study period (source: UK Met Office records, Rothamsted).

temperatures remain at their lowest (below 5 °C) from julian day 310. A secondary peak in the aphid population later in the season was probably due to relatively mild temperatures from julian day 336 to julian day 356. Overall, the complex response of this IBM to temperature change produces results that are comparable to real field populations at a local scale.

This comparison with the field data assumes there is no 'equifinality' (Beven, 1996) in the model (where the model produces results that appear to represent real trends, when in fact they do so for the wrong reasons): rigorous assessment at the conceptual and development stages has ensured this. Examination of individual life-histories indicates that the model aphids resemble those of real aphids in the field by expert judgement. For example, generation times are realistic (three to four generations of aphids are produced across the time period). Rigorous sensitivity analysis gives confidence that key processes (such as temperature change) acting on the aphids in reality are also key to the behaviour of the aphids in the model, and this analysis also ensures that the influence of such parameters on the aphids is realistic.

The comparison with field data also assumes that the field data used represents reality: there is a greater confidence in measurement and observation, yet there is no philosophical foundation that measurements should be more accurate than a model's prediction of 'reality' (Lane and Richards, 2001).

Sensitivity test

The same observable relationships that occur in the system should occur in the model. However, a non-linear system such as this may behave in an unpredictable manner. Some parameters will have more impact on model behaviour, and thus should be most accurately estimated. Parameters to which the model is sensitive may not match those to which the real system is sensitive. Lane et al. (1994) propose that sensitivity analysis performs at least five key functions as part of model validation:

1. Determination of whether theoretically realistic model behaviour is experienced.
2. Analysis of whether the model is sensitive enough to represent perceived 'real' behaviour of the system.
3. Identification of the parameters to which the model is most sensitive and therefore need to be based upon the most robust data.
4. Improve the performance of a model by limiting sensitivity (where appropriate).
5. Give an indication of the error in model prediction.

In its simplest form, sensitivity analysis is conducted by keeping all parameters but the one to be tested constant, then altering the values of the test parameter by a factor across a realistic range of values. This is known as 'factor perturbation' (Lane and Richards, 2001).

However, there are a number of difficulties in conducting sensitivity analysis. Often parameters are inter-correlated, or there may be a non-linear response to parameter perturbation. The model may also vary in sensitivity depending upon the spatial and temporal resolution at which it is tested (Lane and Richards, 2001). For this reason all parameters are kept constant in the sensitivity analysis, whilst key parameters are varied in a controlled manner.

The following section details the sensitivity analysis conducted for the model.

5.3 Sensitivity analysis

Some of the parameters in the model are stochastic or highly variable in time and space. In terms of environmental variables, these are: temperature, wind speed and wind direction. In terms of the aphid agents, these are: immigration, survival, movement and migration. Such variability in a number of parameters simultaneously makes it difficult to isolate the influence of a single parameter on model behaviour during a simulation run, or to determine which parameter is the most sensitive. However, by limiting all variable parameters to single values and then adjusting them one at a time in a sensitivity analysis, the behaviour of the model can be more easily understood. An initial population of 1,000 alate immigrant aphids was used.

This is lower than realistic densities at the scale used (25 × 25m) and was chosen to enable faster processing time.

This section is divided into three parts, corresponding to three means of sensitivity analysis. These analyses tested the behaviour of the model and its ability to simulate realistic population dynamics and the sensitivity of the model parameters. The sensitivity of the model was tested both temporally and spatially. The parameters tested are shown in table 5.1.

The first analysis (section 5.4) investigated temporally and spatially the relative importance of morphology. This used a 'coarse' approach, where the morphology was set to extreme values to determine the effects of this variable on the system. In the second analysis (section 5.5), the extent and cell-size of the region was varied, and the temporal population dynamics across a comparable $1m^2$ area were assessed. This was to test the consistency of the model across scales. In addition, the regional implications of altering cell size in a fragmented landscape were assessed, to look at the spatial effects of changing cell size and if this had implications for the behaviour of the aphids under the same fragmentation scenario at the scale of 25 × 25m. In the third analysis (section 5.6), the temporal sensitivity of the model to perturbing parameters was investigated on a single 25 × 25m cell. A single cell was used as spatial heterogeneity complicates the temporal model analysis (van Nes and Scheffer, 2005). This is termed a 'fine' approach, where small positive and negative changes were made to individual key parameters in the model, to examine the model stability and sensitivity over a hundred iterations. The main parameters that were tested in this analysis were temperature, immigration numbers and survival rate. A series of diffusion experiments were conducted to observe the spatial impact of variation of parameters on a 50 × 50 cell grid. Along with the sensitivity analysis of landscape parameters in section 5.4, this forms preliminary studies for the scenario chapters 7 and 8, which examine the effects of spatial habitat configuration on the diffusion of aphids across the landscape, as part of the investigations into the effects of variation in agronomic practices in time and space on aphid populations.

Parameter	**Control Value**	**Sensitivity Test**
Aphid biology		
Survival rate	90%	±5%
Adult longevity	2.0	±5%, ±20%
Immigration (number of aphids)	1,000	±5%, ±20%
Alate Fecundity	0.722	±5%, ±20%
Apterous Fecundity	0.939	±5%, ±20%
Alate adult daily development rate	0.044	±5%, ±20%
Apterous adult daily development rate	0.047	±5%, ±20%
Alate nymph daily development rate	0.029	±5%, ±20%
Apterous nymph daily development rate	0.044	±5%, ±20%
Aphid movement		
Max movement good habitat	0.02 km	±5%, ±20%
Max movement bad habitat	0.1 km	
Max migratory flight time	3.333 hours	
Min migratory flight time	1.667 hours	
Max aphid flight speed	3.2 kmh^{-1}	
Boundary threshold height	1.0 m	
Environment		
Mean temperature	7°C	±1°C
Wind speed	8 kmh^{-1}	
Wind speed halting take-off	8 kmh^{-1}	
Wind direction	270.0°	
Landscape fragmentation	0%[a]	
Habitat coverage	100%	
Cell size	25×25 m	1 m, 10m, 100 m
Region size	50×50 cells[b]	3 × 3, 5 × 5 cells

Table 5.1: Sensitivity test parameters.

---

[a] The landscape fragmentation index for the single cell and the diffusion sensitivity test landscapes was not necessary, as the landscape was 100% habitat.

[b] This was the size for spatial tests. The sensitivity test of the region size compared a single 25×25 m cell to 3 × 3 and 5 × 5 25×25 m cells

5.4 Coarse sensitivity test

In this analysis, a parameter was simply omitted from the model or set to an extreme value, to determine the importance of the component and its overall effect on the system. The parameters immigration number, temperature, cell size and region size are key to the model (see chapter 4); values for these parameters must exist in order for the model to be established. Coarse sensitivity analysis was carried out on morph determination; although this depends primarily on density values, it was useful to observe the effects of all nymphs developing into apterous adults, or alate adults (initial population of 1,000 alate adults).

5.4.1 Alate versus apterous morphs

The behaviour under controlled model parameters of a population of 100% alate, 100% apterous and 100% alate migrants was compared to a realistically structured 'control' population. Temporally, aphid density decreased rapidly when all nymphs were alate, to a mean low density of 26 aphids within a single 25×25m cell (figure 5.2). The reason for this is the lower development rate of alate nymphs (apterous development rate/1.5), and alate adults. The resultant longer lifespan of alate morphs increased mortality pressure, as the natural population increase was delayed. The delayed population increase is shown by the later second peak in the alate curve at julian day 314, with only 606 aphids in the 25×25m cell compared to 2,417 aphids 10 days earlier for the apterous curve. Alate adults are also less fecund than apterous adults, magnifying the trend. A population of 100% apterous aphids followed a similar trend to the control, showing that in reality very few alates were produced in the sensitivity test model, as the simulation was run at low density (1,000 aphids per 25×25m cell). A population of 100% alate migrants was also tested, as in reality a large proportion of the alate population would move from the cell. This resulted in a rapid decline in density and population extinction after julian day 284 as all alates were mature enough to migrate from the cell by this time.

Similar observations were made by Carter et al. (1982: pp 55-57) in the coarse sensitivity analysis conducted on their model, where lower population increase was experienced when alates were included in the model. The alate population dynamics (both population size and proportion of migrants) are therefore important in determining the timing of the peak aphid population and the date when aphid numbers begin to drop. The effect was amplified as natural enemies and intra-specific competition were not included in the model (Carter et al., 1982).

Spatially (figure 5.3) using a 50 × 50 cell grid of 25×25m cells, it was shown that a purely alate population produced a highly dispersed, low density population, in the direction of the wind (here the aphids cycle around the landscape; hence a 'band' of aphids was produced across the landscape). The population dynamics at the central cell is shown in figure 5.4. This is a more accurate temporal representation than the study conducted with a single cell, as migrants were free to move between cells so cell population dynamics were more realistic than those of the single cell and aphids persisted by moving between cells rather than dying out unrealistically.

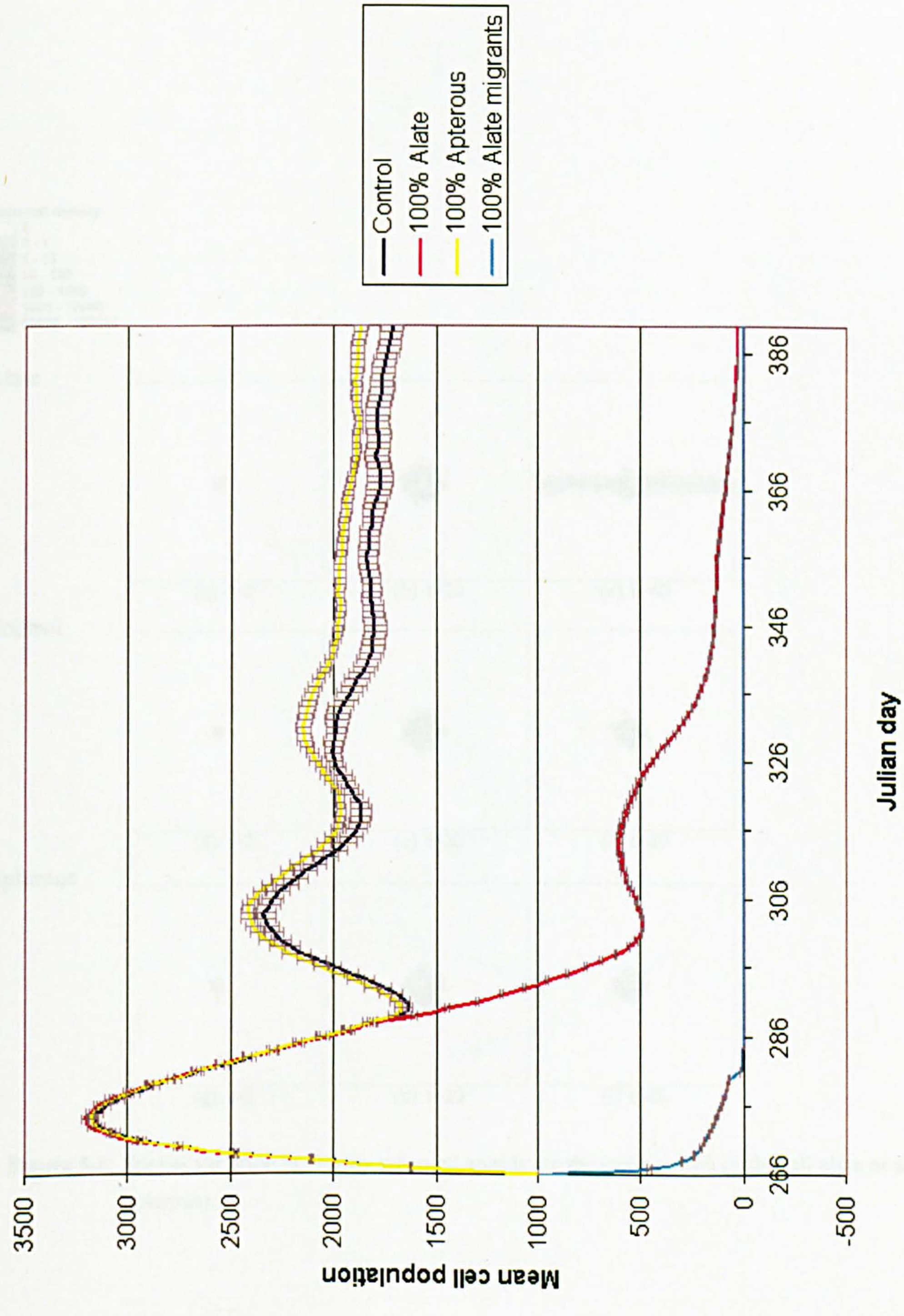
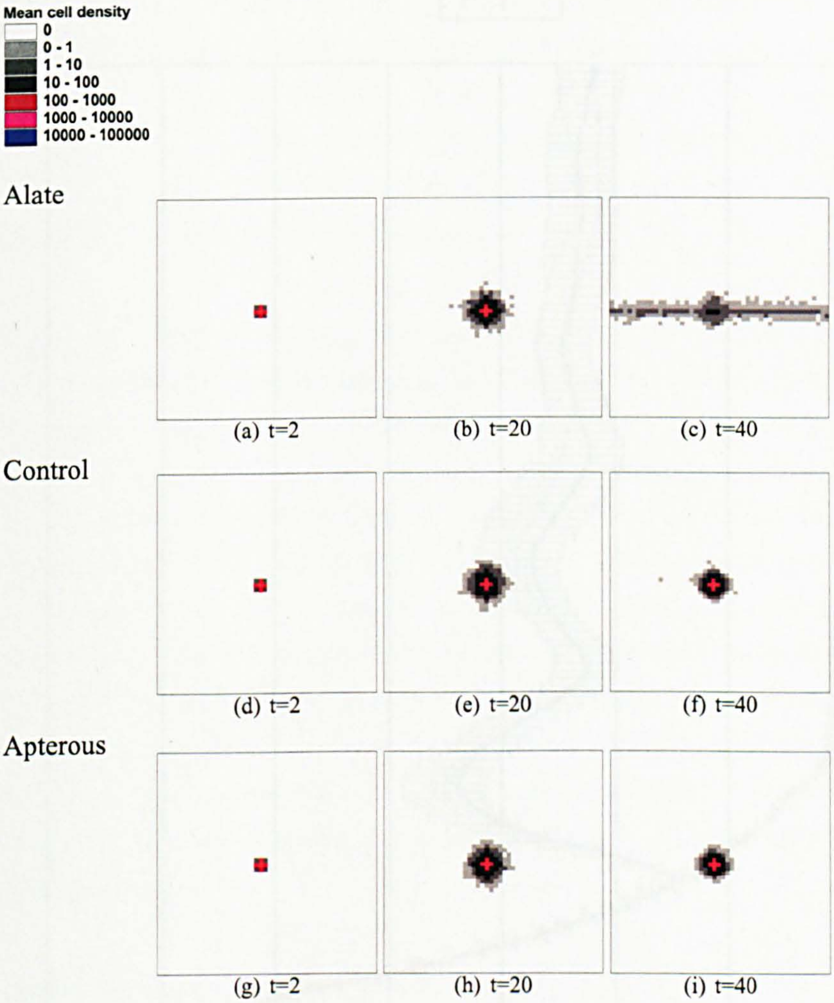


Figure 5.2: Density trends when all aphids are the same morph (either all alate or all apterous).



(a) t=2 (b) t=20 (c) t=40

(d) t=2 (e) t=20 (f) t=40

(g) t=2 (h) t=20 (i) t=40

Figure 5.3: Spatial variation in density when all aphids are the same morph (either all alate or all apterous).

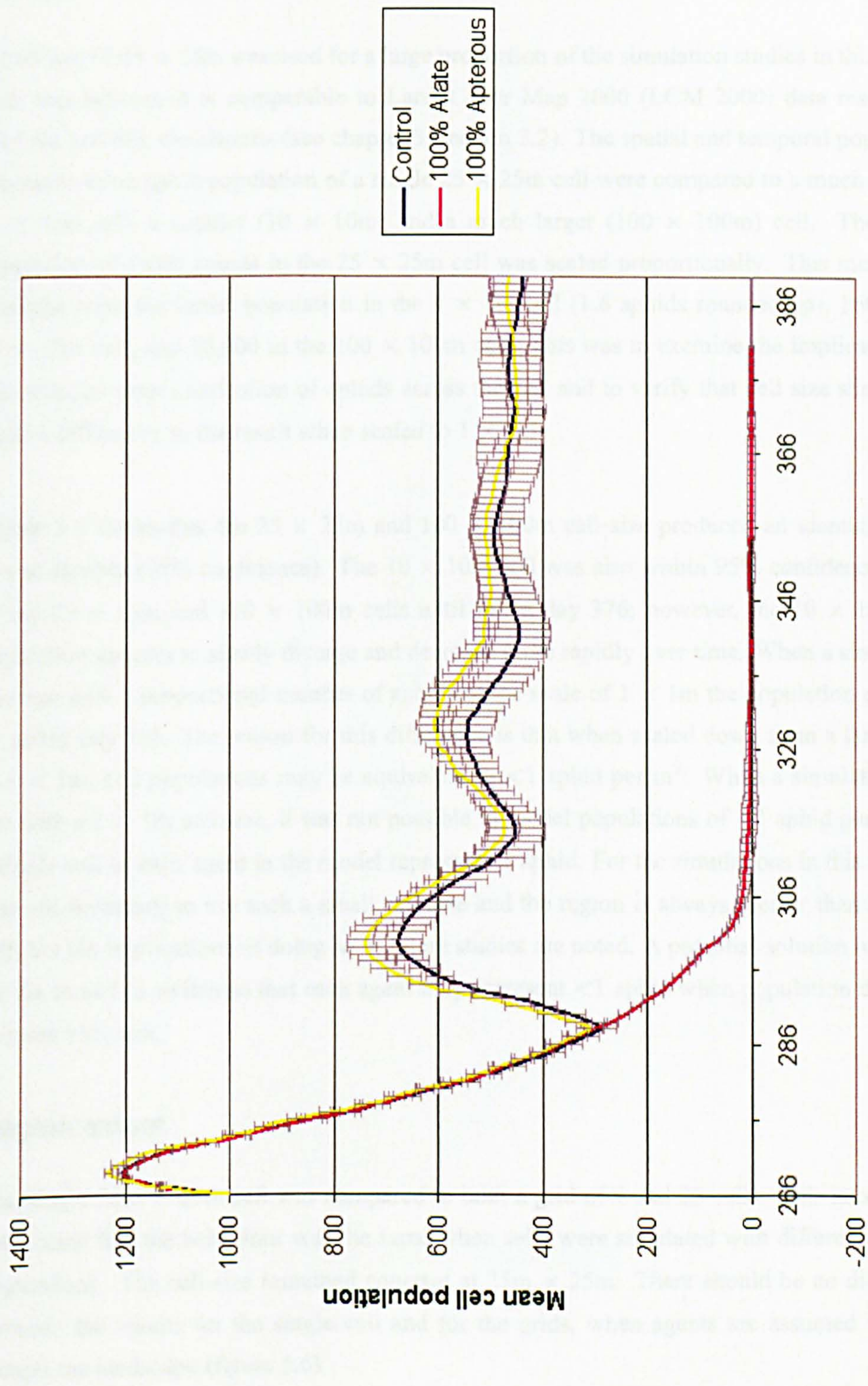


Figure 5.4: Density trends when all aphids are the same morph (either all alate or all apterous) for origin cell of spatial simulation.

5.5 Scaling of cell size and region extent

5.5.1 Cell size

A cell size of 25 × 25m was used for a large proportion of the simulation studies in this thesis. This was because it is comparable to Land Cover Map 2000 (LCM 2000) data resolution, used for realistic simulations (see chapter 3, section 3.2). The spatial and temporal population dynamics of an aphid population of a single 25 × 25m cell were compared to a much smaller (1 × 1m) cell, a smaller (10 × 10m) and a much larger (100 × 100m) cell. The initial population of 1,000 aphids in the 25 × 25m cell was scaled proportionally. This meant that 2 aphids were the initial population in the 1 × 1m cell (1.6 aphids rounded up), 160 in the 10 × 10m cell, and 16,000 in the 100 × 100m cell. This was to examine the implications of assuming an even distribution of aphids across the cell and to verify that cell size should not make a difference to the result when scaled to 1 × 1m.

Figure 5.5 shows that the 25 × 25m and 100 × 100m cell size produced an identical trend to one another (95% confidence). The 10 × 10m cell was also within 95% confidence limits of the 25 × 25m and 100 × 100m cells until julian day 376; however, the 10 × 10m cell population appears to slowly diverge and decrease more rapidly over time. When a simulation was run with a proportional number of aphids at the scale of 1 × 1m the population died out by julian day 316. The reason for this difference is that when scaled down from a larger cell to 1 × 1m, cell populations may be equivalent to <1 aphid per m^2. When a simulation was run with a 1 × 1m cell size, it was not possible to model populations of <1 aphid per m^2 for a single cell as each agent in the model represents 1 aphid. For the simulations in this thesis it was not necessary to use such a small cell size and the region is always greater than a single cell, but the implications of doing so in future studies are noted. A potential solution would be for the model to switch so that each agent may represent <1 aphid when population densities become very low.

5.5.2 Region extent

The single 25m × 25m cell was compared to both a grid of 9 and 25 cells. This allowed for verfication that the behaviour was the same when cells were simulated with different starting populations. The cell size remained constant at 25m × 25m. There should be no difference between the results for the single cell and for the grids, when agents are assumed to cycle around the landscape (figure 5.6).

As hypothesised, figure 5.6 shows that there was little difference between the results for a single cell and for the varying size grids. This ensured that the cycling around the torus was accurate and there was no spatial bias in the rules of this sensitivity analysis.

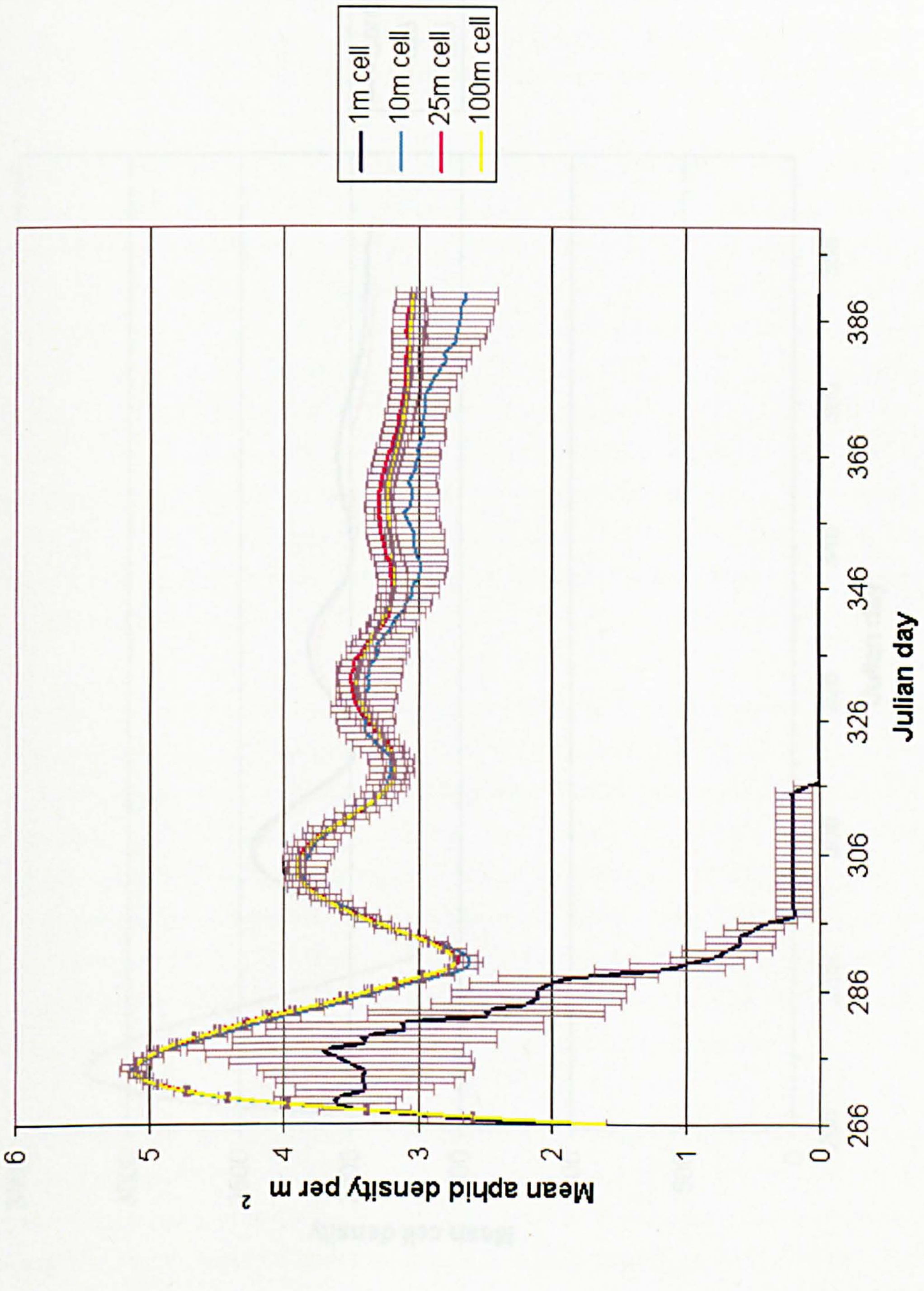


Figure 5.5: Impact of varying cell size on mean aphid density per m^2.

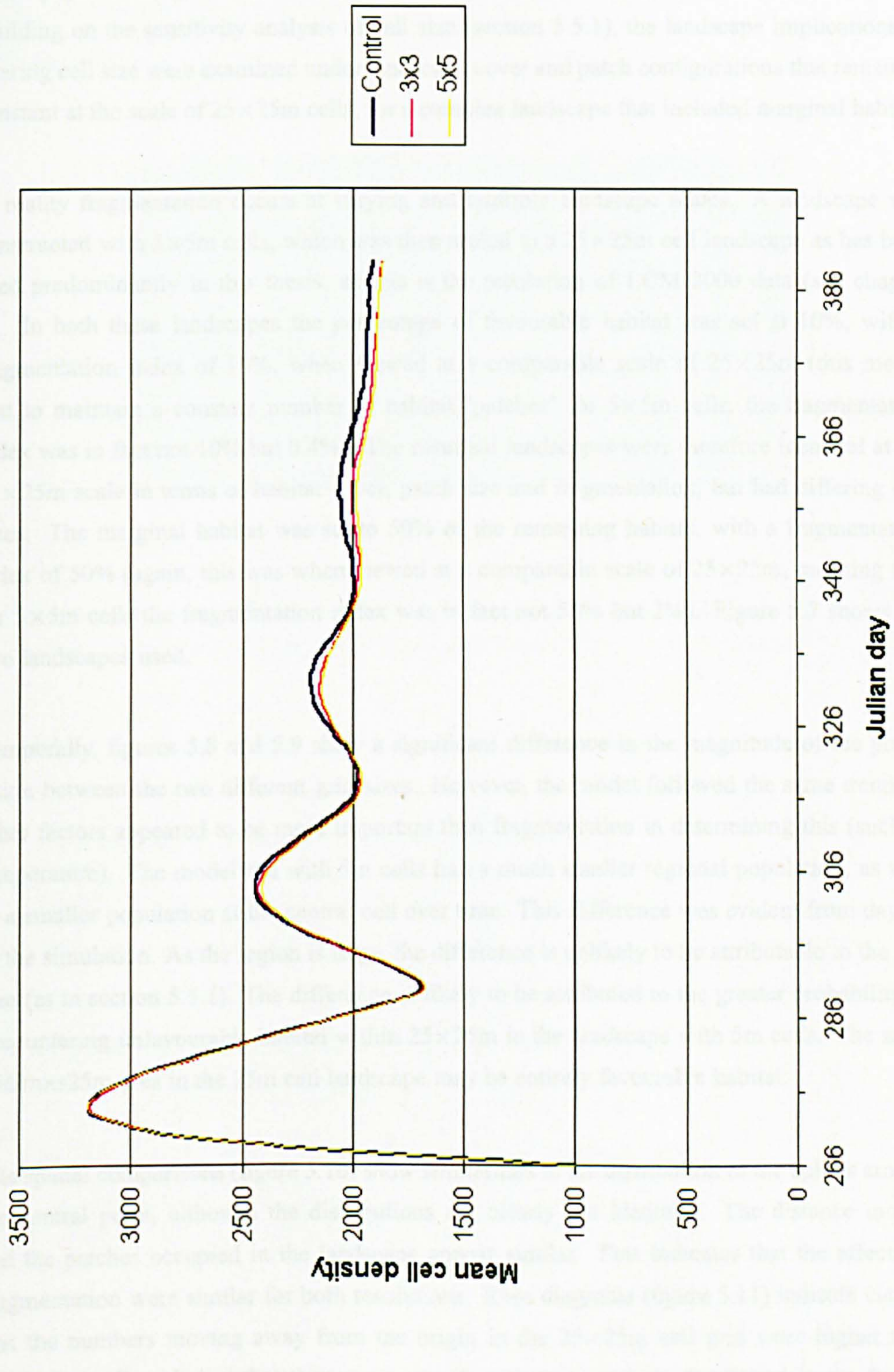


Figure 5.6: Impact of varying grid size on cell density (per 25m × 25 m cell).

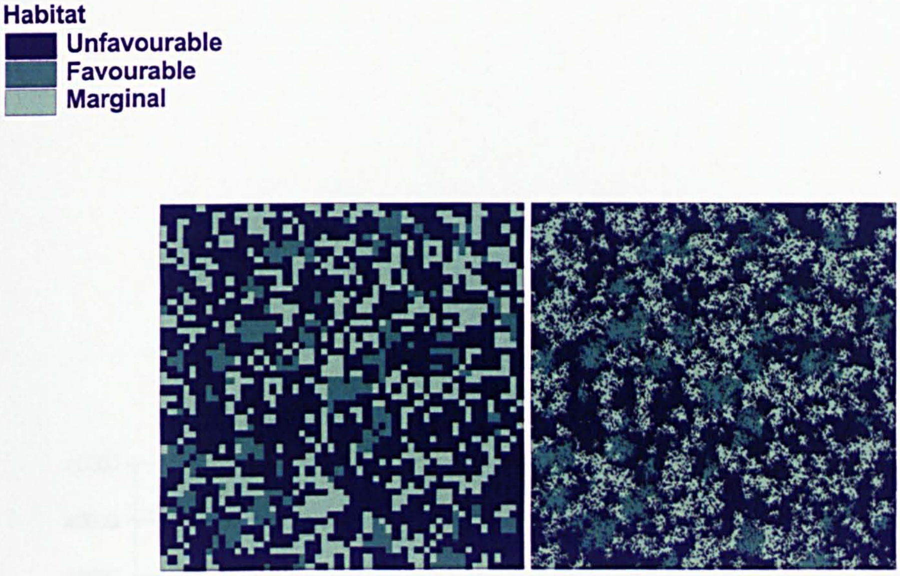
5.5.3 Regional implications of varying cell size

A number of the scenarios presented in this thesis focus on the effects of fragmentation upon aphid population dynamics and movement in theoretical landscapes (see chapter 7 and 8). Building on the sensitivity analysis of cell size (section 5.5.1), the landscape implications of altering cell size were examined under landscape cover and patch configurations that remained constant at the scale of 25×25m cells, for a complex landscape that included marginal habitat.

In reality fragmentation occurs at varying and multiple landscape scales. A landscape was constructed with 5×5m cells, which was then scaled to a 25×25m cell landscape as has been used predominantly in this thesis, as this is the resolution of LCM 2000 data (see chapter 3). In both these landscapes the percentage of favourable habitat was set at 10%, with a fragmentation index of 10%, when viewed at a comparable scale of 25×25m (this means that to maintain a constant number of habitat 'patches' for 5×5m cells, the fragmentation index was in fact not 10% but 0.4%). The resultant landscapes were therefore identical at the 25×25m scale in terms of habitat cover, patch size and fragmentation, but had differing cell sizes. The marginal habitat was set to 50% of the remaining habitat, with a fragmentation index of 50% (again, this was when viewed at a comparable scale of 25×25m, meaning that for 5×5m cells the fragmentation index was in fact not 50% but 2%). Figure 5.7 shows the two landscapes used.

Temporally, figures 5.8 and 5.9 show a significant difference in the magnitude of the population between the two different grid sizes. However, the model followed the same trend, as other factors appeared to be more important than fragmentation in determining this (such as temperature). The model run with 5m cells had a much smaller regional population, as well as a smaller population at the central cell over time. This difference was evident from day 10 of the simulation. As the region is large, the difference is unlikely to be attributable to the cell size (as in section 5.5.1). The difference is likely to be attributed to the greater probability of encountering unfavourable habitat within 25×25m in the landscape with 5m cells. The same $25times25$m area in the 25m cell landscape may be entirely favourable habitat.

The spatial comparisons (figure 5.10) show similarities in the distribution of the aphids around the central point, although the distributions are clearly not identical. The distance moved and the patches occupied in the landscape appear similar. This indicates that the effects of fragmentation were similar for both resolutions. Rose diagrams (figure 5.11) indicate clearly that the numbers moving away from the origin in the 25×25m cell grid were higher than the 5×5m cell grid, and that there were significantly more aphids distributed in the South-easterly and North-westerly directions on the 25×25m cell grid (in the directions of the nearest patches of favourable habitat, which is separated from the origin by unfavourable habitat). However, in the North-easterly and South-westerly directions the low number of aphids that had moved from the origin were similar (there are no nearby favourable habitat patches in these directions). Thus, in a landscape of a 25×25m cell grid, the aphids are able to traverse the lanscape more easily, mainly because their spatial location is averaged across a larger area.



(a) 25×25m cell grid (b) 5×5m cell grid

Figure 5.7: Landscapes with the same habitat cover and fragmentation indices including marginal habitat, but with varying cell size: a) 25×25m cell grid, b) 5×5m cell grid.

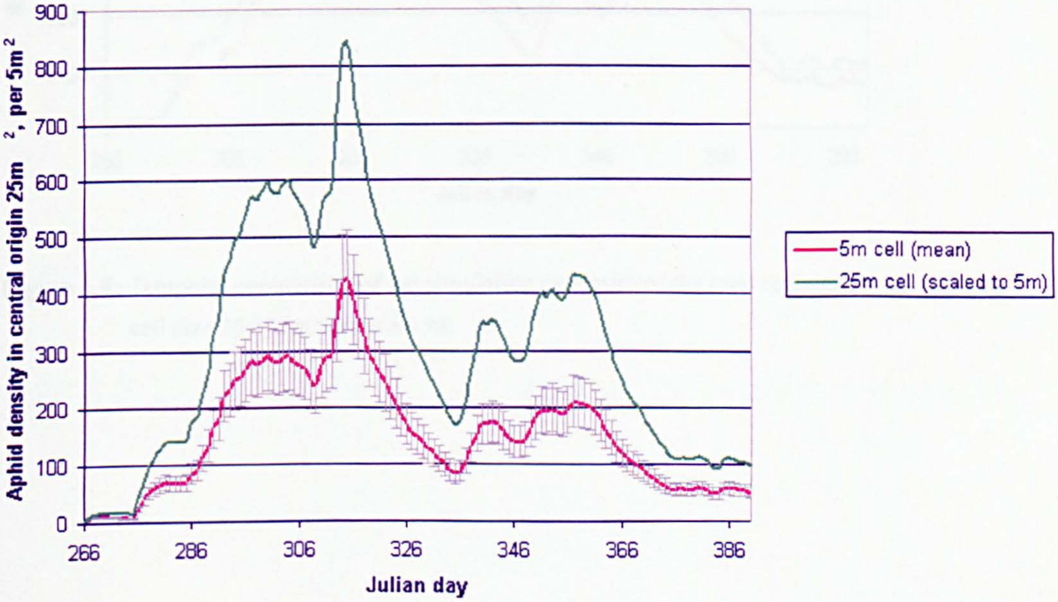


Figure 5.8: Temporal comparison of the population dynamics of the central 25×25m of the region where the immigrant population originates, scaled to 5×5m: for the 25×25m cell grid this is the central cell density scaled down to the population density per 5×5m, and for the 5×5m grid mean population density of the cells in the central 25×25m area.

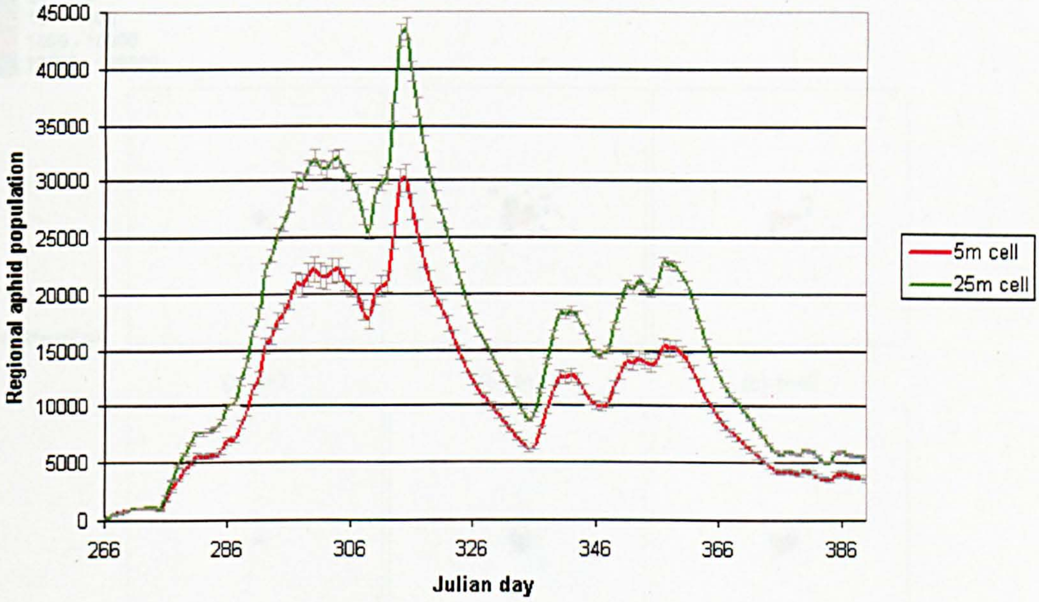
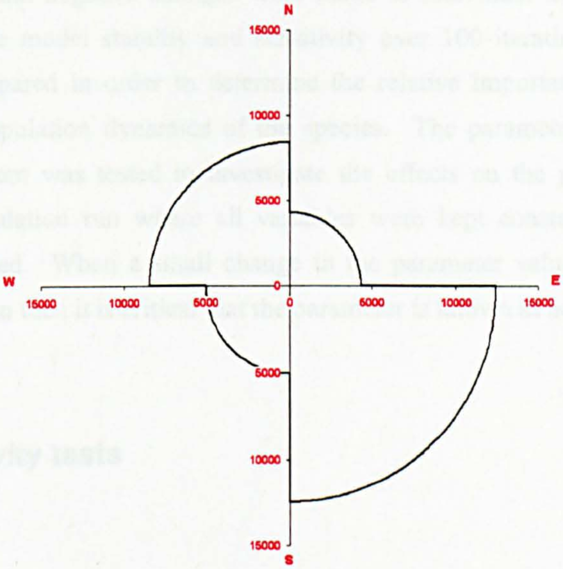


Figure 5.9: Temporal comparison of the population dynamics of the total regional population, for cell size 25 × 25m versus 5 × 5m.

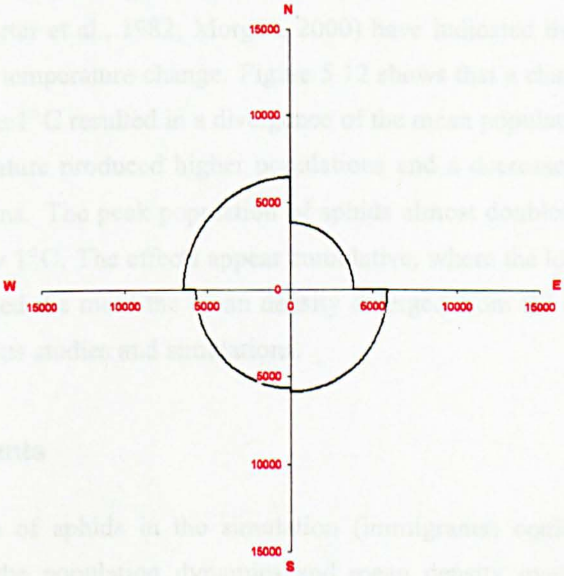
(a) t=2 (b) t=20 (c) t=40

(d) t=2 (e) t=20 (f) t=40

Figure 5.10: Spatial diffusion of aphids from the central 25×25m of the region: comparison between grid sizes.



(a) 25×25m cell grid



(b) 5×5m cell grid

Figure 5.11: Rose diagrams to illustrate the magnitude and direction of the number of aphids that diffuse from the origin at t=40, for each grid size.

5.6 Fine sensitivity test

For the temporal sensitivity tests, a single 25 × 25m cell was used. For the spatial tests, a grid of 50 × 50 25 × 25m cells was used, with a single central origin of the population. Small positive and negative changes were made to individual key parameters in the model, to examine the model stability and sensitivity over 100 iterations. The impact on predictions were compared in order to determine the relative importance of each variable in determining the population dynamics of the species. The parameters tested are shown in table 5.1. Each item was tested to investigate the effects on the population dynamics of *R. padi*, in a simulation run where all variables were kept constant throughout apart from the variable tested. When a small change in the parameter value resulted in a large change in the prediction then it is critical that the parameter is known as accurately as possible.

5.6.1 Temporal sensitivity tests

Temperature

The model's relationship with abiotic variables, in particular temperature (which also alters the intervals between lifestages), is important. Previous studies (Dean, 1974b; Walters and Dixon, 1984) and models (Carter et al., 1982; Morgan, 2000) have indicated that aphid populations are highly sensitive to temperature change. Figure 5.12 shows that a change in the fixed temperature parameter of ±1°C resulted in a divergence of the mean population over time: where an increase in temperature produced higher populations and a decrease in temperature produced lower populations. The peak population of aphids almost doubled when temperatures were increased by only 1°C. The effects appear cumulative, where the longer the temperature difference was sustained the more the mean density diverged from the control. This agrees with findings in previous studies and simulations.

Number of immigrants

The initial population of aphids in the simulation (immigrants) could potentially have a significant effect on the population dynamics and mean density over time. Figure 5.13 shows the variation in mean density over time as initial immigration was varied by ±5% and ±20%. Overall, there was no change to the trend, but reducing initial immigrant numbers produced lower peak densities and higher initial immigrant numbers increased the peak population. The results for ±5% were within the bounds of the error of the control; thus, a variation of the initial population by ±5% made little difference to the population dynamics. Reducing the initial immigration by 20% (200 aphids) resulted in a lower mean cell population by approximately 500 less aphids than the control throughout the simulation. The amplitude of the fluctuations in mean cell density was also reduced as the immigration

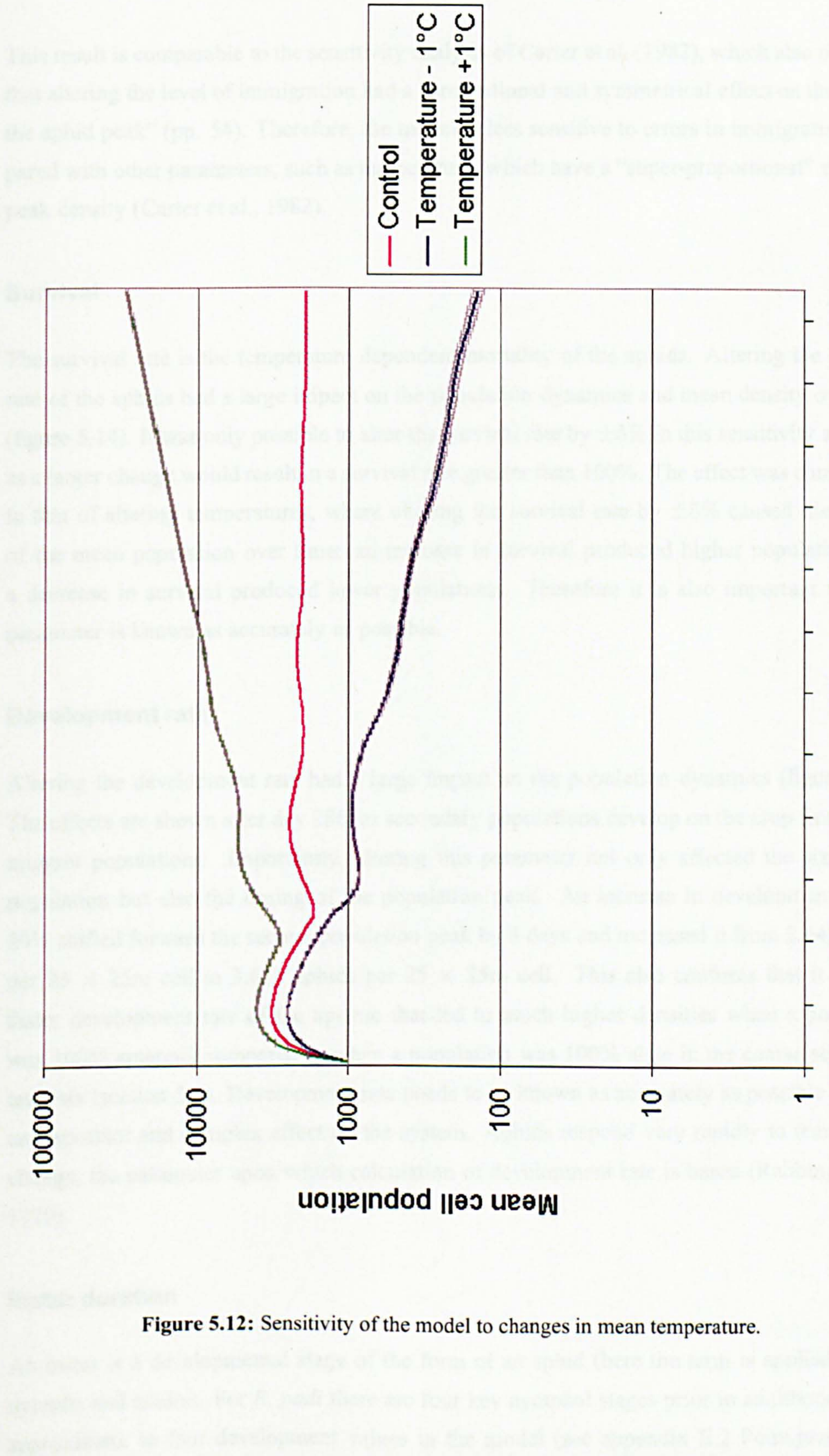


Figure 5.12: Sensitivity of the model to changes in mean temperature.

was reduced by 20%. On the other hand, there appears to be a slight increase in amplitude of the fluctuations in mean cell density when immigration was increased by 20%. Population size was only increased by about 300 aphids compared to the control under this scenario.

This result is comparable to the sensitivity analysis of Carter et al. (1982), which also observed that altering the level of immigration had a "proportional and symmetrical effect on the size of the aphid peak" (pp. 54). Therefore, the model is less sensitive to errors in immigration compared with other parameters, such as temperature, which have a "super-proportional" effect on peak density (Carter et al., 1982).

Survival

The survival rate is the temperature dependent mortality of the aphids. Altering the survival rate of the aphids had a large impact on the population dynamics and mean density over time (figure 5.14). It was only possible to alter the survival rate by $\pm 5\%$ in this sensitivity analysis, as a larger change would result in a survival rate greater than 100%. The effect was comparable to that of altering temperatures, where altering the survival rate by $\pm 5\%$ caused divergence of the mean population over time: an increase in survival produced higher populations and a decrease in survival produced lower populations. Therefore it is also important that this parameter is known as accurately as possible.

Development rate

Altering the development rate had a large impact on the population dynamics (figure 5.15). The effects are shown after day 286, as secondary populations develop on the crop from initial migrant populations. Importantly, altering this parameter not only affected the size of the population but also the timing of the population peak. An increase in development rate of 20% shifted forward the second population peak by 3 days and increased it from 2,445 aphids per 25 $\times$ 25m cell to 3,612 aphids per 25 $\times$ 25m cell. This also confirms that it was the faster development rate of the apterae that led to much higher densities when a population was 100% apterous compared to when a population was 100% alate in the coarse sensitivity analysis (section 5.4). Development rate needs to be known as accurately as possible as it has an important and complex effect on the system. Aphids respond very rapidly to temperature change, the parameter upon which calculation of development rate is based (Rabbinge et al., 1979).

Instar duration

An instar is a developmental stage of the form of an aphid (here the term is applied to both nymphs and adults). For *R. padi* there are four key nymphal stages prior to adulthood, which approximate to four development values in the model (see appendix E.2 Point.java, where

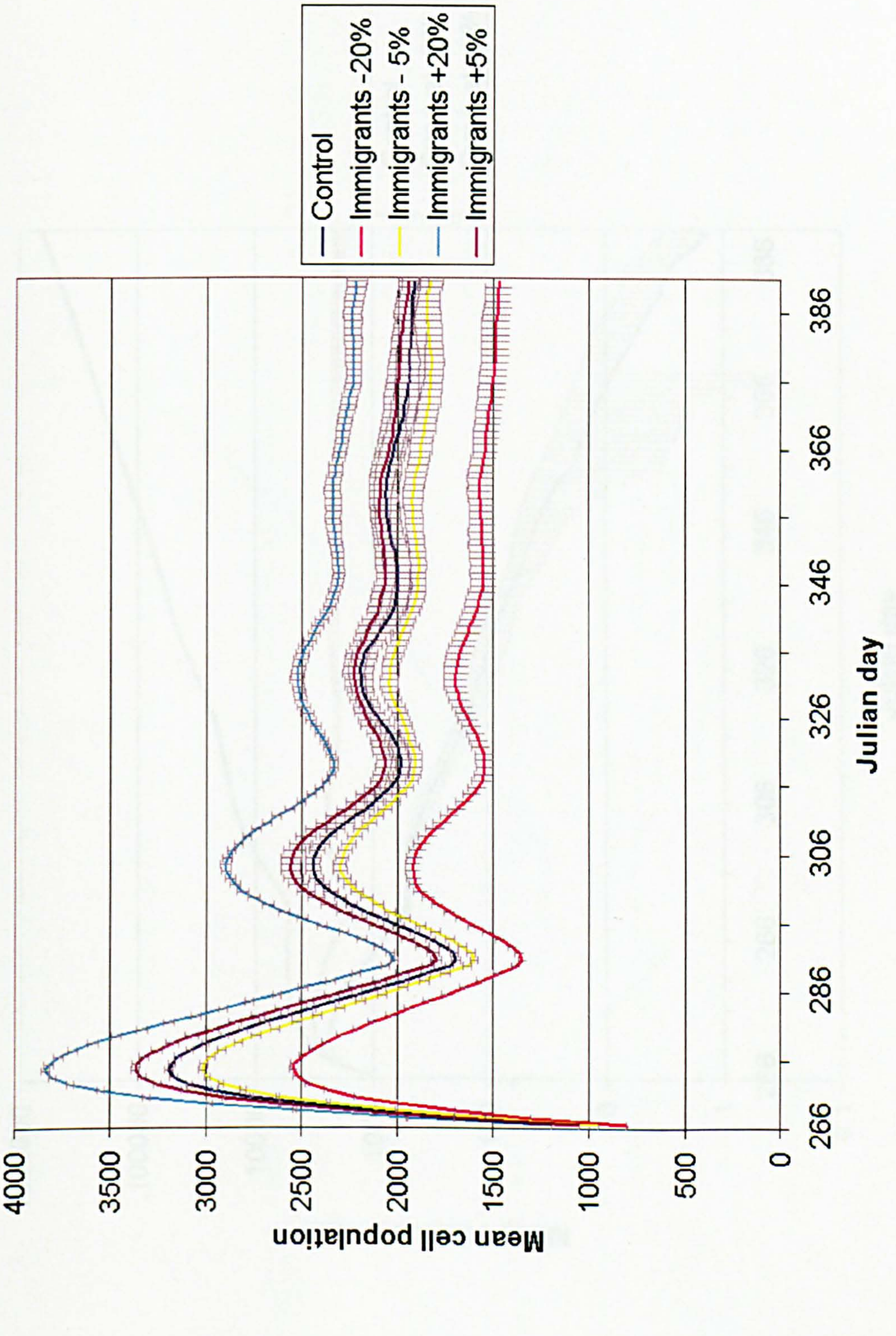


Figure 5.13: Sensitivity of the model to changes in initial immigrant numbers.

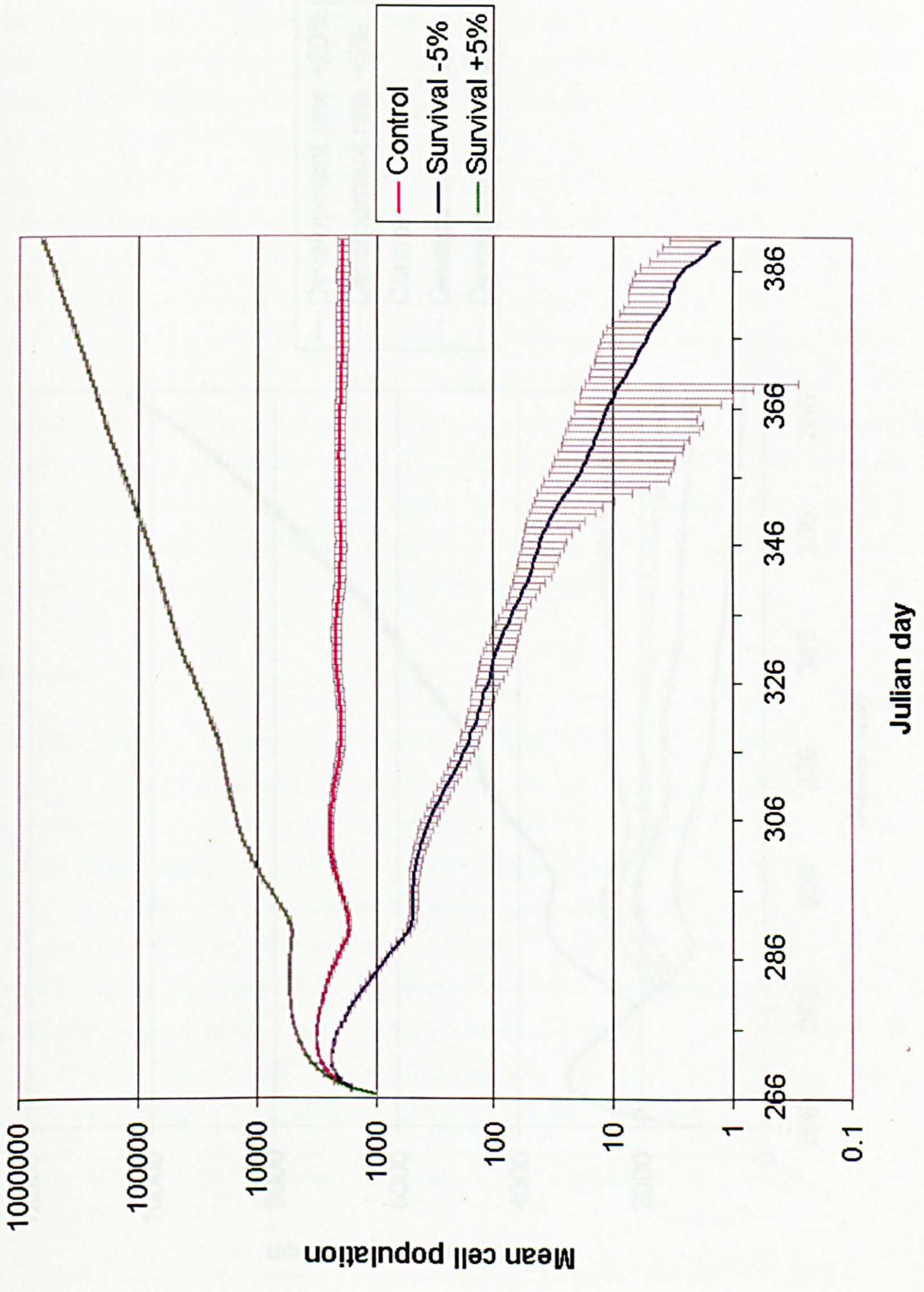


Figure 5.14: Sensitivity of the model to changes in the survival rate.

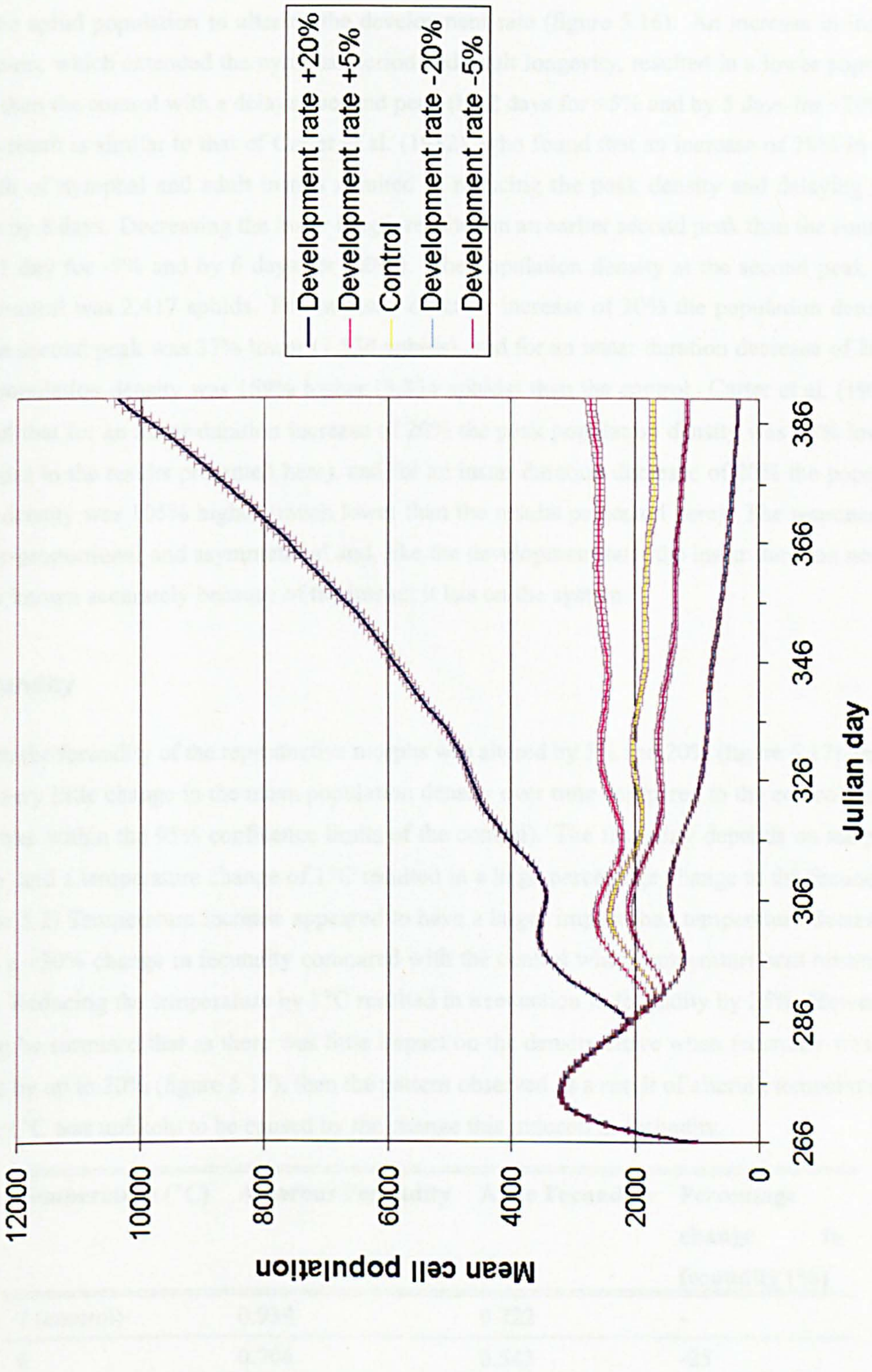


Figure 5.15: Sensitivity of the model to changes in development rate.

these values are given).

Altering the length of all instar durations (thus the total nymphal period) had the inverse effect on the aphid population to altering the development rate (figure 5.16). An increase in instar duration, which extended the nymphal period and adult longevity, resulted in a lower population than the control with a delayed second peak (by 2 days for +5% and by 5 days for +20%). This result is similar to that of Carter et al. (1982), who found that an increase of 20% in the length of nymphal and adult instars resulted in reducing the peak density and delaying the peak by 8 days. Decreasing the instar length resulted in an earlier second peak than the control (by 1 day for -5% and by 6 days for -20%). The population density at the second peak for the control was 2,417 aphids. For an instar duration increase of 20% the population density at the second peak was 37% lower (1,530 aphids), and for an instar duration decrease of 20% the population density was 159% higher (3,834 aphids) than the control. Carter et al. (1982) found that for an instar duration increase of 20% the peak population density was 33% lower (similar to the results presented here), and for an instar duration decrease of 20% the population density was 105% higher (much lower than the results presented here). The response is super-proportional and asymmetrical and, like the development rate, the instar duration needs to be known accurately because of the impact it has on the system.

Fecundity

When the fecundity of the reproductive morphs was altered by 5% and 20% (figure 5.17), there was very little change in the mean population density over time compared to the control (each test was within the 95% confidence limits of the control). The fecundity depends on temperature, and a temperature change of 1°C resulted in a large percentage change to the fecundity (table 5.2) Temperature increase appeared to have a larger impact than temperature decrease, with a +30% change in fecundity compared with the control when temperature was raised by 1°C. Reducing the temperature by 1°C resulted in a reduction in fecundity by 25%. However, it can be surmised that as there was little impact on the density curve when fecundity was altered by up to 20% (figure 5.17), then the pattern observed as a result of altering temperatures by ±1°C was unlikely to be caused by the change this induced in fecundity.

Temperature (°C)	**Apterous Fecundity**	**Alate Fecundity**	**Percentage change in fecundity (%)**
7 (control)	0.939	0.722	-
6	0.706	0.543	-25
8	1.226	0.943	+30

Table 5.2: Relationship between temperature change of 1°C and Fecundity.

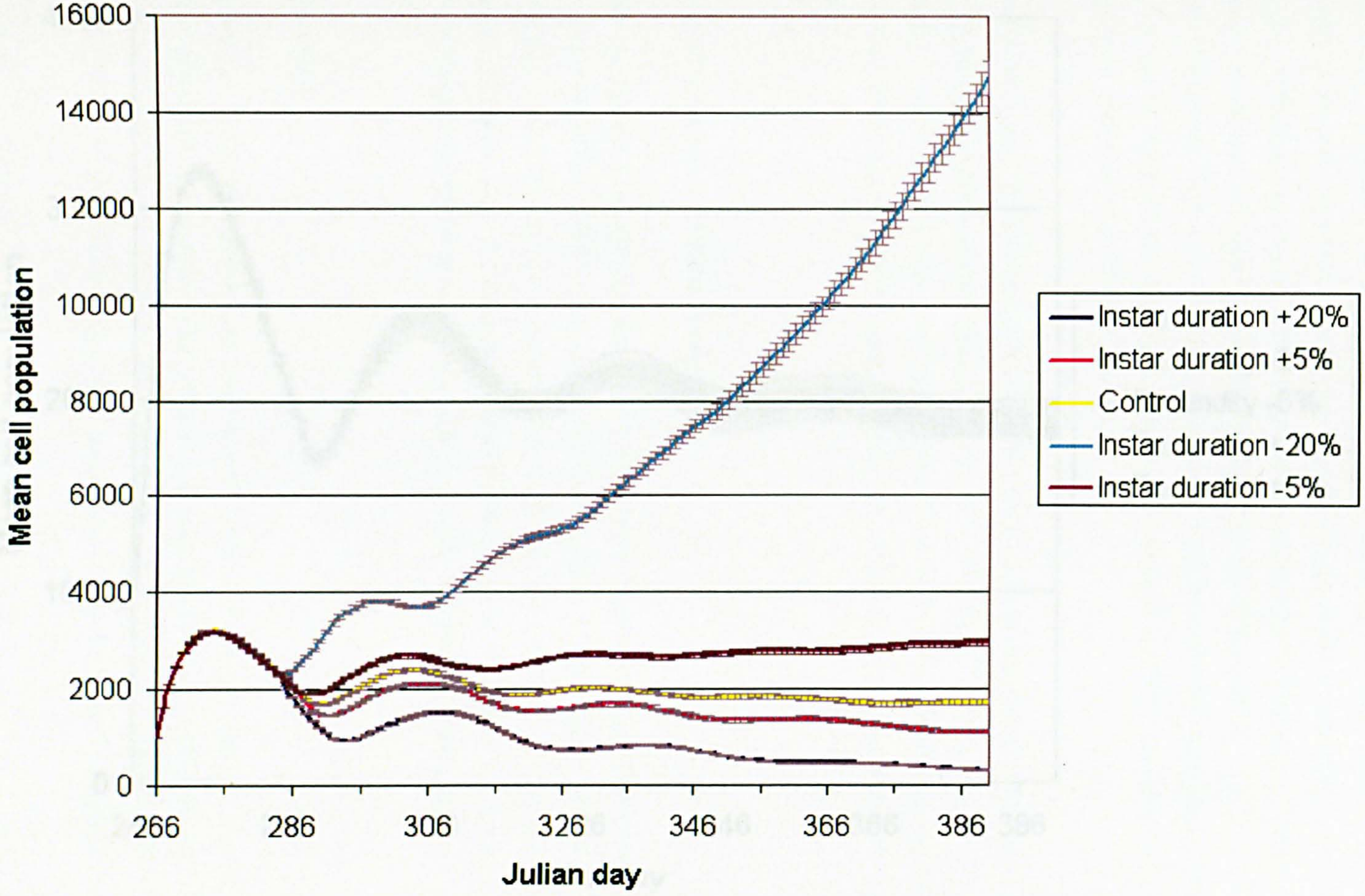


Figure 5.16: Sensitivity of the model to changes in instar duration.

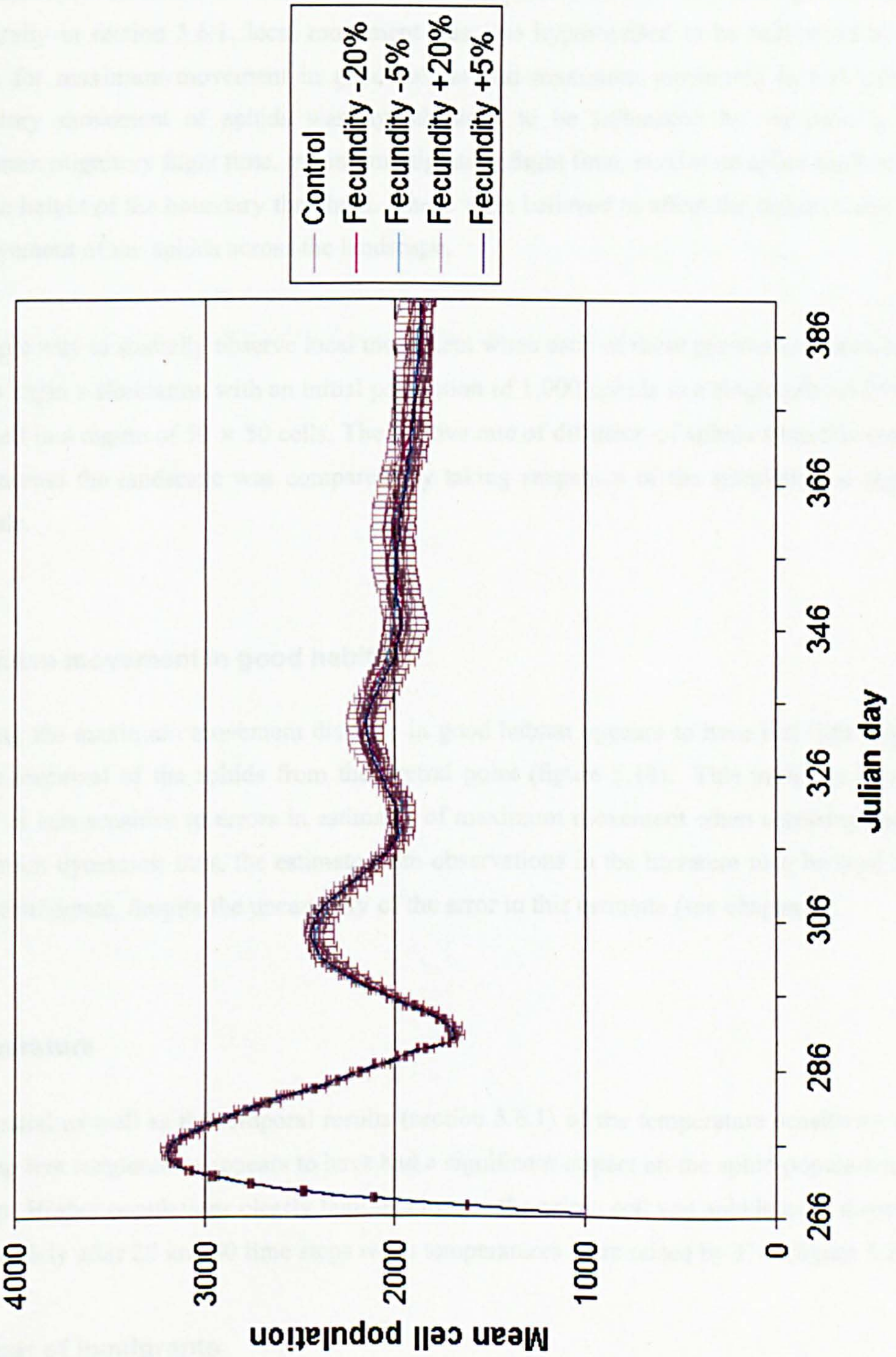


Figure 5.17: Sensitivity of the model to changes in aphid fecundity.

5.6.2 Diffusion: spatial sensitivity tests

Several parameters were thought to have a key influence on the spread of aphids across the landscape. In addition to the parameters temperature, survival and immigration tested temporally in section 5.6.1, local movement was also hypothesised to be influenced by the values for maximum movement in good habitat and maximum movement in bad habitat. Migratory movement of aphids was hypothesised to be influenced by variation in the maximum migratory flight time, minimum migratory flight time, maximum aphid flight speed and the height of the boundary threshold. These were believed to affect the distance and rate of movement of the aphids across the landscape.

A simple way to spatially observe local movement when each of these parameters was altered was to begin a simulation with an initial population of 1,000 aphids in a single central 25m × 25m cell in a region of 50 × 50 cells. The relative rate of diffusion of aphids from this central point across the landscape was compared, by taking snapshots of the simulation at regular intervals.

Maximum movement in good habitat

Altering the maximum movement distance in good habitat appears to have had little impact on the dispersal of the aphids from the central point (figure 5.18). This indicates that the model is less sensitive to errors in estimates of maximum movement when assessing spatial population dynamics; thus, the estimate from observations in the literature may be used with some confidence, despite the uncertainty of the error in this estimate (see chapter 3).

Temperature

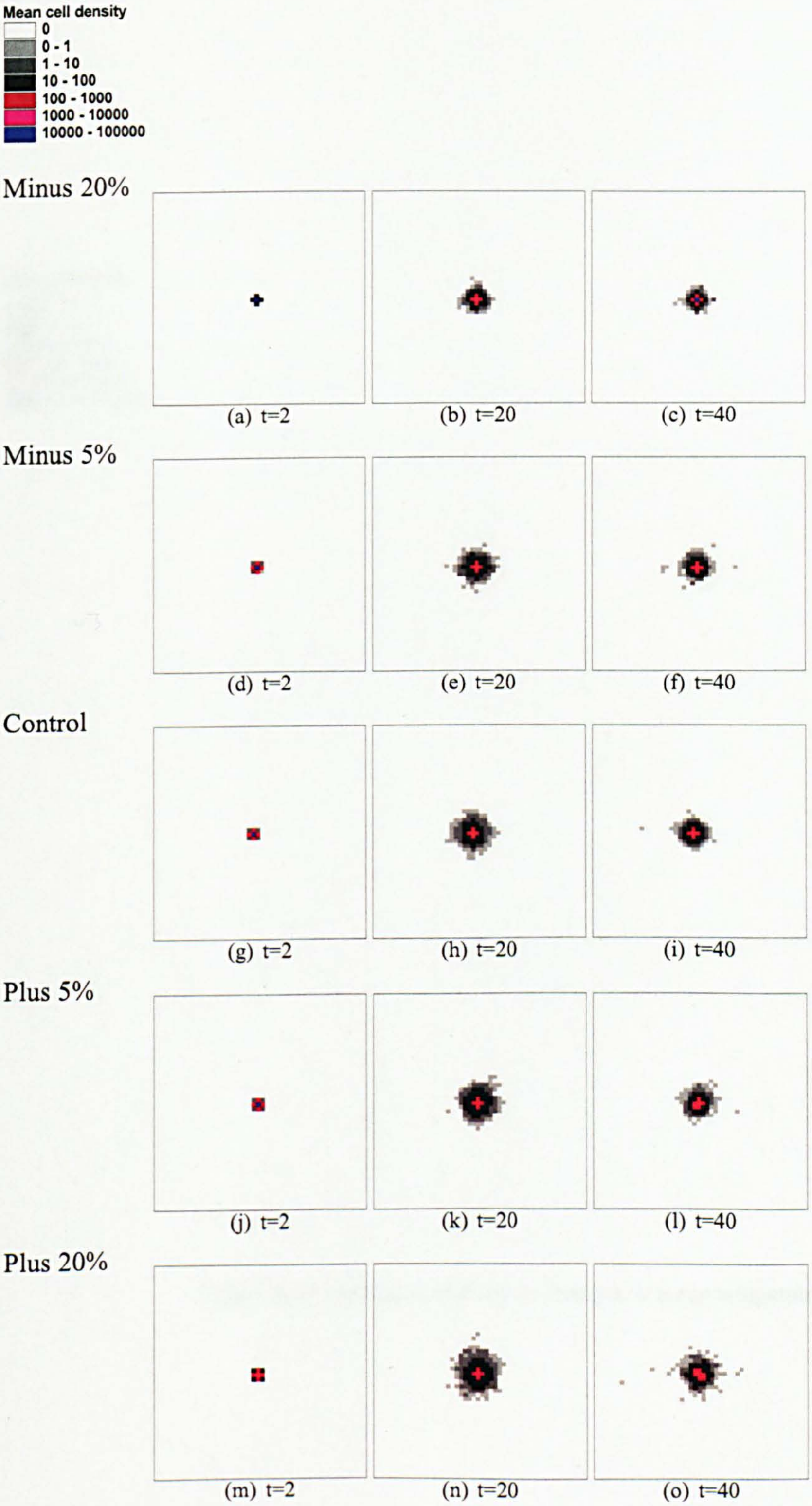
The spatial as well as the temporal results (section 5.6.1) of the temperature sensitivity tests indicate that temperature appears to have had a significant impact on the aphid population dynamics. Higher populations clearly built up close to the origin cell and aphids were dispersed more widely after 20 and 40 time steps when temperatures were raised by 1°C (figure 5.19).

Number of immigrants

At the scale of 25 × 25 m cells, immigration does not appear to have had an impact upon spatial population dynamics (figure 5.20). This confirms that spatially the model is less sensitive to errors in immigration compared with other parameters, as also shown temporally (section 5.6.1).

Survival

Altering the survival rate also had spatial as well as temporal implications. These were more pronounced spatially than the effects of temperature (figure 5.21). When survival was reduced by 5% the aphid density at the central cell drops below 100 aphids per 25×25m, and dispersal from the origin was low. Aphids were more dispersed around the origin and reached much higher densities at the origin (above 10,000 aphids per 25×25m) when the survival rate was increased by 5%. Survival rates were derived by comparing expected and observed population densities of aphids, exposed to the effects of low temperatures under field conditions (Morgan, 2000). Aphid populations are therefore highly transient depending on temperature, and if mortality is increased at key stages, by either natural or other means such as pesticide, control of the population can be achieved with a relatively small increase in mortality. This is explored further in chapter 9.



(a) t=2 (b) t=20 (c) t=40

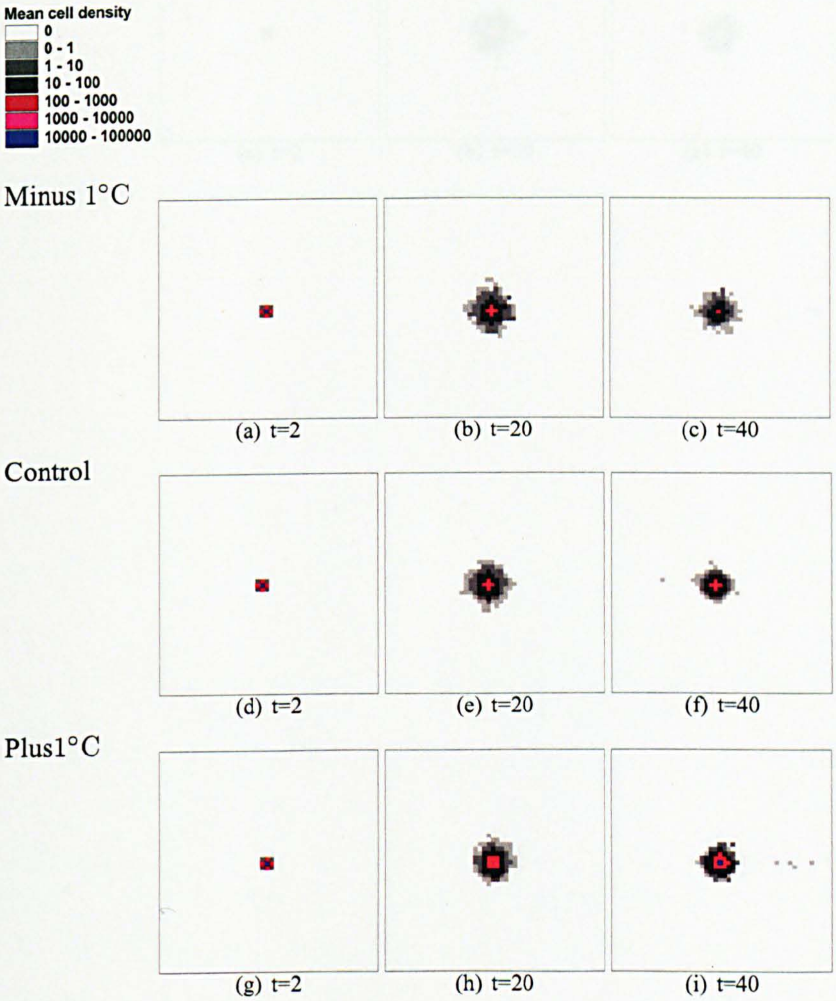
(d) t=2 (e) t=20 (f) t=40

(g) t=2 (h) t=20 (i) t=40

(j) t=2 (k) t=20 (l) t=40

(m) t=2 (n) t=20 (o) t=40

Figure 5.18: Spatial sensitivity to changes in the maximum movement in good habitat.

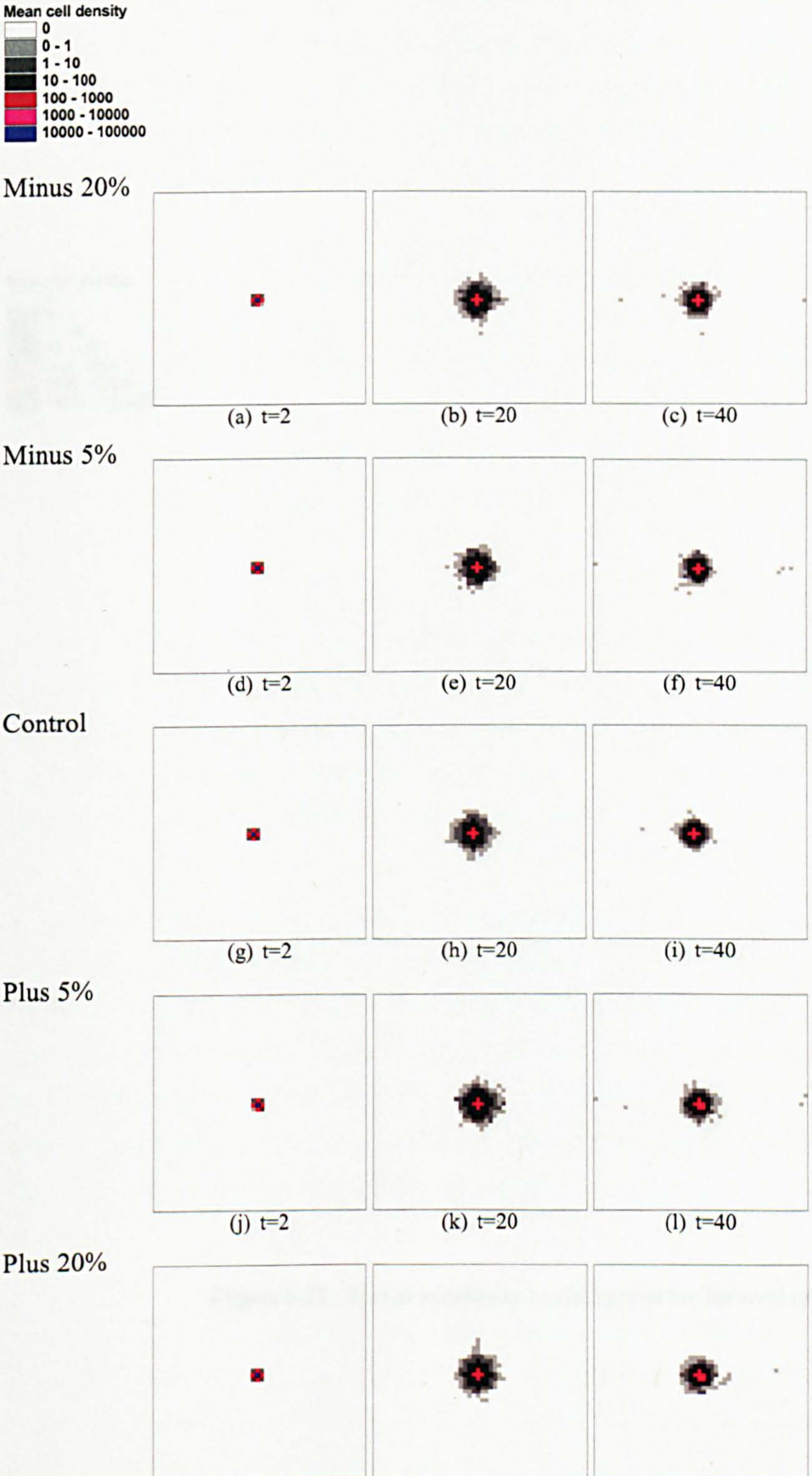


(a) t=2 (b) t=20 (c) t=40

(d) t=2 (e) t=20 (f) t=40

(g) t=2 (h) t=20 (i) t=40

Figure 5.19: Spatial sensitivity to changes in mean temperature.



(a) t=2 (b) t=20 (c) t=40

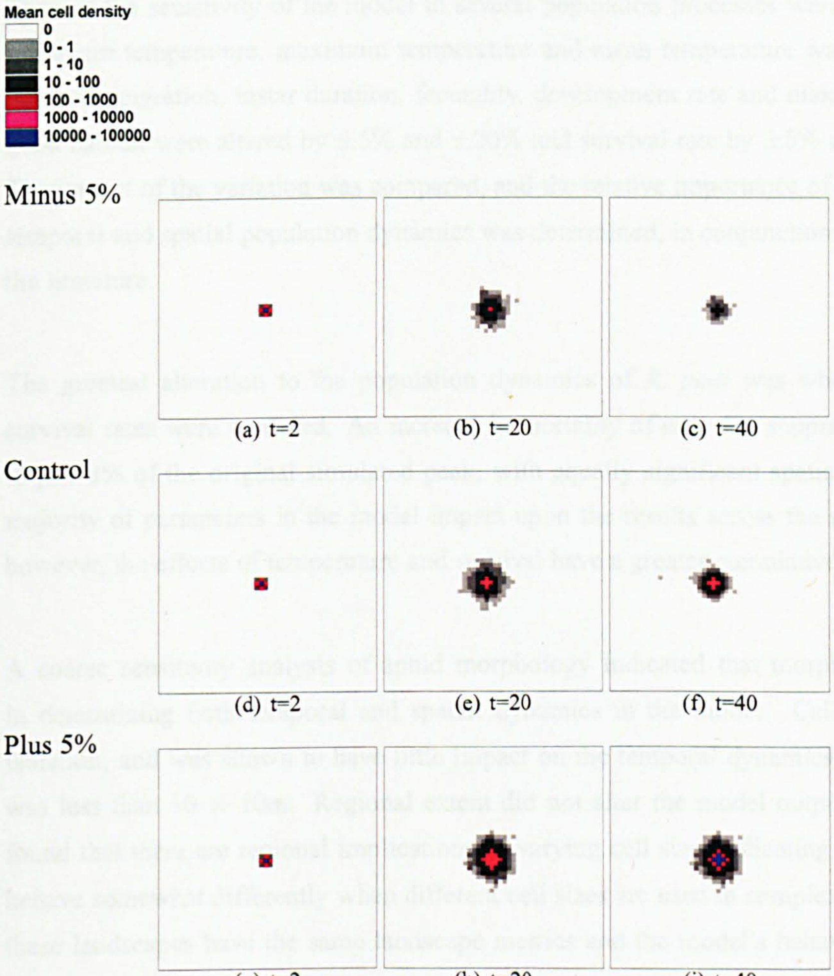
(d) t=2 (e) t=20 (f) t=40

(g) t=2 (h) t=20 (i) t=40

(j) t=2 (k) t=20 (l) t=40

(m) t=2 (n) t=20 (o) t=40

Figure 5.20: Spatial sensitivity to changes in initial immigration numbers.



(a) t=2 (b) t=20 (c) t=40

(d) t=2 (e) t=20 (f) t=40

(g) t=2 (h) t=20 (i) t=40

Figure 5.21: Spatial sensitivity to changes in the survival rate.

5.7 Summary

Continuous model assessment has been undertaken at three key stages in the modelling process: at the conceptual stage, development stage and operational stage. This assessment qualifies the knowledge base of the model (largely presented in earlier chapters), verifies the computation and data, and provides both empirical validation of the model and validation in the form of a sensitivity analysis.

Tests of the sensitivity of the model to several population processes were carried out. Daily minimum temperature, maximum temperature and mean temperature was altered by ±1°C, initial immigration, instar duration, fecundity, development rate and maximum movement in good habitat were altered by ±5% and ±20% and survival rate by ±5% of the control value. The impact of the variation was compared, and the relative importance of each variable to the temporal and spatial population dynamics was determined, in conjunction with evidence from the literature.

The greatest alteration to the population dynamics of *R. padi* was when temperature and survival rates were modified. An increase in mortality of only 5% suppressed peak densities to just 8% of the original simulated peak, with equally significant spatial implications. The majority of parameters in the model impact upon the results across the simulation timeline; however, the effects of temperature and survival have a greater cumulative effect.

A coarse sensitivity analysis of aphid morphology indicated that morphology is important in determining both temporal and spatial dynamics in the model. Cell size was tested in isolation, and was shown to have little impact on the temporal dynamics unless the cell size was less than 10 × 10m. Regional extent did not alter the model output. However, it was found that there are regional implications of varying cell size, indicating that the model may behave somewhat differently when different cell sizes are used in complex landscapes, even if these landscapes have the same landscape metrics and the model's behavioural rules remain constant.

The sensitivity analysis indicated that, overall, the model is robust: it does not become chaotic when parameters were varied and the effects of changing a number of variables were nearly linear. In reality, aphids tend to maximise population growth by producing large numbers of off-spring with a rapid development cycle, with the ability to disperse easily to colonise new environments, but are highly sensitive to environmental conditions. Such organisms are generally termed 'r-strategists' in the ecological literature (MacArthur and Wilson, 1967). Evidence from the sensitivity analysis confirms the model reflects such behaviour: exponential population growth was evident when the limits to the population growth were reduced (i.e. survival rate), or the environmental conditions were made more favourable (e.g. increased temperatures).

Calibration of the model was not performed, as it is more revealing to use the model to investigate why there are discrepancies between the observed and predicted results, thus highlighting key errors in our understanding of landscape-insect relationships.

CHAPTER 6

Techniques for Simulating Large Numbers of Individuals

6.1 Millions of insects: pushing the limits of individual-based simulation?

The construction, parameterisation and functionality of the model are described and illustrated in chapter 4. A key limitation of the model was the restriction on the number of insect agents that can be modelled. This restriction is due to the limits imposed by the processing power and memory availability of the single 2.80 GHz Intel Xeon processor 2097 MB RAM machine used. Figure 6.1 and figure 6.2 show the memory and simulation speed curves respectively of the original model run on this machine, with increasing numbers of initial immigrant aphids, using the sensitivity test control parameters given in chapter 4. These curves can be represented by equations 6.1 and 6.2.

$$m = 2E^{-04}n + 2.8115 \tag{6.1}$$

Where m = maximum memory required (MB) and n = number of immigrant aphids ($r^2 = 1.0$).

$$\mathrm{t} = 2E^{-08}n^2 - 3E^{-04}n + 14.088 \tag{6.2}$$

Where t = run time of simulation and n = number of immigrant aphids ($r^2 = 1.0$).

With a maximum available memory capacity of 1.5GB RAM on the single machine, the theoretical limit to the number of immigrants is n=7,500,000. However, from equation 6.2, this simulation would take approximately 1 million seconds to run (approximately 312 hours or 13 days!). In a simulation using realistic stochastic input variables this immigration limit is likely to be much reduced, as the population peak relative to the immigrant numbers could be much higher.

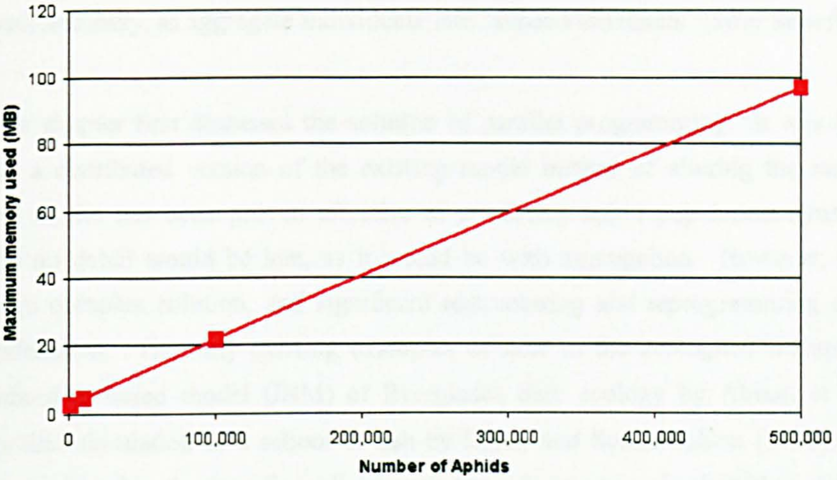


Figure 6.1: Plot of maximum memory used in a simulation run against number of immigrant aphids for the model

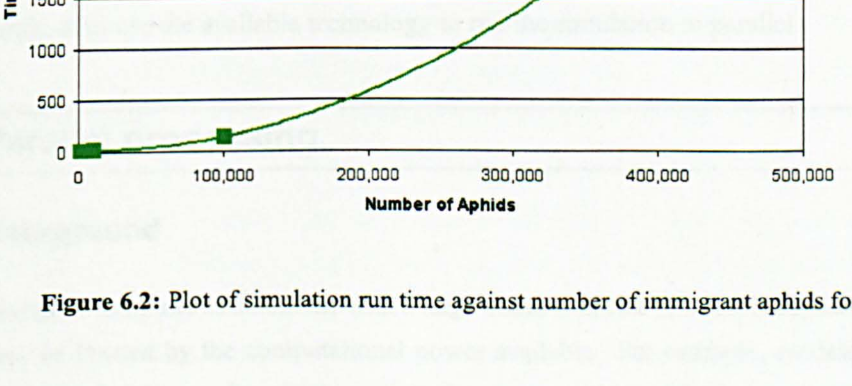


Figure 6.2: Plot of simulation run time against number of immigrant aphids for the model

Although the model is therefore comfortably·able to cope with several hundred thousand insects, insects such as aphids can number in their millions over relatively small areas in the field. There are two choices to alter the model to incorporate more insects; firstly, to parallel program the model to work across a network of powerful computers, so splitting the data load; secondly, to aggregate individuals into 'super-individuals' (after Scheffer et al., 1995).

This chapter first discusses the solution of parallel programming. It was believed better to use a distributed version of the existing model instead of altering the model structure, as this model has been proven effective at predicting aphid population dynamics (chapter 4) and no detail would be lost, as it would be with aggregation. However, this was also the more complex solution, and significant restructuring and reprogramming of the model was undertaken. The only existing examples of note in the ecological literature are a parallel individual-based model (IBM) of Everglades deer ecology by Abbott et al. (1997) and a parallel simulation of a school of fish by Lorek and Sonnenschein (1995). Other examples can be found in the use of parallel agents for reducing genetic algorithm search times (Lefley and McKew, 2004) and performing large scale traffic simulations (Dupuis and Chopard, 2001).

The simplicity of the super-individual approach makes it attractive, particularly as it does not require complex programming and powerful computer systems to implement. It maintains the philosophy of an individual-based approach without reverting to a population model to deal with large numbers of individuals. Implementations of this approach to date are primarily temporal. Spatially, the approach potentially has a large impact on the results of the simulation as the aggregated individuals result in a more clustered population than predicted by a purely individual-based simulation. The extent of this effect is investigated in the second part of this chapter, and the consequences for the use of the super-individual technique in spatial simulation are assessed.

The final section of this chapter compares the aggregated super-individual model with the parallel model, showing the implications of grouping individuals in the model and concluding that the best approach for an individual-based spatially explicit simulation, as presented in this thesis, is to use the available technology to run the simulation in parallel.

6.2 Parallel processing

6.2.1 Background

Complex computer simulations, with a large number of calculations that need to be performed, may be limited by the computational power available. For example, models used to predict atmospheric change for weather prediction use a cubic grid where calculations are made at each vertex. These models can require over a trillion calculations per second for a large geographical area such as the United States and Canada (Pacheco, 1997). Increasing the size of individual computers used in terms of memory and processing power would perhaps make the calculations possible, but they would take a very long time to perform. However, by using multiple processors and distributed memory working simultaneously, the scale of the problem for each individual computer is much reduced, and so simulations can run in a fraction of the time that would be taken to perform the same complex, memory intensive, operations. This is the essence of parallel computing.

Increasing memory by parallel programming

Many computing problems are hindered by their data requirements, which may be far larger than can be accommodated at any individual processing element (Chalmers and Tidmus, 1996). This was a key problem with the simulation developed for this thesis, as the data for each insect object in a landscape scale simulation quickly exceeded the storage capacity of an ordinary PC with up to 2097 MB of RAM. The combined or 'virtual shared' memory of several computers was used to cope with the amount of data needed, using a Sequential-Algorithm Multiple-Data approach (SAMD), where the same algorithm is applied to different data items on different processors. This required significant data management strategies to make the physically distributed systems act as a single memory unit.

Message passing

Although quite simple in concept, distributing a program across several computers (or nodes) in a cluster can be complex, and may involve significant rewriting of code. Message passing (MP) is the principal way by which parallel clusters of machines are programmed. It is a widely-used, powerful and general method of distribution and creating efficient programs (Pacheco, 1997). Key advantages of using MP are an ability to scale to many processors, flexibility, 'future-proofing' of programs and portability (Openshaw and Turton, 2000). However, there is a great deal of effort involved in detecting, enhancing and expressing parallelisation to its full benefit, often meaning that code and algorithms have to be rewritten to be suitable for MP. According to Openshaw and Turton (2000: pp. 179) this need not be too difficult, and 'it frequently leaves you with a feeling of euphoria and achievement when it finally works... it is all rather an immense intellectual challenge of considerable excitement'.

Thus motivated, the model was reprogrammed in parallel using a Message-passing Interface (MPI) for Java, MPIJava (`http://www.hpjava.org`). MPIJava is a library of subroutine specifications, called by Java programs (not a language). Similar interfaces are available for C and Fortran (Pacheco, 1997). MPIJava enables communication between nodes of a distributed memory parallel computer. It is a Java wrapper implementation of MPI, rather than a pure Java implementation. This means that it is written in native code and native methods are called by the programme using the Java Native Interface (JNI). Pure Java implementations have the advantage of being more portable as the whole library is developed in Java, but communication is not as efficient (Taboada et al., 2003). MPIJava uses the open-source native MPI 'LAM' (`http://www.lam-mpi.org/`).

6.2.2 Implementation

Hardware: the Beowulf cluster

The architecture used consisted of a distributed-memory system of parallel computers in a multiple-instruction, multiple-data (MIMD) configuration: where a collection of autonomous processors operate on their own data streams (see figure 6.3). The system comprised of two Beowulf clusters (`http://www.beowulf.org`) at the School of Geography, University of Leeds. A cluster is a set of independent computers combined into a unified system through software and networking. A Beowulf Cluster is "a scalable performance cluster based on commodity hardware, on a private system network, with open source software (Linux) infrastructure" (`http://www.beowulf.org/overview/`). The commodity hardware used in the Leeds cluster comprised thirty machines (nodes), with access and control provided by two master nodes. Each node has dual 2.66 GHz Intel Xeon processors with 1280 MB of dual channel PC2100 Dial-on-Demand Routing (DDR) memory, 40GB 7200rpm internal Integrated Drive Electronics (IDE) disks running over a switched GB network. Common uses of Beowulf clusters include technical applications such as "simulations, biotechnology, and petro-clusters; financial market modelling, data mining and stream processing; and Internet servers for audio and games" (`http://www.beowulf.org/overview/`). Beowulf programs use message passing to construct parallel computations.

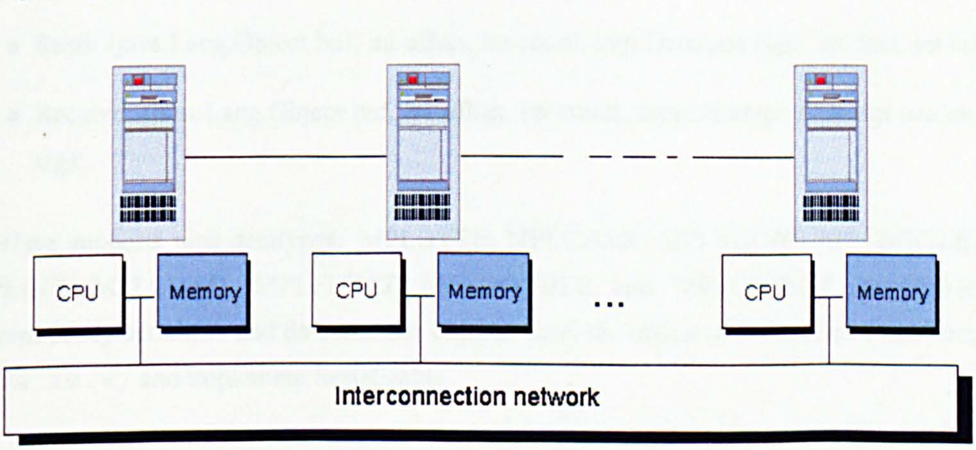


Figure 6.3: A MIMD network with interconnected separate memory and processors (after Pacheco, 1997: pp. 19).

Incorporating the MPI

In order for message passing to be made possible using MPIJava, the MPI package has to be imported at the top of the main class file (`import mpi.*;`). The args String array that is in the main method is used to specify the number of processors to be used to run the program. Thus the following method call is included within the main method to initialise the MPI: `MPI.Init(args);`. This method must be called before any other MPI command is used. Within the main method two key integer variables are set when using MPI. These are:

- `nodeRank = MPI.COMM_WORLD.Rank();`
- `numberOfNodes = MPI.COMM_WORLD.Size();`

`MPI.COMM_WORLD.Rank();` assigns the nodeRank variable to the processor number for each processor used by the program; thus, the program is then able to recognise which parts of the code should be run by any given processor. `MPI.COMM_WORLD.Size();` determines how many processors are being used to run the program, thus allowing data and simulations to be divided equally between processors (a message passing example is given in figure 6.4).

In order to send data between processors the following two commands are used:

- `MPI.COMM_WORLD.Send();`
- `MPI.COMM_WORLD.Recv();`

Each method takes in a similar set of parameters referring to the datatype sent or received, the source node or the destination node.

- Send: (java.Lang.Object buf, int offset, int count, mpi.Datatype type, int dest, int tag)
- Receive: (java.Lang.Object buf, int offset, int count, mpi.Datatype type, int source, int tag);

MpiJava supports nine datatypes: MPI.BYTE, MPI.CHAR, MPI.SHORT, MPI.BOOLEAN, MPI.INT, MPI.LONG, MPI.FLOAT, MPI.DOUBLE and MPI.OBJECT. MPI.OBJECT automatically serializes and de-serializes objects; thus, the object to be serialised must import `java.io.*;` and implement Serializable.

To receive data properly the processor receiving needs to match the tag argument of the processor sending. Also, the memory available for receiving the message, specified by the buffer, count and datatype parameters, must be at least as large as the message sent. To avoid wasting memory this often means that passing arrays between processors is a two stage process, where firstly the size of the array is sent as an MPI.INT datatype and then the array itself is sent.

To send an integer array ``intArray'' from the control node (node 0) to each node, the first example loop in figure 6.4 could be used on node 0. Similarly, to receive an integer array from each of the worker nodes, the second loop could be used on node 0.

```
for (int node =1; node < numberOfNodes; node++) {
        try {
              MPI.COMM_WORLD.Send(intArray, 0, arraySize, MPI.INT, node, 50);
        } catch (MPIException mpiE) {
              mpiE.printStackTrace();
        }
}
for (int node =1; node < numberOfNodes; node++) {
        try {
              MPI.COMM_WORLD.Recv(intArray, 0, arraySize, MPI.INT, node, 50);
        } catch (MPIException mpiE) {
              mpiE.printStackTrace();
        }
}
```

Figure 6.4: Demonstration of the use of a loop with the integer `numberOfNodes` to send and receive messages via the MPI.

Data mapping

It is important to minimise communication between nodes in parallel computing, as communication is computationally expensive due to networking speeds. It is also important to ensure the data are distributed evenly between the nodes, called 'load balancing' (Pacheco, 1997). This ensures that the best use is made of the available resources.

Even distribution of data between nodes was achieved by splitting immigrant agents evenly across the system, with each node containing information on the environment and local densities passed from the control node. The number of immigrants to be added to each node was calculated by a form of 'block mapping' (Pacheco, 1997: pp. 35), which partitioned the number of immigrants into blocks which were then assigned to each node. So, if there were three nodes ($n = 3$) and thirteen immigrants ($i = 13$), the immigrants mapped to each node would be as follows:

$$i_0, i_1, i_2, i_3 \rightarrow n_1$$
$$i_4, i_5, i_6, i_7 \rightarrow n_2$$
$$i_8, i_9, i_{10}, i_{11}, i_{12} \rightarrow n_3$$

As thirteen does not divide evenly by three, the thirteenth agent is added to the final node.

In order to minimise communication between nodes, communication was restricted to passing of small integer arrays containing key summary information: the number of immigrants and the number of births to be created on each node was sent to a given processor from the control node, and the density of agents present in each cell on each processor was returned to the control node. This information was summed on the control node to give local density information which was then passed back to each node for use in model calculations and output to file. The environment object was passed only once from the control node to the other nodes at model setup. From this minimal information passing agents could be created on nodes without the need to pass large agent objects between nodes.

Parallelising Repast

Ironically, a factor that made parallelisation of the model difficult was the use of the Repast toolkit to aid the construction of the model. At present there is no parallel support for Repast; thus, a means to use the toolkit for distributed simulation had to be found. Parallelisation of the key model incrementations was not a simple task, due to these being controlled by Repast classes. The fundamental problem was that Repast controls the main incrementations in a synchronous way on a single processor and there does not appear to be a way to override its classes. That is, Repast cannot be run on multiple nodes without getting multiple models; however, Repast is required to control the processing on each node.

Potential approaches to this problem included switching to a different agent workbench with built-in parallelism, e.g. IBM Aglets (`http://www.research.ibm.com/trl/aglets/`), or combining Repast with other technologies, such as the Distributed Parallel Programming Environment for Java (`http://www.alphaworks.ibm.com/aw.nsf/reqs/dppej`). However, an original, elegant and straightforward solution was devised for the model, which was to run the Repast interface on the control node (including the GUI etc.), while the rest of the model runs independently of Repast, synchronised by the control node code. As the purpose of the parallelisation was to increase the memory capacity, any time delay caused by the timing control of the simulation from node 0 would not be important. Overall, speed improvements would be expected with an increased number of processors.

A similar strategy was employed by Lorek and Sonnenschein (1995) for a non-Repast simulation, with the advantages of enabling larger simulations to be constructed, interactive modification of the model parameters and faster simulation speed (Lorek and Sonnenschein, 1995: pp. 2). As the model iterates, if it is necessary for any node to do anything, a message is passed from node zero to that node to tell it to run through a particular stage. The structure of

Repast, with pre-step, step, and post-step methods, made this a logical solution to isolate the code from the Repast classes on all nodes but node zero. For example at preStep(), node zero sends out messages to the other nodes telling them to perform model methods associated with preStep(). Different code may be run on each node, controlled by if statements predicated on the nodeRank.

Although successful, the drawback of this approach was that the automatic batch run function of Repast could not be used, because the MPI had to be finalised by code on each processor which was done at the end of each simulation. Once shut down, the MPI could not be restarted with Repast because this called non-static code inside the initiation routines of Repast.

In order to shut down the MPI on the non-Repast worker nodes, a function was written into the Repast atEnd() method, to send a 'shutdown' command from the control node to other nodes, to exit the process and to finalise the MPI (`MPI.Finalize();`).

Changes to the model structure

A highly simplified version of the original model was constructed in parallel as a preliminary stage in the model development. In this basic model a fixed number of static agents increase in lifestage at every model step, with a corresponding colour change. The initial agents are established on node zero. Once the model steps forward the agents are distributed from node zero to the other nodes where their lifestage is incremented. The updated agents are then passed back to node zero (figure 6.5).

Passing of agents from one node to another proved very memory intensive; therefore an alternative solution was found. Only the environment object and information on the number of agents to create on each node needed to be passed from node 0 to each of the nodes, and only density information is returned to node 0 for display. Once it was understood how a simple parallel simulation could be constructed efficiently with the Repast framework, the original model was then 'translated' into parallel code. Changes were made to the main Model class, Agent class and the RasterMap class of the original non-parallel model. There was also an addition of a 'Density' class, replacing the 'Point' class of the original model (although quite similar in content). The model flow is illustrated by figure 6.6, which can be compared to the flow diagram of the original model in chapter 4, figure 4.1.

Overall the changes have simplified the model slightly compared to the original. The point class was removed as the model is now structured by iterating through agent lists on each node, rather than agent lists associated with each cell. If it had been decided to split the environment across the nodes (as in the Everglades model by Abbott et al., 1997), the model structure would have been kept the same as the original model. It was better to split the agents

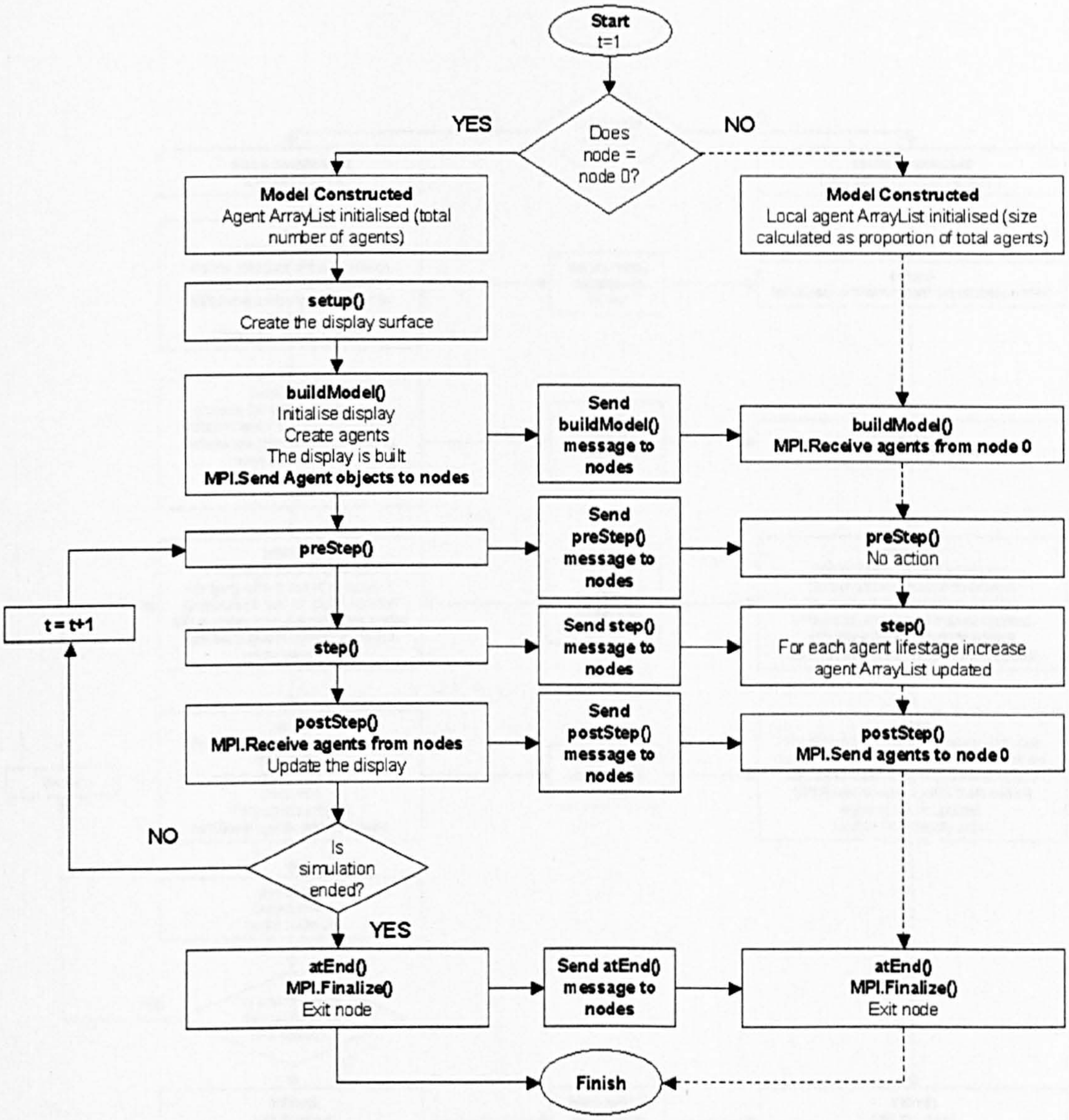


Figure 6.5: Flowchart illustrating the operation of rules at each stage of a model run for a simple Repast model, and the role of message passing to control the program flow between node 0 and the other nodes.

between nodes to avoid issues of transferring agents between processors because they have moved to different parts of the environment (a computationally expensive operation). Moving density summaries of the number of agents at each point is more efficient as there is only one summary array per node per iteration. The environment data were small relative to the agent population size, so an environment object could be established on each node without taking much memory space. A density class in the parallel model replaced the point class of the original model, but the density class is much simpler than the point class containing only information on densities at each cell location. The density class is used only to store density objects on node 0 for display.

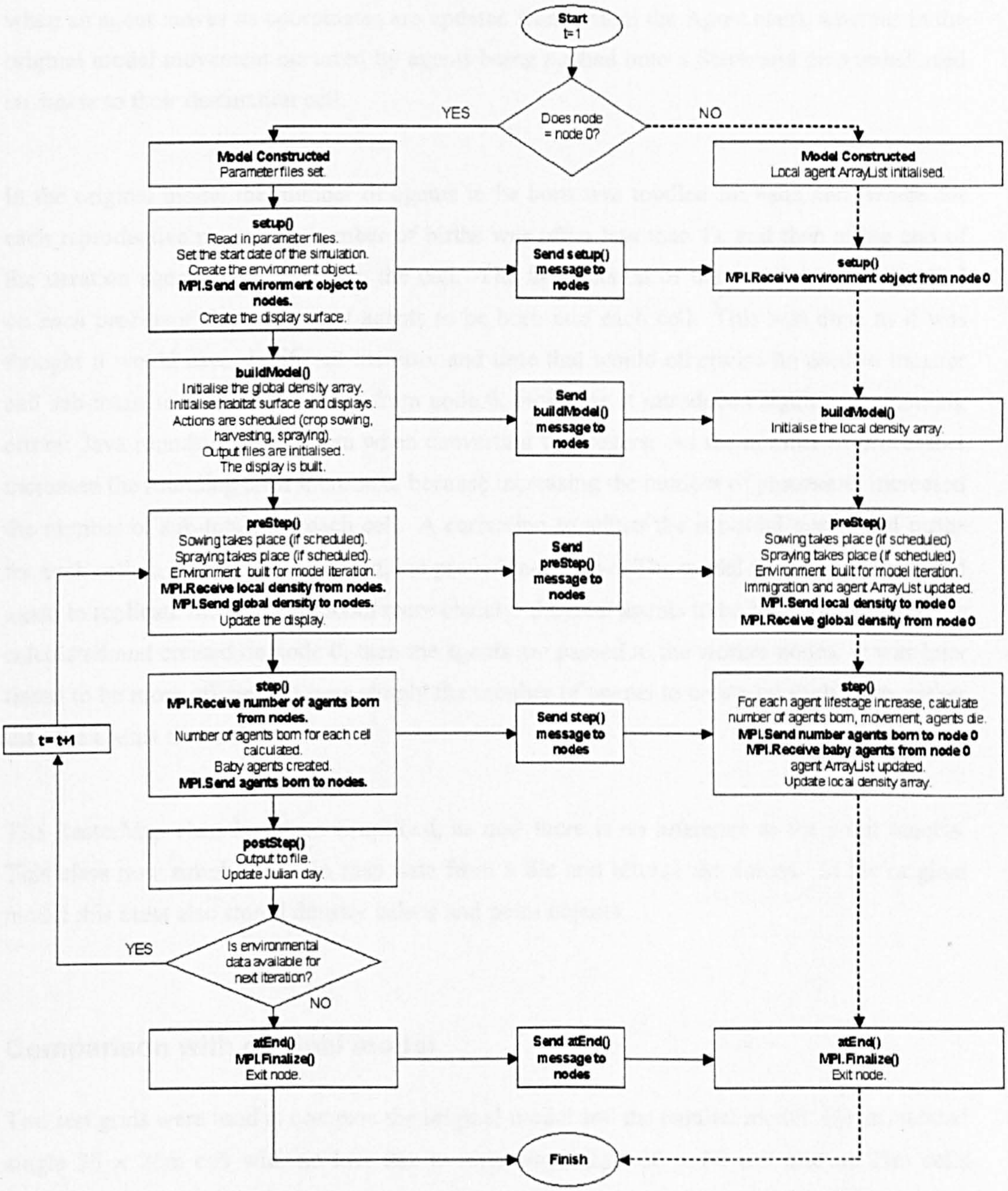


Figure 6.6: Flowchart illustrating the operation of rules at each stage of a model run for the aphid model, and the role of message passing to control the program flow between node 0 and the other nodes.

The main model class was altered significantly, although the basic Repast model structure remains the same. Within each of the Repast methods if-else statements are used to control the program flow between processors, splitting the control of the program and Repast function to node 0 and the model calculations and agents to all other nodes.

The Agent class was altered so that updates to agent lifestage and movement occur for the list of agents on each processor, rather than for the cell the agent is associated with. Thus, when an agent moves its coordinates are updated from within the Agent class, whereas in the original model movement occurred by agents being pushed onto a Stack and then transferred *en masse* to their destination cell.

In the original model the number of agents to be born was totalled for each cell (where for each reproductive parent the number of births was often less than 1), and then at the end of the iteration agents were born into the cell. The first version of the parallel model totalled *on each processor* the number of agents to be born into each cell. This was done as it was thought it would save significant memory and time that would otherwise be used to transfer cell sub-totals to each of the nodes from node 0. However, it introduced significant rounding errors: Java rounds doubles down when converting to integers. As the number of processors increased the rounding error increased, because increasing the number of processors increased the number of sub-totals for each cell. A correction to adjust the sub-total number of births for each cell on each node was tested, but proved ineffective. The model was thus restructured again to replicate the original model more closely: the total agents to be born for each cell are calculated and created on node 0, then the agents are passed to the worker nodes. It was later found to be more efficient to pass simply the number of agents to create on each node, rather than the agents themselves.

The RasterMap class has been simplified, as now there is no reference to the point objects. This class now simply reads in map data from a file and returns the values. In the original model this class also stored density values and point objects.

6.2.3 Comparison with original model

Two test grids were used to compare the original model and the parallel model: (1) an isolated single 25 × 25m cell with no loss due to migration; (2) a 50 × 50 cell grid of 25m cells with an initial population at a single central cell. This was comparable to the grids used in the sensitivity analysis, and the parameters were set to the control parameters used in the sensitivity analysis (chapter 4). The initial population was 1,000 adult alate migrants. For the parallel model, the simulation was run with two nodes - the control node and the 'worker' node, five nodes (four worker nodes) and twenty-five nodes (twenty-four worker nodes). This was to compare the consistency of the output as the model was distributed over a larger number of nodes. The temporal and spatial results are presented below and each simulation was run five times.

Key measures of improvement in performance of the parallel model compared to the original model were taken to be speed and efficiency. The main objective of parallelising the model was to increase the memory available to run the simulation. However, it was also important that the simulation was not much slower than the original (hopefully faster) and that the modifications to the code, including the addition of message passing, did not make high demands on memory. The speed and efficiency of the parallel model was compared to that of the original model, and criteria were developed to define when it was most beneficial to use parallel simulation.

Temporal comparison

In the single cell simulation (figure 6.7), the results for the parallel simulation all fell within the 95% confidence limits of the original simulation, for any number of processors. This means that the two implementations of the aphid model performed virtually identically in this instance.

There was greater temporal disparity between the original model and the parallel model when tested on a 50 × 50 cell landscape of 25 × 25m cells. This was particularly evident when examining population changes at the central cell (figure 6.8): beyond Julian day 350 there was significant divergence of the parallel model outside the 95% confidence limits of the original model. Until this date the parallel implementation was within the confidence limits of the original model. This is due to the stochastic nature of the movement and migration from the central cell (altering cell densities stochastically), where density within the cell has important effects on the future population of the cell. If the simulation is run more than five times this deviation is likely to reduce. Thus, when using the model for spatial investigations in later chapters, a greater number of model runs shall be used to reduce this divergence from the original model. A comparison of the total population within the 50 × 50 cell landscape (figure 6.9) shows that, despite localised cell density variations due to the stochastic nature of migration, the total population of the parallel implementation within the region remains virtually identical to the original model (well within the 95% confidence limits).

Spatial comparison

Comparing simulations spatially at t=2, 20 and 40, there was very little difference between the distributions of the original and parallel model at each time step. Cell populations in all cells were very similar. At t=40 it is also evident that a comparable migration event occurred in both implementations, where a very small number of alates moved a significant distance from the central cell.

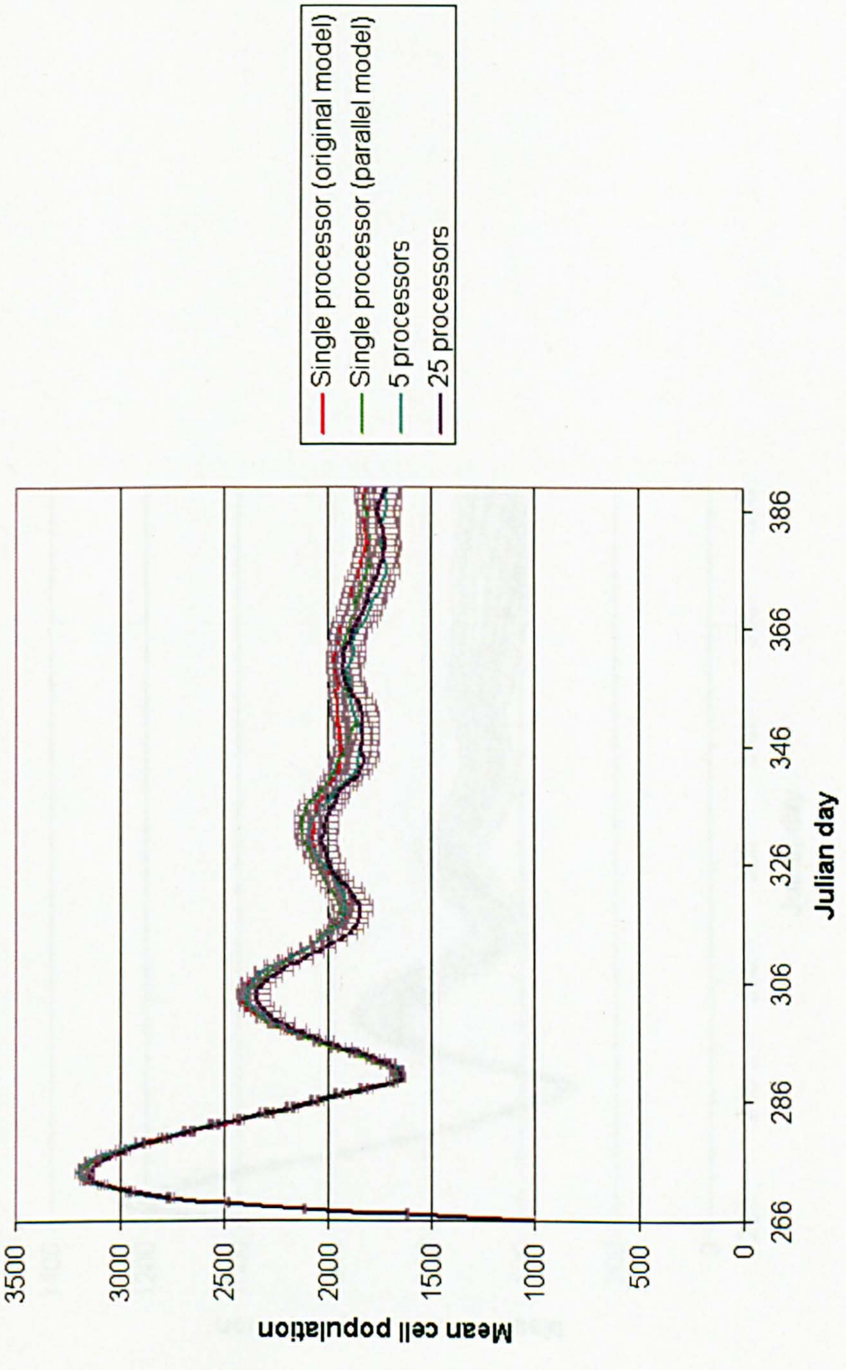


Figure 6.7: Temporal comparison for a single cell between a control (the original model) and the reprogrammed model distributed over 2, 5 and 25 nodes

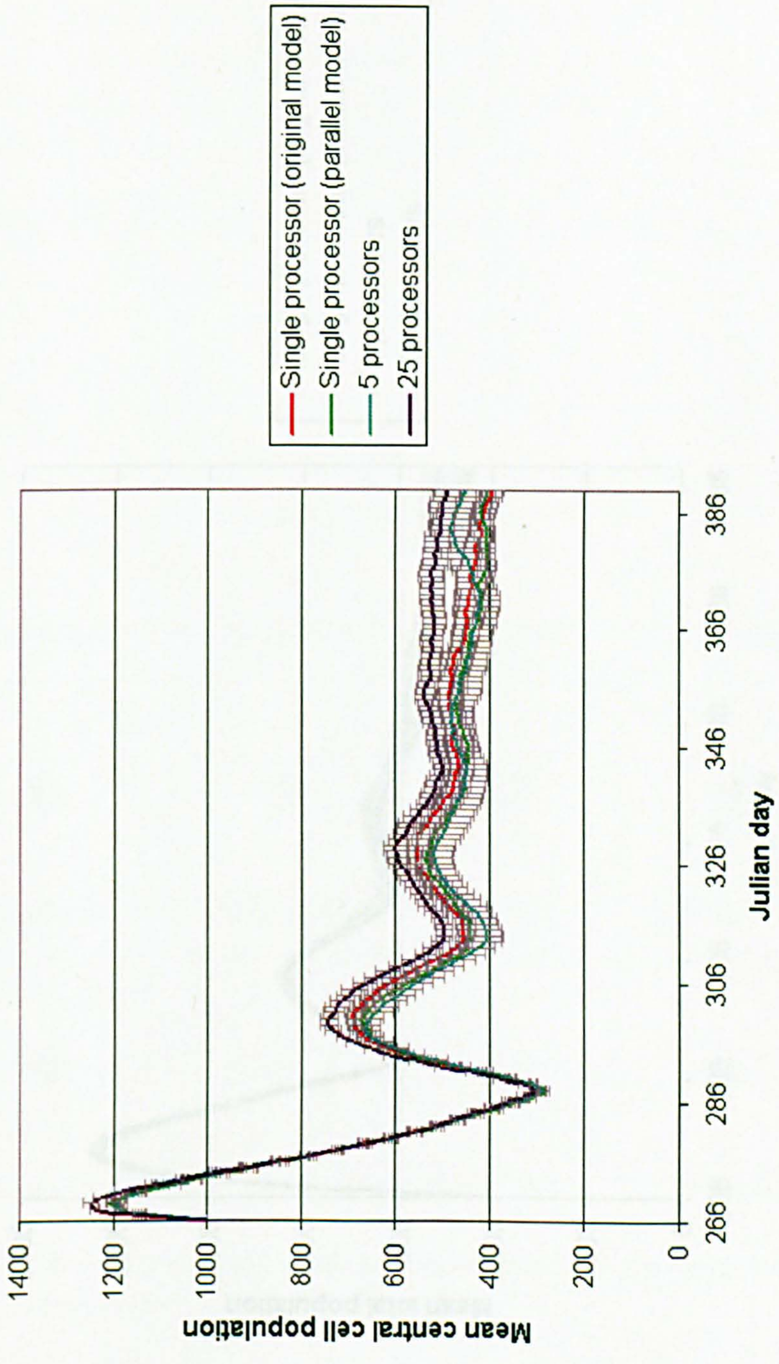


Figure 6.8: Temporal comparison for a central cell in a 50 × 50 cell landscape between a control (the original model) and the reprogrammed model distributed over 2, 5 and 25 nodes

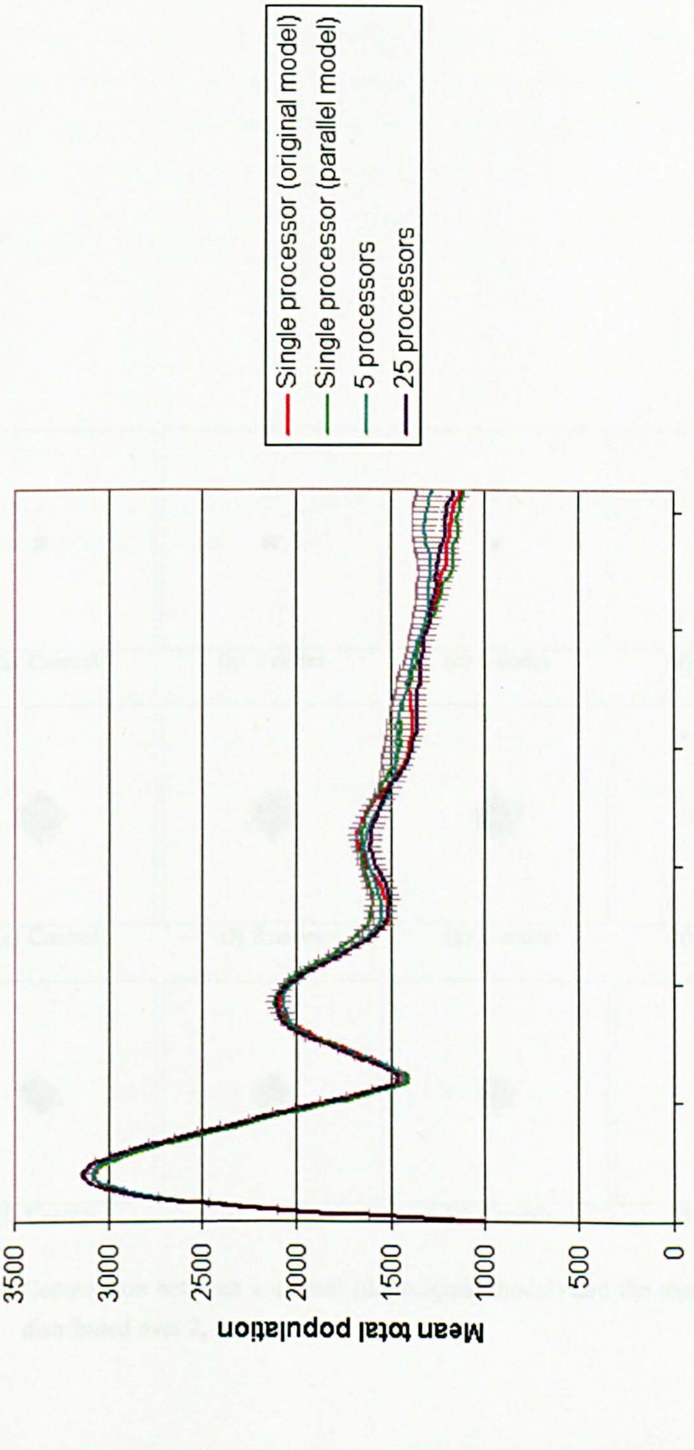
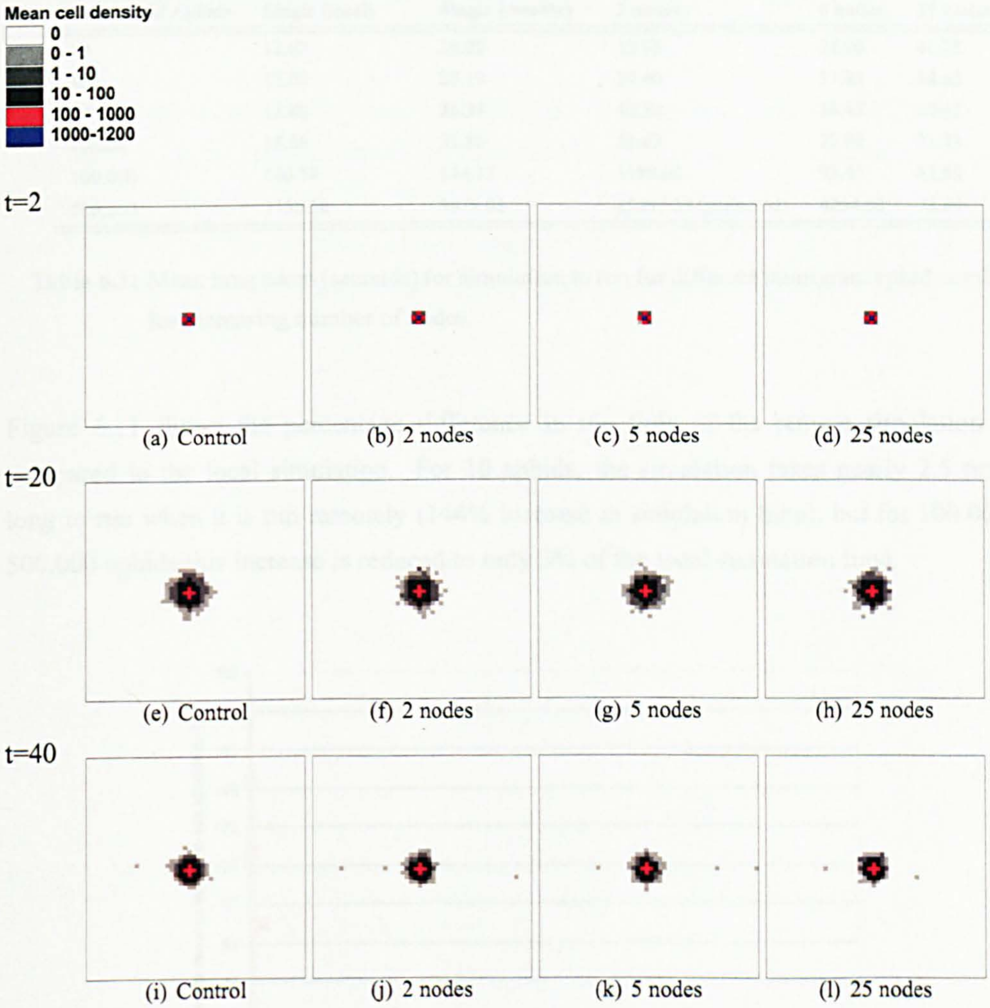


Figure 6.9: Temporal comparison for the total population in a 50×50 cell landscape between a control (the original model) and the reprogrammed model distributed over 2, 5 and 25 nodes



(a) Control (b) 2 nodes (c) 5 nodes (d) 25 nodes

(e) Control (f) 2 nodes (g) 5 nodes (h) 25 nodes

(i) Control (j) 2 nodes (k) 5 nodes (l) 25 nodes

Figure 6.10: Comparison between a control (the original model) and the reprogrammed model distributed over 2, 5 and 25 nodes

Speed

Before comparing the speed of the original model with the parallel model, it should be noted that the speed of the parallel model is affected by a lag time caused by running the simulations remotely across the Internet using X Windows with the X server Exceed `http://www.hummingbird.com`. The time delay is found by running the original model on a node of the cluster (third column 'single (remote)', table 6.1) and then comparing the result with the model run locally (second column 'single (local)', table 6.1).

Number of Aphids	Single (local)	Single (remote)	2 nodes	5 nodes	25 nodes
10	12.02	29.29	32.88	34.00	41.25
100	12.08	25.19	39.40	37.85	38.60
1,000	12.80	26.39	40.52	34.45	30.47
10,000	18.56	31.55	52.63	28.98	31.73
100,000	140.58	144.17	1189.00	93.45	43.65
500,000	3755.30	3879.01	25287.29 (projected)	4264.62	72.80

Table 6.1: Mean time taken (seconds) for simulation to run for different immigrant aphid numbers, for increasing number of nodes.

Figure 6.11 shows the percentage difference in run time of the remote simulation when compared to the local simulation. For 10 aphids, the simulation takes nearly 2.5 times as long to run when it is run remotely (144% increase in simulation time), but for 100,000 and 500,000 aphids this increase is reduced to only 3% of the local simulation time.

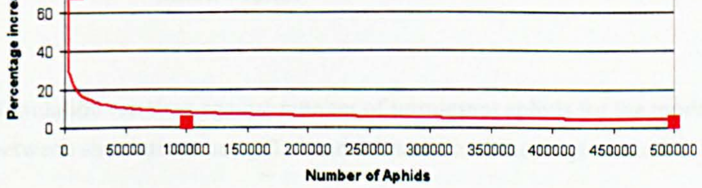


Figure 6.11: Percentage difference between the remote and local simulation run time for the original model on a single processor

Figure 6.12 and table 6.1 show that for simulations run on two nodes (one control node, one worker node), the simulation takes longer to run in parallel compared to the original model, even when the lag time is taken into account. This could be caused by message passing time delay and the modified structure of the code. Simulations run on five nodes are slightly slower than two nodes for 10 and 100 immigrants as message passing volume increases. However, as immigrant numbers increase to 1,000 and above, five nodes become much faster than two, as memory management issues start to override message passing issues. For 100,000 aphids five nodes are faster than the original model, but for 500,000 the original model is faster. This is perhaps related to the balance between the time taken to iterate through agents on each node, and the time taken to communicate between processors. It was found that 25 nodes are also slower than the original model until 100,000 aphid immigrants are input, after which 25 nodes are much faster than the original model (taking only 31% of the time of the original model to run for 100,000 aphids, and only 2% of the time of the original for 500,000 aphids).

By comparing the simulation time to the number of agents per node for the original model (one 'node') and the parallel models, a direct comparison can be made between the speed of the parallel model code in relation to the original model (figure 6.13). The parallel model takes a lot longer to run for the same number of agents per node as the original model (over three times longer); thus, the parallel code itself and the message passing makes the parallel code less efficient per node.

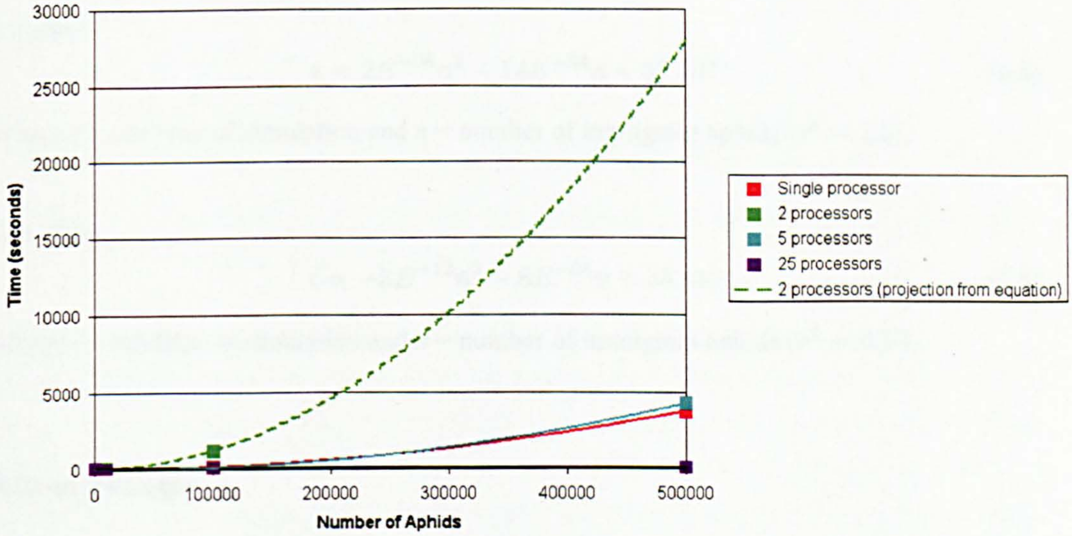


Figure 6.12: Simulation run time against number of immigrant aphids for the model: comparison between simulations using 2, 5 and 25 nodes and the original model

From the curves in figure 6.12 equations to calculate simulation run times for increasing numbers of immigrant aphids can be derived as in section 6.1 (equations 6.3 to 6.5).

2 nodes:

$$\mathrm{t} = 1E^{-07}n^2 - 5E^{-04}n + 37.29 \tag{6.3}$$

Where t = run time of simulation and n = number of immigrant aphids ($r^2 = 1.0$).

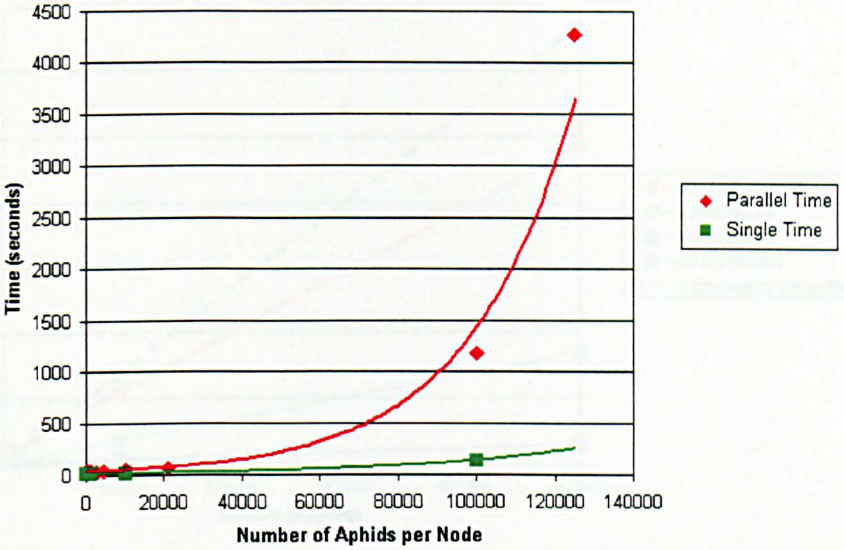


Figure 6.13: Simulation run time against number of immigrant aphids on each node: comparison between the original model (single 'node') and parallel implementations

5 nodes:

$$t = 2E^{-08}n^2 - 14E^{-04}n + 37.107 \tag{6.4}$$

Where *t* = run time of simulation and *n* = number of immigrant aphids ($r^2 = 1.0$).

25 nodes:

$$t = -3E^{-12}n^2 - 8E^{-05}n + 35.447 \tag{6.5}$$

Where *t* = run time of simulation and *n* = number of immigrant aphids ($r^2 = 0.93$).

Memory usage

The mean maximum memory usage on each worker node is taken to be a key measure of the efficiency of the parallel model in comparison to the maximum memory used by the original single processor model (figure 6.14 and table 6.2). The mean maximum memory usage for the worker nodes was calculated as the mean of the peak memory used by each node throughout the simulation. For all the input immigrant numbers, the maximum memory used by each node was so similar that 95% confidence limits derived from the standard error evaluated to ±0.00 in all cases. This shows that the distribution of individuals across the worker nodes is highly efficient, and the load is very well balanced (see section 6.2.2).

The separation of the GUI and model control from the model processing done on the worker nodes means that even for a simulation using only 2 nodes the worker node uses far less memory for simulations below 100,000 immigrants than the original model. However, the

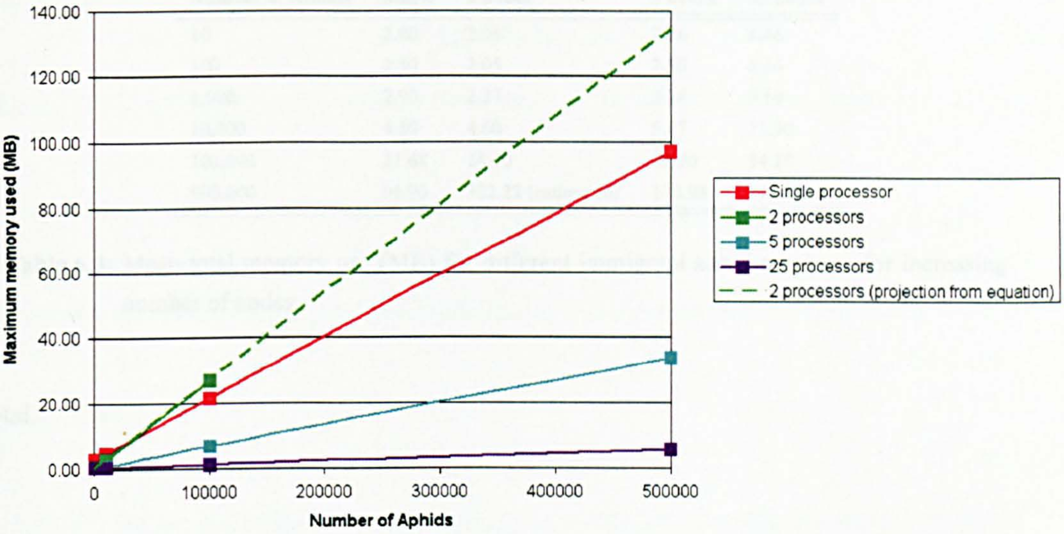


Figure 6.14: Plot of maximum memory used in a simulation run per node against number of immigrant aphids for the model, for different numbers of nodes

Number of Aphids	Single	2 nodes	5 nodes	25 nodes
10	2.80	0.30	0.29	0.29
100	2.80	0.31	0.30	0.29
1,000	2.90	0.53	0.35	0.30
10,000	4.80	2.90	0.91	0.40
100,000	21.68	27.20	6.95	1.35
500,000	96.90	150.26 (projected)	33.86	5.85

Table 6.2: Mean memory use (MB) on each 'worker' node for different immigrant aphid numbers, for increasing number of nodes.

2 node simulation uses more memory on the worker node than the original model when the simulation has 100,000 immigrants or above. This could be due to memory requirements of message passing, the cost of running the Java Virtual Machine (JVM) on each node and density calculations, which becomes more evident as larger numbers of immigrants are input. However, when 5 nodes and 25 nodes are used, the memory requirements of the individual agents on each node are very much reduced. This proves that the parallel simulation is highly efficient when distributed across more than two nodes.

The overall memory use for all nodes combined (including the control node) is shown in table 6.3. The lower total memory used for 2 nodes by the parallel model shows that the model structure of the parallel model is perhaps more efficient than that of the original model when a single processor is used. However, this is only true below 100,000 immigrant aphids, after which the original model is more efficient, perhaps due to memory requirements of message passing and density calculations, as suggested above. More message passing is required with a larger number of nodes, and thus total memory usage is increased when the number of nodes are increased, so that 25 nodes use much more memory to run the simulation than 2 nodes in

Number of Aphids	Single	2 nodes	5 nodes	25 nodes
10	2.80	2.08	2.86	8.66
100	2.80	2.05	2.90	8.66
1,000	2.90	2.27	3.28	9.60
10,000	4.80	4.60	5.47	11.90
100,000	21.68	28.90	29.50	34.10
500,000	96.90	302.22 (estimated)	170.98	142.09

Table 6.3: Mean total memory use (MB) for different immigrant aphid numbers, for increasing number of nodes.

total.

6.2.4 Implications

In order to optimise the use of parallel computing for agent simulations, it is necessary to assess the balance between the advantage of increasing the memory availability (by increasing the number of nodes used) and the cost of communicating between nodes, in relation to the number of individuals simulated. When the number of individuals is low parallel simulations take longer (figure 6.12) and are less efficient (figure 6.14) than a non-parallel model run on a single node. Increasing the number of nodes can reduce the demands on each individual node, but time to communicate between processors can be increased; thus, it was found that for 5 processors the simulation runs more quickly than the original model when 100,000 immigrants are input, but it is slower when there are 500,000 immigrants.

Below are equations derived from the curves in figure 6.14 to estimate the maximum memory required per worker node for a given number of immigrant aphids. These equations can be used to estimate the maximum number of immigrants that can be simulated by a given number of processors with 1GB RAM available (table 6.4). The run time of the simulations with the maximum number of immigrants is also calculated for each number of nodes. For 2 nodes and 5 nodes the run times are huge; 13 days and 94 days respectively. In addition, the original simulation is estimated to have a higher maximum immigrant capacity (section 6.1); thus, running the simulation in parallel across only 2 nodes is highly inefficient. Although there is a higher maximum immigrant number for 5 nodes (double that of the original model), the run time is extremely long. The maximum immigrant capacity of 25 nodes is very high, nearly 100 million aphids. The run time is also acceptable, at approximately 6 hours. Thus, it can be concluded that the most efficient way, both in time and in the number of agents that can be handled, to model large populations in parallel on the Beowulf cluster is clearly to maximise the number of processors used.

2 Processors:

$$m = 3E^{-04}n + 0.2645 \tag{6.6}$$

Where m = maximum memory required (MB) and n = number of immigrant aphids ($r^2 = 1.0$).

5 Processors:

$$m = 7E^{-05}n + 0.2741 \tag{6.7}$$

Where m = maximum memory required (MB) and n = number of immigrant aphids ($r^2 = 1.0$).

25 Processors:

$$m = 1E^{-05}n + 0.2805 \tag{6.8}$$

Where m = maximum memory required (MB) and n = number of immigrant aphids ($r^2 = 1.0$).

Number of Nodes	Maximum number of aphid immigrants	Estimated run time of simulation (seconds)
2	$3.33E^6$	$1.11E^6$
5	$14.28E^6$	$4.06E^6$
25	$99.97E^6$	$22.02E^3$

Table 6.4: The maximum number of immigrants that can be simulated for 2, 5 and 25 processors, and the associated estimated run time (from equations 6.3, 6.4 and 6.5).

6.3 Super-individuals

6.3.1 Background

The super-individual approach to model large populations on an individual basis was proposed by Scheffer et al. (1995), comparable to the earlier 'generalised individuals' of Metz and de Roos (1992). A super-individual approach 'allows zooming from a real individual-by-individual model to a cohort representation or ultimately an all-animals-are-equal view without changing the model formulation' (Scheffer et al., 1995: pp. 161). The simple idea is that individuals in a population can be grouped together into 'super-individuals', thus reducing the number of objects to simulate and therefore reducing the memory and processing power required (figure 6.15). For populations such as aphids where there are high reproductive and mortality rates leading to large juvenile populations, this approach can be very useful (Grimm and Railsback, 2005). It is possible to use the approach to test the effects of grouping individuals, and also to examine the degree to which individual behaviour explained the observed phenomena. A similar approach is used in physical models such as lattice models of fluid dynamics, particle modelling and Lagrangian modelling (e.g. Semi-lagrangian modelling of atmospheric systems and weather prediction Pironneau, 1982; Woods and Barkmann, 1994: Lagrangian Ensemble integration in Plankton modelling).

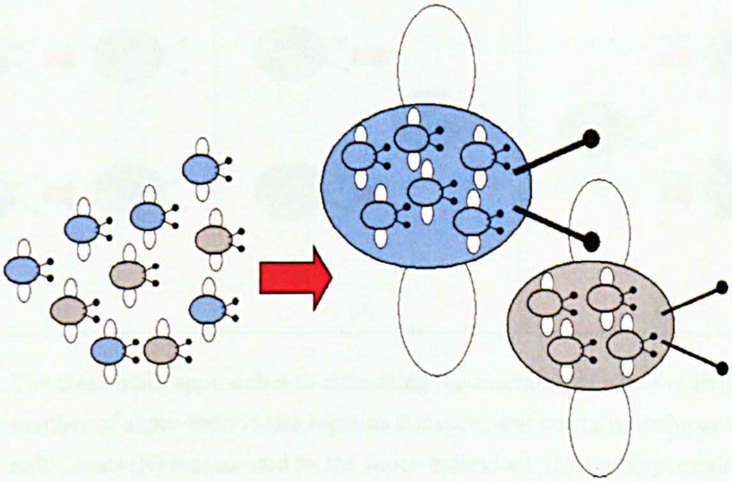


Figure 6.15: Super-individuals: Grouping of individuals into single objects that represent the collective.

6.3.2 Implementation

Combining individuals into a single super-individual

Super-individuals were split by age and morphology, thus only individuals of the same age and morphology would be represented by the same super-individual. Each super-individual represented a fixed number of individuals throughout the course of the simulation.

Adding individual immigrants to super-individuals

Immigrants were added as super-individuals of a fixed size and, as in the original model, these were of uniform age and morphology (adult alates).

Mortality of individuals/super-individuals

Estimating the mortality of super-individuals can be done in a number of ways, all of which are prone to error. The three main approaches are given by Grimm and Railsback (2005: pp. 253) (figure 6.16):

1. The number of super-individuals remains constant, and mortality reduces the number of individuals (N) represented by the super-individual.
2. N is kept relatively constant. By mortality reducing N, then super-individuals are combined when N falls below N/2.
3. Assume that an entire super-individual dies when subject to mortality.

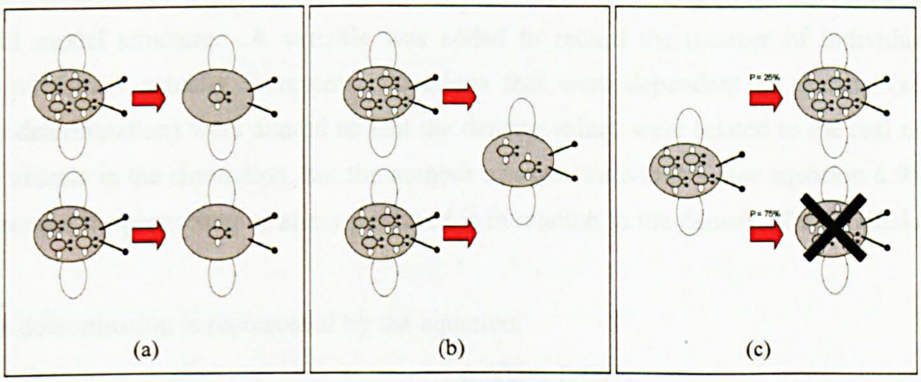


Figure 6.16: The three main approaches to estimating the mortality of super-individuals: (a) The number of super-individuals remains constant, and mortality reduces the number of individuals (N) represented by the super-individual. (b) N is kept relatively constant. By mortality reducing N, then super-individuals are combined when N falls below N/2. (c) Assume that an entire super-individual dies when subject to mortality.

Both approaches 1 and 2 require dynamic updating of the number of individuals represented by the super-individual, but in this way they do maintain more of the original variability of the model. Reducing the number of super-individuals in approach 2 and 3 has computational advantages (the number of super-individuals to iterate is minimised and individual variability is less important so calculations are less complex).

Approach 3 was chosen: super-individuals are subject to the same probability of mortality as individuals and when the super-individual dies all individuals represented by the super-individual die. This approach was chosen because the variability between individuals of the model (particularly age) meant that solution 2 (combining super-individuals) was problematic. Significant error would be introduced as individuals are re-grouped, and the process would be computationally intensive. Solution 1 (maintaining a constant number of super-individuals) would also be problematic to implement as the constant updating and variability of N would be computationally intensive, particularly as the density of individuals is important to a number of model processes. Solution 3 was therefore considered to be the most computationally efficient.

Movement of individuals represented by super-individuals

Movement of super-individuals also followed the same rules as that of individuals; however, this produced clustering of the populations (see section 6.3.3). Several alternative solutions are proposed and investigated, and the implications for spatial modelling using super-individuals are discussed.

Changes to the model structure

The construction of a super-individual simulation involved very little alteration of the original model structure. A variable was added to record the number of individuals all super-individuals actually represent. Equations that were dependent on density (such as morph determination) were altered so that the density values were related to the real number of individuals in the simulation, not the number of super-individuals (see equation 6.9). This was because the proportion of alates produced is in relation to the density of individuals.

Morph determination is represented by the equation:

$$ALPROP = \frac{0.002 + 0.991}{(1 + EXP(-0.076 \times (DENSITY - 67.416)))} \tag{6.9}$$

where ALPROP = the proportion of newly laid nymphs that will become alate and DENSITY = the total number of individual aphids per plant.

An important question to answer before super-individuals could be constructed, however, was whether individuals in the population scale linearly? Would a starting population, when increased ten-fold, produce simulation output that was simply a multiple of the original population? Complex, stochastic processes in the model could perhaps influence this and produce a non-linear relationship. A linear relationship makes it simpler to construct super-individuals and relate them to the individuals they represent: the output associated with super-individuals can be simply divided by the number of individuals the super-individual represents (N) to analyse the effects on the individual.

The initial population of aphids was varied for simulation runs, on an area of 50×50 25m cells. The parameters were set as for the sensitivity analysis (section 5.3). Initial populations were of 10, 100, 1,000 and 10,000 aphids, originating from a single central cell. Each simulation was run thirty times and an average taken to represent the total population trend over time (as several parameters in the model are stochastic). Spatial comparisons were made after 2, 20 and 40 days, by creating surfaces that show the mean density for 30 runs in each cell. Temporal comparisons were made of the population dynamics at the central cell.

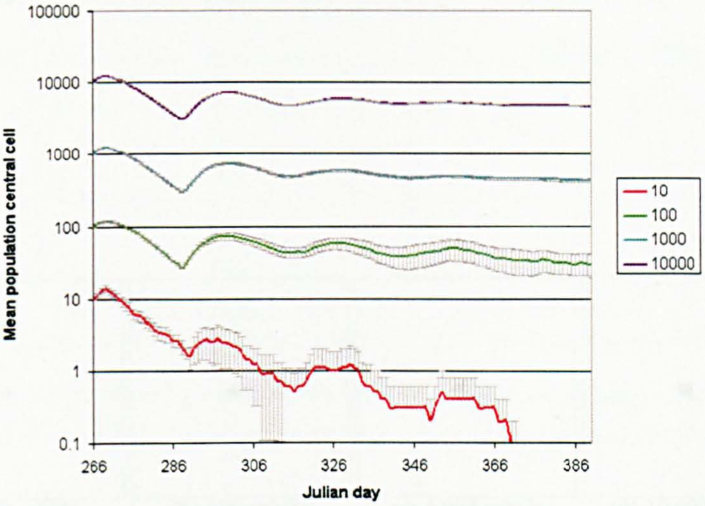
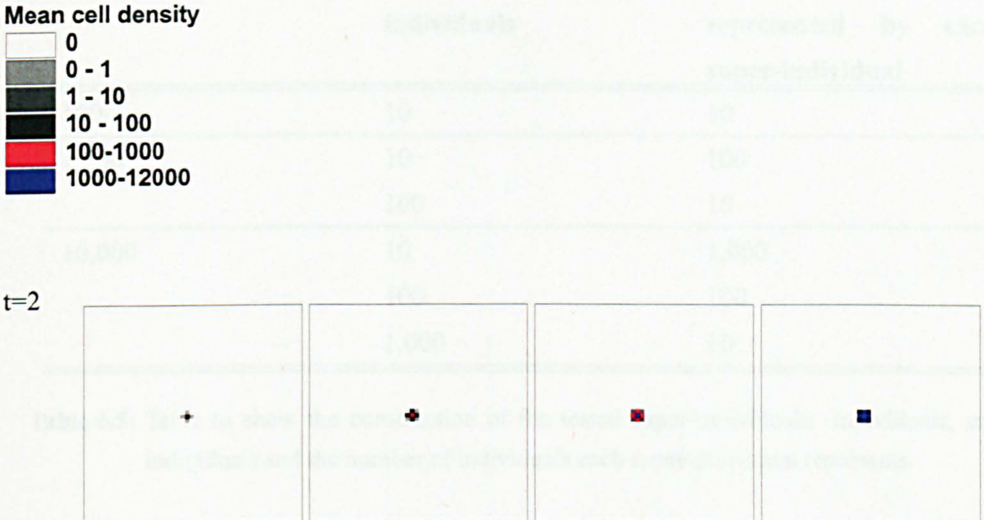


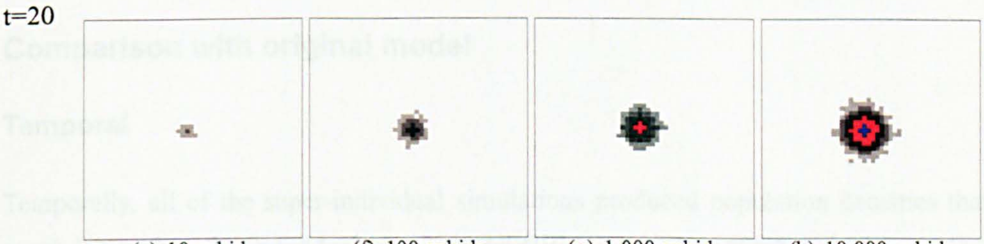
Figure 6.17: Temporal results for individuals.

The population of the central cell was recorded over time (figure 6.17). This shows that the population follows the same trend, irrespective of the initial population density. Also, there is a relationship between these trends which appears to be log-linear. It was found that simply by dividing the population of 10,000 by a factor of 10 (so ÷10, ÷100 etc...), it made a very accurate prediction of the population trend for an initial population of 1,000, 100 etc...respectively. Thus, temporally, it would be quite possible to simply represent larger populations of individuals with super-individuals that could each represent anything up to 1,000 individuals scaled log-linearly.

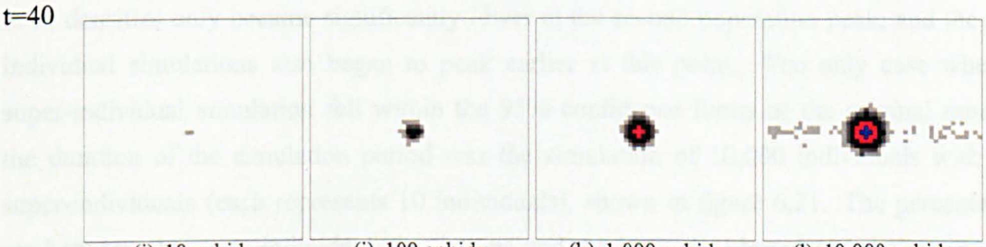
The spatial patterns produced with each initial population after 2 days, 20 days and 40 days are shown (figure 6.18). It is clear from just a brief examination of the differences between the results produced by differing initial populations that denser populations disperse further. Low initial populations are more spatially concentrated. Therefore, although the temporal trends may be comparable, the spatial trends differ significantly. A fundamental reason for this is the importance of density to migration (through morph determination, equation 6.9), which means that denser initial populations become more dispersed over time.



(a) 10 aphids (b) 100 aphids (c) 1,000 aphids (d) 10,000 aphids



(e) 10 aphids (f) 100 aphids (g) 1,000 aphids (h) 10,000 aphids



(i) 10 aphids (j) 100 aphids (k) 1,000 aphids (l) 10,000 aphids

Figure 6.18: Density distribution resulting from increasing the initial aphid population at the origin.

Test super-individuals

To test the super-individual model, populations of 100, 1,000 and 10,000 individuals were represented by varying numbers of super-individuals. Table 6.5 shows the individuals, the tested number of super-individuals and the number of individuals that each super-individual represents.

Number of individuals	Number of super-individuals	Number of individuals represented by each super-individual
100	10	10
1,000	10	100
	100	10
10,000	10	1,000
	100	100
	1,000	10

Table 6.5: Table to show the construction of the tested super-individuals: individuals, super-individuals and the number of individuals each super-individual represents.

6.3.3 Comparison with original model

Temporal

Temporally, all of the super-individual simulations produced population densities that were much lower than the individual-based model (IBM) equivalent (figures 6.19 to 6.21). However, densities only became significantly lower at the second population peak, and the super-individual simulations also began to peak earlier at this point. The only case where the super-individual simulation fell within the 95% confidence limits of the original model for the duration of the simulation period was the simulation of 10,000 individuals with 1,000 super-individuals (each represents 10 individuals), shown in figure 6.21. The percentage error between the super-individual simulations and the individual-based simulations is shown graphically in figure 6.22. This confirms that the only super-individual simulation that is acceptable is the simulation of 10,000 individuals with 1,000 super-individuals. This also shows that when a large number of individuals is represented by very few super-individuals (in this case 10) the error is greatest.

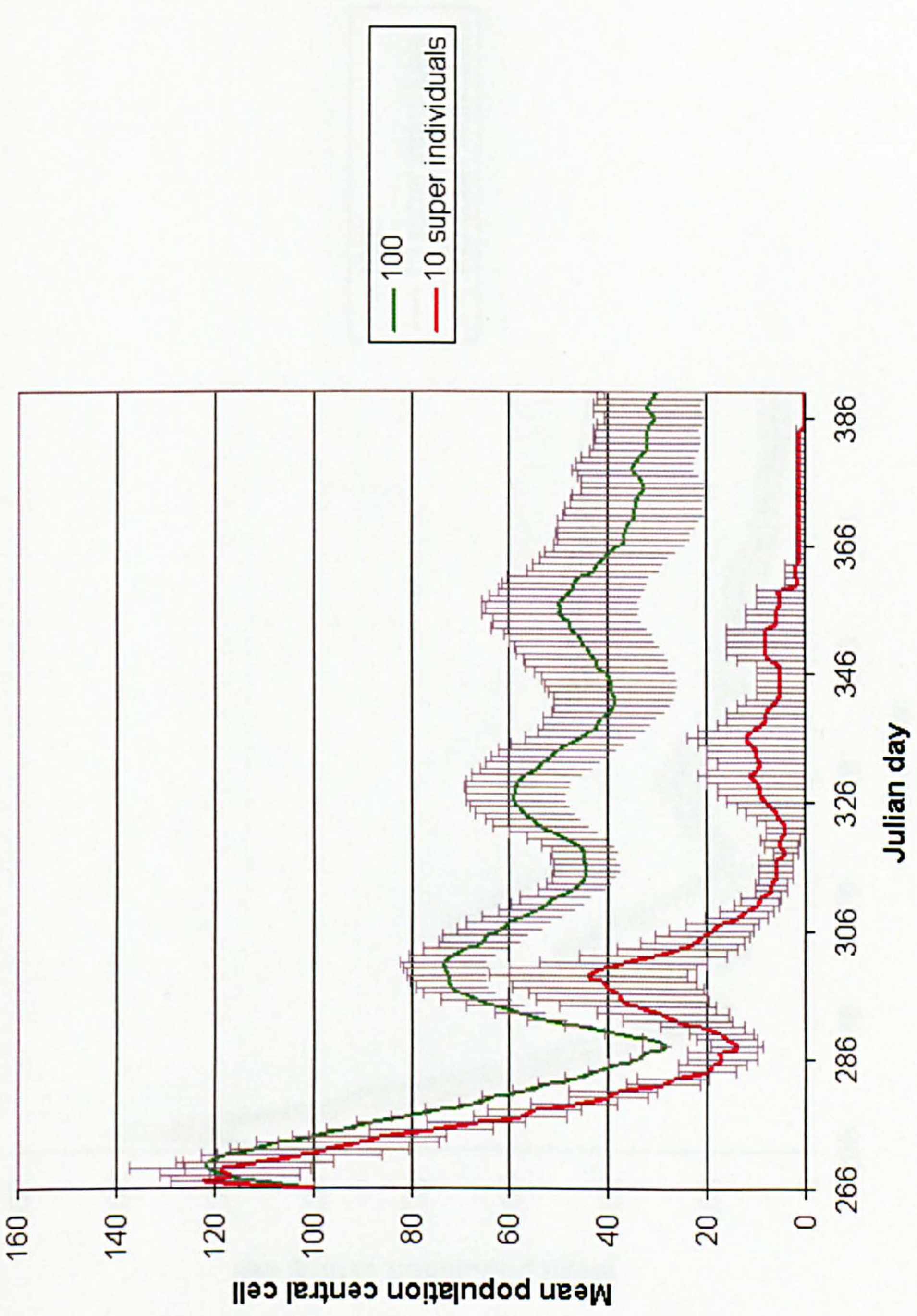


Figure 6.19: 100 individuals: comparison between individual-based simulation and 10 super-individual simulation (each represents 10 individuals).

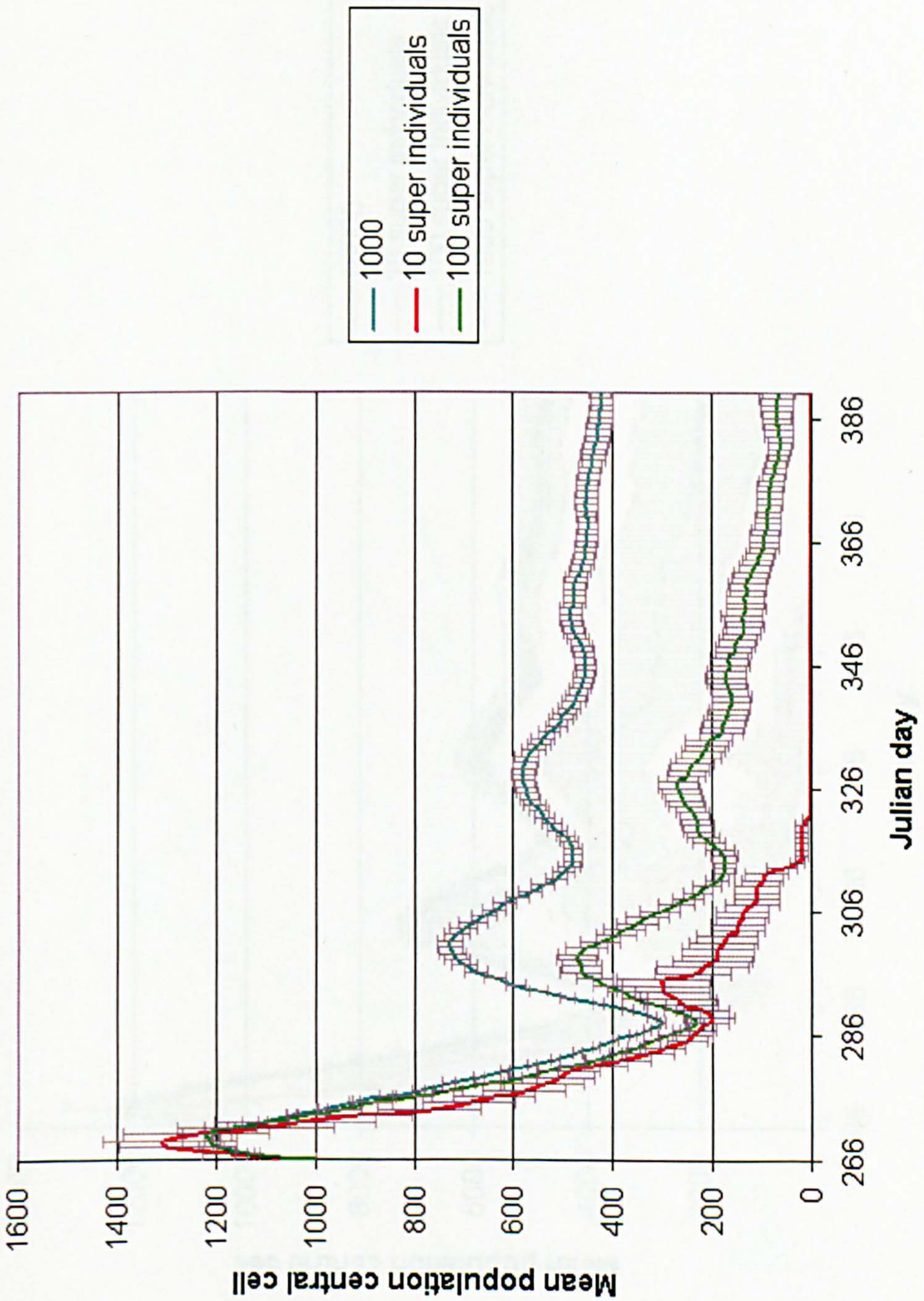


Figure 6.20: 1,000 individuals: comparison between individual-based simulation, 100 super-individual simulation (each represents 10 individuals) and 10 super-individual simulation (each represents 100 individuals).

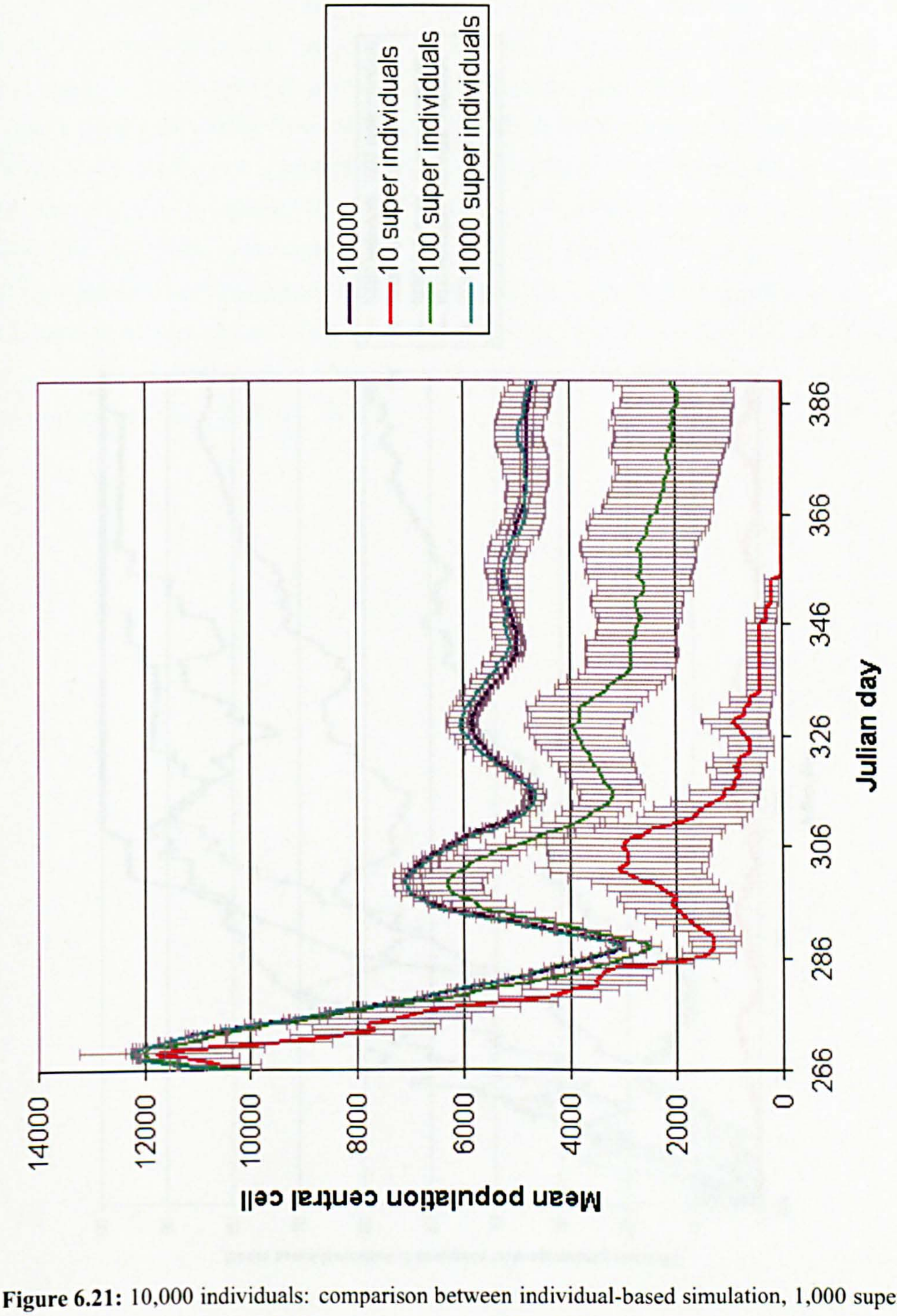


Figure 6.21: 10,000 individuals: comparison between individual-based simulation, 1,000 super-individual simulation (each represents 10 individuals), 100 super-individual simulation (each represents 100 individuals) and 10 super-individual simulation (each represents 1,000 individuals).

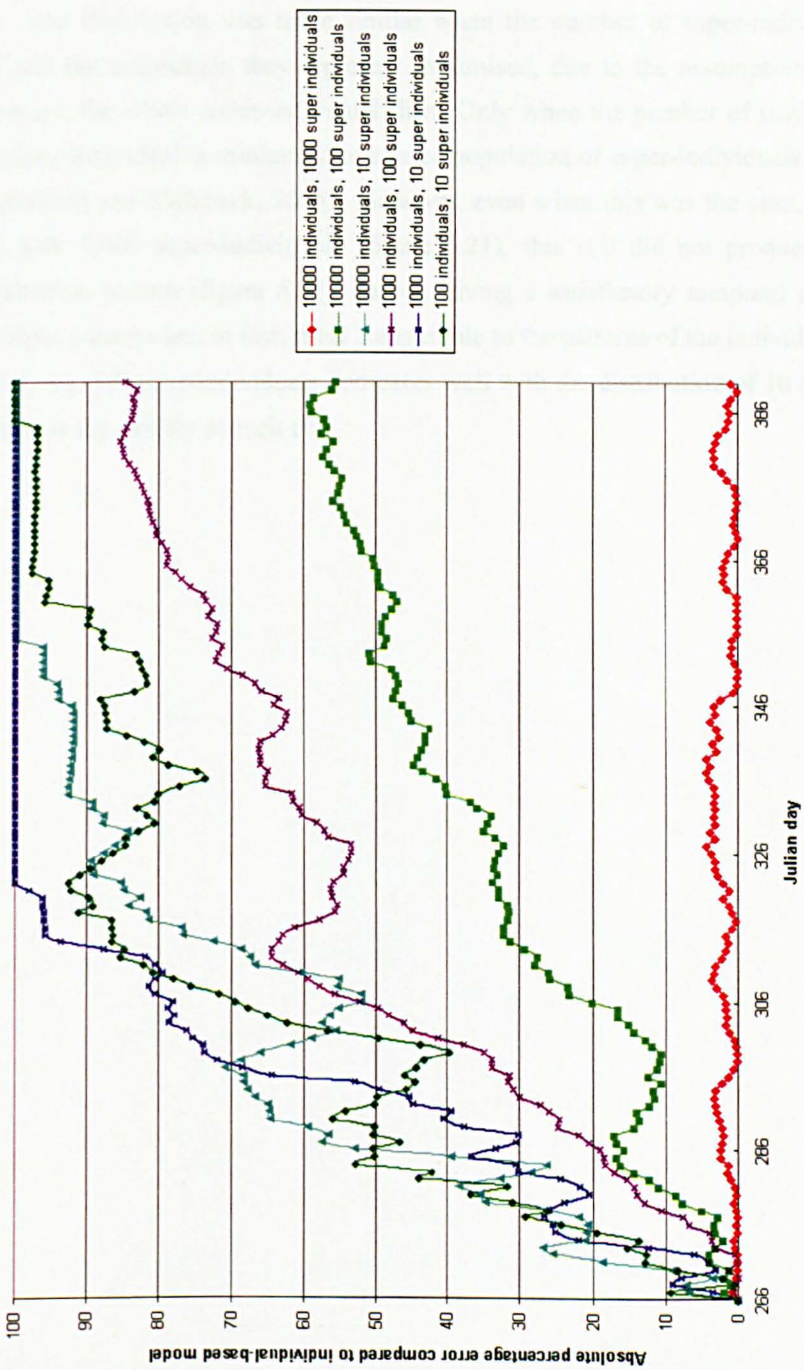
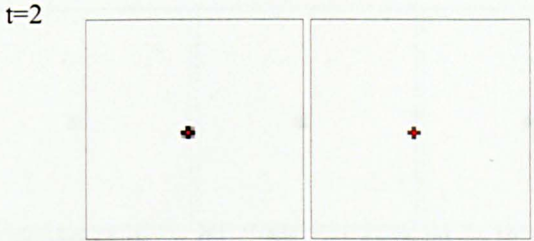


Figure 6.22: Comparison of the (absolute) percentage error between the super-individual simulations and the individual-based simulation, for each test super-individual simulation.

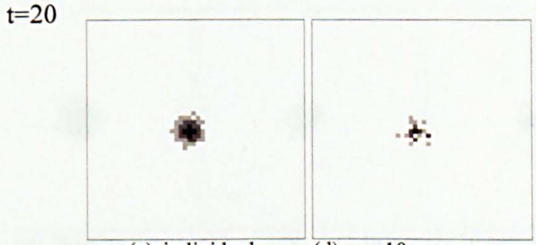
Spatial

Spatial clustering was evident in the spatial distribution (figures 6.23 to 6.25). The super-individuals were contained in fewer cells, closer to the origin, than the individual-based simulation. The distribution was more similar when the number of super-individuals was maximised and the individuals they represent minimised, due to the assumption that when mortality occurs, the whole super-individual dies. Only when the number of individuals (N) within the super-individual is minimised in a large population of super-individuals can this be overcome (Grimm and Railsback, 2005). However, even when this was the case, for 10,000 individuals with 1,000 super-individuals (figure 6.21), this still did not produce a similar spatial distribution pattern (figure 6.25), despite giving a satisfactory temporal result. The super-individual patterns are, in fact, most comparable to the patterns of the individuals for the same number, e.g. 10 super-individuals compares well with the distribution of 10 individuals. The difference is the density at each cell.

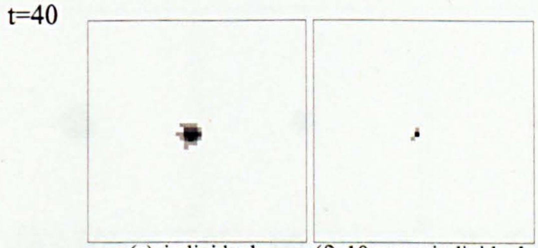
Mean cell density
- ☐ 0
- ▨ 0 - 1
- ▨ 1 - 10
- ■ 10 - 100
- 🟥 100 - 1000
- 🟥 1000 - 10000
- 🟦 10000 - 100000



(a) individuals (b) 10 super-individuals

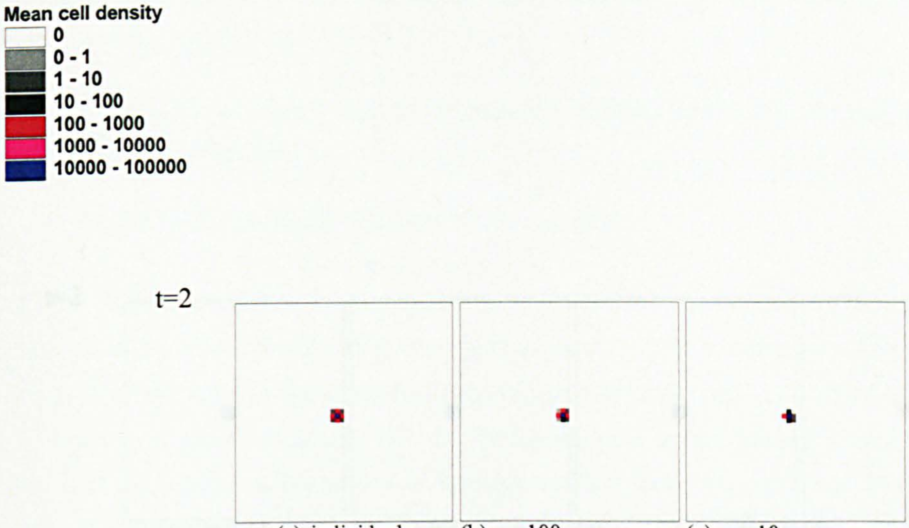


(c) individuals (d) 10 super-individuals

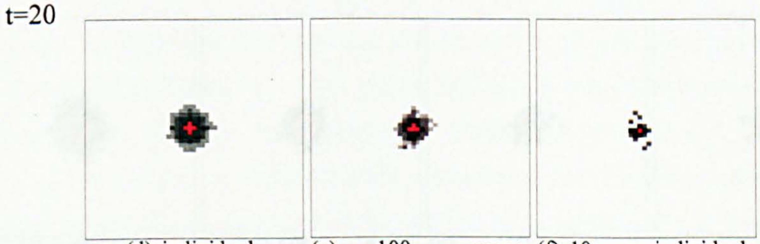


(e) individuals (f) 10 super-individuals

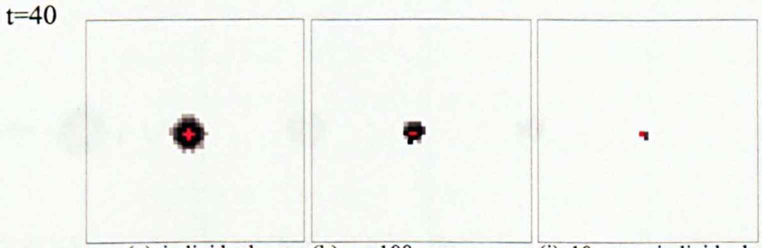
Figure 6.23: 100 individuals, density distribution from individual-based simulation compared to super-individual simulation.



(a) individuals (b) 100 super-individuals (c) 10 super-individuals

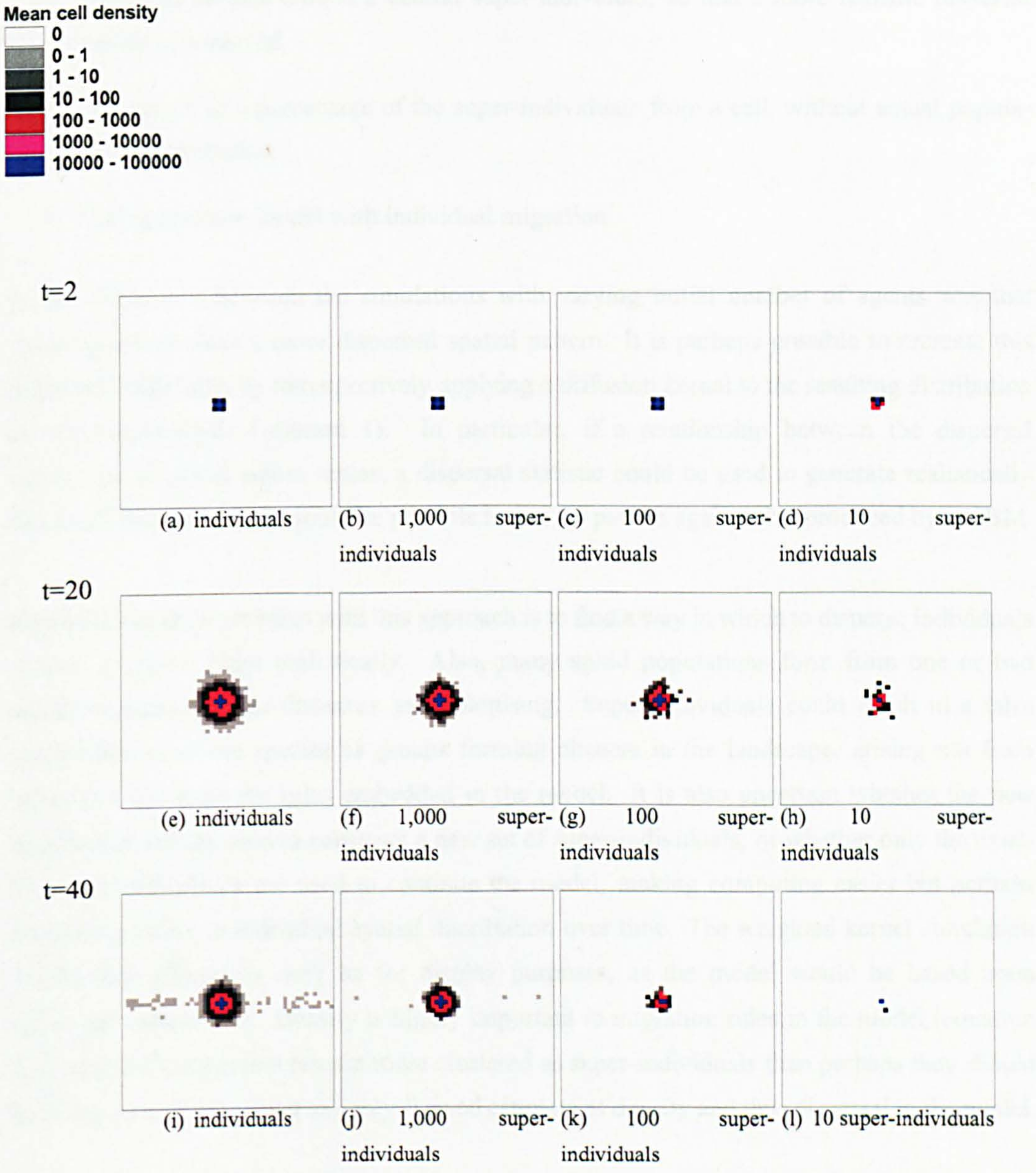


(d) individuals (e) 100 super-individuals (f) 10 super-individuals



(g) individuals (h) 100 super-individuals (i) 10 super-individuals

Figure 6.24: 1,000 individuals, density distribution from individual-based simulation compared to super-individual simulation.



(a) individuals (b) 1,000 super-individuals (c) 100 super-individuals (d) 10 super-individuals

(e) individuals (f) 1,000 super-individuals (g) 100 super-individuals (h) 10 super-individuals

(i) individuals (j) 1,000 super-individuals (k) 100 super-individuals (l) 10 super-individuals

Figure 6.25: 10,000 individuals, density distribution from individual-based simulation compared to super-individual simulation.

6.3.4 Implications

The super-individual approach is not suitable for the spatially-explicit simulation of the aphid model, using the rules presented here. Modifications to the approach could make it a possibility for future work; these include:

1. Weighted kernels around a central super-individual, so that a more realistic dispersal pattern is achieved.
2. Relocation of a percentage of the super-individuals from a cell, without actual population redistribution.
3. Cell population model with individual migration.

A key difference between the simulations with varying initial number of agents was that more agents produce a more dispersed spatial pattern. It is perhaps possible to recreate this dispersal artificially, by retrospectively applying a diffusion kernel to the resulting distribution of super-individuals (solution 1). In particular, if a relationship between the dispersal extent can be found across scales, a dispersal statistic could be used to generate realistically dispersed populations. It would be possible to test the pattern against that produced by an IBM.

However, the main problem with this approach is to find a way in which to disperse individuals around a central point realistically. Also, many aphid populations form from one or two aphids dispersing large distances and colonising. Super-individuals could result in a false representation of the species as groups forming clusters in the landscape, arising not from behaviour but from the rules embedded in the model. It is also uncertain whether the new distribution is then used to construct a new set of super-individuals, or whether only the existing super-individuals are used to continue the model, making computing easier but perhaps increasing errors in individual spatial distribution over time. The weighted kernel simulation would thus effectively only be for display purposes, as the model would be based upon super-individuals only. Density is highly important to migration rules in the model (equation 6.9), and individuals that remain more clustered as super-individuals than perhaps they should be in reality risk producing a highly skewed estimate of density and thus dispersal in the model.

The second solution is the relocation of a percentage of the super-individuals to a random adjacent cell. A series of experiments could be conducted to see which produces the greatest similarity to the real distribution. It would also be necessary to see if the same percentages apply to all super-individual groupings and initial populations. Individuals perhaps do not need to be represented as individuals for the entire simulation and thus a more aggregated representation can be used. In particular, individuals may be grouped as a single population associated with a particular cell. When dispersal occurs, the agents that disperse from this population can then be modelled individually via a 'switch' in the model. They would then be incorporated into, or form, another cell population at their new location. A plant seed dispersal model that follows a similar structure has been constructed by Dr. Hans Baveco (*pers. comm.*).

The third solution would maintain individual-based aspects, but at the same time would also lose some to a cell-based population model, comparable to a cellular automata. It is perhaps possible to view this approach as two models in one, where a switch occurs when necessary to an IBM (when individuals leave a population to migrate). It would involve some significant rewriting and also 're-philosophising' of the model. This approach adds complexity to the model. It is not thought that the increased performance gained by the elimination of the need to store and iterate through all the individuals would be enough to justify the effort of restructuring and 'model switching'.

6.4 Summary

In order to address the limitations on the number of agents imposed by the processing power and memory available, two solutions have been tested: parallel processing and super-individuals. The former approach involved significant restructuring of the model, but increased performance significantly enabling the simulation of at least ten times more agents. The parallel model produced results that are comparable to the original, non-parallel model, leading to the conclusion that for the simulation of very large populations the parallel model is an adequate solution.

Although initially far simpler to implement, the super-individual approach is inappropriate for spatial simulation in the density-dependent form presented here. However, it may be possible to use this approach if the model were to be significantly altered, perhaps most effectively as a super-individual model merged with an IBM, where dispersal can be simulated by switching from a super-individual to an individual-based model when necessary. There is a high risk that the complexity of switching between models or implementing retrospective re-distribution of agents could introduce significant error, or put high demands on the processor or memory, which are already limited. This risk was believed too high to attempt to restructure the super-individual solution for this thesis. Thus, as it was necessary to simulate large populations in realistic simulations in the following chapters, parallel processing has been used.

CHAPTER 7

The Timing of Crop Introduction and Configuration

7.1 Introduction

This chapter, and the two that follow, relate to objective 6 of this thesis: they demonstrate the application of the model to the study of insect pest population outbreaks, both spatially and temporally, in relation to agricultural management practices. The analysis of a number of scenarios constructed with both real and theoretical landscapes allows for an examination of how populations are spatially structured and how changing landscape structure alters the dynamics of pest outbreaks.

The timing of crop sowing and harvesting, along with the spatial configuration of the crop, is known to influence insect population dynamics (Leather et al., 1989; Mann et al., 1996). A simulation approach helps to better understand such landscape-insect interactions and potentially to predict the impact of landscape change. This investigation examines the ways in which the spatial population distribution, and the population structure of *Rhopalosiphum padi* and the temporal population dynamics of the species are influenced by: a) the timing of the host crop sowing; b) the spatial configuration of the host crop and the distance to other host crop patches (host crop 'fragmentation'); c) the proportion of host crop in the landscape (host crop 'cover') and d) the field size of the crop area. This was done with the use of theoretical landscapes. Simplified theoretical binary landscapes were used without marginal habitat (for landscape construction see chapter 3, section 3.2.1). The influence of marginal habitat surrounding the host crop is analysed further in chapter 8, where a more complex theoretical matrix structure is developed and analysed.

The impact of the size, configuration and composition of habitat patches on insect population dynamics forms a central part of metapopulation research (see chapter 2). Such research acknowledges the importance of dispersal between sub-populations; however, metapopulation models create arbitrary divisions between patches and do not facilitate analysis of dispersal events. The purpose of this investigation is to assess the effects of landscape fragmentation and 'patch' configuration upon the population displacement of a generalist species, by individual-based simulation. In particular the investigations focus on movement and survival of the aphids under different landscape configuration scenarios. Proximity to other habitat and the fragmentation of the landscape are key factors in the sustainability of the population within the landscape. Key observations include: within-cell population density changes in favourable habitat; movement frequency and distance; and the direction of movement.

7.2 Method

The model was run under the environmental conditions for the autumn and winter of 1988/9 (see chapter 3). Some parameters were simplified: the wind direction was set as a constant westerly wind (that of the dominant wind direction in the UK) with variable speed for the timing of crop sowing scenarios 7.2.1 and wind was not included in scenarios 7.2.2 and 7.2.3 (meaning that only local movement as a function of habitat response was modelled for the latter scenarios, not migration). Immigration was based upon real numbers in 1988/9 for scenarios 7.2.1, but was simplified to a single immigration event of 100 aphids per 25 × 25 m for the other scenarios. The use of consistent, simplified environmental conditions in these scenarios helped to isolate key relationships between the landscape changes and aphid population response. All other parameters were as described in chapter 4, including details of the movement simulation and movement ranges.

7.2.1 Timing of crop sowing

Crop sowing date can be crucial to the incidence of aphid populations. The sowing date was thought to have the biggest impact on the seasonal population levels as it is the primary determinant of the crop growth stage when the aphids arrive in the crop (Leather et al., 1989; Mann et al., 1996). Crops sown in early September are generally considered to be at highest risk (Foster et al., 2004).

The model was used to assess the effects sowing in early September may have on *R. padi* population dynamics, compared to sowing at a later date with respect to the immigration of the aphids. Thus, scenarios were created where the estimated date of the single initial immigration (julian day 266) was shifted 5 days later and 10 days later, to show how the impacts of the sowing date might vary with the timing of the first immigration. A single field was isolated from a landscape derived from Land Cover Map 2000 (chapter 3), so as to observe the effects of the timing of crop sowing on the population of a single field in isolation. It was only this isolated field into which immigration takes place.

The model does not contain a complex crop simulation; rather, the crop simply exists or it does not. However, it was thought that this would be sufficient to examine the implications of altering the sowing date of the crop. To alter the sowing date of the crop, the scheduler in the modelling toolkit RePast was used and the land-use class was over-ridden. Crops were planted at model tick x (sowing), and the land was declared unsuitable habitat until crops emerge at tick x+30. The land becomes unsuitable after model tick y (harvest).

Before the crop matures the crop field area was considered 'bare unfavourable' habitat. This is comparable to the 'unfavourable' habitat environment, where aphids may survive for up to 3 days without finding more suitable habitat; however, under 'bare' conditions aphids will not reproduce. Thus, in the following analysis, when sowing had not yet taken place across the landscape, areas of crop classified as 'unfavourable' habitat were also considered to be 'bare'. It was assumed in this scenario that sowing and maturity of the crop occured at the same time across the entire region.

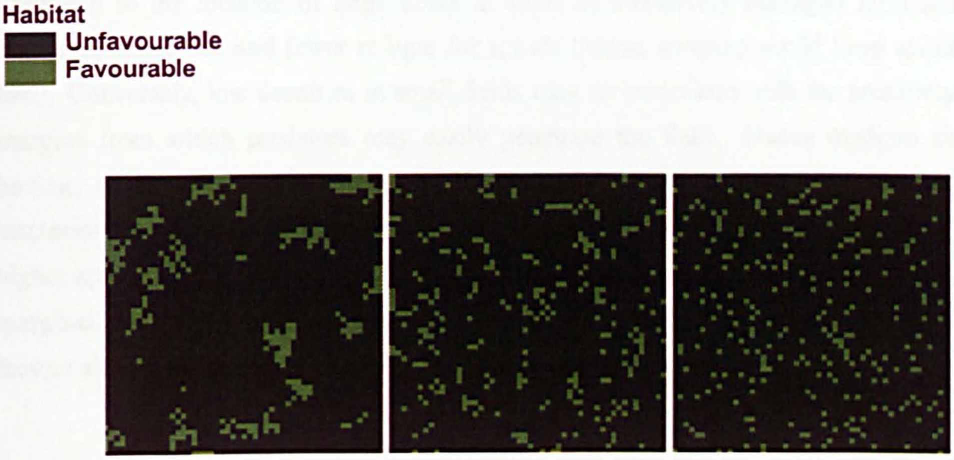
7.2.2 Fragmentation of crop and crop cover

The objective of this part of the study was to observe the changes in the displacement of the aphids as a result of changes in the proportion of habitat in the landscape and the configuration of this habitat. This was done by assessing the diffusion of the aphids across the landscape from a single cell source. This cell was central or near central to the landscape, in favourable habitat. The results focus on the survival of the aphids that leave the origin, and the resultant distance, displacement and pattern of movement from this origin.

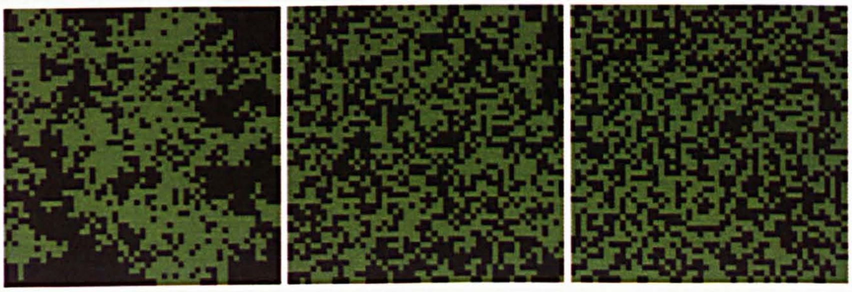
Alteration of landscape configuration was based upon theoretical patterns (for landscape construction, see section 3.2, chapter 3). It was decided to conduct tests using extreme values, as well as test an intermediate/neutral value. The configurations used in the analysis are given in table 7.1 and the corresponding landscape is shown in figure 7.1. Configurations are subsequently abbreviated to their habitat configuration 'c' followed by their fragmentation 'f': for example 10% habitat cover, 50% fragmentation is referred to as "c10f50". For these scenarios no wind was simulated, and the initial number of aphids was kept constant at 100 aphids per 25 × 25m.

Habitat cover	10% fragmentation	50% fragmentation	100% fragmentation
10%	a	b	c
50%	d	e	f
90%	g	h	i

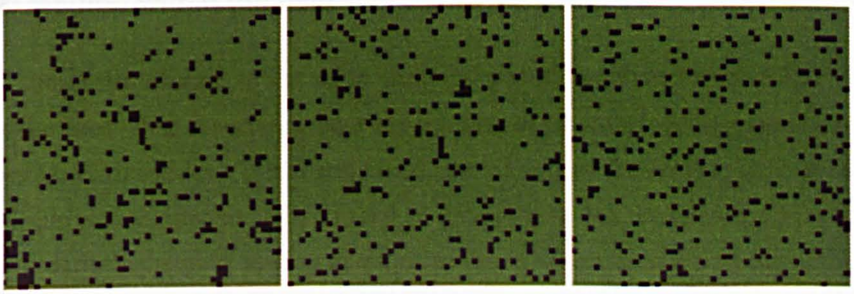
Table 7.1: Habitat configuration matrix for fragmentation and habitat cover, letters refer to landscapes in figure 7.1.



(a) 10% habitat cover, 10% fragmentation (b) 10% habitat cover, 50% fragmentation (c) 10% habitat cover, 100% fragmentation



(d) 50% habitat cover, 10% fragmentation (e) 50% habitat cover, 50% fragmentation (f) 50% habitat cover, 100% fragmentation



(g) 90% habitat cover, 10% fragmentation (h) 90% habitat cover, 50% fragmentation (i) 90% habitat cover, 100% fragmentation

Figure 7.1: Fragmentation versus habitat cover indices and the resultant landscape.

7.2.3 Field size

There is evidence that field size can influence *R. padi* and *Sitobion avenae* density, where fields of 2-4 ha in area have higher aphid density than smaller or larger fields (Foster et al., 2004). In the study of Foster et al. (2004), the lower density of aphids in larger fields was attributed to the location of large fields in areas of intensively managed farmland (where higher pesticide use and fewer refugia for aphids during spraying could keep aphid density low). Conversely, low densities in small fields may be associated with the proximity of field margins from which predators may easily penetrate the field. Hence medium size fields having: good availability of marginal habitat; less intensive pesticide regimes; and greater distances for predators to traverse to penetrate the field, would be well configured to have higher aphid density. The ability of the model to capture some of these factors (availability of marginal habitat, pesticide regimes) will allow further investigations to analyse whether these factors alone explain this relationship (see chapters 8 and 9).

Field size variation corresponds to cover increase whilst fragmentation remains constant. This is because, in these theoretical landscapes, fragmentation corresponds directly to the number of patches in the landscape (Tischendorf and Fahrig, 2000). Thus a constant fragmentation index maintains a constant number of patches in the landscape. These patches have a similar spatial correspondence because the fragmentation is constant.

7.3 Results

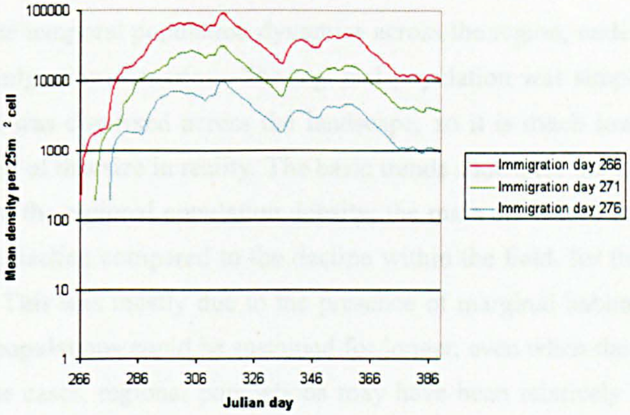
7.3.1 Timing of crop sowing

Temporal results: within the field

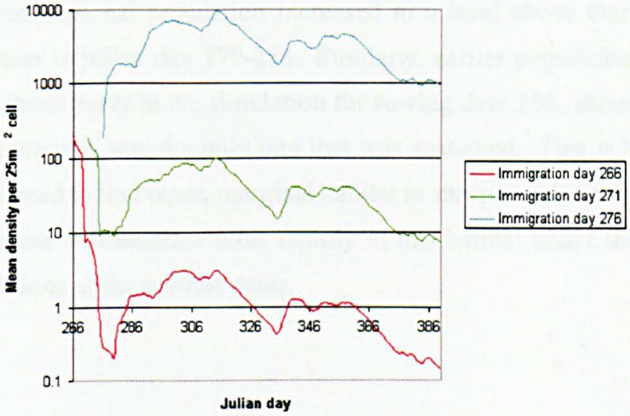
Figure 7.2 shows the changes to the mean temporal population dynamics per 25 × 25m grid cell within the field. Both the sowing date and the date of the initial immigration had an effect on the population dynamics. When combined, the timing of the sowing date in relation to the initial immigration appears quite crucial in determining the population dynamics within the field. An early September sowing date (julian day 236) combined with the earliest initial immigration at julian day 266 resulted in the highest aphid density, where peak density was approximately 90,000 aphids per 25 × 25m (143 aphids per m^2) (figure 7.2a). At this sowing date, later initial immigration resulted in lower peak densities: the peak density for initial immigration 10 days later resulted in a peak population of approximately 10,000 aphids per 25 × 25m (16 aphids per m^2). However, the timing of the peaks and the population trends over time were identical. This indicates that the identical temperature conditions helped maintain a constant population trend over time, further supported by the results of the model sensitivity analysis (chapter 5).

When the sowing date was shifted 10 days later (julian day 246), the absence of crop until julian day 276 had a large impact upon the earliest immigrant population (figure 7.2b). As the aphids did not have any favourable habitat in the landscape for a full ten days, the population rapidly declined and could not reproduce; thus, the population was reduced to very low levels (reaching a peak of only 0.0050 aphids per m^2 after julian day 270). However, later initial immigrant populations were able to survive better, as their arrival coincided with the maturity of the crop. The field population under these conditions reaches a peak of 0.16 aphids per m^2 after julian day 270. The highest population density was reached by initial immigrants arriving 10 days later (julian day 276). Thus, despite the warmer temperatures available to immigrants at julian day 271, they were unable to take advantage of the better environmental conditions (warmer temperatures) without a mature crop upon which they could reproduce. Initial immigrants arriving 10 days late were able to settle on the crop straight away and thus the initial population growth was more rapid as well as much higher than earlier immigrants. The general population trend and peaks were very similar (with peaks at julian day 314 or 315), despite the large differences in population density in the field.

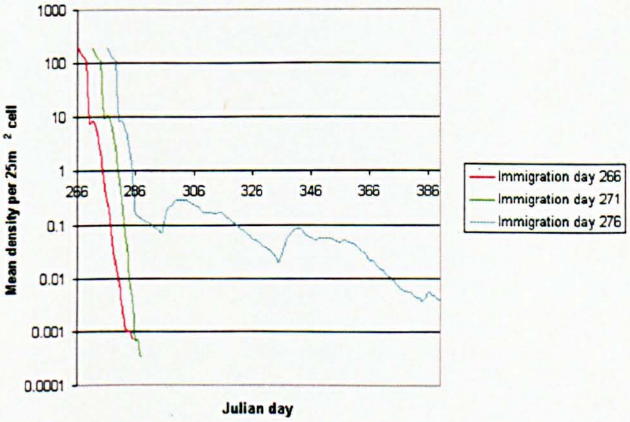
When the sowing date was shifted to late September, julian day 256, under all immigration scenarios aphid densities did not reach high levels (figure 7.2c). The initial immigrant population was now the highest: 0.31 aphids per m^2. The highest densities were maintained when the initial immigration was 10 days late. Although it was another 10 days after this that the crop emerges, the population was able to sustain itself at a high enough density that it can still establish a very low level population that could survive throughout the time period, whilst temperatures were still favourable. The subsequent population peaks were 15 days earlier compared to the earlier sowing dates, the peaks were reached more rapidly and then populations declined more slowly. At this late sowing date, earlier immigrants were not able to sustain populations within the field, as the within-field population was at too low a level once the crop emerges to recover. The population descended from the initial immigrants that were 5 days late declines rapidly and disappears from the field by julian day 288; it was unable to recover even though this coincides with the crop maturity. The population descended from the earliest initial immigrant population declined less rapidly, so that it was eliminated from the field only 3 days earlier (julian day 285), before the crop matures.



(a) sowing date 236



(b) sowing date 246

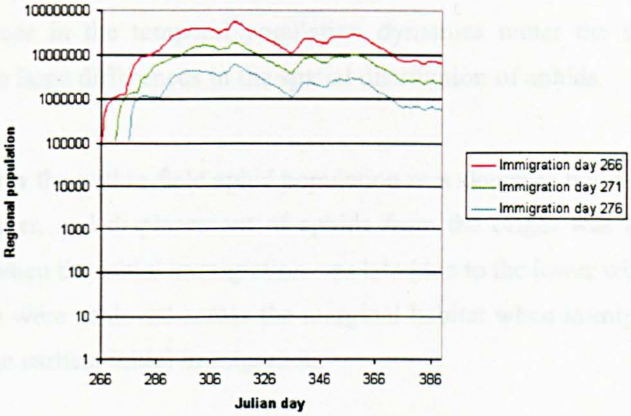


(c) sowing date 256

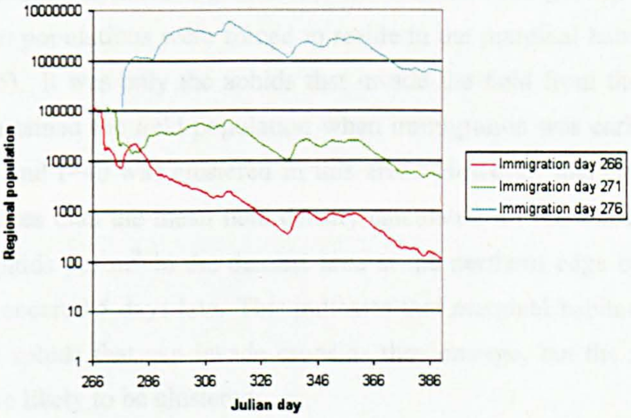
Figure 7.2: Comparison of the mean temporal dynamics per 25m grid cell within the field when the sowing date was (a) julian day **236** (b) julian day **246** and (c) julian day **256**. The simulations were run with initial immigration at julian day 266, 5 days later and 10 days later for each sowing date.

Temporal results: across the region

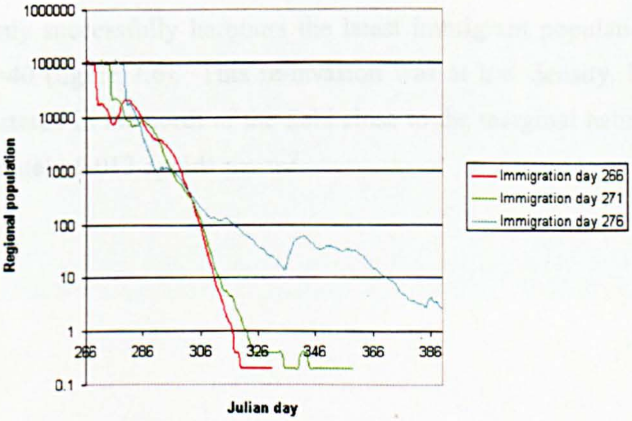
Figure 7.3 shows the temporal population dynamics across the region, under the same sowing date and initial immigration scenarios. The regional population was simply that originating from the field as it was dispersed across the landscape, so it is much lower than would be expected for a region of this size in reality. The basic trends shown for the within field population were similar for the regional population density; the main difference is the slower rate of regional population decline compared to the decline within the field, for the populations that were under stress. This was mostly due to the presence of marginal habitat in the landscape as a whole, where populations could be sustained for longer, even when there was no crop for the aphids. In some cases, regional populations may have been relatively high compared to the other scenarios, whilst the corresponding field population was relatively low. This was the case for the earliest immigrant population when sowing took place at julian date 246, where the earliest immigrant regional population increased to a level above that of the immigrant population 5 days later at julian day 279-283. Similarly, earlier populations increased above those of later populations early in the simulation for sowing date 256, although eventually the latest immigrant population was the only one that was sustained. This is because the immigrants which were forced to find other, marginal habitat to survive early in the simulation when there is no crop present will increase most rapidly in this habitat under the more favourable environmental conditions at the earliest dates.



(a) sowing date 236



(b) sowing date 246



(c) sowing date 256

Figure 7.3: Comparison of the mean temporal dynamics of the regional population when the sowing date was (a) julian day **236** (b) julian day **246** and (c) julian day **256**. The simulations were run with initial immigration at julian day 266, 5 days later and 10 days later for each sowing date.

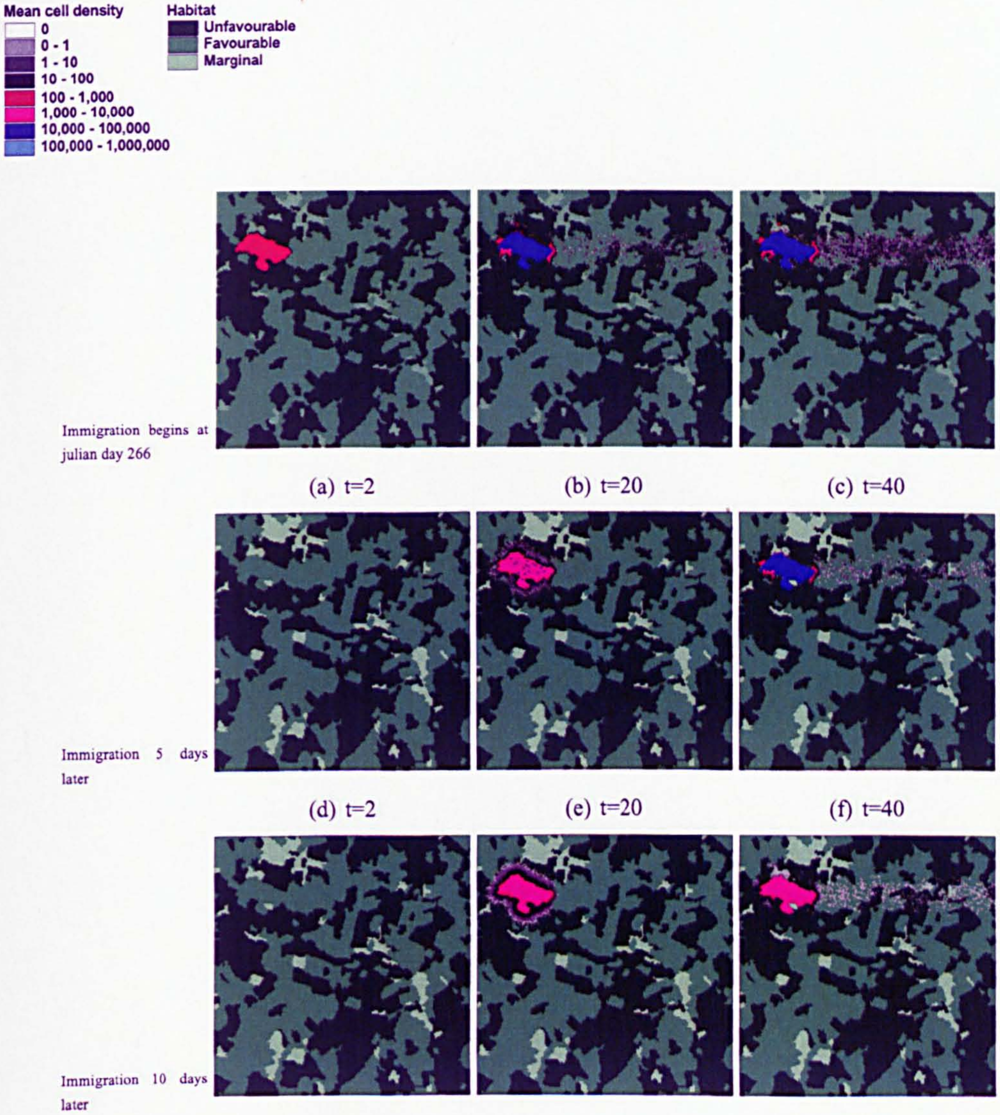
Spatial results

The large differences in the temporal population dynamics under the three crop sowing scenarios reflect the large differences in the spatial distribution of aphids.

Figure 7.4 shows how the within-field aphid population was slower to build up when the initial immigration was later, and displacement of aphids from the origin was lower (particularly evident at day 20) when the initial immigration was late (due to the lower within-field density). Denser populations were achieved across the marginal habitat when immigration was 5 days late compared to the earliest initial immigration.

After sowing on julian day 246, it is clear from the spatial results why the mean field population was sustained at much higher levels when initial immigration was 10 days late, as earlier immigrant populations were forced to reside in the marginal habitat to the north of the field (figure 7.5). It was only the aphids that invade the field from the marginal habitat to the north that sustained the field population when immigration was earlier; thus, the field population at $t=20$ and $t=40$ was clustered in this area. However, this clustered population was at higher densities than the mean field density calculated for the whole field, with up to approximately 2 aphids per m^2 in the densest area at the northern edge of the field at $t=40$ when immigration occured 5 days late. This indicates that marginal habitat may successfully harbour immigrant aphids that can invade crops as they emerge, but the second generation field populations are likely to be clustered.

A similar pattern was observed when the sowing date was latest (julian day 256), where the marginal habitat only successfully harbours the latest immigrant population for re-invasion into the field by $t=40$ (figure 7.6). This re-invasion was at low density, but again the field population was clustered in the north of the field close to the marginal habitat, at a maximum density of approximately 0.013 aphids per m^2.



Immigration begins at julian day 266

(a) t=2 (b) t=20 (c) t=40

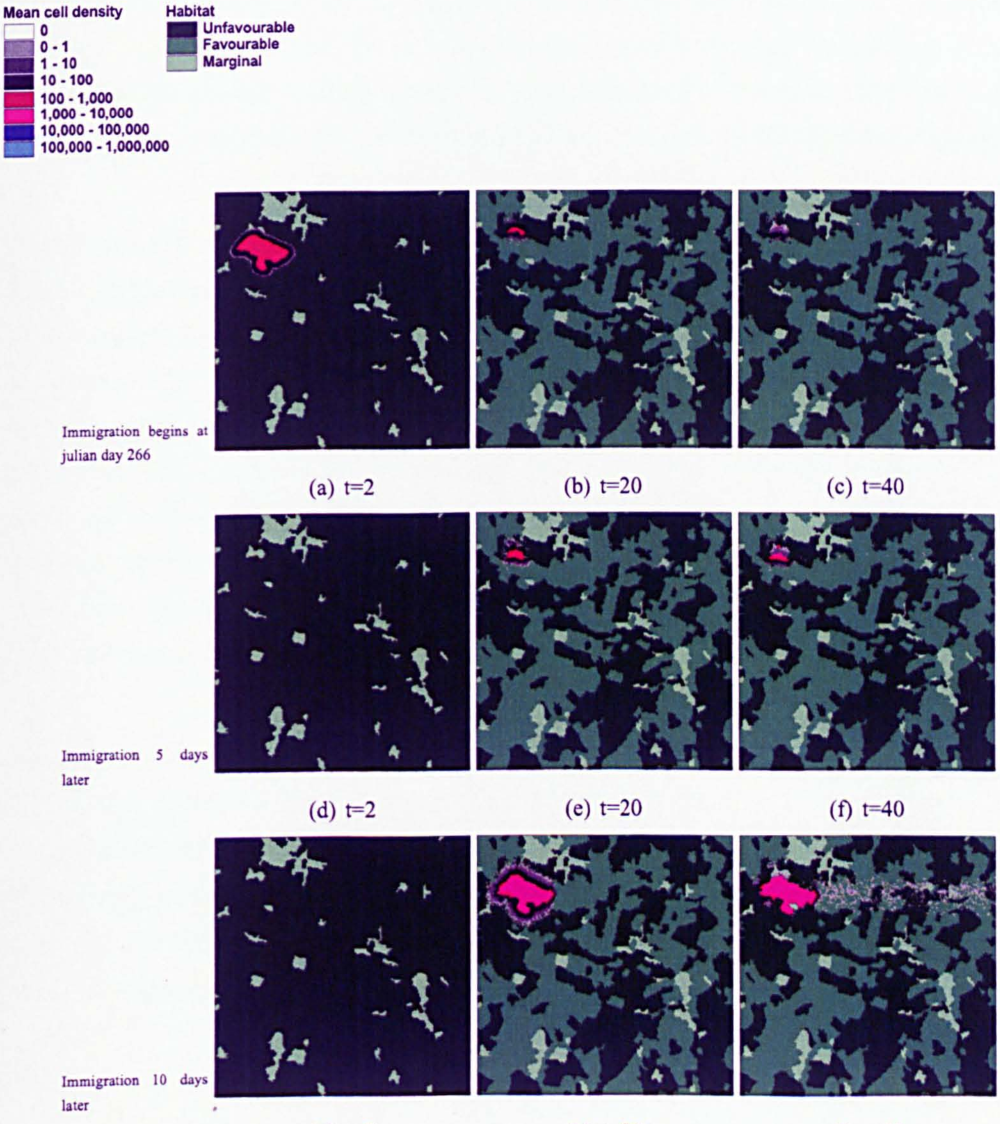
Immigration 5 days later

(d) t=2 (e) t=20 (f) t=40

Immigration 10 days later

(g) t=2 (h) t=20 (i) t=40

Figure 7.4: Spatial population dynamics: sow date = 236.

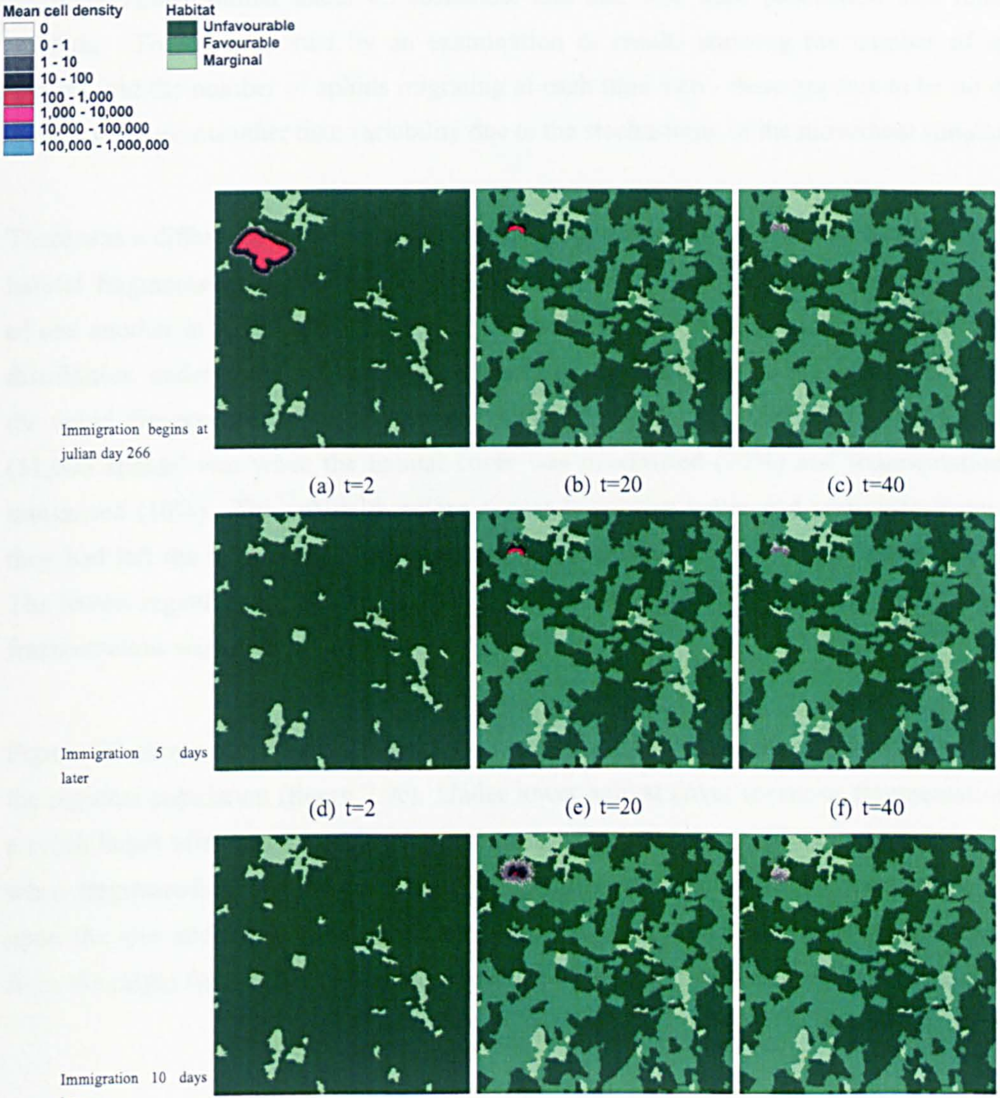


(a) t=2 (b) t=20 (c) t=40

(d) t=2 (e) t=20 (f) t=40

(g) t=2 (h) t=20 (i) t=40

Figure 7.5: Spatial population dynamics: sow date = 246.



(a) t=2 (b) t=20 (c) t=40

(d) t=2 (e) t=20 (f) t=40

(g) t=2 (h) t=20 (i) t=40

Figure 7.6: Spatial population dynamics: sow date = 256.

7.3.2 Fragmentation of crop and crop cover

Temporal results

Figure 7.7 shows the temporal dynamics at the central origin cell within the different landscapes. There was no variability between the population dynamics of the central cell under each landscape scenario. This suggests that the density of the aphids at the origin remained similar under all scenarios, and therefore alate production also remained constant. This is supported by an examination of results showing the number of aphids moving and the number of aphids migrating at each time step - there appears to be no difference in these events other than variability due to the stochasticity of the movement simulations.

There was a difference in the regional population dynamics under different habitat cover and habitat fragmentation scenarios (figure 7.8): the trends were not within the standard error of one another in most cases. The key differences between the regional population size and distribution under each scenario were a result of the influence of the wider landscape on the aphid dispersal and population dynamics (figure 7.8). The highest regional population (51,000 aphids) was when the habitat cover was maximised (90%) and fragmentation was minimised (10%). The aphids therefore appear to survive better and reproduce faster once they had left the origin when there was more favourable habitat that was easily accessible. The lowest regional population was therefore found to be when habitat cover was lowest and fragmentation was highest.

Figure 7.9 shows that when habitat cover was maximised, fragmentation had little effect on the regional population (figure 7.9c). Under lower habitat cover scenarios fragmentation had a much larger effect on the regional population, with the highest regional population reached when fragmentation was lowest. However, spatially, under all scenarios there was an effect upon the size and distribution of the alate population that had moved across the landscape from the origin (see section 7.3.2), despite little difference in the total regional population.

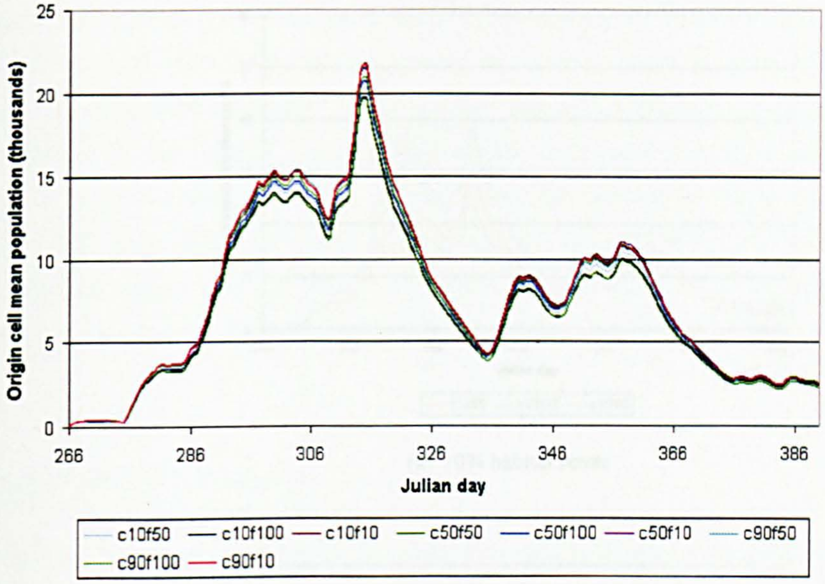


Figure 7.7: Comparison of the temporal population dynamics at the central origin cell under different fragmentation and habitat cover scenarios.

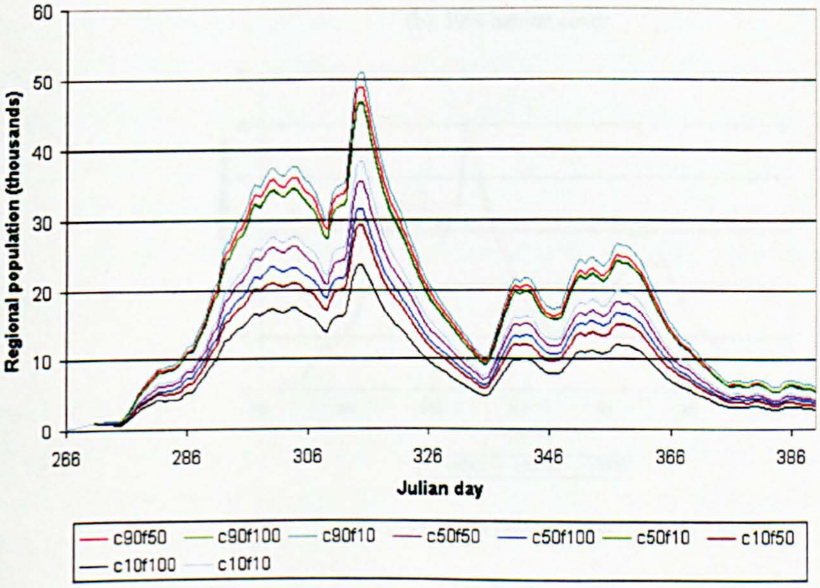
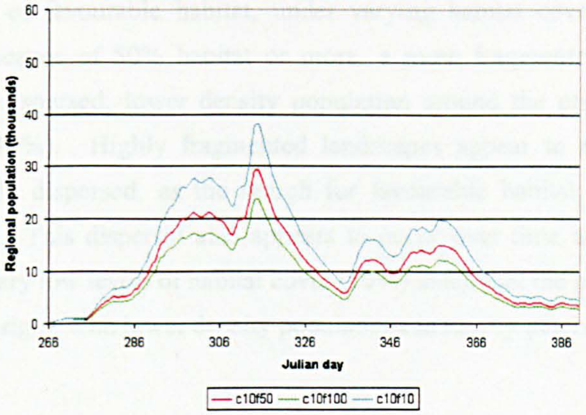
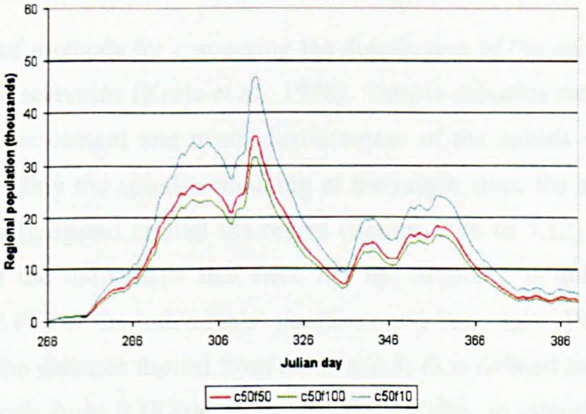


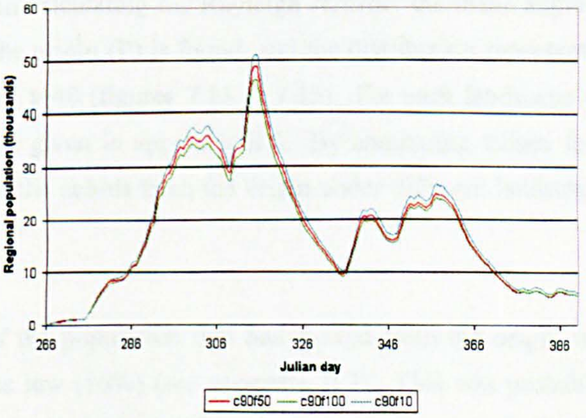
Figure 7.8: Comparison of the temporal population dynamics across the region under different fragmentation and habitat cover scenarios.



(a) 10% habitat cover



(b) 50% habitat cover



(c) 100% habitat cover

Figure 7.9: Comparison of the temporal population dynamics across the region, under different habitat cover scenarios.

Spatial results: overview

Figures 7.10 to 7.12 show the spatial distribution of the aphid population originating from a central cell of favourable habitat, under varying habitat cover and fragmentation scenarios. In landscapes of 50% habitat or more, a more fragmented landscape appears to produce a more dispersed, lower density population around the origin than when fragmentation is low (10%). Highly fragmented landscapes appear to result in populations that are more widely dispersed, as the search for favourable habitat takes aphids further from the origin cell. This dispersal also appears to occur over time as a 'travelling wave' around the origin. Very low levels of habitat cover (10%) mean that the population was highly concentrated at the origin, with lower density populations in nearby patches of isolated habitat.

Spatial results: analysis

There are a number of methods for comparing the distribution of the aphids that have left the origin under different scenarios (Korie et al., 1998). Simple statistics can be used to compare the mean individual movement and mean displacement of the aphids from the origin after 40 days (t=40), excluding the aphids remaining at the origin from the analysis. Clearly, the population remains aggregated around the origin (figures 7.10 to 7.12). Firstly, to quantify any displacement of the individuals that have left the origin, δ is defined as the distance between the centroid (C) of the individuals' positions and the origin (P) (Korie et al., 1998). Second, to quantify the distance moved from the origin P, Δ is defined as the average distance of observed individuals from P (Korie et al., 1998). Thirdly, to quantify the uniformity or otherwise of the circular distribution about P the Rayleigh statistic is used (Fisher et al., 1987: see appendix B.2). In calculating the Rayleigh statistic, the mean angle at which individuals are displaced from the origin (P) is found, and the distribution represented as a rose diagram for each model run at t=40 (figures 7.13 to 7.15). For each landscape configuration, values for these statistics are given in appendix B.3. By comparing values for these statistics, the relative movement of the aphids from the origin under different landscape configurations can be quantified.

The displacement of the population that had moved from the origin was generally highest when crop cover was low (10%) (see appendix B.3). This was probably because the lower proportion of habitat in the landscape forced the aphids to move further from their origin and form a more clustered population once they found favourable habitat. Aphids that choose to move in a direction where less habitat is present (more likely when cover is low) are less likely to survive. At low crop cover, figure 7.13 indicates that displacement from the origin was highly unidirectional, supporting this inference.

The distance moved from the origin by the population seems to be most affected by the fragmentation of the landscape. Highly fragmented landscapes (100%) resulted in aphids moving the greatest distance from the origin. Greater fragmentation equates to more patches within the landscape; thus, although these patches may be small, they enable alate aphids to survive better and to traverse the landscape more easily (see appendix B.3).

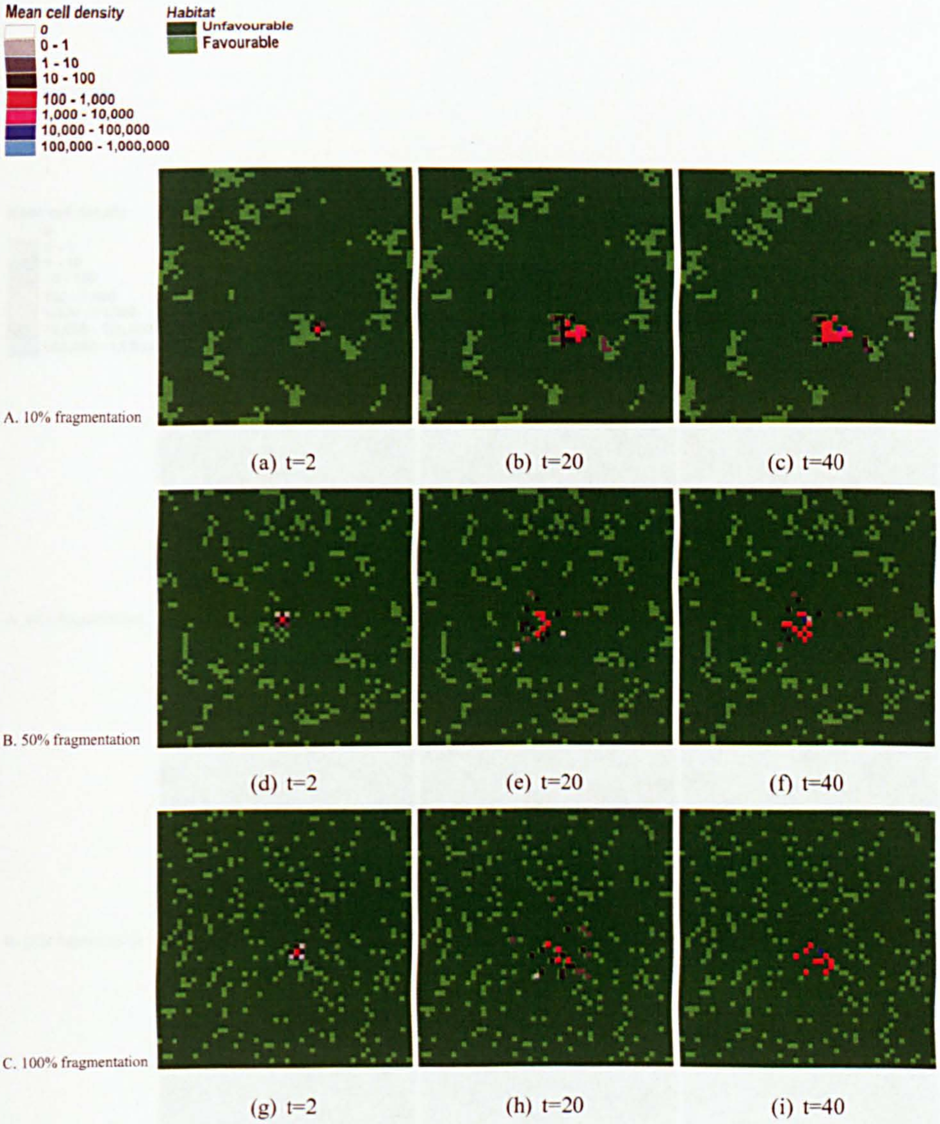
The Rayleigh statistic transformed as per Korie et al. (1998) to $\omega = 3R^2/n$ was compared to a chi-square distribution, and indicated that all distributions are significantly non-uniform. The distributions are not necessarily unimodal, although in most cases it appears that they are (figures 7.13 to 7.15). The only distribution that was definitely not unimodal was c90f50 (figure 7.15C). Thus aphids tended to survive and move in a preferred direction, and are strongly influenced by the presence of favourable habitat, its proximity and configuration (NB wind was excluded in these scenarios).

Spatial results: ANOVA

Using data from the individual simulation runs, a two-way analysis of variance was conducted to quantify variation in the distribution of aphids around the origin, and to prove whether there were significant differences between scenarios and directions in which the population was distributed. The population was split into sectors as for the rose diagrams, and the sectors compared to one another. The use of a two-way analysis of variance identifies whether there was interaction between the scenario and the direction the population was distributed.

The analysis of variance showed that there were significant differences in the numbers of aphids distributed around the origin between scenarios (F=81.76, d.f.=8, $p<0.00$). There was also significant difference between directions of the numbers of aphids distributed around the origin (F=311.10, d.f.=3, $p<0.00$). There was found to be significant interaction between the direction of movement and the habitat configuration scenario (F=57.67, d.f.=24, $p<0.00$). This means that scenario and direction the population was distributed cannot be viewed in isolation, but have a strong interaction with one another to have a combined influence the distribution of the aphids. The direction in which habitat is available may largely depend upon scenario, and appears to vary between scenarios to impact upon the distribution. Figure 7.16 shows the interaction plot of the data, demonstrating the complex interactions between the configuration of the habitat resulting in differing directions of movement and differing distributions of aphids around the origin.

Tukey's pairwise comparisons can be used to assess which scenario-direction combination differs significantly from others (as there are an equal number of replicates across all combinations). This test produces a matrix for each pair of combinations, and proves whether there is a significant difference between them. However, there are 36 combinations, and thus a possible 630 combination pairs. There was therefore no clear trend discernible from these results although individual differences between combinations can be proved. The more general statistics and qualitative assessment made above provide an overall assessment.

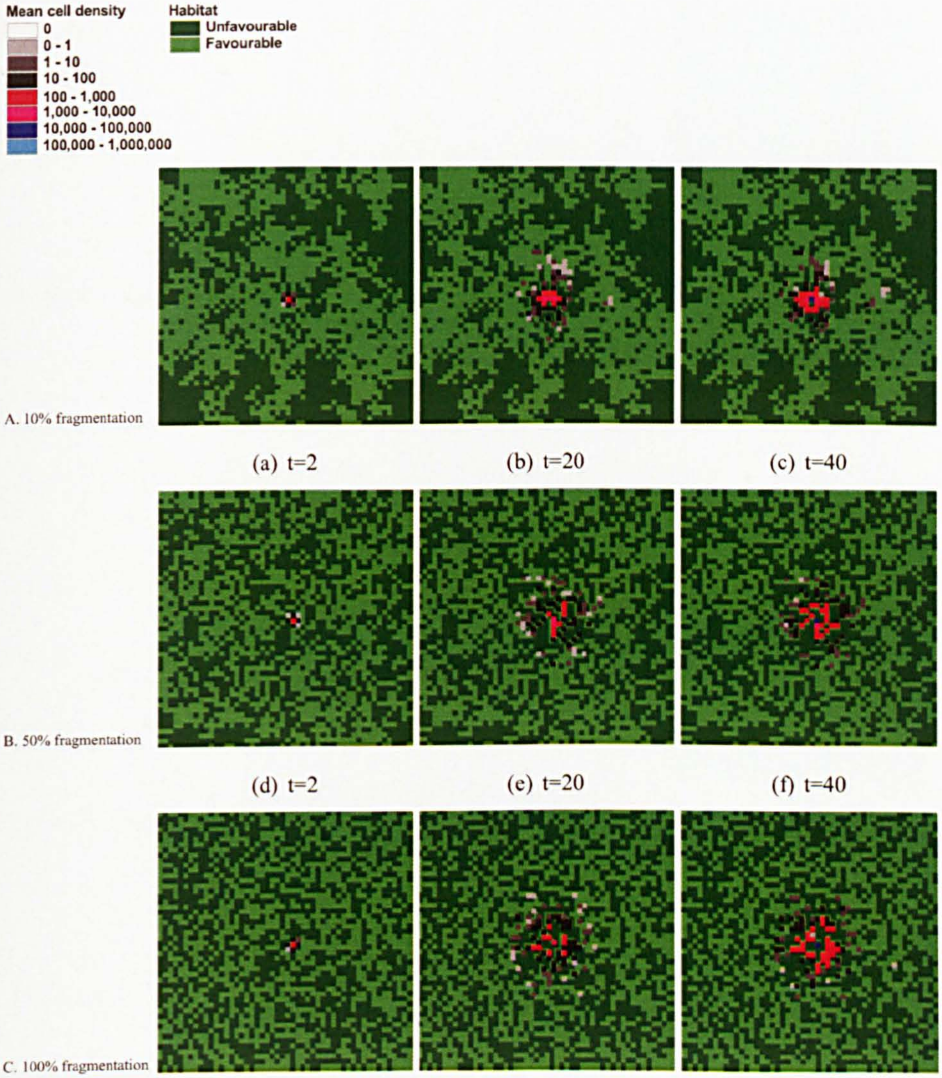


(a) t=2 (b) t=20 (c) t=40

(d) t=2 (e) t=20 (f) t=40

(g) t=2 (h) t=20 (i) t=40

Figure 7.10: Comparison of spatial diffusion of 100 immigrants under different fragmentation scenarios, where habitat cover = 10%.

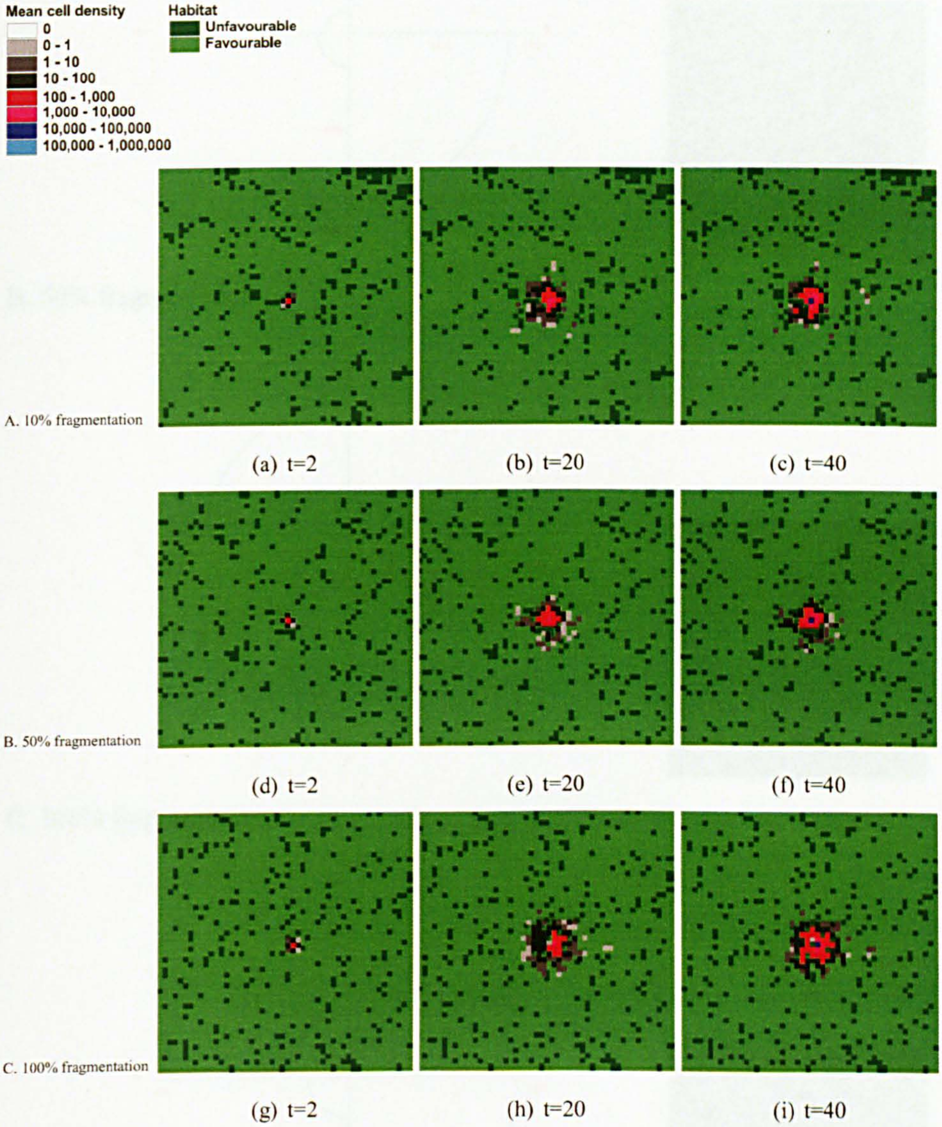


(a) t=2 (b) t=20 (c) t=40

(d) t=2 (e) t=20 (f) t=40

(g) t=2 (h) t=20 (i) t=40

Figure 7.11: Comparison of spatial diffusion of 100 immigrants under different fragmentation scenarios, where habitat cover = 50%.



A. 10% fragmentation

(a) t=2 (b) t=20 (c) t=40

B. 50% fragmentation

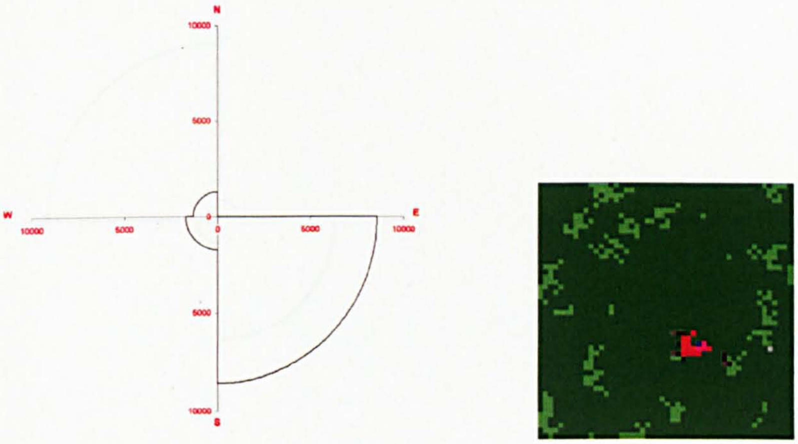
(d) t=2 (e) t=20 (f) t=40

C. 100% fragmentation

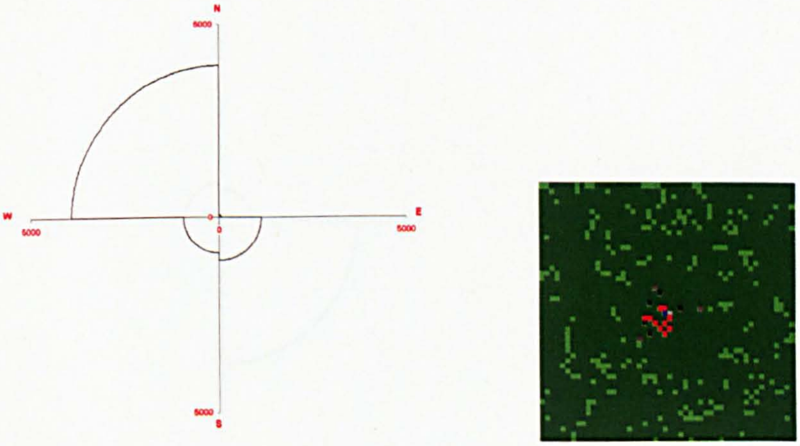
(g) t=2 (h) t=20 (i) t=40

Figure 7.12: Comparison of spatial diffusion of 100 immigrants under different fragmentation scenarios, where habitat cover = 90%.

A. 10% fragmentation



B. 50% fragmentation



C. 100% fragmentation

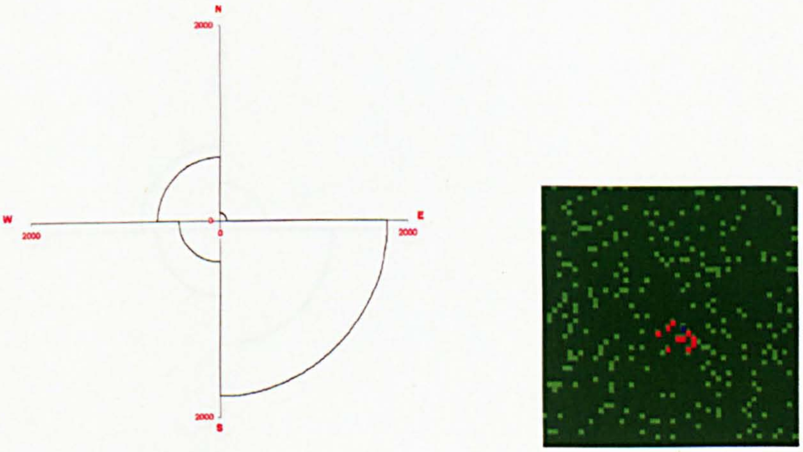
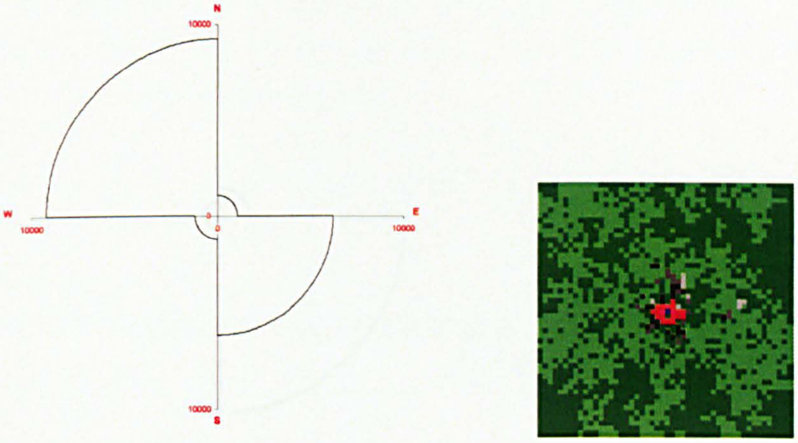
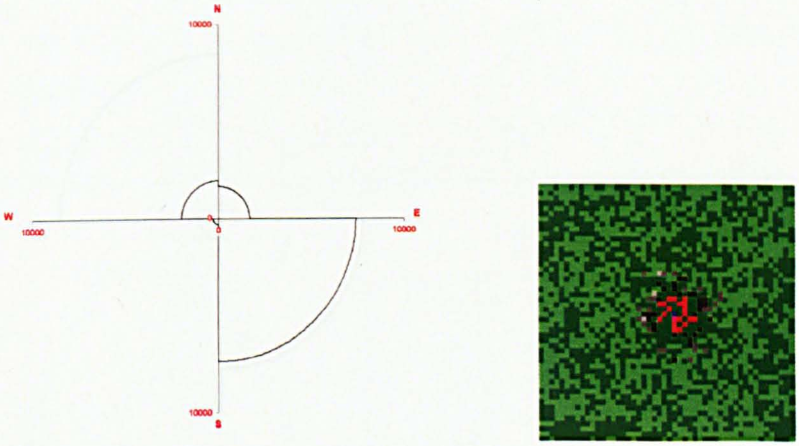


Figure 7.13: Rose diagrams to illustrate the most frequent direction the aphids have moved in after 40 days, for scenarios where habitat cover = 10%. Axes represent numbers of aphids that have moved in each sector.

A. 10% fragmentation



B. 50% fragmentation



C. 100% fragmentation

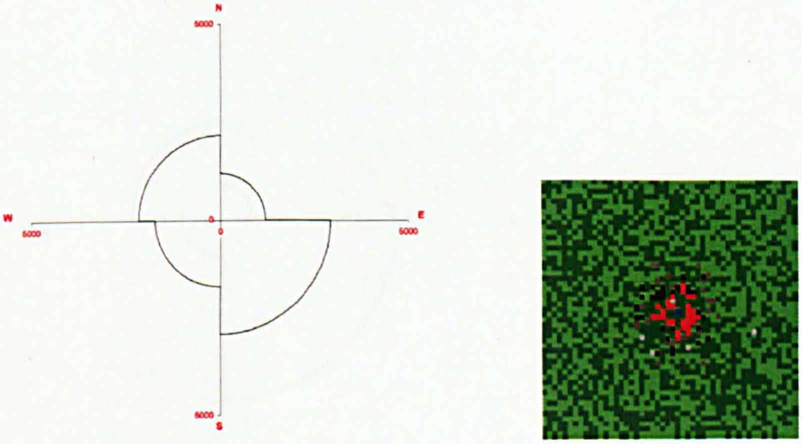
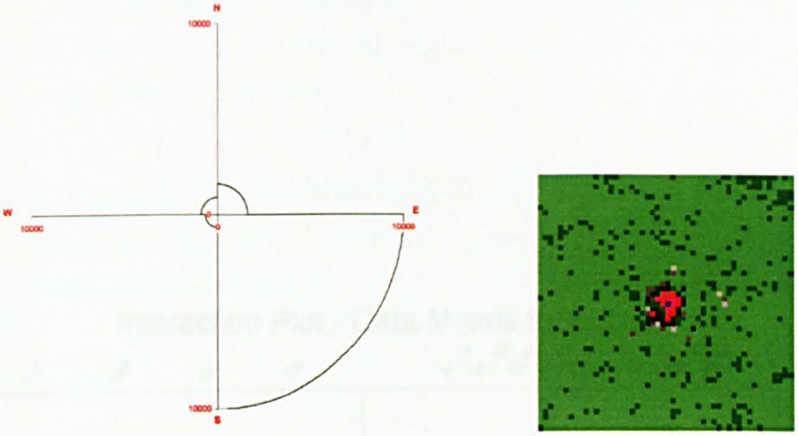
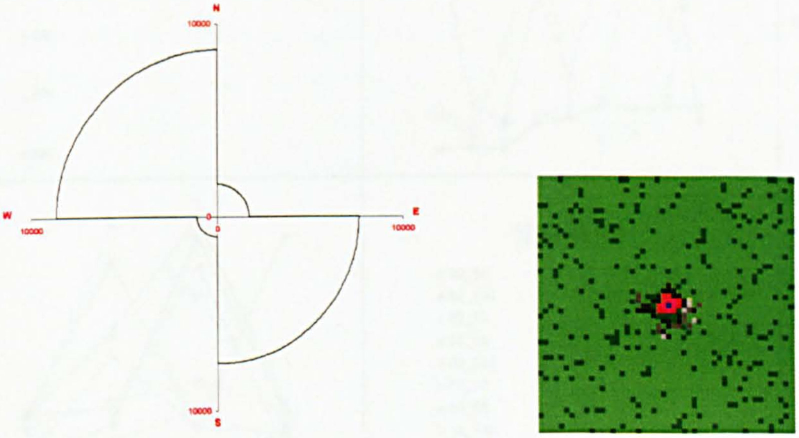


Figure 7.14: Rose diagrams to illustrate the most frequent direction the aphids have moved in after 40 days, for scenarios where habitat cover = 50%. Axes represent numbers of aphids that have moved in each sector.

A. 10% fragmentation



B. 50% fragmentation



C. 100% fragmentation

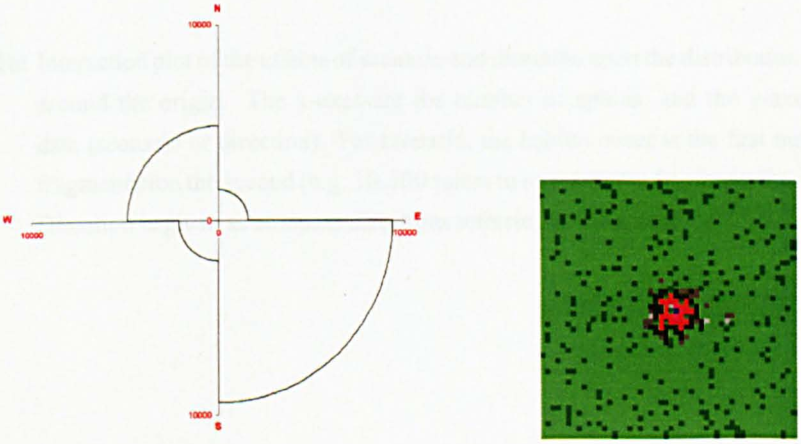


Figure 7.15: Rose diagrams to illustrate the most frequent direction the aphids have moved in after 40 days, for scenarios where habitat cover = 90%. Axes represent numbers of aphids that have moved in each sector.

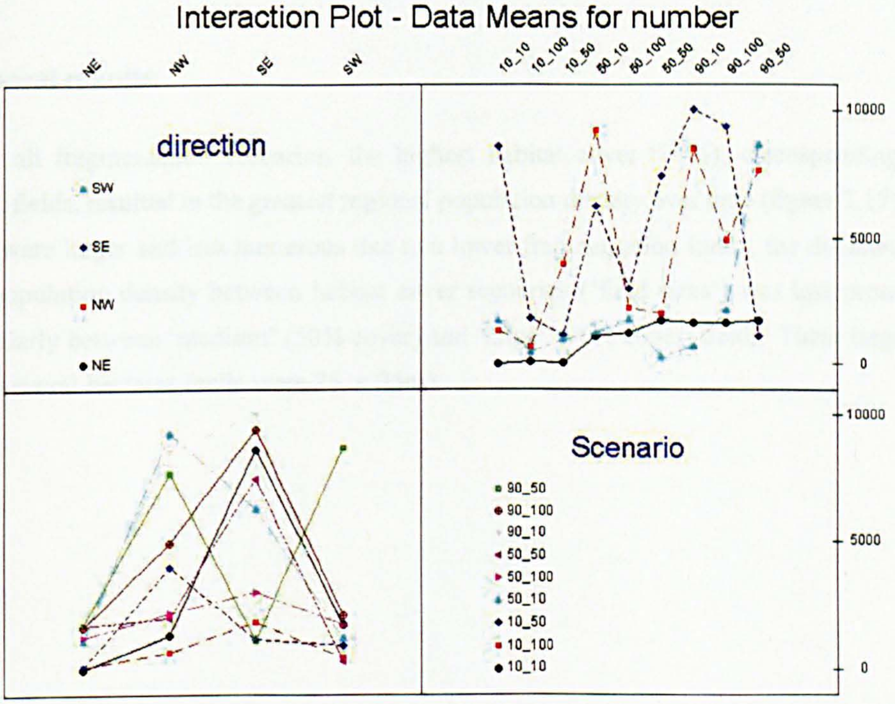


Figure 7.16: Interaction plot of the effects of scenario and direction upon the distribution of aphids around the origin. The x-axes are the number of aphids, and the y-axes are the data (scenario or direction). For scenario, the habitat cover is the first number, the fragmentation the second (e.g. 10_100 refers to cover=10%, fragmentation = 100%). Direction is given as compass directions referring to the sector.

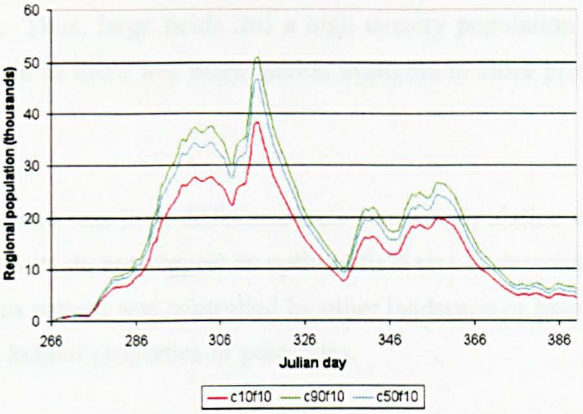
7.3.3 Field size

Alteration of field size was taken to be equivalent to altering habitat cover whilst keeping fragmentation constant (Tischendorf and Fahrig, 2000). The same number of patches will, therefore, exist in the landscape whilst the fragmentation index is constant; the patches increase in size as habitat cover increases.

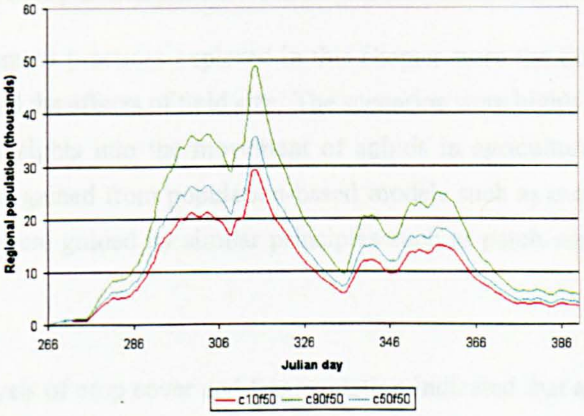
The output used to describe the effects of fragmentation upon species population dynamics in section 7.3.2 is now used to compare the effects of varying habitat cover, under each of 10%, 50% and 100% fragmentation indices.

Temporal results

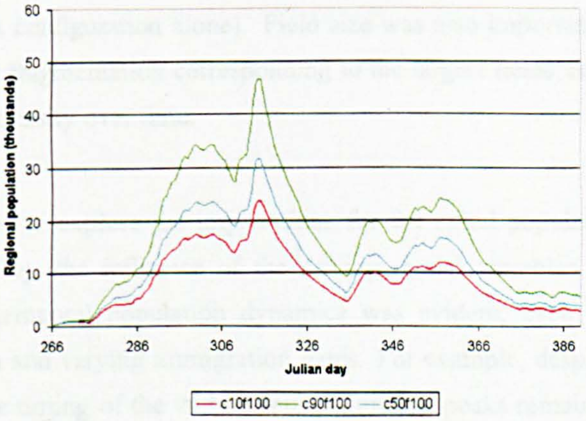
Under all fragmentation scenarios, the highest habitat cover (90%), corresponding to the largest fields, resulted in the greatest regional population density over time (figure 7.17). When fields were larger and less numerous due to a lower fragmentation index, the difference in *R. padi* population density between habitat cover scenarios ('field sizes') was less pronounced, particularly between 'medium' (50% cover) and 'large' (90% cover) fields. These larger fields cover several hectares (cells were 25 × 25m).



(a) 10% fragmentation



(b) 50% fragmentation



(c) 100% fragmentation

Figure 7.17: Comparison of the temporal population dynamics across the region under different habitat cover scenarios (='field size'), where fragmentation is kept constant at a) 10%, b) 50% and c) 100%.

Spatial results

In general, as field size increased (habitat cover increased) the distance the aphids moved from the origin decreased. Thus, large fields had a high density population concentrated closely together near the origin as there was more habitat available in close proximity (see appendix B.3).

As previously noted there was little difference between the population densities at the origin cell. Overall, these results do not suggest an optimal field size as demonstrated by Foster et al. (2004), suggesting this pattern was controlled by other landscape or management factors such as different marginal habitat properties or pesticides.

7.4 Discussion and summary

Agricultural management practices explored in this chapter were the timing of crop sowing, crop configuration and the effects of field size. The scenarios were highly simplified; however, they do give some insights into the movement of aphids in agricultural landscapes, above those which might be gained from population-based models such as metapopulation models, while the scenarios were guided by similar principles such as patch size, fragmentation and distribution.

The theoretical analysis of crop cover and fragmentation indicated that aphids tend to survive and move in a preferred direction, and were strongly influenced by the presence of favourable habitat, its proximity and configuration (wind was excluded from the analysis to demonstrate the effects of habitat configuration alone). Field size was also important, the highest habitat cover and the lowest fragmentation corresponding to the largest fields, resulted in the greatest regional population density over time.

In the scenarios used to explore the implications for the aphid population dynamics of the timing of crop sowing, the influence of the driving model variables such as temperature (chapter 5) in the temporal population dynamics was evident, even under quite extreme habitat manipulation and varying immigration dates. For example, despite variation in initial immigration date the timing of the within-field population peaks remained consistent (but at lower magnitude) when the crop had emerged before the initial immigration of aphids, or was up to five days later emerging. When the crop was ten days late in emerging it had a greater impact on the population dynamics of the within-field population, altering the timing of the population peaks or even eliminating the population altogether, thus over-riding the previously driving effects of temperature.

Overall, it can be surmised that in these scenarios a late sowing date was most effective in reducing aphid populations within the field. This was most effective when initial aphid immigration was relatively early (julian day 266), and least effective if initial immigration was later. If the sowing date was early, the aphid population density was higher overall, and earlier immigrants reached higher densities as they were able to take advantage of the emergent crop and favourable temperature conditions earlier.

The individual-based approach allowed for a spatial examination of the implications of the crop sowing scenarios, indicating the importance of marginal habitat in sustaining the within-field population when immigration preceded crop emergence. The importance of the adjacent marginal habitat was most evident when sowing takes place latest, on julian day 256 (figure 7.6). Populations surviving in the marginal habitat sustain the regional population for longer than the field population, and they clearly provide a source of aphids invading the crop once it had emerged. It was evident spatially, as well as temporally, that the latest sowing of the crop reduced aphid populations most significantly overall, both within-field and landscape-wide. The regional displacement of aphids was much more limited under late sowing and early immigration scenarios.

The manipulation of the habitat fragmentation and habitat cover altered the magnitude of the population peaks, but did not alter the overall population trends. Thus, it can be inferred that manipulating crop configuration and also field size had less impact upon the temporal population dynamics than manipulating the crop sowing regime. However, there are important spatial implications of altering the habitat fragmentation and cover, that are measurable through an individual-based approach. Statistical evidence given shows that there was significant variability in the direction of spread of aphids from a single origin under different fragmentation and habitat cover scenarios. The shape of this distribution varied between scenarios, and there is proven to be significant difference in the overall distribution between these scenarios in each of four compass directions. In general, it was perceived that high fragmentation (100%) with low habitat cover (10-50%) produced a more dispersed population distribution than low fragmentation and high habitat cover. High cover and low fragmentation, equated to large fields, tended to result in less movement from the origin unless in the direction of favourable habitat. This was due to the greater availability of habitat within the range of movement of the aphids when there was more habitat cover that was easily accessible (less fragmented), leading to higher population densities, greater survival and lower dispersal. This information could aid pest control strategies, for example targeted spray application, and would be further enhanced by information from remote sensing surveys.

CHAPTER 8

The Influence of Non-cropped Habitat

8.1 Introduction

Marginal habitat potentially plays an important and complex role in both local and regional population dynamics of pest species such as aphids. Marginal habitat, such as grassland and moorland, has been found to harbour aphids (Foster et al., 2004). Thus, a greater availability of surrounding marginal habitat in the landscape matrix could influence aphid density within the crop. Assuming there are patches of habitat in the landscape that can sustain populations to some degree between fields and between crops, this study examines the ways in which the influences of marginal habitat can be studied and understood for the aphid species *Rhopalosiphum padi*.

More complex theoretical landscapes were developed than the binary landscapes of the previous chapter (see chapter 3, section 3.2.1). The amount and quality of marginal habitat in the matrix were varied, as well as its spatial configuration and fragmentation. In order to move away from the view of agricultural landscapes as binary, common in aspatial models of patch dynamics, complexity was introduced that allows for analyses of population dynamics beyond that which could achieved by a standard metapopulation model. The model is able to analyse the effects of landscapes that are characterised as spatially explicit mosaics of different habitat types. In real agricultural landscapes a binary, patch-based view is less appropriate, as real landscapes are in fact not composed of isolated islands of habitat in a hostile matrix, particularly when viewed from the perspective of a mobile, generalist pest species like an aphid (see section 2.2, chapter 2). To isolate the effects of habitat, in this analysis the wind model was excluded from the simulation runs (meaning that only local movement as a function of habitat response is modelled for these scenarios, not migration).

Using the model, the degree to which marginal habitat is important to aphid population dynamics within crops and across landscapes was examined. The model also helps to understand the way in which differing availability and configuration of marginal habitat influences individuals such that variations in population densities and distributions were observed.

8.2 Method

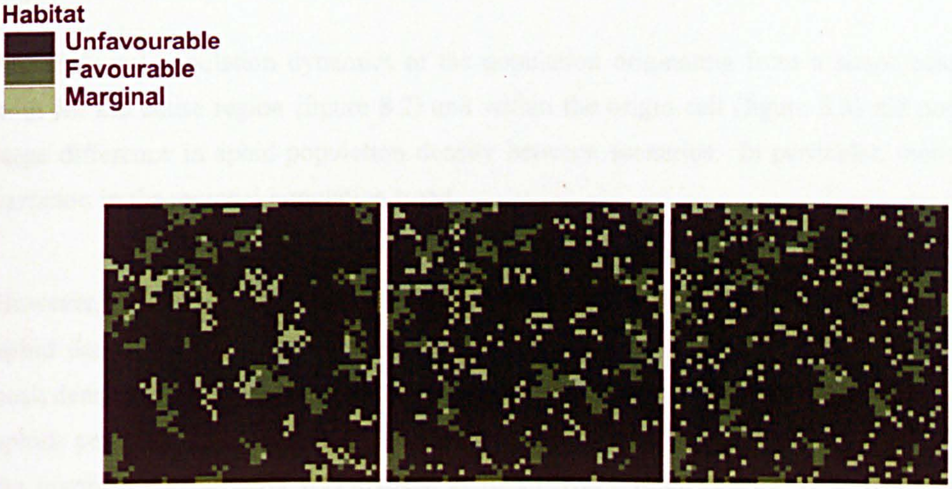
The marginal habitat availability in the model was varied by setting an increasing proportion of the non-habitat in the theoretical landscapes to marginal habitat (see section 3.2, chapter 3). The habitat cover and fragmentation of the favourable habitat remained constant at 10% habitat cover and 10% fragmentation. The percentage of marginal habitat in the remaining area was altered (10%, 50%, 90%), as was the marginal habitat fragmentation (10%, 50%, 100%). Table 8.1 gives the resultant habitat configurations, which are illustrated in figure 8.1. Configurations, when abbreviated, are given as the marginal habitat fragmentation (mf) followed by their marginal habitat configuration (mc): for example 10% marginal habitat cover and 50% marginal fragmentation is referred to as "mf50mc10".

Marginal habitat	10% fragmentation	50% fragmentation	100% fragmentation
10%	a	b	c
50%	d	e	f
90%	g	h	i

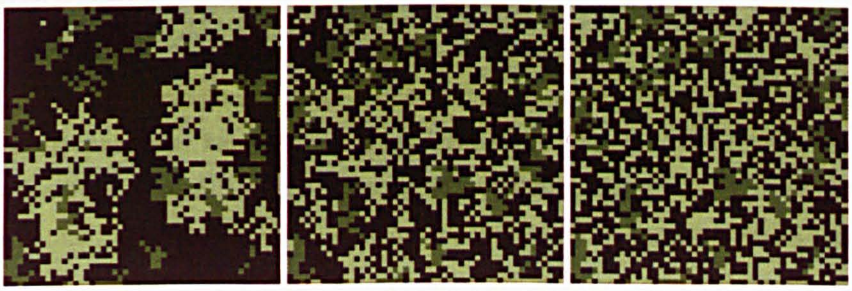
Table 8.1: Habitat configuration matrix for fragmentation and availability of marginal habitat, letters refer to landscapes in figure 8.1.

Two studies were undertaken with these landscapes. The first examined aphid dispersal across the matrix. This concentrated on the population dynamics and distribution following a single immigration event of 100 aphids at a central (or near central) 25 × 25m cell of favourable habitat. In this way the effects of the matrix upon species movement and dispersal could be isolated.

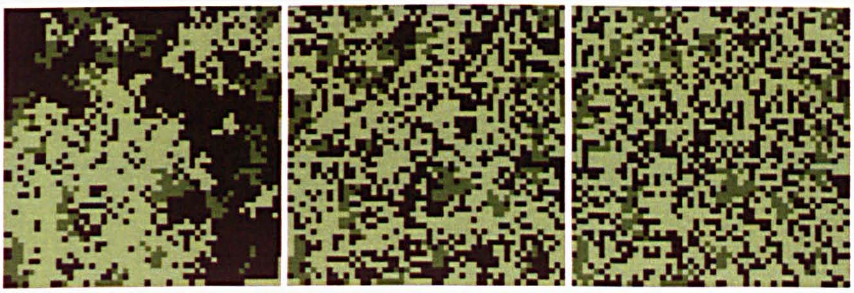
The second study population was initiated as an even distribution of aphids at the same density (100 aphids per 25 × 25m cell) across the landscape. This assessed the regional population dynamic implications of varying marginal habitat distribution.



(a) 10% marginal fragmentation, 10% marginal cover (b) 50% marginal fragmentation, 10% marginal cover (c) 100% marginal fragmentation, 10% marginal cover



(d) 10% marginal fragmentation, 50% marginal cover (e) 50% marginal fragmentation, 50% marginal cover (f) 100% marginal fragmentation, 50% marginal cover



(g) 10% marginal fragmentation, 90% marginal cover (h) 50% marginal fragmentation, 90% marginal cover (i) 100% marginal fragmentation, 90% marginal cover

Figure 8.1: Fragmentation versus Habitat Cover indices and the resultant landscape.

8.3 Results

8.3.1 Aphid dispersal across the matrix from a single origin cell

Temporal results

The temporal population dynamics of the population originating from a single central cell, both for the entire region (figure 8.2) and within the origin cell (figure 8.3) did not show a large difference in aphid population density between scenarios. In particular, there was no variation in the regional population trend.

However, there was some variation between the trends at the central cell, where the greatest aphid densities were achieved when the marginal habitat cover was 50%, with the highest peak density at 50% marginal habitat cover and 100% marginal habitat fragmentation (22,042 aphids per 25 × 25 m field cell). The lowest population density at the central cell was when the marginal habitat cover was highest, at 90% (19,037 when cover = 90%, fragmentation = 100%; 86% of the highest peak density). There was less than 14% difference in the population peak between scenarios.

It can be inferred that marginal habitat fragmentation is not important to the density at the central cell, but marginal habitat cover does have some impact. However, overall, there appears to be little temporal impact upon aphid population density with varying marginal habitat cover and fragmentation when studying a population originating from a single central cell.

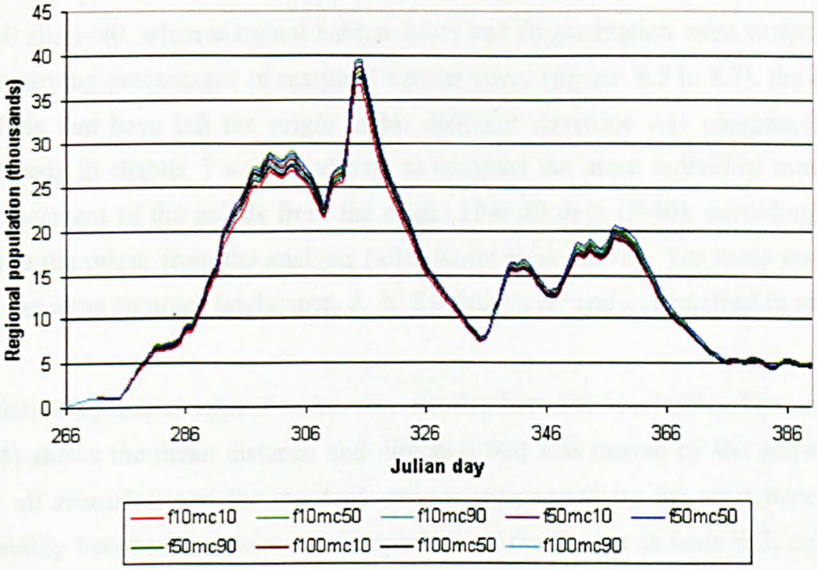


Figure 8.2: Comparison of the temporal population dynamics of the total regional aphid population, under different marginal habitat fragmentation and marginal habitat cover scenarios.

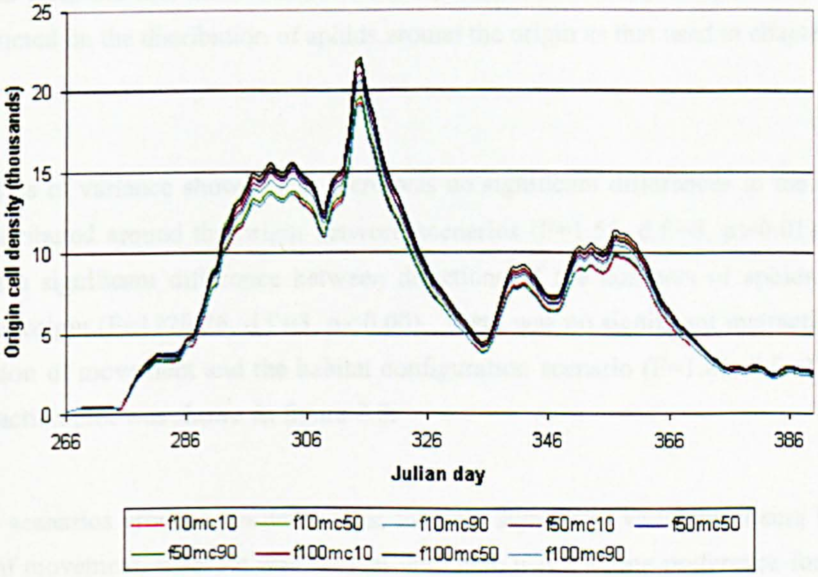


Figure 8.3: Comparison of the temporal population dynamics of the origin cell, under different marginal habitat fragmentation and marginal habitat cover scenarios.

Spatial results: overview

Spatially, there appears to be variation in the distribution of the aphids in relation to the origin cell at t=20 and t=40, when marginal habitat cover and fragmentation were varied. Using the results for varying percentages of marginal habitat cover (figures 8.5 to 8.7), the distribution of the aphids that have left the origin under different scenarios was compared. Methods used previously in chapter 7 were used here to compare the mean individual movement and mean displacement of the aphids from the origin after 40 days (t=40), excluding the aphids remaining at the origin from the analysis (after Korie et al., 1998). The same statistics were used for these more complex landscapes, δ, Δ, Rayleigh's R^2 and ω, described in section 7.3.2.

The population dynamics appear to be very similar between scenarios. The rose diagram (figure 8.4) shows the mean distance and direction that was moved by the aphids from the origin for all scenarios, and the standard error is represented by the error bars; there was little variability between scenarios. A comparison of the figures in table B.3, appendix B.3 indicates strong similarity between the scenarios as most statistics appear to give similar results; the direction of movement was very similar, the number of aphids that move was similar, the mean distance moved was similar, under all scenarios displacement was minimal and the Rayleigh statistics indicate that all distributions were significantly non-uniform.

Spatial results: ANOVA

Using data from the individual simulation runs, a similar two-way analysis of variance test was conducted on the distribution of aphids around the origin as that used in chapter 7, section 7.3.2.

The analysis of variance showed that there was no significant differences in the numbers of aphids distributed around the origin between scenarios (F=1.55, d.f.=8, $p>0.01$). However, there was a significant difference between directions of the numbers of aphids distributed around the origin (F=1220.76, d.f.=3, $p<0.00$). There was no significant interaction between the direction of movement and the habitat configuration scenario (F=1.81, d.f.=24, $p>0.01$). The interaction plot was shown in figure 8.8.

Thus, the scenarios produce similar results; the only significant variability being between directions of movement, where it was evident that there was a strong preference for movement in a south-easterly direction. This was a factor of the crop configuration, as can be seen by referring to the previous chapter (chapter 7), figure 7.13a, which was the same landscape, but without marginal habitat. Thus, overall, there was little impact on the population distribution around the origin of varying marginal habitat cover and fragmentation.

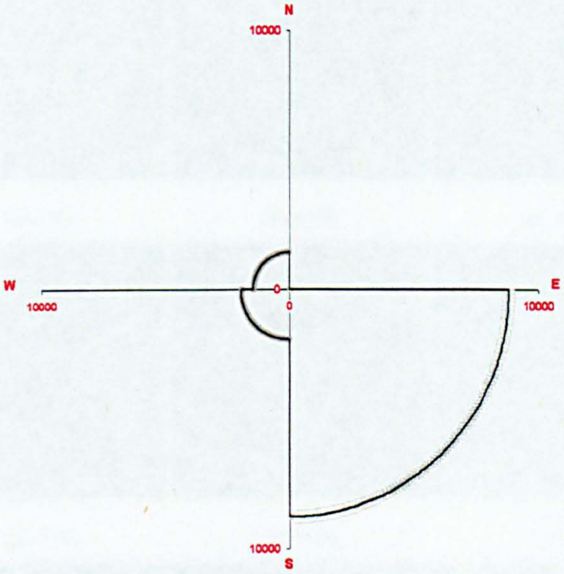
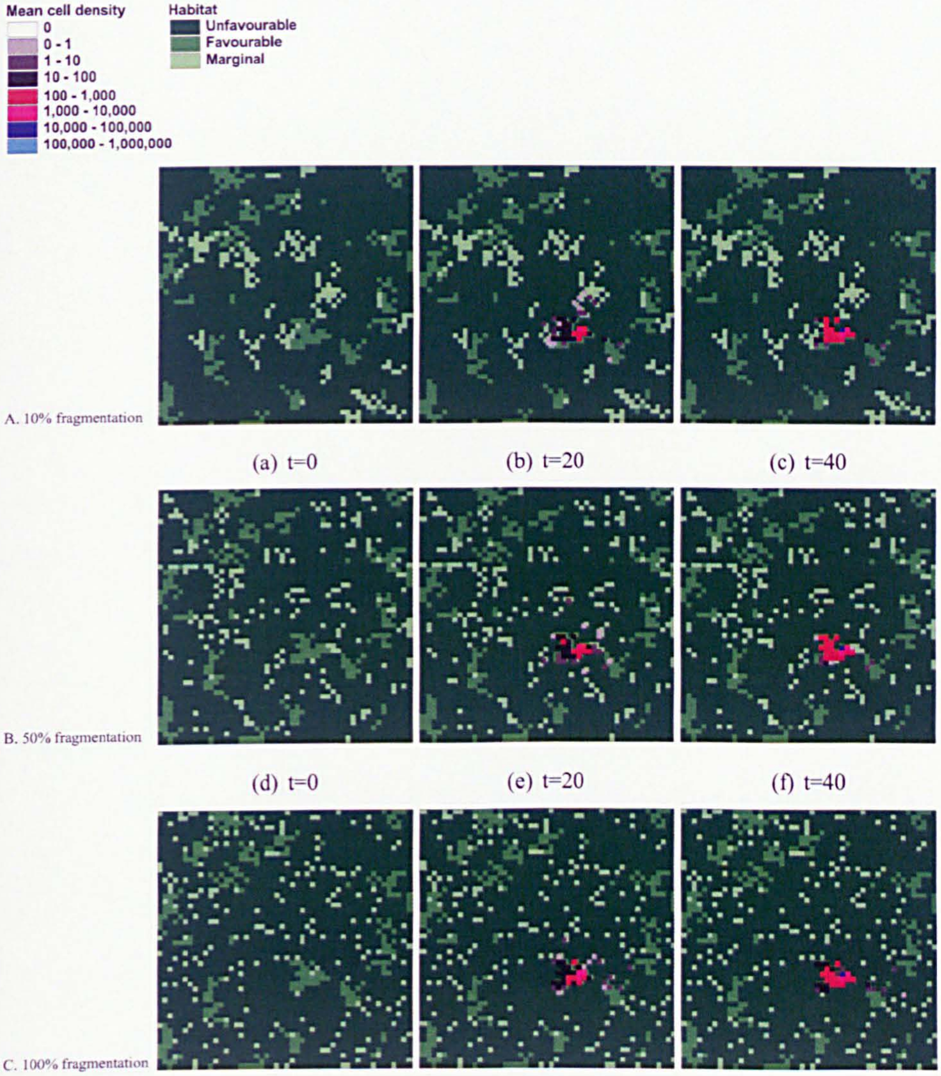


Figure 8.4: Rose diagrams to illustrate the mean most frequent direction the aphids have moved in after 40 days, for all scenarios. Axes represent numbers of aphids that have moved in each sector. Error bars indicate standard error between scenarios.

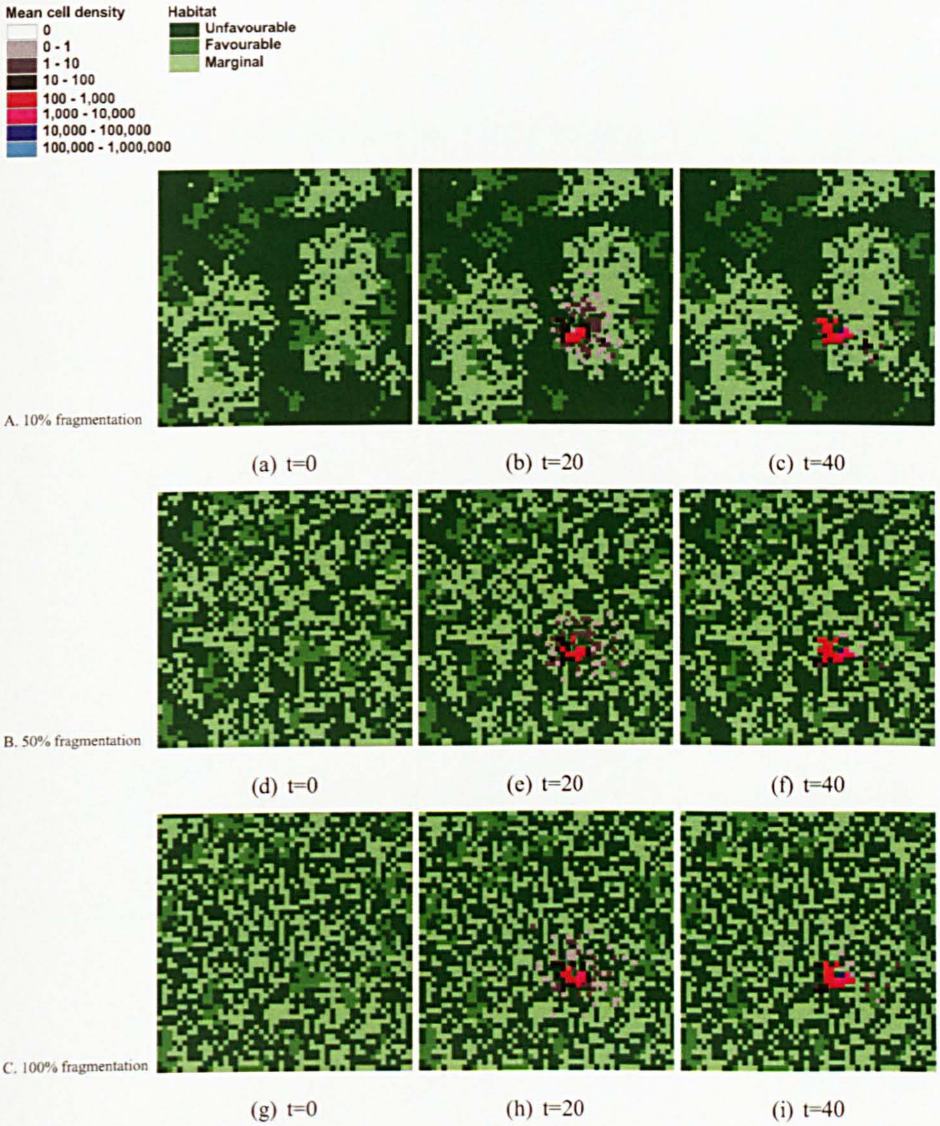


(a) t=0 (b) t=20 (c) t=40

(d) t=0 (e) t=20 (f) t=40

(g) t=0 (h) t=20 (i) t=40

Figure 8.5: Comparison of mean spatial distribution under different marginal habitat fragmentation scenarios, where marginal habitat cover = 10%.

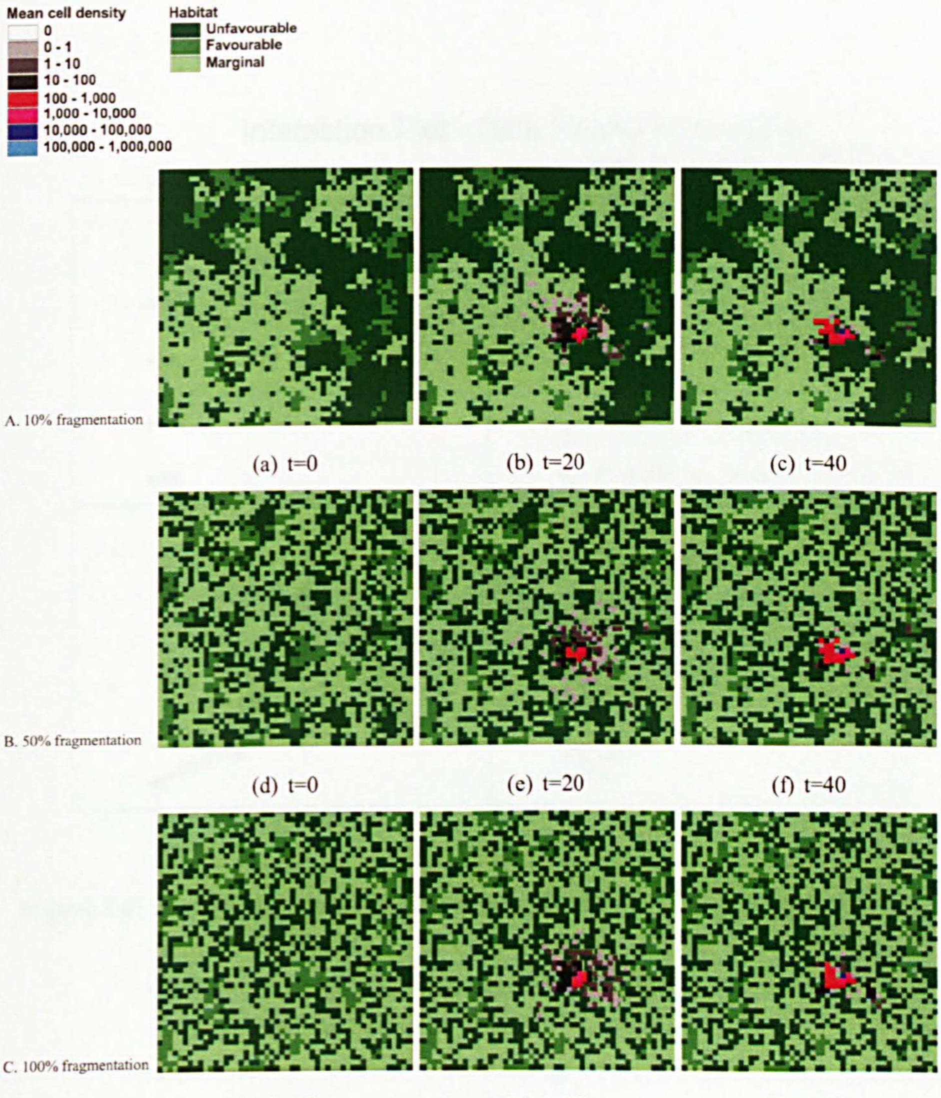


(a) t=0 (b) t=20 (c) t=40

(d) t=0 (e) t=20 (f) t=40

(g) t=0 (h) t=20 (i) t=40

Figure 8.6: Comparison of mean spatial distribution under different marginal habitat fragmentation scenarios, where marginal habitat cover = 50%.



(a) t=0 (b) t=20 (c) t=40

(d) t=0 (e) t=20 (f) t=40

(g) t=0 (h) t=20 (i) t=40

Figure 8.7: Comparison of mean spatial distribution under different marginal habitat fragmentation scenarios, where marginal habitat cover = 90%.

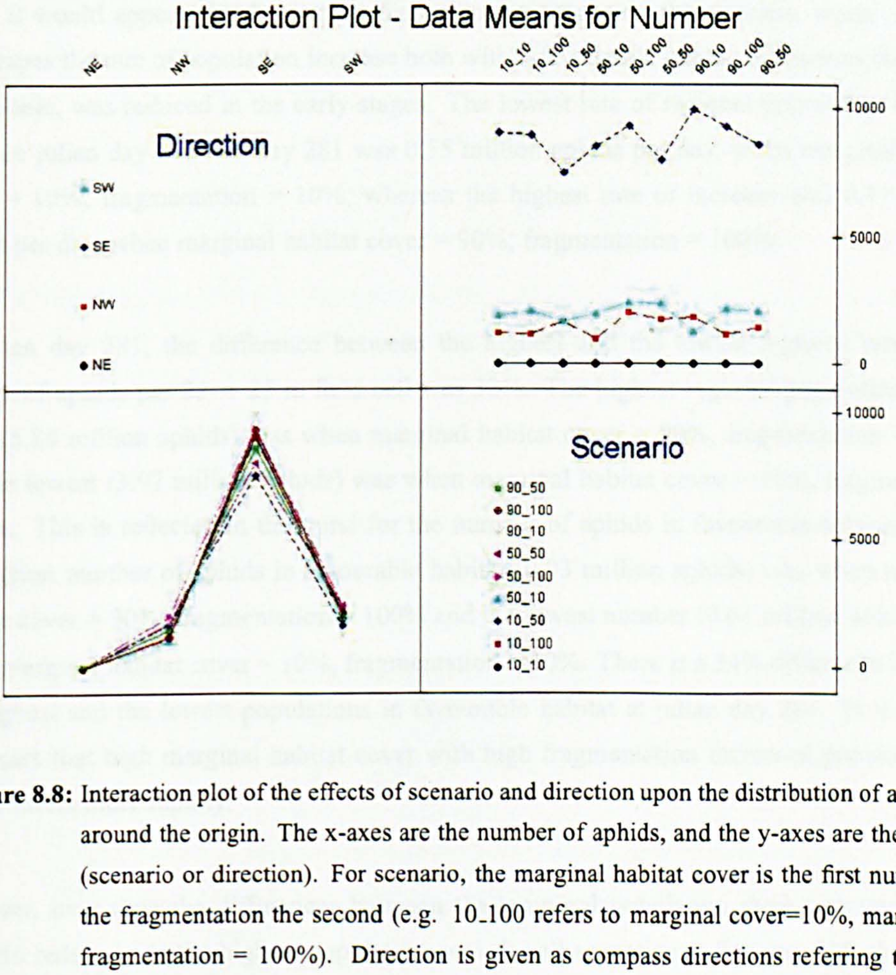


Figure 8.8: Interaction plot of the effects of scenario and direction upon the distribution of aphids around the origin. The x-axes are the number of aphids, and the y-axes are the data (scenario or direction). For scenario, the marginal habitat cover is the first number, the fragmentation the second (e.g. 10_100 refers to marginal cover=10%, marginal fragmentation = 100%). Direction is given as compass directions referring to the sector.

8.3.2 Impacts of the availability and fragmentation of marginal habitat, for immigrants spread across the region

Temporal results

Examining the population dynamics over time for the region (figure 8.9) and also within areas of favourable habitat (figure 8.10) when immigrants were distributed evenly, there appears to be little difference between the habitat configuration scenarios. The greatest variation in population density occurred at julian day 281. This was as the initial immigrant population was moving around the landscape and beginning to produce the second generation of aphids. Thus, it would appear that habitat configuration impacts upon this process, where in some landscapes the rate of population increase both within favourable habitat and across the region as a whole, was reduced in the early stages. The lowest rate of regional population increase between julian day 275 and day 281 was 0.55 million aphids per day, when marginal habitat cover = 10%, fragmentation = 10%, whereas the highest rate of increase was 0.77 million aphids per day, when marginal habitat cover = 90%, fragmentation = 100%.

At julian day 281, the difference between the highest and the lowest regional population density of aphids per 25 × 25 m field cell was 32%. The highest regional population at this point (5.84 million aphids) was when marginal habitat cover = 90%, fragmentation = 100% and the lowest (3.97 million aphids) was when marginal habitat cover = 10%, fragmentation = 10%. This is reflected in the trend for the number of aphids in favourable habitat, where the highest number of aphids in favourable habitat (0.93 million aphids) was when marginal habitat cover = 90%, fragmentation = 100% and the lowest number (0.61 million aphids) was when marginal habitat cover = 10%, fragmentation = 10%. There is a 34% difference between the highest and the lowest populations in favourable habitat at julian day 281. In summary, it appears that high marginal habitat cover with high fragmentation increased populations to higher levels more rapidly.

However, over time the differences between the temporal population dynamics under each scenario reduced. At the highest population peak for all scenarios, julian day 315, the lowest peak was only 5% less than the highest.

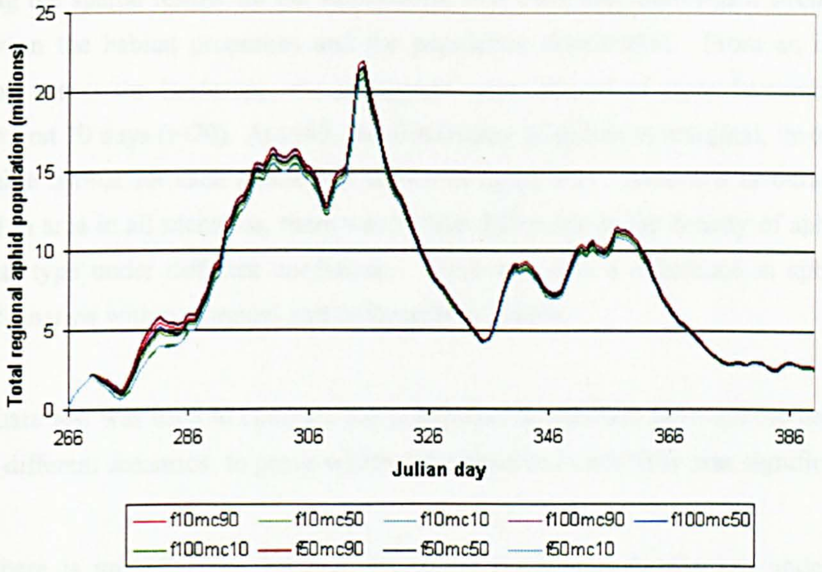


Figure 8.9: Comparison of the temporal population dynamics of the total regional aphid population, under different marginal habitat fragmentation and marginal habitat cover scenarios.

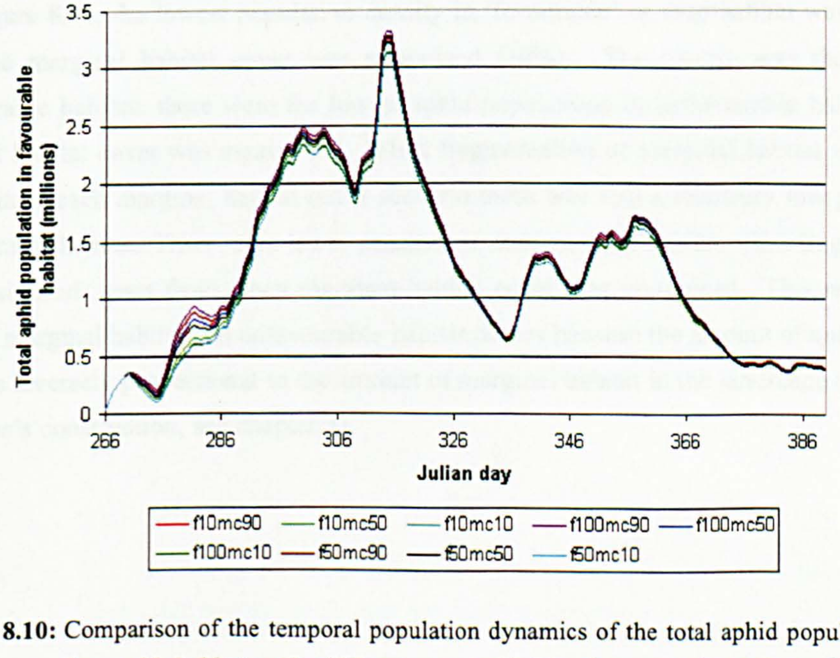


Figure 8.10: Comparison of the temporal population dynamics of the total aphid population in favourable habitat, under different marginal habitat fragmentation and marginal habitat cover scenarios.

Spatial results

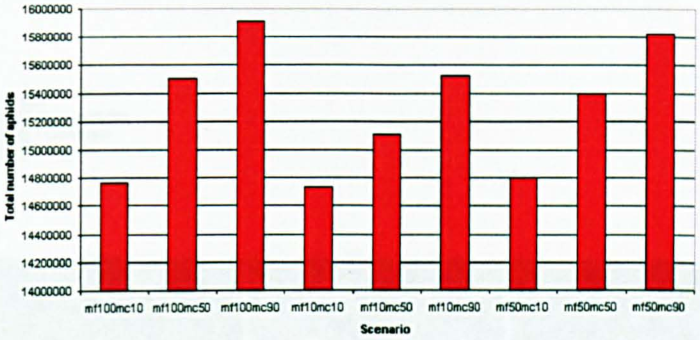
Comparing the spatial results for the simulations, it is clear that there was a strong relationship between the habitat properties and the population distribution. From an initial even distribution across the landscape, the immigrant alates moved to more favourable habitat, within the first 20 days (t=20). At t=40, the distribution of aphids in marginal, favourable and unfavourable habitat for each scenario is shown in figure 8.11. Although favourable habitat was equal in area in all scenarios, there was a clear difference in the density of aphids within this habitat type under different conditions. There was also a difference in aphid density between scenarios within marginal and unfavourable habitat.

A Chi-square test was used to compare the population distribution between the habitat types under the different scenarios, to prove whether the observed variability was significant:

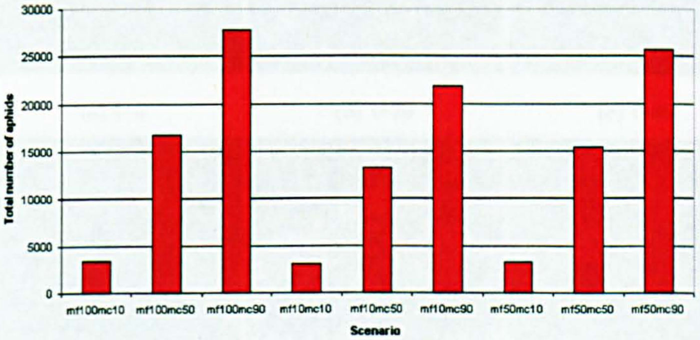
H_0 = There is no difference between the aphid population distribution under different marginal habitat fragmentation and cover scenarios, for each habitat type.
H_1 = There is a significant difference between the aphid population distribution under different marginal habitat fragmentation and cover scenarios, for each habitat type.

The chi-square test on the data (appendix B.1) proved that the null hypothesis could be rejected. There was a significant difference between the aphid population distribution under different marginal habitat fragmentation and cover scenarios, for each habitat type.

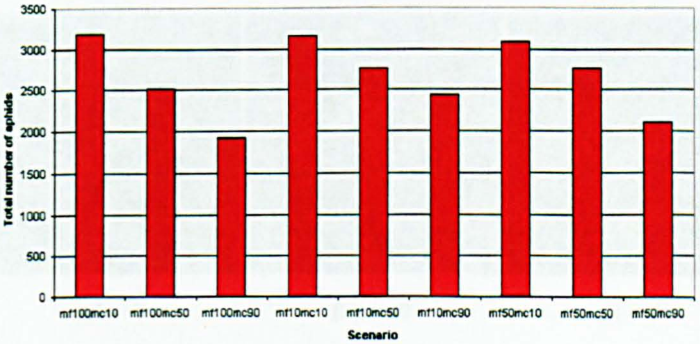
From figure 8.11, the lowest population density in 'favourable' or crop habitat was achieved when the marginal habitat cover was minimised (10%). The inverse was the case for unfavourable habitat: there were the lowest aphid populations in unfavourable habitat when marginal habitat cover was maximised. When fragmentation of marginal habitat was lowest (10%) under each marginal habitat cover scenario there was also a relatively low population in favourable habitat. There were lower densities in unfavourable habitat when fragmentation was maximised, apart from when marginal habitat cover was minimised. This relationship between marginal habitat and unfavourable habitat occurs because the amount of unfavourable habitat is inversely proportional to the amount of marginal habitat in the landscape (due to the landscape's construction, see chapter 3).



(a) Favourable

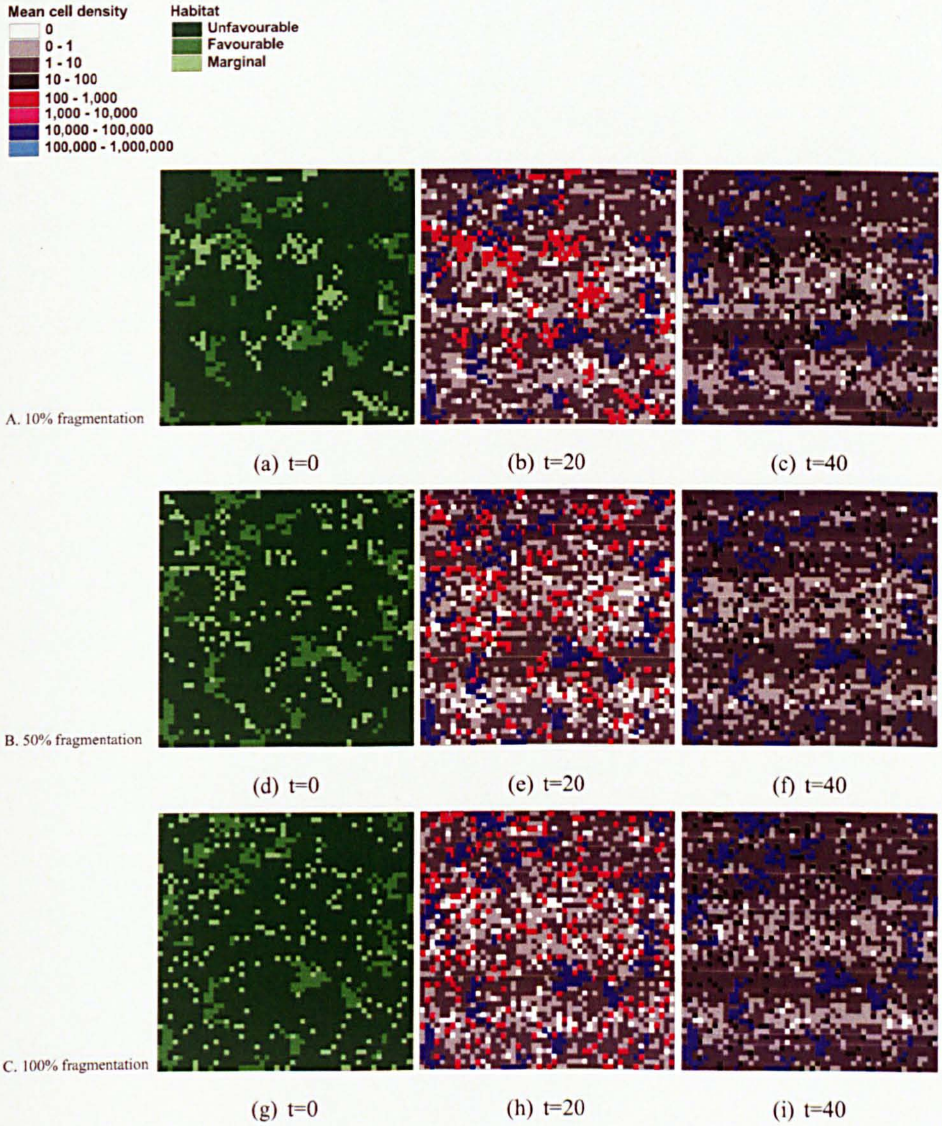


(b) Marginal



(c) Unfavourable

Figure 8.11: The total population at t= 40 in each habitat type, for different marginal habitat cover and marginal habitat fragmentation scenarios.

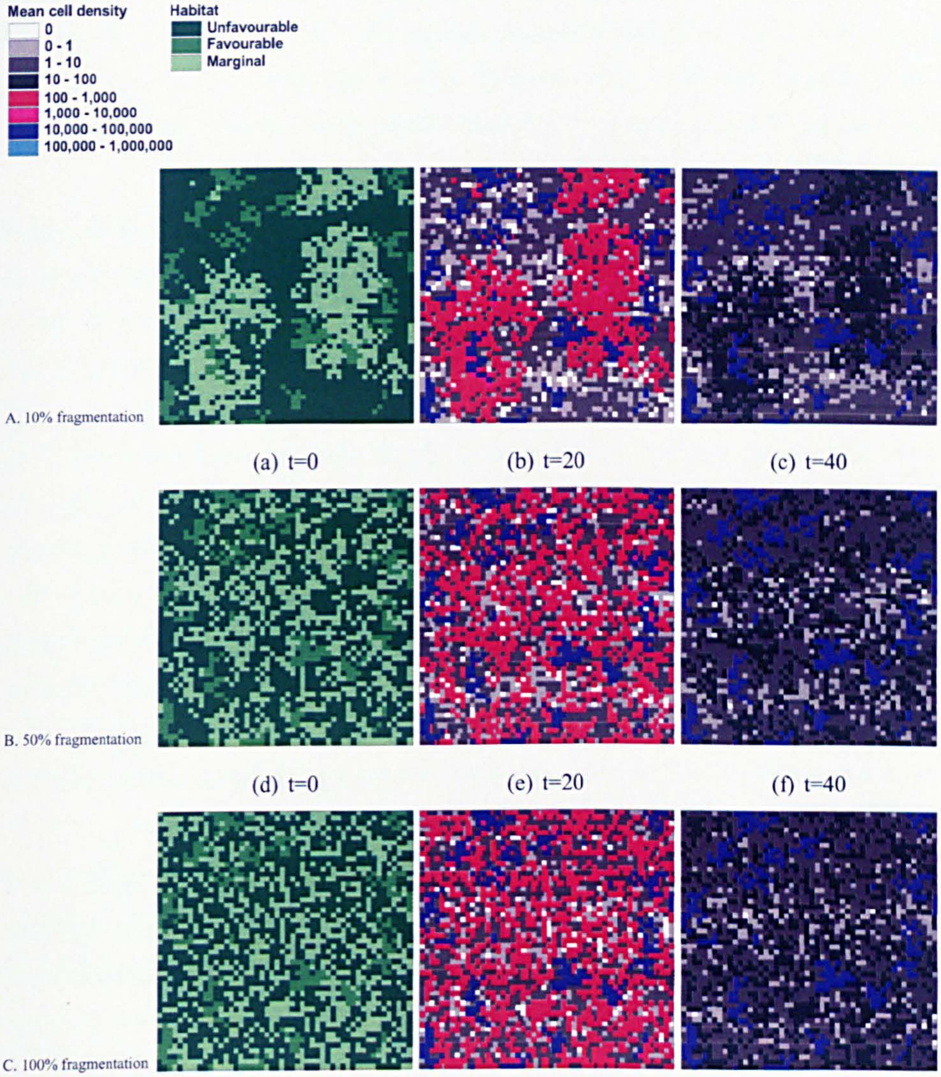


(a) t=0 (b) t=20 (c) t=40

(d) t=0 (e) t=20 (f) t=40

(g) t=0 (h) t=20 (i) t=40

Figure 8.12: Comparison of mean spatial distribution under different marginal habitat fragmentation scenarios, where marginal habitat cover = 10%.

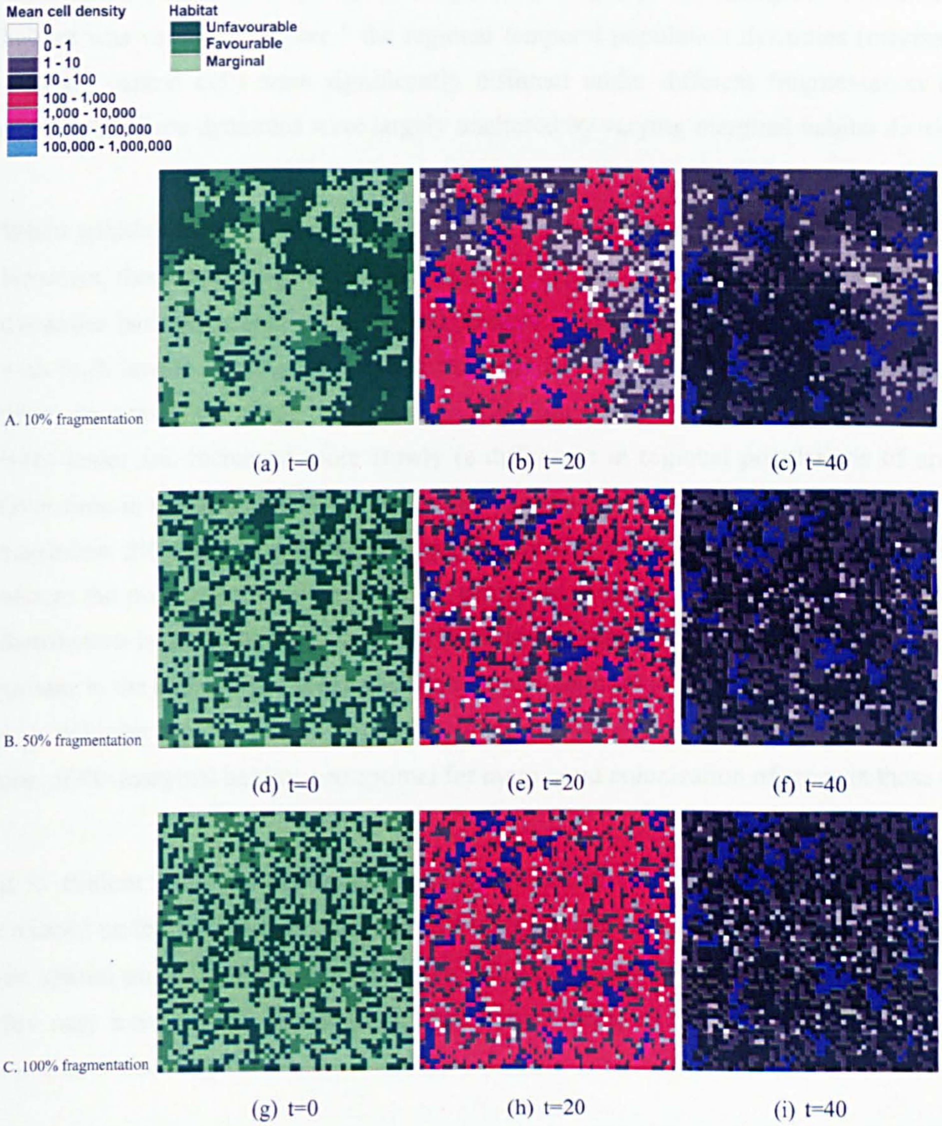


(a) t=0 (b) t=20 (c) t=40

(d) t=0 (e) t=20 (f) t=40

(g) t=0 (h) t=20 (i) t=40

Figure 8.13: Comparison of mean spatial distribution under different marginal habitat fragmentation scenarios, where marginal habitat cover = 50%.



(a) t=0 (b) t=20 (c) t=40

(d) t=0 (e) t=20 (f) t=40

(g) t=0 (h) t=20 (i) t=40

Figure 8.14: Comparison of mean spatial distribution under different marginal habitat fragmentation scenarios, where marginal habitat cover = 90%.

8.4 Discussion and summary

This chapter has explored the impacts of the matrix properties and non-crop habitat management upon the population dynamics of *R. padi*, whilst favourable (crop) habitat distribution has remained constant. Overall, it would appear that marginal habitat distribution does impact upon species population dynamics, but was not as important as the configuration and management of the crop itself (explored in the previous chapter). For example, when the favourable habitat was varied in chapter 7 the regional temporal population dynamics (originating from a single central cell) were significantly different under different fragmentation and cover scenarios. These dynamics were largely unaltered by varying marginal habitat distribution.

When aphids immigrate across the whole region, similar temporal dynamics were observed. However, there was a significant difference in regional population dynamics and within-crop dynamics between scenarios in the first 30 days of the simulation. Populations in landscapes with high levels of marginal habitat cover and fragmentation increased most rapidly in this time, whereas populations in landscapes with low marginal habitat cover and fragmentation were lower and increased more slowly (a difference in regional populations of up to 32%). Over time in the different scenarios the regional populations converged, so at the peak only a maximum 5% difference in regional population size was observed between scenarios. This was as the population eventually settled primarily in favourable habitat (which had the same distribution in all scenarios). It can be inferred that marginal habitat configuration is thus important in the early stages of crop settlement, as it appears to influence the rate at which aphids may colonise a crop. High levels of marginal habitat cover, combined with high fragmentation of the marginal habitat, are optimal for more rapid colonisation of crops in these scenarios.

It is evident from the chi-square tests that although the regional population may remain constant under the different scenarios, the marginal habitat cover and fragmentation influences the spatial population distribution of the aphids. It is possible that over a longer time period this may have further implications than evident here, as this will inevitably impact upon survival in the longer term.

Overall, these scenarios illustrate that the marginal habitat configuration within the matrix can be important to aphid population dynamics in agricultural landscapes, although it is shown to be not as important as the distribution of the crop itself. A key finding is that the scenarios show how the importance of the marginal habitat configuration to regional population dynamics may change over time, as generations of aphids move around the landscape and settle in crops. This is something that is not considered by conventional modelling approaches.

CHAPTER 9

The Timing of Pesticide Spray Application and Configuration

9.1 Introduction

The control of *Rhopalosiphum padi* and the subsequent spread of Barley Yellow Dwarf Virus (BYDV) in the UK is based on the routine application of broad spectrum insecticides. Under CAP reform, pesticides are becoming more closely targeted to minimise costs and environmental effects. By examining individual responses of insects to pesticide application the impacts of pesticide application can be better understood at the population level. The model is used to investigate ways in which pesticide behaviour can be simulated, how pesticide application can be minimised in dosage strength and through spatial and temporal concentration, and the theoretical impact this will have on insect populations at varying landscape scales. This is comparable to the use of individual-based models (IBMs) in epidemiological research to test eradication strategies for the control of infectious diseases (e.g. Eisinger et al., 2005).

The timing and spatial configuration of the spray of pesticide such as deltamethrin impacts upon both pest and beneficial insect species distributions. Important to agricultural management is the recovery rate of species. To demonstrate the ability of the model to assess the impacts of change in pesticide regime, the model is used to examine a series of hypothetical scenarios. This explores the model's potential as a tool for agricultural management planning at both the local and regional scale. Firstly, pesticide behaviour and its impacts on a regional population of *R. padi* were explored with the model, to establish realistic pesticide behaviour. Second, the impacts of altering the timing of a pesticide regime were studied. Thirdly, the influence of spatial variation in pesticide regime was explored with the model.

9.2 Method

9.2.1 Type of pesticide regime

A number of hypothetical 'pesticides' were tested on a theoretical landscape containing marginal habitat and a uniform initial aphid population distribution, calculated from suction trap data (see chapter 3). Wind was included in the simulation, but was kept in a constant westerly direction. The theoretical landscape contained 10% favourable habitat at fragmentation index 10%, with marginal habitat occupying 50% of the remaining landscape, at 50% fragmentation (see chapter 3 for more detail on the construction of this neutral landscape). These pesticides affected the aphid population in different ways, all of which were founded on the response of aphids to pesticides in the field, as studied in the literature and according to expert knowledge. The aim was to determine the variability in the population dynamics dependant upon the 'pesticide' effects on individual aphids, and also to decide which regime may represent the most realisitic population-level response. The primary duration of the pesticide 'effective period' was set at 30 days (one month), based upon the results of Longley et al. (1997), with the pyrethroid insecticide deltamethrin at the recommended field concentration (Decis 2.5% E.C.; 6.25 $gAI^{-1}ha^{-1}$ in 200 l of water).

The 'pesticides' tested were:

- A. All aphids within the field are killed on the day of pesticide application (after reproduction has taken place). No subsequent deaths.
- B. The sprayed area becomes 'unfavourable' habitat for 30 days - this alters the survival and movement of the aphids to the general rules chosen to represent unfavourable habitat, where aphid survival is restricted to a maximum of 3 days if more favourable habitat is not found (see chapter 4).
- C. The toxicity of the pesticide declines over 30 days. This was implemented simplistically, where up to 5 days after spraying mortality is set at 78%: This is based upon the effects of deltamethrin at the recommended field concentration, observed by (Longley et al., 1997). It is then assumed to decline to an estimated 50% mortality in effect until 30 days after spraying.
- D. The toxicity of the pesticide declines over 20 days. This is essentially the same as scenario C, except that the pesticide remains active for only 20 days rather than 30 (less than stated in the literature). This was done to see if a shorter effective period had a differing impact upon population dynamics.
- E. No pesticide.

The results from this study informed the choice of 'pesticide' to be used in the following study on the timing of pesticide spray.

9.2.2 Timing of pesticide spray

The main scenarios in this chapter examined the way in which the density and dispersal of the aphid population of a field selected from the landscape was influenced by four different hypothetical pesticide regimes, over autumn and winter 1985-6. These four regimes were:

1. No pesticide spray.
2. Prophylactic spray 10 days before the first immigration wave (julian day 256).
3. Spray immediately following the first immigration wave (julian day 267).
4. Spray at the second immigration wave (julian day 286).

The effect of the pesticide was to reduce the quality of the habitat to the equivalent of 'bad habitat', where aphids were assumed to die if they remained on bad habitat for 3 days (i.e. pesticide B, section 9.2.1). The pesticide was assumed to be effective for 30 days, distributed uniformly across the field but not to the surrounding areas. The wind direction was assumed to be westerly: a consistent direction was chosen to show clearly the wind-induced movement (see chapter 4, section 4.3.2). The initial immigrant aphid population was calculated from suction trap data as described in chapter 3.

9.2.3 Spatial extent and configuration

This scenario tested whether the population dynamics of an aphid population within a field are altered according to whether or not a pesticide regime was also applied to the surrounding fields. This explores the spatial implications of changes in pesticide applications, and how the spatial variability in the reduction of pesticide use may impact upon within-field pest population dynamics of adjacent crops. It also allows for an exploration of how the spatial configuration of pesticide regimes impact upon the regional pest population.

The landscape was the same as that used to test the pesticide regimes. Using a 50×50 cell theoretical landscape with 25×25m cells, the landscape contained 10% favourable habitat at fragmentation index 10%, with marginal habitat occupying 50% of the remaining landscape, at 50% fragmentation (see chapter 3 for more detail on the construction of this neutral landscape). Initial populations of aphids were established within the fields containing favourable habitat (derived from suction trap estimates, see chapter 3). Simulations were run with:

1. pesticide applied to all fields.
2. pesticide applied randomly to only 50% of fields.
3. no pesticide applied to the fields.

A small theoretical landscape had to be used, as the limits to the number of aphids that could be simulated were reached with the realistic simulation of the population originating from a single field in the real landscape (section 9.2.2). However, it was believed that this demonstration of the use of the model in this way still has significant heuristic value. Pesticide was applied as pesticide B, section 9.2.1. The wind direction was assumed to be westerly: a consistent direction was chosen to show clearly the wind-induced movement (see chapter 4, section 4.3.2).

9.3 Results

9.3.1 Type of pesticide regime

Temporal results

Figure 9.1 shows the significant temporal differences in aphid population dynamics as a result of the different pesticide behaviours. When compared to a control (no pesticide), all regimes reduced the aphid populations below the control for the duration of the simulation. However, it was only pesticide B (equivalent to unfavourable habitat), C and D (declining toxicity over time) that resulted in large, realistic reductions in aphid populations in the field. These hypothetical pesticides behaved very similarly until julian day 330, after which pesticide B and C resulted in population extinction within the field, whereas pesticide D resulted in continued low-level aphid presence. Pesticide A (kills all immediately and then no subsequent impact) was unrealistic, as populations remained at relatively high levels and followed a very similar temporal trend to the control.

The difference between pesticide B and C and pesticide D was the effective period: pesticide B and C were effective for 30 days, whereas pesticide D was effective for only 20 days. The longer effective duration of pesticide B and C was therefore the cause of an extinction of the aphid population, whereas the shorter effective duration of pesticide D explains the survival of some aphids in the field throughout the time period.

Spatial results

It is evident from the spatial results of the varying pesticide behaviours, that pesticide A was not effective in a realistic manner (figure 9.2). The spatial population dynamics of the aphids under pesticide A scenario were comparable to the simulation without pesticide, scenario E (figure 9.3). Pesticide B, C and D followed similar spatial dynamics, with aphids mostly located in marginal habitat throughout the simulation once the favourable habitat had been sprayed. It is evident from figure 9.2C and figure 9.2D that pesticides C and D appear to result in spatially varied populations in these marginal areas, where lower populations were found close to larger areas of favourable habitat. This may also be the case for pesticide B, but is

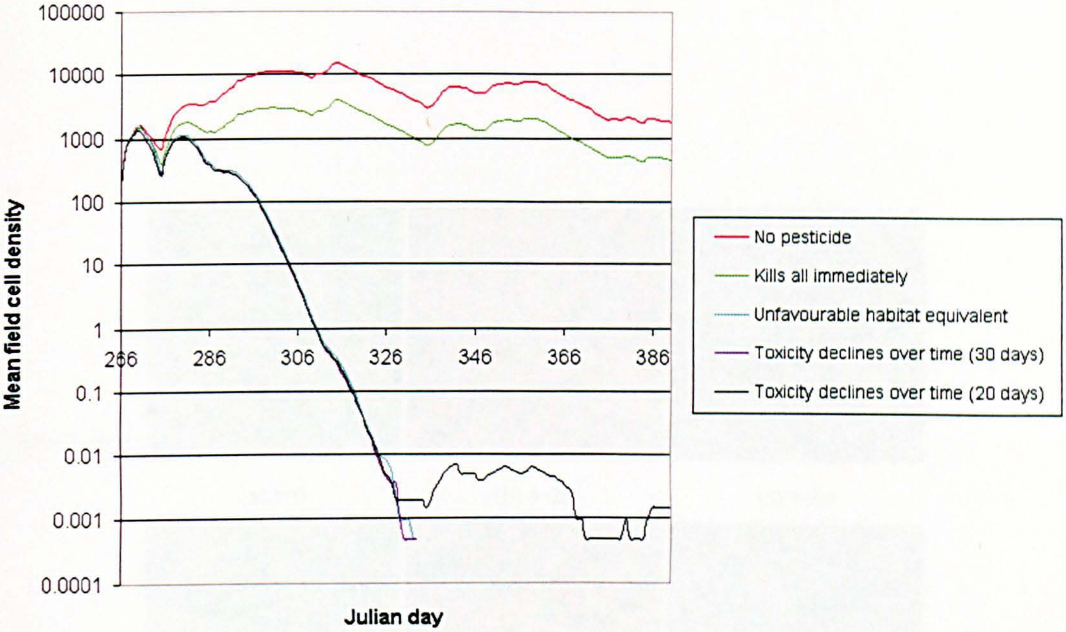
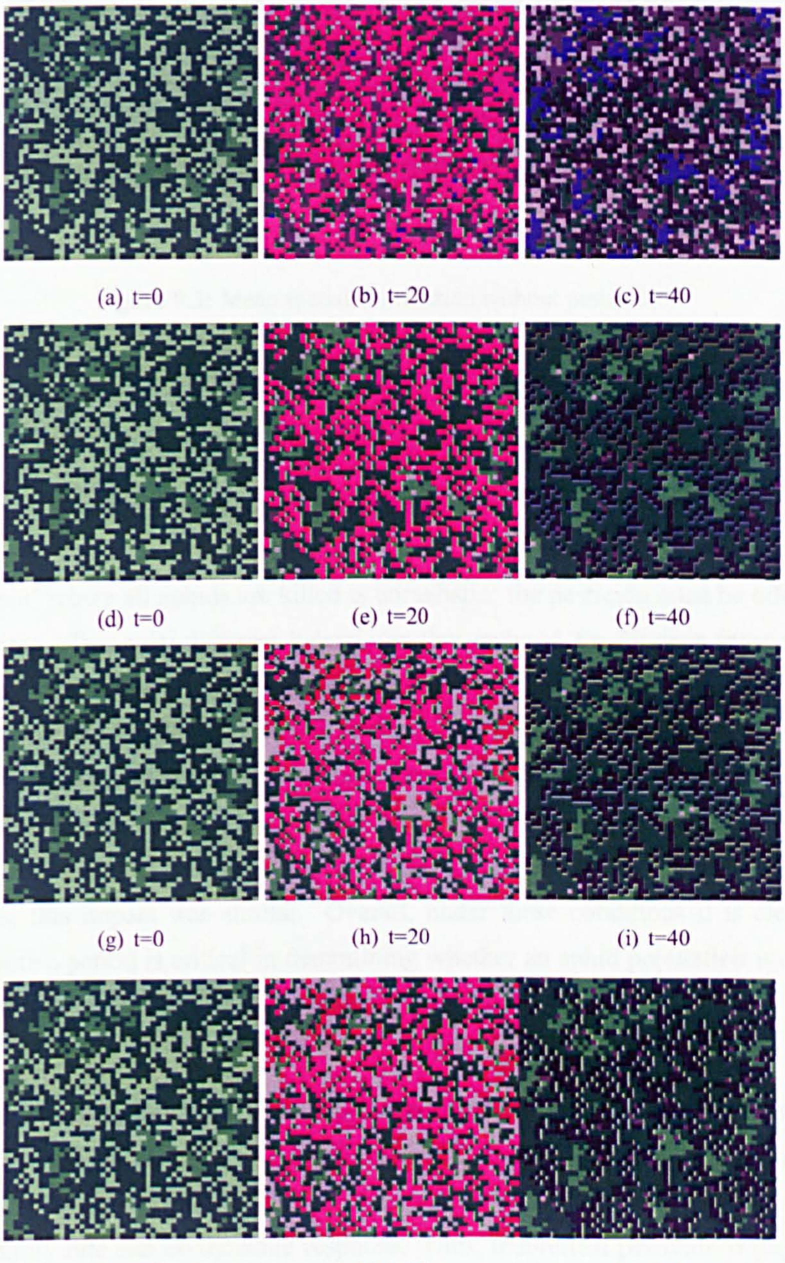


Figure 9.1: Temporal comparison of mean within-field population density per cell, under different pesticide behaviour in constant landscape configuration.

not visible due to the scale used (suggesting marginal populations were higher overall under pesticide B). This shows that within-field pesticide regimes have impacts upon field-edge populations, even when there is no spraying on these areas.

Pesticide B appears to result in just a few clusters of populations within the field at t=20 compared with a more even distribution within the field under the toxicity declining pesticides C and D, however at t=40 the within-field distribution of populations appears more similar. This was perhaps due to the more dynamic impacts of pesticide B (where aphids will be repelled by the pesticide and disperse from the field), compared to the more static 'sudden death' scenario of pesticides C and D.



A. Kills all immediately

(a) t=0 (b) t=20 (c) t=40

B. Equivalent to unfavourable habitat

(d) t=0 (e) t=20 (f) t=40

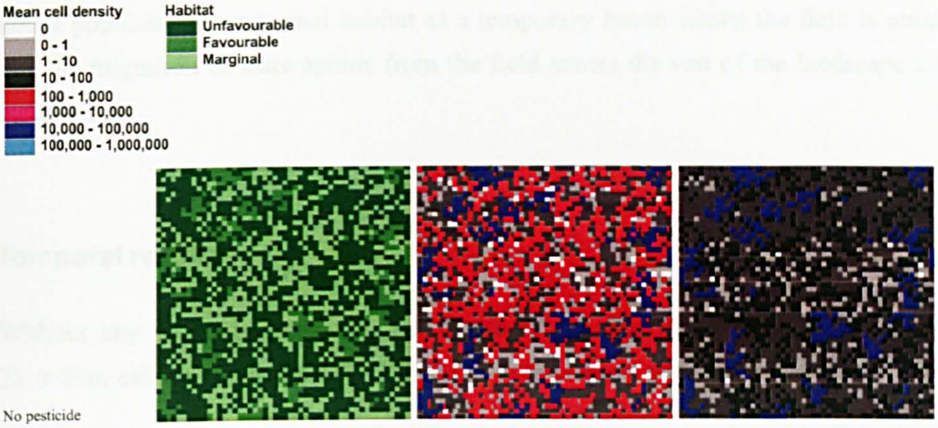
C. Toxicity declines over time (30 days)

(g) t=0 (h) t=20 (i) t=40

D. Toxicity declines over time (20 days)

(j) t=0 (k) t=20 (l) t=40

Figure 9.2: Comparison of mean spatial distribution under different pesticide behaviour in constant landscape configuration.



(a) t=0 (b) t=20 (c) t=40

Figure 9.3: Mean spatial distribution without pesticide.

Summary

The comparison of some theoretical pesticide behaviours was undertaken in order to assess the most realistic way in which to simulate pesticide. It is clear from this study that a single pesticide 'event' where all aphids are killed is unrealistic: the pesticide must be effective over a period of time. Two very different behaviours that endured for 30 days (pesticide B and C) and 20 days (pesticide D) produced very similar temporal and spatial aphid population dynamics. It was the length of effective duration that resulted in aphid populations under pesticide D remaining in the field at low levels whilst under pesticide B and C they became extinct. This similarity between the regimes gives confidence in the results and the use of either approach, as although there was clearly a large impact upon aphid populations for all three regimes, this impact was similar. Overall, under these conditions it is clear that the pesticide effective period is critical in determining whether an aphid population is completely eliminated from the field during the simulation time period.

The advantage of using pesticide B in further scenarios is that this theoretical pesticide's behaviour allows for some insects to survive initial treatment, response of aphids to residual deposits and the repellence/dispersal of aphids from the sprayed field, whereas pesticide C simply assumes a mortality rate and no dynamic response. Thus, theoretical pesticide B (equivalent to unfavourable habitat) was chosen as the basis for the following studies.

9.3.2 The timing of pesticide spray

The model was run thirty times for each simulation. The mean field population per 25 × 25m grid cell over time for each pesticide regime is shown in figure 9.4. The spatial density pattern of dispersal was observed. Both the within-field population dynamics, as well as movement of the initial field population across the landscape, are shown. This includes the movement of the initial population to marginal habitat as a temporary haven whilst the field is uninhabitable, and the migration of alate aphids from the field across the rest of the landscape as densities increased.

Temporal results

Without any pesticide treatment, the within-field population peaked at 89,353 aphids per 25 × 25m cell on julian day 315. All pesticide treatments significantly reduced this peak.

Both prophylactic pesticide treatment and treatment on julian day 267, immediately following the first wave of immigration, had a similar impact. In both cases, this reduced the within-field population peak to 1.4% of the non-treatment scenario peak. Prophylactic treatment reduced the within-field peak to 1,260 aphids per 25 × 25m cell on julian day 270. Similarly, pesticide immediately following the first wave of immigration, reduced the within-field peak to 1,256 aphids per 25 × 25m cell on the same day. Once pesticide effects expired, population density remained low in both scenarios, fluctuating around 5 aphids per 25 × 25m cell, following the same basic trend as without pesticide treatment. There was an earlier resurgence of aphid populations under the prophylactic treatment as this treatment wore off sooner; however, the trend and densities was very similar to that of spray following the first wave of immigration for most of the time period.

Pesticide treatment on julian day 286, following the second wave of immigration, allowed aphid populations to reach higher densities early in the time period, giving a within-field peak of 25,890 aphids per 25×25m cell, on julian day 288. This reduced the within-field population peak to 29% of the non-treatment scenario peak. Subsequently, aphid population levels were reduced dramatically by the pesticide, remaining very low for the rest of the time period at less than 2 aphids per 25×25m cell.

The integral of the mean density curve shown in figure 9.4 indicates the relative aphid pressure within the field across the time period as a result of each pesticide regime (figure 9.5). This indicates that both prophylactic spraying and pesticide spraying at the first wave of immigration were effective at reducing aphid pressure, where spraying at the first wave was marginally more effective, with an area of 6,171 compared to the prophylactic area of 6,333. Spraying at the second wave of immigration was much less effective, with an area of 258,695. This was a large reduction in pressure compared to the curve without pesticide spray, with an area under the curve of 4,095,274.

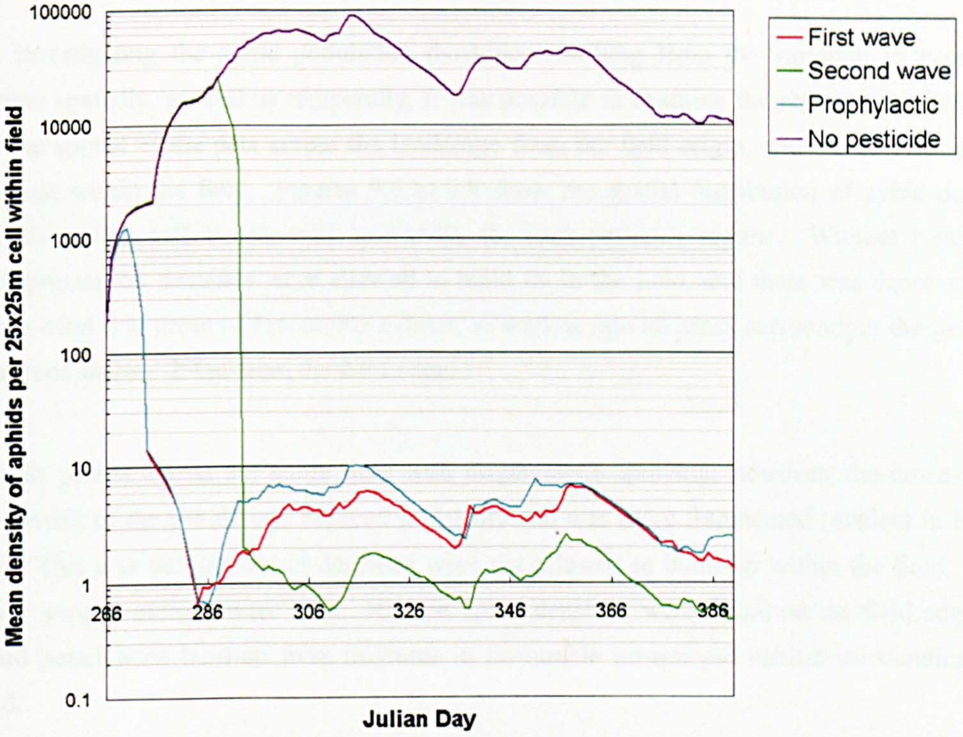


Figure 9.4: Temporal results for pesticide simulation, showing the changes in mean density per 25×25m cell within the field.

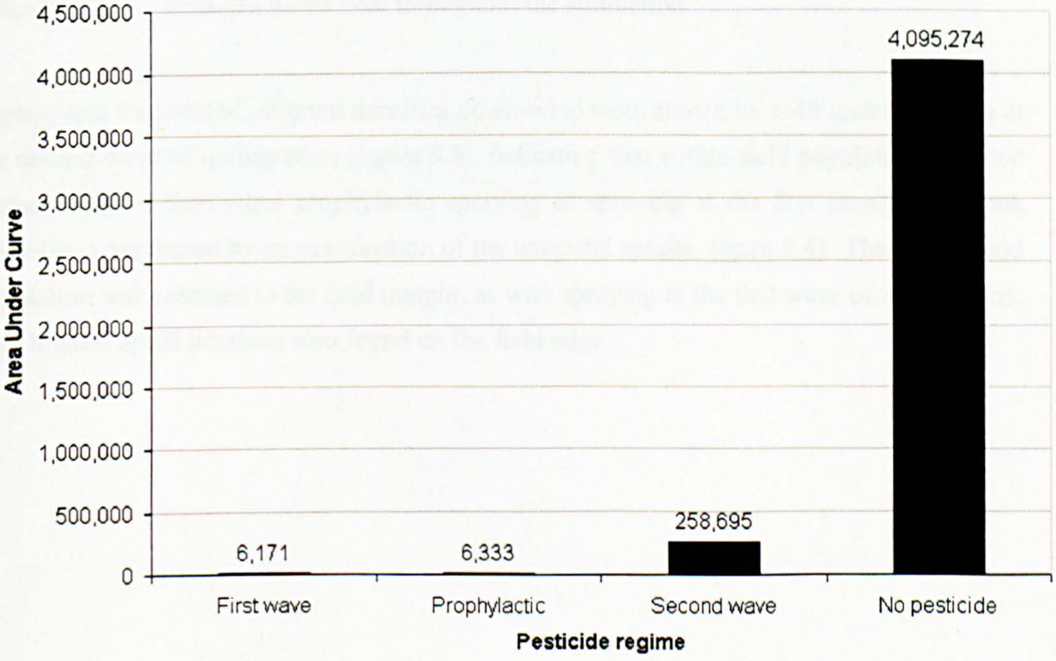


Figure 9.5: Area under curve of mean density per 25×25m cell within field, indicating aphid pressure for the time period.

Spatial results

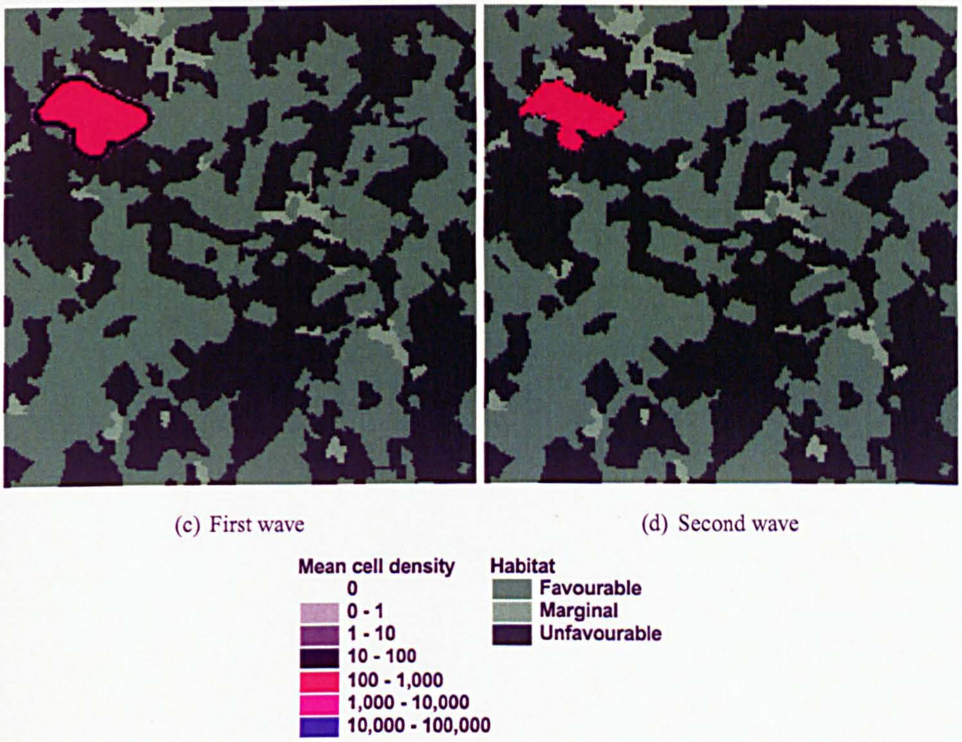
By investigating the aphid population dynamics resulting from the variation in pesticide regime spatially, as well as temporally, it was possible to examine the regional implications for the spread of the pest across the landscape from the field origin, and the distribution of the pest within the field. Figures 9.6 to 9.8 show the spatial distribution of aphid density per 25 × 25m cell at t=2, t=20 and t=40, for each pesticide regime. Without pesticide, high population densities were allowed to build up in the field, and there was dense spread down-wind into areas of favourable habitat, as well as into all areas surrounding the field up to approximately 250m from the field edge.

Aphids persist across the entire field with prophylactic spraying; however, the down-wind movement of the aphids was reduced in density and was more fragmented (evident in figure 9.8). This was because aphid densities were not allowed to build up within the field; thus, fewer winged morphs were born. Highest aphid densities were found on the field edge, as aphid populations built-up from migrants in favourable un-sprayed habitat surrounding the field.

Spray at the first wave of immigration confined the within-field population to the margins, where re-invasion was beginning to occur by t=40 from aphid populations built-up from migrants in favourable un-sprayed habitat surrounding the field. Population density of migrants down-wind was also low, and even more fragmented than with prophylactic spraying, indicating lower densities in the field throughout the simulation.

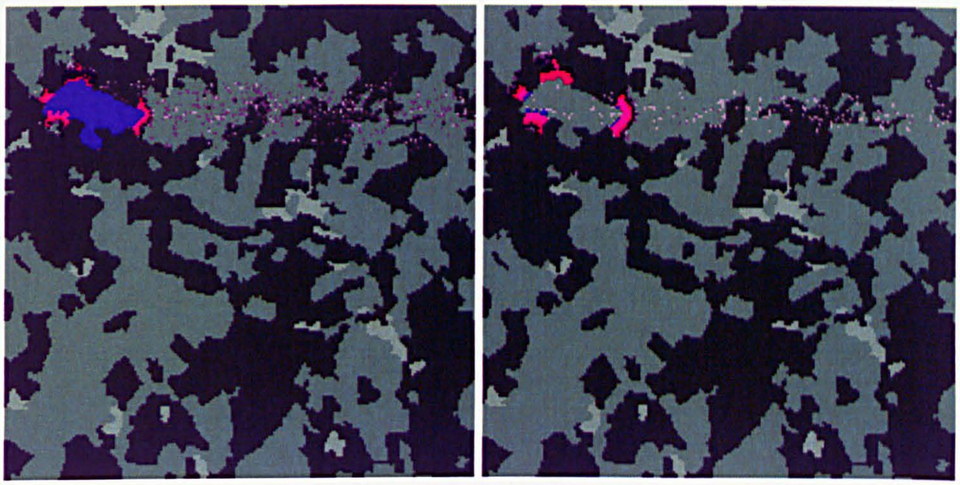
Higher, less fragmented, migrant densities down-wind were shown by t=40 under spraying at the second wave of immigration (figure 9.8), indicating that within-field populations reached higher densities than when prophylactic spraying or spraying at the first immigration took place (as is confirmed by an examination of the temporal results, figure 9.4). The within-field population was confined to the field margin, as with spraying at the first wave of immigration, with highest aphid densities also found on the field edge.

(a) No pesticide (b) Prophylactic



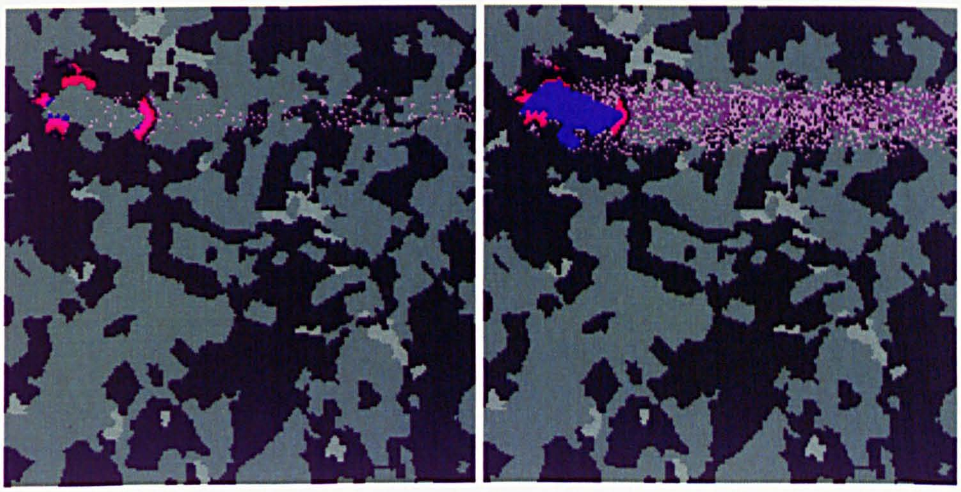
(c) First wave (d) Second wave

Figure 9.6: Comparison of spatial effects of different pesticide regimes, at model iteration 2.



(a) No pesticide

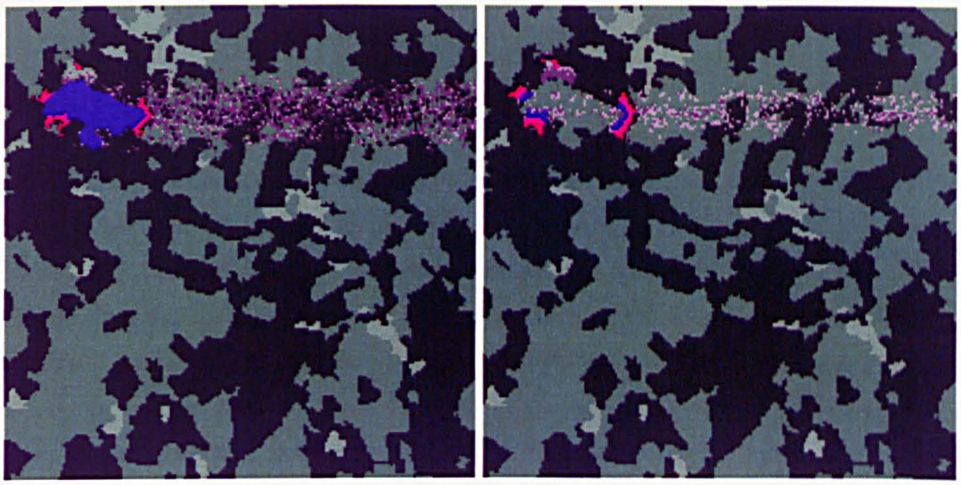
(b) Prophylactic



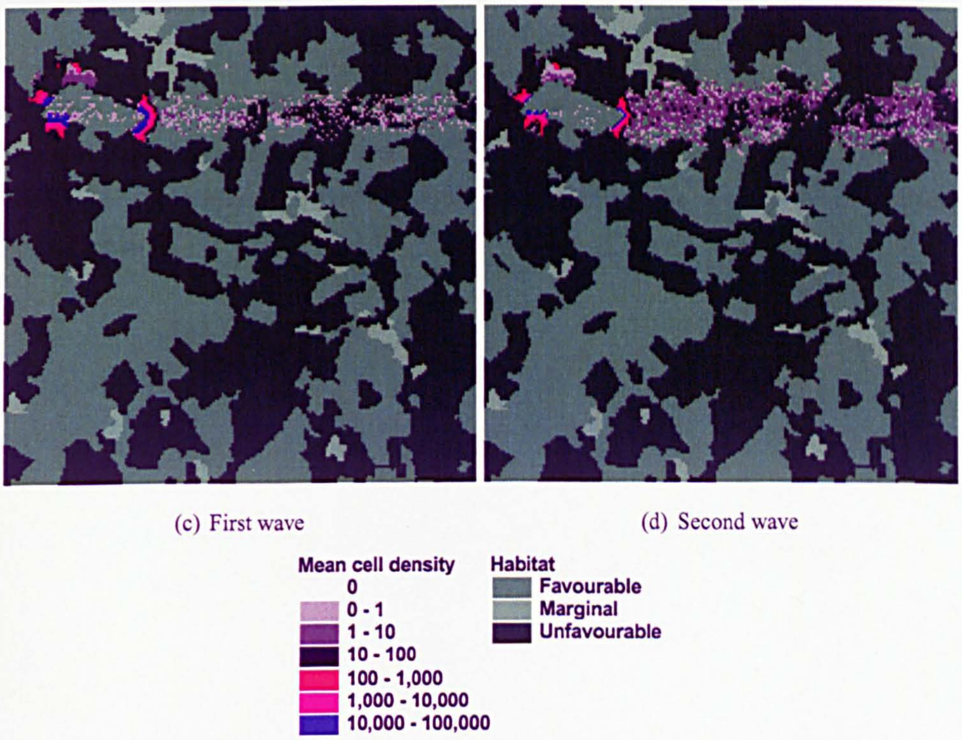
(c) First wave

(d) Second wave

Figure 9.7: Comparison of spatial effects of different pesticide regimes, at model iteration 20.



(a) No pesticide (b) Prophylactic



(c) First wave (d) Second wave

Figure 9.8: Comparison of spatial effects of different pesticide regimes, at model iteration 40.

9.3.3 Spatial extent and configuration

This final model scenario, in which pesticides were applied to all, some (50%) or none of the fields, is highly theoretical. However, it demonstrates some potential use of the model, and gives some indication of the way in which landscape configuration may influence the effects of pesticides upon aphid populations.

Temporal results

There was a large amount of variation between the three scenarios. As expected, no pesticide on fields resulted in the highest within-field populations over time, peaking at over 60,000 aphids per 25 x 25m cell (figure 9.9). When pesticide was applied to all fields, the within-field population died out completely within four days. Each of the scenarios where 50% of fields were randomly selected and sprayed produced quite different population densities within-field to one another. This was perhaps a function of the variability of the field sizes that were selected, as not only was field size believed to have some impact upon aphid population dynamics in itself (see chapter 7) but also a larger proportion of within-field area may be sprayed under some scenarios.

Overall, there was no real variation in the overall temporal trend, other than when pesticide was applied to all fields. This indicates that driving variables such as temperature were more important than the configuration of the landscape in determining the population dynamics in the model scenario.

9.3.3 Spatial extent and configuration

This final model scenario, in which pesticides were applied to all, some (50%) or none of the fields, is highly theoretical. However, it demonstrates some potential use of the model, and gives some indication of the way in which landscape configuration may influence the effects of pesticides upon aphid populations.

Temporal results

There was a large amount of variation between the three scenarios. As expected, no pesticide on fields resulted in the highest within-field populations over time, peaking at over 60,000 aphids per 25 x 25m cell (figure 9.9). When pesticide was applied to all fields, the within-field population died out completely within four days. Each of the scenarios where 50% of fields were randomly selected and sprayed produced quite different population densities within-field to one another. This was perhaps a function of the variability of the field sizes that were selected, as not only was field size believed to have some impact upon aphid population dynamics in itself (see chapter 7) but also a larger proportion of within-field area may be sprayed under some scenarios.

Overall, there was no real variation in the overall temporal trend, other than when pesticide was applied to all fields. This indicates that driving variables such as temperature were more important than the configuration of the landscape in determining the population dynamics in the model scenario.

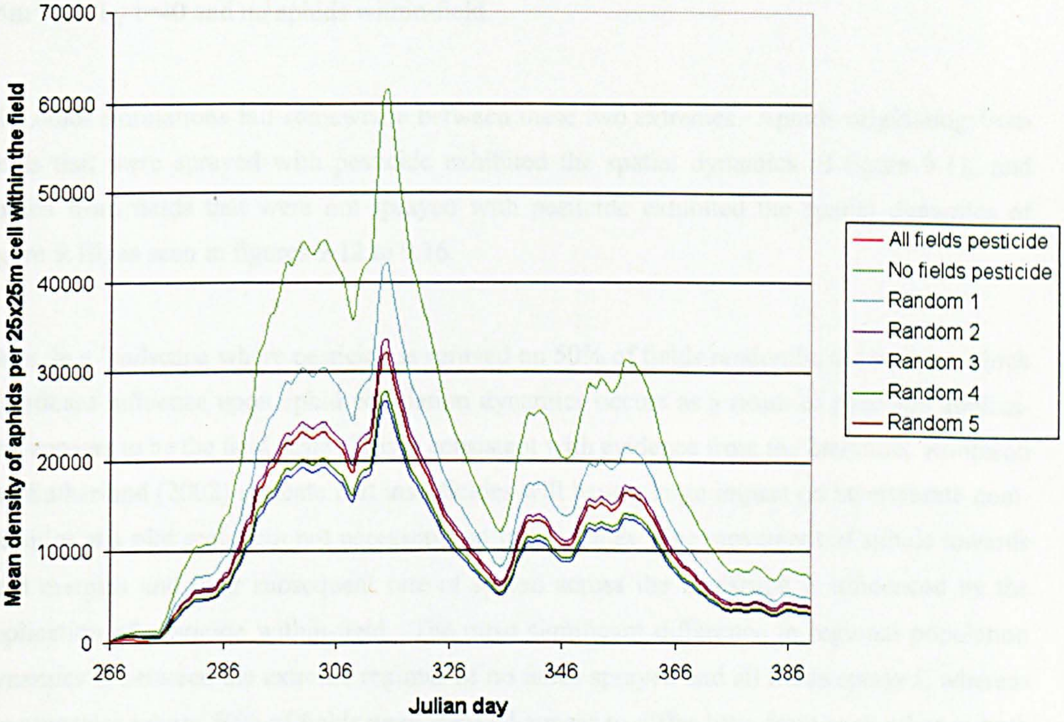


Figure 9.9: Temporal comparison of mean within-field population density per cell with pesticide spray under different landscape configuration scenarios.

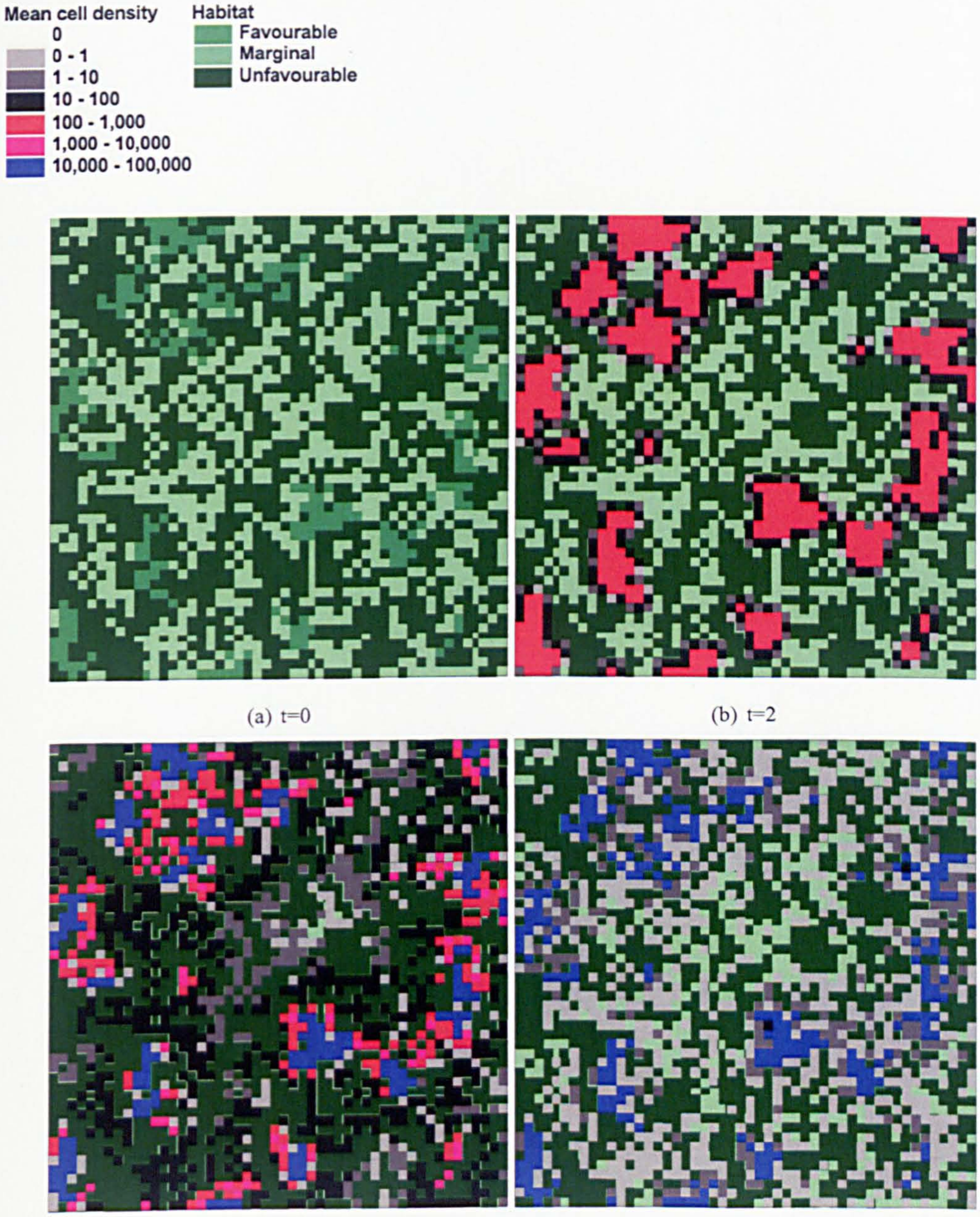
Spatial results

The spatial results are shown in figures 9.10 to 9.16. Figure 9.10 shows the spatial population dynamics across the landscape when no fields were sprayed. Here, the populations do not disperse widely from the field, but remain within-field and at the margins reaching high within-field densities as also shown by the temporal results.

At the other extreme, figure 9.11 shows the spatial dynamics when all fields were sprayed with pesticide. The key difference appears to be the initial rate at which aphids move away from the field, as they were repelled by the pesticide. Populations remain at high levels close to the field (above 100 aphids per $25m^2$ cell) at t=20, decreasing to less than 10 aphids per $25m^2$ cell by t=40 and no aphids within-field.

The other simulations fall somewhere between these two extremes. Aphids originating from fields that were sprayed with pesticide exhibited the spatial dynamics of figure 9.11, and aphids from fields that were not sprayed with pesticide exhibited the spatial dynamics of figure 9.10, as seen in figures 9.12 to 9.16.

Thus, in a landscape where pesticide is sprayed on 50% of fields randomly, the scale at which significant influence upon aphid population dynamics occurs as a result of pesticide application appears to be the field scale. This is consistent with evidence from the literature: Robinson and Sutherland (2002) indicate that insecticides will have a large impact on invertebrate communities at a plot scale, but not necessarily at larger scales. The movement of aphids towards field margins and their subsequent rate of spread across the landscape is influenced by the application of pesticide within-field. The most significant difference in regional population dynamics is between the extreme regimes of no fields sprayed and all fields sprayed, whereas the scenarios where 50% of fields were sprayed appear to differ little from each other in both spatial and temporal population dynamics at the regional scale.



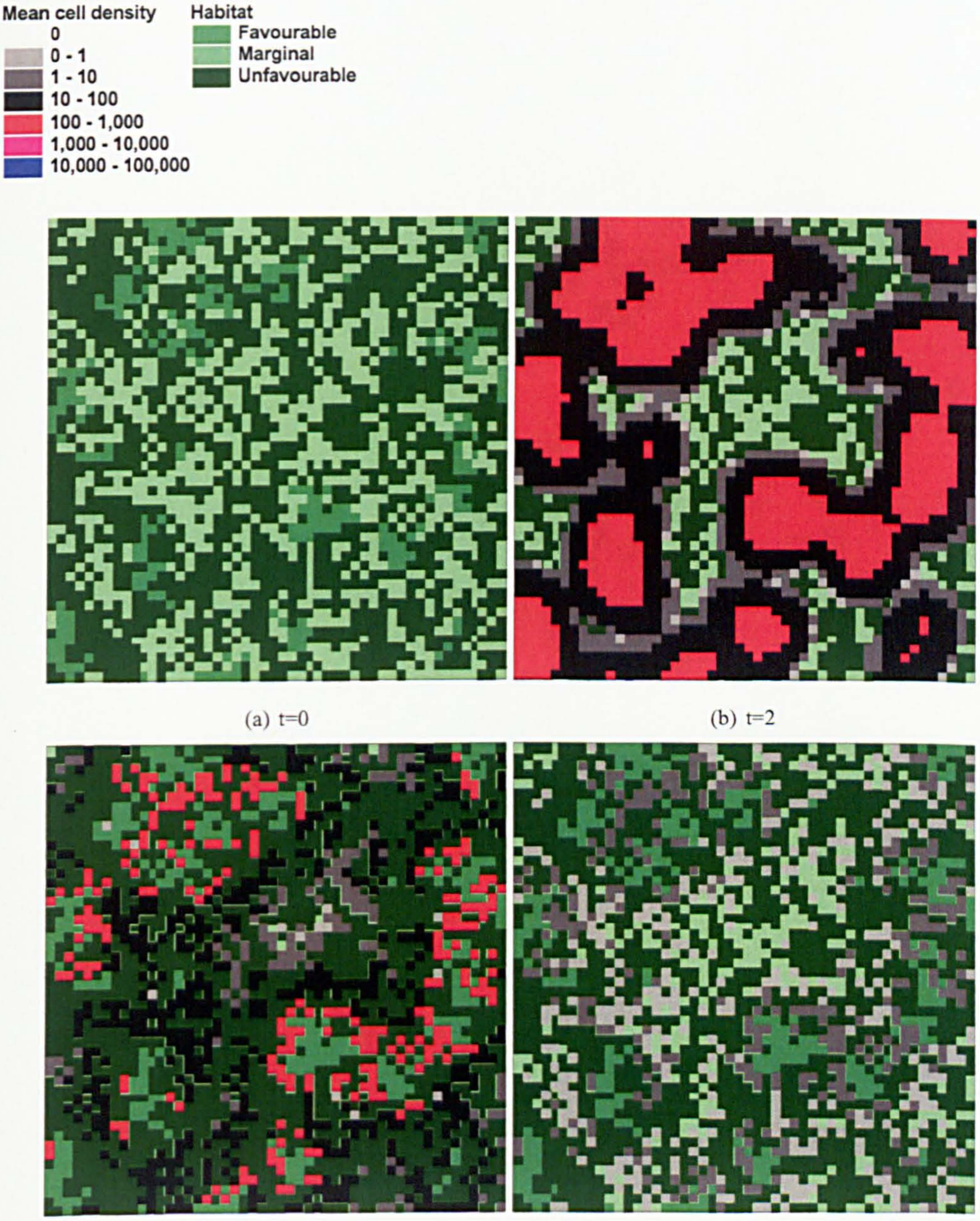
(a) t=0

(b) t=2

(c) t=20

(d) t=40

Figure 9.10: Spatial effects of pesticides spray: no fields sprayed.



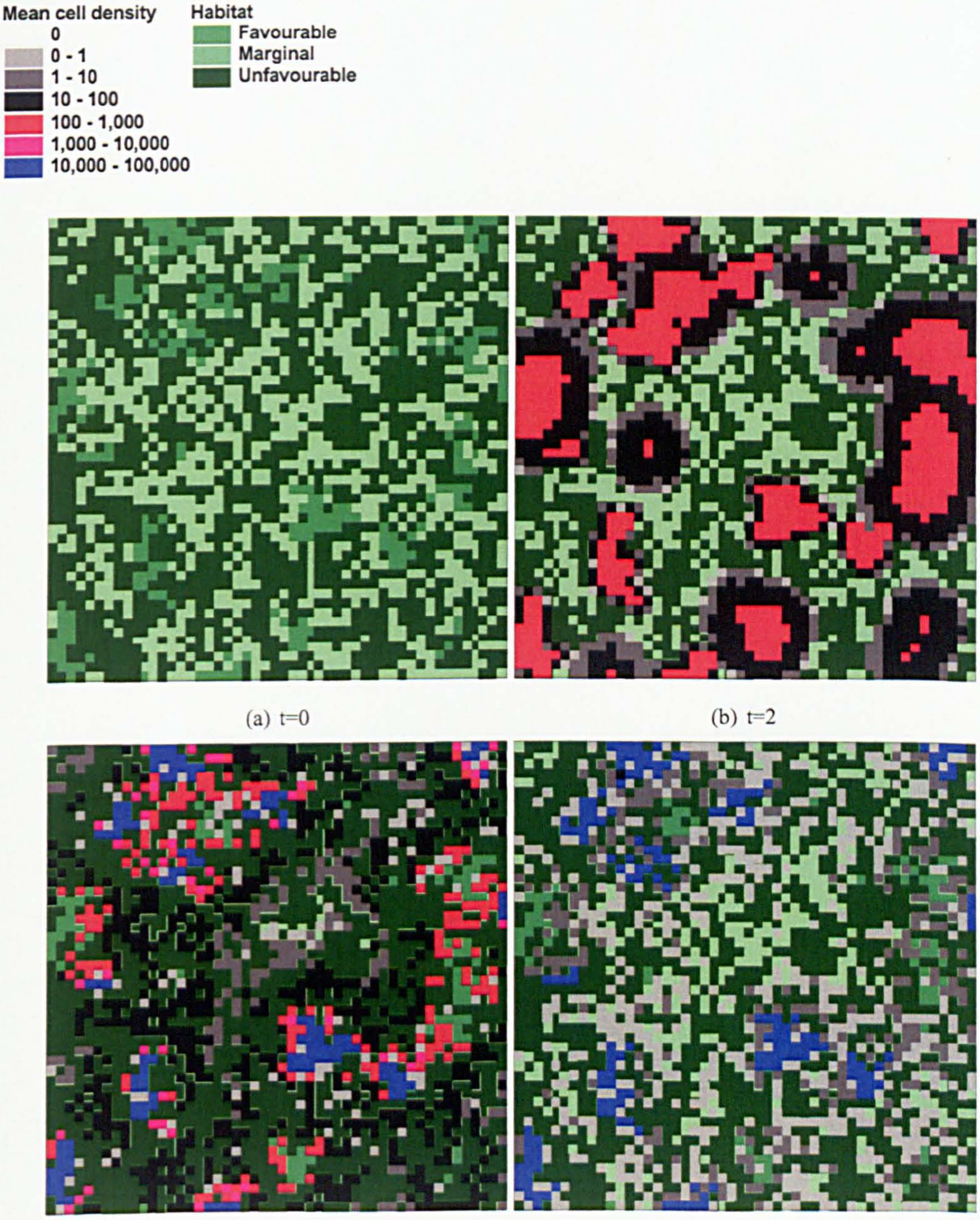
(a) t=0

(b) t=2

(c) t=20

(d) t=40

Figure 9.11: Spatial effects of pesticides spray: all fields sprayed.



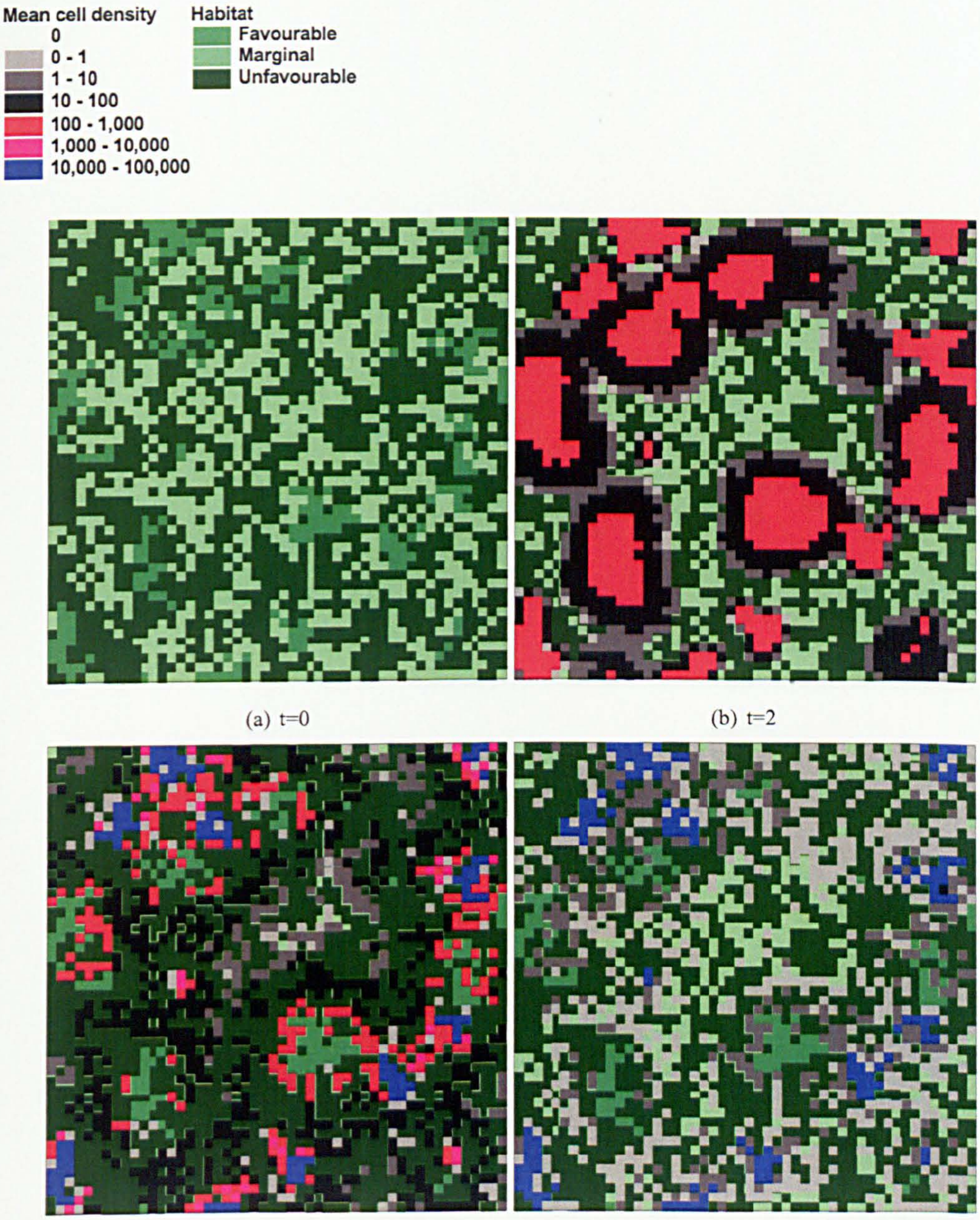
(a) t=0

(b) t=2

(c) t=20

(d) t=40

Figure 9.12: Spatial effects of pesticides spray: 50% of fields sprayed randomly (test 1).



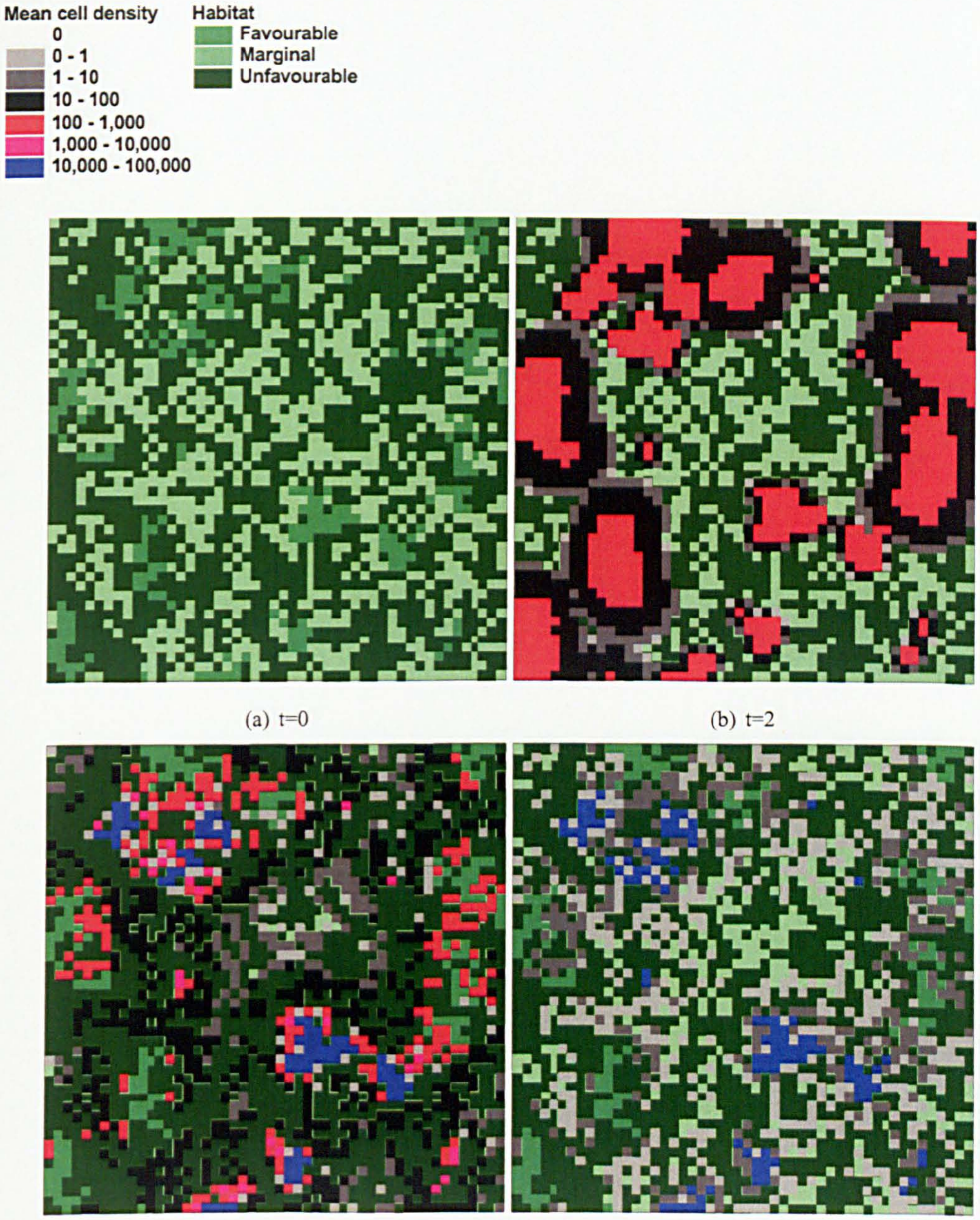
(a) t=0

(b) t=2

(c) t=20

(d) t=40

Figure 9.13: Spatial effects of pesticides spray: 50% of fields sprayed randomly (test 2).



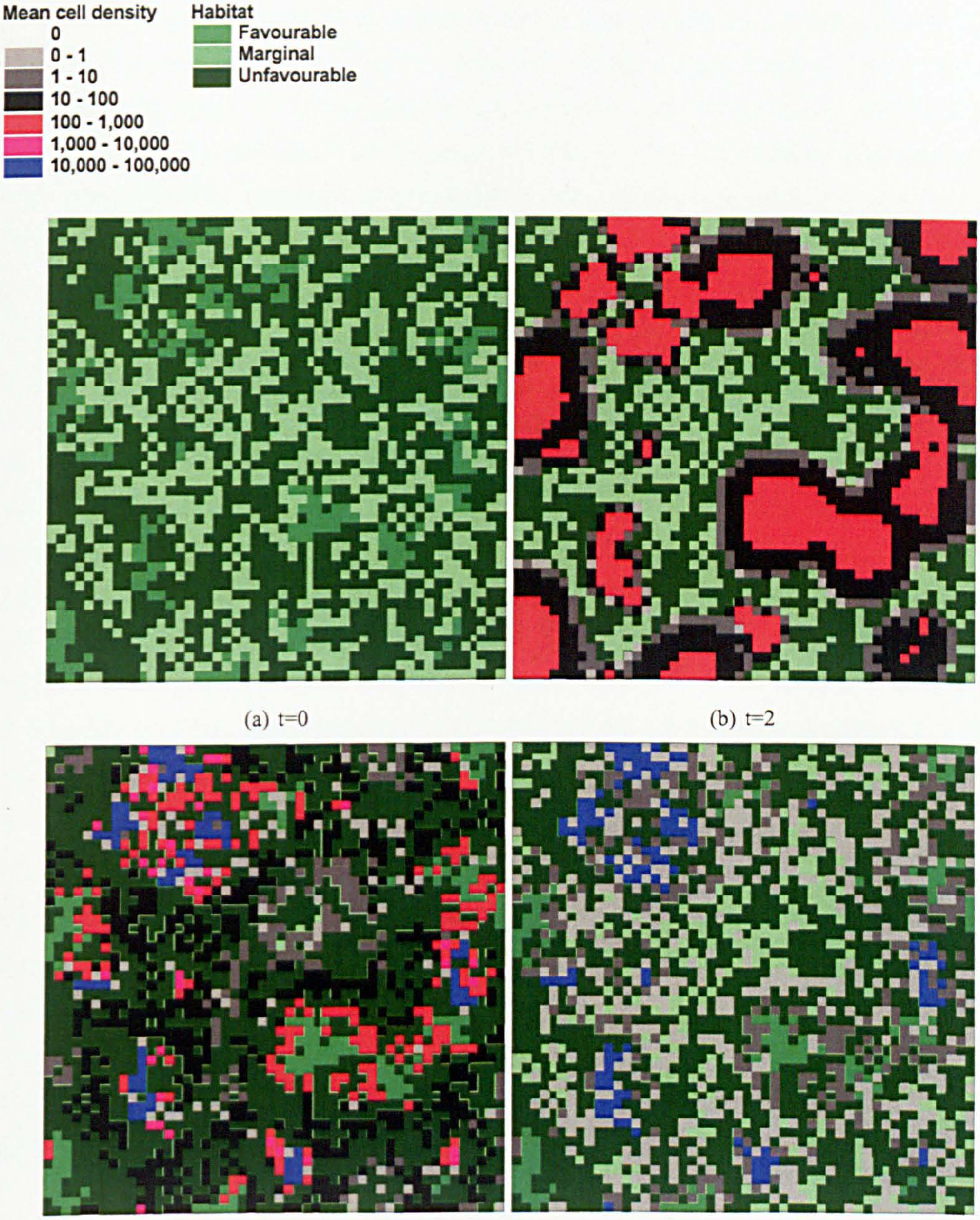
(a) t=0

(b) t=2

(c) t=20

(d) t=40

Figure 9.14: Spatial effects of pesticides spray: 50% of fields sprayed randomly (test 3).



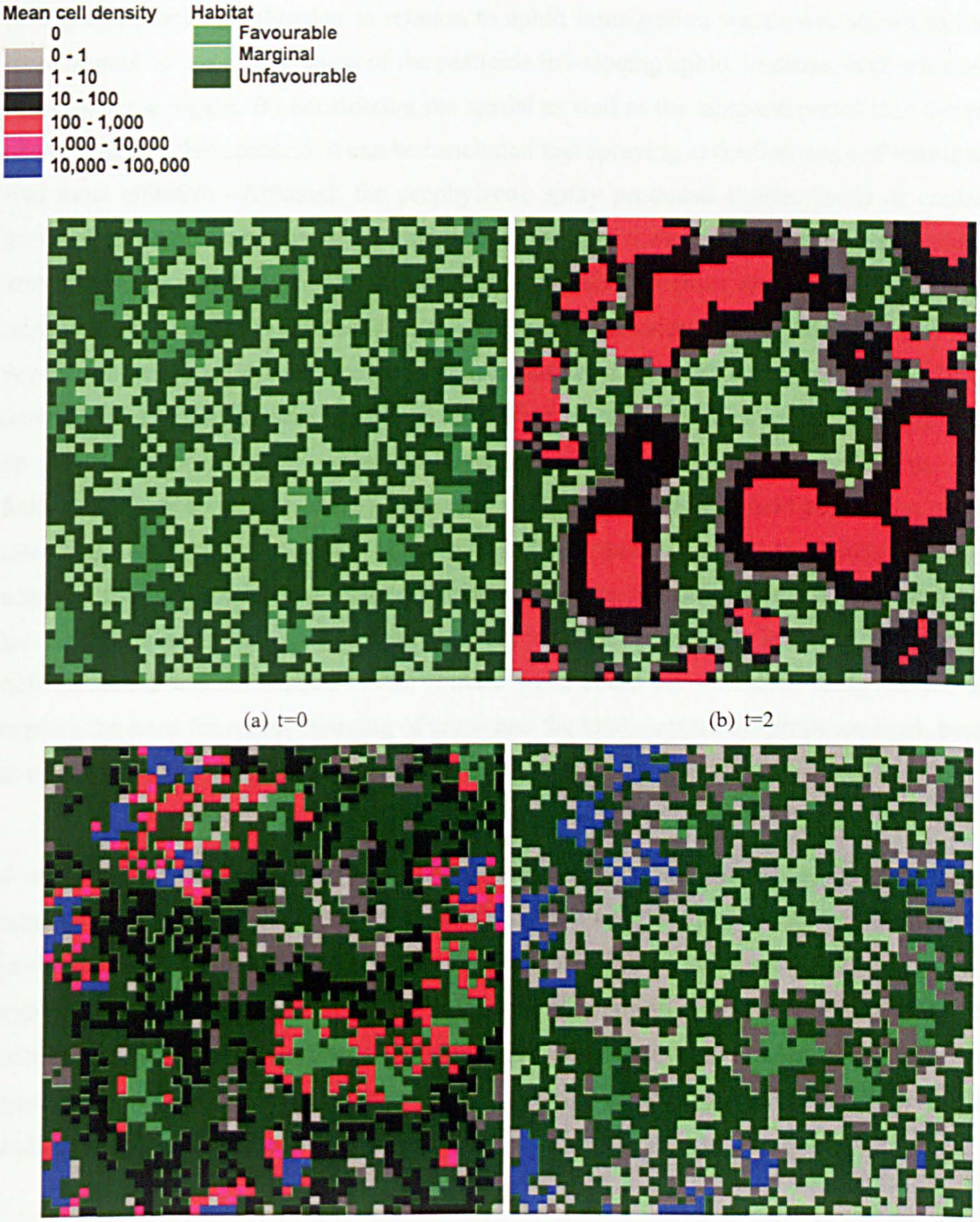
(a) t=0

(b) t=2

(c) t=20

(d) t=40

Figure 9.15: Spatial effects of pesticides spray: 50% of fields sprayed randomly (test 4).



(a) t=0 (b) t=2

(c) t=20 (d) t=40

Figure 9.16: Spatial effects of pesticides spray: 50% of fields sprayed randomly (test 5).

9.4 Summary

The use of the model to conduct an investigation into the effects of different pesticide regimes illustrates the power and potential of this approach to further understanding of the effects of landscape change and the impacts of variation of agronomic practices in time and space. The timing of pesticide application in relation to aphid immigration waves was shown to have a large impact on the effectiveness of the pesticide in reducing aphid densities, both within-field and across the region. By considering the spatial as well as the temporal population dynamics of the aphids in this scenario, it can be concluded that spraying at the first wave of immigration was most effective. Although the prophylactic spray produced similar levels of control in terms of aphid pressure across the time period, the spatial evidence from the model suggests that it allowed aphids to persist across the entire field, whereas spray at the first wave of immigration confined populations to the margins. Considering both temporal and spatial population dynamics allows for greater understanding of the relationship between within-field densities and landscape-wide distributions. When within-field densities are allowed to build up this has regional implications, as larger numbers of migrants spread from the source field and colonise the surrounding landscape more densely. Within-field population recovery rates and distribution also varies between pesticide regimes, and future simulations need to consider the important impacts the pesticide application has on predator populations, which has a subsequent impact on the aphid recovery rate both temporally and spatially within the field (Duffield and Aebischer, 1994). Future work could also consider using the model to explore the need for repeat spraying of crops and the total number of sprays required, perhaps in relation to economic factors.

A simplified landscape gives some preliminary heuristic insights into the capabilities the model may have in analysing the effects of pesticide regime upon aphid population dynamics at the landscape scale. Population dynamics at the field scale and the influence of pesticide spraying appear to be important in determining the overall distribution of aphids in the landscape both within-field and at field margins. However, a realistic investigation at a landscape scale would require significantly more computing resources, or the use of a combined parallel computing and super-individual approach, to simulate.

CHAPTER 10

Discussion

10.1 Introduction

This thesis has presented a novel application of individual-based modelling to pest management at the landscape scale, that moves beyond the paradigm of metapopulations and habitat islands, putting forward an alternative view of agri-landscapes and pest dynamics. The focus has been to use the model to explore the relationships between pest population dynamics and changes in agricultural practices. Multi-agent simulation (MAS) techniques have been applied to construct the simulation model, with the aid of Repast (a socio-economic modelling toolkit used in a novel ecological context). Advanced computing solutions to the problem of large numbers of individuals have been tested, compared and applied: super-individuals and parallel processing.

This chapter concludes the thesis, providing a summary of the research findings and what the thesis has achieved in relation to the aims and objectives identified in chapter 1. The methodology is evaluated, and recommendations for future research are provided.

10.2 Summary of the research findings

The research has aimed to develop and apply a simulation model of pest population dynamics in agri-landscapes, using detailed biological information about insect pest species and their localised behaviour to predict potential outcomes from landscape change scenarios at a larger scale. This aim is originally stated in chapter 1, as are the seven main objectives of the thesis. To tackle these objectives, the research was split into three key phases across the ten chapters.

The first phase covers the background research into agricultural landscape change, pest management, the role of simulation models and the current state of aphid modelling to date, addressed by chapter 2. This meets objectives 1 and 2 of the thesis, in particular highlighting potential areas for ecological simulations to increase understanding of insect pest population dynamics and giving an evaluation of the modelling techniques that could be employed.

The second phase is the model development. This covers several chapters, including data (chapter 3); model development (chapter 4); and model analysis and validation (chapter 5). This meets objectives 2-4 of the thesis. The capabilities of the model are enhanced by using both a mathematical and a computational solution to the problem of large numbers of individuals limited by computing power (chapter 6). This meets objective 5 of the thesis.

Objective 6, the demonstration of the application of the model to study insect pest population outbreaks, both spatially and temporally, in relation to agricultural management practices, is a demanding objective, and chapters 7, 8 and 9 are dedicated to meeting this objective. These chapters examine a number of aspects of how populations are spatially structured and how changing landscape structure alters the dynamics of pest outbreaks. This includes an assessment of the impacts of landscape configuration, pesticide regime, marginal habitat, and crop sowing and timing upon aphid population dynamics.

Below is a summary of the research findings and the primary outcomes of the case studies. This chapter will also provide a critique of the methodology and meet objective 7: to provide recommendations for future work in the form of a research agenda.

10.2.1 Literature review and identification of potential areas for research

The current state of modelling in agricultural pest management, potential areas for ecological simulations to aid research and key simulation techniques employed were presented. Work is firmly seated in the current trend of object-based simulation in many scientific fields: including MAS in socio-economic research as well as individual-based modelling (IBM) in ecology. IBM is a recently emerging field in ecology (since the late 1980s, e.g. DeAngelis and Gross, 1992; Grimm and Railsback, 2005), and has significant parallels in socio-economic science (e.g. Epstein and Axtell, 1996; Gilbert and Troitzsch, 1999; Gimblett, 2002), where MAS has become a popular methodology (originally inspired by biological systems). Technical developments in computer science and artificial intelligence research have helped advance IBM and MAS.

A thorough review of existing models of pest populations in agricultural landscapes showed that a large number of simulation models of insect pests, particularly aphids, have been developed since the late 1960s. Existing pest models can be broadly said to be aspatial, deterministic, stage-structured and largely descriptive rather than predictive. These models use methods such as box-car aphid development, population-level simulations and differential equations. Although capable of describing specific systems, these models are not transferable or predictive to new species and landscape configurations. The contribution of this thesis to the field of aphid simulation has been to provide a flexible, spatial model of pest population dynamics with increased biological realism.

10.2.2 Model development

The development of a model of spatial and temporal population dynamics of a species of cereal aphid, *Rhopalosiphum padi*, in dynamic agricultural landscapes, using IBM modelling techniques is presented. A comprehensive model of pest population dynamics in agricultural landscapes has been established, that is also transferable to other species and environments due to its object-oriented design, and is capable of the inclusion of other species at different trophic levels. The research has highlighted some problems with the use of IBM at a large scale, but also provides a thorough analysis of potential solutions to such problems. The research has proven the methodology to be useful to answer research questions at the landscape scale.

Model assessment has been undertaken at three key stages in the modelling process: at the conceptual stage, development stage and operational stage. Overall, the model is robust: it does not become chaotic when parameters are varied and the effects of changing a number of variables are nearly linear. Evidence from the sensitivity analysis confirms the model reflects real ecosystem behaviour: exponential population growth is evident when the limits to the population growth are reduced (i.e. survival rate), or the environmental conditions are made more favourable (e.g. increased temperatures). The greatest alteration to the population dynamics of *R. padi* was when temperature and survival rates were modified: the majority of parameters in the model impact upon the results across the simulation timeline; however, the effects of temperature and survival have a greater cumulative effect.

The thesis advances the capabilities of the individual-based model (IBM), and evaluates two methods that can be employed to increase the number of individuals that can be simulated: parallel computing and the concept of 'super-individuals'. It was found that parallel computing provides an effective solution, when used on a very large number of individuals across a number of processors (>5). Super-individuals do not conserve the spatial distributions of the model, which makes them unsuitable for spatially-explicit applications such as that presented here.

10.2.3 Demonstration of the application of the model

Studies of plant-insect interactions in an agricultural context have been a key focus of important ecological research in the past 20 years. In particular, they have been fundamental to the development of theories on habitat structure and extinction, such as metapopulation theory. The model has been used to explore spatial and temporal relationships between pest population dynamics and crop management scenarios.

Simulation modelling of plant-insect interactions at the landscape scale is rarely undertaken, and thus this study makes a significant contribution in understanding how landscape change may influence such interactions. IBM applied in this context is an emerging field; most studies at the landscape scale are population-level studies, often based on metapopulation concepts. The model uses IBM to understand how population-level phenomena arise from individual interactions between insects and the crop as well as between insects and other environmental factors (such as pesticide spray and temperature). In the majority of scenarios, the influence of the driving model variables such as temperature (chapter 5) in the temporal population dynamics was evident, even under quite extreme habitat manipulation and varying immigration dates.

It is quite difficult to understand the relative impact of temperature changes on insect population dynamics empirically; however, by adopting a simulation approach the complex interaction between temperature and individual processes can be more easily examined. Changes to the temperatures in the model have both a significant temporal and spatial impact. In general, as the population density was reduced by lower temperatures over time in the scenarios, the aphid population became spatially more sparse. Increased temperatures resulted in more active, widespread dispersal, possibly due to higher population densities and thus greater numbers of alate morphs. Increased mortality had a significant spatial impact, with very limited dispersal of aphids and few individuals found in bad or marginal habitat.

The model has shown that wind plays a significant role in the distribution of individuals and thus their persistence in habitat patches. For example, the importance of wind direction in determining the subsequent distribution of a population originating from a single source is evident in the pesticide case study. In order to better understand spatial population dynamics of insects that are influenced by wind, incorporation of such a simple wind model into a simulation is beneficial.

In chapter 7, the theoretical analysis of crop cover and fragmentation indicated that aphids tend to survive and move in a preferred direction (excluding the effects of wind), and are strongly influenced by the presence of favourable habitat, its proximity and configuration. Field size is also important: the highest habitat cover and the lowest fragmentation (corresponding to the largest fields) resulted in the greatest regional population density over time. In the scenarios based upon a real landscape used to explore the implications for the aphid population dynamics of the timing of crop sowing, it can be surmised that a late sow date is most effective in reducing aphid populations within the field. This is most effective when initial aphid migration is relatively early (julian day 266), and least effective if initial immigration is later.

In chapter 8, theoretical scenarios illustrate how the importance of the marginal habitat configuration to regional population dynamics may change over time, as generations of aphids move around the landscape and settle in crops. This is something that is not considered by conventional modelling approaches. The importance of marginal habitat is also evident in the crop sowing scenarios of chapter 7 and the pesticide spray scenarios of chapter 9.

In chapter 9, the timing of pesticide application in relation to aphid immigration waves was shown to have a large impact on the effectiveness of the pesticide in reducing aphid densities, both within-field and across the region. By considering the spatial as well as the temporal population dynamics of the aphids in this scenario, it can be concluded that spraying at the first wave of immigration was most effective. At present it is not possible to use the model to investigate the landscape-wide implications of pesticide regimes applied across a region; however, a small theoretical study is of heuristic value in identifying the possible scale at which pesticide regime is likely to impact upon pest populations.

10.3 Critique of the methodology

The research methods employed in this thesis bring together three diverse fields of study: computer science, ecology (pest management) and geography, to construct a spatially-explicit IBM of aphid population dynamics in agricultural landscapes. A large body of IBM simulation studies are building up in the literature, but they are somewhat disparate and varied in complexity, analytical ability and utility. Comprehensive model analysis is often lacking as well as solid justification for the use of the methodology, and there is little use of its full potential to simulate at larger spatial scales. It is hoped that the methods demonstrated in this thesis go some way to rectifying these shortcomings.

This model was not intended to be used by farmers as it is largely theoretical, but could be used by policy makers to examine broad-scale implications or theoretical farm-scale outcomes of various land management scenarios. In the exploration of the impacts of the timing of crop introduction and configuration (chapter 7) and also the influence of non-cropped habitat (chapter 8), an individual-based approach enabled:

- The simulation of complex movement across the landscape.
- A comparison of aggregation around an origin under different habitat scenarios.
- Analysis of the relative importance of marginal habitat, including outcomes that could not have been predicted, for example, the spatial dynamic implications of aphids remaining in marginal habitat until crop emerges.
- Identification of the timing and extent of regional movements and dispersal.

The individual-based approach has given the flexibility to manipulate both landscape (e.g. sowing) and aphid dynamics (e.g. timing of immigration) temporally and spatially. The time steps and spatial resolution given here could be easily altered for other simulations; thus, many other scenarios at differing spatial and temporal scales could be established and examined and many further questions explored. Some caution should be exercised in the application of this particular model at small grid sizes (<5m) in small regions, however, as underestimation of aphid numbers occurs as the mean number of aphids per m^2 drops below 1 (see chapter 5, section 5.5). Difficulty was experienced in developing rules for the behaviour of aphids in unfavourable habitat, as there is an absence of research into an aphid's ability to survive poor environmental conditions; thus, an expert estimate of survival time was used.

The key limitation of the methodology as applied in a pest management context has been the inability of the model to cope with realistic aphid densities across a region, which may reach several billion. This was addressed by an exploration of techniques to cope with the simulation of large numbers of individuals (chapter 6). In this chapter it was found that parallel computing provides the most effective solution to deal with large numbers of individuals in a spatial context. However, there are a number of considerations to make when implementing this method. Firstly, a decision whether to split the agents or the environment between processors can relate to the complexity of the environment, the mobility of the agents and the number of interactions between the agents. For this model the aphid agents were distributed across the nodes, as they are complex, highly mobile agents in a relatively simple environment. Thus, information on the whole environment was stored on all nodes. However, if there are complex, spatially-defined interactions between agents, this may be more problematic, as agents may be interacting with other agents that are spatially local in the context of the whole simulation but residing on different processors. Therefore conversely, if agents are not mobile but have complex local interactions, and/or the agents reside in a complex environment, it would be best to split the environment between nodes.

Overall, in parallel computing using MPI any message passing must be minimised, particularly passing of agents, as this can really slow the simulation down and can also occupy significant memory. It was found that a parallel solution is not always more efficient than a non-parallel solution (as was found for lower numbers of nodes, chapter 6).

10.4 Recommendations for future research

The model explores the use of IBM in landscape-scale analyses of aphid population dynamics, in contrast to existing techniques. It signals the arrival of emerging alternative concepts of landscape configuration (contrasting with an often binary, metapopulation view), and a focus on ecological understanding at larger spatial and temporal scales. It is a flexible model, and can be adapted to different scenarios and species easily, can form a basis for an analysis of the relationship of population peaks to the spread of BYDV, as well as putting forward some methodological suggestions for further developments in the field. IBM is moving towards a position in ecology where it can begin to challenge the current theoretical constructs of populations in the next few years.

A large amount of information may be extracted from the simulations, too much to present here. This includes population demographics and details of an individual's behaviour, for example movement trajectories and stimuli. More complex landscapes with varying crop types and matrix properties could be constructed upon which to run the model. However, the model is limited by the extent to which it can be validated at present, and more complex scenarios would only be suitable to construct if it is possible to more rigorously parameterise and validate the model (perhaps including some landscape-scale empirical validation). This is also a limitation when considering the combined effects of landscape change (such as the joint effects of the timing of crop sowing and pesticide spraying). Future research should examine ways in which the impacts on population dynamics of multiple landscape changes may be explored concurrently.

The model could be easily transferred to other, similar, species of aphid, such as *Sitobion avenae*, with a minimum of model restructuring. The first steps to re-parameterising the model are given in appendix D, where equations have been derived from literature sources and experimental data to establish equations to represent development and reproduction in this species (see also appendix A.4). Many of the rules used in the model (for example movement) would also apply to this species (see chapter 3). This would enable comparisons between the two species of aphid, and with further development could be used to investigate why *R. padi* is comparatively rare in the field in the UK, despite large numbers found in suction traps and a prolific distribution in similar environments such as Finland, a question that has puzzled entomologists for some time (Leather et al., 1989).

An example of how complexity may be added to the simulation is to incorporate the factors that induce an aphid to land. The factors that influence aphids to take-off and to undertake particular forms of flight are included in the model, but the factors that induce an aphid to land are not. This further development could explore in more detail the flight and migration of the aphids, and how this influences the local and regional population dynamics of the species *R. padi.*

When an aphid is in a migratory flight phase, research has shown that it will generally not respond to host plants and settle (Klingauf, 1987). A number of papers show that it is the aphid's response to wavelengths of green light (approximately 500-600nm) that induce it to land after migratory maiden flight (Kennedy et al., 1961; Nottingham et al., 1991). This response is highly variable between morphs and is perhaps not so simplistic as a simple change in response to visual stimuli (Kennedy et al., 1961). Wiktelius (1981) found that summer migrants of *R. padi* may respond and land after only 1 minute of flight, although Nottingham et al. (1991) argue that a slightly longer period of up to half an hour is more likely. This is contrasted with the ability of Autumn migrant gynoparae to undertake more prolonged flight: *R. padi* were found to fly for a longer period before responding to green light (110 minutes average) (Nottingham et al., 1991). During non-migratory 'trivial' flight, aphids consistently use visual stimuli to navigate between plants, keeping flight time to a minimum.

Migrant *R. padi* and *S. avenae* have been observed to occur mainly around the crop edge on all sides of the field (contrasting with *M. dirhodum* that has been found to have a more even distribution). This is related to the windbreak effect of hedges encouraging landing at the field margin (Bowden and Dean, 1977). This effect is complex and thus debated in the literature, and so is not included in the simulation model at present, but may be included with a more realistic wind model in future research.

In agriculture there is increasing interest in the options for the biocontrol of pests. Attempts to increase understanding of tri-trophic parasitoid-prey-crop systems, at all spatial scales both through modelling and also field studies would be a useful further development of the work in this thesis. It is probable that modelling will continue to advance and become more central to decision making, as landscapes are likely to undergo significant change: environmentally induced (e.g. Climate change) as well as man-induced (e.g. CAP reform). The potential of IBM to increase understanding of the possible implications of climate change for insect dispersal was illustrated in a recent paper by Walters et al. (2006). The model presented here could be developed further to similarly explore the possible impacts of climate change upon pest population dynamics. Future models will incorporate multi-trophic levels, and research is needed to realise the most beneficial application of an IBM versus a population-level approach, as well as development of methodology to compare and transfer between these approaches to find the most suitable simulation tools for policy advising.

The model constructed in this thesis is a research tool to aid understanding of insect response to landscape change and how individual insect behaviour may lead to population-level dynamics. It is also a broad-scale management tool for testing pest management strategies (e.g. for policy makers). This enables understanding of how changing landscapes may impact upon species population dynamics and provides a theoretical testing ground to explore the implications of landscape configurations and management regimes. Refinement of the model could lead to its use in pest management at a more local field-scale, applied level. This would require significant testing, further validation and the use of field records rather than suction trap data to calculate immigration rates.

10.5 Conclusion

A model has been constructed that can help to answer important questions in pest management relating to: the spatial structure of populations; how landscape structure alters the dynamics of pest outbreaks; and how relationships between pests and the landscape alter with spatial scale. The model scenarios illustrate the use of the model to provide some theoretical understanding of the spatial and temporal dynamics of aphid populations in changing agricultural landscapes, and give heuristic insights into the relationships between the landscape and insect behaviour.

There is increasing awareness of the importance of an IBM methodology in the ecological community, and the application of the methodology is widening (for example it is applied here at a large spatial scale). The method is gradually gaining increasing acceptance in population ecology and is becoming a more formal discipline, in particular in the context of concerns over global warming and agricultural changes where more flexible, predictive approaches are required. Better frameworks now exist for model analysis and development; thus, models may be developed that can be easily transferred to other species/regions, and that may incorporate easily multiple species at different trophic levels. This research attempts to construct such a flexible model.

There is clearly a high degree of uncertainty about what arable landscape changes could actually take place in England, and importantly what effects they are likely to have on species distributions (section 2.2). The model presented in this thesis has been used to explore the potential impacts of some of the suggested scenarios and likely land management trends. However, no general rules can effectively relate species diversity to landscape features, as this relationship depends upon the organism and the region of study (Jeanneret et al., 2003). It is important that future research should not just use the model to develop theoretical landscape scenarios, but also try to analyse how likely changes may combine in a real region to affect the population dynamics of a species.

IBM is yet to rival metapopulation theory as a well-founded, powerful construct in population ecology. However, recent publications such as Grimm and Railsback (2005) are providing foundations for a more coherent and effective approach to the science of individual-based modelling in ecology. The model in this thesis highlights some of the problems (in particular application at a large scale) as well as some of the power of the methodology. The power of the method includes flexibility, a reflection of real ecosystem structure, multi-scale interactions, ease of parameterisation and the possibility to examine individual life histories. This thesis contributes one of the first comprehensive IBM simulation models of pest population dynamics in a spatial context. The model constructed for this thesis has also pushed the limits of IBM in terms of scale, and provided some solutions to these limits. The model has proven that the methodology can be useful to understand landscape-scale population dynamics and generate scenarios, and the approach used is flexible enough to transfer to other species and other systems.

References

Abbott, C. A., M. W. Berry, E. J. Comiskey, L. J. Gross, and H.-K. Luh (1997). Parallel individual-based modeling of everglades deer ecology. *IEEE Computational Science and Engineering 4*(4), 60–78.

Abbott, C. A., M. W. Berry, and J. C. Dempsey (1995, August). Computational models of white-tailed deer in the florida everglades. Technical Report CS-95-296, UTK-CS.

Araya, J. E. and A. Fereres (1991). Cereal aphid survival under flooding conditions. *Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz 98*, 168–173.

Ashby, W. R. (1956). *An Introduction to Cybernetics*. Chapman and Hall, London.

Asher, J., M. Warren, R. Fox, P. Harding, G. Jeffcoate, and S. Jeffcoate (2001). *The Millenium Atlas of Butterflies in Britain and Ireland.* University Press, Oxford.

Bailey, S. M., M. E. Irwin, G. E. Kampmeier, C. E. Eastman, and A. D. Hewings (1995). Physical and biological perturbations: their effect on the movement of apterous *Rhopalosiphum padi* (Homoptera: aphididae) and localized spread of Barley Yellow Dwarf Virus. *Environmental Entomology 24*(1), 24–33.

Bar-Yam, Y. (1999). Complexity rising: From human beings to human civilization, a complexity profile. Electronic publication: `http://www.necsi.org/civilization.html`.

Barlow, N. D. and A. F. G. Dixon (1980). *Simulation of lime aphid population dynamics.* Simulation Monographs. Pudoc, Wageningen. 165 pp.

Barrett, G. W. (2000). The impact of corridors on arthropod populations within simulated agrolandscapes. In B. Ekbom, M. E. Irwin, and Y. Robert (Eds.), *Interchanges of Insects between Agricultural and Surrounding Landscapes*, pp. 71–84. Kluwer Academic Publishers, Dordrecht.

Beecham, J. A. and K. D. Farnsworth (1998). Animal foraging from an individual perspective: an object orientated model. *Ecological Modelling 113*, 141–156.

Beecham, J. A., S. P. Oom, and C. P. D. Birch (2002). HOOFS: A multi-scale, agent-based simulation framework for studying the impact of grazing animals on the environment. In A. Rizzoli and A. Jakeman (Eds.), *Integrated Assessment and Decision Support, Proceedings of the First Biennial Meeting of the International Environmental Modelling and Software Society*, Volume 2, pp. 220–225.

Beissinger, S. R. and M. I. Westphal (1998). On the use of demographic models of population viability in endangered species management. *Journal of Wildlife Management 62*, 821–841.

Benenson, I., I. Omer, and E. Hatna (2001). Entity-based modelling of urban residential dynamics: the case of Yaffo, Tel Aviv. *Environment and Planning B 29*, 491–512.

Benton, T. G., J. A. Vickery, and J. D. Wilson (2003). Farmland biodiversity: is habitat heterogeneity the key? *Trends in Ecology and Evolution 18*(4), 182–188.

Beven, K. J. (1996). Equifinality and uncertainty in geomorphological modelling. In B. L. Rhoads and C. E. Thorn (Eds.), *The Scientific Nature of Geomorphology*, pp. 289–214. Wiley, Chichester.

Bommarco, R. and W. F. Fagan (2002). Influence of crop edges on movement of generalist predators: a diffusion approach. *Agricultural and Forest Entomology 4*, 21–30.

Booch, G. (1994). *Object-Oriented Analysis and Design with Applications*. Addison-Wesley, Reading, MA.

Bousquet, F. and C. Le Page (2004). Multi-agent simulations and ecosystem management: a review. *Ecological Modelling 176*, 313–332.

Bowden, J. and G. J. W. Dean (1977). The distribution of flying insects in and near a tall hedgerow. *Journal of Applied Ecology 14*, 343–354.

Broadbent, L. (1948). Aphis migration and the efficiency of the trapping method. *Annals of Applied Biology 35*(3), 379–394.

Brown, H. I. (1996). The methodological roles of theory in science. In B. L. Rhoads and C. E. Thorn (Eds.), *The Scientific Nature of Geomorphology*, pp. 3–20. Wiley, Chichester.

Burel, F., J. Baudry, Y. Delettre, S. Petit, and N. Morvan (2000). Relating insect movements to farming systems in dynamic landscapes. In B. Ekbom, M. E. Irwin, and Y. Robert (Eds.), *Interchanges of Insects between Agricultural and Surrounding Landscapes*, pp. 5–32. Kluwer Academic Publishers, Dordrecht.

Büssenschütt, M. and C. Pahl-Wostl (2000). A discrete, allometric approach to the modeling of ecosystem dynamics. *Ecological Modelling 126*, 33–48.

Cabeza, M., M. B. Araujo, R. J. Wilson, C. D. Thomas, M. J. R. Cowley, and A. Moilanen (2004). Combining probabilities of occurrence with spatial reserve design. *Journal of Applied Ecology 41*(2), 252–262.

Campbell, C. A. M. and M. S. Ridout (2001). Effects of plant spacing an interplanting with oilseed rape on colonisation of dwarf hops by the damson-hop aphid, *Phorodon humuli*. *Entomologia Experimentalis et Applicata 99*, 211–216.

Carter, N., A. F. G. Dixon, and R. Rabbinge (1982). *Cereal aphid populations: biology, simulation and prediction*. Centre for Agricultural Publishing and Documentation, Wageningen, Netherlands. 91 pp.

Carter, N. and R. Rabbinge (1980). Simulation models of the population development of *Sitobion avenae*. *IOBC/WPRS Bulletin 3*, 93–98.

Chalmers, A. and J. Tidmus (1996). *Practical Parallel Processing: An Introduction to Problem Solving in Parallel*. International Thomson Computer Press, London.

Clough, Y., A. Kruess, D. Kleijn, and T. Tscharntke (2005). Spider diversity in cereal fields: comparing factors at local, landscape and regional scales. *Journal of Biogeography 32*, 2007–2014.

Collier, N. (2003). Repast: An extensible framework for agent simulation. Electronic publication: `http://repast.sourceforge.net/`.

Collier, N., T. Howe, and M. North (2003, June). Onward and upward: The transition to Repast 2.0. In *Proceedings of the First Annual North American Association for Computational Social and Organizational Science Conference, Electronic Proceedings*, Pittsburgh, PA USA.

Collis, J., D. Ndumu, and C. van Buskirk (2000, July). *The Zeus Technical Manual, Release 1.04*. BT. Electronic publication: `http://labs.bt.com/projects/agents/zeus/techmanual/TOC.html`.

Compton, S. (2002). Sailing with the wind: dispersal by small flying insects. In J. M. Bullock, R. E. Kenward, and R. S. Hails (Eds.), *Dispersal Ecology: the 42nd symposium of the British Ecological Society held at the University of Reading, 2-5 April 2001*, pp. 113–133. Blackwell Science, Oxford.

Congleton, W. R., B. R. Pearce, and B. F. Beal (1997). A C++ implementation of an individual/landscape model. *Ecological Modelling 103*, 1–17.

Conroy, M. J., Y. Cohen, F. C. James, Y. G. Matsinos, and B. A. Maurer (1995). Parameter estimation, reliability, and model improvement for spatially explicit models of animal populations. *Ecological Applications 5*(1), 17–19.

Curry, G., B. L. Deuermeyer, and R. M. Feldman (1989). *Discrete Simulation*. Holden-Day, Oakland, CA.

Daniels, M. (1999). Integrating simulation technologies with swarm. Electronic publication: `http://www.santafe.edu/~mgd/anl/anlchicago.html`.

Dean, G. J. (1973a). Aphid colonization of spring cereals. *Annals of Applied Biology 75*, 183–193.

Dean, G. J. (1974a). The overwintering and abundance of cereal aphids. *Annals of Applied Biology 76*, 1–7.

Dean, G. J. W. (1973b). Bionomics of aphids reared on cereals and some Gramineae. *Annals of Applied Biology 73*, 127–135.

Dean, G. J. W. (1974b). Effects of temperature on the cereal aphids *Metopolophium dirhodum* (Wlk.), *Rhopalosiphum padi* (L.) and *Macrosiphum avenae* (F.) (Hem., Aphididae). *Bulletin of Entomological Research 63*, 401–409.

Dean, G. J. W. and N. Wilding (1971). *Entomophthora* infecting the cereal aphids *Metopolophium dirhodium* and *Sitobion avenae*. *Journal of Invertebrate Pathology 18*, 169–176.

DeAngelis, D. and L. Gross (Eds.) (1992). *Individual-based models and approaches in ecology: populations, communities and ecosystems.* Routledge, Chapman and Hall, New York. 544 pp.

DeAngelis, D. L., D. K. Cox, and C. C. Coutant (1980). Cannibalism and size dispersal in young-of-the-year largemouth bass: experiment and model. *Ecological Modelling 8*, 133–148.

DeAngelis, D. L., L. J. Gross, M. A. Huston, W. F. Wolff, D. M. Flemming, E. J. Comiskey, and S. M. Sylvester (1998). Landscape modeling for everglades ecosystem restoration. *Ecosystems 1*, 64–45.

DeAngelis, D. L., K. A. Rose, L. B. Crowder, E. A. Marschall, and D. Lika (1993). Fish cohort dynamics: Application of complementary modeling approaches. *American Naturalist 142*, 604–622.

DEFRA (2000, November). Countryside survey 2000, accounting for nature: Assessing habitats in the UK countryside. Technical report, DEFRA.

Dennis, R. L. H., D. Bart, T. H. Sparks, and E. Pollard (2000). Ecological correlates of island incidence and geographical range among british butterflies. *Biodiversity and Conservation 9*, 343–359.

Dhalival, J. S. and B. Singh (1975). Effects of simulated rain on *Macrosiphum miscanthi*. *Indian Journal of Ecology 3*, 186–187.

Dixon, A. F. G. (1976). Reproductive strategies of the alate morphs of the bird-cherry aphid, *Rhopalosiphum padi* (l.). *Journal of Animal Ecology 45*, 817–830.

Duffield, S. J. and N. J. Aebischer (1994). The effect of spatial scale of treatment with dimethoate on invertebrate population recovery in winter wheat. *Journal of Applied Ecology 31*, 263–281.

Dumoulin, N. (2004, October). Towards a framework for multi-scale models of freshwater fishes. In Eurosis (Ed.), *European Simulation and Modeling Conference 2004 (ESMC2004)*, Paris, France, pp. 175–177. Unesco.

Dunning, J. B. J., D. J. Stewart, B. J. Danielson, B. J. Noon, T. L. Root, R. H. Lamberson, and E. E. Stevens (1995). Spatially explicit population models: current forms and future uses. *Ecological Applications 5*(1), 3–11.

Dupuis, A. and B. Chopard (2001). *Parallel simulation of traffic in Geneva using cellular automata.* Commack, NY, USA: Nova Science Publishers, Inc.

Eisinger, D., H.-H. Thulke, T. Selhorst, and T. Müller (2005). Emergency vaccination of rabies under limited resources: combating or containing? *BMC Infectious Diseases 5*(10), 16.

Ekbom, B. S., S.Wiktelius, and P. Chiverton (1992). Can polyphagous predators control the bird cherry-oat aphid (*Rhopalosiphum padi*) in spring cereals? a simulation study. *Entomologia Experimentalis et Applicata 65*, 215–223.

Elmhagen, B. and A. Angerbjörn (2001). The applicability of metapopulation theory to large mammals. *Oikos 94*, 89–100.

Epstein, J. M. and R. L. Axtell (1996). *Growing Artificial Societies: Social Science from the Bottom Up.* MIT Press, Cambridge, MA.

Fabre, F., J. S. Pierre, C. A. Dedryver, and M. Plantegenest (2006). Barley yellow dwarf disease risk assessment based on bayesian modelling of aphid population dynamics. *Ecological Modelling 193*(3-4), 457–466.

Favis-Mortlock, D. (1998). A self-organizing dynamic systems approach to the simulation of rill initiation and development on hillslopes. *Computers and Geosciences 24*, 353–372.

Ferber, J. (1999). *Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence.* Pearson Education, Harlow.

Ferreira, J. G. (1995). ECOWIN - an object-oriented ecological model for aquatic ecosystems. *Ecological Modelling 79*, 21–34.

Fisher, N. I., T. Lewis, and B. J. J. Embleton (1987). *Statistical analysis of spherical data.* Cambridge University Press, Cambridge.

Flake, G. W. (1998). *The Computational Beauty of Nature: Computer Explorations of Fractals, Chaos, Complex Systems and Adaptation.* The MIT Press, Cambridge MA.

Foster, G. N., S. Blake, S. J. Tones, I. Barker, and R. Harrington (2004). Occurrence of barley yellow dwarf virus in autumn-sown cereal crops in the United Kingdom in relation to field characteristics. *Pest Management Science 60*, 113–125.

Freier, B., H. Triltsch, and D. Rossberg (1996). GTLAUS: A model of wheat-cereal aphid-predator interaction and its use in complex agroecological studies. *Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz 103*(5), 543–554.

Friesland, V. H. (1994). 'LAUS', ein prognosemodell für den getreideblattlausbefall als beispiel aus dem agrarmeteorologischen softwarepaket 'AMBER'. *Nachrichtenblatt des Deutschen Pflanzenschutzdienstes 46*(12), 287–291.

Fuller, R. M., G. M. Smith, J. M. Sanderson, R. A. Hill, and A. G. Thomson (2002). Land Cover Map 2000: a general description of the UK's new vector GIS based on classification of remotely sensed data. *Cartographic Journal 39*, 15–25.

Fuller, R. M., G. M. Smith, J. M. Sanderson, R. A. Hill, A. G. Thomson, R. Cox, N. J. Brown, R. T. Clarke, P. Rothery, and F. F. Gerard (2000). Module 7: Land cover map 2000. In R. Haines-Young, C. Barr, H. Black, D. Briggs, R. Bunce, R. Clarke, A. Cooper, F. Dawson, L. Firbank, R. M. Fuller, M. T. Furse, M. K. Gillespie, R. Hill, M. Hornung, D. C. Howard, T. McCann, M. D. Morecroft, S. Petit, A. R. J. Sier, S. M. Smart, G. M. Smith, A. P. Stott, R. C. Stuart, and J. W. Watkins (Eds.), *Accounting for nature: assessing habitats in the UK countryside*. DETR, London.

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

GFA-RACE and IEEP (2003). A long term policy perspective for sustainable agriculture: the potential environmental impacts of CAP mid term review proposals. Technical Report GRP-P-158, Report to Defra.

Giere, R. N. (1991). *Understanding Scientific Reasoning*. Harcourt Brace Jovanovich College Publishers, New York.

Gilbert, N. and R. D. Hughes (1971). A model of an aphid population –three adventures. *Journal of Animal Ecology 40*(2), 525–534.

Gilbert, N. and K. G. Troitzsch (1999). *Simulation for the Social Scientist*. Open University Press, Buckingham. 288 pp.

Gimblett, H. R. (2002). *Integrating Geographic Information Systems and Agent-based Modeling Techniques for Simulating Social and Ecological Processes*. Santa Fe Institute studies in the science of complexity. Oxford University Press, New York.

Gosselke, U., H. Triltsch, D. Rossberg, and B. Freier (2001). GETLAUS01: the latest version of a model for simulating aphid population dynamics in dependence on antagonists in wheat. *Ecological Modelling 145*, 143–157.

Green, D. G. (1994). Connectivity and complexity in landscapes and ecosystems. *Pacific Conservation Biology 1*(3), 194–200.

Grimm, V. (1999). Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future? *Ecological Modelling 115*, 129–148.

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

Grimm, V. and S. F. Railsback (2005). *Individual-based Modeling and Ecology*. Princeton Series in Theoretical and Computational Biology. Princeton University Press, Princeton. 480 pp.

Grimm, V., T. Wyszomirski, D. Aikman, and J. Uchmański (1999). Individual-based modelling and ecological theory: synthesis of a workshop. *Ecological Modelling 115*, 275–282.

Haine, E. (1955). Aphid take-off in controlled wind speeds. *Nature 175*, 474–475.

Haines-Young, R. H., C. J. Barr, H. I. J. Black, D. J. Briggs, R. G. H. Bunce, R. T. Clarke, A. Cooper, F. H. Dawson, L. G. Firbank, R. M. Fuller, M. T. Furse, M. K. Gillespie, R. Hill, M. Hornung, D. C. Howard, T. McCann, M. D. Morecroft, S. Petit, A. R. J. Sier, S. M. Smart, G. M. Smith, A. P. Stott, R. C. Stuart, and J. W. Watkins (2000). Accounting for nature: assessing habitats in the UK countryside. Technical report, DETR, London.

Hanski, I. (1997). Habitat destruction and metapopulation dynamics. In S. T. A. Pickett, R. S. Ostfeld, M. Shachak, and G. E. Likens (Eds.), *The Ecological Basis of Conservation: Heterogeneity, Ecosystems and Biodiversity*. Chapman and Hall, New York.

Hanski, I. (1999). *Metapopulation Ecology*. Oxford University Press, Oxford.

Hanski, I., M. Kuussaari, and M. Nieminen (1994). Metapopulation structure and migration in the butterfly *Melitaea cinxia*. *Ecology 75*, 747–762.

Hanski, I. and D. Simberloff (1997). The metapopulation approach, its history, conceptual domain and application to conservation. In I. Hanski and M. E. Gilpin (Eds.), *Metapopulation biology: ecology, genetics and evolution*, pp. 5–26. Academic Press.

Hardie, J. (1994). Variation in behavioural migration in aphids. *European Journal of Entomology 91*, 115–120.

Harrewijn, P., H. A. Hoof, and J. P. W. Norrdink (1981). Flight behaviour of the aphid *Myzus persicae* during its maiden flight. *Netherlands Journal of Pant Pathology 87*, 111–117.

Harrington, R. (2002). BYDV: The heat is on. In M. Henry and A. McNab (Eds.), *Barley Yellow Dwarf Disease: recent advances and future strategies*. Proceedings of an International Symposium held at El Batán, Texcoco, Mexico: CIMMYT, Mexico.

Harrison, S. (1994). Metapopulations and conservation. In P. J. Edwards, R. M. May, and N. R. Webb (Eds.), *Large-scale ecology and conservation biology*, pp. 111–128. Blackwell, Oxford.

Henien, K., J. Wegner, and G. Merriam (1998). Population effects of landscape model manipulation on two behaviourally different woodland small mammals. *Oikos 81*, 168–186.

Heppenstall, A. J. (2004). *Application of Hybrid Intelligent Agents to Modelling a Dynamic, Locally Interacting Retail Market.* Phd thesis, University of Leeds, UK.

Hewitt, C. (1977). Viewing control structures as patterns of passing messages. *Artificial Intelligence 8*, 323–364.

Hill, J. K., C. D. Thomas, and O. T. Lewis (1996). Effects of habitat patch size and isolation on dispersal by *Hesperia comma* butterflies: implications for metapopulation structure. *Journal of Animal Ecology 65*, 725–735.

Hogeweg, P. and B. Hesper (1990). Individual-oriented modelling in ecology. *Mathematical and Computer Modelling 13*, 83–90.

Holland, J. M., M. Begbie, T. Birkett, S. Southways, S. R. Thomas, C. J. Alexander, and C. F. G. Thomas (2004). The spatial dynamics and movement of *Pterosticus melanarius* and *Pterosticus madidus* (Carabidae) between and within arable fields in the UK. *International Journal of Ecological and Environmental Science 30*, 35–53.

Holst, N., J. A. Axelsen, J. E. Olesen, and P. Ruggle (1997). Object-oriented implementation of the metabolic pool model. *Ecological Modelling 104*, 175–187.

Holst, N. and P. Ruggle (1997). A physiologically based model of pest-natural enemy interactions. *Experimental and Applied Acarology 21*, 325–341.

Hopkins, T. and D. R. Morse (1996, October). The implementation and visualisation of a large spatial individual-based model using Fortran 90. Technical Report 18-96*, University of Kent, Computing Laboratory, University of Kent, Canterbury, UK.

Hughes, R. D. (1963). Population dynamics of the cabbage aphid, *Brevicoryne brassicae* (L.). *Journal of Animal Ecology 32*, 393–424.

Hughes, R. D. and N. Gilbert (1968). A model of an aphid population - a general statement. *Journal of Animal Ecology 37*(3), 553–563.

Hunter, M. D. (2002). Landscape structure, habitat fragmentation and the ecology of insects. *Agricultural and Forest Entomology 4*, 159–166.

Huston, M., D. DeAngelis, and W. Post (1988). New computer models unify ecological theory. *BioScience 38*(10), 682–691.

Inchiosa, M. E. and M. T. Parker (2002). Overcoming design and development challenges in agent-based modeling using ASCAPE. *PNAS 99*(Suppl 3), 7304–7308.

Jeanneret, P., B. Schüpbach, L. Pfiffner, and T. Walter (2003). Arthropod reaction to landscape and habitat features in agricultural landscapes. *Landscape Ecology 18*(3), 253–263.

Jennings, N. R. (2000). On agent-based software engineering. *Artificial Intelligence 117*, 277–296.

Jin, D. and K. Lee (2001). Impacts and limitations of intelligent agents in electronic commerce. *Lecture Notes in Computer Science 2105*, 33–48.

Johnson, C. G. (1962). Aphid migration. *New Scientist 305*, 622–625.

Kaiser, H. (1979). The dynamics of populations as result of the properties of individual animals. *Fortschritte der Zoologie 25*, 109–136.

Keitt, T. H. and A. R. Johnson (1995). Spatial heterogeneity and anomalous kinetics: Emergent patterns in diffusion limited predatory-prey interaction. *Journal of Theoretical Biology 172*(2), 127–139.

Kennedy, J. S. and C. O. Booth (1963). Free flight of aphids in the laboratory. *Journal of Experimental Biology 40*, 67–85.

Kennedy, J. S., C. O. Booth, and W. J. S. Kershaw (1959). Host finding by aphids in the field II. *Aphis fabae* Scop. (gynoparae) and *Brevicoryne brassicae* L.; with a re-appraisal of the role of host-finding behaviour in virus spread. *Annals of Applied Biology 47*(3), 424–444.

Kennedy, J. S., C. O. Booth, and W. J. S. Kershaw (1961). Host finding by aphids in the field III. visual attraction. *Annals of Applied Biology 49*, 1–21.

Kephart, J., J. Hanson, and A. Greenwald (2000). Dynamic pricing by software agents. *Computer Networks 32*(6), 731–752.

Kinny, D. and M. Georgeff (1996). Modelling and design of multi-agent systems. In J. Muller, M. Wooldridge, and N. Jennings (Eds.), *Intelligent Agents III: LNAI Vol. 1193*. Springer, Berlin.

Kirkby, M. J. (1996). A role for theoretical models in geomorphology? In B. L. Rhoads and C. E. Thorn (Eds.), *The Scientific Nature of Geomorphology*, pp. 257–272. Wiley, Chichester.

Klingauf, A. (1987). Host plant finding and acceptance. In A. K. Minks and P. Harrewijn (Eds.), *Aphids: Their Biology, Natural Enemies and Control*, World Crop Pests, pp. 209–223. Elsevier, Amsterdam.

Korie, S., S. J. Clark, J. N. Perry, M. A. Mugglestone, P. W. Bartlett, E. J. P. Marshall, and J. A. Mann (1998). Analysing maps of dispersal around a single focus. *Environmental and Ecological Statistics 5*(4), 317–344.

LaGra, J. J. (1991). Assessing patch shape in landscape mosaics. *Photogrammetric Engineering and Remote Sensing 57*(3), 285–293.

Lane, S., K. Richards, and J. Chandler (1994). Distributed sensitivity analysis in modelling environmental systems. *Proceedings of the Royal society, Series A 447*, 49–63.

Lane, S. N. and K. S. Richards (2001). The 'validation' of hydrodynamic models: Some critical perspectives. In M. G. Anderson and P. D. Bates (Eds.), *Model Validation: Perspectives in Hydrological Science*, pp. 413–438. Wiley, Chichester.

Laval, P. (1995). Hierarchical object-oriented design of a concurrent, individual-based, model of a pelagic tunicate bloom. *Ecological Modelling 82*, 265–276.

Laval, P. (1996). The representation of space in an object-oriented computational pelagic ecosystem. *Ecological Modelling 88*, 113–124.

Leather, S. R. and A. F. G. Dixon (1982). Secondary host preferences and reproductive activity of the bird cherry-oat aphid, *Rhopalosiphum padi*. *Annals of Applied Biology 101*, 219–228.

Leather, S. R., K. F. A. Walters, and A. F. G. Dixon (1989). Factors determining the pest status of the bird cherry-oat aphid, *Rhopalosiphum padi* (L.) (Hemiptera: Aphididae), in Europe: a study and review. *Bulletin of Entomological Research 79*, 345–360.

Lefley, M. and I. D. McKew (2004). Can a parallel agent approach to genetic algorithms reduce search times. In A. Loffi and J. Garobaldi (Eds.), *Applications and Science in Soft Computing*, pp. 69–74. New York: Cambridge University Press.

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

Lewis, T. (1965). The effect of an artificial windbreak on the distribution of aphids in a lettuce crop. *Annals of Applied Biology 55*, 513–518.

Lindenmayer, D. B., M. A. Burgman, H. R. Akçakaya, R. C. Lacy, and H. P. Possingham (1995). A review of the generic computer programs ALEX, RAMAS/space and VORTEX for modelling the viability of wildlife metapopulations. *Ecological Modelling 82*, 161–174.

Longley, M., P. C. Jepson, J. Izquierdo, and N. Sotherton (1997). Temporal and spatial changes in aphid and parasitoid populations following applications of deltamethrin in winter wheat. *Entomologia Experimentalis et Applicata 83*, 41–52.

Lorek, H. and M. Sonnenschein (1995). Using parallel computers to simulate individual-oriented models in ecology: a case study. In *Proceedings: ESM '95 European Simulation Multiconference, Prague, June 1995*.

Lorek, H. and M. Sonnenschein (1998). Object-oriented support for modelling and simulation of individual-oriented ecological models. *Ecological Modelling 108*, 77–96.

Lorek, H. and M. Sonnenschein (1999). Modelling and simulation software to support individual-based ecological modelling. *Ecological Modelling 115*, 199–216.

Lotka, A. . J. (1925). *Elements of Physical Biology*. Williams & Wilkins Co., Baltimore. Reprinted in 1956: Elements of mathematical biology. Dover Publications, Inc., New York, New York.

Loxdale, H. D., J. Hardie, S. Halbert, R. Foottit, N. A. C. Kidd, and C. I. Carter (1993). The relative importance of short-range and long-range movement of flying aphids. *Biological Review 68*, 291–311.

Loxdale, H. D. and G. Lushai (1999). Slaves of the environment: the movement of herbivorous insects in relation to their ecology and genotype. *Philosophical Transactions of the Royal Society London B 354*, 1479–1498.

MacArthur, R. H. and E. O. Wilson (1967). *The Theory of Island Biogeography*. Princeton University Press, Princeton.

Mann, B. P. and S. D. Wratten (1987). A computer-based advisory system for the control of *Sitobion avenae* and *Metopolophium dirhodum*. *IOBC/WPRS Bulletin 10*, 143–155.

Mann, J. A., R. Harrington, D. Morgan, K. F. A. Walters, I. Barker, S. J. Tones, and G. N. Foster (1996). Towards decision support for control of barley yellow dwarf vectors. In *Brighton Crop Protection Conference: Pests and Diseases*, pp. 179–184.

Mathevet, R., F. Bousquet, C. Le Page, and M. Antona (2003). Agent-based simulations of interactions between duck population, farming decisions and leasing of hunting rights in the Carmargue (Southern France). *Ecological Modelling 165*, 107–126.

Melbourne, B. A., K. F. Davies, C. R. Margules, D. B. Lindenmayer, D. A. Saunders, C. Wissel, and K. Henle (2004). Species survival in fragmented landscapes: where to from here? *Biodiversity and Conservation 13*, 275–284.

Metz, J. A. J. and A. M. de Roos (1992). The role of physiologically structured popualation models within a general individual based model perspective. In D. L. DeAngelis and L. J. Gross (Eds.), *Individual Based Models and Approaches in Ecology: Concepts and Models*, pp. 88–111. Routledge, Chapman and Hall, New York.

Minar, M., R. Burkhart, C. Langton, and M. Askenazy (1996). *The Swarm Simulation System: A Toolkit for Building Multi-agent Simulations*. Santa Fe Institute.

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

Miramontes, O. and O. DeSouza (1996). The nonlinear dynamics of survival and social facilitation in termites. *Journal of Theoretical Biology 181*, 373–380.

Moericke, B. V. (1955). Über die Lebensgewohnheiten der geflügelten Blattläuse (*Aphidina*) unter besonderer Berücksichtigung des Verhaltens beim Landen. *Zeitschrift für Angewandte Entomologie 37*, 29–91.

Moloney, K. A. and S. A. Levin (1996). The effects of disturbance architecture on landscape-level population dynamics. *Ecology 77*(2), 375–394.

Morgan, D. (1990). *Modelling the Epidemiology of Barley Yellow Dwarf Virus.* Phd Thesis, University of Southampton, England.

Morgan, D. (2000). Population dynamics of the bird cherry-oat aphid, *Rhopalosiphum padi* (L.), during the autumn and winter: a modelling approach. *Agricultural and Forest Entomology 2*, 297–304.

Myers, J. H. (1976). Distribution and dispersal in populations capable of resource depletion - a simulation study. *Oecologia 23*, 255–269.

Newnham, R. M. (1964). *The Development of a Stand Model for Douglas Fir.* Phd thesis, University of British Columbia, Canada.

Nottingham, S. F., J. Hardie, and G. M. Tatchell (1991). Flight behaviour of the bird cherry aphid, *Rhopalosophum padi. Physiological Entomology 16*, 223–229.

Openshaw, S. and I. Turton (2000). *High performance computing and the art of parallel programming : an introduction for geographers, social scientists, and engineers.* Routledge, London.

Östman, O., B. Ekbom, and J. Bengtsson (2001). Landscape heterogeneity and farming practice influence biological control. *Basic and Applied Ecology 2*, 365–371.

Pacheco, P. S. (1997). *Parallel Programming with MPI.* Morgan Kauffman Publishers, San Francisco, CA.

Pahl-Wostl, C. (2002, June). Agent based simulation in integrated assessment and resources management. In A. Rizolli and A. Jakeman (Eds.), *Integrated Assessment and Decision Support, Proceedings of the First Biennial Meeting of the International Environmental Modelling and Software Society*, Volume 2, Manno, Switzerland, pp. 239–244. International Environmental Modelling and Software Society.

Palmer, M. W. (1992). The coexistence of species in fractal landscapes. *American Naturalist 139*(2), 375–397.

Parker, M. T. (2001). What is Ascape and why should you care? *Journal of Artificial Societies and Social Simulation 4*(1).

Parrott, L. (2002). Complexity and the limits of ecological engineering. *Transactions of the ASAE 45*(5), 16971702.

Parrott, L. and R. Kok (2000). Incorporating complexity in ecosystem modelling. *Complexity International 7*, 1–19.

Parrott, L. and R. Kok (2001). Use of an object-based model to represent complex features of ecosystems. *InterJournal of Complex Systems Manuscript 371*, 8 pages.

Pierre, J. S. and C. A. Dedryver (1984). Un modèle de régression multiple appliqué à la prévision des pullulations d'un puceron des céréals *Sitobion avenae* F., sur blé d'hiver. *Acta Oecologica, Oecol. applic. 7*, 167–179.

Pironneau, O. (1982). On the transport diffusion algorithm and its applications to the navier-stokes equations. *Numerische Mathematik 38*, 309–332.

Plantegenest, M., J. S. Pierre, C. A. Dedryver, and P. Kindlmann (2001, 410). Assessment of the relative impact of different natural enemies on population dynamics of the grain aphid *Sitobion avenae* in the field. *Ecological Entomology 26*(4), 404.

Portugali, J. (1999). *Self-organization and the City*. Springer, Berlin.

Rabbinge, R., G. W. Ankersmit, and G. A. Pak (1979). Epidemiology and simulation of population development of *Sitobion avenae* in winter wheat. *Netherlands Journal of Plant Pathology 85*(5), 197–220.

Ratnieks, F. L. W. and C. Anderson (1999). Task partitioning in insect societies. *Insectes Sociaux 46*, 95–108.

Renshaw, E. (1991). *Modelling biological populations in space and time*. Number 11 in Cambridge Studies in Mathematical Biology. Cambridge University Press, Cambridge.

Repenning, A., A. Ioannidou, and J. Zola (2000). Agentsheets: End-user programmable simulations. *Journal of Artificial Societies and Social Simulation 3*(3).

Reynolds, J. F. and B. Acock (1997). Modularity and genericness in plant and ecosystem models. *Ecological Modelling 94*, 7–16.

Robert, Y. (1987). Aphids and their environment: Dispersion and migration. In A. K. Minks and P. Harrewijn (Eds.), *Aphids: Their biology, natural enemies and control*, World Crop Pests, pp. 299–313. Elsevier, Amsterdam.

Robinson, R. A. and W. J. Sutherland (2002). Post-war changes in arable farming and biodiversity in Great Britain. *Journal of Applied Ecology 39*(157-176), 157–176.

Room, P. M., J. S. Hanan, and P. Prusinkiewicz (1996). Virtual plants: New perspectives for ecologists, pathologists and agricultural scientists. *Trends in Plant Science 1*, 33–38.

Russell, S. and P. Norvig (1995). *Artificial Intelligence: A Modern Approach* (1st ed.). Prentice-Hall, Englewood Cliffs, NJ.

Rykiel, E. J. (1996). Testing ecological models: the meaning of validation. *Ecological Modelling 90*, 229–244.

Saffre, F., R. Furey, B. Krafft, and J. L. Deneubourg (1999). Collective decision-making in social spiders: Dragline-mediated amplification process acts as a recruitment mechanism. *Journal of Theoretical Biology 198*, 507–517.

Scheffer, M., J. M. Baveco, D. L. DeAngelis, K. A. Rose, and E. H. van Nes (1995). Super-individuals: a simple solution for modelling large populations on an individual basis. *Ecological Modelling 80*, 161–170.

Schelhorn, T., D. O'Sullivan, and M.Haklay (1999). STREETS: An agent-based pedestrian model. CASA Working Paper 9, University College London, Centre for Advanced Spatial Analysis.

Schmeeckle, M. W. and J. M. Nelson (2003). Direct numerical simulation of bedload transport using a local, dynamic boundary condition. *Sedimentology 50*, 279–301.

Sequeira, R. A., R. L. Olson, and J. M. McKinion (1997). Implementing generic, object-oriented models in biology. *Ecological Modelling 94*, 17–31.

Shaw, M. J. P. (1970). Effects of population density on alienicolae of *aphis fabae* Scop. III. the effect of isolation on the development of form and behaviour of alatae in a laboratory clone. *Annals of Applied Biology 65*, 205–212.

Sibly, R. M., H. R. Akcakaya, C. J. Topping, and R. J. O'Connor (2005). Population-level assessment of risks of pesticides to birds and mammals in the UK. *Ecotoxicology 14*(8), 863–876.

Silcock, P., V. Swales, and W. Manley (2003, December). The potential environmental impacts of the CAP reform agreement. DEFRA Report No. GRP-P-172.

Silcock, P., V. Swales, G. Smith, and K. Sealy (2004, April). Impacts of CAP reform agreement on diffuse water pollution from agriculture. DEFRA Report No. GRP-P-175.

Skirvin, D. J., J. N. Perry, and R. Harrington (1997). The effect of climate change on an aphid-coccinellid interaction. *Global Change Biology 3*, 1–11.

Stevens Colella, V., E. Klopfer, and M. Resnick (2001). *Adventures in Modelling: Exploring Complex, Dynamic Systems with StarLogo*. Teachers College Pr, New York.

Sumpter, D. J. T. and D. S. Broomhead (1998). Formalising the link between worker and society in honey bee colonies. *Lecture Notes in Computer Science 1534*, 95–110.

Taboada, G. L., J. Tourino, and R. Doallo (2003). Performance analysis of Java message-passing libraries in fast Ethernet, Myrinet and SCI clusters. In *IEEE International Conference on Cluster Computing CLUSTER*, Hong Kong, pp. 118–126.

Tatchell, G. M. (1989). An estimate of the potential economic-losses to some crops due to aphids in Britain. *Crop Protection 8*, 25–29.

Taylor, L. R. (1974). Insect migration, flight periodicity and the boundary layer. *Journal of Animal Ecology 43*, 225–238.

Taylor, L. R. and J. M. P. Palmer (1972). Aerial sampling. In H. F. van Emden (Ed.), *Aphid Technology*, pp. 189–234. Academic Press, London.

Tenhumberg, B. (2004, June). Predicting predation efficiency of biocontrol agents: linking behavior of individuals and population dynamics. In C. Pahl-Wostl, S. Schmidt, and T. Jakeman (Eds.), *iEMSs 2004 International Congress: "Complexity and Integrated Resources Management"*., Osnabrueck, Germany. International Environmental Modelling and Software Society.

Tenhumberg, B., A. J. Tyre, and B. D. Roitberg (2000). Stochastic variation in food availability influences weight and age at maturity. *Journal of Theoretical Biology 202*(4), 257–272.

Terna, P. (1998). Simulation tools for social scientists: Building agent based models with SWARM. *Journal of Artificial Societies and Social Simulation 1*(2).

Thies, C., I. Roschewitz, and T. Tscharntke (2005). The landscape context of cereal aphid-parasitoid interactions. *Proceedings of the Royal Society of London B 272*, 203–210.

Tischendorf, L. and L. Fahrig (2000). How should we measure landscape connectivity? *Landscape Ecology 15*, 633–641.

Tobias, R. and C. Hofmann (2004). Evaluation of free java-libraries for social-scientific agent based simulation. *Journal of Artificial Societies and Social Simulation 7*(1), 1–26.

Topping, C. J., T. S. Hansen, T. S. Jensen, J. U. Jepsen, F. Nikolajsen, and P. Odderskær (2003). ALMaSS, an agent-based model for animals in temperate European landscapes. *Ecological Modelling 167*, 65–82.

Topping, C. J., R. M. Sibly, H. R. Akcakaya, G. C. Smith, and D. R. Crocker (2005). Risk assessment of UK skylark populations using life-history and individual-based landscape models. *Ecotoxicology 14*(8), 925–936.

Tscharntke, T. and R. Brandl (2004). Plant-insect interactions in fragmented landscapes. *Annual Review of Entomology 49*, 405–430.

Tyler, J. A. and K. A. Rose (1994). Individual variability and spatial heterogeneity in fish population models. *Reviews in Fish Biology and Fisheries 4*(1), 91–123.

Usunoff, E., J. Carrera, and S. . F. Mousavi (1992). An approach to the design of experiments for discriminating among alternative conceptual models. *Advances in Water Resources 15*, 199–214.

van Emden, H. F. (Ed.) (1972). *Aphid Technology*. Academic Press, London.

van Nes, E. H. and M. Scheffer (2005). A strategy to improve the contribution of complex simulation models to ecological theory. *Ecological Modelling 185*, 153–164.

Van Winkle, W., K. A. Rose, B. J. Shuter, B. I. Jager, and B. D. Holcomb (1997). Effects of climatic temperature change on growth, survival, and reproduction of rainbow trout: predictions from a simulation model. *Canadian journal of fisheries and aquatic sciences 54*(11), 2526–2542.

Volterra, V. (1926). Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. *Mem. R. Accad. Naz. dei Lincei 2*, 31–113.

Vorley, W. T. and S. D. Wratten (1985). A simulation-model of the role of parasitoids in the population development of *Sitobion avenae* (Hemiptera: Aphididae) on cereals. *Journal of Applied Ecology 22*(3), 813–823.

Walters, K. F. A. and A. F. G. Dixon (1984). The effect of temperature and wind on the flight activity of cereal aphids. *Annals of Applied Biology 104*, 17–26.

Walters, R. J., M. Hassall, M. G. Telfer, G. M. Hewitt, and J. P. Palutikof (2006). Modelling dispersal of a temperate insect in a changing climate. *Proceedings of the Royal Society B 273*(1597), 2017–2023.

Ward, S. A., S. R. Leather, J. Pickup, and R. Harrington (1998). Mortality during dispersal and the cost of host specificity in parasites: how many aphids find hosts? *Journal of Animal Ecology 67*, 763–773.

Wiegand, T., K. A. Moloney, J. Naves, and F. Knauer (1999). Finding the missing link between landscape structure and population dynamics: a spatially explicit perspective. *The American Naturalist 154*(6), 605–627.

Wiktelius, S. (1981). Diurnal flight periodicities and temperature thresholds for flight for different migrant forms of *Rhopalosiphum padi* L. (Hom., Aphididae). *Zeitschrift für Angewandte Entomologie 92*, 449–457.

Wiktelius, S. and J. Pettersson (1985). Simulations of bird cherry-oat aphid population dynamics: a tool for development strategies for breeding aphid-resistant plants. *Agriculture, Ecosystems and Environment 14*, 159–170.

Wilder, J. W. (1999). A predictive model for gypsy moth population dynamics with model validation. *Ecological Modelling 116*, 165–181.

Williams, C. T. (1980). Low temperature mortality of cereal aphids. *Bulletin WPRS/SROP 3*, 63–66.

With, K. A. (2004). Assessing the risk of invasive spread in fragmented landscapes. *Risk Analysis 24*(4), 803–815.

Woiwod, I. P. (1991). The ecological importance of long-term synoptic monitoring. In L. G. Firbank, N. Carter, J. F. Darbyshire, and G. R. Potts (Eds.), *The Ecology of Temperate Cereal Fields, 32nd Symposium of the British Ecological Society*, pp. 275–304. Blackwell Scientific, Oxford.

Woiwod, I. P. and G. M. Tatchell (1984). Computer mapping of aphid abundance. In *British Crop Protection Conference, Pests and Diseases*, Volume 2, pp. 675–683. BCPC.

Wood, J. (2002). *Java Programming for Spatial Sciences*. Taylor and Francis, London.

Woods, J. D. and W. Barkmann (1994). Simulating plankton ecosystem by the Lagrangian Ensemble Method. *Philosophical Transactions of the Royal Society London B 343*, 27–31.

Wooldridge, M. (1999). Intelligent agents. In G. Weiss (Ed.), *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*, pp. 27–78. MIT Press.

Wratten, S. D. (1977). Reproductive strategy of winged and wingless morphs of the aphids *Sitobion avenae* and *Metopolophium dirhodum*. *Annals of Applied Biology 85*, 319–331.

Ziv, Y. (1998a). The effect of habitat heterogeneity on species-diversity patterns: A community-level approach using an object-oriented landscape simulation model (SHALOM). *Ecological Modelling 111*, 135–170.

Ziv, Y. (1998b). *Effects of Environmental Heterogeneity on Species Diversity: A New Process-Based, Multi-Species, Landscape Simulation Model (SHALOM)*. Ph. D. thesis, Department of Ecology and Evolutionary Biology, University of Arizona.

APPENDIX A

Equations

A.1 General aphid

Alate Exule Immigration derived from suction trap estimates (Taylor and Palmer, 1972).

$$\text{ALTIM} = 0.0237 \times \text{IMM} \tag{A.1}$$

where ALTIM = the number of alates colonising the crop per m^2 per day and IMM = the number of alate exules caught in a suction trap.

Simpson's rule (Barlow and Dixon, 1980: pp 66)

$$\text{DEV} = (\text{DEV}_{\text{max}} + \text{DEV}_{\text{min}} + \text{DEV}_{\text{mean}})/3 \tag{A.2}$$

where DEV = proportion of development experienced per day, and DEV_{max}, DEV_{min} and DEV_{mean} = development rates at the daily maximum, minimum and mean temperatures, respectively.

Survival (Morgan, 2000)

$$\text{SURV} = 0.9511 - 0.0173 \times \text{DDB} \tag{A.3}$$

where SURV = daily survival rate and DDB = the number of day-degrees below 2.8 experienced by the aphids each day.
$(r = 0.703; d.f. = 15; P < 0.01)$

A.2 *Rhopalosiphum padi*

Nymphal development, apterae (Morgan, 2000)

$$\text{APDEV} = -0.015 + 0.291/(1 + \text{EXP}(-0.138 \times (\text{TEMP} - 16.911))) \tag{A.4}$$

where APDEV = development rate of apterous nymphs and TEMP = either daily maximum, minimum or mean temperature.
$(F = 509.6; d.f. = 3.1; P < 0.001)$

Nymphal development, alatae (Morgan, 2000)

$$\text{ALDEV} = \text{APDEV}/1.5 \tag{A.5}$$

where ALDEV = development rate of alate nymphs and APDEV = development rate of apterous nymphs.

Adult development, apterae (Morgan, 2000)

$$\text{ADAPDEV} = 0.0193 + 0.0039 \times \text{TEMP} \tag{A.6}$$

where ADAPDEV = development rate of apterous adults and TEMP = either daily maximum, minimum or mean temperature.
$(r = 0.9887; d.f. = 2; P < 0.02)$

Adult development, alateae (Morgan, 2000)

$$\text{ADALDEV} = 0.0184 + 0.0037 \times \text{TEMP} \tag{A.7}$$

where ADALDEV = development rate of alate adults and TEMP = either daily maximum, minimum or mean temperature.
$(r = 0.9778; d.f. = 2; P < 0.05)$

Reproduction, apterous adults (Morgan, 2000)

$$\text{APFEC} = -0.036 + 5.825/(1 + \text{EXP}(-0.319 \times (\text{TEMP} - 12.03))) \tag{A.8}$$

where APFEC = the number of nymphs laid per day per apterous adult and TEMP = either daily maximum, minimum or mean temperature.
$(F = 227.1; d.f. = 3.1; P < 0.05)$

Reproduction, alate adults (Morgan, 2000)

$$\text{ALFEC} = \text{APFEC}/1.3 \tag{A.9}$$

where ALFEC = the number of nymphs laid per day per alate adult and APFEC = the number of nymphs laid per day per apterous adult.

Morph determination (Morgan, 2000)

$$\text{ALPROP} = 0.002 + 0.991/(1 + \text{EXP}(-0.076 \times ((\text{D}/300) - 67.416))) \tag{A.10}$$

where ALPROP = the proportion of newly laid nymphs that will become alate and D = the number of aphids per m^2.
$(F = 1006.8; d.f. = 3.7; P < 0.001)$

A.3 *Sitobion avenae*

Nymphal development, apterae

$$\text{APDEV} = -0.0238 + 0.1410/(1 + \text{EXP}(-0.1817 \times (\text{TEMP} - 8.77))) \tag{A.11}$$

where APDEV = development rate of apterous nymphs and TEMP = either daily maximum, minimum or mean temperature.
$(F = 397.95; d.f. = 8; P < 0.001)$

It was assumed that alatiform nymphs took 1.5 times longer to develop than apterous nymphs (Hughes, 1963), an assumption also used by (Carter et al., 1982) for *S. avenae*:

Nymphal development, alatae

$$\text{ALDEV} = \text{APDEV}/1.5 \tag{A.12}$$

where ALDEV = development rate of alate nymphs and APDEV = development rate of apterous nymphs.

Adult development, apterae and alateae

$$\text{ADAPDEV} = 0.0030 \times \text{TEMP} \tag{A.13}$$

where ADAPDEV = development rate of apterous adults and TEMP = either daily maximum, minimum or mean temperature.
$(r = 0.9142; d.f. = 3; P < 0.01)$

Reproduction, apterous adults

$$\text{APFEC} = -0.24 + 3.13/(1 + \text{EXP}(-0.37 \times (\text{TEMP} - 6.7))) \tag{A.14}$$

where APFEC = the number of nymphs laid per day per apterous adult and TEMP = either daily maximum, minimum or mean temperature.
$(F = 2.41; d.f. = 4; P < 0.5)$

It was assumed that alatiform aphids are 1.3 times less fecund than apterous aphids (Wratten, 1977):

Reproduction, alate adults

$$\text{ALFEC} = \text{APFEC}/1.3 \tag{A.15}$$

where ALFEC = the number of nymphs laid per day per alate adults and APFEC = the number of nymphs laid per day per apterous adult.

A.4 Wind speed

This equation was derived from an analysis of 1994-2000 wind data obtained from BADC recorded at Rothamsted ($R^2 = 0.4082$):

$$\text{speed} = (0.0001334 \times \text{julianday}^2) - (0.056564 \times \text{julianday}) + 11.778 \tag{A.16}$$

where speed = predicted wind speed and julianday = the julian day of the model iteration at which the wind speed is calculated.

The output from this equation is then perturbed by ±4.0, to give a quasi-random wind speed that relates to the mean wind speed at the time of year.

APPENDIX B

Statistics

B.1 Chi-squared test

Marginal habitat cover and fragmentation	**Good**	**Bad**	**Marginal**	**Total**
f100mc10	3188 (2567.97)	14763590 (1.48E+07)	3451 (14006.89)	14770229
f100mc50	2516(2697.93)	15498450 (1.55E+07)	16767 (14715.76)	15517733
f100mc90	1926 (2771.91)	15913428 (1.59E+07)	27841 (15119.23)	15943195
f10mc10	3158 (2562.98)	14735282 (1.47E+07)	3054 (13979.64)	14741494
f10mc50	2761 (2629.29)	15106948 (1.51E+07)	13181 (14341.32)	15122890
f10mc90	2443 (2703.04)	15522853 (1.55E+07)	21830 (14743.63)	15547126
f50mc10	3084 (2574.31)	14800329 (1.48E+07)	3270 (14041.46)	14806683
f50mc50	2765 (2679.99)	15396277 (1.54E+07)	15461 (14617.86)	15414503
f50mc90	2102 (2755.58)	15821469 (1.58E+07)	25741 (15030.20)	15849312
Total	23943	1.38E+08	130596	1.38E+08

Table B.1: Contingency table for χ^2 test to examine the difference between aphid population distribution under different marginal habitat fragmentation scenarios, for each habitat type.

Chi-Sq = 47816.696
DF = 16, P-Value = 0.000

B.2 Rayleigh statistic

Displacement of the population from the origin can be computed simply as the distance δ between the present focus of the population (P) and the original centroid.

The Rayleigh statistic (Fisher et al., 1987) is used to measure in more detail the uniformity in the circular spread of individuals, particularly when the overall displacement of the population from the origin is small (i.e. δ) (Korie et al., 1998). Thus, for the diffusion of aphids from a single point source in a varied environment, this statistic can be used to measure the impact of habitat configuration on the eveness of the spread of individuals from the source. The null hypothesis is that dispersal is unrestricted and isotropic (the same in all directions) (Korie et al., 1998).

Given coordinates $(x_1, y_1), i = 1 \ldots, n$, direction cosines, $(x'_i, y'_i$,, are defined as:

$x'_i = \cos\beta_i, y'_i = \sin\beta_i; 0 \leq \beta_i = \arctan(y_i/x_i) < 2\pi.$

(Korie et al., 1998)

The Rayleigh statistic is the resultant of the direction cosines:

$R^2 = (\sum_i x'_i)^2 + (\sum_i y'_i)^2.$

for $n \geq 10, \omega = 3R^2/n$ approximates well to a chi-square distribution under the null hypothesis of uniformity, with three degrees of freedom.

B.3 Circular statistics tables

Crop cover (%)	**10**	**10**	**10**	**50**	**50**	**50**	**90**	**90**	**90**
Crop fragmentation (%)	**10**	**50**	**100**	**10**	**50**	**100**	**10**	**50**	**100**
origin x	32	25	28	26	27	27	26	26	27
origin y	32	26	28	26	26	26	26	26	26
mean x	31.96	24.89	27.97	25.99	27.18	26.95	26.04	25.99	27.06
mean y	32.29	26.00	28.24	25.92	26.14	26.04	26.01	25.95	26.08
Mean distance moved from P, Δ, (cells)	1.51	1.71	3.06	1.47	2.29	3.51	1.36	1.30	1.67
Mean distance moved from P, Δ, (m)	38	43	77	37	57	88	34	32	42
Mean displacement, δ, xdirection (cells)	-0.04	-0.11	-0.03	-0.01	0.18	-0.05	0.04	-0.01	0.06
Mean displacement, δ, xdirection (m)	-1	-3	-1	0	5	-1	1	0	2
Mean displacement, δ, ydirection (cells)	0.29	0.00	0.24	-0.08	0.14	0.04	0.01	-0.05	0.08
Mean displacement, δ, ydirection (m)	7	0	6	-2	4	1	0	-1	2
Mean displacement, δ, (cells)	0.29	0.11	0.24	0.08	0.23	0.06	0.04	0.05	0.10
mean displacement, δ, (m)	7	3	6	2	6	2	1	1	3
Uniformity of the distribution around P: Rayleigh R^2	3.73E7	5022045	2006869	9231585	1.86E7	74015	2306303	935790	1.30E7
Rayleigh ω	9633	2494	2040	1562	4922	28	333	147	2181
aphid number (moved)	11600	6042	2951	17726	11346	8015	20757	19080	17939
Mean angle (rad)	1.33	4.39	1.63	0.00	1.24	1.75	-0.05	4.41	0.42

Table B.2: Circular statistics to describe the distribution around an origin under different crop configuration scenarios at simulation day 40 (varying crop cover and crop fragmentation), see chapter 7.

Marginal cover (%)	**10**	**10**	**10**	**50**	**50**	**50**	**90**	**90**	**90**
Marginal fragmentation (%)	**10**	**50**	**100**	**10**	**50**	**100**	**10**	**50**	**100**
origin x	32	32	32	32	32	32	32	32	32
origin y	32	32	32	32	32	32	32	32	32
mean x	31.96	31.98	31.98	32.01	31.99	32.01	32.01	31.98	31.99
mean y	32.29	31.97	31.99	32.00	32.01	32.00	32.00	31.97	32.02
Mean distance moved from P, Δ, (cells)	1.53	1.59	1.53	1.54	1.64	1.57	1.57	1.60	1.60
Mean distance moved from P, Δ, (m)	38	40	38	39	41	39	39	40	40
Mean displacement, δ, xdirection (cells)	-0.04	-0.02	-0.02	0.01	-0.01	0.01	0.01	-0.02	-0.01
Mean displacement, δ, xdirection (m)	-1	0	0	0	0	0	0	0	0
Mean displacement, δ, ydirection (cells)	0.29	-0.03	-0.01	0.00	0.01	0.00	0.00	-0.03	0.02
Mean displacement, δ, ydirection (m)	7	-1	0	0	0	0	0	-1	1
Mean displacement, δ, (cells)	0.29	0.04	0.02	0.01	0.01	0.01	0.01	0.04	0.02
mean displacement, δ, (m)	7	1	1	0	0	0	0	1	1
Uniformity of the distribution around P: Rayleigh R^2	3.77E7	2.10E7	3.81E7	3.89E7	3.17E7	4.01E7	3.79E7	3.77E7	4.59E7
Rayleigh ω	9333	5863	9395	10199	7905	8787	8775	9397	10973
aphid number (moved)	12107	10758	12157	11439	12042	13684	12956	12022	12544
Mean angle (rad)	1.20	1.26	1.23	1.23	1.47	1.40	1.11	1.33	1.30

Table B.3: Circular statistics to describe the distribution around an origin under different marginal habitat configuration scenarios at simulation day 40 (varying marginal habitat cover and marginal habitat fragmentation), see chapter 8.

APPENDIX C

Landscape generation code

The following code creates an artifical landscape with each cell representing one of three habitat types (favourable = 1, marginal = 2, unfavourable =0).

```
/** MakeFragMarginal.java**/
/** @author Hazel Parry **/
import java.io.*;
import java.util.*;

public class MakeFragMarginal {
    /** number of neighbours **/
    int noNeighbours = 0;
    /** the current level of habitat coverage **/
    int currentCover = 0;
    /** x size, y size for ascii file **/
    private int xSize = 250, ySize = 250;
    /** 'no data' value for ascii file */
    private int NODATA_value = -9999;
    /** xllcorner location of ascii file */
    private double xllcorner = 515331.166714;
    /** yllcorner location of ascii file */
    private double yllcorner = 218525.556963;
    /** cell size of ascii file */
    private double cellSize = 0.005;
    private PrintWriter out;
    private String textline;
    private double cover = 0;
    private double frag = 0;
    private double margincover = 0;
    private double marginfrag = 0;


    /** MakeFrag constructor **/
    public MakeFragMarginal(String filename) {
        //  reset();
        // ask user for x-size
        System.out.print("Please enter x size: ");
        BufferedReader inX = new BufferedReader(new InputStreamReader(System.in));
        try
        { textline = inX.readLine(); }
        catch (IOException e)
        { System.err.println("Problem reading keyboard input x size"); }
        try
        { setXSize(new Integer(textline).intValue());}
        catch (NumberFormatException e)
        {System.err.println("Error converting input x to integer.");}

        // ask user for y-size
        System.out.print("Please enter y size: ");
        BufferedReader inY = new BufferedReader(new InputStreamReader(System.in));
```

```
try
{ textline = inY.readLine(); }
catch (IOException e)
{ System.err.println("Problem reading keyboard input y size"); }
try
{ setYSize(new Integer(textline).intValue());}
catch (NumberFormatException e)
{System.err.println("Error converting input y to integer.");}

// ask user for cell size
System.out.print("Please enter cell size (km): ");
BufferedReader inCell = new BufferedReader(new InputStreamReader(System.in));
try
{ textline = inCell.readLine(); }
catch (IOException e)
{ System.err.println("Problem reading keyboard input cell size"); }
try
{ setCellSize(new Double(textline).doubleValue());}
catch (NumberFormatException e)
{System.err.println("Error converting input cell size to integer.");}

// ask user for frag
System.out.print("Please enter fragmentation index (0.0 - 1.0): ");
BufferedReader inFrag = new BufferedReader(new InputStreamReader(System.in));
try
{ textline = inFrag.readLine(); }
catch (IOException e)
{ System.err.println("Problem reading keyboard input frag"); }
try
{ frag = new Double(textline).doubleValue();}
catch (NumberFormatException e)
{System.err.println("Error converting input to double.");}

// ask user for margin frag
System.out.print("Please enter marginal fragmentation index (0.0 - 1.0): ");
BufferedReader inMarginFrag = new BufferedReader(new InputStreamReader(System.in));
try
{ textline = inMarginFrag.readLine(); }
catch (IOException e)
{ System.err.println("Problem reading keyboard input marginal frag"); }
try
{ marginfrag = new Double(textline).doubleValue();}
catch (NumberFormatException e)
{System.err.println("Error converting input to double.");}

// ask user for cover
System.out.print("Please enter habitat coverage (decimal percentage e.g. 0.5): ");
BufferedReader inCover = new BufferedReader(new InputStreamReader(System.in));
try
{ textline = inCover.readLine(); }
catch (IOException e)
{ System.err.println("Problem reading keyboard input cover"); }
try
{ cover = new Double(textline).doubleValue();}
catch (NumberFormatException e)
{System.err.println("Error converting input to double.");}

// ask user for marginal cover %
System.out.print("Please enter marginal habitat coverage (decimal percentage e.g. 0.5): ");
BufferedReader inMarginCover = new BufferedReader(new InputStreamReader(System.in));
try
{ textline = inMarginCover.readLine(); }
catch (IOException e)
{ System.err.println("Problem reading keyboard input margin cover"); }
try
{ margincover = new Double(textline).doubleValue();}
catch (NumberFormatException e)
{System.err.println("Error converting input to double.");}


System.out.print("making Frag pattern (cover = " + cover + "margin cover = " + margincover + ", Frag = "
+ frag + ", Margin Frag = " + marginfrag + ", xSize = " + getXSize() + ", ySize = " + getYSize() + ")...");
/** integer array to store grid data to form ascii file */
int terrain[][] = new int[getXSize()][getYSize()];
makeFragPat(cover, margincover, frag, marginfrag, terrain);
System.out.println("DONE");
System.out.print("Printing to File " + filename + "...");
outArcAscii(filename, terrain);
System.out.println("DONE");
```

```
}

/** Method to make the Frag pattern **/
public void makeFragPat(double cover, double margincover, double frag, double marginfrag, int[] [] terrain) {
    /** target habitat coverage derived from desired decimal percentage cover **/
    double targetCover = (double)getXSize() * (double)getYSize() * cover;
    frag *= frag; //square it
    while (currentCover < targetCover) {
        //pick a random grid square
        int x = (int)(Math.random() * getXSize());
        int y = (int)(Math.random() * getYSize());

        /** if the grid square hasn't already been assigned as habitat **/
        if(terrain[x] [y] != 1) {
            /** if r < frag the cell is assigned to habitat**/
            if (Math.random() < frag) {
                terrain[x] [y] = 1;
                currentCover++;
            }
            else {
                /** if a neighbour is habitat, the cell is assigned to habitat**/
                //check for neighbours
                //north
                if (y + 1 < getYSize() -1)
                    noNeighbours += terrain[x] [y+1];
                //north east
                if (y + 1 < getYSize() -1 && x + 1 < getXSize() -1)
                    noNeighbours += terrain[x+1] [y+1];
                //north west
                if (x -1 > -1 && y + 1 < getYSize() -1)
                    noNeighbours += terrain [x-1] [y+1];

                //South
                if (y -1 > -1)
                    noNeighbours += terrain[x] [y-1];
                //South East
                if (y -1 > -1 && x + 1 < getXSize() -1)
                    noNeighbours += terrain[x+1] [y-1];
                //south west
                if (y -1 > -1 && x -1 > -1)
                    noNeighbours += terrain[x-1] [y-1];

                //West
                if (x -1 > -1)
                    noNeighbours += terrain[x-1] [y];

                //East
                if (x + 1 < getXSize() -1)
                    noNeighbours += terrain[x+1] [y];
                // at least one of the eight neighbours must contain habitat
                if (noNeighbours >= 1) {
                    terrain [x] [y] = 1;
                    currentCover ++;
                    noNeighbours = 0;
                }
            }
        }
    }
    // calculate all the cells still non habitat:
    double targetCoverMargin = (((double)getXSize() * (double)getYSize()) - targetCover) * margincover;
    marginfrag *= marginfrag; //square it
    currentCover = 0;
    while (currentCover < targetCoverMargin) {
        //pick a random grid square
        int x = (int)(Math.random() * getXSize());
        int y = (int)(Math.random() * getYSize());

        /** if the grid square hasn't already been assigned as habitat **/
        if(terrain[x] [y] != 1) {
            /** if r < marginfrag the cell is assigned to habitat**/
            if (Math.random() < marginfrag) {
                terrain[x] [y] = 2;
                currentCover++;
            }
            else {
                /** if a neighbour is habitat, the cell is assigned to habitat**/
                //check for neighbours
                //north
                if (y + 1 < getYSize() -1){
```

```
                        if (terrain[x] [y+1] == 2) {
                            noNeighbours++;
                        }
                    }
                    //north east
                    if (y + 1 < getYSize() -1 && x + 1 < getXSize() -1){
                        if (terrain[x+1] [y+1] == 2) {
                            noNeighbours++;
                        }
                    }
                    //north west
                    if (x -1 > -1 && y + 1 < getYSize() -1){
                        if (terrain [x-1] [y+1] == 2) {
                            noNeighbours++;
                        }
                    }
                    //South
                    if (y -1 > -1){
                        if (terrain[x] [y-1] == 2) {
                            noNeighbours++;
                        }
                    }
                    //South East
                    if (y -1 > -1 && x + 1 < getXSize() -1){
                        if (terrain[x+1] [y-1] == 2) {
                            noNeighbours++;
                        }
                    }
                    //south west
                    if (y -1 > -1 && x -1 > -1){
                        if (terrain[x-1] [y-1] == 2) {
                            noNeighbours++;
                        }
                    }
                    //West
                    if (x -1 > -1) {
                        if (terrain[x-1] [y] == 2){
                            noNeighbours++;
                        }
                    }

                    //East
                    if (x + 1 < getXSize() -1) {
                        if (terrain[x+1] [y] == 2) {
                            noNeighbours++;
                        }
                    }
                    // at least one of the eight neighbours must contain habitat
                    if (noNeighbours >= 1) {
                        terrain [x] [y] = 2;
                        currentCover ++;
                        noNeighbours = 0;
                    }
                }
            }
        }

    }


    /** Writes output to file **/
    public void outArcAscii(String filename, int[] []terrain) {

        try {
            out = new PrintWriter(
            new BufferedWriter(
            new FileWriter( filename ) ) );
            out.println("ncols " + getXSize());
            out.println("nrows " + getYSize());
            out.println("xllcorner " + xllcorner);
            out.println("yllcorner " + yllcorner);
            out.println("cellsize " + getCellSize());
            out.println("NODATA_value " + NODATA_value);

            for (int j = 0; j < getYSize(); j++) {
                for (int i = 0; i < getXSize(); i++) {
                    out.print(terrain[i] [j] + " ");
                }
                out.println("");
```

```
            }
        }

        catch( IOException ioException ) {
            System.out.println("A problem was encountered during file io.");
            ioException.printStackTrace();
        }

        out.flush();
        out.close();
    }

    public void setXSize(int xsize) {
        xSize = xsize;
    }

    public int getXSize() {
        return xSize;
    }

    public void setYSize(int ysize) {
        ySize = ysize;
    }

    public int getYSize() {
        return ySize;
    }

    public void setCellSize(double cellsize) {
        cellSize = cellsize;
    }

    public double getCellSize() {
        return cellSize;
    }

  public static void main( String [] args ) {
        MakeFragMarginal makefragmarginal = new MakeFragMarginal("c:/filelocation/filename.asc");
    }

}
```

APPENDIX D

Reparameterising the model for *Sitobion avenae*

Temperature (°C)	Apterous nymph development time (h)	Mean adult lifespan (days)	Mean number of nymphs per week	Mean reproductive life (days)
0	∞	∞	∞	∞
10	405	17 [a]	33	15
12.5	308	-	-	-
15	260	22	46	18
17.5	245	-	-	-
20	212	22	61	17
22.5	201	-	-	-
25	201	19	33	14
27.5	246	-	-	-

[a] Excluded as distorts relationship

Table D.1: Data used to derive logistic functions for *S. avenae* parameters. Note that development is assumed to cease at 0°C

D.1 Development rate

D.1.1 Nymphs

Using the data of Dean (1974b) and Williams (1980) a logistic function of proportion of nymphal development against temperature was derived (figure D.1, equation D.1), in the same way as described in Morgan (1990, 2000) using GenStat. The data used to derive the function is shown in table D.1.

$$\text{APDEV} = -0.0238 + 0.1410/(1 + \text{EXP}(-0.1817 \times (\text{TEMP} - 8.77))) \qquad \text{(D.1)}$$

where APDEV = development rate of apterous nymphs and TEMP = either daily maximum, minimum or mean temperature.

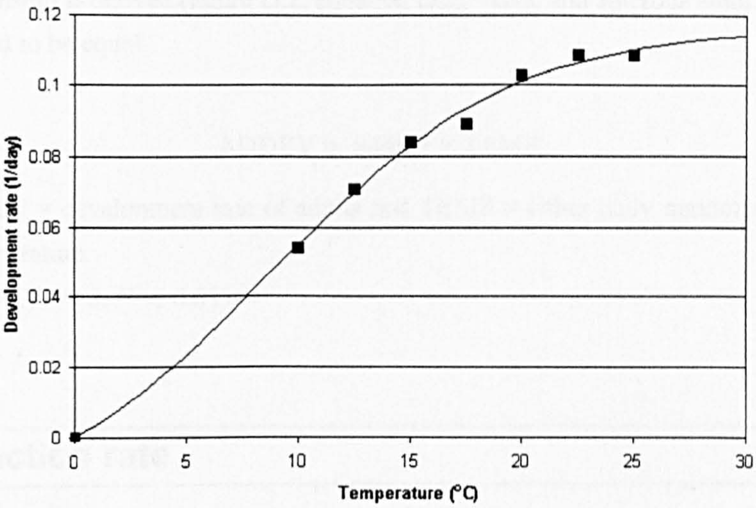


Figure D.1: The effect of temperature on the development rate of *S. avenae* apterous nymphs (Dean, 1974b)

$(F = 397.95; d.f. = 8; P < 0.001)$

It was assumed that alatiform nymphs took 1.5 times longer to develop than apterous nymphs (Hughes, 1963), an assumption also used by (Carter et al., 1982) for *S. avenae*.

D.1.2 Adults

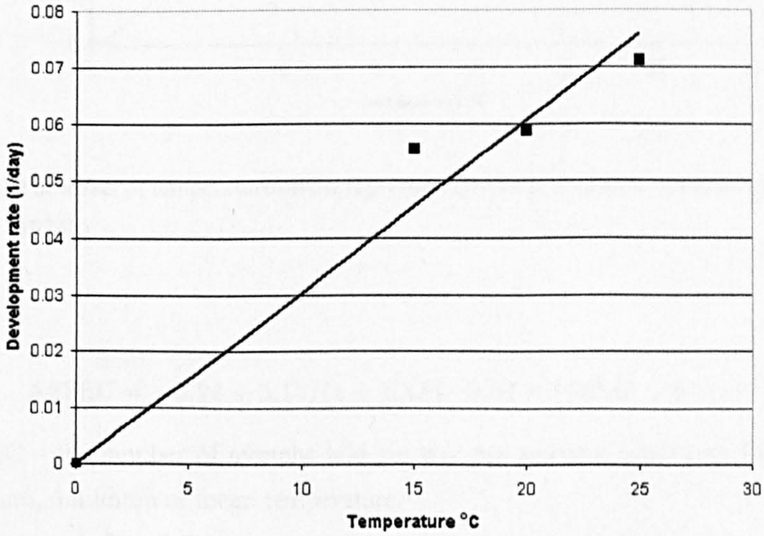


Figure D.2: The effect of temperature on the development rate of *S. avenae* adults (Dean, 1974b)

Development rates for adults were derived in the same way with reference to Dean (1974b) and Williams (1980), however insufficient data was available to generate a logistic function thus a linear-relationship is derived (figure D.2, equation D.2). Alate and apterous adult development were assumed to be equal.

$$\mathrm{ADDEV} = 0.0030 \times \mathrm{TEMP} \qquad \text{(D.2)}$$

where ADDEV = development rate of adults and TEMP = either daily maximum, minimum or mean temperature.
$(r = 0.9589; d.f. = 3; P < 0.01)$

D.2 Reproduction rate

The apterous adult reproduction rate was calculated by using Dean's data for nymphs born per week divided by the mean reproductive life (days) of an adult.

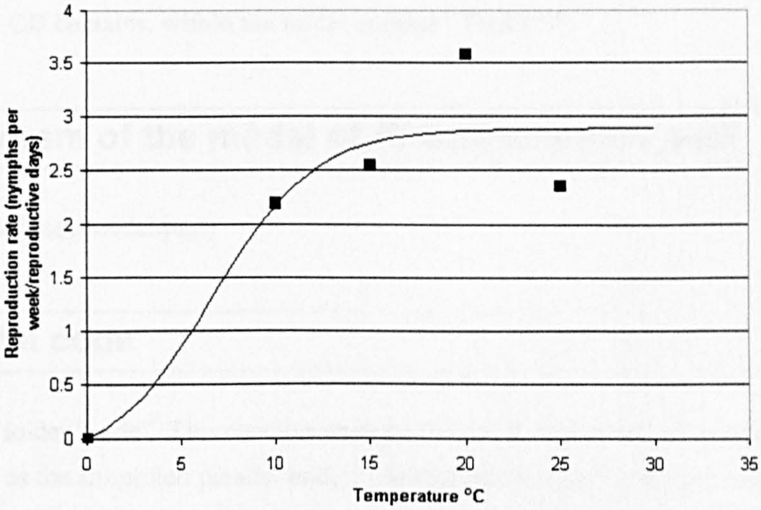


Figure D.3: The effect of temperature on the reproduction rate of *S. avenae* apterous adults (Dean, 1974b)

$$\mathrm{APFEC} = -0.24 + 3.13/(1 + \mathrm{EXP}(-0.37 \times (\mathrm{TEMP} - 6.7))) \qquad \text{(D.3)}$$

where APFEC = the number of nymphs laid per day per apterous adult and TEMP = either daily maximum, minimum or mean temperature.
$(F = 2.41; d.f. = 4; P < 0.5)$

It was assumed that alatiform aphids are 1.3 times less fecund than apterous aphids (Wratten, 1977).

APPENDIX E

Model code and documentation (included on CD)

The included CD contains, within the folder entitled "THESIS":

E.1 UML diagram of the model of *Rhopalosiphum padi*

This is the file insectmodel.png.

E.2 Full model code

This is in the folder "code". This contains code for the full *R. padi* model ("insectmodel" package), as well as the simplified parallel code ("parallelaphidsimple" package) and the full parallel code including pesticide simulation ("parallelfinalpesticide" package). In order for this to run the user must install Repast v. 2.0 included on the CD. Repast v. 3.0 is the current version of Repast (untested with this model), available at `http://repast.sourceforge.net`. The user must also have Java 2 Platform Standard Edition Development Kit (JDK) installed, available at `http://java.sun.com/`, to enable compilation of the code. Batch file examples are given to compile and run the full *R. padi* model. Further tips on running Repast simulations can be found on the Repast website.

E.3 Javadocs

This is in the folder "javadoc". These are full javadocs for the three model code packages included in appendix E.2. These are best accessed by launching the package-summary.html file for each set of code in a web browser such as Microsoft Internet Explorer.